

1 **Title:** MHC-II constrains the natural neutralizing antibody response to the SARS-CoV-2 spike

2 RBM in humans

3 **Authors:** Andrea Castro^{1,2}, Kivilcim Ozturk², Maurizio Zanetti^{3,4}, Hannah Carter^{2,4}

4 **Affiliations:**

5 ¹Biomedical Informatics Program, University of California San Diego, La Jolla, CA 92093, USA

6 ²Department of Medicine, Division of Medical Genetics, University of California San Diego, La

7 Jolla, CA 92093, USA

8 ³The Laboratory of Immunology, Department of Medicine, University of California San Diego,

9 La Jolla, CA 92093, USA

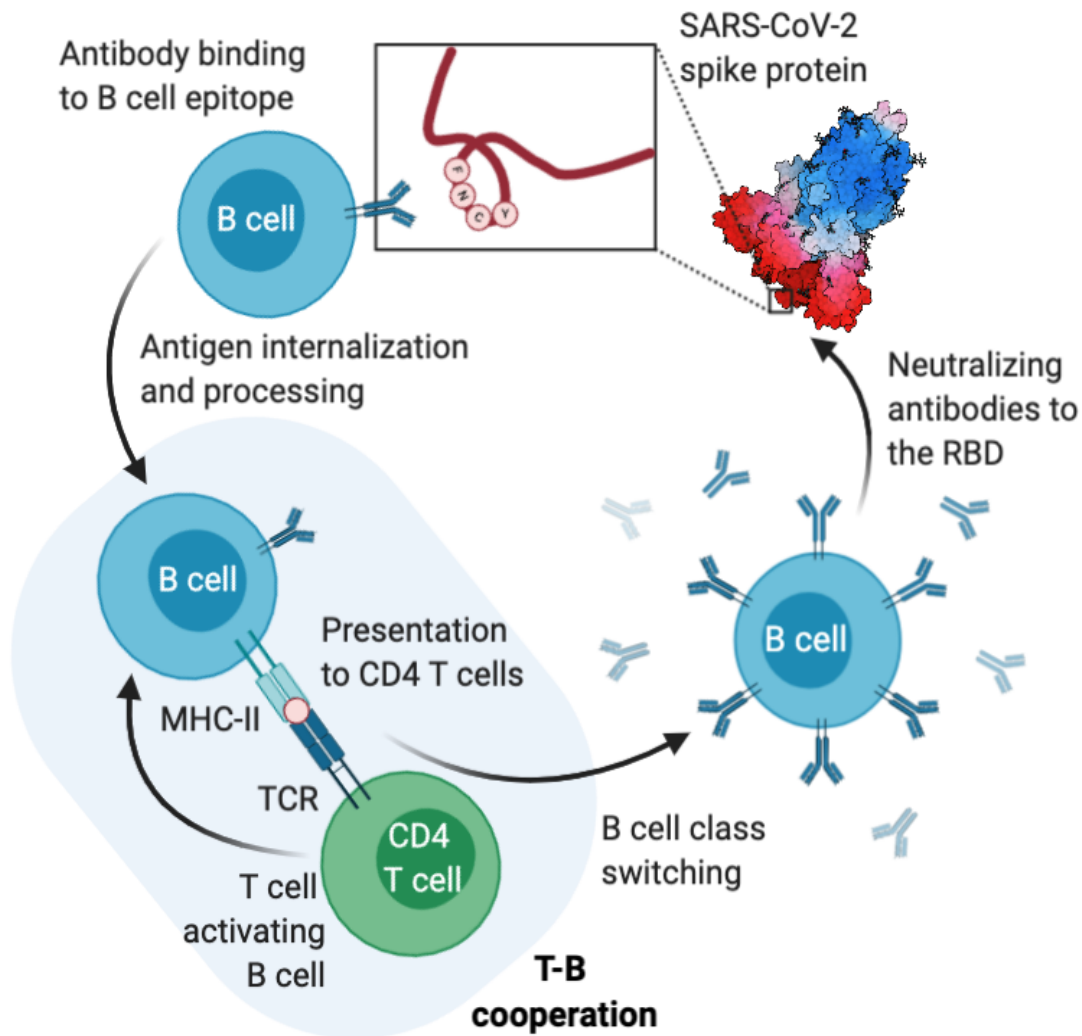
10 ⁴Moores Cancer Center, University of California San Diego, La Jolla, CA 92093, USA

11 **Corresponding authors:** Maurizio Zanetti and Hannah Carter

12 **Corresponding authors emails:** mzanetti@health.ucsd.edu and hkcarter@health.ucsd.edu

13

14 **Graphical abstract**



15

16

17 **Abstract**

18 SARS-CoV-2 antibodies develop within two weeks of infection, but wane relatively rapidly
19 post-infection, raising concerns about whether antibody responses will provide protection upon
20 re-exposure. Here we revisit T-B cooperation as a prerequisite for effective and durable
21 neutralizing antibody responses centered on a mutationally constrained RBM B cell epitope. T-B
22 cooperation requires co-processing of B and T cell epitopes by the same B cell and is subject to

23 MHC-II restriction. We evaluated MHC-II constraints relevant to the neutralizing antibody
24 response to a mutationally-constrained B cell epitope in the receptor binding motif (RBM) of the
25 spike protein. Examining common MHC-II alleles, we found that peptides surrounding this key
26 B cell epitope are predicted to bind poorly, suggesting a lack MHC-II support in T-B
27 cooperation, impacting generation of high-potency neutralizing antibodies in the general
28 population. Additionally, we found that multiple microbial peptides had potential for RBM
29 cross-reactivity, supporting previous exposures as a possible source of T cell memory.

30

31 **Keywords:** COVID-19, SARS-CoV-2, spike protein, RBD, RBM, T-B cooperation, MHC-II,
32 CD4 T cell, T cell help, neutralizing antibody, prior immunological history

33

34 **Introduction**

35 Upon infection with SARS-CoV-2 the individual undergoes seroconversion. In mildly
36 symptomatic patients, seroconversion occurs between day 7 and 14, includes IgM and IgG, and
37 outlasts virus detection with generally higher IgG levels in symptomatic than asymptomatic
38 groups in the early convalescent phase (1). Alarming, the IgG levels in both asymptomatic and
39 symptomatic patients decline during the early convalescent phase, with a median decrease of
40 ~75% within 2–3 months after infection (2). This suggests that the systemic antibody response
41 which follows natural infection with SARS-CoV-2 is rapid but short-lived, with the possibility of
42 no residual immunity after 6-12 months (3) affecting primarily neutralizing antibodies in plasma
43 (4).

44 The generation of an antibody response requires cooperation between a B cell producing
45 specific antibody molecules and a CD4 T cell (helper cell) activated by an epitope on the same

46 antigen as that recognized by the B cell (T-B cooperation) (5). This reaction occurs in the
47 germinal center (6,7). Excluded from this rule are responses against carbohydrates and antigens
48 with repeating motifs that alone cross-link the B cell antigen receptor leading to B cell activation
49 (8). Discovered over 50 years ago (9–11), it also became apparent that T-B cooperation is
50 restricted by Major Histocompatibility Complex class II (MHC-II) molecules (12–14). T-B
51 cooperation plays a key role in the facilitation and strength of the antibody response (10,15) and
52 the size of the antibody response is proportional to the number of Th cells activated by the B cell
53 during T-B cooperation (13,14,16). The importance of T cell help during the activation of
54 antigen specific B cells to protein antigens driving B cell selection is emphasized by recent
55 experiments where the injection of a conjugate of antigen (OVA) linked with an anti-DEC205
56 antibody induced a greater proliferation of DEC205+ relative to DEC205- B cells consistent with
57 a T helper effect on B cell activation (17).

58 T-B cooperation requires that the epitopes recognized by the B and T cell be on the same
59 portion of the antigen (11,18,19) leading to a model requiring the contextual internalization and
60 co-processing of T and B cell epitopes (5) which is consistent with the principle of linked (aka
61 associative) recognition of antigen (20). Studies *in vitro* using human T and B lymphocytes
62 showed that an antigen specific B cell can present antigen to CD4 T cells even if antigen is
63 present at very low concentration ($10^{-11} - 10^{-12}$ M) (21). Presentation of antigen by the B cell also
64 facilitates the cooperation between CD4 T cells of different specificities resulting in enhanced
65 generation of memory CD4 T cells (22). However, T-B cooperation is not the only form of
66 cooperative interaction among lymphocytes as cooperation exists between CD4 T and CD8 T
67 cells (23) and between two CD4 T cells responding to distinct epitopes on the same antigen (24).

68 A model based on coprocessing of T and B epitopes also led to the suggestion that
69 preferential T-B pairing could be based on topological proximity (25–29) so that during BCR-
70 mediated internalization the T cell epitope is protected by the paratope of the BCR. Indeed, a
71 more recent study showed that not only is CD4 T cell help a limiting factor in the development
72 of antibodies to smallpox (vaccinia virus), but that there also exists a deterministic epitope
73 linkage of specificities in T-B cooperation against this viral pathogen (30). Collectively, it
74 appears that T-B pairing and MHC-II restriction are key events in the selection of the antibody
75 response to pathogens and that operationally T-B cooperation and MHC-II restriction are key
76 events in the generation of an adaptive antibody response, suggesting that lack of or defective T-
77 B preferential pairing could result in an antibody response that is suboptimal, short-lived, or
78 both.

79 In SARS-CoV-2, neutralizing antibodies (NAbs) are a key defense mechanism against
80 infection and transmission. NAbs generated by single memory B cell VH/VL cloning from
81 convalescent COVID-19 patients have been extremely useful in defining the fine epitope
82 specificity of the antibody response in COVID-19 individuals. At present, SARS-CoV-2 NAbs
83 can be distinguished into three large categories. 1) Repurposed antibodies, that is, NAbs
84 discovered and characterized in the context of SARS-CoV and subsequently found to neutralize
85 SARS-CoV-2 via cross-reactivity. These antibodies map away from the receptor binding domain
86 (RBD) of the spike protein (31–33). 2) Non-RBD neutralizing antibodies discovered in SARS-
87 CoV-2 patients whose paratope is specific for sites outside the RBD (34). 3) RBD antibodies,
88 including NAbs, derived from SARS-CoV-2 patients that map to a restricted site in the RBD
89 (35–41). Cryo-EM of this third antibody category shows that they bind to residues in or around
90 the four amino acids Phe-Asp-Cys-Tyr (FNCY) in the receptor binding motif (RBM) (residues

91 437-508) which is inside the larger RBD (residues 319-541) at the virus:ACE2 interface (36).
92 Although the RBD has been shown to be an immunodominant target of serum antibodies in
93 COVID-19 patients (42), high potency NAbs are directed against a conserved portion of the
94 RBM on or around the FNCY patch, a sequence only found in the RBD of SARS-CoV-2 and not
95 in other coronaviruses. Indeed while the RBD is mutationally tolerant, the RBM is constrained to
96 the wild-type amino acids (43), implying that the B cell epitope included in this region of the
97 virus:ACE2 interface is resistant to antigenic drift. Thus, we may refer to this site as a key RBM
98 B cell epitope in the generation of potent NAbs.

99 Antibody responses against SARS-CoV-2 depend on CD4 T cell help. Spike-specific
100 CD4 T cell responses have been found to correlate with the magnitude of the anti-RBD IgG
101 response whereas non-spike CD4 T cell responses do not (44). However, spike-specific CD4 T
102 cells reactive with MHC-II peptides proximal to the central B cell epitope represent a minority
103 (~10%) of the total CD4 T cell responses, which are dominated by responses against either the
104 distal portion of the spike protein or other structural antigens (45). Surprisingly, these CD4 T cell
105 responses are largely cross-reactive and originate from previous coronavirus infections (46).

106 As mounting evidence suggests that the NAb response in COVID-19 patients is relatively
107 short-lived, we decided to test the hypothesis that associative recognition of the key RBM B cell
108 epitope and proximal MHC-II-restricted epitopes may be defective with detrimental effects on
109 preferential T-B pairing. Therefore, to quantify the potential effects of T-B cooperation *in vivo*,
110 we analyzed all 15mer putative MHC-II epitopes (+/- 50 amino acid residues) relative to the key
111 RBM B cell epitope for coverage by all known 5,620 human MHC-II alleles and predicted
112 binding affinity. The analysis shows that there exists in general less availability of effective T
113 cell epitopes in close proximity to the key RBM B cell epitope in the human population.

114

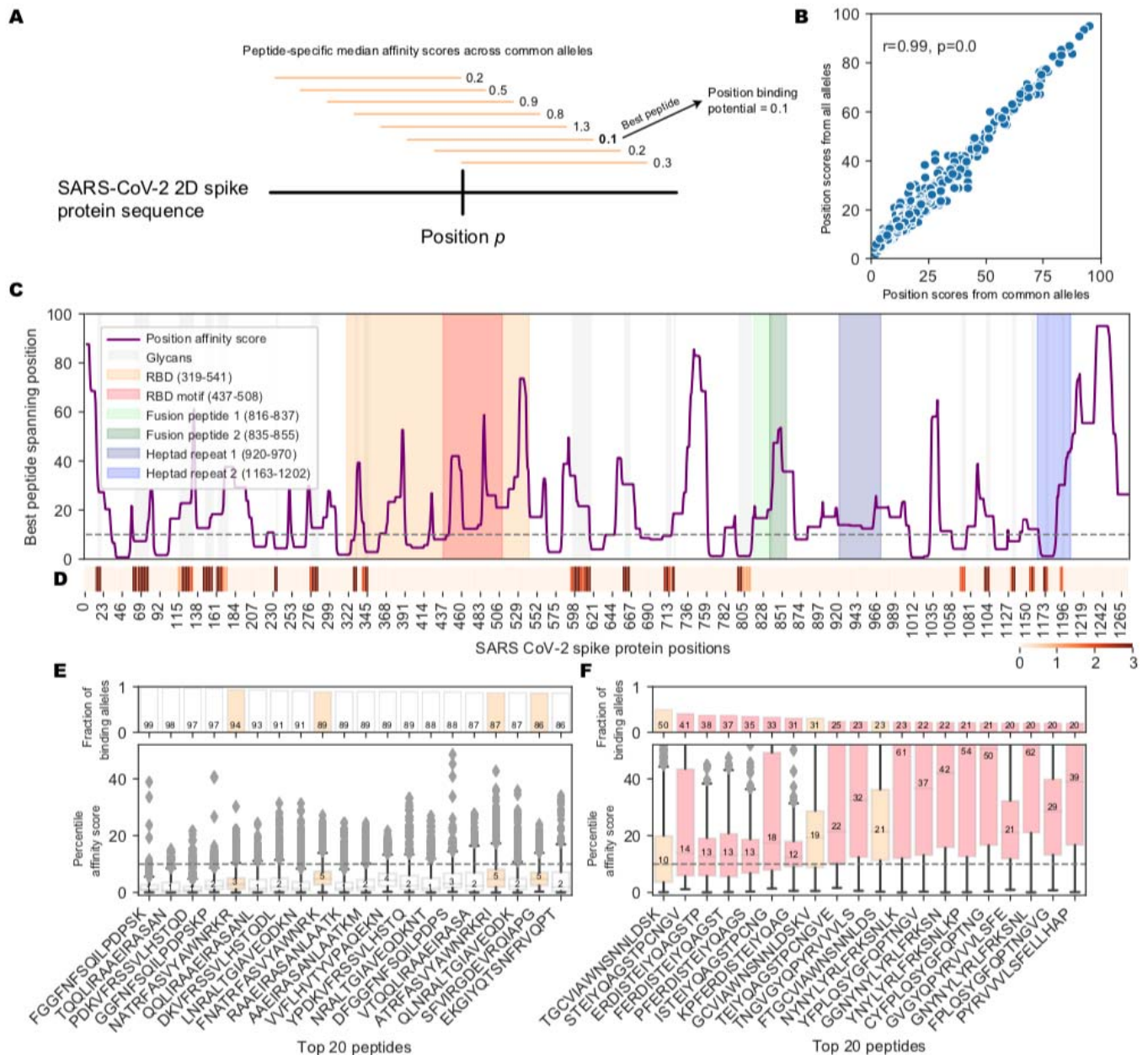
115 **Results**

116 **Topology of a key RBM B cell epitope**

117 Within the 222 amino acid long RBD of the spike protein (residues 319-541), the RBM (residues
118 437-508) is the portion of the spike protein that establishes contact with the ACE2 receptor (**Fig**
119 **1A**). The contact residues span a relatively large surface involving approximately 17 residues
120 (36), among them residues F486, N487, Y489 form a loop, which we term the FNCY patch,
121 which is surface exposed and protrudes up towards the ACE2 receptor from the bulge of the
122 RBD (**Fig 1B-C**). F486 forms hydrophobic interactions with three ACE2 residues (L79, M82,
123 W83). N487 forms hydrogen bonds with Q24 and W83, and Y489 is linked with K31 via a
124 hydrophobic interaction. This makes the amino acid residues in or around the FNCY patch a
125 logical B cell epitope target for antibodies blocking the virus:receptor interaction. In addition,
126 these core residues are mutationally constrained by the ACE2 contact surface (43). Not
127 surprisingly, a set of recently reported potentially neutralizing antibodies generated by single B cell
128 VH/VL cloning from convalescent COVID-19 patients all bear paratopes that include the FNCY
129 patch in their recognition site (34,39–41,47) (**Fig 1D**). While other residues (Q493, N501, and
130 Y505) are also shared between ACE2 and the paratope of these antibodies, they are not as
131 protruding and are on a β -sheet unlike the FNCY patch which is organized in a short loop as a
132 result of the C480:C488 disulfide bond. Thus, blockade of the RBM:ACE2 interaction
133 (neutralization) depends at least in part on a B cell epitope in the RBM that is structurally and
134 functionally critical to the interaction, virus internalization, and cell infectivity.

135

151 MHC-II peptide restriction across the entire SARS-CoV-2 spike protein with respect to common
152 MHC-II alleles in the human population. To assess the potential for effective restriction by
153 MHC-II molecules in a reasonable proportion of the population, we devised a position-based
154 score that assigns each amino acid residue the median affinity of the best overlapping peptide,
155 where median affinity is calculated across the 1911 most common MHC-II alleles (**Fig 2A**),
156 which was highly correlated with scores across all 5620 MHC-II alleles (**Fig 2B**; Pearson
157 $\rho=0.99$, $p<2.2e-308$). While a number of sites along the spike protein are predicted to generate
158 high affinity peptides for most common MHC-II alleles, the region around the FNCY patch was
159 depleted for generally effective binders (**Fig 2C**, Fisher's exact OR=0.21, $p=0.015$, Methods,
160 **Supplemental Fig 1**). Interestingly, the RBM region containing the FNCY patch was free of
161 glycans that could potentially mask the epitope (**Fig 2D**). We further evaluated the distributions
162 of binding affinities for the 20 best-ranked peptides across all sites in the spike protein (**Fig 2E**),
163 and in comparison, the distributions for the best 20 peptides overlapping positions within +/- 50
164 residues of the FNCY patch (**Fig 2F**). In the best case, less than half of the considered MHC-II
165 alleles bound a shared peptide close to the FNCY patch, whereas at other sites there were
166 multiple peptides that could be bound by nearly all of the MHC-II alleles (**Fig 2E**). This
167 suggested overall less availability of effective T cell epitopes in close proximity to the FNCY B
168 cell epitope, which could limit the availability of T cell help during an epitope-specific T-B
169 cooperative interaction in the germinal center.
170



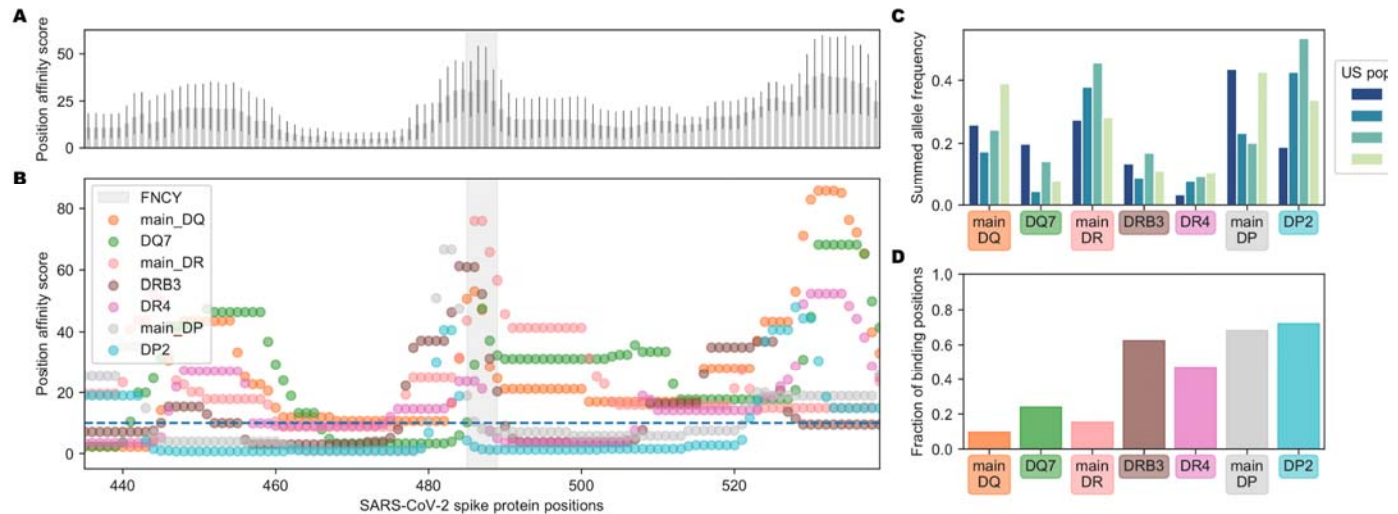
171

172 Figure 2: Landscape of MHC-II binding affinity across spike protein 2D sequence. (A) Overview of the
 173 position affinity score. (B) Scatterplot showing position affinity scores estimated using only common
 174 (>10% frequency, from (48)) MHC-II alleles (x-axis) versus across all MHC-II alleles (y-axis). (C)
 175 Lineplot showing the position affinity scores across common MHC-II alleles (Methods). Annotated
 176 domains from UniProt are highlighted. (D) Heatmap showing amino acid positions that are glycosylated
 177 (49). (E) Barplots (top) and boxplots (bottom) describing the fraction of binding MHC-II alleles and
 178 corresponding affinity percentile rank distributions respectively for the top 20 peptides with the highest
 179 fraction of common binding alleles. The binding threshold of 10 is shown as a dotted line, with values
 180 less than 10 indicating binding. Colors correspond to the regions listed in C. (F) Barplots (top) and

181 boxplots (bottom) describing the fraction of binding MHC-II alleles and corresponding affinity percentile
182 rank distributions respectively for the top 20 peptides within +/-50 amino acids of the FNCY B cell
183 epitope. Colors correspond to the regions listed in C.
184
185

186 To further assess whether population variation in MHC-II MHC alleles might contribute
187 to heterogeneity in potential to generate neutralizing antibodies, we also evaluated the potential
188 of MHC-II supertypes to restrict peptides from neighboring the FNCY patch. Greenbaum *et al.*
189 previously defined 7 supertypes that group MHC-II alleles based on shared binding repertoire.
190 These 7 supertypes account for between 46%-77% of haplotypes and cover over 98% of
191 individuals when all four loci are considered together (50). We revisited our analysis of peptide
192 restriction proximal to the FNCY patch treating each supertype separately. There was
193 considerable variability in potential to effectively present FNCY patch proximal sequences
194 across supertypes (**Fig 3A-B**, $X^2=175$, $p=3.75e-35$, **Supplemental Fig 2**). Only 3 supertypes
195 (DP2, main DP and DR4) commonly presented peptides overlapping the FNCY patch (**Fig 3B**).
196 We were able to obtain population allele frequencies for four populations from the Be The Match
197 registry (51) and Du *et al.* (52). These data show that DR4 is relatively infrequent across the
198 populations evaluated, whereas main DR, main DP, and DP2 are more common (**Fig 3C**), and
199 thus could be more important for MHC-II restriction supportive of neutralizing antibodies. While
200 there were some large population-specific differences in main DP and DP2 supertype
201 frequencies, these frequency estimates are based on a limited population sample and may provide
202 only a rough approximation. In general, DP and DR haplotypes were able to restrict more FNCY
203 patch proximal sequences (**Fig 3D**).

204



205

206 Figure 3: Population variation affecting availability of FNCY proximal T cell epitopes. (A) Barplot
207 showing the aggregated supertype position affinity scores for each position +/- 50 amino acids from the
208 FNCY patch (grey zone). (B) Scatterplot showing the specific supertype position scores for each position
209 +/- 50 amino acids from the FNCY patch (grey zone). The binding threshold of 10 is shown as a dashed
210 blue line, with points below the threshold indicating binding. (C) Barplot showing United States
211 population frequencies, summed across the available alleles in each supertype. (D) Fraction of positions
212 falling below the binding threshold within the region of interest for each supertype.

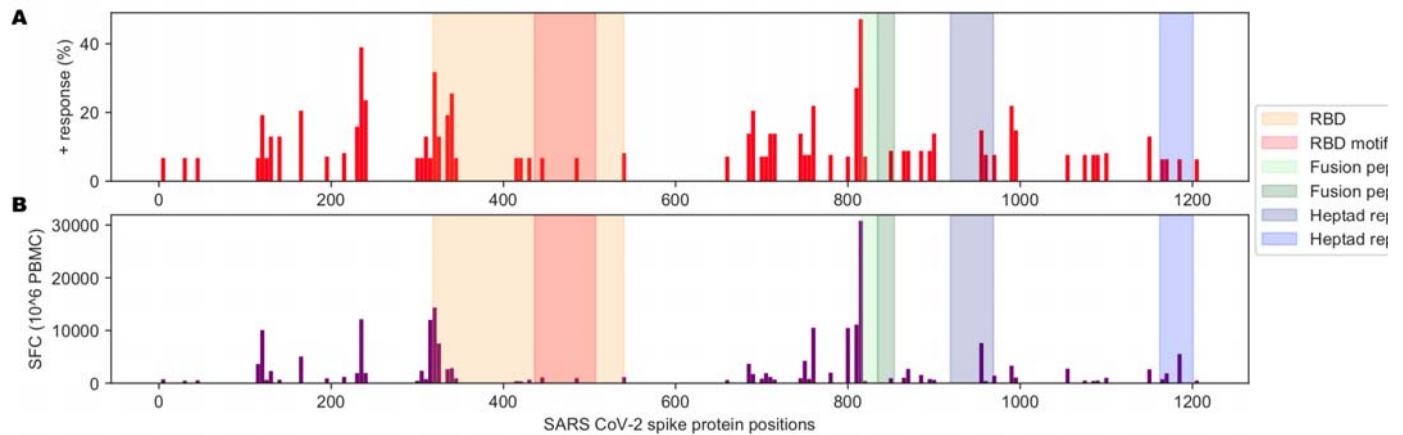
213

214 **Cross-reactivity to a non-coronavirus MHC-II binding peptide as a potential driver of T** 215 **cell responses helping antibody response to the RBM B cell epitope**

216 Interestingly, Mateus *et al.* reported pre-existing CD4 T cell responses to peptides derived from
217 the spike protein using T cells from unexposed individuals, suggesting previous exposures to
218 other human coronaviruses could potentially generate protective immunity toward SARS-CoV-2.
219 Indeed, regions of higher coronavirus homology were associated with more T cell responses in
220 their data (46). This represents the most comprehensive interrogation of the spike protein with
221 response to CD4 T cell responses to date. They screened all 15mers of the spike protein in
222 pooled format and further evaluated 66 predicted MHC-II peptides that generated CD4 T cell
223 responses. Visualizing the landscape of the CD4 T cell responses described in their work by
224 percent positive response (**Fig 4A**) or spot forming cells (**Fig 4B**), we noted relatively few

225 responses proximal to the FNCY patch in the RBM. Accordingly, few other coronaviruses had
226 limited homology to the FNCY region, and none fully included the FNCY patch (**Fig 5A**).

227



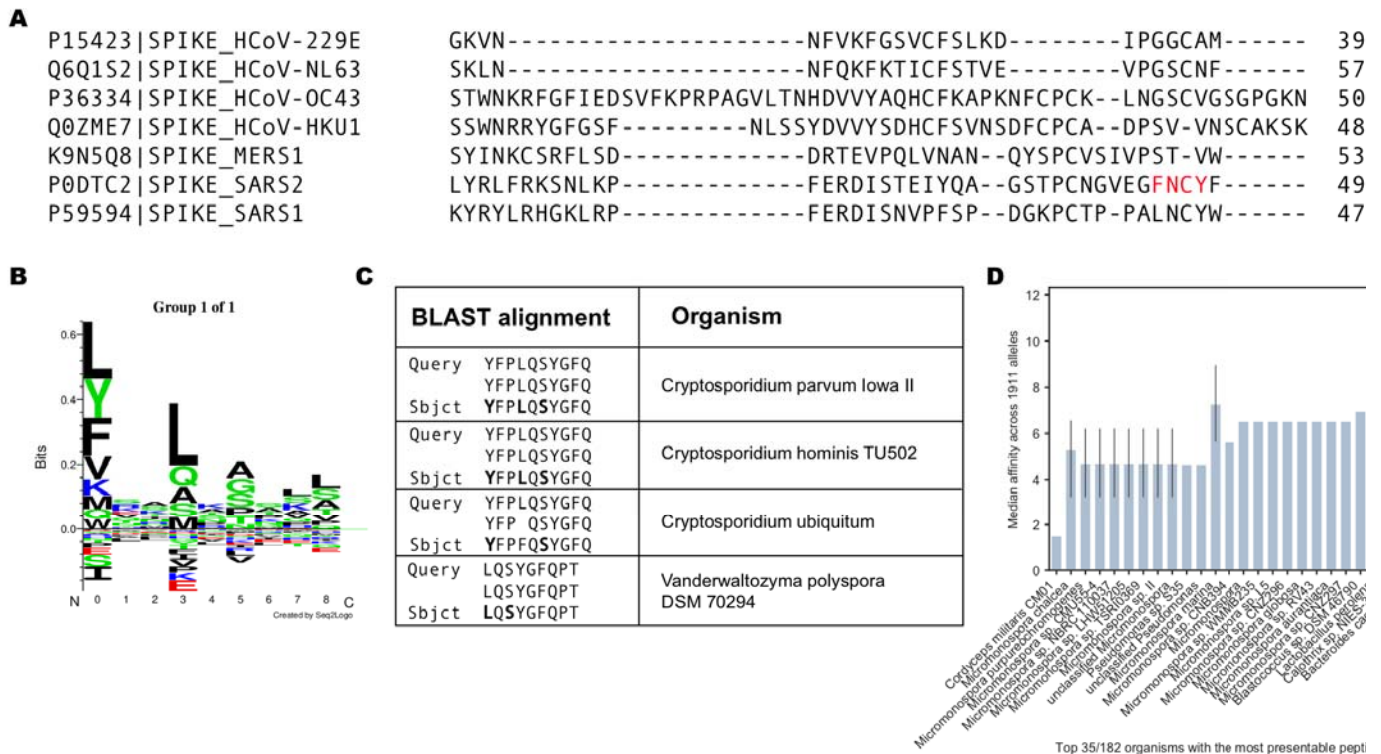
228

229 Figure 4. Immunological history of relevance to SARS-CoV-2. (A) Barplot showing the percentage of
230 positive responses toward SARS-CoV-2 peptides from unexposed individuals. (B) Barplot showing the
231 number of spot-forming cells (SFC) for tested SARS-CoV-2 peptides against PBMCs from unexposed
232 individuals. Data from Table S1 from (46).

233

234 A notable exception in Mateus' results is peptide 486FNCYFPLQSYGFQPT500, which
235 was reported to induce a CD4 T cell response in an unexposed individual. In this case, the
236 peptide was restricted by HLA-DRB1*0101 or HLA-DQA1*0101/DQB1*0501. We found that
237 the peptide sequence had greater *in silico* predicted affinity to HLA-DRB1*0101. To explain the
238 conundrum, we blasted this peptide against the "refseq_protein" database excluding SARS-CoV-
239 2 (Methods). Surprisingly, the sequences with the best homology for this query were not from
240 coronaviruses but rather from common pathogens, first among them parasites of the
241 *Cryptosporidium* genus of apicomplexan parasitic alveolates. These sequences included
242 conserved anchor positions for the HLA-DRB1*0101 allele making it plausible that a prior
243 exposure could account for the formation of a memory CD4 T cell response (**Fig 5B-C**). To
244 further assess the potential for other prior exposures in generating immune memory for

245 sequences proximal to the FNCY patch we blasted all 15mers within +/-30 amino acids of the
 246 FNCY patch and filtered the resulting sequences based on restriction by consensus MHC-II
 247 supertypes (50) (**Supplemental Table 2**). We found peptides associated with multiple microbial
 248 organisms that may meet the criteria to potentially generate CD4 T cell memory relevant to the
 249 RBM of SARS-CoV-2 (**Fig 5D**).



251
 252 Figure 5. Learned immunity to other targets that could support T cell responses to SARS-CoV-2. (A)
 253 Multiple sequence alignment between SARS-CoV-2, SARS1, MERS, and other human coronaviruses,
 254 focusing on the region surrounding the FNCY B cell epitope. (B) SeqLogo plot obtained by clustering
 255 IEDB peptides reported to bind to DRB1*01:01. (C) Top results after blasting the FNCYFPLQSYGFQPT
 256 peptide against all reference proteins. (D) Barplot describing best peptide affinities across MHC-II alleles
 257 of the top 35 unique organisms with one or more peptides matching a peptide with high similarity to
 258 15mers +/-30aa from the FNCY binding epitope based on BLAST analysis. The closer to 0, the greater
 259 the binding potential.

261 **Discussion**

262 SARS-CoV-2 uses the RBD of the spike protein to bind to the ACE2 receptor on target cells.
263 The actual contact with ACE2 is mediated by a discrete number of amino acids that have been
264 visualized by cryo-EM (Lan et al., 2020; Shang et al., 2020). Although several SARS-related
265 coronaviruses share 75% homology and interact with ACE2 on target cells (Ge et al., 2013; Ren
266 et al., 2008; Yang et al., 2015) the RBM in SARS-CoV-2 is unique to this virus. *In vitro* binding
267 measurements show that SARS-CoV-2 RBD binds to ACE2 with an affinity in the low
268 nanomolar range (Walls et al., 2020). Mutations in this motif could be detrimental to the virus's
269 ability to infect ACE2 positive human cells. Since the RBD is an immunodominant site in the
270 antibody response in humans (42) it is not surprising that the paratope of some antibodies
271 isolated from convalescent individuals via single B cell VH/VL cloning, and selected on the
272 basis of high neutralization potency, all seem to bind a surface encompassing the FNCY patch in
273 the RBM (35,37–41,53). Arguably, this motif corresponds to a relevant B cell epitope in the
274 spike protein of SARS-CoV-2 and is a logical target of potent neutralizing antibodies.

275 Although antibodies directed to this site have been isolated by different groups, little is
276 known about their contribution to the pool of antibodies in serum of SARS-CoV-2 infected
277 individuals, but evidence suggests they are likely to be rare. In one study they were found to
278 represent a subdominant fraction of the anti-RBD response (41) while the estimated frequency of
279 antigen-specific B cells ranges from 0.07 to 0.005% of all the total B cells in COVID-19
280 convalescent individuals (54). In a second study, the identification of two ultra-potent NAbs
281 having a paratope involving the FNCY patch required screening of 800 clones from twelve
282 individuals (53). This suggests that a potent NAb response to a mutationally constrained RBM
283 epitope is a rare component of the total anti-virus response consistent, with the observation that
284 there is no correlation between RBM site-specific neutralizing antibodies and serum half-

285 maximal neutralization titer (NT50) (54). Here we show that the core RBM B cell epitope is
286 apparently uncoupled from preferential T-B pairing, a prerequisite for a coordinated activation of
287 B cells against the pathogen. We analyzed MHC-II binding of 15mer peptides in the spike
288 protein upstream (-50 aa) or downstream (+50 aa) of the central RBM B cell epitope and found
289 both low coverage by 1911 common MHC-II alleles and a depletion of binding 15mers proximal
290 to the FNCY patch versus other exposed areas on the spike protein. This could be due to the fact
291 that a sizeable proportion (40%) of CD4 T cells responding to the spike protein are memory
292 responses found in SARS-CoV-2 unexposed individuals (44,55) or other structural protein of
293 SARS-CoV-2 such as the N protein (45). Thus, it is possible that these conserved responses are
294 used as a decoy mechanism to polarize the response away from the RBM. However, this does not
295 rule out the contribution of a bias in frequency of specific B cells in the available repertoire.

296 Corroboration to our hypothesis also comes from Mateus *et al.* (46) who tested sixty-six
297 15mer peptides of the spike protein in SARS-CoV-2 unexposed individuals and found that CD4
298 T cell responses against this narrow RBM site account for only 2/110 (1.8%) of the total CD4 T
299 cell response to 15mer peptides of the spike protein. Surprisingly, a CD4 T cell response against
300 peptide FNCYFPLQSYGFQPT was by CD4 T cells of an unexposed individual. Since this
301 peptide has low homology with previous human coronaviruses, we reasoned that this could either
302 represent a case of TCR cross-reactivity since a single TCR can engage large numbers of unique
303 MHC/peptide combinations without requiring degeneracy in their recognition (56,57).
304 Remarkably, however, a BLAST analysis revealed a 10 amino acid sequence match with
305 proteins from pathogens including those from the *Cryptosporidium* genus, with identity in
306 binding motif and anchor residues (agretope) for the restricting MHC-II allele strongly
307 suggesting peptide cross-reactivity. *Cryptosporidium hominis* is a parasite that causes watery

308 diarrhea that can last up to 3 weeks in immunocompetent patients (58). Additional possibilities
309 for cross-reactivity to the RBM, albeit of a lesser stringency, involve antigens from
310 *Micromonospora*, *Pseudomonas*, *Blastococcus*, *Lactobacillus*, and *Bacteroides* (**Fig 5D**). Thus,
311 it appears as if memory CD4 T cells reactive with peptides in the RBM may reflect the
312 immunological history of the individual that, as evidenced by this case, can be unrelated to
313 infection by other coronaviruses. Interestingly, the great majority (64-88%) of COVID-19
314 positive individuals in homeless shelters in Los Angeles and Boston were found to be
315 asymptomatic (59). This suggests that the status of the immune system, which itself reflects past
316 antigenic exposure, may be a determining factor in the generation of a protective immune
317 response after SARS-CoV-2 infection.

318 The findings reported herein have considerable implications for natural immunity to
319 SARS-CoV-2. The fact that there seems to be an overall suboptimal T-B preferential pairing
320 suggests that B cells that respond to the RBM B cell epitope may receive inadequate T cell help.
321 This is consistent with the observation that in general potent neutralizing antibodies to the RBM
322 undergo very limited somatic mutation (38,53) and are by and large in quasi-germline
323 configuration (60). Since T cell help is also necessary to initiate somatic hypermutation in B cell
324 through CD40 or CD38 signaling in the germinal center (61), it follows that one important
325 implication of our study is that defective T-B pairing may negatively influence the normal
326 process of germinal center maturation of the B cell response in response to SARS-CoV-2
327 infection in a critical way.

328 Which antigens can generate T cell responses depends on the binding specificities of
329 MHC-II molecules, which are highly polymorphic in the human population. We noted a general
330 trend for MHC-II alleles to less effectively present peptides from the RBM region, but also

331 observed some variability across MHC-II supertypes. The main DP and DP2 haplotypes were
332 both common and had the highest potential to present peptides, suggesting that most individuals
333 should carry at least one allele capable of presenting peptides in this region. Which of the two
334 DP haplotypes was more common varied by ancestral population, thus it is possible that
335 differences in the haplotypes could translate to differences in T-B cooperativity levels within
336 groups, though binding affinities for epitopes near the FNCY patch were similar for both. DQ
337 and DR supertypes were less able to present peptides near FNCY, with the exception of DR4,
338 which is among the less common supertypes. Importantly, our analysis was limited to predicted
339 affinity of peptides to MHC-II, and other characteristics such as expression levels, stability or
340 differences in interactions with molecular chaperones likely also contribute to whether FNCY
341 proximal peptides are available to support B-T cooperation (62).

342 In light of our findings, it can be predicted that, in general, a specific RBM antibody
343 response may be short-lived and that residual immunity from a primary infection may not be
344 sufficient to prevent reinfection after 6-9 months. Sporadic cases of re-infection have been
345 reported by the media in Hong Kong and Nevada (63). A third case has been reported in a care-
346 home resident who after the second infection produced only low levels of antibodies (64).
347 Finally, silent re-infections in young workers in a COVID-19 ward who tested positive for the
348 new coronavirus and became reinfected several months later with no symptoms in either instance
349 have been reported (65). It is tempting to speculate that waning antibody levels or a poorly
350 developed specific NAb antibody response to SARS-CoV-2 can potentially put people at risk of
351 reinfection. Other factors to consider are a bias in the available B cell repertoire in the population
352 and the extent to which a defective T-B cooperation influences the longevity of terminally
353 differentiated plasma cells in the bone marrow (66).

354 In summary, we provide evidence that MHC-II constrains the CD4 T cell response for
355 epitopes that are best positioned to facilitate T-B pairing in generating and sustaining a potent
356 neutralizing antibody response against a mutationally constrained RBM B cell epitope.
357 Furthermore, we show that the immunological history of the individual, not necessarily related to
358 infection by other coronaviruses, may confer immunologic advantage. Finally, these findings
359 may have implications for the quality and persistence of a protective, neutralizing antibody
360 response to RBM induced by current SARS-CoV-2 vaccines.

361

362 **Materials and Methods**

363 Data and code are available at https://github.com/cartercompbio/SARS_CoV_2_T-B_co-op.

364

365 *Affinity analysis*

366 NetMHCIIpan version 4.0 was used to predict peptide-MHC-II affinity (69) for generated
367 15mers along the SARS-CoV-2 spike protein.

368

369 *Spike protein analyses*

370 SARS-CoV-2 spike protein sequence and protein regions were obtained from
371 <https://www.uniprot.org/uniprot/P0DTC2>. Glycan data were obtained from (49) and true-positive
372 sites were aggregated across 3 replicates. To assess depletion of effective binders near the FNCY
373 patch, we performed a Fisher's exact test for binding (median affinity across common alleles
374 <10) versus proximity (+/- 50 amino acids) to FNCY for positions free of glycans. We excluded
375 positions within 10 amino acids of a glycan using the data obtained from Watanabe *et al.* and
376 added a pseudocount of 1.

377

378 The SARS1, MERS1, HCoV-229E, HCoV-NL63, HCoV-OC43, and HCoV-HKU1 spike protein
379 sequences were also downloaded from UniProt (P59594, K9N5Q8, P15423, Q6Q1S2, P36334,
380 Q0ZME7, respectively). Multiple sequence alignment was performed on the EMBL-EBI Clustal
381 Omega web server using default parameters (70).

382

383 *Structure analysis*

384 The 6M0J 3D X-ray structure for the protein complex containing the SARS-CoV-2 spike protein
385 RBD (P0DTC2) interaction with ACE2 (Q9BYF1) from (36). The structure figures were
386 prepared using VMD (71).

387

388 *Supertype analysis*

389 Supertypes were obtained from (50). All alpha/beta combinations spanning any of these types
390 were included, resulting in 279 alleles. US supertype frequencies for alleles in DRB1 and DQB1
391 were obtained from the Be the Match registry (51), US frequencies for alleles in DPB1 were
392 obtained from (52) as DPB1 was not available from the Be the Match registry. Available allele
393 frequencies within each supertype were summed for Fig 3C.

394

395 *Motif analysis*

396 All 13-20mer peptides adhering to the following parameters were downloaded from the IEDB
397 (72): MHC-II assay, positive only, DRB1*01:01 allele, linear peptides; and any peptides with
398 post-translational modifications or noncanonical amino acids were removed. The remaining
399 10,117 peptides were input into Gibbs cluster v2.0 (73) using the default MHC-II ligand
400 parameters.

401

402 *BLAST analysis*

403 15mers were generated along a sliding window +/-30 amino acids from the FNCY patch start
404 and end (455-518, 0-index) and input into NCBI BLAST (74) using the ‘refseq_protein’ database
405 and excluding SARS-CoV-2 (taxid:2697049). Identified peptides (**Supplemental table 2**) were
406 then evaluated for binding affinity and any peptide binding to at least one allele was retained for
407 Fig 5D.

408

409 **Acknowledgements**

410 This work was supported by an NIH National Library of Medicine Training Grant
411 T15LM011271 to A.C., an Emerging Leader Award from The Mark Foundation for Cancer
412 Research, grant #18-022-ELA and a CIFAR fellowship to H.C. and RO1 CA220009 to M.Z. and
413 H.C. The graphical abstract was created using BioRender and used the PDB (67) structure 6VXX
414 from (68).

415

416 **Author Contributions**

417 Original concept, M.Z.; project supervision, H.C. and M.Z.; project planning and experimental
418 design, A.C., M.Z., and H.C.; data acquisition, processing, and analysis, A.C. and K.O.;
419 preparation of paper, A.C., M.Z., and H.C.

420

421 **Declaration of Interests**

422 The authors declare no competing interests.

423

424 **References**

- 425 1. Wölfel R, Corman VM, Guggemos W, Seilmaier M, Zange S, Müller MA, et al. Virological
426 assessment of hospitalized patients with COVID-2019. Nature. 2020 May;581(7809):465–
427 9.

- 428 2. Long Q-X, Tang X-J, Shi Q-L, Li Q, Deng H-J, Yuan J, et al. Clinical and immunological
429 assessment of asymptomatic SARS-CoV-2 infections. *Nat Med*. 2020 Aug;26(8):1200–4.
- 430 3. Rydyznski Moderbacher C, Ramirez SI, Dan JM, Grifoni A, Hastie KM, Weiskopf D, et al.
431 Antigen-Specific Adaptive Immunity to SARS-CoV-2 in Acute COVID-19 and
432 Associations with Age and Disease Severity. *Cell* [Internet]. 2020 Sep 16; Available from:
433 <http://dx.doi.org/10.1016/j.cell.2020.09.038>
- 434 4. Prévost J, Gasser R, Beaudoin-Bussièrès G, Richard J, Duerr R, Laumaea A, et al. Cross-
435 sectional evaluation of humoral responses against SARS-CoV-2 Spike. *Cell Rep Med*. 2020
436 Sep 30;100126.
- 437 5. Mitchison NA. T-cell-B-cell cooperation. *Nat Rev Immunol*. 2004 Apr;4(4):308–12.
- 438 6. Jacob J, Kelsoe G, Rajewsky K, Weiss U. Intracloonal generation of antibody mutants in
439 germinal centres. *Nature*. 1991 Dec 5;354(6352):389–92.
- 440 7. Berek C, Berger A, Apel M. Maturation of the immune response in germinal centers. *Cell*.
441 1991 Dec 20;67(6):1121–9.
- 442 8. Zanetti M, Glotz D. Considerations on thymus-dependent and -independent antigens in
443 acquired and natural immunity. *Ann Inst Pasteur Immunol*. 1988 Mar;139(2):192–3.
- 444 9. Claman HN, Chaperon EA, Triplett RF. Thymus-marrow cell combinations. Synergism in
445 antibody production. *Proc Soc Exp Biol Med*. 1966 Aug;122(4):1167–71.
- 446 10. Mitchison NA. The carrier effect in the secondary response to hapten-protein conjugates. I.
447 Measurement of the effect with transferred cells and objections to the local environment
448 hypothesis. *Eur J Immunol*. 1971 Jan;1(1):10–7.
- 449 11. Rajewsky K, Rottländer E, Peltre G, Müller B. The immune response to a hybrid protein
450 molecule; specificity of secondary stimulation and of tolerance induction. *J Exp Med*. 1967
451 Oct 1;126(4):581–606.
- 452 12. Katz DH, Hamaoka T, Dorf ME, Benacerraf B. Cell interactions between histoincompatible
453 T and B lymphocytes. The H-2 gene complex determines successful physiologic
454 lymphocyte interactions. *Proc Natl Acad Sci U S A*. 1973 Sep;70(9):2624–8.
- 455 13. Sprent J. Restricted helper function of F1 hybrid T cells positively selected to heterologous
456 erythrocytes in irradiated parental strain mice. II. Evidence for restrictions affecting helper
457 cell induction and T-B collaboration, both mapping to the K-end of the H-2 complex. *J Exp*
458 *Med*. 1978 Apr 1;147(4):1159–74.
- 459 14. Jones B, Janeway CA Jr. Cooperative interaction of B lymphocytes with antigen-specific
460 helper T lymphocytes is MHC restricted. *Nature*. 1981 Aug 6;292(5823):547–9.
- 461 15. Mitchison NA. The carrier effect in the secondary response to hapten-protein conjugates. II.
462 Cellular cooperation. *Eur J Immunol*. 1971;1(1):18–27.

- 463 16. Janeway CA Jr. Cellular cooperation during in vivo anti-hapten antibody responses. I. The
464 effect of cell number on the response. *J Immunol.* 1975 Apr;114(4):1394–401.
- 465 17. Shulman Z, Gitlin AD, Targ S, Jankovic M, Pasqual G, Nussenzweig MC, et al. T follicular
466 helper cell dynamics in germinal centers. *Science.* 2013 Aug 9;341(6146):673–7.
- 467 18. Celada F, Sercarz EE. Preferential pairing of T-B specificities in the same antigen: the
468 concept of directional help. *Vaccine.* 1988 Apr;6(2):94–8.
- 469 19. Manca F, Kunkl A, Fenoglio D, Fowler A, Sercarz E, Celada F. Constraints in T-B
470 cooperation related to epitope topology on *E. coli* β -galactosidase. I. The fine specificity of
471 T cells dictates the fine specificity of antibodies directed to conformation-dependent
472 determinants. *Eur J Immunol.* 1985;15(4):345–50.
- 473 20. Bretscher P, Cohn M. A theory of self-nonsel self discrimination. *Science.* 1970 Sep
474 11;169(3950):1042–9.
- 475 21. Lanzavecchia A. Antigen-specific interaction between T and B cells [Internet]. Vol. 314,
476 *Nature.* 1985. p. 537–9. Available from: <http://dx.doi.org/10.1038/314537a0>
- 477 22. Kroeger DR, Rudulier CD, Bretscher PA. Antigen presenting B cells facilitate CD4 T cell
478 cooperation resulting in enhanced generation of effector and memory CD4 T cells. *PLoS*
479 *One.* 2013 Oct 14;8(10):e77346.
- 480 23. Cassell D, Forman J. Linked recognition of helper and cytotoxic antigenic determinants for
481 the generation of cytotoxic T lymphocytes. *Ann N Y Acad Sci.* 1988;532:51–60.
- 482 24. Gerloni M, Xiong S, Mukerjee S, Schoenberger SP, Croft M, Zanetti M. Functional
483 cooperation between T helper cell determinants. *Proc Natl Acad Sci U S A.* 2000 Nov
484 21;97(24):13269–74.
- 485 25. Berzofsky JA, Richman LK, Killion DJ. Distinct H-2-linked Ir genes control both antibody
486 and T cell responses to different determinants on the same antigen, myoglobin. *Proc Natl*
487 *Acad Sci U S A.* 1979 Aug;76(8):4046–50.
- 488 26. Berzofsky JA, Schechter AN, Shearer GM, Sachs DH. Genetic control of the immune
489 response to staphylococcal nuclease. III. Time-course and correlation between the response
490 to native nuclease and the response to its polypeptide fragments. *J Exp Med.* 1977 Jan
491 1;145(1):111–22.
- 492 27. Berzofsky JA, Schechter AN, Shearer GM, Sachs DH. Genetic control of the immune
493 response to staphylococcal nuclease. IV. H-2-linked control of the relative proportions of
494 antibodies produced to different determinants of native nuclease. *J Exp Med.* 1977 Jan
495 1;145(1):123–35.
- 496 28. Zanetti M, Sercarz E, Salk J. The immunology of new generation vaccines. *Immunol*
497 *Today.* 1987;8(1):18–25.

- 498 29. Celada F, Kunkl A, Manca F, Fenoglio D, Fowler A, Krzych U, et al. Preferential pairings
499 in T-B encounters utilizing Th cells directed against discrete portions of b-galactosidase and
500 B cells primed with the native enzyme or a hapten epitope. *Regulation of the Immune*
501 *System*. 1984;637–46.
- 502 30. Sette A, Moutaftsi M, Moyron-Quiroz J, McCausland MM, Davies DH, Johnston RJ, et al.
503 Selective CD4+ T cell help for antibody responses to a large viral pathogen: deterministic
504 linkage of specificities. *Immunity*. 2008 Jun;28(6):847–58.
- 505 31. Lv Z, Deng Y-Q, Ye Q, Cao L, Sun C-Y, Fan C, et al. Structural basis for neutralization of
506 SARS-CoV-2 and SARS-CoV by a potent therapeutic antibody. *Science*. 2020 Sep
507 18;369(6510):1505–9.
- 508 32. Pinto D, Park Y-J, Beltramello M, Walls AC, Tortorici MA, Bianchi S, et al. Cross-
509 neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature*. 2020
510 Jul;583(7815):290–5.
- 511 33. Yuan M, Wu NC, Zhu X, Lee C-CD, So RTY, Lv H, et al. A highly conserved cryptic
512 epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *Science*. 2020
513 May 8;368(6491):630–3.
- 514 34. Piccoli L, Park Y-J, Tortorici MA, Czudnochowski N, Walls AC, Beltramello M, et al.
515 Mapping Neutralizing and Immunodominant Sites on the SARS-CoV-2 Spike Receptor-
516 Binding Domain by Structure-Guided High-Resolution Serology. *Cell* [Internet]. 2020 Sep
517 16; Available from: <http://dx.doi.org/10.1016/j.cell.2020.09.037>
- 518 35. Barnes CO, West AP Jr, Huey-Tubman KE, Hoffmann MAG, Sharaf NG, Hoffman PR, et
519 al. Structures of Human Antibodies Bound to SARS-CoV-2 Spike Reveal Common
520 Epitopes and Recurrent Features of Antibodies. *Cell*. 2020 Aug 20;182(4):828–42.e16.
- 521 36. Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, et al. Structure of the SARS-CoV-2 spike
522 receptor-binding domain bound to the ACE2 receptor. *Nature*. 2020 May;581(7807):215–
523 20.
- 524 37. Liu L, Wang P, Nair MS, Yu J, Rapp M, Wang Q, et al. Potent neutralizing antibodies
525 against multiple epitopes on SARS-CoV-2 spike. *Nature*. 2020 Aug;584(7821):450–6.
- 526 38. Rogers TF, Zhao F, Huang D, Beutler N, Burns A, He W-T, et al. Isolation of potent SARS-
527 CoV-2 neutralizing antibodies and protection from disease in a small animal model.
528 *Science*. 2020 Aug 21;369(6506):956–63.
- 529 39. Shi R, Shan C, Duan X, Chen Z, Liu P, Song J, et al. A human neutralizing antibody targets
530 the receptor-binding site of SARS-CoV-2. *Nature*. 2020 Aug;584(7819):120–4.
- 531 40. Wu Y, Wang F, Shen C, Peng W, Li D, Zhao C, et al. A noncompeting pair of human
532 neutralizing antibodies block COVID-19 virus binding to its receptor ACE2. *Science*. 2020
533 Jun 12;368(6496):1274–8.

- 534 41. Zost SJ, Gilchuk P, Case JB, Binshtein E, Chen RE, Nkolola JP, et al. Potently neutralizing
535 and protective human antibodies against SARS-CoV-2. *Nature*. 2020 Aug;584(7821):443–
536 9.
- 537 42. Premkumar L, Segovia-Chumbez B, Jadi R, Martinez DR, Raut R, Markmann A, et al. The
538 receptor binding domain of the viral spike protein is an immunodominant and highly
539 specific target of antibodies in SARS-CoV-2 patients. *Sci Immunol* [Internet]. 2020 Jun
540 11;5(48). Available from: <http://dx.doi.org/10.1126/sciimmunol.abc8413>
- 541 43. Starr TN, Greaney AJ, Hilton SK, Ellis D, Crawford KHD, Dings AS, et al. Deep
542 Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on
543 Folding and ACE2 Binding. *Cell*. 2020 Sep 3;182(5):1295–310.e20.
- 544 44. Grifoni A, Weiskopf D, Ramirez SI, Mateus J, Dan JM, Moderbacher CR, et al. Targets of
545 T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and
546 Unexposed Individuals. *Cell*. 2020 Jun 25;181(7):1489–501.e15.
- 547 45. Le Bert N, Tan AT, Kunasegaran K, Tham CYL, Hafezi M, Chia A, et al. SARS-CoV-2-
548 specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. *Nature*.
549 2020 Aug;584(7821):457–62.
- 550 46. Mateus J, Grifoni A, Tarke A, Sidney J, Ramirez SI, Dan JM, et al. Selective and cross-
551 reactive SARS-CoV-2 T cell epitopes in unexposed humans. *Science* [Internet]. 2020 Aug
552 4; Available from: <http://dx.doi.org/10.1126/science.abd3871>
- 553 47. Yuan M, Liu H, Wu NC, Lee C-CD, Zhu X, Zhao F, et al. Structural basis of a shared
554 antibody response to SARS-CoV-2. *Science*. 2020 Aug 28;369(6507):1119–23.
- 555 48. Dosset M, Castro A, Carter H, Zanetti M. Telomerase and CD4 T Cell Immunity in Cancer.
556 *Cancers* [Internet]. 2020 Jun 25;12(6). Available from:
557 <http://dx.doi.org/10.3390/cancers12061687>
- 558 49. Watanabe Y, Allen JD, Wrapp D, McLellan JS, Crispin M. Site-specific glycan analysis of
559 the SARS-CoV-2 spike. *Science*. 2020 Jul 17;369(6501):330–3.
- 560 50. Greenbaum J, Sidney J, Chung J, Brander C, Peters B, Sette A. Functional classification of
561 class II human leukocyte antigen (HLA) molecules reveals seven different supertypes and a
562 surprising degree of repertoire sharing across supertypes. *Immunogenetics*. 2011
563 Jun;63(6):325–35.
- 564 51. Maiers M, Gragert L, Klitz W. High-resolution HLA alleles and haplotypes in the United
565 States population. *Hum Immunol*. 2007 Sep;68(9):779–88.
- 566 52. Du Z. HLA-DPA1 and HLA-DPB1 Frequencies in the US Populations [Internet]. 2017
567 American Transplant Congress; 2017 Apr 30 [cited 2020 Sep 30]; Chicago, IL. Available
568 from: <https://atcmeetingabstracts.com/abstract/hla-dpa1-and-hla-dpb1-frequencies-in-the-us-populations/>
569

- 570 53. Tortorici MA, Beltramello M, Lempp FA, Pinto D, Dang HV, Rosen LE, et al. Ultrapotent
571 human antibodies protect against SARS-CoV-2 challenge via multiple mechanisms. *Science*
572 [Internet]. 2020 Sep 24; Available from: <http://dx.doi.org/10.1126/science.abe3354>
- 573 54. Robbiani DF, Gaebler C, Muecksch F, Lorenzi JCC, Wang Z, Cho A, et al. Convergent
574 antibody responses to SARS-CoV-2 in convalescent individuals. *Nature*. 2020
575 Aug;584(7821):437–42.
- 576 55. Braun J, Loyal L, Frentsch M, Wendisch D, Georg P, Kurth F, et al. SARS-CoV-2-reactive
577 T cells in healthy donors and patients with COVID-19. *Nature* [Internet]. 2020 Jul 29;
578 Available from: <http://dx.doi.org/10.1038/s41586-020-2598-9>
- 579 56. Birnbaum ME, Mendoza JL, Sethi DK, Dong S, Glanville J, Dobbins J, et al.
580 Deconstructing the peptide-MHC specificity of T cell recognition. *Cell*. 2014 May
581 22;157(5):1073–87.
- 582 57. Selin LK, Cornberg M, Brehm MA, Kim S-K, Calcagno C, Ghersi D, et al. CD8 memory T
583 cells: cross-reactivity and heterologous immunity. *Semin Immunol*. 2004 Oct;16(5):335–47.
- 584 58. Gharpure R, Perez A, Miller AD, Wikswo ME, Silver R, Hlavsa MC. Cryptosporidiosis
585 Outbreaks - United States, 2009-2017. *MMWR Morb Mortal Wkly Rep*. 2019 Jun
586 28;68(25):568–72.
- 587 59. Oran DP, Topol EJ. Prevalence of Asymptomatic SARS-CoV-2 Infection : A Narrative
588 Review. *Ann Intern Med*. 2020 Sep 1;173(5):362–7.
- 589 60. Kreer C, Zehner M, Weber T, Ercanoglu MS, Gieselmann L, Rohde C, et al. Longitudinal
590 Isolation of Potent Near-Germline SARS-CoV-2-Neutralizing Antibodies from COVID-19
591 Patients. *Cell*. 2020 Sep 17;182(6):1663–73.
- 592 61. Bergthorsdottir S, Gallagher A, Jainandunsing S, Cockayne D, Sutton J, Leanderson T, et
593 al. Signals that initiate somatic hypermutation of B cells in vitro. *J Immunol*. 2001 Feb
594 15;166(4):2228–34.
- 595 62. Anczurowski M, Hirano N. Mechanisms of HLA-DP Antigen Processing and Presentation
596 Revisited. *Trends Immunol*. 2018 Dec;39(12):960–4.
- 597 63. Tillett RL, Sevinsky JR, Hartley PD, Kerwin H, Crawford N, Gorzalski A, et al. Genomic
598 evidence for reinfection with SARS-CoV-2: a case study. *Lancet Infect Dis* [Internet]. 2020
599 Oct 12; Available from:
600 <http://www.sciencedirect.com/science/article/pii/S1473309920307647>
- 601 64. Goldman JD, Wang K, Roltgen K, Nielsen SCA, Roach JC, Naccache SN, et al. Reinfection
602 with SARS-CoV-2 and Failure of Humoral Immunity: a case report. *medRxiv* [Internet].
603 2020 Sep 25; Available from: <http://dx.doi.org/10.1101/2020.09.22.20192443>
- 604 65. Gupta V, Bhojar RC, Jain A, Srivastava S, Upadhayay R, Imran M, et al. Asymptomatic
605 reinfection in two healthcare workers from India with genetically distinct SARS-CoV-2.

- 606 Clin Infect Dis [Internet]. 2020 Sep 23; Available from:
607 <http://dx.doi.org/10.1093/cid/ciaa1451>
- 608 66. Slifka MK, Matloubian M, Ahmed R. Bone marrow is a major site of long-term antibody
609 production after acute viral infection. *J Virol*. 1995 Mar;69(3):1895–902.
- 610 67. Sehnal D, Rose AS, Koča J, Burley SK, Velankar S. Mol*: towards a common library and
611 tools for web molecular graphics. In: *MolVa: Workshop on Molecular Graphics and Visual*
612 *Analysis of Molecular Data*, Brno, Czech Republic Eurographics. 2018.
- 613 68. Walls AC, Park Y-J, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, Function,
614 and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell*. 2020 Apr 16;181(2):281–
615 92.e6.
- 616 69. Reynisson B, Alvarez B, Paul S, Peters B, Nielsen M. NetMHCpan-4.1 and NetMHCIIpan-
617 4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution
618 and integration of MS MHC eluted ligand data. *Nucleic Acids Res* [Internet]. 2020;
619 Available from: [https://academic.oup.com/nar/advance-article-](https://academic.oup.com/nar/advance-article-abstract/doi/10.1093/nar/gkaa379/5837056)
620 [abstract/doi/10.1093/nar/gkaa379/5837056](https://academic.oup.com/nar/advance-article-abstract/doi/10.1093/nar/gkaa379/5837056)
- 621 70. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, et al. The EMBL-EBI
622 search and sequence analysis tools APIs in 2019. *Nucleic Acids Res*. 2019 Jul
623 2;47(W1):W636–41.
- 624 71. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph*. 1996
625 Feb;14(1):33–8, 27–8.
- 626 72. Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, et al. The Immune
627 Epitope Database (IEDB): 2018 update. *Nucleic Acids Res*. 2019 Jan 8;47(D1):D339–43.
- 628 73. Andreatta M, Lund O, Nielsen M. Simultaneous alignment and clustering of peptide data
629 using a Gibbs sampling approach. *Bioinformatics*. 2013 Jan 1;29(1):8–14.
- 630 74. Sayers EW, Barrett T, Benson DA, Bolton E, Bryant SH, Canese K, et al. Database
631 resources of the National Center for Biotechnology Information. *Nucleic Acids Res*. 2011
632 Jan;39(Database issue):D38–51.

633
634
635

636 **Figure Titles and Legends**

637

638 Figure 1: Visualization of the FNCY core of the RBM B cell epitope on the SARS-CoV-2 spike
639 protein RBD. (A) 3D structure of the SARS-CoV-2 spike protein RBD (white) binding the
640 ACE2 receptor (green) (PDB: 6M0J) with contact residues highlighted in blue and the FNCY
641 patch highlighted in red. (B-C) Spike protein RBD with ACE2 contact residues and FNCY patch

642 residues labeled in two orientations (front and back). (D) Heatmap of neutralizing antibody
643 contact residues (purple) on the spike protein RBM region (positions 437-508). Black dots
644 indicate ACE2 contact residues and the FNCY patch is highlighted in red. Source data available
645 in Supplemental Table 1.

646
647 Figure 2: Landscape of MHC-II binding affinity across spike protein 2D sequence. (A) Overview
648 of the position affinity score. (B) Scatterplot showing position affinity scores estimated using
649 only common (>10% frequency, from (Dosset et al., 2020)) MHC-II alleles (x-axis) versus
650 across all MHC-II alleles (y-axis). (C) Lineplot showing the position affinity scores across
651 common MHC-II alleles (Methods). Annotated domains from UniProt are highlighted. (D)
652 Heatmap showing amino acid positions that are glycosylated (Watanabe et al., 2020). (E)
653 Barplots (top) and boxplots (bottom) describing the fraction of binding MHC-II alleles and
654 corresponding affinity percentile rank distributions respectively for the top 20 peptides with the
655 highest fraction of common binding alleles. The binding threshold of 10 is shown as a dotted
656 line, with values less than 10 indicating binding. Colors correspond to the regions listed in C. (F)
657 Barplots (top) and boxplots (bottom) describing the fraction of binding MHC-II alleles and
658 corresponding affinity percentile rank distributions respectively for the top 20 peptides within +/-
659 50 amino acids of the FNCY B cell epitope. Colors correspond to the regions listed in C.

660
661 Figure 3: Population variation affecting availability of FNCY proximal T cell epitopes. (A)
662 Barplot showing the aggregated supertype position affinity scores for each position +/- 50 amino
663 acids from the FNCY patch (grey zone). (B) Scatterplot showing the specific supertype position
664 scores for each position +/- 50 amino acids from the FNCY patch (grey zone). The binding
665 threshold of 10 is shown as a dashed blue line, with points below the threshold indicating
666 binding. (C) Barplot showing United States population frequencies, summed across the available
667 alleles in each supertype. (D) Fraction of positions falling below the binding threshold within the
668 region of interest for each supertype.

669
670 Figure 4: Immunological history of relevance to SARS-CoV-2. (A) Barplot showing the
671 percentage of positive responses toward SARS-CoV-2 peptides from unexposed individuals. (B)
672 Barplot showing the number of spot-forming cells (SFC) for tested SARS-CoV-2 peptides
673 against PBMCs from unexposed individuals. Data from Table S1 from (Mateus et al., 2020).

674
675 Figure 5: Learned immunity to other targets that could support T cell responses to SARS-CoV-2.
676 (A) Multiple sequence alignment between SARS-CoV-2, SARS1, MERS, and other human
677 coronaviruses, focusing on the region surrounding the FNCY B cell epitope. (B) SeqLogo plot
678 obtained by clustering IEDB peptides reported to bind to DRB1*01:01. (C) Top results after
679 blasting the FNCYFPLQSYGFQPT peptide against all reference proteins. (D) Barplot describing
680 best peptide affinities across MHC-II alleles of the top 35 unique organisms with one or more

681 peptides matching a peptide with high similarity to 15mers +/-30aa from the FNCY binding
682 epitope based on BLAST analysis. The closer to 0, the greater the binding potential.
683

684

685

686 **Supplemental Table Legends**

687 Supplemental Table 1: SARS-CoV-2 neutralizing antibody residues and references used to
688 generate Fig 1D.

689 Supplemental Table 2: BLAST-identified peptides with affinity, and binding fraction.

690

