nature communications

Article

https://doi.org/10.1038/s41467-022-32232-0

A public antibody class recognizes an S2 epitope exposed on open conformations of SARS-CoV-2 spike

Received: 30 November 2021

Accepted: 22 July 2022

Published online: 04 August 2022

Check for updates

Mathieu Claireaux[1,2,8], Tom G. Caniels[1,2,8], Marlon de Gast[1,2], Julianna Han[3], Denise Guerra[1,2], Gius Kerster[1,2], Barbera D. C. van Schaik[4], Aldo Jongejan[4], Angela I. Schriek[1,2], Marloes Grobben[1,2], Philip J. M. Brouwer[1,2,3], Karlijn van der Straten[1,2,5], Yoann Aldon[1,2], Joan Capella-Pujol[1,2], Jonne L. Snitselaar[1,2], Wouter Olijhoek[1,2], Aafke Aartse[1,2,6], Mitch Brinkkemper[1,2], Ilja Bontjer[1,2], Judith A. Burger[1,2], Meliawati Poniman[1,2], Tom P. L. Bijl[1,2], Jonathan L. Torres[3], Jeffrey Copps[3], Isabel Cuella Martin[1,2], Steven W. de Taeye[1,2], Godelieve J. de Bree[5], Andrew B. Ward[3], Kwinten Sliepen[1,2], Antoine H. C. van Kampen[4], Perry D. Moerland[4], Rogier W. Sanders[1,2,7] ✉ & Marit J. van Gils[1,2] ✉

Delineating the origins and properties of antibodies elicited by SARS-CoV-2 infection and vaccination is critical for understanding their benefits and potential shortcomings. Therefore, we investigate the SARS-CoV-2 spike (S)-reactive B cell repertoire in unexposed individuals by flow cytometry and single-cell sequencing. We show that ~82% of SARS-CoV-2 S-reactive B cells harbor a naive phenotype, which represents an unusually high fraction of total human naive B cells (~0.1%). Approximately 10% of these naive S-reactive B cells share an IGHV1-69/IGKV3-11 B cell receptor pairing, an enrichment of 18-fold compared to the complete naive repertoire. Following SARS-CoV-2 infection, we report an average 37-fold enrichment of IGHV1-69/IGKV3-11 B cell receptor pairing in the S-reactive memory B cells compared to the unselected memory repertoire. This class of B cells targets a previously undefined non-neutralizing epitope on the S2 subunit that becomes exposed on S proteins used in approved vaccines when they transition away from the native pre-fusion state because of instability. These findings can help guide the improvement of SARS-CoV-2 vaccines.

[1]Amsterdam UMC, University of Amsterdam, Department of Medical Microbiology and Infection prevention, Laboratory of Experimental Virology, Amsterdam, the Netherlands. [2]Amsterdam institute for Infection and Immunity, Infectious diseases, Amsterdam, the Netherlands. [3]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037, USA. [4]Bioinformatics Laboratory, Department of Epidemiology and Data Science, Amsterdam UMC, University of Amsterdam, Amsterdam Institute for Infection and Immunity, Amsterdam Institute for Public Health, Amsterdam, the Netherlands. [5]Department of Internal Medicine, Amsterdam UMC, University of Amsterdam, Amsterdam Institute for Infection and Immunity, Amsterdam Institute for Infection and Immunity, Amsterdam, the Netherlands. [6]Department of Virology, Biomedical Primate Research Centre, Rijswijk, The Netherlands. [7]Department of Microbiology and Immunology, Weill Medical College of Cornell University, New York, NY, USA. [8]These authors contributed equally: Mathieu Claireaux, Tom G. Caniels. ✉e-mail: r.w.sanders@amsterdamumc.nl; m.j.vangils@amsterdamumc.nl

The emergence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) at the end of 2019 and its increased spread as a result of novel viral variants has posed considerable danger to global health. Multiple vaccines are now in use that confer high levels of protection. As of February 4th 2022, 61% of the world's population has received at least one dose of a coronavirus disease 2019 (COVID-19) vaccine, illustrating the rapid development and distribution of vaccines[1]. A major goal of these vaccines is to induce neutralizing antibodies (NAbs) and memory B cells that protect against subsequent infection.

Most licensed vaccines aim to induce immunity against SARS-CoV-2 S, a trimeric glycoprotein on the surface of the virion, that is the only known target for NAbs. It consists of an apical S1 subunit encompassing an N-terminal domain (NTD) and the receptor-binding domain (RBD) that is responsible for binding to the ACE2 receptor; and a membrane-proximal S2 subunit which is responsible for fusion of viral and cellular membranes. Coronavirus (CoV) S proteins can suffer from instability and deteriorate into non-native forms leading to altered exposure of antibody epitopes[2–4]. Therefore, some of the approved vaccines, including those from J&J/Janssen, Moderna and Pfizer/BioNTech, but not those from Oxford/AstraZeneca and several others, contain modifications to stabilize the S proteins. These changes include the removal of the furin cleavage site between S1 and S2 and two proline substitutions (K986P/V987P) in S2, resulting in a prefusion stabilized S trimer termed S-2P[5–7]. Although S-2P is more stable than wild-type S (S-WT), subsequent studies revealed that even S-2P suffers from instability issues and displays open conformations[2,4,8]. Several additional stabilization strategies have been described, including the introduction of an additional four prolines (F817P/A892P/A899P/A942P), resulting in S-6P (HexaPro S), which shows considerably increased stability and resistance to heat/freeze cycles compared to S-2P and S-WT[2,4,8,9].

Early in the pandemic many groups identified and isolated potent NAbs from COVID-19 patients that defined important epitopes on S, which have led to emergency use authorization of several monoclonal antibody (MAb) therapies for COVID-19[10]. Most of the potently neutralizing MAbs target the immunodominant RBD and NTD on the apex of S, whereas MAbs against the S2 domain, while generally less potent, tend to be more broadly reactive across sarbecoviruses and endemic human CoVs (HCoVs)[11,12]. B cells that arise during SARS-CoV-2 infection target both S1 and S2 domains, but their ontogeny is often unclear. S1 NAbs are usually poorly cross-reactive with different SARS-CoV-2 variants or other HCoVs and have not undergone extensive somatic hypermutation (SHM), suggesting that they originate from de novo activation of naive B cells[13–17]. In contrast, a subset of S2 NAbs cross-bind and cross-neutralize endemic HCoVs, suggesting that they arise from pre-existing memory B cells, formed during infection with an endemic HCoV and reactivated upon SARS-CoV-2 infection[18]. Thus, it is conceivable that in SARS-CoV-2 infection and similarly, vaccination, the humoral immune response is driven by both de novo activation of naive B cells and reactivation of memory B cells.

Although the NAb response to SARS-CoV-2 infection and vaccination has been studied extensively, non-neutralizing MAbs (non-NAbs) have not been a focal point. While understudied, non-NAbs make up a substantial portion of the antibody repertoire after infection and the majority of vaccine-induced anti-S Abs are also non-neutralizing[12,19,20]. Non-NAbs can contribute to immunity through effector mechanisms such as antibody-dependent cellular cytotoxicity (ADCC) and phagocytosis (ADCP)[11,21–24], while high levels of pro-inflammatory antibodies may contribute to severe disease[25,26]. Non-NAbs elicited by infection target S as well as other viral proteins such as nucleoprotein (N) and ORF8[19], whereas the non-NAbs induced by currently approved vaccines exclusively target S. S instability can contribute to the induction of non-NAbs because their epitopes become exposed on open/aberrant conformations. Extensive work on respiratory syncytial virus (RSV) and human immunodeficiency virus 1 (HIV-1) has shown that the occurrence of post-fusion or open conformations of the surface glycoproteins leads to the induction of non-NAbs, while stabilizing these glycoproteins in the closed pre-fusion conformation reduced the induction of non-NAbs[27–29]. In particular for RSV, the reduced induction of non-NAbs was accompanied by an enhanced NAb response. Elucidating the properties and epitopes of non-NAbs as well as their B cell origins, is therefore highly relevant for both understanding humoral immunity against SARS-CoV-2 and improving vaccines. Here, we characterized the human baseline B cell repertoire that recognizes SARS-CoV-2 S prior to any antigenic SARS-CoV-2 S encounter through infection or vaccination. We found that an unusually high proportion (~0.1%) of naive B cells is able to recognize SARS-CoV2 S and that these naive B cells show a highly enriched usage of heavy chain immunoglobulin V gene (IGHV)1-69 and kappa chain immunoglobulin V gene (IGKV)3-11 in their B cell receptor (BCR). A subset of these naive B cells exhibits polyreactivity and can be activated upon SARS-CoV-2 S encountering. Following SARS-CoV-2 infection this particular pairing was selected and highly enriched in the memory B cell population and almost exclusively targets an undescribed non-neutralizing apical S2 epitope present on aberrant forms of S, suggesting that further stabilization of S might benefit SARS-CoV-2 vaccine responses.

Results

Multiple B cell compartments recognize SARS-CoV-2 S

Recent studies on sera and memory B cells from recovered COVID-19 patients suggest that a fraction of the antibody response against SARS-CoV-2 S could stem from pre-existing cross-reactive memory B cells induced by prior infection with endemic HCoVs[18,30–32]. Therefore, we interrogated the baseline SARS-CoV-2 S-reactive B cell repertoire from ten unexposed and unvaccinated healthy donors (HD01-10), sampled in 2019 and early 2020, before the first official case report of SARS-CoV-2 infection in the Netherlands. We developed a combinatorial B cell staining approach with labeled antigenic probes, allowing for the simultaneous identification of B cells that are reactive to six different pathogens in a single sample (Table S1). SARS-CoV-2 S-reactive B cells were compared to those recognizing pre-encountered antigens, including influenza A virus hemagglutinin ($H1N1_{pdm09}$ HA), RSV fusion protein (RSV F) and tetanus toxoid (TT), as well as those against unencountered antigens HIV-1 envelope glycoprotein (Env) and hepatitis C virus (HCV) envelope glycoprotein (E1E2; Fig. 1A, Fig. S1A–E). B cells reactive to SARS-CoV-2 were present at high frequency in unexposed individuals (median 0.086% of total B cells), which was ~10–30-fold higher compared to the frequency of B cells reactive with other unencountered antigens HIV-1 Env and HCV E1E2 (median 0.007% and 0.0025%, respectively), but in a similar range to the frequency of B cells reactive with previously encountered antigens H1N1 HA (0.066%), RSV F (0.042%) and TT (0.12%, Fig. 1B).

Next, we analyzed the naive and memory subsets of these antigen-reactive B cells, which can be subdivided into four populations based on their surface expression of IgD and CD27: naive ($IgD^+/CD27^-$), classical memory ($IgD^-/CD27^+$), $CD27^-$ memory ($IgD^-/CD27^-$) and unswitched memory ($IgD^+/CD27^+$) (Fig. 1A lower panel, C–D and Fig. S2A, B). Naive B cells reactive with HIV-1 Env and HCV E1E2 were detected at low frequency (~0.001%) (Fig. 1C). In contrast, the proportion of B cells that recognize SARS-CoV-2 S was significantly higher (0.11%) in the naive compartment than those reactive to HIV-1 Env (~100-fold higher), HCV E1E2 (~90-fold), H1N1 HA (~2-fold), RSV F (~4-fold) and TT (~4-fold) (Fig. 1C). This indicates that a de novo B cell response stemming from the naive compartment likely plays a role in SARS-CoV-2 infection.

In the classical memory compartment, B cells reactive against recurring seasonal infections (H1N1 HA and RSV F; 0.084% and 0.095%, respectively) or those elicited against a vaccine component (TT; 0.17%) were significantly more predominant than those against SARS-CoV-2 S

https://doi.org/10.1038/s41467-022-32232-0

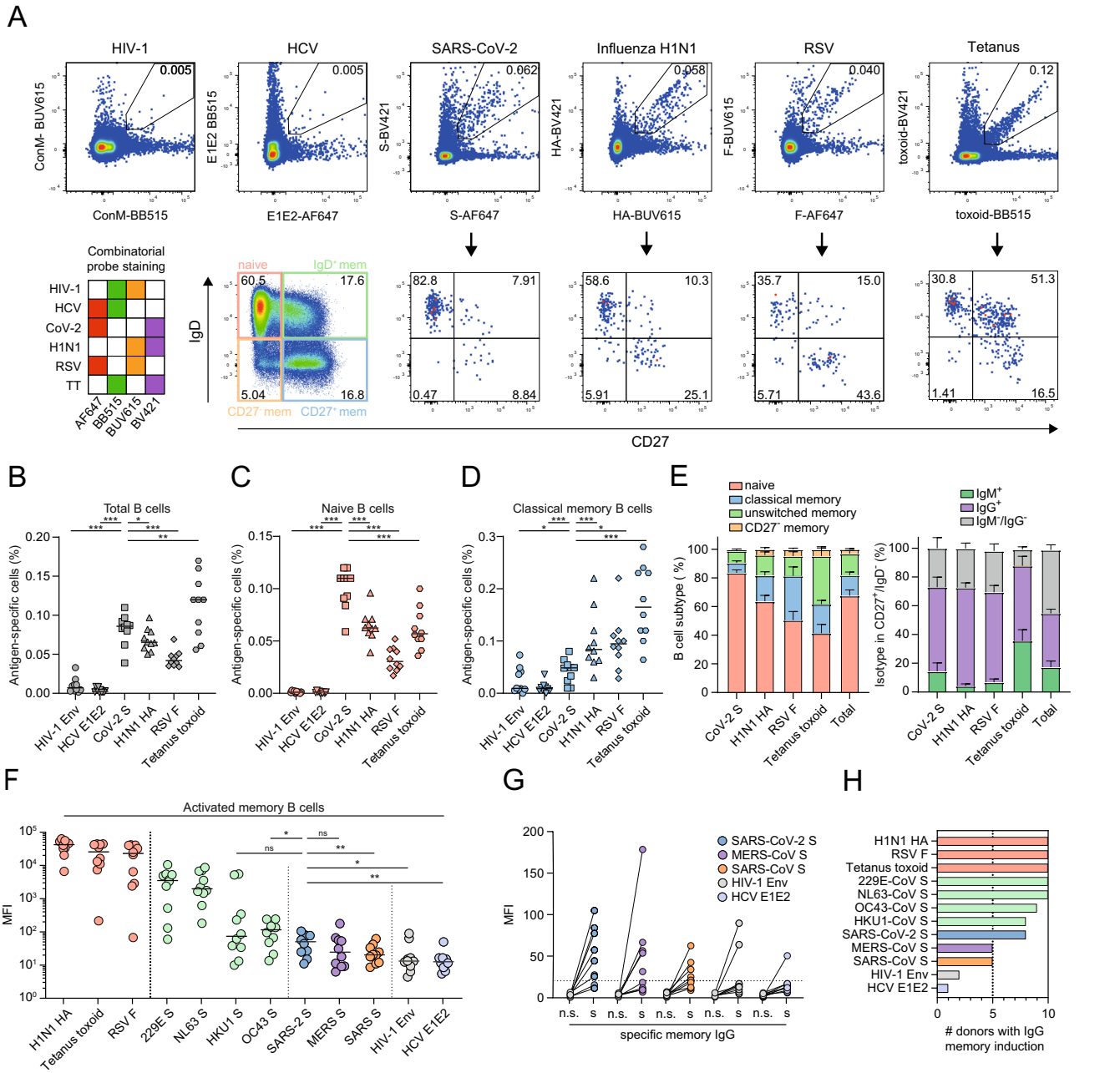

Fig. 1 | Phenotypic characterization of SARS-CoV-2 S-reactive B cells in unexposed individuals. A Combinatorial probe staining and gating strategy for the detection of multiple B cell specificities in a single PBMC sample. Top panel: From live B cells (gating strategy Fig. S1), antigen-reactive B cells are detected as double positive for the binding of the same antigen multimerized with two different fluorochromes. Bottom panel left, matrix depicting each combination of two fluorochrome-coded to the same antigen. Bottom panel left, B cell subsets were determined from antigen-reactive B cells according IgD and CD27 expression ($IgD^+/CD27^-$, naive; $IgD^+/CD27^+$, unswitched IgD^+ memory; $IgD^-/CD27^-$, $CD27^-$ memory; $IgD^-/CD27^+$, classical memory). The numbers inside the boxes represent the frequency (%) of cells in a gate. **B**–**D** Frequency of antigen-reactive B cells for ten HDs (Statistical differences were tested only in comparison to SARS-CoV-2 condition ($N = 10$, median) in total B cells (**B**), naive B cells ($IgD^+/CD27^-$, (**C**), or memory B cells ($IgD^-/CD27^+$, (**D**). Each dot represents one individual HD. The line represents the median frequency. **E** The phenotype of antigen-reactive B cells for each antigen (left panel) and the isotype as detected by flow cytometry in the classical memory B cell population ($IgD^-/CD27^+$, right panel). Bars represent mean ± SD, $n = 10$ HDs. **F** Median fluorescence intensity (MFI) of IgG detected in undiluted supernatants of cultured PBMCs for each antigen as measured by custom Luminex assay. Each dot represents one individual HD ($N = 10$, median). **G** MFI of IgG detected in supernatants of cultured PBMCs as measured by Luminex assay for five antigens. n.s. not-stimulated; s stimulated with resiquimod (TLR7/8 agonist), IL-2 and IL-21 to induce IgG secretion. The dotted line represents the cutoff of detectable IgG (three times above background). **H** Number of HDs that show IgG secretion reactive against all antigens measured when stimulated with cytokines to induce IgG secretion. statistical differences were tested using Wilcoxon test. n.s. not significant; $^*p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$; $^{****}p < 0.0001$.

(0.049%) (Fig. 1D), as expected. However, the SARS-CoV-2 S-reactive memory B cell frequency was significantly higher than that of HIV-1 Env and HCV E1E2 (Fig. 1D, 0.0088% and 0.0094%, respectively). We further confirmed the presence of pre-existing SARS-CoV-2 S-reactive humoral immunity by observing significantly higher plasma Ig levels against SARS-CoV-2 S compared to SARS-CoV S, MERS-CoV S and HIV-1 Env-reactive Igs (Fig. S2E–F). The majority of the SARS-CoV-2 S-binding classical memory B cells were IgG+. In contrast, memory B cells reactive with the vaccine antigen TT were frequently IgM +, while H1N1 HA and RSV F reactive B cells were almost exclusively IgM-, probably

reflecting recurrent antigenic stimulation (Fig. 1E, right panel; Fig. S2C). Similar results were obtained for the unswitched memory compartment, with SARS-CoV-2 having a higher S-reactive B cell frequency (0.042%) than that of HIV-1 Env (0.0083%) and HCV E1E2 (0.0018%), comparable to RSV F (0.034%), but lower than to H1N1 HA (0.076%) and TT (0.26%) (Fig. S2B). Thus, a significant portion of B cells in unexposed individuals is reactive with SARS-CoV-2 S. These B cells are found in multiple B cell compartments, suggesting that both de novo activation of naive B cells as well as reactivation of memory B cells elicited by a previous (HCoV) infection might play a role in the development of antibody responses after SARS-CoV-2 infection and/or vaccination. However, the majority of B cells that recognize SARS-CoV-2 are naive (82% of SARS-CoV-2 S-reactive B cells), whereas B cells against recurring infections and the TT vaccine predominantly have a memory phenotype, as one might expect (Fig. 1E, left panel; Fig. S2A, D).

To further evaluate the pre-existing SARS-CoV-2 reactive memory B cells, peripheral blood mononuclear cells (PBMCs) were stimulated with resiquimod (TLR7/8 agonist), IL-2 and IL-21 in order to promote the differentiation of memory B cells into plasmablasts and subsequently assess SARS-CoV-2 S-reactive Ig secretion as well as to a set of control antigens (Fig. 1F)[33]. Stimulation of memory B cells resulted in the secretion of high antibody levels against viral proteins from previously encountered common seasonal infections (H1N1 HA, RSV F, endemic HCoVs), while low antibody levels could be detected against SARS-CoV S, MERS-CoV S, HIV-1 Env and HCV E1E2 (Fig. 1F). Stimulated memory B cells produced a significantly higher antibody titer against SARS-CoV-2 S than SARS-CoV S, HIV-1 Env and HCV E1E2, in line with memory B cell frequencies as measured by flow cytometry (Fig. 1D). In eight out of ten supernatants from unexposed individuals, the secretion of reactive antibodies against SARS-CoV-2 S was similar to that against endemic HCoVs HKU1-CoV S and OC43-CoV S (Fig. 1F–H). Taken together, while the majority of SARS-CoV-2 S-reactive B cells belong to the naive compartment, the memory B cell compartment of the majority of HDs has the capacity to produce SARS-CoV-2 S-reactive antibodies despite not having previously encountered SARS-CoV-2. One to two donors appear to show some response to HIV-1 Env and HCV E1E2 (Fig. 1G, H), while they were confirmed HIV-1 and HCV-naive, an observation that requires further study but may relate to rare cross-reactivity of virus-reactive antibodies[34].

S-reactive B cells often carry a IGHV1-69/IGKV3-11-derived BCR

Next, we single-cell sorted SARS-CoV-2 S-reactive B cells from four unexposed donors (HD01-HD04). The majority of the sorted cells were $IgD^{+}/CD27^{-}$ and thus belonged to the naive B cell compartment, confirming our previous observations (Fig. 2A). We obtained 132 heavy chain (HC) with 101 paired light chain (LC) sequences from the SARS-CoV-2 S-double positive B cells (Table S2 and Source Data). Of the 132 HC sequences obtained, 101, 20 and 5 were derived from naive, unswitched and classical memory B cells, respectively, while we were unable to attribute six to a defined B cell subset. The BCRs of the SARS-CoV-2 S-reactive naive B cells were virtually identical to their assigned germlines, while BCRs from all classical memory and some unswitched memory S-reactive B cells showed higher mutation rates, consistent with maturation in response to previous antigen exposure (Fig. 2B). In three of the four donors, IGHV1-69 was the most dominant HC V gene used by SARS-CoV-2 S-reactive B cells, while it was the second most dominant in the fourth donor, accounting for 39% of all expressed V genes across the four donors (Fig. 2C, left panel), representing an enrichment of up to tenfold compared to the unselected germline repertoire, where IGHV1-69 is expressed in 4–7% of BCRs[35,36]. In the LC sequences a more moderate fourfold enrichment was observed for the IGVK3 family, which made up 36% of the LCs compared to 8% in an unselected naive repertoire (Fig. 2C, right panel)[35]. When studying HC/LC pairing, we observed that 13/101 (13%) of all recovered pairs from SARS-CoV-2 S-reactive B cells are IGHV1-69/IGKV3-11 (Fig. 2D). While informative, the numbers of BCRs obtained here were quite low. Therefore, the BCR repertoire of SARS-CoV-2 S-reactive B cells was studied more extensively in the next section.

The 101 HC/LC pairs were used to generate MAbs which were screened for binding to SARS-CoV-2 S by ELISA. 22 MAbs that showed significant binding in ELISA, defined as being three times above background, were selected for further evaluation. Of these 22 MAbs, 18 were from naive B cells, three from unswitched memory B cells and one from a classical memory cell. Three naive MAbs had the aforementioned IGHV1-69/IGKV3-11 pairing (MAbs 2F3, 3D4 and 5E10). The majority of the selected MAbs recognized cell surface-expressed full-length SARS-CoV-2 S (Fig. 2E and Fig. S3A), with varying efficiency. The binding was substantially lower compared to the control MAb COVA1-18 obtained from a SARS-CoV-2 convalescent individual[13], consistent with the fact that most of these MAbs were derived from naive B cells, and all were derived from B cells that had never encountered SARS-CoV-2. None of the 22 MAbs displayed any neutralizing activity in a pseudovirus neutralization assay (Fig. S3B). 5C4, 5E10 and 6B10 exhibited strong binding to untransfected cells (Fig. 2E and Fig. S3A), suggesting poly- and/or autoreactive properties. Therefore, the selected 22 MAbs were tested for polyreactivity against human insulin, bacterial lipopolysaccharide (LPS), bovine cardiolipin, calf thymus double-stranded DNA (dsDNA) and flagellin[37], while known polyreactive HIV-1 MAbs were included as controls. Nine MAbs showed some degree of polyreactivity, with the MAbs binding to untransfected HEK293T cells showing the highest degree of polyreactivity (Fig. 2F). In contrast, MAbs that showed a high degree of specificity in the cell-surface binding assay also showed little polyreactivity (e.g. 1C12). In addition, three of nine polyreactive MAbs were strongly cross-reactive against a disintegrin and metalloproteinase with a thrombospondin type 1 motif member 13 (ADAMTS13), an enzyme associated with COVID-19 thrombosis pathology (Fig. 2F)[38,39]. All three were isolated from different HDs and originated from the unswitched memory compartment. Altogether, these results indicate that pre-existing anti-SARS-CoV-2 S-reactive B cells may have a high degree of polyreactivity, expansion of these polyreactive MAbs might play a role in COVID-19 physiopathology[38,40–42].

To confirm that both polyreactive and non-polyreactive MAbs could be activated upon encounter with SARS-CoV-2 S as could occur during infection or vaccination, we engineered Ramos B cells to express a BCR that incorporates the variable regions of MAbs 1C12 and 3C9, both originally derived from naive S-reactive B cells. These MAbs are on opposite ends of the polyreactivity scale (1C12, low; 3C9, high) and show differential binding to S and thus represent diverse B cells that might encounter SARS-CoV-2 S in vivo (Fig. 2F). Both B cell lines recognized SARS-CoV-2 S (Fig. 2G) and were activated by I53-50 nanoparticles presenting 20 SARS-CoV-2 S and soluble SARS-CoV-2 S[13], irrespective of their polyreactive properties (Fig. 2H and Fig. S3C).

IGHV1-69/IGKV3-11 BCRs are enriched in S-reactive naive B cells

To elaborate on the BCR signatures of SARS-CoV-2 S-reactive B cells observed above, we sorted additional S-reactive B cells starting from 40×10^6 PBMCs of each donor (HD01-10) and performed single-cell sequencing using the 10X Genomics platform. We cross-referenced the variable regions with those of the ImMunoGeneTics (IMGT) database to obtain isotype, V(D)J gene usage and SHM, defined as mismatch to their assigned germline sequences[43]. Next, we assigned each B cell to a naive or memory phenotype using k-means clustering based on the expression of surface markers IgD and CD27, defined with feature barcode antibodies, as well as the number of mismatches to its assigned germline V(D)J genes (Fig. 3A). Naive B cells (IgD^{high}) were generally very close to their assigned germlines, whereas memory B cells ($CD27^{high}$) exhibit higher levels of SHM and are generally IgD^{low} (Fig. 3A, left panel). The majority of all SARS-CoV-2 S-reactive B cells in

https://doi.org/10.1038/s41467-022-32232-0

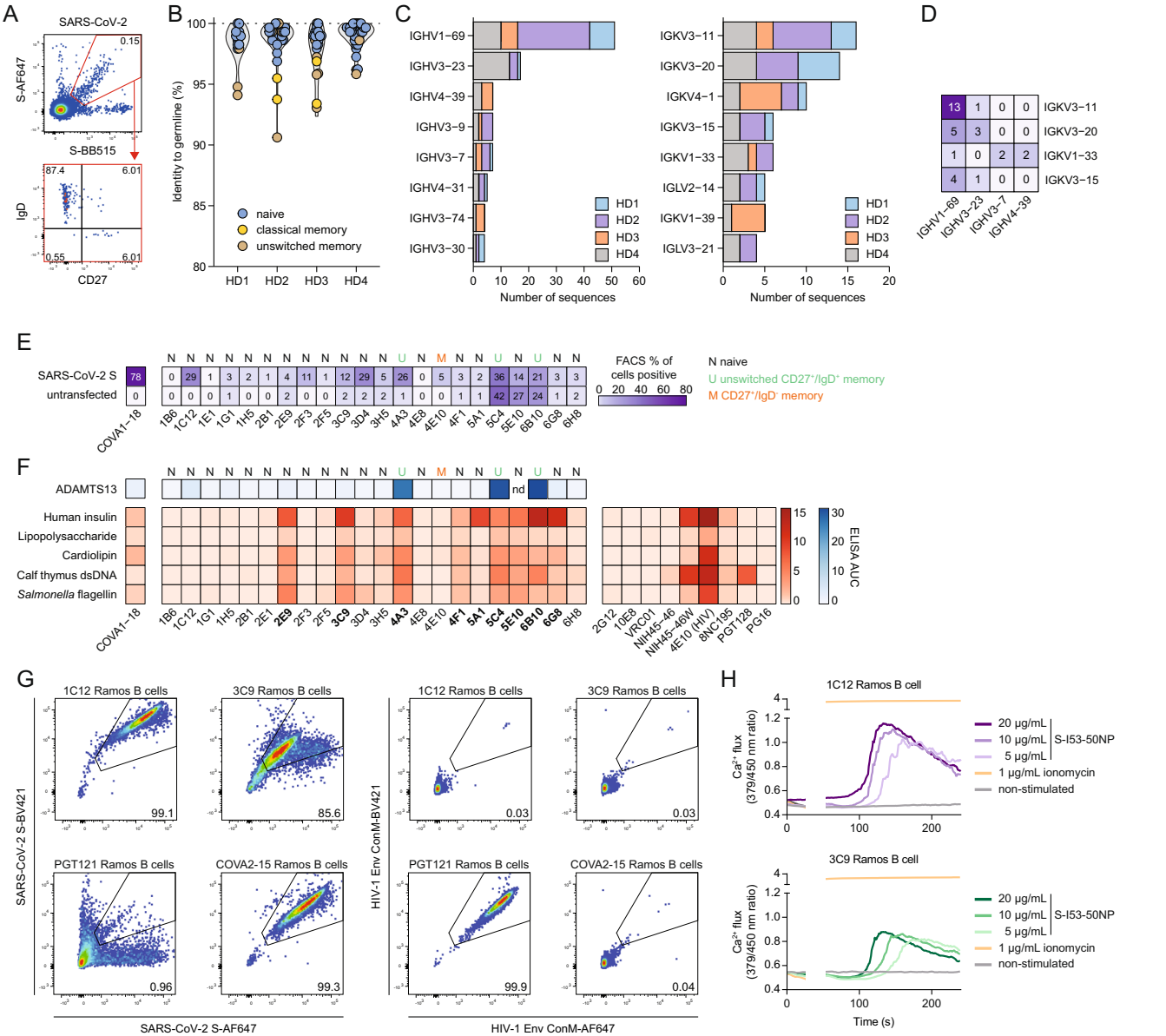

Fig. 2 | Genotypic and phenotypic characterization of SARS-CoV-2 S-reactive monoclonal antibodies from unexposed individuals. A Representative gating strategy of SARS-CoV-2 S-reactive B cells. Antigen-reactive B cells, were associated to naïve, unswitched IgD$^+$ memory, CD27- memory, and classical memory phenotype) (lower panel). **B** Dot plot overlaying a violin plot showing sequence identity (%) to IMGT-annotated germline heavy chain sequences for all isolated BCR heavy chains. Each color represents a B cell phenotype from four healthy donors. HD, healthy donor. **C** Bar plot showing the number of sequences recovered for each immunoglobulin heavy chain V (IGHV, left panel) gene and immunoglobulin kappa/lambda light chain (IGKV/IGLV, right panel) gene. Colored bars represent different HDs. **D** Matrix showing the number of pairs with a certain IGHV (x-axis) and IGKV (*y*-axis). The numbers inside the boxes represent the number of pairs recovered for each pairing. **E** Matrix showing flow cytometric binding assay to SARS-CoV-2 S-transfected or untransfected HEK293T cells for each selected MAb and control MAb COVA1-18. Numbers and colors in the boxes represent the percentage of cells showing binding to a particular MAb. The phenotype as determined in (**A**) is shown on top of the matrix. **F** Matrix depicting area under the curve (AUC) as determined by a polyreactive enzyme-linked immunosorbent assay (ELISA) for each of the antigens shown on the left. The letters on top represent B cell phenotypes as determined in (**A**). Polyreactive MAbs are indicated in bold text. nd not determined. **G** Antigen specificity of Ramos B cells designed to express a 1C12, 3C9, PGT121 or COVA2-15 BCR to SARS-CoV-2 S (left) or HIV-1 Env (right). The numbers inside the boxes represent the frequency (%) of cells in a gate. **H** Ramos B cell activation of 1C12 B cells (top panel) and 3C9 B cells (bottom panel) as measured by calcium (Ca^{2+}) flux assay. A baseline without antigen was established between 0 and 30 s, after which the measurement was interrupted to add the antigen to the B cells (30–50 s). Ionomycin was used at 1 µg/mL as positive control.

HDs belong to the naive compartment (1755 naive vs. 210 memory), consistent with the results from our earlier flow cytometry data (Fig. 1). Similar numbers of naive and memory B cells were obtained for each HD, with the provision that HD05 yielded low B cell numbers overall (Fig. 3A, right panel). Many B cells in the memory cluster had class-switched, while nearly all B cells in the naive compartment only expressed IgD and/or IgM BCRs. Most of the IgM$^+$ memory B cells are IgD^{high}/CD27^{high}, consistent with an unswitched memory phenotype (Fig. 3B).

We further studied the genetic characteristics of SARS-CoV-2 S-reactive naive B cells specifically, as only limited numbers of memory B cells were obtained. Heavy chain complementarity determining region 3 (CDRH3) length analysis revealed that the SARS-CoV-2 S-reactive naive B cell repertoire has on average longer CDRH3s than

https://doi.org/10.1038/s41467-022-32232-0

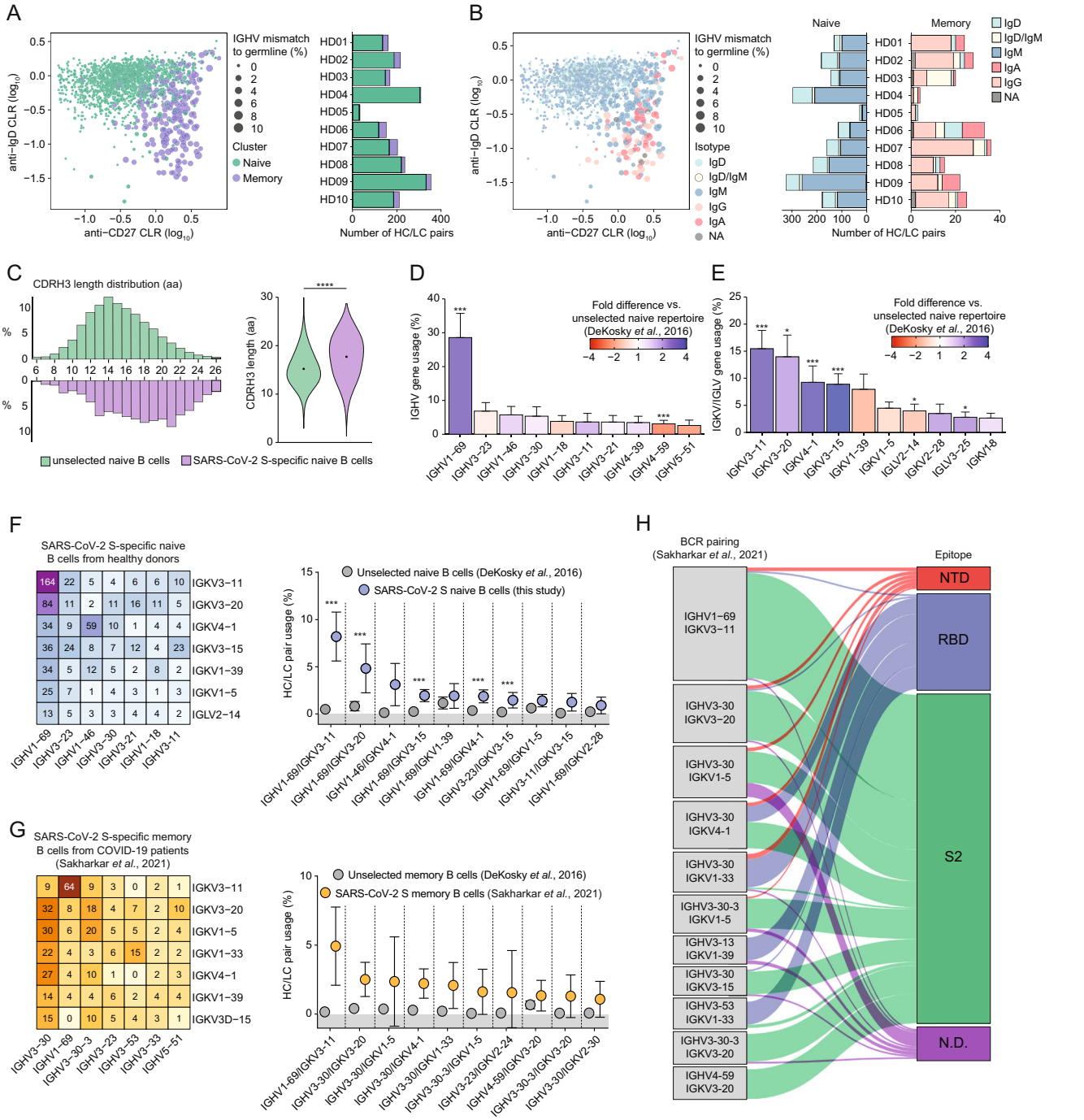

Fig. 3 | A public antibody class dominates the SARS-CoV-2 S-reactive B cell repertoire. A Jittered dot plot showing the phenotype of individual B cells and bar plot showing the number of recovered sequences per HD per cluster, from 10 HDs. Each dot represents an individual B cell, with the size of the dot corresponding to the percentage of mismatches to the IMGT-assigned germline sequence for IGHV genes. **B** Dot plot showing the isotype of individual B cells and bar plot showing the number of recovered sequences per HD per isotype, separated per cluster as determined in (**A**). NA not applicable; if no isotype could be retrieved. **C** Heavy chain complementarity-determining region 3 (CDRH3) lengths in amino acids (aa) for naive SARS-CoV-2 S-reactive B cells (purple) and an unselected naive B cell repertoire[35] (green). Significance was calculated with a Mann–Whitney *U* test. **D**, **E** Bar plots depicting the mean ± SEM IGHV (**D**) and IGKV/IGLV (**E**) gene usage (%) in naive SARS-CoV-2 S-reactive B cells from nine HDs. The colors represent the fold difference over an unselected naive B cell repertoire[35]. A non-parametric Kruskal–Wallis test was performed to compare the SARS-CoV-2 S-reactive naive repertoire to the unselected naive repertoire and is indicated on top of the bars. **F** Matrix showing the number of HC/LC pairs recovered from SARS-CoV-2 S-reactive naive B cells ($n = 9$, left panel) and the frequency of observed BCR pairs compared to an unselected naive repertoire ($n = 3$, right panel, mean ± SD)[35]. A non-parametric Kruskal-Wallis test was performed to compare the SARS-CoV-2 S-reactive naive repertoire to the unselected naive repertoire. **G** As in (**F**), but for SARS-CoV-2 S-reactive memory B cells from COVID-19 patients[20] ($n = 8$, left panel) and the frequency of observed BCR pairs compared to an unselected memory repertoire ($n = 2$, right panel, mean ± SD)[35]. **H** Sankey diagram showing the most frequent BCR pairs in SARS-CoV-2 S-reactive memory B cells from COVID-19 patients and their epitopes from ref. 20. N.D. not determined; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$).

an unselected naive repertoire (Fig. 3C, Mann–Whitney *U* test, 17.7 vs 15.2 amino acids on average, respectively). This result is in line with previous findings on SARS-CoV-2 S-reactive memory B cells after infection[13,36,44].

In accordance with the gene usage of isolated MAbs described above (Fig. 2C), S-reactive naive B cells showed a highly skewed HC gene usage in all HDs, with close to 30% of all sequenced HC genes belonging to the IGHV1-69 family (Fig. 3D and Fig. S4A, B), which is a significant enrichment compared to an unselected naive repertoire (3-fold, $p = 0.009$). IGHV1-46 and IGHV3-11 were also enriched, yet not significantly so ($p > 0.05$). In contrast, IGHV4-59 usage was significantly underrepresented (2.5-fold) compared to an unselected naive repertoire (Fig. 3D, $p = 0.009$). Four kappa light chain genes (i.e., IGKV3-11, IGKV3-20, IGKV4-1, IGKV3-15) were preferentially used compared to an unselected repertoire with IGKV3-11 representing over 15% of the BCR LC repertoire (Fig. 3E and Fig. S4A, B), consistent with our earlier observations (Fig. 2C). Indeed, usage of all three members of the IGKV3 Ig subfamily was significantly increased by at least 2-fold, as was IGKV4-1 (Fig. 3E, $p = 0.009$ for IGKV3-11, IGKV4-1 and IGKV3-15; $p = 0.03$ for IGKV3-15). SARS-CoV-2 S-reactive memory B cells of HDs do not show preferential gene usage compared to the unselected memory B cell repertoire, likely in part due to the low number of memory cells recovered (Fig. S4A–C). In summary, we observed substantial and significant increases in the usage of certain Ig genes, most notably in IGHV1-69, which is commonly involved in the humoral response to viral infections[13,45].

Next, we analyzed whether specific HC/LC pairings were preferred in SARS-CoV-2 S-reactive naive B cells in unexposed individuals. Indeed, we observed significant enrichment of multiple HC/LC pairings, in particular IGHV1-69/IGKV3-11, which was found in 164 out of 1755 total naive HC/LC pairs in S-reactive B cells (9.3%), corresponding to 8.5% on average for the nine donors included (Fig. 3F). These results match our earlier observations (Fig. 2D). Of note, the average CDRH3 length of IGHV1-69/IGKV3-11 paired BCRs was somewhat shorter than other S-reactive naive B cells (15.0 vs 18.0 amino acids). Overall, IGHV1-69 was used in seven out of the ten highest-frequency HC/LC pairings (IGHV1-69 paired with IGKV3-11, IGKV3-20, IGKV3-15, IGKV1-39, IGKV4-1, IGKV1-5, IGKV2-28) and four of these pairings were significantly overrepresented in SARS-CoV-2 S-reactive naive B cells compared to unselected naive B cells (IGHV1-69 paired with IGKV3-11, IGKV3-20, IGKV3-15, IGKV4-1, $p = 0.009$, Fig. 3F).

IGHV1-69/IGKV3-11 BCRs are enriched following SARS-CoV-2 infection

To contextualize these enriched naive clonotypes and verify whether they were relevant to the response following SARS-CoV-2 infection, we analyzed a recently published data set containing 1213 paired BCR sequences from SARS-CoV-2 S-reactive memory B cells isolated 1–5 months after SARS-CoV-2 infection (Fig. 3G)[20]. Strikingly, IGHV1-69/IGKV3-11 was the most frequent pair, representing 5.6% of the SARS-CoV-2 S-reactive memory B cell pool and showing ~37-fold enrichment compared to the unselected memory repertoire. Most of the other pairings in SARS-CoV-2 S-reactive memory B cells involved IGHV3-30 (Fig. 3G). In addition, IGHV1-69/IGKV3-11 pairs from these memory B cells had a similar CDRH3 length as the ones found in naive individuals (15.3 vs 15.0 amino acids), with 36/64 having a length of either 15 or 16 amino acids. Together, these observations highlight that IGHV1-69/IGKV3-11 is not only the predominant pair in the naive SARS-CoV-2 S-reactive B cell repertoire, but also expands and persists after SARS-CoV-2 infection.

Sakharkar and colleagues[20] determined the epitope of these SARS-CoV-2 S-reactive memory B cells, which allowed us to interrogate epitope specificities of certain enriched IGHV/IGKV pairings. From this dataset, most of the expressed MAbs targeted the S2 domain, while the RBD is targeted by the well-documented involvement of IGHV3-53 (Fig. 3H)[44,46]. The enriched IGHV1-69/IGKV3-11 pairing in SARS-CoV-2 S-reactive memory B cells is almost exclusively targeting the S2 domain, with only 4/64 (6%) targeted other epitopes on the S glycoprotein, and accounted for 9.3% of all S2-directed MAbs (Fig. 3H). These results suggest that a large pool of restricted IGHV1-69/IGKV3-11 naive B cells recognizes the S2 domain of SARS-CoV-2 S and that this B cell population is subject to expansion upon infection with SARS-CoV-2.

IGHV1-69/IGKV3-11 B cells target an apical S2 epitope

To characterize the epitopes of the IGHV1-69/IGKV3-11 clonotype, we studied six non-clonal IGHV1-69/IGKV3-11 MAbs isolated from two COVID-19 patients in an earlier study[13], which all showed high identity to germline V genes and bear an average-length CDRH3 (Fig. 4A). Out of the six IGHV1-69/IGKV3-11 MAbs only COVA2-17 had the ability to neutralize SARS-CoV-2 pseudovirus. COVA2-17 stood out in this selection as it used a different D gene compared to the other five (IGHD2-15 versus IGHD3-22) and had a shorter CDRH3 loop (13 amino acids versus 15 or 16) (Fig. 4A). Previous studies revealed that COVA2-17 recognized the NTD, while the epitopes of the other five MAbs were not yet identified[47]. ELISA experiments revealed that the other MAbs bound to SARS-CoV-2 S, and all but COVA1-04 bound soluble S2 (Fig. S5A). Our observation that four out of six IGHV1-69/IGKV3-11 MAbs recognized S2, one recognized NTD, and one unresolved, is in accordance with the findings from ref. 20 (Fig. 3H). Moreover, the observation that 25/164 IGHV1-69/IGKV3-11 BCRs in the naive repertoire bear a 15 or 16 amino acid CDRH3 illustrates these MAbs could have originated from an IGHV1-69/IGKV3-11 naive B cell.

Three of four S2-targeting MAbs (COVA1-07, COVA2-14 and COVA2-18) were then complexed as whole IgG molecule with soluble S2 and imaged by single-particle negative-stain electron microscopy (NS-EM, Fig. 4B). We also attempted to complex S2 with Fabs but these efforts were unsuccessful, possibly because of dependency on avidity for binding, as has been previously observed with other SARS-CoV-2 MAbs[48,49].The low-resolution 3D reconstructions revealed an unusual topology in which the two Fab arms bound individual protomers of S2 (Fig. 4B and Fig. S5B). Due to the flexibility of soluble S2, no density was obtained for the third protomer of the trimeric S2 construct. Strikingly, all three IGHV1-69/IGKV3-11 MAbs, derived from different clonal lineages and from different donors engaged the soluble S2 in a highly similar way. The epitope is buried on the prefusion, "three RBD down" S trimer, which is consistent with the inability of these MAbs to neutralize SARS-CoV-2 (Fig. 4C).

We assessed whether the IGHV1-69/IGKV3-11-derived MAbs COVA1-07, COVA2-14, COVA2-18 and COVA2-32 recognized full-length S expressed on the surface of cells. All four S2 MAbs engaged full-length S and the binding was similar to that of the neutralizing RBD-specific COVA1-16 and COVA1-18 MAbs (Fig. S6A). These data are in line with observations that wild-type full-length S on viral/cell membranes is unstable and deteriorates into open conformations and/or S2 stumps that are in the post-fusion conformation[50]. We also evaluated binding to full-length, cell surface-expressed S of globally circulating SARS-CoV-2 variants-of-concern (VOCs) alpha, beta, gamma and delta (B.1.1.7, B.1.351, P.1 and B.1.617.2, respectively) as well as endemic HCoVs (NL63, OC43, HKU1, 229E), epidemic HCoVs (SARS and MERS) and bat CoVs (SHC014, WIV-1) (Fig. S6A). The four S2 MAbs bound efficiently to the VOCs, SARS-CoV and bat CoV SHC014, but not to any of the other HCoVs tested, indicating that the epitope targeted by these IGHV1-69/IGKV3-11-derived MAbs is conserved across sarbecoviruses, but not beyond the sarbecovirus subgenus (Fig. S6A). Since these S2 MAbs were able to engage cell-surface expressed S protein we investigated their ability to partake in antibody-dependent effector mechanisms in vitro. IGHV1-69/IGKV3-11 MAbs COVA2-14, COVA2-18 and COVA2-32 efficiently triggered antibody-dependent cellular trogocytosis by the monocytic cell line THP-1, as shown by the extraction

https://doi.org/10.1038/s41467-022-32232-0

A

MAb name	IGHV	IGHV identity to germline	IGHD	IGHJ	CDRH3	CDRH3 length (aa)	IGKV	IGKV identity to germline	IGKJ	CDRL3	CDRL3 length (aa)
COVA1-04	IGHV1-69*01	99.65	IGHD3-22*01	IGHJ4*02	CARIPGYDSSGYYRGDYW	16	IGKV3-11*01	98.21	IGKJ3*01	CQQRSNWPPRFTF	11
COVA1-07	IGHV1-69*01	98.26	IGHD3-22*01	IGHJ4*02	CARVGAYDSSGYSNDYW	15	IGKV3-11*01	97.85	IGKJ4*01	CQQRSNWPPRVTF	11
COVA2-14	IGHV1-69*05	99.31	IGHD3-22*01	IGHJ4*02	CARVRYYDSSGYYEDYW	15	IGKV3-11*01	99.28	IGKJ2*01	CQQRSNWPPMYTF	11
COVA2-17	IGHV1-69*01	98.26	IGHD2-15*01	IGHJ3*02/IGHJ4*02	CASFGDDSGDEGVRW	13	IGKV3-11*01	98.92	IGKJ2*01	CQQRSNWPPYTF	10
COVA2-18	IGHV1-69*01	98.96	IGHD3-22*01	IGHJ4*02	CARVYSYDSSGYYLEYW	15	IGKV3-11*01	100	IGKJ5*01	CQQRSNWPPSITF	11
COVA2-32	IGHV1-69*01	97.92	IGHD3-22*01	IGHJ4*02	CARTHSYDNSGQYFDYW	15	IGKV3-11*01	98.21	IGKJ4*01	CQQRSNWPPRLTF	11

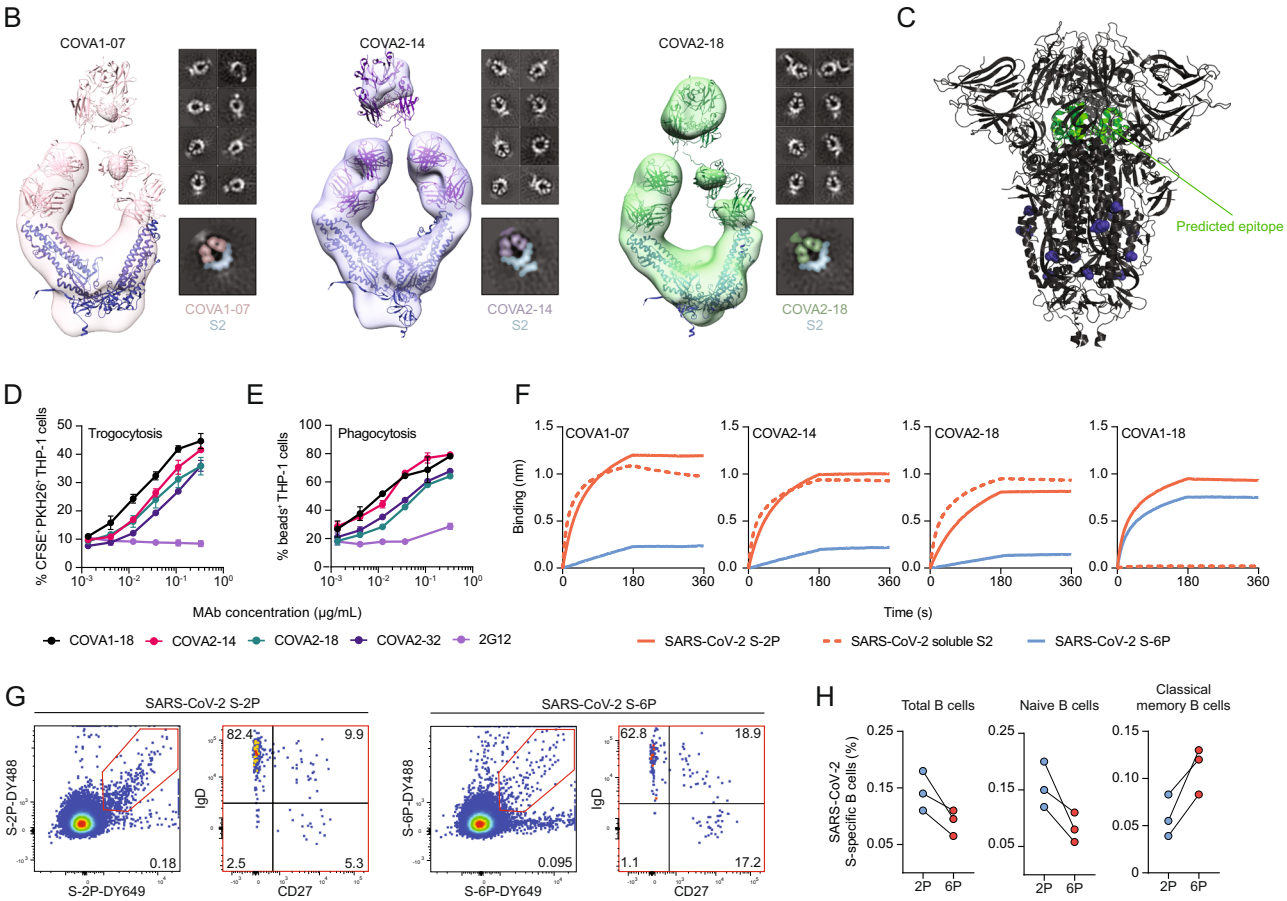

Fig. 4 | A public antibody class recognizes lesser stabilized epitopes on the S2 domain of SARS-CoV-2 S. A Characteristics of IGHV1-69/IGKV3-11 MAbs isolated in a previous study[13]. **B** 3D reconstructions of COVA1-07 (pink), COVA2-14 (purple), and COVA2-18 (green) complexed with soluble S2. The models for IgG (PDB 1HZH) and the S2 domain (generated from PDB 7JJI) are docked into each density map. Individual representative 2D class averages of COVA MAbs in complex with S2 are shown to the right of each 3D reconstruction. **C** The predicted epitope (green) of COVA1-07, COVA2-14 and COVA2-18 MAbs on S2 in the context of SARS-CoV-2 S with 3 RBDs in the down conformation (PDB 7JJI). The purple spheres indicate the location of the additional proline substitutions found in the S6-P construct (F817P, A892P, A899P, A942P). (**D**–**E**) Antibody-dependent cellular trogocytosis (**D**) and antibody-dependent cellular phagocytosis (**E**) assays for COVA MAbs and control MAb 2G12 ($n = 2$ replicates mean ± SD, representative curves are shown of at least 2 independent experiments). **F** Biolayer interferometry (BLI) sensorgrams of COVA1-07, COVA2-14, COVA2-18 and COVA1-18 binding to SARS-CoV-2 S-2P (orange), SARS-CoV-2 soluble S2 (dashed orange) and SARS-CoV-2 S-6P (blue). **G** Representative gating strategy of SARS-CoV-2 S2P- and S6P-reactive B cells. Each dot represents an individual B cell. From antigen-reactive B cells, the phenotype was determined ($IgD^+/CD27^-$, naive; $IgD^+/CD27^+$, unswitched IgD^+ memory; $IgD^-/CD27^-$, $CD27^-$ memory; $IgD^-/CD27^+$, classical memory). The numbers inside the plots represent the frequency (%) of cells in a gate. **H** Connected dot plots showing the frequency of total (left panel), naive (middle panel) and memory (right panel) of SARS-CoV-2 S-2P- or S-6P-reactive B cells (%) in three HDs, Not tested for statistical differences due to low sample number.

of membranous filaments from S-expressing cells (Fig. 4D and Fig. S6B). Similar results were obtained with the RBD-targeting neutralizing MAb COVA1-18, while the HIV-1 specific MAb 2G12 was unable to do so. COVA2-14, COVA2-18 and COVA2-32 also triggered antibody-dependent cellular phagocytosis as shown by the uptake of S-coated beads (Fig. 4E and Fig. S6C).

Some SARS-CoV-2 vaccines currently in use, including the AZD1222 vaccine from Oxford/AstraZeneca, encode for S-WT and are expected to expose aberrant forms of S that expose the IGHV1-69/IGKV3-11-targeted S2 epitope. Other vaccines, including mRNA-1273/Spikevax (Moderna), BNT162b2/Comirnaty (Pfizer/BioNTech), Ad26.CoV2.S (Janssen), and NVX-CoV2372 (Novavax) use a pre-fusion stabilized form of S that includes two proline substitutions (S-2P)[29]. However, several studies have shown that even S-2P is unstable and can deteriorate into aberrant conformations, albeit less efficiently so than S-WT[2,4]. Indeed, we found that COVA1-07, COVA2-14 and COVA2-17 bound efficiently to S-2P as measured by bio-layer interferometry (BLI). In fact, the binding kinetics to S-2P were similar to those of S2, revealing that the non-neutralizing S2 epitope is well exposed on S-2P. Next, we tested a more stabilized S trimer that incorporates four extra prolines compared to S-2P, amounting to six extra prolines compared to S-WT (S-6P or HexaPro S)[2]. S-6P is more resistant to heat/freeze cycles while its increased stabilization did not impact the conformation of S2. BLI experiments revealed a large decrease of the binding of

the S2 MAbs COVA1-07, COVA2-14 and COVA2-18 to S-6P compared to S-2P while RBD-targeting NAb COVA1-18 bound equally well to both proteins (Fig. 4F). Since the additional prolines are distal from the epitope targeted, we surmise that the loss of binding of these MAbs to S6-P is caused by effects of these prolines on S conformation (Fig. 4C). Together, these data suggest that IGHV1-69/IGKV3-11 MAbs might target open and/or aberrant conformations of S2 that are more frequently present on S-2P compared to S-6P.

To better understand the population of B cells that target open or aberrant conformations of SARS-CoV-2 S in unexposed individuals, we compared the reactivity of B cells of three HDs to S-2P and S-6P by flow cytometry (Fig. 4G). Overall, the frequency of S-2P-reactive B cells was 1.6-fold higher than that of S-6P-reactive cells (Fig. 4G, H left panel). Strikingly, the loss of reactivity toward S-6P is explained by the lower frequency of naive and unswitched memory B cells reactive to S-6P compared to S-2P (~2-fold) (Fig. 4H middle panel and Fig. S6D, E), whereas the opposite was true for the classical memory B cell population (Fig. 4H right panel, and Fig. S6D). This implies that naive B cells and unswitched memory B cells but not classical memory B cells often target open and/or aberrant S glycoprotein conformations in SARS-CoV-2-naive individuals. Moreover, we observed that MAbs derived from S-reactive B cells isolated from these unexposed individuals show a significantly reduced reactivity with S-6P compared to S-2P, confirming the propensity of naive and unswitched memory B cells to target open and/or aberrant conformations of S (Fig. S6F). Taken together, these results show that in unexposed individuals, a major part of the SARS-CoV-2 S-reactive B cell repertoire targets non-neutralizing epitopes on S2 exposed on open and/or aberrant forms of SARS-CoV-2 S which can trigger antibody responses upon natural infection and vaccination.

Discussion

In this study, we aimed to elevate our understanding of the origin and development of the humoral response against SARS-CoV-2 S following infection or vaccination. To this end, we characterized the S-reactive baseline B cell repertoire prior to SARS-CoV-2 exposure. We observed that an unexpectedly large proportion of B cells is reactive with SARS-CoV-2 S. Recognition in unexposed individuals is predominantly mediated by the naive B cell compartment, but substantial numbers of classical and unswitched SARS-CoV-2 S-reactive B cells are also present. The naive SARS-CoV-2 S-reactive B cells show a significant and meaningful enrichment in their BCR genes, with a public IGHV1-69/IGKV3-11 pairing present in 9.3% of S-reactive naive B cells. A recent study found no skewing in germline BCR pairings targeting the SARS-CoV-2 receptor binding motif (RBM) in unexposed individuals, consistent with observations that BCRs using IGHV1-69/IGKV3-11 pairing target (an) epitope(s) outside of the RBD[51].This enriched genotype was observed after vaccination and after infection, where it increased in frequency over time[19,20] as well as in a database attempting to collect all available anti-S MAb sequences (CoVAbDab), suggesting that the vast majority of patients developed this memory B cell subset[52]. The molecular characterization of IGHV1-69/IGKV3-11 derived MAbs, which do not have any neutralizing capacity, revealed that this class of MAbs targets an epitope on the apex of the S2 domain. This was recently corroborated by a study showing that alanine substitutions in the apical S2 epitope affect the binding of IGHV1-69/IGKV3-11 MAbs[45]. In addition, increased stabilization of the spike prevent the binding of IGHV1-69/IGKV3-11 MAbs suggesting that this epitope may be exposed on open and/or aberrant conformations of the S glycoprotein. Another recent study, using hydrogen deuterium exchange mass spectrometry, showed that prefusion stabilized S protein can adopt an alternative, more open conformation that could explain the exhibition of this epitope[53]. While it is currently unclear why specifically this IGHV1-69/IGKV3-11 pairing targets this epitope, IGHV1-69 encodes a hydrophobic CDRH2 which has been shown to be involved in recognition of viral glycoproteins[54].

We also identified a rare pre-existing memory B cell pool able to recognize SARS-CoV-2 S in unexposed individuals. Because of the scarcity of this population, the derived BCR sequence data set was too restricted to be linked with memory B cell clonotypes identified in COVID-19 patients. The detection of this pre-existing memory population is in line with previous studies showing a boosting of endemic HCoV S-reactive antibodies in sera following SARS-CoV-2 infection[31] and the isolation of HCoV cross-reactive antibodies from memory B cells showing a high degree of maturation[12,20,55]. Molecular characterization of the 210 paired BCR sequences from SARS-CoV-2 S-reactive memory B cells reported in our study could help identify additional cross-reactive epitopes. In addition, we report heterogeneity of this pre-existing memory B cell pool, with some donors showing only a limited memory B cell response toward SARS-CoV-2 S. A subset of SARS-CoV-2 S-reactive pre-existing memory B cells has been shown to cross-react with commensal microbiota antigens, which may at least partially explain the origin of this population[48]. Further studies should assess the role of these pre-existing memory B cells and their implications for disease outcome.

The immunodominance of the IGHV1-69/IGKV3-11 class of non-NAbs recognizing a conserved open and/or aberrant S conformation in SARS-CoV-2 infection raises questions about their role in COVID-19 outcome, and whether we should aim to promote or prevent their generation during vaccination. The open and/or aberrant conformations of S that are targeted by these non-NAbs, are present on cells expressing membrane-associated S protein and are therefore likely available on infected cells as well[56]. Consistent with this, these non-NAbs can mediate antibody-dependent effector mechanisms and thereby may play a role in the clearance of SARS-CoV-2. Indeed, recent studies highlighted the contribution of Fc effector function in the protection, clearance or limitation of symptoms in COVID-19, in which one of the studied Abs was a non-neutralizing targeting S2[11,57]. Moreover, it has been shown that Fc effector functions are a key determinant of protection after vaccination in a macaque model[58]. We report that the S2-targeting IGHV1-69/IGKV3-11 MAbs elicit effector functions and that they cross-react with multiple distinct S glycoproteins, including variants of concern and SARS-CoV-like viruses isolated from bats, but not endemic HCoVs. Overall, such antibody responses could help to mount a quicker adaptive immune response against recent variants and potential de novo SARS-CoV-like outbreaks. However, it should be noted that in vitro Fc effector function does not necessarily translate into in vivo protection, as was recently shown in a SARS-CoV-2 challenge experiment that used an IGHV1-69/IGKV3-11 MAb[45].

In contrast, exposure of immunodominant non-NAb epitopes might disfavor the induction of NAbs, which currently represent the most convincing correlate of protection for SARS-CoV-2 vaccines[59–61]. Thus, an argument can be made to engineer away non-NAb epitopes such as the one described here, in order to favor the induction of NAbs. Several studies have shown that stabilization of S with two proline substitutions (S-2P) promotes stronger NAb responses and/or protection following immunization in animal models[62–64]. As a result, S-2P was incorporated in the mRNA-1273/Spikevax (Moderna), BNT162b2/Comirnaty (Pfizer/BioNTech), Ad26.CoV2.S (Janssen), and NVX-CoV2372 (Novavax) vaccines. However, even S-2P is not entirely stable[2,4] and can expose non-NAb epitopes on S2. We show that a substantial fraction of naive B cells can still engage S-2P whereas the binding of non-NAbs of the IGHV1-69/IGKV3-11 class to S is abrogated when using the further stabilized S-6P. Ongoing vaccine trials with S-6P should shed light on whether preventing the induction of IGHV1-69/IGKV3-11 non-NAbs benefits the protective capacity of vaccines.

Finally, S2-targeting non-NAbs could participate in the physiopathology of COVID-19. Indeed, it has been shown that chronic (over) activation of the immune system can lead to the development of

autoimmune disease through the expansion of auto- and/or polyreactive clones[65]. In fact, this phenomenon has been observed in SARS-CoV and SARS-CoV-2 infection[55], with auto-antibodies to self-antigens such as clotting factor ADAMTS13 contributing to thrombotic thrombocytopenia[40].

In conclusion, we characterized the baseline SARS-CoV-2 S-reactive B cell repertoire. We found that a public antibody class using IGHV1-69 and IGKV3-11 very commonly recognizes a non-neutralizing epitope on SARS-CoV-2 S2 and that this antibody class is selected after SARS-CoV-2 infection and vaccination. These findings may help guide further improvements in SARS-CoV-2 vaccine design.

Methods

Sample collection

Our research complies with all relevant ethical regulations in the Netherlands. No ethical approval was necessary for the current study. Blood was collected from healthy blood donors ($n = 10$) by a Dutch blood bank (Sanquin, Amsterdam, The Netherlands) between March 2019 and February 2020, prior to the first official case of COVID-19 in The Netherlands. Plasma samples were collected after centrifugation and peripheral blood mononuclear cells (PBMCs) were isolated by density gradient centrifugation using Ficoll-Paque Plus. Isolated PBMCs were cryopreserved at −80 °C for future usage. Donors consent was waived due to anonymized donation of blood for blood donation, blood products and research by the donors to the Dutch national blood bank.

Protein design and purification

All soluble proteins, including SARS-CoV-2 S-2P and S-6P[2,13], S glycoproteins from variants of concern[66] and endemic HCoVs[31], influenza A hemagglutinin (H1N1pdm2009, A/Netherlands/602/2009, GenBank: CY039527.2[67], RSV prefusion stabilized F (DS-Cav1)[27], hepatitis C virus E1E2 and HIV-1 ConM Env[68] constructs with avi-tag and/or hexahistidine (his)-tag and/or strep-tag were expressed and purified as previously described[13]. After purification, avi-tagged proteins were biotinylated with a BirA500 biotin-ligase reaction kit according to the manufacturer's instruction (Avidity). Tetanus toxoid was purchased from Creative Biolabs (Vcar-Lsx003) and aspecifically biotinylated using EZ-Link Sulfo-NHS-LC-Biotinylation Kit (ThermoFisher) according to the manufacturer's instruction.

Probe preparation and staining

Biotinylated protein antigens were individually multimerized with fluorescently labeled streptavidin (BB515, BD Biosciences; BUV615, BD Biosciences; AF647, Biolegend; BV421, Biolegend) as described previously[13]. Briefly, biotinylated proteins and fluorescently labeled streptavidin were mixed at a 2:1 protein to fluorochrome molar ratio and incubated at 4 °C for 1 h. Unbound streptavidin conjugates were quenched with 10 uM biotin (Genecopoiea) for at least 10 min. Individual labeled proteins were then equimolarly mixed at a final concentration of 45.5 pM. $2–4 \times 10^7$ previously frozen PBMC samples were first depleted for T cells using CD3 selection kit II (StemCell) according to the manufacturer's instruction. Enriched B cells were then stained in Eppendorf tubes with 50–100 µL of antigen probe cocktail for 30 min at 4 °C, subsequently washed with FACS buffer (PBS supplemented with 1 mM EDTA and 2% fetal calf) and stained with the Live/DEAD dye together with MAbs coupled with fluorophores for FACS (Table S1) for an additional 30 min at 4 °C. Stained samples were washed twice and acquired on the BD LSRFortessa™ for cell analysis and ARIA-SORP-II 4 lasers for cell sorting. Analysis was performed using DIVA and FlowJo 10 software (BD Biosciences).

Combinatorial probe staining strategy

A combinatorial probe staining strategy was set up to perform simultaneous identification of multiple B cell specificities in a single sample, in the context of limited parameters/channel, which has been used previously with pMHC tetramers[69]. Conventionally, antigen-specific B cells are detected by the binding of two different fluorochrome-coded to the same protein. This method of double probe staining has been commonly used to reveal the fine specificity of B cells and avoid artefact of non-specific B cells binding to the fluorophore itself. Contrastingly, this combinatorial probe staining strategy uses all possible combinations of two fluorophores to increase the number of specificities that can be detected. The number of different B-cell specificities that can be detected equals N(N-1)/2, where N is the number of different fluorescent labels. In our study, we were able to detect 6 different antigen-specificities using 4 distinct fluorophores. Probes were labeled and used as previously described, in the following manner: SARS-CoV-2 S (AF647, BV421), H1N1 HA (BUV615, BV421), RSV F (AF647, BUV615), HCV E1E2 (AF647, BB515), HIV-1 ConM Env (BB515, BUV615) and Tetanus toxoid (BB515, BV421). Probe staining was used in combination with LIVE/DEAD dye (ThermoFisher) and labeled antibodies as previously described and acquired by FACS. For the analysis, the lymphocyte population was first gated based on the morphology (FSC-A/SSC-A) and doublets were removed. Next, dead cells and remaining $CD4^+$ cells to avoid artefact binding of HIV-1 probes were first excluded within a dump channel and live antigen-specific B cells were studied in the $CD19^+$population. To remove potential cross-reactive B cells to streptavidin, each probe combination was first gated on cells double negative for the two other channels (Fig. S1).

B cell stimulation

Polyclonal stimulation was performed according to the published method from ref. 33. Briefly, the freshly thawed PBMC were resuspended in RPMI 1640 (Gibco) supplemented with 10% fetal bovine serum, 100 U/mL penicillin/streptomycin, 2 mM L-Glutamine, 1 mM sodium pyruvate, 8 mM HEPES (RPMI10) (Life Technologies, Grand Island, NY, USA). R848 (MedChemExpress), recombinant human IL-2, and IL-21 (ImmunoTool) were used at a final concentration of 1 µg/mL, 100 units/mL and 50 ng/mL, respectively. PBMCs were cultured at 2×10^6 cells/mL, in 24-well Greiner CELLSTAR suspension culture plates and were incubated at 37 °C, 5% CO2 for 9 days. Harvested supernatants were then used for subsequent experiments.

Luminex assay

Luminex assays were performed as described previously[31]. In short, expressed glycoproteins were covalently coupled to Luminex Magplex beads using a ratio of 75 µg protein to 12.5 million beads for SARS-CoV-2 S and equimolarly for all other proteins. Next, 1:10.000-diluted plasma or undiluted stimulated B cell supernatant was added to the protein-coated beads overnight at 4 °C. The following day, beads were washed with TBS/0.05% Tween using a magnetic separator and resuspended in 1.3 µg/mL goat anti-human IgG-PE and incubated for 2 h at room temperature (Southern Biotech). After washing, the beads were resuspended in Magpix drive fluid (Luminex). The plates were read on a Magpix (Luminex). Specific binding was attributed to median fluorescence intensity (MFI) levels >3× above background.

B cell sorting for MAb isolation

SARS-CoV-2 S biotinylated proteins were individually multimerized with fluorescently labeled AF647 and BV421 streptavidin and used for staining as previously described[13]. Enriched B cells were stained with probes and antibodies conjugated to fluorophore and LIVE/DEAD dye (ThermoFisher) (Table S1). The lymphocyte population was first gated based on the morphology (FSC-A/SSC-A) and doublets were removed. Dead cells and non-B cells were first excluded within a dump channel ($CD3^-$/$CD14^-$/$CD16^-$). Live B cells ($CD19^+$) that were double positive for the SARS-CoV-2 S protein (AF647 and BV421) were single cell-sorted using index sorting into a 96-well plate containing a lysis buffer. The lysis buffer consisted of 20 U RNAse inhibitor (Invitrogen), first strand

SuperScript III buffer (Invitrogen), 1.25 µl of 0.1 M DTT (Invitrogen), in a total volume of 20 µL. The plates with the sorted single cells were stored at −80 °C for at least 1 h before performing reverse transcriptase (RT)-PCR.

B cell receptor variable region amplification, cloning and MAb expression

The BCR heavy chain and light chain variable regions were amplified and cloned into an IgG1 expression vector using reverse transcriptase polymerase chain reaction (RT-PCR) as described previously[13]. The V(D)J variable regions of antibodies were then cloned into human IgG1 expression vectors. The cloning was performed by mixing 2 µL of home-made Gibson enzyme mix containing 5X isothermal reaction buffer (0,5 g PEG-8000, 1 M Tris/HCl at pH 7.5, 1 M $MgCl_2$, 1 M DTT, 100 mM of each dNTP, 50 mM NAD and MQ), 1U/µL T5 exonuclease, 2U/µL Phusion polymerase and MQ, together with 1 uL of enzymatically digested expression vector (50 ng/uL) and 1 uL of PCR product. This mixture was then transformed into chemically competent *Escherichia coli*. After DNA purification, the sequences were verified by Sanger sequencing. For small scale expression, adherent HEK293T cells cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal calf serum and a mixture of penicillin/streptomycin (100 U/mL and 100 µg/mL, respectively) were transfected with purified DNA as described previously[13]. Supernatants of transfected cells were harvested 48 h post-transfection.

Preliminary ELISA screening of MAb supernatants

Preliminary enzyme-linked immunosorbent assays (ELISA) were conducted with clarified supernatants from HEK293T cells transfections. In order to maintain equimolar binding, SARS-CoV-2 RBD- and S-his-tagged proteins were diluted in casein to 1 µg/mL and 4.8 µg/mL, respectively. Antigens were immobilized for 2 h at RT on Ni-NTA 96-well plates. After three washes with TBS, binding of supernatants was allowed for 2 h at RT. Horseradish peroxidase (HRP)-labeled secondary antibody (goat anti-human IgG 1:3000) in casein was then incubated for 1 h at RT. After five washes with 1xTBS/0.05% Tween-20, 100 µl of freshly-made developing solution (1% 3,3',5,5'-tetramethylbenzidine, 0.01% hydrogen peroxide, 0.1 M sodium acetate and 0.1 M citric acid) were added to the plates. The reaction was terminated by adding 50 ul of 0.8 M sulfuric acid to each well and optical density at 450 nm (OD450) was measured.

Large-scale expression of monoclonal antibodies

MAbs that showed binding in the preliminary screening ELISA assays were selected and produced at a larger scale. Briefly, 250 mL suspensions of HEK293F cells were maintained in FreeStyle medium and transfected with 19.5 µg of the two HC and LC plasmids in a 1:1 ratio, together with 117 µL of PEImax (1 mg/ml). The produced MAbs were then harvested and purified after 5 days. For antibodies purification, cell suspensions were centrifuged for 30 min at 4000 rpm, and then filtered using 0.22 µm pore size SteriTop filters (Millipore), followed by a run of the supernatants over a 1 ml protein G beads column. 18 ml of elution buffer (0.1 M glycine pH 2.5) were used to elute the antibodies into 2 ml of neutralization buffer (1 M TRIS pH 8.7). By using 100 kDa VivaSpin20 columns, the purified antibodies were then concentrated and buffer exchanged to PBS. Concentration was measured on a NanoDrop 2000 (ThermoFisher).

Binding assays to cell surface expressed CoV-S by Flow cytometry

HEK293T cells were transfected with full length S plasmid DNA (SARS-CoV-2 WT and variants, other epidemic and endemic CoVs, and Bat-CoVs) using Lipofectamine2000 (Invitrogen). Briefly, 0.5×10^6 cells/well were plated in a 6-well plate. After 24 h, 4 µg DNA and 10 µL lipofectamine were mixed, incubated and added to each well. After 48 h, cells were harvested and pooled, and 5×10^4 cells were incubated in RPMI with 50 µg/mL of purified SARS-CoV-2 S-reactive MAbs for 1 h at RT. Cells were subsequently washed twice with PBS and stained for 30 min on ice and in the dark, in 50 µL of FACS buffer containing 1:1000 diluted PE- conjugated goat anti-human IgG (Biolegend). Cells were then washed twice with FACS buffer, fixed with 2% PFA and subsequently analyzed on BD LSRFortessa™.

Polyreactivity ELISA

Polyreactivity ELISAs were performed as described elsewhere[37]. High-binding half-area 96-well plates (Costar) were coated with 2 µg/mL *Salmonella enterica* serovar *Typhimurium* flagellin (Invitrogen), 10 µg/mL calf thymus double-stranded DNA (ThermoFisher), 5 µg/mL human insulin (Sigma-Aldrich) and 10 µg/mL *Escherichia coli* lipopolysaccharide (Sigma-Aldrich) in PBS and stored at RT overnight. Separate plates were coated with 10 µg/mL bovine cardiolipin (Sigma-Aldrich) in 99% ethanol and allowed to air-dry overnight. The following day, plates were washed with demineralized water and blocked with PBS/0.05% Tween/1 mM ethylenediaminetetraacetic acid (EDTA) for 1 h. MAbs were serially diluted five-fold (starting concentration 10 µg/mL) and binding to the plates was allowed for 2 h at RT. After washing, HRP-labeled secondary antibody (goat anti-human IgG 1:3000) in PBS was then incubated for 1 h at RT. Plates were washed and developed as described before. All experiments were performed in duplicate and are representative of two independent experiments.

Generation of Ramos B cells expressing custom B cell receptors

The B cell specific expression plasmid was constructed by exchanging the gl2-1261 gene of the pRRL EuB29 gl2-1261 IgGTM.BCR.GFP.WPRE plasmid[70] with the variable heavy and light chain genes of 1C12, 3C9, PGT121 and COVA2-15 using Gibson assembly (Integrated DNA Technologies). The production of lentivirus in HEK293T cells and the subsequent transduction was conducted as described elsewhere[71]. In short, lentiviruses were produced by co-transfecting the expression plasmid with pMDL, pVSV-g and pRSV-Rev into HEK293T cells using lipofectamine 2000 (Invitrogen). Two days post transfection, IgM-negative Ramos B cells cultured in RPMI10 were transduced with filtered (0.45 µm) and concentrated (100 kDa molecular weight cutoff, GE Healthcare) HEK293T supernatant. Seven days post-transduction, BCR-expressing B cells were sorted on IgG and green fluorescent protein (GFP) double-positivity using a FACS Aria-II SORP (BD Biosciences). B cells were then expanded and cultured indefinitely.

Ramos cell binding assay

Antigen specificity of the generated 1C12, 3C9, PGT121 and COVA2-15 Ramos B cells to SARS-CoV-2 S and HIV-1 ConM Env was detected as previously described using labeled probes and flow cytometry. Briefly, SARS-CoV-2 S and ConM v7 Env proteins multimerized with fluorescently labeled AF647 and BV421 streptavidin were used to stain 5×10^5cells together with live/DEAD dye, IgM-BV605, IgG PE-Cy7, and the antigen-probe cocktail (Table S1). Stained samples were subsequently washed twice with FACS buffer and acquired on the BD LSRFortessa™ for cell analysis. Analysis was performed using FlowJo v10.7. Ramos cells were first gated based on the morphology (FSC-A/SSC-A) and doublets were removed. Live cells were selected and subsequently gated on IgM-, GFP + and IgG + . Antigen-specific Ramos cells were double positive for the SARS-CoV-2 S (AF647 and BV421) and ConM Env (AF647 and BV421).

Calcium flux assay

B cell activation experiments of Ramos B cells were performed as previously described[72]. In short, 4×10^6 cells/mL in RPMI10 were loaded with 1.5 µM of the calcium indicator Indo-1 (Invitrogen) for 30 min at 37 °C, washed with Hank's Balance Salt Solution supplemented with 2 mM $CaCl_2$, followed by another incubation of 30 min at 37 °C.

Antigen-induced Ca^{2+} influx of B cells was monitored on a LSR Fortessa (BD Biosciences) by measuring the 379/450 nm emission ratio of Indo-1 fluorescence upon UV excitation. Following 30 s of baseline measurement, aliquots of 1×10^6 cells/mL were then stimulated for 210 s at RT with either 20 µg/mL, 10 µg/mL or 5 µg/mL of SARS-CoV-2 S or the equimolar amount presented on I53-50NPs. Ionomycin (Invitrogen) was added to a final concentration of 1 µg/µL to determine the maximum Indo-1-fluorescence. Kinetic analyses were performed using FlowJo v10.7.

Flow cytometry for 10X Genomics

Enriched B cells from 4×10^7 PBMCs of 10 donor were stained with Human TruStain FcX Fc Blocking Reagent (BioLegend, 422302) for 10 min at 4 °C. For each donor a mix of Abs linked to feature barcodes, containing one specific hashtag barcode together with CD27 and IgD, was centrifuged at 14,000 *g* at 4 °C for 10 min and supernatant harvested. Barcoded antibody mix, anti-CD19-AF700, Live/DEAD dye, and labeled SARS-CoV-2 S (AF647 and BV421) were added to the cells and stained for 30 min at 4 °C. Cells were then washed twice and resuspended in FACS buffer. Live B cells positives for the SARS-CoV-2 S protein (AF647 and BV421) were bulk sorted, for a total of 10,700 cells, and used for downstream 10X Genomics analysis.

10X Genomics library construction

Single-cell embedding in gel beads-in-emulsion (GEMs) and generation of barcoded cDNA was performed on a 10X Genomics Chromium Controller, following the 10X Genomics v1.1 single-cell V(D)J Next GEM with Feature Barcoding technology for cell surface protein chemistry. In short, single cell suspensions were kept at 4 °C and processed within 2 h upon completion of single cell sorting. Cells were then filtered to prevent clumping and counted to assess cell integrity and concentration. Out of 10,000 cells that were sorted, approximately 8000 cells were loaded and the resulting sequencing libraries were prepared following standard 10X Genomics protocols, generating a transcriptome, cell surface protein marker and a V(D)J library from each experiment. The cDNA libraries were paired-end sequenced on an Illumina NovaSeq S4 with a 2 × 150 bp Illumina kit.

Computational analyses for single cell sequencing data

Demultiplexing of raw base call (BCL) files, alignment, read filtering, barcode and UMI counting of the gene expression and feature barcode (10 hashtag oligos (HTOs) to discriminate individual donors and two antibody-derived tags (ADTs) for IgD/CD27 phenotyping) libraries was performed with the 10X Genomics Cell Ranger analysis pipeline using 'cellranger count'. Sequence assembly of the V(D)J library was performed using 'cellranger vdj'. During sequencing, Read 1 was assigned 28 base pairs and was used for identification of the Illumina library barcodes, cell barcodes and UMIs. Read 2 was used to map to the human reference transcriptome (GRCh38, version 3.0.0) and the CellRanger human V(D)J reference (GRCh38, version 5.0), respectively. Filtering of empty barcodes was done following standard CellRanger procedure. Resulting filtered feature-barcode matrices were imported in R (version 4.0.3) using the Seurat package (version 4.0.2)[73]. HTO and ADT UMI counts were normalized using centered log ratio (CLR) transformation. The Seurat function 'MULTIseqDemux' was used to demultiplex samples from the 10 HDs using the HTOs and identify singlets, doublets and negative cells[74].

RESEDA (https://bitbucket.org/barbera/reseda/) was used to further analyze the BCR repertoire. Sequences assembled by CellRanger were aligned to the IMGT gene database[75] with BWA (default settings)[76]. Variants were called with samtools mpileup[77] and VarScan[78] with minimum coverage set to 1 read to avoid missing mutations from antibody parts covered by a single sequence read. Consequently, any differences with respect to the IMGT sequences were considered as SHM. The CDR3 sequences were determined by translating the nucleotide sequences to peptide sequences and searching for conserved motifs in the V and J genes. The information from RESEDA was merged with the results obtained with Seurat.

Clustering and BCR repertoire analysis

In total, 1965 paired heavy/light chain BCRs were recovered, which could be assigned reliably to a single HD using the HTOs. K-means clustering with three determinants (IgD CLR/CD27 CLR/total number of mutations in HC/LC V region) was used to cluster B cells. The optimal number of clusters (k = 2) was determined using the average silhouette method and was used to assign a phenotype (memory/naive). IMGT-assigned CDRH3 lengths were compared to those of a large dataset of naive B cells available online[36]. For comparison of the expression of variable regions, paired HC/LC BCRs from naive and memory B cells were separately compared to naive and memory B cells in previously published data sets, respectively[20,35]. Analysis and handling of datasets was performed using R version 4.0.3.

Negative-stain electron microscopy

MAbs in molar excess were complexed with S2 for 30 min at RT. Immune complexes were diluted to ~20 µg/mL and deposited on glow-discharged, carbon-coated 400 mesh copper grids (Electron Microscopy Sciences). Sample was blotted from the grid with filter paper followed by two successive additions of 2% w/v uranyl formate stain with blotting. Grids were imaged on a Tecnai T20 (FEI) electron microscope with a CMOS 4k camera (TVIPS) at 200 kV, 62,000× nominal magnification, and 1.77 Å/pixel. Micrographs were collected using Leginon, particles were picked using Difference of Gaussians picker, and particles were cleaned through successive rounds of reference-free 2D classification in Relion 3.0[79–82]. Particles were also processed in CryoSPARC2 and reconstructed using Ab Initio[83]. Figures were made in UCSF Chimera and Adobe Photoshop[84].

Biolayer interferometry

Nickel–nitrilotriacetic acid (Ni-NTA) biosensors (ForteBio) were loaded with 20 µg/mL his-tagged S glycoproteins (S-2P and S-6P) or equimolar amounts of his-tagged soluble S2 in running buffer (PBS, 0.02% Tween 20, and 0.1% bovine serum albumin) for 300 s. Subsequently, the biosensors were washed in a well containing running buffer to remove excess protein, followed by transfer to a well containing MAb (50 µg/mL) in running buffer for 180 s to measure association. Next, the biosensors were moved to a well containing running buffer for 180 s to measure dissociation of the S–MAb complexes. All measurements were performed on an Octet K2 (ForteBio).

Antibody-dependent cellular trogocytosis

HEK293F cells (Invitrogen) at a density of 1×10^6cells/mL were transfected using SARS-CoV-2 S plasmid and PEImax (1 µg/µl) in a 3:1 ratio in OptiMEM. HEK293F cells were harvested 72 h after transfection and their plasma membrane was stained with 10 µM PKH26 (Sigma-Aldrich) dye in PBS, for 20 min (RT) with periodic mixing. Cells were washed twice with PBS and taken up in culture medium. THP-1 effector cells (ATCC) were stained intracellularly with 0.05 µM carboxyfluorescein succinimidyl ester (CFSE, ThermoFisher) in PBS and incubate 20 min (RT) with periodic mixing. Cells were washed twice with PBS and taken up in culture medium.). PKH26 stained HEK293F cells were opsonized for 30 min at 37 °C, with serial MAb dilutions. 2G12-IgG1, specific for HIV-1 gp120, was used as a negative control. After incubation, cells were washed and THP-1 cells were added to the HEK293F cells at a 2:1 effector:target ratio. Plates were centrifuged shortly to promote cell to cell contact and incubated 1 h at 37 °C. Afterwards, cells were washed and resuspended in PBS/2% FCS. Flow cytometry was used to measure the double positive, $PKH26^+CFSE^+$THP-1 cells. ADCT was calculated by the fraction of THP-1 cells that received membrane fragments from the HEK293F cells. To exclude measurement of antibody-independent

trogocytosis, cells were gated on stained HEK293F and THP-1 cells in the absence of antibodies.

Antibody-dependent cellular phagocytosis

Fluorescent neutravidin beads (Invitrogen) were incubated with biotinylated SARS-CoV-2 S-2P or RBD protein overnight at 4 °C. Beads were subsequently centrifuged shortly and washed twice with PBS/2% BSA to remove unbound antigen and block the remaining hydrophobic sites on the microspheres. The coated beads were resuspended in PBS/2% BSA and 0.1 µL of the original suspension was placed in every well of a V-bottom 96 well plate and incubated (2 h at 37 °C) with serial MAb dilutions. 2G12-IgG1, specific for HIV-1 gp120, was used as a negative control. After incubation, plates were washed and 5×10^4 THP-1 effector cells (ATCC) were added to each well in a final volume of 100 µL of RPMI10. Subsequently, plates were centrifuged shortly to promote beads to cell contact before incubation (5 h at 37 °C). After incubation, the cells were washed, resuspended in PBS/2% FCS and analyzed by flow cytometry.

Data analysis and visualization

Data visualization and statistical analyses were performed in GraphPad Prism software (version 8.3), while sequence handling, analysis and visualization were performed in RStudio (version 1.3.1093, R 4.0.3). A non-parametric Wilcoxon signed-rank test was performed to assess statistical differences for paired samples, while a Mann–Whitney *U* test was performed for unpaired samples. A Mann–Whitney *U* test to compare CDRH3 lengths was performed in RStudio 1.3.1093 (R 4.0.3). All statistical comparative tests were performed using a two-sided hypothesis test. All FACS data analyses were performed in FlowJo 10.8.0. Significance is denoted as $^{****}p < 0.0001$; $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; ns not significant.

Reporting summary

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

Data availability

Source data are provided as a Source Data File accompanying this manuscript. All data is readily available in the main text and supplementary materials. Supplementary information files related to single cell RNAseq analysis will be available at NCBI GEO under the accession number GSE196820. NS-EM reconstructions are deposited to the Electron Microscopy Data Bank under the accession number EMD26217-EMD26220. All reasonable requests for materials used in this study should be directed to and will be fulfilled under an MTA by Prof. Rogier W. Sanders (r.w.sanders@amsterdamumc.nl) and Dr. Marit J. van Gils (m.j.vangils@amsterdamumc.nl). Source data are provided with this paper.

Code availability

The code used in this study is available at the public GitHub repository https://github.com/caniels/hdpaper/.

References

1. Ritchie, H. et al. Coronavirus Pandemic (COVID-19). *Our World in Data* (2020).
2. Hsieh, C.-L. et al. Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **369**, 1501–1505 (2020).
3. Berger, I. & Schaffitzel, C. The SARS-CoV-2 spike protein: balancing stability and infectivity. *Cell Res.* **30**, 1059–1060 (2020).
4. Juraszek, J. et al. Stabilizing the closed SARS-CoV-2 spike trimer. *Nat. Commun.* **12**, 1–8 (2021).
5. Sadoff, J. et al. Safety and efficacy of single-dose Ad26.COV2.S vaccine against Covid-19. *N. Engl. J. Med.* https://doi.org/10.1056/NEJMoa2101544 (2021).
6. Baden, L. R. et al. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *N. Engl. J. Med.* **384**, 403–416 (2021).
7. Polack, F. P. et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *N. Engl. J. Med.* **383**, 2603–2615 (2020).
8. Henderson, R. et al. Controlling the SARS-CoV-2 spike glycoprotein conformation. *Nat. Struct. Mol. Biol.* **27**, 925–933 (2020).
9. McCallum, M., Walls, A. C., Bowen, J. E., Corti, D. & Veesler, D. Structure-guided covalent stabilization of coronavirus spike glycoprotein trimers in the closed conformation. *Nat. Struct. Mol. Biol.* **27**, 942–949 (2020).
10. Taylor, P. C. et al. Neutralizing monoclonal antibodies for treatment of COVID-19. *Nat. Rev. Immunol.* **21**, 382–393 (2021).
11. Shiakolas, A. R. et al. Cross-reactive coronavirus antibodies with diverse epitope specificities and Fc effector functions. *Cell Rep. Med.* **2**, 100313 (2021).
12. Amanat, F. et al. SARS-CoV-2 mRNA vaccination induces functionally diverse antibodies to NTD, RBD, and S2. *Cell* **184**, 3936–3948.e10 (2021).
13. Brouwer, P. J. M. et al. Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. *Science* **369**, 643–650 (2020).
14. Kreer, C. et al. Longitudinal isolation of potent near-germline SARS-CoV-2-neutralizing antibodies from COVID-19 patients. *Cell* **182**, 843–854.e12 (2020).
15. Zost, S. J. et al. Potently neutralizing and protective human antibodies against SARS-CoV-2. *Nature* **584**, 443–449 (2020).
16. Rogers, T. F. et al. Isolation of potent SARS-CoV-2 neutralizing antibodies and protection from disease in a small animal model. *Science* **369**, 956–963 (2020).
17. Robbiani, D. F. et al. Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* **584**, 437–442 (2020).
18. Song, G. et al. Cross-reactive serum and memory B-cell responses to spike protein in SARS-CoV-2 and endemic coronavirus infection. *Nat. Commun.* **12**, 1–10 (2021).
19. Dugan, H. L. et al. Profiling B cell immunodominance after SARS-CoV-2 infection reveals antibody evolution to non-neutralizing viral targets. *Immunity* **54**, 1290–1303.e7 (2021).
20. Sakharkar, M. et al. Prolonged evolution of the human B cell response to SARS-CoV-2 infection. *Sci. Immunol.* **6**, eabg6916 (2021).
21. Ullah, I. et al. Live imaging of SARS-CoV-2 infection in mice reveals neutralizing antibodies require Fc function for optimal efficacy. *bioRxiv* https://doi.org/10.1101/2021.03.22.436337 (2021).
22. Chan, C. E. Z. et al. The Fc-mediated effector functions of a potent SARS-CoV-2 neutralizing antibody, SC31, isolated from an early convalescent COVID-19 patient, are essential for the optimal therapeutic efficacy of the antibody. *PLoS One* **16**, e0253487 (2021).
23. Winkler, E. S. et al. Human neutralizing antibodies against SARS-CoV-2 require intact Fc effector functions for optimal therapeutic protection. *Cell* **184**, 1804–1820.e16 (2021).
24. Schäfer, A. et al. Antibody potency, effector function, and combinations in protection and therapy for SARS-CoV-2 infection in vivo. *J. Exp. Med.* **218**, e20201993 (2021).
25. Chakraborty, S. et al. Proinflammatory IgG Fc structures in patients with severe COVID-19. *Nat. Immunol.* **22**, 67–73 (2021).
26. Larsen, M. D. et al. Afucosylated IgG characterizes enveloped viral responses and correlates with COVID-19 severity. *Science* **371**, eabc8378 (2021).
27. McLellan, J. S. et al. Structure-based design of a fusion glycoprotein vaccine for respiratory syncytial virus. *Science* **342**, 592–598 (2013).
28. de Taeye, S. W. et al. Immunogenicity of stabilized HIV-1 envelope trimers with reduced exposure of non-neutralizing epitopes. *Cell* **163**, 1702–1715 (2015).

29. Sanders, R. W. & Moore, J. P. Virus vaccines: proteins prefer prolines. *Cell Host Microbe* **29**, 327–333 (2021).
30. Ng, K. W. et al. Preexisting and de novo humoral immunity to SARS-CoV-2 in humans. *Science* **370**, 1339–1343 (2020).
31. Grobben, M. et al. Cross-reactive antibodies after SARS-CoV-2 infection and vaccination. *bioRxiv* https://doi.org/10.1101/2021.05.26.21256092 (2021).
32. Tong, P. et al. Memory B cell repertoire for recognition of evolving SARS-CoV-2 spike. *Cell* **184**, 4969–4980.e15 (2021).
33. Franke, F., Kirchenbaum, G. A., Kuerten, S. & Lehmann, P. V. IL-21 in conjunction with anti-CD40 and IL-4 constitutes a potent polyclonal B cell stimulator for monitoring antigen-specific memory B cells. *Cells* **9**, 433 (2020).
34. Williams, W. B., Han, Q. & Haynes, B. F. Cross-reactivity of HIV vaccine responses and the microbiome. *Curr. Opin. HIV AIDS* **13**, 9–14 (2018).
35. DeKosky, B. J. et al. Large-scale sequence and structural comparisons of human naive and antigen-experienced antibody repertoires. *Proc. Natl Acad. Sci. USA* **113**, E2636–E2645 (2016).
36. Briney, B., Inderbitzin, A., Joyce, C. & Burton, D. R. Commonality despite exceptional diversity in the baseline human antibody repertoire. *Nature* **566**, 393–397 (2019).
37. Guthmiller, J. J. et al. Polyreactive broadly neutralizing B cells are selected to provide defense against pandemic threat influenza viruses. *Immunity* **53**, 1230–1244.e5 (2020).
38. Mancini, I. et al. The ADAMTS13-von Willebrand factor axis in COVID-19 patients. *J. Thromb. Haemost.* **19**, 513–521 (2021).
39. Doevelaar, A. A. N. et al. von willebrand factor multimer formation contributes to immunothrombosis in coronavirus disease 2019. *Crit. Care Med.* **49**, e512–e520 (2021).
40. Doevelaar, A. A. N. et al. Generation of inhibitory autoantibodies to ADAMTS13 in Coronavirus disease 2019. *bioRxiv* https://doi.org/10.1101/2021.03.18.21253869 (2021).
41. Ruhe, J. et al. Acquired thrombotic thrombocytopenic purpura after first vaccination dose of BNT162b2 mRNA COVID-19 vaccine. *Ann. Hematol.* https://doi.org/10.1007/s00277-021-04584-y (2021).
42. de Bruijn, S., Maes, M.-B., De Waele, L., Vanhoorelbeke, K. & Gadisseur, A. First report of a de novo iTTP episode associated with an mRNA-based anti-COVID-19 vaccination. *J. Thromb. Haemost.* **19**, 2014–2018 (2021).
43. Lefranc, M.-P. et al. IMGT®, the international ImMunoGeneTics information system® 25 years on. *Nucleic Acids Res.* **43**, D413–D422 (2015).
44. Barnes, C. O. et al. SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* **588**, 682–687 (2020).
45. Chen, E. C. et al. Convergent antibody responses to the SARS-CoV-2 spike protein in convalescent and vaccinated individuals. *Cell Rep.* **36**, 109604 (2021).
46. Wu, N. C. et al. An alternative binding mode of IGHV3-53 antibodies to the SARS-CoV-2 receptor binding domain. *Cell Rep.* **33**, 108274 (2020).
47. Rosa, A. et al. SARS-CoV-2 can recruit a heme metabolite to evade antibody immunity. *Sci. Adv.* **7**, eabg7607 (2021).
48. Liu, H. et al. Cross-neutralization of a SARS-CoV-2 antibody to a functionally conserved site is mediated by avidity. *Immunity* **53**, 1272–1280.e5 (2020).
49. Maisonnasse, P. et al. COVA1-18 neutralizing antibody protects against SARS-CoV-2 in three preclinical models. *Nat. Commun.* **12**, 6097 (2021).
50. Ke, Z. et al. Structures and distributions of SARS-CoV-2 spike proteins on intact virions. *Nature* **588**, 498–502 (2020).
51. Feldman, J. et al. Naive human B cells engage the receptor binding domain of SARS-CoV-2, variants of concern, and related sarbecoviruses. *Sci. Immunol.* **6**, eabl5842 (2021).
52. Raybould, M. I. J., Kovaltsuk, A., Marks, C. & Deane, C. M. CoV-AbDab: the coronavirus antibody database. *Bioinformatics* **37**, 734–735 (2020).
53. Costello, S. M. et al. The SARS-CoV-2 spike reversibly samples an open-trimer conformation exposing novel epitopes. *bioRxiv* 2021.07.11.451855 (2021) https://doi.org/10.1101/2021.07.11.451855.
54. Chen, F., Tzarum, N., Wilson, I. A. & Law, M. VH1-69 antiviral broadly neutralizing antibodies: genetics, structures, and relevance to rational vaccine design. *Curr. Opin. Virol.* **34**, 149–159 (2019).
55. Huang, K.-Y. A. et al. Breadth and function of antibody response to acute SARS-CoV-2 infection in humans. *PLoS Pathog.* **17**, e1009352 (2021).
56. Letarov, A. V., Babenko, V. V. & Kulikov, E. E. Free SARS-CoV-2 spike protein S1 particles may play a role in the pathogenesis of COVID-19 infection. *Biochemistry* **86**, 257–261 (2021).
57. Beaudoin-Bussières, G. et al. An anti-SARS-CoV-2 non-neutralizing antibody with Fc-effector function defines a new NTD epitope and delays neuroinvasion and death in K18-hACE2 mice. *bioRxiv* 2021.09.08.459408 (2021) https://doi.org/10.1101/2021.09.08.459408.
58. Gorman, M. J. et al. Fab and Fc contribute to maximal protection against SARS-CoV-2 following NVX-CoV2373 subunit vaccine with Matrix-M vaccination. *Cell Rep. Med.* **2**, 100405 (2021).
59. Corbett, K. S. et al. Immune correlates of protection by mRNA-1273 vaccine against SARS-CoV-2 in nonhuman primates. *Science* **373**, eabj0299 (2021).
60. Earle, K. A. et al. Evidence for antibody as a protective correlate for COVID-19 vaccines. *Vaccine* **39**, 4423–4428 (2021).
61. Feng, S. et al. Correlates of protection against symptomatic and asymptomatic SARS-CoV-2 infection. *Nat. Med.* https://doi.org/10.1038/s41591-021-01540-1 (2021).
62. Amanat, F. et al. Introduction of two prolines and removal of the polybasic cleavage site lead to higher efficacy of a recombinant spike-based SARS-CoV-2 vaccine in the mouse model. *MBio* **12**, e02648-20 (2021).
63. Yu, J. et al. DNA vaccine protection against SARS-CoV-2 in rhesus macaques. *Science* **369**, 806–811 (2020).
64. Klasse, P. J., Nixon, D. F. & Moore, J. P. Immunogenicity of clinically relevant SARS-CoV-2 vaccines in nonhuman primates and humans. *Sci Adv* **7**, eabe8065 (2021).
65. Wong, A. K. H. et al. Broad auto-reactive IgM responses are common in critically ill patients, including those with COVID-19. *Cell Rep. Med.* **2**, 100321 (2021).
66. Caniels, T. G. et al. Emerging SARS-CoV-2 variants of concern evade humoral immune responses from infection and vaccination. *Sci. Adv.* **7**, eabj5365 (2021).
67. Aartse, A. et al. Influenza a virus hemagglutinin trimer, head and stem proteins identify and quantify different hemagglutinin-specific B cell subsets in humans. *Vaccines (Basel)* **9**, 717 (2021).
68. Sliepen, K. et al. Structure and immunogenicity of a stabilized HIV-1 envelope trimer based on a group-M consensus sequence. *Nat. Commun.* **10**, 1–16 (2019).
69. Hadrup, S. R. et al. Parallel detection of antigen-specific T-cell responses by multidimensional encoding of MHC multimers. *Nat. Methods* **6**, 520–526 (2009).
70. McGuire, A. T., Glenn, J. A., Lippy, A. & Stamatatos, L. Diverse recombinant HIV-1 Envs fail to activate B cells expressing the germline B cell receptors of the broadly neutralizing anti-HIV-1 antibodies PG9 and 447-52D. *J. Virol.* **88**, 2645–2657 (2014).
71. ter Brake, O., Konstantinova, P., Ceylan, M. & Berkhout, B. Silencing of HIV-1 with RNA interference: a multiple shRNA approach. *Mol. Ther.* **14**, 883–892 (2006).
72. Brouwer, P. J. M. et al. Two-component spike nanoparticle vaccine protects macaques from SARS-CoV-2 infection. *Cell* **184**, 1188–1200.e19 (2021).

73. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).
74. McGinnis, C. S. et al. MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. *Nat. Methods* **16**, 619–626 (2019).
75. Giudicelli, V., Chaume, D. & Lefranc, M.-P. IMGT/GENE-DB: a comprehensive database for human and mouse immunoglobulin and T cell receptor genes. *Nucleic Acids Res.* **33**, D256–D261 (2005).
76. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
77. Li, H. et al. The sequence alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
78. Koboldt, D. C. et al. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics* **25**, 2283–2285 (2009).
79. Suloway, C. et al. Automated molecular microscopy: the new Leginon system. *J. Struct. Biol.* **151**, 41–60 (2005).
80. Voss, N. R., Yoshioka, C. K., Radermacher, M., Potter, C. S. & Carragher, B. DoG Picker and TiltPicker: software tools to facilitate particle selection in single particle electron microscopy. *J. Struct. Biol.* **166**, 205–213 (2009).
81. Lander, G. C. et al. Appion: an integrated, database-driven pipeline to facilitate EM image processing. *J. Struct. Biol.* **166**, 95–102 (2009).
82. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, e42166 (2018).
83. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
84. Pettersen, E. F. et al. UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

Acknowledgements

We thank Single Cell Discoveries (Utrecht, the Netherlands) for help with single-cell RNA sequencing. We thank Drs. Li Wu and Vineet N. Kewal Ramani from the NIH AIDS Reagent Program for kindly providing Ramos B cells and Dr. Andrew McGuire for kindly sharing the pRRL.EuB29 lentiviral vector that was used to transduce Ramos B cells. We thank Prof. S Marieke van Ham and Dr. George Elias (Sanquin, the Netherlands) as well as Dr. Juan J. Garcia Vallejo (Amsterdam UMC, the Netherlands) for their help in developing the combinatorial probe staining strategy. Characterization of the BCR repertoire using RESEDA was carried out on the Dutch national e-infrastructure with the support of SURF Cooperative. We also acknowledge all funders, Netherlands Organization for Scientific Research (NWO) Vici grant (RWS). Bill & Melinda Gates Foundation, Collaboration for AIDS Vaccine Discovery (CAVD) grant INV-002022 (RWS), Amsterdam UMC AMC Fellowship (MJvG), Bill & Melinda Gates Foundation, COVID-19 Wave 2 mAbs grant INV-024617 (MJvG), Fondation Dormeur, Vaduz (MJvG, RWS), Bill & Melinda Gates Foundation, grant INV-004923 (ABW), Netherlands Organization for Scientific Research (NWO) Veni grant (VI.Veni.192.114) (MC).

Author contributions

Conceptualization: M.C., T.G.C., R.W.S., M.J.v.G.. Funding acquisition: M.C., A.B.W., A.H.C.v.K., R.W.S., M.J.v.G. Investigation: M.C., T.G.C., M.d.G., J.H., D.G., G.K., B.D.C.v.S., A.J., A.I.S., M.G., P.J.M.B., K.v.d.S., Y.A., J.C.-P., J.L.S., W.O., I.B., J.A.B., M.P., T.P.L.B., J.L.T., J.C., I.C.M., S.W.d.T. Methodology: M.C., T.G.C., M.d.G., J.H., B.D.C.v.S., A.J., A.H.C.v.K., P.D.M., R.W.S., M.J.v.G. Project administration: M.C., T.G.C., M.J.v.G. Resources: A.A., M.B., G.J.d.B. Supervision: M.C., T.G.C., A.B.W., K.S., A.H.C.v.K., P.D.M., R.W.S., M.J.v.G.. Writing – original draft: M.C., T.G.C., R.W.S., M.J.v.G.. Writing – review & editing: all authors.

Competing interests

T.G.C., P.J.M.B., K.v.d.S., G.J.d.B., R.W.S. and M.J.v.G. are inventors on a patent application related to this work filed by Amsterdam UMC (PCT/EP2021/062558, filed on 11 May 2021). The other authors declare that they have no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-32232-0.

Correspondence and requests for materials should be addressed to Rogier W. Sanders or Marit J.van Gils.

Peer review information *Nature Communications* thanks Other anonymous Reviewer(s) to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022