medRxiv preprint doi: https://doi.org/10.1101/2021.03.19.21253620; this version posted March 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

1 **SARS-CoV-2 genome sequencing from COVID-19 in Ecuadorian patients: a whole**
2 **country analysis**

3 Sully Márquez[1], Belén Prado-Vivar[1,2], Juan José Guadalupe[4], Mónica Becerra-Wong[1]
4 Bernardo Gutierrez[4,5], Clinical COVID-19 Ecuador Consortium, Juan Carlos
5 Fernández-Cadena[8], Derly Andrade-Molina[8], Gabriel Morey-Leon [9], Miguel Moncayo[1],
6 Rommel Guevara[1], Josefina Coloma[7], Gabriel Trueba[1], Michelle Grunauer[3,6], Verónica
7 Barragán[1], Patricio Rojas-Silva[1], Paúl Cárdenas[1,2*]

8

9 [1]Universidad San Francisco de Quito, COCIBA, Instituto de Microbiología
10 [2]Universidad San Francisco de Quito, Centro de Bioinformática
11 [3]Universidad San Francisco de Quito, COCSA, Escuela de Medicina
12 [4]Universidad San Francisco de Quito, COCIBA, Laboratorio de Biotecnología Vegetal
13 [5]Departament of Zoology, University of Oxford
14 [6]Unidad de Cuidados Intensivos, Hospital de los Valles, Quito
15 [7]School of Public Health, University of California, Berkeley
16 [8] Universidad Espíritu Santo, Laboratorio de Omicas
17 [9] Universidad de Guayaquil
18
19
20
21
22

23 *Correspondence: pacardenas@usfq.edu.ec, tel.: +593-22971700 ext. 1517, fax number
24 +593-22971700

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.03.19.21253620; this version posted March 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

25

26 **Article Summary Line**

27 We report 119 sequences of SARS-CoV-2 across all Ecuadorian provinces, 20 different
28 lineages were found until January 17th, including B.1.1.7.

29 **Running Title**

30 Genomic surveillance of SARS-Cov-2 across Ecuador.

31 **Abstract**

32 SARS-CoV-2, the etiological agent of COVID-19, was first described in Wuhan, China
33 in December 2019 and has now spread globally. Ecuador was the second country in South
34 America to confirm cases and Guayaquil was one of the first cities in the world to
35 experience high mortality due to COVID-19. The aim of this study was to describe the
36 lineages circulating throughout the country and to compare the mutations in local
37 variants, to the reference strain. In this work we used the MinION platform (Oxford
38 Nanopore Technologies) to sequence the whole SARS-CoV-2 genomes of 119 patients
39 from all provinces of Ecuador, using the ARTIC network protocols. Our data from lineage
40 assignment of the one hundred and nineteen whole genomes revealed twenty different
41 lineages. All genomes presented differences in the S gene compared to the Wuhan
42 reference strain, being the D614G amino acid replacement the most common change. The
43 B.1.1.119 lineage was the most frequent and was found in several locations in the Coast
44 and Andean region. Three sequences were assigned to the new B.1.1.7 lineage. Our work
45 is an important contribution to the understanding of the epidemiology of SARS- CoV-2
46 in Ecuador and South America.

medRxiv preprint doi: https://doi.org/10.1101/2021.03.19.21253620; this version posted March 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

47

48

49 **Highlights**

50 - All 119 genomes showed mutations compared to the reference strain, which could be
51 important to understand the virulence, severity and transmissibility of the virus.

52 - Until January 17, three sequences were assigned to the new B.1.1.7 lineage.

53 - Our findings suggest that there were at least twenty independent introductions of
54 SARS-CoV-2 to Ecuador.

55

56 **Key words**: SARS-CoV-2, MinION sequencing, COVID-19, whole genome
57 sequencing, Ecuador.

58

59 **INTRODUCTION**

60 SARS-CoV-2, the etiologic agent of COVID-19, has spread globally reaching all
61 continents (1). Ecuador was the second country in South America, after Brazil, to report
62 its first confirmed case, on February 29, 2020 (2). The number of cases has grown ever
63 since, reaching 241,567 laboratory qPCR confirmed cases and 14,668 deaths by Jan 28,
64 2021 (3). The pandemic has had a severe impact in the Ecuadorian population. It is
65 suggested that up to 838.35 people per million inhabitants may have died since the first
66 COVID-19 case. Despite of this being the official data, mortality rate could be bias due
67 to scarce testing.

medRxiv preprint doi: https://doi.org/10.1101/2021.03.19.21253620; this version posted March 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

68 Understanding the local epidemiology of SARS-CoV-2 can be greatly enhanced
69 from understanding the viral evolution, establish the origin of variants (4), transmission
70 of variants across different regions (5), the genetic diversity of variants in the population
71 and to identify notable mutations (6). Additionally, key biological aspects such as
72 virulence, transmissibility and infectivity of the circulating lineages can be investigated,
73 when combined with clinical data from the patients (5). The importance of genomic
74 surveillance is underscored with the current emergence of variants with potential
75 increased transmission or disease severity like the B.1.1.7, P.1 and B.1.351. As a global
76 effort, over 751,513 genome sequences of SARS-CoV-2 have been reported to date
77 globally and are available in GenBank and the GISAID (Global Initiative on Sharing All
78 Influenza Data) repositories (7).

79 Previous to this work there is no information about the SARS-CoV-2 lineages circulating
80 in Ecuador. Here, we report one hundred and nineteen SARS-CoV-2 whole genome
81 sequences sampled from every province of the country in the coastal, Andean and
82 Amazonian regions plus the Galapagos archipielago. We identify lineages that circulated
83 among the population from March 2020 through January 2021. We also describe the
84 mutations present in these genomes as they could influence virulence, transmission and
85 infectivity. The sequencing was performed using the portable MinION platform (Oxford
86 Nanopore Technologies).

87 MATERIALS AND METHODS

88 Epidemiological information and sample collection

89 Nasopharyngeal swabs (NS) or broncho-alveolar lavage (BAL) samples were collected
90 from patients in public third-level hospitals located in different provinces of Ecuador.
91 Sample positivity for SARS-CoV-2 using standard RT-qPCR protocols, was officially

medRxiv preprint doi: https://doi.org/10.1101/2021.03.19.21253620; this version posted March 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

92 reported to hospitals by the Ecuadorian Ministry of Public Health (MSP) and National
93 Institute of Public Health and Research (INSPI). The use of samples was approved by the
94 Bioethics Committee of Universidad San Francisco de Quito (CEISH No. 1234). The
95 BAL samples were collected in a sterile tube with 2X DNA/RNA Shield (Zymo), and NS
96 were immersed in 1X DNA/RNA Shield (Zymo), to ensure virus inactivation and
97 preservation of the genetic material. Samples were transported immediately at 4ºC to the
98 Institute of Microbiology at USFQ (IM-USFQ) in a sealed container with all the
99 biosecurity and containment measures recommended by the CDC of the USA
100 (https://www.fda.gov/media/134922/download).

101 **RNA extraction**

102 The genetic material from samples was extracted in a biosafety type II chamber with
103 HEPA filters in the Virology Laboratory at IM-USFQ. The SV Total RNA Isolation
104 System (Promega, USA) and Quick RNA viral, w/zymo-Spin IC (Zymo, USA) kits were
105 used to extract RNA from samples for whole genome sequencing. A pre-digestion step
106 was added to the RNA extraction protocol of the BAL samples as follows. Before nucleic
107 acid extraction, 280 µl of the BAL samples were predigested with 360 µl of PureLink™
108 Genomic Lysis Buffer and 20 µl of proteinase K. The mix was incubated at 55ºC for 10
109 minutes vortexing every 5 minutes (Life Technologies, USA). All RNA extractions were
110 performed following manufacturer instructions. Retro-transcription was carried out using
111 the Protocol of the Public Health England Genomics Lab (8,9) at the USFQ
112 Bioinformatics Center and ARTIC protocol (10).

113 **Viral whole genome sequencing**

114 The Primer Scheme (V1 and V3) developed by the ARTIC network for nCoV-2019 was
115 employed to generate an amplicon tiling path across the viral genome (10,11). The final

medRxiv preprint doi: https://doi.org/10.1101/2021.03.19.21253620; this version posted March 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

116 product of multiplex PCR was quantified using Qubit dsDNA HS (High Sensitivity)
117 Assay Kit (Thermo Scientific, Invitrogen, Carlsbad, CA, USA). cDNA MinION library
118 preparation was performed using the Rapid Barcoding kit (SQK-RBK004), native
119 barcoding kit (NB-114) with ligation sequencing kit (LSK-109) following manufacturer
120 instructions and then loaded into a MinION flow cell (FLO-MIN 106). Basecalling of
121 FAST5 files was performed using Guppy (version 3.4.5) (12) (Oxford Nanopore
122 Technologies). Also, the RAMPART software (v1.0.5) from the ARTIC Network
123 (https://github.com/artic-network/rampart) was used to monitor sequencing in real-time.
124 Sequence quality scoring, demultiplexing and adapter removal was performed with the
125 NanoPlot (13) and Porechop (version 0.2.4) algorithms, respectively
126 (https://github.com/rrwick/Porechop). The ARTIC Network bioinformatics pipeline was
127 used for variant calling, and the reads were mapped against the reference strain Wuhan-
128 Hu-1 (GenBank accession number MN908947), to generate consensus genomes. Tablet
129 alignment viewer (version 1.19.09.3) (https://ics.hutton.ac.uk/tablet) was used to
130 visualize the mapped sequence. The online tool NextClade (14) was used to determine
131 the genomes clades. Then, genomes were uploaded to the CoV-GLUE resource (15), to
132 determine the mutations, epidemiological linkage of circulating SARS-CoV-2 variants
133 and primer mismatches. Lineage classification was carried out with Pangolin online
134 software (https://github.com/cov-lineages/pangolin). Visualization of sequences in a
135 phylogenetic tree was performed using the following strategy. The 119 Ecuadorian SARS
136 COV 2 genomes of the present study were aligned with eight hundred and sixty one
137 genomes from GISAID using NextAlign tool. A maximum likelihood phylogenetic
138 reconstruction was performed with GTR substitution model and 1000 bootstrap
139 resampling using IQtree 2.1.1. The phylogenetic tree was visualized with iTOL tool.

140

medRxiv preprint doi: https://doi.org/10.1101/2021.03.19.21253620; this version posted March 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

141 **RESULTS**

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

143 Sequencing results demonstrated that there were numerous SNPs in the ORF1a, ORF3a,
144 ORF 7a, ORF7b, ORF8, S, M and N genes **(Figure 1 and Supplementary Table 1);**
145 interestingly, the number of SNPs in the E gene was null. Our results confirmed the
146 presence of several hotspot mutations in different viruses infecting the Ecuadorian
147 communities. Phylogenetic analysis showed that Ecuadorian genomes were clustered
148 with genomes from other countries, suggesting multiple entries **(Supplementary figure**
149 **1)**

150 Twenty lineages were identified in which the dominant lineages were: B.1 (n=31)
151 followed by B.1.1.119 (n=30), B.1.1.1 (n=10), B.1.1.207 (n=8), B.1.1.10 (n=6) and
152 B.1.1.67 (n=5) (see Table S1 for all lineages). Interestingly some lineages were found
153 only in 4 of the 24 provinces: A.1, B.1.9, B.1.371, B (Pichincha), A (Santo Domingo),
154 B.1.308 (Zamora-Chinchipe) and B.1.6, B.1.325 (Imbabura). Lineages B.1.1.119 and B.1
155 were found in Awa and Waorani Amazonian indigenous communities respectively
156 **(Table 1)**, **(Supplementary Table 1)**.

157 Three of the analyzed sequences showed the aminoacid changes T183I, A890D, I1412T,
158 P323L, L493F, N501Y, T553I, A570D, D614G, P681H, T716I, S982A, D1118H, Q27,
159 R52I, Y73C, D3L, R203K, G204R and S235F placing them in the B.1.1.7 lineage. All
160 sequences showed mutations at the S gene, being the most prevalent A23403G (95.79%
161 sequences) conferring the D614G aminoacid change. **(Supplementary Table 1) (Table**
162 **2).**

medRxiv preprint doi: https://doi.org/10.1101/2021.03.19.21253620; this version posted March 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

163 Primer mismatches were analyzed with seven COVID-19 diagnostic kits. All sequences
164 showed changes at RdRP gene of Charité RT-PCR kit in positions 15469-15494, 15505-
165 15530, and15431-15452.Mismatches were also found for the National Institute of
166 Infectious Diseases Japan, CDC USA and CDC China RT-PCR diagnostic kits,
167 **Supplementary Table 2.**

168

169 **DISCUSSION**

170 Ecuador is one of the countries that has been severely affected by SARS-COV2
171 (1). Despite of the non-pharmaceutical interventions for preventing its dissemination, the
172 virus continued to spread in all provinces. As most Latin American countries, Ecuador
173 has had limited success in lowering transmission curves, mostly due to limited testing and
174 socio-economic reasons(16). Public and private hospitals in Ecuador have continuously
175 been flooded with patients with severe COVID-19 during the last year. In this study, we
176 analyzed 119 SARS-CoV-2 whole genome sequences from all 24 provinces, using the
177 third-generation sequencing MinION platform (Oxford Nanopore).

178

179 A total of twenty lineages were found in the 24 provinces. Unique lineages were
180 detected only in Pichincha, Imbabura (Andes), Santo Domingo (Foothills) and Zamora-
181 Chinchipe (Southern Amazon) provinces, however we cannot stablish that these lineages
182 are not present in other provinces due to the limited genomic surveillance in the country
183 Additionally, HGSQ-USFQ-007 and HGSQ-USFQ-010 genomes showed a distinct set
184 of mutations. A mutation which may be increasing the virulence due to inhibition of the
185 interferon response *in vitro* was found in these genomes (17)
186 SARS-COV2 is constantly changing through mutations, that led to the
187 occurrence of new variants. Mutations in the S gene that confer a selective advantage to

medRxiv preprint doi: https://doi.org/10.1101/2021.03.19.21253620; this version posted March 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

188 the virus had become more common. An example is A23403G mutation (D614G amino
189 acid change) which was present in 95.79% of sequences from this study (18).
190
191 However, from all the variants that have been detected worldwide, only a small
192 number are of public health concern. Three variants of concern (VOC) had emerged:
193 B.1.1.7 in England, B.1.351 in South Africa, and P.1 in Brazil considering their
194 improvements in viral replication mechanisms, transmissibility, and capacity to evade the
195 immune host response. All these factors could have a direct impact on the efficacy of
196 vaccines (18,19). In Ecuador B.1.1.7 was identified in three sequences from Los Ríos
197 province, introduced by travelers from other countries Currently local authorities have
198 reported that this variant is showing community-acquired transmission (20).
199
200 The presence of multiple lineages of SARS-COV-2 in Ecuador could be explained
201 by independent introductions from different countries. The genomic surveillance results
202 offer a high-resolution picture of the virus spread in the community probably due to the
203 frequent movement of people between provinces. The current pandemic is unlikely to be
204 the last, and it is, therefore, essential to improve the response capacity of our public health
205 systems and to implement and strengthen continuous scientific research programs. Only
206 in this way, we will be able to better understand these types of threats, act based on
207 evidence, and thus reduce their impact. Community-acquired transmission is widespread
208 in Ecuador and more samples from community-acquired infections are needed to inform
209 analyses of the local epidemiology of the virus.
210 Medical and scientific efforts throughout the country made possible to compile at least
211 one sequence from each province, however we must join efforts with other universities
212 and state laboratories to increase genomic surveillance in the country

medRxiv preprint doi: https://doi.org/10.1101/2021.03.19.21253620; this version posted March 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

213 We are aware that the small number of samples reported in this study might not capture
214 the full picture of what is happening in Ecuador. However, revealing the information
215 contained in the SARS-CoV-2 virus genome sequences is a robust tool to understand the
216 epidemiology of COVID-19 locally. The present study is a steppingstone to understand
217 the importance of carrying local genomic surveillance for infectious diseases circulating
218 in the country.

219

220 **ACKNOWLEDGEMENTS**

221 This work was funded by Universidad San Francisco de Quito through Emergency
222 Research Funds and COCIBA grants (P.R-S.) and CADDE project
223 (www.caddecentre.org/). P.C. is funded by NIH FIC D43TW010540 Global Health
224 Equity Scholars. Lourdes Torres, Diego Quiroga and the people of USFQ-Proyectos for
225 their support on the project development.

226 We acknowledge the submitting authors and institutions of the sequences of SARS-CoV-
227 2 from GISAID Database.

228 **CONSORTIA**

229 The members of Clinical COVID-19 Ecuador Consortium in alphabetical order are
230 Alejandra Ramones, Alexandra Tino, Andrea Carrera, Andrea Macias, Anita Garcia,
231 Carlos Guerrero, Carlos Mena, Carlos Tobar, Carolina Proaño-Bolaños, Damaris
232 Zandoya, Dayron Brossard, Diana Zambrano, Diego Egas, Eddy Chavez, Edison Chavez,
233 Edison Ligña, Edy Quizhpe, Eulalia Pazmiño, Fabian Aguilar, Fausto Maldonado,
234 Francisco Córdova, Francisco Mora, Francisco Rodriguez, Franklin Espinoza, Freddy

medRxiv preprint doi: https://doi.org/10.1101/2021.03.19.21253620; this version posted March 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

235 Iza, Fredy Loor, Gabriel Morey, Geovanny Cazorla, Giovanna Moran, Grace Salazar,
236 Hermelinda Paguay, Jonathan Araujo, Jorge Luis Velez, Jorge Montaño, Jorge Reyes,
237 Juan Gaviria, Juan Zuñiga, Karina Barragan, Karolina Pacheco, Katherine Apunte,
238 Khurram Mahbbob, Killen Briones-Claudette, Killen Briones-Zamora, Ligia Briceño,
239 Manuel Jaramillo, Manuel Jibaja, Marcelo Ortiz, Marcos Di Stefano, Maureen Mosquera,
240 Milton Tobar, Nabih Dahik, Nina Espinoza de los Monteros, Ninfa Henriquez, Patricio
241 Reyes, Rene Bracho, Rosario Erazo, Sonia Sislema, Stalin Castillo, Stephanie Arregui,
242 Tania Guayasamin, Yeimy Rojas and Yomara Napa.

243 **AUTHOR CONTRIBUTIONS**

244 Conceptualization: M.J., G.T., M.G., P.C. Methodology: S.M., B.P.-V., J.J.G., M.T., P.C.
245 Software: B.P.-V., V.B., M.B., P.C. Validation: B.P.-V., V.B. and P.C. Formal analysis
246 and investigation: V.B., P.R.-S., M.B, P.C. Writing - Original Draft: P.R.-S. and P.C.
247 Writing - Review & Editing: All authors. Visualization: V.B. and P.C. and M.A
248 Supervision: G.T. Funding acquisition: M.G, G.T.,J.C & P.C. The Clinical COVID-19
249 Ecuador Consortium collected samples and provided epidemiological information for all
250 the sequences.

251 **Funding:** The current project has been funded by USFQ through its Emergency grants,
252 and CADDE through donations of reagents.

253 **Conflict of Interest:** None of the authors report conflicts of interest.

254 **Ethical approval:** Ethical approval for all samples was given by CEISH-USFQ
255 (Comité de Ética de Investigación en Seres Humanos-USFQ): IE-JP067-2020-CEISH-
256 USFQ. Additionally by the Ministerio de Salud Publica Comité Expedito (021-2020)

medRxiv preprint doi: https://doi.org/10.1101/2021.03.19.21253620; this version posted March 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

257 **Informed consent:** Every patient or relative (in case of severely ill patients) has signed
258 the informed consent.

259 **Biographical Sketch:** SM is a graduate student at the Instituto de Microbiología in
260 Universidad San Francisco de Quito and interested in virology research.

261

262 **References**

263 1. Mercatelli D, Giorgi FM. Geographic and Genomic Distribution of SARS-CoV-2
264 Mutations. Front Microbiol. 2020;11(July):1–13.

265 2. National Geographic en español. Ecuador confirma su primer caso de
266 coronavirus [Internet]. 2020 [cited 2020 Mar 30]. Available from:
267 https://www.ngenespanol.com/el-mundo/ecuador-confirma-su-primer-caso-de-
268 coronavirus/

269 3. Ministerio de Salud Pública Ecuador. Ministerio de Salud Pública Ecuador.
270 SITUACIÓN NACIONAL POR COVID-19- INFOGRAFIA N°325 [Internet].
271 2021 [cited 2021 Jan 28]. Available from:
272 https://www.gestionderiesgos.gob.ec/wp-
273 content/uploads/2021/01/INFOGRAFIA-NACIONALCOVID19-COE-
274 NACIONAL-08h00-17012021-v2.pdf

275 4. Andersen KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF. The proximal
276 origin of SARS-CoV-2. Nat Med [Internet]. 2020;26(4):450–2. Available from:
277 https://doi.org/10.1038/s41591-020-0820-9

278 5. Kupferschmidt K. Genome analyses help track coronavirus' moves. Science
279 [Internet]. 2020;367(6483):1176–7. Available from:

medRxiv preprint doi: https://doi.org/10.1101/2021.03.19.21253620; this version posted March 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

280 http://europepmc.org/abstract/MED/32165562

281 6. Shen Z, Xiao Y, Kang L, Ma W, Shi L, Zhang L, et al. Genomic Diversity of
282 Severe Acute Respiratory Syndrome-Coronavirus 2 in Patients With Coronavirus
283 Disease 2019. Clin Infect Dis [Internet]. 2020 Jul 28;71(15):713–20. Available
284 from: https://pubmed.ncbi.nlm.nih.gov/32129843

285 7. GISAID. GISAID - EpiCoV. 2020.

286 8. Greninger AL, Naccache SN, Federman S, Yu G, Mbala P, Bres V, et al. Rapid
287 metagenomic identification of viral pathogens in clinical samples by real-time
288 nanopore sequencing analysis. Genome Med [Internet]. 2015;7(1):1–13.
289 Available from: http://dx.doi.org/10.1186/s13073-015-0220-9

290 9. Kafetzopoulou LE, Efthymiadis K, Lewandowski K, Crook A, Carter D, Osborne
291 J, et al. Assessment of metagenomic Nanopore and Illumina sequencing for
292 recovering whole genome sequences of chikungunya and dengue viruses directly
293 from clinical samples. Euro Surveill [Internet]. 2018 Dec;23(50):1800228.
294 Available from: https://pubmed.ncbi.nlm.nih.gov/30563591

295 10. Quick J. nCoV-2019 sequencing protocol. protocol.io. 2020;

296 11. Quick J, Grubaugh ND, Pullan ST, Claro IM, Smith AD, Gangavarapu K, et al.
297 Multiplex PCR method for MinION and Illumina sequencing of Zika and other
298 virus genomes directly from clinical samples. Nat Protoc [Internet].
299 2017;12(6):1261–76. Available from: https://doi.org/10.1038/nprot.2017.066

300 12. Wick RR, Judd LM, Holt KE. Performance of neural network basecalling tools
301 for Oxford Nanopore sequencing. Genome Biol [Internet]. 2019;20(1):129.
302 Available from: https://doi.org/10.1186/s13059-019-1727-y

medRxiv preprint doi: https://doi.org/10.1101/2021.03.19.21253620; this version posted March 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

303 13. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. NanoPack:
304 visualizing and processing long read sequencing data. Bioinformatics.
305 2018;34(15):2666–9.

306 14. Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, et al.
307 NextStrain: Real-time tracking of pathogen evolution. Bioinformatics.
308 2018;34(23):4121–3.

309 15. Singer JB, Thomson EC, McLauchlan J, Hughes J, Gifford RJ. GLUE: A flexible
310 software system for virus sequence data. BMC Bioinformatics. 2018;19(1):1–18.

311 16. Garcia PJ, Alarcón A, Bayer A, Buss P, Guerra G, Ribeiro H, et al. COVID-19
312 Response in Latin America. Am J Trop Med Hyg. 2020;103(5):1765–72.

313 17. Konno Y, Kimura I, Uriu K, Fukushi M, Irie T, Koyanagi Y, et al. SARS-CoV-2
314 ORF3b is a potent interferon antagonist whose activity is further increased by a
315 naturally occurring elongation variant. bioRxiv [Internet]. 2020 Jan
316 1;2020.05.11.088179. Available from:
317 http://biorxiv.org/content/early/2020/05/12/2020.05.11.088179.abstract

318 18. Plante JA, Liu Y, Liu J, Xia H, Johnson BA, Lokugamage KG, et al. Spike
319 mutation D614G alters SARS-CoV-2 fitness. Nature [Internet].
320 2020;(September). Available from: http://dx.doi.org/10.1038/s41586-020-2895-3

321 19. Walensky RP, Walke HT, Fauci AS. SARS-CoV-2 Variants of Concern in the
322 United States—Challenges and Opportunities. JAMA [Internet]. 2021 Feb 17;
323 Available from: https://doi.org/10.1001/jama.2021.2294

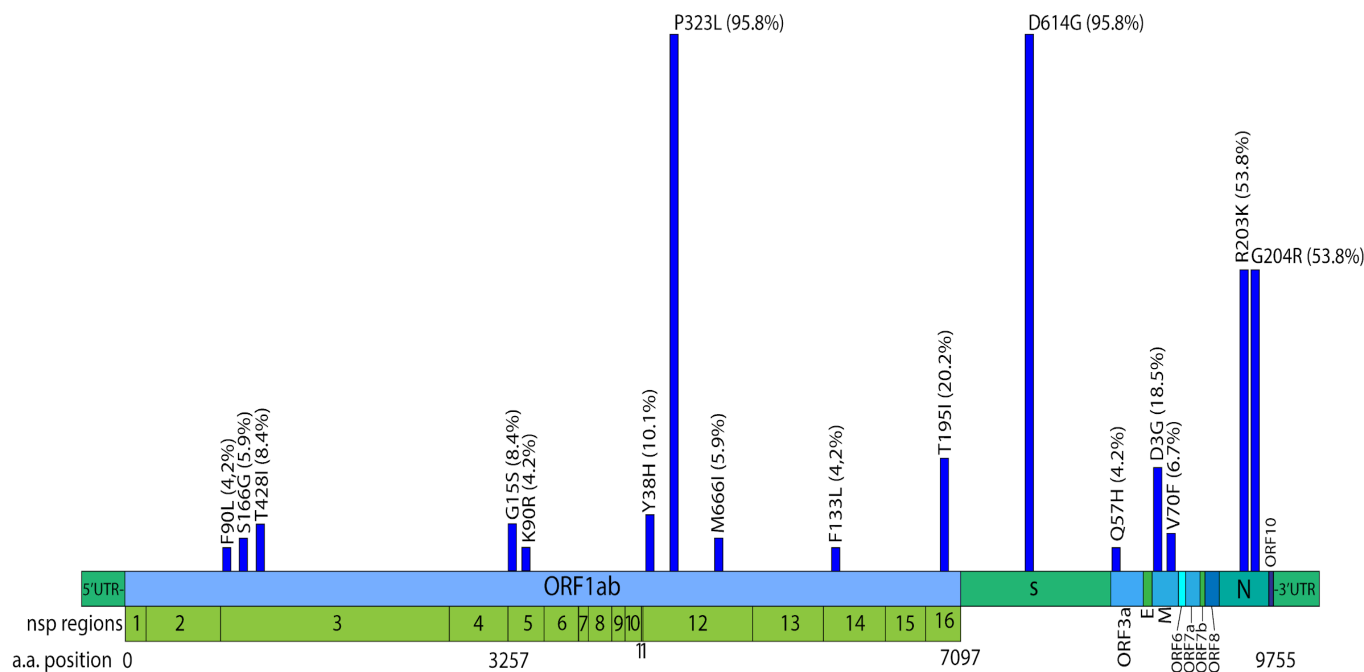
324 20. MSP. Estudios de Secuenciación Genómica del Inspi confirman la presencia de
325 nueva variación del SARS-COV-2 en Ecuador. 2021; Available from:

medRxiv preprint doi: https://doi.org/10.1101/2021.03.19.21253620; this version posted March 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

326 https://www.salud.gob.ec/estudios-de-secuenciacion-genomica-del-inspi-
327 confirman-la-presencia-de-nueva-variacion-del-sars-cov-2-en-ecuador/

328

329 **Figures and tables legends:**



330

331 **Figure 1.** Genome annotation and frequency of the amino acid changes identified in at
332 least 4% of the sequences. The amino acid changes are defined in comparison with the
333 Wuhan-Hu-1 reference genome (GenBank accession number MN908947).

334

335 **Table 1.** Lineages distribution in the 24 provinces of Ecuador.

Province	**Total population***	**Total confirmed cases****	**Cases per 100,00 hab**	**Lineages identified***

medRxiv preprint doi: https://doi.org/10.1101/2021.03.19.21253620; this version posted March 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Azuay	881,394	15,383	1,745.30	B.1.1.1, B.1.223, B.1 (n=4)
Bolívar	209,933	2,866	1,365.20	B.1.1.1, B.1.1.119 (n=3)
Cañar	281,396	3,321	1,180.19	B.1.1.119 (n=1)
Carchi	186,869	4,725	2,528.51	B.1.1.119 (n=1)
Cotopaxi	488,716	4,076	777.86	B.1, B.1.197 (n=2)
Chimborazo	524,004	6,84	1,399.59	B.1.1.207 (n=3)
Esmeraldas	643,654	10,023	1,400.35	B.1.223, B.1.1.119, B.1 (n=6)
El Oro	715,751	5,361	832.90	B.1.1.1, B.1 (n=4)
Galápagos	33,042	1,211	3,665.03	B.1 (n=1)
Guayas	4,387,434	30,882	703.87	B.1.1.1, B.1.1.119, B.1.1.10, B.1 (n=6)
Imbabura	476,257	7,188	1,509.27	B.1.1.207, B.1, B.1.67, B.1.1.119, B.1.223, B.1, B.1.1.1, B.1.6, B.1.1.31, B.1.1.10, B.1.325 (n=23)
Loja	521,154	8,694	1,668.22	B.1, B.1.1, B.1.1.119 (n=5)
Los Ríos	921,763	5,713	619.79	B.1.1.119, B.1.1.207, B.1, B.1.67, B.1.223, B.1.1.1, B.1.1.7 (n=14)
Manabí	1,562,079	16,94	1,084.45	B.1.1.119, B.1 (n=3)
Morona Santiago	196,535	4,112	2,092.25	B.1.1.119, B.1.1 (n=2)
Napo	133,705	1,888	1,412.06	B.1.1.119, B.1 (n=3)
Orellana	161,338	2,167	1,343.14	B.1.223, B.1 (n=4)
Pastaza	114,202	2,501	2,189.98	B.1.1.119 , B.1 (n=2)
Pichincha	3,228,233	84,367	2,613.41	B, B.1.1.119, B.1.14, B.1.371, B.1,

medRxiv preprint doi: https://doi.org/10.1101/2021.03.19.21253620; this version posted March 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

				B.1.67, A.1, B.1.223, B.1.9, (n=14)
Santa Elena	401,178	2,979	742.56	B.1 (n=1)
Santo Domingo de los Tsáchilas	458,58	6,61	1,441.41	A, B.1.1.31, B.1.1.10, B.1.1.159 (n=7)
Sucumbíos	230,503	3,183	1,380.89	B.1, B.1.1.159 (n=3)
Tungurahua	590,6	8,775	1,485.78	B.1.1.119, B.1.197, B.1.1.10 (n=3)
Zamora Chinchipe	120,416	1,762	1,463.26	B.1, B.1.308, B.1.223 (n=4)
Total	**17,510,643**	**241,567**	**1,379.54**	**20 lineages (N=119)**

336 *Data based on the last estimations for 2020 from INEC (Instituto Nacional de
337 Estadísticas y Censos), **Data from the official Ecuadorian portal
338 CoronavirusEcuador.com on January, 3 2021
339 https://www.coronavirusecuador.com/datos-provinciales/
340
341 **Table 2**. Amino Acid replacement in gene S found in SARS-CoV2 genomes from
342 Ecuador.

Aminoacid replacement in S protein	no. of genomes with replacement (%)
A288S	2 (1,68)
D614G	114 (95.80)
E1207V	1 (0.84)
L18F	1 (0.84)
N1187S	1 (0.84)

medRxiv preprint doi: https://doi.org/10.1101/2021.03.19.21253620; this version posted March 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

A570D	3 (2.52)
D1118H	3 (2.52)
G1167A	4 (3.36)
N501Y	3 (2.52)
Non-codon-aligned deletion	3 (2.52)
P681H	3 (2.52)
S982A	3 (2.52)
T553I	3 (2.52)
T716I	3 (2.52)

343

344 **Supplementary table 1**. Mutations found in four SARS CoV2 Genomes from Ecuador
345 compared to Wuhan-Hu-1 (GenBank accession number MN908947)

346 **Supplementary table 2**. Primer mismatches analyzed with seven COVID 19
347 diagnostic kits.

348 **Supplementary figure 1.** Phylogenetic analysis of SARS-COV2 genomes. The
349 phylogenetic tree was generated using the software IQ tree 2.1.1 and was visualized and
350 annotated using the iTOL tool. Ecuador sequences are highlighted in red.

351

Tree scale: 0.001