

1 **Real-time RT-PCR Allelic Discrimination Assay for Detection of N501Y Mutation in the**
2 **Spike Protein of SARS-CoV-2 Associated with Variants of Concern.**

3 **Running Title: Assay for detection of S:N501Y in SARS-CoV-2**

4 Mariana Abdulnoor HBSc^{a,b,d,e}, AliReza Eshaghi PhD^{b,c}, Stephen J. Perusini MSc^b, Antoine
5 Corbeil MD^{a,b}, Kirby Cronin BSc^b, Nahuel Fittipaldi PhD^{a,b*}, Jessica D. Forbes PhD^b, Jennifer L.
6 Guthrie PhD^a, Julianne V. Kus PhD^{a,b}, Anna Majury PhD^a, Tony Mazzulli MD^{a,c}, Roberto G.
7 Melano PhD^{a,b}, Romy Olsha MD^a, Ashleigh Sullivan MPH^a, Vanessa Tran PhD^{a,b}, Samir N.
8 Patel PhD^{a,b}, Vanessa G. Allen MDCM^{a,b}, Jonathan B. Gubbay MBBS^{a,b,d#}

9 Affiliations:

- 10 a. Public Health Ontario, Toronto, Ontario Canada
- 11 b. Department of Laboratory Medicine and Pathobiology, University of Toronto,
12 Toronto, Ontario, Canada
- 13 c. Department of Microbiology, Sinai Health/University Health Network, Toronto,
14 Ontario, Canada
- 15 d. Division of Infectious Diseases, Department of Paediatrics, The Hospital for Sick
16 Children, Toronto, Ontario, Canada
- 17 e. Co-first authorship: Mariana Abdulnoor, AliReza Eshaghi. Mariana Abdulnoor wrote
18 up the manuscript, AliReza Eshaghi conducted experiments and wrote up the
19 manuscript. Order was determined alphabetically by last name

20 # Corresponding Author, Email address: jonathan.gubbay@oahpp.ca

21 *Nahuel Fittipaldi's new address is Université de Montréal, St-Hyacinthe St-Hyacinthe, Québec,
22 Canada

23

24

25 **Abstract:**

26 The N501Y amino acid mutation caused by a single point substitution A23063T in the spike
27 gene of SARS-CoV2 is possessed by the three most common variants of concern - B.1.1.7,
28 B.1.351, and P.1. A rapid screening tool using this mutation is important for surveillance during
29 the COVID-19 pandemic.

30 We developed and validated a single nucleotide polymorphism real-time reverse transcription
31 polymerase chain reaction assay using allelic discrimination of the spike gene N501Y mutation to
32 screen for potential variants of concern and differentiate them from wild-type SARS-CoV-2. A
33 total of 160 clinical specimens positive for SARS-CoV-2 were characterized as mutant (N501Y)
34 or wild-type by Sanger sequencing and were subsequently tested with the N501Y single
35 nucleotide polymorphism real time reverse transcriptase polymerase chain reaction assay. Our
36 assay compared to sequencing, the gold standard for SNP detection and lineage identification,
37 demonstrated clinical sensitivity of 100% for all 57 specimens displaying N501Y mutant, which
38 were confirmed by Sanger sequencing to be typed as A23063T, including one specimen with
39 mixed signal for wildtype and mutant. Clinical specificity was 100% in all 103 specimens typed
40 as wild-type, with A23063 identified as wild-type by Sanger sequencing. The identification of
41 circulating SARS-CoV-2 lineages carrying an N501Y mutation is critical for surveillance
42 purposes. Current identification methods rely primarily on Sanger sequencing or whole genome
43 sequencing which are time-consuming, labor-intensive and costly. The assay described herein is
44 an efficient tool for high-volume specimen screening for SARS-CoV-2 VOCs and for selecting
45 specimens for confirmatory Sanger or whole genome sequencing.

46

47

48 **Introduction:**

49 Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the viral etiology of the
50 coronavirus disease 2019 (COVID-19) pandemic. Several SARS-CoV-2 variants of concern
51 (VOC) have been identified which are associated with increased transmissibility, increased
52 virulence, changes in clinical disease presentation and decreased effectiveness of public health
53 measures, diagnostics, vaccines, and therapeutics (1). In 2020, three novel VOCs of SARS-
54 CoV-2 emerged independently (2-4).

55
56 The VOC first detected in the UK represents the B.1.1.7 lineage (also known as 20I/501Y.V1
57 and VOC202012/010) and contains three key mutations in S gene 69-70del, N501Y and P681H.
58 Recently, E484K has been observed in some B.1.1.7 genome sequences (5). B.1.1.7 is associated
59 with increased transmissibility of 50-70% (6), and increased disease severity and risk of death
60 (7). Cases of B.1.1.7 lineage have been found in 114 countries (8). In Ontario, Canada, 2,165
61 cases of B.1.1.7 have been confirmed as of April 5, 2021(9).

62
63 A second SARS-CoV-2 VOC representing the B.1.351 lineage (also known as 20H/501Y.V2
64 and VOC-202012/02) emerged in October 2020 in South Africa, and rapidly has become the
65 dominant circulating strain in that country (3). This VOC has multiple mutations in the receptor
66 binding domain (RBD) of the S protein, including K417N, E484K, and N501Y. The B.1.351
67 lineage is associated with increased transmissibility (10) and immune evasion (11,12). It is
68 unknown how this VOC influences disease severity, hospitalizations and deaths. B.1.351 has
69 been detected in 68 countries (13) including Canada, where in Ontario there are 71 confirmed
70 cases, as of April 5, 2021(9).

71
72 A third emerging VOC is the P.1 lineage (also known as B.1.1.28.1, 20J/501Y.V3, and VOC-
73 202101/02), first detected in Japan from travelers returning from Brazil (14). This VOC shares
74 the S gene N501Y single nucleotide polymorphism (SNP) with B.1.1.7 and the E484K SNP with
75 B.1.351; however, all three lineages arose independently (4). The P.1 lineage may have higher
76 inherent transmissibility compared to previous lineages (14). This lineage is associated with
77 antigenic escape (12). The P.1 lineage has been identified in 36 countries (15) including Canada,
78 where there are 106 confirmed cases in Ontario as of April 5, 2021 (9).

79
80 These SARS-CoV-2 VOCs will likely hamper public health efforts to contain the spread of the
81 virus. Thus, surveillance is important to control the spread of these lineages. Current methods
82 rely on genome sequencing, either by Sanger sequencing or whole genome sequencing (WGS).
83 These approaches, although of high accuracy, are labour-intensive and are not amenable to rapid
84 turnaround times as they require extensive laboratory processing and bioinformatic analysis.
85 They usually take several days to complete, therefore testing capacity is quickly outpaced as the
86 number of specimens to be analyzed increases with transmission.

87
88 The use of real-time reverse transcription PCR (rRT-PCR) to detect SNPs of relevance would
89 facilitate rapid preliminary identification of potential VOCs and initiation of public health
90 measures. Herein, we report a laboratory developed SNP rRT-PCR protocol to detect the
91 presence of N501Y, which can act as a screening tool for the three circulating VOCs in Ontario
92 as they all share this SNP. Specimens with N501Y detected are subsequently submitted for WGS
93 to confirm the presence of a VOC and identify the SARS-CoV-2 VOC lineage.

94

95 **Methods:**

96 **Clinical Specimens:**

97 Public Health Ontario (PHO) Laboratory is the reference microbiology laboratory for the
98 province of Ontario, Canada. As part of the response to detect VOCs in Ontario, PHO
99 Laboratory developed indications for VOC screening, which were similar to Canadian guidelines
100 (16). Clinical specimens that met these criteria with a cycle threshold (Ct) value ≤ 30 underwent
101 Sanger sequencing of a 698 bp (nucleotide positions 22516 to 23214) S gene fragment, which
102 includes the locations of key RBD mutations (K417N/T, E484K, N501Y). Specimens with
103 mutations detected underwent further sequencing to confirm if a VOC was present. In addition, a
104 subset of all SARS-CoV-2 positive specimens underwent WGS as part of the provincial genomic
105 surveillance program, which included VOC screening.

106

107 A convenience sample of 160 SARS-CoV-2 positive clinical specimens collected between
108 November 20, 2020 and January 20, 2021, submitted to PHO Laboratory from across the
109 province, that had undergone VOC screening by partial S gene sequencing, as described above,
110 were used in the optimization and validation of the assay. No clinical information was available
111 when sequencing was conducted. However, the presence of N501Y was known prior to running
112 specimens in our assay as it was previously sequenced. All SARS-CoV-2 positive specimens
113 with N501Y available at PHO Laboratory were included. Specimen types included
114 nasopharyngeal swabs, throat swabs, and bronchoalveolar lavage fluid.

115

116 **rRT-PCR Design and Optimization:**

117 Primer and probe sequences were selected using an alignment of spike nucleotide sequences of a
118 wild-type SARS-CoV2 reference genome (GenBank accession: MN908947.3) and a
119 representative of the VOC B.1.1.7 lineage (GSAID accession: EPI_ISL_601443). Selected
120 primer and probe sequences were aligned to 9000 gene sequences downloaded from GISAID
121 with submission date between December 1, 2020 through December 6, 2020 to ensure they were
122 targeting a conserved region. In addition, a BLAST search of the primer and probe sequences
123 against the whole genome of the reference sequence (MN908947.3) was conducted to ensure the
124 absence of homology to other parts of the SARS-CoV-2 genome. Primers and probes selected for
125 the detection of SNP A23063T are reported in **Supplementary Table 1**. Primer and probes were
126 synthesized by LGC, Biosearch Technologies (Middlesex, United Kingdom). Optimization was
127 performed using previously characterized wild-type and B.1.1.7 specimens. No B.1.351 or P.1
128 lineage SARS-CoV-2 were detected in Ontario at the time of assay development.

129
130 An additional smaller validation was conducted using primers and probes from a second vendor,
131 Thermo Fisher Scientific (Massachusetts, United States), to allow redundancy of reagent
132 suppliers.

133

134 **Total Nucleic Acid Extraction:**

135 Specimens underwent nucleic acid extraction using either Perkin Elmer chemagic™ 360
136 automated systems (PerkinElmer, Massachusetts, United States) (input volume 300 µl; eluate
137 volume 60 µl) or MGISP-960RS platform (MGI Tech Co., Guangdong, China) (input volume
138 180 µl; eluate volume 45 µl) according to the manufacturer's recommendations.

139

140 **N501Y SNP rRT-PCR Assay:**

141 The PCR amplification was performed using the Taqpath™ 1-Step Multiplex Master Mix on the
142 Applied Biosystems™ QuantStudio™ 5 Real-Time PCR system (Thermo Fisher Scientific,
143 Massachusetts, United States).

144

145 A total volume of 10µl reaction mix contained 2.5µl of 4X TaqPath™ 1-Step Multiplex Master
146 Mix, 1.0µl of primer mix, 1.0µl of probe mix, 2.5µl of RNase-free water, and 3µl of viral RNA
147 extract. The default thermocycling profile suggested for 4X TaqPath™ 1-Step Multiplex Master
148 Mix was used including 20 minutes at 25° C, 10 minutes at 53°C, 2 minutes at 95°C and 45
149 cycles of 95°C for 3 seconds then 60°C for 30 seconds. The controls used in each run were an
150 extraction negative control that consisted of a nuclease-free water processed along with SARS-
151 CoV-2 specimens during the extraction process, a PCR negative control that consisted of
152 nuclease-free water in place of nucleic acid in the reaction mix, and a positive RNA
153 amplification control consisting of combined wild-type and B.1.1.7 SARS-CoV-2 RNA obtained
154 from clinical specimens positive by in-house end-point PCR and sequencing. During
155 optimization, it was determined that nonspecific signals occurred above a Ct of 37, therefore
156 only results with Ct ≤37 were considered positive for either probe target. The assay does not
157 have an indeterminate range. The time interval between sequencing and running the N501Y rRT-
158 PCR assay ranged from one to 14 days.

159

160 **PCR Amplification and Sequencing of Partial Spike Gene:**

161 Sanger sequencing was chosen as the comparator method as it is a highly accurate methodology
162 to determine the presence of SNPs, is the gold standard for identifying SNPs, and was the

163 method in use at PHO Laboratory for VOC detection. A detailed explanation of the Sanger
164 sequencing protocol used is provided in **Supplementary appendix A**. Specimens with Ct values
165 ≤ 35 with a good quality Sanger chromatogram, containing evenly-spaced peaks each with one
166 colour, and the least amount of baseline noise, were considered positive. Sanger sequencing
167 failed when Ct values were >35 and the chromatogram contained mixed bases and was messy.

168

169 **Validation:**

170 Primary outcome measures were clinical sensitivity, clinical specificity, and precision.

171

172 **Accuracy – Clinical Sensitivity and Specificity**

173 A panel of 160 SARS-CoV-2 positive clinical specimens that had been characterized by partial
174 Sanger sequencing of the S gene were used for validation of the N501Y SNP assay. The result
175 obtained by our rRT-PCR N501Y SNP assay was compared to the known sequence results
176 previously obtained by partial S gene sequencing and used to calculate sensitivity and specificity.
177 Sensitivity of $>90\%$ and specificity $>95\%$ were considered acceptable.

178

179 **Intra-run repeatability**

180 Intra-run repeatability was evaluated by running three replicates of 10 N501Y VOCs (including
181 1 with a mixed population) and 5 wild-type samples in the same rRT-PCR run and comparing Ct
182 values across the replicates. The threshold for acceptance between runs was a Ct difference of
183 <3.3 .

184

185 **Inter-run reproducibility**

186 Inter-run reproducibility was evaluated by rerunning the panel used for intra-run repeatability on
187 a second day. This panel was also sent out to six other microbiology laboratories in Ontario to
188 validate the assay prior to implementation. The threshold for acceptance was >95% agreement
189 between labs.

190

191 **Ethical Statement:**

192 The PHO Ethics Review Board has determined that this project did not require research ethics
193 committee approval, as it describes analyses that were completed at PHO Laboratory as part of
194 routine clinical respiratory testing and surveillance during the COVID-19 pandemic in Ontario
195 and are, therefore, considered public health practice and exempt from this requirement.

196

197 **Results:**

198 **Accuracy - Clinical Sensitivity and Specificity:**

199 To analyze the performance of the N501Y SNP assay, a panel of 160 specimens including 57
200 specimens with N501Y detected by Sanger sequencing, and 103 wild-type were tested. One of
201 the 57 N501Y positive specimens contained a mixed result on Sanger sequencing, indicating the
202 presence of both wild-type and mutant strains. All specimens with a known A23063T
203 substitution in the Sanger chromatogram produced a positive signal for N501Y with the rRT-
204 PCR SNP assay (57/57, 100%), and were confirmed to be B.1.1.7 lineage by Sanger sequencing
205 analysis. All samples with known A23063 wild-type characterization by Sanger showed positive
206 signal for wild-type with the rRT-PCR SNP assay (103/103, 100%). The positive predictive
207 value was 100%% (57/57) and the negative predictive value was 100% (103/103) (**Table 1**).

208

209 The mean E gene and N501Y SNP rRT-PCR assay Ct values of all VOC specimens in this study
210 were 19.15 (S.D. 3.76) and 21.88 (S.D. 3.80), respectively. The mean Ct values of E gene and
211 N501 SNP rRT-PCR for wild-type specimens were 20.87 (S.D. 5.38) and 23.75 (S.D. 5.52),
212 respectively. (**Supplementary Table 2**).

213

214 **Comparison of two primer and probe sets:**

215 Our N501Y SNP assay was validated using primer and probe sets from two vendors, LGC,
216 Biosearch Technologies (Middlesex, United Kingdom) and Thermo Fisher Scientific
217 (Massachusetts, United States). A subset of 87 specimens, including 14 mutants, 72 wild-type
218 and one mixed population were tested (**Table 1**). We observed comparable results for both
219 primer and probe sets, as there were no significant differences in average Ct values or number of
220 N501Y specimens detected.

221

222 **Precision – intra-run repeatability and inter-run reproducibility:**

223 The analysis of intra-run repeatability was done with a panel of nine N501Y samples, five wild-
224 type, one mixed N501Y/N501 sample, and five SARS-CoV-2 negative samples. The panel was
225 run on three replicates. The results show 100% agreement (**Supplementary Table 3**) with an
226 average Ct 26.17 (average SD 0.19, average coefficient of variation 0.71%) for the mutant
227 specimens and average Ct 26.76 (average SD 0.16, average coefficient of variation 0.6%) for
228 wild-type specimens.

229

230 The analysis of inter-run reproducibility was done with replicates of a panel of nine N501Y
231 samples, five wild-type samples, one mixed N501Y/N501 sample, and five negative samples.

232 Among the 10 runs obtained from seven laboratories, the results showed 100% agreement with
233 an average Ct 26.04 (SD 0.98, coefficient of variation 4%) for mutant specimens and an average
234 Ct 27.39 (SD 1.29, coefficient of variation 5%) for wild-type specimens.

235

236 **Discussion:**

237 Here we describe the validation of a N501Y SNP rRT-PCR assay to provide rapid screening of
238 SARS-CoV-2 positive specimens for VOCs. Currently, three VOCs (B.1.1.7, B.1.351 and P.1)
239 are circulating in Ontario and surveillance of these VOCs is of considerable public health
240 importance. The World Health Organization recommends VOC surveillance to help control the
241 spread of VOCs (1). In our workflow, positive SARS-COV-2 specimens are screened for the
242 N501Y SNP using the described rRT-PCR assay then potentially submitted for WGS to confirm
243 VOC identification and determine the VOC lineage. There were no adverse events from
244 performing the N501Y SNP rRT-PCR assay as individuals had already been classified as having
245 the mutation or not based on previous sequencing. Our assay is efficient, and clinically validated
246 with a high degree of sensitivity and specificity for detection of the N501Y SNP.

247

248 Sanger sequencing or WGS of VOCs require longer turnaround times, are expensive, and some
249 regions may not have access to sequencing facilities or the ability to scale-up their sequencing
250 capabilities. A two-factor screening process for B.1.1.7 detection has been employed to aid with
251 surveillance efforts (17). The protocol detects the 69-70del in the S gene of B.1.1.7 as this
252 deletion causes S-gene target failure (SGTF) in certain assays (e.g. Thermo Fisher's TaqPath
253 COVID-19 Combo Kit). Once SGTF is detected, specimens are sequenced by WGS. SGTF has
254 been used as a proxy for the B.1.1.7 VOC (6). This method is limited to detecting VOCs carrying
255 69-70del. A multiplex rRT-PCR to screen for VOC was published by Vogel et al. (18) that

256 targets the aa3675-3677 deletion in ORF1a gene as well as the aa69-70del in the S gene. The
257 ORF1a deletion of aa3675-3677 is present in B.1.17, B.1.351, and P.1 and the aa69-70del in S
258 gene is used to differentiate B.1.17 from B.1.351 and P.1. However, the ORF1a deletion is not
259 found in all B.1.351 VOCs, as there is a monophyletic clade that does not contain this deletion.
260 Another group (19) published a rRT-PCR assay to detect only the N501Y SNP and requires a
261 melting curve to confirm the specimens as wild-type if the test is negative. Our procedure differs
262 as it is able to detect both the wild-type and the SNP, avoiding the additional step of analyzing
263 melting-curves and thus providing easier interpretation of the results.

264
265 To understand the provincial prevalence of VOCs in Ontario, the majority of SARS-CoV-2
266 positive specimens reported on January 20, 2021 across the province were sent for screening
267 using our rRT-PCR N501Y SNP assay (20). Our rRT-PCR N501Y SNP assay has now been
268 instituted for surveillance testing across Ontario, allowing notification to public health at least
269 one to seven days earlier than relying on Sanger sequencing or WGS. On March 11, 2021, 1,138
270 specimens were screened with our N501Y rRT-PCR assay, 897/899 (99.8%) specimens that
271 tested positive for N501Y SNP were confirmed through WGS and 238/239 (99.6%) were
272 confirmed wildtype through WGS (data extracted from PHO laboratory information system).
273 Using this assay, we've now documented a rapid rise in VOC cases with roughly 60% of cases
274 on April 5, 2021 identified as VOC or containing N501Y and/or E484K SNPs (9).

275
276 Because our assay only detects the N501Y SNP, new emerging lineages that do not carry this
277 SNP, such as B.1.525, will be missed (21). The assay is also unable to differentiate between
278 VOCs that share the N501Y SNP; therefore, an additional sequencing step is required to

279 delineate the lineage. An additional SNP assay target detecting E484K was incorporated into this
280 assay in a multiplex format as of March 22, 2021, and assays to detect E417T/N S gene
281 mutations are currently under development in our laboratory as these could provide further
282 characterization of currently known VOCs without requirement of WGS. Even with additional
283 targets, definitive lineage identification is not possible based on SNP assays alone and requires
284 sequencing. Our validation was conducted with specimens characterized as B.1.1.7 but when
285 implemented as a screening tool we were able to detect B.1.351 and P.1 VOCs. Our validation
286 dataset contained a small sample size of 57 positive specimens; however, clinical sensitivity and
287 specificity were further confirmed in our data review post-implementation. A limit of detection
288 was not performed, however, comparison of rRT-PCR SNP Ct values to E gene Ct values using
289 the Corman et al (22) assay, shows the SNP assay is approximately 1 log less sensitive (3-4
290 cycles higher). We did not conduct analytical specificity tests with a cross reactivity panel due to
291 limited availability of SARS-CoV-2 VOC strains. In our analysis we did not assess baseline
292 demographics, clinical characteristics of the patients and severity of disease, as limited
293 information is provided to PHO Laboratory, which provides testing as a reference laboratory.
294 Our assay is not intended for diagnostic purposes, but rather as a screening tool to aid with faster
295 identification of VOCs through differentiation of N501 and N501Y.

296
297 Our assay is a quick and simple tool that can be applied to screening SARS-CoV-2 positive
298 specimens from priority groups at greater risk for VOC infection such as international travellers,
299 suspected SARS-CoV-2 reinfection, infection post COVID-19 vaccination, and outbreaks.
300 Ontario is screening all positive specimens, which can be in excess of 1000 specimens per day,
301 and the N501Y SNP rRT-PCR assay has allowed us to screen a high volume of specimens. Due

302 to high specificity and positive predictive value, any specimen that is N501Y positive can be
303 presumed to be a VOC while awaiting characterization by WGS. Early, high-throughput
304 screening with PHO Laboratory's N501Y screening assay allows effective public health
305 measures to be put in place to limit the spread of VOCs in the community.

306

307 Conflict of Interest Disclosure: All authors declare no competing interests.

308

309 Financial Disclosure: This study was funded by PHO Laboratory.

310

311 Acknowledgement: We acknowledge the collaboration of members of the Ontario Provincial
312 COVID-19 Diagnostic Network. We acknowledge the staff of Virus Detection and Molecular
313 Diagnostics, Public Health Ontario Laboratory, for diagnostic testing and N501Y screening of
314 SARS-CoV-2 specimens. We thank the Public Health Ontario Laboratory research group for
315 conducting SARS-CoV-2 whole genome sequencing.

316

317

318

319

320

321

322

323

324

325 **References:**

- 326 1. COVID-19 Weekly epidemiological update. World Health Organization. Published
327 February 25, 2021. [https://www.who.int/publications/m/item/covid-19-weekly-](https://www.who.int/publications/m/item/covid-19-weekly-epidemiological-update)
328 [epidemiological-update](https://www.who.int/publications/m/item/covid-19-weekly-epidemiological-update)
- 329 2. Rambaut A, Loman N, Pybus O, et al. Preliminary genomic characterisation of an
330 emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations.
331 Virological. Published December 2020. [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)
332 [characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-](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)
333 [of-spike-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).
- 334 3. Tegally H, Wilkinson E, Giovanetti M, et al. Emergence and rapid spread of a new
335 severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with
336 multiple spike mutations in South Africa. *medRxiv*. 2020.
337 <http://doi.org/10.1101/2020.12.21.20248640>.
- 338 4. Faria NR, Claro IM, Candido D, et al. Genomic characterisation of an emergent
339 SARS-CoV-2 lineage in Manaus: preliminary findings. Virological. Published Jan 12,
340 2021. Accessed February 6, 2021. [https://virological.org/t/genomic-characterisation-](https://virological.org/t/genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-manaus-preliminary-findings/586)
341 [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).
- 342 5. Volz E, Mishra S, Chand M, et al. Transmission of SARS-CoV-2 Lineage B.1.1.7 in
343 England: Insights from linking epidemiological and genetic data. *medRxiv*. 2021.
344 <https://doi.org/10.1101/2020.12.30.20249034>.
- 345 6. Davies NG, Barnard RC, Jarvis CI, Kucharski AJ, Munday J, Pearson CAB, et al.
346 Estimated transmissibility and severity of novel SARS-CoV-2 variant of concern

- 347 202012/01 in England. *medRxiv*. 2020.
- 348 <https://www.medrxiv.org/content/10.1101/2020.12.24.20248822v1>
- 349 7. Horby P, Huntley C, Davies N, Edmunds J, Ferguson N, Medley G, Semple C.
- 350 NERVTAG note on B.1.1.7 severity. GOV.UK. Published January 21, 2021.
- 351 [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachme](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/961037/NERVTAG_note_on_B.1.1.7_severity_for_SAGE_77_1_.pdf)
- 352 [nt_data/file/961037/NERVTAG_note_on_B.1.1.7_severity_for_SAGE_77_1_.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/961037/NERVTAG_note_on_B.1.1.7_severity_for_SAGE_77_1_.pdf)
- 353 8. B.1.1.7. SARS-CoV-2 Lineages. Published February 20, 2021. [https://cov-](https://cov-lineages.org/global_report_B.1.1.7.html)
- 354 [lineages.org/global_report_B.1.1.7.html](https://cov-lineages.org/global_report_B.1.1.7.html)
- 355 9. Daily epidemiologic summary: COVID-19 in Ontario: January 15, 2020 to April 5,
- 356 2021. Ontario Agency for Health Promotion and Protection (Public Health Ontario).
- 357 [https://www.publichealthontario.ca/-/media/documents/ncov/epi/covid-19-daily-epi-](https://www.publichealthontario.ca/-/media/documents/ncov/epi/covid-19-daily-epi-summary-report.pdf?la=en)
- 358 [summary-report.pdf?la=en](https://www.publichealthontario.ca/-/media/documents/ncov/epi/covid-19-daily-epi-summary-report.pdf?la=en)
- 359 10. Risk Assessment: Risk related to the spread of new SARS-CoV-2 variants of concern
- 360 in the EU/EEA-first update. European Centre for Disease Prevention and Control.
- 361 Published January 21, 2021. [https://www.ecdc.europa.eu/en/publications-data/covid-](https://www.ecdc.europa.eu/en/publications-data/covid-19-risk-assessment-spread-new-variants-concern-eueea-first-update)
- 362 [19-risk-assessment-spread-new-variants-concern-eueea-first-update](https://www.ecdc.europa.eu/en/publications-data/covid-19-risk-assessment-spread-new-variants-concern-eueea-first-update)
- 363 11. Wibmer CK, Ayres F, Hermanus T, Madzivhandila M, Kgagudi P, Oosthuysen B, et
- 364 al. SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor
- 365 plasma. *Nat Med*. 2021; <https://doi.org/10.1038/s41591-021-01285-x>
- 366 12. Comprehensive mapping of mutations to the SARS-CoV-2 receptor-binding domain
- 367 that affect recognition by polyclonal human serum antibodies. *Cell Host & Microbe*.
- 368 2021; 29(3):463-476.

- 369 13. B.1.351. SARS-CoV-2 Lineages. Published February 20, 2020. <https://cov->
370 [lineages.org/global_report_B.1.351.html](https://cov-lineages.org/global_report_B.1.351.html)
- 371 14. Sabino EC, Buss LF, Carvalho MPS, Prete Jr CA, Crispim MAE, Fraiji NA, et al.
372 Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. *The*
373 *Lancet*. 2021; 6736(21)00183-00185. [https://doi.org/10.1016/ S0140-6736\(21\)00183-](https://doi.org/10.1016/S0140-6736(21)00183-5)
374 5.
- 375 15. P.1. SARS-CoV-2 Lineages. Published February 20, 2021. <https://cov->
376 [lineages.org/global_report_P.1.html](https://cov-lineages.org/global_report_P.1.html)
- 377 16. SARS-CoV-2 (COVID-19 Virus) Variant of Concern (VoC) Surveillance. Ontario
378 Agency for Health Protection and Promotion (Public Health Ontario).
379 [https://www.publichealthontario.ca/en/laboratory-services/test-information-](https://www.publichealthontario.ca/en/laboratory-services/test-information-index/covid-19-voc)
380 [index/covid-19-voc](https://www.publichealthontario.ca/en/laboratory-services/test-information-index/covid-19-voc)
- 381 17. Bal A, Destras G, Gaymard A, Stefic K, Marlet J, Eymieux S, et al. Two-step
382 strategy for the identification of SARS-CoV-2 variant of concern 202012/01 and
383 other variants with spike deletion H69–V70, France, August to December 2020.
384 *Eurosurveillance*. 2021;26(3):pii=2100008. [https://doi.org/10.2807/1560-](https://doi.org/10.2807/1560-7917.ES.2021.26.3.2100008)
385 [7917.ES.2021.26.3.2100008](https://doi.org/10.2807/1560-7917.ES.2021.26.3.2100008)
- 386 18. Vogels CBF, Breban M, Alpert T, Petrone ME, Watkins AE, Hodcroft EB, et al. PCR
387 assay to enhance global surveillance for SARS-CoV-2 variants of concern. *medRxiv*,
388 2021. <https://doi.org/10.1101/2021.01.28.21250486>.
- 389 19. Durner J, Burggraf S, Czibere L, Tehrani A, Watts DC, Becker M. Fast and cost-
390 effective screening for SARS-CoV-2 variants in a routine diagnostic setting. *Dental*
391 *Materials*. 2021. <https://doi.org/10.1016/j.dental.2021.01.015>.

392 20. Summary report: SARS-CoV-2 variants of concern: results of point-prevalence study.
393 Ontario Agency for Health Protection and Promotion (Public Health Ontario).
394 [https://www.publichealthontario.ca/-/media/documents/ncov/voc/2021/02/sars-cov-2-](https://www.publichealthontario.ca/-/media/documents/ncov/voc/2021/02/sars-cov-2-variants-point-prevalence.pdf?la=en)
395 [variants-point-prevalence.pdf?la=en](https://www.publichealthontario.ca/-/media/documents/ncov/voc/2021/02/sars-cov-2-variants-point-prevalence.pdf?la=en)

396 21. B.1.525. SARS-CoV-2 Lineages. Published February 20, 2021. [https://cov-](https://cov-lineages.org/global_report_B.1.525.html)
397 [lineages.org/global_report_B.1.525.html](https://cov-lineages.org/global_report_B.1.525.html)

398 22. Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DKW, et al.
399 Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR.
400 *Eurosurveillance*. 2020;25(3):pii=2000045. [https://doi.org/10.2807/1560-](https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045)
401 [7917.ES.2020.25.3.2000045](https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045)

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427 **Tables:**

428

429 Table 1: Summary of accuracy testing data for PHO Laboratory SARS- CoV-2 N501Y SNP

430 rRT-PCR assay

	Sanger Mutant	Sanger Wild-type	Total
SNP (N501Y) detected	57	0	57
SNP (N501) not detected	0	103	103
Total	57	103	160

431 There was one specimen with a mixed population of wild-type and mutant, this specimen was

432 counted as positive for N501Y.

433

434 Table 2: Results of validation experiments with primer and probe sets from two vendors, LGC

435 Biosearch Technologies and Thermo Fisher Scientific.

	LGC Biosearch Technologies		Thermo Fisher Scientific	
	Wild-type (n=72)	Mutant (n=14)	Wild-type (n=72)	Mutant (n=14)
Average Ct Value (S.D.)	27.44 (6.86)	22.11 (6.92)	27.18 (6.80)	22.75 (7.14)

Median Ct Value	27.43	21.41	27.07	21.82
------------------------	-------	-------	-------	-------

436

437 Supplementary Table 1: Primers and probes used in the PHO Laboratory SARS-CoV-2 N501Y

438 SNP rRT-PCR assay and Sanger sequencing

Name	Sequence (5'-3')	Position	Final Concentration (nM)
SNP assay primers and probes			
501F	GAAGGTTTTAATTGTTACTTT C	23012- 23033	1000
501R	AAACAGTTGCTGGTGCATGT	23116- 23136	1000
FAM-Y501 (T23063)	CCAACCCACTTATGGTGTG	23053- 23072	250
HEX-N501 (A23063)	CCAACCCACTAATGGTGTG		500

S gene fragment RT-PCR and sequencing primers			
VOCF	AGAGTCCAACCAACAGAATC TATTGT	22517- 22542	400
VOCR	ACACCTGTGCCTGTAAACCA T	23192- 23214	400

439

440 Supplementary Table 2: Individual specimen cycle threshold (Ct) values for E gene (22) and

441 PHO Laboratory rRT-PCR N501Y assay.

Specimen ID	E gene Ct	SNP assay		SNP results	Sanger Sequencing			Result
		N501	N501 Y		K417N	E484K	N501 Y	
					G2281 3T	G2301 2A	A2306 3T	
1	18.81	22	ND ^a	Neg	G	G	A	Wild-type
2	22.93	27	ND	Neg	G	G	A	Wild-type
3	23.33	26.9	ND	Neg	G	G	A	Wild-type
4	15.08	17.8	ND	Neg	G	G	A	Wild-type
5	18.16	21.6	ND	Neg	G	G	A	Wild-type
6	19.37	22.3	ND	Neg	G	G	A	Wild-type
7	13.04	16.7	ND	Neg	G	G	A	Wild-type
8	15.03	19	ND	Neg	G	G	A	Wild-type

9	24.14	27.7	ND	Neg	G	G	A	Wild-type
10	29.78	33.4	ND	Neg	G	G	A	Wild-type
11	19.22	22.9	ND	Neg	G	G	A	Wild-type
12	28.95	32.1	ND	Neg	G	G	A	Wild-type
13	20.24	ND	24.9	Pos	G	G	T	N501Y
14	15.73	ND	20.4	Pos	G	G	T	N501Y
15	18.29	ND	22.2	Pos	G	G	T	N501Y
16	19.00	ND	22.9	Pos	G	G	T	N501Y
17	18.24	ND	22.5	Pos	G	G	T	N501Y
18	22.59	ND	26.7	Pos	G	G	T	N501Y
19	18.97	ND	22.7	Pos	G	G	T	N501Y
20	14.6	18.3	ND	Neg	G	G	A	Wild-type
21	9.1	12.4	ND	Neg	G	G	A	Wild-type
22	18.9	22.7	ND	Neg	G	G	A	Wild-type
23	13.3	16.9	ND	Neg	G	G	A	Wild-type
24	25.4	ND	29.4	Pos	G	G	T	N501Y
25	12.8	15.9	ND	Neg	G	G	A	Wild-type
26	17.1	20.8	ND	Neg	G	G	A	Wild-type
27	17.2	ND	21.1	Pos	G	G	T	N501Y
28	17.7	ND	22.1	Pos	G	G	T	N501Y
29	27.9	ND	32.2	Pos	G	G	T	N501Y
30	22.4	ND	26.6	Pos	G	G	T	N501Y
31	12.7	ND	17	Pos	G	G	T	N501Y
32	13.2	ND	18.5	Pos	G	G	T	N501Y
33	34.90	35.8	ND	Neg	G	G	A	Wild-type

34	27.10	29.9	ND	Neg	G	G	A	Wild-type
35	19.97	22.0	ND	Neg	G	G	A	Wild-type
36	23.65	25.9	ND	Neg	G	G	A	Wild-type
37	28.23	30.5	ND	Neg	G	G	A	Wild-type
38	24.89	28.8	ND	Neg	G	G	A	Wild-type
39	23.13	25.6	ND	Neg	G	G	A	Wild-type
40	19.35	22.4	ND	Neg	G	G	A	Wild-type
41	24.49	26.9	ND	Neg	G	G	A	Wild-type
42	16.30	19.1	ND	Neg	G	G	A	Wild-type
43	30.41	33.3	ND	Neg	G	G	A	Wild-type
44	21.20	23.8	ND	Neg	G	G	A	Wild-type
45	26.96	29.7	ND	Neg	G	G	A	Wild-type
46	27.99	30.4	ND	Neg	G	G	A	Wild-type
47	25.96	28.6	ND	Neg	G	G	A	Wild-type
48	23.29	25.8	ND	Neg	G	G	A	Wild-type
49	31.79	33.7	ND	Neg	G	G	A	Wild-type
50	20.76	23.4	ND	Neg	G	G	A	Wild-type
51	21.56	24.8	ND	Neg	G	G	A	Wild-type
52	22.73	24.7	ND	Neg	G	G	A	Wild-type
53	17.08	19.5	ND	Neg	G	G	A	Wild-type
54	16.34	18.8	ND	Neg	G	G	A	Wild-type
55	27.65	30.3	ND	Neg	G	G	A	Wild-type
56	18.93	21.3	ND	Neg	G	G	A	Wild-type
57	21.52	24.3	ND	Neg	G	G	A	Wild-type
58	18.10	21.0	ND	Neg	G	G	A	Wild-type

59	16.46	19.0	ND	Neg	G	G	A	Wild-type
60	19.09	21.7	ND	Neg	G	G	A	Wild-type
61	20.80	23.6	ND	Neg	G	G	A	Wild-type
62	21.42	ND	23.7	Pos	G	G	T	N501Y
63	21.7	24.7	ND	Neg	G	G	A	Wild-type
64	20.1	22.7	ND	Neg	G	G	A	Wild-type
65	17.6	20.1	ND	Neg	G	G	A	Wild-type
66	22.8	25.7	ND	Neg	G	G	A	Wild-type
67	19.8	22.6	ND	Neg	G	G	A	Wild-type
68	19.7	22.1	ND	Neg	G	G	A	Wild-type
69	16.8	19.3	ND	Neg	G	G	A	Wild-type
70	18.8	21.3	ND	Neg	G	G	A	Wild-type
71	19.9	22.4	ND	Neg	G	G	A	Wild-type
72	23.6	ND	25.6	Pos	G	G	T	N501Y
73	24.7	ND	27.0	Pos	G	G	T	N501Y
74	22.1	ND	24.1	Pos	G	G	T	N501Y
75	17.0	ND	18.8	Pos	G	G	T	N501Y
76	21.9	ND	23.2	Pos	G	G	T	N501Y
77	24.3	26.5	ND	Neg	G	G	A	Wild-type
78	17.33	ND	20.12	Pos	G	G	T	N501Y
79	18.23	ND	21.74	Pos	G	G	T	N501Y
80	22.37	ND	24.91	Pos	G	G	T	N501Y
81	13.54	18.98	ND	Neg	G	A	A	Wild-type
82	19.63	ND	22.08	Pos	G	G	T	N501Y
83	23.34	ND	25.60	Pos	G	G	T	N501Y

84	19.65	ND	21.96	Pos	G	G	T	N501Y
85	14.29	ND	17.09	Pos	G	G	T	N501Y
86	16.77	ND	19.95	Pos	G	G	T	N501Y
87	21.30	ND	23.09	Pos	G	G	T	N501Y
88	22.16	ND	23.75	Pos	G	G	T	N501Y
89	20.03	ND	22.45	Pos	G	G	T	N501Y
90	16.44	ND	18.46	Pos	G	G	T	N501Y
91	14.93	ND	16.85	Pos	G	G	T	N501Y
92	15.04	ND	16.86	Pos	G	G	T	N501Y
93	16.38	ND	21.15	Pos	G	G	T	N501Y
94	13.71	ND	15.44	Pos	G	G	T	N501Y
95	17.89	ND	19.75	Pos	G	G	T	N501Y
96	27.11	31.80	ND	Neg	G	G	A	Wild-type
97	28.46	32.66	ND	Neg	G	G	A	Wild-type
98	20.26	24.98	ND	Neg	G	G	A	Wild-type
99	26.55	31.17	ND	Neg	G	G	A	Wild-type
100	25.55	30.62	ND	Neg	G	G	A	Wild-type
101	20.63	25.58	ND	Neg	G	G	A	Wild-type
102	16.87	ND	19.57	Pos	G	G	T	N501Y
103	16.04	ND	17.5	Pos	G	G	T	N501Y
104	13.64	15.7	ND	Neg	G	G	A	Wild-type
105	21.50	22.9	ND	Neg	G	G	A	Wild-type
106	18.10	19.4	ND	Neg	G	G	A	Wild-type
107	18.16	19.4	ND	Neg	G	G	A	Wild-type
108	16.52	18.6	ND	Neg	G	G	A	Wild-type

109	21.79	23.7	ND	Neg	G	G	A	Wild-type
110	20.81	22.4	ND	Neg	G	G	A	Wild-type
111	19.32	20.9	ND	Neg	G	G	A	Wild-type
112	12.65	15.2	ND	Neg	G	G	A	Wild-type
113	21.08	23.4	ND	Neg	G	G	A	Wild-type
114	22.37	24.2	ND	Neg	G	G	A	Wild-type
115	19.81	21.9	ND	Neg	G	G	A	Wild-type
116	18.91	20.9	ND	Neg	G	G	A	Wild-type
117	20.64	23.6	ND	Neg	G	G	A	Wild-type
118	20.99	23.3	ND	Neg	G	G	A	Wild-type
119	17.45	19.9	ND	Neg	G	G	A	Wild-type
120	17.30	ND	20.0	Pos	G	G	T	N501Y
121	20.92	ND	23.3	Pos	G	G	T	N501Y
122	15.60	17.9	ND	Neg	G	G	A	Wild-type
123	32.45	34.2	ND	Neg	G	G	A	Wild-type
124	23.01	25.5	ND	Neg	G	G	A	Wild-type
125	25.93	28.1	ND	Neg	G	G	A	Wild-type
126	12.08	14.5	ND	Neg	G	G	A	Wild-type
127	13.66	16.3	ND	Neg	G	G	A	Wild-type
128	12.11	14.4	ND	Neg	G	G	A	Wild-type
129	23.63	25.4	ND	Neg	G	G	A	Wild-type
130	27.52	29.4	ND	Neg	G	G	A	Wild-type
131	24.40	ND	26.7	Pos	G	G	T	N501Y
132	28.29	28.70	29.3	Mix	MIX base			MIX base
133	22.44	23.9	ND	Neg	G	G	A	Wild-type

134	24.06	25.6	ND	Neg	G	G	A	Wild-type
135	18.42	20.7	ND	Neg	G	G	A	Wild-type
136	17.93	19.7	ND	Neg	G	G	A	Wild-type
137	23.90	ND	25.4	Pos	G	G	T	N501Y
138	20.66	ND	22.9	Pos	G	G	T	N501Y
139	27.82	ND	30.4	Pos	G	G	T	N501Y
140	18.76	ND	20.8	Pos	G	G	T	N501Y
141	15.49	ND	17.4	Pos	G	G	T	N501Y
142	15.30	ND	17.5	Pos	G	G	T	N501Y
143	16.07	ND	18.5	Pos	G	G	T	N501Y
144	18.51	ND	21.0	Pos	G	G	T	N501Y
145	14.70	ND	17.0	Pos	G	G	T	N501Y
146	14.11	ND	16.7	Pos	G	G	T	N501Y
147	18.9	ND	21.7	Pos	G	G	T	N501Y
148	20.1	ND	22.4	Pos	G	G	T	N501Y
149	18.7	ND	20.6	Pos	G	G	T	N501Y
150	14.7	ND	15.8	Pos	G	G	T	N501Y
151	11.9	14.5	ND	Neg	G	G	A	Wild-type
152	15.0	18.3	ND	Neg	G	G	A	Wild-type
153	32.2	36.2	ND	Neg	G	G	A	Wild-type
154	23.7	27.9	ND	Neg	G	G	A	Wild-type
155	20.8	24.1	ND	Neg	G	G	A	Wild-type
156	17.0	19.8	ND	Neg	G	G	A	Wild-type
157	34.5	36.0	ND	Neg	G	G	A	Wild-type
158	15.1	17.8	ND	Neg	G	G	A	Wild-type

159	10.9	13.5	ND	Neg	G	G	A	Wild-type
160	ND	37.9	ND	Neg	G	G	A	Wild-type

442 a.ND= Not Detected. Mix means probes for N501 and N501Y were detected when the specimen
 443 was run in our assay and the Sanger sequencing results showed signals for both N501 and
 444 N501Y.

445

446 Supplementary Table 2: Summary of intra-run repeatability

447

Intra- run						Interpretation
CT		Average	SD	CV		
20.2	20.19	19.96	20.12	0.136	0.007	N501Y Variant
28.99	28.69	28.57	28.75	0.216	0.008	Wildtype
Und ^a	Und	Und	NA	NA	NA	undetermined
29.43	29.49	29.38	29.43	0.055	0.002	N501Y Variant
24.96	25	24.85	24.94	0.078	0.003	N501Y Variant
Und	Und	Und	NA	NA	NA	undetermined
23.55	23.68	23.3	23.51	0.193	0.008	Wildtype
30.4	30.38	30	30.26	0.225	0.007	N501Y Variant
21.57	21.57	21.31	21.48	0.15	0.007	N501Y Variant
Und	Und	Und	NA	NA	NA	undetermined
23.82	23.59	23.77	23.73	0.121	0.005	Wildtype
22.93	23.03	22.98	22.98	0.05	0.002	N501Y Variant
28.99	29.11	28.59	28.9	0.272	0.009	N501Y Variant
Und	Und	Und	NA	NA	NA	undetermined

Und	Und	Und	NA	NA	NA	undetermined
24.35	24.36	24.06	24.26	0.17	0.007	Wildtype
29.96	30.05	30.05	30.02	0.052	0.002	Wildtype
24.21	24.35	24.03	24.2	0.16	0.007	N501Y Variant
30.36 ^b	29.84	30.46	30.22	0.333	0.011	N501Y Variant
30.49 ^b	30.12	30.3	30.3	0.185	0.006	Wildtype
29.73	28.9	28.97	29.2	0.46	0.016	N501Y Variant

448

- a. Und: Undetermined (not positive for SARS-COV2)
- b. Specimen with mixed population of N501 and N501Y.

449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476

477 **Supplementary Appendix A:**

478

479 **PCR Amplification and Sequencing of Partial Spike Gene Methodology:**

480 Primers were designed to amplify a 698 bp S gene fragment spanning the region of interest (genome positions:
481 22516 to 23214). Reverse transcription and amplification of the partial S gene was performed using the OneStep
482 RT-PCR kit (QIAGEN, Hilden, Germany). A 25 µl reaction mixture contained 5 µl of 5X QIAGEN OneStep RT-
483 PCR Mix, 1 µl (200 µM) of dNTPs, 1 µl of OneStep RT-PCR enzyme mix, 1 µl of each forward (VOCF) and
484 reverse (VOCR) primers at a final concentration of 0.4 µM (**Supplementary Table 1**), 11µl of water and 5µl of
485 RNA. Reactions were carried out on the Applied Biosystems™ SimpliAmp® Thermal Cycler (Thermo Fisher
486 Scientific, Massachusetts, United States) with the following conditions: 1 cycle at 50 °C for 30 minutes and 1 cycle
487 at 95 °C for 15 minutes, followed by 40 cycles of amplification at 94 °C for 30 seconds, 56 °C for 30 seconds, and
488 72 °C for 1 minute, and a final extension cycle of 72 °C for 10 minutes.

489

490 Applied Biosystems™ BigDye™ Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific, Massachusetts,
491 United States) was used for bidirectional sequencing of PCR fragments following manufacturer's recommendations.
492 Each reaction contained 0.5 µl of Terminator Ready Reaction Mix, 3.75 µl of 5X buffer, 1µl of primer mix and 1 µl
493 amplified PCR product, and 13.75µl of water in a final volume of 20 µl. Default cycling conditions for the
494 BigDye™ Terminator reaction were used as follows: 96°C for 1 minute followed by 35 cycles of 96°C for 10
495 seconds, 50°C for 5 seconds, and 60°C for 4 minutes. BigDye™ reaction products were purified using an EDTA/
496 ethanol precipitation method and resuspended in 20µl Hi-Di™ Formamide (Thermo Fisher Scientific,
497 Massachusetts, United States). Nucleotide sequences were determined using the Applied Biosystems™ ABI 3130xl
498 Genetic Analyzer (Thermo Fisher Scientific, Massachusetts, United States) and analyzed using FinchTV software
499 (Geospiza, Inc.).

500