

1 **Identification of differences in the magnitude and specificity of SARS-CoV-**
2 **2 nucleocapsid antibody responses in naturally infected and vaccinated**
3 **individuals**

4 Pradeep D. Pushpakumara¹, Chandima Jeewandara¹, Farha Bary¹, Deshan Madushanka¹,
5 Lahiru Perera¹, Inoka Sepali Aberathna¹, Thashmi Nimasha¹, Jeewantha Jayamali¹, Thushali
6 Ranasinghe¹, Heshan Kuruppu¹, Saubhagya Danasekara¹, Ananda Wijewickrama², Graham
7 S. Ogg³, Gathsaurie Neelika Malavige^{1,3*}

8
9
10 ¹Allergy Immunology and Cell Biology Unit, Department of Immunology and Molecular
11 Medicine, University of Sri Jayewardenepura, Nugegoda, Sri Lanka; ²National Institute of
12 Infectious Diseases, Angoda, Sri Lanka; ³MRC Human Immunology Unit, MRC Weatherall
13 Institute of Molecular Medicine, University of Oxford, Oxford, United Kingdom

14
15 Correspondence should be addressed to:

16 Prof. Neelika Malavige DPhil (Oxon), FRCP (Lond), FRCPath (UK)

17 Allergy Immunology and Cell Biology Unit, Department of Immunology and Molecular
18 Medicine, Faculty of Medical Sciences, University of Sri Jayawardanapura, Sri Lanka. Tel
19 +94 (0) 772443193; Email: gathsaurie.malavige@ndm.ox.ac.uk

20

21

22

23 **Abstract**

24 **Background:** As there are limited data on B cell epitopes for the nucleocapsid protein in
25 SARS-CoV-2, we sought to identify the immunodominant regions within the N protein,
26 recognized by patients with varying severity of natural infection with the Wuhan strain (WT),
27 delta, omicron and in those who received the Sinopharm vaccines, which is an inactivated,
28 whole virus vaccine.

29 **Methods:** Using overlapping peptides representing the N protein, with an in-house ELISA,
30 we mapped the immunodominant regions within the N protein, in seronegative (n=30), WT
31 infected (n=30), delta infected (n=30), omicron infected+vaccinated (n=20) and Sinopharm
32 (BBIBP-CorV) vaccinees (n=30). We then investigated the sensitivity and specificity of these
33 immunodominant regions and analysed their conservation with other SARS-CoV-2 variants
34 of concern, seasonal human coronaviruses and bat Sarbecoviruses. We then investigated the
35 kinetics of responses to these regions in those with varying severity of acute COVID-19.

36 **Results:** We identified four immunodominant regions aa 29-52, aa 155-178, aa 274 to 297
37 and aa 365 to 388, were highly conserved within SARS-CoV-2 and the bat coronaviruses.
38 The magnitude of responses to these regions varied based on the infecting SARS-CoV-2
39 variants, with WT infected individuals predominantly recognizing aa155 to 178 regions, delta
40 infected individuals and vaccinated+omicron infected individuals predominantly recognizing
41 regions aa 29 to 52 and aa 274 to 294 regions. Sinopharm vaccinees recognized all four
42 regions, with the magnitude of responses significantly lower than other groups. >80% of
43 individuals gave responses above the positive cut-off threshold to many of the four regions,
44 with some differences with individuals who were infected with different VoCs. These regions
45 were found to be 100% specific, as none of the seronegative individuals gave any responses.

46 **Conclusions:** N-protein specific responses appear to be detectable in over 90% of those who
47 were naturally infected or vaccinated with a whole virus inactivated vaccine, with responses
48 mainly directed against four regions of the protein, which were highly conserved. As these
49 regions were highly specific with high sensitivity, they have a potential to be used to develop
50 diagnostic assays and to be used in development of vaccines.

51

52 **Keywords:** SARS-COV-2; nucleocapsid protein; ELISA; overlapping peptides;
53 immunodominant; conservation; sarbecoviruses; variants; Sinopharm

54

55

56

57

58

59

60

61

62

63

64

65

66

67 **Introduction**

68 The SARS-CoV-2 virus continues to evolve, giving rise to more immune evasive and more
69 transmissible variants, which continue to drive outbreaks globally [21]. Although variants
70 such as omicron (BA.1) were thought to initially cause milder illness, the sub-lineages that
71 subsequently emerged such as BA.2, were associated with more severe disease in certain
72 populations [30]. In fact, BA.2 outbreaks in the United States and in Hong Kong resulted in
73 several fold higher mortality rates than seen during the delta outbreaks in many countries
74 [13]. Many factors could contribute to the differences in mortality rates and hospitalization
75 rates during different outbreaks in different countries such as co-morbidities, age, vaccination
76 rates of a population, the proportion of individuals naturally infected, COVID-19 control
77 measures, better treatment modalities, infra-structure to manage hospitalized patients and
78 seasonal changes [2,5,18]. Among all the factors that have contributed to a reduction in
79 mortality rates, COVID-19 vaccines, are likely to be one of the single most important factors
80 that were responsible for this reduction [2,29].

81

82 While neutralization antibodies (Nabs) have shown to associate with protection against
83 severe disease when infected with the SARS-CoV-2 [1,12], the mRNA COVID-19 vaccines
84 appear to induce higher levels of Nabs compared to other vaccines [15]. However, there is
85 emerging evidence that nucleocapsid (N) protein specific antibody responses may be
86 protective based on data in animal models [8]. Indeed, a high frequency of polyfunctional T
87 cell responses specific for certain epitopes within the N protein was found to associate with
88 milder illness [22] and N protein specific antibody responses were detected earlier in
89 infection and were present at detectable levels in a larger proportion of individuals compared
90 to spike protein specific antibody responses [6].

91 The N protein is one of the most abundant, highly conserved RNA-binding proteins, which
92 plays an important role in the packing of the SARS-CoV-2 genome [3]. It plays an important
93 role in the regulation of the virus replication cycle, inhibits interferon response and induced
94 apoptosis [3]. The N protein, which spans 419 amino acids, consists of five domains and all
95 five have shown to bind to RNA [6]. The region starting from the 388 amino acid position
96 was found to induce a high frequency of immune responses in patients with acute COVID-19
97 (from the Wuhan strain) and was found to be 100% specific to detect infection with SARS-
98 CoV-2. Although the N protein is an important T cell and antibody target, the main
99 immunodominant regions within this protein, targeted by antibodies has not been extensively
100 studied. For instance, although the N protein is highly conserved, as it is an important
101 antibody target, certain mutations in SARS-CoV-2 variants of concern (VoC), can give rise to
102 differences in the magnitude of antibody responses to certain regions. Therefore, we sought
103 to identify the immunodominant regions within the N protein, recognized by patients with
104 varying severity of natural infection with the Wuhan strain (WT), delta, omicron and in those
105 who received the Sinopharm vaccines, which is an inactivated, whole virus vaccine.

106

107

108

109

110

111

112

113 **Methods**

114 **Participants for identification of immunodominant regions within the N protein**

115 Blood samples from healthy adult volunteers who were either vaccinated or naturally infected
116 with SARS-CoV-2 were obtained following informed written consent. Serum separated from
117 blood samples were used to assess antibody responses to the overlapping N peptides in the
118 following groups of individuals.

119 A. SARS-CoV-2 seronegative negative individuals (n=30) prior to COVID-19
120 vaccination (negative)

121 B. Unvaccinated individuals who were naturally infected (n=30) with the SARS-CoV-2
122 wild type/Wuhan strain (WT) from day 14 to 21 from day of onset of symptoms. 5/30
123 of them had severe illness and 25/30 had mild. Clinical disease severity was classified
124 according to the WHO COVID-19 disease severity classification [31]. (WT)

125 C. Unvaccinated individuals who were naturally infected with the SARS-CoV-2 delta
126 variant (n=30), 7 to 21 days from the onset of symptoms. All individuals had mild
127 infection. (delta)

128 D. Those who were vaccinated or who possibly had prior infection (infection status
129 unknown) in those who were subsequently infected with omicron (n=20) days 14 to
130 21 since onset of illness. (Omicron+vaccinated)

131 E. Sinopharm (BBIBP-CorV) vaccine recipients 2 weeks post second dose (n=30)
132 (Sinopharm)

133

134 Ethics statement

135 Blood samples were obtained following informed written consent. Ethics approval was
136 obtained from the Ethics Review Committee of University of Sri Jayewardenepura.

137 **Participants for assessing the kinetics of antibody responses to immunodominant**
138 **regions of N protein**

139 After identification of immunodominant regions within the N protein, the responses to these
140 regions were further assessed in the groups of individuals described above. However, smaller
141 numbers were included in the analysis due to limitations in the sample volume available.

142 A. SARS-CoV-2 seronegative negative individuals (n=15) prior to COVID-19 vaccination

143 B. Unvaccinated individuals who were naturally infected (n=12) with the SARS-CoV-2 wild
144 type/Wuhan strain (WT) from day 14 to 21 since onset of illness

145 C. Unvaccinated individuals who were naturally infected with the SARS-CoV-2 delta variant
146 (n=12), with mild illness, from day 7 to 14 since onset of illness.

147 D. Those who were vaccinated (different vaccines) or who possibly had prior infection
148 (infection status unknown) in those who were infected with omicron (n=22), 14 to 21 days
149 since onset of illness. All participated individuals had mild infection.

150 E. Sinopharm (BBIBP-CorV) vaccine recipients 2 weeks post second dose (n=12)

151 F. Uninfected individuals who received COVID-19 vaccines, which only contain the SARS-
152 CoV-2 spike protein, 3 months since obtaining the second dose. AZD1222 (ChAdOx1)
153 (n=10), Moderna (mRNA-1273) (n=10) and Sputnik V (Gam-COVID) (n=10). This was to
154 assess the specificity of the responses to the immunodominant N peptides.

155

156 **Participants with acute infection due to SARS-CoV-2 WT virus**

157 Adult patients who were acute infected with the SARS-CoV-2 virus and had mild illness
158 (n=16) or severe illness (n=9) during acute stage (<7 days since onset of symptoms) and

159 during late infection (21 to 28 days since onset of symptoms) were recruited following
160 informed written consent, to compare antibody responses against four immunodominant
161 regions between individuals with mild and severe disease during early and late stages of the
162 illness. Clinical disease severity was classified according to the WHO COVID-19 disease
163 severity classification [31].

164

165 **N protein peptide array**

166 Overlapping peptides representing the N protein of SARS-CoV-2 virus (USA-WA1/2020
167 strain of SARS-CoV-2; QHO60601) was obtained through BEI Resources, NIAID, NIH:
168 Peptide Array, SARS-Related Coronavirus 2 Nucleocapsid (N) Protein, NR-52404. The
169 whole peptide array consists of 59 overlapping peptides, which overlap by 10aa to 17aa and
170 10aa with the adjacent peptide. All peptides were dissolved in appropriate solvent mentioned
171 by the manufacturer. Initially, all 59 peptides were pooled in to 4 pools. Namely, pool 1
172 (peptide 1 to 15), pool 2 (peptide 16 to 30), pool 3 (peptide 31-45) and pool 4 (peptide 46-
173 59). Antibody responses to the peptides which gave the highest responses were further
174 assessed.

175

176 **Identification of SARS-CoV-2 serostatus of the participants**

177 The Wantai SARS-CoV-2 total antibody ELISA (Beijing Wantai Biological Pharmacy
178 Enterprise, China) was used to identify the presence of antibodies (IgM, IgG and IgA) to the
179 receptor binding domain (RBD) of the virus. The specificity of this assay in the Sri Lankan
180 population was found to be 100% [17] SARS-CoV-2. Those who tested negative for the
181 presence of total antibodies to the RBD by this assay, were considered to be seronegative.

182

183 In those who had received the spike protein contained vaccines, the presence of
184 asymptomatic infection with the virus was assessed by the presence of N protein specific
185 antibodies. This was done by using the Elecsys® Anti-SARS-CoV-2
186 electrochemiluminescence immunoassay (Cat: 09 203 095 190, Roche Diagnostics,
187 Germany) using the Cobas e 411 analyzer (Roche Diagnostics, Germany). A Cutoff index
188 (COI) ≥ 1.0 was interpreted as reactive and COI < 1.00 was considered non-reactive as
189 indicated by the manufacturer.

190

191 **Measuring ACE2 blocking antibodies by the surrogate virus neutralizing test (sVNT)**

192 ACE2 blocking antibodies were measured using the sVNT assay which measures the
193 percentage of inhibition of binding of the RBD of the S protein to recombinant ACE2
194 (Genscript Biotech, USA). Inhibition percentage $\geq 25\%$ in a sample was considered as
195 positive for Nabs in the Sri Lankan population as previously described [26].

196

197 **Identification of SARS-CoV-2 variants in individuals who were naturally infected with** 198 **SARS-CoV-2**

199 In this study we recruited individuals infected with the WT, delta and omicron. All those who
200 were considered to be infected with the WT had a confirmed SARS-CoV-2 infection (PCR
201 positive) between in March to May 2020, when other VoC were not detected. Infection with
202 either delta or omicron was identified by carrying out genomic sequencing using either the
203 Oxford Nanopore (ONT) or the Illumina platforms as previously [23].

204

205

206

207 **In-house ELISA to determine IgG antibody responses to the SARS-CoV-2 overlapping**
208 **peptides of the N protein**

209 Ninety-six-well microtitre plates (Thermofisher, USA, Pierce™ Cat: 15031) were coated
210 with the overlapping peptide representing the different pools and incubated overnight
211 at 4 °C. The peptides were diluted in bicarbonate/carbonate coating buffer (pH 9.6) and the
212 final concentration of each peptide was 1 µg/100 µl. The plates were blocked with PBS with
213 2% (w/v) bovine serum albumin (Sigma Aldrich, Germany, Cat: A7030) and incubated for 2
214 hours at room temperature and washed before incubation with serum samples diluted 1:500 in
215 1% BSA. After an incubation of 30 min at room temperature, the plates were washed, and
216 incubated with biotinylated goat anti-human IgG antibody (Mabtech, Sweden, Cat: 3820-4-
217 250) diluted 1:1000 in 1% BSA. After a 30-minute incubation at room temperature, the plates
218 were washed and further incubated with Streptavidin–HRP (Mabtech (Sweden) Cat: 3310-9)
219 diluted 1:1000 in 1% BSA solution for 30 minutes. After washing the plates, the TMB
220 ELISA substrate solution (Mabtech, Sweden, Cat: 3652-F10) was added at 100 µl/well and
221 the plates were incubated in the dark for 10 minutes at room temperature. The reaction was
222 stopped by adding 2M H₂SO₄ (Sigma Aldrich, Germany, Cat: 339741) and absorbance values
223 were read at 450nm. Optic density (OD) values above the mean ± 3SD of the OD values of
224 the sera from SARS-CoV-2 seronegative individuals was considered as a positive response
225 for a particular peptide or a pool of peptides.

226

227 **Statistical Analysis**

228 GraphPad Prism version 9 was used for statistical analysis. As the data were not normally
229 distributed, differences in means were compared using the Mann-Whitney U test (two tailed).
230 The descriptive statistics including the mean and frequencies were used to compare antibody
231 responses of individual peptides. Kruskal-Wallis test was used to determine the differences
232 between the antibody levels (indicated by the OD value) in the four peptide pools (pool 1, 2,
233 3, and 4). and four immunodominant regions (P5/6, P23/24, P40/41, and P53/54). If Kruskal-
234 Wallis test was significant, a post hoc test (Dunn test) was done to identify which group or
235 groups different from others. Spearman's correlation coefficient was used to determine the
236 correlation between antibody responses against immunodominant regions of N protein and
237 neutralizing antibodies.

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252 **Results**

253 **Identification immunodominant regions within the N protein of SARS-CoV-2**

254 We initially tested the four pools of the SARS CoV-2 overlapping peptides of the N protein
255 in the cohorts A to E. WT, delta, omicron infected and sinopharm vaccinated individual's
256 antibody responses were significantly different in the four overlapping pools of peptides
257 (Figure 1A to 1D). The number of individuals included in each of the cohorts that tested
258 positive for the different pools is shown in table 1. Omicron infected + vaccinated individuals
259 (who were vaccinated) had the highest positivity rates (>75%) for all four peptide pools.

260

261 Of those in who were infected prior to being vaccinated, the WT infected individuals had the
262 highest positivity rates for pool 2, while delta infected individuals gave highest antibody
263 responses to pool 3 and 4. Sinopharm vaccinees had the highest positivity rates for pool 1.
264 However, Sinopharm vaccinees had overall lower positivity rates and magnitude of responses
265 for all four peptide pools compared to naturally infected individuals.

266

267 **Table 1: The number of individuals in different cohorts who gave a positive response to**
268 **different overlapping peptide pools of the N protein**

SARS-CoV-2 Variant	Number of individuals who gave a positive antibody response				
	Pool 1	Pool 2	Pool 3	Pool 4	At least one pool

WT (n=30)	9 (30%)	15 (50%)	2 (6.7%)	3 (10%)	18/30 (60%)
Delta (n=30)	10 (33.3%)	9 (30%)	17 (56.7%)	16 (53.3%)	22/30 (73.3%)
Omicron + vaccinated (n=20)	16 (80%)	15 (75%)	18 (90%)	16 (80%)	18/20 (90%)
Sinopharm (n=30)	11 (36.7%)	7 (23.3%)	5 (16.7%)	4 (13.3%)	14/30 (46.7%)

269

270 **Mapping of antibody responses in the different cohorts to identify immunodominant**
 271 **regions of N protein**

272 As the WT infected individuals (cohort B) had the highest responses to pool 2, Sinopharm
 273 vaccinees (cohort E) to pool 1, delta infected individuals (cohort C) to pool 3 and 4, we
 274 proceeded to map the immunodominant regions within these different pools of overlapping
 275 peptides, by testing antibody responses to these individual peptides separately.

276

277 In cohort D and E, the highest responses were observed for the two overlapping peptides 5
 278 and 6 of pool 1 (peptides 1 to 15) (Supplementary Figure 1A and 1B). Cohort B (WT infected
 279 individuals) and D (omicron infected+ vaccinated) had the highest responses to overlapping
 280 peptides 23 and 24 of pool 2 (Peptide 16 to 30) (Supplementary Figure 1C and 1D).
 281 Individuals from cohort C (delta infected) and D, had the highest responses to overlapping
 282 peptide 40 and 41 of pool 3 (Peptide 31 to 45) (Supplementary Figure 1E 1F). In cohort C
 283 and D, the highest responses were observed for peptide 53 and peptide 54 of pool 4 (Peptide
 284 46 to 59) (Supplementary Figure 1G and 1H). Based on these results, overlapping peptides 5

285 and 6 of pool 1, overlapping peptides 23 and 24 of pool 2, overlapping peptides 40 and 41 of
286 pool 3 and overlapping peptides 53 and 54 of pool 4, were the immunodominant regions
287 within the N protein.

288

289 **Characterizing antibody responses to the immunodominant regions identified within N** 290 **protein**

291 In order to further characterize the antibody responses to the above immunodominant regions
292 within the N protein, the overlapping peptides 5 and 6, 23 and 24, 40 and 41 and 53 and 54
293 were pooled together from four different pools. The antibody responses for these regions
294 were assessed in all cohort (cohort A to E), to identify responses in all individuals for these
295 pools. Although the WT infected individuals gave the highest antibody responses to P23/24
296 (Figure 2A and B), there was no significant difference ($p=0.057$) in the magnitude of
297 responses for the four immunodominant peptide pools in this group. Similarly, delta infected
298 individuals had a similar magnitude of responses for the four pools ($p=0.53$). Omicron
299 infected+vaccinated (cohort D) had the highest responses to P40/41 and P53/54, while
300 Sinopharm vaccinees had the highest responses to P5/6 (Figure 2A and B).

301

302 The positivity rates for each of the four immunodominant regions in these cohorts is shown in
303 table 2. Overall, all cohorts had >80% positivity rates for all 4 regions, except lower
304 positivity rates in delta infected and omicron infected+ vaccinated individuals for P23/24
305 region. In contrast, all WT infected individuals gave a positive response for this region, while
306 they had low positivity rates (66.7%) for P53/54.

307

308 We also assessed specificity of the antibody responses to the four immunodominant regions
 309 of the N protein (P5/6, P23/24, P40/41, and P53/54), by investigating responses in those who
 310 received COVID-19 vaccines with only the spike protein. Accordingly, responses were
 311 assessed in those who received AZD1222 (ChAdOx1) (n=10), Moderna (mRNA-1273)
 312 (n=10), and Sputnik V (Gam-COVID-Vac) (n=10). None of the SARS-CoV-2 seronegative
 313 individuals (cohort A) or those who received ChAdOx1, Gam-COVID-Vac or mRNA-1273
 314 responded to the peptides P53/54, while one individual (1/30) had a positive response to the
 315 peptides P5/6. Two individuals (2/30) responded to the peptides P23/24 and P40/41 (Table 2,
 316 Figure 2C). Therefore, while the specificity of peptide P53/54 was 100% in detecting SARS-
 317 CoV-2 N protein specific responses following or vaccination (whole virus vaccine) or
 318 infection, the specificity of peptides P5/6 was 96.7% while for peptides P23/24 and P40/41 it
 319 was 93.3%.

320

321 **Table 2: The number of individuals in different cohorts who gave a positive response to**
 322 **immunodominant regions within the N protein**

SARS-CoV-2 Variant	Number of individuals who gave a positive antibody response				
	P5/6	P23/24	P40/41	P53/54	At least for one pair
WT (n=12)	10 (83.3%)	12 (100%)	11 (91.7%)	8 (66.7%)	12/12 (100%)
Delta (n=12)	11 (91.7%)	9 (75%)	11 (91.7%)	10 (83.3%)	12/12 (100%)
Omicron+ vaccinated (n=22)	20 (90.9%)	16 (72.7%)	20 (90.9%)	18 (81.8%)	20/22 (90.9%)

Sinopharm (n=12)	12 (100%)	11 (91.7%)	11 (91.7%)	11 (91.7%)	12/12(100%)
Total (n=58)	53 (91.4%)	48 (82.8%)	53 (91.4%)	47 (81.0%)	56/58 (96.5%)
AstraZeneca (n=10)	0	1 (10%)	1 (10%)	0	1/10 (10%)
Sputnik V (n=10)	0	0	1 (10%)	0	1/10 (10%)
Moderna (n=10)	1 (10%)	1 (10%)	0	0	2/10 (20%)
Total (n=30)	1 (3.3%)	2 (6.7%)	2 (6.7%)	0	4/30 (13.3%)

323 **Conservation of immunodominant regions of the N protein of SARS-CoV-2 with**
 324 **seasonal human coronavirus and SARS-CoV-2 variants of concern (VoC)**

325 As the consensus peptide sequence may not be representative of the infecting subtype, we
 326 determined the conservation within these four immunodominant regions within the different
 327 SARS-CoV-2 variants (Alpha, QVX37034.1; Beta, QWW93444.1; Gamma, QXF23757.1;
 328 Delta, UKA47847.1; Omicron, (BA.1 (SriLanka/aicbu4450/2022), BA.2
 329 (SriLanka/aicbu4463/2022) and BA.5 (USA/CA-CDPH-FS27225444/2022) and also the
 330 cross reactivity with other seasonal human coronaviruses (OC43, QBP84763.1; HKU1,
 331 ABG77571.1; NL63, YP_003771.1). We used Jalview software [28] and tools available at
 332 European Bioinformatics Institute (EBI) (www.ebi.ac.uk, 22 March 2022) to determine
 333 conservation between identified four immunodominant regions of the wild type SARS-CoV-2
 334 (YP_009724397.2) The four regions in which the conservation and cross reactivity were
 335 assessed are as follows;

336 P5/P6: ²⁹NGERSGARSKQRRPQGLPNNTASW⁵²

337 P23/P24:¹⁵⁵AAIVLQLPQGTTLPKGFYAEGSRG¹⁷⁸

338 P40/P41:²⁷⁴FGRRGPEQTQGNFGDQELIRQGTG²⁹⁷)

339 P53/54: ³⁶⁵PTEPKKDKKKKADETQALPQRQKK³⁸⁸

340 All regions showed <50% sequence identity with the three seasonal human coronaviruses
341 (Supplementary table 1 Figure 3A to 3D). P5/6 showed <20% sequence identity with OC43,
342 HKU1, and NL63, with P53/54 showing <10% sequence identity with these viruses
343 (Supplementary table 1). In contrast, the regions P23/24 and P40/41 showed >45% of
344 sequence identity with OC43 and HKU1 but not with NL63. All four immunodominant
345 regions were found to be conserved in both alpha and beta variants (Supplementary table 1
346 and Figure 3E to H), while there was 95.8% sequence identity with delta (single amino acid
347 replacement) in the P53/54 region and 95.8% in the P40/41 region with the gamma variant
348 (single amino acid replacement). All three omicron sub-lineages (BA.1, BA.2 and BA.5) P5/6
349 have a 3 amino acid deletion within the regions represented by P5/6 and therefore, a sequence
350 identity of 87.5% (Figure 3E). P23/24 regions was 100% conserved in all five VoC (Figure
351 3F).

352

353 As the SARS-CoV-2 virus continues to further evolve and due to the future threat of other bat
354 coronaviruses spilling over and causing future pandemics, many Pan-Sarbecovirus vaccines
355 are currently under development [9,19]. Therefore, we proceeded to find out the conservation
356 of these four regions with 4 bat coronaviruses, RS4081, WIV1, RatG13 and Rf1 (Figure 4)
357 and Supplementary table 1. These four regions showed >91% sequence identity with all the
358 four bat coronaviruses.

359

360 **Antibody responses to the four immunodominant regions in patients with varying**
361 **severity of COVID-19 due to the WT virus**

362 We then sought to compare antibody responses between mild and severe disease during early
363 and late stage of the infection in individuals who had mild illness (n=16) or severe illness
364 (n=9) during acute stage (<7 days since onset of symptoms) and during late infection (21 to
365 28 days since onset of symptoms). Those with severe illness had significantly higher antibody
366 responses to P23/24, P40/41 and P53/54 during the first week of illness compared to those
367 with mild illness (Figure 5A). During late infection, those with severe disease had
368 significantly higher antibody responses to all four regions than individuals with mild illness
369 (Figure 5B).

370

371 We then sought to explore if antibody responses to these immunodominant regions, correlate
372 with neutralizing antibody (Nab) responses, by comparing ACE2 blocking antibodies (which
373 were shown to correlate with Nabs), by using a surrogate SARS-CoV-2-neutralizing antibody
374 assay. The ACE2 blocking antibodies did not correlate with antibody response against all
375 four regions (Spearman's $r=0.17$, $p=0.02$) (Figure 5C).

376

377

378

379

380

381

382

383

384

385

386

387 **Discussion**

388 In this study we have identified four immunodominant regions within the N protein, which
389 gave a high frequency of responses in those infected with the WT virus, delta, omicron and
390 those vaccinated with Sinopharm. Overall, >80% of individuals gave responses above the
391 positive cut-off threshold to many of the four regions, with some differences with individuals
392 who were infected with different VoCs. These regions were found to be 100% specific, as
393 none of the seronegative individuals gave any responses. However, 6.7% to 10% of
394 individuals who had received the spike proteins vaccines and therefore, should not have
395 responses to these regions also responded. Although they had tested negative for infection
396 with SARS-CoV-2 by the commercial N protein-specific antibody assay, it is possible that
397 they could still have been naturally infected. Due to the high sensitivity and specificity of
398 these regions, they have a potential to be used to identify N protein specific antibody
399 responses, especially regions aa 29 to 52 (P5/6) and aa 365 to 388 (P40/41), for which >90%
400 of individuals responded to. However, as we used overlapping peptides to identify potential
401 epitopes, we are likely to have missed conformational epitopes which could be key antibody
402 recognition sites.

403

404 The five domains of the N protein have shown to bind to RNA and carry out multiple
405 functions including RNA interference, regulating virus replication and host immune evasion
406 [6]. Of the four regions identified here, one is within the N terminal domain (aa 29-52, P5/6),

407 one in the RNA binding domain (aa 155-178, P23/24), one within the dimerization domain
408 (aa 274 to 297, P40/41) and one in the C-terminal domain (aa365 to 388/ P53/54). Some
409 previous studies had identified immunodominant regions of the N protein, in llamas
410 (domesticated South American camel), and had identified antibodies that bind to highly
411 conserved regions within the N protein [32]. One of these antibodies had shown to bind to the
412 C-terminal domain (aa 49-174) and two antibodies to the N-terminal domain (aa 247-364 and
413 aa 365-419)[32]. The four immunodominant regions that were identified here, also fall within
414 these three regions. Another study, which screened for B cell epitopes within the N protein
415 using mouse models, identified aa 401 to 408 as the main antibody target. Our data show that,
416 the predominant B cell epitopes identified within the N protein differs based on the infecting
417 SARS-CoV-2 variant. For instance, those who were infected with the WT virus,
418 predominantly recognized the aa155 to 178 regions (P23/24), whereas those who were
419 infected with delta had the highest responses to the aa 29 to 52 and aa 274 to 294 regions.
420 However, the responses to these overlapping peptides were assessed using the sequence of a
421 Wuhan virus strain isolated from USA in 2020 and the epitope recognition could be different,
422 based on the sequence of peptides used. In those who were infected with different omicron
423 sub-lineages (BA.1 and BA.2) also had the highest responses to the aa 29 to 52 and aa 274 to
424 294 regions. Therefore, the predominant B cell epitope recognition, appears to differ based on
425 the variant of infection. Although Sinopharm vaccinees responded to all four regions, the
426 magnitude of the responses was significantly lower than following natural infection. This is
427 possibly due to natural infection inducing more robust responses to the N protein than
428 following inactivated vaccines containing the whole protein.

429

430 The mortality rates and hospitalization rates have varied widely throughout the COVID-19
431 pandemic in different countries, with many countries in Europe, and United States reporting

432 higher mortality rates and hospitalization rates than some countries in Africa and Asia,
433 despite higher rates of vaccination [13,14]. These differences could be attributed to reporting
434 of COVID-19 deaths and limitations in testing, as many countries in sub-Saharan Africa have
435 reported high excess mortality rates [10]. Sri Lanka experienced high mortality rates during
436 the delta outbreak during the months from June to October 2021 prior to vaccination [24].
437 However, mortality rates have been significantly less (0.85/ million individuals in Sri Lanka)
438 during the massive omicron wave (BA.2), than many European countries and the United
439 States (mortality rates 4.03/million individuals in Europe and 5.4/million individuals in
440 United States) [13]. Only 18% of Sri Lankans had received an mRNA booster dose, when the
441 omicron variant was rapidly spreading in Sri Lanka [14]. The lower mortality rates seen in Sri
442 Lanka during the omicron outbreak were unlikely to be due to under reporting or limited
443 testing as the excess mortality rates in Sri Lanka were found to be less than the excess
444 mortality rates reported in Europe and North America [10]. Sinopharm/BBIBP-CorV was the
445 most widely used vaccine, in Sri Lanka with 12 million (70.6%) individuals receiving this
446 vaccine by end of December 2021 [9]. Sinopharm/BBIBP-CorV vaccine was found to be less
447 immunogenic than the mRNA-1273, AZD1222 and Sputnik V, 3 months post second dose, in
448 a head-to-head comparison in the Sri Lankan population, based on ACE2 blocking antibodies
449 and antibodies to the receptor binding domain of the spike protein [15]. However, as
450 Sinopharm/BBIBP-CorV is an inactivated vaccine, it did induce T cell and antibody
451 responses to the N protein [16]. In addition, although the N protein was thought to be
452 localized to the cytosol, it was recently shown that this protein was expressed on the surface
453 of infected cells [20]. As the N protein has shown to bind to several different types of
454 chemokines, antibodies against the N protein could also inhibit chemotaxis of leucocytes
455 [20]. Furthermore, antibodies bound to N protein were shown to activate FcR expressing
456 innate immune cells, further contributing to the phagocytosis and apoptosis of infected

457 cells[20]. Although there could be many reasons for the differences in mortality rates for
458 different variants, it is possible that antibody responses to the N protein, offered additional
459 protection in Sinopharm vaccinees, which should be further investigated.

460

461 Due to the rapidly evolving nature of SARS-CoV-2 and emergence of more immune evasive
462 omicron sub-lineages, there is a global effort to develop a pan-Sarbecovirus vaccine [7,9,11].
463 However, many pan-Sarbecovirus vaccines only use the spike protein as the immunogen and
464 explore the immune responses to the spike protein [19,25], while only a few vaccines also
465 include the N protein [9]. In this study we show that the four immunodominant regions
466 identified here, were highly conserved regions within SARS-CoV-2 and the bat
467 coronaviruses. Although the neutralizing antibody responses for many bat Sarbecoviruses has
468 been investigated [27], there are no data if antibodies targeting the main B cell epitopes
469 within the N protein, also cross-neutralize the most frequent bat Sarbecoviruses, which would
470 be important. While vaccination, especially with an mRNA booster dose induced high levels
471 of Nabs and protected individuals from severe disease [4], natural infection and vaccination
472 induced a high magnitude of durable immune responses [2]. In fact, it was shown that two
473 doses of an mRNA vaccine and natural infection gave similar immune responses as three
474 doses of a mRNA vaccine, while the immune responses induced by natural infection were
475 longer lasting [2]. Therefore, in order to induce persistent and broad immune responses,
476 antibody and T cell responses to the N protein may play an important role in addition to
477 Nabs.

478

479 **Conclusions**

480 We have identified four immunodominant regions within the N protein of SARS-CoV-2,
481 which are highly conserved in the SARS-CoV-2 variants and also show high conservation in
482 bat Sarbecoviruses. Responses to these regions were highly specific and elicited responses in
483 > 90% of naturally infected individuals or those who received a whole virus inactivated
484 vaccine to at least two of the regions. As these regions were highly specific with high
485 sensitivity, they have a potential to be used to develop diagnostic assays and to be used in
486 development of vaccines.

487

488 **List of abbreviations**

489 Nabs: Neutralization antibodies

490 N protein: Nucleocapsid protein

491 VoC: variants of concern

492 sVNT: surrogate virus neutralizing test

493 RBD: receptor binding domain

494 WT: Wuhan strain of SARS-CoV-2

495 **Ethics approval and consent to participate**

496 Ethics approval was obtained by the Ethics Review Committee of the University of Sri

497 Jayewardenepura. All individuals gave informed, written consent.

498 **Consent for publication**

499 Not applicable

500 **Availability of data and materials**

501 All data is available in the manuscript and figures.

502 **Competing interests**

503 Authors have no competing interests.

504

505 **Funding**

506 We are grateful to Allergy Immunology and Cell Biology Unit, University of Sri
507 Jayewardenepura, the NIH, USA (grant number 5U01AI151788-02), World bank, Sri Lanka
508 Covid 19 Emergency Response and Health Systems Preparedness Project (ERHSP)
509 of Ministry of Health Sri Lanka funded by World Bank and the UK Medical Research
510 Council for funding.

511 **Authors' contributions**

512 Conceptualization of the study: PDP, CJ, GNM

513 Data curation: ISA, TN, JJ, TR, HK, SD

514 Project administration: CJ, AW, GNM

515 Experiments: PDP, FB, DM, LP

516 Data analysis: PDP

517 Funding: CJ, GSO, GNM

518 Writing the manuscript: PDP, GNM

519 Reviewing the manuscript: CJ, GSO

520

521 **Acknowledgements**

522 None

523

524 **References**

- 525 1. Addetia A., Crawford K.H.D., Dingens A., Zhu H., Roychoudhury P., Huang M.L.,
526 Jerome K.R., Bloom J.D. and Greninger A.L. Neutralizing Antibodies Correlate with
527 Protection from SARS-CoV-2 in Humans during a Fishery Vessel Outbreak with a
528 High Attack Rate. *Journal of clinical microbiology* 58(11), 2020.
- 529 2. Altarawneh H.N., Chemaitelly H., Ayoub H.H., Tang P., Hasan M.R., Yassine H.M.,
530 Al-Khatib H.A., Smatti M.K., Coyle P., Al-Kanaani Z., Al-Kuwari E., Jeremijenko
531 A., Kaleeckal A.H., Latif A.N., Shaik R.M., Abdul-Rahim H.F., Nasrallah G.K., Al-
532 Kuwari M.G., Butt A.A., Al-Romaihi H.E., Al-Thani M.H., Al-Khal A., Bertollini R.
533 and Abu-Raddad L.J. Effects of Previous Infection and Vaccination on Symptomatic
534 Omicron Infections. *The New England journal of medicine*, 2022.
- 535 3. Bai Z., Cao Y., Liu W. and Li J. The SARS-CoV-2 Nucleocapsid Protein and Its Role
536 in Viral Structure, Biological Functions, and a Potential Target for Drug or Vaccine
537 Mitigation. *Viruses* 13(6), 2021.
- 538 4. Barda N., Dagan N., Cohen C., Hernan M.A., Lipsitch M., Kohane I.S., Reis B.Y. and
539 Balicer R.D. Effectiveness of a third dose of the BNT162b2 mRNA COVID-19
540 vaccine for preventing severe outcomes in Israel: an observational study. *Lancet*
541 398(10316):2093-2100, 2021.
- 542 5. Bechman K., Yates M., Mann K., Nagra D., Smith L.J., Rutherford A.I., Patel A.,
543 Periselneris J., Walder D., Dobson R.J.B., Kraljevic Z., Teo J.H.T., Bernal W., Barker

- 544 R., Galloway J.B. and Norton S. Inpatient COVID-19 mortality has reduced over
545 time: Results from an observational cohort. *PloS one* 17(1):e0261142, 2022.
- 546 6. Burbelo P.D., Riedo F.X., Morishima C., Rawlings S., Smith D., Das S., Strich J.R.,
547 Chertow D.S., Davey R.T. and Cohen J.I. Sensitivity in Detection of Antibodies to
548 Nucleocapsid and Spike Proteins of Severe Acute Respiratory Syndrome Coronavirus
549 2 in Patients With Coronavirus Disease 2019. *The Journal of infectious diseases*
550 222(2):206-213, 2020.
- 551 7. Cao Y., Yisimayi A., Jian F., Song W., Xiao T., Wang L., Du S., Wang J., Li Q.,
552 Chen X., Yu Y., Wang P., Zhang Z., Liu P., An R., Hao X., Wang Y., Wang J., Feng
553 R., Sun H., Zhao L., Zhang W., Zhao D., Zheng J., Yu L., Li C., Zhang N., Wang R.,
554 Niu X., Yang S., Song X., Chai Y., Hu Y., Shi Y., Zheng L., Li Z., Gu Q., Shao F.,
555 Huang W., Jin R., Shen Z., Wang Y., Wang X., Xiao J. and Xie X.S. BA.2.12.1, BA.4
556 and BA.5 escape antibodies elicited by Omicron infection. *Nature*, 2022.
- 557 8. Chansaenroj J., Yorsaeng R., Posuwan N., Puenpa J., Wanlapakorn N., Sudhinaraset
558 N., Sripramote M., Chalongsriyalert P., Jirajariyavej S., Kiatpanabhikul P., Saiyarin
559 J., Soudon C., Thienfaidee O., Palakawong Na Ayuthaya T., Brukesawan C.,
560 Chirathaworn C., Intharasongkroh D., Chaiwanichsiri D., Issarasongkhram M.,
561 Kitphati R., Mungaomklang A., Nagavajara P. and Poovorawan Y. Long-term
562 specific IgG response to SARS-CoV-2 nucleocapsid protein in recovered COVID-19
563 patients. *Sci Rep* 11(1):23216, 2021.
- 564 9. Coleon S., Wiedemann A., Surenaud M., Lacabaratz C., Hue S., Prague M.,
565 Cervantes-Gonzalez M., Wang Z., Ellis J., Sansoni A., Pierini C., Bardin Q., Fabregue
566 M., Sharkaoui S., Hoest P., Dupaty L., Picard F., El Hajj M., Centlivre M., Ghosn J.,
567 French C.C.S.G., Thiebaut R., Cardinaud S., Malissen B., Zurawski G., Zarubica A.,

- 568 Zurawski S.M., Godot V. and Levy Y. Design, immunogenicity, and efficacy of a
569 pan-sarbecovirus dendritic-cell targeting vaccine. *EBioMedicine* 80:104062, 2022.
- 570 10. Collaborators C.-E.M. Estimating excess mortality due to the COVID-19 pandemic: a
571 systematic analysis of COVID-19-related mortality, 2020-21. *Lancet*
572 399(10334):1513-1536, 2022.
- 573 11. Fountain-Jones N.M., Appaw R.C., Carver S., Didelot X., Volz E. and Charleston M.
574 Emerging phylogenetic structure of the SARS-CoV-2 pandemic. *Virus Evol*
575 6(2):veaa082, 2020.
- 576 12. Garcia-Beltran W.F., Lam E.C., Astudillo M.G., Yang D., Miller T.E., Feldman J.,
577 Hauser B.M., Caradonna T.M., Clayton K.L., Nitido A.D., Murali M.R., Alter G.,
578 Charles R.C., Dighe A., Branda J.A., Lennerz J.K., Lingwood D., Schmidt A.G.,
579 Iafrate A.J. and Balazs A.B. COVID-19-neutralizing antibodies predict disease
580 severity and survival. *Cell* 184(2):476-488 e411, 2021.
- 581 13. Hannah Ritchie E.M., Lucas Rodés-Guirao, Cameron Appel, Charlie Giattino,
582 Esteban Ortiz-Ospina, Joe Hasell, Bobbie Macdonald, Diana Beltekian and Max
583 Roser (2020). *Coronavirus Pandemic (COVID-19)*. 02/24/2022 ed, 2022.
- 584 14. Hannah Ritchie E.O.-O., Diana Beltekian, Edouard Mathieu, Joe Hasell, Bobbie
585 Macdonald, Charlie Giattino, Cameron Appel and Max Roser. 2021. *Coronavirus*
586 *(COVID-19) Vaccinations*. In *COVID-19 vaccine doses administered per 100 people*,
587 Apr 1, 2021 Our World in Data <<https://ourworldindata.org/covid-vaccinations>>.
588 Accessed
- 589 15. Jeewandara C., Aberathna I.S., Danasekara S., Gomes L., Fernando S., Guruge D.,
590 Ranasinghe T., Gunasekera B., Kamaladasa A., Kuruppu H., Somathilake G.,
591 Jayamali J., Jayathilaka D., Wijayatilake H.D.K., Pushpakumara P.D., Harvie M.,
592 Nimasha T., de Silva S.D.G., Wijayamuni R., Schimanski L., Rijal P., Tan J.,

- 593 Townsend A., Ogg G.S. and Malavige G.N. Comparison of the Immunogenicity of
594 five COVID-19 vaccines in Sri Lanka. *Immunology*, 2022.
- 595 16. Jeewandara C., Aberathna I.S., Pushpakumara P.D., Kamaladasa A., Guruge D.,
596 Wijesinghe A., Gunasekera B., Tanussiya S., Kuruppu H., Ranasinghe T., Dayarathne
597 S., Dissanayake O., Gamalath N., Ekanayake D., Mpdj J., Jayathilaka D.,
598 Dissanayake M., Madusanka D., Jayadas T.T., Mudunkotuwa A., Somathilake G.,
599 Harvie M., Nimasha T., Danasekara S., Wijayamuni R., Schimanski L., Rijal P., Tan
600 T.K., Dong T., Townsend A., Ogg G.S. and Malavige G.N. Immune responses to
601 Sinopharm/BBIBP-CorV in individuals in Sri Lanka. *Immunology*, 2022.
- 602 17. Jeewandara C., Guruge D., Jayathilaka D., Deshan Madhusanka P.A., Pushpakumara
603 P.D., Tanussiya Ramu S., Sepali Aberathna I., Saubhagya Rasikangani Danasekara
604 D.R., Pathmanathan T., Gunatilaka B., Malavige S., Dias Y., Wijayamuni R., Ogg
605 G.S. and Malavige G.N. Transmission dynamics, clinical characteristics and sero-
606 surveillance in the COVID-19 outbreak in a population dense area of Colombo, Sri
607 Lanka April- May 2020. *PloS one* 16(11):e0257548, 2021.
- 608 18. Kifer D., Bugada D., Villar-Garcia J., Gudelj I., Menni C., Sudre C., Vuckovic F.,
609 Ugrina I., Lorini L.F., Posso M., Bettinelli S., Ughi N., Maloberti A., Epis O.,
610 Giannattasio C., Rossetti C., Kalogjera L., Persec J., Ollivere L., Ollivere B.J., Yan
611 H., Cai T., Aithal G.P., Steves C.J., Kantele A., Kajova M., Vapalahti O., Sajantila A.,
612 Wojtowicz R., Wierzbza W., Krol Z., Zaczynski A., Zycinska K., Postula M., Luksic
613 I., Civljak R., Markotic A., Brachmann J., Markl A., Mahnkopf C., Murray B.,
614 Ourselin S., Valdes A.M., Horcajada J.P., Castells X., Pascual J., Allegri M.,
615 Primorac D., Spector T.D., Barrios C. and Lauc G. Effects of Environmental Factors
616 on Severity and Mortality of COVID-19. *Front Med (Lausanne)* 7:607786, 2020.

- 617 19. Liu Z., Chan J.F., Zhou J., Wang M., Wang Q., Zhang G., Xu W., Chik K.K., Zhang
618 Y., Wang Y., Yuen K.Y., Lu L. and Jiang S. A pan-sarbecovirus vaccine induces
619 highly potent and durable neutralizing antibody responses in non-human primates
620 against SARS-CoV-2 Omicron variant. *Cell research* 32(5):495-497, 2022.
- 621 20. Lopez-Munoz A.D., Kosik I., Holly J. and Yewdell J.W. Cell surface SARS-CoV-2
622 nucleocapsid protein modulates innate and adaptive immunity. *Sci Adv*
623 8(31):eabp9770, 2022.
- 624 21. Markov P.V., Katzourakis A. and Stilianakis N.I. Antigenic evolution will lead to new
625 SARS-CoV-2 variants with unpredictable severity. *Nature reviews*, 2022.
- 626 22. Peng Y., Felce S.L., Dong D., Penkava F., Mentzer A.J., Yao X., Liu G., Yin Z., Chen
627 J.L., Lu Y., Wellington D., Wing P.A.C., Dominey-Foy D.C.C., Jin C., Wang W.,
628 Hamid M.A., Fernandes R.A., Wang B., Fries A., Zhuang X., Ashley N., Rostron T.,
629 Waugh C., Sopp P., Hublitz P., Beveridge R., Tan T.K., Dold C., Kwok A.J., Rich-
630 Griffin C., Dejnirattisa W., Liu C., Kurupati P., Nassiri I., Watson R.A., Tong O.,
631 Taylor C.A., Kumar Sharma P., Sun B., Curion F., Revale S., Garner L.C., Jansen K.,
632 Ferreira R.C., Attar M., Fry J.W., Russell R.A., Consortium C., Stauss H.J., James
633 W., Townsend A., Ho L.P., Klenerman P., Mongkolsapaya J., Screaton G.R.,
634 Dendrou C., Sansom S.N., Bashford-Rogers R., Chain B., Smith G.L., McKeating
635 J.A., Fairfax B.P., Bowness P., McMichael A.J., Ogg G., Knight J.C. and Dong T. An
636 immunodominant NP105-113-B*07:02 cytotoxic T cell response controls viral
637 replication and is associated with less severe COVID-19 disease. *Nature immunology*
638 23(1):50-61, 2022.
- 639 23. Ranasinghe D., Jayadas T.T.P., Jayathilaka D., Jeewandara C., Dissanayake O.,
640 Guruge D., Ariyaratne D., Gunasinghe D., Gomes L., Wijesinghe A., Wijayamuni R.
641 and Malavige G.N. Comparison of different sequencing techniques for identification

- 642 of SARS-CoV-2 variants of concern with multiplex real-time PCR. *PloS one*
643 17(4):e0265220, 2022.
- 644 24. Ranasinghe D., Jayathilaka D., Jeewandara C., Gunasinghe D., Ariyaratne D., Jayadas
645 T.T.P., Kuruppu H., Wijesinghe A., Bary F.F., Madhusanka D., Pushpakumara P.D.,
646 Guruge D., Wijayamuni R., Ogg G.S. and Malavige G.N. Molecular Epidemiology of
647 AY.28 and AY.104 Delta Sub-lineages in Sri Lanka. *Front Public Health* 10:873633,
648 2022.
- 649 25. Su S., Li W. and Jiang S. Developing pan-beta-coronavirus vaccines against emerging
650 SARS-CoV-2 variants of concern. *Trends in immunology* 43(3):170-172, 2022.
- 651 26. Tan C.W., Chia W.N., Qin X., Liu P., Chen M.I., Tiu C., Hu Z., Chen V.C., Young
652 B.E., Sia W.R., Tan Y.J., Foo R., Yi Y., Lye D.C., Anderson D.E. and Wang L.F. A
653 SARS-CoV-2 surrogate virus neutralization test based on antibody-mediated blockage
654 of ACE2-spike protein-protein interaction. *Nature biotechnology* 38(9):1073-1078,
655 2020.
- 656 27. Tan C.W., Chia W.N., Young B.E., Zhu F., Lim B.L., Sia W.R., Thein T.L., Chen
657 M.I., Leo Y.S., Lye D.C. and Wang L.F. Pan-Sarbecovirus Neutralizing Antibodies in
658 BNT162b2-Immunized SARS-CoV-1 Survivors. *The New England journal of*
659 *medicine* 385(15):1401-1406, 2021.
- 660 28. Waterhouse A.M., Procter J.B., Martin D.M., Clamp M. and Barton G.J. Jalview
661 Version 2--a multiple sequence alignment editor and analysis workbench.
662 *Bioinformatics* 25(9):1189-1191, 2009.
- 663 29. Watson O.J., Barnsley G., Toor J., Hogan A.B., Winskill P. and Ghani A.C. Global
664 impact of the first year of COVID-19 vaccination: a mathematical modelling study.
665 *The Lancet infectious diseases*, 2022.

- 666 30. Whitaker M., Elliott J., Bodinier B., Barclay W., Ward H., Cooke G., Donnelly C.A.,
667 Chadeau-Hyam M. and Elliott P. Variant-specific symptoms of COVID-19 among
668 1,542,510 people in England. medRxiv:2022.2005.2021.22275368, 2022.
- 669 31. WHO. Clinical management of severe acute respiratory infection when novel
670 coronavirus (2019-nCoV) infection is suspected: interim guidance: WHO; 2020 28th
671 January 2020.
- 672 32. Ye Q., Lu S. and Corbett K.D. Structural Basis for SARS-CoV-2 Nucleocapsid
673 Protein Recognition by Single-Domain Antibodies. *Frontiers in immunology*
674 12:719037, 2021.
- 675
- 676
- 677
- 678
- 679
- 680
- 681
- 682
- 683
- 684
- 685
- 686
- 687

688

689

690 **Figure legends**

691 **Figure 1: Antibody responses for the four overlapping pools of peptide of the N protein** 692 **of SARS-CoV-2**

693 IgG antibody responses were measured by an in-house ELISA for pool 1 (A), pool 2 (B),
694 pool 3 (C) and pool 4 (D) containing overlapping peptides of the N protein in SARS-CoV-2
695 seronegative (non-vaccinated) individuals (n=30), WT infected individuals (n=30), delta
696 infected individuals (n=30), omicron infected and vaccinated individuals (n=20), and
697 Sinopharm vaccinees (n=30). Kruskal-Wallis test was used to determine the differences
698 between the levels of antibody responses in four peptide pools (pool 1, 2, 3, and 4). Dotted
699 line shows the cutoff value (OD value) of a positive response for the each of the peptide
700 pools. The error bars indicate the median and the interquartile ranges. Black: seronegatives ;
701 Brown: WT infected individuals ; Blue: Delta infected individuals ; Green: Omicron +
702 vaccinated; Red: Sinopharm vaccinees.

703

704 **Figure 2: Characterizing antibody responses to the immunodominant regions identified** 705 **within N protein**

706 IgG antibody responses to the four immunodominant regions (P5/6, P23/24, P40/41, and
707 P53/54) in the N protein were measured by an in-house ELISA in SARS-CoV-2 in SARS-
708 CoV-2 seronegative (non-vaccinated) individuals (n=15), WT infected individuals (n=12),
709 delta infected individuals (n=12), omicron infected and vaccinated individuals (n=22), and
710 Sinopharm vaccinees (n=12) (A). The magnitude of antibody responses to these regions in

711 the above cohorts were compared with each other (B). In order to determine specificity,
712 antibody responses were measured in individuals who were vaccinated (uninfected) with
713 AZD1222 (n=10), Moderna (n=10), and Sputnik V (n=10). Kruskal-Wallis test was used to
714 determine the differences between the levels of antibody responses in four peptide pools
715 (pool 1, 2, 3, and 4). The error bars indicate the median and the interquartile ranges.

716

717 **Figure 3: Analysis of conservation of immunodominant regions of the N protein of**
718 **SARS-CoV-2 with seasonal human coronavirus and SARS-CoV-2 variants of concern**
719 **(VoC)**

720 The cross reactivity of the four immunodominant regions P5/6 (A), P23/24 (B), P40/41 (C),
721 and P53/54 (D) with three seasonal human corona viruses (OC43, HKU1, and NL63) were
722 determined The conservation within these four immunodominant regions were also assessed
723 for the five VoCs (alpha, beta, gamma, delta, and omicron (BA.1, BA.2, and BA.5). P5/6 (E),
724 P23/24 (F), P40/41 (G), and P53/54 (H). Matching (sequence identity 100%) respective
725 immunodominant regions were highlighted purple color.

726

727 **Figure 4: Analysis of conservation of immunodominant regions of the N protein of**
728 **SARS-CoV-2 with bat coronaviruses**

729 The cross reactivity of the four immunodominant regions P5/6 (A), P23/24 (B), P40/41 (C),
730 and P53/54 (D) RS4081, WIV1, RatG13 and Rf1 were analyzed. Matching (sequence identity
731 100%) respective immunodominant regions were highlighted purple color.

732

733 **Figure 5: Antibody responses to the four immunodominant regions in patients with**
734 **varying severity of COVID-19 due to the WT virus**

735 Antibody responses to the four immunodominant regions were measured by an in-house
736 ELISA, in patients infected with the WT of SARS-CoV-2 with mild (n=16) and severe
737 disease (n=9) during early illness (<7 days since onset of symptoms) (A) and late illness (21
738 to 28 days since onset of symptoms) (B). The antibody responses to the four regions were
739 correlated with ACE2 blocking antibodies measured by the surrogate virus neutralization test,
740 and the ACE2 blocking antibodies did not correlate with the levels of the four
741 immunodominant regions (C). The Mann-Whitney U test (two-tailed) was used to determine
742 the differences in antibody levels between those with mild and severe disease. All tests were
743 two sided. The error bars indicate the median and the interquartile ranges.

744

745 **Supplementary figure legends**

746 **Supplementary Figure 1: Mapping of antibody responses in the different cohorts to**
747 **identify immunodominant regions within the pools of overlapping peptides.**

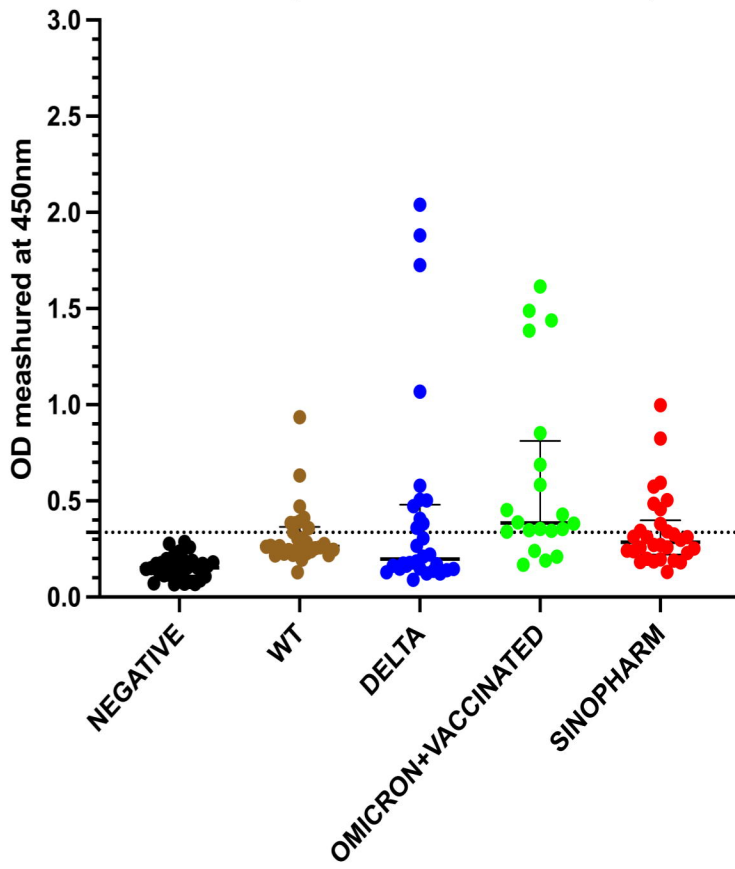
748 IgG antibody responses were measured by an in-house ELISA for individual overlapping
749 peptides of pool 1 (peptide 1 to 15) in omicron + vaccinated (A) and Sinopharm vaccinees
750 (B), pool 2 (peptide 16 to 30) in WT infected (C), and omicron+ vaccinated (D), pool 3
751 (peptide 31 to 45) delta infected (E) and omicron + vaccinated (F) and in pool 4 in delta
752 infected(G) and omicron+ vaccinated (H). 10 individuals were included in each cohort to
753 identify individual antibody responses. The error bars indicate the median and the
754 interquartile ranges.

755

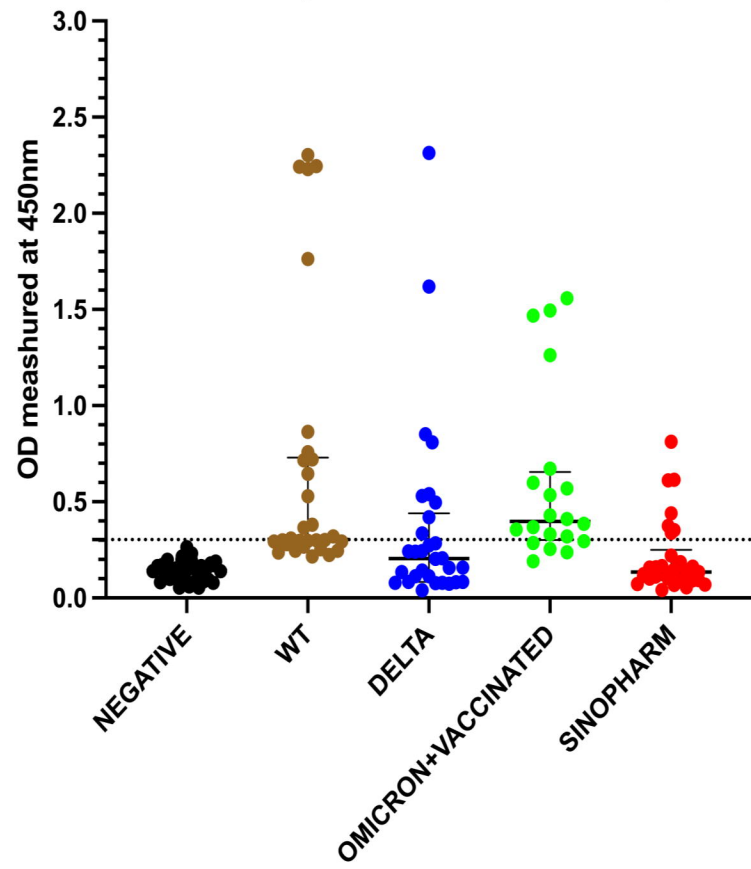
756

757

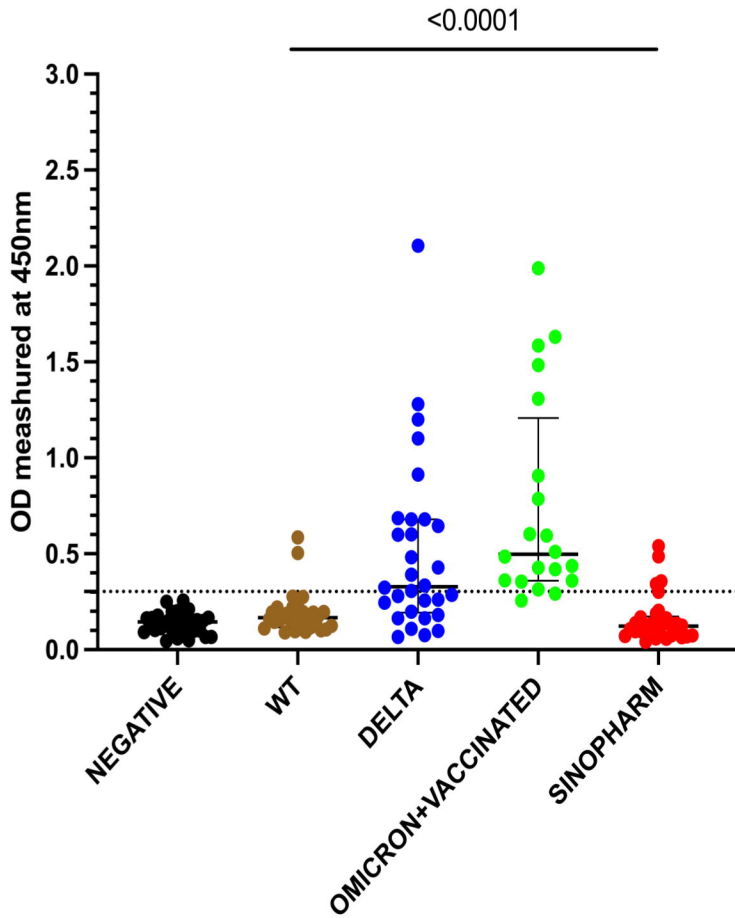
A



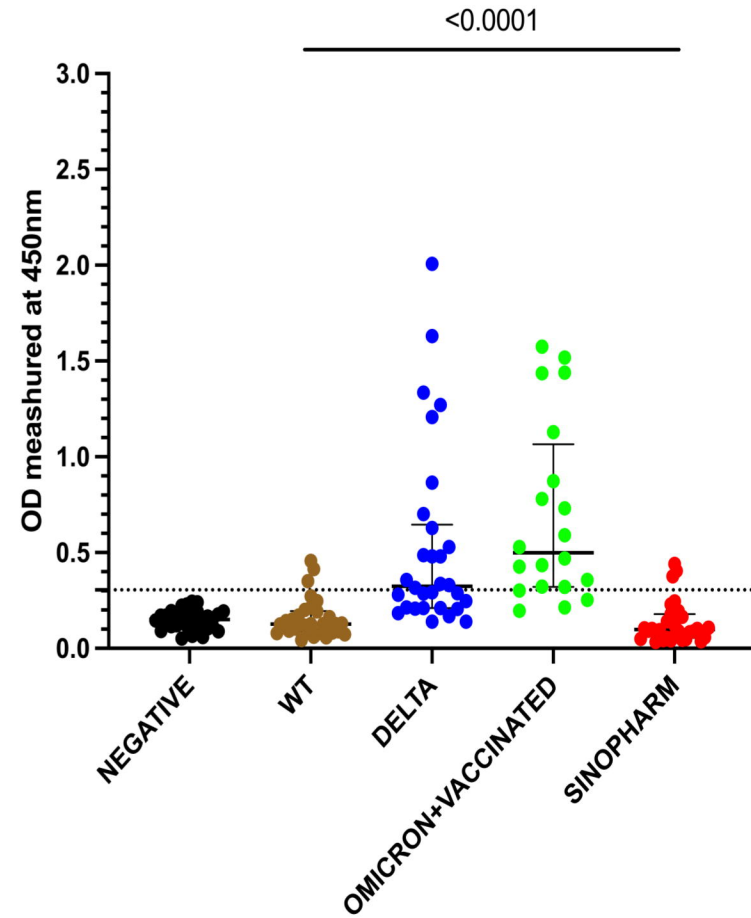
B

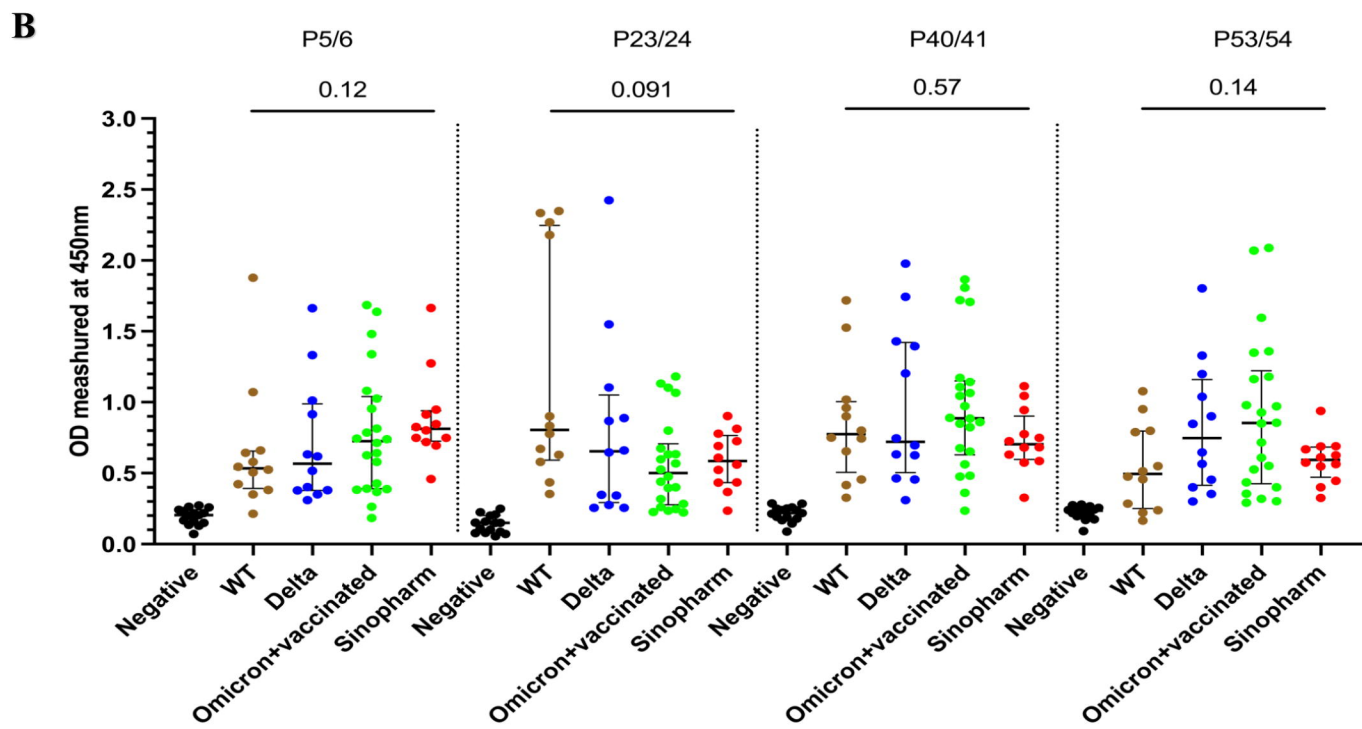
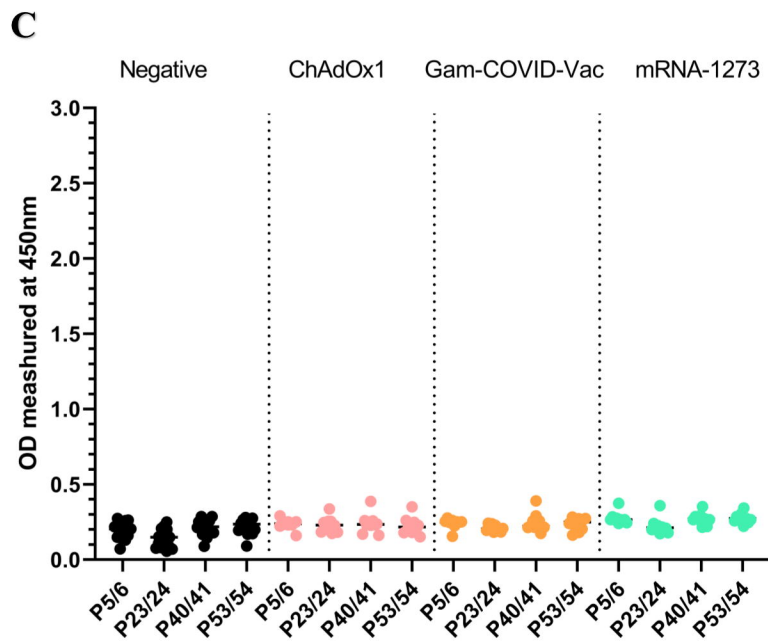
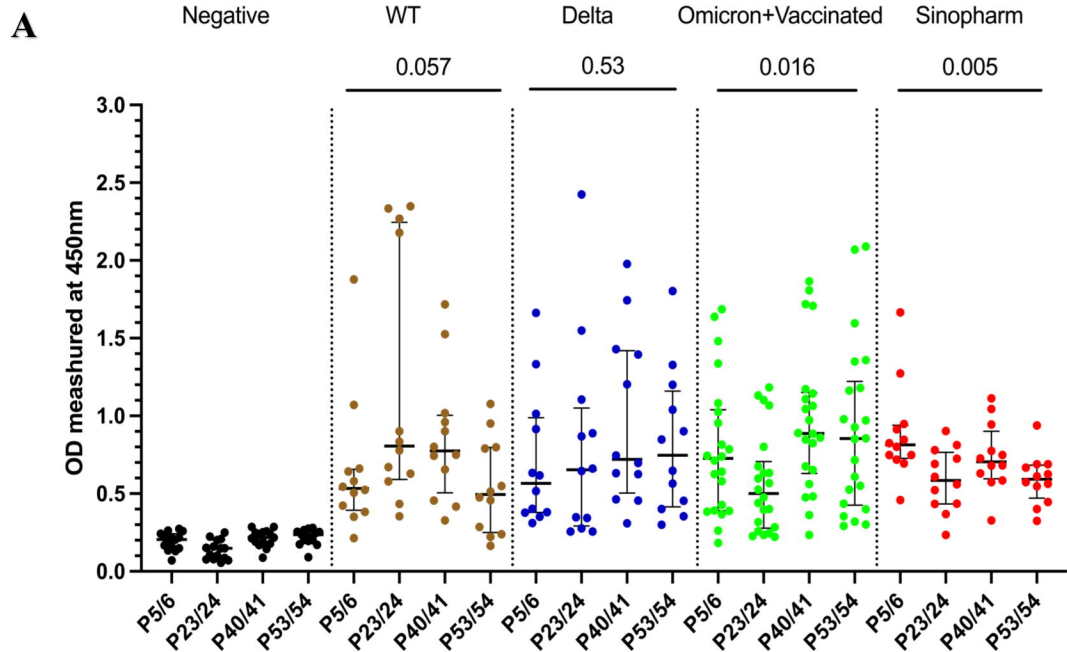


C



D





A

YP_009724397.2_Wild_type/29-52 29 **NGERSGARSKQRRPQGLPNN** **TASW**
 QBP84763.1_OC43/35-65 35 **NVQTRGRRAQPKQTSTSQQP**SGGNVVPYY**SW**
 ABG77571.1_HKU_1/36-64 36 **QTFNRGRKTQPKFTVSTQPQ** . . GNTIPHY**SW**
 YP_003771.1_NL63/7-21 7 **ADDRAARKKFP**PP **SF**

B

YP_009724397.2_Wild_type/155-171 155 **AAIVLQLPQGTTLPKGFYAEGSRG**
 QBP84763.1_OC43/170-192 170 **EAIPTRFNPGTVLNQGIYIEKS** . G
 ABG77571.1_HKU_1/169-191 169 **EAIPTRFNPGTILNQGIYVEKS** . G
 YP_003771.1_NL63/123-146 123 **LEPKFSIALPPELSVVEFEDRS**NN

C

YP_009724397.2_Wild_type/274-300 274 **FGRRGPEQTQGNFGDQELIRQGT**DQKK
 QBP84763.1_OC43/286-309 286 **FGKRGFNI** . . . **NFGGGEM**LKLGTSQGE
 ABG77571.1_HKU_1/284-307 284 **FGKRGFSI** . . . **NFGNAEM**LKLGTNPEL
 YP_003771.1_NL63/250-271 250 **FGPRDFNH** . . . **NMGDS**DLVQNGVDK . .

D

YP_009724397.2_Wild_type/365-381 365 **PTEPKKDKKKKADETQALPQRQKK**
 QBP84763.1_OC43/382-405 382 **QQDGMNMSPK**PQRQRGLKNGQGE
 ABG77571.1_HKU_1/380-403 380 **QNTVSGSLSPK**PQRKRGVKQSP
 YP_003771.1_NL63/338-359 338 **MQSQSSHVAQNTV**LNASIPESK . .

E

medRxiv preprint doi: <https://doi.org/10.1101/2023.01.05.23284247>; this version posted January 7, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY-NC 4.0 International license](https://creativecommons.org/licenses/by-nc/4.0/).

YP_009724397.2_Wild_type/29-52 29 **NGERSGARSKQRRPQGLPNNTASW**
 QVX37034.1_Alpha/29-52 29 **NGERSGARSKQRRPQGLPNNTASW**
 QWW93444.1_Beta/29-52 29 **NGERSGARSKQRRPQGLPNNTASW**
 QXF23757.1_Gamma/29-52 29 **NGERSGARSKQRRPQGLPNNTASW**
 UKA47847.1_Delta/29-52 29 **NGERSGARSKQRRPQGLPNNTASW**
 aicbu4463/2022_BA2/29-49 29 **NG** . . . **GARSKQRRPQGLPNNTASW**
 aicbu4450/2022_BA1/29-49 29 **NG** . . . **GARSKQRRPQGLPNNTASW**
 CDPH-FS27225444/2022_BA5/29-49 29 **NG** . . . **GARSKQRRPQGLPNNTASW**

F

YP_009724397.2_Wild_type/57-80 57 **AAIVLQLPQGTTLPKGFYAEGSRG**
 QVX37034.1_Alpha/57-80 57 **AAIVLQLPQGTTLPKGFYAEGSRG**
 QWW93444.1_Beta/57-80 57 **AAIVLQLPQGTTLPKGFYAEGSRG**
 QXF23757.1_Gamma/57-80 57 **AAIVLQLPQGTTLPKGFYAEGSRG**
 UKA47847.1_Delta/57-80 57 **AAIVLQLPQGTTLPKGFYAEGSRG**
 aicbu4463/2022_BA2/54-77 54 **AAIVLQLPQGTTLPKGFYAEGSRG**
 aicbu4450/2022_BA1/54-77 54 **AAIVLQLPQGTTLPKGFYAEGSRG**
 CDPH-FS27225444/2022_BA5/54-77 54 **AAIVLQLPQGTTLPKGFYAEGSRG**

G

YP_009724397.2_Wild_type/176-199 176 **FGRRGPEQTQGNFGDQELIRQGT**D
 QVX37034.1_Alpha/176-199 176 **FGRRGPEQTQGNFGDQELIRQGT**D
 QWW93444.1_Beta/176-199 176 **FGRRGPEQTQGNFGDQELIRQGT**D
 QXF23757.1_Gamma/176-199 176 **FGRRGPEQTQGNFGDQEL**TRQGT
 UKA47847.1_Delta/176-199 176 **FGRRGPEQTQGNFGDQELIRQGT**D
 aicbu4463/2022_BA2/173-196 173 **FGRRGPEQTQGNFGDQELIRQGT**D
 aicbu4450/2022_BA1/173-196 173 **FGRRGPEQTQGNFGDQELIRQGT**D
 CDPH-FS27225444/2022_BA5/173-196 173 **FGRRGPEQTQGNFGDQELIRQGT**D

H

YP_009724397.2_Wild_type/267-290 267 **PTEPKKDKKKKADETQALPQRQKK**
 QVX37034.1_Alpha/267-290 267 **PTEPKKDKKKKADETQALPQRQKK**
 QWW93444.1_Beta/267-290 267 **PTEPKKDKKKKADETQALPQRQKK**
 QXF23757.1_Gamma/267-290 267 **PTEPKKDKKKKADETQALPQRQKK**
 UKA47847.1_Delta/267-290 267 **PTEPKKDKKKKA**Y**ETQALPQRQKK**
 aicbu4463/2022_BA2/264-287 264 **PTEPKKDKKKKADETQALPQRQKK**
 aicbu4450/2022_BA1/264-287 264 **PTEPKKDKKKKADETQALPQRQKK**
 CDPH-FS27225444/2022_BA5/264-287 264 **PTEPKKDKKKKADETQALPQRQKK**

A

ATO98129.1_RS4081/30-53
AGZ48841.1_WIV1/30-53
QHR63308.1_RaTG13/29-52
ABD75315.1_Rf1/29-52
YP_009724397.2_Wild_Type/29-52

NGGRNGARPKQRRPQGLPNNTASW
NGGRNGARPKQRRPQGLPNNTASW
NGERSGARPKQRRPQGLPNNTASW
DGRSGARPKQRRPQGLPNNTASW
NGERSGARSKQRRPQGLPNNTASW

B

ATO98129.1_RS4081/156-179
AGZ48841.1_WIV1/156-179
QHR63308.1_RaTG13/155-178
ABD75315.1_Rf1/155-178
YP_009724397.2_Wild_Type/155-179

AATVLQLPQGTTLPKGFYAEGSRG
AATVLQLPQGTTLPKGFYAEGSRG
AAIVLQLPQGTTLPKGFYAEGSRG
AAIVLQLPQGTTLPKGFYAEGSRN
AAIVLQLPQGTTLPKGFYAEGSRG

C

ATO98129.1_RS4081/275-298
AGZ48841.1_WIV1/275-298
QHR63308.1_RaTG13/274-297
ABD75315.1_Rf1/274-297
YP_009724397.2_Wild_Type/274-298

FGRRGPEQTQGNFGDQDLIRQGTD
FGRRGPEQTQGNFGDQDLIRQGTD
FGRRGPEQTQGNFGDQELIRQGTD
FGRRGPDQTQGNFGDQELIRQGTD
FGRRGPEQTQGNFGDQELIRQGTD

D

ATO98129.1_RS4081/366-389
AGZ48841.1_WIV1/366-389
QHR63308.1_RaTG13/365-388
ABD75315.1_Rf1/365-388
YP_009724397.2_Wild_Type/365-389

PTEPKKDKKKKTDEAQPLPQRQKK
PTEPKKDKKKKTDEAQPLPQRQKK
PTEPKKDKKKKADETQALPQRQKK
PTEPKKDKKKKTDEAQPLPQRQKK
PTEPKKDKKKKADETQALPQRQKK

