Signatures of omicron-like adaptation in early SARS-CoV-2 variants and chronic infection

Mark Tsz Kin Cheng[1,2**], Mazharul Altaf[1,2**], Jesu Castin[3,4], Ann-Kathrin Reuschl[5], Benjamin L. Sievers[1,2], Kimia Kamelian[1,2], Dejan Mesner[5], Rebecca B. Morse[1,2], Adam Abdullahi[1,2], Bo Meng[1,2], Kata Csiba[1,2], Cambridge NIHR Bioresource, Steven A. Kemp[1,2], Darren Martin[5], Clare Jolly[5], Christopher Ruis[5,6,7], Lipi Thukral[3,4*], Ravindra K. Gupta[1,2,8*]

1. Department of Medicine, University of Cambridge, Cambridge, UK.
2. Cambridge Institute of Therapeutic Immunology & Infectious Disease (CITIID), Cambridge, UK.
3. CSIR-Institute of Genomics and Integrative Biology, Mathura Road, New Delhi-110025, India
4. Academy of Scientific and Innovative Research (AcSIR), Ghaziabad- 201002, India
5. Division of Infection and Immunity, University College London, UK
6. Division of Computational Biology, Institute of Infectious Disease and Molecular Medicine, University of Cape Town
7. Victor Philip Dahdaleh Heart & Lung Research Institute, University of Cambridge, Cambridge, UK
8. Cambridge Centre for AI in Medicine, University of Cambridge, Cambrdge, UK
9. Department of Veterinary Medicine, University of Cambridge, Cambridge, UK
10. Africa Health Research Institute, Durban, KZN, South Africa

*Corresponding authors:

Ravindra K. Gupta rkg20@cam.ac.uk ,

Lipi Thukral lipi.thukral@igib.res.in

Mark Cheng tkmc2@cam.ac.uk

**authors contributed equally.

Lead contact: Ravindra K. Gupta rkg20@cam.ac.uk

SUMMARY

Omicron emergence represented a seismic event in the COVID-19 pandemic, demonstrating what was essentially antigenic shift in a virus that cannot reassort its genome as influenza can. Understanding the success of Omicron is essential, and yet we have little understanding of the biological underpinnings of its ability to accommodate diverse mutations and bring together deleterious mutations to generate a highly successful new serotype. Persistent SARS-CoV-2 infections are a likely source of new variants and may provide valuable insight into past and future evolutionary trajectories, in particular those involving allosteric interactions that defy genotype to phenotype prediction. Here we observe upper airway specific evolution of SARS-CoV-2 demonstrating fusion peptide (FP) domain mutation S:P812S adjacent to the S2' cleavage site that emerged during a chronic infection in an immunocompromised individual. Indeed, this mutation had previously emerged in an ancestral B lineage as well as the delta variant lineage and transmitted successfully in populations globally, though remains uncharacterised. P812S in spike pseudotyped virus particles did not impact entry efficiency across cell lines expressing endogenous ACE2 and TMPRSS2. However, efficiency of spike cleavage at S1/S2 was reduced and molecular dynamics simulation demonstrated altered S1/S2 loop conformations that possibly impacted furin mediated cleavage. Consistent with impaired S1/S2 cleavage, and reminiscent of Omicron BA.1, cell-cell fusogenicity was severely impaired by introduction of P812S. The mutation also introduced significant perturbations to the FP region and also affected protomer-protomer packing. P812S conferred evasion of a neutralising monoclonal antibody targeting the fusion peptide, consistent with significant structural rearrangements in the FP region. Finally, P812S bearing viruses showed evasion of polyclonal neutralising antibodies in sera from vaccinated individuals at 32°C (simulating upper respiratory tract) and to a lesser extent at 37°C. Thus we report a novel mutational adaptation to the upper airway allowing enhanced immune evasion to fusion peptide targeting neutralising antibodies that also incurs a defect in ability to induce syncytia. These data shed light on the balance between upper airway adaptation/immune evasion by SARS-CoV-2, ability to induce syncytia formation, and disease severity given the established link between syncytia and severe COVID-19.

Introduction

Omicron emergence represented a seismic event in the COVID-19 pandemic, demonstrating what was essentially antigenic shift in a virus that cannot reassort its genome as influenza can.

60 Understanding the success of Omicron is essential, and yet we have little understanding of the
61 biological underpinnings of its ability to accommodate diverse mutations and bring together
62 deleterious mutations to generate a highly successful new serotype. A key defining feature of
63 Omicron in addition to immune evasion was the loss of ability of its spike to carry out cell-cell
64 fusion.[1,2] We do not have a clear understanding of which mutations or combinations are responsible
65 given the multiple spike mutations observed across the protein. It is thought that this loss of syncytia
66 formation was associated with the lower pathogenicity of Omicron, and that gradual reacquisition of
67 spike mediated cell-cell fusogenicity by certain omicron subvariants correlates with increased
68 severity.[3]

69 In virus producer cells, full length spike is cleaved by host furin to S1/S2 subunits which then
70 reassociate through electrostatic interactions. The primary receptor for SARS-CoV-2 is ACE2 and
71 following engagement with ACE2, the serine protease TMPRSS2 at the target cell membrane is able
72 to cleave spike at the S2' site, liberating the fusion peptide to initiate viral-cell fusion at the plasma
73 membrane. The Omicron variant appears less dependent on TMPRSS2 and has a preference for
74 endocytic entry where cathepsin carries out S2' cleavage within endosomes.[2] In addition we have
75 observed a cleavage defect in virus producer cells and impaired cell-cell fusion.[2] Given the broad
76 expression of ACE2 across diverse tissues, it is unsurprising that infection of multiple organs occurs
77 during COVID-19 infection. Local tissue environments likely exert specific selective pressure for
78 replication. Viral divergence in different anatomical compartments has been reported, including
79 respiratory compartments such as nasopharynx, sputum, trachea, lungs,[4] and other compartments
80 including plasma, heart, spleen, urine, and gut,[5,6] consistent with differential selection pressure
81 imposed by niches[12–15]. ACE2 and TMPRSS2 expression gradients exist between the upper- (URT) and
82 lower- respiratory tract (LRT), and efficiency of TMPRSS2 use has varied with VOC. Omicron showed
83 adaptation towards the URT[2,11] and greater transmission potential. This tropism shift has been
84 associated with reduced intrinsic disease severity.[12] Affinity towards the URT also leads to
85 differential exposure to environmental and immune selective pressures, creating a unique
86 mutational spectrum which can be leveraged to determine the niche of pathogens.[5,11] Another key
87 difference is the difference in temperature, which has an impact on spike fusogenicity and thus entry
88 efficiency.[13]

89 SARS-CoV-2 adaptation in persistent infection of immunocompromised hosts has been
90 demonstrated in myriad studies.[14–17] COVID-19 remains a particularly relevant threat for
91 immunocompromised individuals, as COVID-19 vaccines have been shown to induce lower
92 seroconversion rates,[18] reduced neutralisation activity,[19] and shorter durations of protection as
93 compared to immune-competent individuals[20]. Here we leverage an intensively sampled case of

chronic infection to report that a single mutation can have diverse phenotypic impacts and recapitulate some of Omicron's key features. P812S in the fusion peptide region and also adjacent to the S2' cleavage site was observed to arise in a sublineage of the Delta variant and in an early B lineage, yet remains uncharacterised[21]. We show that P812S causes changes in the fusion peptide (as one might expect), but also unexpected changes in the S1/S2 cleavage site that impairs cleavage. Moreover, the mutation severely impairs the ability of the spike to mediate cell-cell fusion and thereby we do not see syncytia formation. Syncytia formation is a hallmark of severe COVID-19.[22,23] Importantly and unexpectedly, the mutation also confers a degree of escape from neutralising antibodies directed at the fusion peptide region. This is reflected in the reduced neutralisation of P812S bearing viruses by sera from vaccinated individuals, especially at low temperatures associated with the nasopharynx. Indeed, we first isolated the mutation on an early B lineage (pre-omicron virus) in a chronically infected individual from the upper airway and not in samples from the tracheal region - indicating this mutation may be an adaptation to the upper airway. These properties are reminiscent of Omicron and its adaptation to upper airway transmission with reduced severity correlating with impaired syncytia formation. Our work shows how careful genotypic and phenotypic study of chronic infections can shed light on the evolutionary biology of COVID-19 and .

Results

Clinical course of chronic infection

We previously reported an individual treated with rituximab for a lymphoma, who had a chronic COVID-19 infection with an early Wuhan-Hu-1+D614G SARS-CoV-2 variant (B.1.1.1) between April and August of 2020.[14] The individual was treated with three courses of remdesivir / convalescent plasma and we previously documented immune escape conferred by D796H and viral infectivity enhancement conferred by delH69/V70. We obtained simultaneous samples from nose & throat swabs (NT) and lower endotracheal aspirates (ETA), though previously only data on throat swabs were analysed. ETA samples were only available starting from day 93 because the individual deteriorated and required intensive care admission and ventilatory support due to acute respiratory distress syndrome (ARDS)around day 90. From day 93 until day 101, five NT and four ETA samples were taken in total. The day 100 NT sample was removed from intra-host diversity analysis due to poor Illumina coverage and no available Ct value.

We first examined deep sequence data for the two anatomically proximate compartments of the nasopharynx and trachea. Using Shannon's entropy as a measurement of viral diversity, we observed

126 rises in diversity in NT samples when remdesivir or convalescent plasma were administered as the
127 infection progressed, but no statistically significant difference in genetic diversity within the ETA
128 compared to the NT samples **(Supplementary Figure 1a-d)**. Consistent with previous studies linking
129 Ct values to viral diversity, the increase in viral diversity was accompanied by an increase in viral
130 load.[5] **(Supplementary Figure 1d)**

131

132 **Nucleotide spectral analysis reveals possible mixed upper and lower respiratory tract viral**
133 **populations**

134 The relative frequency of G>T mutations has previously been demonstrated to correlate with SARS-
135 CoV-2 replication niche, with replication in the lung resulting in higher frequencies of G>T
136 mutations.[11] Omicron variants, which are more URT tropic than pre-Omicron variants,[2] have lower
137 G>T frequencies than pre-Omicron variants. **(Supplementary Figure 2a)** We therefore aimed to
138 determine whether there was a difference in G>T frequencies between the NT and ETA isolates from
139 the same patient. As expected given the short distance between these sites, we did not observe a
140 difference between these niches **(Supplementary figure 2b)**; however, we noticed a clade of patient
141 isolates that exhibited elevated numbers of G>T mutations **(Figure 1b,d)**. This clade contains both NT
142 and ETA sequences from late in the course of infection (after day 93), suggesting spread of viruses
143 within the patient across the respiratory tract. The elevation in G>T mutations amongst isolates in
144 this clade suggests that this spread occurred from the LRT to the URT, consistent with deteriorating
145 lung disease. This mixing of viral populations as the patient deteriorated clinically could have been
146 related to the physical conduit provided by the endotracheal tube or air movement associated with
147 mechanical ventilation.

148

149 **Multiple nonsynonymous spike and non-spike mutations arise following therapies for COVID-19**

150 We further examined the changes in mutation frequencies between the NT samples compared to
151 the ETA samples. Although the long branch with S:S13I, S:W64G, and S:P330S emerged, there is
152 limited divergence in variant frequencies between NT and ETA samples on day 93 immediately after
153 remdesivir administration and two units of convalescent plasma infusion. On day 95, divergent
154 mutations arose for Spike (Δ69/70, D796H, P812S) and other genes, including ORF1ab (E1013V),
155 Envelope (T19I), Nucleocapsid (Q28R), and Membrane (H125Y) **(Figure 1a, Supplementary figure 1e,**
156 **Supplementary figure 2c)**. This is likely due to selection pressure imposed by the convalescent
157 plasma and/or remdesivir treatment.[17] Of note, the spike mutation **P812S** emerged on day 101 at a

158 frequency of 88.0% on day 101 in the NT sample, but not in the ETA sample. **P812S** was
159 accompanied by a 98.1% frequency of Δ69/70 and a 89.7% frequency of 796H (Figure 1a,
160 Supplementary figure 1d,3), suggesting linkage of these three mutations on the same viral genomes.

161 At the time of this study (7th April 2024), P812S had a low prevalence in the Global Initiative on
162 Sharing All Influenza Data (GISAID) database (<0.5%, n =12380) **(Table 1, supplementary figure 3)**.
163 P812S was most abundantly found in Delta variant viruses (n = 5744), but was also found in the
164 Alpha variant (n = 1819) and the major Omicron lineages BA.1 (n=610), BA.2 (n=602), BA.4/5
165 (n=1181), and XBB (n=507) viruses **(Table 1, Supplementary Figure 3)**. Though P812S is a minority in
166 these major lineages, over 99.6% of AY.4.10 sublineage sequences (1801 out of 1808) and 99.2% of
167 B.1.463 sublineage sequences (122 out of 123) contained the P812S mutation **(Figure 1c)**, indicating
168 it was a defining mutation in these lineages and arguing for successful transmission of P812S-bearing
169 viruses.[24–26] Indeed, from the start of the pandemic to August 2023, spike codon 812 had an average
170 dN/dS ratio of 2.47 (p=0.184), suggesting positive selection (i.e., selection favouring changes away
171 from the P amino acid state). (https://observablehq.com/@spond/sars-cov-2-global-genomic-selection-2019-
172 aug-2023). Statistical evidence (p<0.01 with the IFEL method) of positive selection at codon 812 was
173 most marked between May to November 2021 when the Delta variant was most prevalent,
174 corroborating the beneficial effect of diverging on a Delta genomic background from P812 to
175 alternative amino acids such as P812S under at least some epidemiological conditions.
176 **(Supplementary figure 4).**

177 We stratified by variant and proportionally subsampled 247 SARS-CoV-2 sequences containing P812S
178 and inferred a maximum-likelihood tree **(Supplementary Figure 3; GISAID EPI_SET_240114wm)**. We
179 annotated the sequences based on the accompanying mutations and noticed that, apart from the
180 D614G mutation that arose early in the pandemic, there were no mutations consistently found to
181 occur with P812S.

182 As P812S was found only in the NT sample, we hypothesised that P812S conferred an advantage in
183 that niche. The proximity to the S2' cleavage site and fusion peptide, in addition to the exposed
184 position of P812 when Spike is in the open position **(PDB: 6XM4, Figure 2a)** prompted us to study
185 the infectivity, cleavage, cell-cell fusion and immune evasion properties of P812S mutants.

186

187 **P812S leads to a general decrease in S1/S2 cleavage and variable cell entry efficiency**

188 As P812S is directly upstream (N-terminal) to the S2' cleavage site (R815/S816) and the S2 viral
189 fusion peptide (FP, 816-855),[27–29] we hypothesised that P812S may directly impact spike glycoprotein

190 S2 cleavage and viral entry. To test this, we introduced the P812S mutation into Wuhan-Hu-1 +
191 D614G (WT), WT with Δ69/70 + 796H mutations, as well as Delta B.1.617.2. We further incorporated
192 Delta and Delta+P812S in the panel because P812S was predominantly found in Delta and previous
193 genomic-epidemiological studies showed P812S[21] further increases the likelihood of Delta
194 breakthrough infections[30,31]. We then performed infectivity assays using pseudoviruses with and
195 without the P812S mutations on cell lines that varied in ACE2 and TMPRSS2 expression levels. We
196 observed little impact of P812S in cells expressing endogenous levels of receptors **(Figure 2b)**.
197 However, P812S was associated with a significant deficit in cell entry efficiency in cell lines where
198 ACE2 or TMPRSS2 were overexpressed **(Figure 2b, Supplementary Figure 5).** P812S-associated entry
199 defects persisted across 32°C, 37°C, and 39°C, a temperature which approximates pyrexia**.**
200 **(Supplementary Figure 6)**.

201 As the cleavage of S1/S2 has been shown to correlate with infectivity/entry/transmission efficiency,
202 to explore this further, we performed western blot for full length (FL) and S2 spike on 293T producer
203 cell lysate and purified virions.[32] We noticed that P812S was associated with suboptimal S1/S2
204 cleavage in both producer cell lysate and virion, demonstrated by decreased ratio of S2 to FL spike
205 (**Figure 2c**). The P812S-induced reduction of S2 in cell lysate and released virions suggests that P812S
206 may play a role in stabilising the S protein and slowing S1 shedding and it appears that cleaved spike
207 is preferentially incorporated into viral particles as we observed previously.[32] A reduction in S1/S2
208 cleavage is reminiscent of the change observed from Delta to Omicron variant[2], which also showed
209 tropism shift away from cells expressing TMPRSS2 and impaired cell-cell fusion in Omicron.[11,33] We
210 next sought to confirm whether P812S confers an entry route preference shift, as seen in the
211 transition of global sequences from the TMPRSS2-reliant plasma membrane entry favoured by the
212 Delta variant to cathepsin-reliant endosomal entry favoured by the Omicron variant (**Table 2**). P812S
213 was associated with a <2-fold resistance to both E64d (cathepsin inhibitor, inhibiting endosomal
214 entry) and camostat (serine protease inhibitor, including TMPRSS2, inhibiting plasma membrane
215 entry). In summary we did not observe substantial shift in efficiency of receptor or coreceptor usage
216 for P812S.

217 GBP5 is an interferon inducible protein that inhibits furin cleavage and was shown to impact HIV Env
218 cleavage by furin[34,35] as well as SARS-CoV-2.[36] Given we observed an S1/S2 cleavage defect for P812S
219 bearing spike, we wished to explore what impact GBP5 might have on furin cleavage of the P812S
220 spike. To test this, we co-expressed HA tagged GBP5 along with spikes bearing a range of mutations
221 and packaging/reporter genome plasmids. We isolated PV and infected target A549-A2T2 lung cells
222 lines. We observed that GBP5 reduced infectivity of each virus tested, and that the effect was

223 smaller for PV bearing P812S, consistent with cleavage that was already impaired at S1/2. Western
224 blot analysis reflected these findings (**Figure 2d**).

225

226 **Significant conformational rearrangements in the fusion peptide and S1/S2 cleavage site regions**
227 **associated with P812S**

228 To probe structural changes in P812S Mutant in SARS-CoV-2 spike (S) protein, we generated two
229 molecular dynamics simulations of glycosylated wildtype and P812S spike variant based on
230 experimentally resolved cryo-EM structure 7A94 with 1RBD up bound to ACE2 (**Supplementary**
231 **Figure 6**). The global arrangement of the SARS-CoV-2 spike (S) protein, which is structurally stabilized
232 through the packing interactions of its three protomers**,** was significantly perturbed upon the
233 introduction of the mutation. The protomer area formed by three CH domains increased from 86.8
234 A^2 to 91.8 A^2 upon the introduction of P812S mutation, leading to altered spatial organization of the
235 protomers (**Supplementary Figure 6b**). Comparative structural analysis revealed increased dynamic
236 fluctuations, with pronounced flexibility observed in the mutant, primarily originating from the NTD,
237 FP and S1-S2 regions (**Supplementary Figure 6c**). Notably, the fusion peptide (FP) region of the SARS-
238 CoV-2 spike (S) protein, a short, hydrophobic segment within the S2 subunit, is positioned just
239 downstream of the S2' cleavage site (R815) and spans residues 816–833. To gain deeper insights into
240 the structural changes surrounding this region, we extended our analysis to encompass residues
241 788–855, allowing us to capture broader conformational dynamics that may impact fusion efficiency
242 and viral entry. Furthermore, the region of S1/S2 proteolytic cleavage site (residues 669–696) also
243 showed greater structural compaction in the P812S mutant compared to the WT**,** suggesting
244 potential alterations in cleavage accessibility and spike activation dynamics. To understand the
245 residues that are responsible for the local perturbation at the FP region (**Supplementary Figure 6**),
246 we monitored the backbone conformational changes and computed rotamer shift with respect to
247 WT (explained in methods). The shift is majorly elevated in the loop flanking FP helix, which is
248 comprising of residues 830 to 855, as shown in Supplementary Figure 6e-f. These structural
249 modelling studies provide insight into our observations of impaired S1/S2 cleavage, consistent with
250 our understanding that proteolytic cleavage at the S2' site may be impacted by P812S mutation.

251

252 **P812S confers evasion of serum neutralising antibodies**

253 Multiple broadly neutralising antibodies have been identified to interact with amino acids on the N-
254 terminal side of the FP (816-655), including P812.[37–39] In addition, P812 has been identified as a

255 highly exposed epitope.[40] Therefore, we hypothesised that P812S may confer a degree of immune
256 escape due to selective pressure imposed by convalescent plasma treatment. We tested the
257 neutralisation sensitivity of Wuhan-Hu-1 + D614G (WT), WT+P812S, WT+Δ69/70+D796H, and WT+
258 Δ69/70+D796H+P812S using a previously reported pseudovirus (PV) system.[41] . We tested 20 stored
259 vaccinee sera (average age 66.0, 55% M) one month after second-dose AstraZeneca ChAdOx1
260 COVID-19 vaccine.[42,43]

261 We confirmed previous observations[14] showing modest immune escape by the double mutant
262 Δ69/70+D796H, with a 1.2-fold decrease in geometric mean serum neutralising titres between WT
263 and WT+Δ69/70+D796H **(Figure 3)**. P812S on a WT+Δ69/70+D796H backbone was associated with a
264 further loss of sensitivity to neutralising serum antibodies **(Figure 3)**. This suggests that the immune
265 evasion capabilities conferred by D796H are enhanced by P812S.

266 To further understand the impact of the P812S mutation, we produced Delta pseudovirions with and
267 without the P812S mutation. Using sera one month after two doses of AstraZeneca ChAdOx1-S, both
268 baseline serum neutralising titres against Delta and Delta+P812S were lower than WT (Table 3).
269 There was a significant decrease in serum neutralisation GMT of the Delta P812S mutant compared
270 to Delta, consistent with immune evasion **(Figure 3)**. In summary, at 37°C, in both the Delta and
271 WT+Δ69/70+D796H backbones, P812S incorporation conferred a degree of resistance to serum
272 neutralisation. To better simulate the upper airway environment, we tested PV expressing
273 WT+Δ69/70+D796H or Delta +/- P812S against sera/monoclonal antibody at 32°C.[44] The serum
274 geometric mean titre for 32°C increased for all viruses by between x1.5 to x2.0 fold compared to
275 37°C. (Table 3) At 32°C, P812S also conferred some resistance to vaccinee sera by x1.4 fold in the
276 background of WT+Δ69/70+D796H and x1.3 fold in the Delta background, respectively **(Figure 3).**

277

278 Finally, we wished to explore the impact of P812S on susceptibility to Sotrovimab treatment, given it
279 is was the only licensed therapeutic monoclonal during the chronic infection case. Sotrovimab
280 (derived from S-309) (GSK) is a monoclonal class 3 antibody which targets a highly conserved
281 epitope within the Spike receptor binding domain (RBD).[45] In agreement with previous literature,
282 Sotrovimab demonstrated similar neutralisation titres with serum against WT and Delta, [46] and
283 P812S had limited effect on sotrovimab susceptibility, with IC50 fold differences less than two fold in
284 the presence versus absence of P812S **(Table 4, Supplementary Figure 7)**.

285

286 **P812S confers evasion of neutralising antibodies targeting the fusion peptide**

287 To gain some understanding of the mechanism of P812S on immune evasion, two monoclonal
288 antibodies were tested. We used fusion-peptide targeting recombinant monoclonal antibodies
289 *fp.006* and *fp.007* from Bianchini and colleagues.[29] In brief, *fp.006* recognises a conserved fusion
290 peptide epitope that is only exposed upon ACE2 binding. IC50 reduction conferred by P812S for WT
291 (x1.22) and Delta (x2.14) is modest for the ACE2-binding dependent fp.006, compared to the IC50
292 increase for Δ69/70+D796H (x8.88). This suggests that in on the background of Δ69/70+D796H the
293 P812S decreased sensitivity of the FP epitope to neutralising antibodies in sera, but not on a WT Wu-
294 1 background.

295 Antibody *fp.007* binding affinity and neutralising capacity is independent of ACE2 binding. *fp.007*
296 demonstrated a significant 5.58- and 2.96-fold decrease in IC50 upon P812S introduction on both
297 WT and Delta backbones, respectively, with little difference upon P812S introduction to the
298 Δ69/70+D796H double mutant. **(Figure 3, Table 3)**. Taken together it appears that FP antibodies are
299 likely to have driven selection of P812S especially as P812S induced significant conformational
300 changes in the FP region.[47] In order to probe this further, we performed MD simulations of the
301 Δ69/70+D796H double mutant with and without P812S (**Figure 4a**). The average inter-protomer
302 distance increased to 15.2 Å in the presence of P812S along with background mutations (**Figure 4b**).
303 Although the perturbations were not as high as observed in P812S without background mutations
304 (**Supplementary Figure 6**), we captured elevated fluctuations at the FP and S1/S2 regions. We
305 computed ΔRMSF ($RMSF_{WT}$-$RMSF_{MUT}$) to assess the changes in the fluctuation patterns upon
306 mutation. In the FP region, the P812S, along with background mutations, introduces local instability
307 with up to 1 Å increase in RMSF (**Figure 4c-d**). We could capture relatively increased fluctuations in
308 the S1/S2 region (up to ~3Å) in P812S + background mutations. While the presence of P812S
309 imparted instability to FP and S1/S2 sites, the lack of P812S (Δ69/70+D796H double mutant) showed
310 no major changes in the fluctuations (**Figure 4e-g**). These observations support the observed escape
311 from FP targeting antibody and the reduced S1/S2 cleavage seen for the triple mutant.

312

313 **P812S severely impairs cell-cell fusion**

314 The ability to induce syncytia via cell-cell fusion is a known property of type I virus glycoproteins,
315 including HIV-1 Env, RSV F and SARS CoV-2 S proteins.[48] This process requires processing of S1/S2 in
316 the producer cell and activity of a serine protease in the target cell to cleave S2' and release the
317 fusion peptide. Syncytia are thought to be a correlate of disease severity in COVID-19.[49] We
318 previously showed that Omicron BA.1 shifted biologically from Delta, demonstrating severely
319 reduced cell-cell fusion activity in vitro, and that this appeared to correlate with S1/S2 cleavage.

320 Subsequent Omicron variants such as BA.4 have regained some fusogenicity.[50] Given the impact of
321 P812S on both S1/2 cleavage and on FP structural dynamics, we measured the impact of P812S on
322 cell-cell fusion activity using a split GFP system. As expected, Delta spike fusion was highest (**Figure**
323 **5**), followed by the double mutant D796H and del69/70 and then WT. Surprisingly, we observed a
324 severe defect induced by the P812S mutation on each background.

325

326 **Discussion**

327 Persistent SARS-CoV-2 infections are a likely source of new variants and may provide valuable insight
328 into past and future evolutionary trajectories, in particular those involving allosteric interactions that
329 defy genotype to phenotype prediction. Here we observe upper airway specific evolution of SARS-
330 CoV-2 demonstrating fusion peptide (FP) domain mutation S:P812S adjacent to the S2' cleavage site
331 that emerged during a chronic infection in an immunocompromised individual. This mutation had
332 also previously emerged in an ancestral B lineage as well as the delta variant lineage and transmitted
333 successfully. P812S in spike pseudotyped virus particles did not impact entry efficiency across cell
334 lines expressing endogenous ACE2 and TMPRSS2. However, efficiency of spike cleavage at S1/S2 was
335 reduced and molecular dynamics simulation demonstrated altered S1/S2 loop conformations that
336 likely impacted furin mediated cleavage. We noted that GBP5, an interferon inducible restriction
337 factor that inhibits furin mediated cleavage, had a smaller impact on inhibition of P812S bearing
338 pseudotyped viruses, consistent with a pre-existing S1/S2 cleavage defect in P812S bearing spikes.

339

340 Consistent with this observation and reminiscent of Omicron BA.1, cell-cell fusogenicity was severely
341 impaired by introduction of P812S. The mutation also introduced significant perturbations to the FP
342 region and protomer-protomer packing. On the other hand, P812S conferred evasion of a
343 neutralising monoclonal antibody targeting the fusion peptide, consistent with significant structural
344 rearrangements in the FP region. As the S2 subunit and FP region are strong candidates for pan-
345 sarbecoviruses broadly neutralising antibody (bnAb) therapy and vaccines,[29,37,38,51] our work shines
346 light on potential functional deficits associated with immune escape mutations.

347

348 Finally, P812S bearing viruses showed evasion of polyclonal neutralising antibodies in sera from
349 vaccinated individuals at 32°C (simulating upper respiratory tract) and to a lesser extent at 37°C. We
350 propose that the selection of P812S in the upper airway compartment is an immune evasion

351 adaptation that came at the cost of cell-cell fusion and entry efficiency in some cells overexpressing
352 ACE2/TMPRSS2, possibly mediated through impaired S1/2 cleavage in P812S-bearing spikes.

353

354 Interestingly, this URT adaptation occurred on the background of within-host evolution suggestive of
355 adaptation to the LRT based on mutational spectra analysis. Thus, it is plausible that the LRT
356 population migrated to the URT given the intense inflammation in the lower airway in the present
357 case. Although we were unable to separate URT and LRT isolates in mutational spectra due to the
358 similarity of sequences at the consensus level from our patient, comparison of patient sequences
359 and global sequences of pre-Omicron and Omicron lineages is suggestive of progression from lower
360 to upper airway. Our work demonstrates the importance of studying compartmentalisation of viral
361 populations, especially in persistent infection. Monitoring these mutations that have epistatic
362 effects could potentially assist in prediction of mutations in future circulating strains. The driving
363 force for mutations in the immunocompromised host is significantly different and allows for unique
364 combinations, especially given suboptimal antibody-driven selection pressure.

365

366 It is interesting and notable that a single mutation can have multiple unexpected phenotypic
367 impacts. These observations provide clues as to the events leading to the emergence of the highly
368 successful Omicron variant, highlighting the role of allosteric impacts of immune escape mutations –
369 for example on cleavage sites - and the importance of mutational context. Indeed, the dogma that
370 viruses become less pathogenic over time is widely quoted, and yet the evidence for this is weak. If
371 syncytia formation indeed is a correlate of disease severity, our data provide one route as to how
372 such a process might occur for SARS-CoV-2, although ultimately the 'holy grail' is a step by step
373 understanding of how the Omicron variant selected individual mutations.

374

375 **Limitations of the study**

376 Our study was based on lentiviruses pseudotyped with different SARS-CoV-2 spikes. Although
377 consistency has been observed between pseudotyped virus system and the live-virus system for the
378 study of virus entry,[42,52] there may still be potential effectors only present in the live-virus system.[2]
379 Confirmation of our finding in a live-virus system would be desirable.

380

381 **Acknowledgement**

382 This work was funded by a Wellcome Senior Fellowship to RKG (WT108082AIA). This research was
383 supported by the NIHR Cambridge Biomedical Research Centre (NIHR203312). A-KR and CJ are
384 funded by a Wellcome Investigator award to CJ (223065). LT is thankful to the funding support from
385 the DBT/Wellcome Trust India Alliance (IA/21/2/505925). The views expressed are those of the
386 authors and not necessarily those of the NIHR or the Department of Health and Social Care. We
387 thank the valuable discussion with Dr Maximillian Muenchhoff on SARS-CoV-2 HLA and CD8 T cell
388 epitope evasion analysis. We also thank Dr Filippo Bianchini, Ms Jasmine Chantal Cantergiani, and Dr
389 Davide F. Robianni for sending samples and expression plasmids of fp.006 and fp.007 antibodies.

390

391 **Author Contributions**

392 Conceptualization, M.T.K.C. and R.K.G.;

393 Methodology, M.T.K.C., D.M.,C.R.,B.M. M.A;

394 Investigation, M.T.K.C., A.M., D.M., C.R., B.L.S., K.K., R.B.M., M.A., K.C. and A-K.R.

395 Writing – Original Draft, M.T.K.C.;

396 Writing – Review & Editing, M.T.K.C, M.A, D.M., C.R., R.K.G. A-K.R. and C.J.

397 Funding Acquisition, R.K.G;

398 Resources, R.K.G;

399 Supervision, B.M., S.A.K., R.K.G., C.J.

400 **Declaration of interests**

401 The authors declare no competing interests.

402

403 **Methods**

404 **Clinical-sample collection and sequencing**

405 Sequencing of SARS-CoV-2 viral DNA extracted from serial samples from the LRT (sputum and
406 endotracheal aspirate), and URT (throat and nasal swab) was described previously in Kemp et al.[14] In
407 short, nucleic acid extraction was done from 500 µL sample using the easyMAG platform
408 (Biomerieux). Sequencing amplicons were prepared with the ARTICnetwork v.3 protocol and diluted

409 to 2 ng μl^{-1} and 25 µl (50 ng) were used as input for each library preparation reaction, before
410 sequencing using the Illumina MiSeq platform.

411 Sera from recovered patients in the COVIDx study were used for testing of neutralization activity by
412 SARS-CoV-2 mutants.

413 **Sequence data processing and analysis**

414 Illumina sequencing data from Kemp et al. was re-analysed.[14] Bam files were downloaded from
415 Kemp et al. (BioProject PRJNA682013), mapped to Wuhan-Hu-1 (NC_045512.2) by minimap v2.28
416 (arguments -ax sr),[53] sorted and indexed by samtools v1.20,[54] and variants were called by BCFtools
417 v1.18 functions mpileup (minimum base quality q30), and call (arguments --multiallelic-caller --
418 variants-only).[55] Mutations identified by BCFtools were filtered by a minimum depth cutoff of 20x
419 before annotations by snpEff v5.2 using the publicly available NC_045512.2 database.[56]

420 **Genetic diversity and viral load analysis**

421 The genetic diversity of each sample was estimated using Shannon Entropy based on the frequency
422 of each iSNV, assuming that all iSNVs are independent from each other.

$$H = -\sum_{i}^{n} P(i) log_e P(i)$$

423 Where *P(i)* is the iSNV percentage frequency at gene loci *i*.

424 Shannon Entropy of NT samples collected prior to day 93, NT samples collected on day 93 and after,
425 and ETA samples collected on day 93 and after were compared using student's t-test. For Ct values,
426 Welch's t-test was performed.

427 **Phylogenetic analysis and lineage assignment**

428 Excluding duplicates, and low coverage sequences, we searched GISAID for all complete coverage
429 SARS-CoV-2 sequences with the Spike_P812S substitution on the 7th April 2024. All sequence
430 encoding Spike-S812 were aligned to MN908947.3 by mafft (v7.525, --6merpair --keeplength --
431 addfragments), and then were assigned lineages using Pangolin (V4.3).[57] The major lineages were
432 grouped by Scorpio (V0.3.19). We stratified the sequences based on their major lineage designations
433 and then performed stratified random sampling with proportional allocation to major lineages to
434 obtain a subset of 250 P812S-containing sequences using the R package Rsample v1.2.1.

435 We constructed a maximum-likelihood tree from patient X1 and the 250 sequence sample using IQ-
436 Tree v2.2.5 employing the GTR+F+I model of nucleotide substitution (model with the lowest

437 Bayesian Information Criterion identified by IQ-Tree ModelFinder) with clustering support assessed
438 by 1000 rounds of ultrafast bootstrapping.[58]

439 **Signals of natural selection in the SARS-CoV-2 S-gene**

440 Mutational sites found in the patient, the Alpha variant, and the Delta variant, together with sites of
441 neutralisation escape mutations were chosen for natural selection analysis using the internal-branch
442 fixed effects (IEFL) method implemented in HYPHY.[59] All globally sampled near-full length SARS-CoV-
443 2 sequences are sampled between December 2019 and March 2022 were binned in three-monthly
444 intervals (labelled according to the last month of sampling within a bin) and analysed separately
445 within the context of the entire Spike open reading frame: an analysis setup needed for both
446 phylogenetic context and for informing the nucleotide and codon substitution models upon which
447 the IEFL method relies. Only codon substitutions mapped to internal phylogenetic tree branches
448 were considered in these analyses since mutations mapping to terminal branches would have likely
449 experienced insufficient durations of selection to yield meaningful evolutionary information. The
450 approximate dominant-variant era during the COVID-19 pandemic where indicated based on results
451 from Amodio and colleagues.[60]

452 **Structure Analysis**

453 Coordinates for the SARS-CoV-2 Spike trimer complex with two RBD down and one RBD up was from
454 PDB: 6XM4-Chain A to C, and domains definition from multiple references were used.[29,61–63] Graphics
455 describing the structures were made in UCSF Chimera X V1.8.[64]

456 **Mutational Spectrum analysis**

457 We calculated the mutational spectrum of mutations within each patient sample using the mutation
458 files generated above. The korimuto pipeline within MutTui v2.0.2[65] was run on each sample
459 independently using NC_045512.2 as the reference genome. The proportion of G>T mutations
460 within each sample was calculated as the number of G>T mutations divided by the total number of
461 mutations.

462 **Generation of P812S Spike mutants**

463 Amino acid substitutions were introduced into a codon-optimised pCDNA3.1_SARS-CoV-2 WT,
464 Δ69/70, Delta **(B.1.617.2, EPI_ISL_2378732)**, and Omicron (**EPI_ISL_7418017**) Spike Plasmid as
465 previously described,[2,14,31] using the QuikChange Lightning Site-Directed mutagenesis kit, as per
466 manufacturer's protocol (Agilent Technologies).

467

468 **Molecular dynamics simulations**

469 The spike protein in '1-RBD-up' state in complex with ACE-2 was considered for molecular dynamics
470 simulation. The D614G spike was considered as the wild type (WT) protein. Along with D614 spike,
471 three different mutant systems were prepared – i) WT+ P812S, ii) WT+D796H+del 69/70 and iii)
472 WT+D796H+del 69/70+P812S. The spike protein was glycosylated as given in Casalino et al., 2020.
473 The systems were solvated with TIP3P water and further neutralized with the counter ions. All the
474 simulations were carried out with Gromacs version 2021.3 incorporating the CHARMM36 all-atom
475 force field (Abraham et al., 2015). A grid spacing of 0.16 nm was used along with the fourth-order
476 cubic interpolation for Particle Mesh Ewald (PME) summation. The van der Waals cut-off was set to
477 1.2 nm. The systems were initially subjected to energy minimization with steepest descent method.
478 Temperature and pressure were maintained at 310 K and 1 bar using V-rescale thermostat and
479 Parrinello-Rahman barostat, respectively. The numerical integration of equation of motion was
480 computed every 2 fs and the coordinates of the system were saved 100 ps. Each system was
481 simulated for 1 μs. The structures were visualized and rendered with PyMOL (Schrödinger, LLC,
482 2015). All basic distance and RMSD/RMSF calculations were performed in Python with MDAnalysis
483 library.

484

485 **Computing backbone rotamer shift with respect to WT**

486 To track the residues that has changed their backbone dihedral conformations, we computed a
487 comparative metric called rotamer shift. We discretized the observed backbone dihedral angle into
488 *cis* and *trans* states as given in Singh et al., 2017. We quantify the rotamer shift for the residue
489 $R_i$with the following equation.

$$Rotamer\ Shift_{R_{i\ WT\ vs\ MUT}} = \frac{\sum_{r=0}^{N=n(r)} |f(r)_{WT} - f(r)_{MUT}|}{N}$$

490 The r denotes the rotamer state of the dihedral. For backbone dihedrals, the rotamer states include
491 Cis and Trans. Thus, N is the total number of rotamers for a dihedral. Each of the rotamer fractions
492 f(r) (fraction of a rotamer r of a residue throughout the trajectory) is computed for WT and the
493 mutant and subtracted. The rotamer shift is then defined as the average of absolute value of these
494 differences in the rotamer fractions. The resulting shift varies from 0 (no changes in rotamer
495 fractions) to 1 (complete change in rotamer fractions). We computed rotamer shift in ψ and ϕ
496 angles for each residue. The maximum of two was finally considered to represent the dihedral
497 change occurred at a residue. The dihedral angles were calculated with the gromacs command *gmx*

498 *rama*. The rotamer assignment and the shift calculations were performed with the in-house Python
499 script.

500

501 **Pseudovirus preparation**

502 Pseudovirus were prepared by transfecting HEK293Tv cells with 1µg of p8.91 HIV-1 gag-pol
503 expression vector, 1.5 µg of pCSFLW, and 1µg of plasmid expressing the respective spike protein
504 using a 3:1 ratio of Fugene HD transfection agent (Promega).[41,66] After 48 hours of incubation at 10%
505 CO_2, 37°C, viral supernatant was collected and filtered through a .45 µm filter and stored at -80°C
506 until further use. The 50% tissue-culture infectious dose ($TCID_{50}$) of SARS-CoV-2 pseudovirus was
507 determined using the Steady-Glo Luciferase assay system (Promega).

508 For GBP5 sensitivity assays, pseudoviruses were prepared as described previously.[36] In brief,
509 HEK293T cells were transfected the indicated spike plasmids, p.8.91 and pCSLW in the presence of
510 40ng GBP5 vector[34,35] or a corresponding empty vector control using Fugene6 (Promega). After 48-
511 72h, culture supernatants were collected and 0.45 µm filtered.

512

513 **Standardization of Virus Input by reverse transcriptase activity using SYBR Green-Based Product**
514 **Enhanced RT-PCR Assay (SG-PERT)**

515 The reverse transcriptase activity of pseudovirus preparations was determined by quantitative
516 polymerase chain reaction (qPCR) using a SYBR green-based product-enhanced PCR assay (SG-PERT)
517 as previously described.[67] Briefly, undiluted supernatants were lysed in a 1:1 ratio in a 2x lysis
518 solution (40% glycerol v/v, 0.25% Trition X-100 v/v, 50 mM KCl, Ribolock RNase inhibitor 0.8 U/mL,
519 TrisHCl 100mM, buffered to pH 7.4) for 10 minutes at room temperature.

520 **Western blot analysis**

521 For virus pellets, the filtered viral supernatant was centrifuged at 15,000 rpm for 120 min to pellet
522 virions. For cell lysates, transfected HEK293Tv cells were resuspended in PBS, before centrifugation
523 at 1500 rpm for 5 minutes. The cell pellet was then lysed in 400 µL cell lysis buffer (Cell Signaling
524 Technology Cat#9803) and sonicated for 1 minute.

525 All samples are further lysed and reduced in Laemmli reducing buffer (1M Tris-HCl (pH6.8), SDS,
526 100% glycerol, β-mercaptoethanol and bromophenol blue). This was followed by electrophoresis on
527 SDS 4-12% Bis-Tris protein gels (Thermo Fisher Scientific) under reducing conditions. Proteins on the
528 gel are transferred by electroblotting onto polyvinylidene difluoride (PVDF) membranes. The SARS-

529 CoV-2 spike proteins and pseudotype virus capsid were visualized using a ChemiDoc MP imaging
530 system (Biorad) or Odyssey Cxl Infrared Imager (Licor) using a rabbit anti-Spike polyclonal (Thermo
531 Fisher Scientific, Cat# PA1-41165, RRID:AB_1087210, LOT:WE3286564, 1 in 4000), mouse anti-p24
532 monoclonal (BEI Resources; NIAID Cat# ARP-3537, RRID:AB_3086785, 1 in 3000), mouse anti-Spike
533 monoclonal (GeneTex 1A9, 1:1000), mouse anti-HA-tag (Biolegend, 16B12 1:2000), rabbit anti-p24
534 (CFAR ARP432, 1:2000) and mouse anti-alpha-tubulin (Sigma DM1A, 1:1000).

535 Chemiluminescence was detected using Bio-Rad Clarity Western ECL Substrate (Cat#1705061), or
536 Cytiva Amersham ECL Prime (Cat#RPN2232), and the band intensity is quantified using Fiji Software
537 (ImageJ v.153k).

538 **EC50 of entry inhibitors in A549-ACE2-TMPRSS2 cells with PV**

539 A549-ACE2/TMPRSS2 (A549-A2T2) cells were treated with either E64d (Tocris) or camostat (Sigma-
540 Aldrich) for 2 h at each drug concentration before the addition of a comparable amount of input
541 pseudotyped viruses (approximately 500,000 RLU). The cells were then left for 48 h before addition
542 of lysis buffer + luciferase substrate (Promega) and read on a Glomax plate reader (Promega).
543 Normalisation was performed using an inhouse script, and EC_{50} values were calculated using a
544 nonlinear regression model of inhibitor dilution vs. normalized response with a variable slope in
545 GraphPad Prism 10.

546 **Vaccinee serum and monoclonal antibody pseudotype virus (PV) neutralization assay**

547 Previous work have shown that IC_{50} obtained from spike pseudotype assays quantitatively correlate
548 with neutralisation activity measured using an authentic SARS-CoV-2 neutralization assay.[52]

549 Virus neutralisation assays were conducted on HeLa cells stably expressing ACE2 (HeLa-A2) using
550 SARS-CoV-2 spike pseudotyped virus expressing firefly luciferase as previously described.[14]

551 Sera were first heat-inactivated at 55°C for an hour. Serial dilutions of sera or purified monoclonal
552 antibodies are then incubated with pseudotyped virus for 1h at 37°C. Freshly trypsinized and
553 quenched HEK293T-A2T2 or HeLa-A2 cells were subsequently added to each well. After 48h
554 incubation at 37°C and 5% CO_2, luminescence was measured using a Bright-Glo luciferase assay
555 system (Promega) and neutralisation was calculated relative to cell-only and cell-and-virus-only
556 controls. Normalisation was performed using an inhouse script, and IC_{50} values were calculated using
557 a nonlinear regression model of inhibitor dilution vs. normalized response with a variable slope in
558 GraphPad Prism 10.

559

Table

Table 1: Lineage of P812S-containing sequences on GISAID

scorpio_call	Aggregate Group	frequency	percentage
Alpha (B.1.1.7-like)	Alpha (B.1.1.7-like)	1818	14.68%
B.1.1.7-like+E484K	Alpha (B.1.1.7-like)	1	0.01%
Delta (AY.4-like)	Delta (AY.4-like)	2903	23.45%
Delta (AY.4.2-like)	Delta (AY.4-like)	21	0.17%
Delta (B.1.617.2-like)	Delta (B.1.617.2-like)	2820	22.78%
Omicron (BA.1-like)	Omicron (BA.1-like)	610	4.93%
Omicron (BA.2-like)	Omicron (BA.2-like)	602	4.86%
Omicron (BA.5-like)	Omicron (BA.4/5-like)	1127	9.10%
Omicron (BA.4-like)	Omicron (BA.4/5-like)	54	0.44%
Omicron (XBB.1-like)	Omicron (XBB.x-like)	210	1.70%
Omicron (XBB.1.5-like)	Omicron (XBB.x-like)	162	1.31%
Omicron (XBB-like)	Omicron (XBB.x-like)	109	0.88%
Omicron (XBB.1.16-like)	Omicron (XBB.x-like)	26	0.21%
Gamma (P.1-like)	Other	167	1.35%
A.23.1-like	Other	61	0.49%
Iota (B.1.526-like)	Other	20	0.16%
Epsilon (B.1.429-like)	Other	18	0.15%
Beta (B.1.351-like)	Other	16	0.13%
Epsilon (B.1.427-like)	Other	7	0.06%
Mu (B.1.621-like)	Other	7	0.06%
Lambda (C.37-like)	Other	6	0.05%
B.1.617.1-like	Other	4	0.03%
Eta (B.1.525-like)	Other	1	0.01%

Theta (P.3-like)	Other	1	0.01%
Zeta (P.2-like)	Other	1	0.01%
Unassigned	Unassigned	1590	12.84%
Probable Omicron (Unassigned)	Unassigned	10	0.08%
Omicron (Unassigned)	Unassigned	8	0.06%
	Total	12380	

All SARS-CoV-2 isolates bearing P812S as of 7th April 2024, aggregated by Scorpio assigned lineage.

Table 2 – half maximal effective concentration (EC50) of camostat and E64d

	EC50 (µM)	
Spike presented on pseudotyped Virus	**camostat**	**E64d**
WT	16.85	1.99
WT+P812S	23.58	3.24
Δ69/70+D796H	17.69	1.79
Δ69/70+D796H+P812S	23.76	4.50
Delta	19.45	2.11
Delta + P812S	23.59	3.87

Table 3 – Geometric mean half-maximal neutralising titre (GMT) of AZD1222 one month post second dose vaccinee sera of P812S mutants

	GMT	
Spike presented on pseudotyped Virus	**32 °C**	**37°C**
WT	N/A	339.3
WT+P812S	N/A	347.1
Δ69/70+D796H	603.2	293.2
Δ69/70+D796H+P812S	179.8	221.7
Delta	429.8	199.9
Delta + P812S	116.1	132.3

Geometric mean half-maximal neutralising titre (GMT) of AZD1222 vaccinee sera (one month after two doses) at 32°C and 37°C against WT, WT+P812S, Δ69/70+D796H, Δ69/70+D796H+P812S, Delta, and Detla+P812S.

Table 4 – half-maximal inhibitory concentration (IC50) of SARS-CoV-2 Spike targeting monoclonal antibodies

	IC50 (ng/mL)			
Spike presented on pseudotyped Virus	**Sotrovimab (32°C)**	**Sotrovimab (37°C)**	**fp.006**	**fp.007**
WT	NA	46.8	1173	953
WT+P812S	NA	31.03	960	170.8
Δ69/70+D796H	37.30	45.38	127.8	218.3
Δ69/70+D796H+P812S	14.79	55.04	1135	288.4
Delta	8.60	45.09	18286	17195
Delta + P812S	9.71	79.56	8545	5818

IC50 of Sotrovimab fp.006, fp.007 against WT, WT+P812S, Δ69/70+D796H, Δ69/70+D796H+P812S, Delta, and Detla+P812S.

583 **Citations**

584 1. Willett, B.J., Grove, J., MacLean, O.A., Wilkie, C., De Lorenzo, G., Furnon, W., Cantoni, D.,
585 Scott, S., Logan, N., Ashraf, S., et al. (2022). SARS-CoV-2 Omicron is an immune escape variant with
586 an altered cell entry pathway. Nat Microbiol *7*, 1161–1179. https://doi.org/10.1038/s41564-022-
587 01143-7.

588 2. Meng, B., Abdullahi, A., Ferreira, I.A.T.M., Goonawardane, N., Saito, A., Kimura, I.,
589 Yamasoba, D., Gerber, P.P., Fatihi, S., Rathore, S., et al. (2022). Altered TMPRSS2 usage by SARS-CoV-
590 2 Omicron impacts infectivity and fusogenicity. Nature *603*, 706–714.
591 https://doi.org/10.1038/s41586-022-04474-x.

592 3. Tamura, T., Ito, J., Uriu, K., Zahradnik, J., Kida, I., Anraku, Y., Nasser, H., Shofa, M., Oda, Y.,
593 Lytras, S., et al. (2023). Virological characteristics of the SARS-CoV-2 XBB variant derived from
594 recombination of two Omicron subvariants. Nat Commun *14*, 2800. https://doi.org/10.1038/s41467-
595 023-38435-3.

596 4. El Moussaoui, M., Bontems, S., Meex, C., Hayette, M.-P., Lejeune, M., Hong, S.L., Dellicour,
597 S., Moutschen, M., Cambisano, N., Renotte, N., et al. (2024). Intrahost evolution leading to distinct
598 lineages in the upper and lower respiratory tracts during SARS-CoV-2 prolonged infection. Preprint,
599 https://doi.org/10.1101/2024.03.03.24303157 https://doi.org/10.1101/2024.03.03.24303157.

600 5. Wang, Y., Wang, D., Zhang, L., Sun, W., Zhang, Z., Chen, W., Zhu, A., Huang, Y., Xiao, F., Yao,
601 J., et al. (2021). Intra-host variation and evolutionary dynamics of SARS-CoV-2 populations in COVID-
602 19 patients. Genome Medicine *13*, 30. https://doi.org/10.1186/s13073-021-00847-5.

603 6. Van Cleemput, J., van Snippenberg, W., Lambrechts, L., Dendooven, A., D'Onofrio, V., Couck,
604 L., Trypsteen, W., Vanrusselt, J., Theuns, S., Vereecke, N., et al. (2021). Organ-specific genome
605 diversity of replication-competent SARS-CoV-2. Nature Communications *12*, 6612.
606 https://doi.org/10.1038/s41467-021-26884-7.

607 7. Buckland, M.S., Galloway, J.B., Fhogartaigh, C.N., Meredith, L., Provine, N.M., Bloor, S.,
608 Ogbe, A., Zelek, W.M., Smielewska, A., Yakovleva, A., et al. (2020). Treatment of COVID-19 with
609 remdesivir in the absence of humoral immunity: a case report. Nature Communications *11*, 6385.
610 https://doi.org/10.1038/s41467-020-19761-2.

611 8. Voloch, C.M., da Silva Francisco Jr, R., de Almeida, L.G.P., Brustolini, O.J., Cardoso, C.C.,
612 Gerber, A.L., Guimarães, A.P. de C., Leitão, I. de C., Mariani, D., Ota, V.A., et al. (2021). Intra-host
613 evolution during SARS-CoV-2 prolonged infection. Virus Evolution *7*, veab078.
614 https://doi.org/10.1093/ve/veab078.

615 9. Al Khatib, H.A., Benslimane, F.M., Elbashir, I.E., Coyle, P., Al Maslamani, M., Al-Khal, A., Al
616 Thani, A.A., and Yassine, H.M. (2020). Within-Host Diversity of SARS-CoV-2 in COVID-19 Patients
617 With Variable Disease Severities. Frontiers in cellular and infection microbiology *10*, 575613.
618 https://doi.org/10.3389/fcimb.2020.575613.

619 10. Rueca, M., Bartolini, B., Gruber, C.E., Piralla, A., Baldanti, F., Giombini, E., Messina, F.,
620 Marchioni, L., Ippolito, G., Di Caro, A., et al. (2020). Compartmentalized Replication of SARS-Cov-2 in
621 Upper vs. Lower Respiratory Tract Assessed by Whole Genome Quasispecies Analysis.
622 Microorganisms *8*. https://doi.org/10.3390/microorganisms8091302.

11. Ruis, C., Peacock, T.P., Polo, L.M., Masone, D., Alvarez, M.S., Hinrichs, A.S., Turakhia, Y., Cheng, Y., McBroome, J., Corbett-Detig, R., et al. (2023). A lung-specific mutational signature enables inference of viral and bacterial respiratory niche. Microbial Genomics *9*. https://doi.org/10.1099/mgen.0.001018.

12. Wolter, N., Jassat, W., Walaza, S., Welch, R., Moultrie, H., Groome, M., Amoako, D.G., Everatt, J., Bhiman, J.N., Scheepers, C., et al. (2022). Early assessment of the clinical severity of the SARS-CoV-2 omicron variant in South Africa: a data linkage study. The Lancet *399*, 437–446. https://doi.org/10.1016/S0140-6736(22)00017-4.

13. Dufloo, J., and Sanjuán, R. (2024). Temperature impacts SARS-CoV-2 spike fusogenicity and evolution. mBio *15*, e03360-23. https://doi.org/10.1128/mbio.03360-23.

14. Kemp, S.A., Collier, D.A., Datir, R.P., Ferreira, I.A.T.M., Gayed, S., Jahun, A., Hosmillo, M., Rees-Spear, C., Mlcochova, P., Lumb, I.U., et al. (2021). SARS-CoV-2 evolution during treatment of chronic infection. Nature *592*, 277–282. https://doi.org/10.1038/s41586-021-03291-y.

15. Harari, S., Tahor, M., Rutsinsky, N., Meijer, S., Miller, D., Henig, O., Halutz, O., Levytskyi, K., Ben-Ami, R., Adler, A., et al. (2022). Drivers of adaptive evolution during chronic SARS-CoV-2 infections. Nature medicine *28*, 1501–1508. https://doi.org/10.1038/s41591-022-01882-4.

16. Weigang, S., Fuchs, J., Zimmer, G., Schnepf, D., Kern, L., Beer, J., Luxenburger, H., Ankerhold, J., Falcone, V., Kemming, J., et al. (2021). Within-host evolution of SARS-CoV-2 in an immunosuppressed COVID-19 patient as a source of immune escape variants. Nature Communications *12*, 6405–6405. https://doi.org/10.1038/s41467-021-26602-3.

17. Choi, B., Choudhary, M.C., Regan, J., Sparks, J.A., Padera, R.F., Qiu, X., Solomon, I.H., Kuo, H.H., Boucau, J., Bowman, K., et al. (2020). Persistence and Evolution of SARS-CoV-2 in an Immunocompromised Host. The New England journal of medicine *383*, 2291–2293. https://doi.org/10.1056/nejmc2031364.

18. Tartof, S.Y., Slezak, J.M., Puzniak, L., Hong, V., Xie, F., Ackerson, B.K., Valluri, S.R., Jodar, L., and McLaughlin, J.M. (2022). Immunocompromise and durability of BNT162b2 vaccine against severe outcomes due to omicron and delta variants. The Lancet Respiratory Medicine *10*, e61–e62. https://doi.org/10.1016/S2213-2600(22)00170-9.

19. Tartof, S.Y., Slezak, J.M., Puzniak, L., Hong, V., Frankland, T.B., Ackerson, B.K., Takhar, H.S., Ogun, O.A., Simmons, S.R., Zamparo, J.M., et al. (2022). Effectiveness of a third dose of BNT162b2 mRNA COVID-19 vaccine in a large US health system: A retrospective cohort study. The Lancet Regional Health - Americas *9*, 100198. https://doi.org/10.1016/j.lana.2022.100198.

20. Ainsley Ryan Yan Bin Lee, Shi Yin Wong, Louis Yi Ann Chai, Soo Chin Lee, Matilda Xinwei Lee, Mark Dhinesh Muthiah, Sen Hee Tay, Chong Boon Teo, Benjamin Kye Jyn Tan, Yiong Huak Chan, et al. (2022). Efficacy of covid-19 vaccines in immunocompromised patients: systematic review and meta-analysis. BMJ *376*, e068632. https://doi.org/10.1136/bmj-2021-068632.

21. Fibke, C.D., Joffres, Y., Tyson, J.R., Colijn, C., Janjua, N.Z., Fjell, C., Prystajecky, N., Jassem, A., and Sbihi, H. (2022). Spike Mutation Profiles Associated With SARS-CoV-2 Breakthrough Infections in Delta Emerging and Predominant Time Periods in British Columbia, Canada. Frontiers in public health *10*, 915363. https://doi.org/10.3389/fpubh.2022.915363.

22. Bussani, R., Schneider, E., Zentilin, L., Collesi, C., Ali, H., Braga, L., Volpe, M.C., Colliva, A., Zanconati, F., Berlot, G., et al. (2020). Persistence of viral RNA, pneumocyte syncytia and thrombosis

665 are hallmarks of advanced COVID-19 pathology. EBioMedicine *61*, 103104.
666 https://doi.org/10.1016/j.ebiom.2020.103104.

667 23. Braga, L., Ali, H., Secco, I., Chiavacci, E., Neves, G., Goldhill, D., Penn, R., Jimenez-Guardeño,
668 J.M., Ortega-Prieto, A.M., Bussani, R., et al. (2021). Drugs that inhibit TMEM16 proteins block SARS-
669 CoV-2 spike-induced syncytia. Nature *594*, 88–93. https://doi.org/10.1038/s41586-021-03491-6.

670 24. Gangavarapu, K., Latif, A.A., Mullen, J.L., Alkuzweny, M., Hufbauer, E., Tsueng, G., Haag, E.,
671 Zeller, M., Aceves, C.M., Zaiets, K., et al. (2022). Outbreak.info genomic reports: scalable and
672 dynamic surveillance of SARS-CoV-2 variants and mutations. medRxiv, 2022.01.27.22269965.
673 https://doi.org/10.1101/2022.01.27.22269965.

674 25. Shruti Khare, C.G. (2021). GISAID's Role in Pandemic Response. China CDC Weekly *3*, 1049.
675 https://doi.org/10.46234/ccdcw2021.255.

676 26. Karthik Gangavarapu, A.A.L. (2023). S:P812S Mutation Report Outbreak.info.

677 27. Benton, D.J., Wrobel, A.G., Xu, P., Roustan, C., Martin, S.R., Rosenthal, P.B., Skehel, J.J., and
678 Gamblin, S.J. (2020). Receptor binding and priming of the spike protein of SARS-CoV-2 for membrane
679 fusion. Nature *588*, 327–330. https://doi.org/10.1038/s41586-020-2772-0.

680 28. Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S.,
681 Schiergens, T.S., Herrler, G., Wu, N.-H., Nitsche, A., et al. (2020). SARS-CoV-2 Cell Entry Depends on
682 ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell *181*, 271-280.e8.
683 https://doi.org/10.1016/j.cell.2020.02.052.

684 29. Bianchini, F., Crivelli, V., Abernathy, M.E., Guerra, C., Palus, M., Muri, J., Marcotte, H., Piralla,
685 A., Pedotti, M., De Gasparo, R., et al. Human neutralizing antibodies to cold linear epitopes and
686 subdomain 1 of the SARS-CoV-2 spike glycoprotein. Science Immunology *0*, eade0958.
687 https://doi.org/10.1126/sciimmunol.ade0958.

688 30. Kemp, S.A., Cheng, M.T.K., Hamilton, W.L., Kamelian, K., Chauhan, H., Dikid, T., Gogia, H.,
689 Lall, H., Ponnusamy, K., Verma, K., et al. (2022). Transmission of B.1.617.2 Delta variant between
690 vaccinated healthcare workers. Scientific reports *12*, 10492. https://doi.org/10.1038/s41598-022-
691 14411-7.

692 31. Mlcochova, P., Kemp, S.A., Dhar, M.S., Papa, G., Meng, B., Ferreira, I.A.T.M., Datir, R., Collier,
693 D.A., Albecka, A., Singh, S., et al. (2021). SARS-CoV-2 B.1.617.2 Delta variant replication and immune
694 evasion. Nature *599*, 114–119. https://doi.org/10.1038/s41586-021-03944-y.

695 32. Meng, B., Kemp, S.A., Papa, G., Datir, R., Ferreira, I.A.T.M., Marelli, S., Harvey, W.T., Lytras,
696 S., Mohamed, A., Gallo, G., et al. (2021). Recurrent emergence of SARS-CoV-2 spike deletion
697 H69/V70 and its role in the Alpha variant B.1.1.7. Cell Reports *35*.
698 https://doi.org/10.1016/j.celrep.2021.109292.

699 33. Hui, K.P.Y., Ng, K.-C., Ho, J.C.W., Yeung, H.-W., Ching, R.H.H., Gu, H., Chung, J.C.K., Chow,
700 V.L.Y., Sit, K.-Y., Hsin, M.K.Y., et al. (2022). Replication of SARS-CoV-2 Omicron BA.2 variant in ex vivo
701 cultures of the human upper and lower respiratory tract. eBioMedicine *83*, 104232.
702 https://doi.org/10.1016/j.ebiom.2022.104232.

703 34. Krapp, C., Hotter, D., Gawanbacht, A., McLaren, P.J., Kluge, S.F., Stürzel, C.M., Mack, K.,
704 Reith, E., Engelhart, S., Ciuffi, A., et al. (2016). Guanylate Binding Protein (GBP) 5 Is an Interferon-

705 Inducible Inhibitor of HIV-1 Infectivity. Cell Host Microbe *19*, 504–514.
706 https://doi.org/10.1016/j.chom.2016.02.019.

707 35. Braun, E., Hotter, D., Koepke, L., Zech, F., Groß, R., Sparrer, K.M.J., Müller, J.A., Pfaller, C.K.,
708 Heusinger, E., Wombacher, R., et al. (2019). Guanylate-Binding Proteins 2 and 5 Exert Broad Antiviral
709 Activity by Inhibiting Furin-Mediated Processing of Viral Envelope Proteins. Cell Rep *27*, 2092-
710 2104.e10. https://doi.org/10.1016/j.celrep.2019.04.063.

711 36. Mesner, D., Reuschl, A.-K., Whelan, M.V.X., Bronzovich, T., Haider, T., Thorne, L.G., Ragazzini,
712 R., Bonfanti, P., Towers, G.J., and Jolly, C. (2023). SARS-CoV-2 evolution influences GBP and IFITM
713 sensitivity. Proc Natl Acad Sci U S A *120*, e2212577120. https://doi.org/10.1073/pnas.2212577120.

714 37. Dacon, C., Tucker, C., Peng, L., Lee, C.-C.D., Lin, T.-H., Yuan, M., Cong, Y., Wang, L., Purser, L.,
715 Williams, J.K., et al. (2022). Broadly neutralizing antibodies target the coronavirus fusion peptide.
716 Science *377*, 728–735. https://doi.org/10.1126/science.abq3773.

717 38. Low, J.S., Jerak, J., Tortorici, M.A., McCallum, M., Pinto, D., Cassotta, A., Foglierini, M., Mele,
718 F., Abdelnabi, R., Weynand, B., et al. (2022). ACE2-binding exposes the SARS-CoV-2 fusion peptide to
719 broadly neutralizing coronavirus antibodies. Science *377*, 735–742.
720 https://doi.org/10.1126/science.abq2679.

721 39. Chen, Y., Zhao, X., Zhou, H., Zhu, H., Jiang, S., and Wang, P. (2023). Broadly neutralizing
722 antibodies to SARS-CoV-2 and other human coronaviruses. Nature Reviews Immunology *23*, 189–
723 199. https://doi.org/10.1038/s41577-022-00784-3.

724 40. Harvey, W.T., Carabelli, A.M., Jackson, B., Gupta, R.K., Thomson, E.C., Harrison, E.M.,
725 Ludden, C., Reeve, R., Rambaut, A., COVID-19 Genomics UK (COG-UK) Consortium, et al. (2021).
726 SARS-CoV-2 variants, spike mutations and immune escape. Nat Rev Microbiol *19*, 409–424.
727 https://doi.org/10.1038/s41579-021-00573-0.

728 41. Naldini, L., Blömer, U., Gage, F.H., Trono, D., and Verma, I.M. (1996). Efficient transfer,
729 integration, and sustained long-term expression of the transgene in adult rat brains injected with a
730 lentiviral vector. Proceedings of the National Academy of Sciences *93*, 11382–11388.
731 https://doi.org/10.1073/pnas.93.21.11382.

732 42. Collier, D.A., Ferreira, I.A.T.M., Kotagiri, P., Datir, R.P., Lim, E.Y., Touizer, E., Meng, B.,
733 Abdullahi, A., Baker, S., Dougan, G., et al. (2021). Age-related immune response heterogeneity to
734 SARS-CoV-2 vaccine BNT162b2. Nature *596*, 417–422. https://doi.org/10.1038/s41586-021-03739-1.

735 43. Collier, D.A., De Marco, A., Ferreira, I.A.T.M., Meng, B., Datir, R.P., Walls, A.C., Kemp, S.A.,
736 Bassi, J., Pinto, D., Silacci-Fregni, C., et al. (2021). Sensitivity of SARS-CoV-2 B.1.1.7 to mRNA vaccine-
737 elicited antibodies. Nature *593*, 136–141. https://doi.org/10.1038/s41586-021-03412-7.

738 44. Lindemann, J., Leiacker, R., Rettinger, G., and Keck, T. (2002). Nasal mucosal temperature
739 during respiration. Clin Otolaryngol *27*, 135–139. https://doi.org/10.1046/j.1365-2273.2002.00544.x.

740 45. Barnes, C.O., Jette, C.A., Abernathy, M.E., Dam, K.-M.A., Esswein, S.R., Gristick, H.B.,
741 Malyutin, A.G., Sharaf, N.G., Huey-Tubman, K.E., Lee, Y.E., et al. (2020). SARS-CoV-2 neutralizing
742 antibody structures inform therapeutic strategies. Nature *588*, 682–687.
743 https://doi.org/10.1038/s41586-020-2852-1.

744 46. Cox, M., Peacock, T.P., Harvey, W.T., Hughes, J., Wright, D.W., Willett, B.J., Thomson, E.,
745 Gupta, R.K., Peacock, S.J., Robertson, D.L., et al. (2023). SARS-CoV-2 variant evasion of monoclonal

746 antibodies based on in vitro studies. Nature Reviews Microbiology *21*, 112–124.
747 https://doi.org/10.1038/s41579-022-00809-7.

748 47. Meng, B., Datir, R., Choi, J., Bradley, J.R., Smith, K.G.C., Lee, J.H., Gupta, R.K., Baker, S.,
749 Dougan, G., Hess, C., et al. (2022). SARS-CoV-2 spike N-terminal domain modulates TMPRSS2-
750 dependent viral entry and fusogenicity. Cell Reports *40*, 111220.
751 https://doi.org/10.1016/j.celrep.2022.111220.

752 48. Li, X., Yuan, H., Li, X., and Wang, H. (2023). Spike protein mediated membrane fusion during
753 SARS-CoV-2 infection. J Med Virol *95*, e28212. https://doi.org/10.1002/jmv.28212.

754 49. Rajah, M.M., Bernier, A., Buchrieser, J., and Schwartz, O. (2022). The Mechanism and
755 Consequences of SARS-CoV-2 Spike-Mediated Fusion and Syncytia Formation. J Mol Biol *434*,
756 167280. https://doi.org/10.1016/j.jmb.2021.167280.

757 50. Planas, D., Staropoli, I., Michel, V., Lemoine, F., Donati, F., Prot, M., Porrot, F., Guivel-
758 Benhassine, F., Jeyarajah, B., Brisebarre, A., et al. (2024). Distinct evolution of SARS-CoV-2 Omicron
759 XBB and BA.2.86/JN.1 lineages combining increased fitness and antibody evasion. Nat Commun *15*,
760 2254. https://doi.org/10.1038/s41467-024-46490-7.

761 51. Ng, K.W., Faulkner, N., Finsterbusch, K., Wu, M., Harvey, R., Hussain, S., Greco, M., Liu, Y.,
762 Kjaer, S., Swanton, C., et al. (2022). SARS-CoV-2 S2–targeted vaccination elicits broadly neutralizing
763 antibodies. Sci. Transl. Med. *14*, eabn3715. https://doi.org/10.1126/scitranslmed.abn3715.

764 52. Schmidt, F., Weisblum, Y., Muecksch, F., Hoffmann, H.-H., Michailidis, E., Lorenzi, J.C.C.,
765 Mendoza, P., Rutkowska, M., Bednarski, E., Gaebler, C., et al. (2020). Measuring SARS-CoV-2
766 neutralizing antibody activity using pseudotyped and chimeric viruses. J Exp Med *217*, e20201181.
767 https://doi.org/10.1084/jem.20201181.

768 53. Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics *34*,
769 3094–3100. https://doi.org/10.1093/bioinformatics/bty191.

770 54. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G.,
771 and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–
772 2079. https://doi.org/10.1093/bioinformatics/btp352.

773 55. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A.,
774 Keane, T., McCarthy, S.A., Davies, R.M., et al. (2021). Twelve years of SAMtools and BCFtools.
775 GigaScience *10*, giab008. https://doi.org/10.1093/gigascience/giab008.

776 56. Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and
777 Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide
778 polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3.
779 Fly (Austin) *6*, 80–92. https://doi.org/10.4161/fly.19695.

780 57. O'Toole, Á., Scher, E., Underwood, A., Jackson, B., Hill, V., McCrone, J.T., Colquhoun, R., Ruis,
781 C., Abu-Dahab, K., Taylor, B., et al. (2021). Assignment of epidemiological lineages in an emerging
782 pandemic using the pangolin tool. Virus Evolution *7*. https://doi.org/10.1093/ve/veab064.

783 58. Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A.,
784 and Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in
785 the Genomic Era. Molecular Biology and Evolution *37*, 1530–1534.
786 https://doi.org/10.1093/molbev/msaa015.

787 59. Kosakovsky Pond, S.L., and Frost, S.D. (2005). Not so different after all: a comparison of
788 methods for detecting amino acid sites under selection. Mol Biol Evol *22*, 1208–1222.
789 https://doi.org/10.1093/molbev/msi105.

790 60. Amodio, E., Genovese, D., Fallucca, A., Ferro, P., Sparacia, B., D'Azzo, L., Fertitta, A., Maida,
791 C.M., and Vitale, F. (2022). Clinical Severity in Different Waves of SARS-CoV-2 Infection in Sicily: A
792 Model of Smith's "Law of Declining Virulence" from Real-World Data. Viruses *15*.
793 https://doi.org/10.3390/v15010125.

794 61. Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S., Zhang, Q., Shi, X., Wang, Q., Zhang, L., et al.
795 (2020). Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor.
796 Nature *581*, 215–220. https://doi.org/10.1038/s41586-020-2180-5.

797 62. Kadam, S.B., Sukhramani, G.S., Bishnoi, P., Pable, A.A., and Barvkar, V.T. (2021). SARS-CoV-2,
798 the pandemic coronavirus: Molecular and structural insights. J Basic Microbiol *61*, 180–202.
799 https://doi.org/10.1002/jobm.202000537.

800 63. Zhang, J., Xiao, T., Cai, Y., and Chen, B. (2021). Structure of SARS-CoV-2 spike protein. Curr
801 Opin Virol *50*, 173–182. https://doi.org/10.1016/j.coviro.2021.08.010.

802 64. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H.,
803 and Ferrin, T.E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and
804 developers. Protein Sci *30*, 70–82. https://doi.org/10.1002/pro.3943.

805 65. Ruis, C., Tonkin-Hill, G., Floto, R.A., and Parkhill, J. (2023). Calculating and applying pathogen
806 mutational spectra using MutTui. bioRxiv, 2023.06.15.545111.
807 https://doi.org/10.1101/2023.06.15.545111.

808 66. Gupta, R.K., Kohli, A., McCormick, A.L., Towers, G.J., Pillay, D., and Parry, C.M. (2010). Full-
809 length HIV-1 Gag determines protease inhibitor susceptibility within in-vitro assays. AIDS *24*, 1651–
810 1655. https://doi.org/10.1097/QAD.0b013e3283398216.

811 67. Vermeire, J., Naessens, E., Vanderstraeten, H., Landi, A., Iannucci, V., Van Nuffel, A., Taghon,
812 T., Pizzato, M., and Verhasselt, B. (2012). Quantification of Reverse Transcriptase Activity by Real-
813 Time PCR as a Fast and Accurate Method for Titration of HIV, Lenti- and Retroviral Vectors. PloS one
814 *7*, e50859. https://doi.org/10.1371/journal.pone.0050859.

815

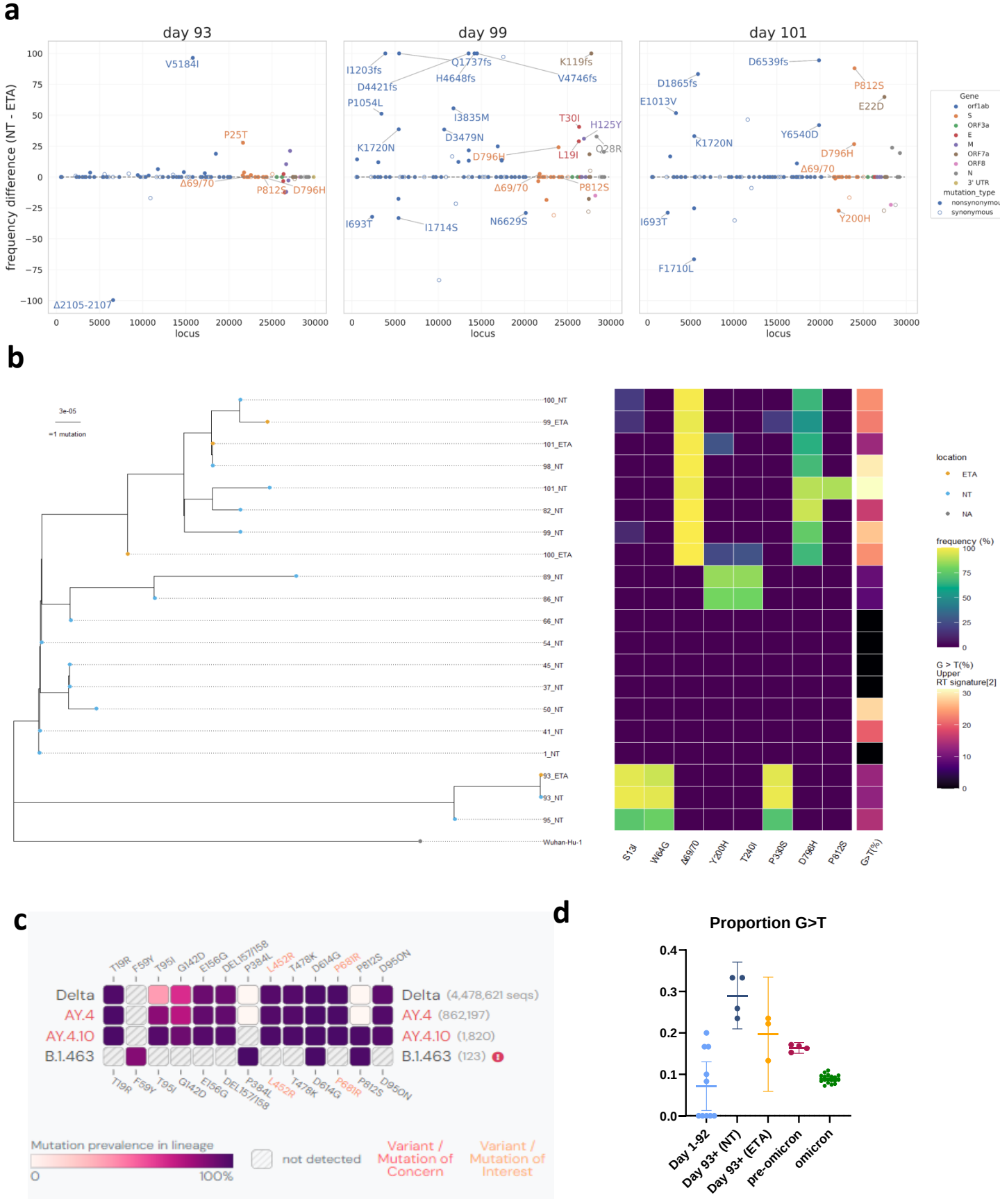


Figure 1: Genetic changes over time and across niches during chronic COVID-19. a) The absolute difference in percentage prevalence of mutations in patient X1 plotted against the SARS-CoV-2 genomic map. Nonsynonymous mutations (filled) with >20% difference are annotated. P812S is found on day 101 NT compartment. b) Maximum-likelihood phylogenetic tree of patient X1 with the day of sampling and location (blue=nose and throat swabs, orange= endotracheal aspirate) indicated. Middle panel shows the Spike mutation frequency foundin each sample, and right panel the proportion of G to T mutation in each sample. c) Global prevalence data taken from Outbreak.info on 7th April 2024. left: Global prevalence of S:P812S per PANGO lineage. Right: Spike mutations which occurs at greater than 75% in one of Delta (B.1.617.2),AY.4, AY.4.10, B.1463. d) Proportion of G to T mutations (a lower airway mutation signature) in patient between day 1 to 92, day 93 onwardsfrom NT samples, day 93 onwards from ETA samples, and pre-omicron and omicron variants. Day 93 and Day 95 were excluded as they are separated from the main clade. The data for omicron and pre-omicron variants are aggregated data from available sequences of the UShER SARS-CoV-2 phylogenetic tree. Data are plotted with mean and 95% CI.

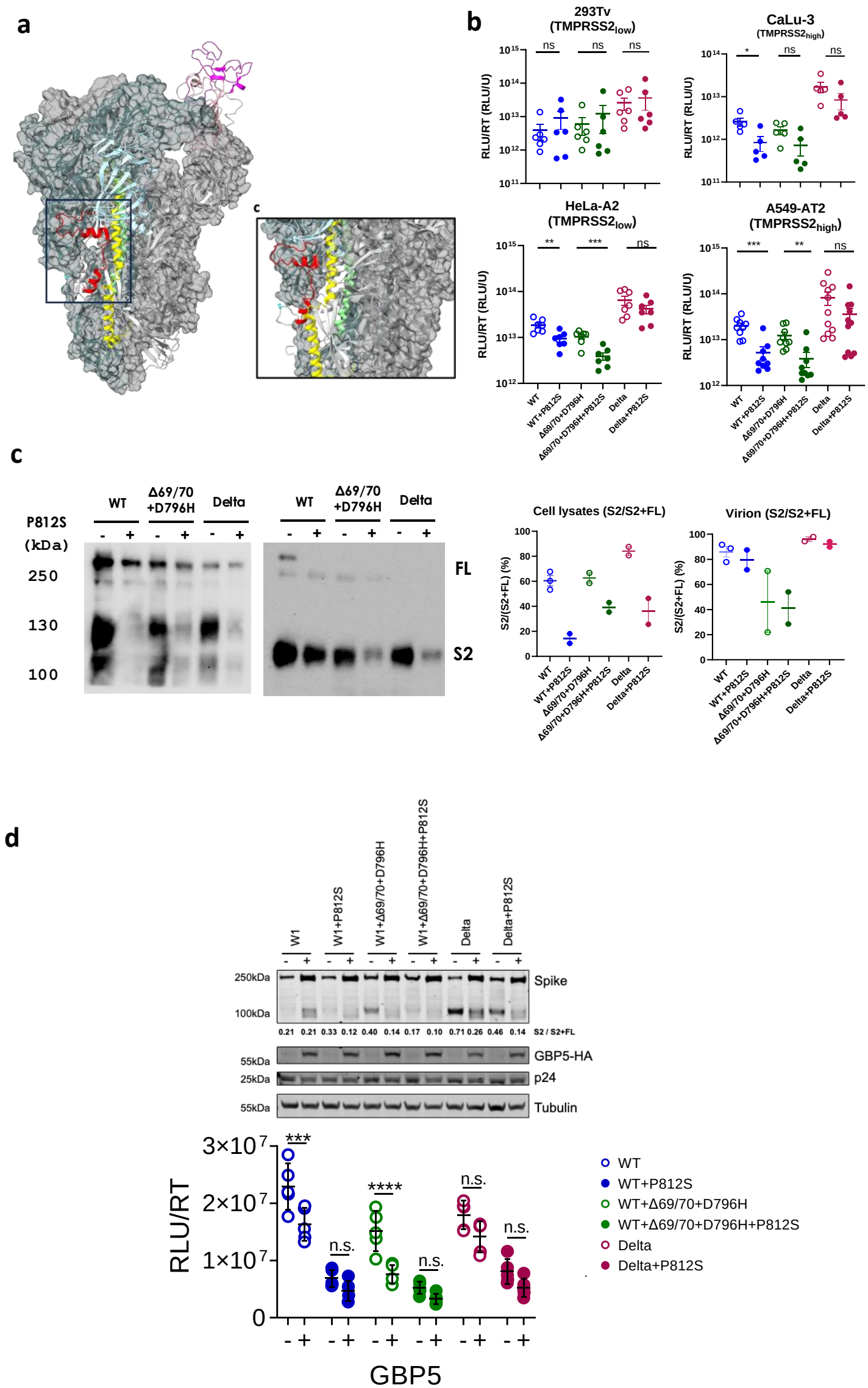


Figure 2: P812S bearing spikes maintain infectivity in cell lines expressing endogenous levels of ACE2/TMPRSS2 but show impaired S1S2 cleavage and reduced sensitivity to GBP5 a) Structure of SARS-CoV-2 Spike (PDB: 6XM4) where two subunits are in the "RBD down" conformation (gray surfaces) and one subunit is in the "RBD up" conformation (coloured ribbon). Domains are coloured as follows: NTD (light blue), RBD (pink), RBM (magenta), P812 (cyan), Fusion peptide (red), HR1 (yellow), HR2 (green). b) PV entry of WT (Wuhan-Hu-1+D614G), WT+P812S, WT+Δ69/70+796H, WT+Δ69/70+796H+P812S, Delta, Delta+P812S introduced to ACE2-high, TMPRSS2-low 293Tv-A2; Hela-A2; and ACE2-high, TMPRSS2-high A549-AT2, Calu-3 lung cells at 37°C. Data are supplemented by mean ± standard error of n = 3–8 technical replicates. Statistical analysis was performed using unpaired student's t-test. c) Western blot analysis of spike cleavage in producer HEK293Tv cell lysate and isolated pseudovirus. S2 to FL spike ratio were analysed by densitometry (ImageJ), and S2/S2+FL cleavage proportion was calculated for virions and cell lysates. d) GBP sensitivity of Spike-Gag-PVs (produced in the presence or absence of GBP5 on ACE2/TMPRSS2-A549. Top panel shows western blot of transfected producer cell lysates probed with antibodies to S2, HA, p24 and tubulin. Bottom panel shows PV infectivity (RLU/ RT U) and is representative of two independent experiments with 5-6 measurements each Statistical analysis was performed using one way ANOVA with Bonferroni correction.

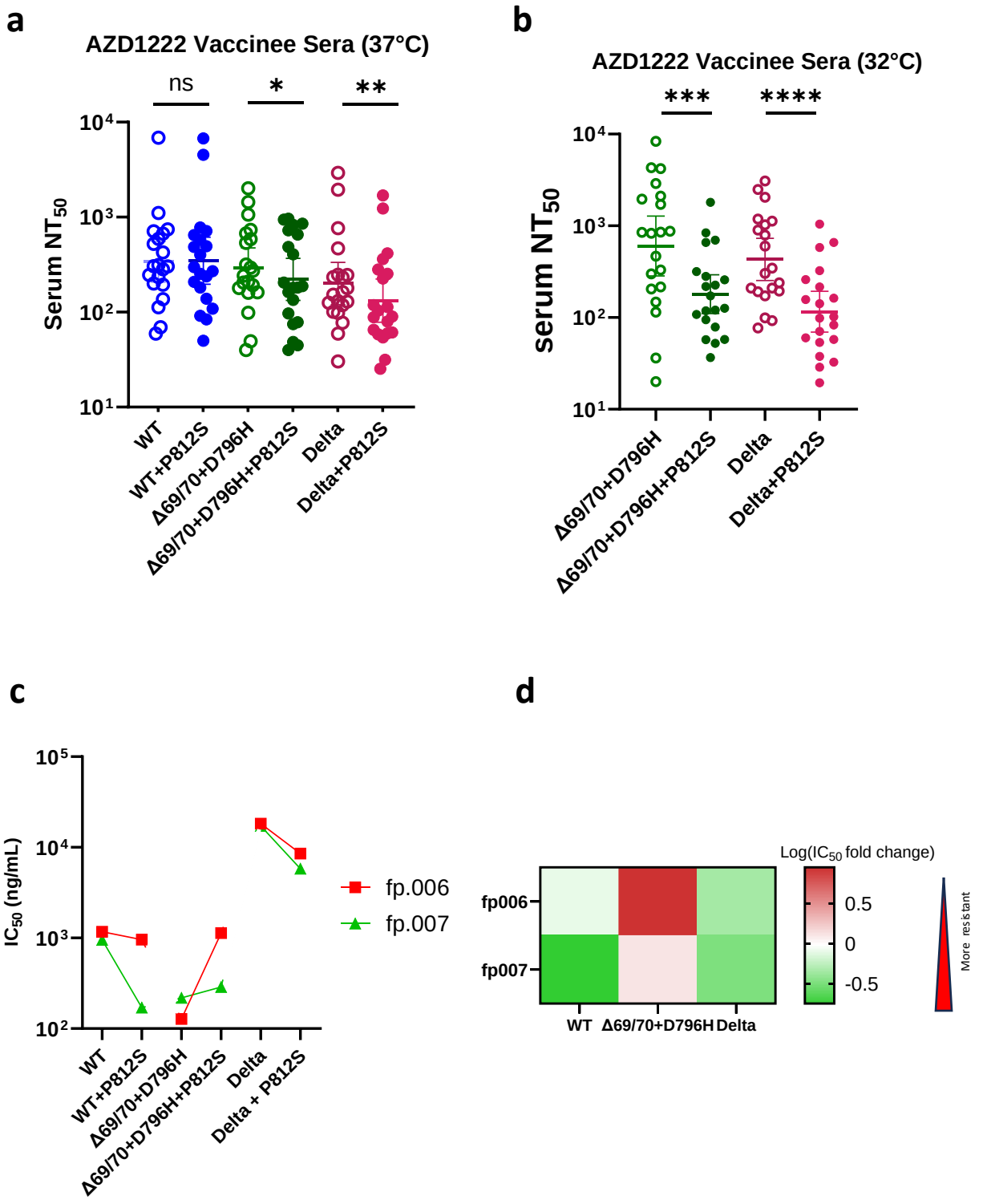


Figure 3: P812S confers escape from polyclonal sera from vaccinated individuals and neutralising antibodies against the fusion peptide a, b) Neutralization by second dose adenoviral vector (AZD1222) vaccinee-elicited sera against of WT (Wuhan-Hu-1+D614G), WT+P812S, WT+Δ69/70+796H, WT+Δ69/70+796H+P812S, Delta, Delta+P812S pseudotyped virus. Pseudovirus was co-incubated with sera for one hour at 37°C (a) or 32°C (b). Reciprocal geometric mean titre shown with 95% confidence interval. Statistical analysis was performed using ratio paired t test corrected for multiple comparison using the Holm-Šídák method. ; *P < .05, **P < .001, ***P < .0001. **c**) IC50 of fp.006 (red), and fp.007 (green) incubated with pseudovirus pseudoviruses corresponding to the indicated mutations. Results are representative of 2 technical replicates. d) The log-transformed fold change induced by adding the P812S to each spike backbone is represented in the heatmap.

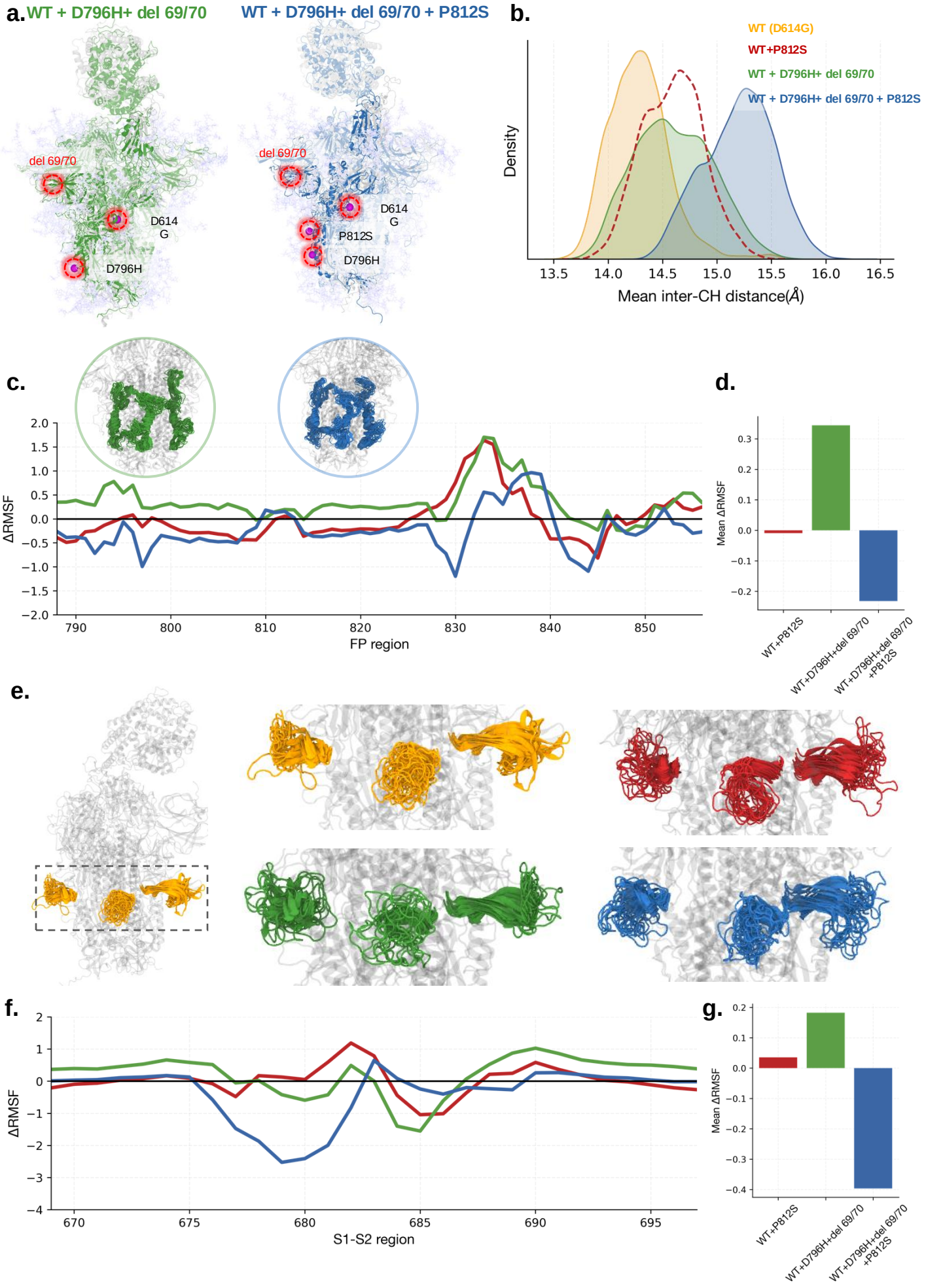


Figure 4. Presence of P812S imparts instability to the S1/S2 cleavage and fusion peptide regions.
a) The snapshots highlight the background double mutant (D796H+ del 69/70) with (right) and without (left) P812S. The position of mutations is highlighted. b) The average inter-CH distance distributions are plotted for WT and all three mutants. c) ΔRMSF ($RMSF_{WT}$-$RMSF_{MUT}$) is plotted for residues of the FP region. The snapshots highlight the conformational ensemble of FP region seen in the background mutations with and without P812S substitution. d) The average ΔRMSF from the FP region is plotted for all the mutants. e) The snapshots highlight the dynamics of S1/S2 region in WT and three mutations. f) ΔRMSF is plotted for S1/S2 region highlighting all three mutants. g) The average ΔRMSF underlying S1/S2 residues is plotted for all the mutants.

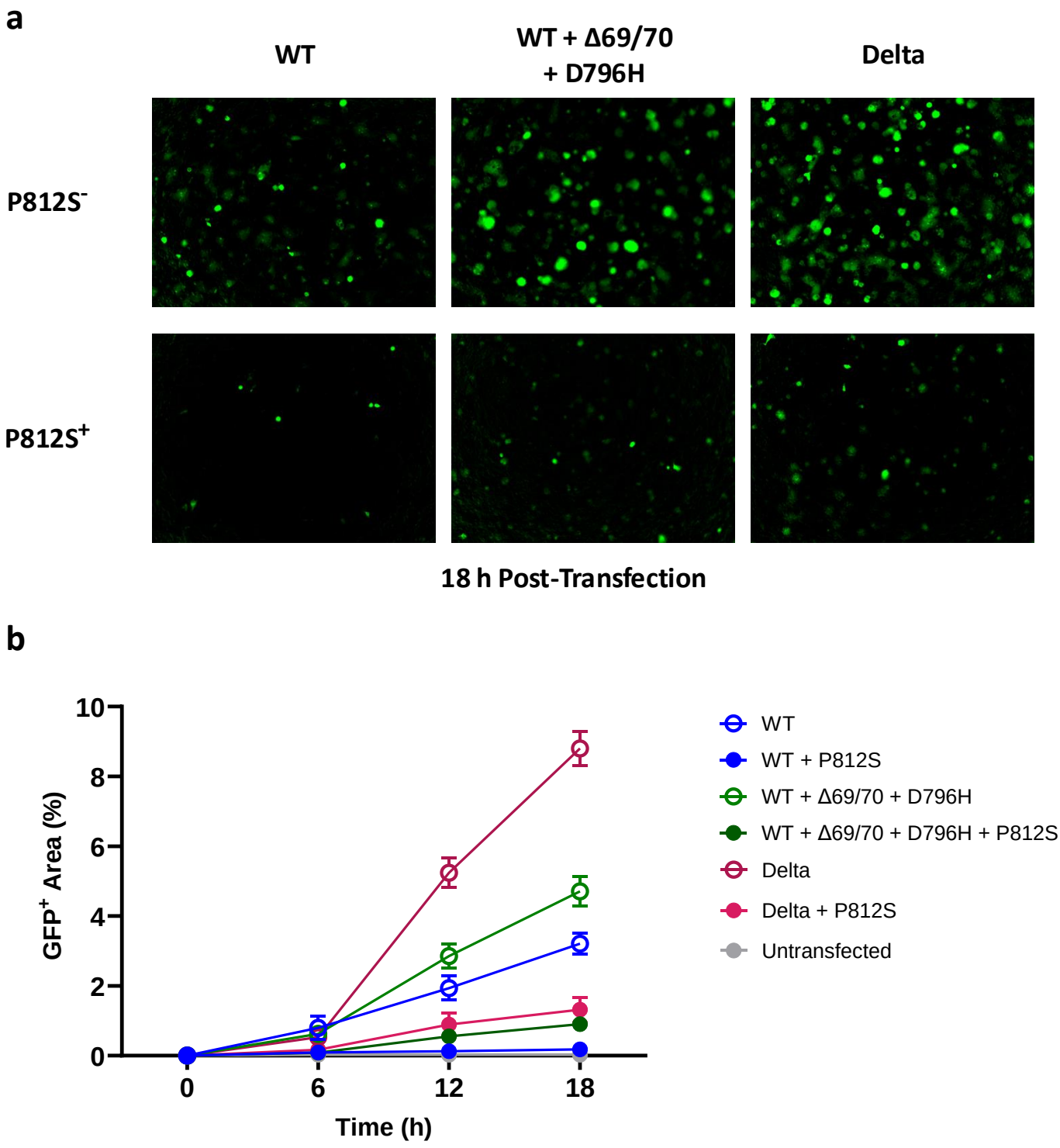
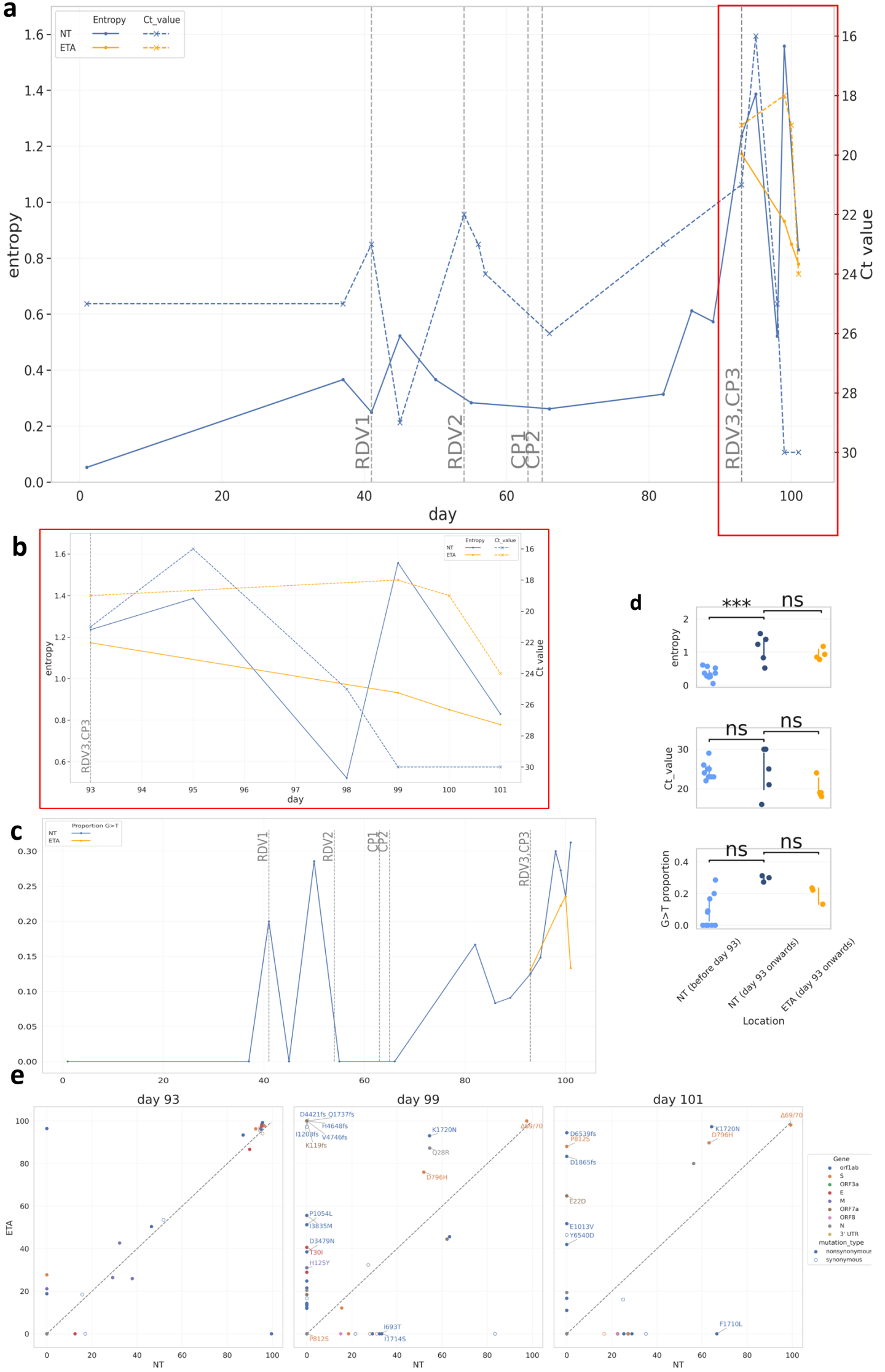
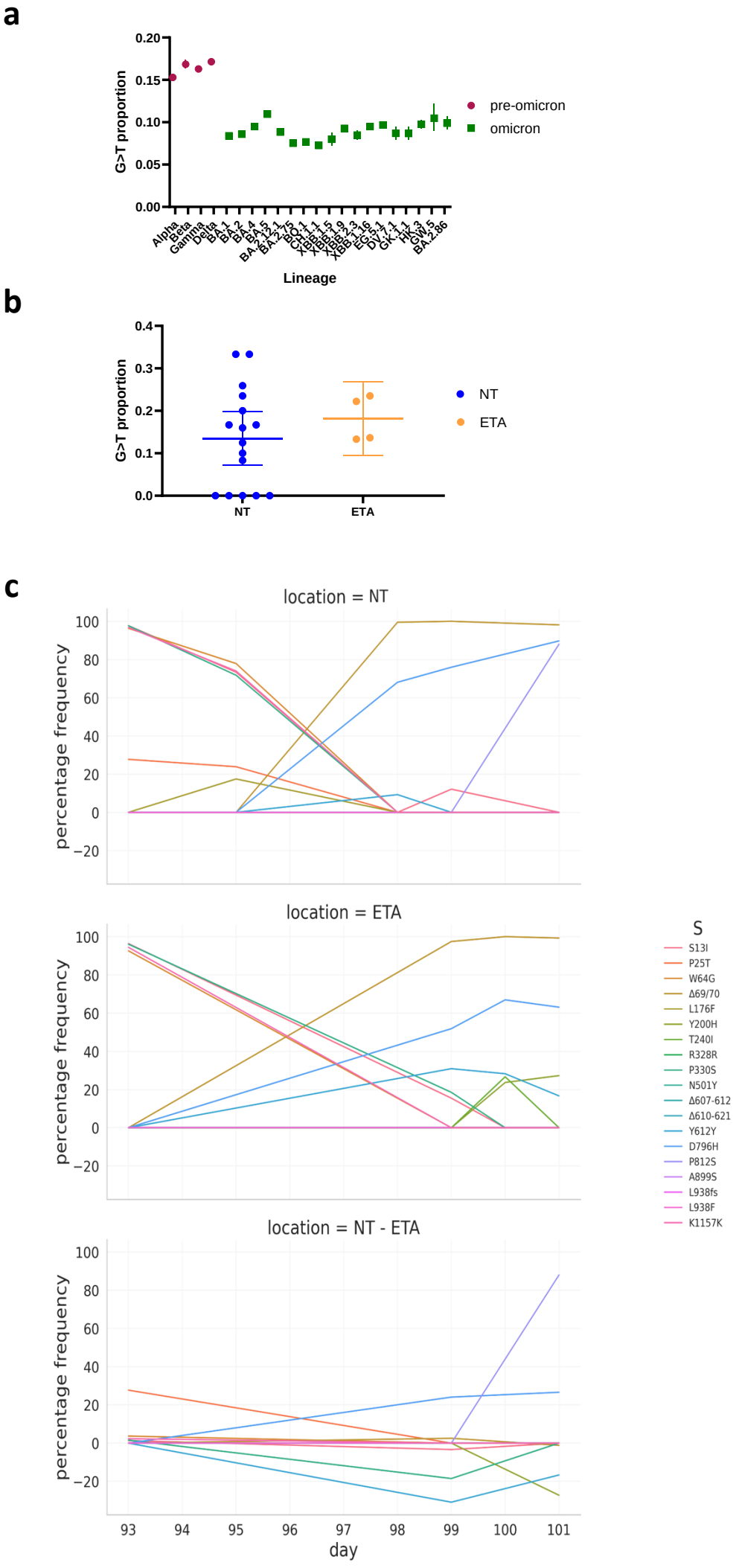


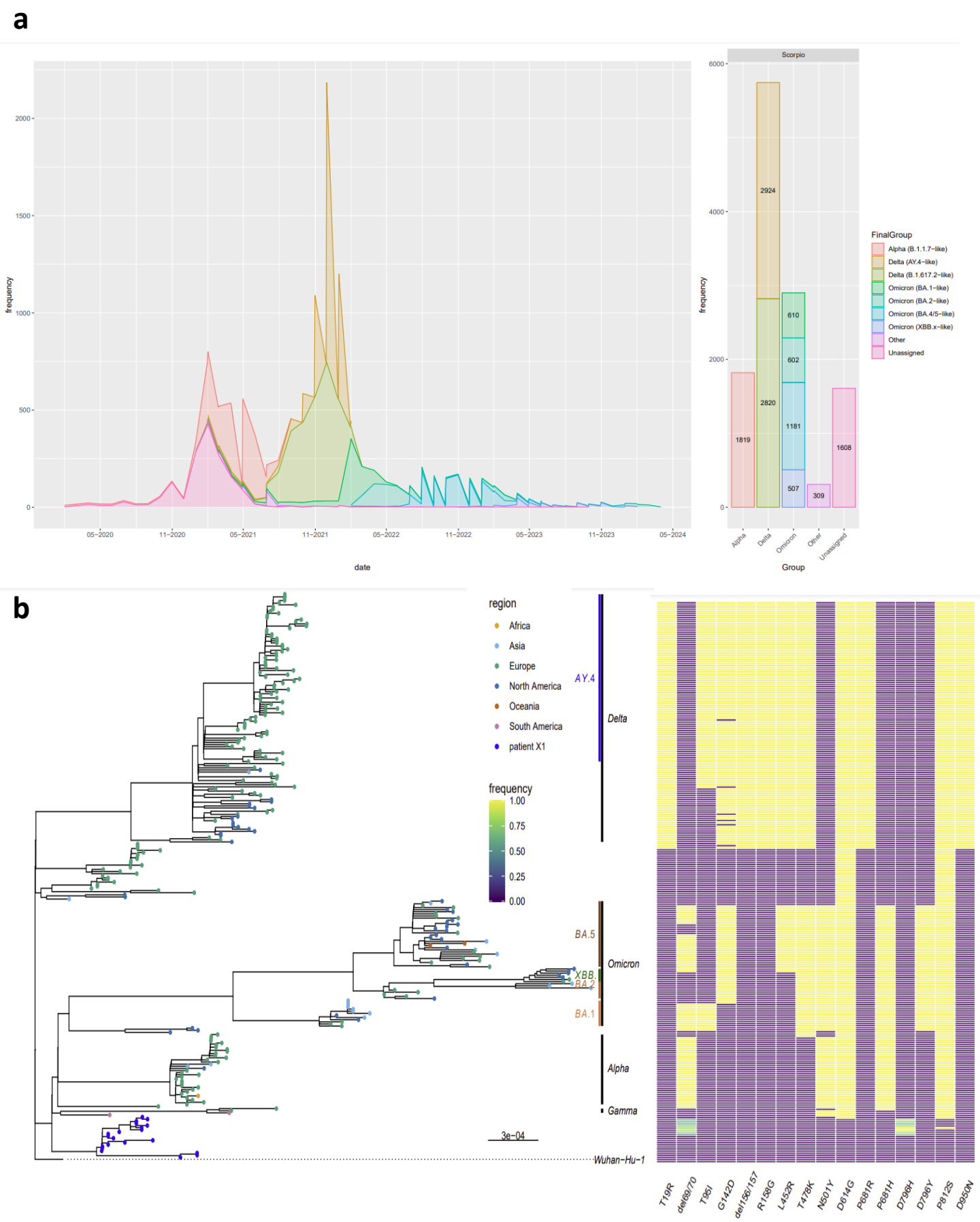
Figure 5: **Cell-cell fusion kinetics of P812S bearing SARS-CoV-2 spike proteins**. Vero-GFP1-10 co-incubated with 293T-GFP11 cells transfected with DNA plasmids encoding SARS-CoV-2 spike proteins bearing a range of mutations. a) Representative images of spike-mediated cell-cell fusion in WT (Wuhan-Hu-1+D614G), WT+P812S, WT+Δ69/70+796H, WT+Δ69/70+796H+P812S, Delta, Delta+P812S at 18 hours post-transfection. Images were captured using the IncuCyte SX5 Live-Cell Analysis System (Sartorius). b) Quantification of spike-mediated cell-cell fusion showing the percentage of GFP+ area to the total cell area at 0-, 6-, 12-, and 18-hours post-transfection. Cell-cell fusion was determined as the proportion of green area to the total phase area and was calculated with the IncuCyte analysis software (Sartorius). Data are mean ± s.e.m. from four fields of views at each time point and are representative of two independent experiments.



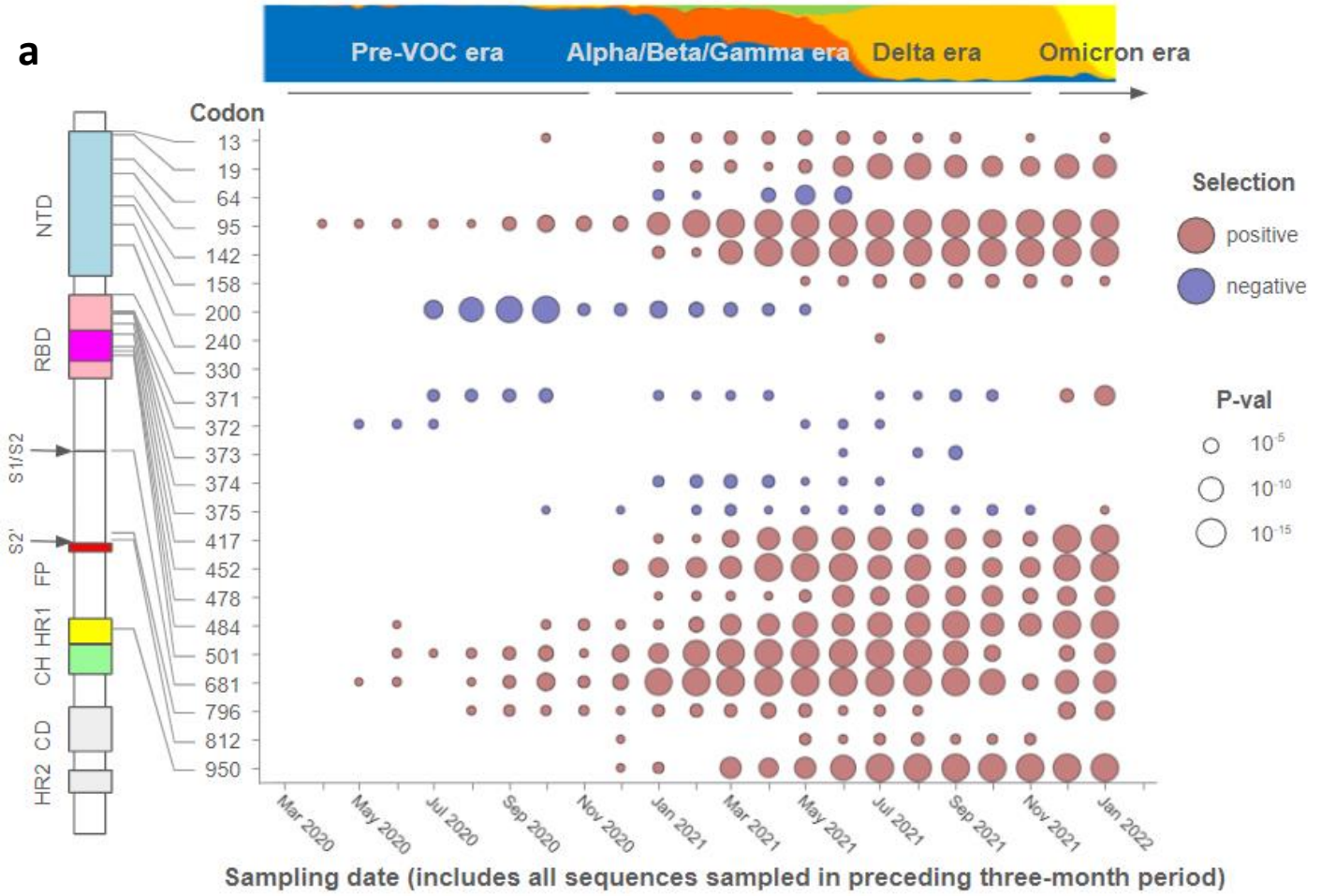
Supplementary Figure 1: Viral load and Intra-host evolution dynamics. a) Shannon's entropy (solid line, left y-axis) and Ct value (dotted line, right y-axis) of nose and throat (N+T, blue) and endotracheal aspirate (ETA, yellow) samples, zoomed in and rescaled from day 93 to day 101 (b). CP, convalescent plasma; RDV, Remdesivir. c) proportion of G to T mutation from day 1 to day 101. Day 93 and Day 95 were excluded as they are separated from the main clade. d) Dot plot (with 95% confidence interval) of Shannon's entropy (top) and Ct value (bottom) for NT samples from day 1-92, NT samples from day 93-101, and ETA samples from day 93 to 101. Student's t-test is performed for entropy whilst Welch correction was performed for Ct value in lieu of unequal variance. e) The percentage prevalence in N+T plotted against ETA. Nonsynonymous mutations (filled) with >30% difference are annotated. The grey reference line represents equivalent prevalence between NT and ETA.



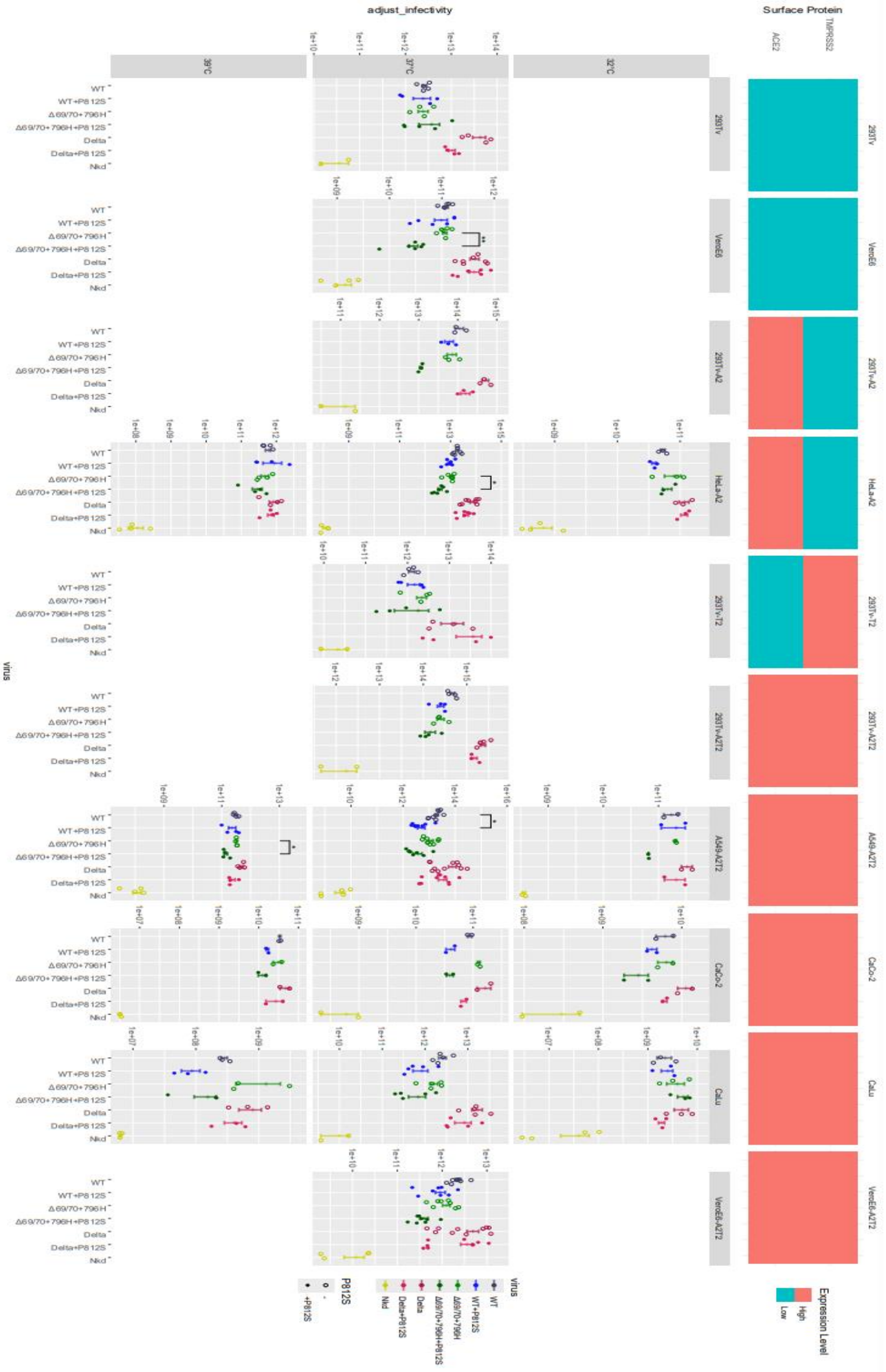
Supplementary Figure 2: Breakdown of mutational spectra and viral iSNV dynamics in the nasal and endotracheal compartments. The G to T proportion of a) globally circulating lineages in chronological order and b) of patient sample broken down by NT and ETA samples across all timepoints. C) Spike mutation frequency in N+T (top), ETA (middle), and absolute percentage difference (bottom) between day 93 and 101.



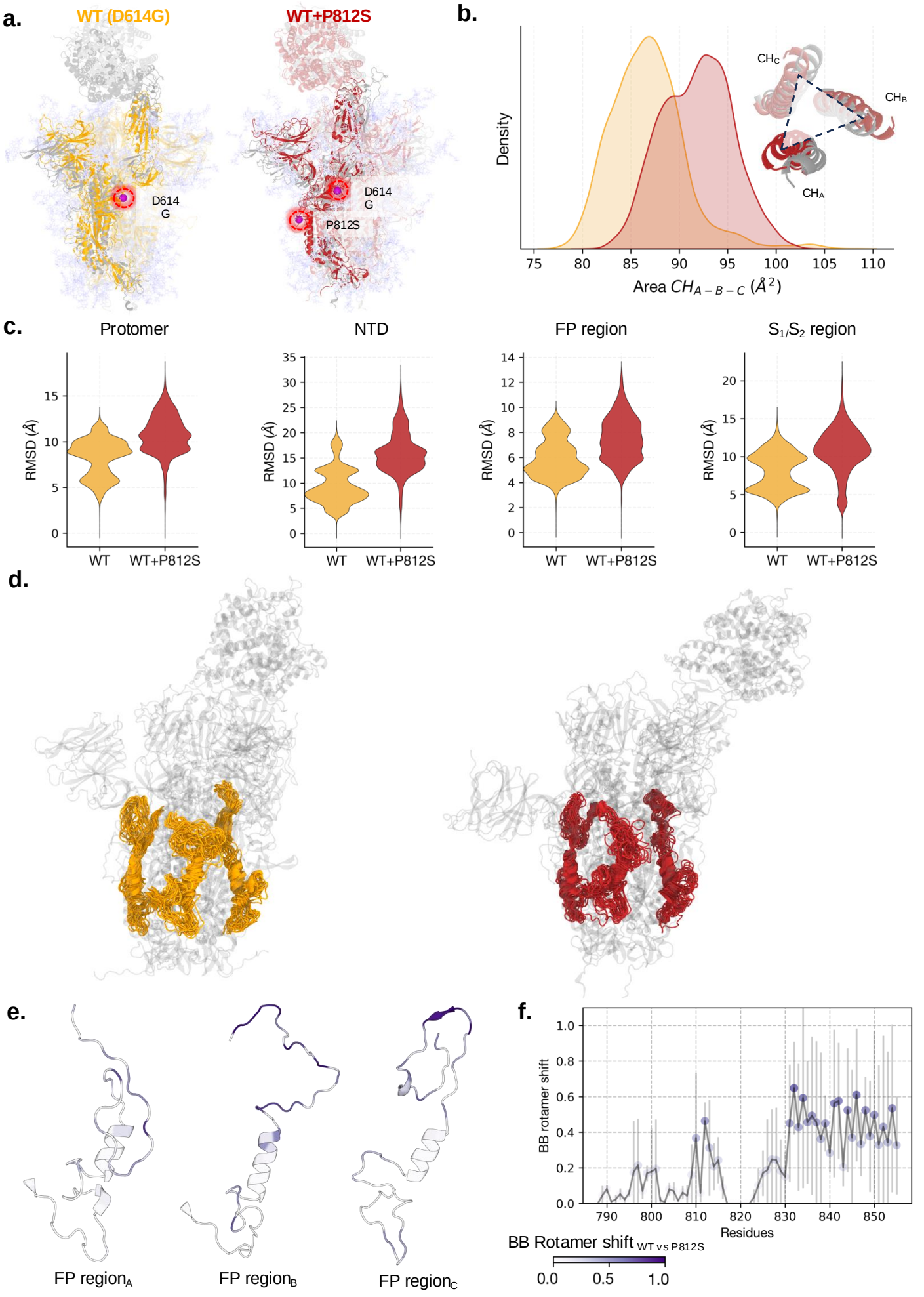
Supplementary Figure 3: Global Context of P812S containing SARS-CoV-2 global variants. a) The global prevalence of SARS-CoV-2 sequences with the S:P812S mutation in different variants defined by Scorpio as of 7th April 2024. b) Left panel - Maximum Likelihood tree (model: GTR+F+I) of patient X1 and 247 variant-stratified subsampled sequences. Tips are labelled by region. Right panel – Aligned heatmap of Spike mutations of all sequences.



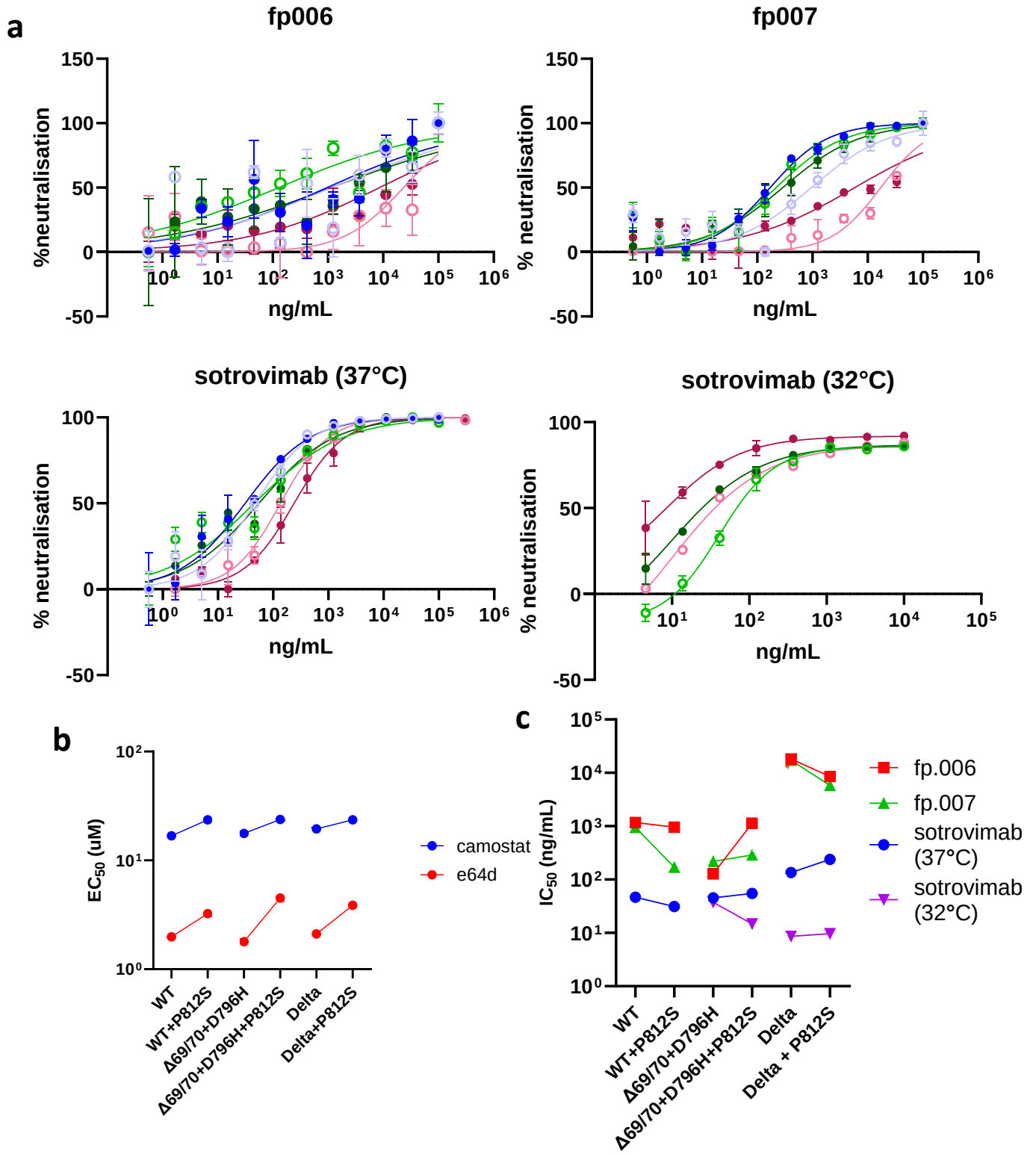
Supplementary Figure 4: Signals of natural selection at a sample of SARS-CoV-2 S-Gene codon sites during the early phases of the COVID-19 pandemic. Evidence of positive (red circles) and negative (blue circles) was detected using the IFEL method using all globally sampled near-full length SARS-CoV-2 sequences in the three months preceding the indicated sampling dates with the sizes of circles indicating the degree of statistical support for the detected selection signals (see key on the right). The approximate era during the COVID-19 pandemic (after Amodio et al) is indicated at the top. The non-variant of concern (VOC) lineages indicated in blue, the Alpha VOC in Red, the Beta VOC in dark orange, the Gamma VOC in green, the Delta VOC in orange and the Omicron VOC in yellow. The approximate location in the Spike protein of amino acids encoded by the various codon sites is indicated on the left: NTD = N-terminal domain; RBD = Receptor binding domain; FP = Fusion peptide; HR1 = Heptad repeat 1; CH = Central helix region, CD = Connector domain; HR2 = Heptad repeat 2.



Supplementary Figure 5: P812S has a detrimental effect on cell entry in additional cell lines across 32°C, 37°C, and 39°C. PV entry of WT (Wuhan-Hu-1+D614G), WT+P812S, WT+Δ69/70+796H, WT+Δ69/70+796H+P812S, Delta, Delta+P812S introduced to cells of varying levels of ACE2 and TMPRSS levels pre-incubated at 32, 37, and 39 degrees. Data are supplemented by mean ± standard error of n = 3–8 replicates. Statistical analysis was performed using unpaired student's t-test with Benjamini-Hochberg correction.



Supplementary Figure 6. The P812S mutant impacts the dynamics at both the S1/S2 and FP/S2' cleavage sites of Spike protein.
a) Snapshot of WT (D614G) and P812S spike trimers in complex with ACE-2. Both 0ns (grey) and 1μs timepoints superimposed on each other. b) The probability density plot of the triangle area formed by the centers of three CH domains is compared between WT and P812S mutant. The snapshot highlights the eagle-eye view of inter-CH orientation in WT (grey) and P812S mutant (dark red). c) The violin plots compare the distribution of RMSD calculated for the entire protomer, NTD, FP and S1/S2 regions respectively. D) The snapshots highlight the ensemble of FP conformations in WT (left) and P812S mutant (right). e) The backbone rotamer shift calculated between WT and P812S is mapped on the individual FP regions of the spike trimer. f) The corresponding rotamer shift values are plotted for FP residues. The values are averaged across three chains.



Supplementary Figure 7: Individual curves of monoclonal antibodies. a) Individual inhibitor curves of monoclonal antibodies. b) EC50 of E64d (cathepsin inhibitor, blue) and camostat (TMPRSS2 inhibitor, red) in A549-ACE2/TMPRSS2 using PV expression respective Spike. c) IC50 of sotrovimab incubated with pseudovirus at 32°C (purple), sotrovimab at 37°C (blue), fp.006 (red), and fp.007 (green) neutralizing pseudoviruses corresponding to the indicated mutations. Results are representative of 2 technical replicates Data points are representative of 2 technical replicates. Error bars represent SD.