

1           **Genome sequencing of the first SARS-CoV-2 reported from**  
2                                   **patients with COVID-19 in Ecuador.**

3

4   Sully Márquez<sup>1</sup>, Belén Prado-Vivar<sup>1,2</sup>, Juan José Guadalupe<sup>4</sup>, Bernardo  
5   Gutierrez<sup>4,6</sup>, Manuel Jibaja<sup>5</sup>, Milton Tobar<sup>5</sup>, Francisco Mora<sup>7</sup>, Juan Gaviria<sup>7</sup>, María  
6   García<sup>7</sup>, Franklin Espinosa<sup>7</sup>, Edison Ligña<sup>7</sup>, Jorge Reyes<sup>7,9</sup>, Verónica Barragán<sup>1</sup>,  
7   Patricio Rojas-Silva<sup>1</sup>, Gabriel Trueba<sup>1</sup>, Michelle Grunauer<sup>3,8</sup>, Paúl Cárdenas<sup>1,2\*</sup>

8

9   <sup>1</sup>Universidad San Francisco de Quito, COCIBA, Instituto de Microbiología

10   <sup>2</sup>Universidad San Francisco de Quito, Centro de Bioinformática

11   <sup>3</sup>Universidad San Francisco de Quito, Escuela de Medicina, COCSA

12   <sup>4</sup>Universidad San Francisco de Quito, COCIBA, Laboratorio de Biotecnología Vegetal

13   <sup>5</sup>Unidad de Cuidados Intensivos, Hospital Eugenio Espejo, Quito

14   <sup>6</sup>Department of Zoology, University of Oxford

15   <sup>7</sup>Hospital General Sur de Quito, IESS

16   <sup>8</sup>Unidad de Cuidados Intensivos, Hospital de los Valles, Quito

17   <sup>9</sup> Universidad Central del Ecuador, Facultad Ciencias Químicas

18

19

20

21   \*Correspondence: [pacardenas@usfq.edu.ec](mailto:pacardenas@usfq.edu.ec), tel.: +593-22971700 ext. 1517

22

23

24

25

26

27

28

29

30

31

32

33

34

35 **Abstract**

36 SARS-CoV-2, the etiological agent of COVID-19 was first described in Wuhan in  
37 December 2019 and has now spread globally. Ecuador was the second country in  
38 South America to report confirmed cases. The first case reported in Quito, the capital  
39 city of Ecuador, was a tourist who came from the Netherlands and presented  
40 symptoms on March 10th, 2020 (index case). In this work we used the MinION platform  
41 (Oxford Nanopore Technologies) to sequence the metagenome of the bronchoalveolar  
42 lavage (BAL) from this case reported, and subsequently we sequenced the whole  
43 genome of the index case and other three patients using the ARTIC network protocols.  
44 Our data from the metagenomic approach confirmed the presence of SARS-CoV-2  
45 coexisting with pathogenic bacteria suggesting coinfection. Relevant bacteria found in  
46 the BAL metagenome were *Streptococcus pneumoniae*, *Mycobacterium tuberculosis*,  
47 *Staphylococcus aureus* and *Chlamydia* spp. Lineage assignment of the four whole  
48 genomes revealed three different origins. The variant HEE-01 was imported from the  
49 Netherlands and was assigned to B lineage, HGSQ-USFQ-018, belongs to the B.1  
50 lineage showing nine nucleotide differences with the reference strain and grouped with  
51 sequences from the United Kingdom, and HGSQ-USFQ-007 and HGSQ-USFQ-010  
52 belong to the B lineage and grouped with sequences from Scotland. All genomes show  
53 mutations in their genomes compared to the reference strain, which could be important  
54 to understand the virulence, severity and transmissibility of the virus. Our findings also  
55 suggest that there were at least three independent introductions of SARS-CoV-2 to  
56 Ecuador.

57 **Key words:** SARS-CoV2, MinION sequencing, COVID-19, metagenomic analysis,  
58 whole genome

59

60 **IMPORTANCE** COVID-19, an infectious disease caused by SARS-CoV-2, has spread  
61 globally including Latin American countries including Ecuador. The first strain of SARS-  
62 CoV-2 sequenced was from Wuhan, which is considered as the reference strain. There  
63 were no data about the SARS-CoV-2 lineages in Ecuador, and the purpose of this  
64 study was to find out the origin of the different lineages circulating in the population. We  
65 also were interested in the mutations present in these genomes as they can influence  
66 virulence, transmission and infectivity.

67

## 68 **INTRODUCTION**

69 SARS-CoV-2, the etiologic agent of COVID-19, has spread globally reaching all  
70 continents<sup>1,2</sup>. Ecuador was the second country in South America, after Brazil, to report  
71 its first case on 29 February 2020<sup>3</sup>. The number has grown since, reaching 39,098  
72 laboratory-confirmed cases and 3,358 deaths until May 31st, 2020<sup>4</sup>. The pandemic  
73 has had a severe impact in Ecuador, as death excess estimates suggest that up to  
74 182.44 people per million inhabitants may have died since the first COVID19 case<sup>5</sup>,  
75 making Ecuador the country with the highest mortality rate in South America.

76 Quito, the capital city of Ecuador, was the second city to report COVID-19  
77 cases after Guayaquil. After a severe wave of cases in Guayaquil during the early  
78 stages of the outbreak, Quito currently reports the largest number of COVID-19 daily  
79 cases in the country. The first case in Quito was diagnosed initially in the Sucumbíos  
80 province, while the patient was doing tourism, and was subsequently brought to a third-  
81 level hospital in the capital. The patient is originally from the Netherlands and was  
82 diagnosed by the Health Ministry on March 10th, 2020. His condition deteriorated while  
83 in Sucumbíos, prompting his transfer to Quito where he was admitted into the Intensive  
84 Care Unit of the Eugenio Espejo Hospital. After 32 days, the patient recovered and  
85 returned to his native country. Currently, local authorities have confirmed 3676 cases<sup>4</sup>  
86 and widespread community transmission is actively happening.

87           Understanding the local epidemiology of SARS-CoV-2 can greatly be enhanced  
88 from understanding the virus evolution. It is therefore relevant to invest in SARS-CoV-2  
89 genome sequencing as it could help to establish the origin of the pathogen <sup>6</sup>, its  
90 transmission across different regions <sup>7</sup>, the genetic diversity of the circulating strains in  
91 the population and to identify notable mutations <sup>8</sup>. Additionally, key biological aspects  
92 such as virulence, transmissibility and infectivity of the circulating strains can be  
93 investigated, when combined with clinical data from the patients <sup>7</sup>. Around 35,473  
94 genome sequences of SARS-CoV-2 have been reported from countries around the  
95 world and are available in GenBank and the GISAID (Global Initiative on Sharing All  
96 Influenza Data) repository <sup>9</sup>.

97           In this work we describe the first four SARS-CoV-2 whole genome sequences  
98 obtained from Ecuador. We also report the metagenomic analysis of the first COVID-19  
99 patient diagnosed in Quito. Both approaches were performed using the portable  
100 MinION platform (Oxford Nanopore Technologies).

## 101   **MATERIALS AND METHODS**

### 102   **Ethics**

103 This study used excess respiratory samples after being officially confirmed with positive  
104 COVID19 diagnosis. Ethical approval for all samples was given by CEISH-USFQ  
105 (Comité de ética de investigación en seres humanos-USFQ): IE-JP067-2020-CEISH-  
106 USFQ.

### 107   **Epidemiological information and sample collection**

108 Samples from the four patients were collected in Hospital Eugenio Espejo (HEE) or  
109 Hospital IESS Sur (HGSQ), two public third-level hospitals at Quito, Ecuador. Sample  
110 positivity for SARS-CoV-2, using standard RT-PCR approach, was officially reported to  
111 hospitals by the Ecuadorian Ministry of Public Health (MSP) and National Institute of

112 Public Health and Research (INSPI). Sample HEE-01 was a bronchi-alveolar lavage  
113 (BAL) from a patient of Dutch origin in his late fifties who presented symptoms during a  
114 visit to the Sucumbios province and then transferred to the HEE intensive care unit  
115 (ICU) in Quito. This patient was the first confirmed COVID-19 case in Quito. Samples  
116 HGSQ-USFQ-018, HGSQ-USFQ-007 and HGSQ-USFQ-010 were all nasopharyngeal  
117 swabs. HGSQ018 was collected from a 27 years old male patient who presented  
118 severe COVID-19 symptoms after a visit to Guayaquil and was admitted to the  
119 pneumology unit at HGSQ. Sample HGSQ-USFQ-007 was collected from a 40 years  
120 old male patient that had no history or contact with a COVID-19 patient, presented  
121 severe COVID19 symptoms and was admitted to ICU at HGSQ. Sample HGSQ-USFQ-  
122 010 was collected from a 39 years old male patient that had contact with patient  
123 HGSQ-USFQ-007 and was also admitted to ICU. The BAL sample was collected in a  
124 sterile tube with 2X DNA/RNA Shield (Zymo), and nasopharyngeal swabs were  
125 submerged in 1X DNA/RNA Shield (Zymo) to ensure virus inactivation and to preserve  
126 the genetic material. Samples were transported immediately at 4°C to the Institute of  
127 Microbiology at USFQ (IM-USFQ) in a sealed container with all the biosecurity and  
128 containment measures recommended by the CDC of the USA  
129 (<https://www.fda.gov/media/134922/download>).

### 130 **RNA extraction**

131 The genetic material from samples was extracted in a biosafety type II chamber with  
132 HEPA filters in the Virology Laboratory of the Microbiology Institute (IM) at USFQ. Two  
133 different sequencing approaches were performed with the extracted RNA:  
134 metagenomics with sample HEE01 and whole genome sequencing with samples  
135 HEE01, HGSQ-USFQ-018, HGSQ-USFQ-007 and HGSQ-USFQ-010. QIAamp® Viral  
136 RNA Extraction Kit (Qiagen) was used to extract RNA from the HEE1 sample that was  
137 used for the metagenomics approach. Briefly, 250 µl of BAL sample and 560 µl of AVL  
138 lysis buffer were incubated for 1 hour, then manufacturer RNA kit instructions were

139 followed. Sample was eluted in a final volume of 70 µl. RNA was immediately purified  
140 using RNA Clean and Concentrator kit (Zymo), 14 µl of purified ARN were used for  
141 retrotranscription of RNA to cDNA following the RNA Viral Metagenomics MinION One-  
142 Pot Sequencing Protocol from the genomics department of Public Health England<sup>10,11</sup>.

143 SV Total RNA Isolation System kit (Promega, USA) was used to extract RNA from  
144 samples used for the whole genome sequencing approach. A predigestion step was  
145 added to the RNA extraction protocol of the bronchoalveolar fluid (BAL) sample as  
146 follows. Before nucleic acid extraction, 280µl of the Bronchoalveolar fluid (BAL) sample  
147 was predigested with 360µl of PureLink™ Genomic Lysis Buffer and 20 µl of  
148 proteinase K. The mix was incubated at 55°C for 10 minutes, with vortexing every 5  
149 minutes (Life Technologies, USA). All RNA extractions were performed following  
150 manufacturer instructions and eluted in a final volume of 50 µl. Total RNA was  
151 quantified using QuBit (Thermo Fisher Scientific) with a Qubit RNA Assay Kit (Thermo  
152 Scientific, Invitrogen, Carlsbad, CA, USA). Retrotranscription of RNA to cDNA was  
153 carried out using the Protocol of the Public Health England Genomics Lab<sup>10,11</sup> at the  
154 USFQ Bioinformatics Center.

#### 155 **Viral whole genome sequencing approach**

156 Primer Scheme approach developed by the ARTIC network for nCoV-2019 (Robertson,  
157 2020) using the V1 primer sets to generate an amplicon tiling path across the viral  
158 genome<sup>12</sup>. cDNA MinION library preparation was performed using the Rapid  
159 Barcoding kit (SQK-RBK004) following manufacturer instructions and then loaded into  
160 a MinION flow cell (FLO-MIN 106). Base calling of FAST5 files was performed using  
161 Guppy (version 3.4.5) (Wick, et al 2019), and the RAMPART software (v1.0.5) from the  
162 ARTIC Network (<https://github.com/artic-network/rampart>) was used to monitor  
163 sequencing in real-time. Sequence quality scoring, demultiplexing and adapter removal  
164 was performed with the NanoPlot<sup>13</sup> and Porechop algorithms respectively. The ARTIC

165 Network bioinformatics pipeline was used for variant calling, and the reads were  
166 mapped against the reference strain Wuhan-Hu-1 (GenBank accession number  
167 MN908947), to generate consensus genomes. The genome was uploaded to the CoV-  
168 GLUE resource <sup>14</sup>, to determine the epidemiological linkage of circulating SARS-CoV-2  
169 variants.

170 **Metagenomic approach and bioinformatic analysis.** cDNA MinION library  
171 preparation was performed using the Rapid Barcoding kit (SQK-RBK004) following  
172 manufacturer instructions. The resulting library was loaded on an Oxford MinION  
173 flowcell (FLO-MIN 106) and sequenced using the MinKNOW version 4.05, for 24  
174 hours. Basecalling and quality control analysis was performed using Guppy (version  
175 3.4.5) and NanoPlot (version 1.29.0) respectively<sup>13</sup>. Adapters and barcodes were  
176 removed from the MinION reads using Porechop (version 0.2.4)  
177 (<https://github.com/rrwick/Porechop>). Taxonomic classification of the sequences was  
178 performed using the Kaiju platform<sup>15</sup>. A SARSCoV-2 consensus sequence was  
179 obtained by mapping reads against the reference strain Wuhan-Hu-1 (GenBank  
180 accession number MN908947) using minimap2 (version 2.14-r883)<sup>16</sup>. Finally,  
181 Samtools (version 1.9) (<http://samtools.github.io>) and Tablet alignment viewer (version  
182 1.19.09.3) (<https://ics.hutton.ac.uk/tablet>) were used to visualize the mapped  
183 sequence.

## 184 **RESULTS**

185 **Metagenomic sequencing confirms SARS-CoV-2 infection in patient HEE-01 and**  
186 **possible bacterial coinfection.**

187 The metagenomic analysis found a total of 206,111 DNA sequences with 43,603,091  
188 bases. Viral sequences represented 0.9%, of which 83% corresponded to non-  
189 assigned coronavirus and 17% were identified as SARS coronaviruses (Figure 1). The  
190 sequences assigned to coronavirus were extracted and mapped using the Wuhan-Hu-1

191 reference genome (GenBank accession number MN908947). A sequence similarity of  
192 99.68% was found with this sequence with 100% query coverage. Additionally, a  
193 phylogenetic tree was generated with the sequences found and the reference strains  
194 used in GenBank NCBI. The phylogenetic alignment grouped the query sequence with  
195 the ORF1AB segment of the virus polyprotein (it encodes replication genes). The  
196 metagenome sequences are publicly available at  
197 <https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA613094> (Accession:  
198 PRJNA613094).

199 Several bacterial and eukaryotic sequences related to the patient's respiratory  
200 microbiota were identified by metagenomics. The most relevant taxa found were  
201 *Streptococcus pneumoniae*, *Mycobacterium tuberculosis*, *Staphylococcus aureus* and  
202 *Chlamydia* spp. We did not identify any particular clinically relevant fungus. To  
203 complement the metagenomic approach we carried on with the whole genome  
204 sequencing of the current sample and additionally with 3 more additional samples from  
205 severely ill patients diagnosed in March.

## 206 **Complete genome sequencing revealed multiple entries of the virus to Ecuador**

207 Whole genome sequencing of the first case of COVID-19 in Quito revealed it to be an  
208 imported case from Europe (strain named HEE-01). Compared with the Wuhan-Hu-1  
209 reference strain, HEE-01 shows only two nucleotide differences that include one amino  
210 acid replacement. HEE-01 was assigned to the B lineage and grouped with sequences  
211 from the Netherlands (Figure 2). Strain HGSQ-USFQ-018 was obtained from a case  
212 detected in Quito, but the infection was presumably acquired in Guayaquil. It shows  
213 nine nucleotide differences with the reference genome (Table 1). This genome is  
214 placed in the B.1 lineage and grouped with genomes from other European countries  
215 (Figure 2). Finally, HGSQ-USFQ-007, a community acquired infection in Quito, and the  
216 sequence from an epidemiologically-linked contact (HGSQ-USFQ-010) were also



217 placed in the B lineage. Both sequences have multiple differences compared to the  
218 reference strain, including 17 amino acids replacement and one amino acid deletion  
219 (Table 1).

220

## 221 **DISCUSSION**

222 The mitigation of the spread of SARS-CoV-2 has been unsuccessful in most  
223 South American countries. Countries like Ecuador, where the population, despite  
224 shocking examples in other cities like Guayaquil, continues to react in disbelief at an  
225 imminent risk. It is now when scientific research is needed to guide the authorities to  
226 develop correct public health policies. In this study, we were able to uncover the first 4  
227 SARS-CoV-2 whole genome sequences from Ecuador using the portable low-cost  
228 MinION platform (Oxford Nanopore). Genome sequences, from the first COVID19  
229 case reported in Quito, have few differences compared to the reference strain (Wuhan-  
230 Hu-1), and is grouped within lineage B near sequences from patients in the  
231 Netherlands. This finding confirms that the patient most likely acquired the infection in  
232 his country of origin. Metagenome sequencing of the patient's BAL revealed possible  
233 bacterial coinfection as several of the microorganisms are not considered members of  
234 the core lower airways bacterial microbiota<sup>17</sup>. As consequence, the encountered  
235 bacteria were probably contributing to the patient's severe condition. Metagenomic  
236 analysis could be crucial in COVID-19 cases because it adds valuable information  
237 about concurrent pulmonary pathogens that can aggravate the clinical condition<sup>18-20</sup>.

238 acquired the infection at the community level, either in Guayaquil or Quito. All  
239 three patients presented severe respiratory symptoms. Sequence recovered from a  
240 patient that got infected in Guayaquil grouped in lineage B.1, suggesting a second  
241 introduction of the virus into the country. Finally, the virus sequences recovered from  
242 two patients infected in Quito had marked differences with the previous sequences,

243 suggesting a third introduction of the virus to the country. These sequences were  
244 identical between them, which was expected since the patients were family members,  
245 developed symptoms at the similar time and lived together. Furthermore, the distinct  
246 set of mutations from these sequences was analyzed in a separate study which reports  
247 a potential virulence factor not previously described, which is associated with inhibition  
248 of the interferon response *in vitro*<sup>21</sup>. Community acquired transmission is now  
249 widespread in Ecuador and more samples from community acquired infections are  
250 needed in order to inform analyses of the local epidemiology of the virus. Comparing  
251 these genomes will allow us to provide a more accurate estimate of the arrival times  
252 and routes of the virus into certain areas, and to better understand how it is locally  
253 spreading and evolving. Revealing the information contained in the SARS-CoV-2 virus  
254 genome sequences is a robust tool to understand the epidemiology of COVID-19  
255 locally and contribute to the global panorama.

256         The current pandemic is unlikely to be the last, and it is therefore essential to  
257 improve the response capacity of our public health systems and to implement and  
258 strengthen continuous scientific research programs. Only in this way we will be able to  
259 better understand these types of threats, act based on evidence, and thus reduce their  
260 impact.

## 261 **ACKNOWLEDGEMENTS**

262 This work was funded by Universidad San Francisco de Quito and CADDE project  
263 ([www.caddecentre.org/](http://www.caddecentre.org/)). P.C. is funded by NIH FIC D43TW010540 Global Health  
264 Equity Scholars.

265 We acknowledge the submitting authors and institutions of the sequences of SARS  
266 CoV 2 from GISAID Database.

## 267 **AUTHOR CONTRIBUTIONS**

268 Conceptualization: M.J., G.T., M.G., P.C. Methodology: S.M., B.P.-V., J.J.G., M.T.,  
269 P.C. Software: B.P.-V., V.B., P.C. Validation: B.P.-V., V.B. and P.C. Formal analysis  
270 and investigation: V.B., P.R.-S. and P.C. Writing - Original Draft: P.R.-S. and P.C.  
271 Writing - Review & Editing: All authors. Visualization: V.B. and P.C. and M.A.  
272 Supervision: G.T. Funding acquisition: M.G, G.T. & P.C.

273 **REFERENCES**

- 274 1. Rodríguez-Morales AJ, Gallego V, Escalera-Antezana JP, et al. COVID-19 in  
275 Latin America: The implications of the first confirmed case in Brazil. *Travel Med*  
276 *Infect Dis*. Published online 2020. doi:10.1016/j.tmaid.2020.101613
- 277 2. Rodríguez-Morales AJ, MacGregor K, Kanagarajah S, Patel D, Schlagenhauf P.  
278 Going global – Travel and the 2019 novel coronavirus. *Travel Med Infect Dis*.  
279 Published online 2020. doi:10.1016/j.tmaid.2020.101578
- 280 3. Ecuador confirma su primer caso de coronavirus. National Geographic en  
281 español. Published 2020. Accessed March 30, 2020.  
282 [https://www.ngenespanol.com/el-mundo/ecuador-confirma-su-primer-caso-de-](https://www.ngenespanol.com/el-mundo/ecuador-confirma-su-primer-caso-de-coronavirus/)  
283 [coronavirus/](https://www.ngenespanol.com/el-mundo/ecuador-confirma-su-primer-caso-de-coronavirus/)
- 284 4. SITUACIÓN NACIONAL POR COVID-19 (CORONAVIRUS) - INFOGRAFÍA  
285 N°094. Ministerio de Salud Pública Ecuador. Published 2020.  
286 [https://www.gestionderiesgos.gob.ec/wp-](https://www.gestionderiesgos.gob.ec/wp-content/uploads/2020/05/INFOGRAFIA-NACIONALCOVI-19-COE-NACIONAL-31052020-08h002.pdf)  
287 [content/uploads/2020/05/INFOGRAFIA-NACIONALCOVI-19-COE-NACIONAL-](https://www.gestionderiesgos.gob.ec/wp-content/uploads/2020/05/INFOGRAFIA-NACIONALCOVI-19-COE-NACIONAL-31052020-08h002.pdf)  
288 [31052020-08h002.pdf](https://www.gestionderiesgos.gob.ec/wp-content/uploads/2020/05/INFOGRAFIA-NACIONALCOVI-19-COE-NACIONAL-31052020-08h002.pdf)
- 289 5. Max Roser, Hannah Ritchie EO-O and JH. Coronavirus Pandemic (COVID-19).  
290 Our World in Data. Published 2020. <https://ourworldindata.org/coronavirus>
- 291 6. Andersen KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF. The proximal origin  
292 of SARS-CoV-2. *Nat Med*. Published online 2020. doi:10.1038/s41591-020-  
293 0820-9
- 294 7. Kupferschmidt K. Genome analyses help track coronavirus' moves. *Science (80-*  
295 *).* Published online 2020. doi:10.1126/science.367.6483.1176
- 296 8. Shen Z, Xiao Y, Kang L, et al. Genomic diversity of SARS-CoV-2 in Coronavirus  
297 Disease 2019 patients. *Clin Infect Dis*. Published online 2020.  
298 doi:10.1093/cid/ciaa203
- 299 9. GISAID. GISAID - EpiCoV. Published 2020. [www.epicov.org](http://www.epicov.org)
- 300 10. Greninger AL, Naccache SN, Federman S, et al. Rapid metagenomic  
301 identification of viral pathogens in clinical samples by real-time nanopore  
302 sequencing analysis. *Genome Med*. Published online 2015. doi:10.1186/s13073-  
303 015-0220-9
- 304 11. Kafetzopoulou LE, Efthymiadis K, Lewandowski K, et al. Assessment of  
305 metagenomic Nanopore and Illumina sequencing for recovering whole genome  
306 sequences of chikungunya and dengue viruses directly from clinical samples.  
307 *Eurosurveillance*. Published online 2018. doi:10.2807/1560-  
308 7917.ES.2018.23.50.1800228
- 309 12. Quick J, Grubaugh ND, Pullan ST, et al. Multiplex PCR method for MinION and  
310 Illumina sequencing of Zika and other virus genomes directly from clinical  
311 samples. *Nat Protoc*. Published online 2017. doi:10.1038/nprot.2017.066
- 312 13. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. NanoPack:  
313 Visualizing and processing long-read sequencing data. *Bioinformatics*. Published  
314 online 2018. doi:10.1093/bioinformatics/bty149
- 315 14. Singer JB, Thomson EC, McLauchlan J, Hughes J, Gifford RJ. GLUE: A flexible  
316 software system for virus sequence data. *BMC Bioinformatics*. Published online  
317 2018. doi:10.1186/s12859-018-2459-9

- 318 15. Menzel P, Ng KL, Krogh A. Fast and sensitive taxonomic classification for  
319 metagenomics with Kaiju. *Nat Commun*. Published online 2016.  
320 doi:10.1038/ncomms11257
- 321 16. Li H. Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics*.  
322 Published online 2018. doi:10.1093/bioinformatics/bty191
- 323 17. Hilty M, Burke C, Pedro H, et al. Disordered microbial communities in asthmatic  
324 airways. *PLoS One*. 2010;5(1):e8578. doi:10.1371/journal.pone.0008578
- 325 18. Moore SC, Penrice-Randal R, Alruwaili M, et al. Amplicon based MinION  
326 sequencing of SARS-CoV-2 and metagenomic characterisation of  
327 nasopharyngeal swabs from patients with COVID-19. *medRxiv*. Published online  
328 2020. doi:10.1101/2020.03.05.20032011
- 329 19. Arabi YM, Al-Omari A, Mandourah Y, et al. Critically Ill patients with the middle  
330 east respiratory syndrome: A multicenter retrospective cohort study. *Crit Care*  
331 *Med*. Published online 2017. doi:10.1097/CCM.0000000000002621
- 332 20. Chen N, Zhou M, Dong X, et al. Epidemiological and clinical characteristics of 99  
333 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive  
334 study. *Lancet*. Published online 2020. doi:10.1016/S0140-6736(20)30211-7
- 335 21. Konno Y, Kimura I, Urie K, et al. SARS-CoV-2 ORF3b is a potent interferon  
336 antagonist whose activity is further increased by a naturally occurring elongation  
337 variant. *bioRxiv*. Published online 2020. doi:10.1101/2020.05.11.088179
- 338
- 339

340 **Figures and tables legends:**

341

342 **Figure 1. Krona chart summarizing the percentage of sequences assigned to**  
343 **Coronavirus in the metagenome.** Coronavirus sequences represented 0.036% of all  
344 microorganisms.

345

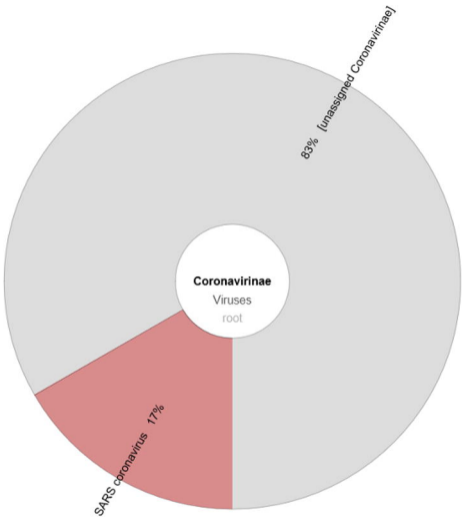
346

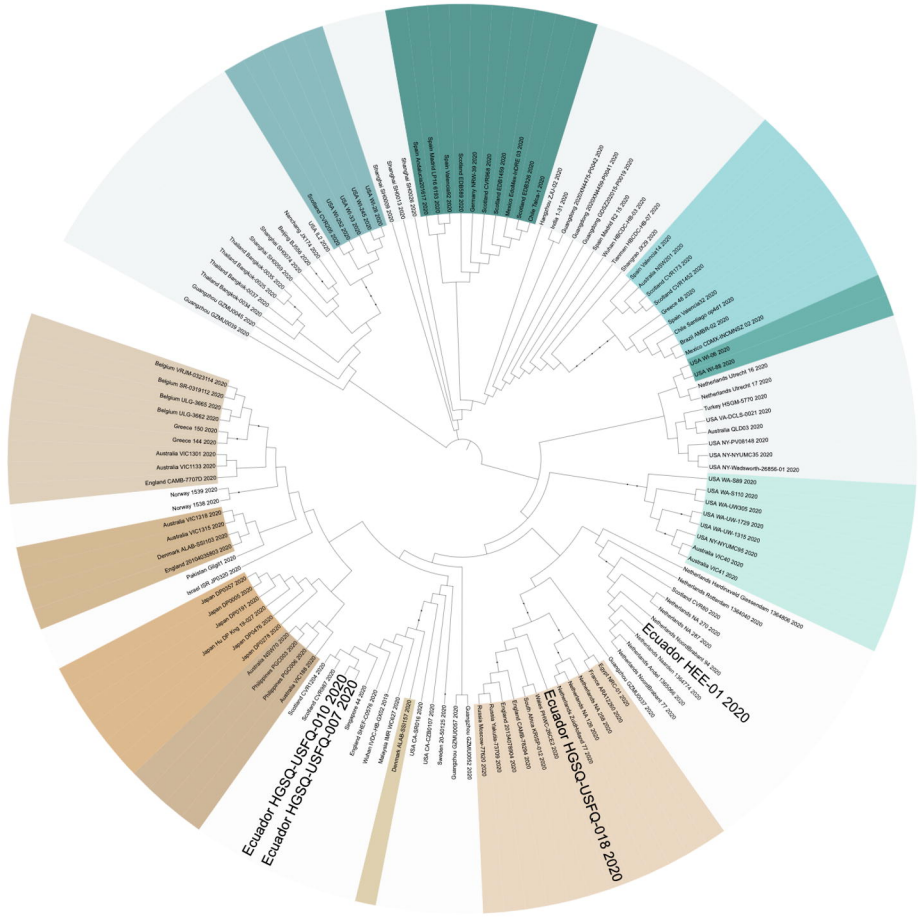
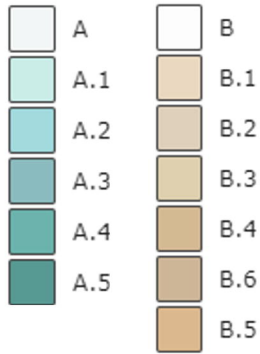
347 **Figure 2. Maximum parsimony phylogenetic tree.** Ecuadorian sequences are  
348 highlighted in bold. Lineage assignments are highlighted in similar colors: lineage A  
349 (Green color palette) and lineage B (Brown color palette).

350

351 **Table1.** Mutations found in four SARS-CoV-2 Genomes from Ecuador compared to  
352 Wuhan-Hu-1 (GenBank accession number MN908947)

353





**Table1. Mutations found in four SARS-CoV-2 Genomes from Ecuador compared to Wuhan-Hu-1 (GenBank accession number MN908947).**

Genome Sequence ID	Genome Sequence GISAID ID	Lineage	Gene	Mutation	Amino acid	Type
HEE-01	EPI_ISL_417482	B	ORF1ab	T514C	-	-
			S	A25182T	E1207V	Transversions
HGSQ-USFQ-018	EPI_ISL_422563	B.1	ORF1ab	C3037T	no	Transition
			ORF1ab	C4010T	L431F	Transition
			ORF1ab	C14408T	P323L	Transition
			ORF1ab	C18555T	no	-
			S	A23403G	D614G	Transition
			ORF3a	A25983G	no	-
			N	G28881A	R203K	Transition
			N	G28882A	no	-
			N	G28883C	G204R	Transversions
HGSQ-USFQ-007	EPI_ISL_422564	B	ORF1ab	A1996C	-	-
	and		ORF1ab	C1997T	no	Transition
HGSQ-USFQ-010	EPI_ISL_422565	B	ORF1ab	T11737G	N255K	Transversions
			ORF1ab	T20497A	N255K	Transversions
			S	G22424T	A288S	Transversions
			S	C24418T	no	Transition
			ORF3a	A25505C	Q38P	Transversions
			ORF3a	G25647T	L85F	Transversions
			ORF3a	T25676C	L95S	Transition
			ORF3a	C25678T	L96F	Transition
			ORF3a	C25692T	no	Transition
			ORF3a	T25695G	-	-
			ORF3a	C25708T	L106F	Transition
			ORF3a	C25710T	no	Transition
			ORF3a	T25711C	Y107H	Transition
			ORF3a	C25728T	no	Transition
			ORF3a	A25741G	S117G	Transition
			ORF3a	A25756G	R122E	Transition
			ORF3a	G25757A	no	Transition
			ORF3a	A25764T	-	-
			ORF3a	T25779G	-	-
			ORF3a	G25781A	C130*	Transition
			ORF3a	C25782A	-	-
			ORF3a	G25879A	V163T	Transition
			ORF3a	T25880C	no	Transition
			ORF3a	C25883A	T164N	Transversions
			ORF3a	T25885G	S165A	Transversions
			ORF8	A28035G	R48G	Transition
			ORF8	A28037G	no	Transition
			ORF8	G28038T	V49L	Transversions
			ORF8	C28076T	no	Transition
			ORF8	T28078C	V62A	Transition

Lineages were assigned using Pangolin COVID-19 Lineage Assigner ([pangolin.cog-uk.io](https://pangolin.cog-uk.io)).