ARTICLE

Mapping mutations to the SARS-CoV-2 RBD that escape binding by different classes of antibodies

Allison J. Greaney [1,2], Tyler N. Starr [1,3], Christopher O. Barnes[4], Yiska Weisblum[5], Fabian Schmidt[5], Marina Caskey[6], Christian Gaebler[6], Alice Cho[6], Marianna Agudelo[6], Shlomo Finkin[6], Zijun Wang[6], Daniel Poston[5], Frauke Muecksch[5], Theodora Hatziioannou [5], Paul D. Bieniasz [3,5], Davide F. Robbiani[6,7], Michel C. Nussenzweig[3,6], Pamela J. Bjorkman [4] & Jesse D. Bloom [1,3✉]

Monoclonal antibodies targeting a variety of epitopes have been isolated from individuals previously infected with SARS-CoV-2, but the relative contributions of these different antibody classes to the polyclonal response remains unclear. Here we use a yeast-display system to map all mutations to the viral spike receptor-binding domain (RBD) that escape binding by representatives of three potently neutralizing classes of anti-RBD antibodies with high-resolution structures. We compare the antibody-escape maps to similar maps for convalescent polyclonal plasmas, including plasmas from individuals from whom some of the antibodies were isolated. While the binding of polyclonal plasma antibodies are affected by mutations across multiple RBD epitopes, the plasma-escape maps most resemble those of a single class of antibodies that target an epitope on the RBD that includes site E484. Therefore, although the human immune system can produce antibodies that target diverse RBD epitopes, in practice the polyclonal response to infection is skewed towards a single class of antibodies targeting an epitope that is already undergoing rapid evolution.

[1] Basic Sciences Division and Computational Biology Program, Fred Hutchinson Cancer Research Center, Seattle, WA, USA. [2] Department of Genome Sciences & Medical Scientist Training Program, University of Washington, Seattle, WA, USA. [3] Howard Hughes Medical Institute, Chevy Chase, MD, USA. [4] Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA. [5] Laboratory of Retrovirology, The Rockefeller University, New York, NY, USA. [6] Laboratory of Molecular Immunology, The Rockefeller University, New York, NY, USA. [7] Institute for Research in Biomedicine, Universita della Svizzera italiana (USI), Bellinzona, Switzerland. ✉email: jbloom@fredhutch.org

Control of the SARS-CoV-2 pandemic will depend on widespread population immunity acquired through infection or vaccination. But a little over a year into the pandemic, a proliferating number of new viral lineages are rising in frequency[1–6]. These emerging lineages have mutations at <1% of all residues in the viral spike, and at no more than 3 of the ~200 residues in the spike receptor-binding domain (RBD)—yet these handfuls of mutations often substantially erode and in some cases even ablate the polyclonal neutralizing antibody response elicited by infection[7–16].

A substantial fraction of the neutralizing activity of polyclonal antibody response to SARS-CoV-2 infection is due to antibodies that target the RBD[17–21], although antibodies that target the NTD also contribute to neutralization[7–9,22–24]. Structural and binding competition studies have shown that the most potently neutralizing anti-RBD antibodies target several distinct epitopes on the RBD's receptor-binding motif[17,19,25–27]. However, the contributions of these different classes of RBD-targeting antibodies to the overall activity of the polyclonal antibody response remain less clear. It is therefore important to systematically determine both how viral mutations impact each antibody class, and how these antibody-specific effects shape the overall effects of viral mutations in a polyclonal context.

Here, we comprehensively map RBD mutations that reduce binding by structurally characterized representatives of three classes of neutralizing monoclonal antibodies that target the RBD's receptor-binding motif, as well as polyclonal plasmas from convalescent individuals from whom some of the antibodies were isolated[21,25,28,29]. We make these measurements by using a deep mutational scanning approach to systematically map how all RBD amino-acid mutations affect binding to yeast-displayed RBDs[30] The resulting escape maps allow us to systematically compare how RBD mutations affect binding by the monoclonal antibodies, and we find that the antibodies cluster in the space of viral escape in a way that largely recapitulates prior classifications based on structural analyses of the antibody epitopes. However, some of the potently neutralizing monoclonal antibodies contribute very little to the escape maps of the polyclonal plasmas, even for individuals from whom the antibodies were isolated. Instead, the plasma-escape maps usually most resemble a single antibody class (class 2 in the Barnes et al. classification[25]) that targets the face of the receptor-binding ridge that is accessible in both up and down RBD conformations. Unfortunately, a mutation that escapes this antibody class (E484K) is present in many emerging viral lineages, including B.1.351, P.1, P.2, and B.1.526[1,2,4–6]. We suggest that the skewing of the RBD-targeting polyclonal response toward a single antibody class is a factor in enabling a small number of viral mutations to sometimes substantially erode neutralizing antibody immunity.

Results

Mapping all mutations that escape binding by key classes of RBD-targeting monoclonal antibodies. Most potent neutralizing antibodies against the SARS-CoV-2 RBD target the receptor-binding motif, where they compete for binding to ACE2[17,25,31]. Antibodies targeting the RBD have been divided into four major classes based on structural analyses of their epitopes; two with epitopes overlapping with the ACE2-binding site (class 1 and class 2), potent neutralizers that do not directly bind to the ACE2 contact surface (class 3), and antibodies that target a cryptic epitope outside of the receptor-binding motif and are generally less potent (class 4)[25]. We focused our studies on several antibodies representative of the first three classes of potently neutralizing receptor-binding motif-targeting antibodies. Class 1 antibodies bind the face of the receptor-binding motif that is

accessible only when the RBD is in the up conformation (Fig. 1a); the antibodies from this class in our study are C105 and LY-CoV016. Class 2 antibodies bind a face of the receptor-binding ridge that is accessible in both the up and downconformations (Fig. 1a); the antibodies from this class in our study are C144, C002, and C121. Class 1 and 2 antibodies compete with ACE2 for RBD binding and have some overlap in their structural footprints particularly at ACE2 contact sites at the top of the receptor-binding ridge. Class 3 antibodies bind the opposite side of the receptor-binding motif (including the 443–450 loop), which like the class 2 epitope is accessible in both the up and down-conformations, but has less overlap with the ACE2-binding footprint[25] (Fig. 1a); the antibodies from this class in our study are C135 and C110. All antibodies were isolated from humans previously infected with SARS-CoV-2[21,28] and have high-resolution structures[25,28,29].

We used a previously described yeast-display deep mutational scanning approach[30] to map all amino-acid mutations to the RBD that escape binding by each antibody. Briefly, we incubated libraries of yeast expressing nearly all possible single amino-acid mutations in the RBD[32] with each antibody, and then used fluorescence-activated cell sorting (FACS) to enrich for cells expressing RBD mutants that escaped antibody binding (Fig. S1). We used deep sequencing to quantify each RBD mutation's "escape fraction," which is the fraction of cells with that mutation that fall into the FACS antibody-escape bin. These escape fractions range from 0 (no cells with the mutation fall into the antibody-escape bin) to 1 (all cells with the mutation fall into the antibody-escape bin). The escape fractions are well-correlated between independent libraries, and we report the average of duplicate measurements throughout (Supplementary Data 1 and Fig. S2). We represent the escape maps as logo plots, where the height of each letter is proportional to its escape fraction (Fig. 1b). The escape map for LY-CoV016 has been previously described[33]; all other data using this assay were newly generated in this study. Interactive versions of these maps are available at https://jbloomlab.github.io/SARS-CoV-2-RBD_MAP_Rockefeller.

The escape maps show that antibodies within the same class are generally escaped by mutations at similar sites in the RBD, with large differences between classes (Fig. 1b). For example, some class 1 but not class 2 antibodies were escaped by mutations to sites K417, D420, N460, and A475 (Fig. 1b). Class 2 but not class 1 antibodies, on the other hand, were escaped by mutations at sites E484, F490, and Q493 (Fig. 1b). Class 3 antibodies, which predominantly bind to the part of the receptor-binding motif opposite that of class 1 and class 2 antibodies (Fig. 1a), tend to be escaped by a different set of mutations, including those at sites R346, K444, and G446–N450 (Fig. 1b).

However, our maps emphasize that the antibody classes are approximate groupings, with overlap at some sites between classes. For example, class 1 and class 2 antibodies overlap in their binding footprints at the top of the receptor-binding ridge that contacts ACE2 (Fig. 1a), and mutations to some of these overlapping sites (i.e., L455, F456, F486, and Y489) escape binding by some antibodies from both classes (Fig. 1b). Similarly, the class 3 antibody C110 and the class 2 antibodies can both be escaped by mutations to site F490. There are also differences in escape mutations within antibody classes. For instance, the class 3 antibody C135 is strongly affected by mutations at site R346 but not L452, whereas the opposite is true for C110 (Fig. 1b). Other examples can be found for the class 2 antibodies: C002 is escaped by some mutations to site L452, whereas C144 and C121 are less affected by these mutations (Fig. 1b).

Our escape maps make it possible to group the antibodies using a more continuous approach, without strict divisions into three classes. Specifically, we used multidimensional scaling to
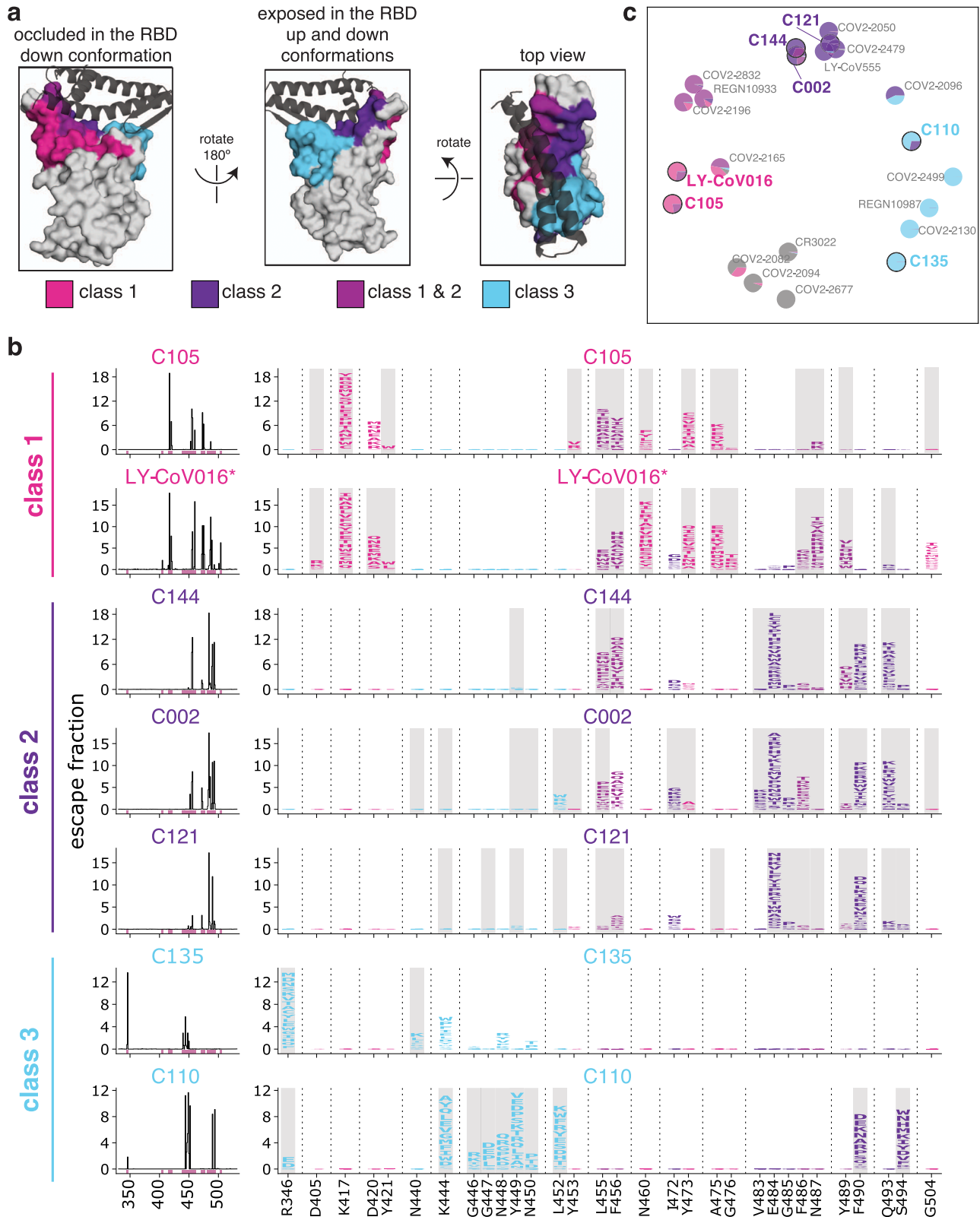
*LY-CoV016 escape mapping was previously performed in Starr, et al. (2021)

project the escape maps into a two-dimensional space of binding escape (note that this computational technique has previously been used to visualize the antigenic relationships among strains of influenza virus[34]). We did this for the seven antibodies shown in Fig. 1b, as well as 15 other antibodies for which we have previously determined escape maps[30,33,35,36], including class 4

antibodies, which bind to the RBD outside the receptor-binding motif and are generally less potently neutralizing[17,19,26,31]. Antibodies with similar escape mutations are located close to one another in the multidimensional scaling projection, and antibodies with very distinct escape mutations are far apart (Fig. 1c). The projection shows that while there is some clustering

Fig. 1 Maps of mutations to the RBD that escape binding by three classes of monoclonal antibodies that target the receptor-binding motif. a Epitopes for each of the three antibody classes[25]. ACE2 is shown as a gray cartoon. Some sites fall under both class 1 and class 2 and are shown as an intermediate pink-purple. **b** Escape maps for monoclonal antibodies from each of the three classes. The line plots at left indicate the summed effect of all mutations at each site in the RBD, with larger values indicating a greater reduction in antibody binding. The logo plots at right show the effects of individual mutations at key sites (indicated by purple highlighting on the x axis of the line plots). In these logo plots, the height of each letter is that mutation's escape fraction, so larger letters indicate mutations that cause a greater reduction in antibody binding. Sites in the logo plots are colored by RBD epitope. Sites that contact antibodies (non-hydrogen atoms within 4 Å in high-resolution structures) are highlighted with gray backgrounds. For C110, sites 444, 446, and 447 are unresolved in the structure but are likely in close contact with the antibody, and so are highlighted in gray. The data for LY-CoV016 were previously reported[33] and are replotted here. **c** Multidimensional scaling projection of the escape maps, such that antibodies with similar escape mutations are drawn close together. Each antibody is represented as a pie chart colored according to the amount of escape in each RBD epitope. Antibodies for which escape maps are shown in panel (**b**) have black outlines and colored names. The other antibodies were profiled previously[30,33,35,36]. Escape maps for all class 1, 2, and 3 antibodies in the plot are shown in Fig. S3; for the class 4 antibody-escape maps, see Greaney et al.[30]. Colors used in all panels: sites are colored according to epitopes, as defined in Greaney et al.[25]. Bright pink for class 1, dark purple for class 2, medium pink-purple for class 1/2 overlap sites (455, 456, 486, 487, 489), cyan for class 3, and gray for all other RBD sites. Escape maps colored according to mutation effects on RBD expression and ACE2 binding are in Fig. S4. The ACE2-bound RBD structure in (**a**) is from PDB 6M0J[64].

by structural class in space of escape, the antibodies are continuously distributed. For instance, the class 3 antibody C110 also selects escape at some class 2 sites, such as F490 and S494, and so C110 is located somewhat between the class 2 and 3 antibodies in the multidimensional scaling projection (Fig. 1c).

RBD mutations reduce antibody binding at only a subset of contact sites. We took advantage of the availability of high-resolution structures to compare the sites of escape mutations to the structural contacts between the antibodies and RBD. Most mutations that escape antibody binding are at sites in the RBD that directly contact the antibody; these sites are highlighted in gray in Fig. 1b, with a structural contact defined as any non-hydrogen atom within 4 Å. To visualize the escape mutations in a structural context, we mapped the extent of escape at each site to the structure of the antibody-bound RBD[25,28,29] (Figs. 2a and S5A; interactive, zoomable versions at https://jbloomlab.github.io/SARS-CoV-2-RBD_MAP_Rockefeller). All sites at which mutations strongly escape binding are in direct (<4 Å) or proximal (4–8 Å) contact with antibody in the resolved structures (Fig. 2b).

However, not all antibody-contact sites had mutations that strongly escaped antibody binding (Figs. 2b and S5A). There are several explanations. First, our approach maps functional antibody-escape mutants that retain proper RBD folding and bind to ACE2 with ≥1% the affinity of the unmutated RBD[32] (see "Methods"). Only 2304 of the 3819 possible amino-acid mutations to the RBD meet these criteria, and some sites have no tolerated mutations. For instance, G416 and R457 are both in the structural epitope of C105, but these sites have no tolerated mutations and thus do not appear in the escape map (sites with no tolerated mutations are indicated in dark gray in Figs. 2a and S5). Second, sometimes mutations at antibody-contact sites simply do not strongly disrupt antibody binding[37]. For instance, site F486 is in structural contact with both C105 and LY-CoV016 and has many well-tolerated mutations, but mutations at this site more strongly affect the binding of LY-CoV016 than C105 (Figs. 1b, 2a and S5A). Other examples include site R346, where nearly all mutations escape C135 but only charge-reversal mutations escape C110 (Figs. 1b, 2a and S5A). Similarly, at site Q493, C144, and C002 are escaped by many mutations, but C121 is only escaped by Q493K/R (Figs. 1b, 2a and S5A).

For some of the class 2 antibodies, the antibody makes a quaternary contact with an adjacent RBD in the context of spike trimer[25] (Fig. S5B). Our yeast-display system assays antibody binding to isolated RBD, and so does not map escape mutations to quaternary contact sites and cannot inform on their importance for antibody binding.

The escape maps of polyclonal plasmas often differ from those of monoclonal antibodies isolated from the same individual. The six antibodies newly mapped in this study were isolated from four different SARS-CoV-2 convalescent individuals (Fig. 3a). Plasma was collected from these individuals at the same time that blood was collected for antibody isolation[21]. Because our escape-mutant mapping approach works for polyclonal sera or plasmas in addition to monoclonal antibodies[18], we mapped mutations that reduced binding by each of the four plasmas plus one plasma sample without corresponding antibodies (Fig. 3b; the plasmas are prefixed with "COV-" to distinguish them from the antibodies which are prefixed with "C").

The plasma-escape maps shared many commonalities across all five individuals (Fig. 3b), and generally resembled those from our prior study of a larger cohort of convalescent individuals[18]. In particular, mutations to sites F456 and E484 reduced binding for all five plasma samples (Fig. 3b). Mutations to site E484 are of special note as the E484K mutation is present in the emerging B.1.351, P.1, P.2, and B.1.526 SARS-CoV-2 lineages[1,2,4–6] and can reduce the neutralization titer of convalescent plasma by threefold or more[8,10–12,18]. Mutations to other sites, such as G446–N450 and F486 also reduce binding for some of the plasma samples profiled here (Fig. 3b), consistent with our prior study of a larger cohort[18]. These findings suggest that while there is some heterogeneity among which mutations reduce binding of different individuals' polyclonal plasma antibodies, there are also sites that are commonly targeted and should be monitored for antigenic evolution.

While the escape maps for the different plasma samples shared broad similarities, they often starkly differed from the escape maps of monoclonal antibodies isolated from the same individuals (Fig. 3b, compare plasma-escape maps in logo plots with overlay bars showing sites of escape for antibodies from the same individual). For instance, mutations to R346 had the largest effects on binding by the class 3 antibody C135, but had little effect on the same individual's plasma (COV-72). Similarly, mutations at K417 had the largest effects on binding by the class 1 antibody C105, but had little effect on the corresponding plasma (COV-107). Conversely, mutations to site G496 reduced binding by the COV-21 plasma, but did not strongly affect any of the monoclonal antibodies in this study (Fig. 3b). Overall, the correlations between the sites at which mutations escaped binding for the monoclonal antibodies and their corresponding polyclonal plasmas were highest for the class 2 antibodies, and lower for the other antibody classes (Fig. 3c).

Class 2 antibodies contribute the most to the RBD escape maps of polyclonal plasmas. To more broadly compare how antibodies
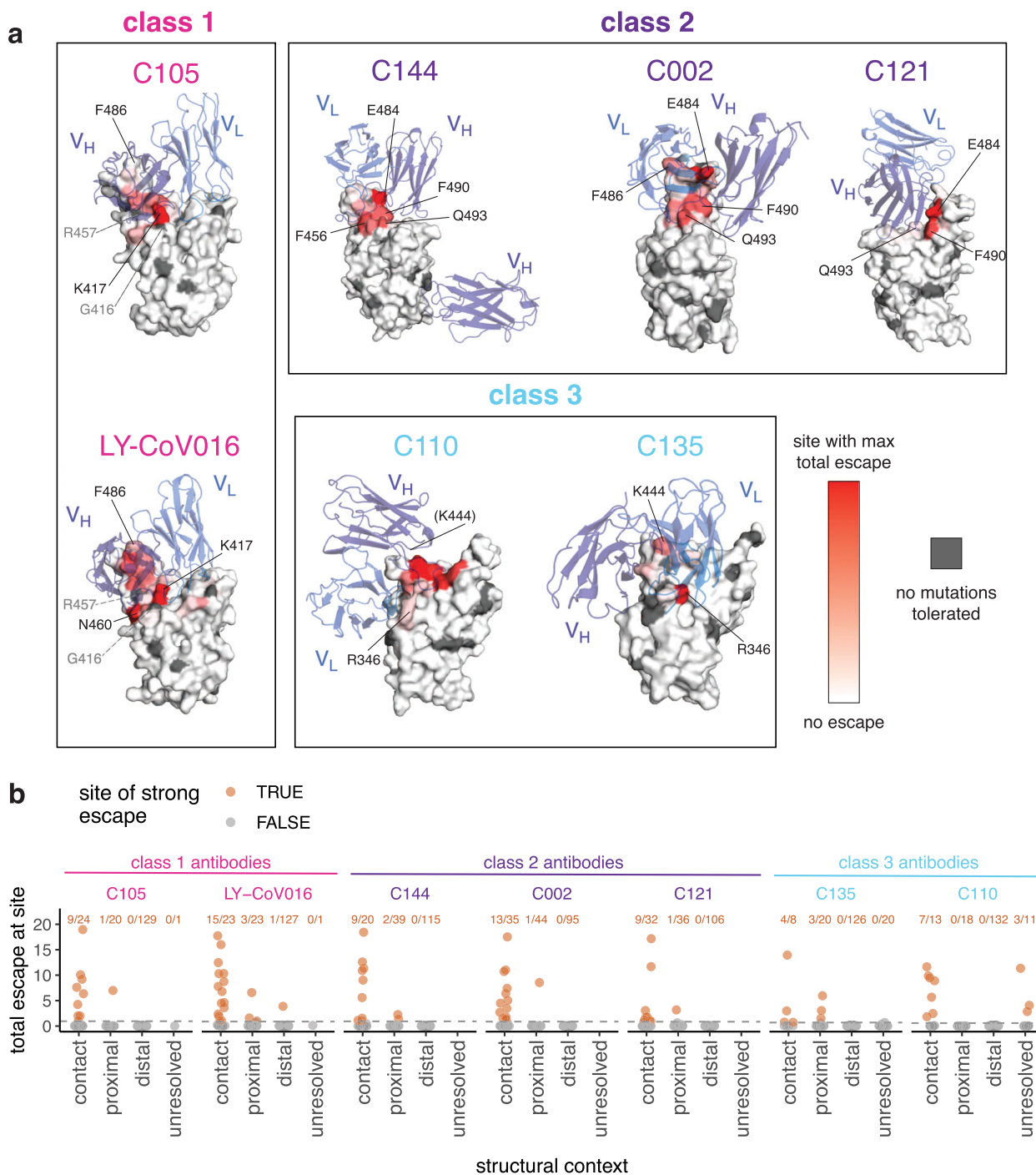
Fig. 2 Mutations that escape antibody binding are usually in the direct structural footprint. a The total escape at each site is mapped onto the surface of the Fab-bound RBD, with white indicating no escape and red indicating the site with the most escape from that antibody. Sites where no mutations are tolerated for RBD folding or ACE2 binding are indicated in dark gray. For C105 and LY-CoV016, gray labels with dashed lines indicate example contact sites with no tolerated mutations. For C110, the general area where site 444 (unresolved in structure) would be located is indicated. **b** Total escape at each site in the RBD, with sites classified according to whether they are an antibody contact (within 4 Å), antibody-proximal (4 to 8 Å), antibody-distal (>8 Å), or unresolved in the Fab-spike trimer structure. The text indicates the number of sites in each structural category that are sites of strong escape, (n/total) shown in orange. See "Methods" for details and PDB accessions.

of different classes contribute to the convalescent plasma-escape maps, we used multidimensional scaling to project 22 antibodies and 28 polyclonal plasmas into a two-dimensional space of binding-escape mutations (Fig. 4a; the projection shows the 22 antibodies in Fig. 1c, 5 plasmas from Fig. 3, and 23 plasmas from a previously characterized larger cohort[18]). The plasmas from both cohorts cluster together in the space of binding escape, but

far from some of the antibodies (Fig. 4A). In particular, while the plasmas are often clustered towards the middle of the plot, suggesting contributions from multiple antibody classes, most plasmas are positioned closest to the class 2 antibodies in the space of binding escape (Fig. 4a; an interactive version that includes additional antibodies and vaccine sera is available at https://jbloomlab.github.io/SARS2_RBD_Ab_escape_maps/).
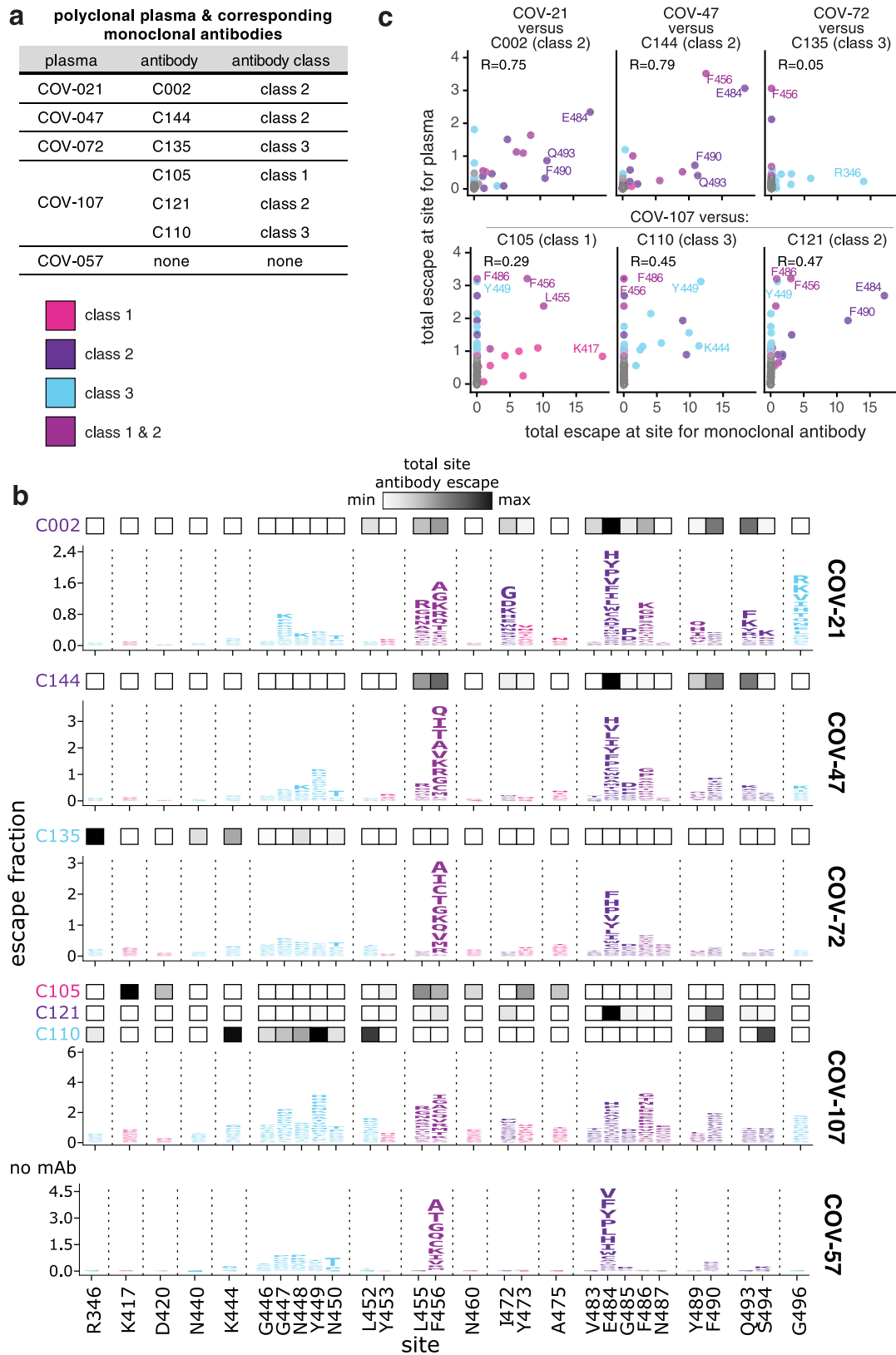
Fig. 3 The mutations that reduce binding of polyclonal plasmas often differ from those that reduce binding by monoclonal antibodies isolated from the same individual. a Table indicating which plasmas and antibodies were derived from the same individual. **b** Escape maps for the polyclonal plasma antibodies, as in Fig. 1b. The y axis is scaled separately for each plasma (see "Methods"). When there are monoclonal antibodies isolated from the same individual, the total monoclonal antibody escape at each site is shown using the heat maps above the escape maps, with white indicating no effect and black indicating strong escape. **c** Correlation of plasma and monoclonal antibody escape for each plasma/antibody pair from the same individual. Each point in the scatter plots is a site, with the x axis indicating the total escape at that site for the antibody and the y axis indicating the total escape at that site for the plasma. Key sites are labeled. Pearson's R shown above each plot. Colors in **b**, **c** reflect antibody classes as in Fig. 1.
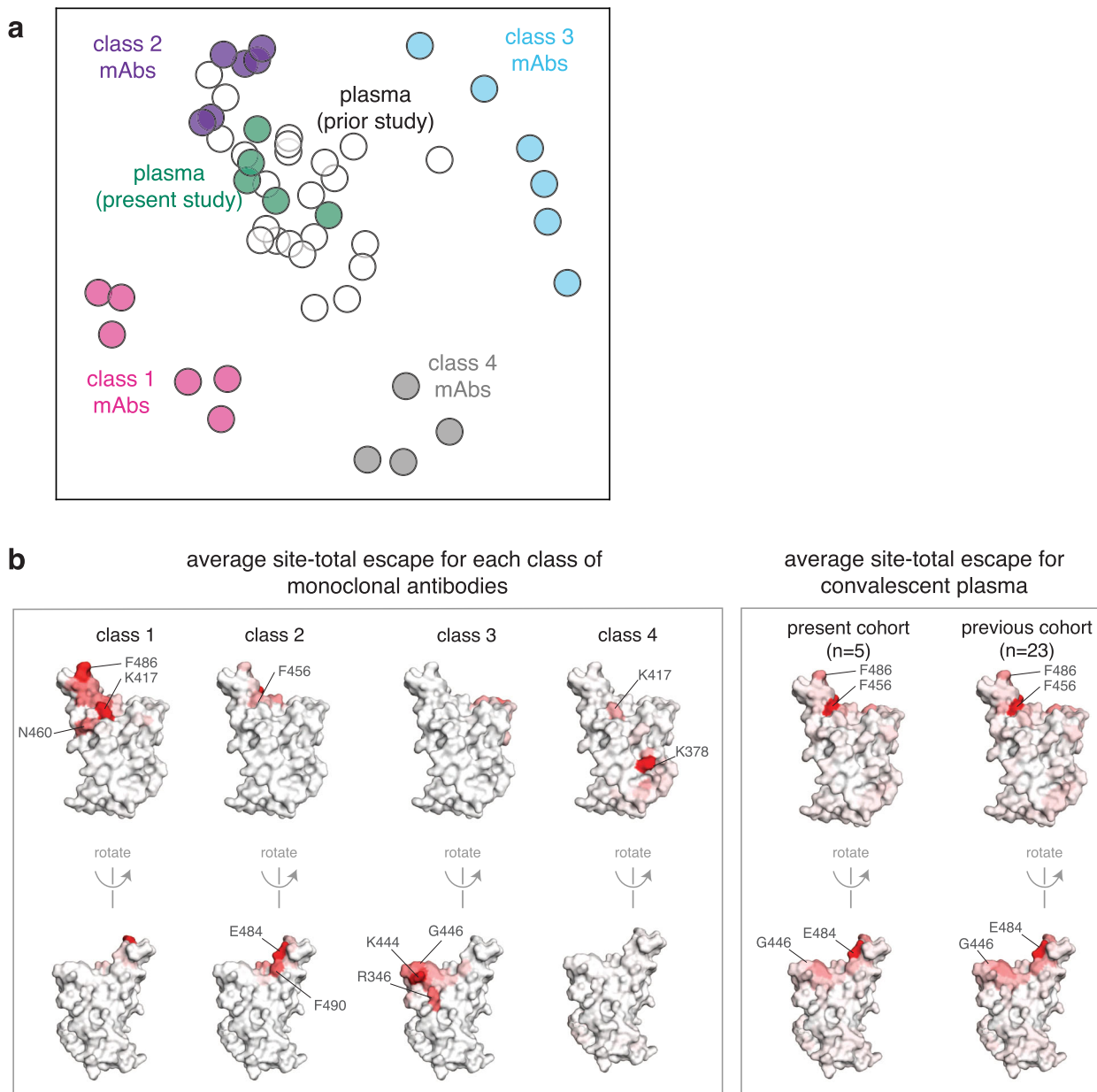
Fig. 4 The escape maps of convalescent polyclonal plasmas most resemble class 2 antibodies. a Multidimensional scaling projection of the escape maps of polyclonal plasmas and monoclonal antibodies of each class. Antibodies or plasmas that are nearby in the plot have their binding affected by similar RBD mutations. The antibodies are those in Fig. 1c, colored according to antibody class, as in Fig. 1. The five plasmas newly mapped in this study are shown in green, and the previously mapped 23 plasmas[18] are shown in white. **b** Structural projection of sites where mutations reduce binding by each class of monoclonal antibodies (left) or polyclonal plasmas (right). The RBD surface coloring is scaled from white to red, with white indicating no escape, and red indicating the site with the greatest average site-total escape for all antibodies or plasmas in that group. Mutations to sites such as E484, F456, and F486 have some of the largest effects on binding by polyclonal plasmas and class 2 antibodies. An interactive version of (**a**) that includes additional antibodies and vaccine sera is available at https://jbloomlab.github.io/SARS2_RBD_Ab_escape_maps/.

To visualize the escape maps in terms of the RBD's three-dimensional structure, we projected onto the surface of the RBD the total escape at each site averaged across all antibodies in a class or all convalescent plasmas in a cohort (Fig. 4b). Again, the polyclonal plasmas most closely resembled the class 2 antibodies. For instance, mutations to site E484 greatly reduced the binding of both class 2 antibodies and polyclonal plasmas (Fig. 4b). Most of the class 1 antibody contributions to the plasma-binding-escape maps came from sites shared with class 2 antibodies, such as F456 (Fig. 4b), although note that mutations to F456 often do not strongly reduce plasma neutralization[18]. Consistent with the lesser contributions of class 1 antibodies, mutations at the class

1 site K417 had little effect on plasma binding, and others have found that the K417N mutation alone has a minimal-to-modest effect on plasma neutralization[10,13,15,16]. Sites of escape from class 3 antibodies (e.g., G446) had visible effects on the polyclonal plasmas, but again less so than for class 2 antibody sites (Fig. 4b). However, in a prior larger study[18], we found that mutations to the 443–450 loop strongly reduced binding of plasmas from a minority of individuals, consistent with a few plasmas falling closer to class 3 than class 2 antibodies in the space of binding-escape mutations (Fig. 4a). Overall, these results show that while multiple antibody specificities contribute, class 2 antibodies usually dominate convalescent polyclonal plasmas—although
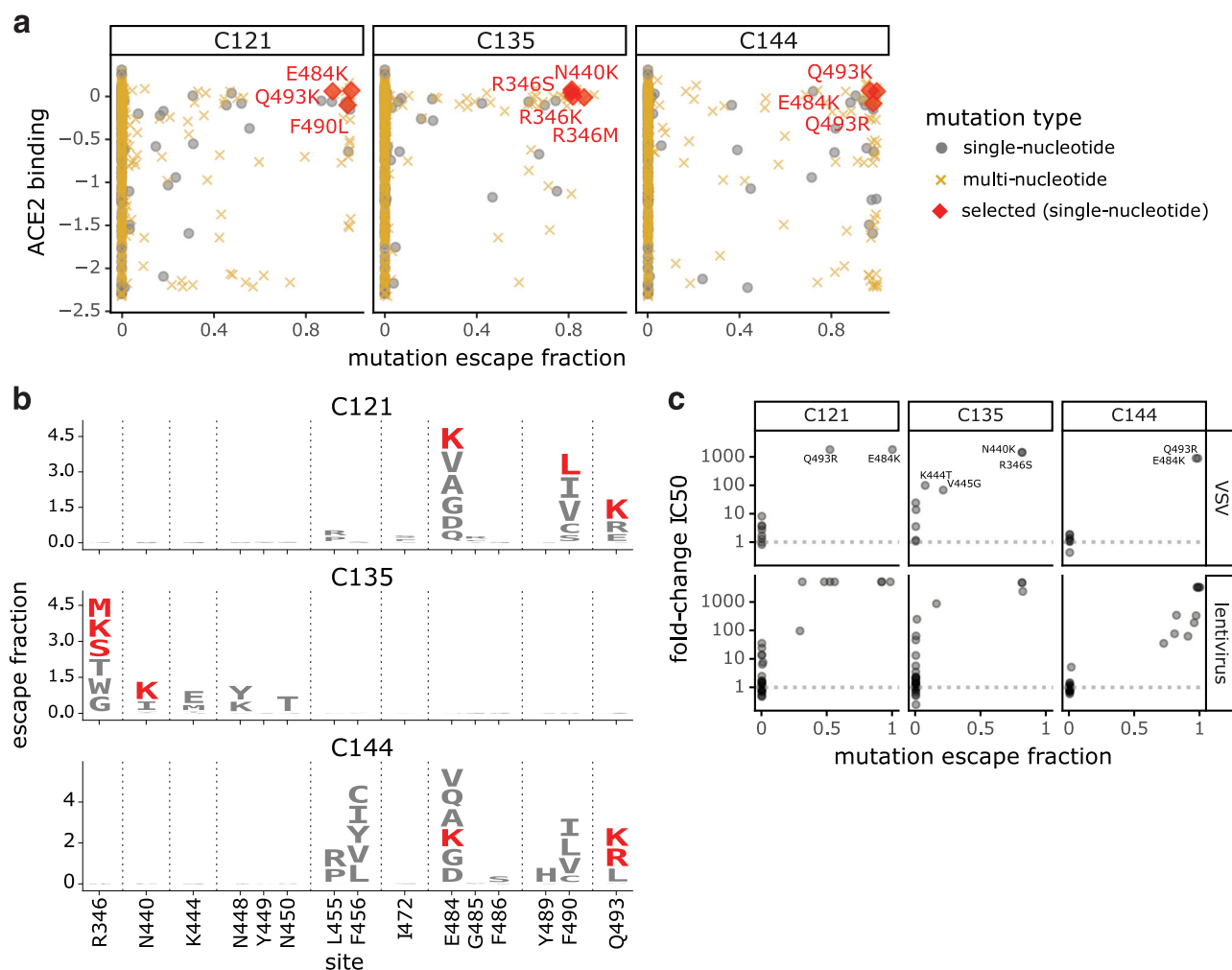
Fig. 5 Escape maps predict mutations that are selected during viral growth in the presence of monoclonal antibodies. a Mutations selected when chimeric VSV encoding the SARS-CoV-2 spike was grown in the presence of each of the three indicated antibodies by Weisblum et al.[7]. Each point represents a different amino-acid mutation, with the x axis indicating how strongly the mutation escapes antibody binding (measured in the current study) and the y axis indicating how well the mutant binds to ACE2 (measured in Starr et al.[32]). The red diamonds indicate the mutations selected in VSV-spike by Weisblum et al., the gray circles indicate all other amino-acid mutations accessible by a single-nucleotide change, and the gold x's indicate amino-acid mutations that require multiple nucleotide changes to the codon. **b** Logo plots showing the effects of only single-nucleotide accessible amino-acid mutations on antibody binding. Mutations selected in VSV-spike virus by Weisblum et al. are colored red. **c** The correlation of the effects of mutations on antibody binding measured in the current study and effects on viral neutralization previously measured by Weisblum et al.[7] using chimeric VSV (top) or lentiviral particles (bottom). The x axis shows the escape fraction measured in the current study, and the y axis shows the fold change in inhibitory concentration 50% (IC50) for viral neutralization caused by that mutation, such that larger numbers correspond to greater reductions in neutralization sensitivity. For effects of all antibody- and plasma-binding-escape mutations on ACE2 binding and RBD expression, see Fig. S4. For each mutation's escape fraction compared to fold-change IC$_{50}$ against each monoclonal antibody or polyclonal plasma tested in Weisblum et al.[7], see Fig. S6.

once a virus has accumulated mutations in class 2 epitopes (as has already occurred in some emerging lineages[1,2,4,6]), then class 1 or 3 antibodies might dominate for the remaining anti-RBD-antibody activity.

Escape maps are consistent with the RBD mutations that arise when the virus is grown in the presence of monoclonal antibodies. We assessed how well our escape maps predicted the actual antibody-escape mutations that arose when the virus was grown in the presence of the antibodies. Prior work selected viral escape mutants by passaging chimeric vesicular stomatitis virus (VSV) encoding the SARS-CoV-2 spike in the presence of several of the monoclonal antibodies we mapped in this study[7]. We hypothesized that the mutations selected during viral passage would reduce antibody binding without impairing ACE2-binding affinity. Accordingly, we examined how all of the selected

mutations affected both antibody binding (as measured in the current study) and ACE2 affinity (as measured in our prior deep mutational scanning[32]). Figure 5a, b shows that in every case, the antibody-escape mutations selected in the virus were indeed among the single-nucleotide-change accessible amino-acid mutations that mediated the strongest escape from antibody binding without strongly impairing ACE2 affinity. Conversely, mutations that escaped antibody binding but were deleterious for ACE2 binding or RBD expression (e.g., E484V/A and mutations to sites 455 and 456) were not selected in the viral passaging. Therefore, our escape maps can be used in conjunction with prior data on the functional effects of RBD mutations to largely predict which escape mutations will arise when the virus is grown in the presence of antibodies.

Weisblum et al. also tested many RBD point mutations for their effects on neutralization of chimeric VSV or lentiviral
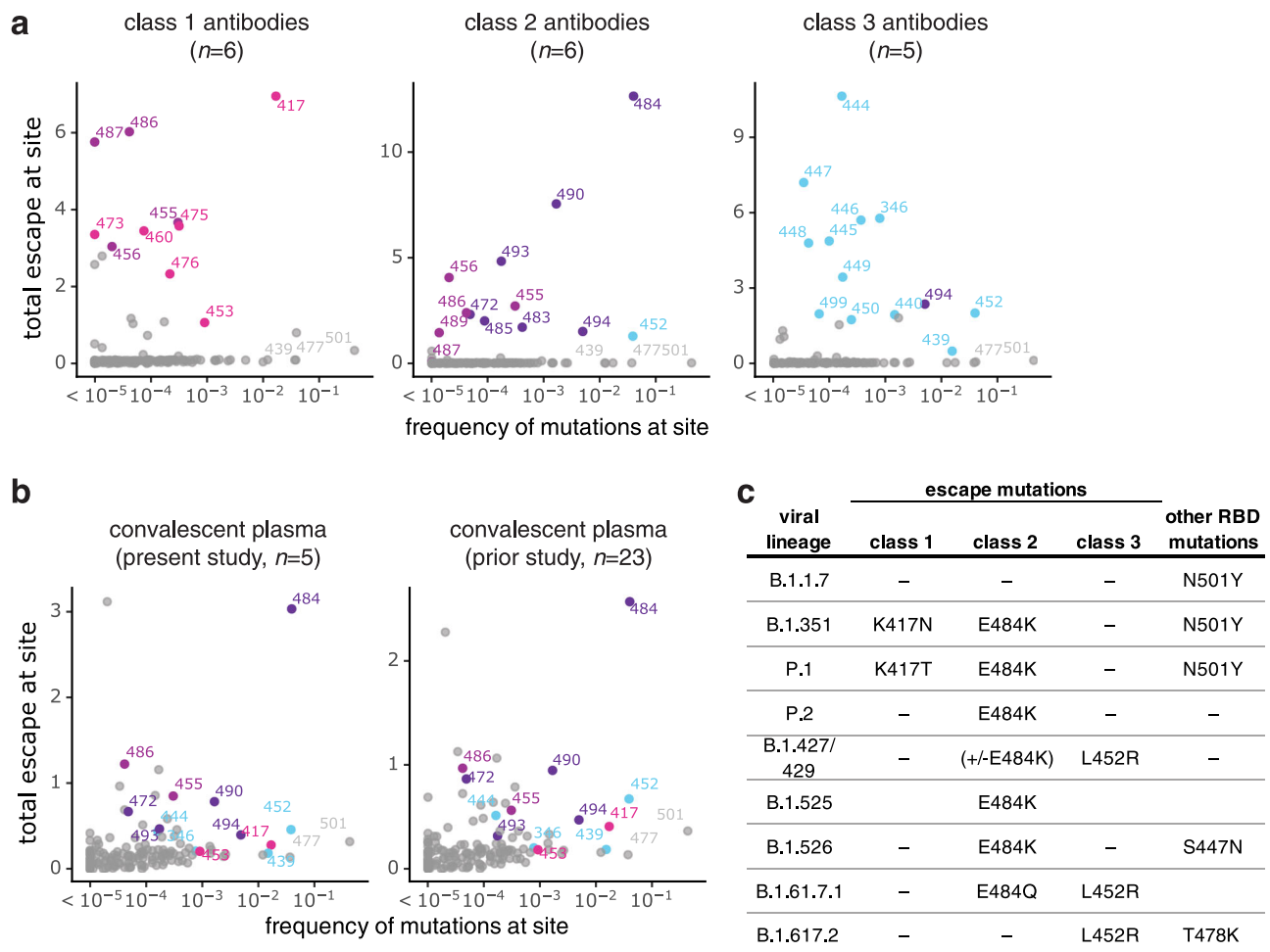
Fig. 6 Mutations that escape binding by antibodies and plasmas among sequenced SARS-CoV-2 isolates. a The total escape at each site averaged across antibodies in each class versus frequency of mutations at each site in GISAID sequences as of May 11, 2021. **b** The total escape at each site averaged across the polyclonal plasmas versus frequency of mutations at each site in GISAID sequences. Left: plasma samples profiled in this study, right: plasma samples profiled previously[18]. **c** Antibody-escape mutations found in emerging viral lineages. RBD mutations in each lineage are assigned to the class of antibody they most strongly escape (e.g., E484K most strongly escapes class 2 antibodies but may also affect some class 1 antibodies). Other RBD mutations present in each viral lineage with negligible effects on binding of these antibodies are listed at right. For numbers of antibodies in each class, see n indicated in each panel in **a** and **b**. In each plot, key sites are labeled and colored according to RBD epitope using the same color scheme as in Fig. 1.

particles by the antibodies studied here[7]. There was generally good agreement between our escape maps and these previously measured effects of mutations on viral neutralization. Nearly all mutations with a >100-fold reduction in neutralization also had large effects in our escape maps, although in a few cases mutations with more moderate effects on neutralization were not prominent in the escape maps (Figs. 5c and S6A). Previously, we and others have reported that single point mutations can reduce the neutralization of some plasmas by >10-fold, although other plasma are largely unaffected by any single mutation[7,8,18]. For the plasmas in this study, prior work found that no tested mutation had such large effects on neutralization[7]. However, the class 2 antibody-escape mutation E484K did reduce neutralization by COV47 plasma by approximately fivefold[7], concordant with the prominence of site 484 in that plasma's escape map (Figs. 3b and S6B).

Mutations that reduce binding by class 1, 2, and 3 antibodies are present in emerging viral lineages. To assess the extent that SARS-CoV-2 has already acquired mutations that reduce binding by each antibody class, we compared the total escape at each site averaged across all antibodies of that class to the frequency of

mutations at the site among sequenced SARS-CoV-2 isolates in GISAID as of May 11, 2021[38]. Figure 6a shows that mutations at sites targeted by each antibody class are present at appreciable frequencies among sequenced SARS-CoV-2 isolates. In particular, sites K417 and E484 are the strongest sites of escape for class 1 and class 2 antibodies, respectively—and mutations at both these sites are present in a substantial number of sequenced viral isolates (Fig. 6a). In contrast, mutations that most strongly escape class 3 antibodies are currently not as prevalent among sequenced isolates (Fig. 6a). While mutations have been observed at site K444 (the strongest site of escape from class 3 antibodies), these are at a lower frequency than site 417 or 484 mutations (Fig. 6a). Mutations at the class 3 site L452 are present at a higher frequency, but L452 is only a moderate site of escape for this antibody class. For instance, mutations to L452 escape binding of C135 but do not strongly escape the other class 3 antibodies. However, L452 mutations also escape binding of some class 2 antibodies. As expected from the fact that class 2 antibodies dominate convalescent polyclonal plasmas, the natural frequency versus escape plots for the plasmas closely resemble those for class 2 antibodies, with mutations at site E484 having the largest effect (Fig. 6b).

We also examined the presence of escape mutations for each antibody class in some key emerging viral lineages (Fig. 6c). All these emerging viral lineages except B.1.1.7[14,16,39,40] have a mutation that escapes some antibodies from at least one class. The class 2 antibody escape mutation E484K is present in the B.1.351, P.1, P.2, and B.1.526 lineages (Fig. 6c)[1,2,4–6]. Two of these lineages, B.1.351 and P.1, also have a class 1 antibody escape mutation, K417N or K417T, respectively (Fig. 6c). The B.1.427/429 and B.1.617 lineages carry a class 3 antibody escape mutation (L452R; Fig. 6c). No viral lineages currently combine mutations that escape all three antibody classes—but the future emergence of a viral lineage that has accumulated mutations that escape all three antibody classes would be a worrying development and should be monitored for closely.

Discussion

We comprehensively mapped all mutations that escape binding by three major classes of antibodies targeting the SARS-CoV-2 RBD and compared these escape maps to those for convalescent polyclonal plasmas. We find that while multiple antibody specificities contribute, a single antibody class (class 2) largely shapes how RBD mutations affect binding by polyclonal plasmas, even for individuals from whom potent neutralizing antibodies of other classes were isolated. A similarly unequal immunodominance of different epitopes has also been described for other viral antigens, such as influenza hemagglutinin[41,42]. However, the immunodominance of different epitopes can vary depending on the antigen formulation and delivery[41], and there is preliminary evidence that some vaccines could elicit antibody responses to SARS-CoV-2 with immunodominance hierarchies that are less skewed than those elicited by infection[10,16,43,44].

The importance of class 2 antibodies in the RBD-targeting portion of polyclonal plasma could be due in part to the twin facts that their epitope is exposed in both up- and downconformations of the RBD, and that such antibodies are often generated from frequently observed germline genes (including *VH1-2*, *VH3-53*, *VH1-69*)[21,45–51]. Consistent with our work here showing that class 2 antibodies shape how mutations affect polyclonal plasma binding, mutations at the site that most strongly affects binding by this antibody class (E484) have arisen multiple times in emerging viral lineages[1,2,4–6]. Therefore, our results show the importance of thinking about antigenic evolution in the context of different classes of antibodies that recognize different epitopes on the RBD.

Our work also sheds light on the extent to which it is functionally meaningful to subdivide anti-RBD antibodies into distinct classes based on their structurally defined epitopes. These structurally defined classes are inherently approximate groupings since even antibodies with superficially similar structural epitopes bind to their antigens in subtly distinct ways[25]. Our comprehensive maps of binding-escape mutations capture these subtle differences and show that antibodies in the same structurally defined class can be differentially affected by the same mutation. Using the escape maps, we can visualize how the antibodies are related in terms of how their binding is functionally impacted by mutations at different RBD sites. These visualizations show that the arrangement of antibodies in the space of viral escape is indeed continuous, but that the class definitions based on structural analyses capture the high-level features of this arrangement since the structural footprint of an antibody largely determines which mutations most impact its binding. However, we suggest that in some cases, more continuous visualizations of the arrangements of antibodies in the space of viral escape such as the ones we present here could have benefits over structural classification schemes for analyzing the impacts of viral mutations.

There are several limitations to our study. First, we did not examine the effects of mutations or deletions to the NTD, which can also affect neutralization[7,9,22–24,26]. Our experiments assayed binding of antibodies to monomeric yeast-expressed RBD, and so cannot capture mutational effects on spike trimer conformation or antibodies with quaternary epitopes[25]. In addition, our experiments were designed to identify mutations with large effects on antibody binding, and may overlook mutations with more subtle effects. Finally, the N-linked glycans on yeast-expressed proteins are more mannose-rich than those on mammalian-expressed proteins[52].

Framing viral antigenic evolution in terms of different classes of antibody epitopes is useful for interpreting the impacts of viral mutations and forecasting where in the RBD new escape mutations may arise in the future. Consistent with our work showing that mutations at class 2 antibody epitopes generally cause the largest reductions in the binding of polyclonal anti-RBD plasma antibodies elicited by infection, many emerging SARS-CoV-2 lineages have already acquired a mutation (E484K) at the site that most potently escapes antibodies of this class. Once viruses have escaped class 2 antibodies, antibodies of other classes (i.e., class 1 and 3) will contribute to most remaining RBD-targeted antibody immunity. In this respect, it is noteworthy that some of the most prominent emerging viral lineages with E484K have also acquired a mutation (K417N/T in B.1.351 and P.1, respectively) at the site that most potently escapes binding by class 1 antibodies. Moreover, the B.1.427/429 and B.1.617.2 viral lineages contain a moderate class 3 escape mutation (L452R)[3,53], and B.1.617.1 contains both E484Q and L452R[54]. Other clusters of sequences containing mutations to class 2 and 3 epitopes (E484K, R346K) have been reported[55]. Fortunately, no major viral lineages currently contain mutations to all three epitopes. However, we suggest the appearance of such a variant would be a worrying development and should be monitored closely.

Methods

SARS-CoV-2 convalescent human plasma samples. Plasma samples were previously described and collected as part of a prospective longitudinal cohort study of SARS-CoV-2 convalescent individuals in New York, NY[21]. Plasma samples profiled in this study were obtained 21–35 days post-symptom onset[21]. Samples were obtained upon written consent from community participants under protocols approved by the Institutional Review Board of the Rockefeller University (DRO-1006). All samples were heat-inactivated prior to use by treatment at 56 °C for 60 min. Prior to use in each assay, plasma samples were centrifuged for 15 min at 2000 × *g* to pellet platelets.

Neutralizing monoclonal antibodies binding the SARS-CoV-2 spike RBD. Antibodies were isolated from individuals in the cohort above as previously described[21]. Briefly, the IGH, IGL, and IGK genes were sequenced using IgG-specific primers from single-cell sorted RBD+, CD20+ memory B cells (CD3−CD8−CD14−CD16−CD20+ Ova−RBD-PE+RBD-AF647+). Recombinant monoclonal antibodies were produced and purified, as previously described[56,57]. Antibodies were produced with a human IgG1 heavy chain and human IgK (C002, C110, C135) or human IgL2 (C105, C121, C144) constant regions. These antibodies were previously structurally characterized in complex with SARS-CoV-2 S trimer[25,29]. A subset was functionally characterized in refs. [7,58]. The PBD accessions for the antibody-S complex structures are 6XCM and 6XCN for C105, 7C01 for LY-CoV016, 7K8S and 7K8T for C002, 7K8X and 7K8Y for C121, 7K90 for C144, 7K8Z for C135, and 7K8V for C110[25,28,29]. The protein sequences for these antibodies are available at the aforementioned PDB accession numbers.

RBD deep mutational scanning library. Monoclonal antibody and polyclonal clonal plasma selection experiments were performed in biological duplicate using a deep mutational scanning approach[30] with previously described duplicate yeast-displayed mutant RBD libraries[32]. These libraries were generated in the RBD background of the SARS-CoV-2 isolate Wuhan-Hu-1 (Genbank accession number MN908947, residues N331-T531) via NNS codon tiling PCR mutagenesis, which introduced an average of 2.7 amino-acid mutations per library variant. RBD variants were linked to unique 16-nucleotide barcode sequences to facilitate downstream sequencing and bottlenecked to library sizes of ~100,000 uniquely barcoded variants. The libraries contain 3804 of the 3819 possible amino-acid mutations,

with >95% present as single mutants on at least one barcode in the libraries. We previously used these libraries to measure the effect of all RBD mutations on yeast-surface RBD expression and ACE2 affinity[32]. These libraries were sorted to eliminate variants that lose ACE2 binding prior to mapping the antibody-escape variants[30].

FACS sorting of yeast libraries to select mutants with reduced binding by polyclonal plasmas. Antibody labeling and selection were performed essentially as described by Greaney et al.[30]. Specifically, 9 OD aliquots of RBD libraries were thawed and grown overnight at 30 °C 275 rpm in 45 mL of SD-CAA (6.7 g/L Yeast Nitrogen Base, 5.0 g/L Casamino acids, 1.065 g/L MES, and 2% w/v dextrose). Libraries were diluted to an OD of 0.67 in SG-CAA + 0.1% dextrose (SD-CAA with 2% w/v galactose and 0.1% w/v dextrose in place of 2% dextrose), and incubated for 16–18 h at room temperature with mild agitation to induce RBD surface expression. For each antibody selection, 20 OD units of induced cells were washed twice with PBS–BSA (0.2 mg/mL), and incubated in 4 mL of PBS–BSA with monoclonal antibody or plasma for 1 h at room temperature with gentle agitation. Incubations were performed with 400 ng/mL for each monoclonal antibody (C105, C144, C002, C121, C135, or C110) or with a sub-saturating dilution of polyclonal plasma such that the amount of fluorescent signal due to plasma antibody binding to RBD was approximately equal across plasmas (COV-021, 1:500; COV-047, 1:200; COV-057, 1:50; COV-072, 1:200; COV-107, 1:80). Labeled cells were washed with ice-cold PBS–BSA followed by secondary labeling for 1 h at 4 °C in 2.5 mL of 1:200 PE-conjugated goat anti-human-IgG (Jackson ImmunoResearch 109-115-098) to label for bound monoclonal antibody or 1:200 Alexa-647-conjugated goat anti-human-IgA+IgG+IgM (Jackson ImmunoResearch 109-605-064) to label for bound plasma antibodies and 1:100 FITC-conjugated anti-Myc (Immunology Consultants Lab, CYMC-45F) to label for RBD surface expression. Labeled cells were washed twice with PBS–BSA and resuspended in 2.5 mL of PBS. Yeast expressing the unmutated SARS-CoV-2 RBD were prepared in parallel to library samples, labeled at the same 400 ng/mL and 100× reduced 4 ng/mL antibody concentrations for the monoclonal antibodies, and with 1× and 10× reduced plasma concentrations for the polyclonal plasmas.

Yeast cells expressing RBD variants with substantially reduced antibody binding were selected via fluorescence-activated cell sorting (FACS) on a BD FACSAria II, using FACSDiva software, version 6.1. For monoclonal antibody selections, FACS selection gates were drawn to capture 95% of yeast expressing unmutated SARS-CoV-2 RBD labeled at 100-fold reduced antibody concentration relative to library samples. For polyclonal plasma selections, FACS selection gates were drawn to capture 2.8–5% of the RBD mutants with the lowest amount of plasma binding for their degree of RBD expression (Fig. S1A–C). Nearly zero (<0.1%) and 0.2 to 27.2% of cells expressing unmutated RBD fell into this gate when stained with 1× and 0.1× the concentration of plasma, respectively. For each sample, approximately 10 million RBD + cells (range 8.7e6–1.5e7 cells) were processed on the cytometer, with between 1.5e6 and 2.0e6 monoclonal antibody-escaped cells and 3.2e5 and 5.3e5 plasma-escaped cells collected per sample. Antibody-escaped cells collected per sample into SD-CAA supplemented with 1% w/v BSA and grown overnight in 1.5mL SD-CAA + 100 U/mL penicillin + 100 μg/mL streptomycin at 30 °C 275 rpm.

DNA extraction and Illumina sequencing. Plasmid samples were prepared from up to 7.5 OD units (4e7 CFUs) of overnight cultures of antibody-escaped cells, and 30 OD units (1.6e8 CFUs) of pre-selection yeast populations (Zymoprep Yeast Plasmid Miniprep II) per the manufacturer instructions, with the addition of a −80 °C freeze–thaw step prior to cell lysis. The 16-nucleotide barcode sequences identifying each RBD variant were amplified by PCR and prepared for Illumina sequencing exactly as described previously[32]. Specifically, a primer with the sequence 5′-AATGATACGGCGACCACCGAGA-3′ was used to anneal to the Illumina P5 adaptor sequence, and the PerkinElmer NextFlex DNA Barcode adaptor primers with the sequence 5′-CAAGCAGAAGACGGCATACGA GATxxxxxxxxGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3′ (where xxxxxxxx indicates the sample index sequence) were used to anneal to the Illumina P7 adaptor sequence and append sample indexes for sample multiplexing. Barcodes were sequenced via 50 bp single-end reads on an Illumina HiSeq 2500, targeting at least 2.5× as many sequencing reads as FACS-selected cells, and pre-sort reference populations of at least 2.5e7 reads.

Analysis of mutant library deep sequencing and computation of per-mutant escape fractions. Escape fractions were computed as described in Greaney et al.[30], with minor modifications as noted below. Specifically, we used the dms_variants package (https://jbloomlab.github.io/dms_variants/, version 0.8.5) to process Illumina sequences into counts of each barcoded RBD variant in each pre-sort and antibody-escape population using the barcode/RBD look-up table from Starr et al.[32]. Markdown renderings of these steps in the computational analysis are at https://github.com/jbloomlab/SARS-CoV-2-RBD_MAP_Rockefeller/blob/main/results/summary/aggregate_variant_counts.md and https://github.com/jbloomlab/SARS-CoV-2-RBD_MAP_Rockefeller/blob/main/results/summary/counts_to_cells_ratio.md.

For each antibody selection, we then computed the escape fraction for each barcoded variant using the deep sequencing counts for each variant in the original and antibody-escape populations and the total fraction of the library that escaped antibody binding via the formula provided in Greaney et al.[30]. These escape fractions represent the estimated fraction of cells expressing that specific variant that falls in the antibody escape bin, so a value of 0 means the variant is always bound by antibody and a value of 1 means that it always escapes antibody binding. We then applied a computational filter to remove variants with low sequencing counts or highly deleterious mutations that might cause antibody escape simply by leading to poor expression of properly folded RBD on the yeast cell surface. Specifically, we ignored all variants with pre-selection sequencing counts that were lower than the counts for the 99th percentile of the stop-codon-containing variants because stop-codon variants are largely purged by the earlier sorts for RBD expressing and ACE2-binding variants and so any residual presence provides an indication of low-count noise. Next, we removed any variants that had poor RBD expression or ACE2 binding, or contained mutations that individually cause poor RBD expression and ACE2 binding to eliminate misfolded or non-expressing RBDs. Specifically, we removed variants that had (or contained mutations with) ACE2-binding scores < −2.35 or expression scores < −1, using the variant- and mutation-level deep mutational scanning scores[32]. Note that these filtering criteria are slightly more stringent than those used in Greaney et al. (2021)[30] but are identical to those used in Greaney et al. and Starr et al[18,33]. The ACE2-binding cutoff of −2.35 is used to represent the binding of RaTG13 to human ACE2[32], which possesses the lowest known affinity capable of mediating cell entry[59]. The RBD expression cutoff of −1 is used to eliminate mutations that have as large an expression deficit as mutations to core disulfide residues. 2034 of the 3819 possible RBD amino-acid mutations passed these filtering steps and were included in our escape maps. All previously reported escape-mapping data[18,30,33,36] were reanalyzed in this study with the parameters listed above. A markdown rendering of the computation of the variant-level escape fractions and the variant filtering is at https://github.com/jbloomlab/SARS-CoV-2-RBD_MAP_Rockefeller/blob/main/results/summary/counts_to_scores.md.

Because some library variants contain multiple amino-acid mutations, we next deconvolved variant-level escape scores into escape fraction estimates for single mutations using global-epistasis models[60] implemented in the dms_variants package, as detailed at (https://jbloomlab.github.io/dms_variants/dms_variants.globalepistasis.html). In this fitting, we excluded variants that contained mutations not seen as either single mutants or in at least two multiple-mutant variants. We then computed the estimated effect of each mutation as the impact of that mutation on the observed phenotype scale transformation of its latent phenotype as computed using the global-epistasis models, and applied a floor of zero and a ceiling of 1 to these escape fractions. All of the above analysis steps were performed separately for each of the duplicate mutant libraries. We then only retained mutations that passed all of the above filtering and were measured in both libraries or had at least two single-mutant variant measurements in one library. The reported scores throughout the paper are the average across the libraries; these scores are also in Supplementary Data 1. Correlations in final single-mutant escape scores are shown in Fig. S2. A markdown rendering of the computation that computes these mutation-level escape fractions is at https://github.com/jbloomlab/SARS-CoV-2-RBD_MAP_Rockefeller/blob/main/results/summary/scores_to_frac_escape.md.

For plotting and analyses that required identifying RBD sites of strong escape (e.g., choosing which sites to show in logo plots in Fig. 1a or Fig. 3b or label in Fig. 2b), we considered a site to mediate strong escape if the total escape (sum of mutation-level escape fractions) for that site exceeded the median across sites by >10-fold, and was at least 10% of the maximum for any site. A markdown rendering of the identification of these sites of strong escape is at https://github.com/jbloomlab/SARS-CoV-2-RBD_MAP_Rockefeller/blob/main/results/summary/call_strong_escape_sites.md.

Comparison of mutation escape fractions to previously measured neutralization concentrations. In Fig. 5c and Fig. S6, mutation-level antibody-escape fractions measured in this study are compared to previously measured neutralization titers (inhibitory concentration 50%, IC$_{50}$) of the same monoclonal antibodies and polyclonal plasmas against some RBD point-mutants[7]. The numerical IC50 values were extracted from figures in Weisblum et al.[7] using the WebPlotDigitizer tool v4.4 (https://apps.automeris.io/wpd/). We convert those previously measured IC50s to fold-change IC50 relative to wildtype RBD. The numerical IC50s found in Weisblum et al.[7] are also tabulated at https://github.com/jbloomlab/SARS-CoV-2-RBD_MAP_Rockefeller/blob/main/experimental_data/data/weisblum_ic50.csv.

Analysis of mutations in circulating human SARS-CoV-2 strains. For the analysis in Fig. 6, all 765,455 spike sequences on GISAID[38] as of May 11, 2021 were downloaded and aligned via mafft[61], version 7.471. Sequences from non-human origins and sequences containing gap or ambiguous characters were removed, as were sequenced with extremely high numbers of RBD mutations relative to other sequences, leaving 679,502 retained sequences. All RBD amino-acid mutations were enumerated compared to the reference Wuhan-Hu-1 SARS-CoV-2 RBD sequence (Genbank MN908947, residues N331-T531). We acknowledge all contributors to the GISAID EpiCoV database for their sharing of sequence data (all contributors

listed at https://github.com/jbloomlab/SARS-CoV-2-RBD_MAP_Rockefeller/blob/main/data/gisaid_hcov-19_acknowledgement_table.pdf).

Data visualization. The static logo plots in the paper were created using dmslogo (https://jbloomlab.github.io/dmslogo/) version 0.6.2; a markdown rendering of the code that creates these logo plots is at https://github.com/jbloomlab/SARS-CoV-2-RBD_MAP_Rockefeller/blob/main/results/summary/escape_profiles.md. For each plasma, the y axis is scaled to be the greatest of (a) the maximum site-wise escape metric observed for that plasma, (b) 20× the median site-wise escape fraction observed across all sites for that plasma, or (c) an absolute value of 1.0 (to appropriately scale plasmas that are not "noisy" but for which no mutation has a strong effect on plasma binding).

In Fig. S4, mutations are colored by prior deep mutational scanning measurements of yeast-displayed RBD ACE2 affinity and RBD expression from Starr et al.[32], which are available at https://github.com/jbloomlab/SARS-CoV-2-RBD_DMS/blob/master/results/single_mut_effects/single_mut_effects.csv.

The multidimensional scaling in Figs. 1c and 4a that projects the antibodies into a two-dimensional space of escape mutations was performed using the Python scikit-learn package, version 0.23.2. We computed the similarity and dissimilarity in the escape maps between each pair of antibodies, then performed metric multidimensional scaling with two components on the dissimilarity matrix exactly as defined in Greaney et al.[30]. The Kruskal stress (i.e., stress(1) =

$$\sqrt{\sum_{h,i} \left(d_{hi} - \hat{d}_{hi} \right)^2 / \sum_{h,i} d_{hi}^2} \quad (1)$$ where $d_{h,i}$ is the ordinated distance between samples h and i, and $\hat{d}$ is the distance predicted from the regression) for the multidimensional scaling is 0.212 in Fig. 1c and 0.186 in Fig. 4a. These are moderately large stress values (the Kruskal stress can range from 0 to 1), and indicate that the high dimensional differences among the antibodies can only be modestly well approximated in two dimensions. In Fig. 1c, the multidimensional scaling shows antibodies as pie charts colored proportionally to the total squared site escape that falls into that RBD structural region. The code that generates these logo plot visualizations is available at https://github.com/jbloomlab/SARS-CoV-2-RBD_MAP_Rockefeller/blob/main/results/summary/mds_escape_profiles.md.

The interactive visualizations of the escape maps and their projections on the RBD-antibody structures available at https://jbloomlab.github.io/SARS-CoV-2-RBD_MAP_Rockefeller/ were created using dms-view (https://dms-view.github.io/docs/)[62].

The static structural views (Figs. 2a and S5) in the paper were rendered in PyMOL using antibody-bound RBD structures. PBDs are as follows: C105, 6XCN; LY-CoV016, 7C01; C110, 7K8V; C135, 7K8Z; C144, 7K90; C002, 7K8S; C121, 7K8Y. In Fig. S5B, 7K8T is used instead of 7K8S to illustrate the quaternary antibody epitope. For identifying contact sites to highlight in Fig. 1b logo plots or to classify sites in Fig. 2b as contact sites (within 4A of antibody) or antibody-proximal sites within 4–8 A, the following PDBs were used: 6XCM and 6XCN for C105, 7K8S and 7K8T for C002, 7K8X and 7K8Y for C121, 7K90 for C144, 7K8Z for C135, and 7K8V for C110[25,29]. Structural distances were computed using the bio3d package version 2.4.0 in R version 3.6.2[63]. Surface representations of the RBD for non-antibody-bound structures utilize PDB 6M0J[64].

In many of the visualizations, the RBD sites are categorized by epitope region (class 1, class 2, or class 3), defined by Barnes et al.[25] and colored accordingly. We define the class 1 epitope as residues 403 + 405 + 406 + 417 + 420 + 421 + 453 + 455-460 + 473-476 + 486 + 487 + 489 + 504, the class 2 epitope to be residues 455 + 456 + 472 + 483-487 + 489 + 490 + 491 + 492 + 493 + 494, and the class 3 epitope to be residues 345 + 346 + 437-452 + 496 + 498-501. There are five residues that overlap between the class 1 and class 2 epitopes: 455 + 456 + 486 + 487 + 489.

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

Data availability

The raw Illumina sequencing data generated in this study have been deposited in the NCBI SRA database under accession code BioProject: PRJNA639956, BioSample SAMN18148595. The data are available without restriction. The processed Illumina sequencing counts for the escape mapping are available at https://github.com/jbloomlab/SARS-CoV-2-RBD_MAP_Rockefeller/blob/main/results/counts/variant_counts.csv. The raw data tables of mutant escape fractions are in Supplementary Data 1: https://github.com/jbloomlab/SARS-CoV-2-RBD_MAP_Rockefeller/blob/main/results/supp_data/all_samples_raw_data.csv. These data are the escape fractions shown in Figs. 1–6 and Figs. S2–S6. The following publicly available datasets were used in the analysis in this study: The PDB accessions for the antibody-S complex structures are 6XCM and 6XCN for C105, 7C01 for LY-CoV016, 7K8S and 7K8T for C002, 7K8X and 7K8Y for C121, 7K90 for C144, 7K8Z for C135, and 7K8V for C110[25,28,29]. The protein sequences for these antibodies are available at the aforementioned PDB accession numbers. Surface representations of the RBD for non-antibody-bound structures utilize PDB 6M0J[64]. The unmutated SARS-CoV-2 RBD sequence is that from isolated Wuhan-Hu-1 (Genbank accession number MN908947, residues N331-T531). The analysis in Fig. 6 uses all 765,455 spike sequences on the GISAID EpiCoV database (https://www.gisaid.org/) as of

May 11, 2021. Together, these are the minimal set of data required to replicate the analysis. All data are available without restriction.

References

1. Faria, N. R. et al. Genomic characterisation of an emergent SARS-CoV-2 lineage in Manaus: preliminary findings. *Science*. 372, 815–821 (2021).
2. Voloch, C. M. et al. Genomic characterization of a novel SARS-CoV-2 lineage from Rio de Janeiro, Brazil. *J. Virol.* https://doi.org/10.1128/JVI.00119-21 (2021).
3. Zhang, W. et al. Emergence of a novel SARS-CoV-2 variant in Southern California. *J. Am. Med. Assoc.* 325, 1324–1326 (2021).
4. West, A. P. et al. Detection and characterization of the SARS-CoV-2 lineage B.1.526 in New York. Preprint at *bioRxiv* https://doi.org/10.1101/2021.02.14.431043 (2021).
5. Annavajhala, M. K. et al. A novel SARS-CoV-2 variant of concern, B.1.526, identified in New York. Preprint at *medRxiv* https://doi.org/10.1101/2021.02.23.21252259 (2021).
6. Tegally, H. et al. Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature* 592, 438–443 (2021).
7. Weisblum, Y. et al. Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants. *eLife* 9, e61312 (2020).
8. Liu, Z. et al. Identification of SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. *Cell Host Microbe* https://doi.org/10.1016/j.chom.2021.01.014 (2021).
9. Andreano, E. et al. SARS-CoV-2 escape in vitro from a highly neutralizing COVID-19 convalescent plasma. Preprint at *bioRxiv* https://doi.org/10.1101/2020.12.28.424451 (2020).
10. Wang, Z. et al. mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. *Nature* 592, 616–622 (2021).
11. Cele, S. et al. Escape of SARS-CoV-2 501Y.V2 from neutralization by convalescent plasma. *Nature* 593, 142–146 (2021).
12. Wibmer, C. K. et al. SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. *Nat. Med.* 27, 622–625 (2021).
13. Chen, R. E. et al. Resistance of SARS-CoV-2 variants to neutralization by monoclonal and serum-derived polyclonal antibodies. *Nat. Med.* 27, 717–726 (2021).
14. Planas, D. et al. Sensitivity of infectious SARS-CoV-2 B.1.1.7 and B.1.351 variants to neutralizing antibodies. *Nat. Med.* 27, 917–924 (2021).
15. Kuzmina, A. et al. SARS-CoV-2 spike variants exhibit differential infectivity and neutralization resistance to convalescent or post-vaccination sera. *Cell Host Microbe* 29, 522–528 (2021).
16. Wang, P. et al. Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. *Nature* 593, 130–135 (2021).
17. Piccoli, L. et al. Mapping neutralizing and immunodominant sites on the SARS-CoV-2 spike receptor-binding domain by structure-guided high-resolution serology. *Cell* 183, 1024–1042.e21 (2020).
18. Greaney, A. J. et al. Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. *Cell Host Microbe* 29, 463–476 (2021).
19. Dejnirattisai, W. et al. The antigenic anatomy of SARS-CoV-2 receptor binding domain. *Cell* 184, 2183–2200.e22 (2021).
20. Steffen, T. L. et al. The receptor binding domain of SARS-CoV-2 spike is the key target of neutralizing antibody in human polyclonal sera. Preprint at *bioRxiv* https://doi.org/10.1101/2020.08.21.261727 (2020).
21. Robbiani, D. F. et al. Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* 584, 437–442 (2020).
22. Kemp, S. et al. Recurrent emergence and transmission of a SARS-CoV-2 Spike deletion ΔH69/V70. Preprint at *bioRxiv* https://doi.org/10.1101/2020.12.14.422555 (2020).

23. Voss, W. N. et al. Prevalent, protective, and convergent IgG recognition of SARS-CoV-2 non-RBD spike epitopes. *Science* **372**, 1108–1112 (2021).

24. McCarthy, K. R. et al. Recurrent deletions in the SARS-CoV-2 spike glycoprotein drive antibody escape. *Science* **371**, 1139–1142 (2021).

25. Barnes, C. O. et al. SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* **588**, 682–687 (2020).

26. Liu, L. et al. Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. *Nature* **584**, 450–456 (2020).

27. 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).

28. Shi, R. et al. A human neutralizing antibody targets the receptor-binding site of SARS-CoV-2. *Nature* **584**, 120–124 (2020).

29. Barnes, C. O. et al. Structures of human antibodies bound to SARS-CoV-2 spike reveal common epitopes and recurrent features of antibodies. *Cell* **182**, 828–842 (2020).

30. Greaney, A. J. et al. Complete mapping of mutations to the SARS-CoV-2 spike receptor-binding domain that escape antibody recognition. *Cell Host Microbe* **29**, 44–57 (2021).

31. Zost, S. J. et al. Potently neutralizing and protective human antibodies against SARS-CoV-2. *Nature* **584**, 443–449 (2020).

32. Starr, T. N. et al. Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding. *Cell* **182**, 1295–1310 (2020).

33. Starr, T. N. et al. Prospective mapping of viral mutations that escape antibodies used to treat COVID-19. *Science* **371**, 850–854 (2021).

34. Smith, D. J. et al. Mapping the antigenic and genetic evolution of influenza virus. *Science* **305**, 371–376 (2004).

35. Dong, J. et al. Genetic and structural basis for recognition of SARS-CoV-2 spike protein by a two-antibody cocktail. Preprint at bioRxiv https://doi.org/10.1101/2021.01.27.428529 (2021).

36. Starr, T. N., Greaney, A. J., Dingens, A. S. & Bloom, J. D. Complete map of SARS-CoV-2 RBD mutations that escape the monoclonal antibody LY-CoV555 and its cocktail with LY-CoV016. *Cell Rep. Med.* **2**, 100255 (2021).

37. Cunningham, B. C. & Wells, J. A. Comparison of a structural and a functional epitope. *J. Mol. Biol.* **234**, 554–563 (1993).

38. Elbe, S. & Buckland-Merrett, G. Data, disease and diplomacy: GISAID's innovative contribution to global health. *Glob. Chall.* **1**, 33–46 (2017).

39. Shen, X. et al. SARS-CoV-2 variant B.1.1.7 is susceptible to neutralizing antibodies elicited by ancestral spike vaccines. *Cell Host Microbe* **29**, 529–539 (2021).

40. Wang, L. et al. Antibodies with potent and broad neutralizing activity against antigenically diverse and highly transmissible SARS-CoV-2 variants. Preprint at bioRxiv https://doi.org/10.1101/2021.02.25.432969 (2021).

41. Angeletti, D. et al. Defining B cell immunodominance to viruses. *Nat. Immunol.* **18**, 456–463 (2017).

42. Angeletti, D. & Yewdell, J. W. Understanding and manipulating viral immunity: antibody immunodominance enters center stage. *Trends Immunol.* **39**, 549–561 (2018).

43. Greaney, A. J. et al. The SARS-CoV-2 mRNA-1273 vaccine elicits more RBD-focused neutralization, but with broader antibody binding within the RBD. Preprint at bioRxiv https://doi.org/10.1126/scitranslmed.abi9915 (2021).

44. Chen, R. E. et al. Resistance of SARS-CoV-2 variants to neutralization by monoclonal and serum-derived polyclonal antibodies. *Nat. Med.* https://doi.org/10.1038/s41591-021-01294-w (2021).

45. Yuan, M. et al. Structural basis of a shared antibody response to SARS-CoV-2. *Science* **369**, 1119–1123 (2020).

46. Cao, Y. et al. Potent neutralizing antibodies against SARS-CoV-2 identified by high-throughput single-cell sequencing of convalescent patients' B cells. *Cell* **182**, 73–84 (2020).

47. Kreer, C. et al. Longitudinal isolation of potent near-germline SARS-CoV-2-neutralizing antibodies from COVID-19 patients. *Cell* **182**, 843–854 (2020).

48. Yuan, M. et al. Structural and functional ramifications of antigenic drift in recent SARS-CoV-2 variants. *Science* https://doi.org/10.1126/science.abh1139 (2021).

49. Brouwer, P. J. M. et al. Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. *Science* **369**, 643–650 (2020).

50. Wec, A. Z. et al. Broad neutralization of SARS-related viruses by human monoclonal antibodies. *Science* **369**, 731–736 (2020).

51. 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).

52. Hamilton, S. R. et al. Production of complex human glycoproteins in yeast. *Science* **301**, 1244–1246 (2003).

53. Deng, X. et al. Transmission, infectivity, and neutralization of a spike L452R SARS-CoV-2 variant. *Cell* https://doi.org/10.1016/j.cell.2021.04.025 (2021).

54. Public Health England. *SARS-CoV-2 Variants of Concern and Variants Under Investigation in England*. (Public Health England, 2021).

55. de Oliveira, T. et al. A novel variant of interest of SARS-CoV-2 with multiple spike mutations detected through travel surveillance in Africa. Preprint at bioRxiv https://doi.org/10.1101/2021.03.30.21254323 (2021).

56. Klein, F. et al. Enhanced HIV-1 immunotherapy by commonly arising antibodies that target virus escape variants. *J. Exp. Med.* **211**, 2361–2372 (2014).

57. Schoofs, T. et al. Broad and potent neutralizing antibodies recognize the silent face of the HIV envelope. *Immunity* **50**, 1513–1529 (2019).

58. 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).

59. Letko, M., Marzi, A. & Munster, V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B beta coronaviruses. *Nat. Microbiol.* **5**, 562–569 (2020).

60. Otwinowski, J., McCandlish, D. M. & Plotkin, J. B. Inferring the shape of global epistasis. *Proc. Natl Acad. Sci. USA* **115**, E7550–E7558 (2018).

61. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evolution* **30**, 772–780 (2013).

62. Hilton, S. et al. dms-view: interactive visualization tool for deep mutational scanning data. *J. Open Source Softw.* **5**, 2353 (2020).

63. Grant, B. J., Rodrigues, A. P. C., ElSawy, K. M., McCammon, J. A. & Caves, L. S. D. Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics* **22**, 2695–2696 (2006).

64. Lan, J. et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**, 215–220 (2020).

Author contributions

Conceptualization: A.J.G., C.O.B., M.C.N., P.J.B. and J.D.B.; methodology: A.J.G., T.N.S. and J.D.B.; investigation: A.J.G.; software: A.J.G., T.N.S. and J.D.B.; formal analysis: A.J.G. and J.D.B.; VSV escape data: Y.W., F.S. and D.P.; resources: M.C.N., D.F.R., M.C., C.G., A.C., M.A., S.F., Z.W. and F.M.; writing—original draft: A.J.G. and J.D.B.; writing—review and editing: all authors; supervision: T.H., P.D.B., M.C.N., P.J.B. and J.D.B.

Competing interests

Subsequent to completion and submission of the initial version of this study, J.D.B. began consulting for Moderna on viral evolution and epidemiology. J.D.B. has the potential to receive a share of IP revenue as an inventor on a Fred Hutch optioned technology/patent (application WO2020006494) related to deep mutational scanning of viral proteins. The Rockefeller University has filed a provisional patent application related to SARS-CoV-2 monoclonal antibodies on which D.F.R. and M.C.N. are inventors. The Rockefeller University has applied for a patent relating to the replication-competent VSV/SARS-CoV-2 chimeric virus on which Y.W, F.S., T.H. and P.B. are inventors (US patent 63/036,124). The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-24435-8.

Correspondence and requests for materials should be addressed to J.D.B.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

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.