1

Molecular analysis of a public cross-neutralizing antibody response to SARS-CoV-2

Meng Yuan[1], Yiquan Wang[2], Huibin Lv[2,3], Ian A. Wilson[1,4], Nicholas C. Wu[1,5,6,7,§]

[1] Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037, USA

[2] Department of Biochemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[3] HKU-Pasteur Research Pole, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China

[4] The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037, USA

[5] Center for Biophysics and Quantitative Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[6] Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[7] Carle Illinois College of Medicine, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[§] Correspondence: nicwu@illinois.edu (N.C.W.)

22   **ABSTRACT**

23   As SARS-CoV-2 variants of concerns (VOCs) continue to emerge, cross-neutralizing antibody

24   responses become key towards next-generation design of a more universal COVID-19 vaccine.

25   By analyzing published data from the literature, we report here that the combination of germline

26   genes IGHV2-5/IGLV2-14 represents a public antibody response to the receptor-binding domain

27   (RBD) that potently cross-neutralizes all VOCs to date, including Omicron and its sub-lineages.

28   Detailed molecular analysis shows that the complementarity-determining region H3 sequences of

29   IGHV2-5/IGLV2-14-encoded RBD antibodies have a preferred length of 11 amino acids and a

30   conserved HxIxxI motif. In addition, these antibodies have a strong allelic preference due to an

31   allelic polymorphism at amino-acid residue 54 of IGHV2-5, which locates at the paratope. These

32   findings have important implications for understanding cross-neutralizing antibody responses to

33   SARS-CoV-2 and its heterogenicity at the population level as well as the development of a

34   universal COVID-19 vaccine.

35 **MAIN**

36 The effectiveness of COVID-19 vaccines has been challenged by the evolution of diverse SARS-

37 CoV-2 variants in the past two years. The recent emergence of Omicron and its sub-lineages

38 BA.2, BA.2.12.1, BA.4, and BA.5 further highlights the urgent need for a more broadly protective

39 vaccine. An ideal COVID-19 vaccine should elicit high titers of neutralizing antibodies that are

40 potent against antigenically distinct variants. However, many potent neutralizing antibodies only

41 have limited cross-reactivity for variants other than the immunizing strain. For example, a major

42 class of antibodies to the receptor-binding domain (RBD) that are encoded by IGHV3-53/3-66 are

43 highly potent against the ancestral Hu-1 strain, but most of them lose their activity against many

44 other variants [1, 2]. Similarly, Beta-specific antibodies can be elicited without cross-neutralizing

45 activity against ancestral or other variants [3]. On the other hand, antibodies to S2 are typically

46 broadly reactive but have weak neutralizing activity [4-6]. Nevertheless, a few RBD antibodies

47 exhibit marked neutralization potency and breadth, as exemplified by those to the RBS-D epitope

48 [2].

49

50 One representative RBS-D antibody is LY-CoV1404 (also known as Bebtelovimab), which is a

51 monoclonal therapeutic antibody from Eli Lilly. LY-CoV1404 is encoded by IGHV2-5/IGLV2-14

52 and can cross-neutralize the ancestral Hu-1 strain as well as all known variants of concern (VOCs),

53 including Omicron and circulating sub-lineages [7, 8]. In fact, the binding mode of LY-CoV1404 is

54 identical to the cross-neutralizing antibody 2-7, which is also encoded by IGHV2-5/IGLV2-14 [9].

55 More recently, Veesler and colleagues reported another potently cross-neutralizing antibody with

56 similar sequences and binding mode as LY-CoV1404 [10]. As IGHV2-5 was shown to be an

57 important contributor to cross-neutralizing antibody response [11], the observations above

58 stimulated a systematic analysis of IGHV2-5/IGLV2-14-encoded RBD antibodies to SARS-CoV-

59 2.

60

61 In our previous study, we assembled a dataset of ~8,000 antibodies to SARS-CoV-2 spike (S)

62 protein [12]. This dataset contains seven IGHV2-5/IGLV2-14-encoded RBD antibodies, including

63 LY-CoV1404, from six different donors [7, 13-17]. In addition, four additional IGHV2-5/IGLV2-14-

64 encoded RBD antibodies were reported in a recent study [18]. Our analysis here is therefore

65 based on a total of 11 IGHV2-5/IGLV2-14-encoded RBD antibodies from at least seven donors.

66 Three of these 11 antibodies have available information for the complete nucleotide sequence,

67 nine have complete amino-acid sequence information, 10 have amino-acid sequence information

68 for the complementarity-determining regions (CDRs) H3 and L3, and four have structure

69 information. Neutralizing data from previous studies have demonstrated that these IGHV2-

70 5/IGLV2-14-encoded RBD antibodies have high cross-neutralizing activity [7, 18-20], some of

71 which remain potent against Omicron (**Figure 1A**). Previous studies have also shown that they

72 compete with ACE2 for RBD binding [7, 18, 20, 21] (**Figure S1**).

73

74 Next, we performed a structural analysis to uncover the sequence determinants of IGHV2-

75 5/IGLV2-14-encoded antibodies for RBD engagement. For antibody residues, the Kabat

76 numbering scheme is used unless otherwise stated. All four IGHV2-5/IGLV2-14-encoded RBD

77 antibodies with available structural information exhibit the same binding mode to the RBD (**Figure

78 1B**). As observed in LY-CoV1404, most amino-acid side chains in the paratope are germline-

79 encoded and form key interactions with the RBD (**Figure 1C**). For example, germline-encoded V_H

80 S32 in the CDR H1 fits into a polar pocket in the RBD. In addition, germline-encoded V_H Y52, D54,

81 D56, and R58 in the CDR H2 form an extensive network of H-bonds and electrostatic interactions

82 with the RBD. Furthermore, two key paratope residues in the light chain V_L Y32 and Y91 are also

83 germline-encoded. These observations demonstrate that the RBD-binding determinants are

84 encoded in the germline sequences of IGHV2-5 and IGLV2-14. Consistently, several IGHV2-

85 5/IGLV2-14-encoded RBD antibodies have very few somatic hypermutations (SHMs) (**Table S1**).

86   For example, S24-223 has only one SHM, and COV2-2268 and 2-7 have only four each. Of note,

87   none of their SHMs overlap.

88

89   Additional sequence analysis indicated that IGHV2-5/IGLV2-14-encoded RBD antibodies had a

90   strong allelic preference towards IGHV2-5*02. Eight out of 11 IGHV2-5/IGLV2-14-encoded RBD

91   antibodies could be assigned to IGHV2-5*02, while the allele usage for the other three was

92   ambiguous (**Figure 1D and Table S1**). In contrast, analysis of the B cell repertoire in 13 healthy

93   donors [22, 23] showed that alleles IGHV2-5*01 and IGHV2-5*02 were both commonly used, with

94   a frequency of 33% and 64%, respectively, among all IGHV2-5 antibodies (**Figure 1D and Table**

95   **S2**). The lack of IGHV2-5*01 among IGHV2-5/IGLV2-14-encoded RBD antibodies is likely due to

96   an allelic polymorphism at residue 54. IGHV2-5*01 and IGHV2-5*02 have Asn and Asp,

97   respectively, at residue 54. V_H D54 in IGHV2-5/IGLV2-14-encoded RBD antibodies plays an

98   important role in RBD binding through a salt bridge with RBD K444 and an H-bond with RBD

99   N450 (**Figure 1C**). Replacing the Asp at V_H residue 54 by Asn would convert the salt bridge with

100   RBD K444 to a H-bond, which would likely significantly weaken the binding energy. Consistently,

101   all eight of the nine IGHV2-5/IGLV2-14-encoded RBD antibodies with sequence information

102   available have an Asp at V_H residue 54, whereas the remaining one has a Glu at V_H residue 54

103   (**Table S1**). These findings provide a mechanistic basis for the allelic preference against IGHV2-

104   5*01, despite its prevalence in the human population. Coincidentally, an almost identical

105   observation was observed in an IGHV2-5-encoded HIV antibody, in which V_H D54 results in much

106   stronger binding than V_H N54 [24].

107

108   Lastly, we analyzed the CDR H3 sequences of the IGHV2-5/IGLV2-14-encoded RBD antibodies.

109   Among 10 IGHV2-5/IGLV2-14-encoded RBD antibodies with CDR H3 sequence information

110   available, eight had a CDR H3 length of 11 amino acids (IMGT numbering) and came from at

111   least five patients (**Figure 2A**). The CDR H3 sequences from these eight antibodies shared a

112   motif HxIxxI or conserved variations of it, including HxIxxL and HxVxxI (**Figures 2A and 2B**). The

113   HxIxxI motif consisted of V_H H95, I97, and I100 (Kabat numbering) and is uncommon among the

114   CDR H3 sequences of IGHV2-5-encoded antibodies in the human antibody repertoire (**Figure**

115   **2B**). V_H H95, I97, and I100 in the HxIxxI motif play critical roles in stabilizing the loop conformation

116   as well as RBD binding (**Figure 2C**). V_H H95 forms two intramolecular H-bonds to stabilize the

117   CDR H3 loop. The first H-bond involves the side chain of V_H Y52, which in turn H-bonds with RBD

118   V445 amide nitrogen. The second H-bond involves the backbone carbonyl of V_H I100. In addition,

119   V_H H95 also forms van der Waals interaction with RBD V445. V_H I97 at the tip of the CDR H3 loop

120   inserts into a hydrophobic pocket formed by RBD V445 and P499, as well as the aliphatic portion

121   of RBD N440. V_H I100 helps position V_L Y91 to interact with RBD V445 and P499. As shown by

122   IgBlast analysis [25], the HxIxxI motif is largely encoded by N-nucleotide addition, although V_H

123   I97 may sometimes be encoded by an IGHD gene (**Figure 2D**). Of note, while CDR H3 of XG005

124   has 12 amino acids (**Figure 2A**), it adopts a similar conformation to those with 11 amino acids

125   (**Figure S2**). Overall, IGHV2-5/IGLV2-14-encoded RBD antibodies with a CDR H3 length of 11

126   amino acids have convergent CDR H3 sequences, and thus can be classified as a public

127   clonotype.

128

129   Due to the continuous evolution of SARS-CoV-2 VOCs, identification of cross-neutralizing human

130   monoclonal antibodies has been a global research focus. IGHV1-58/IGKV3-20-encoded RBD

131   antibodies are perhaps the most well-characterized public antibody clonotype that is cross-

132   neutralizing against multiple SARS-CoV-2 VOCs [3, 12, 26-29]. However, recent studies have

133   shown that many IGHV1-58/IGKV3-20-encoded RBD antibodies have minimal neutralizing

134   activity against Omicron and its sub-lineages due to mutations Q493R and F486V on the RBD

135   [30-32]. In comparison, IGHV2-5/IGLV2-14-encoded RBD antibodies, which mostly retain

136   potency against Omicron and its sub-lineages (**Figure 1A**) [7, 8, 18-20, 33], have higher

137   neutralization breadth. Since IGHV2-5/IGLV2-14-encoded RBD antibodies are also a public

138   antibody clonotype, they further substantiate the rationale and strategy for development of a more

139   universal COVID-19 vaccine.

140

141   Nevertheless, some individuals may have difficulties generating an IGHV2-5/IGLV2-14-encoded

142   RBD antibody response, due to the alleles that they possess (**Figure 1D**). Since there is no known

143   copy number variation for IGHV2-5 [34], each person should carry two copies of IGHV2-5 in the

144   genome. If both copies are IGHV2-5*01 allele, the person may not have the suitable B cell

145   germline clone to produce a IGHV2-5/IGLV2-14-encoded RBD antibody response. In fact, donor

146   112 in the 13 healthy donors that were analyzed in this study is very likely to be IGHV2-5*01

147   homozygous, since 94% of its IGHV2-5-encoded antibodies were assigned to IGHV2-5*01 (**Table**

148   **S2**). Moreover, the conserved HxIxxI motif in CDR H3 of IGHV2-5/IGLV2-14-encoded RBD

149   antibodies is mostly encoded by random N-nucleotide addition. As a result, B cell germline clones

150   that can produce IGHV2-5/IGLV2-14-encoded RBD antibodies may be relatively rare. These

151   results may provide a genetic basis for heterogenicity in the cross-neutralizing antibody response

152   among different individuals. While allelic preference has previously been described for

153   neutralizing antibodies to other viruses [24, 35-37], its clinical implications for COVID-19 remain

154   to be fully explored.

155   **MATERIALS AND METHODS**

156   **Dataset collection**

157   The information on antibodies S24-223, P2B-1E4, 2-7, LY-CoV1404, XG005, XG031, and COV2-

158   2268 were compiled in our previous study [12], whereas the information on XGv042, XGv264,

159   XGv265, and XGv266 were compiled in CoV-AbDab [38]. Neutralization data of each monoclonal

160   antibody were collected from the original papers (**Table S1**). Somatic hypermutations were

161   identified by IgBlast [25].

162

163   **Allele assignment of IGHV2-5/IGLV2-14-encoded RBD antibodies**

164   For antibodies P2B-1E4, XG005, and XG031, the allele information was obtained from the original

165   publications [14, 16]. For other antibodies, IgBlast was used to assign the allele of each antibody

166   [25]. Nucleotide sequence, if available, was used as input for IgBlast. Otherwise, protein

167   sequence was used. If an antibody showed equally likely to be encoded by two or more alleles,

168   the allele assignment would be classified as "ambiguous". All "ambiguous" allele assignments in

169   this study came from antibodies that do not have nucleotide sequence information available,

170   namely XGv264, XGv265, and XGv266. Of note, while IgBlast showed that XGv266 was equally

171   likely to be encoded by IGHV2-5*01 and IGHV2-5*02, we postulated that XGv266 should be

172   assigned to IGHV2-5*02 at the nucleotide level. Specifically, XGv266 had a Glu at V_H residue 54,

173   which was one nucleotide change from the Asp codon used (IGHV2-5*02) but two nucleotides

174   away from Asn (IGHV2-5*01). However, IgBlast did not utilize codon information for allele

175   assignment when the amino-acid sequence was used as input.

176

177   **Analysis of allele usage in published antibody repertoire**

178   Published antibody repertoire sequencing datasets from 13 healthy donors [22, 23] were

179   downloaded from cAb-Rep [39]. Putative germline gene alleles for each antibody sequence in

180   these repertoire sequencing datasets from healthy donors were identified by IgBLAST [25].

181

Analysis of CDR H3 sequences

Sequence alignment was performed using MAFFT [40]. Antibody sequences in the human antibody repertoire were downloaded from the Observed Antibody Space [41]. IGHV2-5 antibodies as well as their CDR H3 sequences were identified using IgBLAST [25]. Sequence logos were generated by WebLogo [42]. Putative germline sequences and segments in the V-D-J junctions were identified by IgBLAST [25].

Code Availability

Custom codes for all analyses have been deposited to https://github.com/nicwulab/IGHV2-5_RBD_Abs.

AUTHOR CONTRIBUTIONS

M.Y. and N.C.W. conceived and designed the study. All authors performed data analysis. M.Y. and N.C.W. wrote the paper and all authors reviewed and/or edited the paper.

COMPETING INTERESTS

The authors declare no competing interests.

REFERENCES

1. 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 (2021)

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

3. Reincke SM, *et al.* SARS-CoV-2 Beta variant infection elicits potent lineage-specific and cross-reactive antibodies. *Science* **375**, 782-787 (2022)

4. Zhou P, *et al.* A human antibody reveals a conserved site on beta-coronavirus spike proteins and confers protection against SARS-CoV-2 infection. *Sci Transl Med* **14**, eabi9215 (2022)

5. Pinto D, *et al.* Broad betacoronavirus neutralization by a stem helix-specific human antibody. *Science* **373**, 1109-1116 (2021)

6. Li W, *et al.* Structural basis and mode of action for two broadly neutralizing antibodies against SARS-CoV-2 emerging variants of concern. *Cell Rep* **38**, 110210 (2022)

7. Westendorf K, *et al.* LY-CoV1404 (bebtelovimab) potently neutralizes SARS-CoV-2 variants. *Cell Reports* **39**, 110812 (2022)

8. Cao Y, *et al.* BA.2.12.1, BA.4 and BA.5 escape antibodies elicited by Omicron infection. *bioRxiv*, doi: 10.1101/2022.04.30.489997, (2022)

9. Kramer KJ, *et al.* Potent neutralization of SARS-CoV-2 variants of concern by an antibody with an uncommon genetic signature and structural mode of spike recognition. *Cell Rep* **37**, 109784 (2021)

10. Park YJ, *et al.* Imprinted antibody responses against SARS-CoV-2 Omicron sublineages. *bioRxiv*, doi: 10.1101/2022.05.08.491108, (2022)

11. Andreano E, *et al.* Hybrid immunity improves B cells and antibodies against SARS-CoV-2 variants. *Nature* **600**, 530-535 (2021)

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

13. Dugan HL, *et al.* Profiling B cell immunodominance after SARS-CoV-2 infection reveals antibody evolution to non-neutralizing viral targets. *Immunity* **54**, 1290-1303.e1297 (2021)

14. Ju B, *et al.* Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature* **584**, 115-119 (2020)

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

240   16.   Zhou Y, *et al.* Enhancement versus neutralization by SARS-CoV-2 antibodies from a
241   convalescent donor associates with distinct epitopes on the RBD. *Cell Rep* **34**, 108699
242   (2021)

243   17.   Zost SJ, *et al.* Rapid isolation and profiling of a diverse panel of human monoclonal
244   antibodies targeting the SARS-CoV-2 spike protein. *Nat Med* **26**, 1422-1427 (2020)

245   18.   Wang K, *et al.* Memory B cell repertoire from triple vaccinees against diverse SARS-
246   CoV-2 variants. *Nature* **603**, 919-925 (2022)

247   19.   Liu L, *et al.* Striking antibody evasion manifested by the Omicron variant of SARS-CoV-
248   2. *Nature* **602**, 676-681 (2022)

249   20.   Liu Z, *et al.* An ultrapotent pan-β-coronavirus lineage B (β-CoV-B) neutralizing antibody
250   locks the receptor-binding domain in closed conformation by targeting its conserved
251   epitope. *Protein Cell*, doi: 10.1007/s13238-021-00871-6, (2021)

252   21.   Cerutti G, *et al.* Structural basis for accommodation of emerging B.1.351 and B.1.1.7
253   variants by two potent SARS-CoV-2 neutralizing antibodies. *Structure* **29**, 655-663.e654
254   (2021)

255   22.   Briney B, Inderbitzin A, Joyce C, Burton DR. Commonality despite exceptional diversity
256   in the baseline human antibody repertoire. *Nature* **566**, 393-397 (2019)

257   23.   Soto C, *et al.* High frequency of shared clonotypes in human B cell receptor repertoires.
258   *Nature* **566**, 398-402 (2019)

259   24.   Alam SM, *et al.* Differential reactivity of germ line allelic variants of a broadly neutralizing
260   HIV-1 antibody to a gp41 fusion intermediate conformation. *J Virol* **85**, 11725-11731
261   (2011)

262   25.   Ye J, Ma N, Madden TL, Ostell JM. IgBLAST: an immunoglobulin variable domain
263   sequence analysis tool. *Nucleic Acids Res* **41**, W34-40 (2013)

264   26.   Dejnirattisai W, *et al.* The antigenic anatomy of SARS-CoV-2 receptor binding domain.
265   *Cell* **184**, 2183-2200.e2122 (2021)

266   27.   Robbiani DF, *et al.* Convergent antibody responses to SARS-CoV-2 in convalescent
267   individuals. *Nature* **584**, 437-442 (2020)

268   28.   Tortorici MA, *et al.* Ultrapotent human antibodies protect against SARS-CoV-2 challenge
269   via multiple mechanisms. *Science* **370**, 950-957 (2020)

270   29.   Wang L, *et al.* Ultrapotent antibodies against diverse and highly transmissible SARS-
271   CoV-2 variants. *Science* **373**, eabh1766 (2021)

272   30.   Dong J, *et al.* Genetic and structural basis for SARS-CoV-2 variant neutralization by a
273   two-antibody cocktail. *Nat Microbiol* **6**, 1233-1244 (2021)

274   31.   Starr TN, *et al.* SARS-CoV-2 RBD antibodies that maximize breadth and resistance to
275   escape. *Nature* **597**, 97-102 (2021)

276 32.   FDA. Fact sheet for healthcare providers: emergency ese authorization for Evusheld
277   (tixagevimab co-packaged with cilgavimab). doi, (2021)

278 33.   Yamasoba D*, et al.* Sensitivity of novel SARS-CoV-2 Omicron subvariants, BA.2.11,
279   BA.2.12.1, BA.4 and BA.5 to therapeutic monoclonal antibodies. *bioRxiv*, doi:
280   10.1101/2022.05.03.490409, (2022)

281 34.   Lefranc MP, Lefranc G. Immunoglobulins or antibodies: IMGT bridging genes, structures
282   and functions. *Biomedicines* **8**, 319 (2020)

283 35.   Ying T*, et al.* Junctional and allele-specific residues are critical for MERS-CoV
284   neutralization by an exceptionally potent germline-like antibody. *Nat Commun* **6**, 8223
285   (2015)

286 36.   Lang S*, et al.* Antibody 27F3 broadly targets influenza A group 1 and 2 hemagglutinins
287   through a further variation in V_H1-69 antibody orientation on the HA stem. *Cell Rep* **20**,
288   2935-2943 (2017)

289 37.   Avnir Y*, et al.* IGHV1-69 polymorphism modulates anti-influenza antibody repertoires,
290   correlates with IGHV utilization shifts and varies by ethnicity. *Sci Rep* **6**, 20842 (2016)

291 38.   Raybould MIJ, Kovaltsuk A, Marks C, Deane CM. CoV-AbDab: the coronavirus antibody
292   database. *Bioinformatics* **37**, 734-735 (2021)

293 39.   Guo Y*, et al.* cAb-Rep: a database of curated antibody repertoires for exploring antibody
294   diversity and predicting antibody prevalence. *Front Immunol* **10**, 2365 (2019)

295 40.   Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7:
296   improvements in performance and usability. *Mol Biol Evol* **30**, 772-780 (2013)

297 41.   Kovaltsuk A*, et al.* Observed Antibody Space: a resource for data mining next-
298   generation sequencing of antibody repertoires. *J Immunol* **201**, 2502-2509 (2018)

299 42.   Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator.
300   *Genome Res* **14**, 1188-1190 (2004)

301 43.   Lan J*, et al.* Structure of the SARS-CoV-2 spike receptor-binding domain bound to the
302   ACE2 receptor. *Nature* **581**, 215-220 (2020)
303
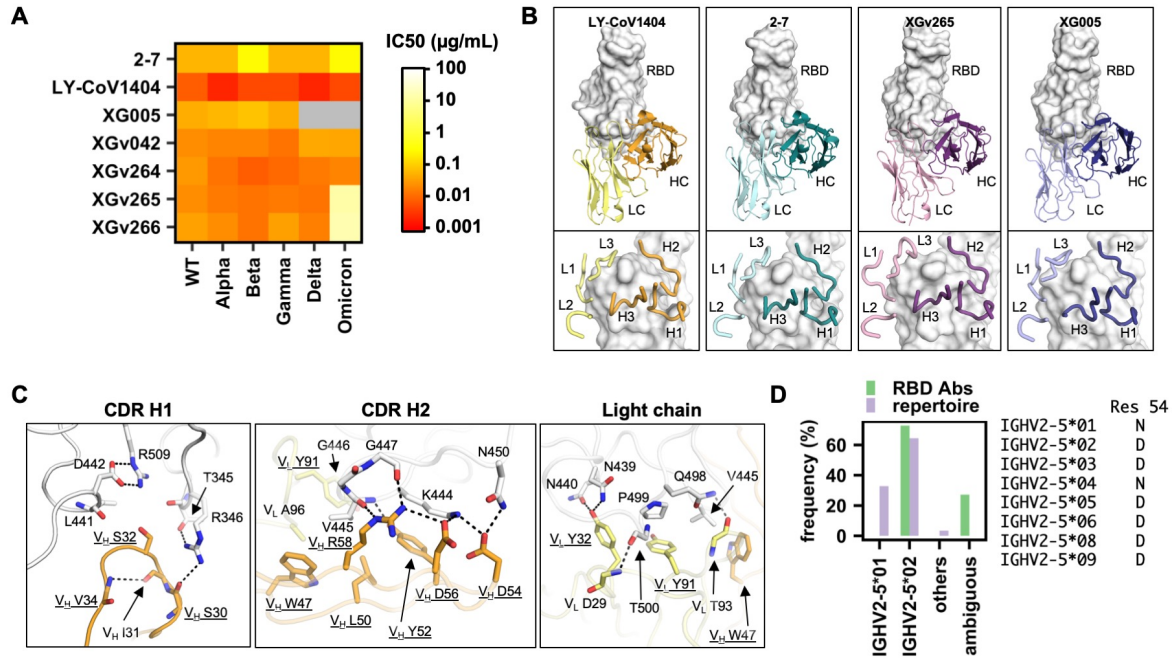
Figure 1. IGHV2-5/IGLV2-14 is a public antibody response with strong allelic preference.
(A) The half maximal inhibitory concentration (IC$_{50}$) of IGHV2-5/IGLV2-14-encoded RBD antibodies against different SARS-CoV-2 VOCs in pseudovirus assays. Data were taken from previous studies [7, 18-20]. **(B)** Four IGHV2-5/IGLV2-14-encoded RBD antibodies have structure information available. Their binding modes to RBD (white surface) are compared. Upper panels: heavy chain (HC) and light chain (LC) of each antibody are shown. Bottom panels: zoom-in views with six CDR loops of each antibody shown. LY-CoV1404: PDB 7MMO [7]. 2-7: PDB 7LSS [21]. XGv265: PDB 7WEE [18]. XG005: PDB 7V26 [20]. **(C)** Key interactions between LY-CoV1404 and RBD are shown. Hydrogen bonds and salt bridges are represented by black dashed lines. All germline-encoded residues are underlined. Heavy chain is in orange, light chain in yellow, and RBD is in white. **(D)** IGHV allele usage of the 11 IGHV2-5/IGLV2-14-encoded RBD antibodies (RBD Abs) is compared with that of IGHV2-5-encoded antibodies in published repertoire sequencing datasets from 13 healthy donors [22, 23]. The amino-acid identity at residue 54 of each IGHV2-5 allele is indicated.

319
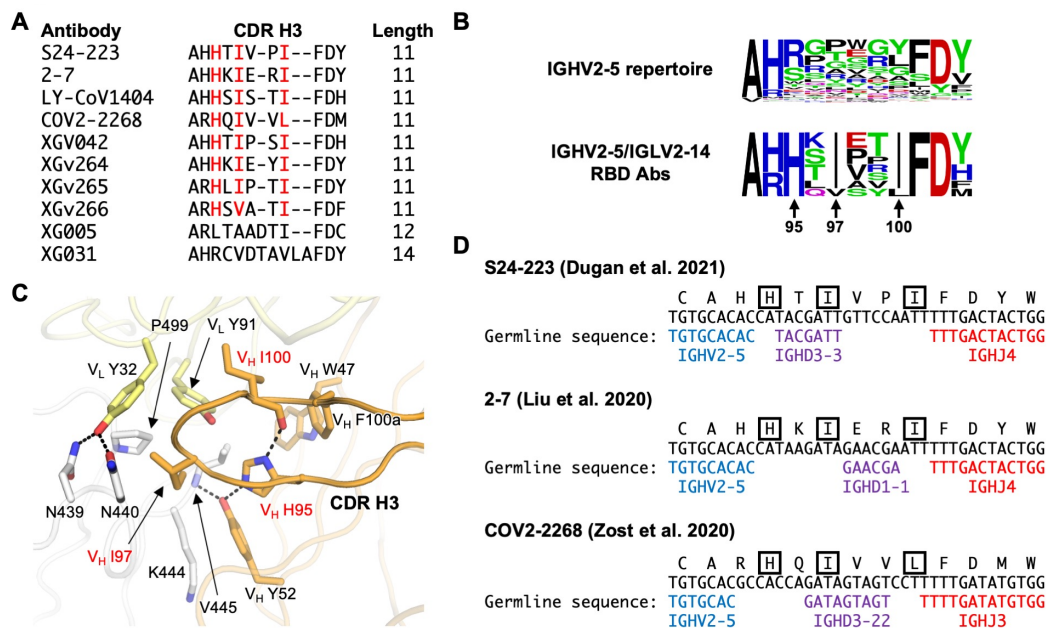
**Figure 2. HxIxxI is a common motif in IGHV2-5/IGLV2-14-encoded RBD antibodies. (A)** CDR H3 sequences (IMGT numbering) from IGHV2-5/IGLV2-14-encoded RBD antibodies are aligned. Residues of interest are highlighted in red. **(B)** CDR H3 sequences (IMGT numbering) of IGHV2-5/IGLV2-14-encoded RBD antibodies and IGHV2-5-encoded antibodies in the human antibody repertoire are shown as sequence logos. Only those antibodies with a CDR H3 length of 11 amino acids are included in this analysis. Residues of interest are labeled. Sequences of IGHV2-5-encoded antibodies in the human antibody repertoire were downloaded from the Observed Antibody Space [41]. A total of 9,197 IGHV2-5-encoded antibodies in the human antibody repertoire were analyzed here. Of note, while Kabat numbering was used for the residue position, IMGT numbering was used for defining CDR H3 length. **(C)** Interaction between the CDR H3 of LY-CoV1404 and RBD is shown. PDB 7MMO [7]. Hydrogen bonds are represented by black dashed lines. Heavy chain is in orange, light chain in yellow, and RBD is in white. Residues of interest are highlighted in red. **(D)** Amino-acid and nucleotide sequences of the V-D-J junction of three IGHV2-5/IGLV2-14-encoded RBD antibodies are shown. Putative germline sequences and segments are indicated. Residues of interest are boxed.
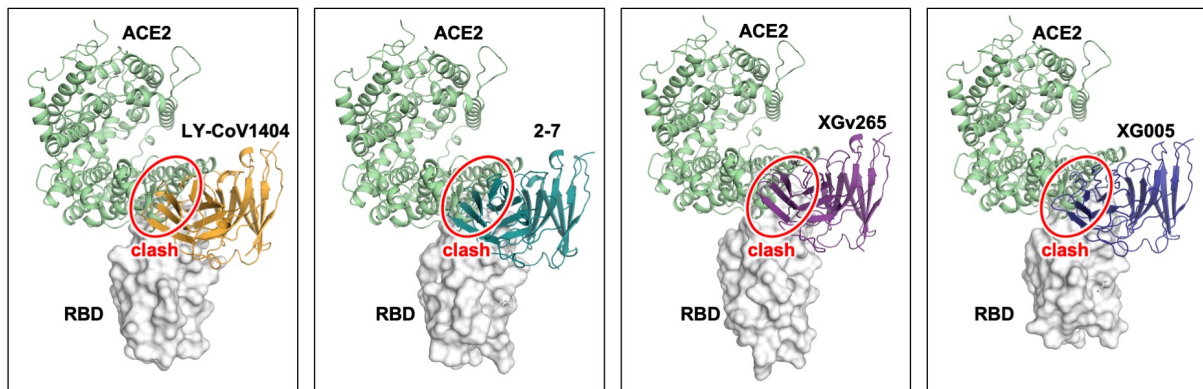
335
336 **Figure S1. IGHV2-5/IGLV2-14 antibodies would clash with ACE2 binding.** Structures of
337 antibody/RBD complexes are superimposed onto the RBD/ACE2 complex structure (PDB 6M0J)
338 [43]. RBD is represented by a white surface. ACE2 is shown as green cartoon. LY-CoV1404: PDB
339 7MMO [7]. 2-7: PDB 7LSS [21]. XGv265: PDB 7WEE [18]. XG005: PDB 7V26 [20].
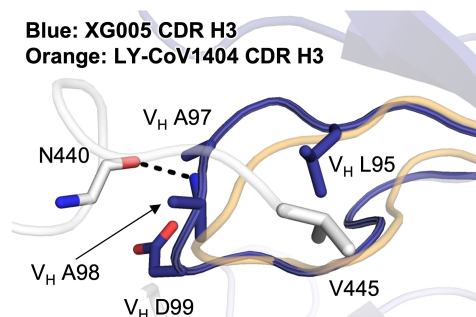
Figure S2. Interactions between RBD and CDR H3 of XG005. A cryo-EM structure of SARS-CoV-2 spike protein in complex with XG005 (PDB 7V26) that was reported in a previous study [20] is shown. A hydrogen bond between XG005 and the RBD is represented by a black dashed line. The CDR H3 of LY-CoV1404 (PDB 7MMO) [7] is also shown here as a transparent orange cartoon to demonstrate the relative positions of the CDR H3 loops from these two antibodies after superimposition.