AI recognizes convergent somatic hypermutation signatures to allow the discovery of variant-resilient broadly neutralizing antibodies

Qihong Yan[1,#], Xiaohan Huang[1,#], Banghui Liu[2,3,#], Yicheng Yao[4,5,#], Huiran Zheng[1,#], Yidong Song[6,7,#], Fandi Wu[6], Zhixiao He[8], Shaowei Li[8], Fangli Chen[1], Chuanying Niu[2,9], Zimu Li[2,5], Hang Yuan[2], Yuting Lin[1], Lan Chen[1], Na Zang[10], Juxue Xiao[1], Tingting Wang[8], Canjie Chen[1], Haoshi Bai[1], Jia Li[1], Mengyu Wu[1], Jingxian Zhao[1,5,*], Jianhua Yao[6,*], Yuedong Yang[7,*], Airu Zhu[1,*], Xiaoli Xiong[2,*], Jincun Zhao[1,5,8,*]

[1]State Key Laboratory of Respiratory Disease, National Clinical Research Center for Respiratory Disease, National Center for Respiratory Medicine, Joint International Research Laboratory of Respiratory Health, Guangdong Research Center of Basic Excellence for Respiratory Medicine, Guangzhou Institute of Respiratory Health, the First Affiliated Hospital of Guangzhou Medical University, Guangzhou, Guangdong Province, 510182, China.

[2]State Key laboratory of Respiratory disease, Guangdong Provincial Key laboratory of Stem cell and Regenerative Medicine, Guangdong-Hong Kong Joint laboratory for Stem cell and Regenerative Medicine, Guangzhou institutes of Biomedicine and health, Chinese Academy of Sciences, Guangzhou, Guangdong Province, 510530, China.

[3]Institute of Urban Agriculture, Chinese Academy of Agricultural Sciences, Chengdu, Sichuan Province, 610213, China.

[4]School of Life Sciences, Sun Yat-sen University, Guangzhou, Guangdong Province, 510275, China.

[5]Guangzhou National Laboratory, Guangzhou, Guangdong Province, 510005, China.

[6]Tencent, AI for Life Sciences Lab, Shenzhen, Guangdong Province, 518057, China.

[7]School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou, Guangdong Province, 510275, China.

[8]Shanghai Institute for Advanced Immunochemical Studies, School of Life Science and Technology, ShanghaiTech University, Shanghai, 201210, China.

[9]Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, Anhui Province 230026, China.

[10]Children's Hospital of Chongqing Medical University, National Clinical Research Center for Child Health and Disorders, Ministry of Education Key Laboratory of Child Development and Disorders, Chongqing, 400014, China

[#]Co-first author

*Correspondence:

zhaojingxian@gird.cn (Jingxian Zhao),

jianhua.yao@gmail.com (Jianhua Yao)

yangyd25@mail.sysu.edu.cn (Yuedong Yang),

zhuairu14@163.com (Airu Zhu),

xiong_xiaoli@gibh.ac.cn (Xiaoli Xiong),

zhaojincun@gird.cn (Jincun Zhao)

Summary

Elucidating the mechanisms by which broadly neutralizing antibodies (bnAbs) develop to confer durable immunity in humans is pivotal for the rational design of next-generation vaccines and therapeutics. VH3-53/3-66-encoded public antibodies are widely elicited in the population after the COVID-19 pandemic. Here, we isolated 15 VH3-53/3-66-encoded bnAbs from elite neutralizers who experienced sequential SARS-CoV-2 Omicron breakthrough infections. These bnAbs exhibit exceptional potency and breadth against circulating variants and, as members of a long-lived public antibody lineage, likely contribute to durable humoral protection. Genetic analysis of VH3-53/3-66-encoded bnAbs reveals a convergent somatic hypermutation pattern at seven positions in and around the CDR loops that emerges with repeated viral exposures and distinguishes them from non-bnAbs. Grafting these combined mutations onto clinically escaped antibodies broadens their breadth and restores neutralization activity. Structural analysis shows that these convergent mutations cooperatively remodel CDRs to bind and tolerate virus receptor-binding domain mutations. Using an AI-based antibody language model trained on data linking somatic hypermutations to binding affinity and neutralization, the model is able to identify mutational patterns predictive of breadth, enabling the discovery of a rare bnAb with protective activity against the latest variants from early-pandemic antibody repertoires. This demonstrates the feasibility of establishing an AI-empowered pipeline to identify mutation-tolerant bnAbs from early-pandemic antibody repertoires to fight fast-evolving newly introduced viruses.

Introduction

Broadly neutralizing antibodies (bnAbs) have shown promises in combating highly mutable RNA viruses such as influenza, HIV, and SARS-CoV-2[1-3]. The capacity of broadly neutralizing antibodies (bnAbs) to neutralize diverse viral strains makes them invaluable guides for vaccine and therapeutic design; however, their rational elicitation or discovery remains a major challenge. For most RNA viruses, bnAbs emerge only after multiple antigen exposures and prolonged affinity maturation, and they typically constitute an exceedingly small fraction of the overall antibody repertoire[4]. Moreover, the molecular rules that govern their development remain incompletely understood, severely constraining efforts to generate bnAbs early during outbreaks of emerging pathogens.

During the COVID-19 pandemic, the continued and rapid emergence of immune escape variants has highlighted these limitations. Despite intensive global vaccination and infection-driven immunity, the establishment of long-lasting, cross-variant neutralizing antibody responses has remained challenging. Shared ("public") neutralizing antibodies targeted immunodominant epitopes on the spike receptor-binding domain (RBD) have been widely identified from convalescents or vaccinee donors, particularly those encoded by VH1-2, VH1-58, VH1-69, and VH3-53/3-66 germlines[5-7]. Among these, VH3-53/3-66 antibodies are the most prevalent and target a class 1 RBD epitope with potent neutralization against the ancestral strain[8]. However, this immunodominant class has also been highly vulnerable to antigenic drift. VH3-53/3-66 antibodies isolated early in the pandemic, such as B-38[9], COVA2-04[10], CC12.1[11], and clinically approved LY-CoV016 (CB6)[12], P2C-1F11[13], and BD-604[14], were rapidly escaped by variants carrying K417N or K417T substitutions that are features of Beta, Gamma, and Omicron sublineages[11,15,16]. Intriguingly, despite extensive viral evolution, a few isolated VH3-53/3-66 antibodies retain cross-variant neutralizing activity. They appear to achieve this through the acquisition of somatic hypermutations (SHMs) that optimize binding to divergent RBDs[17-19]. Longitudinal analyses of convalescents and vaccinees reveal that persistent antigenic stimulation, through prolonged viral RNA or antigen presence, as well as sequential variant exposures[20], drives continued affinity maturation of pre-existing VH3-53/3-66 memory B cells[21]. This

process promotes somatic hypermutations associated with expanded antibody breadth and increased resilience to viral escape[22,23]. Some evidence indicates that matured VH3-53/3-66 antibodies carry stereotypic SHMs such as F27V/I, T28I, S31R, and Y58F, which remodel the paratope to tolerate RBD variation[19,24,25,26-28]. BD55-1205, for instance, a VH3-66-encoded neutralizing antibody isolated from a WT SARS-CoV-2 convalescent, maintains potent neutralization against all major circulating variants through three unique residues (R31, P53, and R102) in its heavy-chain complementarity-determining regions (CDRs) that introduce additional polar interactions at the binding interface[19]. It remains unclear whether convergent somatic hypermutation (SHM) trajectories drive the natural evolution of broad-spectrum VH3-53/3-66 antibodies. Defining such convergent pathways may reveal how cross-variant breadth is acquired and provide a molecular blueprint for broadly neutralizing antibody generation.

The ability to rapidly identify bnAbs that can withstand future antigenic drift represent a major unmet need in pandemic preparedness. The pace of therapeutic antibody discovery still lags behind viral evolution, despite remarkable progress in antibody engineering. The rapid advancement of artificial intelligence (AI), particularly protein language models (PLMs) trained on large-scale immunoglobulin repertoires, offers a potential transformative route to close this gap[24,29-31]. These models can learn the intrinsic grammar of antibody sequence evolution, enabling fitness prediction, structure inference, and sequence generation with unprecedented speed and precision[32-35].

Here, we mechanistically characterize fifteen VH3-53/3-66-encoded, broad-spectrum human monoclonal antibodies isolated from elite neutralizers with repeated Omicron breakthrough infections. These antibodies exhibit exceptional potency and breadth across all major SARS-CoV-2 variants. Integrated genetic and structural analyses show that repeated Omicron exposures drive convergent somatic hypermutations (SHMs) that expand neutralization breadth; introducing these SHMs into clinically escaped VH3-53/3-66 antibodies restores neutralization.

We reasoned that the convergent SHM patterns observed in naturally matured VH3-53/3-66 antibodies encapsulate the molecular logic of bnAb derivation. Accordingly, we show that an antibody language model trained on SHM-informed repertoires can learn mutational signatures predictive of breadth, enabling in silico identification of potent bnAbs before their mutation-tolerant features are experimentally selected. This strategy yields the AI-identified bnAb R102-9, which neutralizes all circulating variants and confers robust protection in vivo. By coupling SHM-convergence analysis with AI-driven inference, we establish a predictive antibody discovery workflow that transforms bnAb discovery from retrospective observation into a proactive, data-guided process.

Results

SARS-CoV-2 spike accumulates sequential mutations on the VH3-53/3-66-targeted epitope.

Canonical class 1 VH3-53/3-66 antibodies neutralize SARS-CoV-2 by blocking ACE2-RBD binding (Figure 1A). Structural analysis identified 20 amino acid residues on the WT RBD that directly participate in ACE2 interaction (Figure 1B). To investigate immune evasion of VH3-53/3-66 antibodies by viral variants, we analyzed epitope profiles of 79 structurally characterized VH3-53/3-66 antibodies, revealing remarkable epitope convergence (Figure S1). Our comprehensive mapping identified 28 RBD residues exhibiting substantial antibody interactions (defined by mean buried surface area >10 $Å^2$; Figure 1C). Evolutionary tracking across major circulating variants identified 34 mutated sites within the RBD (Figure 1D). Notably, we observed 10 overlapping residues among: (1) ACE2-binding sites, (2) VH3-53/3-66 epitopes, and (3) variant mutation hotspots (Figure 1E). Nine of the ten overlapping residues (R403, K417, L455, F456, N460, A475, F486, Q493, and N501) have been experimentally validated to mediate different levels of neutralization escape[25,36-38]. Strikingly, phylogenetic mapping revealed progressive accumulation of these escape mutations across successive variants within the VH3-53/3-66 epitope (Figure 1F). The earliest emerged B.1.1.7 (Alpha) variant carried only the N501Y mutation, followed by B.1.351 (Beta) acquiring K417N at the epitope core. Omicron BA.1/BA.2

155 acquired Q493R over Beta's profile while BA.4/BA.5 gained F486V. BQ.1 and XBB evolved
156 N460K on BA.4/5 backgrounds, with XBB initially carrying F486S before its descendants
157 acquired F486P. Subsequent XBB-derived lineages (EG.5, HV.1, and HK.3) sequentially
158 acquired F456L and L455F. The XBB sub-lineage JD.1.1 emerged with the most extensive
159 set of seven mutations including distinctive A475V. BA.2.86 emerged with five mutations
160 within the VH3-53/3-66-targeted epitope, followed by JN.1 and XDV.1 evolved L455S and
161 F456L; while the recently circulating KP.3.1.1, LP.8.1 and the currently dominant NB.1.8.1
162 all acquired Q493E, culminating in an eight-mutation escape profile (Figure 1F). Early in
163 the pandemic, the K417N mutation was the primary signature identified to confer the
164 escape of VH3-53/3-66 antibodies; however, our analysis demonstrates a continued
165 accumulation of escape mutations within the VH3-53/3-66 epitope. The shift from isolated
166 single-amino acid changes early in the pandemic to multi-mutation constellations
167 underscores sustained evolutionary pressure on this epitope. Therefore, we hypothesize
168 that pre-existing memory B cells utilizing VH3-53/3-66 are still capable of cross-recognizing
169 their convergent epitope even in the most recent SARS-CoV-2 variants.
170

171 **The VH3-53/3-66 antibody lineage constitutes a predominant component of durable**
172 **humoral immunity in elite neutralizers.**

173 To efficiently isolate VH3-53/3-66 antibodies with broad neutralization breadth and
174 tolerance to variant escape, we focused on screening elite neutralizers for monoclonal
175 antibody isolation. From a cohort of 68 convalescent individuals with three doses of
176 inactivated vaccine followed by BA.5 and XBB breakthrough infections, we identified 4
177 subjects with high neutralizing titers ($FRNT_{50}$ > 2,000) against WT, BA.5, XBB.1, and EG.5
178 strains (Figure S2A). Using Avi-tag-biotinylated JD.1.1 and KP.3.1.1 RBD probes, we
179 isolated a total of 3715 JD.1.1/KP.3.1.1-specific memory B cells through fluorescence-
180 activated cell sorting (FACS) (Figure S2B). We profiled sorted cells by 5' directed scRNA-
181 seq for both paired V(D)J and mRNA sequencing, recovering matched single-cell V(D)J
182 and transcriptome profiles in 2124 cells. According to their transcriptomes, cells were
183 clustered into four distinct clusters by Uniform Manifold Approximation and Projection for

184 Dimension Reduction (UMAP) (Figure 2A-B). We also quantified their antibody isotype
185 frequency and SHM levels (Figure 2C-D). Cluster 1 expressed Ig genes with little to no
186 SHM or CSR (class switch recombination) and gene signatures associated with naive B
187 cells (*TCL1A, BACH2, IGHD*), suggesting that this subset was composed of naive-like B
188 cells or very recently activated B cells (*CD69, CD83*) (Figure 2B-D). Clusters 2-4 with
189 patterns of higher SHM and CSR were charactered as memory B cells (MBCs) (Figure 2B-
190 D). The classical MBC gene signature (*CD27*) was mapped to clusters 2 and 3 (Figure 2B).
191 However, cluster 2 contained higher frequency of un-class-switched B cells with lower SHM
192 levels, while cluster 3 were highly mutated and were largely class-switched to IgG1
193 subclass (Figure 2C-D). Cluster 4 exhibited high expression of *ITGAX* (encoding CD11c)
194 and *ZEB2*, both of which are strongly linked to double-negative 2 (DN2) B cells, as well as
195 the inhibitory receptor gene *FCRL5* (Fc receptor-like 5) (Figure 2B).
196
197 We synthesized and expressed 486 mAbs from none-IGHD and none-IGHM-expressing B
198 cells to increase the chance of identifying neutralizing antibodies. These B cells represent
199 nearly 23% of all B cells obtained by scRNA-seq (Figure 2E). Totally, 207 of the selected
200 mAbs from cells isolated using JD.1.1/KP.3.1.1 probes exhibited specific binding to at least
201 one tested RBD (WT, JD.1.1, or KP.3.1.1) by biolayer interferometry (BLI) (Figure 2F).
202 Notably, more than 91% (189/207) of the RBD-specific B cells are identified within the
203 classical MBC cluster 3 (Figure 2F), which displays the highest level of somatic
204 hypermutation (SHM) among all subsets (Figure 2D). Moreover, *BCL2A1*, a key anti-
205 apoptotic regulator, is highly expressed in cluster 3, suggesting this B cell subset may
206 underlie long-lived humoral immunity (Figure 2B). The SHM levels were significantly higher
207 in RBD-specific monoclonal antibodies compared to unspecific counterparts (Figure 2G).
208 Out of the 207 RBD-specific antibodies, a total of 118 candidates achieved binding in BLI
209 assays against WT, JD.1.1, and KP.3.1.1 RBD at the same time (Figure 2H). Analysis of
210 germline genes showed that *IGHV2-5*, *IGHV3-30*, *IGHV3-53*, *IGHV3-66*, *IGHV5-51*, and
211 *IGHV7-4-1* for the heavy chain were predominantly used among the 118 cross-reactive
212 mAbs (Figure 2I). To narrow the range of candidates, we evaluated their neutralizing

activities against WT, JD.1.1, and KP.3.1.1 SARS-CoV-2 pseudoviruses. Although these mAbs were isolated using JD.1.1/KP.3.1.1 probes, 67% (79/118) of which showed effective neutralization with a geometric mean 50% inhibitory concentration (GEM IC_{50}) of 74.8 ng/mL against WT (Figure 2J), indicating that most of these were derived from recalled memory B cells. Finally, we identified 60 mAbs that could effectively neutralize three strains (Figures 2J, S2C, Table S1). We next tested if the cross-neutralizing mAbs were enriched for Ig genes VH3-53 and VH3-66, features that are preferentially found among class 1 nAbs. Our analysis revealed that 15 (25%) of the 60 cross-neutralizing antibodies utilized the VH3-53/3-66 gene segments, representing the second most prevalent genotype after *IGHV7-4-1* (Figure S2C). In addition, all 15 VH3-53/3-66-derived antibodies exhibited canonical class 1 characteristics, featuring short HCDR3 loops of 8-12 amino acids in length (Figure S2D). To investigate the breadth of 15 VH3-53/3-66-derived antibodies, we tested their neutralization activities against 19 representative SARS-CoV-2 pseudoviruses. Notably, 13 of the 15 VH3-53/3-66 antibodies demonstrated potent neutralization against all 19 pseudoviruses tested, with IC_{50} values <50 ng/mL for most strains (Figure 2K). The exceptions were ZL38 and ZL191, which showed reduced efficacy against the emerging JD.1.1, LB.1, KP.3.1.1, LP.8.1.1 and NB.1.8.1 variants (Figure 2K). Interestingly, four of the 15 VH3-53/3-66-class mAbs were derived from $IGHA1^{+}$ B cells, suggesting preferential induction of these broadly neutralizing antibodies at mucosal sites (Figure S2D). Taken together, we identified 15 V3-53/3-66 public antibodies with broad neutralizing capacity against SARS-CoV-2, which dominate the long-lived memory B cell population and exhibit a mucosa-associated antibody isotype expression, suggesting their dual origin from systemic and mucosal immune responses.

Repeated Omicron breakthrough infections drive convergent somatic hypermutations within VH3-53/3-66 antibodies

We next directly compared the *in vitro* neutralization potency and breadth of the 15 mAbs with the previously published VH3-53/3-66 antibodies, using a diverse panel of strains from XBB and JN.1 sub-lineages (Figure 3A). A panel of 32 representative VH3-53/3-66

242 neutralizing antibodies derived from individuals with different exposure histories were
243 included. Overall, antibodies derived from either wild-type (WT) infection or vaccine
244 recipients exhibited the most restricted neutralization breadth, with a median 50% inhibitory
245 concentration (median IC_{50}) of 61.78-10000 ng/mL. In contrast, monoclonal antibodies that
246 isolated from individuals with single Omicron breakthrough infection demonstrated
247 improved cross-neutralization, with median IC_{50} of 33.21-1315 ng/mL. Most notably,
248 antibodies that we isolated from elite neutralizers with sequential BA.5 and XBB
249 breakthrough infections displayed the most exceptional pan-variant neutralization breadth,
250 with median IC_{50} of 14.57-112.8 ng/mL (Figure 3A, Table S2). Among all the previously
251 published antibodies tested, BD55-1205 is the only one that demonstrated neutralizing
252 activity against all the viruses tested with a median IC_{50} of 485.1 ng/ml (range 52.35-791.7
253 ng/mL) (Figure 3A). These findings suggest that repeated heterologous antigen exposure
254 significantly promotes the development of SARS-CoV-2 broadly neutralizing antibodies, a
255 process we propose to be closely associated with the accumulation of somatic
256 hypermutations.
257
258 To understand the development of broad neutralizing activity, we further characterized the
259 somatic hypermutation features of these 15 mAbs compared to other published VH3-53/3-
260 66 public antibodies. We collected a total of 708 published VH3-53/3-66 antibodies that
261 characterized as class 1 features from CoV-AbDab[39]. As previously reported, VH3-53/3-66
262 antibodies isolated early in the pandemic generally have no to limited mutations, with more
263 than 60% antibodies have 0–5 mutations and less than 10% antibodies have more than 10
264 mutations (Figures 3B-C, S3A). Following single Omicron BA.1/BA.2/BA.5 breakthrough
265 infection (BTI), VH3-53/3-66 antibodies exhibited significantly increased SHMs in both
266 heavy and light chains. Repeated Omicron breakthrough infection (BA.5+XBB BTI)
267 induced substantial further accumulation of SHMs (Figures 3B-C, S3A).
268
269 To investigate potential enrichment of specific SHMs, we analyzed the mutation pattern at
270 each residue position across the FR1-FR3 regions of VH3-53/3-66 antibody heavy chains.

271 Notably, we identified seven prominent convergent somatic mutations in antibody CDR1
272 and CDR2 regions, with this enrichment pattern becoming more pronounced following both
273 primary and secondary Omicron breakthrough infections (Figure 3D-E). The convergent
274 SHM profile comprised both some of previous documented substitutions F27I, T28I, S31R,
275 V50L, S53P, and Y58F, as well as a newly observed prevalent mutation G26E. Notably,
276 although convergent somatic hypermutation patterns become apparent when antibody
277 sequences induced by WT infection or vaccination are cross-compared, these mutations
278 rarely co-occur within the same antibody that was isolated early in the pandemic (Figures
279 3D, S3B). Following two Omicron breakthrough infections, we observed a marked increase
280 in the proportion of VH3-53/3-66 antibodies harboring 6-7 concurrent somatic mutations
281 (Figure 3E, S3B). Strikingly, all 15 neutralizing antibodies isolated in this study exhibited
282 near-complete convergence at the seven critical paratope residues, indicative of strong
283 antigen-driven selection during affinity maturation (Figure S3B). Furthermore, we analyzed
284 the paired light chain repertoire of VH3-53/3-66 antibodies, revealing an enrichment of
285 VK1-33 following breakthrough infections (Figure S3C). Using the same analytical
286 approach, we identified three convergent SHMs (30_L-32_L) clustered together within the
287 LCDR1 paratope of VK1-33-encoded light chains ($P<0.001$ for each site, χ^2 test) (Figure
288 S3D-E). Taken together, these results demonstrate that repeated antigen exposure
289 markedly enriches convergent somatic mutations in both heavy and light chains of VH3-
290 53/3-66 mAbs.
291

292 **Evolution selected convergent somatic hypermutations expand potency and**
293 **breadth of VH3-53/3-66 antibodies**

294 To precisely define the function of the convergent SHMs in the isolated cross-neutralization
295 of antibodies, we reverted the heavy chains of the matured ZL54, ZL58, ZL112, and ZL611
296 antibodies to their germline forms (ZL54gl, ZL58gl, ZL112gl, and ZL611gl). While the
297 germline versions could neutralize WT, albeit 30.3- to 83.5-folds less potently than the
298 mature antibodies, they lacked neutralizing activity against the tested variants (Figures 4A,
299 S4A), confirming the essential role of affinity maturation introduced residues for

300 neutralization breadth. To understand the roles of the newly observed novel dominant
301 mutation G26E, we reverted 26E to a germline 26G in matured ZL54, ZL58, ZL112, and
302 ZL611 (Figure 4A-B). Neutralization assays demonstrated that four engineered antibody
303 mutants (ZL54-E26G, ZL58-E26G, ZL112-E26G, and ZL611-E26G) showed compromised
304 neutralization potency against most circulating SARS-CoV-2 variants compared to their
305 parental antibodies (2.5- to 10.3-fold increase in mean IC_{50}; $P < 0.01$). Most notably,
306 neutralization was most severely impaired against recently emerging Omicron subvariants,
307 including KP.3.1.1, LP.8.1.1 and NB.1.8.1 (3.5- to 72-fold IC_{50} increase), highlighting
308 G26E's role in conferring continued cross-neutralization (Figures 4A-B, S4B-D).
309
310 Oppositely, we introduced the seven somatic hypermutations into the germline version of
311 previously clinically approved mAbs BD-604, P2F-1C11, and LY-CoV488, to generate their
312 artificially affinity matured versions designated as BD-604(7), P2F-1C11(7), and LY-
313 CoV488(7), to investigated whether the identified SHM signature would confer increased
314 breadth to neutralize Omicron subvariants (Figure 4C). Neutralization results show that
315 these specific mutations indeed conferred the three antibodies with enhanced neutralizing
316 activity against Omicron subvariants, including BA.1, BA.4/5, BA.2.75, BQ.1.1, XBB.1.9.1
317 (Figure 4C). Of note, these mutations even enabled BD-604 to neutralize EG.5, HK.3, HV.1,
318 JD.1.1 and JN.1. Overall, the neutralizing activity of the artificially affinity matured mAbs
319 display 380-, 13.1-, 2.7-folds improvement compared to their parent antibodies,
320 respectively (Figure 4D). These results show that the efficacy of clinically escaped
321 therapeutics antibodies can be restored by the introduction of the seven convergent
322 hotspot mutations, increasing their neutralizing breadth.
323
324 To structurally define the contribution of SHM-introduced residues to RBD binding by VH3-
325 53/3-66 antibodies, we determined a cryo-EM structure of the Omicron LP.8.1.1 spike (S-
326 6P) in complex with the ZL58 Fab, a representative VH3-53/3-66 antibody. We included
327 the ZL525 Fab, an R1-32-like public antibody known to induce RBD opening[7], to enhance
328 RBD structural resolution through increased molecular mass. We only observed the 1:3:3

329 S-trimer:ZL58:ZL525 complex, with all three RBDs adopting an "up" conformation (Figures
330 4E, S5-6). Based on the locally refined RBD-Fab map at 2.93 Å resolution (Figures S5-6),
331 we built an atomic model of the complex, enabling detailed characterization of the ZL58
332 binding interface (Figure 4F). As expected, ZL58 is a classical class 1 antibody that binds
333 to the apical head of RBD, overlapping the receptor-binding motif. Buried surface area
334 analysis reveals that ZL58 engages the RBD primarily through HCDR1, HCDR2, HCDR3,
335 LCDR1, and LCDR3, with the heavy and light chains contributing 689 Å² and 192 Å²,
336 respectively. The binding interface consists of a network of hydrophilic and hydrophobic
337 interactions. Hydrophilic contacts, including 16 hydrogen bonds and 2 salt bridges, involve
338 RBD residues $R403_S$, $D420_S$, $Y421_S$, $S455_S$, $K460_S$, $Y473_S$, $N474_S$, $A475_S$, $G476_S$, $N477_S$,
339 $K478_S$, $N487_S$, $Y489_S$, and $H505_S$ that interact with HCDR residues $E26^H$, $A28^H$, $R31^H$,
340 $N32^H$, $Y33^H$, $P53^H$, $G54^H$, $S56^H$, $R97^H$, and LCDR residues $D28^L$ and $Y32^L$. Hydrophobic
341 contacts involve RBD residues $G416_S$, $Y421_S$, $L456_S$, $Y473_S$, $A475_S$, $G476_S$, and $Y489_S$
342 packing against HCDR residues $I27^H$, $A28^H$, $Y33^H$, $Y52^H$, $P53^H$, $F58^H$, and $P100^H$. Notably,
343 many of these interactions are mediated by residues introduced through somatic
344 hypermutation, including $E26^H$, $I27^H$, $A28^H$ and $R31^H$ in HCDR1, and $P53^H$ and $F58^H$ in
345 HCDR2, underscoring the indispensable role of SHM in developing ZL58's breadth. In
346 particular, the newly observed dominant mutation $G26E^H$ introduced a salt bridge to $K478_S$,
347 and a network of charged hydrogen bonding interactions at the RBD binding interface
348 (Figure 4F). Collectively, our structural and functional analyses demonstrate that these
349 convergent somatic mutations contribute substantially to the exceptional neutralization
350 breadth, while the HCDR3 loop and paired light chain also likely constitute the determinants
351 of broad-spectrum recognition. Nevertheless, these IGHV hypermutations not only
352 strengthen neutralizing activity but may also confer resilience against emerging antigenic
353 variants.
354

355 **Artificial intelligence-driven discovery of rare and resilient VH3-53/3-66 bnAbs**

356 The existence of bnAbs such as BD55-1205 underscores the feasibility of identifying rare
357 and resilient bnAbs from individuals convalesced or vaccinated during the early phases of

358 a pandemic. We hypothesized that the breadth of neutralization increases with the
359 accumulation of convergent somatic mutations within hotspot paratope residues. To test
360 this, we deeply profiled the antibody repertoires from a cohort of 85 WT vaccine recipients
361 for rare and resilient VH3-53/3-66-encoded bnAbs. Using the WT S-6P probe, we isolated
362 14,342 antigen-specific B cells from the 85 donors (Figure S7A) and identified 785
363 monoclonal antibodies with class 1 features that derived from VH3-53/3-66 germlines
364 across ten distinct B cell subsets (Figure S7B-C). We then developed a deep learning
365 workflow incorporating pre-trained language models with dedicated encoders for antigen
366 and antibody sequences (Figure 5A-C). Antigen features were embedded using ESM2[40],
367 while antibody heavy and light chains were processed using immunoglobulin-specific
368 RoFormer[41] models trained on 1.2 billion heavy and 210 million light chain sequences from
369 the OAS database[42]. These features were further integrated using convolutional layers and
370 subsequently processed through stacked fully connected (FC) layers to predict the final
371 binding affinity. The model was fine-tuned on a curated dataset of 3,897 RBD-binding
372 antibodies with variant-specific affinity measurements, enabling accurate and variant-
373 aware binding affinity prediction. This model was applied to screen the 785 VH3-53/3-66
374 antibodies isolated from vaccinees exposed only to ancestral WT antigen. The top 15
375 candidates, selected based on high predicted affinity for WT, BA.1, and KP.3.1.1, were
376 prioritized for experimental validation (Figure 5D). All the top 15 mAbs showed potent
377 neutralizing activities for both WT and BA.1 strain, and 7 of which cross-neutralize the
378 recent circulating JN.1 variant (Figure 5E). One lead candidate, R102-9, shows potent
379 neutralization to all the five tested strains (Figure 5E). Further deeper validation reveals
380 that R102-9 effectively neutralized all the 19 tested variants, with a mean IC_{50} of 87.8 ng/mL
381 (Figure 5F). The neutralization potency and breadth of R102-9 was superior than BD55-
382 1205 and comparable with the 15 elite VH3-53/3-66 mAbs (Figures 5F, 3A). R102-9 also
383 potently inhibited authentic SARS-CoV-2 WT, XBB.1.5, XBB.1.16, EG.5 and JN.1 isolates
384 with IC_{50} values ranging from 2.6 ng/mL to 37.7 ng/mL (Figure S7D).
385
386 **Accumulated SHMs conferred broad breadth and potency to R102-9**

387 Genetic analysis indicates that R102-9 is encoded by VH3-66, JH4, VK3-15 and JK1, and
388 accumulates ten SHMs in the variable region of the VH3-66 heavy chain. Six of these
389 mutations (F27I^{H}, T28I^{H}, S31R^{H}, V50L^{H}, S53P^{H}, and Y58F^{H}) correspond to previously
390 identified convergent SHMs (Figure 6A), suggesting that our AI modeling is able to
391 recapitulate key SHM features of broadly neutralizing VH3-53/3-66 monoclonal antibodies.
392 Interestingly, introduction of a single G26E^{H} mutation into R102-9 further enhances its
393 neutralization potency (Figure 6B). Notably, the somatically engineered variant, designated
394 R102-9(G26E), exhibits improved potency and breadth compared to both SA55[3] and Vir-
395 7229[43] (Figure 6C). To structurally understand R102-9's exceptional breadth, we
396 determined a cryo-EM structure of the Omicron JD.1.1 S-trimer (S-6P) in complex with
397 R102-9 Fab and C092 Fab, another R1-32-like antibody we previously studied[7] (Figures
398 6D, S8-9). Similar to the structure of Omicron LP.8.1 S-trimer bound to ZL58 Fab, we
399 observed a 1:3:3 S-trimer:R102-9:C092 complex and resolved the antibody bound RBD to
400 2.97 Å resolution through local refinement (Figure 6D, S8-9). Our structure reveals that
401 R102-9 recognizes an epitope nearly identical to that of ZL58, engaging in an extensive
402 network of hydrogen bonds and hydrophobic interactions, with multiple key contacts
403 contributed by the six SHM-derived, predominantly hydrophobic, residues including F27I,
404 T28I, S31R, V50L, S53P, and Y58F (Figure 6E). Taken together, our findings confirm that
405 the predictive power of AI platform is able to identify rare, broadly neutralizing antibodies
406 with exceptional resilience to viral escape from a large collection of mAbs isolated from
407 donors exposed to only WT antigens early in pandemic. These findings demonstrate that,
408 when properly trained, the AI platform can recognize weakened convergent SHM
409 signatures within the noisy immune repertoires of early-pandemic donors to identify rare
410 broadly neutralizing antibodies.
411
412 **R102-9 protects from SARS-CoV-2 Beta and JN.1 infection**
413 We next evaluated whether R102-9 protects against SARS-CoV-2 infection in vivo, using
414 BALB/c mice challenged with the Beta variant, which carries the K417N substitution highly
415 detrimental to germline VH3-53/3-66 antibodies. BALB/c mice permit Beta infection with a

416 higher viral load and severer pulmonary pathology in the lungs compared with those
417 infected with Alpha, Delta, and Omicron[44,45]. In both prophylaxis and therapy study, each
418 mouse received a single intraperitoneal injection of either 10 mg/kg high dose, 5 mg/kg
419 intermediate dose or 2 mg/kg low dose of R102-9 (Figure 7A), respectively. Another two
420 group of mice was injected with a positive control SA55[3] and negative control Influenza-
421 specific mAb MEDI8852[46] at a high dose of 10 mg/kg. 12 hours before or 24 hours post
422 the mAbs injection, each animal was challenged intranasally with $5x10^4$ focus-forming units
423 (FFUs) of live SARS-CoV-2 Beta strain. The mice were sacrificed for analysis at 3 days
424 post-infection (dpi) when high viral loads and acute lung injury were consistently observed
425 in the negative control group (Figure 7A). In both therapeutic and prophylactic studies,
426 R102-9 conferred robust protection at all tested doses (high, intermediate, and low).
427 Notably, viral titers, determined by FRNT, were undetectable in the lungs of all R102-9-
428 treated mice, whereas animals receiving MEDI8852 exhibited measurable pulmonary viral
429 loads (Figure 7B-C). These findings demonstrate R102-9's superior efficacy in suppressing
430 Beta variant replication within murine lung tissues. Next, to understand whether or not
431 R102-9 protects against infection-induced lung injury in BALB/c mice, we performed
432 pathological analysis on lung specimens. Compared with MEDI8852-treated mice, mice
433 treated with high or intermediate dose of R102-9 showed mild interstitial alveoli
434 inflammation with minor septal infiltration and congestion (Figure 7D-E). There was limited
435 apparent peribronchiolar infiltration, and the bronchiolar epithelia appeared normal. Mice
436 treated with the low dose showed limited areas of bronchiolar epithelial cell swelling and
437 detachment (Figure 7D-E) together with mild peribronchiolar infiltration and mild alveolar
438 septal infiltration. In contrast, control mice showed large patchy areas of alveolar wall and
439 alveolar space involvement by inflammatory infiltrates and exudation (Figure 7D-E).
440
441 To further evaluate the protective efficacy of R102-9 against emerging variants, we
442 challenged K18-hACE2 mice with 10,000 FFU of the JN.1 variant (Figure S10A). Mice
443 received a low dose of R102-9 (2 mg/kg) either 12 hours before or after infection.
444 MEDI8852 and SA55 were included as positive and negative controls, respectively. Both

445 prophylactic and therapeutic administration of R102-9 markedly reduced pulmonary viral
446 titers (Figure S10B-C), and histopathological analysis revealed attenuated lung
447 inflammation and improved tissue histological features in treated animals (Figure S10D-E).
448 These results demonstrate that the AI-discovered antibody R102-9 exhibits robust
449 protective efficacy in animal model.
450
451 **Discussion**
452 This study delineates the molecular evolution of durable, broadly neutralizing VH3-53/3-66
453 antibodies in elite neutralizers with sequential Omicron breakthrough infections. Repeated
454 antigenic exposures drive convergent somatic hypermutations that markedly expand
455 breadth and potency, preserving neutralization even against the most recent immune-
456 evasive variants (e.g., JD.1.1, LP.8.1.1, NB.1.8.1). Our findings extend prior reports that
457 VH3-53/3-66 antibodies are public clonotypes[11,47,48] by showing how recurrent antigenic
458 exposures sculpt their affinity maturation into pan-variant neutralizers. Overall, despite the
459 high variability of the VH3-53/3-66–targeted epitope, we show that the immune system can
460 still generate broad and potent neutralizing antibodies directed against this epitope.
461
462 Previous studies showed that VH3-53/3-66 antibodies isolated early in the pandemic (e.g.,
463 CB6, P2C-1F11) had limited breadth owing to minimal SHMs, showing vulnerability to
464 K417N/T substitutions[49]. Although convergent SHMs that improved affinity and breadth
465 against early Omicron variants were reported[26-28,50], these antibodies fail to effectively
466 neutralize more recent escape mutants (Figure 3A). In contrast, our analysis indicates that
467 sequential Omicron exposures drive accumulation of additional key SHMs (G26E, F27I,
468 T28I, S31R, V50L, S53P, Y58F) that collectively strengthen paratope-RBD interactions.
469 Structurally, ZL58 and R102-9 illustrate how these SHMs introduce compensatory contacts
470 at escaped epitopes, paralleling VacBB-551[23] and 19-77[51] yet exhibiting greater resilience
471 to A475V/L455F/F456L in JD.1.1 and LP.8.1.1. These results reveal two layers of epistasis:
472 mutations within SARS-CoV-2 spike epitopes show strong genetic interactions [52], and the
473 SHMs accrued during affinity maturation of VH3-53/3-66 broadly neutralizing antibodies

474 likewise follow a cooperative, epistatic pattern. These bidirectional genetic interactions
475 provide compelling evidence for co-evolutionary dynamics between viral antigens and
476 neutralizing antibodies.
477
478 Critically, we show that germline-reverted antibodies lose breadth, while introducing these
479 convergent SHMs into clinically escaped antibodies (BD-604, P2C-1F11, LY-CoV488)
480 restores neutralization against Omicron lineages (380-fold potency increase for BD-604).
481 This strategy diverges from conventional engineering approaches by leveraging natural
482 selection patterns[51]. Restoring the activity of escaped antibodies through SHM grafting
483 provides a rapid strategy to update clinical antibody candidates against emerging variants.
484 For vaccines, our data suggest that iterative Omicron exposures, akin to prime-boost
485 regimens with divergent spikes[53], foster bnAb maturation. Notably, mucosal IgA^+ VH3-53/3-
486 66 clones were enriched in elite neutralizers, hinting at synergistic systemic/mucosal
487 responses, an underexplored area in current vaccine design[54].
488
489 While most VH3-53/3-66 antibodies isolated early in the pandemic lack breadth, rare
490 exceptions such as BD55-1205 suggest that some clones were pre-adapted to future
491 variants[19]. Our AI-guided discovery of R102-9 advances bnAb identification by showing
492 that machine learning can extract critical SHM signatures from early repertoires, a
493 capability not achieved previously[19,55]. R102-9 neutralizes all 19 tested variants (mean IC_{50}:
494 87.8 ng/mL) and protects mice against Beta and JN.1 challenge, outperforming BD55-1205
495 and matching elite antibodies induced by sequential Omicron breakthrough infections. This
496 approach diverges from prior epitope-focused AI tools[56,57] by linking SHM patterns to
497 function, enabling rapid mining of resilient bnAbs without extensive laborious experimental
498 screening[58]. Structurally, R102-9's mode of action parallels ZL58, underscoring SHM
499 signatures as predictive features of breadth. This study establishes a new workflow for
500 pandemic preparedness: mining early immune repertoires prospectively for rare, resilient
501 bnAbs rather than waiting for elite neutralizers to emerge.
502

In summary, we chart a molecular roadmap for the evolution of broadly protective VH3-53/3-66 antibodies under sequential exposures, identifying seven key convergent SHMs that collectively overcome SARS-CoV-2 immune evasion. Integrating repertoire analysis, antibody engineering, structural biology, and AI, we show that these naturally selected patterns can restore escaped therapeutic antibodies and uncover rare, potent bnAbs from early-pandemic repertoires. This work elucidates the coevolution between a public antibody class and the viral surface antigen and offers actionable guidance for variant-resilient countermeasures, establishing a practical, machine learning–enabled antibody discovery workflow for pandemic preparedness.

Limitations of the study

While this study provides key insights into the evolution of broad SARS-CoV-2 neutralization, several limitations should be noted. First, due to the limited number of B cells that could be sorted from each individual, our B cell sequencing was performed on pooled samples rather than analyzing cells from each donor separately. This approach may obscure individual-specific immune responses and limit our ability to track clonal evolution at the single-donor level. Second, current analytical constraints prevented precise identification of somatic hypermutations in the CDR3 region. As a result, our study primarily focused on SHM patterns in CDR1 and CDR2, which may not fully capture the mutational landscape contributing to antibody breadth and potency. Third, the AI-based prediction framework, while promising, was specifically trained and validated for the VH3-53/3-66 antibody lineage, expansion to other antibody classes (e.g., VH1-69 or VH1-58) will be necessary to determine its generalizability. These limitations highlight important directions for future research, including single-donor sequencing approaches, improved CDR3 mutation analysis methods, and development of lineage-agnostic prediction models. Fourth, the SHM convergence signal is relatively weak among antibody sequences isolated early in the pandemic, and obtaining sufficiently large datasets for a given clonal type may be challenging.

532 **Resource availability**

533 **Lead contact**

534 Further information, requests for resources, and reagents should be directed to and will be
535 fulfilled by the lead contact, Jincun Zhao (jincun_zhao@gird.cn).

536

537 **Materials availability**

538 All expression plasmids and proteins generated in this study are available upon request.

539

540 **Data and code availability**

541 Single-cell RNA-seq raw data sequencing data have been deposited at National Genomics
542 Data Center and are available under the accession BioProject ID PRJCA049739. Cryo-
543 EM density maps for the SARS-CoV-2 Omicron LP.8.1 spike in complex with ZL58 and
544 ZL525 Fabs, and the Omicron JD.1.1 spike in complex with R102-9 and C092 Fabs have
545 been deposited in the Electron Microscopy DataBank (EMDB) with accession codes EMD-
546 66901, EMD-66902, EMD-66903, and EMD-66904 respectively. The related atomic
547 models have been deposited in Protein Data Bank (PDB) under accession codes 9XI9,
548 9XIB, 9XIC, and 9XID respectively. All maps and models are publicly available as of the
549 date of publication. Accession numbers are listed in the key resources table. The code
550 used in AI-based antibody discovery is available at: https://github.com/YidongSong/Ab-Ag-
551 AFF/tree/main. Any additional information required to reanalyze the data reported in this
552 work paper is available from the lead contact upon request.

553

554 **Methods**

555 **Cells and viruses**

556 Expi293F cells (Thermo Fisher Scientific, A14527) were maintained in Expi293F
557 Expression medium (Thermo Fisher Scientific, A1435101) at 37 °C by shaking at 120 rpm
558 under a humidified atmosphere with 8% CO_2. Human embryonic kidney (HEK) 293T (ATCC,
559 Cat# CRL-3216), HEK293T-hACE2, Vero E6 (ATCC, Cat# CRL-1586) and Huh-7 cells
560 (JCRB, Cat# 0403) were maintained in 10% FBS supplemented Dulbecco's Modified

561 Eagle's Medium (DMEM) at 37°C, 5% CO_2. The authentic SARS-CoV-2 viruses, including
562 WT, BA.5, XBB.1, XBB.1.5, XBB.1.16, EG.5, and JN.1 were isolated from COVID-19
563 patients and preserved in Guangzhou Customs District Technology Center BSL-3
564 Laboratory. Experiments related to authentic SARS-CoV-2 viruses were conducted in
565 Guangzhou Customs District Technology Center BSL-3 Laboratory.
566
567 **COVID-19 patient and vaccine recipient enrollment**
568 A cohort of 68 individuals with confirmed SARS-CoV-2 infection was enrolled between
569 December 2022 and December 2023. All participants experienced sequential breakthrough
570 infections with the BA.5 and XBB subvariants during the two major epidemic waves in this
571 period. Each individual had previously received three doses of an inactivated COVID-19
572 vaccine (Sinopharm/CoronaVac). SARS-CoV-2 infection was confirmed by RT–PCR. An
573 additional cohort of 85 vaccine recipients (WT-vaccinated group) was included. This group
574 comprised 43 individuals who received three doses of an inactivated vaccine
575 (Sinopharm/CoronaVac), 11 who received two doses of an inactivated vaccine
576 (Sinopharm/CoronaVac) followed by one adenoviral vector vaccine (CanSino), 18 who
577 received two doses of an inactivated vaccine (Sinopharm/CoronaVac) followed by one
578 mRNA vaccine (BioNTech), and 13 who received three doses of an mRNA vaccine
579 (BioNTech). All participants provided written informed consent prior to enrollment. The
580 study was conducted under the approval and supervision of the Ethics Committee of the
581 First Affiliated Hospital of Guangzhou Medical University (Ethics IDs: ES-2020-65, ES-
582 2023-015-01, ES-2021-78).
583
584 **Protein expression and purification**
585 The gene encoding the SARS-CoV-2 spike (S) protein extracellular domain (residues 14–
586 1211) was cloned into the pcDNA3.1(+) mammalian expression vector. The construct
587 included an N-terminal mu-phosphatase signal peptide, an "R" substitution at the furin
588 cleavage site (residues 682–685), a C-terminal TEV protease cleavage site, a T4 fibritin
589 trimerization motif, and a His_6 tag, designated "S-R", as previously described[59]. For

590 structural studies, a stabilized S-protein variant ("S-GSAS/6P") carrying six proline
591 substitutions (F817P, A892P, A899P, A942P, K986P, and V987P) was engineered and
592 termed "S-6P"[60]. The RBD (residues 319–541) was constructed with an N-terminal mu-
593 phosphatase signal peptide and a C-terminal Avi-tag followed by a 6×His tag, and similarly
594 cloned into pcDNA3.1(+). All constructs were transiently transfected into Expi293F cells
595 using polyethylenimine. Proteins were purified by immobilized metal affinity
596 chromatography (IMAC) under previously established conditions[7,61], aliquoted, flash-
597 frozen in liquid nitrogen, and stored at −80°C.
598
599 **B cell sorting and single-cell RNA sequencing**
600 B cells were enriched from peripheral blood mononuclear cells (PBMCs) using an EasySep
601 Pan-B Cell Magnetic Enrichment Kit (STEMCELL Technologies) and incubated with 2
602 mg/mL of Avi-tagged JD.1.1 or KP.3.1.1 RBD on ice for 30 minutes. Cells were
603 subsequently washed with PBS containing 2% fetal bovine serum (FBS; Hyclone) and
604 stained with a cocktail of anti-human CD3 (PerCP-Cy5.5; BioLegend, Cat# 410710) and
605 anti-human CD19 (FITC; BioLegend, Cat# 348205) antibodies. To identify RBD-specific B
606 cells, PE- and APC-conjugated anti-Avi-tag antibodies (BioLegend, Cat# 637309 and
607 #637307) were included, enabling the gating of double-positive cells. RBD-specific double-
608 positive B cells were gated as $DRAQ7^{-}CD3^{-}CD19^{+}$ and sorted on a FACSAria II cell sorter.
609 Sorted B cells were processed using the following 10x Genomics kits: Chromium Next
610 GEM Single Cell 5′ Kit v3, Library Construction Kit, Chromium Next GEM Chip K Single
611 Cell Kit, Chromium Single Cell Human BCR Amplification Kit, and Dual Index Kit TT Set A.
612 Following GEM generation and barcoding, cDNA was synthesized via GEM reverse
613 transcription (RT) and purified using bead-based cleanup. The purified cDNA was amplified
614 for 10–14 cycles and cleaned with SPRIselect beads. cDNA quality and concentration were
615 assessed using a 4200 TapeStation System (Agilent Technologies). BCR enrichment was
616 performed on full-length cDNA. Libraries for gene expression and BCR analysis were
617 prepared according to the Chromium Next GEM Single Cell 5′ Reagent Kits v3 User Guide,
618 with PCR cycle numbers adjusted based on cDNA concentration. Final libraries were

619 sequenced on a NovaSeq S4 instrument (Illumina), with a target median depth of 50,000
620 read pairs per cell for gene expression libraries and 5,000 read pairs per cell for BCR
621 libraries.
622
623 **Single-cell RNA-seq analysis**
624 Single-cell RNA sequencing and BCR sequencing data were processed with Cell Ranger
625 (v7.0.0) and aligned to the GRCh38-2020 human reference genome provided by 10x
626 Genomics. Count matrices were analyzed in R (v4.0.2) using the Seurat package (v5.2.1).
627 Cells were filtered based on the following criteria: mitochondrial gene content < 20% and
628 feature count < 4,000. The three sequencing runs were integrated using log-normalized
629 counts and a canonical correlation analysis approach with 2,000 highly variable genes.
630 The combined object was subjected to principal component analysis (PCA); the top 30
631 principal components were used for Uniform Manifold Approximation and Projection
632 (UMAP) embedding and neighbor identification. Clustering was performed at a resolution
633 of 0.6. For BCR analysis, filtered contig annotations from Cell Ranger V(D)J were imported
634 into R and processed using the scRepertoire package (v1.1.3). For mutation analysis,
635 heavy and light chains from monoclonal antibodies and single-cell BCRs were annotated
636 with IgBLAST (v1.14.0) using IMGT human reference genes. Resulting files were parsed
637 with Change-O (v0.4.6). Mutation frequency was computed as previously described[62] using
638 the *calcObservedMutations* function from SHazaM (v1.0.2), by counting nucleotide and
639 amino acid mismatches relative to the germline sequence within the variable region up to
640 the CDR3.
641
642 **Expression of monoclonal antibody**
643 Heavy- and light-chain variable region genes (VH and VL) were cloned into human IgG1
644 expression vectors using the CloneExpress II One-Step Cloning Kit (Vazyme, Cat# C112-
645 02). HEK293F cells were cultured to a density of 1×10^6 cells/mL and co-transfected with
646 equimolar amounts of heavy- and light-chain plasmids using the EZ Cell Transfection
647 Reagent (Life-iLab Biotech, Cat# AC04L092). Transfected cells were maintained in CD 293

648 TGE medium (ACRO, Cat# CM-1156) supplemented with 10% CD Feed X (ACRO, Cat#
649 CF-1116-12) at 37 °C under 5% CO_2 with shaking at 120 rpm. Six days post-transfection,
650 culture supernatants were harvested, clarified by centrifugation, and filtered through 0.22-
651 μm membranes (Merck Millipore, Cat# SLGP033NS). Filtered supernatants were
652 incubated with Protein A Resin (Genscript, Cat# L00210) for 2 hours at room temperature
653 for affinity purification. After extensive washing, bound antibodies were eluted with 0.1 M
654 sodium citrate (pH 3.25) and immediately neutralized with 1 M Tris-HCl (pH 8.8). Purified
655 antibodies were concentrated in PBS using Amicon Ultra centrifugal filters (Merck Millipore,
656 Cat# UFC810096) and stored at −80 °C.
657
658 **Pseudovirus production and neutralization assay**
659 The pseudovirus was constructed based on a VSV-G pseudotyped backbone in which the
660 native G gene was replaced by a firefly luciferase (Fluc) reporter gene and the SARS-CoV-
661 2 spike protein was incorporated. The SARS-CoV-2 WT and variant spike genes were
662 codon-optimized and cloned into the pcDNA3.1 vector to generate the expression plasmid
663 pcDNA3.1-SARS-CoV-2-spikes. HEK293T cells were trypsinized, resuspended at a
664 concentration of 5–7 × 10^5 cells/mL, and seeded into a T75 flask with 15 mL of cell
665 suspension. After overnight incubation at 37 °C under 5% CO_2, cells reached 70–90%
666 confluence. The medium was then removed, and cells were infected with 15 mL of G*ΔG-
667 VSV pseudotyped virus (Kerafast) at a titer of 7.0 × 10^4 $TCID_{50}$/mL. Concurrently, 30 μg of
668 the spike expression plasmid was transfected using Lipofectamine 3000 (Invitrogen,
669 L3000015) according to the manufacturer's instructions, followed by 6–8 h of incubation at
670 37 °C with 5% CO_2. After transfection, the supernatant was aspirated, and cells were gently
671 washed twice with PBS containing 1% FBS. Fresh complete DMEM (15 mL) was added,
672 and cells were cultured for 24 h. The supernatant was collected, and another 15 mL of
673 fresh complete DMEM was added for a further 24 h incubation. The two batches of
674 supernatant were combined, centrifuged at 1000 × g for 10 min, filtered through a 0.45-μm
675 membrane, aliquoted, and stored at −80 °C for subsequent use. For pseudovirus
676 neutralization assays, Huh-7 cells were employed. Test antibodies were serially diluted six

677 times in complete DMEM in 96-well plates. Pseudovirus was added at a concentration of
678 1.3 × 10^4 $TCID_{50}$/mL per well, and the mixture was incubated for 1 h at 37 °C under 5%
679 CO_2. Then, 100 μL of trypsinized Huh-7 cells (2 × 10^5 cells/mL, 0.25% Trypsin-EDTA, Gibco)
680 were added to each well. After 24 h of incubation, the supernatant was carefully removed
681 leaving 100 μL per well, and 100 μL of luciferase substrate (PerkinElmer, 6066769) was
682 added. Following a 2-minute incubation at room temperature, 150 μL of the lysate was
683 transferred to a white opaque 96-well microplate (PerkinElmer), and luminescence was
684 measured using an Ensight multimode plate reader (PerkinElmer). All samples were
685 assayed in duplicate. The half-maximal inhibitory concentration (IC_{50}) was determined by
686 four-parameter nonlinear regression analysis using GraphPad Prism version 8.0.
687
688 **SARS-CoV-2 authentic virus neutralization assay**
689 Antibodies were serially diluted in DMEM and mixed with 200 focus-forming units (FFU) of
690 authentic SARS-CoV-2 viruses, including WT, BA.5, XBB.1, XBB.1.5, XBB.1.16, EG.5, and
691 JN.1. Following incubation at 37 °C for 1 h, the antibody–virus mixtures were applied to
692 Vero E6 cells cultured in 96-well plates and incubated for 1 h at 37 °C under 5% CO_2. After
693 removal of the inoculum, each well was overlaid with 100 μL of pre-warmed 1.6%
694 carboxymethylcellulose and cultured for 24 h. The overlay was then aspirated, and cells
695 were fixed with 4% paraformaldehyde (Biosharp, Cat# BL539A) and permeabilized with
696 0.2% Triton X-100 (Sigma, Cat# T8787). Subsequently, cells were incubated for 1 h at
697 37 °C with a human anti-SARS-CoV-2 nucleocapsid protein monoclonal antibody
698 (generated by in-house screening). After three washes with 0.15% PBST, an HRP-
699 conjugated goat anti-human secondary antibody (Jackson ImmunoResearch, Cat# 609-
700 035-213) was applied and incubated for 1 h at 37 °C. Following three additional washes
701 with 0.15% PBST, foci were developed using TrueBlue Peroxidase Substrate (KPL, Cat#
702 50-78-02) and quantified with an ELISPOT reader (Cellular Technology Ltd.). The 50% foci
703 reduction neutralization titer ($FRNT_{50}$) was determined by four-parameter nonlinear
704 regression analysis in GraphPad Prism 8.0.
705

706 **Binding affinity prediction, Virtual screening of the VH3-53/3-66 antibody library**

707 To predict antibody–antigen binding affinities, we employed pre-trained language models
708 tailored to the distinct properties of antigens and antibodies (Figure 5A). Antigen
709 sequences were encoded using ESM2[40] to capture conserved evolutionary features. For
710 antibody sequences, we developed specialized language models trained on
711 immunoglobulin-specific data. As illustrated in Figure 5B, we preprocessed unlabeled
712 antibody sequences using a Unique Amino Acid (UAA) Tokenizer, and pre-trained the
713 RoFormer model[41] on 1.2 billion heavy chain and 210 million light chain sequences from
714 the OAS database[42] via masked language modeling, yielding separate pre-trained models
715 for heavy and light chains. Features were extracted from multiple chains using pre-trained
716 models, and binding affinity was predicted through the convolutional neural network (CNN)
717 module comprising consecutive convolutional layers and stacked fully connected (FC)
718 layers (Figure 5C). The model was fine-tuned on a custom-curated dataset integrating in-
719 house and public sources[16,18,52,63-68], comprising 3,897 RBD-binding antibodies with
720 quantitative affinity measurements across multiple SARS-CoV-2 variants, including WT,
721 Beta, Delta, BA.1, BA.5, and the recently emerged KP.3.1.1 strain. This enabled the model
722 to learn variant-specific binding profiles with high accuracy. We then applied this optimized
723 framework to virtually screen a library of 785 VH3-53/3-66-encoded monoclonal antibodies.
724 Each antibody was computationally paired with target antigens, and binding affinities were
725 predicted to rank candidates. Iterative screening across variants enabled identification of
726 antibodies with strong and broad neutralizing potential.

727

728 **Cryo-EM sample preparation and data collection**

729 For the LP.8.1 S-6P:ZL58 Fab:ZL525 Fab complex, LP.8.1 S-6P at 5.1 mg/ml was mixed
730 with ZL58 Fab and ZL525 Fab at 1:1:1 molar ratio. For the JD.1.1 S-6P:R102-9 Fab:C092
731 Fab complex, JD.1.1 S-6P at 6 mg/ml was mixed with R102-9 Fab and C092 Fab at 1:1:1
732 molar ratio. After incubation for 1 min at room temperature, 3 μL of each mixture
733 supplemented with 0.1% octyl-glucoside (Sigma-Aldrich, Cat# V900365) was applied to
734 freshly glow-discharged (30 s at 15 mA) holey carbon grids (Quantifoil, Cu R1.2/R1.3). The

735 grids were blotted for 2.5 s with a blot force of 4, and then plunge-frozen into liquid ethane
736 using a Vitrobot Mark IV (Thermo Fisher Scientific) at 18 °C and 100% humidity. Cryo-EM
737 grids were loaded onto a 300 keV Titan Krios electron microscope equipped with a Falcon
738 4 direct electron detector with SelectrisX energy filter (slit width 10 eV). Movie stacks were
739 automatically collected using EPU software with the electron event representation (EER)
740 mode at a nominal magnification of ×160,000 with a pixel size of 0.73 Å and a defocus
741 range from -0.6 to -2.0 μm. Each stack was recorded and exposed at a dose rate of 6.53
742 e^-/pixel/s for 4.08 s, resulting in a total dose of ~50 $e^-/Å^2$.
743
744 **Cyro-EM data processing**
745 Data processing was performed using cryoSPARC v4.2.0. After patch motion correction,
746 contrast transfer function (CTF) estimation and removal of poor-quality micrographs,
747 696,375 and 955,134 particles were picked from 6,117 micrographs of the LP.8.1 S-
748 6P:ZL58 Fab:ZL525 Fab dataset and 6,579 micrographs of the JD.1.1 S-6P:R102-9
749 Fab:C092 Fab dataset, respectively. Subsequent 2D classification was used to remove
750 low-quality particles, and well-defined particles were selected for ab-initio reconstruction
751 and heterogeneous refinement. For the JD.1.1 S-6P:R102-9 Fab:C092 Fab dataset,
752 additional non-uniform refinement was applied. Particles from classes exhibiting high-
753 resolution features were used to train the topaz picker, followed by particle re-extraction,
754 2D classification, ab-initio reconstruction, heterogeneous refinement, and non-uniform
755 refinement. This workflow yielded 3:3:3 LP.8.1 S-6P:ZL58 Fab:ZL525 Fab and 3:3:3 JD.1.1
756 S-6P:R102-9 Fab:C092 Fab complexes at global resolutions of 2.75 Å and 2.76 Å,
757 respectively. To improve the resolution of the RBD-Fab binding interface, a focused mask
758 covering the RBD-Fab region was generated from the globally refined maps. Using this
759 mask, 3D classification was performed to select particles with well-defined local features,
760 followed by local refinement, yielding focused maps at 2.93 Å and 2.97 Å, respectively.
761 Map resolutions were estimated at the 0.143 criterion of the phase-randomization-
762 corrected Fourier shell correlation (FSC) curve calculated between two independently
763 refined half-maps multiplied by a soft-edged solvent mask. The data processing

764 procedures are summarized in Figures S5, S6, S8, and S9. The estimated B-factors of
765 maps are listed in Table S3.
766
767 **Model building and structural analysis**
768 A previously determined structure of the SARS-CoV-2 WT S-trimer in complex with 3 YB8-
769 259 Fabs and 3 R1-32 Fabs (PDB: 8HC4) and a structure of the SARS-CoV-2 WT S1 in
770 complex with a YB9-258 Fab and an R1-32 Fab (PDB: 8HC5)[50] were used as starting
771 models and fitted into final refined maps in UCSF Chimera v1.15[69]. Iterative model building
772 and real space refinement were carried out in Coot v0.9.8.1[70] and PHENIX v1.19.1[71].
773 Model refinement statistics are summarized in Table S3. Interface analyses were
774 performed in PISA[72]. Structure figures were generated in UCSF Chimera v1.15.
775
776 **SARS-CoV-2 infection and treatment in mice**
777 In vivo prophylactic and therapeutic activities of R102-9 against SARS-CoV-2 Beta variant
778 were assessed in BALB/c mice. Infection studies were performed on 6 to 8 weeks-old
779 female mice, in animal biosafety level 3 (BSL-3) facilities of Guangzhou Customs District
780 Technology Center. Animal work was approved by the Animal Experimentation Ethics
781 Committee of the First Affiliated Hospital of Guangzhou Medical University (IACUC number:
782 20241225). Briefly, anesthetized mice were inoculated intranasally with 5×10^4 FFU of
783 SARS-CoV-2 Beta variant. 12 hours before inoculation (prophylactic group) or 24 hours
784 post inoculation (therapeutic group), mice received an intraperitoneal injection with 2, 5, or
785 10 mg/kg of R102-9. Mice injected with 10mg/kg MEDI8852 or SA55 were challenged with
786 the same dose of SARS-CoV-2 Beta as controls. Lungs were harvested at 3 days post
787 infection for virus titration and hematoxylin and eosin (H&E) staining. In the JN.1 protection
788 experiment, we intranasally challenged K18-hACE2 transgenic mice with 10,000 FFU of
789 the JN.1 variant under light anesthesia. R102-9 antibody was administered intramuscularly
790 at a dose of 2 mg/kg either 12 hours before (prophylactic group) or 12 hours after
791 (therapeutic group) viral challenge.
792

Statistical analysis

All statistical analyses were performed using R (v4.4.0) or GraphPad Prism (v8.0). Details of the specific tests used are provided in the corresponding figure legends. Participant characteristics were summarized as counts (n), medians with interquartile ranges (IQR), or frequencies with percentages. Chi-square and Fisher's exact tests were used to assess differences in categorical variables. Quantitative variables were compared using the Wilcoxon matched-pairs test for within-group analyses and the Mann–Whitney U test for comparisons between independent groups.

Acknowledgments

We thank the Huimin Feng at the Cryo-EM Facility of GIBH-CAS and Mei Li at Guangzhou Laboratory Bio-imaging Technology Platform for help with cryo-EM sample preparation and data collection. We thank the Drug Discovery Platform of Guangzhou National Laboratory for our ITC experiments, and we would be grateful to Tianyong Tu for his help with sample preparation. This study was supported by the National Natural Science Foundation of China (82025001, 82495200 and 82495203 to J.Z., 82341085 and 32570199 to X.X., and 82201932 to Q.Y.), the R&D Program of Guangzhou Laboratory (EKPG21-30-2 to J.Z. and SRPG22-002 to X.X.), the Basic Research Project of Guangzhou institutes of Biomedicine and health, Chinese Academy of Sciences (GIBHBRP24-02 to X.X.), the Natural Science Fund of Guangdong Province (2025A1515011245 to B.L.), and the Young Doctoral Starting Sail Project of the Guangzhou Municipal Science and Technology Bureau (2024A04J4195 to B.L.).

Author contributions

J.Z. and Q.Y. conceived and initiated the study; X.X. conceived and initiated the structural studies; X.H., Y.Y., H.Z, Z.H., and S.L. performed antibody isolation， expressed, and purified RBDs and antibodies; X.H., Y.Y., H.Z, Z.H., F.C.,Y.L., L.C., N.Z., J.X., C.C., and H.B. performed the pseudovirus neutralization assays; H.X. performed authentic virus neutralization assays and animal challenge assays under the supervision of A.Z.; Q.Y., T.

W., J.L, and M.W. performed the bioinformatic analysis; B.L., C.N., and Z.L. collected cryo-EM data under the supervision of X.X.; B.L., C.N., and Z.L processed cryo-EM data under the supervision of X.X.; B.L. C.N., Z.L., H.Y., and X.X. analyzed cryo-EM structures; Y.S. and F.W. performed the AI-based antibody prediction under the supervision of J.Y. and YD.Y.; Y.Q., X.H., B.L., Y.Y., H.Z., and Y. S. prepared the figures; Q.Y., B.L., and Y.S. wrote the paper with input from all authors. JX.Z., J.Y., YD.Y., A.Z., X.X., and J.Z. supervised the research.

Declaration of interests

The authors declare no competing interests.

References

1 Sok, D. & Burton, D. R. Recent progress in broadly neutralizing antibodies to HIV. *Nat Immunol* **19**, 1179-1188, doi:10.1038/s41590-018-0235-7 (2018).

2 Sun, X. *et al.* Broadly neutralizing antibodies to combat influenza virus infection. *Antiviral Res* **221**, 105785, doi:10.1016/j.antiviral.2023.105785 (2024).

3 Cao, Y. *et al.* Rational identification of potent and broad sarbecovirus-neutralizing antibody cocktails from SARS convalescents. *Cell Reports* **41**, doi:10.1016/j.celrep.2022.111845 (2022).

4 Wu, X. *et al.* Maturation and Diversity of the VRC01-Antibody Lineage over 15 Years of Chronic HIV-1 Infection. *Cell* **161**, 470-485, doi:10.1016/j.cell.2015.03.004 (2015).

5 Gruell, H. *et al.* Antibody-mediated neutralization of SARS-CoV-2. *Immunity* **55**, 925-944, doi:10.1016/j.immuni.2022.05.005 (2022).

6 Yuan, M. & Wilson, I. A. Structural Immunology of SARS-CoV-2. *Immunol Rev* **329**, e13431, doi:10.1111/imr.13431 (2025).

7 He, P. *et al.* SARS-CoV-2 Delta and Omicron variants evade population antibody response by mutations in a single spike epitope. *Nat Microbiol* **7**, 1635-1649, doi:10.1038/s41564-022-01235-4 (2022).

8 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).

9 Wu, Y. *et al.* A noncompeting pair of human neutralizing antibodies block COVID-19 virus binding to its receptor ACE2. *Science*, 1274-1278 (2020).

10 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, doi:10.1016/j.celrep.2020.108274 (2020).

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

12 Shi, R. *et al.* A human neutralizing antibody targets the receptor-binding site of SARS-

CoV-2. *Nature* **584**, 120-124, doi:10.1038/s41586-020-2381-y (2020).

13 Ju, B. *et al.* Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature* **584**, 115-119, doi:10.1038/s41586-020-2380-z (2020).

14 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).

15 Yuan, M. *et al.* Structural and functional ramifications of antigenic drift in recent SARS-CoV-2 variants. *Science* **373**, 818-823 (2021).

16 Cao, Y. *et al.* Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. *Nature* **602**, 657-663, doi:10.1038/s41586-021-04385-3 (2022).

17 Sheward, D. J. *et al.* Structural basis of broad SARS-CoV-2 cross-neutralization by affinity-matured public antibodies. *Cell Rep Med* **5**, 101577, doi:10.1016/j.xcrm.2024.101577 (2024).

18 Muecksch, F. *et al.* Affinity maturation of SARS-CoV-2 neutralizing antibodies confers potency, breadth, and resilience to viral escape mutations. *Immunity* **54**, 1853-1868.e1857, doi:10.1016/j.immuni.2021.07.008 (2021).

19 Jian, F. *et al.* Viral evolution prediction identifies broadly neutralizing antibodies to existing and prospective SARS-CoV-2 variants. *Nat Microbiol* **10**, 2003-2017, doi:10.1038/s41564-025-02030-7 (2025).

20 Cevik, M. *et al.* SARS-CoV-2, SARS-CoV, and MERS-CoV viral load dynamics, duration of viral shedding, and infectiousness: a systematic review and meta-analysis. *Lancet Microbe* **2**, e13-e22, doi:10.1016/S2666-5247(20)30172-5 (2021).

21 Yan, Q. *et al.* Deep immunoglobulin repertoire sequencing depicts a comprehensive atlas of spike-specific antibody lineages shared among COVID-19 convalescents. *Emerging Microbes & Infections* **13**, doi:10.1080/22221751.2023.2290841 (2024).

22 Kim, W. *et al.* Germinal centre-driven maturation of B cell response to mRNA vaccination. *Nature* **604**, 141-145, doi:10.1038/s41586-022-04527-1 (2022).

23 Ju, B. *et al.* Omicron BQ.1.1 and XBB.1 unprecedentedly escape broadly neutralizing antibodies elicited by prototype vaccination. *Cell Reports* **42**, doi:10.1016/j.celrep.2023.112532 (2023).

24 Wang, Y. *et al.* A large-scale systematic survey reveals recurring molecular features of public antibody responses to SARS-CoV-2. *Immunity* **55**, 1105-1117.e1104, doi:10.1016/j.immuni.2022.03.019 (2022).

25 Li, L. *et al.* Breakthrough infection elicits hypermutated IGHV3-53/3-66 public antibodies with broad and potent neutralizing activity against SARS-CoV-2 variants including the emerging EG.5 lineages. *PLoS Pathog* **19**, e1011856, doi:10.1371/journal.ppat.1011856 (2023).

26 Tan, T. J. C. *et al.* Sequence signatures of two public antibody clonotypes that bind SARS-CoV-2 receptor binding domain. *Nat Commun* **12**, 3815, doi:10.1038/s41467-021-24123-7 (2021).

27 Clark, S. A. *et al.* SARS-CoV-2 evolution in an immunocompromised host reveals shared neutralization escape mechanisms. *Cell*, doi:10.1016/j.cell.2021.03.027 (2021).

28 Wang, M. *et al.* A key F27I substitution within HCDR1 facilitates the rapid maturation of P2C-1F11-like neutralizing antibodies in a SARS-CoV-2-infected donor. *Cell Rep* **40**,

111335, doi:10.1016/j.celrep.2022.111335 (2022).

29 Kim, J., McFee, M., Fang, Q., Abdin, O. & Kim, P. M. Computational and artificial intelligence-based methods for antibody development. *Trends Pharmacol Sci* **44**, 175-189, doi:10.1016/j.tips.2022.12.005 (2023).

30 Akbar, R. *et al.* Progress and challenges for the machine learning-based design of fit-for-purpose monoclonal antibodies. *MAbs* **14**, 2008790, doi:10.1080/19420862.2021.2008790 (2022).

31 Wang, Y. *et al.* An explainable language model for antibody specificity prediction using curated influenza hemagglutinin antibodies. *Immunity* **57**, 2453-2465 e2457, doi:10.1016/j.immuni.2024.07.022 (2024).

32 Hie, B. L. *et al.* Efficient evolution of human antibodies from general protein language models. *Nat Biotechnol* **42**, 275-283, doi:10.1038/s41587-023-01763-2 (2024).

33 Watson, J. L. *et al.* De novo design of protein structure and function with RFdiffusion. *Nature* **620**, 1089-1100, doi:10.1038/s41586-023-06415-8 (2023).

34 He, H. *et al.* De novo generation of SARS-CoV-2 antibody CDRH3 with a pre-trained generative large language model. *Nat Commun* **15**, 6867, doi:10.1038/s41467-024-50903-y (2024).

35 Wilman, W. *et al.* Machine-designed biotherapeutics: opportunities, feasibility and advantages of deep learning in computational antibody discovery. *Brief Bioinform* **23**, doi:10.1093/bib/bbac267 (2022).

36 Guo, C. *et al.* Antigenic and virological characteristics of SARS-CoV-2 variants BA.3.2, XFG, and NB.1.8.1. *Lancet Infect Dis* **25**, e374-e377, doi:10.1016/S1473-3099(25)00308-1 (2025).

37 Wang, Q. *et al.* Antibody evasion by SARS-CoV-2 Omicron subvariants BA.2.12.1, BA.4 and BA.5. *Nature*, doi:10.1038/s41586-022-05053-w (2022).

38 Wang, Q. *et al.* Alarming antibody evasion properties of rising SARS-CoV-2 BQ and XBB subvariants. *Cell* **186**, 279-286 e278, doi:10.1016/j.cell.2022.12.018 (2023).

39 Raybould, M. I. J., Kovaltsuk, A., Marks, C. & Deane, C. M. CoV-AbDab: the coronavirus antibody database. *Bioinformatics* **37**, 734-735, doi:10.1093/bioinformatics/btaa739 (2021).

40 Lin, Z. *et al.* Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* **379**, 1123-1130, doi:10.1126/science.ade2574 (2023).

41 Su, J. *et al.* RoFormer: Enhanced transformer with Rotary Position Embedding. *Neurocomputing* **568**, doi:10.1016/j.neucom.2023.127063 (2024).

42 Olsen, T. H., Boyles, F. & Deane, C. M. Observed Antibody Space: A diverse database of cleaned, annotated, and translated unpaired and paired antibody sequences. *Protein Sci* **31**, 141-146, doi:10.1002/pro.4205 (2022).

43 Rosen, L. E. *et al.* A potent pan-sarbecovirus neutralizing antibody resilient to epitope diversification. *Cell* **187**, 7196-7213 e7126, doi:10.1016/j.cell.2024.09.026 (2024).

44 Halfmann, P. J. *et al.* SARS-CoV-2 Omicron virus causes attenuated disease in mice and hamsters. *Nature* **603**, 687-692, doi:10.1038/s41586-022-04441-6 (2022).

45 Chen, Q. *et al.* Comparative characterization of SARS-CoV-2 variants of concern and mouse-adapted strains in mice. *J Med Virol* **94**, 3223-3232, doi:10.1002/jmv.27735 (2022).

46 Kallewaard, N. L. *et al.* Structure and Function Analysis of an Antibody Recognizing All Influenza A Subtypes. *Cell* **166**, 596-608, doi:10.1016/j.cell.2016.05.073 (2016).

47 Yan, Q. *et al.* Germline IGHV3-53-encoded RBD-targeting neutralizing antibodies are commonly present in the antibody repertoires of COVID-19 patients. *Emerg Microbes Infect* **10**, 1097-1111, doi:10.1080/22221751.2021.1925594 (2021).

48 Robbiani, D. F. *et al.* Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* **584**, 437-442, doi:10.1038/s41586-020-2456-9 (2020).

49 Zhang, Q. *et al.* Potent and protective IGHV3-53/3-66 public antibodies and their shared escape mutant on the spike of SARS-CoV-2. *Nat Commun* **12**, 4210, doi:10.1038/s41467-021-24514-w (2021).

50 Yu, H. *et al.* Somatically hypermutated antibodies isolated from SARS-CoV-2 Delta infected patients cross-neutralize heterologous variants. *Nature Communications* **14**, doi:10.1038/s41467-023-36761-0 (2023).

51 Wang, Q. *et al.* Optimizing a human monoclonal antibody for better neutralization of SARS-CoV-2. *Nat Commun* **16**, 6195, doi:10.1038/s41467-025-61472-z (2025).

52 Witte, L. *et al.* Epistasis lowers the genetic barrier to SARS-CoV-2 neutralizing antibody escape. *Nat Commun* **14**, 302, doi:10.1038/s41467-023-35927-0 (2023).

53 Wang, Q. *et al.* XBB.1.5 monovalent mRNA vaccine booster elicits robust neutralizing antibodies against XBB subvariants and JN.1. *Cell Host Microbe* **32**, 315-321 e313, doi:10.1016/j.chom.2024.01.014 (2024).

54 Kwon, D. I. *et al.* Mucosal unadjuvanted booster vaccines elicit local IgA responses by conversion of pre-existing immunity in mice. *Nat Immunol* **26**, 908-919, doi:10.1038/s41590-025-02156-0 (2025).

55 Shan, S. *et al.* Deep learning guided optimization of human antibody against SARS-CoV-2 variants with broad neutralization. *Proc Natl Acad Sci U S A* **119**, e2122954119, doi:10.1073/pnas.2122954119 (2022).

56 Clifford, J. N., Richardson, E., Peters, B. & Nielsen, M. AbEpiTope-1.0: Improved antibody target prediction by use of AlphaFold and inverse folding. *Sci Adv* **11**, eadu1823, doi:10.1126/sciadv.adu1823 (2025).

57 Shanker, V. R., Bruun, T. U. J., Hie, B. L. & Kim, P. S. Unsupervised evolution of protein and antibody complexes with a structure-informed language model. *Science* **385**, 46-53, doi:10.1126/science.adk8946 (2024).

58 Desautels, T. A. *et al.* Computationally restoring the potency of a clinical antibody against Omicron. *Nature* **629**, 878-885, doi:10.1038/s41586-024-07385-1 (2024).

59 Walls, A. C. *et al.* Unexpected Receptor Functional Mimicry Elucidates Activation of Coronavirus Fusion. *Cell* **176**, 1026-1039 e1015, doi:10.1016/j.cell.2018.12.028 (2019).

60 Hsieh, C. L. *et al.* Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **369**, 1501-1505, doi:10.1126/science.abd0826 (2020).

61 Xiong, X. *et al.* A thermostable, closed SARS-CoV-2 spike protein trimer. *Nat Struct Mol Biol* **27**, 934-941, doi:10.1038/s41594-020-0478-5 (2020).

62 Cui, A. *et al.* A Model of Somatic Hypermutation Targeting in Mice Based on High-Throughput Ig Sequencing Data. *J Immunol* **197**, 3566-3574, doi:10.4049/jimmunol.1502263 (2016).

63 Cao, Y. *et al.* BA.2.12.1, BA.4 and BA.5 escape antibodies elicited by Omicron infection.

Nature **608**, 593-602, doi:10.1038/s41586-022-04980-y (2022).

64 Cao, Y. *et al.* Imprinted SARS-CoV-2 humoral immunity induces convergent Omicron RBD evolution. *Nature*, doi:10.1038/s41586-022-05644-7 (2022).

65 Gaebler, C. *et al.* Evolution of antibody immunity to SARS-CoV-2. *Nature* **591**, 639-644, doi:10.1038/s41586-021-03207-w (2021).

66 Wang, Z. *et al.* Naturally enhanced neutralizing breadth against SARS-CoV-2 one year after infection. *Nature* **595**, 426-431, doi:10.1038/s41586-021-03696-9 (2021).

67 Wang, Z. *et al.* Memory B cell responses to Omicron subvariants after SARS-CoV-2 mRNA breakthrough infection in humans. *J Exp Med* **219**, doi:10.1084/jem.20221006 (2022).

68 Kaku, C. I. *et al.* Recall of preexisting cross-reactive B cell memory after Omicron BA.1 breakthrough infection. *Sci Immunol* **7**, eabq3511, doi:10.1126/sciimmunol.abq3511 (2022).

69 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).

70 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).

71 Afonine, P. V. *et al.* Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol* **74**, 531-544, doi:10.1107/S2059798318006551 (2018).

72 Krissinel, E. Macromolecular complexes in crystals and solutions. *Acta Crystallographica Section D* **67**, 376-385, doi:doi:10.1107/S0907444911007232 (2011).

Figures

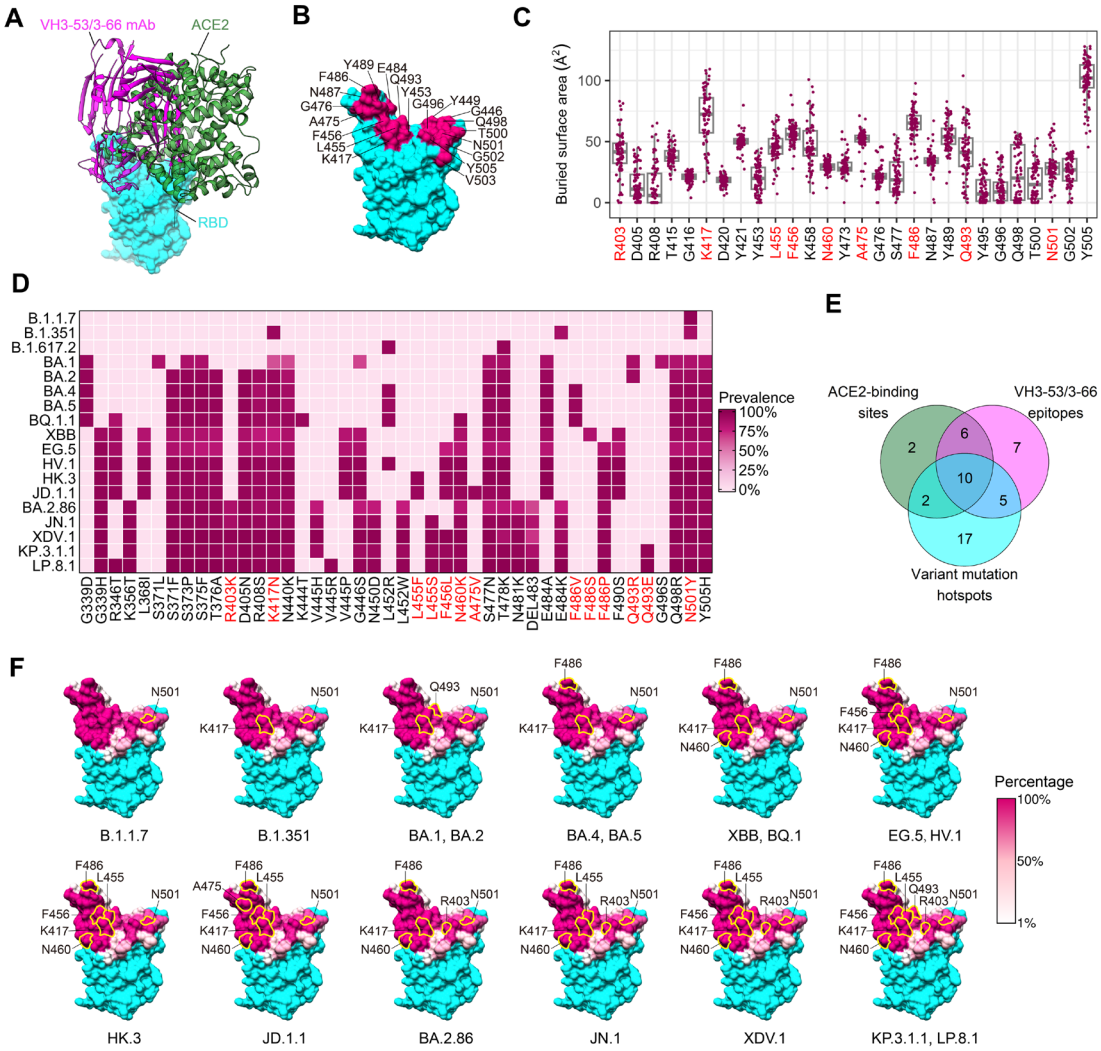


Figure 1 SARS-CoV-2 spike accumulates sequential mutations on the VH3-53/3-66-targeted epitope.

(**A**) Comparison of VH3-53/3-66 mAbs (BD-604 as representative, PDB: 7CHF) and ACE2 (PDB: 6M0J) binding to the RBD. VH3-53/3-66 mAbs, ACE2 and RBD are colored in magenta, green and cyan, respectively.

(**B**) The epitope residues of ACE2 binding sites, binding interface of ACE2 highlighted in magenta.

(**C**) The buried surface area of 79 VH3-53/3-66 mAbs at each binding sites. Nine residues that have been experimentally validated to mediate antibody neutralization escape were highlighted with red.

(**D**) The mutation profile on RBD of 18 SARS-CoV-2 variants. Nine residues that have been

1029 experimentally validated to mediate antibody neutralization escape were highlighted with
1030 red.
1031 (**E**) Venn-plot showing the overlapped residues among ACE2-binding sites, VH3-53/3-66
1032 epitopes, and variant mutation hotspots.
1033 (**F**) Accumulation of mutations within the epitope of VH3-53/3-66 mAbs across SARS-CoV-
1034 2 variants. Footprints of VH3-53/3-66 epitopes on RBD, and the key escape mutations
1035 within the epitope was labeled. Epitope residues on the RBD are colored according to the
1036 proportion of VH3-53/3-66 mAbs interacting with each positions.
1037
1038

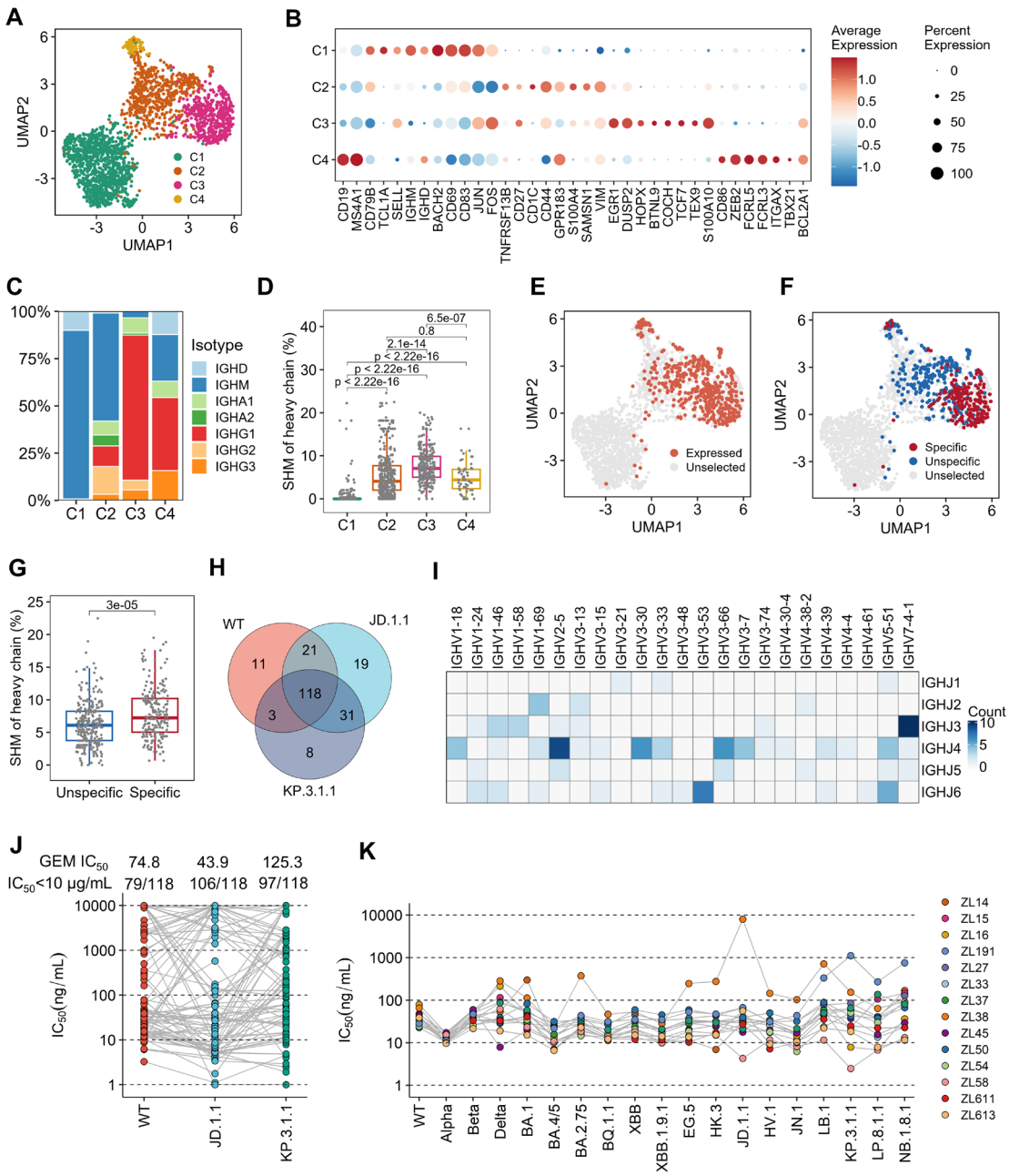


Figure 2 Isolation of broad neutralizing mAbs from SARS-CoV-2 elite neutralizers. (**A**) Uniform manifold approximation projection (UMAP) embedding of scRNA-seq transcriptome profiles of the 2124 JD.1.1 and KP.3.1.1 probe+ B cells, colored by cluster. (**B**) Dotplot showing the marker genes of the four B cell clusters. (**C**) Comparison of isotype frequency of the four B cell clusters. (**D**) Comparison of somatic hypermutation rates among the four B cell clusters. (**E**) UMAP highlighting cells from which mAb was synthesized and expressed. The synthesized and expressed mAbs are colored with red dots and unselected mAbs with grey

1048 dots.

1049 (**F**) UMAP showing the RBD-specific mAbs with red dots, RBD-unspecific mAbs with blue

1050 dots, and unselected mAbs with grey dots.

1051 (**G**) Comparison of somatic hypermutation rates between RBD-specific and unspecific

1052 mAbs.

1053 (**H**) Venn-plot showing the number of mAbs with WT, JD.1.1 and KP.3.1.1 cross-reactivity.

1054 (**I**) IGHV and IGHJ germline gene usage of the 118 mAbs with WT, JD.1.1 and KP.3.1.1

1055 cross-reactivity.

1056 (**J**) Neutralization activities of the 118 mAbs to WT, JD.1.1 and KP.3.1.1. Geometric mean

1057 of IC_{50} values and numbers of nAbs (IC_{50} < 10 μg/mL) are annotated above each group of

1058 points.

1059 (**K**) Neutralization activities of the 15 VH3-53/3-66 mAbs to SARS-CoV-2 variants.

1060

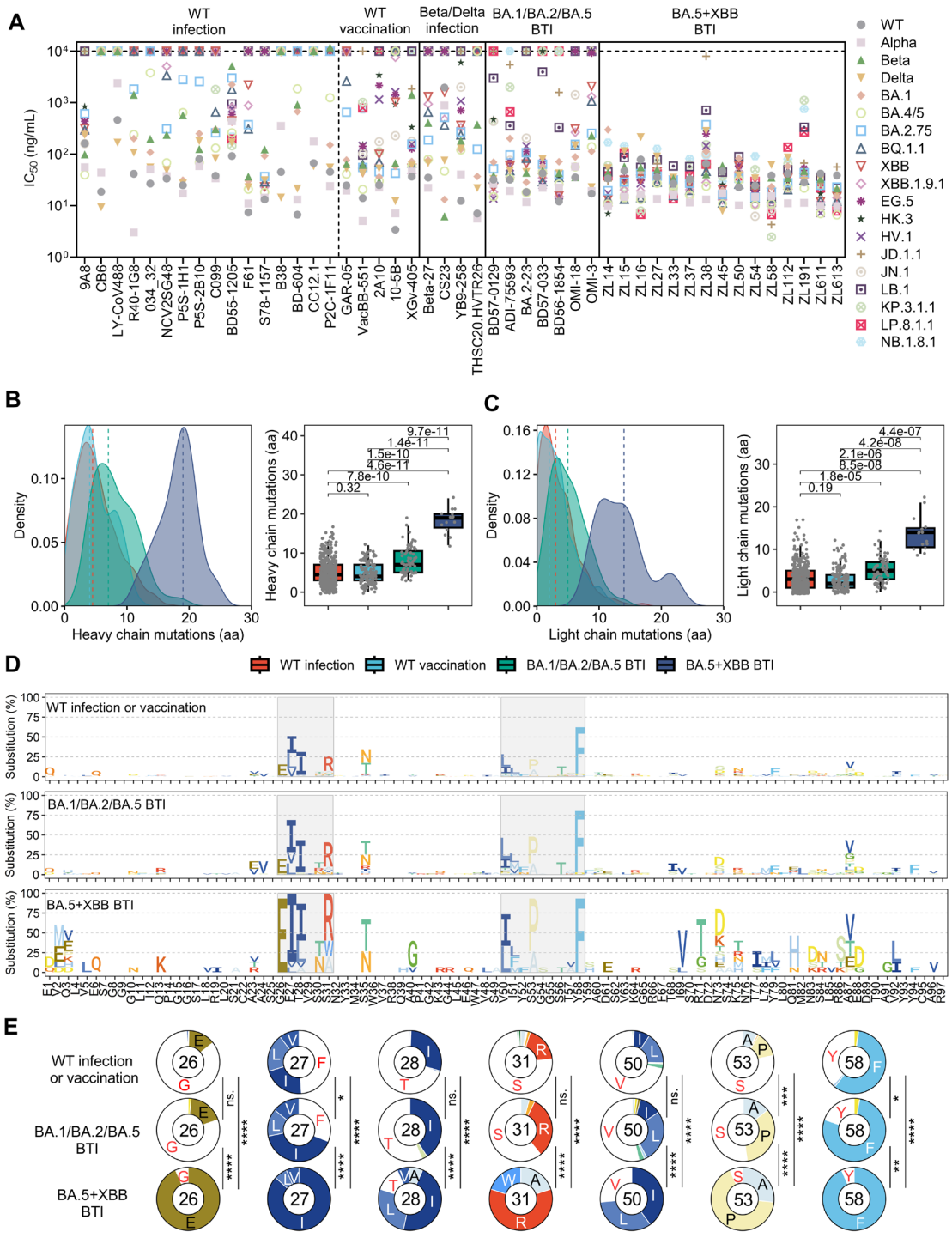


Figure 3 Comparison of the neutralizing activities and somatic hypermutations of VH3-53/3-66 mAbs.

(**A**) Comparison of neutralizing activities (IC_{50}) of VH3-53/3-66 mAbs isolated from different infection or vaccination background. BTI, breakthrough infection.

(**B-C**) Comparison of somatic hypermutation levels within heavy (**B**) or light (**C**) chain variable region of VH3-53/3-66 mAbs isolated from different infection or vaccination

1068 background.

1069 (**D**) Somatic mutation profiles of VH3-53/3-66 mAbs isolated from different infection or

1070 vaccination background.

1071 (**E**) Pie charts indicate the frequency of germline (white) and mutated residues at position

1072 26, 27, 28, 31, 50, 53, and 58 of heavy chain variable region of VH3-53/3-66 mAbs.

1073

1074

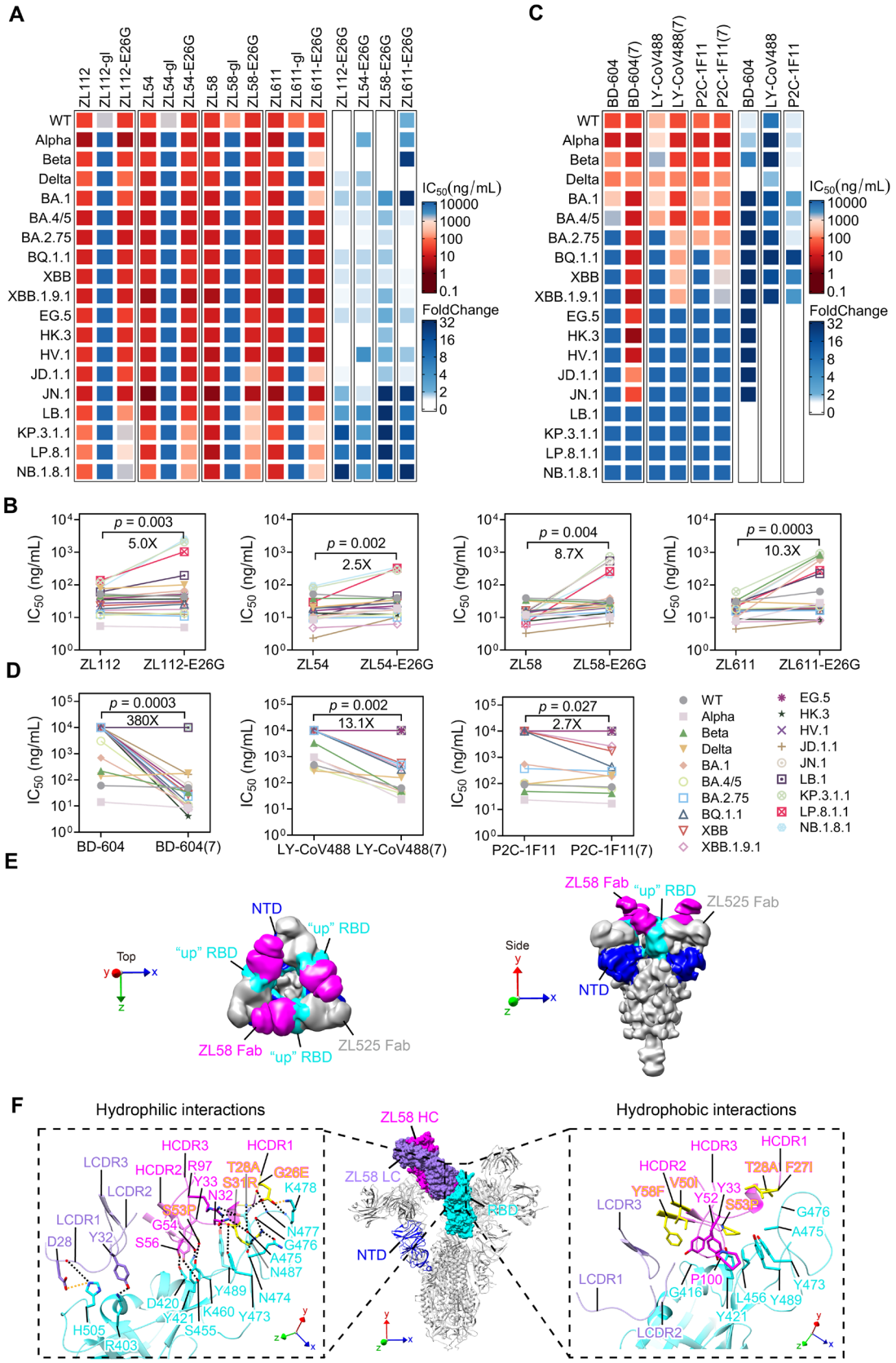


Figure 4 Convergent SHM extent the neutralization and potency and breadth of VH3-53/3-66 mAbs

(**A**) Heatmap (left panel) showing the neutralizing activity of ZL112, ZL54, ZL58, and ZL611, as well as their mutants against 19 major SARS-CoV-2 strains. Heatmap (right panel) showing the foldchange of IC_{50} of E26G versions compared to ZL112, ZL54, ZL58, and ZL611, respectively.

(**B**) Comparison of the IC_{50} values between ZL112, ZL54, ZL58, or ZL611 and their E26G mutants.

(**C**) Heatmap (left panel) showing the neutralizing activity of BD-604, LY-CoV488, and P2C-1F11, as well as their mutants against 19 major SARS-CoV-2 strains. Heatmap (right panel) showing the foldchange of IC_{50} of SHM-introduction versions compared to BD-604, LY-CoV488, and P2C-1F11, respectively.

(**D**) Comparison of the IC_{50} values between BD-604, LY-CoV488, and P2C-1F11 and their SHM-introduction mutants.

(**E**) Structures (low-pass filtered to 12 Å) of Omicron LP.8.1 spike (S-6P) in complex with ZL58 Fab and ZL525 Fab. ZL58 Fab, NTD and RBD are highlighted in magenta, blue and cyan, respectively; the rest of the spike and ZL525 Fab are colored gray.

(**F**) Epitope of ZL58 on the RBD of Omicron LP.8.1 spike. Detailed interactions between ZL58 and RBD are shown in the dashed boxes. CDR loops are indicated, and selected interacting residues in antibody-RBD interface are shown. Somatically hypermutated residues are highlighted in yellow and labeled with the changes indicated. Backbone carbonyl oxygens and amide nitrogens are indicated by red and blue dots, respectively. Yellow and black dashed lines indicate salt bridge and hydrogen bond interactions, respectively.

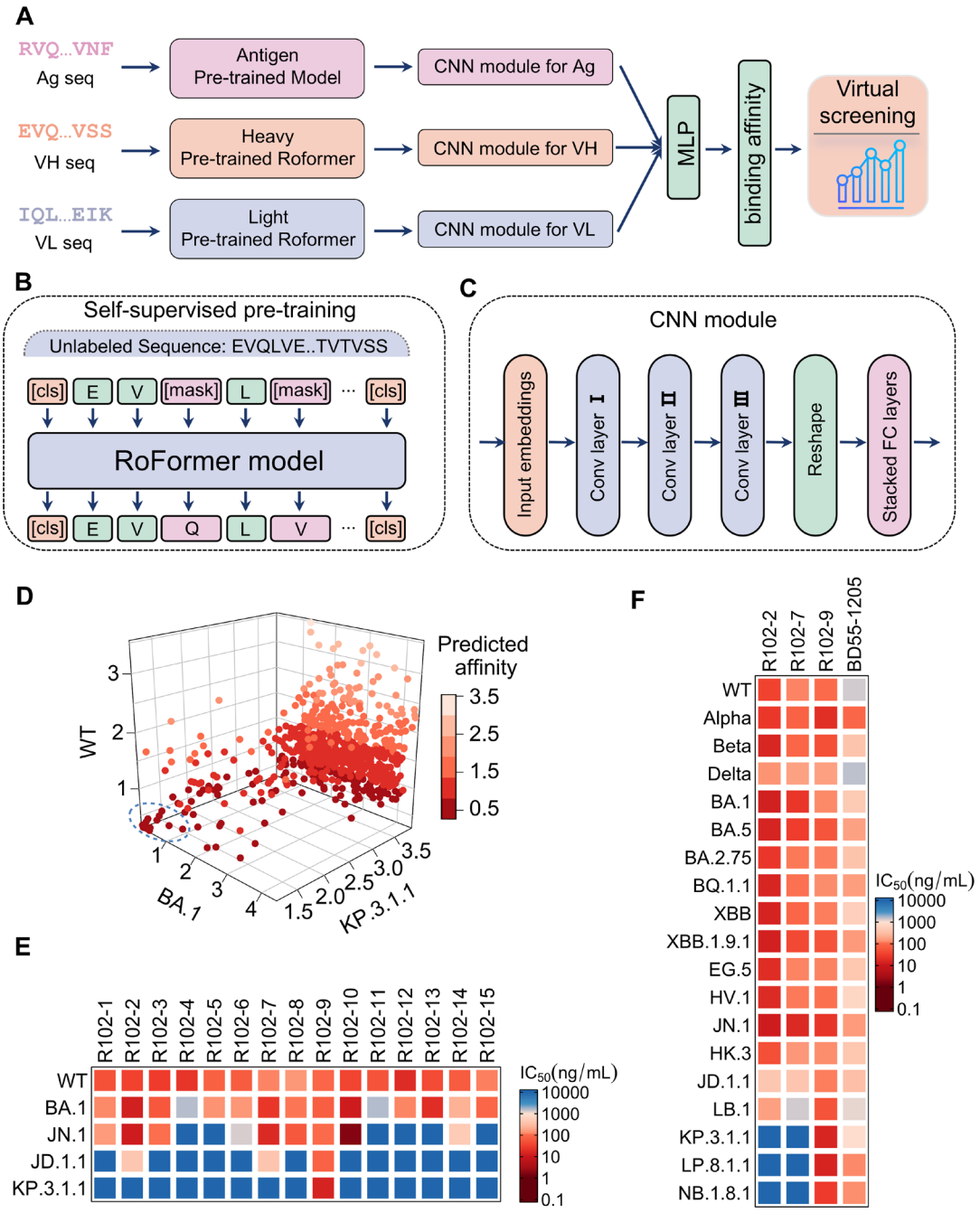


Figure 5 AI-based discovery of VH3-53/3-66 bnAbs.

(**A**) Binding affinity prediction and virtual screening were conducted using pre-trained antibody language models and ESM2. The antigen pre-trained model refers to ESM2, which was used to extract evolutionary information from antigen sequences. Heavy and light pre-trained Roformers denote the language models specifically pre-trained on antibody heavy and light chains, respectively, and were employed to capture comprehensive sequence-level features. The extracted features were further processed

through dedicated CNN modules for in-depth feature learning, and subsequently used to predict antibody–antigen binding affinity. Based on the affinity prediction model, we scored each antibody in the library and selected those with the highest predicted affinities for wet-lab validation.

(**B**) Schematic of self-supervised pre-training. First, tokens in the input sequences were randomly masked, and the RoFormer model was trained to recover the masked regions. In this process, a Unique Amino Acid Tokenizer was employed, which assigns a unique token to each distinct amino acid.

(**C**) The CNN module. The input embeddings were processed through three consecutive convolutional layers for in-depth feature extraction, and the final embeddings were generated through stacked fully connected (FC) layers.

(**D**) Predicted affinity of 785 VH3-53/3-66 mAbs that bind to WT, BA.1 and KP.3.1.1. The blue dashed line highlights the top 15 antibodies.

(**E**) Heatmap showing the neutralizing activity of top15 mAbs against WT, BA.1, JN.1, JD.1.1 and KP.3.1.1.

(**F**) Heatmap showing the neutralizing activity of R102-2, R102-7, R102-9, and BD55-1205 against 19 major SARS-CoV-2 strains.

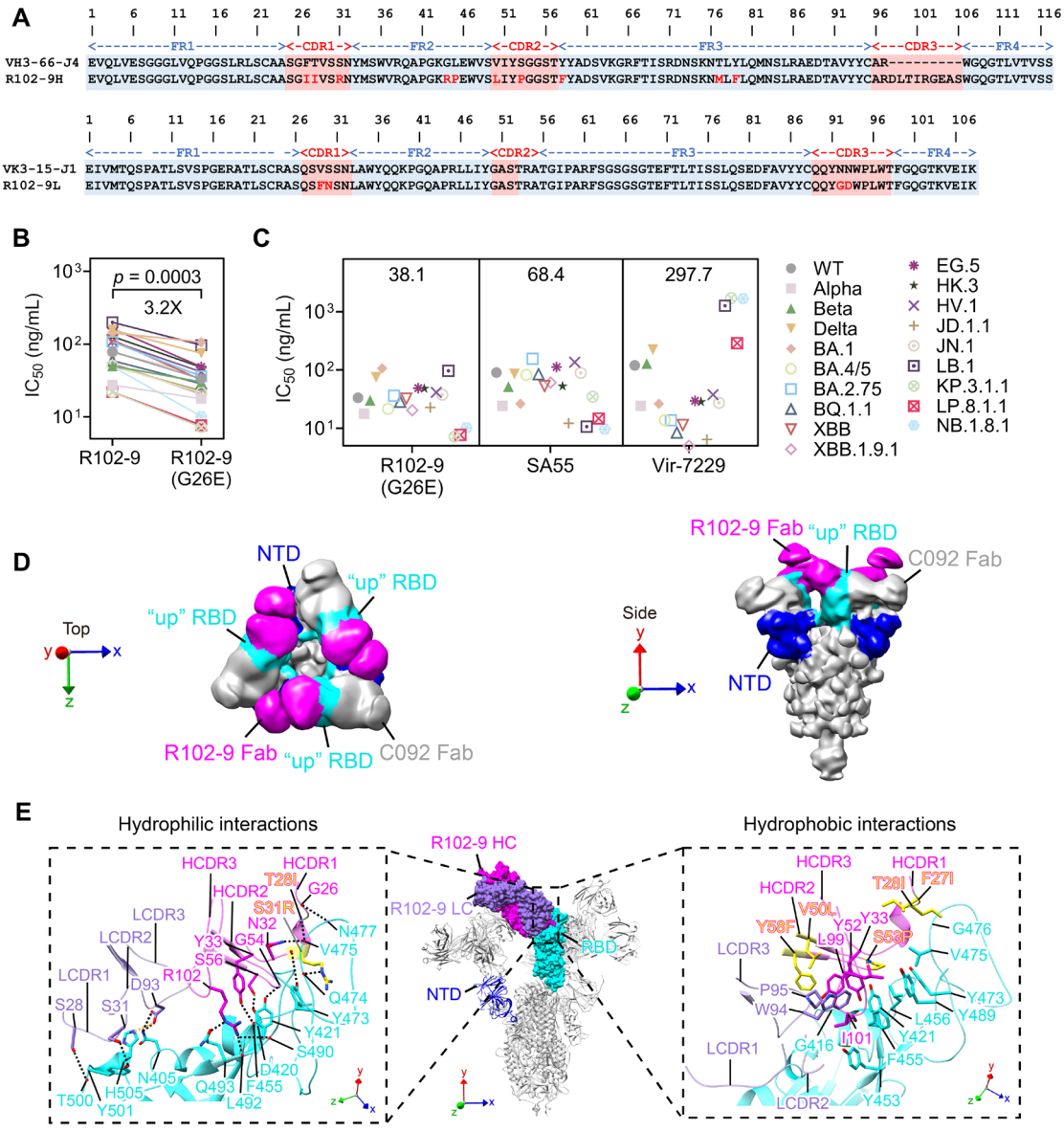


Figure 6 Structural basis of the broad neutralization of R102-9.

(**A**) Sequence analysis of R102-9. Somatic hypermutations are highlighted with red.

(**B**) Comparison of neutralization activity IC_{50} values between R102-9 and R102-9(G26E).

(**C**) Comparison of neutralization activity IC_{50} values across R102-9(G26E), SA55, and Vir-7229. GEM of IC_{50} values were labeled on top.

(**D**) Structures (low-pass filtered to 12 Å) of Omicron JD.1.1 spike (S-6P) in complex with R102-9 Fab and C092 Fab. R102-9 Fab, NTD and RBD are highlighted in magenta, blue and cyan, respectively; the rest of the spike and C092 Fab are colored gray.

(**E**) Epitope of R102-9 on the RBD of Omicron JD.1.1 spike. Detailed interactions between R102-9 and RBD are shown in the dashed boxes. CDR loops are indicated, and selected

interacting residues in antibody-RBD interface are shown. Somatically hypermutated residues are highlighted in yellow and labeled with the changes indicated. Backbone carbonyl oxygens and amide nitrogens are indicated by red and blue dots, respectively. Yellow and black dashed lines indicate salt bridge and hydrogen bond interactions, respectively.

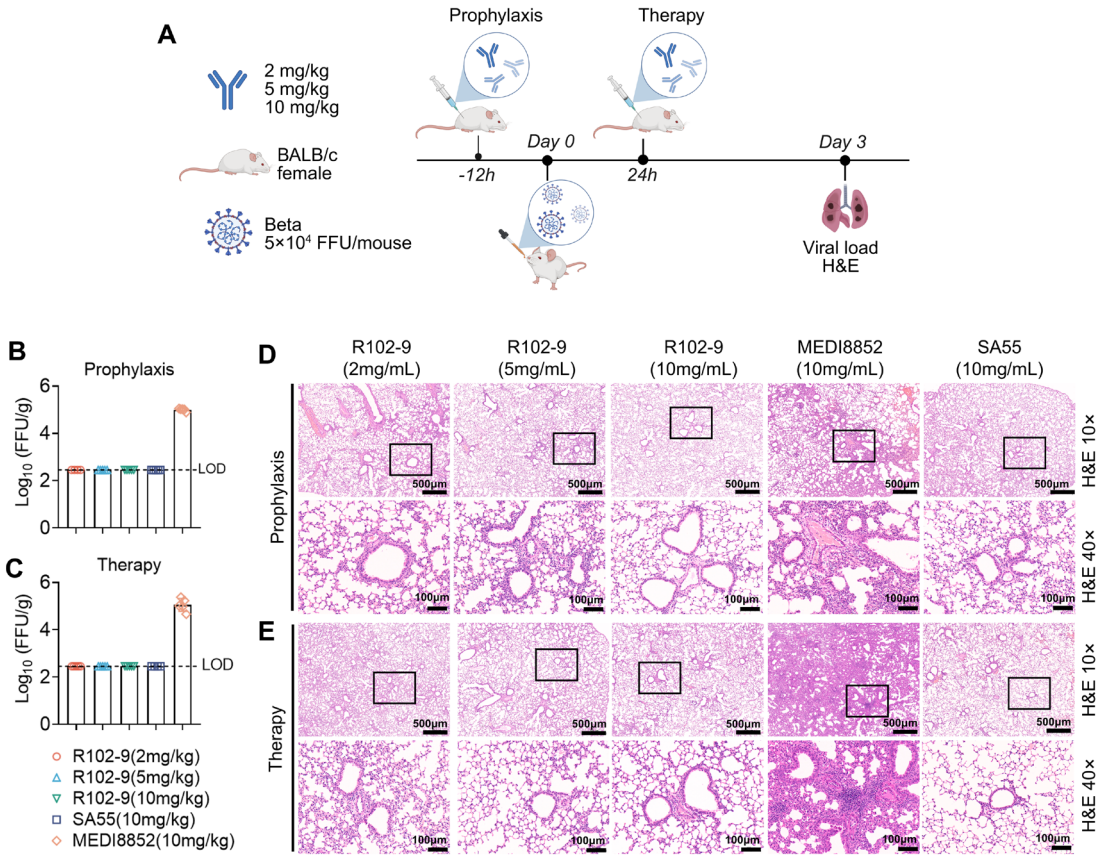


Figure 7 The prophylactic and therapeutic effects of R102-9 against Beta.

(**A**) Flowchart of the experimental design of testing the prophylactic and therapeutic activity of R102-9 to protect against challenge with Beta SARS-CoV-2 in BALB/c mice.

(**B-C**) Infectious virions were tested by viral fluorescent assay in lung tissue homogenates of prophylactic (**B**) and therapeutic (**C**) assay. FFUs per g of tissue extractions were compared between different groups in $\log_{10}$-transformed units. LOD, limit of detection.

(**D-E**) Representative images of mouse lung tissues of prophylactic (**D**) and therapeutic (**E**) by H&E staining.

Supplementary figures

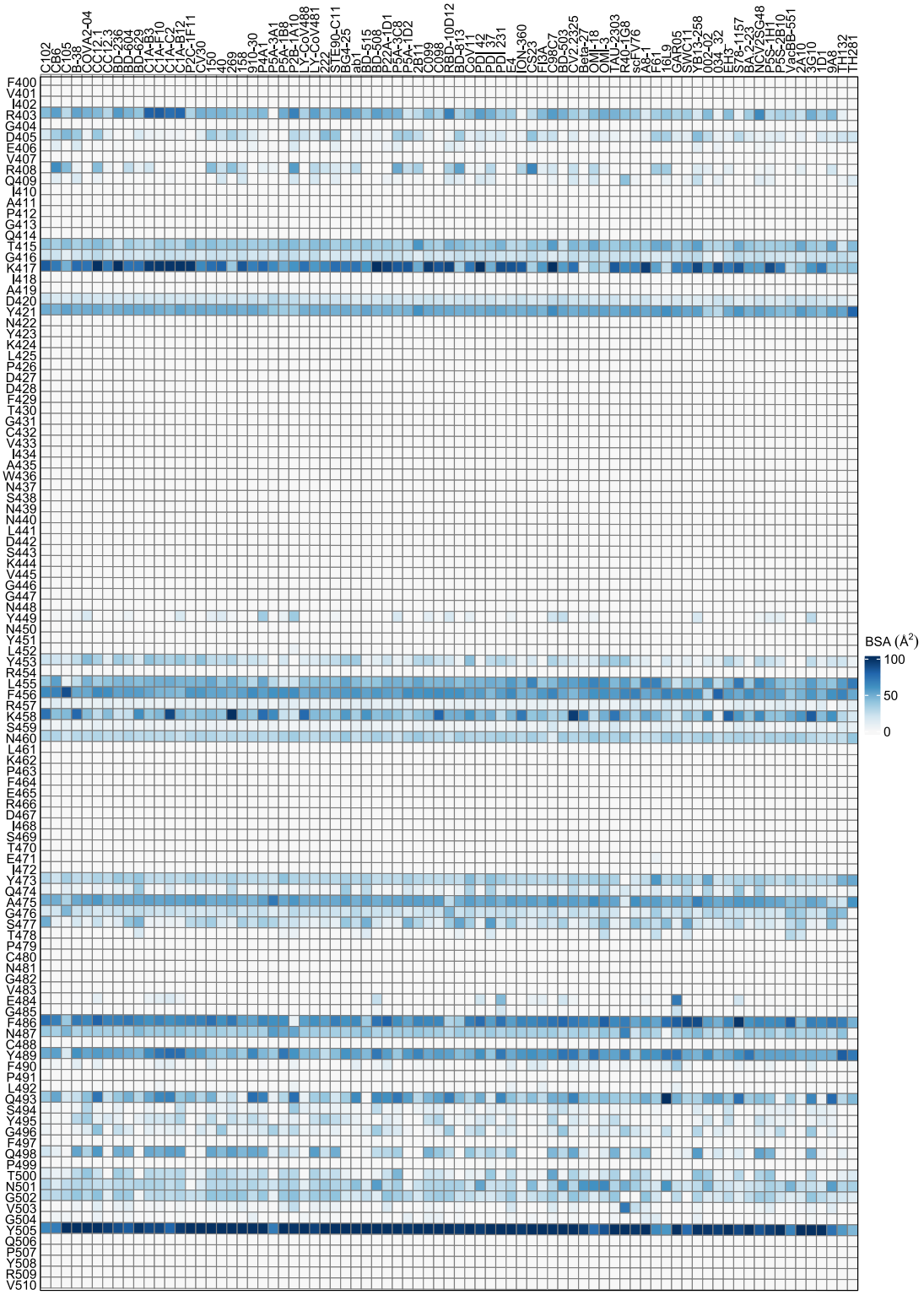
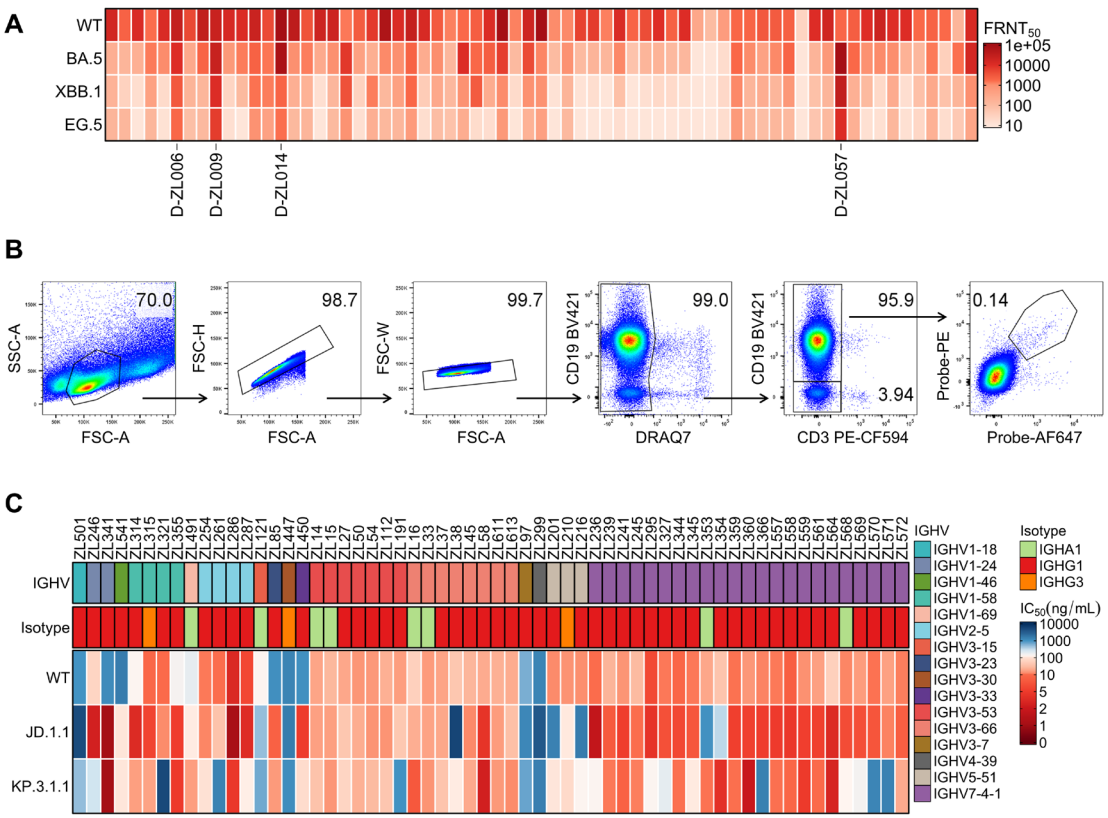


Figure S1 Heatmap showing the buried surface area of 79 class 1 VH3-53/3-66 mAbs at each epitope sites on RBD.



D

mAb_Id	IGHV	IGHJ	IGHC	HCDR3	HCDR3 length	SHM.H (%)	IGKV	IGKJ	LCDR3	LCDR3 length	SHM.L (%)
ZL14	IGHV3-53	IGHJ6	IGHA1	ARNLRERGGLDV	12	17.58	IGKV1-39	IGKJ4	QQNSEIPET	9	11.44
ZL15	IGHV3-53	IGHJ6	IGHA1	ARDLREAGGLDV	12	15.68	IGKV1-39	IGKJ4	HQNSAFPET	9	5.17
ZL16	IGHV3-66	IGHJ5	IGHA1	ARTVRDES	8	11	IGKV1-33	IGKJ1	HQYDNLPFT	9	7.75
ZL27	IGHV3-53	IGHJ6	IGHG1	ARDLQERGGLDV	12	8.53	IGKV1-39	IGKJ4	HQSSTLPET	9	7.3
ZL33	IGHV3-66	IGHJ4	IGHA1	ARQGPNGRNDH	11	11.38	IGKV1-33	IGKJ4	HQYDSLPPT	9	5.92
ZL37	IGHV3-66	IGHJ4	IGHG1	ARDVGGMRLDL	11	10.6	IGKV1-33	IGKJ3	QQFDSLPQT	9	11.65
ZL38	IGHV3-66	IGHJ5	IGHG1	ARDVGGKRLDH	11	10.6	IGKV1-33	IGKJ3	HQSDNLPQT	9	9.22
ZL45	IGHV3-66	IGHJ4	IGHG1	ARDFGDQYSNS	11	11.68	IGKV3-15	IGKJ1	QHYDSRPRT	9	6.59
ZL50	IGHV3-53	IGHJ6	IGHG1	VRNLVSLGMDV	11	9.62	IGKV1-9	IGKJ2	QQVDSDLYT	9	5.19
ZL54	IGHV3-53	IGHJ6	IGHG1	ARDLRQVGGLDV	12	12.97	IGKV1-39	IGKJ2	HQYSYTPET	9	12.28
ZL58	IGHV3-66	IGHJ4	IGHG1	ARQGPNGRNDF	11	12.59	IGKV1-33	IGKJ4	HQYDSFPPT	9	8.3
ZL112	IGHV3-53	IGHJ6	IGHG1	ARDLVYYGMDV	11	7.51	IGKV1-9	IGKJ2	QQVDGSLYT	9	4.81
ZL191	IGHV3-53	IGHJ6	IGHG1	ARDLREAGGLDV	12	9.56	IGKV1-39	IGKJ4	HQNSAIPET	9	7.01
ZL611	IGHV3-66	IGHJ4	IGHG1	VRDFGGRRLVD	11	10.65	IGKV1-33	IGKJ3	HQSDNLPQT	9	6.64
ZL613	IGHV3-66	IGHJ4	IGHG1	ARQAPSGRYDS	11	12.37	IGKV1-33	IGKJ4	HQYDTLPPS	9	8.71

Figure S2 Isolation of broad neutralizing mAbs from elite neutralizers.

(**A**) Heatmap showing the neutralizing activity of serum of 68 individuals with BA.5 and XBB breakthrough infections.

(**B**) The flow cytometry gating strategy for sorting antigen-specific B cells.

(**C**) Heatmap showing the IC_{50} values of 60 cross-neutralizing mAbs against WT, JD.1.1 and KP.3.1.1. The top annotations indicate the IGHV and IGHC (isotype) germline segments.

(**D**) The summary of genetic characters of the 15 VH3-53/3-66 mAbs.

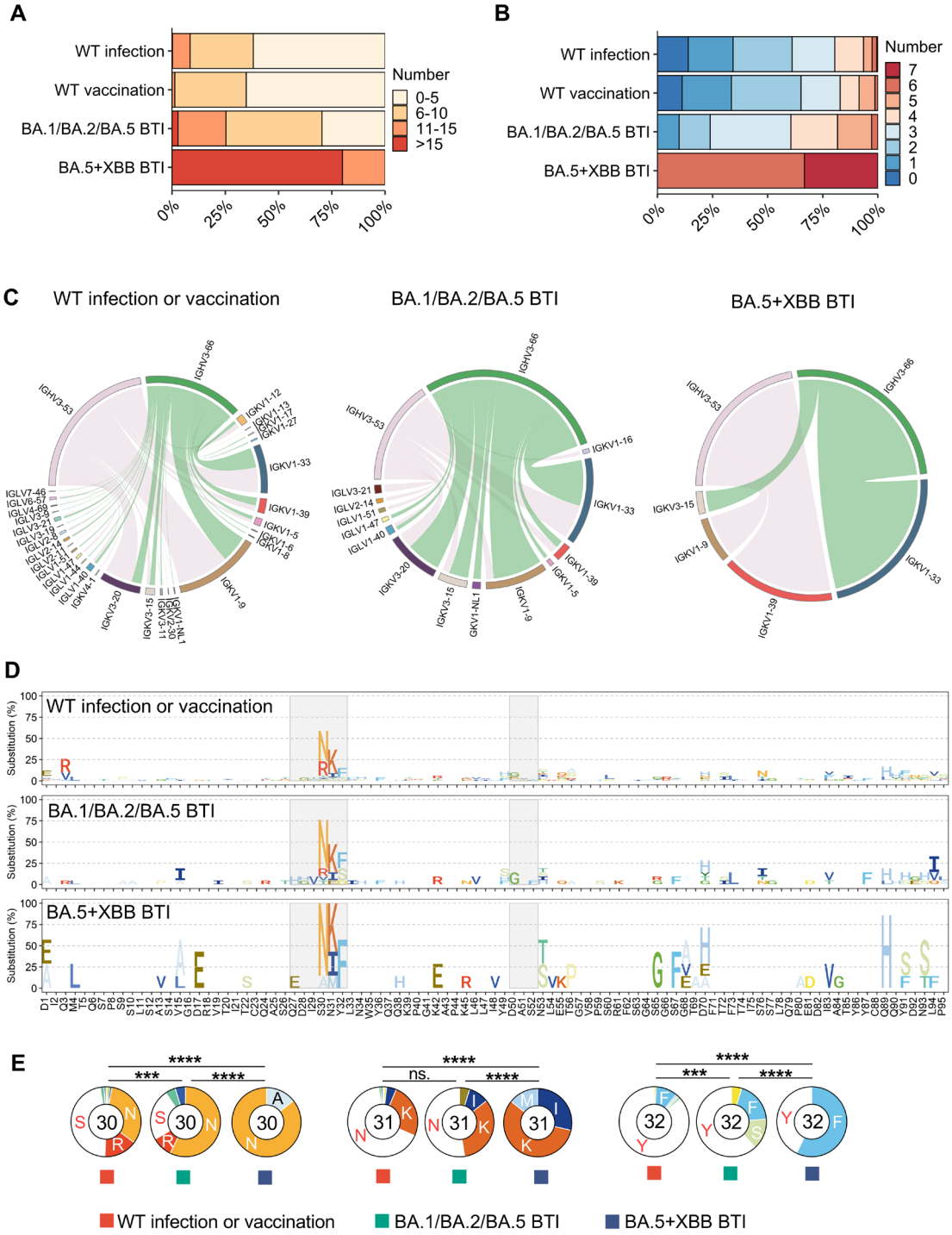


Figure S3 Comparison of the IGLV usage and somatic hypermutations of VH3-53/3-66 mAbs.

(**A**) Comparison of the frequency distribution of total number of somatic hypermutations across VH3-53/3-66 mAbs that isolated from different infection or vaccination background. (**B**) Comparison of the frequency distribution of number of somatic hypermutations among

1177 position G26, F27, T28, S31, V50, S53, and Y58 across VH3-53/3-66 mAbs that isolated
1178 from different infection or vaccination background.
1179 (**C**) Circos plot showing the pairing light chain usage of VH3-53/3-66 mAbs that isolated
1180 from different infection or vaccination background.
1181 (**D**) Somatic mutation profiles of IGKV1-33 light chain variable region of VH3-53/3-66 mAbs
1182 that isolated from different infection or vaccination background.
1183 (**E**) Pie charts indicate the frequency of germline (white) and mutated residues at position
1184 30-32 of IGKV1-33 light chain variable region of VH3-53/3-66 mAbs.
1185

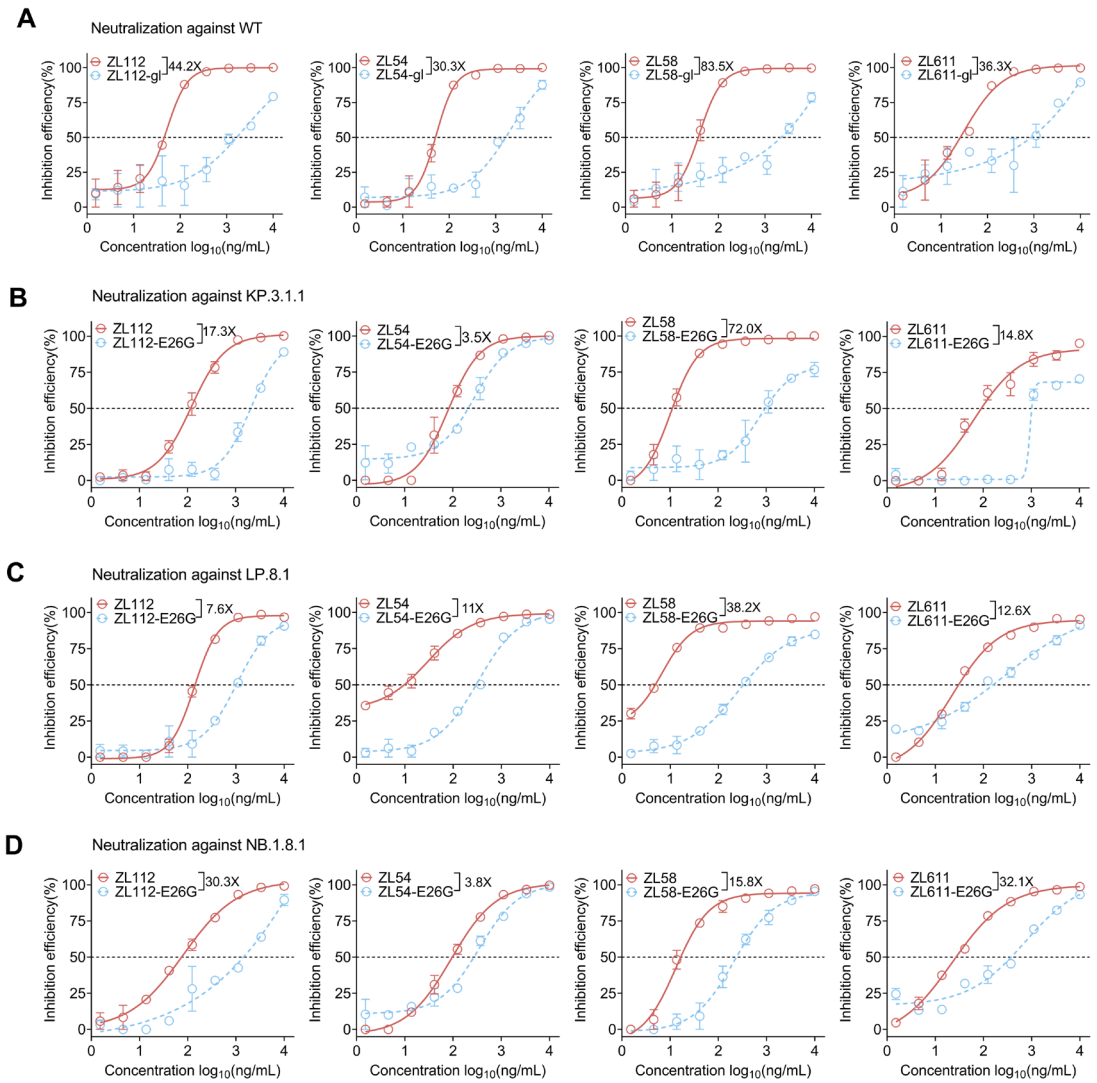


Figure S4 Effects of somatic hypermutation on neutralization breadth.

(**A**) Neutralizing activity of prototype and germline versions of ZL112, ZL54, ZL58, and ZL611 against WT SARS-CoV-2.

(**B-D**) Neutralizing activity of prototype and E26G mutant versions of ZL112, ZL54, ZL58, and ZL611 against KP.3.1.1 (**B**), LP.8.1.1 (**C**), and NB.1.8.1 (**D**).

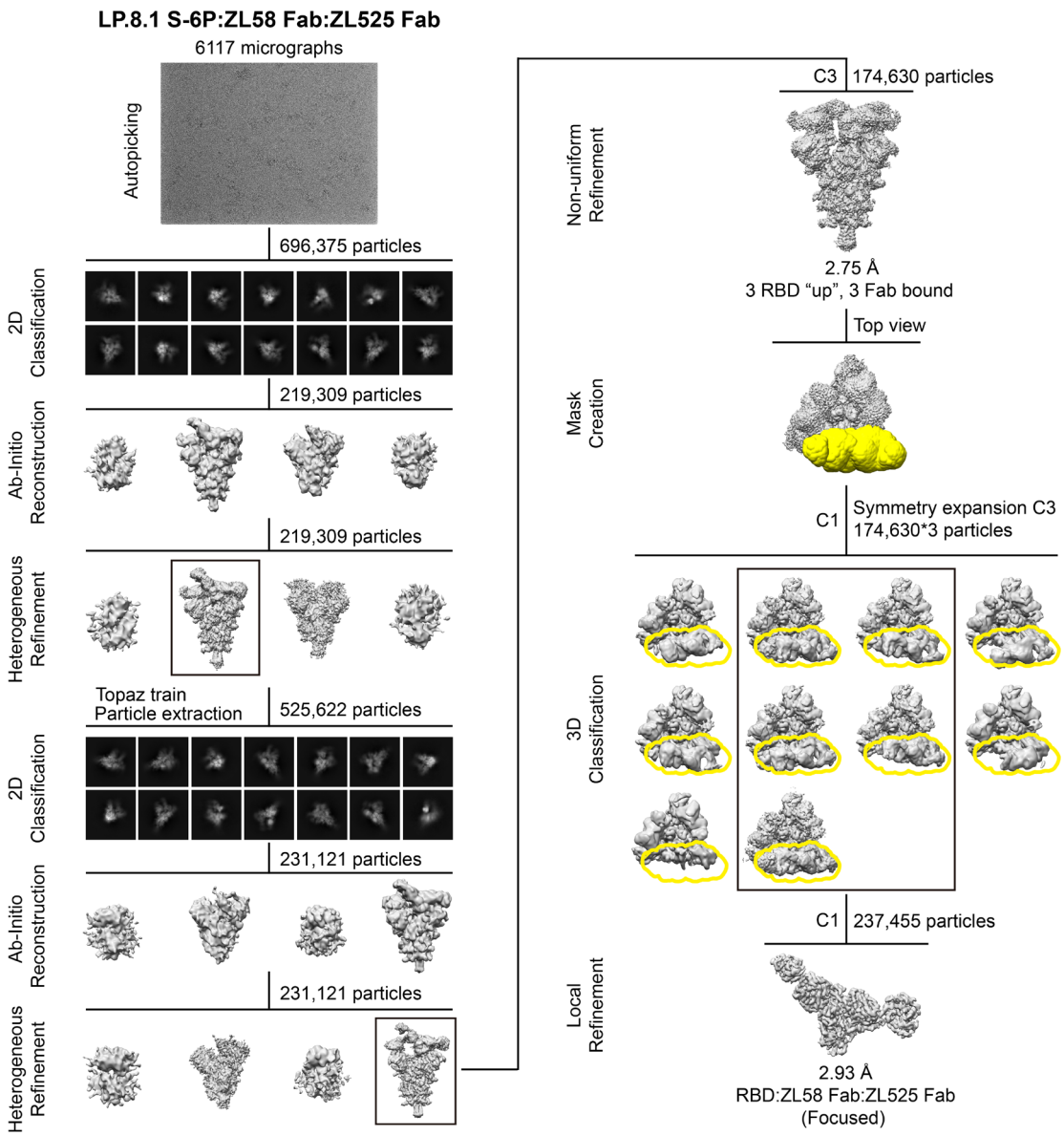


Figure S5 Cryo-EM data processing flow-chart for the structure of LP.8.1 S-6P:ZL58 Fab:ZL525 Fab complex. After 2D classification, ab-initio reconstruction and heterogeneous refinement were performed. High-quality particles were used to train the topaz picker, followed by particle re-extraction, 2D classification, ab-initio reconstruction, heterogeneous refinement, and non-uniform refinement. The structure of the RBD-Fab portion was further improved with focused refinement. Selected classes are highlighted with black boxes, and focused mask is shown in yellow.

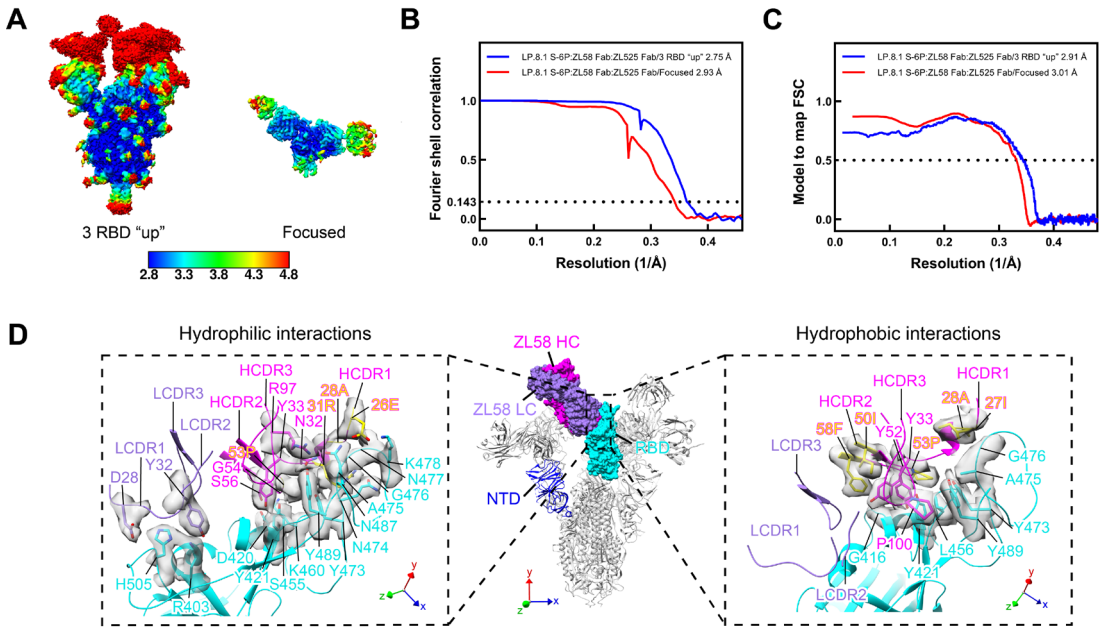


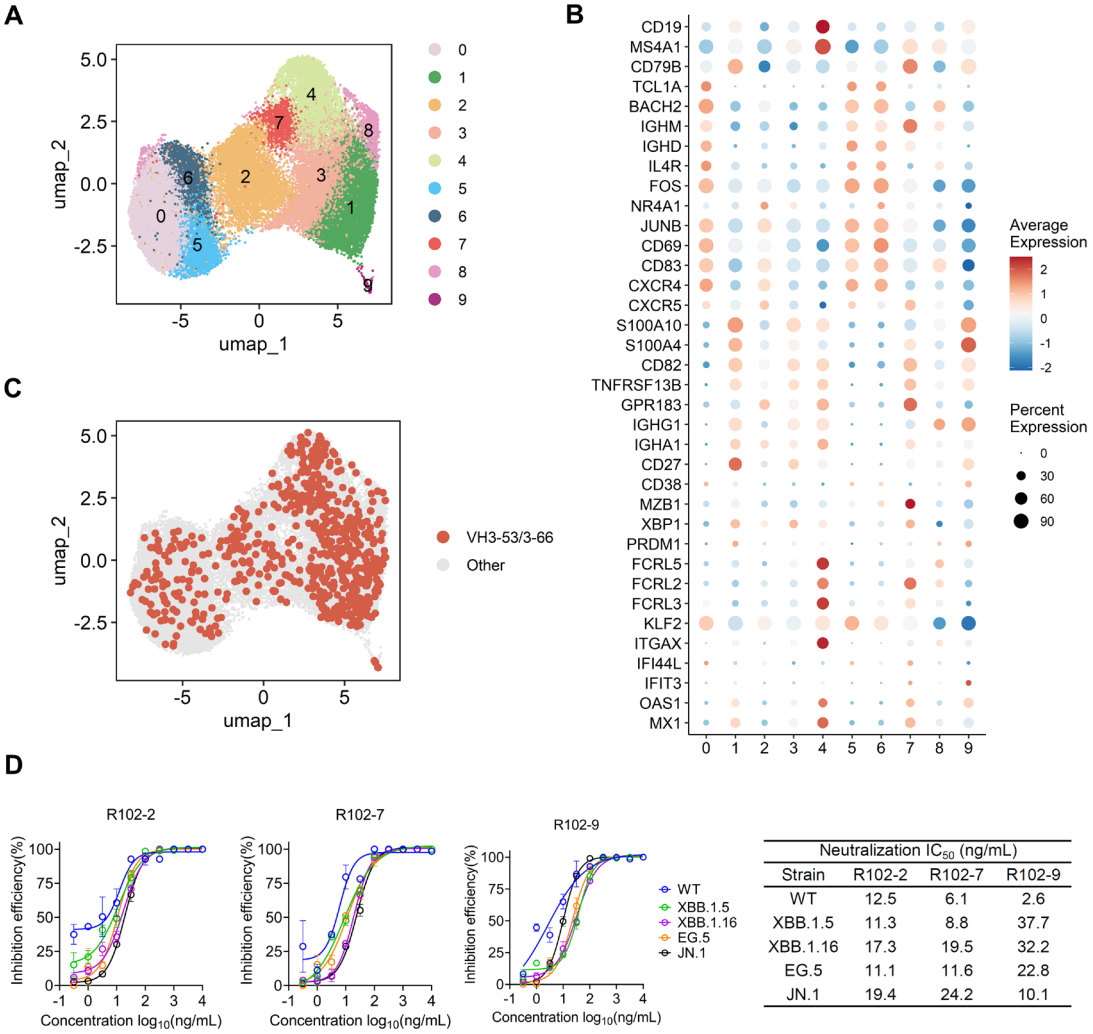
Figure S6 Resolution assessment of the structure of LP.8.1 S-6P:ZL58 Fab:ZL525 Fab complex.

(**A**) Local resolution map for the LP.8.1 S-6P:ZL58 Fab:ZL525 Fab complex.

(**B**) Global resolution assessment by Fourier shell correlation at the 0.143 criterion.

(**C**) Correlations of model vs map by Fourier shell correlation at the 0.5 criterion.

(**D**) Representative densities of the ZL58 epitope and ZL58 CDR loops in the focused map.



Neutralization IC_{50} (ng/mL)			
Strain	R102-2	R102-7	R102-9
WT	12.5	6.1	2.6
XBB.1.5	11.3	8.8	37.7
XBB.1.16	17.3	19.5	32.2
EG.5	11.1	11.6	22.8
JN.1	19.4	24.2	10.1

Figure S7 AI-based discovery of VH3-53/3-66 bnAbs from WT vaccinated donors.

(**A**) Uniform manifold approximation projection (UMAP) embedding of scRNA-seq transcriptome profiles of the WT S-6P probe^{+} B cells, colored by cluster.

(**B**) Dotplot showing the marker genes of the ten B cell clusters.

(**C**) UMAP highlighting cells from which mAb were selected for AI virtual screening. The selected mAbs are colored with red dots and unselected mAbs with grey dots.

(**D**) Authentic virus neutralizing activity of R102-2, R102-7, and R102-9 against WT, XBB.1.5, XBB.1.16, EG.5, and JN.1.

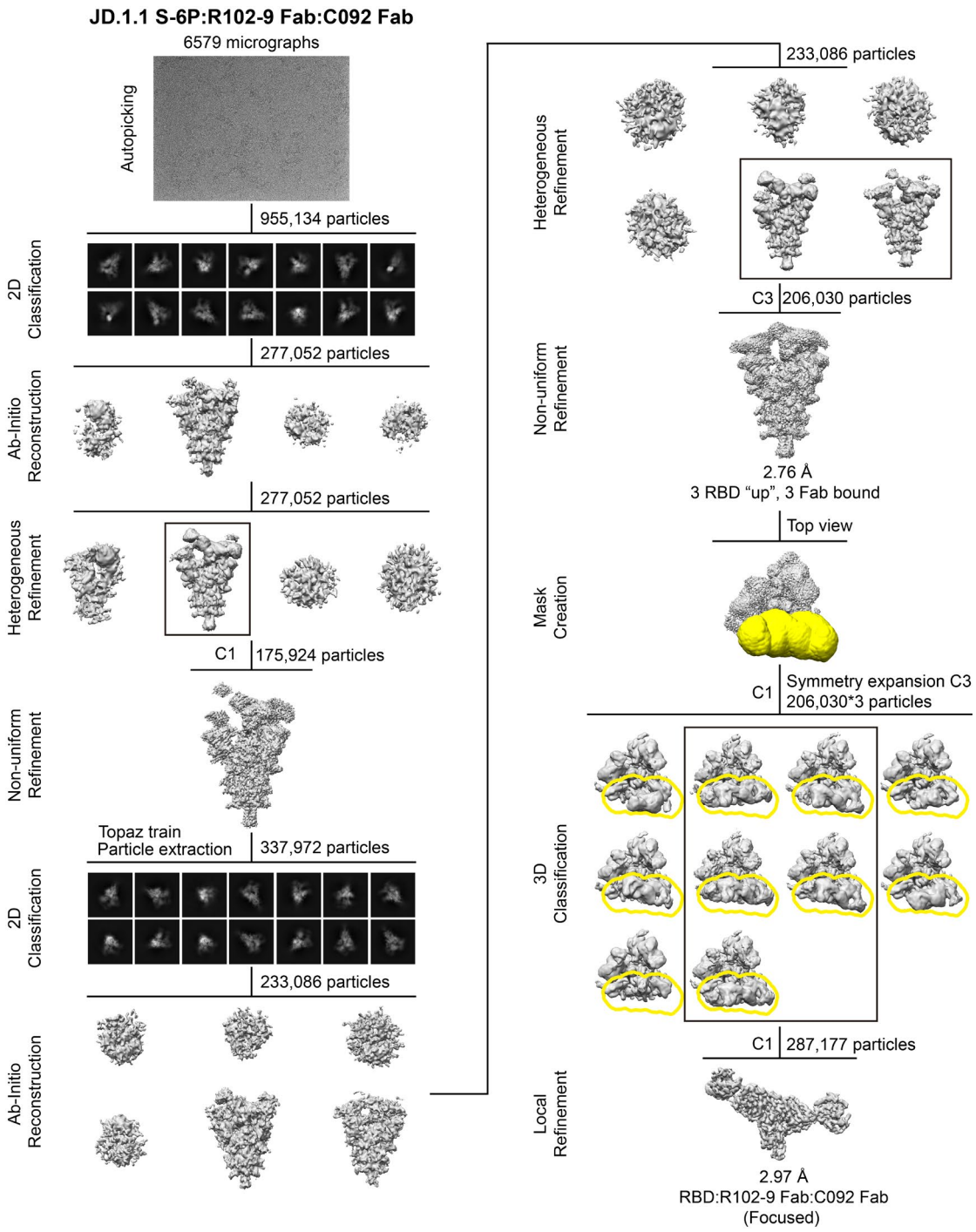


Figure S8 Cryo-EM data processing flow-chart for the structure of JD.1.1 S-6P:R102-9 Fab:C092 Fab complex. After 2D classification, ab-initio reconstruction, heterogeneous refinement and non-uniform refinement were performed. High-quality particles were used to train the topaz picker, followed by particle re-extraction, 2D classification, ab-initio reconstruction, heterogeneous refinement, and non-uniform refinement. The structure of the RBD-Fab portion was further improved with focused refinement. Selected classes are highlighted with black boxes, and focused mask is shown in yellow.

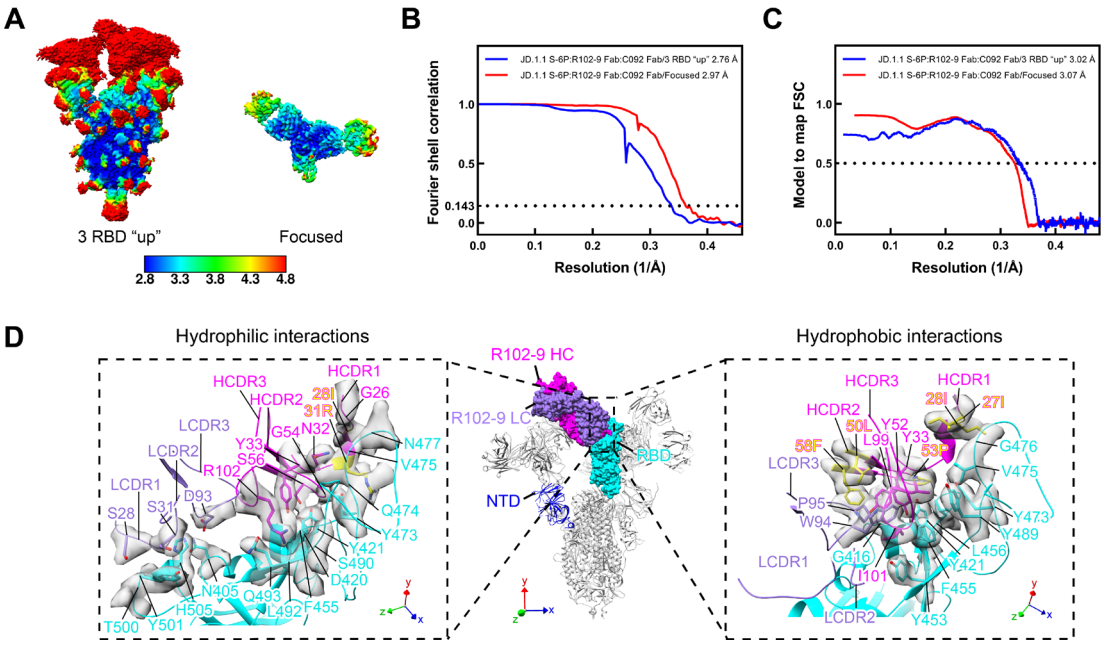


Figure S9 Resolution assessment of the structure of JD.1.1 S-6P:R102-9 Fab:C092 Fab complex.

(**A**) Local resolution map for the JD.1.1 S-6P:R102-9 Fab:C092 Fab complex.

(**B**) Global resolution assessment by Fourier shell correlation at the 0.143 criterion.

(**C**) Correlations of model vs map by Fourier shell correlation at the 0.5 criterion.

(**D**) Representative densities of the R102-9 epitope and R102-9 CDR loops in the focused map.

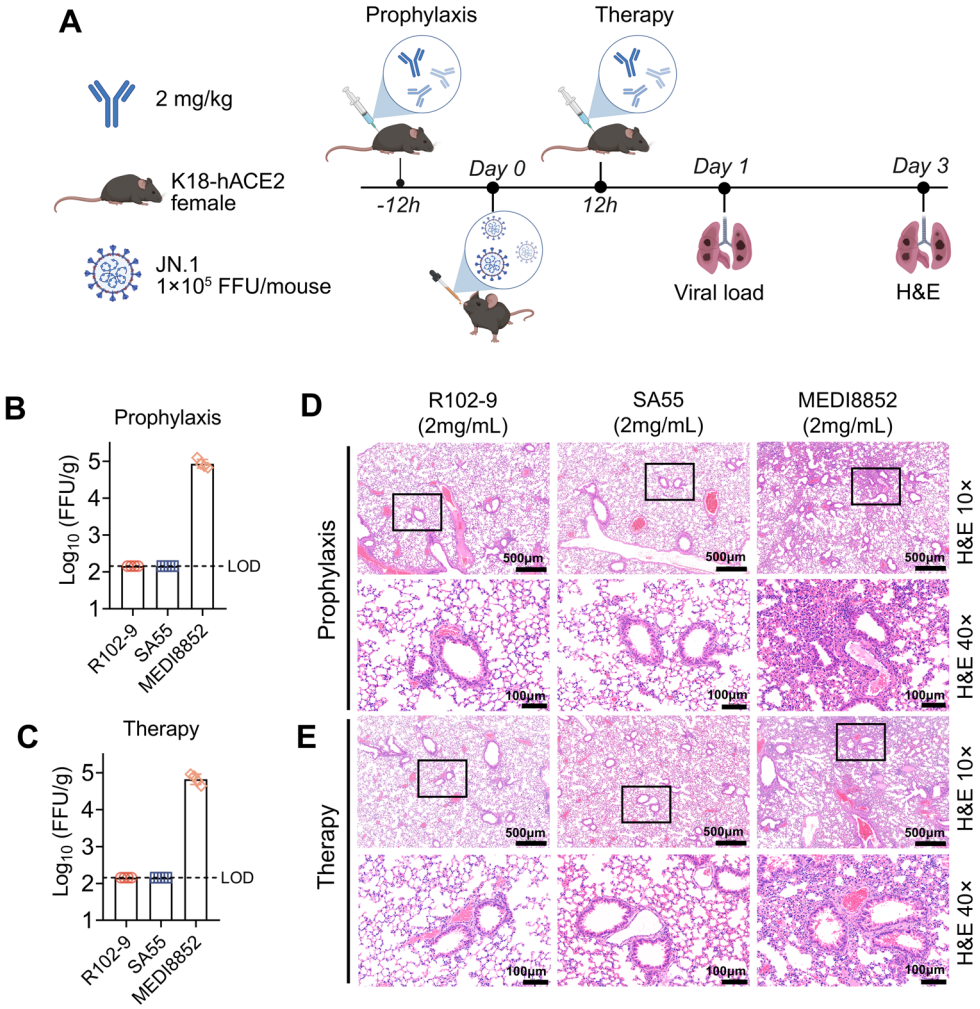


Figure S10 The prophylactic and therapeutic effects of R102-9 against JN.1.

(**A**) Flowchart of the experimental design of testing the prophylactic and therapeutic activity of R102-9 to protect against challenge with JN.1 SARS-CoV-2 in K18-hACE2 mice. (**B-C**) Infectious virions were tested by viral fluorescent assay in lung tissue homogenates of prophylactic (**B**) and therapeutic (**C**) assay. FFUs per g of tissue extractions were compared between different groups in $\log_{10}$-transformed units. LOD, limit of detection. (**D-E**) Representative images of mouse lung tissues of prophylactic (**D**) and therapeutic (**E**) by H&E staining.