

1 **Genome sequencing and analysis of an emergent SARS-CoV-2 variant characterized by multiple spike**
2 **protein mutations detected from the Central Visayas Region of the Philippines**

3

4 Francis A. Tablizo¹, Kenneth M. Kim¹, Carlo M. Lapid¹, Marc Jerrone R. Castro¹, Maria Sofia L. Yangzon¹,
5 Benedict A. Maralit², Marc Edsel C. Ayes³, Eva Maria Cutiongco-de la Paz^{4,7}, Alethea R. De Guzman⁵, Jan
6 Michael C. Yap¹, Jo-Hannah S. Llames², Shiela Mae M. Araiza², Kris P. Punayan², Irish Coleen A. Asin²,
7 Candice Francheska B. Tambaoan², Asia Louisa U. Chong², Karol Sophia Agape R. Padilla², Rianna Patricia
8 S. Cruz¹, El King D. Morado¹, Joshua Gregor A. Dizon¹, Razel Nikka M. Hao⁶, Arianne A. Zamora⁵, Devon
9 Ray Pacial⁵, Juan Antonio R. Magalang⁵, Marissa Alejandria⁷, Celia Carlos⁷, Anna Ong-Lim⁷, Edsel Maurice
10 Salvaña⁷, John Q. Wong⁷, Jaime C. Montoya⁷, Maria Rosario Singh-Vergeire⁷ and Cynthia P. Saloma^{4,7*}

11

12 * To whom correspondence should be addressed. Email address: cpsaloma@up.edu.ph

13

14 ¹ Core Facility for Bioinformatics, Philippine Genome Center, University of the Philippines System

15 ² DNA Sequencing Core Facility, Philippine Genome Center, University of the Philippines System

16 ³ Clinical Genomics Laboratory, Philippine Genome Center, University of the Philippines System

17 ⁴ Philippine Genome Center, University of the Philippines System

18 ⁵ Epidemiology Bureau, Department of Health, Philippines

19 ⁶ Disease Prevention and Control Bureau, Department of Health, Philippines

20 ⁷ Inter-Agency Task Force on Emerging Infectious Diseases (IATF) Task Force on COVID-19 Variants,

21 Department of Health, Philippines

22 **ABSTRACT**

23 The emergence of SARS-CoV-2 variants of concern such as the B.1.1.7, B.1.351 and the P.1 have
24 prompted calls for governments worldwide to increase their genomic biosurveillance efforts. Globally,
25 quarantine and outbreak management measures have been implemented to stem the introduction of
26 these variants and to monitor any emerging variants of potential clinical significance domestically. Here,
27 we describe the emergence of a new SARS-CoV-2 lineage, mainly from the Central Visayas region of the
28 Philippines. This emergent variant is characterized by 13 lineage-defining mutations, including the co-
29 occurrence of the E484K, N501Y, and P681H mutations at the spike protein region, as well as three
30 additional radical amino acid replacements towards the C-terminal end of the said protein. A three-
31 amino acid deletion at positions 141 to 143 (LGV141_143del) in the spike protein was likewise seen in a
32 region preceding the 144Y deletion found in the B.1.1.7 variant. A single amino acid replacement, K2Q,
33 at the N-terminus of ORF8 was also shared by all 33 samples sequenced. The mutation profile of this
34 new virus variant warrants closer investigation due to its potential public health implications. The
35 current distribution of this emergent variant in the Philippines and its transmission are being monitored
36 and addressed by relevant public health agencies to stem its spread in nearby islands and regions in the
37 country.

38

39

40 **INTRODUCTION**

41 On March 11, 2020, the coronavirus disease 2019 (COVID-19), caused by the Severe Acute Respiratory
42 Syndrome Coronavirus 2 (SARS-CoV-2), was declared by the World Health Organization (WHO) as a
43 global pandemic, causing over 115 million infections and over 2.5 million deaths as of March 2, 2021.
44 The global uncontrolled spread of SARS-CoV-2 has given rise to various viral lineages circulating
45 worldwide, with at least three (3) variants being flagged as variants of concern.

46
47 Early in December 2020, the United Kingdom reported the emergence of the B.1.1.7 variant
48 (*20I/501Y.V1 Variant of Concern (VOC) 202012/0*), defined by 17 amino acid changes, including eight (8)
49 mutations in the spike protein (Rambaut et al., 2020). In addition to the B.1.1.7, two (2) more variants of
50 concern were reported: the B.1.351 lineage (South African Variant *20H/501Y.V2*) and the more recent
51 P.1 lineage (*20J/501Y.V3*) of Brazilian origin. All three lineages of concern share the spike N501Y
52 mutation, whereas the P.1 and B.1.351 lineages additionally share two (2) other mutations in the spike
53 protein (K417N/T and E484K) (COVID-19 Genomics UK Consortium [CoG-UK], 2021). Even though the set
54 of mutations/deletions found in the B.1.1.7, B.1.351, and P.1 lineages appear to have emerged
55 independently, these viruses have all been associated with a rapid increase in cases in their respective
56 places of origin (Faria et al., 2021; Rambaut et al., 2020a; Tegally et al., 2020). The WHO has encouraged
57 countries globally to increase routine systematic sequencing of SARS-CoV-2 viruses to better understand
58 SARS-CoV-2 transmission and to monitor possible emergence of new variants (World Health
59 Organization [WHO], 2021).

60
61 The Philippines is among the countries adversely hit by the COVID-19 pandemic, with over 580,000
62 confirmed cases and over 12,000 deaths attributed to the disease as of March 2, 2021. Although border
63 restrictions are in place to prevent the spread of the virus, local and international travel is still ongoing,
64 primarily for returning overseas Filipinos (ROFs) from various countries around the world. The
65 announcement of the UK variant last December 2020 prompted the government to conduct a
66 nationwide genomic surveillance program to detect SARS-CoV-2 variants, particularly among ROFs in
67 their ports of entry, as well as from local case clusters in different regions of the archipelago.

68
69 In this study, we report the detection of 33 Philippine SARS-CoV-2 cases that were found to share 13

70 lineage-defining mutations that may have potential biological significance, including the co-occurrence
71 of the E484K, N501Y, and P681H amino acid replacements at the spike protein region. The mutation
72 profile observed for this set of viruses has not been described in currently known SARS-CoV-2 lineages of
73 concern and were detected primarily from the Central Visayas Region of the country during a period
74 that coincides with a recorded surge in COVID-19 cases in the area.

75

76

77 RESULTS

78 The Philippine Department of Health, through its surveillance efforts, monitors and analyzes COVID-19
79 case data since the start of the pandemic. A decline in cases were seen after cases peaked in August of
80 2020. From January 21, the Department saw an increase in cases particularly in key areas in the country,
81 including Central Visayas. While an increase in cases was not unexpected given the increased mobility,
82 crowding, and gatherings during the holiday season, the cases in Central Visayas continued to rise as
83 other areas started to plateau around February (**Figure 1**). To further enrich the surveillance data
84 gathered from this area, samples from Central Visayas were prioritized for whole genome sequencing of
85 the virus. In total, we sequenced 60 samples from this region.

86

87 We report here the detection of 33 Philippine SARS-CoV-2 infections caused by viruses initially classified
88 to be under lineage B.1.1.28, but were found to have a co-occurrence of the spike protein E484K,
89 N501Y, and P681H mutations. The samples were collected from January 16 to February 02, 2021 and
90 most were among the sequenced samples from Central Visayas, with a few others from laboratory
91 submissions from the National Capital Region (NCR) and Davao Region. **Table 1** shows the breakdown of
92 sample counts per region of laboratory facility.

93

94

95

96 The three (3) co-occurring mutations have all been identified as significant SNPs in lineages of concern.

97 The E484K and N501Y co-occur in the P.1 lineage (Brazilian Variant) and the B.1.351 lineage (South

98 African Variant). In addition, N501Y and P681H are both defining single nucleotide polymorphisms

99 (SNPs) for the B.1.1.7 lineage (UK Variant) (CoG-UK, 2021). However, a profile of other high-frequency

100 mutations in the Philippine samples indicate that they do not share any of the other defining SNPs from

101 any of these lineages (**Figure 2**) (Faria et al., 2021; Rambaut et al., 2020a; Tegally et al., 2020). All the

102 samples also carry the globally prevalent spike D614G mutation.

103

104 Apart from the spike protein mutations previously associated with SARS-CoV-2 lineages of concern, ten

105 additional amino acid replacements were also observed in the 33 Philippine cases, including the

106 following: L3201P, D3681E, L3930F and P4715L in the ORF1ab polyprotein; E1092K, H1101Y and V1176F

107 in the spike protein; K2Q in the ORF8 protein product; and R203K and G204R in the nucleocapsid

108 protein. Three (3) other mutations were found in 32 of the 33 Philippine samples, which include the

109 D1554G and A5692V amino acid replacements at the ORF1ab polyprotein, as well as a three-amino acid

110 deletion at the spike protein at positions 141 to 143 (LGV141_143del). Five (5) synonymous mutations

111 were also seen: F924F, S2433S and N3928N in ORF1ab; and spike protein G593G and S875S. A list of the

112 commonly observed mutations for this viral group is shown in **Table 2** and the designated signature

113 mutations are also indicated. The significance of these mutations, particularly the combination of

114 mutations in the spike and other regions of the viral genome, on the transmissibility, pathogenicity, and

115 immunogenicity of this emergent variant remains to be studied.

116

117 Among the three known lineages of concern (*i.e.* B.1.1.7, P.1, B.1.351), the B.1.1.7 viruses have been the

118 most frequently detected so far in the Philippines, majority of which were collected from the main ports
119 of entry among ROFs and in a cluster of cases in the Cordillera Administrative Region (CAR), north of the
120 capital Manila. A handful of the B.1.351 viruses have also been detected, but none of the P.1 lineage as
121 of this writing. Infections from SARS-CoV-2 viruses belonging to the B.1.1.28 lineage appear to be more
122 widespread, with cases detected in six (6) different regions. However, the majority of the B.1.1.28 cases
123 in Central Visayas exhibited the signature mutations described for the emergent variant, designated
124 here as PHL-B.1.1.28 (**Figure 3A**).

125
126 Phylogenetic tree reconstruction of the depicted local SARS-CoV-2 cases, together with representative
127 sequences of the B.1.1.7, B.1.1.28, and P.1 lineages obtained from GISAID, shows that PHL-B.1.1.28
128 viruses form a distinct monophyletic clade that is separate from other lineages (**Figure 3B**).

129

130

131 **DISCUSSION**

132 In this study, we detected 33 Philippine SARS-CoV-2 infections, majority of which were collected from
133 laboratory facilities in Central Visayas, National Capital Region, and Davao Region (**Table 1**), caused by
134 viruses initially classified under lineage B.1.1.28, but with the co-occurrence of the E484K, N501Y, and
135 P681H spike protein mutations. These three (3) mutations have all been identified as defining mutations
136 in lineages that have attracted global concern with evidence of biological significance.

137

138 The E484K mutation, which is a defining SNP of the B.1.351 and P.1 lineages, has been identified as an
139 immune escape mutation associated with a loss of neutralizing activity by vaccine-elicited antibodies
140 (Collier et al., 2021). The N501Y mutation, which co-occurs with E484K in the B.1.351 and P.1 lineages
141 and is also found in B.1.1.7 viruses, is associated with higher transmissibility and increased binding

142 affinity to the ACE2 receptor (Leung et al., 2021; Luan et al., 2021). The P681H mutation on the other
143 hand that is found in the B.1.1.7 lineage, is immediately adjacent to the furin cleavage site at the S1/S2
144 boundary of the spike protein. The interaction of this site with furin and the ACE2 receptor is
145 implicated in viral transmission and systemic infection (Hasan et al., 2020; Wang et al, 2020). However,
146 none of the other defining SNPs for these lineages are found in the Philippine samples, suggesting that
147 these three (3) mutations may have arisen independently. All the samples also harbor the globally
148 prevalent D614G mutation, which has been previously associated with increased transmissibility and
149 higher viral loads (Hou et al., 2020; Lorenzo-Redondo et al., 2020).

150
151 The mutation profile (**Figure 2**) for this group of viruses further indicates that they have a characteristic
152 combination of mutations not observed in other known SARS-CoV-2 lineages of concern. Apart from
153 E484K, N501Y, D614G, and P681H, we found three (3) other amino acid replacements located towards
154 the C-terminal end of the spike protein in all of the 33 samples: E1092K, H1101Y, and V1176F (**Table 2**).
155 Although much is unknown about the actual effects of these mutations in terms of viral transmissibility,
156 pathogenicity, and its immunogenicity, they all represent radical amino acid replacements with possible
157 implications in protein structure and activity — E1092K being a change from a negatively charged
158 residue to a positively charged one; the H1101Y is a change from a positively charged residue to a
159 hydrophobic one, and the V1176F from a smaller aliphatic residue to a relatively larger aromatic one.
160 Interestingly, a single amino acid replacement (K2Q) at the N-terminal end of the ORF8 protein product
161 was also observed in all samples, which presents a possible significant change from a positively charged
162 residue to a polar uncharged one.

163
164 Moreover, we found a deletion of three (3) amino acids at the spike protein residues 141 to 143
165 (LGV141_143del) in 32 of the 33 samples included in this report. This deletion is adjacent to the tyrosine

166 residue deletion at the spike protein position 144 (Y144del), which is one of the hallmark mutations of
167 viruses in the B.1.1.7 lineage. Deletions in this region of the spike protein, including Δ 144/145 and Δ 141-
168 144, have been previously shown to alter the protein's antigenicity, with one particular monoclonal
169 antibody (4A8 MAb) against SARS-CoV-2 unable to detect mutants harboring the said deletions in a Vero
170 E6 (human kidney epithelial) cell line experiment (McCarthy et al., 2021).

171
172 Phylogenetic reconstruction of the 33 Philippine samples with representative samples from the B.1.1.7
173 and B.1.1.28 lineages, including the P.1 lineage which is derived from B.1.1.28, indicate that they
174 represent an emergent lineage within the global phylogenetic tree (**Figure 3B**). The phylogeny reveals
175 that these samples, referred to here as PHL-B.1.1.28, cluster into a well-defined monophyletic clade
176 (red) that diverges from the B.1.1.28 lineage (yellow). Notably, the tree topology shows that this clade
177 does not derive from the P.1 lineage, indicating that the E484K and N501Y mutations in samples may
178 represent recurrent mutations at these sites, rather than inheritance from the latter. The tree also
179 indicates a recent common ancestor for PHL-B.1.1.28, consistent with the emergence of a novel SARS-
180 CoV-2 variant.

181
182 The geographic distribution of the variant described here further supports its position to be separate
183 from the B.1.1.28 lineage. While other samples assigned to B.1.1.28 in the Philippines are relatively
184 abundant in other regions and rare in Central Visayas, the reverse is true for samples bearing the
185 signature mutations of PHL-B.1.1.28 (**Figure 3A**), suggesting that the emergence of this variant took
186 place within a geographically separate area of the Philippines centered in Cebu Island. This observation
187 is particularly notable because the Central Visayas region has been experiencing a sharp spike in COVID-
188 19 cases since early January 2021 according to statistics compiled by the Republic of the Philippines
189 Department of Health (**Figure 1**). One possible cause of this spike could be the increased mobility of

190 people post-holiday season or the emergence of a novel variant with higher transmissibility during this
191 period, consistent with the collection dates for the described samples in late January to early February.
192 Although the earliest collection (January 16, 2021) recorded for this group of viruses was from a sample
193 submitted by a testing laboratory situated at the NCR, we note that hardly any samples were collected
194 from Central Visayas for sequencing prior to January 31, highlighting the fact that much still remains
195 hidden about the evolutionary history of this emergent variant.

196

197

198 **CONCLUSIONS**

199 Presented in this report are several lines of evidence suggesting the emergence of a novel variant of
200 SARS-CoV-2 in the Philippines meriting further investigation, represented by 33 samples collected in
201 areas mainly from Central Visayas. First, all of these samples, collectively referred to here as PHL-
202 B.1.1.28, have three (3) spike protein mutations that are considered as defining SNPs for other lineages
203 of concern, and are associated with changes to viral behavior such as increased transmissibility or
204 antigenic escape. Second, they exhibit a characteristic mutation profile not found in other lineages,
205 comprising 13 signature non-synonymous mutations including others in the spike protein that are also
206 implicated in altered protein structure or activity. Third, the samples cluster into a distinct monophyletic
207 clade within the global tree that diverges from B.1.1.28, similar to the P.1 lineage. And fourth, these
208 samples were collected mainly at a time and location that was experiencing a sharp spike in COVID-19
209 cases, raising the possibility of a more transmissible SARS-CoV-2 variant as a potential cause.

210

211

212 **METHODOLOGY**

213 Nasopharyngeal swabs from SARS-CoV-2 infected individuals were collected and subjected to

214 automated RNA extraction using the PANAMAX™48 Automated Nucleic Acid Extraction Unit with
215 GenAmplify™ BeaDxMax™ Viral DNA/RNA Extraction Kit. Enrichment and sequencing of the extracted
216 SARS-CoV-2 genetic material were subsequently performed using the Illumina COVIDSeq workflow. The
217 resulting sequence reads were mapped to the reference SARS-CoV-2 genome sequence (NCBI Accession
218 No. NC_045512.2) using minimap2 v2.17 (Li, 2018) with preset parameters for accurate genomic short-
219 read mapping (-x sr). The output mapping file was further processed using Samtools v1.10 (Li et al.,
220 2009) to fill the mate coordinates and insert sizes field (fixmate), sort according to reference coordinates
221 (sort), and remove read duplicates (markdup). Mapping and coverage statistics were obtained using
222 Samtools (flagstat and coverage) as well.

223
224 The consensus sequence generation and intrahost variant calling were primarily done using the tool iVar
225 v1.3 (Grubaugh et al., 2019) with default parameters. Briefly, primer clipping and quality trimming was
226 implemented using the ivar trim function. Intrahost variants were then detected using a combination of
227 Samtools mpileup and ivar variants, with the parameters described in the iVar manual page
228 (<https://andersen-lab.github.io/ivar/html/manualpage.html>). Following the same manual, the detected
229 intrahost variants were used to identify and exclude sequence reads associated with mismatched primer
230 indices using the getmasked and removereads functions of iVar. The consensus sequence assemblies
231 were then generated using Samtools mpileup and ivar consensus commands with default parameters
232 (minimum quality = 20, minimum coverage = 10x).

233
234 From the consensus sequence assemblies, SARS-CoV-2 lineages were assigned using the tool Pangolin
235 v2.3.0 (<https://github.com/cov-lineages/pangolin>; Rambaut et al., 2020b). Variants relative to the
236 generated consensus sequences were also identified using MUMmer v4.0 (Kurtz et al., 2004) as
237 implemented in the annotation tool RATT (Otto et al., 2011). Summary mutation tables and plots were

238 generated using in-house scripts written in Python.

239

240 Phylogenetic trees were produced by initially generating a multiple sequence alignment of the local
241 SARS-CoV-2 consensus sequences using the tool MAFFT v7.407 (Katoch et al., 2002), supplemented with
242 publicly available sequences from the EpiCoV database of the Global Initiative for Sharing All Influenza
243 Data (GISAID) platform (Shu and McCauley, 2017). The resulting alignment was automatically trimmed
244 using trimAl (Capella-Gutierrez et al., 2009). Phylogenetic trees were then reconstructed using a local
245 instance of the Nextstrain analysis platform (Hadfield et al., 2018).

246

247

248 REFERENCES

249 Capella-Gutierrez, S., Silla-Martinez, J., & Gabaldon, T. (2009) trimAl: a tool for automated alignment
250 trimming in large-scale phylogenetic analyses. *Bioinformatics*. 25(15):1972-1973.

251 Collier, D., De Marco, A., Ferreira, I., Meng, B., Datir, R., Walls, A. C., ... & CITIID-NIHR BioResource
252 COVID-19 Collaboration. (2021). SARS-CoV-2 B.1.1.7 sensitivity to mRNA vaccine-elicited,
253 convalescent and monoclonal antibodies. medRxiv. doi:
254 <https://doi.org/10.1101/2021.01.19.21249840>

255 COVID-19 Genomics UK Consortium (CoG-UK). (2021) COG-UK report on SARS-CoV-2 Spike mutations of
256 interest in the UK 15th January 2021. Available from: [https://www.cogconsortium.uk/wp-](https://www.cogconsortium.uk/wp-content/uploads/2021/01/Report-2_COG-UK_SARS-CoV-2-Mutations.pdf)
257 [content/uploads/2021/01/Report-2_COG-UK_SARS-CoV-2-Mutations.pdf](https://www.cogconsortium.uk/wp-content/uploads/2021/01/Report-2_COG-UK_SARS-CoV-2-Mutations.pdf).

258 Faria, N. R., Claro, I. M., Candido, D., Moyses Franco, L. A., Andrade, P. S, Coletti, T. M., ... on behalf of
259 CADDE Genomic Network. (2021) Genomic characterisation of an emergent SARS-CoV-2 lineage
260 in Manaus: preliminary findings. *Virological*. Available from: [https://virological.org/t/genomic-](https://virological.org/t/genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-manaus-preliminary-findings/586)
261 [characterisation-of-an-emergent-sars-cov-2-lineage-in-manaus-preliminary-findings/586](https://virological.org/t/genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-manaus-preliminary-findings/586).

- 262 Grubaugh, N. D., Gangavarapu, K., Quick, J., Mtterson, N. L., de Jesus J. G., Main, B. J., ... & Andersen,
263 K.G. (2019). An amplicon-based sequencing framework for accurately measuring intrahost virus
264 diversity using PrimalSeq and iVar. *Genome Biology*. 20(8). [https://doi.org/10.1186/s13059-018-](https://doi.org/10.1186/s13059-018-1618-7)
265 1618-7.
- 266 Hadfield, J., Megill, C., Bell, S. M., Huddleston, J., Potter, B., Callender, C., Sagulenko, P., Bedford, T., &
267 Neher, R. A. (2018) Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics*.
268 34(23):4121-4123.,
- 269 Hasan, A., Paray, B. A., Hussain, A., Qadir, F. A., Attar, F., Aziz, F. M., ... & Falahati, M. (2020). A review on
270 the cleavage priming of the spike protein on coronavirus by angiotensin-converting enzyme-2 and
271 furin. *Journal of Biomolecular Structure and Dynamics*, 1-9.
- 272 Hou, Y. J., Chiba, S., Halfmann, P., Ehre, C., Kuroda, M., Dinnon, K. H., ... & Baric, R. S. (2020). SARS-CoV-2
273 D614G variant exhibits efficient replication ex vivo and transmission in vivo. *Science*, 370(6523),
274 1464-1468.
- 275 Katoh, K., Rozewicki, J., & Yamada, K. D. (2002) MAFFT: a novel method for rapid multiple sequence
276 alignment based on fast Fourier transform. *Nucleic Acids Research*. 30(14):3059-3066.
- 277 Kurtz, S., Phillippy, A., Delcher, A. D., Smoot, M., Shumway, M., Antonescu, C., & Salzberg, S. L. (2004)
278 Versatile and open software for comparing large genomes. *Genome Biology*. 5: R12.
- 279 Leung, K., Shum, M. H., Leung, G. M., Lam, T. T., & Wu, J. T. (2021). Early transmissibility assessment of
280 the N501Y mutant strains of SARS-CoV-2 in the United Kingdom, October to November 2020.
281 *Eurosurveillance*, 26(1), 2002106.
- 282 Li, H. (2018) Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 34(18): 3094-3100.
- 283 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R.
284 (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 25(16):2078-9.
- 285 Lorenzo-Redondo, R., Nam, H. H., Roberts, S. C., Simons, L. M., Jennings, L. J., Qi, C., ... & Ozer, E. A.

- 286 (2020). A clade of SARS-CoV-2 viruses associated with lower viral loads in patient upper airways.
287 *EBioMedicine*, 62, 103112.
- 288 Luan, B., Wang, H., & Huynh, T. (2021). Molecular Mechanism of the N501Y Mutation for Enhanced
289 Binding between SARS-CoV-2's Spike Protein and Human ACE2 Receptor. *bioRxiv*, 2021-01. doi:
290 <https://doi.org/10.1101/2021.01.04.425316>
- 291 McCarthy, K. R., Rennick, L. J., Nambuli, S., Robinson-McCarthy, L. R., Bain, W. G., Haidar, G., & Duprex,
292 W. P. (2021) Recurrent deletions in the SARS-CoV-2 spike glycoprotein drive antibody escape.
293 *Science*. eabf6950. DOI: 10.1126/science.abf6950.
- 294 Otto, T. D., Dillon, G. P., Degraeve, W. S., & Berriman, M. (2011) RATT: Rapid Annotation Transfer Tool.
295 *Nucleic Acids Research*. 39(9): e57.
- 296 Rambaut, A., Loman, N., Pybus, O., Barclay, W., Barrett, J., Carabelli, A., Connor, T., Robertson, D. L.,
297 Volz, E., on behalf of COVID-19 Genomics Consortium UK. (2020a) Preliminary genomic
298 characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike
299 mutations. *Virological*. Available from: [https://virological.org/t/preliminary-genomic-](https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563)
300 [characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-](https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563)
301 [mutations/563](https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563).
- 302 Rambaut, A., Homes, E.C., O'Toole, A., Hill, V., McCrone, J. T., Ruis, C., du Plessis, L. & Pybus, O. G.
303 (2020b) A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic
304 epidemiology. *Nature Microbiology*. 5: 1403-1407.
- 305 Tegally, H., Wilkinson, E., Giovanetti, M., Iranzadeh, A., Fonseca, V., Giandhari, J., ... & de Oliveira, T.
306 (2020) Emergence and rapid spread of a new severe acute respiratory syndrome-related
307 coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. *medRxiv*. doi:
308 <https://doi.org/10.1101/2020.12.21.20248640>.
- 309 Wang, Q., Qiu, Y., Li, J. Y., Zhou, Z. J., Liao, C. H., & Ge, X. Y. (2020). A unique protease cleavage site

310 predicted in the spike protein of the novel pneumonia coronavirus (2019-nCoV) potentially related
311 to viral transmissibility. *Virologica Sinica*, 1-3.

312 World Health Organization (WHO). (2021) Genomic sequencing of SARS-CoV-2: a guide to
313 implementation for maximum impact on public health. Available from:
314 <https://www.who.int/publications/i/item/9789240018440>.

315

316

317 **ACKNOWLEDGEMENT**

318 This study was funded by the Republic of the Philippines Department of Health, the Philippine Council
319 for Health Research and Development of the Department of Science and Technology, and the University
320 of the Philippines System. The authors would also like to thank the contributing institutions to the
321 Philippine Genomic Biosurveillance Network for their logistical support.

322

323

324 **COMPETING INTERESTS**

325 The authors declare no competing interests.

326

327

328 **ETHICS APPROVAL**

329 The methods used in this study has been approved by the Single Joint Research Ethics Board of the
330 Department of Health, Republic of the Philippines, with approval code SJREB-2021-11, as part of a larger
331 research program entitled “A retrospective study on the national genomic surveillance of COVID-19
332 transmission in the Philippines by SARS-CoV-2 genome sequencing and bioinformatics analysis”.

333

334

335 **DATA AVAILABILITY**

336 Genome sequences of the 33 Philippine SARS-CoV-2 samples reported here are deposited at the

337 EpiCoVTM database of the Global Initiative for Sharing All Influenza Data (GISAID) with accession codes

338 EPI_ISL_1122426 to EPI_ISL_ 1122458. The acknowledgement table for other GISAID sequences used in

339 this study can be found in **Supplementary Table S1**.

340 LIST OF TABLES

341 **Table 1.** Regions in the Philippines where the emergent SARS-CoV-2 variant was detected.

Region of Laboratory Facility	# Cases	Date of Collection	Remarks
Central Visayas	29	January 31 – February 02, 2021	Mostly from the Island of Cebu
National Capital Region	3	January 16 – February 01, 2021	Earliest recorded collection
Davao Region	1	January 23, 2021	

342

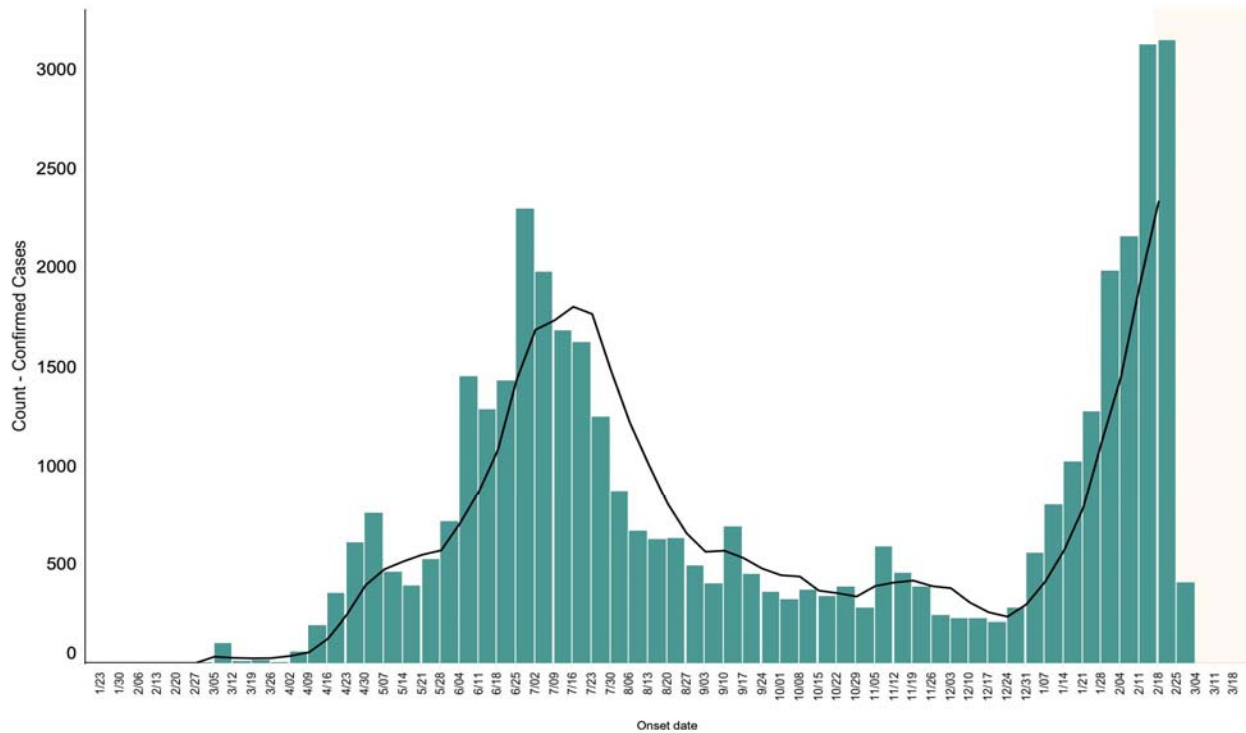
343

344 **Table 2.** Mutations observed in at least 32 of the 33 Philippine SARS-CoV-2 samples comprising an
 345 emergent virus variant. The column “Signature” signifies the mutations found in this group of viruses
 346 that are likely to have functional implications or are not prevalent in other lineages.

Affected Gene	Nucleotide Change	Amino Acid Change	# Cases	Signature
Non-synonymous Mutations				
ORF1ab	4962A>G	D1554G	32	Yes
ORF1ab	9867T>C	L3201P	33	Yes
ORF1ab	11308C>A	D3681E	33	Yes
ORF1ab	12053C>T	L3930F	33	Yes
ORF1ab	14408C>T	P4715L	33	
ORF1ab	17339C>T	A5692V	32	Yes
S	21981_21989del	LGV141_143del	32	Yes
S	23012G>A	E484K	33	Yes
S	23063A>T	N501Y	33	Yes
S	23403A>G	D614G	33	
S	23604C>A	P681H	33	Yes
S	24836G>A	E1092K	33	Yes
S	24863C>T	H1101Y	33	Yes
S	25088G>T	V1176F	33	Yes
ORF8	27897A>C	K2Q	33	Yes
N	28881_28882GG>AA	R203K	33	
N	28883G>C	G204R	33	
Synonymous Mutations				
ORF1ab	3037C>T	F924F	33	
ORF1ab	7564C>T	S2433S	33	
ORF1ab	12049C>T	N3928N	33	
S	23341T>C	G593G	33	
S	24187T>A	S875S	33	

348 LIST OF FIGURES

349



350

351 **Figure 1. Epidemic curve in Central Visayas, Philippines.** The figure above shows the distribution of
352 reported COVID-19 confirmed cases from the Central Visayas region as to their date of onset of illness as
353 of 02 March 2021. A steep rise in cases have been reported since the January 21, 2021.

354

355

356

357

358

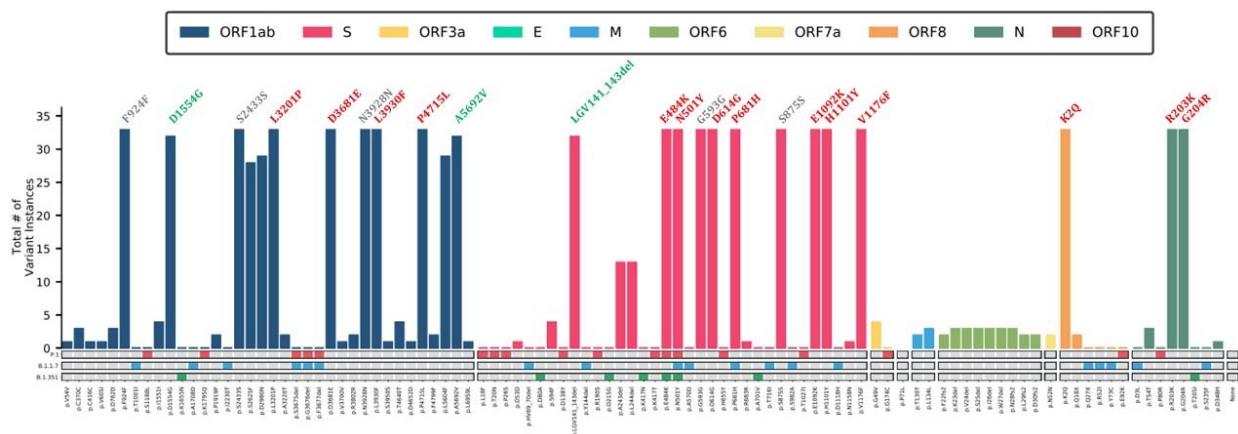
359

360

361

362

363



364

365 **Figure 2. Mutation profile of 33 Philippine SARS-CoV-2 samples comprising an emergent viral lineage.**

366 A total of 14 amino acid replacements were observed in all samples (labeled in red), including seven (7)

367 spike protein mutations. Among the spike protein mutations, four (4) have been previously associated

368 with lineages of concern (*i.e.*, E484K, N501Y, D614G, and P681H) while three (3) additional replacements

369 were observed towards the C-terminal region of the protein (*i.e.*, E1092K, H1101Y, and V1176F).

370 Interestingly, a single amino acid replacement at the N-terminus of ORF8 (*i.e.*, K2Q) was also found in all

371 samples. Three (3) other mutations were seen in 32 of the 33 samples (labeled in green), including a

372 three-amino acid deletion at the spike protein positions 141 to 143. Lastly, five (5) synonymous

373 mutations (labeled in gray) were also detected in all of the cases. For comparison, the defining

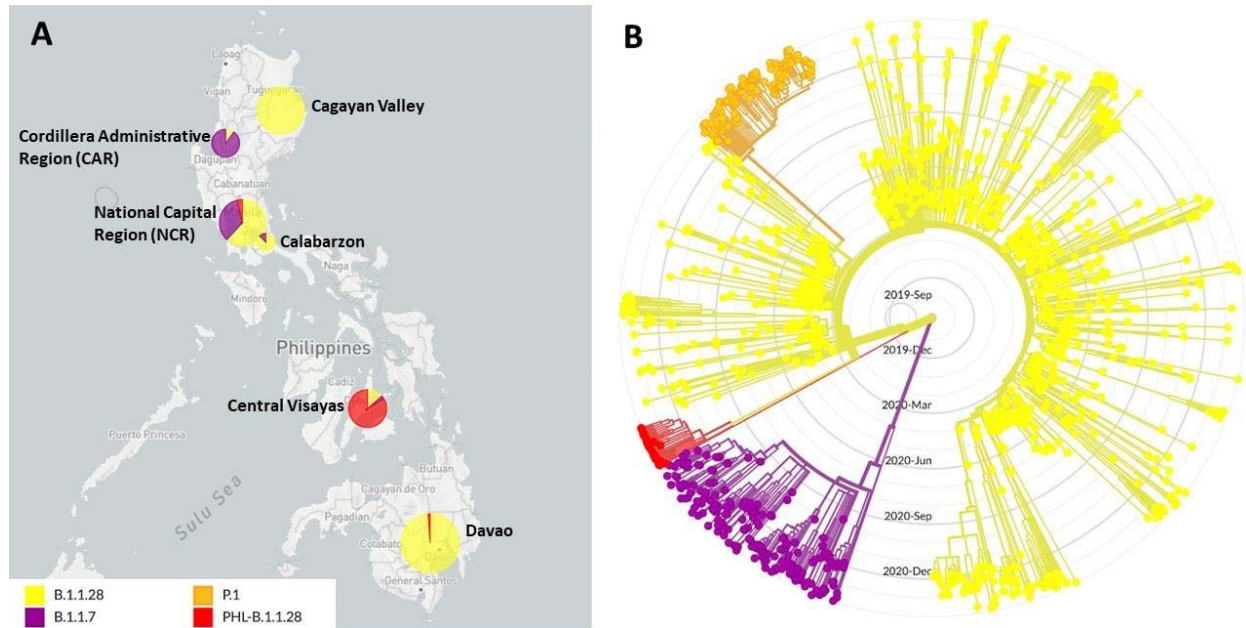
374 mutations for the P.1, B.1.1.7, and B.1.351 lineages of concern are marked below in red, blue, and green

375 boxes, respectively. None of the samples exhibit mutation profiles similar to any of the three (3) known

376 lineages of concern.

377

378



379

380 **Figure 3. Geographic distribution and phylogenetic clustering of an emergent SARS-CoV-2 lineage in**
381 **the Philippines designated as PHL-B.1.1.28. (A)** Geographic distribution of viruses under lineages
382 B.1.1.28 (Yellow), B.1.1.7 (Purple), P.1 (Orange) and the emergent variant PHL-B.1.1.28 (Red) in the
383 Philippines. Most of the samples detected to have the signature mutations of PHL-B.1.1.28 were from
384 the Central Visayas region. Note that viruses from lineage P.1 have not been detected in the country to
385 date. **(B)** Phylogenetic tree of B.1.1.28, B.1.1.7, P.1, and PHL-B.1.1.28 sequences showing that the
386 emergent variant forms a monophyletic clade separate from other lineages included in the analysis. The
387 B.1.1.7 and B.1.1.28 clades include samples collected both globally and in the Philippines.

388