

SARS-CoV-2–specific antibody rearrangements in prepandemic immune repertoires of risk cohorts and patients with COVID-19

Lisa Paschold, ... , Christoph Schultheiß, Mascha Binder

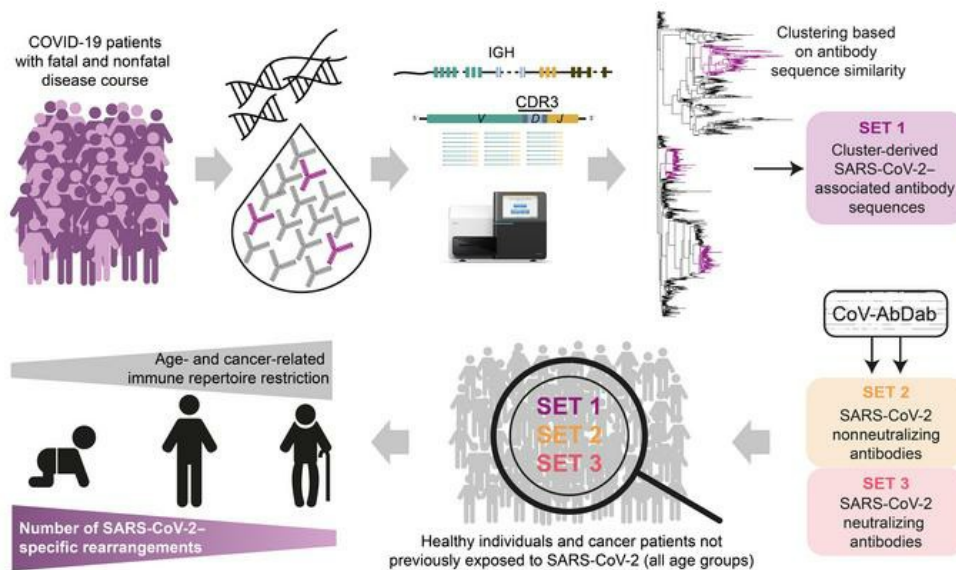
J Clin Invest. 2021;131(1):e142966. <https://doi.org/10.1172/JCI142966>.

Concise Communication

COVID-19

Immunology

Graphical abstract



Find the latest version:

<https://jci.me/142966/pdf>



SARS-CoV-2-specific antibody rearrangements in pre-pandemic immune repertoires of risk cohorts and patients with COVID-19

Lisa Paschold,¹ Donjete Simnica,¹ Edith Willscher,¹ Maria J.G.T. Vehreschild,² Jochen Dutzmann,³ Daniel G. Sedding,³ Christoph Schultheiß,¹ and Mascha Binder¹

¹Department of Internal Medicine IV, Oncology/Hematology, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany. ²Department of Internal Medicine, Infectious Diseases, University Hospital Frankfurt, Goethe University Frankfurt, Frankfurt am Main, Germany. ³Mid-German Heart Center, Department of Cardiology and Intensive Care Medicine, University Hospital, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany.

A considerable fraction of B cells recognize severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) with germline-encoded elements of their B cell receptor, resulting in the production of neutralizing and nonneutralizing antibodies. We found that antibody sequences from different discovery cohorts shared biochemical properties and could be retrieved across validation cohorts, confirming the stereotyped character of this naive response in coronavirus disease 2019 (COVID-19). While neutralizing antibody sequences were found independently of disease severity, in line with serological data, individual nonneutralizing antibody sequences were associated with fatal clinical courses, suggesting detrimental effects of these antibodies. We mined 200 immune repertoires from healthy individuals and 500 repertoires from patients with blood or solid cancers – all acquired prior to the pandemic – for SARS-CoV-2 antibody sequences. While the largely unmutated B cell rearrangements occurred in a substantial fraction of immune repertoires from young and healthy individuals, these sequences were less likely to be found in individuals over 60 years of age and in those with cancer. This reflects B cell repertoire restriction in aging and cancer, and may to a certain extent explain the different clinical courses of COVID-19 observed in these risk groups. Future studies will have to address if this stereotyped B cell response to SARS-CoV-2 emerging from unmutated antibody rearrangements will create long-lived memory.

Introduction

Severe acute respiratory syndrome coronavirus 2-specific (SARS-CoV-2-specific) B cell immunity and antibody responses are being actively investigated, since early neutralization of the virus may play a decisive role in the clinical course of coronavirus disease 2019 (COVID-19) and further insights may be crucial for the development of vaccines and potential antibody therapeutics (1–3). Deep sequencing analyses have shown converging responses with preferential immunoglobulin heavy chain variable-jointing gene (IGHV-J) usage and not more than minimal somatic hypermutation despite postswitch isotypes (4–7). This peculiar pattern of a naive B cell response was found across different cohorts and therefore deemed an overarching immunological principle in COVID-19. It showed that even B cell receptor rearrangements or antibodies with unmutated variable regions could specifically recognize SARS-CoV-2.

The naive character of the B cell response indicates that the presence of certain germline recombinations could be advanta-

geous in response to SARS-CoV-2, thus allowing rapid antibody formation for efficient viral clearance and thereby preventing severe hyperinflammatory disease courses. In the work presented here, we analyzed the representation of SARS-CoV-2-specific antibody rearrangements in 68 B cell repertoires from patients with different disease courses of COVID-19 and 700 repertoires from individuals not previously exposed to the virus.

Results and Discussion

We compiled 3 sets of SARS-CoV-2-specific rearrangements to mine B cell repertoires from patients with COVID-19 and individuals not previously exposed to the virus. Set 1 included converging sequence clusters derived from antibody-positive patients with active COVID-19 infection (previously published cohort, ref. 5, <https://gateway-covid19.ireceptor.org/home>; study: IR-Binder-000001). Set 2 contained sequences from nonneutralizing antibodies ($n = 53$) and set 3 contained sequences from neutralizing SARS-CoV-2-specific antibodies from 10 independent studies ($n = 74$; set 2 and set 3 sequences were both retrieved from CoV-AbDab on July 3, 2020, ref. 8). Since sequences from set 1 were not functionally validated, we aimed at condensing these to the most relevant clones by using the B cell repertoire inductive lineage and immunosequence annotator (BRILIA), an algorithm to identify clonal lineages connected through an evolutionary selection process (9). This analysis showed 3–4 times more clonal

Authorship note: LP and DS contributed equally to this work.

Conflict of interest: The authors have declared that no conflict of interest exists.

Copyright: © 2021, American Society for Clinical Investigation.

Submitted: August 3, 2020; **Accepted:** October 14, 2020; **Published:** January 4, 2021.

Reference information: *J Clin Invest.* 2021;131(1):e142966.

<https://doi.org/10.1172/JCI142966>.

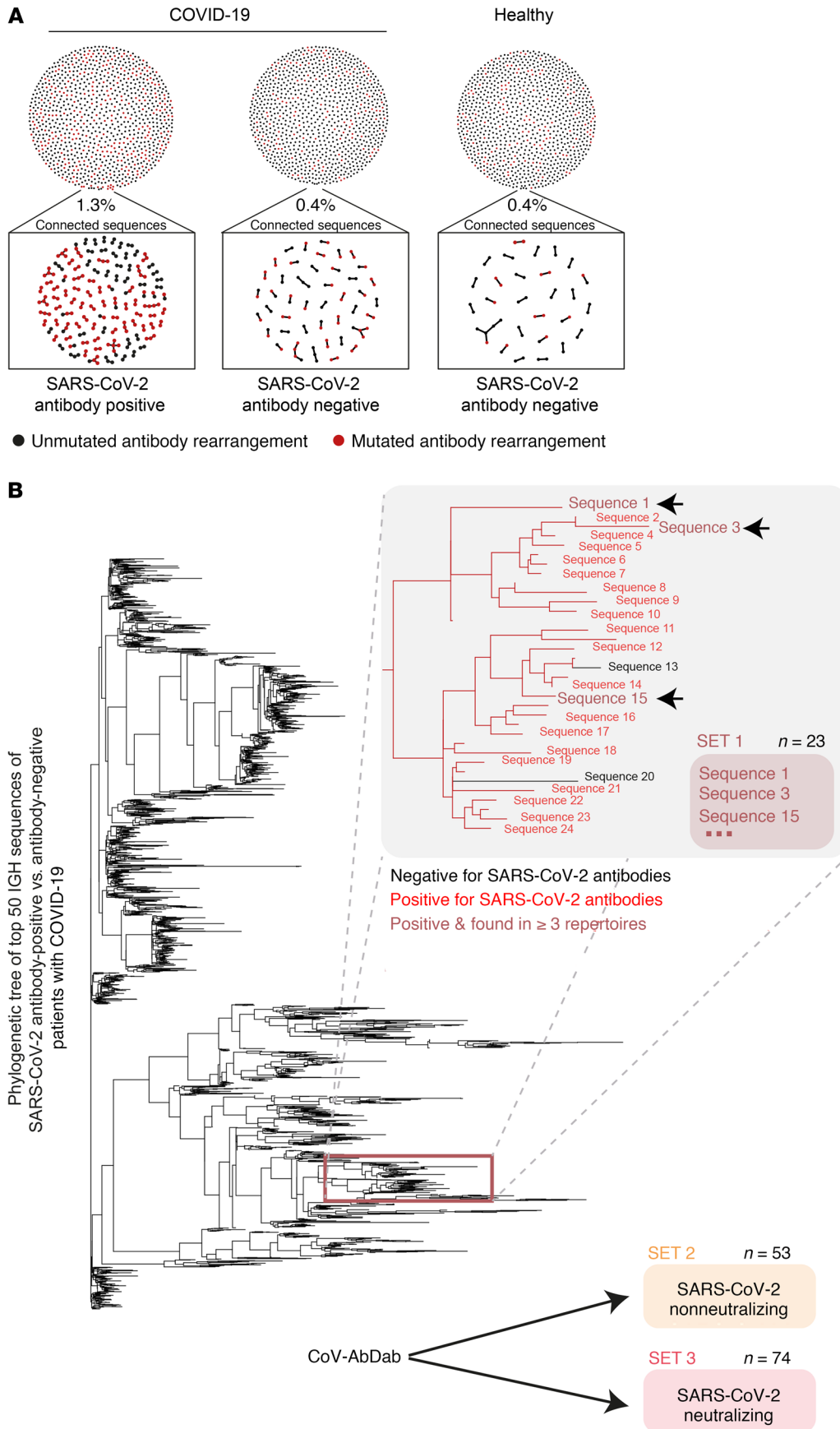


Figure 1. SARS-CoV-2-specific antibody clusters. (A) Analysis of clonal B cell receptor lineages following evolutionary trajectories from B cell repertoires of patients with COVID-19. To determine connected antibody sequences, 20,000 clonotypes from B cell repertoires of antibody-positive patients, antibody-negative patients, and healthy individuals were analyzed using the BRILIA algorithm (9). The upper circle shows the overall view on connected and unconnected sequences, the lower circles show the percentage of connected sequences as well as the type of branching observed. Red dots indicate sequences deviating more than 2% from the respective germline sequence, indicating significant somatic hypermutation. (B) Cluster analysis of B cell repertoires of patients with COVID-19 to identify SARS-CoV-2-specific antibody rearrangements. Top 50 sequences of 34 antibody-positive and 12 antibody-negative repertoires were subjected to sequence alignments and hierarchical trees were generated using the FastTreeMP (10) algorithm. Subtrees containing more than 20 sequences were selected (with fewer than 10% from antibody-negative repertoires) (red). Of these, sequences that occurred in a minimum of 3 repertoires from patients with COVID-19 with a Levenshtein distance of 2 or less were selected for sequence set 1 (dark red and bold). Sequences from set 2 and set 3 were derived from CoVAbDab (8) and comprised sequences derived from published nonneutralizing antibodies (set 2) and published neutralizing antibodies (set 3).

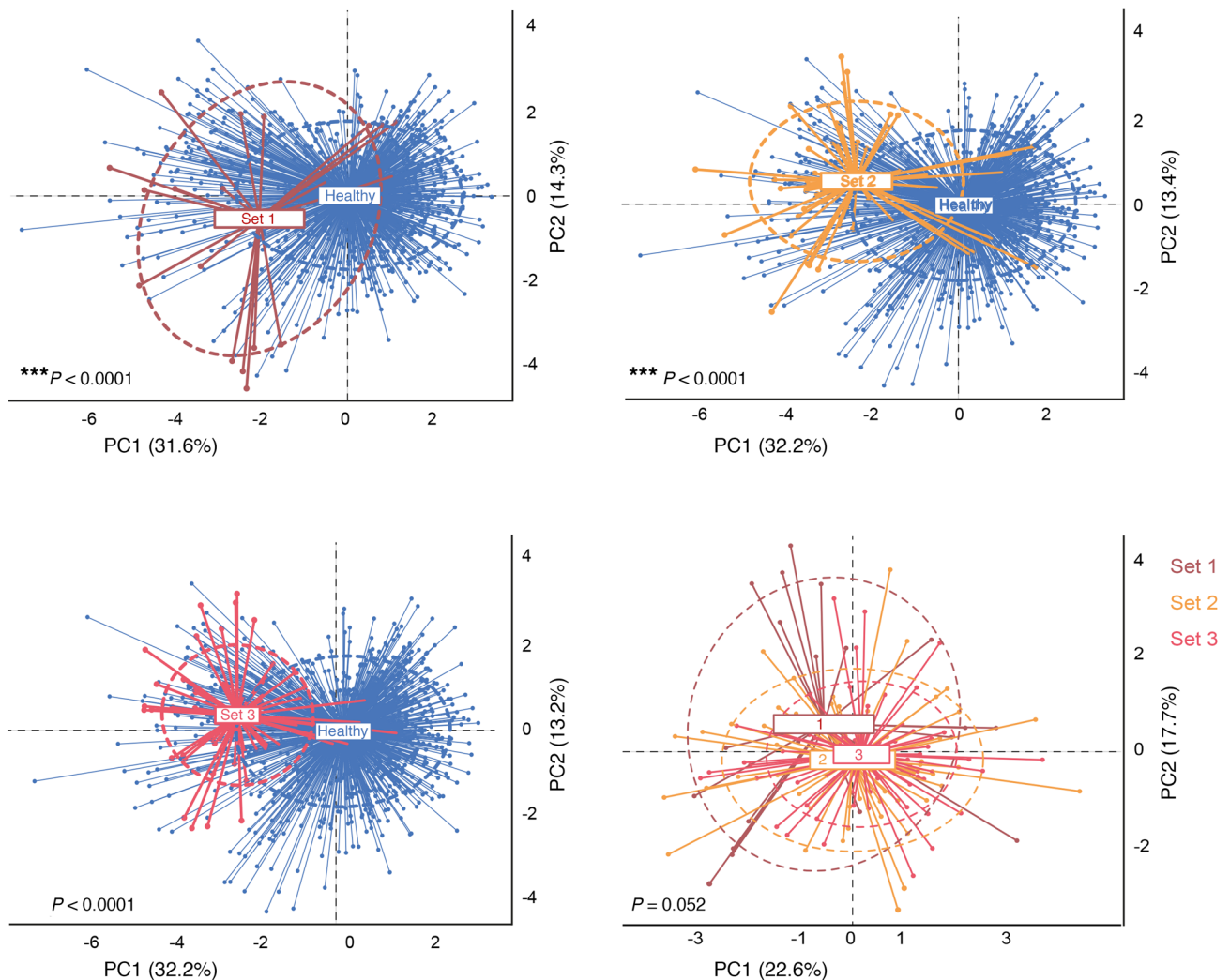


Figure 2. Biophysical properties of SARS-CoV-2-specific antibody rearrangements. The 10 Kidera factors (11) describing fundamental biophysical properties were determined for all SARS-CoV-2 antibody sequences of sets 1–3 along with control sequences from 20 healthy donor B cell repertoires. Principal component analyses are shown for each of the sequence sets versus controls or versus each other. Pillai’s test of MANOVA of principal components 1 and 2 was the statistical test used.

lineages in antibody-positive patients with COVID-19 than in controls, suggesting some affinity maturation in B cell responses to SARS-CoV-2 (Figure 1A). However, clonal lineages were not shared between patients, mostly unbranched with many connected sequences showing greater than 98% identity to germline. This confirmed that somatic hypermutation does to some extent take place, but is not essential in COVID-19. We used a stringent clustering algorithm (FastTreeMP; ref. 10) to identify globally similar rearrangements that were shared among at least 3 COVID-19 repertoires, resulting in a total of 23 sequences allocated to set 1. The majority thereof were unmutated or low-level mutated. The selection process for set 1–3 sequences, including the cluster dendrogram for set 1, is shown in Figure 1B. Set 1–3 sequences are listed in Supplemental Table 1 (supplemental material available online with this article; <https://doi.org/10.1172/JCI142966DS1>).

To explore protein similarity, we analyzed the Kidera factor profile of set 1–3 sequences, reflecting biophysical properties (11). Sets 1–3 differed significantly from healthy control repertoires by

principal component analysis; however, there was some overlap (Figure 2). No statistical differences between sets 1–3 were evident, but set 3 showed most condensed properties potentially reflecting similar epitope specificity since this category encompassed only sequences from neutralizing antibodies. Overall, this analysis showed similarity in biochemical properties of SARS-CoV-2 antibodies shared also by a number of rearrangements from repertoires of healthy individuals not previously exposed to the virus.

We reasoned that restricted repertoire metrics as potentially occurring in old age or cancer (both risk factors for severe courses of COVID-19; refs. 12, 13) may limit the chances to rapidly generate effective B cell responses due to a reduced prevalence of SARS-CoV-2-specific rearrangements in the unexposed repertoire. While T cell repertoire restriction is a well-known phenomenon in aging and cancer (14), data on age- or cancer-specific effects on certain B cell metrics are scarce (15). To study these metrics in risk groups, B cell immune repertoires from 700 blood samples collected before the current pandemic were used. These

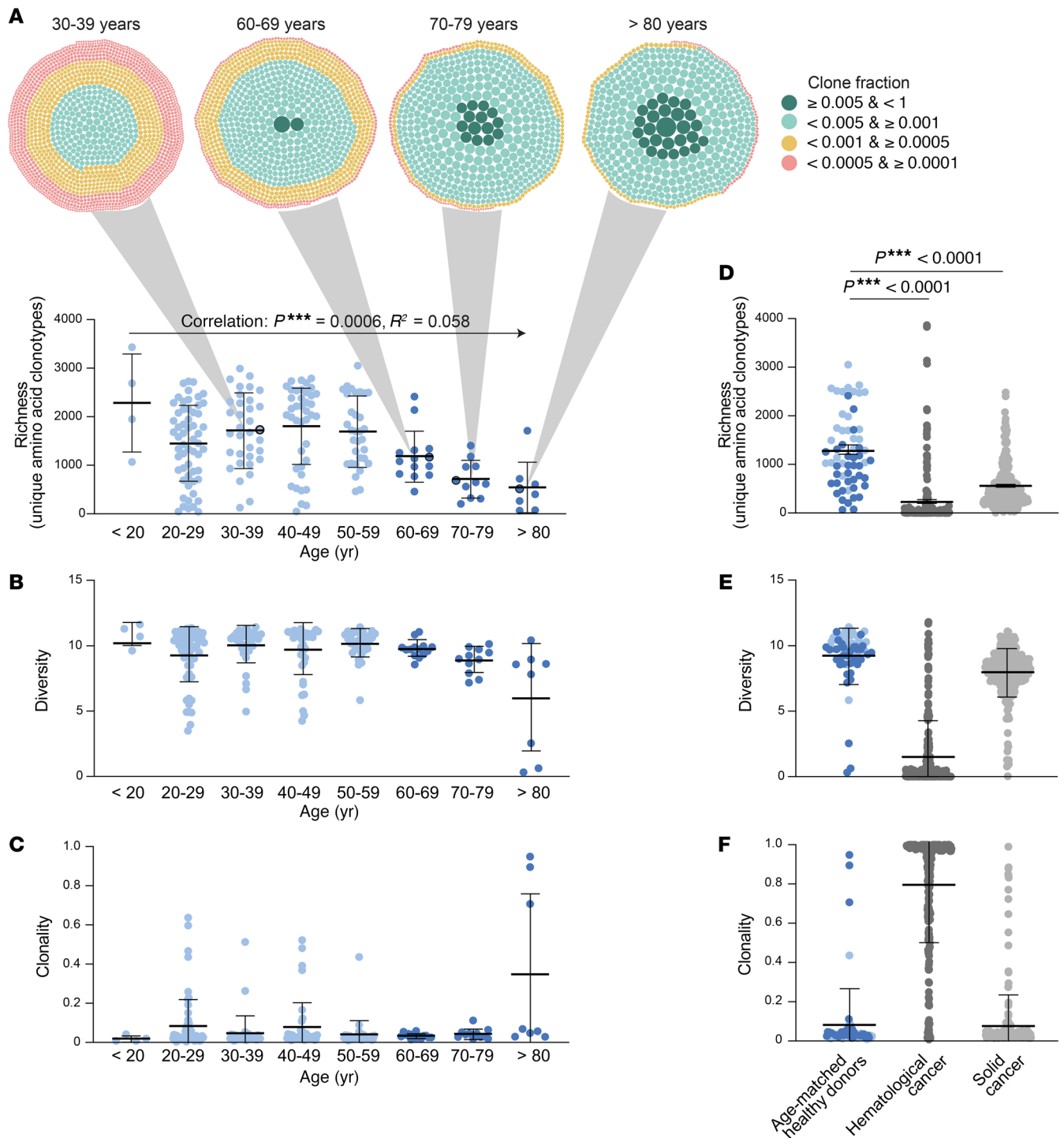


Figure 3. B cell immune repertoire metrics in individuals at risk for severe courses of COVID-19. (A–C) Age-dependent changes in B cell repertoire richness (A), diversity (B), and clonality (C). A cohort of 200 healthy individuals of all age groups is shown. Richness, diversity, and clonality derived from NGS of each individual B cell repertoire are plotted according to age group. Richness and clonal distribution in a representative repertoire of each age group are visualized in A (top). One bubble represents one amino acid clonotype, the size of the bubble is proportional to clone fraction. (D–F) Cancer-dependent changes in B cell repertoire metrics. Richness (D), diversity (E), and clonality (F) derived from NGS of B cell repertoires from 500 patients with solid tumors or blood cancer are compared with age-matched healthy individuals. Bars correspond to mean \pm SD. Statistical test: unpaired, 2-tailed *t* test.

comprised a total of 27.6 million productive reads with 669,981 unique B cell receptor clonotypes. In the 200 repertoires from healthy individuals of all age groups, we found loss in repertoire richness and diversity from 60–70 years of age onwards (Figure 3,

A–C). Moreover, 500 repertoires from patients with cancer (236 with hematological malignancies, predominantly B lineage neoplasias, 264 with solid cancers, predominantly of gastrointestinal origin; Supplemental Table 2) were studied. Patients with cancer

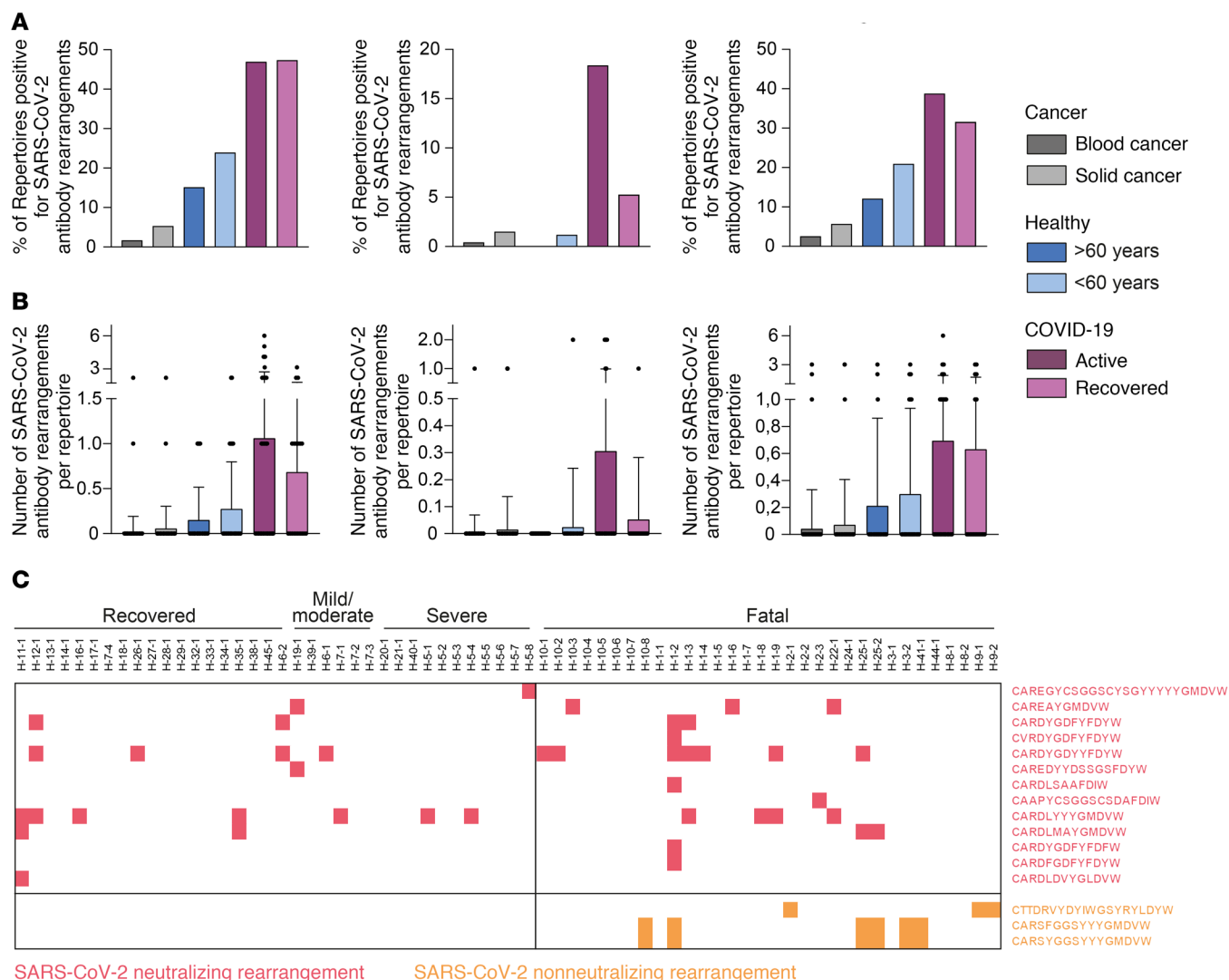


Figure 4. Mining of immune repertoires from patients with COVID-19 and individuals not previously exposed to the virus for SARS-CoV-2-specific antibody rearrangements. (A and B) Representation of SARS-CoV-2-specific sequences from sets 1–3 in immune repertoires of patients with active COVID-19, recovered patients, healthy individuals older than 60 years or younger than 60 years, and patients with blood cancer or solid tumors. (A) The percentage of repertoires with at least one sequence hit. (B) The mean number of sequence hits per repertoire. Bars correspond to mean + SD. (C) Heatmap showing the distribution of rearrangements from nonneutralizing and neutralizing (sets 2 and 3) antibodies found in patients with COVID-19. The heatmap includes all antibody rearrangements from sets 2 and 3 that were at least found in one repertoire of a patient with active COVID-19 disease from our cohort (<https://gateway-covid19.ireceptor.org/home>; study: IR-Binder-000001).

showed even more perturbations compared with age-matched controls with lower richness and diversity as well as higher clonality, independent of prior treatment (Figure 3, D–F). Next, we asked if the repertoire restrictions led to a decreased likelihood to identify SARS-CoV-2-specific unmutated rearrangements in elderly healthy individuals and patients with cancer. To test this, we mined the 200 B cell repertoires from healthy individuals younger than 60 years or older than 60 years, as well as the 500 repertoires from patients with cancer for set 1–3 sequences. The distribution pattern of set 1 and set 3 sequences in the different cohorts was relatively comparable, with most sequence hits found in patients with active COVID-19 disease or after recovery (previously published recovered cohort, ref. 5, <https://gateway-covid19.ireceptor.org/home>; study: IR-Binder-000001). Immune rep-

ertoires of nonexposed individuals showed a lower likelihood to contain SARS-CoV-2-specific antibody rearrangements than those of patients with COVID-19, with the lowest numbers of sequence hits in patients 60 years of age or older as well as in patients with cancer (Figure 4, A and B). Sequences from nonneutralizing antibodies from set 2 were less abundant than sequences from neutralizing ones in COVID-19 and control cohorts. We found a strong association of sequences from nonneutralizing antibodies with active COVID-19 disease, a cohort that contained a lot of severe or fatal cases (Figure 4, A and B).

These data showed that SARS-CoV-2-specific B cells can be identified among healthy individuals and patients with cancer without prior SARS-CoV-2 exposure, as shown recently (16). However, the likelihood for a repertoire to contain such sequences

(at least in a frequency making these clones detectable by our sampling and sequencing approach) decreases with age and is generally reduced in cancer. Next, we searched for associations of set 2 and 3 sequences with the severity of COVID-19. Sixteen antibody hits were found in patients with active disease; only 3 thereof were nonneutralizing (Figure 4C). Thirteen rearrangements from neutralizing antibodies were spread both across fatal and nonfatal cases of COVID-19, consistent with the finding that even fatal cases typically develop neutralizing antibodies. Three rearrangements from nonneutralizing antibodies were exclusively found in individuals with fatal disease courses. This may indicate detrimental properties of these nonneutralizing antibodies in patients with COVID-19, although — due to the lack of prospective validation — we cannot exclude that this association was found by chance or that these B cells and antibodies represent innocent bystanders that arise only during prolonged antigenic exposure.

Taken together, our data indicate that repertoire restriction by age and cancer — not only functional impairment of adaptive immunity — may explain unfavorable clinical courses of COVID-19 in these risk groups that need special attention. Still, a lot of open questions remain. This pertains to the important question of how precursor frequencies are ultimately connected to antibody titers upon infection with SARS-CoV-2. Elderly patients with severe infections generally show very high antibody titers in the course of the disease (4), a finding that at first glance appears to contradict the data presented here. The high titers are likely due to failed viral clearance and continuous B cell stimulation, and show that achieving such titers per se is not protective. Future studies will have to show if failure to induce early and rapid B cell responses (which are potentially favored by a high frequency of SARS-CoV-2-reactive B cell precursors) may be one of the decisive factors determining the clinical course. Moreover, our data on the naive character of the B cell response in natural infection and the potential prognostic relevance of specific SARS-CoV-2 antibody rearrangements raise the general question of what sort of immune memory COVID-19 will cause and how this may relate to vaccination responses.

Methods

Patients and samples. In total, 68 previously published COVID-19 patient samples were analyzed (5). These included blood samples from patients with mild/moderate disease (without intensive care treatment) and/or after recovery ($n = 25$) as well as samples from patients with severe or even fatal disease courses ($n = 43$). Moreover, we used pre-pandemic blood samples from 200 healthy donors of all age groups and 500 blood samples of patients with blood or solid cancer for the data mining analyzes. The latter included 236 samples from patients with blood cancer (predominantly B lineage neoplasia) and 264 samples from patients with solid cancer (predominantly of gastrointestinal origin).

NGS immunosequencing and data analysis. B cell repertoires were analyzed by next-generation immunoglobulin heavy chain sequencing (NGS) as previously described (5, 17, 18). Annotation of immunoglobulin heavy (IGH) loci rearrangements was computed with the MiXCR framework (version 3.0.8) (19). As reference for sequence alignment, the IMGT library (version 3) was used. Only productive

sequences with 2 or more read counts were considered for further analysis. In order to account for differences in repertoire read depths, we normalized each repertoire to a read depth of 10 000 reads. Each unique complementarity-determining region 3 (CDR3) nucleotide sequence was defined as one clone. All analyses and data plotting were performed using R version 3.5.1. Repertoire richness as the number of unique amino acid clonotypes and Shannon diversity (20) index were computed using package *tcR* (21). Clonality was calculated as $1 - (H' / \log_2(R))$ with H' being the Shannon diversity index and R the richness of a distinct sample (22). Clonal distributions in representative repertoires were visualized using package *bubbles* (23).

Assembly of B cell clonal lineages with BRILIA. Using 20,000 random IGH sequences from antibody-positive patients with COVID-19 and antibody-negative healthy donors, clonal lineages were assembled with BRILIA version 3.5.9 (9). Schematic visualization of 1000 sequences in the form of petriplots was done using the package *igraph* and the Fruchterman-Reingold layout (24).

Clustering algorithm for set 1 sequences. Clustering of B cell clones was performed for the most abundant 50 clones of each repertoire ($n = 34$ antibody-positive vs. $n = 12$ antibody-negative) using approximately maximum likelihood trees, a concept shown to successfully cluster B cell sequences in the context of immunization (25). Sequence similarities were inferred with *FastTreeMP* (10) and the data were visualized using R version 3.5.1 and package *ape* (26). Sequences of set 1 were deduced from subtrees containing at least 20 sequences, 18 or more of which were from antibody-positive repertoires. These sequences were filtered for the prevalence of clonotypes with identical IGHV-J genes and CDR3 amino acid sequences with a Levenshtein distance of 2 inches or less in at least 3 different antibody-positive repertoires. Twenty-three sequences matched these criteria and formed set 1.

Analysis of Kidera factor profiles. The Kidera factors were originally developed by dimension reduction on a multivariate analysis of 188 physical properties of the 20 amino acids (11), thereby summarizing the major biochemical determinants of protein structure formation. The Kidera factor profiles, including all 10 factors of the most abundant 50 clones in 20 healthy donor repertoires as well as 150 COVID-19-specific B cell clones (sets 1-3), were computed using *BRepertoire* version 1.2.0 (27). For this analysis, the IGH sequence from framework region (FR) 2 to CDR3 was used. PCA and plots were generated using R packages *ade4* (28) and *factoextra* (29).

B cell clonotype search algorithm. We searched 768 IGH repertoires for the SARS-CoV-2 rearrangements of sets 1, 2, and 3. Sequences with identical CDR3 with a Levenshtein distance of 2 or less and identical IGHV-J gene usage were targeted.

Data availability. IGH repertoire data from patients with COVID-19 are available from the iReceptor Public Archive (IPA) as part of the AIRR Data Commons using the iReceptor Gateway (gateway.ireceptor.org; iReceptor Study ID IR-Binder-000001). Moreover, raw fastq sequencing files are deposited in the European Nucleotide Archive (ENA, deposition number PRJEB38339).

Statistics. Differences in NGS metrics were studied by unpaired, 2-sided Student's *t* test and PCA differences by Pillai-Bartlett test of MANOVA. A *P* value of less than 0.05 was considered significant ($*P < 0.05$, $**P < 0.01$, $***P < 0.001$). All statistical analyses were performed using R version 3.5.1 and GraphPad Prism 8.0.2 (GraphPad Software).

Study approval. Blood collection of patients with COVID-19 was performed under institutional review board approval numbers 2020-039 and 11/17. Blood collection of prepandemic samples was performed under the approval of the ethics committee of the medical association Hamburg (project number PV4767) and the ethics committee of the medical faculty of Martin-Luther-University Halle-Wittenberg (project number 2014-75). Written informed consent was received from participants prior to inclusion in the study. The study was performed in accordance with the Declaration of Helsinki of 1975.

Author contributions

MB, LP, and DS conceived and designed the research project. DGS, MB, CS, MJGTV, and JD supplied critical material. LP, DS, CS, and EW established the study methods. LP, DS, CS performed the experiments. MB, LP, DS, EW, and CS interpreted and analyzed data. MB, LP, and DS drafted the manuscript.

Acknowledgments

We thank the following investigators and advisors for contributing patients and samples, technical assistance, supporting lab facilities, biobanking, and for ethics discussions and other input: Nuray Akyuez, Samiya Al-Robaiy, Johann Fabian Eberhardt, Stephan Eisenmann, Lena Engels, Katrin Hoffmann, Kerstin Körber-Fehrl, Beatrice Ludwig-Kraus, Sebastian Nuding, Marta Siedlecki, Christoph Wosiek, Tanja Zeller, and Lisa von Wenserski. We thank all patients and donors for donating their blood after recovery from COVID-19 for this project. This project was partially funded by the German Research Foundation (CRC 841, to MB) as well as by the Martin-Luther-University Halle (Saale).

Address correspondence to: Mascha Binder, Department of Internal Medicine IV, Oncology/Hematology, Martin Luther University Halle-Wittenberg, Ernst-Grube-Straße 40, 06120 Halle (Saale), Germany. Phone: 49.345.557.4972; Email: mascha.binder@uk-halle.de.

- McKechnie JL, Blish CA. The innate immune system: fighting on the front lines or fanning the flames of COVID-19? *Cell Host Microbe*. 2020;27(6):863–869.
- Puelles VG, et al. Multiorgan renal tropism of SARS-CoV-2. *N Engl J Med*. 2020;383(6):590–592.
- Shi Y, et al. COVID-19 infection: the perspectives on immune responses. *Cell Death Differ*. 2020;27(5):1451–1454.
- Robbiani DF, et al. Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature*. 2020;584(7821):437–442.
- Schultheiß C, et al. Next-generation sequencing of T and B cell receptor repertoires from COVID-19 patients showed signatures associated with severity of disease. *Immunity*. 2020;53(2):442–455.e4.
- Seydoux E, et al. Analysis of a SARS-CoV-2-infected individual reveals development of potent neutralizing antibodies with limited somatic mutation. *Immunity*. 2020;53(1):98–105.e5.
- Barnes CO, et al. Structures of human antibodies bound to SARS-CoV-2 spike reveal common epitopes and recurrent features of antibodies. *Cell*. 2020;182(4):828–842.e16.
- Raybould MIJ, Kovaltsuk A, Marks C, Deane CM. CoV-AbDab: the Coronavirus Antibody Database. *Bioinformatics*. 2020;btaa739.
- Lee DW, Khavrutskii IV, Wallqvist A, Bavari S, Cooper CL, Chaudhury S. BRILIA: integrated tool for high-throughput annotation and lineage tree assembly of B-cell repertoires. *Front Immunol*. 2016;7:681.
- Price MN, Dehal PS, Arkin AP. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One*. 2010;5(3):e9490.
- Kidera A, Konishi Y, Oka M, Ooi T, Scheraga HA. Statistical analysis of the physical properties of the 20 naturally occurring amino acids. *J Protein Chem*. 1985;4:23–55.
- Verity R, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infect Dis*. 2020;20(6):669–677.
- Tian J, et al. Clinical characteristics and risk factors associated with COVID-19 disease severity in patients with cancer in Wuhan, China: a multicentre, retrospective, cohort study. *Lancet Oncol*. 2020;21(7):893–903.
- Simnica D, et al. T cell receptor next-generation sequencing reveals cancer-associated repertoire metrics and reconstitution after chemotherapy in patients with hematological and solid tumors. *Oncoimmunology*. 2019;8(11):e1644110.
- Tabibian-Keissar H, et al. Aging affects B-cell antigen receptor repertoire diversity in primary and secondary lymphoid tissues. *Eur J Immunol*. 2016;46(2):480–492.
- Kreer C, et al. Longitudinal isolation of potent near-germline SARS-CoV-2-neutralizing antibodies from COVID-19 patients. *Cell*. 2020;182(4):843–854.e12.
- Schultheiß C, et al. Next-generation immunosequencing reveals pathological T cell architecture in autoimmune hepatitis [published online July 21, 2020]. *Hepatology*. <https://doi.org/10.1002/hep.31473>.
- Simnica D, et al. High-throughput immunogenetics reveals a lack of physiological T cell clusters in patients with autoimmune cytopenias. *Front Immunol*. 2019;10:1897.
- Bolotin DA, et al. MiXCR: software for comprehensive adaptive immunity profiling. *Nat Methods*. 2015;12(5):380–381.
- Shannon CE. The mathematical theory of communication. 1963. *MD Comput*. 1997;14(4):306–317.
- Nazarov VI, et al. tcR: an R package for T cell receptor repertoire advanced data analysis. *BMC Bioinformatics*. 2015;16:175.
- Pielou EC. Species-diversity and pattern-diversity in the study of ecological succession. *J Theor Biol*. 1966;10(2):370–383.
- Cheng J, Bostock M, Heer J. d3 Bubble Chart htmlwidget. <https://github.com/jcheng5/bubbles>. Accessed June 16, 2020.
- Gabor C, Tamas N. The igraph software package for complex network research. *InterJournal Complex Systems*. 2006;1695.
- VanDuijn MM, Dekker LJ, van IJcken WFJ, Sillevius Smitt PAE, Luider TM. Immune repertoire after immunization as seen by next-generation sequencing and proteomics. *Front Immunol*. 2017;8:1286.
- Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*. 2019;35(3):526–528.
- Margreitter C, Lu HC, Townsend C, Stewart A, Dunn-Walters DK, Fraternali F. BReptoire: a user-friendly web server for analysing antibody repertoire data. *Nucleic Acids Res*. 2018;46(W1):W264–W270.
- Stephane D, Anee-Beatrice D. The ade4 package: implementing the duality diagram for ecologists. *J Stat Soft*. 2007;22(4):<https://doi.org/10.18637/jss.v022.i04>.
- Kassambara A, Mundt F. factoextra: Extract and visualize the results of multivariate data analyses. <https://rpkgs.datanovia.com/factoextra/>. Accessed November 2, 2020.