

1   **SARS-CoV-2 genomic characterization and clinical manifestation of**  
2   **the COVID-19 outbreak in Uruguay**

3

4   Victoria Elizondo<sup>1,\*,#</sup>, Gordon W. Harkins<sup>2,#</sup>, Batsirai Mabvakure<sup>3,#</sup>, Sabine  
5   Smidt<sup>2</sup>, Paul Zappile<sup>4</sup>, Christian Marier<sup>4</sup>, Matthew Maurano<sup>5</sup>, Victoria  
6   Perez<sup>1,6</sup>, Natalia Mazza<sup>1</sup>, Carolina Beloso<sup>1,7</sup>, Silvana Ifran<sup>1</sup>, Mariana  
7   Fernandez<sup>1</sup>, Andrea Santini<sup>1</sup>, Veronica Perez<sup>1</sup>, Veronica Estevez<sup>1</sup>, Matilde  
8   Nin<sup>1</sup>, Gonzalo Manrique<sup>1</sup>, Leticia Perez<sup>1</sup>, Fabiana Ross<sup>1</sup>, Susana Boschi<sup>1</sup>,  
9   Maria Noel Zubillaga<sup>1</sup>, Raquel Balleste<sup>1</sup>, Simon Dellicour<sup>8,9</sup>, Adriana  
10   Heguy<sup>4,5,\*</sup>, Ralf Duerr<sup>5,\*</sup>.

11

12   <sup>1</sup> *Laboratorio de Biología Molecular, Asociación Española Primera en Salud,*  
13   *Montevideo, Uruguay*

14   <sup>2</sup> *South African Medical Research Council Capacity Development Unit, South African*  
15   *National Bioinformatics Institute, University of the Western Cape, Bellville, South*  
16   *Africa*

17   <sup>3</sup> *Department of Medicine, Johns Hopkins School of Medicine, Baltimore, Maryland,*  
18   *United States of America*

19   <sup>4</sup> *Genome Technology Center, Office for Science and Research, NYU Langone Health,*  
20   *New York, New York, United States of America*

21   <sup>5</sup> *Department of Pathology, NYU Grossman School of Medicine, New York, New York,*  
22   *United States of America*

23 <sup>6</sup>*Departamento de Desarrollo Biotecnológico, Instituto de Higiene, Facultad de*  
24 *Medicina, Udelar, Montevideo, Uruguay*

25 <sup>7</sup> Departamento de Biodiversidad y Genética. Instituto de Investigaciones Biológicas  
26 Clemente Estable, Montevideo, Uruguay

<sup>27</sup> <sup>8</sup> *Spatial Epidemiology Lab. (SpELL), Université Libre de Bruxelles, Bruxelles, Belgium*

<sup>9</sup> Department of Microbiology, Immunology and Transplantation, Rega Institute, Leuven, Belgium

30      <sup>#</sup> *Shared first authors*

31

32 \* Corresponding authors

33 Victoria Elizondo, Ph.D. E-mail: [totiep@gmail.com](mailto:totiep@gmail.com)

34 Adriana Heguy, Ph.D. E-mail: [Adriana.Heguy@nyulangone.org](mailto:Adriana.Heguy@nyulangone.org)

35 Ralf Duerr, MD, Ph.D. E-mail: [Ralf.Duerr@nyulangone.org](mailto:Ralf.Duerr@nyulangone.org)

36

37 **SARS-CoV-2 genomic characterization and clinical manifestation of**  
38 **the COVID-19 outbreak in Uruguay**

39 **Abstract**

40 COVID-19 is a respiratory illness caused by severe acute respiratory syndrome  
41 coronavirus 2 (SARS-CoV-2) and declared by the World Health Organization a global  
42 public health emergency. Among the severe outbreaks across South America, Uruguay  
43 has become known for curtailing SARS-CoV-2 exceptionally well. To understand the  
44 SARS-CoV-2 introductions, local transmissions, and associations with genomic and  
45 clinical parameters in Uruguay, we sequenced the viral genomes of 44 outpatients and  
46 inpatients in a private healthcare system in its capital, Montevideo, from March to May  
47 2020. We performed a phylogeographic analysis using sequences from our cohort and  
48 other studies that indicate a minimum of 23 independent introductions into Uruguay,  
49 resulting in five major transmission clusters. Our data suggest that most introductions  
50 resulting in chains of transmission originate from other South American countries, with  
51 the earliest seeding of the virus in late February 2020, weeks before the borders were  
52 closed to all non-citizens and a partial lockdown implemented. Genetic analyses suggest  
53 a dominance of S and G clades (G, GH, GR) that make up >90% of the viral strains in  
54 our study. In our cohort, lethal outcome of SARS-CoV-2 infection significantly  
55 correlated with arterial hypertension, kidney failure, and ICU admission (FDR < 0.01),  
56 but not with any mutation in a structural or non-structural protein, such as the spike  
57 D614G mutation. Our study contributes genetic, phylodynamic, and clinical correlation  
58 data about the exceptionally well-curbed SARS-CoV-2 outbreak in Uruguay, which  
59 furthers the understanding of disease patterns and regional aspects of the pandemic in  
60 Latin America.

61

62 **Keywords:** SARS-CoV-2; COVID-19 outbreak; Uruguay; South America; full genome  
63 amplicon sequencing; spike D614G genetic mutation; phylogeographic BEAST  
64 analysis; clinical correlations

## 65      **Introduction**

66            The novel coronavirus SARS-CoV-2, first discovered in Wuhan, China, in  
67   December 2019, rapidly extended throughout the globe, and among the >35 million  
68   confirmed positive people from >210 affected countries and territories, >1 million have  
69   died from the rapidly-spreading SARS-CoV-2 virus as of October 5th, 2020 [1].

70            Even though South America was mostly spared in the early months of the  
71   pandemic and was the last continent where it spread, it was severely hit with the arrival  
72   of the fall season to the Southern hemisphere. The virus is currently ravaging Latin  
73   America, with Brazil, Colombia, Argentina, Peru, and Mexico among the ten countries  
74   with the highest numbers of cases worldwide. In contrast, Uruguay, a small country  
75   located south of Brazil, has succeeded in maintaining a very low number of total cases  
76   (2,145) through the closing of its borders, a partial lockdown, and an early Test, Trace  
77   and Isolate (TETRIS) strategy [2]. Uruguay has a population of ~3.3 million  
78   inhabitants, of which more than a third (~1.3 million) live in the capital city of  
79   Montevideo, according to the last census in 2011 [3]. Montevideo is Uruguay's most  
80   interconnected city due to the presence of its international airport and harbor. The first  
81   positive case was officially registered on March 13th, and almost seven months later,  
82   245,000 samples (>7% of the total population of Uruguay) had been tested. In March  
83   2020, after the first four positive cases of SARS-CoV-2 were reported in Montevideo,  
84   the government issued a responsible voluntary quarantine in the country, involving the  
85   closure of schools, public entities, and businesses, urging the population to stay home  
86   [4]. On March 24th, land, maritime, and air borders were closed, allowing only  
87   Uruguayan citizens to enter the country [5]. As of October 5th, 2020, 2,145 patients  
88   have tested positive, 1,831 recovered, 243 are active cases, and 48 patients died [6].

89            Viral genome sequencing (genomic surveillance) is a powerful approach to  
90    determine the origin of pathogen introductions into a certain location and trace and track  
91    the virus's subsequent spread and evolution. It has been utilized to complement other  
92    epidemiological parameters determined by testing, contact tracing and implementation  
93    of other public health measures such as lockdowns [2,7-12].

94            Although preliminary phylogenetic analyses of Uruguayan SARS-CoV-2  
95    sequences have previously been communicated on pre-print websites [2,7], there  
96    remains a critical lack of knowledge regarding SARS-CoV-2 mutation, dispersal, and  
97    transmission patterns, and whether statistical associations exist among viral genomic,  
98    demographic, and clinical features. For example, a detailed phylogeographic analysis  
99    has not been undertaken to investigate the global dispersal dynamics with respect to  
100   Uruguay.

101          In the present study, we describe and build on genomic surveillance of a private  
102   nonprofit healthcare system in Uruguay's capital, Montevideo, where SARS-CoV-2  
103   cases have been identified since March 17th. We have sequenced and characterized 44  
104   positive cases spanning March to late May. We build on the existing Nextstrain  
105   framework, a broadly used and universally recognized analytical platform for the  
106   analysis of global SARS-CoV-2 viral genome sequence data that has accumulated since  
107   the start of the pandemic [13] to analyze newly sequenced SARS-CoV-2 genomes  
108   sampled in Uruguay, and combine it with Bayesian Evolutionary Analysis of Sampling  
109   Trees (BEAST).

110          An important aim of this study is to identify the different SARS-CoV-2  
111   introduction events in Uruguay, located in a geographical region of the pandemic that  
112   has been understudied so far. Specifically, we aim at (i) identifying and investigating  
113   the importance of independent introduction events in establishing the COVID-19

114 epidemic in Uruguay, (ii) analyzing the spatial distribution of the resulting clades in the  
115 capital city, Montevideo, (iii) looking for phylogenetic clusters within sampled  
116 institutions (hospitals, nursing homes), iv) determining whether statistically significant  
117 correlations exist among mutations, demographic, and clinical parameters.

118

## 119 **Materials and Methods**

### 120 *Bioethics, sample collection & RNA extraction*

121 The molecular lab at the Asociación Española Primera en Salud (AEPS) is a  
122 fully accredited clinical lab, regulated by the Uruguayan Ministry of Health (Ministerio  
123 de Salud Pública, MSP). Sampling was done according to the regional IRB guidelines  
124 and the recommendations of the MSP. Our study was evaluated by the commission of  
125 bioethics and ethics in research of the AEPS, headed by Fernando García [14-16]. IRB  
126 approval was waived, because the genetic analysis was restricted to the virus and not the  
127 host, and clinical correlation analyses were done on fully de-identified samples (UY ID  
128 and GISAID IDs were given by the Uruguayan and New York research teams,  
129 respectively, and are not traceable to any medical record number). The participants  
130 provided informed oral consent, and the data were analyzed anonymously. For  
131 participants <18 years of age, formal written or verbal consent was obtained from the  
132 parent/guardian at sampling, and data were kept anonymously for the entire study. The  
133 molecular clinical lab at AEPS is accredited for diagnostic RT-PCR tests for SARS-  
134 CoV-2 (COVID-19) in Uruguay. Naso-oropharyngeal swabs were collected in viral  
135 transport media and RNA was extracted using the QIAasympnhy<sup>®</sup> DSP Virus/Pathogen  
136 Mini or Midi kit (Qiagen), respectively, and confirmatory qualitative commercial RT-  
137 PCR kits were used for diagnosis and screening (depending on critical availability  
138 during the outbreak) (**Table 1**). Full details in **Supplemental Methods**.

139

140

141 Table 1. Overview of three different SARS CoV2 detection kits included in this study

Real-time RT-PCR System	Country	Regulatory Status	Target Gene(s)	Limit of Detection (copies/reaction)	Run Time (hours)	PCR Machine
LightMix Modular SARS and Wuhan CoV E/RdRp/N-gene (TIB MOLBIOL)	Germany	RUO	E, N and RdRp	5-10	1.2	LightCycler® z480 System, ROCHE
GeneFinder COVID-19 Plus Real Amp Kit (Healthcare)	Korea	HealthCanada BrazilANVISA SingaporeHSA CE-IVD	E, N and RdRp	10	2.2	CFX96 Touch Real-Time PCR Detection System, BIO-RAD
RealStar SARS- CoV-2RT-PCR kit 1.0 (Altona)	Germany	US FDAEUA CE-IVD	E, S	NA	2.25	CFX96 Touch Real-Time PCR Detection System, BIO-RAD

142

143 ***Library Preparation, Sequencing, and Read Processing***

144 To amplify the viral genomes in preparation for sequencing, we used the Swift

145 Normalase Amplicon Panel (SNAP) SARS-CoV-2 Panel (Swift Biosciences, Whole

146 viral genome single tube NGS assay, cat# SN-5XCOV296). The libraries were run on

147 an Illumina NovaSeq 6000 300 cycle flow cell, as paid end 150, using dual indices.

148 Both positive and negative control samples were also run during library prep but not

149 sequenced. The negative control sample was water, and the positive control sample

150 consisted of SARS-CoV-2 genome (Twist Biosciences, cat# 102024) serially diluted to

151 100 viral copies mixed into 50 ng of Universal Human Reference RNA (Agilent, cat#

152 18091050). Sequencing reads were demultiplexed with Illumina bcl2fastq v2.20

153 requiring a perfect match to indexing barcode sequences, and aligned to the reference

154 SARS-CoV-2 genome (NC\_045512.2, wuhCor1). Only samples with >23,000 bp

155 unmasked sequences were further analyzed, and variants were called using bcftools

156 v1.9. Details of the library generation and sequencing read processing are in the

157 **Supplemental Methods.**

158

159 ***Statistics***

160 D'Agostino & Pearson normality tests were performed to assess whether data  
161 values followed Gaussian distribution and whether parametric or nonparametric  
162 statistical tests were indicated (GraphPad Prism v.8). As a result, correlation analyses  
163 were done using nonparametric Spearman rank tests. Correlation coefficients ( $r$ ),  $P$ -  
164 values, adjusted  $P$ -values, and  $q$  values were calculated in Prism.  $P$ -values were  
165 adjusted for multiple comparisons using Holm-Sidak ( $\alpha = 0.05$ ). The false discovery  
166 rate (FDR) of  $q$  was calculated at 0.5%, 1% and 5%. Correlation analyses with a sample  
167 size of 44 had 80.7% power ( $\alpha = 0.05$ ) to distinguish correlation coefficients that differ  
168 by 0.4 standard deviation units (G\*Power 3.1.9.4).

169

170 ***Genetic analysis***

171 Sequence retrieval and multiple sequence alignment. SARS-CoV-2 reference  
172 sequences were downloaded from GISAID EpiCoV and combined with our Uruguayan  
173 study sequences in MEGA v.5.2 software, also used for sequence quality analysis,  
174 capping, and data refinement, if applicable [17]. Sequence alignments were performed  
175 using MAFFT v7.471, FFT-NS-2 method [18]. Highlighter analyses were performed on  
176 MAFFT-aligned full SARS-CoV-2 sequences from our Uruguayan study cohort with  
177 reference sequence Wuhan-Hu-1 as a master using the Highlighter tool provided by the  
178 Los Alamos HIV sequence database [19].

179

180 ***Inference of a time-scaled phylogeny***

181 To infer our SARS-CoV-2 time-scaled phylogenetic tree, we selected global  
182 reference sequences used for the Nextstrain analysis specific for South America as of  
183 August 6th, 2020. This dataset consisted of 1747 sequences sampled between December

184 26th, 2019 and July 17th, 2020, from Africa (40), Asia (194), Europe (112), North  
185 America (45), Oceania (17), and South America (128). We added 74 SARS-CoV-2  
186 Uruguayan sequences to generate an initial dataset containing 1821 viral whole-genome  
187 sequences (Table S2). Using the Nextstrain metadata to identify the accessions of  
188 interest, we then downloaded the latest whole-genome sequence alignment from the  
189 GISAID database.

190 We aligned the whole genome sequences using MAFFT v7.471 [18] and  
191 manually edited these by trimming the 5' and 3' untranslated regions and removing any  
192 gap only sites and low-quality sequences. This resulted in one low-quality Uruguayan  
193 sequence being removed, and a total of 73 remaining. The manually-edited alignment  
194 was then used to construct a maximum likelihood tree with ultra-fast bootstraps of 1000  
195 replicates in IQ-TREE version 2.0.3 [20], using the GTR+F+I nucleotide substitution  
196 model selected by Bayesian information criterion using model test implemented in IQ-  
197 TREE. TempEst [21] was used to check for outlier sequences in the tree resulting in the  
198 removal of a further ten sequences, to make up a final data set of 1810 sequences. The  
199 tree was dated using TreeTime version 0.7.6 [22], specifying a clock rate of  $8 \times 10^{-4}$   
200 substitutions per site per year to replicate the original Nextstrain workflow analysis as  
201 faithfully as possible.

202

### 203 ***Phylogeography analyses***

204 To obtain an estimate of the number of independent introductions of SARS-  
205 CoV-2 into Uruguay, we performed a preliminary phylogeography analysis using the  
206 asymmetric rates discrete diffusion model implemented in BEAST version 1.10.4 [23],  
207 adopting the fixed tree approach [11]. This approach greatly reduces the computational  
208 time required to perform the analyses by annotating the phylogeography analyses onto a

209 fixed time-scaled phylogenetic tree. Our model considered two discrete ancestral  
210 location states, i.e., Uruguay and non-Uruguay, and specified a Markov chain Monte-  
211 Carlo (MCMC) length of 1 million steps, sampling every 200 steps to produce a  
212 posterior distribution of trees containing 5,000 trees. The time-scaled maximum clade  
213 credibility tree (MCC) tree from the discrete model phylogeography analysis conducted  
214 in BEAST was then identified using TreeAnnotator, available as part of the BEAST  
215 package and visualized in FigTree v.1.4.4 [24]. The BEAST log files were inspected for  
216 convergence using Tracer version 1.7.1 [25]. All model parameters achieved effective  
217 sample size (ESS) values >200 indicating sufficient mixing and convergence to  
218 stationary.

219 To estimate the potential regional source(s) of the independent introductions into  
220 Uruguay, we then replicated this phylogeography analysis, but this time considered  
221 seven discrete ancestral location states, including Uruguay, Africa, South America,  
222 North America, Oceania, Asia, and Europe.

223

#### 224 ***Software Scripts and Visualization***

225 See Supplemental Methods.

226

#### 227 **Results**

##### 228 ***Study population and clinical parameters***

229 A total of 44 diagnosed positive COVID-19 participants were included in this  
230 study, 25 men and 19 women with a comparable mean age (and range) of 54 (15-92)  
231 and 59 (24-89) years, respectively. Sixty eight percent of the participants did not require  
232 hospitalization, 18% were hospitalized or received in-home care (9% on ventilation and  
233 5% on ICU), and 14% were deceased. Whereas the early COVID-19 positive cases in

234 March were predominantly clade S infections, clade G viruses subsequently became  
235 dominant in April and May 2020. (**Figure 1 A-C**).

236 Our cohort's cases came from 15 neighborhoods in the capital, Montevideo. The  
237 majority of the samples came from the central neighborhoods of Cordón and Pocitos,  
238 where the two biggest outbreaks within our cohort occurred (**Figure 1B and Figure**  
239 **S1**). Only two participants (from the Montevideo neighborhoods Pocitos and Prado)  
240 reported a possibility of travel-related infection after returning from Japan.

241

242 ***Diverse SARS-CoV-2 mutation profiles with increased prevalence of spike D614G***  
243 ***variants and associated mutations in the evolving epidemic***

244 The SARS-CoV-2 genetic variants in our study group are very diverse and  
245 comprise sequences from clades S, V, G, GH, and GR (**Figures 1A and 2A, Table S1**).  
246 We found 446 mutations in our Uruguayan study sequences compared to the reference  
247 Wuhan.Hu.1 sequence [26], including three gaps. 313 were SNPs (non-amino acid  
248 changing), scattered across 60 positions, and 130 were amino acid-changing, scattered  
249 across 32 positions of the open reading frames (**Figure 2A, B**). We observed between 5-  
250 12 single nucleotide polymorphisms (SNP) resulting in 1-7 amino acid (AA)  
251 replacements per viral genome (**Figure 2B, Figures S2 and S3**). In chronologically  
252 sorted mutation/highlighter plots, various mutation patterns appeared in the first 2/3 of  
253 the study period, whereas in the last 1/3, the mutation patterns became more  
254 homogenous. In March and early April, ORF8 mutation L84S and the associated  
255 C8782T single nucleotide polymorphism (SNP) were most abundant (>1/3 of  
256 sequences) (**Figure 2A, C, Figures S2 and S3**). Both mutations have become known as  
257 clade S-defining mutations (mostly in sublineage A.5) [27]. In addition, we found three  
258 more SNPs to be significantly associated, namely C17470T, C25521T, and C26088T

259 that all together build a strongly significant, positive correlation cluster composed of  
260 five mutations (**Figure S4**). Interestingly, after April, most sequences belonged to clade  
261 G (sublineage B.1), defined by the spike protein's D614G mutation. 57% of our  
262 cohort's sequences contain the D614G mutation co-occurring with SNPs C241T,  
263 C3037T, and C14408T, the latter causing the AA replacement P323L in the RNA-  
264 dependent RNA polymerase (RdRP) gene of nsp12 (**Figure 2A, C, Figures S2 and S4**).  
265 As known from other global studies [27], clade S and G-associated mutations are  
266 largely exclusive, resulting in strongly inverse correlation clusters (**Figure S4**).  
267

#### 268 ***Phylogenetic assessment of the regional SARS-CoV-2 outbreak***

269 To determine phylogenetic and epidemiologic characteristics of the Uruguayan  
270 SARS-CoV-2 outbreak, we performed a comprehensive set of phylogenetic analyses,  
271 including maximum-likelihood trees, haplotype networks, Nextstrain-based  
272 phylogenetic placements, and Bayesian phylogeographic analyses using BEAST  
273 (**Figures 3 and 4, Figures S5-S8**). Focusing on all available Uruguayan SARS-CoV-2  
274 sequences that passed our internal and GISAID's quality assessment (n=73), we  
275 observed an intermixing of our (44) and other (29) Uruguayan study sequences, both in  
276 maximum-likelihood IQ trees and in genetic-distance-based haplotype networks  
277 (**Figure S5**). Consistent with the mutational analysis of our internal data set (**Figure 2**),  
278 the phylogenetic analysis of the publicly available Uruguayan sequences reveals a  
279 predominance of clades S and G (G, GH, and GR), the former occupying most of the  
280 upper half of the maximum-likelihood tree (**Figure S5A**) and the right half of the  
281 haplotype network (**Figure S5B**), the latter the respective opposite halves, highlighted  
282 by green and red arrows for spike D614G and ORF8 L84S key mutations, respectively.  
283 The allocation of neighborhood data on the phylogenetic tree indicated regionally

284 clustered appearances of phylogenetically related viruses, as most evident for Carrasco,  
285 Pocitos, Malvin, Reducto (all in Montevideo), and Rivera (**Figure S5A**).

286

287 ***Phylogeographic BEAST analysis reveals 23 introductions into Uruguay mostly from***  
288 ***surrounding South American countries, resulting in five clusters***

289 To assess global introductions and the regional spread of SARS-CoV-2 in  
290 Uruguay, we performed discrete phylogeographic analysis using BEAST with all 73  
291 Uruguayan sequences, complemented with 1737 global reference sequences, based on  
292 the global subsampling dataset suggested by Nextstrain (**Table S2**). Genetic distance-  
293 based haplotype networks indicated genetic relationships of Uruguayan sequences with  
294 those from other South American countries, Europe and Asia (**Figure S6**).

295 By explicitly considering the time and sampling locations of our sequences,  
296 BEAST analyses revealed further important details about the evolutionary relationships  
297 of our sequences. The time-scaled maximum clade credibility tree (MCC) generated by  
298 the discrete phylogeographic analysis of our 1810 SARS-CoV-2 sequences is presented  
299 in **Figure S7** and shows a minimum of 23 independent introductions of the SARS-CoV-  
300 2 virus occurred. Collectively, these introduction events included representatives from  
301 seven of the GISAID clades (G, GH, GR, L, O, S and V) circulating worldwide and all  
302 were imported to the capital city Montevideo with the exception of one imported to  
303 Rivera.

304 The results of the discrete phylogeography analysis we performed that  
305 considered seven ancestral state locations further revealed that 18 of the 23 independent  
306 introductions are inferred to have originated from other South American countries. The  
307 remainder include two independent introductions of GISAID clade GR viruses from  
308 Asia, two from North America, and one clade V virus from Oceania (**Figure 3**). The

309 viral sequences from the only two participants in our cohort that reported a possibility  
310 of travel-related infection after returning from Japan are inferred to have been imported  
311 from Asia and grouped on the tree within separate clades of Asian sequences. The  
312 estimated time to the most recent common ancestor (TMRCA), represented by the age  
313 of the root node of the entire tree, is in mid-December 2019 (9<sup>th</sup>-21<sup>st</sup> of December),  
314 which is in agreement with the estimated origin of the pandemic in the Hubei Province  
315 in China, sometime between October and December, and the first contracted case in  
316 China recorded in mid-November [28-31].

317 Only five of the 23 independent virus introductions into Uruguay resulted in  
318 monophyletic clades with more than two sampled sequences in the country. These  
319 clades comprised 21, 18, 5, 4, and 3 sequences, respectively (**Figures 3 and 4, Figure**  
320 **S8**), suggesting that these viral outbreaks were maintained by community transmission  
321 once introduced into Montevideo.

322 To investigate the timing of the introduction of the viruses that founded these  
323 five main clades circulating in Montevideo, we estimated the time of their most recent  
324 common ancestor (TMRCA), acknowledging that the actual introduction events likely  
325 occurred even before the corresponding TMRCAs.

326 The TMRCA of the first main clade of 21 sampled sequences, highlighted in red  
327 in **Figure 4**, was estimated to fall between the 2<sup>nd</sup> and 5<sup>th</sup> of March, 2020 and involved  
328 the importation of a GISAID clade S virus. Viruses within this transmission cluster  
329 were restricted to Montevideo, where they were distributed among nine of the local  
330 neighborhoods, including two Hospitals and one research institute (**Figure S8**). A North  
331 American sequence from Mexico was positioned basally to this clade on the tree  
332 identifying this country as the most plausible source location.

333           The TMRCA of the second main clade, which comprised of 18 virus sequences,  
334         highlighted in yellow in **Figure 4**, was estimated to fall between the 13<sup>th</sup> and 17<sup>th</sup> March  
335         2020 and involved the import of a GISAID clade G virus. Viruses within this  
336         transmission cluster were restricted to the city of Montevideo, where they were  
337         distributed among nine neighborhoods and included samples from two nursing homes  
338         and one hospital (**Figure S8**). The viral sequences in this clade were inferred to have a  
339         South American origin.

340           The TMRCA of the third main clade of five sampled sequences, highlighted in  
341         blue in **Figure 4** and corresponding to a GISAID clade GR, was estimated to fall  
342         between March 20<sup>th</sup> and May 26th, 2020. This transmission cluster consisted of five  
343         viruses from the Hospital de Rivera in Rivera, a small city situated on the border with  
344         Brazil (**Figure S8**). This clade was inferred to have originated in South America and  
345         groups with a sequence from Brazil on the MCC tree suggesting this was the source  
346         location for this viral introduction

347           The TMRCA of the fourth main clade with four sampled sequences, highlighted  
348         in purple in **Figure 4**, was estimated to fall between March 14<sup>th</sup> and May 20<sup>th</sup>, 2020 and  
349         involved the introduction of a GISAID clade G virus into Montevideo (**Figure 4**). This  
350         clade comprised four viruses from within two Montevideo neighborhoods (**Figure S8**).  
351         The virus responsible for this introduction was also inferred to have originated in South  
352         America.

353           The TMRCA of the fifth main clade of three sequences, highlighted in green in  
354         **Figure 4**, was estimated to have occurred between 22<sup>nd</sup> and the 24<sup>th</sup> March 2019 and  
355         involved the introduction of a GISAID clade GR virus into Montevideo. This clade  
356         comprised three viruses from within a single Montevideo neighborhood (**Figure S8**).  
357         Collectively, sequences from the four main clades described above were sampled

358 between March and May 2020 and were distributed among 17 of the 57 neighborhoods  
359 or *barrios* in Montevideo.

360 All of these five independent introduction events that formed sustained  
361 transmission chains identified here were estimated to have occurred close to the time of  
362 the first officially diagnosed case in Uruguay on March 13<sup>th</sup>, 2020 and the putative date  
363 of origin of the pandemic in Uruguay on March 7<sup>th</sup>, which was thought to be introduced  
364 by a single female traveler who arrived in Montevideo on a flight from Italy and  
365 subsequently attended a wedding reception in the city that was attended by over 500  
366 guests.

367

368 ***Distribution of SARS-CoV-2 within Uruguay***

369 Within these five clades we identified a total of ten sequence clusters (sequences  
370 from the same institution that group together on the tree), spread across six of the seven  
371 health institutions from which we had more than one sample (**Figure S8**). These  
372 included one cluster of five GISAID clade GR sequences in the Hospital de Rivera (red  
373 dots), two clusters of two clade G sequences in the Hospital Vilardebó (green dots),  
374 three clusters each comprised of two clade S sequences in the Institut Pasteur (brown  
375 dots), and two clusters from the Asociación Española Primera en Salud (blue dots)  
376 comprising two S and two G clade sequences (**Figure S8**).

377

378 ***Clinical correlations separate from mutational correlation clusters***

379 The availability of study participants' clinical and demographic data combined  
380 with mutational data of the infecting SARS-CoV-2 viruses enabled us to perform  
381 comprehensive correlation analyses (**Figures 5 and 6, Figures S9-S11**). Overall,  
382 Spearman rank correlation analyses revealed six prominent correlation clusters of

383 significant positive or mixed correlations (**Figure 5, Figure S9**). Four of the highlighted  
384 clusters in **Figure 5** (clusters 1-4) are dominated by mutations that form mixed clusters  
385 with demographic parameters such as sampling location or treating healthcare  
386 institution and a few other parameters. Clusters 5 and 6 are different in forming separate  
387 clinical correlation clusters together with the parameters “sex” (cluster 5) or “age”  
388 (cluster 6) without essential associations to virus mutations. Cluster 5 reveals a  
389 significant association of female sex with clinically asymptomatic courses of the disease  
390 and a lower risk of developing fever. Cluster 6 indicates a network of positive  
391 correlations among age, lethal outcome, and five clinical parameters. In addition to the  
392 highlighted clusters, we observed some smaller clusters, mainly composed of mutations.  
393 The tight network of positive clinical correlations and a more outspread correlation  
394 network of clades with regional appearances and selected demographic parameters are  
395 shown in greater detail in **Figure 6A**. Specifically, statistical analyses of lethal outcome  
396 as study parameter revealed significant positive correlations with arterial hypertension  
397 (AHT), kidney failure and ICU admission complemented by borderline-significant  
398 associations with additional clinical parameters (hospitalization, diabetes mellitus II,  
399 and obesity) and age, but no association with any specific mutation (**Figure 6B, Figure**  
400 **S10**). Accordingly, the spike D614G mutation and clade G-related viruses, in  
401 consequence, are not associated with any clinical parameters, severity or lethality.  
402 D614G only correlates with co-occurring/inversely occurring mutations, treating  
403 healthcare institutions, and time since sampling started (**Figure 6C, Figure S11**).  
404 Fatality rates among clade G, GR, and S-infected individuals were comparable at 18%  
405 (3/17), 14% (1/7), and 12% (2/17), respectively.

406

407 **Discussion**

408           The Uruguayan epidemic is characterized by an early clade S dominance that  
409           was subsequently replaced by clade G variants (**Figures 1 and 2**), which matches the  
410           global trend [27,32,33]. Uruguayan clade S viruses are characterized by the key  
411           mutations T28144C, causing the AA replacement L84S in ORF8, and C8782T, in  
412           agreement with global clade S strains [27]. In our data set, we observed three additional  
413           co-occurring SNPs in ORF1b and ORF3 that were present in 15 out of 17 clade S  
414           variants, i.e., C17470T, C25521T, and C26088T. These mutations are less common and  
415           are presumably characteristic of the regional outbreak. Further studies need to show  
416           whether this subclade will be fixed in regional and/or supraregional epidemics, and  
417           whether founder effects or functional features accounted for the early clade S  
418           dominance over the original clade as well as the subsequent fluctuating prevalence and  
419           decline. According to the CoV-GLUE database, ORF8 L84S is the 8th most prevalent  
420           AA replacement to date [34]. After a controversial debate about the ancestry of L and S  
421           types and the functional impact of L84S [35-38], the still limited amount of data  
422           indicates that L84S might confer selection advantage and render the virus more virulent  
423           based on destabilizing the immuno- and replication-modulatory protein ORF8 and  
424           mitigating binding of ORF8 to human complement C3b [35,37-39].

425           End of March/early April 2020, we observed a subsequent switch in dominance  
426           from clade S to clade G-variants (G, GR, and GH) (**Figures 1 and 2**). It positioned  
427           Uruguay somewhere in the middle in the asynchronous transition process from spike  
428           D614 to G614 virus predominance, i.e., between the early European and the mostly late  
429           Asian countries [32,33]. Diverse structural and functional assays strongly suggest that  
430           the spike D614G mutation renders SARS-CoV-2 more infectious by stabilizing its  
431           structure, i.e., through impact on the, compared to SARS-CoV-1, even more fragile  
432           metastable SARS-CoV-2 spike protein, thus reducing the shedding of the S1 subunit.

433 Furthermore, D614G triggers higher spike numbers on the virion surface and induces a  
434 more open, receptor-binding domain (RBD)-up spike conformation toward a receptor-  
435 binding and fusion-competent state [33,35,37,40-44]. In D614G spikes, binding to the  
436 angiotensin-converting enzyme 2 (ACE2) receptor is not increased, and SARS-CoV-2  
437 viruses do not acquire D614G escape mutations *in vitro* under neutralizing antibody  
438 immune pressure [40,45]. Instead, D614G increases neutralization susceptibility of  
439 SARS-CoV-2, which assures high sensitivity to vaccination-induced neutralizing  
440 antibodies [46,47].

441 Although D614G serves as the clade G-defining mutation with likely effects on  
442 virus infectivity/transmissibility, D614G is governed by a very strict co-appearance with  
443 C241T, C3037T, and C14408T, both in Uruguayan and global G variants (**Figure 5**,  
444 **Figure S5**) [27]. Notably, in addition to the D614G-causing A23403G mutation in  
445 spike, C14408T is responsible for the AA replacement P323L in the RdRp gene. Based  
446 on their central roles in viral entry and replication, their co-evolution is of particular  
447 interest, and a mutual contribution to the selective advantage of G-haplotypes is  
448 assumed [33]. Interestingly, none of the viruses harboring a single clade G mutation  
449 prevailed to achieve epidemiological relevance, e.g., D614G alone or P323L alone have  
450  $\leq 0.3$  global prevalence, whereas D614G and P323L together have  $\sim 70\%$  global  
451 prevalence as of August 2020 [33]. P323L, although not located in the active center,  
452 possibly influences RdRp fidelity through allosteric effects at the nsp12 interface with  
453 the nsp8 cofactor and might increase the viral mutation rate [33,39,48 ]. Thus, the  
454 strong correlation patterns between key mutations in our Uruguayan data set, mirroring  
455 global patterns, allows us hypothesize that coupled mutations, such as D614G in spike  
456 and P323L in RdRp might synergize for the epidemiological success of the virus. It  
457 may allow a fine balance between efficient transmission (e.g., by D614G, even in

458 asymptomatic cases) and limited clinical presentation (e.g., by P323L, decreasing the  
459 production of viral RNA) to eventually attenuate an aggressive virus that, as shown for  
460 MERS and SARS-CoV-1, is more vulnerable to viral clearance with lack of long-term  
461 epidemiological success. Further studies are needed to determine how the increasing  
462 diversity of mutation patterns influence the fitness and reproduction of viral  
463 populations, the susceptibility/evasion to immune responses and treatment, and how  
464 these mutations are selected in the human body or during transmission.

465 Beyond the functional relevance of emerging and transmitted mutations, they  
466 serve as a fine tool to dissect population phylodynamics, transmission chains and  
467 epidemiologic clusters. The number of independent introduction events into a particular  
468 country as a proportion of the total number of sequences in the data set is considered a  
469 rough measure of the relative influence of intercontinental and international travel on  
470 the subsequent epidemiological dynamics within that country. For Uruguay, with 23  
471 identified independent migration events out of a total of 73 viral sequences in our  
472 dataset, this proportion is relatively low if compared to other countries such as Belgium  
473 (331/740) [11] but higher than others like Brazil (>100/490) [8]. It possibly indicates  
474 that the relative influence of intercontinental and international travel has been less  
475 important in driving the dispersal dynamics of the Uruguay outbreak compared to other,  
476 larger or more connected countries like Belgium. The town of Rivera, situated on the  
477 border with Brazil, has been a main concern for the government and many outbreaks  
478 occurred along the border that were rapidly brought under control. In line with our  
479 BEAST data that suggested a single transmission from Brazil as the main source of the  
480 Rivera infections, the first Rivera outbreak was reported to have begun when a COVID-  
481 19-positive Uruguayan was diagnosed on May 7<sup>th</sup>, 2020. Our BEAST analysis strongly  
482 supports the former hypothesis from the Uruguayan Government and Ministry of Health

483 (MSP) that a local metallurgic worker that used to travel to Brazil everyday had been  
484 infected with SARS-CoV-2 in Brazil and introduced this strain to Rivera [49,50].

485 There are two main reasons for why the extent of the geographical distribution  
486 and the density of viruses in each neighborhood within Montevideo and Rivera  
487 responsible for the five main clades of SARS-CoV-2 circulating in Uruguay (**Figure 4**,  
488 **Figure S8**) represent an underestimate of the true values of both these variables. First,  
489 our 73 Uruguayan sequenced viral genomes represent only a relatively small fraction of  
490 the total number of infections that occurred in the actual outbreak seeded by these  
491 viruses, with estimates obtained from contact tracing efforts suggesting that collectively  
492 these viruses infected at least 364 individuals [2]. Secondly, 39 (53%) of our Uruguayan  
493 samples were either collected in hospitals (Hospital de Rivera, Asociacion Espanola -  
494 AEPS, or Hospital Vilardebó) where the infected patient was treated or where the  
495 infection was acquired, in nursing homes, or in research institutes (Institut Pasteur)  
496 where samples were processed. In these cases, the home address of the patient samples  
497 is not publicly available. This precluded the adoption of the continuous diffusion  
498 phylogeography model [51] that makes use of the actual geographic coordinates of the  
499 samples to infer the transmission links among the sampled locations in the various  
500 Montevideo neighborhoods, and instead, limited our analysis to describing the spatial  
501 relationships among the sampled SARS-CoV-2 sequences.

502 Our discrete phylogeographic analysis provided a new perspective to the  
503 believed origin of the Uruguayan SARS-CoV-2 outbreak from overseas by a single  
504 traveler returning from Europe [52-54]. While overall similarities of Uruguayan  
505 sequences with European and Asian sequences were observed in mutation patterns and  
506 genetic distance-based haplotype networks, our phylogeographic analyses using BEAST  
507 indicate that Uruguayan introductions that resulted in outbreaks were mostly restricted

508 to neighboring South American countries (**Figure 3**), which stands in contrast to the  
509 suggestions of a recent preprint article [7]. Instead, our data support the idea that the  
510 outbreaks that were seeded after return from overseas travel, were contained  
511 successfully through social distancing, mask wearing, rigorous testing, contact tracing,  
512 and partial lockdown [2].

513 Having determined the regional SARS-CoV-2 mutation patterns and  
514 phylogeographic spread, the question remained whether and to what extent genomic  
515 features are coupled to demographic and/or clinical parameters. Our correlation data  
516 revealed significant clusters of co-occurring or mutually excluding mutations with  
517 regionally accumulated appearances of clades/mutation clusters (**Figures 4 and 5**,  
518 **Figure S5**). In contrast, the large bulk of clinical parameters clustered separately  
519 without major influence from viral mutations (**Figures 5 and 6, Figure S10**), which is  
520 in line with a recent publication that reported no significant impact of genetic variation  
521 on clinical outcome [55]. More specifically, we studied pairwise correlations with the  
522 spike D614G mutation, which, because of perfectly matching mutation patterns, also  
523 represents correlation analyses for the RdRp P323L mutation or G-related clades (G,  
524 GR, and GH). In an FDR-adjusted analysis, the presence of D614G mutation was  
525 coupled to other viral mutations, treatment of the infected patients in a regional  
526 healthcare institution, and late sampling, but not to clinical parameters (**Figures 5 and**  
527 **6, Figure S11**). There have been controversial reports of D614G being associated with  
528 higher fatality rates and/or severe illness in a few data sets [56,57], whereas more recent  
529 data suggests no correlations of D614G with clinical outcome [33], the latter supporting  
530 our findings. Our analysis of associations with clinical parameters pointed at significant  
531 associations between lethal outcome and arterial hypertension, kidney failure, and ICU

532 admission, which mirrors clinical studies on associated factors or predictors of disease  
533 severity/progression [58,59].

534 In sum, our characterization of Uruguayan SARS-CoV-2 phylogenetics,  
535 mutation patterns, and their correlation with demographic and clinical parameters did  
536 not identify critical viral attenuations or clinical peculiarities that can primarily account  
537 for the exceptionally well curbed regional COVID-19 epidemic [2]. It instead suggests  
538 that socio-epidemiologic mitigation strategies managed to curtail COVID-19 to  
539 restricted regional transmission clusters in Uruguay.

540 We hope that these findings contribute to define the South American COVID-19  
541 outbreak better, to optimize and develop efficient, fast, and low-cost mitigation  
542 strategies and diagnostic pipelines for Uruguay and other countries, and to assist  
543 physicians dealing with strategies for this and future emerging infections.

544

#### 545 **Acknowledgments**

546 The authors thank all participants who agreed to participate in this study and the  
547 healthcare personnel in Uruguay for their dedication to COVID-19 patients' care. We  
548 wish to acknowledge the support of New York University's Data Services, Bobst  
549 Library, and, in particular, the expertise shared by Christopher Schwarz and Senior  
550 Academic Technology Specialist Denis Rubin in software script development. We also  
551 acknowledge the support of the NYU Langone Health High-Performance Computing  
552 resource and the Laura and Isaac Perlmutter Cancer Center, which partially supports the  
553 Genome Technology Center. We thank the scientists across the world who deposited  
554 SARS-CoV-2 sequences in GISAID, especially those who deposited 29 Uruguayan  
555 sequences that supplemented our 44 sequences. We would also like to thank Flavia  
556 Camacho for administrative assistance.

557

558 **Disclosure Statement**

559       The authors declared no potential conflicts of interest with respect to the  
560 research, authorship, and/or publication of this article.

561

562

563 **Funding**

564       All the work at the Laboratorio de Biología Molecular AEPS was supported by  
565 internal funding from the Asociación Española Primera en Salud. SD is supported by  
566 the *Fonds National de la Recherche Scientifique* (FNRS, Belgium). RD was partially  
567 supported by the NIH grant 1R01AI122953. AH, CM, and PZ are supported by the  
568 Genome Technology Center, and in part by the Cancer Center Support Grant  
569 P30CA016087 at the Laura and Isaac Perlmutter Cancer Center.

570

571 **Contributors**

572       VE1, AH, and RD conceived research goals, experiments, and analyses. GWH,  
573 BM, SS, VP, and RD performed formal analyses. VE1, RB, and AH acquired funding  
574 for the project. PZ, CM, VP, NM, CB, SI, MF, VP, VE2, GM, AS, FR, MN, LP, SB,  
575 and MNZ performed experiments and/or were involved in sample acquisition. BM,  
576 MM, GWH, SD, AH, and RD developed methodologies, designed, and/or implemented  
577 computer codes. VE1, GWH, MM, SD, RB, AH, and RD supervised the research  
578 activity and validated the research output. GWH, BM, SD, VP, and RD prepared figures  
579 and tables. VE1, GWH, AH, and RD wrote the manuscript. All authors reviewed and  
580 edited the manuscript.

581

582    **References**

- 583    1. CSSE. Center for Systems Science and Engineering (CSSE) at Johns Hopkins  
584       University 2020. Available from:  
585       <https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6>
- 586    2. Moreno P, Moratorio GA, Iraola G, et al. An effective COVID-19 response in  
588       South America: the Uruguayan Conundrum. medRxiv.  
589       2020:2020.07.24.20161802.
- 590    3. INE. Global Health Data Exchange Washington2020 [cited 2020 September 8].  
591       Available from: <http://ghdx.healthdata.org/organizations/national-institute-statistics-uruguay>
- 593    4. IMPO. Decreto No. 93/020: Declaracion de Estado de Emergencia Nacional  
594       Sanitaria como Consecuencia de la Pandemia Originada por el Virus COVID-19  
595       (Coronavirus) Uruguay: Registro Nacional de Leyes y Decretos; 2020 [cited  
596       2020 September 4]. Available from:  
597       <https://www.impo.com.uy/bases/decretos/93-2020>
- 598    5. IMPO. Decreto No. 104/020: Autorizacion del Ingreso al Pais Unicamente de  
599       Ciudadanos Uruguayos y Extranjeros Residentes Provenientes del Exterior  
600       Uruguay: Registro Nacional de Leyes y Decretos; 2020 [cited 2020 September  
601       4]. Available from: <http://www.impo.com.uy/bases/decretos/104-2020>
- 602    6. SNE. Visualizador de casos coronavirus COVID-19 en Uruguay 2020 [cited  
603       2020 September 15]. Available from: <https://www.gub.uy/sistema-nacional-emergencias/pagina-embebida/visualizador-casos-coronavirus-covid-19-uruguay>

- 605 7. Salazar C, Díaz-Viraqué F, Pereira-Gómez M, et al. Multiple introductions,  
606 regional spread and local differentiation during the first week of COVID-19  
607 epidemic in Montevideo, Uruguay. bioRxiv. 2020:2020.05.09.086223.  
608 8. Candido DS, Claro IM, de Jesus JG, et al. Evolution and epidemic spread of  
609 SARS-CoV-2 in Brazil. Science. 2020 Sep 4;369(6508):1255-1260.  
610 9. Maurano MT, Ramaswami S, Westby G, et al. Sequencing identifies multiple,  
611 early introductions of SARS-CoV2 to New York City Region. medRxiv.  
612 2020:2020.04.15.20064931.  
613 10. Gonzalez-Reiche AS, Hernandez MM, Sullivan MJ, et al. Introductions and  
614 early spread of SARS-CoV-2 in the New York City area. Science. 2020 Jul  
615 17;369(6501):297-301.  
616 11. Dellicour S, Durkin K, Hong SL, et al. A phylodynamic workflow to rapidly  
617 gain insights into the dispersal history and dynamics of SARS-CoV-2 lineages.  
618 bioRxiv. 2020:2020.05.05.078758.  
619 12. Lemieux J, Siddle KJ, Shaw BM, et al. Phylogenetic analysis of SARS-CoV-2  
620 in the Boston area highlights the role of recurrent importation and  
621 superspreading events. medRxiv. 2020:2020.08.23.20178236.  
622 13. Hadfield J, Megill C, Bell SM, et al. Nextstrain: real-time tracking of pathogen  
623 evolution. Bioinformatics. 2018 Dec 1;34(23):4121-4123.  
624 14. PÚBLICA MDS. Decreto N° 93/020 de 13 de marzo de 2020 2020. Available  
625 from: <https://www.gub.uy/ministerio-salud-publica/sites/ministerio-salud-publica/files/documentos/noticias/Procedimiento%20diagn%C3%B3stico%20Covid-19%20en%20PIAS.pdf>  
626  
627  
628 15. PÚBLICA MDS. COVID-19 Vigilancia y diagnóstico laboratorial 2020.  
629 Available from: <https://www.gub.uy/ministerio-salud-publica/sites/ministerio-salud-publica/files/documentos/noticias/Procedimiento%20diagn%C3%B3stico%20Covid-19%20en%20PIAS.pdf>

- 630           salud-
- 631           [publica/files/documentos/noticias/03 MSP COVID 19 VIGILANCIA DIAG](#)
- 632           [NOSTICO\\_FINAL.pdf](#)
- 633     16. PÚBLICA MDS. Plan Nacional de Contingencia para la Infección (COVID-19)  
634           por el nuevo Coronavirus (SARS CoV2) 2020. Available from:  
635           [https://www.gub.uy/ministerio-salud-publica/sites/ministerio-salud-publica/files/documentos/noticias/01\\_MSP\\_COVID\\_19\\_PLAN\\_NACIONAL\\_CONTINGENCIA\\_GESTION\\_RESPUESTA.pdf](https://www.gub.uy/ministerio-salud-publica/sites/ministerio-salud-publica/files/documentos/noticias/01_MSP_COVID_19_PLAN_NACIONAL_CONTINGENCIA_GESTION_RESPUESTA.pdf)
- 638     17. Tamura K, Peterson D, Peterson N, et al. MEGA5: molecular evolutionary  
639           genetics analysis using maximum likelihood, evolutionary distance, and  
640           maximum parsimony methods. Mol Biol Evol. 2011 Oct;28(10):2731-9.
- 641     18. Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence  
642           alignment, interactive sequence choice and visualization. Brief Bioinform. 2019  
643           Jul 19;20(4):1160-1166.
- 644     19. tools LANL. Highlighter tool 2020. Available from: <http://www.hiv.lanl.gov/>
- 645     20. Nguyen LT, Schmidt HA, von Haeseler A, et al. IQ-TREE: a fast and effective  
646           stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol  
647           Evol. 2015 Jan;32(1):268-74.
- 648     21. Rambaut A, Lam TT, Max Carvalho L, et al. Exploring the temporal structure of  
649           heterochronous sequences using TempEst (formerly Path-O-Gen). Virus Evol.  
650           2016 Jan;2(1):vew007.
- 651     22. Saguleko P, Puller V, Neher RA. TreeTime: Maximum-likelihood  
652           phylogenetic analysis. Virus Evol. 2018 Jan;4(1):vex042.
- 653     23. Suchard MA, Lemey P, Baele G, et al. Bayesian phylogenetic and phylogenetic  
654           data integration using BEAST 1.10. Virus Evol. 2018 Jan;4(1):vey016.

- 655 24. Rambaut A, Drummond AJ. Tracer v1.5. 2009.
- 656 25. Rambaut A, Drummond AJ, Xie D, et al. Posterior Summarization in Bayesian  
657 Phylogenetics Using Tracer 1.7. *Syst Biol.* 2018 Sep 1;67(5):901-904.
- 658 26. Wu F, Zhao S, Yu B, et al. A new coronavirus associated with human  
659 respiratory disease in China. *Nature.* 2020 Mar;579(7798):265-269.
- 660 27. Mercatelli D, Giorgi FM. Geographic and Genomic Distribution of SARS-CoV-  
661 2 Mutations. *Front Microbiol.* 2020;11:1800.
- 662 28. Bryner J. 1st known case of coronavirus traced back to November in China  
663 2020. Available from: [https://www.livescience.com/first-case-coronavirus-  
664 found.html](https://www.livescience.com/first-case-coronavirus-found.html)
- 665 29. Nie Q, Li X, Chen W, et al. Phylogenetic and phylodynamic analyses of SARS-  
666 CoV-2. *Virus Res.* 2020 Oct 2;287:198098.
- 667 30. van Dorp L, Acman M, Richard D, et al. Emergence of genomic diversity and  
668 recurrent mutations in SARS-CoV-2. *Infect Genet Evol.* 2020 Sep;83:104351.
- 669 31. Andersen KG, Rambaut A, Lipkin WI, et al. The proximal origin of SARS-  
670 CoV-2. *Nat Med.* 2020 Apr;26(4):450-452.
- 671 32. Korber B, Fischer WM, Gnanakaran S, et al. Tracking Changes in SARS-CoV-2  
672 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell.*  
673 2020 Aug 20;182(4):812-827 e19.
- 674 33. Ilmjärv S, Abdul F, Acosta-Gutiérrez S, et al. Epidemiologically most successful  
675 SARS-CoV-2 variant: concurrent mutations in RNA-dependent RNA  
676 polymerase and spike protein. *medRxiv.* 2020:2020.08.23.20180281.
- 677 34. CoV-GLUE. CoV-GLUE - Amino acid replacements 2020. Available from:  
678 <http://cov-glue.cvr.gla.ac.uk/#replacement>

- 679 35. Wang R, Chen J, Gao K, et al. Characterizing SARS-CoV-2 mutations in the  
680 United States. Res Sq. 2020 Aug 11.
- 681 36. Kim JS, Jang JH, Kim JM, et al. Genome-Wide Identification and  
682 Characterization of Point Mutations in the SARS-CoV-2 Genome. Osong Public  
683 Health Res Perspect. 2020 Jun;11(3):101-111.
- 684 37. Liu Q, Zhao S, Hou Y, et al. Ongoing natural selection drives the evolution of  
685 SARS-CoV-2 genomes. medRxiv. 2020:2020.09.07.20189860.
- 686 38. Chaw SM, Tai JH, Chen SL, et al. The origin and underlying driving forces of  
687 the SARS-CoV-2 outbreak. J Biomed Sci. 2020 Jun 7;27(1):73.
- 688 39. Portelli S, Olshansky M, Rodrigues CHM, et al. COVID-3D: An online resource  
689 to explore the structural distribution of genetic variation in SARS-CoV-2 and its  
690 implication on therapeutic development. bioRxiv. 2020:2020.05.29.124610.
- 691 40. Zhang L, Jackson CB, Mou H, et al. The D614G mutation in the SARS-CoV-2  
692 spike protein reduces S1 shedding and increases infectivity. bioRxiv.  
693 2020:2020.06.12.148726.
- 694 41. Li Q, Wu J, Nie J, et al. The Impact of Mutations in SARS-CoV-2 Spike on  
695 Viral Infectivity and Antigenicity. Cell. 2020 Sep 3;182(5):1284-1294 e9.
- 696 42. Yurkovetskiy L, Wang X, Pascal KE, et al. Structural and Functional Analysis  
697 of the D614G SARS-CoV-2 Spike Protein Variant. bioRxiv.  
698 2020:2020.07.04.187757.
- 699 43. Mansbach RA, Chakraborty S, Nguyen K, et al. The SARS-CoV-2 Spike  
700 Variant D614G Favors an Open Conformational State. bioRxiv.  
701 2020:2020.07.26.219741.

- 702 44. Ogawa J, Zhu W, Tonnu N, et al. The D614G mutation in the SARS-CoV2  
703 Spike protein increases infectivity in an ACE2 receptor dependent manner.  
704 bioRxiv. 2020:2020.07.21.214932.
- 705 45. Baum A, Fulton BO, Wloda E, et al. Antibody cocktail to SARS-CoV-2 spike  
706 protein prevents rapid mutational escape seen with individual antibodies.  
707 Science. 2020 Aug 21;369(6506):1014-1018.
- 708 46. Weissman D, Alameh M-G, de Silva T, et al. D614G Spike Mutation Increases  
709 SARS CoV-2 Susceptibility to Neutralization. medRxiv.  
710 2020:2020.07.22.20159905.
- 711 47. Sahin U, Muik A, Derhovanessian E, et al. Concurrent human antibody and T<sub>H</sub>1  
712 type T-cell responses elicited by a COVID-19 RNA vaccine. medRxiv.  
713 2020:2020.07.17.20140533.
- 714 48. Pachetti M, Marini B, Benedetti F, et al. Emerging SARS-CoV-2 mutation hot  
715 spots include a novel RNA-dependent-RNA polymerase variant. J Transl Med.  
716 2020 Apr 22;18(1):179.
- 717 49. LaRed21. Primer caso autóctono de COVID-19 en Rivera: Hombre se habría  
718 contagiado en Brasil 2020. Available from:  
719 <https://www.lr21.com.uy/comunidad/1427746-primer-caso-autoctono-de-covid-19-en-rivera-hombre-se-habria-contagiado-en-brasil>
- 721 50. Subrayado. Metalworker from Rivera, who works in Livramento, first  
722 autochthonous case of Covid-19 2020. Available from:  
723 <https://www.subrayado.com.uy/obrero-metalurgico-rivera-que-trabaja-livramento-primer-caso-autoctono-covid-19-n626618>
- 725 51. Lemey P, Rambaut A, Welch JJ, et al. Phylogeography takes a relaxed random  
726 walk in continuous space and time. Mol Biol Evol. 2010 Aug;27(8):1877-85.

- 727 52. Asadu C. SOCIALDISTANCING: How one wedding guest infected 44  
728 Uruguayans. The Cable. 2020. Available from: <https://www.thecable.ng/s-o-c-i-a-l-d-i-s-t-a-n-c-i-n-g-how-one-wedding-guest-infected-44-uruguayans>
- 729 53. Núñez I. "Estuve en un casamiento con 500 personas", relató diagnosticada con  
730 coronavirus en Uruguay. El País. 2020. Available from:  
731 <https://www.elpais.com.uy/informacion/salud/estuve-casamiento-personas-relato-diagnosticada-coronavirus-uruguay.html>
- 732 54. El Observador. "Me decían que no era un caso grave", dijo una de las  
733 uruguayas con coronavirus. El Observador. 2020. Available from:  
734 <https://www.elobservador.com.uy/nota/-me-decian-que-no-era-un-caso-grave-dijo-una-de-las-uruguayas-con-coronavirus-2020313194329>
- 735 55. Zhang X, Tan Y, Ling Y, et al. Viral and host factors related to the clinical  
736 outcome of COVID-19. Nature. 2020 Jul;583(7816):437-440.
- 737 56. Becerra-Flores M, Cardozo T. SARS-CoV-2 viral spike G614 mutation exhibits  
738 higher case fatality rate. Int J Clin Pract. 2020 May 6:e13525.
- 739 57. Toyoshima Y, Nemoto K, Matsumoto S, et al. SARS-CoV-2 genomic variations  
740 associated with mortality rate of COVID-19. J Hum Genet. 2020 Jul 22.
- 741 58. Petrilli CM, Jones SA, Yang J, et al. Factors associated with hospital admission  
742 and critical illness among 5279 people with coronavirus disease 2019 in New  
743 York City: prospective cohort study. BMJ. 2020 May 22;369:m1966.
- 744 59. Perez-Guzman PN, Daunt A, Mukherjee S, et al. Clinical characteristics and  
745 predictors of outcomes of hospitalized patients with COVID-19 in a multi-ethnic  
746 London NHS Trust: a retrospective cohort study. Clin Infect Dis. 2020 Aug 7.
- 747  
748  
749  
750  
751

752

753

754

755 **Figure 1. Virologic, demographic, and clinical parameters of the Uruguayan study**

756 **cohort.** (A) SARS-CoV-2 clade distribution over the ~3 month study period from  
757 March to May. (B) Demographic, virologic, and sample collection data are shown in a  
758 multicategorical alluvial diagram with the display of relatedness among features of two  
759 neighboring nodes. (C) Clinical parameters of study participants, shown in alluvial  
760 representation as in B.

761

762 **Figure 2. SARS-CoV-2 mutation patterns over time and mutation hotspots along**

763 **the genome.** (A) Highlighter plot showing mutations (mut) of Uruguayan study  
764 sequences compared to the reference Wuhan.Hu.1 sequence as master (on top).  
765 Mutations are shown as ticks, color-coded according to the legend to the right. Study  
766 sequences are sorted along the y-axis according to sampling time, with the earliest  
767 sequences on top and most recent sequences at the bottom. A SARS-CoV-2 genome  
768 map with the three reading frames' coding genes is shown for orientation on top. All  
769 single-nucleotide polymorphisms (SNPs) are summarized in orange, and all SNPs  
770 resulting in amino acid (AA) replacements are summarized in blue at the bottom of the  
771 plot at the respective alignment positions. SNPs that are prevalent in >30% of study  
772 sequences are highlighted by orange or blue (if AA replacement) diamonds on top of  
773 the plot. (B) Mirror bar chart summarizing the number of SNPs (orange) and AA  
774 replacements (blue) per study sample, aligned with the study sample IDs in A. (C)  
775 Lollipop plot summarizing the frequency of SARS-CoV-2 mutations in the Uruguayan  
776 study cohort (n=44), using the same color code as in B. A SARS-CoV-2 genome map  
777 with base-pair positions is shown at the bottom. The bubbles' y-coordinates indicate  
778 mutation frequencies, which are also shown inside the bubbles for mutations with >10%  
779 prevalence. Mutation details are shown in orange (SNPs, bp mutation) or blue and black

780 (AA replacements, bp mutation and aa mutation/affected protein region) for mutations  
781 >30% prevalence.

782

783 **Figure 3. Time-scaled phylogenetic tree to identify the source regions of the**  
784 **sequences in the imported Uruguayan clusters.** We employed a discrete state  
785 phylogeography diffusion model in BEAST with seven ancestral location states (Africa,  
786 Asia, Europe, North America, Oceania, South America, and Uruguay) to identify the  
787 most probable source locations for the sequences in the 23 previously identified  
788 international introductions into Uruguay. The branches of the trees are color-coded  
789 according to each geographic region. A color gradient along the branches indicates  
790 historic introduction events between locations. The introductions into Uruguay are  
791 highlighted by black arrows and circles with consecutive numbering according to the  
792 introduction event (color-code of circle outline: probable source continent). The times  
793 of most recent common ancestors (TMRCAs) of Uruguayan (UY) sequences and their  
794 sampling period are indicated as ranges along the x-axis timeline. Branches that are not  
795 involved in introduction events are collapsed to facilitate visualization. The introduction  
796 of the spike D614G mutation is indicated by a red arrowhead. The two major  
797 Uruguayan clades are highlighted by brackets, and GISAID clades are indicated.  
798

799 **Figure 4. Visualization of the evolutionary relationships and spatial distribution of**  
800 **SARS-CoV-2 samples in the five Uruguayan clusters.** A time-scaled maximum clade  
801 credibility tree (MCC) was generated by the discrete phylogeographic analysis of 1810  
802 SARS-CoV-2 genomic sequences. Uruguayan sequences are shown as colored circles,  
803 both in the phylogenetic tree and in the Uruguayan maps. The five main Uruguayan  
804 clusters are color-coded according to the legend (clades indicated in brackets). The

805 remaining Uruguayan sequences, which are based on introduction events that did not  
806 form subsequent transmission chains within Uruguay, are shown as gray circles.  
807

808 **Figure 5. Correlation network analysis of virologic, demographic, and clinical**  
809 **parameters among Uruguayan study samples/participants.** (A) In the correlation  
810 network plot, nodes represent clinical, demographic, laboratory, mutational, and  
811 personal parameters, and red and blue edges represent positive and negative correlations  
812 between connected parameters, respectively. Only significant correlations ( $P < 0.05$ ) are  
813 displayed between parameters with at least two positive events. Edge width corresponds  
814 to the strength of the correlation. Nodes are color-coded based on the grouping in  
815 clinical, demographic, laboratory, mutation, or personal parameters according to the  
816 legend to the upper right, and node size corresponds to the degree of relatedness of  
817 correlations. The six most prominent mixed correlation clusters are encircled with  
818 dashed lines and shown in greater detail as correlograms in the dashed boxes with  
819 matching numbers (1-6). (B) In the correlograms, squares are sized and color-coded  
820 according to the magnitude of the correlation coefficient ( $r$ ). The color code of  $r$  values  
821 is shown to the right (red colors represent positive, blue colors negative correlations  
822 between two parameters). Asterisks indicate statistically significant correlations (\* $P <$   
823  $0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.005$ ). Correlation analysis was done using nonparametric  
824 Spearman rank tests. MVD: Montevideo, HI: healthcare institution, ICU: intensive care  
825 unit, AHT: arterial hypertension, DM II: diabetes mellitus type II, COPD: chronic  
826 obstructive pulmonary disease.

827  
828 **Figure 6. Clade and clinical correlation analysis and study parameters associated**  
829 **with lethal outcome and spike D614G mutation.** (A) In the edge bundling correlation

830 plot, red and blue edges represent positive and negative correlations between connected  
831 parameters, respectively. Only significant correlations ( $P < 0.05$ ) are displayed, and all  
832 parameters have at least two positive events. Nodes are color-coded based on the  
833 grouping into clades, clinical, laboratory, neighborhood, and personal data according to  
834 the legend at the bottom. Node size corresponds to the degree of relatedness of  
835 correlations. Surrounding circle segments highlight a strong clinical cluster and a less  
836 pronounced clade/sampling location cluster. **(B)** Volcano plot of parameters associated  
837 with lethal outcome. The full data set (see **Figure 5** and **Figure S9**) was screened for  
838 parameters with false discovery rates (FDR) of  $q < 0.01$  (red, considered significant) and  
839  $0.01 < q < 0.05$  (pale red, considered borderline significant). **(C)** Volcano plot of  
840 parameters associated with the presence of D614G mutation in spike proteins of study  
841 participants' infecting SARS-CoV-2 viruses. The same display was used as in B.  
842 Correlation analysis was done using nonparametric Spearman rank tests. All parameters  
843 that achieved  $P < 0.05$  correlations are labeled. Parameters inheriting the same dot are  
844 boxed. MVD: Montevideo, HI: healthcare institution.

845

846

847

848

849

850

851

852

853

854 **Supporting Information**

855 **Table S1. Study samples, demographic, and clinical parameters of COVID-19**  
856 **participants.**

857 Abbreviations: GF: GeneFinder, na: not available, NOS\_C: naso- and oropharyngeal  
858 swabs combination, TIB: TIB ROCHE, TS: tracheal secretions. hCoV-19/Uruguay/UY-  
859 NYUMC932/2020|EPI\_ISL\_480430|2020-04-14 was removed from analyses because  
860 of > 35% undefined base pairs (N's) in full genome sequence.

861

862 **Table S2. Origin and number of global study sequences used for Bayesian**  
863 **analyses.**

864

865

866 **Figure S1. Map of SARS-CoV-2 sample collection in our Uruguayan study cohort.**  
867 Map of Montevideo, Uruguay, with sampling locations indicated by dots that are  
868 colored by sample numbers according to the legend to the right. Sampling density was  
869 most significant around the center of Montevideo, indicated by a gray-to-red density  
870 gradient.

871

872 **Figure S2. SARS-CoV-2 amino acid replacements over time.**

873 Highlighter plot showing amino acid (AA) replacements of Uruguayan study sequences  
874 compared to the reference Wuhan.Hu.1 sequence as master (on top). Mutations are  
875 shown as ticks, color-coded according to the legend at the bottom. Study sequences are  
876 sorted along the y-axis according to sampling time, with the earliest sequences on top  
877 and most recent sequences at the bottom. A SARS-CoV-2 genome map is shown on top  
878 with protein regions consecutively assembled as done for the AA alignment. The gray

879 tones of the protein bars relate to the reading frames of their coding genes. All occurring  
880 AA replacements are summarized at the bottom of the plot (compressed). AA  
881 replacements that are prevalent in >30% of study sequences are indicated by diamonds  
882 on top of the plot in the color of the replacing amino acid. A bar chart is shown to the  
883 right summarizing the number of AA replacements per study sample, aligned with the  
884 highlighter results and sample IDs on the left.

885

886 **Figure S3. Summary of SARS-CoV-2 BP and AA mutations in the cohort and per**  
887 **subject.**

888 **Left:** Bar diagrams showing mutation counts per site in the entire cohort. All single  
889 nucleotide polymorphisms (SNPs) are shown on the upper left, with total counts  
890 indicated on the y-axis and as numerical values on top of each bar. The base pair  
891 mutations are indicated on the x-axis together with the amino acid (AA) replacement, if  
892 applicable. On the lower left, all AA replacements are shown separately. **Right:** Bar  
893 diagrams showing cohort-wide mutation counts per genome. SNPs are shown on the  
894 upper right, AA replacements on the lower right. The number of mutations per genome  
895 is listed on the x-axis, and the occurrences in the cohort on the y-axis and as numerical  
896 values on top of each bar.

897

898 **Figure S4. Uruguayan SARS-CoV-2 mutation clusters.**

899 Correlogram of associations among mutations observed in our Uruguayan study cohort  
900 (n=44) with squares sized and color-coded according to the magnitude of the correlation  
901 coefficient ( $r$ ). The color code of  $r$  values is shown to the right; red colors represent  
902 positive, blue colors negative correlations between two connected parameters on the x-  
903 and y-axes. Asterisks indicate statistically significant correlations (\* $P < 0.05$ , \*\* $P <$

904 0.01, \*\*\*P < 0.005). The correlogram is shown with hierarchical clustering according to  
905 the dendrogram at the bottom. The color-strip indicates gene relatedness of mutations  
906 according to the color code in the legend. Correlation analysis was done using non-  
907 parametric Spearman rank tests.

908

909 **Figure S5. Phylogenetic and mutation network analysis of Uruguayan SARS-CoV-  
910 2 viruses.**

911 **A.** Maximum-likelihood IQ-tree of 73 Uruguayan SARS-CoV-2 sequences and  
912 Wuhan.Hu.1 reference sequence, run with 1000 Bootstrap replications. Branch symbols  
913 and taxa (ID colored ranges) are colored according to country and study source, as  
914 explained in the figure legend. The introductions of key mutations are shown by red and  
915 green triangles. For each sample, reported/estimated site of infection/sampling location  
916 (est: estimated, hos: treating hospital), clades, and time since global outbreak are  
917 indicated by the circular color strip around the tree according to the legend. Clustered  
918 appearances of clades at specific sites/neighborhoods are highlighted by arrows and  
919 labeled. **B.** Genetic distance-based haplotype network analysis of Uruguayan SARS-  
920 CoV-2 sequences. Circles represent populations of sequences with identical mutation  
921 patterns (haplotypes) as compared to Wuhan.Hu.1 as reference (gaps and missing data  
922 not considered). The circles are sized and colored relative to the number and source  
923 countries of contributing sequences, respectively. One sample ID annotates each  
924 haplotype population representatively. Ticks on the connecting lines indicate  
925 discriminating mutations between haplotypes. Two major branch-defining mutations are  
926 highlighted as red and green ticks according to the legend. Red and green polygons  
927 encircle haplotype populations carrying respective key mutations (spike D614G and  
928 ORF8 L84S).

929

930 **Figure S6. Haplotype network analyses of Uruguayan SARS-CoV-2 sequences**

931 **among global reference strains.**

932 Genetic distance-based haplotype network analysis of 73 Uruguayan SARS-CoV-2  
933 sequences among 609 subsampled, global SARS-CoV-2 sequences. Circles represent  
934 populations of sequences with identical mutation patterns (haplotypes) as compared to  
935 Wuhan.Hu.1 as reference (gaps and missing data not considered). The circles are sized  
936 and colored relative to the number and source countries of contributing sequences,  
937 respectively. One sample ID annotates each haplotype population representatively.  
938 Ticks on the connecting lines indicate discriminating mutations between haplotypes.  
939 The node carrying the Wuhan-Hu-1 reference sequence is labeled, and haplotypes  
940 sharing the spike D614G mutation among the sequences of the reticular network are  
941 encircled by a dashed, dark red polygon. The location of major and minor clusters of  
942 Uruguayan sequences is highlighted by large and small blue arrows, respectively.

943

944 **Figure S7. Time-scaled phylogenetic tree to identify Uruguayan clusters.**

945 A cluster is defined as a phylogenetic clade corresponding to an independent  
946 introduction into Uruguay. Uruguayan sequence tree branches are colored light blue and  
947 non-Uruguayan sequence branches gray. The two main clusters are highlighted by  
948 brackets, and their GISAID clades are indicated. The introduction of the spike D614G  
949 mutation is indicated by a red arrowhead.

950

951 **Figure S8. A visualization of the evolutionary relationships and spatial distribution**  
952 **of our SARS-CoV-2 samples in hospitals and nursing homes in Montevideo.**

953 Time-scaled maximum clade credibility tree (MCC) generated by the discrete  
954 phylogeographic analysis of 1810 SARS-CoV-2 sequences. According to the legend,  
955 Uruguayan sequences are shown as colored circles both in the phylogenetic tree and in  
956 the Uruguayan maps, color-coded based on their relationship to hospitals and nursing  
957 homes. Samples that are not associated with a hospital or nursing home are shown as  
958 gray circles.

959

960 **Figure S9. Correlation analysis of viral, demographic, and clinical parameters of**  
961 **Uruguayan study samples/participants.**

962 Correlogram of associations among indicated parameters as present in our Uruguayan  
963 study cohort (n=44) with squares sized and color-coded according to the magnitude of  
964 the correlation coefficient ( $r$ ). The color code of  $r$  values is shown to the right; red  
965 colors represent positive, blue colors negative correlations between two connected  
966 parameters on the x- and y-axes. Asterisks indicate statistically significant correlations  
967 (\* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.005$ ). The correlogram is shown with hierarchical  
968 clustering according to the dendrogram at the bottom. The color-strip indicates group  
969 relatedness of parameters according to the color code in the legend. Correlation analysis  
970 was done using nonparametric Spearman rank tests. MVD: Montevideo, HI: healthcare  
971 institution.

972

973 **Figure S10. Statistics of correlations with lethal outcome of SARS-CoV-2 infection**  
974 **in our Uruguayan study cohort.**

975 Correlation statistics between lethal outcome and virus mutations are shown on the **left**,  
976 and between lethal outcome and clinical and demographic parameters in the **middle**. A  
977 summary of significant correlations, according to  $q < 0.01$ , is shown on the **upper right**.

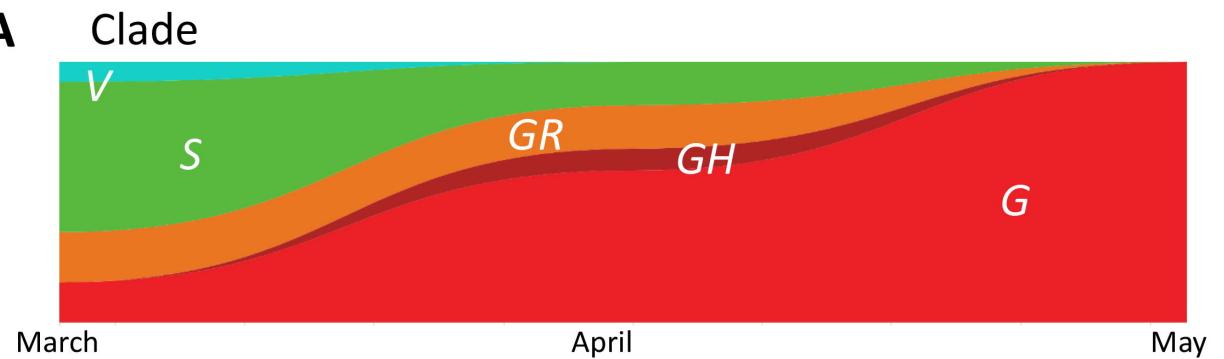
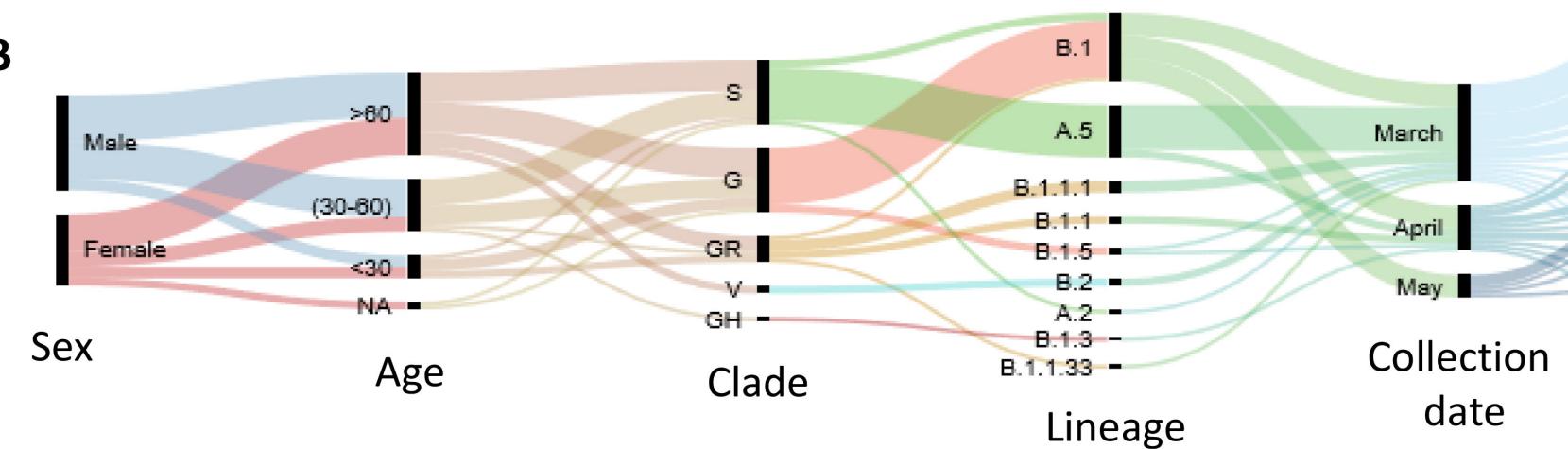
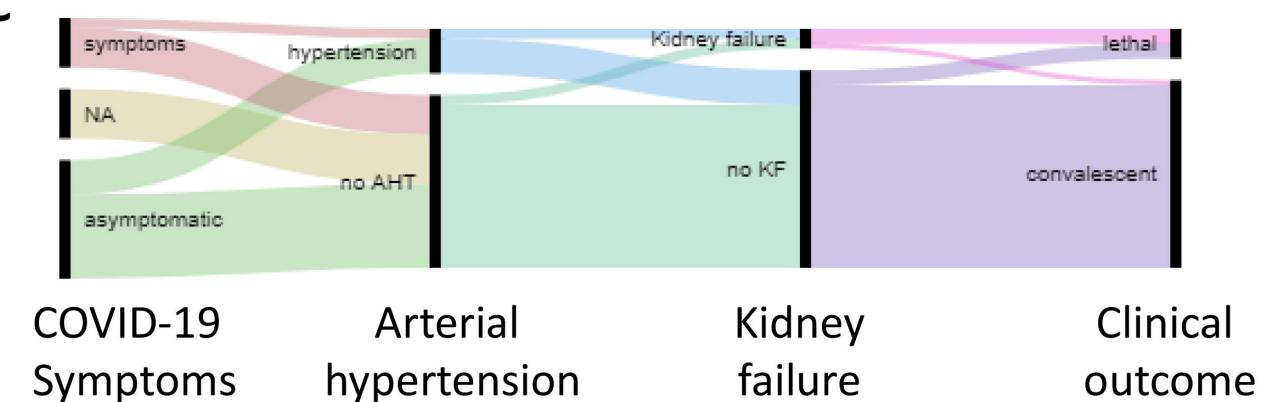
978 For each correlation, correlation values  $r$ ,  $P$  values, adjusted  $P$  values, and false  
979 discovery rates (FDR)  $q$  values are displayed together with significant discovery  
980 assessment. The distribution of  $P$  values across the data set is illustrated in a dot plot of  
981 ranked  $P$  values in the **lower right**. All significant results are highlighted in red. MVD:  
982 Montevideo, HI: healthcare institution, ICU: intensive care unit, AHT: arterial  
983 hypertension, DM II: diabetes mellitus type II, COPD: chronic obstructive pulmonary  
984 disease.

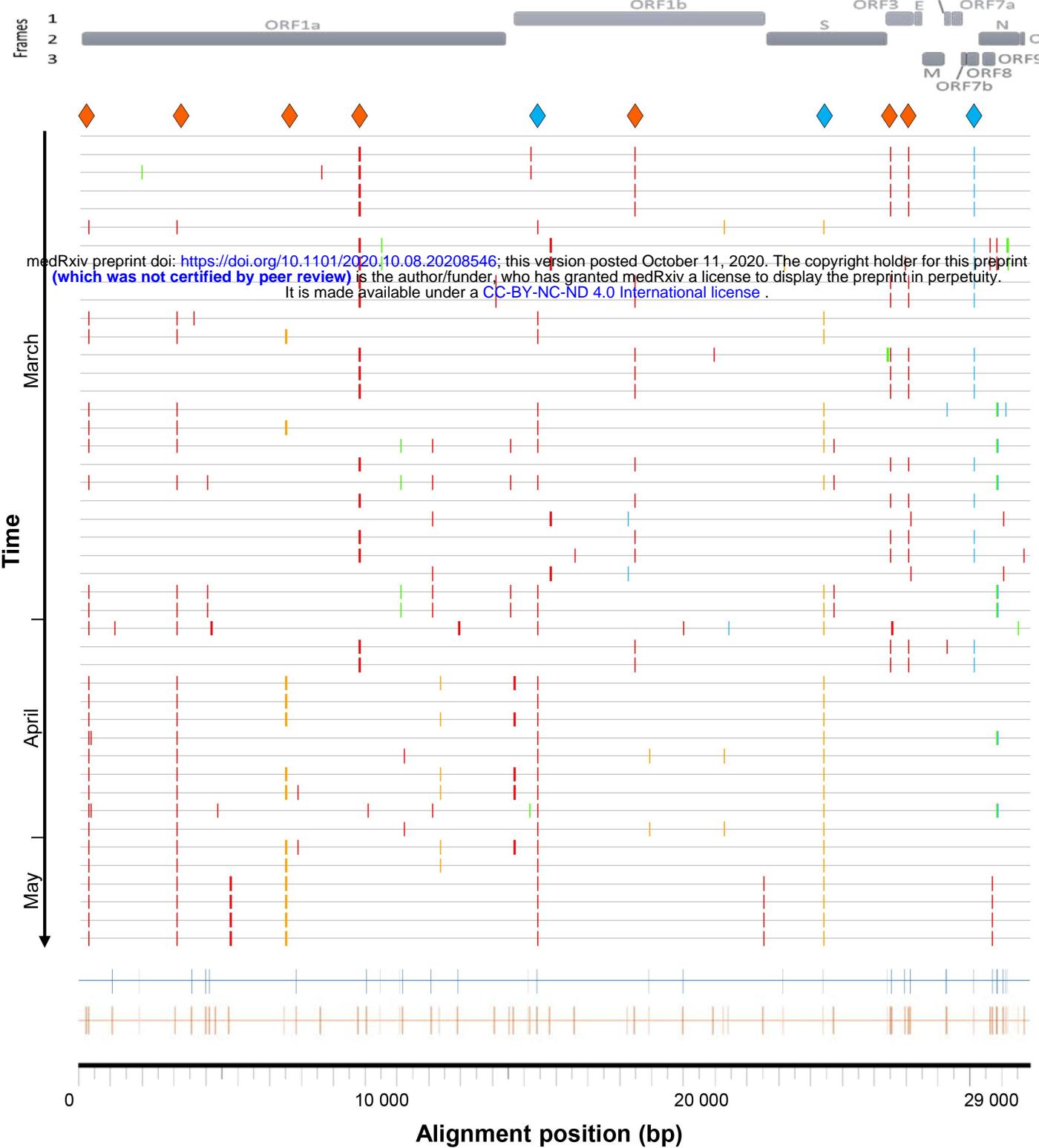
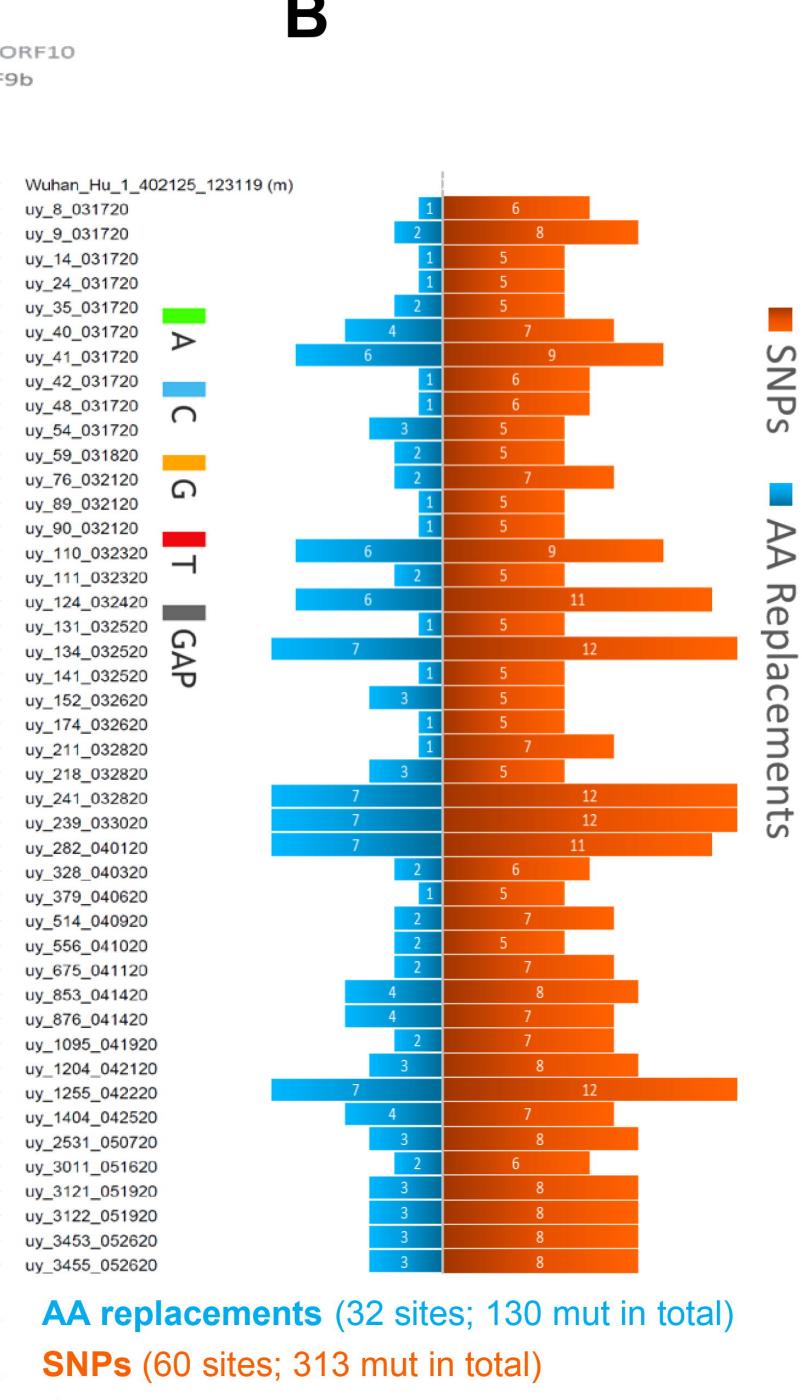
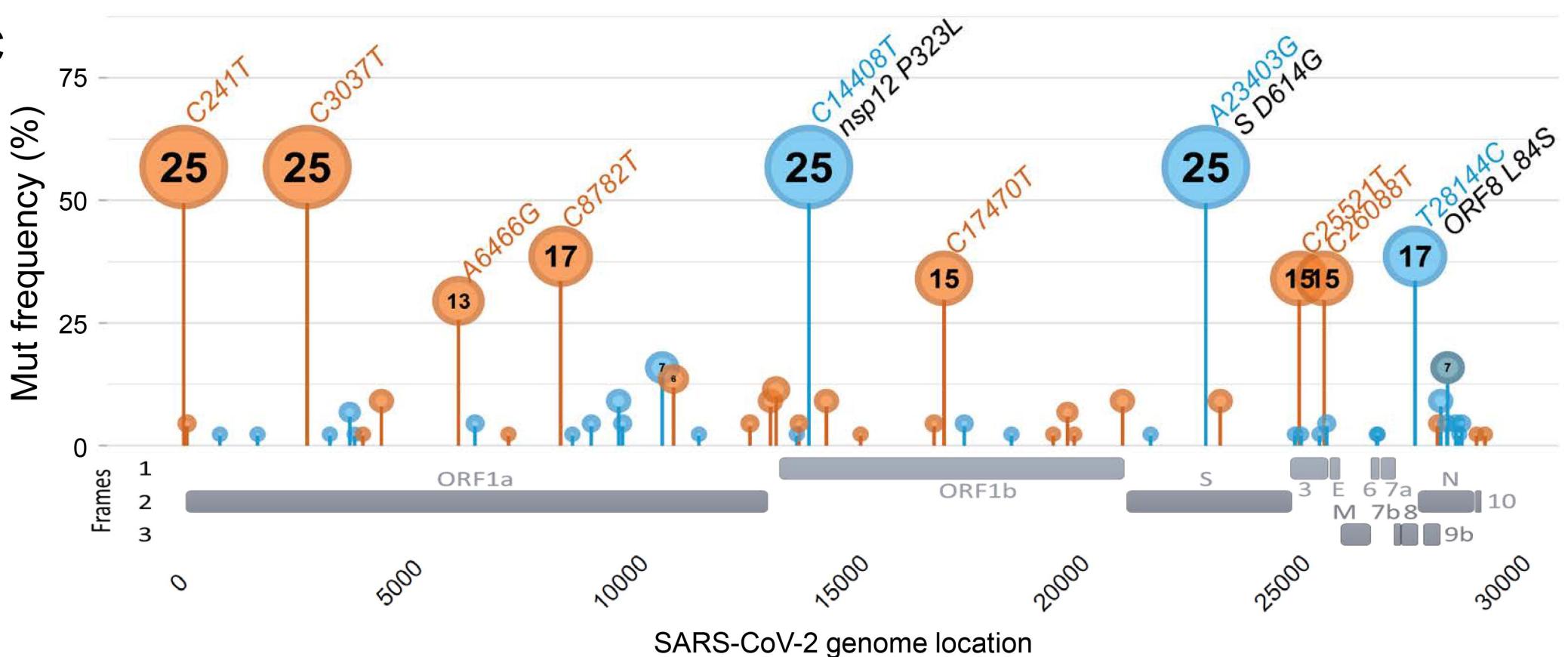
985

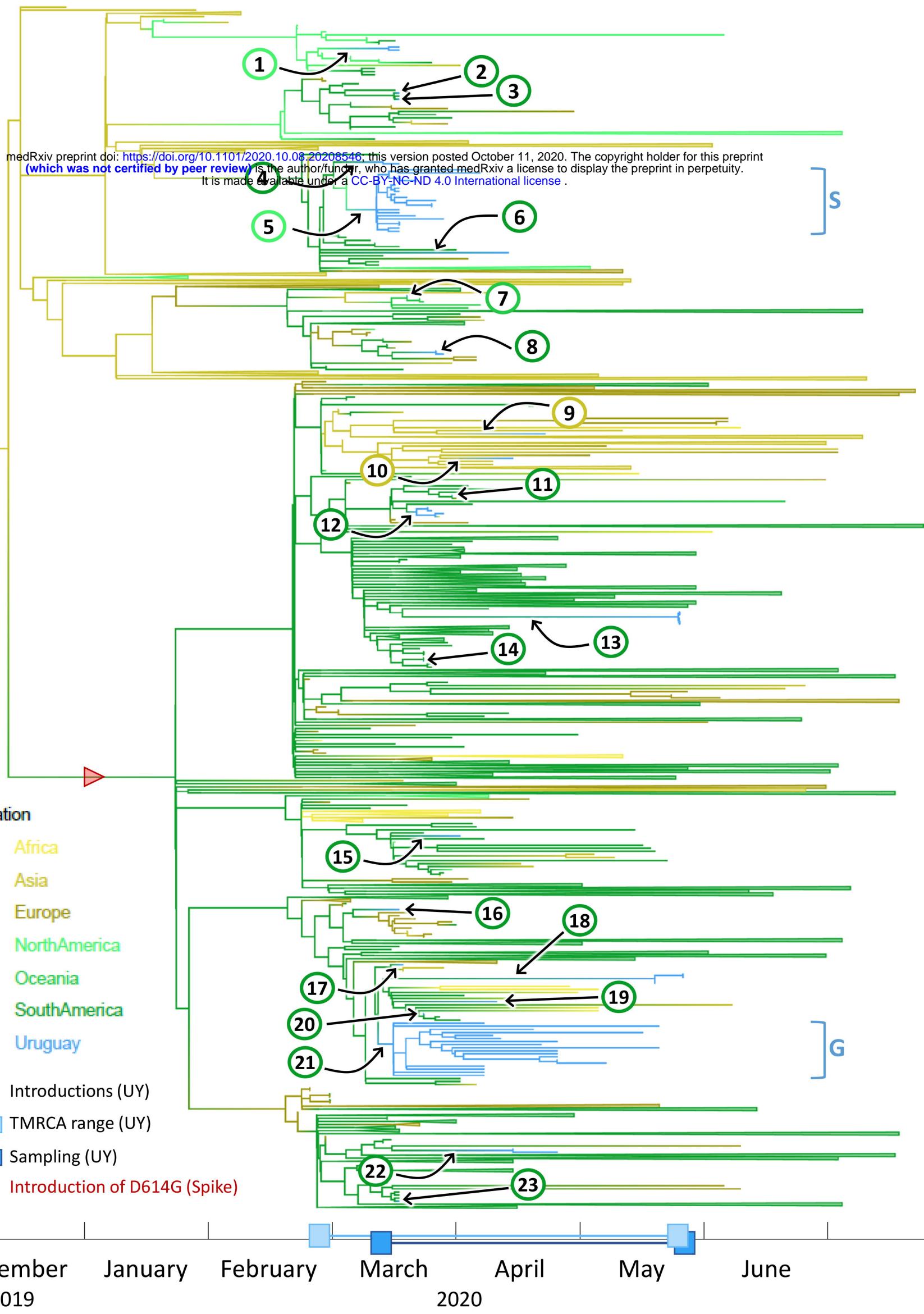
986 **Figure S11. Statistics of correlations with D614G spike mutation in infecting  
987 viruses of our Uruguayan study cohort.**

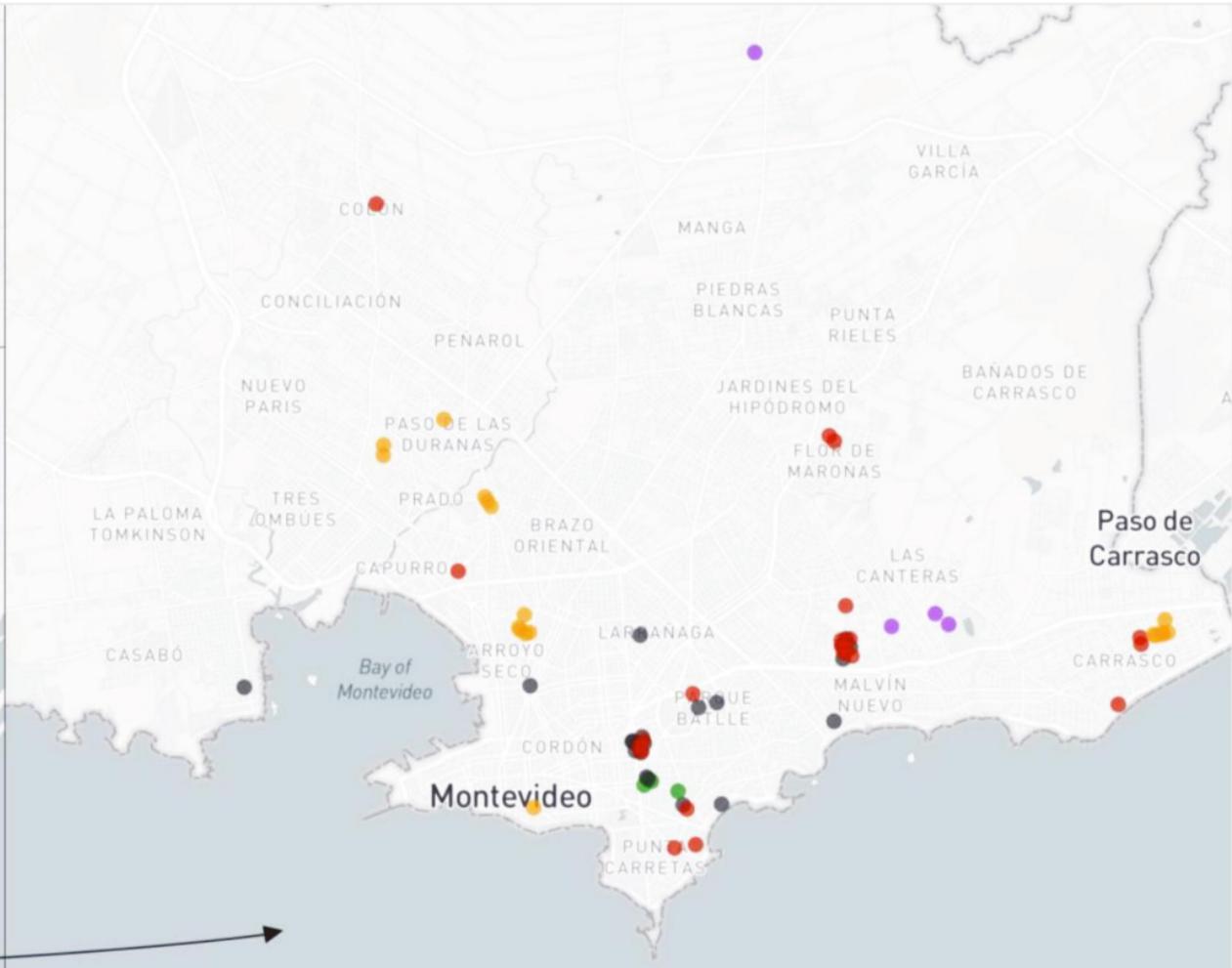
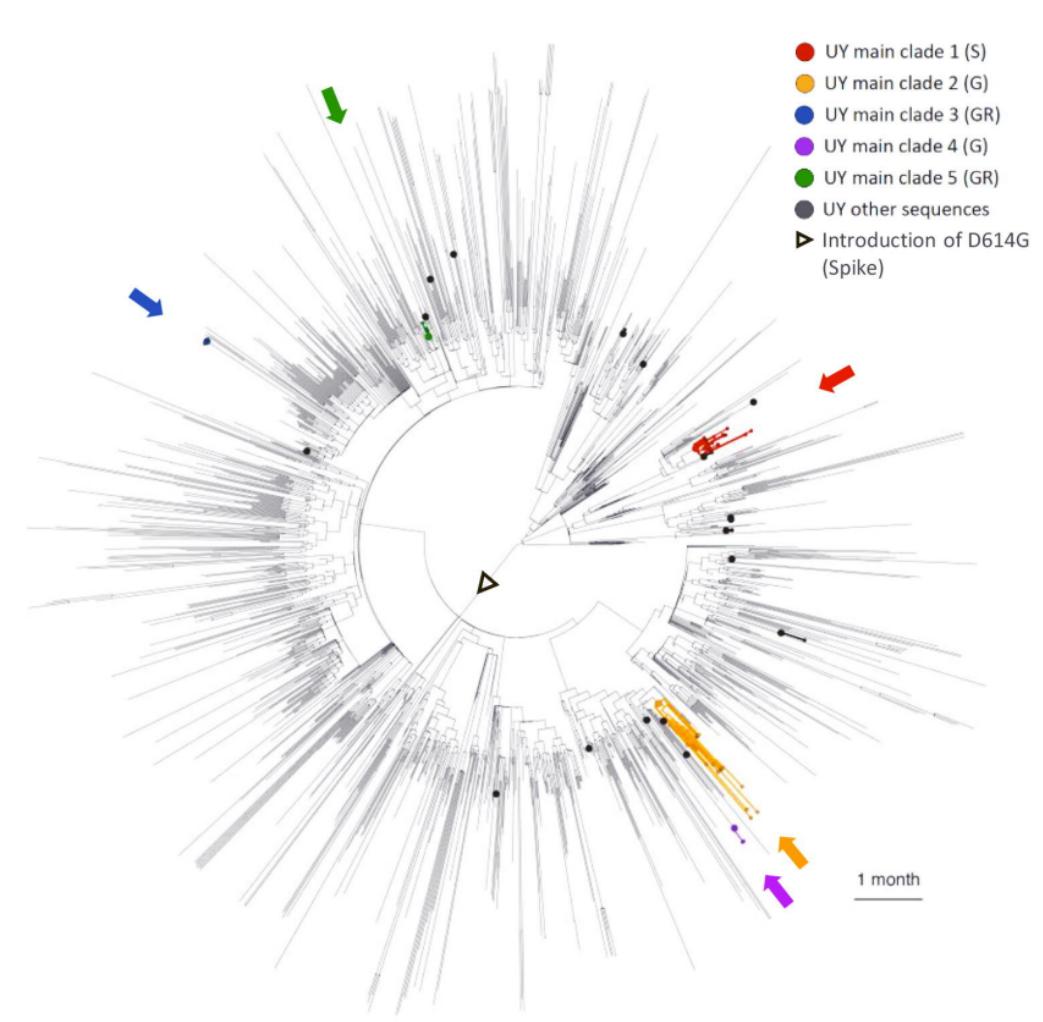
988 Correlation statistics between the presence of D614G spike mutation and other SARS-  
989 CoV-2 mutations are shown on the **left**, and between D614G spike mutation and clinical  
990 and demographic parameters in the **middle**. A summary of significant correlations,  
991 according to  $q < 0.01$ , is shown on the **upper right**. For each correlation, correlation  
992 values  $r$ ,  $P$  values, adjusted  $P$  values, and false discovery rates (FDR)  $q$  values are  
993 displayed together with assessment of significant discovery. The distribution of  $P$   
994 values across the data set is illustrated in a dot plot of ranked  $P$  values in the **lower**  
995 **right**. All significant results are highlighted in red. MVD: Montevideo, HI: healthcare  
996 institution, ICU: intensive care unit, AHT: arterial hypertension, DM II: diabetes  
997 mellitus type II, COPD: chronic obstructive pulmonary disease.

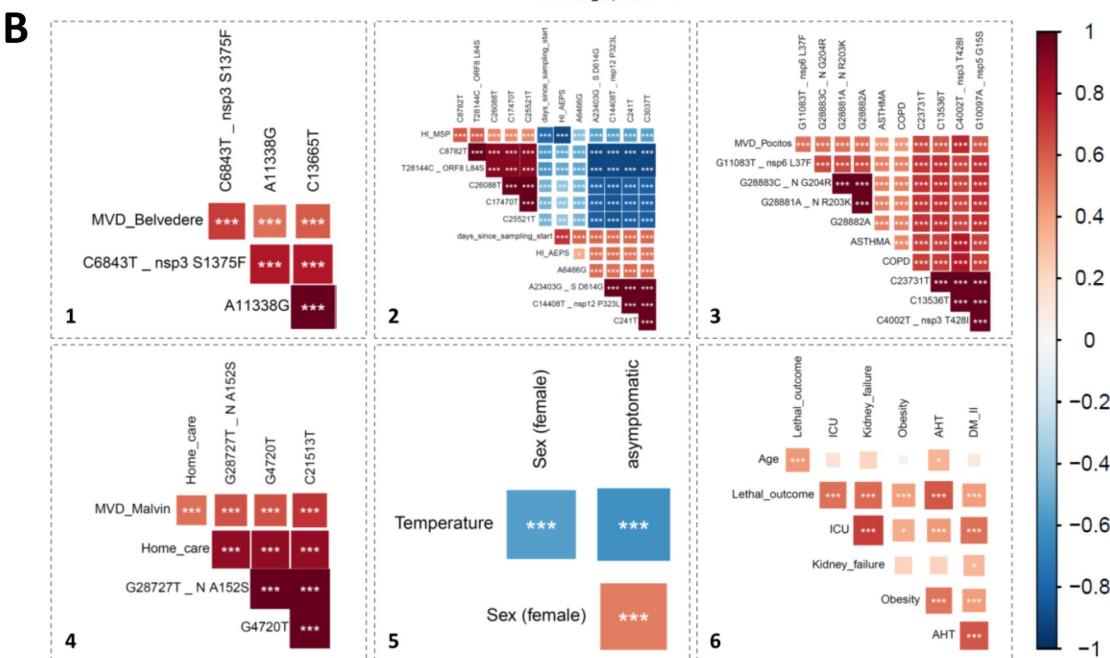
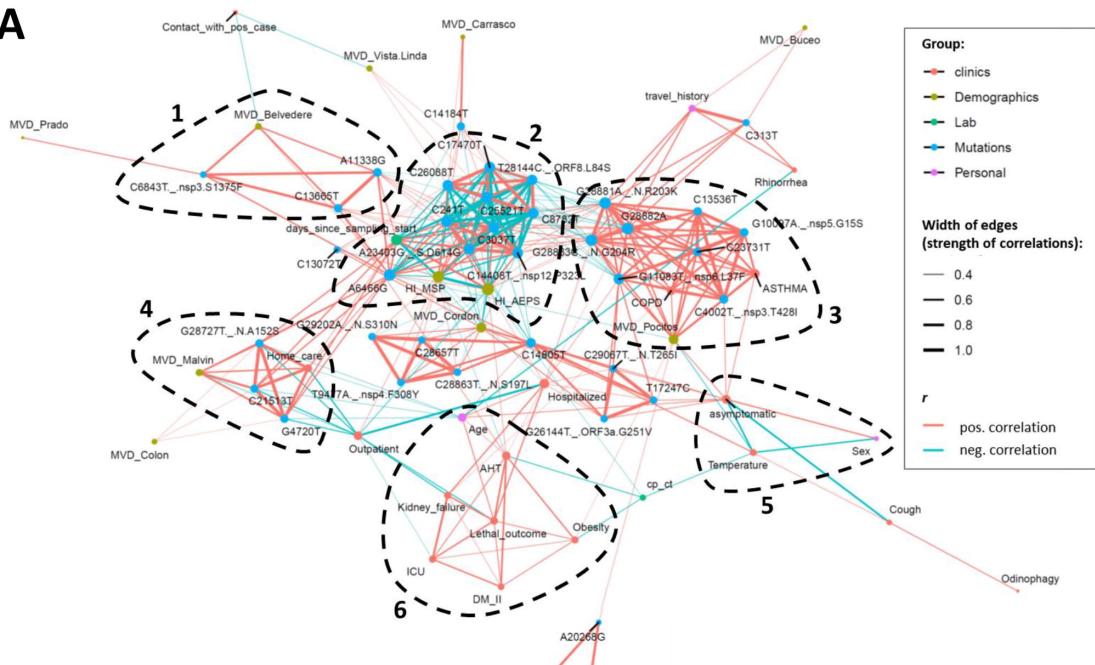
998

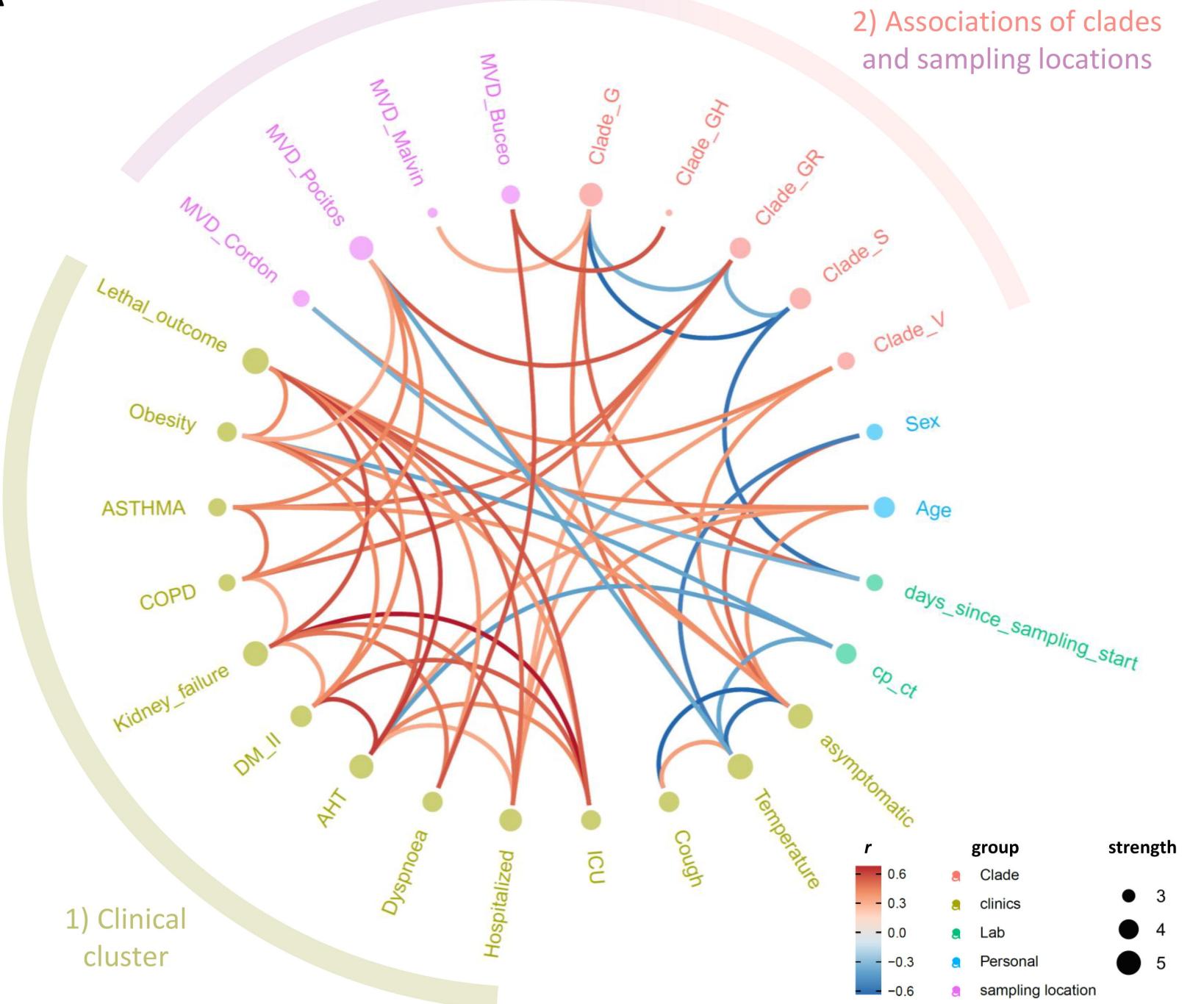
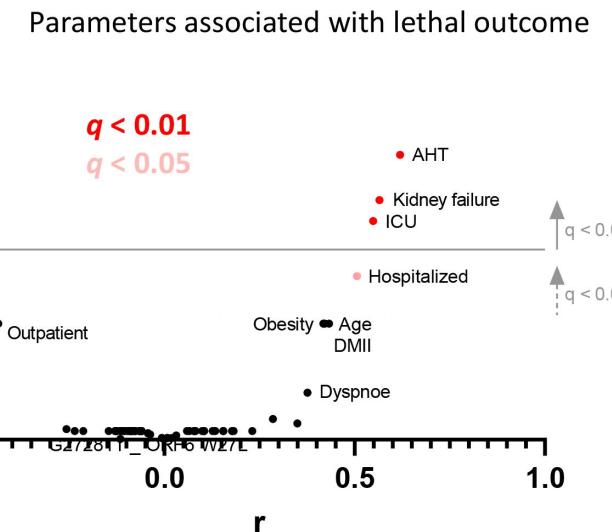
**A****B****C**

**A****B****C**







**A****B****C**