

1 **Dynamics of a dual SARS-CoV-2 strain co-infection on a prolonged viral shedding**
2 **COVID-19 case: insights into clinical severity and disease duration**

3

4 Nicole Pedro MSc^{1,2,3}, Cláudio N. Silva MD^{4,5}, Ana C. Magalhães MSc^{1,2,3}, Bruno Cavadas
5 PhD^{1,2}, Ana M. Rocha MSc^{1,2}, Ana C. Moreira PhD^{1,3,6}, Maria S. Gomes PhD^{1,3,6}, Diogo
6 Silva MD⁴, Joana Sobrinho-Simões MD⁴, Angélica Ramos MSc^{4,7}, Maria J. Cardoso MD⁴,
7 Rita Filipe MD^{4,5}, Pedro Palma MD⁴, Filipa Ceia, MD^{4,5}, Susana Silva, MD^{4,5}, João T.
8 Guimarães MD-PhD^{4,5,7}, António Sarmiento MD-PhD^{1,4,5}, Verónica Fernandes PhD^{1,2}, Luisa
9 Pereira PhD^{1,2*}, Margarida Tavares MD-MPH^{1,4,5,7*}

10

11 1 i3S - Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal

12 2 Ipatimup - Instituto de Patologia e Imunologia Molecular, Universidade do Porto, 4200-135

13 Porto, Portugal

14 3 ICBAS - Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, Porto,

15 Portugal

16 4 CHUSJ - Centro Hospitalar Universitário S. João, Porto, Portugal

17 5 FMUP – Faculdade de Medicina da Universidade do Porto, Porto, Portugal

18 6 IBMC - Instituto de Biologia Molecular e Celular, Universidade do Porto, Porto, Portugal

19 7 EPIUnit - Instituto de Saúde Pública, Universidade do Porto, Porto, Portugal

20

21 *These authors contributed equally

22

23 **Keywords:** COVID-19 and SARS-CoV-2; viral shedding; co-infection; viral whole-genome
24 sequencing; polygenic risk score

25

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

26 **Running title:** SARS-CoV-2 strain co-infection on a severe case

27

28 Corresponding author:

29 Luisa Pereira

30 i3S, R. Alfredo Allen 208, 4200-135 Porto, Portugal

31 luisap@ipatimup.pt

32

33

34 **Abstract**

35 **Objectives.** A few molecularly proven SARS-CoV-2 cases of symptomatic reinfection are
36 currently known worldwide, with a resolved first infection followed by a second infection after
37 a 48 to 142-day intervening period. We report a multiple-component study of a clinically severe
38 and prolonged viral shedding COVID-19 case in a teenager Portuguese female. She had two
39 hospitalisations, a total of 19 RT-PCR tests, mostly positive, and criteria for releasing from
40 home isolation at the end of 97 days.

41 **Methods.** The viral genome was sequenced in seven serial samples and in the diagnostic
42 sample from an infected close relative. A human genome-wide array (>900K) was screened on
43 the seven samples, and in vitro culture was conducted on isolates from three late samples.

44 **Results.** The patient had co-infection by two SARS-CoV-2 strains, affiliated in distinct
45 clades and diverging by six variants. The 20A lineage was absolute at the diagnosis (shared
46 with a cohabitating relative), but nine days later the 20B lineage had 3% frequency, and two
47 months later the 20B lineage had 100% frequency. The 900K profiles confirmed the identity
48 of the patient in the serial samples, and allowed us to infer that she had polygenic risk scores
49 for hospitalization and severe respiratory disease within the normal distributions for a
50 Portuguese population cohort.

51 **Conclusions.** The early-on dynamic co-infection was the probable cause for the severity of
52 COVID-19 in this otherwise healthy young patient, and for her prolonged SARS-CoV-2
53 shedding profile.

54

55 **Introduction**

56 The new Coronavirus disease 2019 (COVID-19), caused by the severe acute respiratory
57 syndrome coronavirus 2 (SARS-CoV-2), was identified in Wuhan, China [1] and rapidly
58 disseminated at a global scale. While the world is currently in the countdown for a vaccine,
59 whether or not SARS-CoV-2 infection induces long-term protective immunity remains an open
60 question. Closely related with the immunity issue is the puzzling observation of persistent viral
61 shedding detectable in real time reverse transcriptase–polymerase chain reaction (RT-PCR)
62 tests, even after symptom resolution [2]. This extended or recurrent positivity could be due to:
63 (1) a reactivation of the virus after a period of clinical latency; (2) SARS-CoV-2 reinfection;
64 (3) or simply RT-PCR tests detecting viral remains and not necessarily active viral particles.
65 To untangle between the two first possibilities, it is necessary to perform viral whole-genome
66 sequencing and detect genetically distinct strains of SARS-CoV-2 in each of the disease
67 episodes; of course, reinfections by the same strain may remain unnoticed [3]. To check for the
68 last possibility, the inclusion of infectivity studies might help understand if the virus retains
69 both viability and integrity [4].

70 At the moment [5], a few molecularly proven SARS-CoV-2 symptomatic reinfections have
71 been published worldwide, namely in Hong-Kong [6], Belgium [7], Ecuador [8] and USA [9].
72 After being considered recovered, these four patients presented a second infection by a
73 genetically distinct SARS-CoV-2 strain after a 48 to 142-day intervening period. The latter two
74 patients displayed a more severe disease in the second episode, in accordance to the hypothesis
75 of an antibody-dependent enhancement (ADE) to SARS-CoV-2 [10]. The number of
76 reinfection cases must be higher than currently reported as they are easily missed when
77 asymptomatic. So far, two asymptomatic reinfections were detected in Indian healthcare
78 workers, whom were routinely screened in their workplace [11]. Although these two cases were
79 asymptomatic in the first and second infections, the viral load was higher in the second one, as

80 inferred from the RT-PCR cycle threshold. Another theoretical possibility is the occurrence of
81 reinfection while the first infection was not yet cured, a situation best described as co-infection.
82 This hypothesis was advanced by the authors for the USA reinfection case [9], but not
83 thoroughly studied.

84 In this work, we report the first molecularly proven dynamic early-on co-infection by two
85 genetically distinct SARS-CoV-2 strains. This was a prolonged viral shedding case (97-day
86 long), with a first severe disease manifestation, followed by a short-second hospitalisation
87 episode, in an otherwise healthy young female.

88

89 **Methods**

90 *RT-PCR and Antibody Testing*

91 For detection of viral RNA by RT-PCR, each tube sample contained a nasopharyngeal and an
92 oropharyngeal swab immersed in virus preservation solution. RNA was extracted with the
93 Qiacube extractor by using the spin-column Qiamp virus minikit (Qiagen, Hilden, Germany).
94 The reported RT-PCR results for the SARS-CoV-2 E gene were obtained with the
95 LightCycler® Multiplex RNA Virus Master (Roche Life Science, Penzberg, Germany) at a
96 LightCycler® 480 Instrument II (Roche Life Science Penzberg, Germany), and including a
97 RNA extraction control (LightMix® Modular EAV RNA Extraction Control; Tib Molbiol,
98 Berlin, Germany). Positive and negative controls were routinely included with each batch of
99 tests. Relative quantification of the sample crossing points (cp) was automatically inferred with
100 the LightCycler software.

101 The serological test used was a chemiluminescent microparticle immunoassay for qualitative
102 detection of IgG against SARS-CoV-2 nucleoprotein (Abbott Diagnostics, Chicago, USA).
103 Serum samples were run on the Abbott Architect instrument following the manufacturer's
104 instructions. The amount of IgG antibodies to SARS-CoV-2 in each sample was determined
105 by comparing its chemiluminescent relative light unit (RLU) to the calibrator RLU (index S/C).
106 A signal/cut-off (S/CO) ratio of ≥ 1.4 was interpreted as reactive and an S/CO ratio of < 1.4 was
107 interpreted as non-reactive.

108 The research protocol was approved by the Ethics Committee of the CHUSJ and University of
109 Porto Medical School (process number: 102-20). The patient and her close relative provided
110 written informed consent for publication.

111

112 *Viral Whole Genome-Sequencing and Phylogenetic Analysis*

113 Seven serial extracted viral RNA samples from the patient and the diagnosis sample from her
114 close relative were also used for the viral whole-genome sequencing. The protocol consisted
115 in: reverse-transcription with the SuperScript™ VILO™ cDNA synthesis Kit (Thermo Fisher
116 Scientific, Waltham, MA, USA); PCR enrichment of the SARS-CoV-2 genome and five human
117 gene expression controls with the Ion AmpliSeq™ SARS-CoV-2 Research Panel; library
118 construction with the Ion AmpliSeq™ Library Kit Plus; library quantification and size range
119 verification at the 2200 TapeStation Automated Electrophoresis System, using the High
120 Sensitivity DNA ScreenTape (Agilent Technologies, Santa Clara, CA, USA); next-generation
121 sequencing (NGS) on the Ion S5XL system with the Ion 530™ chip; raw data extracted with
122 the Ion Torrent pipeline. The bioinformatic pipeline consisted in: alignment of the raw data
123 versus the reference genome (accession number NC_045512.2) with BWA tool; variant calling
124 with three tools, FreeBayes, BCFtools and GATK, with editing of variants identified in at least
125 two; variant annotation with SnpEff; consensus sequence was inferred with Bcftools [12].
126 Phylogenetic analysis and lineage/clade affiliation were done by merging the consensus
127 sequences with publically available SARS-CoV-2 whole genomes from ViPR [13] with Mafft
128 [14] (first 130bp and last 50bp were masked). The rooted phylogenetic tree was obtained with
129 IQ-TREE 2 [15], and visualised using Interactive Tree Of Live version 4 [16]. The relative
130 viral load in the samples (in molarity) was corrected from the TapeStation-based library
131 quantification by removing the proportion of human reads (included in the panel and spurious
132 ones).

133

134 *Genome-Wide Array Screening and Calculation of the Polygenic Risk Score*

135 The genotyping of over 900,000 SNPs (900K) was attained with the Axiom™ Precision
136 Medicine Diversity Array and the GeneTitan Multi-Channel (MC) Instrument (Thermo Fisher
137 Scientific, Waltham, MA, USA). The laboratorial procedure was adapted to allow the use of

138 samples extracted from the nasopharyngeal swabs, by increasing the time of the whole genome
139 amplification step from 24h to 48h. Genotyping was inferred with Array Power Tool, and
140 PADRE algorithm [17] was used for accurate estimate of shared ancestry (in this case,
141 identity). These data were also used to calculate the polygenic risk score (PRS) in the
142 Portuguese population cohort (n=198) to contextualise the patient score. The PMDA variants
143 were uploaded into the Michigan Imputation Server
144 (<https://imputationserver.sph.umich.edu/index.html#!>) and imputed based on the Haplotype
145 Reference Consortium panel. The PRS values were calculated from the significant odd ratios
146 (p-value<10⁻⁵) reported online by the COVID-19 host genetics initiative
147 (<https://www.covid19hg.org/>) [18], for two phenotype cohorts from data released in 30th
148 September 2020 (COVID19-hg GWAS meta-analyses round 4 (alpha)): A2_ALL dataset of
149 “very severe respiratory confirmed covid” (n=2,972) vs. population (n=284,472); and B2_ALL
150 dataset of “hospitalised covid” (n=6,492) vs. population (n=1,012,809). Linked variants were
151 removed in plink using the flags --clump-r2 0.4 --clump-kb 250; we ended up with 91 variants
152 for the first dataset and 79 for the second. Plink was also used to estimate PRS values via an
153 additive model, as the sum of the risk alleles, weighted by the effect size estimates from the
154 genome-wide association study.

155

156 *In Vitro Culture of the Virus*

157 To ascertain if the virus was still viable in the samples after the second hospitalization, in vitro
158 culture in Vero cells (ATCC CCL-81; ATCC, Manassas, VA, USA) was performed, followed
159 by SARS-CoV-2 spike antibody (GeneTex, Irvine, CA, USA) immunofluorescence detection.
160 Images were acquired on the IN Cell Analyzer 2000 (Cytiva, Marlborough, MA, USA).
161 Following a first inoculation of the samples for 96h, the resulting supernatant was transferred
162 to 96 wells and inoculated for 24h, 48h and 72h in duplicates, to assess residual virus particles

163 not detected in the first culture. We included a recently-diagnosed sample from another patient
164 to guarantee the assay was able to detect the viral particles.

165

166

167 **Results**

168 *Clinical features*

169 A previously healthy young female presented to the local hospital emergency department in
170 the beginning of March reporting a 9-day history of sustained fever, dry cough, pleuritic chest
171 pain and vomiting. She was hemodynamically stable (105/63 mmHg blood pressure), but
172 tachypnoeic (28 cpm respiratory rate), hypoxic to 88% on room air and febrile to 101.3°F. Her
173 chest computed tomography (CT) scan revealed extensive bilateral subpleural ground-glass
174 opacities (GGO) with areas of air-space consolidation, and she was admitted for aetiological
175 investigation and treatment (Supplementary Figure S1). A nasopharyngeal swab performed
176 during the initial workup detected SARS-CoV-2 RNA (Supplementary Table S1) and the
177 patient was transferred to our referral centre for Emerging Infectious Diseases. At admission
178 lymphopenia, a mild increased level of C-reactive protein and normal prothrombin and
179 activated partial thromboplastin times were seen (Supplementary Table S2). Due to worsening
180 respiratory status and increasing supplementary oxygen demands, she was placed on High-
181 Flow Nasal Oxygen (HFNO). After a 12-hour HFNO trial without improvement, she was
182 admitted to our Infectious Diseases ICU on the 12th day of symptoms, beginning an off-label
183 5-day hydroxychloroquine course.

184 After six days of inpatient care, she complained of left upper limb pain with signs consistent
185 with deep vein thrombosis associated with indwelling peripheral venous catheter. She also
186 complained of worsening bilateral pleuritic chest pain. Accordingly, D-dimer levels increased
187 to 2.47 µg/mL and chest angio-CT scan showed lower left lung lobe peripheral infarction
188 (Supplementary Figure S2). Anticoagulation with low-molecular-weight-heparin was started,
189 the supplementary oxygen demands gradually decreased and she was discharged home 9 days
190 after admission completely asymptomatic with an oxygen saturation of 99% on room air. RT-
191 PCR tests of nasopharyngeal specimens remained positive for SARS-CoV-2 RNA at discharge.

192 The patient was asked to continue isolating at home until attaining two consecutive negative
193 tests.

194 Nearly two months after discharge, she was re-admitted with headaches, fever, myalgia and
195 right pleuritic chest pain. Her vital signs were: 110/52 mmHg blood pressure, 112 bpm heart
196 rate and 98.7°F body temperature. Her peripheral blood oxygen saturation was 97% on room
197 air, and all of the blood work values were within the normal range, non-immunocompromised
198 (Supplementary Tables S1 and S3). After an extensive microbiologic workup (Supplementary
199 Table S2), other infectious aetiologies were ruled out. The chest CT scan (Supplementary
200 Figure S3) revealed improved aeration of the lungs and resolving GGO features. Repeated
201 SARS-CoV-2 RT-PCR was still positive. The symptoms resolved in 2 days. Anti-SARS-CoV-
202 2 IgG antibodies were detected and reactive by immunoassay 77 days (relative light unit (RLU)
203 index sample/calibrator (S/C) of 7.16), 91 days (index S/C of 6.89) and again 226 days after
204 diagnosis, almost 8 months after first symptoms (index S/C of 2.26).

205 Only one household contact of the patient developed symptomatic COVID-19. In this case, the
206 symptoms were mild, limited to a one-day low grade fever, and scarce dry cough starting 10
207 days after the patient's symptom onset.

208

209 *Viral Load, Viability and Genome Analysis*

210 The relative viral load on the nasopharyngeal samples along the disease course (Figure 1),
211 inferred from the library quantification, showed a dramatic decrease (650 times lower) between
212 the time of diagnosis and day nine when the patient was first discharged from the hospital. The
213 viral load increased again by the 2nd inpatient stay (inferred from the cp values) and was still
214 relatively high two and three weeks after the second discharge (10-25 times higher than at first
215 discharge). Nonetheless, it was impossible to retrieve integral and infectious SARS-CoV-2

216 viral particles from these samples through in vitro culture (Figure 2), despite some of these
217 samples being still positive in the RT-PCR.

218 The sequencing of the viral genome in the serial samples from the patient revealed very
219 interesting results (Figure 1 and 3). At the diagnosis (sample P1.1), she displayed 100% of a
220 20A affiliated haplotype, bearing the basal C241T-C3037T-C14408T-A23403G variants (in
221 relation to the reference sequence), and additional C3140T-G24077T-C295557 variants. The
222 G24077T variant has been extensively observed in Lombardy, Italy (<https://www.gisaid.org/>)
223 [19], while the C3140T-C295557 variants were highly observed in Portugal, by the time of this
224 patient infections, especially so around her geographical region (reported online by the
225 National Institute of Health Ricardo Jorge; <https://insaflu.insa.pt/covid19/>). This viral
226 haplotype was also obtained from the RNA extracted from the diagnostic sample of the close
227 relative. Transmission is supposed to have occurred from the patient to the close relative. Other
228 cohabitants never presented symptoms, and three and half months later they had no serological
229 evidence of past SARS-CoV-2 infection. However, her asymptomatic boyfriend showed
230 reactivity to specific anti-SARS-CoV-2 IgG antibodies.

231 The second serial sample (P1.2), at day nine after diagnosis and coinciding with the first
232 hospital discharge, already revealed a mixed viral profile compatible with a 3% co-infection
233 by a basal 20B affiliated haplotype. This 3%-20B haplotype is defined by G28881A-G28882A-
234 G28883C variants, and shares C241T-C3037T-C14408T-A23403G variants with the 20A
235 haplotype (hence, these shared variants had 100% frequency in P1.2). Thus, the two infecting
236 haplotypes diverged by six SNPs. At least 12 days after the second hospital discharge (samples
237 P1.3 to P1.7), the 20B lineage had totally replaced the initially dominant 20A lineage, including
238 a sample with a negative RT-PCR result (P1.5). The samples with lower viral load (P1.2, P1.6
239 and P1.7) had other variants (Supplementary Table S4), at variable frequencies, in addition to
240 the ones defining the lineages backbones. The variants that accumulate in the patient samples

241 with lower viral load, or in other words, farther apart from the events that led to hospitalisations,
242 can be due to two factors: (1) a lower resolution of our sequencing; (2) real representatives of
243 the intra-host genomic diversity and plasticity of SARS-CoV-2 [20], a phenomenon known as
244 quasispecies [21].

245

246 *Genome-Wide Array Screening and Polygenic Risk Score*

247 We are certain that these RNA samples belong to the patient, and were not mismatched in the
248 hospital or in the laboratories, through the characterization of an array containing 900K human
249 variants. The 900K-profile was shared between the seven samples, and is unique of the patient,
250 as she has no monozygotic twin. The 900K genotyping in the patient also allowed to estimate
251 the PRS values for “very severe respiratory confirmed covid” and “hospitalised covid”, based
252 on the incipient evidence that is being collected by the COVID-19 host genetics initiative [18].
253 The patient values were contextualised in a Portuguese population cohort (Figure 3;
254 Supplementary Tables S5-S7). The PRS values for the patient were within the normal
255 distributions observed in the population cohort, almost matching the mean values.

256

257 **Discussion**

258 We illustrate a severe presentation of COVID-19 in a young healthy patient with prolonged
259 viral shedding of SARS-CoV-2. This case was extremely challenging in terms of clinical
260 diagnosis, and only the molecular study allowed us to shed light into its classification as the
261 first proven SARS-CoV-2 co-infection.

262 The affiliation of the two haplotypes in distinct clades, which emerged at different time points
263 along the pandemic (see <https://nextstrain.org/sars-cov-2/>) [22] renders it unlikely that they
264 evolved intra-host from one-another. The co-infection with a virus belonging to a different
265 clade had to occur in the 19-day period between beginning of symptoms and detection of the
266 two strains. However, we hypothesize that co-infection was already present from disease onset,
267 partially explaining the severe disease course of this healthy young female, and we simply
268 failed to detect the second strain in the first diagnostic sample.

269 The episode that led to the second hospitalisation could be speculated as corresponding to the
270 moment when the 20B strain became dominant. Unfortunately, we have no samples from this
271 period allowing us to confirm this speculation. It is a fact that the viral load was higher (10-25
272 times) in the samples after this episode comparing with those from the date of the first
273 discharge. But even so, we found no in vitro evidence that the patient was contagious at this
274 period. These results agree with previous findings [23, 24] that a later-on positive RT-PCR
275 does not imply the presence of active viral particles.

276 All this evidence seems to imply that the early-on co-infection by two SARS-CoV-2 strains
277 was the cause of the severe disease displayed by this young and healthy female. Furthermore,
278 the PRS evidence testified that she had no genetic predisposition for hospitalisation or severe
279 COVID-19, when comparing with a Portuguese population cohort. The dynamics of exchange
280 between dominant strains may have contributed to such prolonged viral shedding case. Most

281 probably the close relative was only infected by one strain, hence displaying the milder disease,
282 although we could not test it due to the unavailability of samples.

283 For sure, in the near future, several other cases of co-infection and reinfection will be identified
284 as the pandemic continues. These cases will help to clarify if a worse disease course can be
285 caused by the overlapping or sequential infection by different SARS-CoV-2 strains. The
286 authors of the paper reporting the first reinfection in the USA patient [9], who presented with
287 a second infection symptomatically more severe than the first, pointed to the possibility of co-
288 infection instead of reinfection, implying that they simply did not detect overlapping
289 “specimens A and B” in the April 2020 sample. In this case, as the patient was considered
290 recovered after the first hospitalisation, he was not isolated, rendering the hypothesis of
291 reinfection more probable. Reinfection would also agree with the more severe second COVID-
292 19 episode presented by this patient. Both this case and the one we present here highlight the
293 importance of performing more molecular studies on virus transmission dynamics. Either way,
294 co-infection or reinfection, taking into account the possibility that the immunity driven by a
295 specific SARS-CoV-2 strain does not protect against another strain, but can instead lead to a
296 more severe disease pattern, is of extreme relevance in public health.

297

298 **Transparency declaration**

299 **Conflict of interest:** No reported conflicts of interest.

300 **Financial support.** The Portuguese Foundation for Science and Technology, FCT, funded this
301 project through the Research4COVID19 projects 198_596862267, 617_613735895,
302 429_613538368 and 186_596855206. FCT also financed the PhD grant to NP
303 (SFRH/BD/136299/2018) and post-doc grant to VF (SFRH/BPD/114927/2016). i3S is
304 supported by FEDER – Fundo Europeu de Desenvolvimento Regional funds through the
305 COMPETE 2020 – Operational Program for Competitiveness and Internationalization (POCI),
306 Portugal 2020, and by Portuguese funds through FCT/Ministério da Ciência, Tecnologia e
307 Inovação in the framework of the project ‘Institute for Research and Innovation in Health
308 Sciences’ (POCI-01-0145-FEDER-007274).

309 **Acknowledgements.** We wish to thank the patients who agreed to participate in this study. The
310 authors acknowledge the support of the i3S Scientific Platforms BioSciences Screening and
311 Genomics, members of the national infrastructure PPBI-Portuguese Platform of Bioimaging
312 (PPBI-POCI-01-0145-FEDER-022122), PT-OPENSREEN, GenomePT project (POCI-01-
313 0145-FEDER-022184). And the authors also acknowledge the free information provided
314 online by the COVID-19 host genetics initiative (<https://www.covid19hg.org/>).

315 **Access to data:** Supplementary materials are available at Clinical Microbiology and Infection
316 online. Extensive files listing the variant calls for the viral whole-genome sequences can be
317 downloaded from <https://portal.i3s.up.pt/docs/bcavadas/PEDROetal.zip>. Given that the
318 genome-wide profiles from the patient samples allow individual identification, researchers can
319 request the data from the corresponding author, after justifying its need for the advancement of
320 research.

321 **Contribution:** NP performed the viral sequencing and array genotyping, with the technical
322 supervision of AMR; BC established the bioinformatics pipelines, and together with NP and

323 VF analysed the raw data; ACMagalhães, ACMoreira and MSG established the in vitro cultures
324 and performed this assay with the patient samples; CNS and MT identified the patient case,
325 contributed to clinical care and collected the clinical data; RF, PP, FC, SS and AS contributed
326 to clinical care; DS, JSS and JTG were responsible for the RT-PCR diagnosis, while AR and
327 MJC performed the serological tests, at the hospital environment; LP and MT conceptualised
328 the study, and together with NP, VF and CNS wrote the manuscript that was later edited by all
329 co-authors.

330

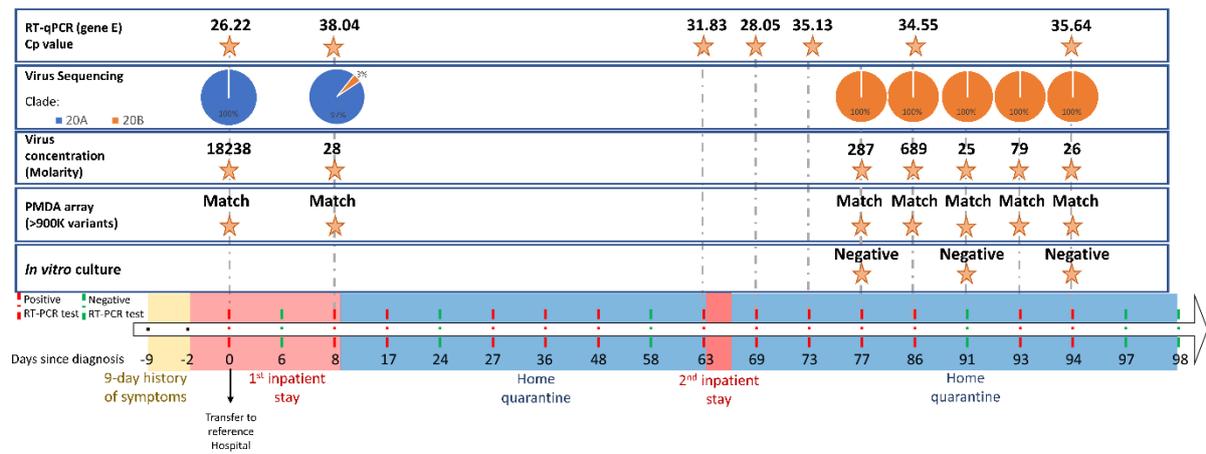
331 **References**

- 332 [1] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan,
333 X. Ma, D. Wang, W. Xu, G. Wu, G.F. Gao, W. Tan, A Novel Coronavirus from Patients with Pneumonia
334 in China, 2019, *N. Engl. J. Med.* 382(8) (2020) 727-733.
- 335 [2] B. Yuan, H.Q. Liu, Z.R. Yang, Y.X. Chen, Z.Y. Liu, K. Zhang, C. Wang, W.X. Li, Y.W. An, J.C. Wang, S.
336 Song, Recurrence of positive SARS-CoV-2 viral RNA in recovered COVID-19 patients during medical
337 isolation observation, *Sci. Rep.* 10(1) (2020) 11887.
- 338 [3] A.Y. Kim, R.T. Gandhi, Re-infection with SARS-CoV-2: What Goes Around May Come Back Around,
339 *Clin. Infect. Dis.* published online Oct 09 (2020).
- 340 [4] A. Gidari, M. Nofri, L. Saccarelli, S. Bastianelli, S. Sabbatini, S. Bozza, B. Camilloni, I. Fusco-Moffa, C.
341 Monari, E. De Robertis, A. Mencacci, D. Francisci, Is recurrence possible in coronavirus disease 2019
342 (COVID-19)? Case series and systematic review of literature, *Eur. J. Clin. Microbiol. Infect. Dis.*
343 published online Oct 10 (2020).
- 344 [5] A. Iwasaki, What reinfections mean for COVID-19, *Lancet Infect. Dis.* published online Oct 12
345 (2020).
- 346 [6] K.K. To, I.F. Hung, J.D. Ip, A.W. Chu, W.M. Chan, A.R. Tam, C.H. Fong, S. Yuan, H.W. Tsoi, A.C. Ng,
347 L.L. Lee, P. Wan, E. Tso, W.K. To, D. Tsang, K.H. Chan, J.D. Huang, K.H. Kok, V.C. Cheng, K.Y. Yuen,
348 COVID-19 re-infection by a phylogenetically distinct SARS-coronavirus-2 strain confirmed by whole
349 genome sequencing, *Clin. Infect. Dis.* published online Aug 25 (2020).
- 350 [7] J. Van Elslande, P. Vermeersch, K. Vandervoort, T. Wawina-Bokalanga, B. Vanmechelen, E.
351 Wollants, L. Laenen, E. André, M. Van Ranst, K. Lagrou, P. Maes, Symptomatic SARS-CoV-2 reinfection
352 by a phylogenetically distinct strain, *Clin. Infect. Dis.* published online Sept 5 (2020).
- 353 [8] B. Prado-Vivar, M. Becerra-Wong, J. Guadalupe, S. Marquez, B. Gutierrez, P. Rojas-Silva, M.
354 Grunauer, G. Trueba, V. Barragan, P. Cardenas, COVID-19 re-infection by a phylogenetically distinct
355 SARS-CoV-2 variant, first confirmed event in South America, *SSRN* published online Sept 8 (2020).
- 356 [9] R. Tillett, J. Sevinsky, P. Hartley, H. Kerwin, N. Crawford, A. Gorzalski, C. Laverdure, S. Verma, C.
357 Rossetto, D. Jackson, M. Farrell, S. van Hooser, M. Pandori, Genomic evidence for reinfection with
358 SARS-CoV-2: a case study, *Lancet Infect. Dis.* published online Oct 12 (2020).
- 359 [10] A.M. Arvin, K. Fink, M.A. Schmid, A. Cathcart, R. Spreafico, C. Havenar-Daughton, A. Lanzavecchia,
360 D. Corti, H.W. Virgin, A perspective on potential antibody-dependent enhancement of SARS-CoV-2,
361 *Nature* 584(7821) (2020) 353-363.
- 362 [11] V. Gupta, R.C. Bhojar, A. Jain, S. Srivastava, R. Upadhayay, M. Imran, B. Jolly, M.K. Divakar, D.
363 Sharma, P. Sehgal, G. Ranjan, R. Gupta, V. Scaria, S. Sivasubbu, Asymptomatic reinfection in two

- 364 healthcare workers from India with genetically distinct SARS-CoV-2, *Clin. Infect. Dis.* published online
365 Sep 23 (2020).
- 366 [12] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin,
367 The Sequence Alignment/Map format and SAMtools, *Bioinformatics* 25(16) (2009) 2078-9.
- 368 [13] B.E. Pickett, E.L. Sadat, Y. Zhang, J.M. Noronha, R.B. Squires, V. Hunt, M. Liu, S. Kumar, S. Zaremba,
369 Z. Gu, L. Zhou, C.N. Larson, J. Dietrich, E.B. Klem, R.H. Scheuermann, ViPR: an open bioinformatics
370 database and analysis resource for virology research, *Nucleic Acids Res.* 40(Database issue) (2012)
371 D593-8.
- 372 [14] K. Katoh, J. Rozewicki, K.D. Yamada, MAFFT online service: multiple sequence alignment,
373 interactive sequence choice and visualization, *Brief Bioinform* 20(4) (2019) 1160-1166.
- 374 [15] B.Q. Minh, H.A. Schmidt, O. Chernomor, D. Schrempf, M.D. Woodhams, A. von Haeseler, R.
375 Lanfear, IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era,
376 *Mol. Biol. Evol.* 37(5) (2020) 1530-1534.
- 377 [16] I. Letunic, P. Bork, Interactive Tree Of Life (iTOL) v4: recent updates and new developments,
378 *Nucleic Acids Res.* 47(W1) (2019) W256-W259.
- 379 [17] J. Staples, D.J. Witherspoon, L.B. Jorde, D.A. Nickerson, J.E. Below, C.D. Huff, PADRE: Pedigree-
380 Aware Distant-Relationship Estimation, *Am. J. Hum. Genet.* 99(1) (2016) 154-62.
- 381 [18] C.-H.G. Initiative, The COVID-19 Host Genetics Initiative, a global initiative to elucidate the role of
382 host genetic factors in susceptibility and severity of the SARS-CoV-2 virus pandemic, *Eur. J. Hum.*
383 *Genet.* 28(6) (2020) 715-718.
- 384 [19] S. Elbe, G. Buckland-Merrett, Data, disease and diplomacy: GISAID's innovative contribution to
385 global health, *Glob Chall* 1(1) (2017) 33-46.
- 386 [20] T. Karamitros, G. Papadopoulou, M. Bousali, A. Mexias, S. Tsiodras, A. Mentis, SARS-CoV-2 exhibits
387 intra-host genomic plasticity and low-frequency polymorphic quasispecies, *J. Clin. Virol.* 131 (2020)
388 104585.
- 389 [21] E. Domingo, J. Sheldon, C. Perales, Viral quasispecies evolution, *Microbiol. Mol. Biol. Rev.* 76(2)
390 (2012) 159-216.
- 391 [22] J. Hadfield, C. Megill, S.M. Bell, J. Huddleston, B. Potter, C. Callender, P. Sagulenko, T. Bedford,
392 R.A. Neher, Nextstrain: real-time tracking of pathogen evolution, *Bioinformatics* 34(23) (2018) 4121-
393 4123.
- 394 [23] R. Wölfel, V.M. Corman, W. Guggemos, M. Seilmaier, S. Zange, M.A. Müller, D. Niemeyer, T.C.
395 Jones, P. Vollmar, C. Rothe, M. Hoelscher, T. Bleicker, S. Brünink, J. Schneider, R. Ehmann, K.
396 Zwirgmaier, C. Drosten, C. Wendtner, Virological assessment of hospitalized patients with COVID-
397 2019, *Nature* 581(7809) (2020) 465-469.
- 398 [24] W. Li, Y.Y. Su, S.S. Zhi, J. Huang, C.L. Zhuang, W.Z. Bai, Y. Wan, X.R. Meng, L. Zhang, Y.B. Zhou, Y.Y.
399 Luo, S.X. Ge, Y.K. Chen, Y. Ma, Virus shedding dynamics in asymptomatic and mildly symptomatic
400 patients infected with SARS-CoV-2, *Clin. Microbiol. Infect.* 26(11) (2020) 1556 e1-1556 e6.

401

402 **Figures**

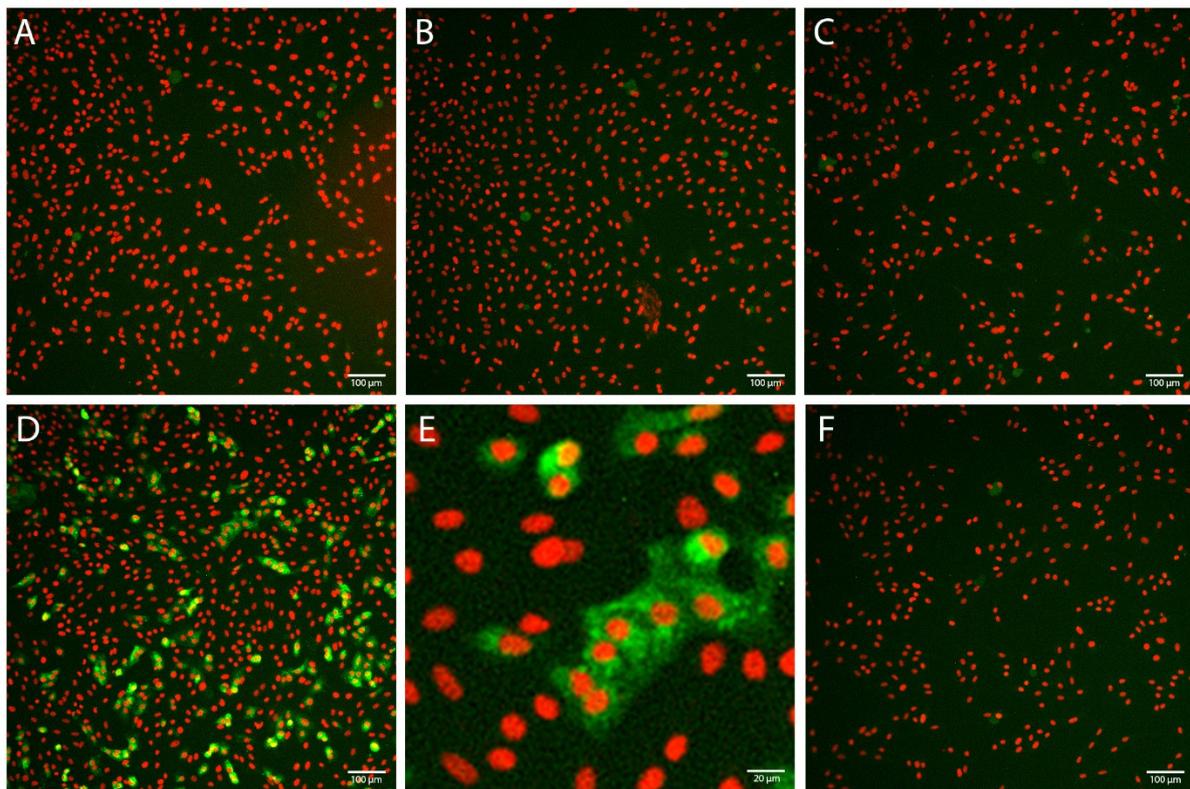


403

404 **Figure 1.** Timeline of the reported COVID-19 case and main results of the multi-component
 405 study conducted on the serial samples. The timeline bar indicates the periods of: symptom onset
 406 (yellow); inpatient stays (pink; the black arrow indicates the moment when the patient was
 407 transferred to the referral centre for Emerging Infectious Diseases); and, home quarantine
 408 (blue). It also includes the dates of all the 19 RT-PCR tests performed (green for negative result
 409 and red for positive result). The molecular tests were focused either on the virus (RT-PCR
 410 crossing point, proportion of clade/lineage affiliation inferred from viral whole-genome
 411 sequencing, virus concentration on the sequencing library, and in vitro culture) or the host
 412 genome-wide diversity (array containing >900K variants).

413

414



415

416 **Figure 2.** Results from the in vitro culture of SARS-CoV-2 from the nasopharyngeal samples.

417 Images from immunofluorescence staining with the antibody against SARS-CoV-2 spike

418 protein (green) and the marker for the nucleus (red), at 48h after the second inoculation in Vero

419 cells. Images were acquired in an IN Cell Analyzer 2000, with 10x objective. Negative results

420 in samples from the patient at 77 (A), 91 (B) and 94 (C) days after diagnosis. Positive result

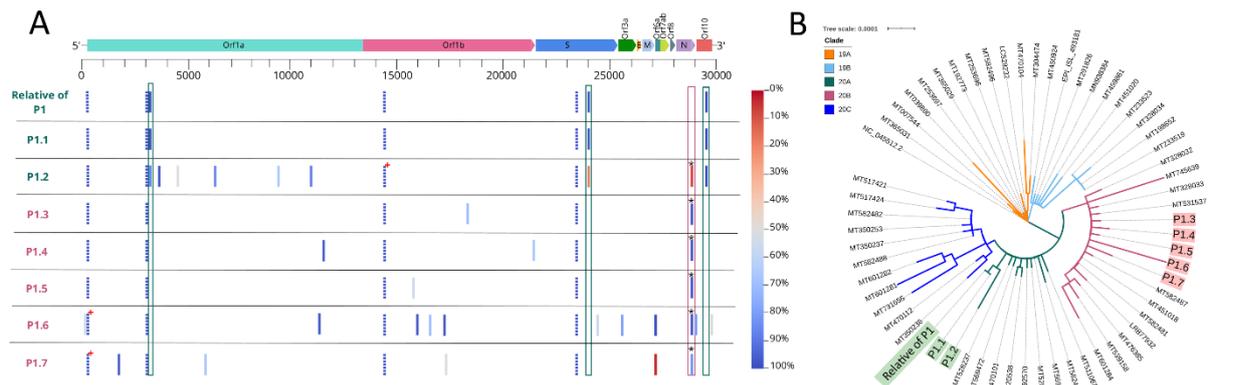
421 (D) and higher detail (E) for a recently-diagnosed sample from another patient. Negative result

422 (F) for control of non-infected cells. Bars represent 100μm or 20μm.

423

424

425

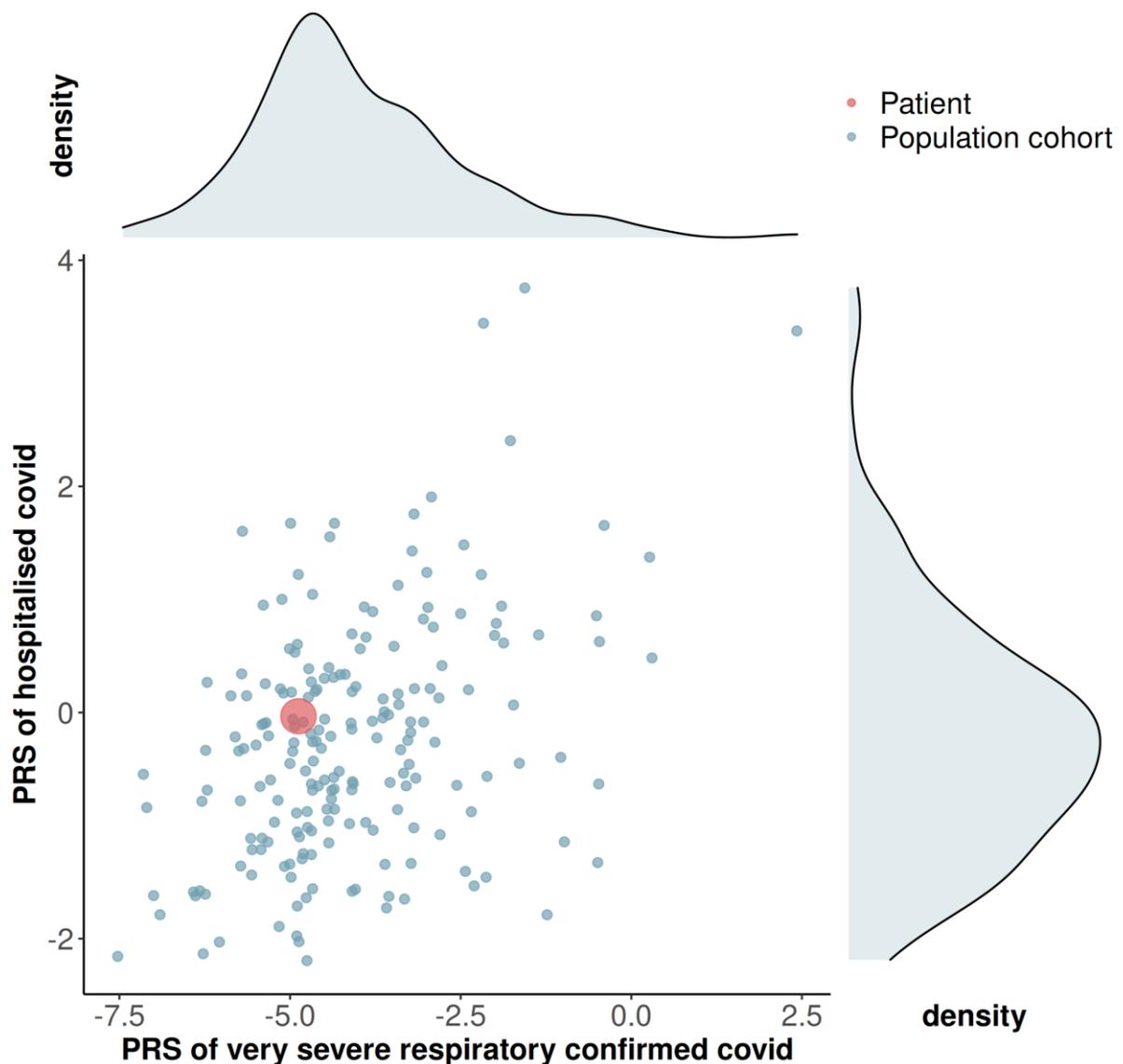


426

427 **Figure 3.** Detailed sequence diversity of the SARS-CoV-2 isolates from the various samples.

428 A- Representation of the SARS-CoV-2 genome (genes are indicated) and the variants (only
 429 those with 50% frequency in at least one of the eight samples) sequenced in the isolates from
 430 the patient serial samples (P1.1 to P1.7) and her relative diagnostic sample. The colour code
 431 reflects the main 20A (in green) or 20B (in pink) SARS-CoV-2 strain. The gradient bar
 432 indicates the frequency of the variants. For easier visualization, the shared variants between
 433 20A and 20B lineages are indicated in dotted bars. The boxes highlight either the specific 20A
 434 variants (in green) or the 20B-defining variants (in pink; the asterisk calls the attention to the
 435 three sequential variants, G28881A-G28882A-G28883C). The red-cross indicates missing
 436 positions, matching known regions of SARS-CoV-2 genome that are difficult to sequence. B-
 437 Phylogenetic tree of the main SARS-CoV-2 clades known so far (19A, 19B, 20A 20B and 20C)
 438 and the sequences reported here (following the colour scheme of A).

439



440

441 **Figure 4.** Scatter plot of the polygenic risk score (PRS) values and respective density plots for
442 “very severe respiratory confirmed covid” and “hospitalised covid” phenotypes for the
443 Portuguese population cohort (n=198), and the scatter point value for the patient (in red). The
444 PRS values were calculated based on odd ratios reported online by the COVID-19 host genetics
445 initiative (<https://www.covid19hg.org/>) [18].