Nasopharyngeal specific selection and emergence of SARS-CoV-2 spike fusion peptide mutation P812S during chronic infection

Tentative author list:

Mark Tsz Kin Cheng[1,2], Mazharul Altaf[1,2], Darren Martin[3], Christopher Ruis[4,5,6], Benjamin L. Sievers[1,2], Kimia Kamelian[1,2], 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] & Ravindra K. Gupta[1,2,*]

1. Department of Medicine, University of Cambridge, Cambridge, UK.
2. Cambridge Institute of Therapeutic Immunology & Infectious Disease (CITIID), Cambridge, UK.
3. Division of Computational Biology, Institute of Infectious Disease and Molecular Medicine, University of Cape Town
4. Victor Philip Dahdaleh Heart & Lung Research Institute, University of Cambridge, Cambridge, UK
5. Cambridge Centre for AI in Medicine, University of Cambridge, Cambrdge, UK
6. Department of Veterinary Medicine, University of Cambridge, Cambridge, UK

Corresponding authors: Ravindra K. Gupta rkg20@cam.ac.uk and Mark Cheng tkmc2@cam.ac.uk

Lead contact: rkg20@cam.ac.uk

SUMMARY

Persistent SARS-CoV-2 infections are a likely source of new variants of concern. In a fatal infection in an immunocompromised patient, we find association between viral load and diversity, and abrupt diversification between contemporaneous nasopharyngeal and endotracheal aspirate samples following remdesivir and convalescent plasma treatment. Shared high proportion of G>T mutations suggests spread of viruses from the lower to the upper tract as the patient deteriorated. We identified the S:P812S spike mutation adjacent to the fusion peptide unique to the nasopharyngeal sample site. We find P812S reduced S1/S2 cleavage and decreased entry efficiency in cell lines with a range of ACE2 and TMPRSS2 expression levels. This reduction of infectivity contrasted with reduced susceptibility to neutralising antibodies in sera from vaccinated individuals conferred by P812S at both 37°C and 32°C (simulating upper tract). Thus, S:P812S is a specific adaptation during SARS-CoV-2 intrahost evolution, allowing immune evasion at lower temperatures observed in the upper respiratory tract at the expense of target cell entry efficiency.

Introduction

SARS-CoV-2 evolution and adaptation in persistent infection of immunocompromised hosts has been demonstrated in myriad studies.[1–4] COVID-19 remains a particularly relevant threat for

immunocompromised individuals, as COVID-19 vaccines have been shown to induce lower seroconversion rates,[5] reduced neutralisation activity,[6] and shorter durations of protection as compared to immune-competent individuals[7]. Added complexity arises in managing viral infections in immunocompromised individuals as immune status and disease presentation can vary greatly. This heterogenous group includes individuals with cancers, human immunodeficiency virus (HIV) infections, transplants, and others using immunomodulatory therapies including B cell depletion.

The primary receptor for SARS-CoV-2 is ACE2. The serine protease TMPRSS2 cleaves spike at the S2', liberating the fusion peptide to initiate viral-cell fusion at the plasma membrane. Newer variants appear less dependent on TMPRSS2 and have preference for endocytic entry where cathepsin carries out S2' cleavage.[8] Given the broad expression of ACE2 across diverse tissues, it is unsurprising that infection of multiple organs occurs during COVID-19 infection. Local tissue environments likely exert specific selective pressure for replication. Viral divergence in different anatomical compartments has been reported, including respiratory compartments such as nasopharynx, sputum, trachea, lungs, and other compartments including plasma, heart, spleen, urine, and gut,[9,10] consistent with differential selection pressure imposed by different niches.[10,11]

ACE2 and TMPRSS2 expression gradients exist between the upper- (URT) and lower- respiratory tract (LRT), and efficiency of TMPRSS2 use has varied with VOC. This phenomenon is demonstrated in chronic infections[9,11–14] and at the population level, with Omicron showing affinity towards the URT[8,15] and greater transmission potential. This tropism shift has been associated with reduced intrinsic disease severity.[16] Affinity towards the URT also leads to differential exposure to environmental and immune selective pressures, creating a unique mutational spectrum which can be leveraged to determine the niche of pathogens.[9,15]

Results

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.[1] 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 rises in diversity in NT samples when remdesivir or convalescent plasma were administered as the infection progressed, but no statistically significant difference in genetic diversity within the ETA compared to the NT samples **(Supplementary Figure 1a-d)**. Consistent with previous studies linking Ct values to viral diversity, the increase in viral diversity was accompanied by an increase in viral load.[9] **(Supplementary Figure 1d)**

The relative frequency of G>T mutations has previously been demonstrated to correlate with SARS-CoV-2 replication niche, with replication in the lung resulting in higher frequencies of G>T mutations.[15] Omicron variants, which are more URT tropic than pre-omicron variants,[8] have lower G>T frequencies than pre-omicron variants. **(Supp Figure 2a)** We therefore aimed to determine whether there was a difference in G>T frequencies between the NT and ETA isolates from the same patient. As expected given the short distance between these sites, we did not observe a difference between these niches **(Supplementary figure 2b)**; however, we noticed a clade of patient isolates that exhibited elevated numbers of G>T mutations **(Figure 1b,d)**. This clade contains both NT and ETA sequences from late in the course of infection (after day 93), suggesting spread of viruses within the patient across the respiratory tract. The elevation in G>T mutations amongst isolates in this clade suggests that this spread occurred from the LRT to the URT, consistent with deteriorating lung disease. This mixing of viral populations as the patient deteriorated clinically could have been related to the physical conduit provided by the endotracheal tube or air movement associated with mechanical ventilation.

We further examined the changes in mutation frequencies between the NT samples compared to the ETA samples. While there is limited divergence in variant frequencies between NT and ETA samples on day 93, after remdesivir administration and two units of convalescent plasma infusion on day 95, divergent mutations arose for Spike (Δ69/70, D796H, P812S) and other genes, including ORF1ab (E1013V), Envelope (H125Y), Nucleocapsid (Q28R) and Membrane (H125Y) **(Figure 1a, supplementary figure 1e, Supplementary figure 3)**. This is likely due to selection pressure imposed by the convalescent plasma and/or remdesivir treatment.[4]

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

At the time of this study (7th April 2024), P812S has a low prevalence in the Global Initiative on Sharing All Influenza Data (GISAID) database (<0.5%, n =12380) **(Table 1, supplementary figure 4a)**.

105 P812S was most abundantly found in Delta variant viruses (n = 5744), but was also found in the
106 Alpha variant (n = 1819) and the major Omicron lineages BA.1 (n=610), BA.2 (n=602), BA.4/5
107 (n=1181), and XBB (n=507) viruses **(Table 1, Supp Figure 4a)**. Though P812S is a minority in these
108 major lineages, over 99.6% of AY.4.10 sublineage sequences (1801 out of 1808) and 99.2% of
109 B.1.463 sublineage sequences (122 out of 123) contained the P812S mutation **(Figure 1c)**, indicating
110 it was a defining mutation in these lineages and arguing for successful transmission of P812S-bearing
111 viruses.[17–19] Indeed, from the start of the pandemic to August 2023, spike codon 812 had an average
112 dN/dS ratio of 2.47 (p=0.184), suggesting positive selection (i.e., selection favouring changes away
113 from the P amino acid state). (https://observablehq.com/@spond/sars-cov-2-global-genomic-selection-2019-
114 aug-2023). Statistical evidence (p<0.01 with the IFEL method) of positive selection at codon 812 was
115 most marked between May to November 2021 when the Delta variant was most prevalent,
116 corroborating the beneficial effect of diverging on a Delta genomic background from P812 to
117 alternative amino acids such as P812S under at least some epidemiological conditions.
118 **(Supplementary figure 5a)**

119 We stratified by variant and proportionally subsampled 247 SARS-CoV-2 sequences containing P812S
120 and inferred a maximum-likelihood tree **(Supplementary Figure 4b; GISAID EPI_SET_240114wm)**.
121 We annotated the sequences based on the accompanying mutations and noticed that, apart from
122 the D614G mutation that arose early in the pandemic, there were no mutations consistently found
123 to occur with P812S.

124 As P812S was found only in the NT sample, we hypothesised that P812S conferred an advantage in
125 that niche. The proximity to the S2' cleavage site and fusion peptide, in addition to the exposed
126 position of P812 when Spike is in the open position **(PDB: 6XM4, Supplementary Figure 5b)**
127 prompted us to study the infectivity, cleavage, and immune evasion properties of P812S mutants.

128 **P812S leads to a general decrease in spike-mediated entry and cleavage efficiency across a range**
129 **of temperatures**

130 As P812S is directly upstream (N-terminal) to the S2' cleavage site (R815/S816) and the S2 viral
131 fusion peptide (FP, 816-855),[20–22] we hypothesised that P812S may directly impact spike glycoprotein
132 S2 cleavage and viral entry. To test this, we introduced the P812S mutation into Wuhan-Hu-1 +
133 D614G (WT), WT with Δ69/70 + 796H mutations, as well as Delta B.1.617.2. We then performed
134 infectivity assays using viruses with and without the P812S mutations on cell lines that varied in
135 ACE2 and TMPRSS2 expression levels. Across $ACE2_{high}$/$TMPRSS2_{low}$, $ACE2_{high}$/$TMPRSS2_{high}$, **(Figure 2a)**,
136 $ACE2_{low}$/$TMPRSS2_{low}$, and $ACE2_{low}$/$TMPRSS2_{high}$ cells, P812S was associated with a significant deficit in
137 cell entry efficiency **(Supplementary Figure 6).** This P812S-associated entry defect persisted across
138 32°C, 37°C, and 39°C, a temperature which approximates pyrexia**. (Supplementary Figure 6)**.

139 As the cleavage of S1/S2 has been shown to correlate with infectivity/entry efficiency, to explore this
140 further, we performed western blot for full length (FL) and S2 spike on 293T producer cell lysate and
141 purified virions.[23] We noticed that P812S was associated with suboptimal S1/S2 cleavage in both
142 producer cell lysate and virion, demonstrated by decreased ratio of S2 to FL spike (**Figure 2b**). The
143 P812S-induced reduction of S2 in cell lysate and released virions suggests that P812S plays a role in
144 stabilising the S protein and slowing S1 shedding. In addition, reduced susceptibility to camostat
145 (serine protease inhibitor, including TMPRSS2) suggests that P812S alters S2' cleavage allosterically,
146 thereby reducing entry via plasma membrane fusion. On the other hand, P812S does not change
147 sensitivity to E64d (a cathepsin inhibitor), implying that there is no change to the endosomal entry
148 pathway (**Figure 2c, Supplementary Figure 7a**). This reduction in S1/S2 cleavage and TMPRSS2 usage

149 is reminiscent of the change observed from delta to omicron variant[8], which also saw a preference
150 towards replication in the upper respiratory tract.[15,24]

151 **Context-specific vaccinee sera and antibody evasion associated with P812S**

152 Multiple broadly neutralising antibodies have been identified to interact with amino acids on the N-
153 terminal side of the FP (816-655), including P812.[25–27] Therefore, we hypothesised that P812S may
154 confer a degree of immune escape due to selective pressure imposed by convalescent plasma
155 treatment.

156 We tested the neutralisation sensitivity of Wuhan-Hu-1 + D614G (WT), WT+P812S,
157 WT+Δ69/70+D796H, and WT+ Δ69/70+D796H+P812S using a previously reported pseudovirus (PV)
158 system.[28] We further incorporated Delta and Delta+P812S in the panel because P812S was
159 predominantly found in Delta and previous genomic-epidemiological studies showed P812S[29] further
160 increases the likelihood of Delta breakthrough infections[30,31]. We tested 20 stored vaccinee sera
161 (average age 66.0, 55% M) one month after second-dose AstraZeneca ChAdOx1 COVID-19 vaccine.[32]

162 We confirmed previous observations[1] showing modest immune escape by the double mutant
163 Δ69/70+D796H, with a 1.2-fold decrease in geometric mean serum neutralising titres between WT
164 and WT+Δ69/70+D796H **(Figure 2e)**. P812S on a WT+Δ69/70+D796H backbone was associated with
165 a reversal to WT levels of antibody sensitivity **(Figure 2e)**. This suggests that the immune evasion
166 capabilities conferred by D796H are partially abolished in the presence of P812S. To further
167 understand the impact of the P812S mutation, we produced Delta pseudovirions with and without
168 the P812S mutation. Using sera one month after two doses of AstraZeneca ChAdOx1-S, both
169 baseline serum neutralising titres against Delta and Delta+P812S were lower by 3.1- and 2-fold than
170 WT, respectively. There was a significant 1.5-fold increase in serum neutralisation GMT of the Delta
171 P812S mutant compared to Delta **(Figure 2e)**. In summary, at 37°C, in both the Delta and
172 WT+Δ69/70+D796H backbones, P812S incorporation increased NT50, indicating increased
173 susceptibility to serum neutralisation. **(Table 2)**

174 Given the difference in vaccine sera evasion provided by P812S on different spike backbones, we
175 used antibodies targeting different spike regions to infer key structural changes essential to P812S
176 antibody evasion. Sotrovimab (derived from S-309) (GSK) is a licensed monoclonal antibody
177 treatment which targets a highly conserved epitope within the Spike receptor binding domain
178 (RBD).[33] We wished to explore the impact of P812S on susceptibility to Sotrovimab treatment, given
179 its emergence in previous VOC. In agreement to previous literature, Sotrovimab demonstrated
180 similar neutralisation titres with serum against WT and Delta, [34] and P812S had limited effect on
181 sotrovimab susceptibility, eliciting IC50 fold changes less than two fold **(Figure 2d, Table 3,**
182 **Supplementary Figure 7b)**.

183 To better simulate the upper airway environment, we tested PV expressing WT+Δ69/70+D796H or
184 Delta +/- P812S against sera/monoclonal antibody at 32°C.[35] The serum geometric mean titre for
185 32°C increased for all viruses by between x1.5 to x2.0 fold compared to 37°C. (Table 2) At 32°C,
186 P812S also reduced geometric mean titre to vaccinee sera by x1.4 fold in the background of
187 WT+Δ69/70+D796H and x1.3 fold in the Delta background, respectively **(Figure 2f).** For Sotrovimab,
188 WT+Δ69/70+D796H was x2.5 fold more sensitive than WT+Δ69/70+D796H+P812S **(Figure 2d)**.

189 To further elucidate the impact of the P812S mutation on immune evasion, two additional
190 monoclonal antibodies were added to the panel. We used fusion-peptide targeting recombinant
191 monoclonal antibodies *fp.006* and *fp.007* from Bianchini and colleagues.[22] In brief, *fp.006* recognises

192 a conserved fusion peptide epitope that is only exposed upon ACE2 binding. On the other hand,
193 antibody *fp.007* binding affinity and neutralising capacity is independent of ACE2 binding.

194 *fp.007* demonstrated a significant 5.58- and 2.96-fold decrease in IC50 upon P812S introduction on
195 both WT and Delta backbones, respectively, but a marginal increase in IC50 levels upon P812S
196 introduction to the Δ69/70+D796H double mutant. **(Figure 2d, Table 3)** On the other hand, the IC50
197 reduction conferred by P812S for WT (x1.22) and Delta (x2.14) is modest for the ACE2-binding
198 dependent fp.006, compared to the IC50 increase for Δ69/70+D796H (x8.88).

199 This data is suggestive that P812S exposes the fusion peptide site regardless of ACE2 binding, and
200 the addition of Δ69/70+D796H in our patient confers epistatic changes that are evasive to FP-specific
201 antibodies. This anti-FP-antibody evasion is context specific and is partially abrogated in the
202 Δ69/70+D796Y-inclusive context of the Delta variant.

203 **Discussion**

204 In summary, we demonstrate the emergence of P812S in a chronic infection case in the upper
205 respiratory tract (URT) after remdesivir and convalescent plasma treatment. We propose that the
206 selection of P812S is an immune evasion adaptation that came at the cost of cell entry efficiency,
207 possibly mediated through impaired S1/2 cleavage in P812S-bearing spikes. Interestingly, this URT
208 adaptation occurred on the background of within-host evolution suggestive of adaptation to the LRT
209 based on mutational spectra analysis. Thus, it is plausible that the LRT population migrated to the
210 URT given the intense inflammation in the lower airway in the present case. Although we were
211 unable to separate URT and LRT isolates in mutational spectra due to the similarity of sequences at
212 the consensus level from our patient, comparison of patient X1 sequences and global sequences of
213 pre-Omicron and Omicron lineages is suggestive of progression from lower to upper airway. Our
214 work demonstrates the importance of studying compartmentalisation of viral populations, especially
215 in persistent infection. Monitoring these mutations that have epistatic effects could potentially
216 assist in prediction of mutations in future circulating strains.

217 In addition, the emergence of short-lived variants in an immunocompromised host can provide
218 insight into context-specific adaptations. We explored the example of P812S, which grants immune
219 evasion from vaccinee polyclonal sera and fusion peptide (FP)-specific antibodies, only when in the
220 context of the patient's mutations Δ69/70+D796H. This immune evasion is at the price of S1/S2
221 cleavage, S2' cleavage, TMPRSS2 utilisation, and infectivity, which is expected given the location of
222 P812 within S2 and its proximity to the FP.

223 The effect of P812S on the neutralisation activity of *fp.006* and *fp.007* suggests the immune evasion
224 of P812S is through the alteration of immunogenicity of the cryptic FP epitope exposed after ACE2
225 binding. This would explain the minimal changes in geometric mean titre (GMT) in vaccinee sera, as
226 FP-specific antibodies are rare in the general population. Given P812S reduces the sensitivity to
227 ACE2-binding dependent fp.006 and increases the sensitivity to ACE2-binding independent fp.007
228 WT+Δ69/70+D796H and the opposite is true in WT and Delta, we speculate that there are unique
229 changes to the FP structure of WT+Δ69/70+D796H pre- and post-ACE2 binding.

230 **Limitations of the study**

231 Our study was based on lentiviruses pseudotyped with different SARS-CoV-2 spikes. Although
232 significant consistency has been observed between pseudotyped virus system and the live-virus
233 system for the study of virus entry, there may still be potential effectors only present in the live-virus
234 system.[8] Confirmation of our finding in a live-virus system would be desirable.

Acknowledgement

This work was funded by a Wellcome Senior Fellowship to RKG (WT108082AIA). This research was supported by the NIHR Cambridge Biomedical Research Centre (NIHR203312). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. We thank the valuable discussion with Dr Maximillian Muenchhoff on SARS-CoV-2 HLA and CD8 T cell epitope evasion analysis. We also thank Dr Filippo Bianchini, Ms Jasmine Chantal Cantergiani, and Dr Davide F. Robianni for sending samples and expression plasmids of FP.006 and FP.007 antibodies.

Author Contributions

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

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

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

Writing – Original Draft, M.T.K.C.;

Writing – Review & Editing, M.T.K.C, D.M., C.R., R.K.G.;

Funding Acquisition, R.K.G;

Resources, R.K.G;

Supervision, B.M., S.A.K., R.K.G.

Declaration of interests

The authors declare no competing interests.

Methods

Clinical-sample collection and sequencing

Sequencing of SARS-CoV-2 viral DNA extracted from serial samples from the LRT (sputum and endotracheal aspirate), and URT (throat and nasal swab) was described previously in Kemp et al.[1] In short, nucleic acid extraction was done from 500 µl sample using the easyMAG platform (Biomerieux). Sequencing amplicons were prepared with the ARTICnetwork v.3 protocol and diluted to 2 ng μl^{-1} and 25 µl (50 ng) were used as input for each library preparation reaction, before sequencing using the Illumina MiSeq platform.

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

Sequence data processing and analysis

Illumina sequencing data from Kemp et al. was re-analysed.[1] Bam files were downloaded from Kemp et al. (BioProject PRJNA682013), mapped to Wuhan-Hu-1 (NC_045512.2) by minimap v2.28 (arguments -ax sr),[36] sorted and indexed by samtools v1.20,[37] and variants were called by bcftools v1.18 functions mpileup (minimum base quality q30), and call (arguments --multiallelic-caller --variants-only).[38] Mutations identified by BCFtools were filtered by a minimum depth cutoff of 20x before annotations by snpEff v5.2 using the publicly available NC_045512.2 database.[39]

Genetic diversity and viral load analysis

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

274 $$H = -\sum_{i}^{n} P(i) log_e P(i)$$

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

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

279 **Phylogenetic analysis and lineage assignment**

280 Excluding duplicates, and low coverage sequences, we searched GISAID for all complete coverage
281 SARS-CoV-2 sequences with the Spike_P812S substitution on the 7th April 2024. All sequence
282 encoding Spike-S812 were aligned to MN908947.3 by mafft (v7.525, --6merpair --keeplength --
283 addfragments), and then were assigned lineages using Pangolin (V4.3).[40] The major lineages were
284 grouped by Scorpio (V0.3.19). We stratified the sequences based on their their major lineage
285 designations and then performed stratified random sampling with proportional allocation to major
286 lineages to obtain a subset of 250 P812S-containing sequences using the R package Rsample v1.2.1.

287 We constructed a maximum-likelihood tree from patient X1 and the 250 sequence sample using IQ-
288 Tree v2.2.5 employing the GTR+F+I model of nucleotide substitution (model with the lowest
289 Bayesian Information Criterion identified by IQ-Tree ModelFinder) with clustering support assessed
290 by 1000 rounds of ultrafast bootstrapping.[41]

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

292 Mutational sites found in the patient, the Alpha variant, and the Delta variant, together with sites of
293 neutralisation escape mutations were chosen for natural selection analysis using the internal-branch
294 fixed effects (IEFL) method implemented in HYPHY.[42] All globally sampled near-full length SARS-CoV-
295 2 sequences are sampled between December 2019 and March 2022 were binned in three-monthly
296 intervals (labelled according to the last month of sampling within a bin) and analysed separately
297 within the context of the entire Spike open reading frame: an analysis setup needed for both
298 phylogenetic context and for informing the nucleotide and codon substitution models upon which
299 the IEFL method relies. Only codon substituons mapped to internal phylogenetic tree branches were
300 considered in these analyses since mutations mapping to termnail branches would have likely
301 experienced insufficient durations of selection to yield meaningful evolutionary information. The
302 approximate dominant-variant era during the COVID-19 pandemic where indicated based on results
303 from Amodio and colleagues.[43]

304 **Structure Analysis**

305 Coordinates for the SARS-CoV-2 Spike trimer complex with two RBD down and one RBD up was from
306 PDB: 6XM4-Chain A to C, and domains definition from multiple references were used.[22,44–46] Graphics
307 describing the structures were made in UCSF Chimera X V1.8.[47]

308 **Mutational Spectrum analysis**

309 We calculated the mutational spectrum of mutations within each patient sample using the mutation
310 files generated above. The korimuto pipeline within MutTui v2.0.2[48] was run on each sample

311 independently using NC_045512.2 as the reference genome. The proportion of G>T mutations
312 within each sample was calculated as the number of G>T mutations divided by the total number of
313 mutations.

314 **Generation of P812S Spike mutants**

315 Amino acid substitutions were introduced into a codon-optimised pCDNA3.1_SARS-CoV-2 WT,
316 Δ69/70, Delta **(B.1.617.2, EPI_ISL_2378732)**, and Omicron (**EPI_ISL_7418017**) Spike Plasmid as
317 previously described,[1,8,31] using the QuikChange Lightning Site-Directed mutagenesis kit, as per
318 manufacturer's protocol (Agilent Technologies).

319 **Pseudovirus preparation**

320 Pseudovirus were prepared by transfecting HEK293Tv cells with 1µg of P8.91 HIV-1 gag-pol
321 expression vector, 1.5 µg of pCSFLW, and 1µg of plasmid expressing the respective spike protein
322 using a 3:1 ratio of Fugene HD transfection agent (Promega).[28,49] After 48 hours of incubation at 10%
323 CO_2, 37°C, viral supernatant was collected and filtered through a .45 µm filter and stored at -80°C
324 until further use. The 50% tissue-culture infectious dose ($TCID_{50}$) of SARS-CoV-2 pseudovirus was
325 determined using the Steady-Glo Luciferase assay system (Promega).

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

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

333 **Western blot analysis**

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

338 All samples are further lysed and reduced in Laemmli reducing buffer (1M Tris-HCl (pH6.8), SDS,
339 100% glycerol, β-mercaptoethanol and bromophenol blue). This was followed by electrophoresis on
340 SDS 4-12% Bis-Tris protein gels (Thermo Fisher Scientific) under reducing conditions. Proteins on the
341 gel are transferred by electroblotting onto polyvinylidene difluoride (PVDF) membranes. The SARS-
342 CoV-2 spike proteins and pseudotype virus capsid were visualized using a ChemiDoc MP imaging
343 system (Biorad) using a rabbit anti-Spike polyclonal (Thermo Fisher Scientific, Cat# PA1-41165,
344 LOT:WE3286564 : 1 in 4000) and mouse anti-p24 monoclonal (NIH-ART3537: 183-H12-5C).

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

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

349 A549-ACE2/TMPRSS2 (A549-A2T2) cells were treated with either E64d (Tocris) or camostat (Sigma-
350 Aldrich) for 2 h at each drug concentration before the addition of a comparable amount of input
351 pseudotyped viruses (approximately 500,000 RLU). The cells were then left for 48 h before addition

352 of lysis buffer + luciferase substrate (Promega) and read on a Glomax plate reader (Promega).
353 Normalisation was performed using an inhouse script, and EC_{50} values were calculated using a
354 nonlinear regression model of inhibitor dilution vs. normalized response with a variable slope in
355 GraphPad Prism 10.

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

357 Previous work have shown that IC_{50} obtained from spike pseudotype assays quantitatively correlate
358 with neutralisation activity measured using an authentic SARS-CoV-2 neutralization assay.[51]

359 Virus neutralisation assays were conducted on HeLa cells stably expressing ACE2 (HeLa-A2) using
360 SARS-CoV-2 spike pseudotyped virus expressing firefly luciferase as previously described.[1]

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

369

370

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 – Geometric mean titre of AZD1222 one month post second dose vaccinee sera

Spike presented on pseudotyped Virus	Geometric mean titre	
	32 °C	37°C
WT	N/A	339.3
WT+P812S	N/A	347.1
Δ69/70+D796H	373.9	293.2
Δ69/70+D796H+P812S	274.0	221.7
Delta	302.5	199.9
Delta + P812S	229.2	132.3

Geometric mean titre 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 3 – IC50 of SARS-CoV-2 Spike targeting monoclonal antibodies

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

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

383 **Citations**

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

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

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

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

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

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

406 7. Ainsley Ryan Yan Bin Lee, Shi Yin Wong, Louis Yi Ann Chai, Soo Chin Lee, Matilda Xinwei Lee,
407 Mark Dhinesh Muthiah, Sen Hee Tay, Chong Boon Teo, Benjamin Kye Jyn Tan, Yiong Huak Chan, et al.
408 (2022). Efficacy of covid-19 vaccines in immunocompromised patients: systematic review and meta-
409 analysis. BMJ *376*, e068632. https://doi.org/10.1136/bmj-2021-068632.

410 8. Meng, B., Abdullahi, A., Ferreira, I.A.T.M., Goonawardane, N., Saito, A., Kimura, I.,
411 Yamasoba, D., Gerber, P.P., Fatihi, S., Rathore, S., et al. (2022). Altered TMPRSS2 usage by SARS-CoV-
412 2 Omicron impacts infectivity and fusogenicity. Nature *603*, 706–714.
413 https://doi.org/10.1038/s41586-022-04474-x.

414 9. Wang, Y., Wang, D., Zhang, L., Sun, W., Zhang, Z., Chen, W., Zhu, A., Huang, Y., Xiao, F., Yao,
415 J., et al. (2021). Intra-host variation and evolutionary dynamics of SARS-CoV-2 populations in COVID-
416 19 patients. Genome Medicine *13*, 30. https://doi.org/10.1186/s13073-021-00847-5.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

467 24. 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,
468 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
469 cultures of the human upper and lower respiratory tract. eBioMedicine *83*, 104232.
470 https://doi.org/10.1016/j.ebiom.2022.104232.

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

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

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

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

485 29. Fibke, C.D., Joffres, Y., Tyson, J.R., Colijn, C., Janjua, N.Z., Fjell, C., Prystajecky, N., Jassem, A.,
486 and Sbihi, H. (2022). Spike Mutation Profiles Associated With SARS-CoV-2 Breakthrough Infections in
487 Delta Emerging and Predominant Time Periods in British Columbia, Canada. Frontiers in public health
488 *10*, 915363. https://doi.org/10.3389/fpubh.2022.915363.

489 30. Kemp, S.A., Cheng, M.T.K., Hamilton, W.L., Kamelian, K., Chauhan, H., Dikid, T., Gogia, H.,
490 Lall, H., Ponnusamy, K., Verma, K., et al. (2022). Transmission of B.1.617.2 Delta variant between
491 vaccinated healthcare workers. Scientific reports *12*, 10492. https://doi.org/10.1038/s41598-022-
492 14411-7.

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

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

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

503 34. Cox, M., Peacock, T.P., Harvey, W.T., Hughes, J., Wright, D.W., Willett, B.J., Thomson, E.,
504 Gupta, R.K., Peacock, S.J., Robertson, D.L., et al. (2023). SARS-CoV-2 variant evasion of monoclonal
505 antibodies based on in vitro studies. Nature Reviews Microbiology *21*, 112–124.
506 https://doi.org/10.1038/s41579-022-00809-7.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

549 49. Gupta, R.K., Kohli, A., McCormick, A.L., Towers, G.J., Pillay, D., and Parry, C.M. (2010). Full-
550 length HIV-1 Gag determines protease inhibitor susceptibility within in-vitro assays. AIDS *24*, 1651–
551 1655. https://doi.org/10.1097/QAD.0b013e3283398216.

552 50. Vermeire, J., Naessens, E., Vanderstraeten, H., Landi, A., Iannucci, V., Van Nuffel, A., Taghon,
553 T., Pizzato, M., and Verhasselt, B. (2012). Quantification of Reverse Transcriptase Activity by Real-
554 Time PCR as a Fast and Accurate Method for Titration of HIV, Lenti- and Retroviral Vectors. PloS one
555 *7*, e50859. https://doi.org/10.1371/journal.pone.0050859.

556 51. Schmidt, F., Weisblum, Y., Muecksch, F., Hoffmann, H.-H., Michailidis, E., Lorenzi, J.C.C.,
557 Mendoza, P., Rutkowska, M., Bednarski, E., Gaebler, C., et al. (2020). Measuring SARS-CoV-2
558 neutralizing antibody activity using pseudotyped and chimeric virusesSARS-CoV-2 neutralizing
559 antibody activity. Journal of Experimental Medicine *217*. https://doi.org/10.1084/jem.20201181.

560

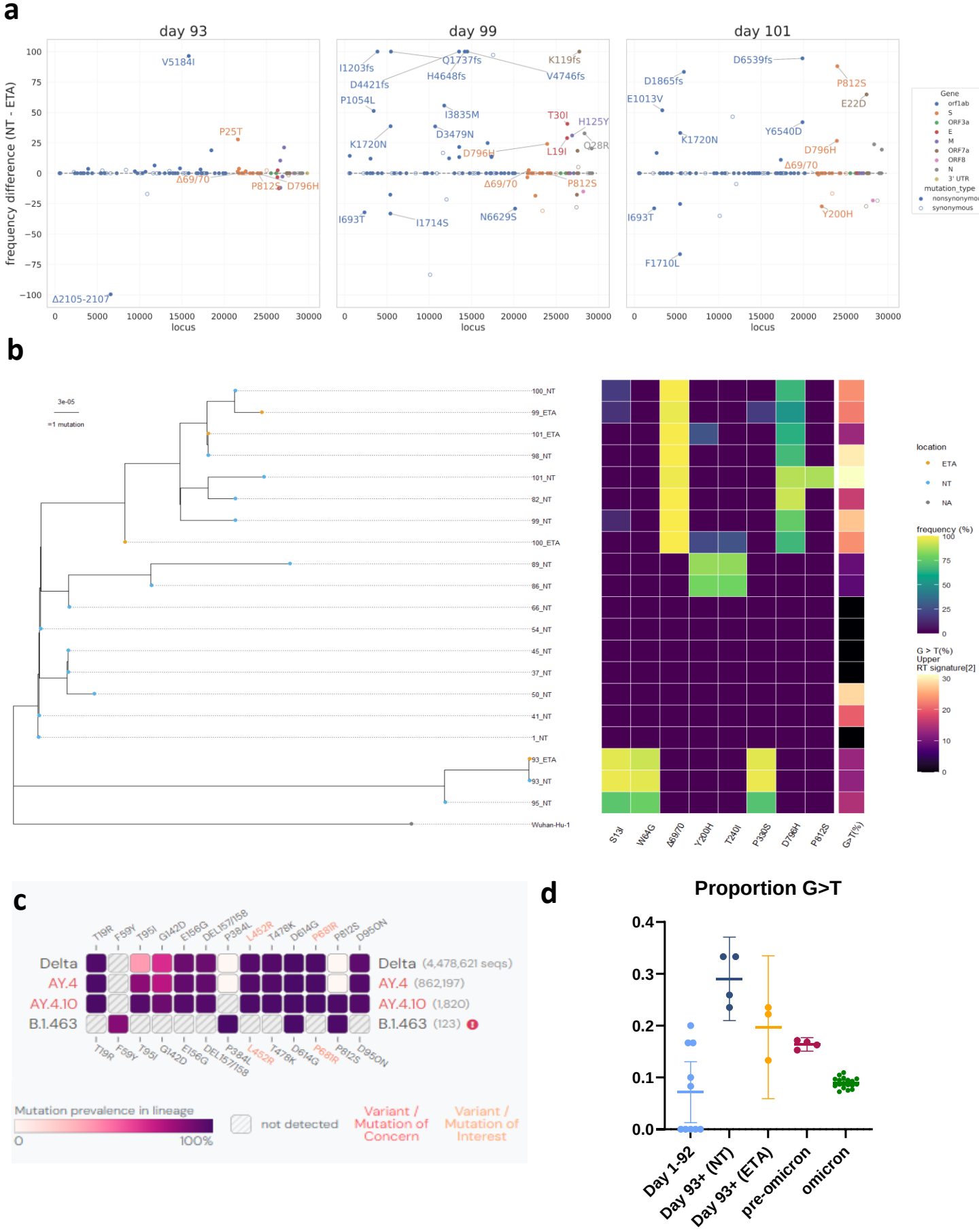

Figure 1: Genetic changes over time and across niches during chronic COVID-19

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 found in 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 onwards from 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.

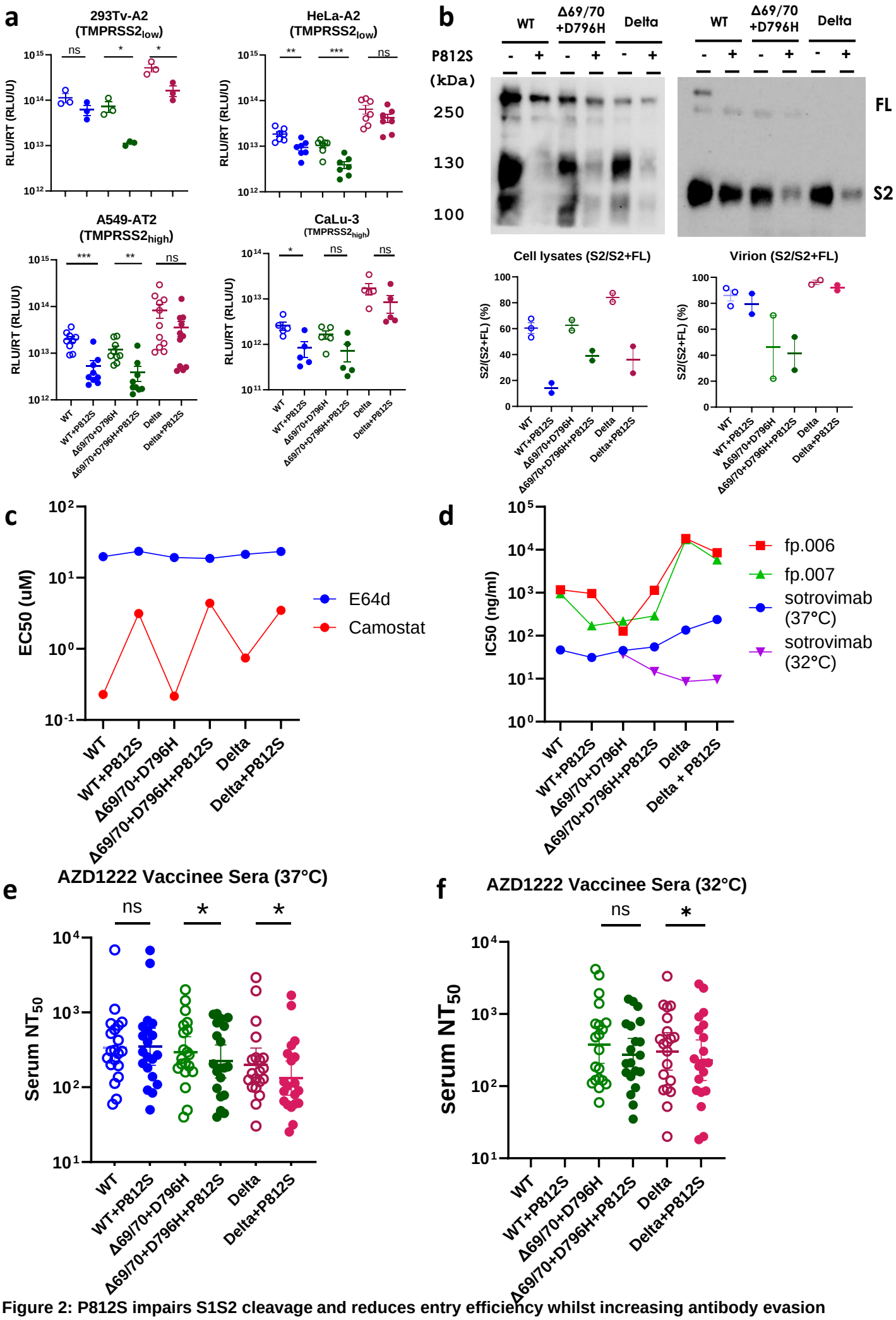

Figure 2: P812S impairs S1S2 cleavage and reduces entry efficiency whilst increasing antibody evasion

a) 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.

b) 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.

c) EC50 of E64d (cathepsin inhibitor, blue) and camostat (TMPRSS2 inhibitor, red) in A549-ACE2/TMPRSS2 using PV expression respective Spike.

d) 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.

e, f) 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 (e) or 32°C (f). 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.