Comprehensive Epitope Mapping of Broad Sarbecovirus Neutralizing Antibodies

Yunlong Cao[1,2,#,*], Ayijiang Yisimayi[2,3,#], Fanchong Jian[2,4,#], Tianhe Xiao[2,5,#], Weiliang Song[2,3,#], Jing Wang[2,3,#], Shuo Du[3,#], Zhiying Zhang[3,#], Pulan Liu[3,#], Xiaohua Hao[6,#], Qianqian Li[7,#], Peng Wang[1], Ran An[2], Yao Wang[1], Jing Wang[1], Haiyan Sun[1], Lijuan Zhao[1], Wen Zhang[6], Dong Zhao[6], Jiang Zheng[1], Lingling Yu[1], Can Li[1], Na Zhang[1], Rui Wang[1], Xiao Niu[2], Sijie Yang[2,8], Xuetao Song[1], Linlin Zheng[1], Zhiqiang Li[8,9], Qingqing Gu[1], Fei Shao[1], Weijin Huang[7], Youchun Wang[7], Ronghua Jin[6,*], Junyu Xiao[3,8,9,*], Xiaoliang Sunney Xie[1,2,*]

[1]Changping Laboratory, Beijing, P.R. China.

[2]Biomedical Pioneering Innovation Center (BIOPIC), Peking University, Beijing, P.R. China.

[3]School of Life Sciences, Peking University, Beijing, P.R. China.

[4]Joint Graduate Program of Peking-Tsinghua-NIBS, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China.

[5]College of Chemistry and Molecular Engineering, Peking University, Beijing, P.R. China.

[6]Beijing Ditan Hospital, Capital Medical University, Beijing, P.R. China.

[7]Division of HIV/AIDS and Sex-transmitted Virus Vaccines, Institute for Biological Product Control, National Institutes for Food and Drug Control (NIFDC), Beijing, P.R. China.

[8]Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, P.R. China

[9]Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, P.R. China.

[*]Correspondence: Yunlong Cao (yunlongcao@pku.edu.cn); Ronghua Jin (ronghuajin@ccmu.edu.cn); Junyu Xiao (junyuxiao@pku.edu.cn); Xiaoliang Sunney Xie (sunneyxie@biopic.pku.edu.cn)

[#]These authors contributed equally.

Abstract

Constantly emerging SARS-CoV-2 variants, such as Omicron BA.1, BA.1.1 and BA.2, pose a severe challenge to COVID-19 control[1-10]. Broad-spectrum antibody therapeutics and vaccines are needed for defending against future SARS-CoV-2 variants and sarbecovirus pandemics[11-14]; however, we have yet to gain a comprehensive understanding of the epitopes capable of inducing broad sarbecovirus neutralization. Here, we report the identification of 241 anti-RBD broad sarbecovirus neutralizing antibodies isolated from 44 SARS-CoV-2 vaccinated SARS convalescents. Neutralizing efficacy of these antibodies against D614G, SARS-CoV-1, Omicron variants (BA.1, BA.1.1, BA.2), RATG13 and Pangolin-GD is tested, and their binding capability to 21 sarbecovirus RBDs is measured. High-throughput yeast-display mutational screening was further applied to determine each antibody's RBD escaping mutation profile, and unsupervised epitope clustering based on escaping mutation hotspots was performed[7,15-18]. A total of 6 clusters of broad sarbecovirus neutralizing antibodies with diverse breadth and epitopes were identified, namely Group E1 (S309[19], BD55-3152 site), E3 (S2H97[20] site), F1 (CR3022[21], S304[22] site), F2 (DH1047[23], BD55-3500 site), F3 (ADG-2[24], BD55-3372 site) and B' (S2K146[25] site). Members of E1, F2 and F3 demonstrate the highest neutralization potency; yet, Omicron, especially BA.2, has evolved multiple mutations (G339D, N440K, T376A, D405N, R408S) to escape antibodies of these groups. Nevertheless, broad sarbecovirus neutralizing antibodies that survived Omicron would serve as favorable therapeutic candidates. Furthermore, structural analyses of selected drug candidates propose two non-competing antibody pairing strategies, E1-F2 and E1-F3, as broad-spectrum antibody cocktails. Together, our work provides a comprehensive epitope map of broad sarbecovirus neutralizing antibodies and offers critical instructions for designing broad-spectrum vaccines.

61 **Main**

62 The recent emergence and global spreading of severe acute respiratory syndrome
63 coronavirus 2 (SARS-CoV-2) variant Omicron has posed a critical challenge to the
64 efficacy of COVID-19 vaccines and neutralizing antibody therapeutic[2-5]. Due to
65 multiple mutations to the spike protein and its receptor-binding domain (RBD),
66 Omicron can cause severe neutralizing antibody evasion[5-9]. The continuous evolution
67 of Omicron, such as BA.1.1 and BA.2, may lead to an even stronger immune escape[2,8,10].
68 For pandemic preparedness, pan-sarbecovirus therapeutics and vaccines that are
69 effective against future SARS-CoV-2 variants and other sarbecoviruses are urgently
70 needed[11-14]. Several broad sarbecovirus neutralizing antibodies have been reported[20,23-
71 34]; however, since the number of antibodies studied has been limited, we still lack a
72 comprehensive understanding of the RBD epitopes that could induce broad
73 sarbecovirus neutralizing antibodies.

74 SARS convalescents who were subsequently vaccinated against SARS-CoV-2 may
75 carry a large number of broad-spectrum neutralizing antibodies[19,23,35]. To study their
76 antibody repertoire, we recruited 44 SARS convalescents and collected their plasma
77 and peripheral blood mononuclear cells (PBMCs) after SARS-CoV-2 vaccination
78 (Supplementary Table 1). The plasma of these individuals exhibited high neutralization
79 ability against both SARS-CoV-1 and SARS-CoV-2 pseudovirus (Extended Data Fig
80 1a,b). Using flow cytometry, we sorted the memory B cells that could cross-bind to
81 both SARS-CoV-1 and SARS-CoV-2 RBD (Extended Data Fig. 2). Then, we used high-
82 throughput single-cell VDJ sequencing to obtain the B-cell repertoire of these memory
83 B cells[36,37], and expressed the selected IgG antibodies *in vitro* as human IgG1
84 monoclonal antibodies. In total, we obtained 1409 SARS-CoV-1 and SARS-CoV-2
85 RBD binding IgG1 antibodies, where 269 showed SARS-CoV-1 and SARS-CoV-2
86 cross-neutralizing capability (Supplementary Table 2).

87 To study the epitope distribution of those bivalent antibodies, we analyzed their
88 escaping mutation profiles using high-throughput yeast display mutant screening that

covers all possible single residue substitutions in the wildtype RBD background[7,15-18]. We successfully retained the escaping mutation profiles of 241 SARS-CoV-1/SARS-CoV-2 cross-neutralizing antibodies as well as 85 non-neutralizing bivalent binding antibodies. Combined with previously characterized antibodies isolated from SARS-CoV-2 convalescents and vaccinees[36,38,39], we collected a library of 715 human SARS-CoV-2 IgG1 antibodies with corresponding escaping mutation profiles (Supplementary Data 1). Neutralizing efficacy of these antibodies against sarbecovirus, including D614G, SARS-CoV-1, Omicron variants (BA.1, BA.1.1, BA.2), RATG13 and Pangolin-GD, is tested, and their binding capability to 21 sarbecovirus RBD is measured through ELISA (Supplementary Table 2-3). Based on the escaping mutation profiles, the antibodies could be unsupervised clustered into 10 epitope groups (Fig. 1a, Extended Data Fig. 3, Supplementary Data 2), as described previously[7]. Group A-D recapitulates our previous antibody clustering[7], in which the members mainly target the ACE2-binding motif[11,22,40-42]. Here, the larger collection of bivalent binding antibodies expands the previous Group E and F into E1-E3 and F1-F3, which respectively target the front and backside of RBD (Fig. 1b).

By projecting each antibody's experimental measurements against each sarbecovirus onto the t-distributed stochastic neighbor embedding (t-SNE) dimension, we surprisingly found that antibodies of the same cluster have unified sarbecovirus neutralization potency and binding spectra (Fig. 1c-f, Extended Data Fig. 4a-e). Moreover, the antibodies' neutralization mechanism also tends to cluster based on the epitope distribution. Antibodies of group E1-E3 and F1 do not compete with ACE2 (Fig. 1g), which results in relatively weaker neutralizing potency compared to ACE2-blocking antibody (Extended Data Fig. 5a), except that group E1 antibodies still possess high neutralization potency (Fig. 1c), suggesting a rather unique neutralization mechanism. In total, 6 clusters of antibodies were found to exhibit broad sarbecovirus neutralizing ability with diverse breadth, namely Group E1, E3, F1, F2, F3 and a subcluster of Group B (Fig 1a, 1e-f).

117 Group E1 antibodies display potent neutralizing activities against human infecting
118 sarbecoviruses, including both SARS-CoV-2 variants and SARS-CoV-1 variants (Fig.
119 1c, Extended Data Fig. 5b). Their epitope is fully exposed regardless of the up and
120 down conformations of RBD and may involve a mixed protein and carbohydrate,
121 specifically the N-linked glycan on N343, as for S309[19]. Indeed this unique feature is
122 shared among Group E antibodies, evidenced by structural analyses of BD55-3152,
123 BD55-3546 and BD55-5840 in complex with spike proteins using cryo-electron
124 microscopy (cryo-EM) (Fig. 2a, Extended Data Fig. 6a-b). Importantly, members of
125 Group E1 are generally sensitive to the changes of G339, E340, T345 and especially
126 R346, revealed by their escaping mutation profiles and structures (Fig. 2a, d). Thus,
127 most E1 antibodies could not bind to clade 2 and 3 sarbecoviruses because of the
128 changes of RBD antigenic sites corresponding to G339, E340 and R346, as calculated
129 by multiple sequence alignment (MSA) (Fig. 1f, Fig. 2d). Importantly, Omicron causes
130 considerable antibody escaping for Group E1 (Fig. 1d); however, a proportion of E1
131 antibodies that could tolerate G339D and N440K mutations could retain potent
132 neutralizing ability against BA.1, BA.1.1, and BA.2, making them good therapeutics
133 candidates (Extended Data Fig. 7).
134
135 Interestingly, despite the importance of R346 to E1 antibodies, the additional R346K
136 carried by BA.1.1 does not readily affect their efficacy (Fig. 1d). This is foreseeable, as
137 Arg and Lys possess similar chemical properties. In fact, SARS-CoV-1 features a Lys
138 at the corresponding site (K333) (Fig. 2d). The structure of BD55-3152 in complex with
139 the Omicron spike reveals that a CDRL2 Asp (D50) interacts with R346 (Fig. 2a),
140 whereas the structure of BD55-3152 complexed with the SARS-CoV-1 spike shows
141 that the same Asp also coordinates K333 (Extended Data Fig. 8). Nevertheless, most
142 E1 antibodies lost binding and neutralization ability towards clade 1a/1b sarbecoviruses
143 circulating among animals, including the bat coronavirus RaTG13 and the pangolin
144 coronavirus Pangolin-GD (Fig. 1e, Extended Data Fig. 5f-g). This is largely because of
145 the Thr substitution in these sarbecoviruses (Fig. 2d). In contrast to R346K, R346S/T
146 would greatly compromise the binding activities of E1 antibodies.

147

148 Group E2 antibodies are directed to the front chest of RBD. Previously, we solved cryo-
149 EM structures of several neutralizing antibodies in this group, including BD-744[43] (Fig.
150 2b). These neutralizing antibodies target a relatively flat region partially exposed in the
151 down RBDs. The E2 antibodies are sensitive to mutations of R346, A348, A352, K356,
152 R357, L452, and I468, and these sites are largely not conserved in clade 1a/2/3
153 sarbecoviruses (Fig. 2d). Therefore, the E2 Antibodies usually do not have broad
154 sarbecovirus neutralizing ability (Fig. 1e-f, Extended Data Fig. 4a-e). The L452R
155 mutation in the Delta variant also substantially abrogates their activities (Fig. 2d).
156 Nonetheless, most E2 antibodies can still neutralize Omicron BA.1, since Omicron
157 BA.1 do not have mutations at the sites mentioned above (Fig. 1d, Extended Data Fig.
158 5c-e); however, the R346K carried by BA.1.1 would escape E2 antibodies as well (Fig
159 2d), such as Brii-198.

160

161 Group E3 antibodies, such as S2H97[20], bind to the left flank of RBD (Fig. 2c). As this
162 epitope region is deeply buried within the spike trimer, the E3 antibodies can only bind
163 when the RBD adopts a not only up but also wide open conformation. Likely because
164 of this, the neutralizing activities of group E3 antibodies are relatively modest.
165 Critically sites recognized by E3 antibodies include R357, T393, Y396, D428, K462,
166 S514, E516 and L518, all well conserved in most sarbecoviruses (Fig. 2d); thus, the E3
167 antibodies display great neutralization breadth (Fig. 1f, Extended Data Fig. 4a-e).
168 Interestingly, though not escaped by Omicron variants, their overall neutralizing
169 activities are reduced (Fig. 1d). The neutralization mechanisms of S2H97 and other
170 antibodies targeting the cryptic sites in RBD have been partly attributed to their ability
171 to disrupt the spike trimer's prefusion state, as these antibodies demand the extensive
172 opening of the RBD to engage their respective binding sites[20,44-46]. Their reduced
173 activities towards Omicron could thus be caused by the more stabilized prefusion
174 structure of Omicron spike[47].

175

176 Group F1, F2, and F3 antibodies cover a continuous surface on the backside of RBD

177 and can only bind to the up RBDs (Fig. 1b). Members of the F1 group, such as
178 CR3022[21] and S304[22], also require a wide open RBD to engage and do not directly
179 block ACE2, therefore displaying weak neutralizing activities in general (Fig. 1c).
180 Escape hotspots for this group include SARS-CoV-2 residues 383-386, 390, and 391
181 (Fig. 3a, d). Most of these sites are conserved in all sarbecoviruses, except an alanine
182 on 384 in clade 1a, which is also tolerated (Fig. 3d). Due to their intrinsic weak
183 neutralizing activities and the triple mutations on S371, S373, S375 of Omicron, we
184 found that the F1 antibodies virtually have no neutralizing power against Omicron
185 variants (Fig. 1d, Extended Data Fig. 5c-e).
186
187 The epitopes for group F2 antibodies are shifted upward compared with F1 (Fig. 3b).
188 We solved the cryo-EM structure of two representative antibodies in this group, BD55-
189 1239 and BD55-3500, both in complexes with the Omicron spike (Fig. 3b, Extended
190 Data Fig. 6c). Structural analyses showed that although these antibodies do not directly
191 target the ACE2-binding residues, both project towards where ACE2 would be lodged
192 and therefore exclude ACE2 through steric hindrance, like DH1047[23] (Fig. 3b). Group
193 F2 antibodies can be escaped by RBD mutation involving T376, K378, and R408 (Fig.
194 3d). Indeed, these residues are all at the heart of BD55-1239's and BD55-3500's
195 epitopes. These sites are fairly conserved across sarbecoviruses (Fig. 3b, d), and the
196 neutralization breadth of group F2 antibodies is comparable to that of E3 and F1 (Fig.
197 1e-f). A notable exception is the corresponding K378Q substitution in BM48-31, which
198 might compromise the binding of many F2 antibodies (Extended Data Fig. 4d). Despite
199 the extraordinary breadth and potency against sarbecovirus, Group F2 antibodies suffer
200 great neutralization efficacy reduction against Omicron, mostly due to the triple
201 mutations on S371, S373, S375 (Fig 1d). Importantly, mutation T376A and R408S
202 harbored by Omicron BA.2 completely abolished the neutralizing capacity of F2
203 antibodies (Fig 1d, Extended Data Fig. 5e).
204
205 Group F3 Antibodies reach further towards the ACE2-binding site, such as S2X259[26]
206 and ADG-2[24] (Fig. 1b). Cryo-EM structures of BD55-4637 complexed with the

Omicron spike and BD55-3372 with the Delta spike reveal that F3 antibodies interact with several ACE2-binding residues (Fig. 3c, Extended Data Fig. 6d), and therefore directly compete with ACE2. Major escape sites for this group of antibodies include D405, R408, V503, G504, and Y508 (Fig. 3d). Given the fact that D405, G504, and Y508 are not conserved in clade 2 sarbecoviruses, F3 antibodies cannot bind to non-ACE2 utilizing clade 2 sarbecovirus (Extended Data Fig. 4e), but showed good neutralization breadth against clade 1a/b/2 sarbecovirus (Fig. 1e-f). Similar to Group E1, a proportion of F3 antibodies showed potent neutralization against Omicron BA.1, despite the Y505H mutation carried by Omicron (Fig. 1d, 3d). Moreover, several elite members of BA.1 tolerating F3 antibodies could also overcome the D405N and R408S mutations of BA.2, making them good therapeutic drug candidates (Fig 1d, Extended Data Fig. 7).

Notably, a unique subcluster of Group B antibodies also showed broad-spectrum sarbecovirus neutralizing capability (Fig 1c). Most Group B antibodies are SARS-CoV-2 specific since their major escaping mutations consist of E484 and F486, which are not conserved in sarbecovirus clade (Extended Data Fig. 9); however, the rare sub-cluster B', featured by S2K146[25], displayed skewed escaping mutation profiles toward N487 and Y489, which are highly conserved in clade 1a/1b/3 (Extended Data Fig. 9), making members of B' exhibits similar breadth as F3 antibodies. Sadly, most B' antibodies failed to neutralize Omicron, except for S2K146[6] (Fig. 1d).

The above analyses indicate that each epitope group's mutational escape hotspots are closely related to the antibodies' sarbecovirus reactivity breadth. To further extrapolate this observation, we simulated each antibody's sarbecovirus RBD binding spectrum based on its escaping mutation profile and the corresponding MSA results of sarbecovirus RBDs (Extended Data Fig. 10a-e). Surprisingly, the simulated spectrum well matches that obtained from ELISA (Extended Data Fig. 4a-e), suggesting that the antibody's breadth is mostly governed by the degree of conservation of its major mutational escaping sites, and the escaping mutation profile of an antibody could be

237 used to predict its sarbecovirus binding spectrum.

238

239 Importantly, even broad sarbecovirus neutralizing antibodies are largely escaped by
240 Omicron (Fig. 1b), which is consistent with that Omicron could greatly reduce the
241 neutralization efficacy of vaccinated SARS convalescents' plasma (Extended Data Fig.
242 1c). This supports the speculation that Omicron does not originate from zoonotic
243 spillover but immune selection pressure. Indeed, multiple top antibody-escaping
244 mutations inferred from yeast display have appeared on Omicron variants, including
245 K417N, E484A, Q490R, G446S, G339D, R346K, N440K, T376A, D405N and R408S
246 (Supplementary Data 3). Especially BA.2, which has evolved multiple mutations on the
247 sarbecovirus conserved amino acids (D405N, R408S and T376A) that could cause
248 pinpoint escapes of most F2 and partial F3 broad-spectrum neutralizing antibodies (Fig.
249 1d, Extended Data Fig. 3). Nevertheless, BA.2 lacks the G446S mutation, and thus part
250 of group D antibodies that target the linear 440-449 loop retained their neutralization
251 capability against BA.2, such as REGEN-1098748 and AZD106149 (Fig. 1d).

252

253 We have identified a large panel of potent sarbecovirus neutralizing antibodies that stay
254 effective to Omicron variants, mainly belonging to Group E1, F2 and F3. To further
255 examine their potential as antibody therapeutics, we analyzed these antibodies'
256 authentic virus neutralization potency against Omicron (BA.1) as well as their binding
257 affinity against sarbecovirus RBDs, along with well-recognized broad-spectrum
258 neutralizing antibody serving as controls (Fig. 4a, Supplementary Data 4). Several
259 antibodies, such as BD55-5840 (E1), BD55-3546 (E1), BD55-5640 (F2), BD55-5514
260 (F3) and BD55-5483 (F3), stand out as promising drug candidates (Fig. 4a).

261

262 In addition, pairing non-competing antibodies into cocktails is a promising strategy to
263 reduce the chance of mutation-induced antibody evasion. As the epitopes of the E1
264 antibodies are located on the opposite sides of RBD compared to F2 and F3, we
265 envision that many E1 antibodies can function in combinations with an F2 or F3
266 counterpart. Indeed, several plausible strategies can be designed based on the structural

267 information presented in this study. For example, overlaying the targeting RBDs in the
268 BD55-5840 (E1) and BD55-3500 (F2) cryo-EM structures demonstrate that they can
269 readily bind to the Omicron RBD simultaneously (Fig. 4b), whereas BD55-3546 (E1)
270 can bind concurrently with BD55-3372 (F3) on the Delta RBD (Fig. 4c). From a
271 mechanistic point of view, E1 antibodies do not directly impede ACE2 but can engage
272 both the up and down RBDs, whereas F2/F3 antibodies only bind to the up RBDs but
273 promptly blocks ACE2. Due to these complementary properties, they might synergize
274 in an ideal way to achieve maximized breadth and potency.
275
276 Our results could also guide the development of broad sarbecovirus vaccines.
277 According to our analyses, certain neutralizing antibodies, such as those belonging to
278 E1, F2, F3 epitope groups, should be the ideal inducing targets of broad sarbecovirus
279 vaccines, since their neutralization potency and breadth far exceeds the antibodies in
280 other epitope groups. Also, the overall mean escaping mutation profile for each epitope
281 group has great implications for rational vaccine antigen design (Supplementary Data
282 3). It could offer the exact mutation combination that should be used to avoid
283 stimulating a certain antibody group, such as a vaccine booster dose using a spike
284 protein with K417D/F456K/A475R mutations should completely block the stimulation
285 of Group A antibodies. Together, our results provide a comprehensive understanding of
286 the RBD epitopes that can induce broad sarbecovirus neutralizing antibodies, which
287 offers critical instructions for sarbecovirus antibody therapeutics and the antigen design
288 of broad-spectrum sarbecovirus vaccines.
289

290 **Reference**

291 1 Viana, R. *et al.* Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa.
292 *Nature*, doi:10.1038/s41586-022-04411-y (2022).
293 2 Lu, L. *et al.* Neutralization of SARS-CoV-2 Omicron variant by sera from BNT162b2 or Coronavac
294 vaccine recipients. *Clin Infect Dis*, doi:10.1093/cid/ciab1041 (2021).
295 3 Cele, S. *et al.* Omicron extensively but incompletely escapes Pfizer BNT162b2 neutralization.
296 *Nature*, doi:10.1038/s41586-021-04387-1 (2021).
297 4 Collie, S., Champion, J., Moultrie, H., Bekker, L. G. & Gray, G. Effectiveness of BNT162b2 Vaccine

298 against Omicron Variant in South Africa. *N Engl J Med* **386**, 494-496,
299 doi:10.1056/NEJMc2119270 (2022).
300 5 Dejnirattisai, W. *et al.* SARS-CoV-2 Omicron-B.1.1.529 leads to widespread escape from
301 neutralizing antibody responses. *Cell*, doi:10.1016/j.cell.2021.12.046 (2022).
302 6 Cameroni, E. *et al.* Broadly neutralizing antibodies overcome SARS-CoV-2 Omicron antigenic
303 shift. *Nature*, doi:10.1038/s41586-021-04386-2 (2021).
304 7 Cao, Y. *et al.* Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies.
305 *Nature*, doi:10.1038/s41586-021-04385-3 (2021).
306 8 Liu, L. *et al.* Striking Antibody Evasion Manifested by the Omicron Variant of SARS-CoV-2.
307 *Nature*, doi:10.1038/s41586-021-04388-0 (2021).
308 9 Planas, D. *et al.* Considerable escape of SARS-CoV-2 Omicron to antibody neutralization. *Nature*,
309 doi:10.1038/s41586-021-04389-z (2021).
310 10 Lyngse, F. P. *et al.* Transmission of SARS-CoV-2 Omicron VOC subvariants BA.1 and BA.2:
311 Evidence from Danish Households. *medRxiv* (2022).
312 11 Corti, D., Purcell, L. A., Snell, G. & Veesler, D. Tackling COVID-19 with neutralizing monoclonal
313 antibodies. *Cell* **184**, 3086-3108, doi:10.1016/j.cell.2021.05.005 (2021).
314 12 Cohen, A. A. *et al.* Mosaic nanoparticles elicit cross-reactive immune responses to zoonotic
315 coronaviruses in mice. *Science* **371**, 735-741, doi:10.1126/science.abf6840 (2021).
316 13 Martinez, D. R. *et al.* Chimeric spike mRNA vaccines protect against Sarbecovirus challenge in
317 mice. *Science* **373**, 991-998, doi:10.1126/science.abi4506 (2021).
318 14 Walls, A. C. *et al.* Elicitation of broadly protective sarbecovirus immunity by receptor-binding
319 domain nanoparticle vaccines. *Cell* **184**, 5432-5447 e5416, doi:10.1016/j.cell.2021.09.015
320 (2021).
321 15 Starr, T. N. *et al.* Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals
322 Constraints on Folding and ACE2 Binding. *Cell* **182**, 1295-1310 e1220,
323 doi:10.1016/j.cell.2020.08.012 (2020).
324 16 Greaney, A. J. *et al.* Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding
325 domain that affect recognition by polyclonal human plasma antibodies. *Cell Host Microbe* **29**,
326 463-476 e466, doi:10.1016/j.chom.2021.02.003 (2021).
327 17 Greaney, A. J. *et al.* Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding
328 Domain that Escape Antibody Recognition. *Cell Host Microbe* **29**, 44-57 e49,
329 doi:10.1016/j.chom.2020.11.007 (2021).
330 18 Starr, T. N., Greaney, A. J., Dingens, A. S. & Bloom, J. D. Complete map of SARS-CoV-2 RBD
331 mutations that escape the monoclonal antibody LY-CoV555 and its cocktail with LY-CoV016.
332 *Cell Rep Med* **2**, 100255, doi:10.1016/j.xcrm.2021.100255 (2021).
333 19 Pinto, D. *et al.* Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody.
334 *Nature* **583**, 290-295, doi:10.1038/s41586-020-2349-y (2020).
335 20 Starr, T. N. *et al.* SARS-CoV-2 RBD antibodies that maximize breadth and resistance to escape.
336 *Nature* **597**, 97-102, doi:10.1038/s41586-021-03807-6 (2021).
337 21 Yuan, M. *et al.* A highly conserved cryptic epitope in the receptor binding domains of SARS-
338 CoV-2 and SARS-CoV. *Science* **368**, 630-633, doi:10.1126/science.abb7269 (2020).
339 22 Piccoli, L. *et al.* Mapping Neutralizing and Immunodominant Sites on the SARS-CoV-2 Spike
340 Receptor-Binding Domain by Structure-Guided High-Resolution Serology. *Cell* **183**, 1024-1042
341 e1021, doi:10.1016/j.cell.2020.09.037 (2020).

342 23 Martinez, D. R. *et al.* A broadly cross-reactive antibody neutralizes and protects against
343 sarbecovirus challenge in mice. *Sci Transl Med* **14**, eabj7125,
344 doi:10.1126/scitranslmed.abj7125 (2022).
345 24 Rappazzo, C. G. *et al.* Broad and potent activity against SARS-like viruses by an engineered
346 human monoclonal antibody. *Science* **371**, 823-829, doi:10.1126/science.abf4830 (2021).
347 25 Park, Y. J. *et al.* Antibody-mediated broad sarbecovirus neutralization through ACE2 molecular
348 mimicry. *Science* **375**, 449-454, doi:10.1126/science.abm8143 (2022).
349 26 Tortorici, M. A. *et al.* Broad sarbecovirus neutralization by a human monoclonal antibody.
350 *Nature* **597**, 103-108, doi:10.1038/s41586-021-03817-4 (2021).
351 27 Lv, H. *et al.* Cross-reactive Antibody Response between SARS-CoV-2 and SARS-CoV Infections.
352 *Cell Rep* **31**, 107725, doi:10.1016/j.celrep.2020.107725 (2020).
353 28 Jette, C. A. *et al.* Broad cross-reactivity across sarbecoviruses exhibited by a subset of COVID-
354 19 donor-derived neutralizing antibodies. *Cell Rep* **36**, 109760,
355 doi:10.1016/j.celrep.2021.109760 (2021).
356 29 Liu, H. *et al.* A combination of cross-neutralizing antibodies synergizes to prevent SARS-CoV-2
357 and SARS-CoV pseudovirus infection. *Cell Host Microbe* **29**, 806-818 e806,
358 doi:10.1016/j.chom.2021.04.005 (2021).
359 30 Wec, A. Z. *et al.* Broad neutralization of SARS-related viruses by human monoclonal antibodies.
360 *Science* **369**, 731-736, doi:10.1126/science.abc7424 (2020).
361 31 Burnett, D. L. *et al.* Immunizations with diverse sarbecovirus receptor-binding domains elicit
362 SARS-CoV-2 neutralizing antibodies against a conserved site of vulnerability. *Immunity* **54**,
363 2908-2921 e2906, doi:10.1016/j.immuni.2021.10.019 (2021).
364 32 Wang, P. *et al.* A monoclonal antibody that neutralizes SARS-CoV-2 variants, SARS-CoV, and
365 other sarbecoviruses. *Emerg Microbes Infect* **11**, 147-157,
366 doi:10.1080/22221751.2021.2011623 (2022).
367 33 Vanshylla, K. *et al.* Discovery of ultrapotent broadly neutralizing antibodies from SARS-CoV-2
368 elite neutralizers. *Cell Host Microbe* **30**, 69-82 e10, doi:10.1016/j.chom.2021.12.010 (2022).
369 34 Onodera, T. *et al.* A SARS-CoV-2 antibody broadly neutralizes SARS-related coronaviruses and
370 variants by coordinated recognition of a virus-vulnerable site. *Immunity* **54**, 2385-2398 e2310,
371 doi:10.1016/j.immuni.2021.08.025 (2021).
372 35 Tan, C. W. *et al.* Pan-Sarbecovirus Neutralizing Antibodies in BNT162b2-Immunized SARS-CoV-
373 1 Survivors. *N Engl J Med* **385**, 1401-1406, doi:10.1056/NEJMoa2108453 (2021).
374 36 Cao, Y. *et al.* Potent Neutralizing Antibodies against SARS-CoV-2 Identified by High-Throughput
375 Single-Cell Sequencing of Convalescent Patients' B Cells. *Cell* **182**, 73-84 e16,
376 doi:10.1016/j.cell.2020.05.025 (2020).
377 37 Scheid, J. F. *et al.* B cell genomics behind cross-neutralization of SARS-CoV-2 variants and SARS-
378 CoV. *Cell* **184**, 3205-3221 e3224, doi:10.1016/j.cell.2021.04.032 (2021).
379 38 Cao, Y. *et al.* Humoral immune response to circulating SARS-CoV-2 variants elicited by
380 inactivated and RBD-subunit vaccines. *Cell Res* **31**, 732-741, doi:10.1038/s41422-021-00514-9
381 (2021).
382 39 Greaney, A. J. *et al.* Mapping mutations to the SARS-CoV-2 RBD that escape binding by different
383 classes of antibodies. *Nat Commun* **12**, 4196, doi:10.1038/s41467-021-24435-8 (2021).
384 40 Dejnirattisai, W. *et al.* The antigenic anatomy of SARS-CoV-2 receptor binding domain. *Cell* **184**,
385 2183-2200 e2122, doi:10.1016/j.cell.2021.02.032 (2021).

41 Barnes, C. O. *et al.* SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* **588**, 682-687, doi:10.1038/s41586-020-2852-1 (2020).
42 Yuan, M. *et al.* Structural and functional ramifications of antigenic drift in recent SARS-CoV-2 variants. *Science* **373**, 818-823, doi:10.1126/science.abh1139 (2021).
43 Du, S. *et al.* Structures of SARS-CoV-2 B.1.351 neutralizing antibodies provide insights into cocktail design against concerning variants. *Cell Res* **31**, 1130-1133, doi:10.1038/s41422-021-00555-0 (2021).
44 Li, T. *et al.* Cross-neutralizing antibodies bind a SARS-CoV-2 cryptic site and resist circulating variants. *Nat Commun* **12**, 5652, doi:10.1038/s41467-021-25997-3 (2021).
45 Huo, J. *et al.* Neutralization of SARS-CoV-2 by Destruction of the Prefusion Spike. *Cell Host Microbe* **28**, 445-454 e446, doi:10.1016/j.chom.2020.06.010 (2020).
46 Wrobel, A. G. *et al.* Antibody-mediated disruption of the SARS-CoV-2 spike glycoprotein. *Nat Commun* **11**, 5337, doi:10.1038/s41467-020-19146-5 (2020).
47 Cui, Z. *et al.* Structural and functional characterizations of infectivity and immune evasion of SARS-CoV-2 Omicron. *Cell*, doi:10.1016/j.cell.2022.01.019.
48 Hansen, J. *et al.* Studies in humanized mice and convalescent humans yield a SARS-CoV-2 antibody cocktail. *Science* **369**, 1010-1014, doi:10.1126/science.abd0827 (2020).
49 Dong, J. *et al.* Genetic and structural basis for SARS-CoV-2 variant neutralization by a two-antibody cocktail. *Nat Microbiol* **6**, 1233-1244, doi:10.1038/s41564-021-00972-2 (2021).

Methods

Antibody isolation and recombinant production

SARS-CoV-1 and SARS-CoV-2 RBD cross-binding memory B cells were isolated from PBMC of SARS-CoV-1 convalescents who received SARS-CoV-2 vaccine. Briefly, CD19+ B cells were isolated from PBMC with EasySep™ Human CD19 Positive Selection Kit II (STEMCELL, 17854). B cells were then stained with FITC anti-human CD19 antibody (BioLegend, 392508), FITC anti-human CD20 antibody (BioLegend, 302304), Brilliant Violet 421™ anti-human CD27 antibody (BioLegend, 302824), PE/Cyanine7 anti-human IgM antibody (BioLegend, 314532), biotinylated Ovalbumin (SinoBiological) conjugated with Brilliant Violet 605™ Streptavidin (BioLegend, 405229), SARS-CoV-1 biotinylated RBD protein (His & AVI Tag) (SinoBiological, 40634-V27H-B) conjugated with PE-streptavidin (BioLegend, 405204), SARS-CoV-2 biotinylated RBD protein (His & AVI Tag) (SinoBiological, 40592-V27H-B)

421 conjugated with APC-streptavidin (BioLegend, 405207), and 7-AAD (Invitrogen, 00-
422 6993-50). CD19/CD20+, CD27+, IgM-, OVA-, SARS-COV-1 RBD+, and SARS-CoV-
423 2 RBD+ were sorted with MoFlo Astrios EQ Cell Sorter (Beckman Coulter). FACS
424 data were analyzed using FlowJo™ v10.8 (BD Biosciences).

425 Sorted SARS-CoV-1 and SARS-CoV-2 RBD cross-binding B cells were then processed
426 with Chromium Next GEM Single Cell V(D)J Reagent Kits v1.1 following the
427 manufacturer's user guide (10x Genomics, CG000208). Briefly, Cells sorted were
428 resuspend in PBS after centrifugation. Gel beads-in-emulsion (GEMs) were obtained
429 with 10X Chromium controller and then subjected to reverse transcription (RT). After
430 GEM-RT clean up, RT products were subject to preamplification. After amplification
431 and purification with SPRIselect Reagent Kit (Beckman Coulter, B23318) of RT
432 products, B cell receptor (BCR) sequence (paired V(D)J) were enriched with 10X BCR
433 primers. After library preparation, libraries were sequenced by Novaseq 6000 platform
434 running Novaseq 6000 S4 Reagent Kit v1.5 300 cycles (Illumina, 20028312) or
435 NovaSeq XP 4-Lane Kit v1.5 (Illumina, 20043131).

436 After sequencing and data processing, monoclonal antibodies were expressed as
437 recombinant human IgG1. Briefly, HEK293F cells (Thermo Fisher, R79007) were
438 transiently transfected with heavy and light chain expression vectors. The secreted
439 monoclonal antibodies from cultured cells were purified by protein A affinity
440 chromatography. The specificities of these antibodies were determined by ELISA.

441 **Antibody sequence analysis**

442 The antibody sequences obtained from 10X Genomics V(D)J sequencing were aligned
443 to GRCh38 reference and assembled as immunoglobulin contigs by the Cell Ranger
444 (v6.1.1) pipeline. Non-productive contigs and B cells that had multiple heavy chain or
445 light chain contigs were filtered out of the analysis. V(D)J gene annotation was
446 performed using NCBI IgBlast (v1.17.1) with the IMGT reference. Mutations on V(D)J
447 nucleotide sequences were calculated by using the igpipeline, which compared the

448 sequences to the closest germline genes and counted the number of different nucleotides.
449 For antibodies from public sources whose original sequencing nucleotide sequences
450 were not all accessible, the antibody amino acid sequences were annotated by
451 IMGT/DomainGapAlign[50] (v4.10.2) with default parameters. The V-J pairs were
452 visualized by R package circlize (v0.4.10).

453 **High-throughput antibody-escape mutation profiling**

454 The previously described high-throughput MACS (magnetic-activated cell sorting)-
455 based antibody-escape mutation profiling system was used to characterize RBD
456 escaping mutation profile for neutralizing antibodies. Briefly, duplicate RBD mutant
457 libraries were constructed based on Wuhan-Hu-1 RBD sequence (GenBank:
458 MN908947, residues N331-T531), theoretically containing 3819 possible amino acid
459 mutations. Each RBD mutant was barcoded with a unique 26-neuclotide (N26) and only
460 ACE2 binding variants were enriched for downstream experiment.

461 For antibody escape profiling, yeast libraries were induced overnight for RBD
462 expression and washed followed by two rounds of Protein A antibody based negative
463 selection and MYC-tag based positive selection to enrich RBD expressing cells. Protein
464 A antibody conjugated products were prepared following the protocol of Dynabeads
465 Protein A (Thermo Fisher, 10008D) and incubated with induced yeast libraries at room
466 temperature. MYC-tag based positive selection was performed according to the
467 manufacturer's protocol (Thermo Fisher, 88843).

468 After three rounds of sequential cell sorting, the obtained yeast cells were recovered
469 overnight. Plasmids were extracted from pre- and post-sort yeast populations by 96-
470 Well Plate Yeast Plasmid Preps Kit (Coolaber, PE053). The extracted plasmids were
471 then used to amplify N26 barcode sequences by PCR. The final PCR products were
472 purified with AMPure XP magnetic beads (Beckman Coulter, A63882) and 75bp single-
473 end sequencing was performed on an Illumina Nextseq 500 platform.

Processing of deep mutational scanning data

Single-end Illumina sequencing reads were processed as previously described. Briefly, reads were trimmed into 16 or 26 bp and aligned to the reference barcode-variant dictionary with dms_variants package (v0.8.9). Escape scores of variants were calculated as $F\times(n_{X,ab} / N_{ab}) / (n_{X,ref} / N_{ref})$, where $n_{X,ab}$ and $n_{X,ref}$ is the number of reads representing variant X, and N_{ab} and N_{ref} are the total number of valid reads in antibody-selected (ab) and reference (ref) library, respectively. F is a scale factor defined as the 99th percentiles of escape fraction ratios. Variants detected by less than 6 reads in the reference library were removed to avoid sampling noise. Variants containing mutations with ACE2 binding below -2.35 or RBD expression below -1 were removed as well, according to data previously reported. Finally, global epistasis models were built using dms_variants package to estimate mutation escape scores. For most antibodies, at least two independent assays are conducted and single mutation escape scores are averaged across all experiments that pass quality control.

Antibody clustering and visualization

Site total escape scores, defined as the sum of escape scores of all mutations at a particular site on RBD, were used to evaluate the impact of mutations on each site for each antibody. Each of these scores is considered as a feature of a certain antibody and used to construct a feature matrix $\mathbf{A}_{N\times M}$ for downstream analysis, where N is the number of antibodies and M is the number of features (valid sites). Informative sites were selected using sklearn.feature_selection.VarianceThreshold of scikit-learn Python package (v0.24.2) with the variance threshold as 0.1. Then, the selected features were L2-normalized across antibodies using sklearn.preprocessing.normalize. The resulting matrix is referred as $\mathbf{A'}_{N\times M'}$, where M' is the number of selected features. The dissimilarity of two antibodies *i, j* is defined as $1-\mathrm{Corr}(\mathbf{A'}_i,\mathbf{A'}_j)$, where $\mathrm{Corr}(\mathbf{x},\mathbf{y})$ is the Pearson's correlation coefficient of vector **x** and **y**, i.e. $\mathrm{Corr}(\mathbf{x},\mathbf{y})= \mathbf{x}\cdot\mathbf{y}/|\mathbf{x}||\mathbf{y}|$. We used sklearn.manifold.MDS to reduce the number of features from M' to D=20 with multidimensional scaling under the above metric. Antibodies are clustered into 10 epitope groups using sklearn.cluster.KMeans of scikit-learn in the resulting D-

504 dimensional feature space. Finally, these D-dimensional representations of antibodies
505 were further embedded into two-dimensional space for visualization with t-SNE using
506 sklearn.manifold.TSNE of scikit-learn. All t-SNE plots were generated by R package
507 ggplot2 (v3.3.3).
508

509 **Pseudovirus neutralization assay**

510 SARS-CoV-2 spike (GenBank: MN908947), Pangolin-GD spike (GISAID:
511 EPI_ISL_410721), RaTG13 spike (GISAID: EPI_ISL_402131), SARS-COV-1 spike
512 (GenBank: AY278491), Omicron BA.1 spike (A67V, H69del, V70del, T95I, G142D,
513 V143del, Y144del, Y145del, N211del, L212I, ins214EPE, G339D, S371L, S373P,
514 S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R,
515 N501Y, Y505H, T547K, D614G, H655Y, N679K, P681H, N764K, D796Y, N856K,
516 Q954H, N969K, L981F), BA.2 spike (GISAID: EPI_ISL_7580387), BA.1.1 spike
517 (BA.1+R346K), plasmid is constructed into pcDNA3.1 vector. G*ΔG-VSV virus (VSV
518 G pseudotyped virus, Kerafast) is used to infect 293T cells (American Type Culture
519 Collection [ATCC], CRL-3216), and spike protein expressing plasmid was used for
520 transfection at the same time. After culture, the supernatant containing pseudovirus was
521 harvested, filtered, aliquoted, and frozen at −80°C for further use.

522 Pseudovirus detection of PCoV-GD and RaTG13 was performed in 293T cells
523 overexpressing human angiotensin-converting enzyme 2 (293T-hACE2 cells). Other
524 pseudovirus neutralization assays were performed using the Huh-7 cell line (Japanese
525 Collection of Research Bioresources [JCRB], 0403).

526 Monoclonal antibodies or plasma were serially diluted (5-fold or 3-fold) in DMEM
527 (Hyclone, SH30243.01) and mixed with pseudovirus in 96-well plates. After incubation
528 at 5% CO_2 and 37°C for 1 h, digested Huh-7 cell (Japanese Collection of Research
529 Bioresources [JCRB], 0403) or 293T-hACE2 cells
530 (AmericanTypeCultureCollection[ATCC],CRL-3216) were seeded. After 24 hours of
531 culture, supernatant was discarded and D-luciferin reagent (PerkinElmer, 6066769) was

532 added to react in the dark, and the luminescence value was detected using a microplate
533 spectrophotometer (PerkinElmer, HH3400). IC50 was determined by a four-parameter
534 logistic regression model.

535 **Authentic virus neutralization**

536 SARS-CoV-2 and SARS-CoV-2 variants neutralization by plasma were performed
537 using cytopathic effect (CPE) assay. Plasma samples were inactivated at 56°C for 0.5h.
538 Plasma samples were serially diluted (2 fold) in the 96-well plates with cell culture
539 medium and mixed with 100TCID50 virus, and then incubated at 37°C, 5% CO_2 for 2
540 hours. About 10^4 digested Vero cells (ATCC, CCL-81) were seeded in plates and
541 cultured for 5 days. CPE of each well was recorded under microscopes. Assays for each
542 sample were replicated and the neutralization titer was calculated by the Spearman-
543 Karber method. All experiments were performed in a Biosafety Level 3 laboratory.

544 **ELISA**

545 To detect the broad-spectrum binding of the antibodies among Sarbecovirus, we
546 entrusted SinoBiological Technology Co., Ltd. to synthesize a panel of 20 sarbecovirus
547 RBDs (Supplementary Table 3). According to the sequence of 20 RBDs, a set of nested
548 primers was designed. The coding sequences were obtained by the overlap-PCR with a
549 6xHis tag sequence to facilitate protein purification. The purified PCR products were
550 ligated to the secretory expression vector pCMV3 with CMV promoter, and then
551 transformed into E. coli competent cells XL1-blue. Monoclones with correct
552 transformation were cultured and expanded, and plasmids were extracted. Healthy
553 HEK293 cells were passaged into a new cell culture and grown in suspension at 37 °C,
554 120 RPM, 8% CO2 to logarithmic growth phase and transfected with the recombinant
555 constructs by using liposomal vesicles as DNA carrier. After transfection, the cell
556 cultures were followed to assess the kinetics of cell growth and viability for 7 days.
557 The cell expression supernatant was collected, and after centrifugation, passed through
558 a Ni column for affinity purification. The molecular size and purity of eluted protein
559 was confirmed by SDS-PAGE. Production lot numbers and concentration information

560 of the 20 Sarbecovirus proteins are shown in Suppplemenatary Table 4. The WT RBD
561 in the article is SARS-CoV-2 (2019-nCoV) Spike RBD-His Recombinant Protein
562 (SinoBiological, 40592-V08H).
563
564 A panel of 21 sarbecovirus RBDs (supplementary table3) in PBS was pre-coated onto
565 ELISA Plates (NEST, 514201) at 4°C overnight. The plates were washed and blocked.
566 Then 1μg/ml purified antibodies or serially diluted antibodies were added and incubated
567 at room temperature (RT) for 20min. Next, Peroxidase-conjugated AffiniPure Goat
568 Anti-Human IgG (H+L) (JACKSON, 109-035-003) was applied and incubated at RT
569 for 15min. Tetramethylbenzidine (TMB) (Solarbio, 54827-17-7) was added onto the
570 plates. The reaction was terminated by 2 M H_2SO_4 after 10min incubation. Absorbance
571 was measured at 450 nm using Ensight Multimode Plate Reader (PerkinElmer,
572 HH3400). ELISA OD450 measurements at different antibody concentration for a
573 particular antibody-antigen pair are fit to the model $y=Ac^n/(c^n + E^n)$ using R package
574 mosaic (v1.8.3), where y is OD450 values and c is corresponding antibody
575 concentration. A, E, n are parameters, where E is the desired EC_{50} value for the specific
576 antibody and antigen.
577
578 **Antibody-ACE2 competition for RBD**
579 mFC-WT-RBD (Sino Biological, 40592-V05H) protein in PBS was immobilized on
580 the ELISA plates at 4°C overnight. The coating solution was removed and washed
581 three times by PBST and the plates were then blocked for 2 h. After blocking, the
582 plates were washed five times, and the mixture of ACE2-his (Sino Biological, 10108-
583 105H) and serially diluted competitor antibodies was added followed by 30min
584 incubation at RT. Then anti-his-HRP (Proteintech, HRP-66005) was added into each
585 well for another 20min incubation at RT. After washing the plates for five times,
586 Tetramethylbenzidine (TMB) (Solarbio, 54827-17-7) was added into each well. After
587 10 min, the reaction was terminated by 2M H_2SO_4. Absorbance was measured at 450
588 nm using Ensight Multimode Plate Reader (PerkinElmer, HH3400). The ACE2
589 competition coefficient is calculated as (B-A)/B, where B is the OD450 value under

590 0.15ug/ml antibody concentration and A is the OD450 value under 3ug/ml antibody
591 concentration.
592
593 **Biolayer Interferometry**
594 Biolayer interferometry assays were performed on Octet® RED96 Protein Analysis
595 System (Fortebio) according to the manufacturer's instruction. The kinetics assays
596 were conducted with protein A biosensor (ForteBio 18-5010). (1) Sensor check:
597 sensors immersed 10 min in buffer alone (buffer ForteBio 18-1105). (2) Baseline:
598 sensors immersed 30 s in buffer alone. (3) Loading: sensors immersed 300 s with
599 antibody at 2 μg/ml to capture Ab, with a threshold of 0.4 nm. (4) Baseline 2: sensors
600 immersed 120 s in buffer alone. (5) Antigen association: sensors immersed 60 s with
601 serial dilutions of RBD or its variants at 75-5 nM. (6) Antigen dissociation: sensors
602 immersed 600 s in buffer alone. (7) Sensor regeneration: sensors immersed 30 s in
603 regeneration buffer (10 mM Glycine-HCl, pH 1.5), then immersed 30 s in buffer.
604 Repeat for 2 more times. Data were collected with Octet Acquisition 9.0 (Fortebio)
605 and analyzed by Octet Analysis 9.0 (Fortebio) and Octet Analysis Studio 12.2
606 (Fortebio).
607
608 **Protein expression and purification for cryo-EM study**
609 The S6P expression construct encoding the SARS-CoV-2 spike ectodomain (residues
610 1-1208) with six stabilizing Pro substitutions (F817P, A892P, A899P, A942P, K986P,
611 and V987P) and a "GSAS" substitution for the furin cleavage site (residues 682–685)
612 was previously described[51]. The Delta specific mutations (T19R, G142D, 156del,
613 157del, R158G, L452R, T478K, D614G, P681R, D950N) were introduced into this
614 construct using site-directed mutagenesis. The S6P expression construct containing the
615 Omicron mutations (A67V, H69del, V70del, T95I, G142D, V143del, Y144del, Y145del,
616 N211del, L212I, ins214EPE, G339D, S371L, S373P, S375F, K417N, N440K, G446S,
617 S477N, T478K, E484A, Q493R, G496S, Q498R, N501Y, Y505H, T547K, D614G,

618 H655Y, N679K, P681H, N764K, D796Y, N856K, Q954H, N969K, L981F) were
619 assembled from three synthesized DNA fragments. The expression construct encoding
620 the SARS-CoV spike ectodomain (residues 1-1195)[52] was kindly provided by Prof. X.
621 Wang (Tsinghua university), and two stabilizing Pro substitutions (K968P, V969P) was
622 engineered into this construct using mutagenesis. For protein production, these
623 expression plasmids, as well as the plasmids encoding the antigen-binding fragments
624 (Fabs) of the antibodies described in this paper, were transfected into the HEK293F
625 cells using polyethylenimine (Polysciences). The conditioned media were harvested
626 and concentrated using a Hydrosart ultrafilter (Sartorius), and exchanged into the
627 binding buffer (25 mM Tris, pH 8.0, and 200 mM NaCl). Protein purifications were
628 performed using the Ni-NTA affinity method, followed by gel filtration
629 chromatographies using either a Superose 6 increase column (for the spike proteins) or
630 a Superose 200 increase column (for the Fabs). The final buffer used for all proteins is
631 20 mM HEPES, pH 7.2, and 150 mM NaCl.
632
633 **Cryo-EM data collection, processing, and structure building**
634 The samples for cryo-EM study were prepared essentially as previously described[51,53].
635 All EM grids were evacuated for 2 min and glow-discharged for 30 s using a plasma
636 cleaner (Harrick PDC-32G-2). Four microliters of spike protein (0.8 mg/mL) was
637 mixed with the same volume of Fabs (1 mg/mL each), and the mixture was immediately
638 applied to glow-discharged holy-carbon gold grids (Quantifoil, R1.2/1.3) in an FEI
639 Vitrobot IV (4 °C and 100% humidity). Data collection was performed using either a
640 Titan Krios G3 equipped with a K3 direct detection camera, or a Titan Krios G2 with a
641 K2 camera, both operating at 300 kV. Data processing was carried out using
642 cryoSPARC[54]. After 2D classification, particles with good qualities were selected for
643 global 3D reconstruction and then subjected to homogeneous refinement. To improve
644 the density surrounding the RBD-Fab region, UCSF Chimera[55] and Relion[56] were used
645 to generate the masks, and local refinement was then performed using cryoSPARC.
646 Coot[57] and Phenix[58] were used for structural modeling and refinement. Figures were
647 prepared using USCF ChimeraX[59] and Pymol (Schrödinger, LLC.).

Reference

50 Ehrenmann, F. & Lefranc, M. P. IMGT/DomainGapAlign: IMGT standardized analysis of amino acid sequences of variable, constant, and groove domains (IG, TR, MH, IgSF, MhSF). *Cold Spring Harb Protoc* **2011**, 737-749, doi:10.1101/pdb.prot5636 (2011).

51 Du, S. *et al.* Structurally Resolved SARS-CoV-2 Antibody Shows High Efficacy in Severely Infected Hamsters and Provides a Potent Cocktail Pairing Strategy. *Cell* **183**, 1013-1023 e1013, doi:10.1016/j.cell.2020.09.035 (2020).

52 Gui, M. *et al.* Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a prerequisite conformational state for receptor binding. *Cell Res* **27**, 119-129, doi:10.1038/cr.2016.152 (2017).

53 Du, S. *et al.* Structures of SARS-CoV-2 B.1.351 neutralizing antibodies provide insights into cocktail design against concerning variants. *Cell research*, doi:10.1038/s41422-021-00555-0 (2021).

54 Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296, doi:10.1038/nmeth.4169 (2017).

55 Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612, doi:10.1002/jcc.20084 (2004).

56 Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, doi:10.7554/eLife.42166 (2018).

57 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).

58 Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* **75**, 861-877, doi:10.1107/S2059798319011471 (2019).

59 Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci* **30**, 70-82, doi:10.1002/pro.3943 (2021).

Declaration of interests

X.S.X. and Y.C. are listed as inventors on the provisional patent applications of BD series antibodies. X.S.X. and Y.C. are founders of Singlomics Biopharmaceuticals. The remaining authors declare no competing interests.

Corresponding authors

Correspondence to Yunlong Cao or Ronghua Jin or Junyu Xiao or Xiaoliang Sunney Xie. Request for materials described in this study should be directed to Yunlong Cao and Xiaoliang Sunney Xie.

Data availability

Logo plots of escape maps for antibodies in this study are available in Supplementary Data 1. Processed mutation escape scores can be downloaded at https://github.com/jianfcpku/SARS-CoV-2-RBD-DMS-broad. Raw Illumina and PacBio sequencing data are available on NCBI Sequence Read Archive BioProject PRJNA787091. We used vdj_GRCh38_alts_ensembl-5.0.0 as the reference of V(D)J alignment, which can be obtained from https://support.10xgenomics.com/single-cell-vdj/software/downloads/latest. IMGT/DomainGapAlign is based on the built-in lastest IMGT antibody database, and we let the "Species" parameter as "Homo sapiens" while kept the others as default. Public deep mutational scanning datasets involved in the study from literature could be downloaded at https://media.githubusercontent.com/media/jbloomlab/SARS2_RBD_Ab_escape_maps/main/processed_data/escape_data.csv.

Code availability

Codes for analyzing SARS-CoV-2 escaping mutation profile data are available at https://github.com/sunneyxielab/SARS-CoV-2-RBD-Abs-HTDMS. R and Python scripts for reproducing figures in this manuscript are available at https://github.com/jianfcpku/SARS-CoV-2-RBD-DMS-broad.

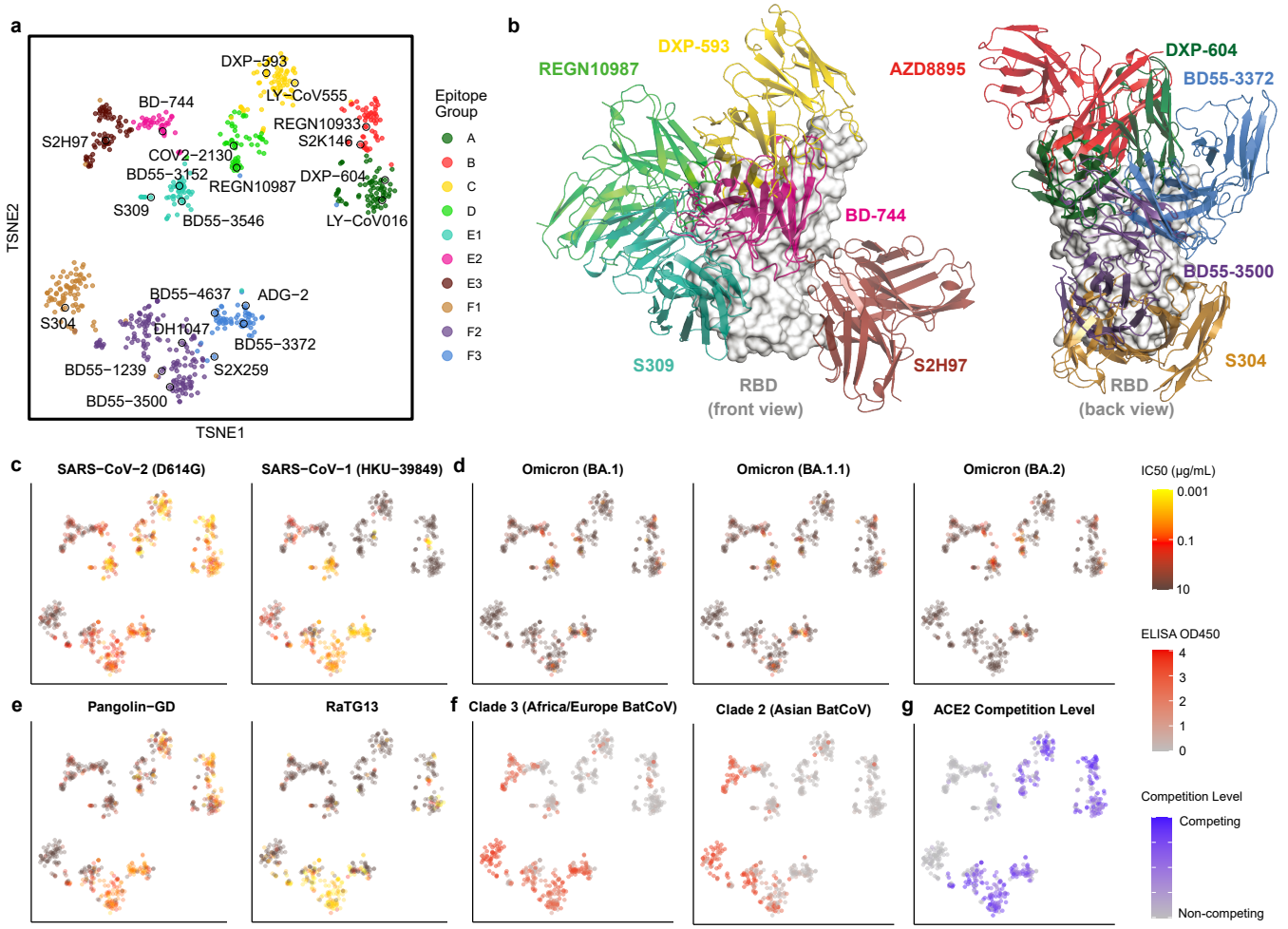


Fig. 1 | Comprehensive epitope and functional mapping of broad sarbecovirus neutralizing antibodies.

a, t-distributed stochastic neighbor embedding (t-SNE) and unsupervised k-means clustering of neutralizing antibodies against SARS-CoV-2 based on antibody escape maps characterized by MACS-based high-throughput deep mutational screening. **b,** Representative antibody structures of ten epitope groups. c-e, Neutralization of sarbecovirus (spike-pseudotyped VSV) by 715 RBD antibodies. Color bars indicate IC50 values (μg/mL). **c,** SARS-CoV-2 D614G and SARS-CoV-1 HKU-39849. **d,** Omicron variants BA.1, BA.1.1 and BA.2. **e,** SARS-CoV-2 related sarbecovirus Pangolin-GD and RaTG13. **f,** ELISA reactivity to various sarbecovirus RBD of different clades of 715 RBD antibodies. ELISA OD450 is measured using 0.3 μg/mL antigen and 1 μg/mL antibody. Shades of red show the average OD450 of BM48-31 and BtKY72 (for clade 3, left), and the average of YN2013, Shaanxi2011, SC2018, Rp3, ZXC21, ZC45 and Anlong112 (for clade 2, right). **g,** The ACE2 blocking activity for 715 RBD antibodies. Shades of blue show the competition level measured through competing ELISA. All pseudovirus neutralization assays and ELISA measure-

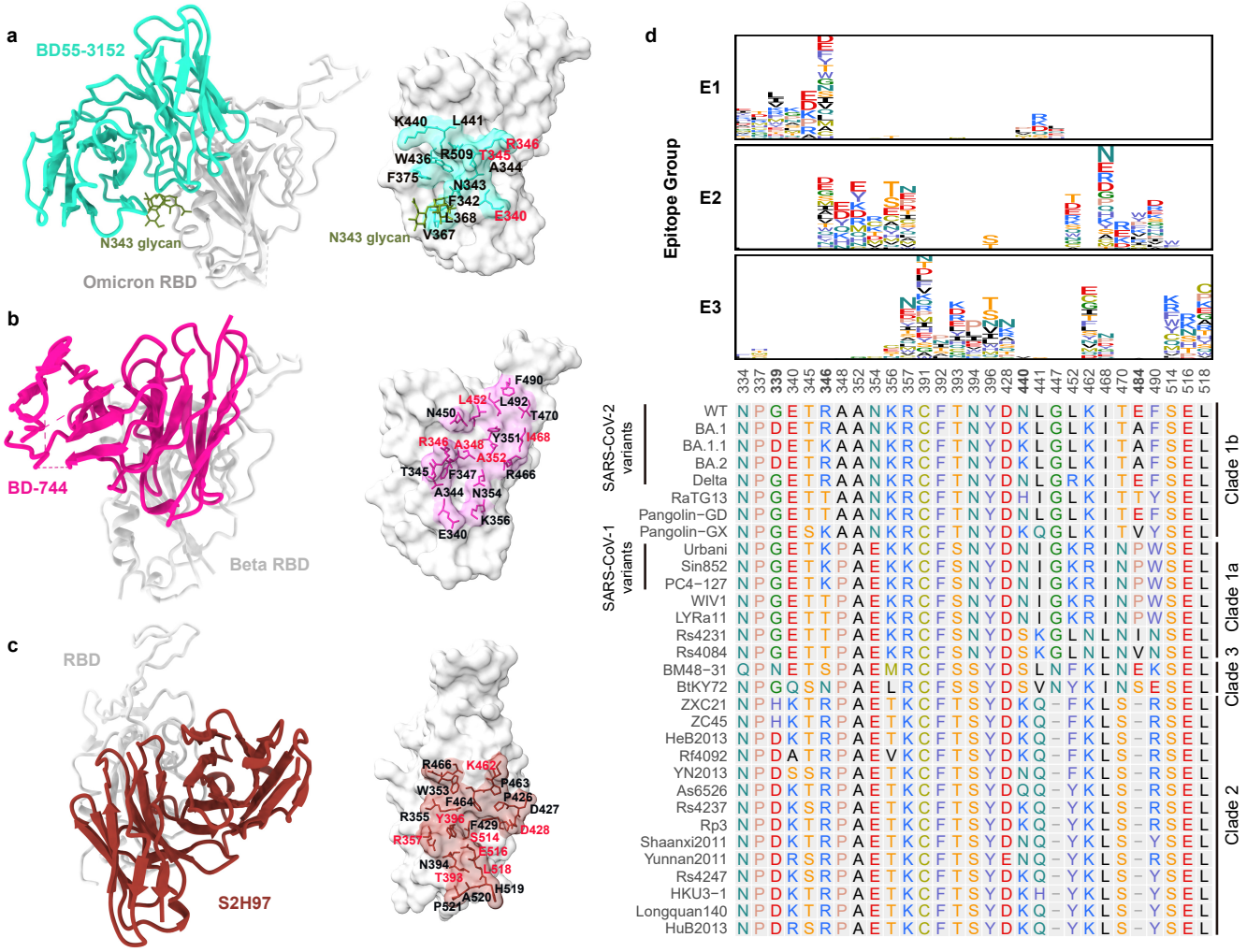


Fig. 2 | Structural and escaping mutation analyses of group E1-E3 antibodies.
a-c, High-resolution cryo-electron microscopy antibody structures of representative epitope group E1-E3 neutralizing antibodies. **a,** BD55-3152 (group E1) in complex of SARS-CoV-2 Omicron RBD complex. **b,** BD-744 (group E2) in complex of SARS-CoV-2 Beta RBD complex (PDB: 7EY0). **c,** S2H97 (group E3) in complex of SARS-CoV-2 RBD complex (PDB: 7M7W). Residues on the binding interface are marked. Residues highlighted in red indicate featuring escaping hotspots of the representative epitope groups. **d,** Averaged escape maps of antibodies in epitope group E1-E3, and corresponding multiple sequence alignment (MSA) of various sarbecovirus RBDs. Height of each amino acid in the escape maps represents its mutation escape score. Residues are colored corresponding to their chemical properties. Mutated sites in Omicron variants are marked in bold.

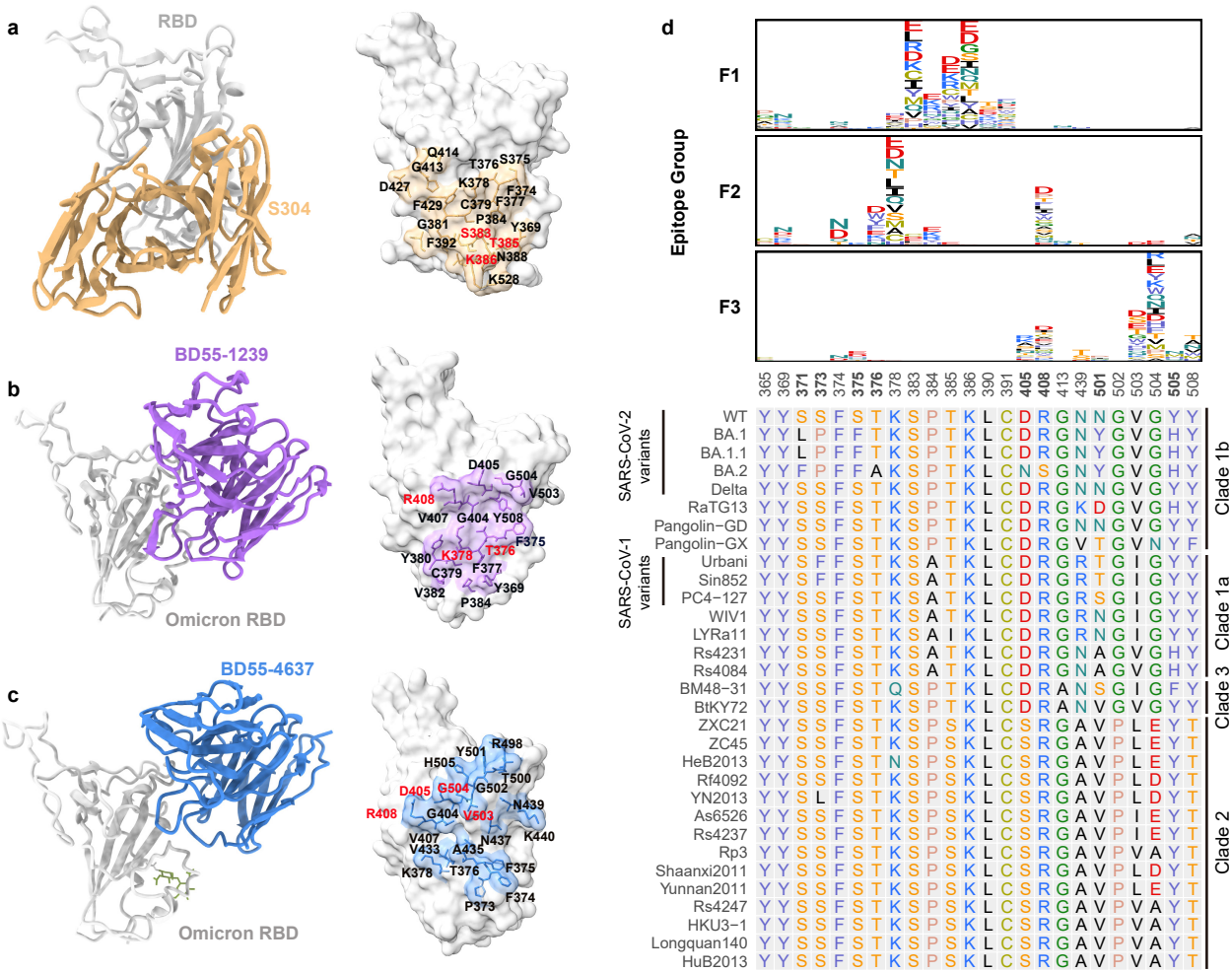
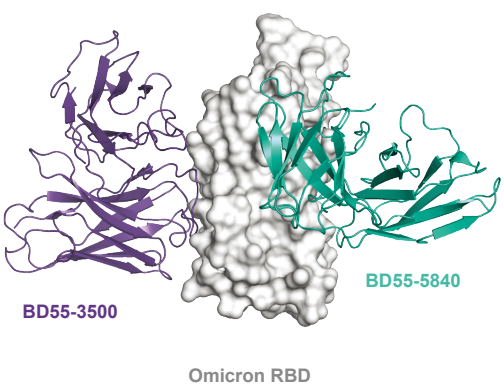


Fig. 3 | Structural and escaping mutation analyses of group F1-F3 antibodies.
a-c, High-resolution cryo-electron microscopy antibody structures of representative epitope group F1-F3 neutralizing antibodies. **a,** S304 (group F1) in complex of SARS-CoV-2 RBD complex (PDB: 7JW0). **b,** BD55-1239 (group F2) in complex of SARS-CoV-2 Omicron RBD complex. **c,** BD55-4637 (group F3) in complex of SARS-CoV-2 Omicron RBD complex. Residues on the binding interface are marked. Residues highlighted in red indicate featuring escaping hotspots of the representative epitope groups. **d,** Averaged escape maps of antibodies in epitope group F1-F3, and corresponding multiple sequence alignment (MSA) of various sarbecovirus RBDs. Height of each amino acid in the escape maps represents its mutation escape score. Residues are colored corresponding to their chemical properties. Mutated sites in Omicron variants are marked in bold.

a

		Antibody	S309	BD55-5840	BD55-3546	BD55-5549	BD55-5585	S2H97	BD55-5263	BD55-5640	BD55-5242	BD55-3500	BD55-5700	S2X259	BD55-3372	BD55-5514	BD55-5483	BD55-5558	ADG-2
		Epitope Group	E1	E1	E1	E1	E1	E3	E3	F2	F2	F2	F2	F3	F3	F3	F3	F3	F3
IC50 (ng/mL)	Pseudovirus	D614G	78.3	0.9	1.1	1.1	2.1	571.7	244.2	14.3	15.4	105.3	73.3	125.2	6.8	10.7	13.8	15.7	13.4
		SARS-CoV-1	31.3	5.6	4.6	24.3	6.5	6743.2	51.1	14.3	4.4	52.2	10.5	46.2	13.8	4.4	5.8	4.0	1.7
		BA.1	355.7	4.4	5.8	26.5	13.2	2407.6	1785.4	69.2	25.3	578.2	66.9	1847.0	20.3	1.7	6.5	16.0	1467.0
		BA.1.1	314.0	4.5	3.5	14.1	4.8	2231.2	1922.4	79.3	44.5	1218.8	92.7	2341.2	14.0	3.0	5.2	12.5	990.6
		BA.2	944.4	16.1	35.9	57.6	163.8	2015.1	1771.8	*	*	*	*	*	105.1	18.7	13.3	51.8	*
		Pangolin-GD	*	295.7	1860.0	*	2099.9	97.6	295.0	18.3	17.3	8.7	20.3	15.2	3.2	57.0	7.9	13.4	5.0
		RaTG13	*	*	*	*	*	834.2	219.1	2.5	0.8	3.1	2.1	1.1	*	37.5	*	1.9	*
	Authentic virus	Wuhan-Hu-1	78.1	12.3	13.8	58.0	15.5	1365.4	1088.2	55.2	124.0	44.9	210.3	456.2	22.4	62.0	26.3	55.2	40.6
		Omicron	928.8	22.4	52.6	116.1	89.7	7071.1	7937.0	196.9	359.0	6729.5	787.5	*	58.0	49.2	49.2	156.3	3715.0
RBD ELISA EC50(ng/mL)	SARS-CoV-2 clade(1b)	Wuhan-Hu-1	56.1	13.8	15.6	15.1	20.0	18.1	26.5	78.7	8.7	9.6	21.6	41.8	10.7	10.7	9.7	15.7	25.2
		Pangolin-GD	78.6	54.1	130.2	*	*	56.8	33.1	73.4	7.5	9.0	23.2	65.6	8.9	10.1	9.6	14.9	18.4
		Pangolin-GX	708.9	*	*	*	*	52.9	123.5	62.3	44.5	13.4	23.1	*	*	*	*	*	*
		RaTG13	144.7	*	*	*	*	41.2	129.2	66.3	27.1	15.6	24.7	61.2	*	35.4	*	58.3	*
	SARS-CoV-1 clade(1a)	BJ01	41.7	9.9	17.3	20.7	26.4	39.0	26.9	45.4	11.1	8.3	11.5	79.0	7.8	7.1	8.5	24.7	15.3
		PC4-127	52.9	12.5	15.1	14.6	29.6	36.7	12.9	44.1	7.1	7.3	13.5	64.9	8.3	5.6	7.3	10.0	9.9
		Sin852	71.4	16.4	29.2	28.8	43.5	1627.4	15.2	47.5	12.5	11.2	14.9	*	8.1	9.6	10.5	10.0	42.5
		WIV1	57.7	32.8	247.3	272.9	*	78.1	10.6	45.9	7.9	11.2	11.5	116.0	9.5	8.1	10.4	12.1	17.8
		LYRa11	113.5	39.4	854.2	*	*	91.4	51.7	64.6	21.1	19.2	18.0	98.9	19.1	11.3	15.5	32.8	24.6
		Rs7327	53.7	41.8	20.9	241.9	*	91.1	15.1	72.7	8.3	8.9	14.7	82.0	9.7	13.4	15.6	26.2	18.2
		GZ-C	116.2	13.6	43.5	33.6	34.2	34.3	23.7	52.5	14.2	13.7	13.2	97.3	15.4	10.3	12.1	18.1	17.0
		Urbani	139.3	33.4	464.6	58.7	91.1	1221.0	48.3	81.0	18.4	13.5	29.5	64.8	10.1	13.0	14.3	61.7	16.6
		Rs4231	99.6	*	*	*	*	*	119.7	81.9	17.1	9.4	17.8	38.1	7.2	50.3	38.3	38.8	24.8
	Africa/Europe BatCoV clade(3)	BM48-31	*	*	*	*	*	64.7	585.4	190.5	86.8	40.7	*	82.9	55.9	27.3	17.5	118.4	*
		BtKY72	*	*	*	*	*	75.3	60.9	57.6	14.7	12.6	21.2	77.1	14.0	13.8	12.9	27.0	52.9
	Asia BatCoV clade(2)	Shaanxi2011	*	*	*	*	*	41.9	113.3	81.7	34.4	24.9	30.8	*	*	*	*	*	*
		YN2013	*	*	*	*	*	44.2	39.5	83.1	40.1	21.7	66.4	*	*	*	*	*	*
		Rp3	*	*	*	*	*	37.3	1.5	76.1	43.6	35.8	19.7	*	*	*	*	*	*
		Rs4247	*	*	*	*	*	22.9	87.7	*	31.7	26.4	30.5	323.8	*	*	*	*	*
		ZC45	*	*	*	*	*	44.3	47.9	73.1	30.3	118.3	46.1	*	*	*	*	*	*
		Anlong112	*	*	*	*	*	39.6	41.0	139.8	49.6	39.1	149.2	*	*	*	*	*	*
		SC2018	*	*	*	*	*	42.0	44.8	174.6	28.2	15.7	60.4	*	*	*	*	*	*
		ZXC21	*	*	*	*	*	24.4	72.6	*	28.9	51.8	63.4	*	*	*	*	*	*

b



c

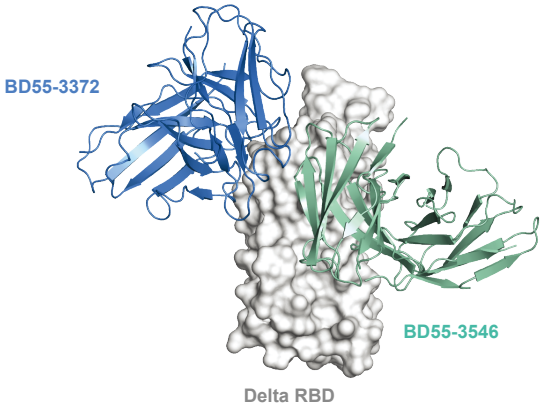
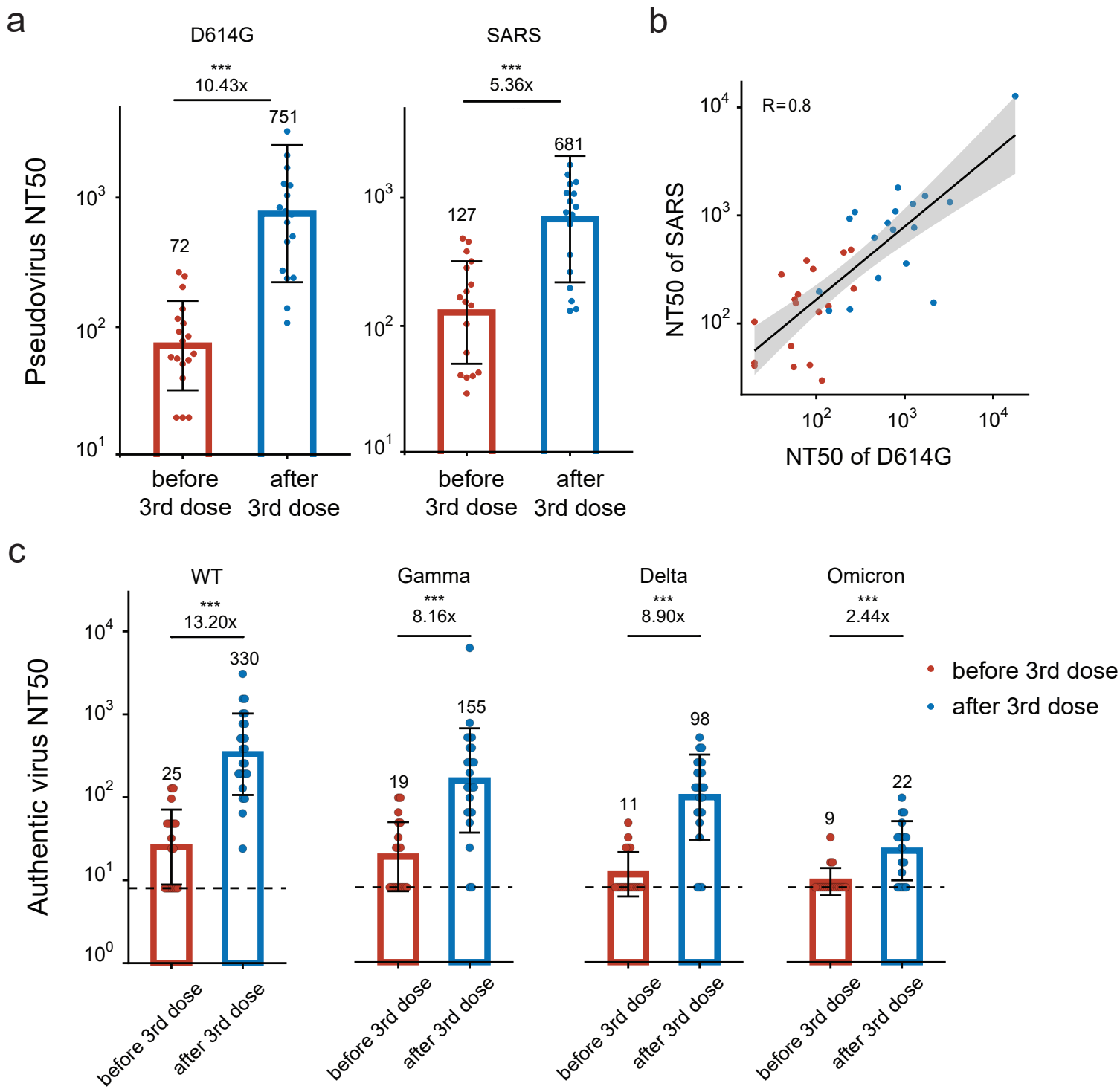


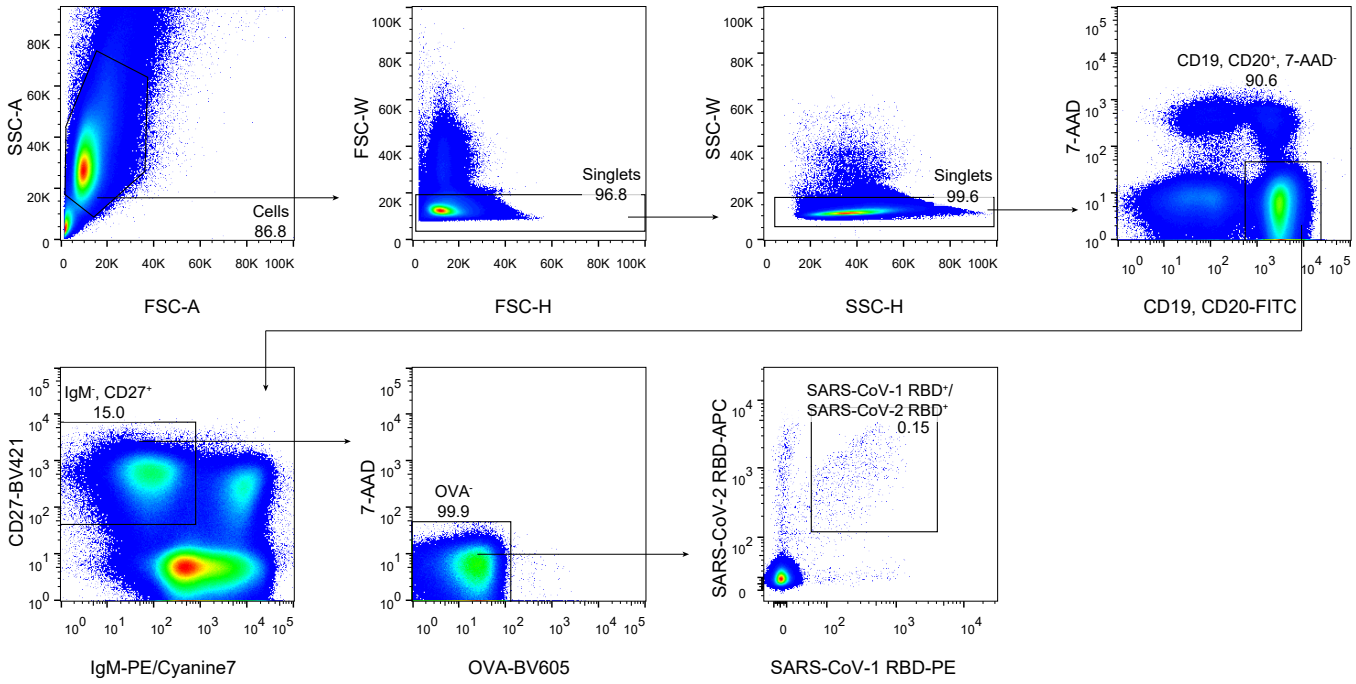
Fig. 4 | Potency and breadth of selected sarbecovirus neutralizing antibody drug candidates.

a, The sarbecovirus neutralization and binding capability (half-maximum effective concentration, EC50) of selected potent broad-spectrum neutralizing antibodies. *: >10,000 ng/mL for IC50, or >2,000 ng/mL for EC50. **b,** Structural model of the non-competitive broad sarbecovirus neutralizing antibodies BD55-3500 (group F2, in complex of Omicron RBD) and BD55-5840 (group E1, in complex of Omicron RBD). **c,** Structural model of the non-competitive broad sarbecovirus neutralizing antibodies BD55-3372 (group F3, in complex of Delta RBD) and BD55-3546 (group E1, in complex of Delta RBD).



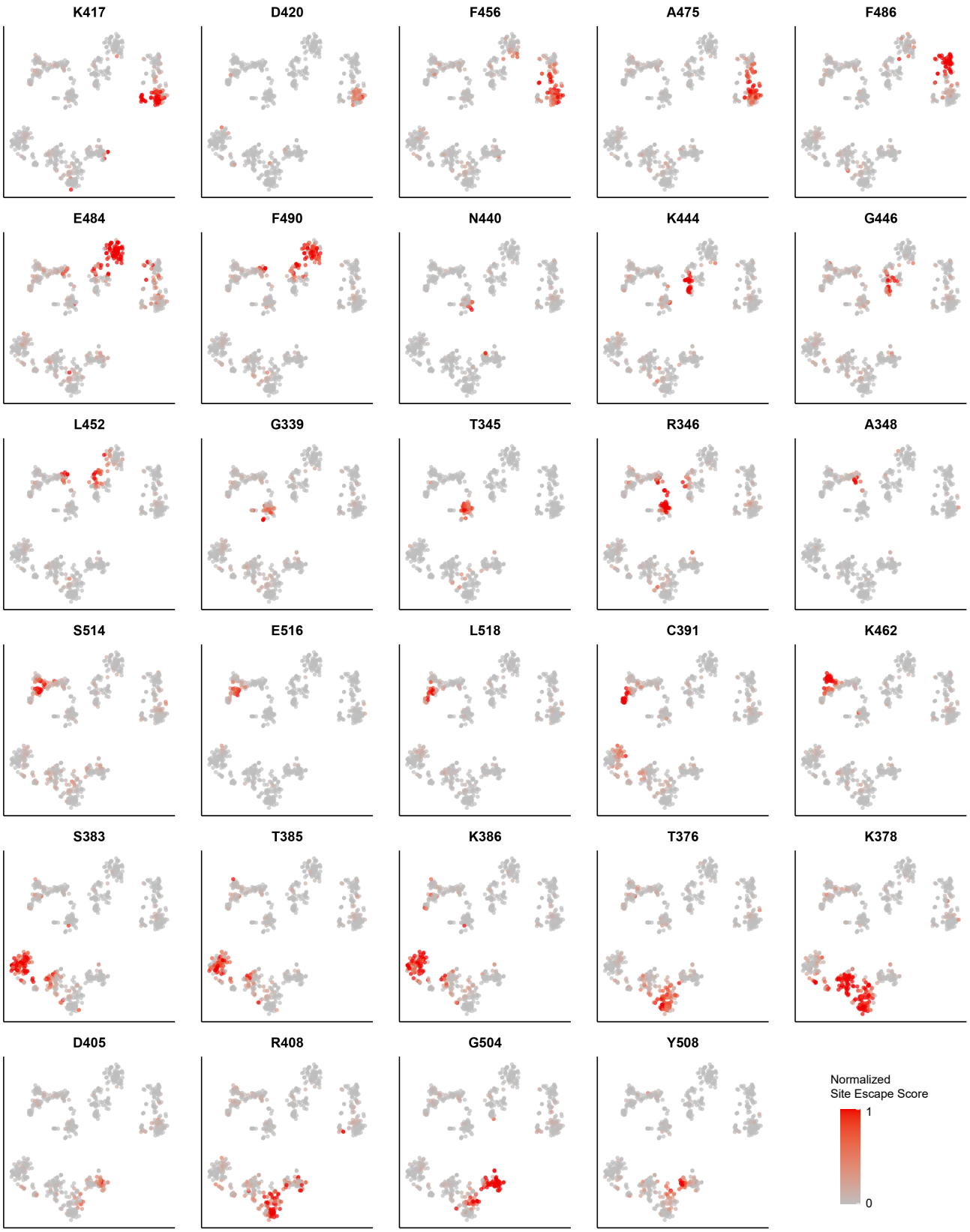
Extended Data Fig. 1 | Neutralization titre of plasma specimens from SARS convalescents vaccinated with 3 dose of SARS-CoV-2 vaccine.

a, Half-maximal neutralization titres (NT50) of plasma against SARS-CoV-2 with D614G mutant (D614G) and SARS-CoV-1 (SARS) pseudovirus, for SARS convalescents before and after the third dose vaccination (n=18). **b,** Scatter plot showing the correlation between plasma NT50 against D614G pseudovirus and SARS pseudovirus for SARS convalescents before (red) and after (blue) the third dose vaccination (n=18) . Pearson's correlation coefficient is labeled. **c,** NT50 of plasma against different SARS-CoV-2 variants authentic virus, for SARS convalescents before and after the third dose vaccination (n=18). Statistical significance in a, c was determined by two-tailed Wilcoxon signed rank test(***p<0.001, **p<0.01, *p<0.05). NT50 values are displayed as geometric mean ± s. d. in the log10 scale.

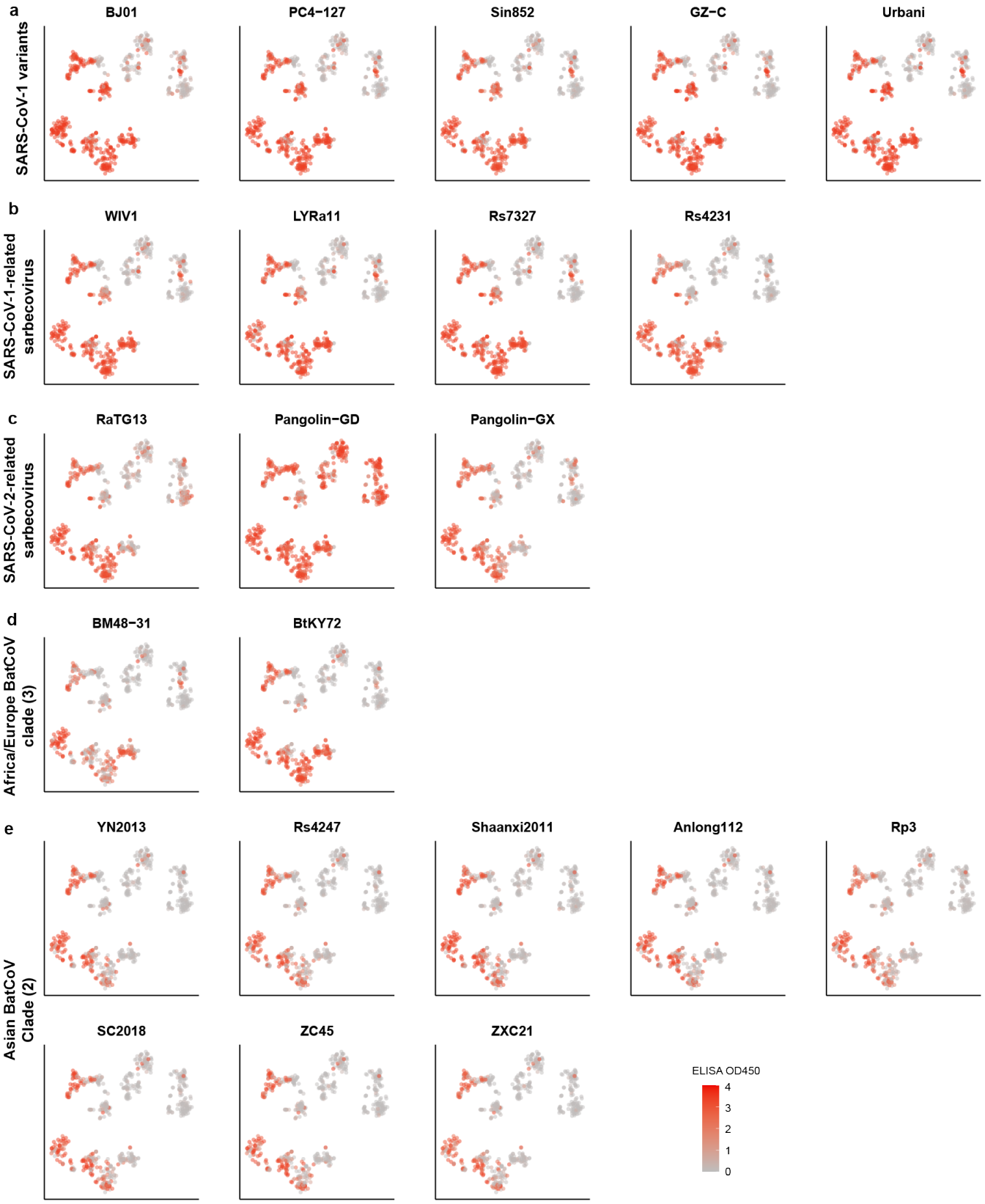


Extended Data Fig. 2 | FACS of SARS-CoV-1 RBD and SARS-CoV-2 RBD cross-binding B cells.

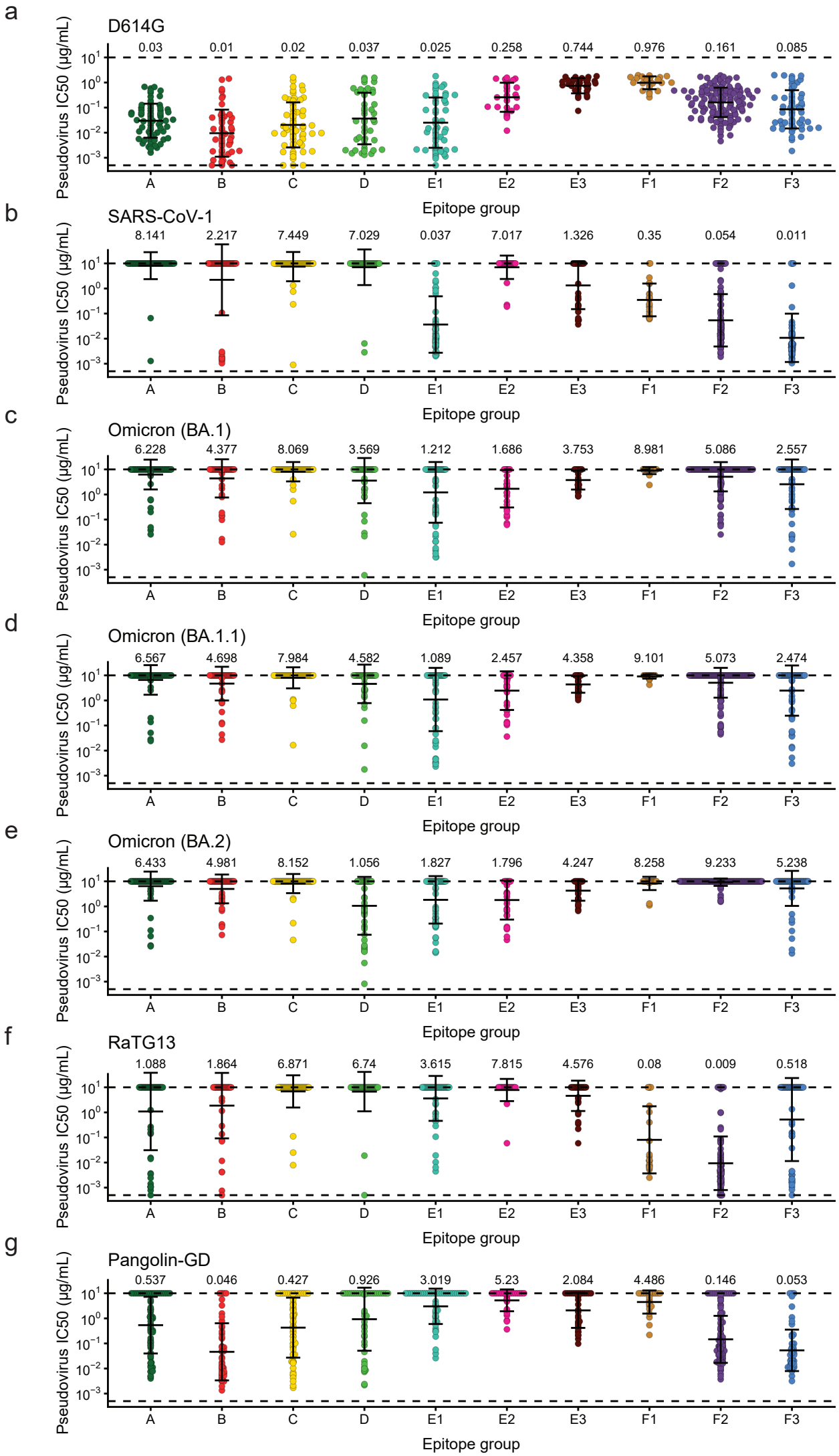
The gating strategy for sorting SARS-CoV-1-RBD+/SARS-CoV-2-RBD+ single memory B cells. Numbers next to outlined areas indicate percentage of cells in the gate. Sorting of the PBMCs from SARS convalescents that received 3 doses of the SARS-CoV-2 vaccine are shown.



Extended Data Fig. 3 | Escape scores projection of top RBD escaping hotspots. Shades of red indicate normalized site total escape scores of the representative residues for each antibody inferred from yeast display deep mutational screening.

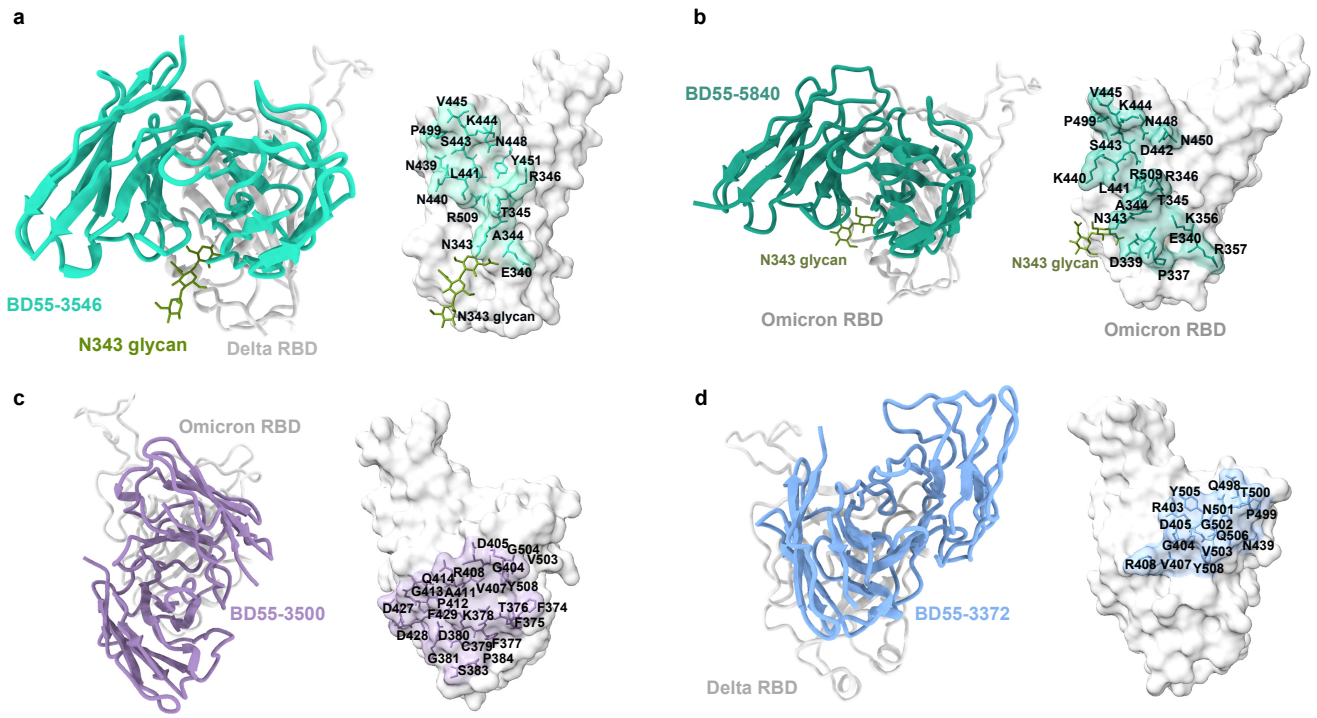


Extended Data Fig. 4 | Projection of ELISA reactivity against 22 sarbecovirus RBD. **a-e,** Shades of red indicate ELISA OD450 for each antibody against various sarbecovirus clades. (**a,** SARS-CoV-1 variants. **b,** SARS-CoV-1-related sarbecovirus. **c,** SARS-CoV-2 related sarbecovirus. **d,** Africa/Europe batcoronavirus. **e,** Asian non-ACE2-utilizing batcoronavirus.)


a
D614G
0.03 0.01 0.02 0.037 0.025 0.258 0.744 0.976 0.161 0.085
b
SARS-CoV-1
8.141 2.217 7.449 7.029 0.037 7.017 1.326 0.35 0.054 0.011
c
Omicron (BA.1)
6.228 4.377 8.069 3.569 1.212 1.686 3.753 8.981 5.086 2.557
d
Omicron (BA.1.1)
6.567 4.698 7.984 4.582 1.089 2.457 4.358 9.101 5.073 2.474
e
Omicron (BA.2)
6.433 4.981 8.152 1.056 1.827 1.796 4.247 8.258 9.233 5.238
f
RaTG13
1.088 1.864 6.871 6.74 3.615 7.815 4.576 0.08 0.009 0.518
g
Pangolin-GD
0.537 0.046 0.427 0.926 3.019 5.23 2.084 4.486 0.146 0.053
Pseudovirus IC50 (μg/mL)
10^{1} 10^{0} 10^{-1} 10^{-2} 10^{-3}
A B C D E1 E2 E3 F1 F2 F3
Epitope group

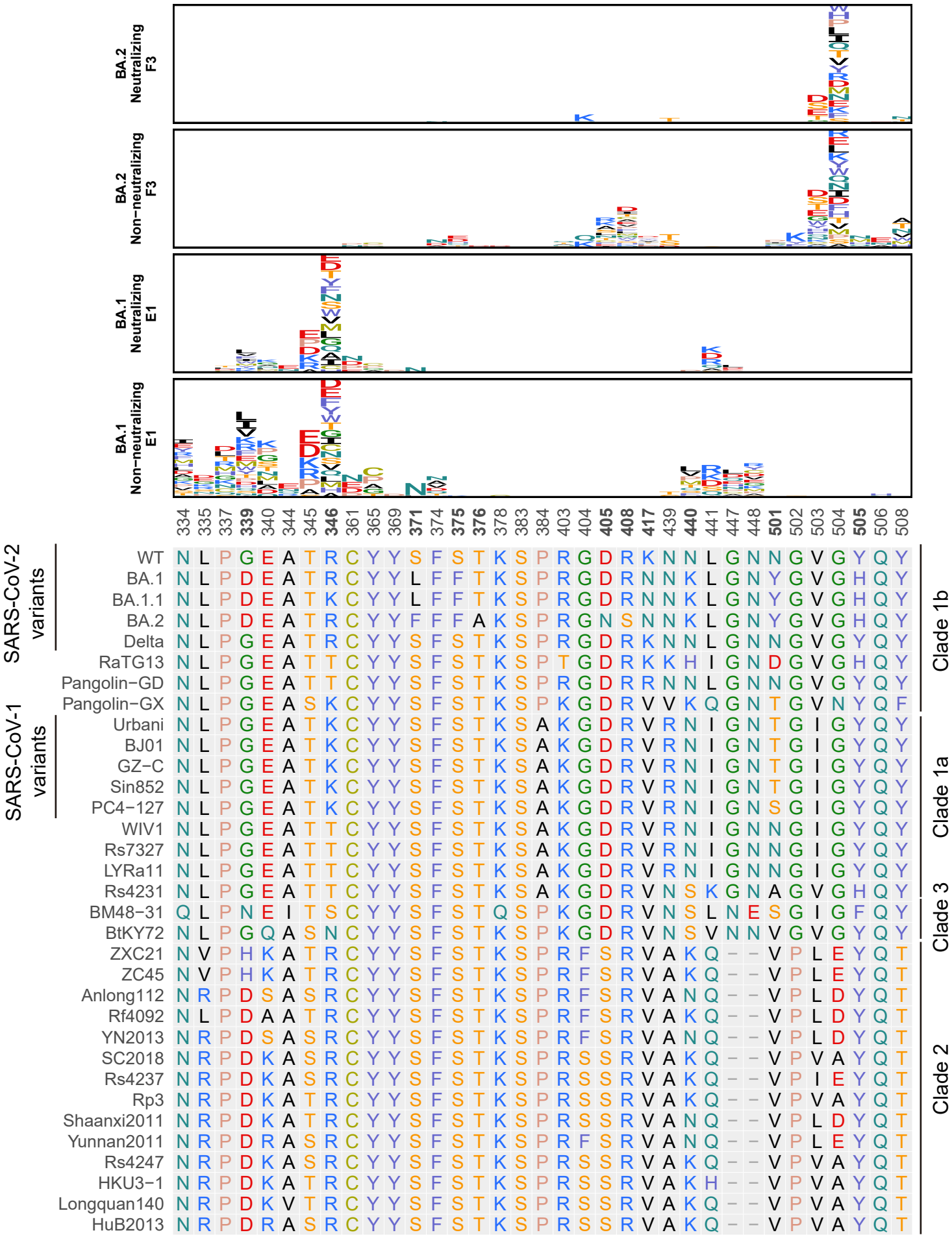
Extended Data Fig. 5 | Neutralization potency for NAbs of each epitope group.

a-g, Half maximal inhibitory concentration (IC50) of antibodies against different sarbecovirus determined by pseudovirus neutralization assay. The IC50 values are shown as geometric mean ± s. d. in log10 scale. Dashed lines show the detection limit, which is from 0.0005 μg/mL to 10μg/mL. Number of antibodies n = 68, 49, 61, 45, 53, 27, 31, 23, 126, 57 for epitope group A, B, C, D, E1, E2, E3, F1, F2, F3, respectively (except for **f**). **a**, SARS-CoV-2 with D614G mutation. **b**, SARS-CoV-1 (HKU-39849). **c**, SARS-CoV-2 Omicron (BA.1). **d**, SARS-CoV-2 Omicron (BA.1.1). **e**, SARS-CoV-2 Omicron (BA.2). **f**, RaTG13, n = 50, 38, 47, 41, 52, 27, 31, 23, 125, 56 for epitope group A, B, C, D, E1, E2, E3, F1, F2, F3, respectively. **g**, Pangolin-GD.



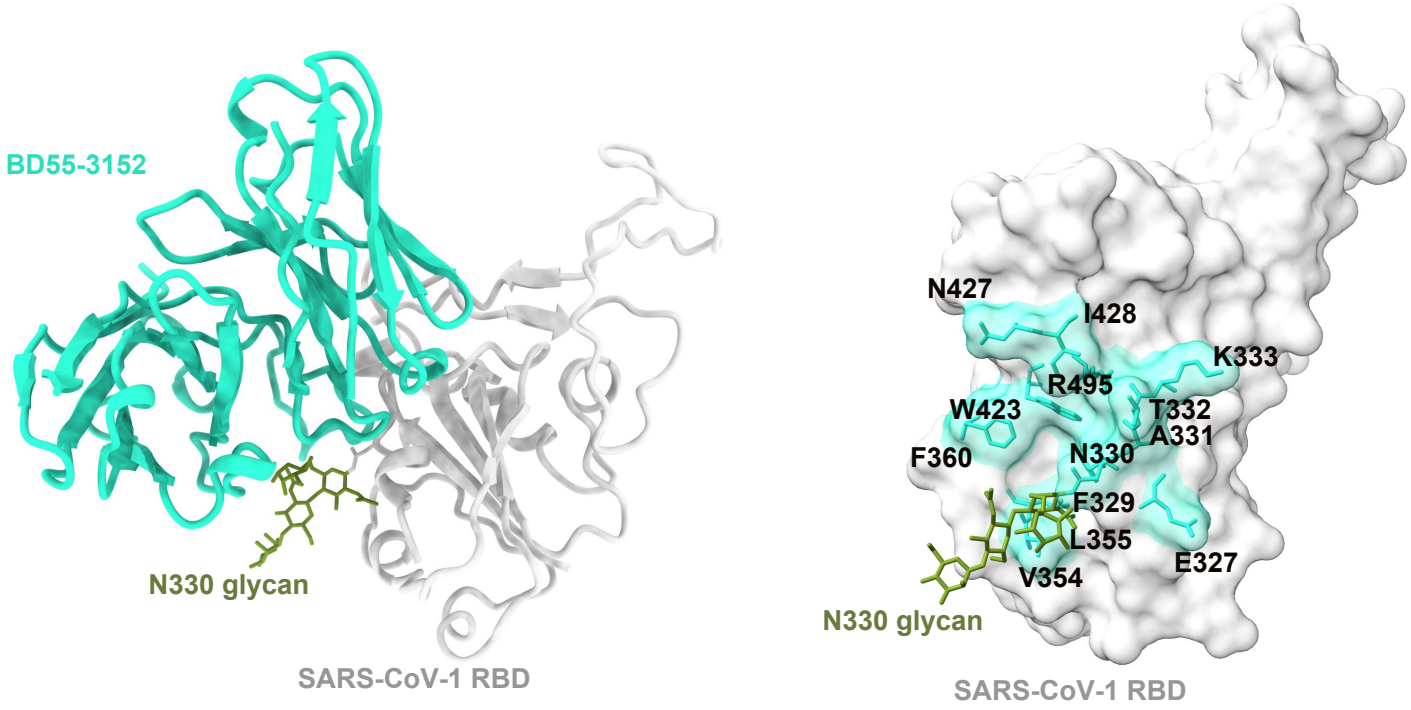
Extended Data Fig. 6 | Structures of individual broad sarbecovirus neutralizing antibody drug candidates.

a, BD55-3546 Fab in complex of Delta RBD. **b,** BD55-5840 Fab in complex of Omicron RBD. **c,** BD55-3500 Fab in complex of Omicron RBD. **d,** BD55-3372 Fab in complex of Delta RBD. Residues on the binding interface are marked.

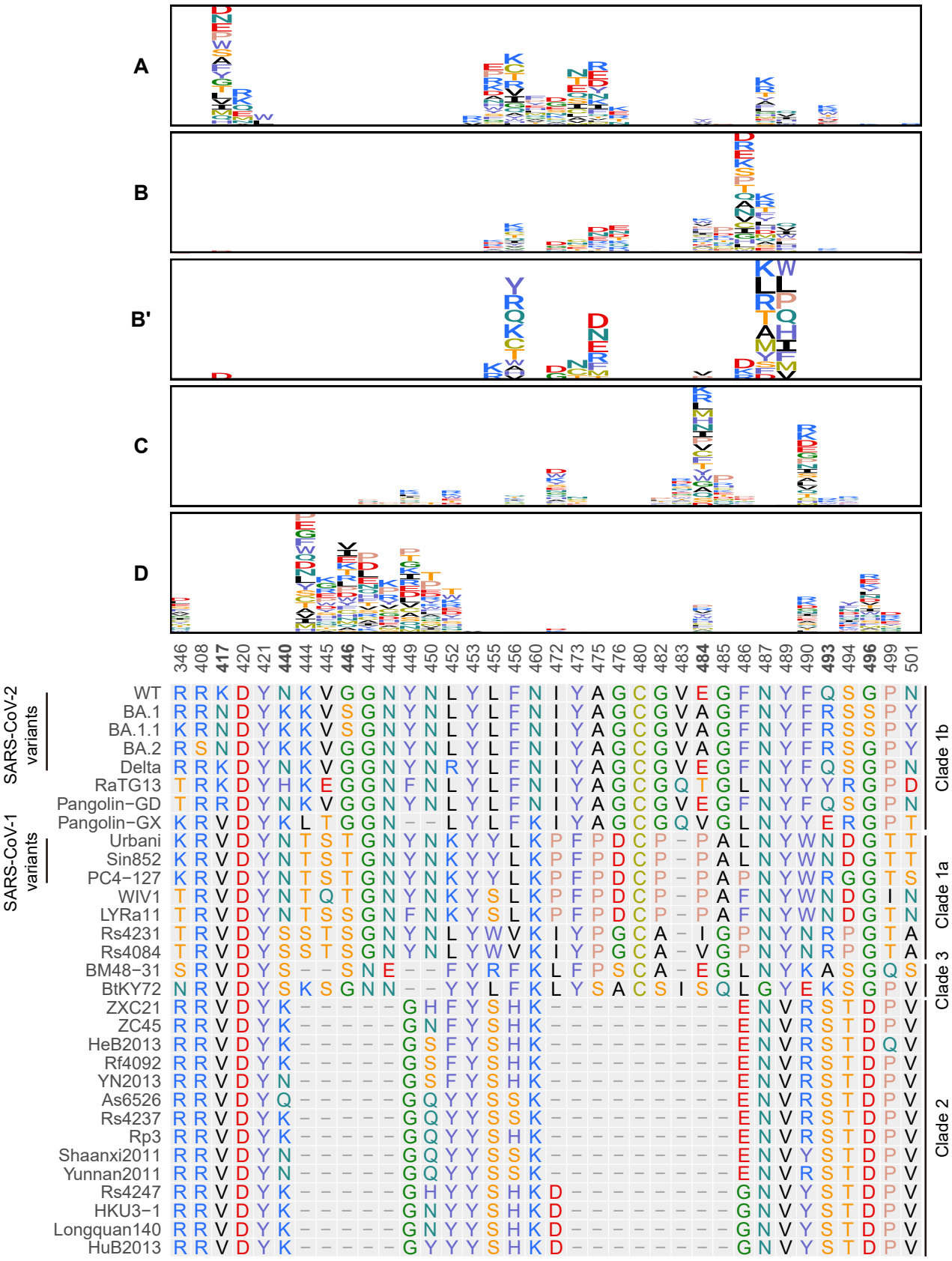


Extended Data Fig. 7 | Averaged escape maps of BA.1-neutralizing and non-neutralizing antibodies in epitope group E1 and F3.

Height of each amino acid in the escape maps represents its mutation escape score. Residues are colored corresponding to their chemical properties. Mutated sites in Omicron variants (including BA.1, BA.1.1 and BA.2) are marked in bold.

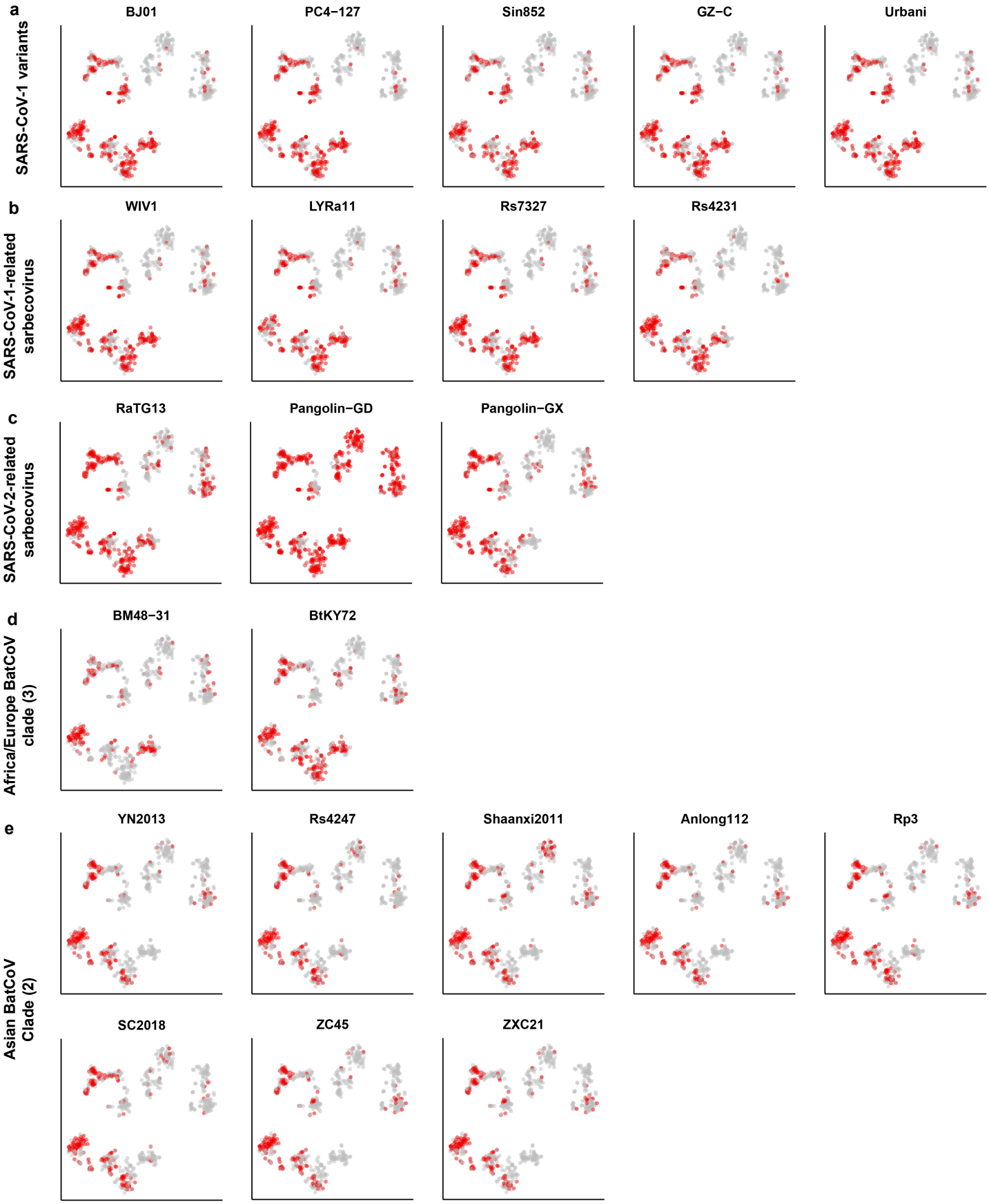


Extended Data Fig. 8 | Cryo-EM structure of BD55-3152 Fab in complex of SARS-CoV-1 RBD.
Residues on the binding interface are marked.



Extended Data Fig. 9 | Averaged escape maps of antibodies in epitope group A-D.

Height of each amino acid in the escape maps represents its mutation escape score. Residues are colored corresponding to their chemical properties. Mutated sites in Omicron variants are marked in bold.



Extended Data Fig. 10 | Projection of DMS predicted escaping of sarbecovirus RBD. **a-e,** Gray points represent antibodies predicted to be escaped from DMS data. Deletions and insertions on RBD are not considered. (**a,** SARS-CoV-1 variants. **b,** SARS-CoV-1-related sarbecovirus. **c,** SARS-CoV-2 related sarbecovirus. **d,** Africa/Europe batcoronavirus. **e,** Asian non-ACE2-utilizing batcoronavirus.)