

1 **Rapid base-specific calling of SARS-CoV-2 variants of concern using** 2 **combined RT-PCR melting curve screening and SIRPH technology**

3 **Running title: Rapid calling of SARS-CoV-2 variants**

4 Tierling Sascha^{1§}, Kattler Kathrin^{1§}, Vogelgesang Markus², Pfuhl Thorsten², Lohse Stefan², Lo Porto
5 Christina¹, Schmitt Beate¹, Salhab Abdulrahman¹, Smola Sigrun^{2*}, Walter Jörn^{1*}

6

7 ¹ Department of Genetics, Saarland University, Saarbrücken, Saarland, 66123, Germany

8 ² Institute of Virology, Saarland University Medical Center, Homburg, Saarland, 66421, Germany

9 [§] shared first-authorship, ^{*} shared senior-authorship

10 Corresponding author: Prof. Dr. Jörn Walter – Tel: +49 (0) 681 302 4367; Fax: +49 (0) 681 302 2703;

11 Email: j.walter@mx.uni-saarland.de.

12

13

14 **ABSTRACT (86 words)**

15 The emergence of novel variants of concern of SARS-CoV-2 demands a fast and reliable detection of
16 such variants in local populations. Here we present a cost-efficient and fast workflow combining a pre-
17 screening of SARS-CoV-2 positive samples using RT-PCR melting curve analysis with multiplexed IP-
18 RP-HPLC-based single nucleotide primer extensions (SIRPH). The entire workflow from positive SARS-
19 CoV-2 testing to base-specific identification of variants requires about 24 h. We applied the sensitive
20 method to monitor the local VOC outbreaks in a few hundred positive samples collected in a confined
21 region of Germany.

22

23

24

25 **Keywords (3-10):** SARS-CoV-2, Variants of concern, SNUPE, SIRPH, mutation screening

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

26

27 **Background (1999 words)**

28 During the ongoing COVID-19 pandemic recent reports of new SARS-CoV-2 variants with changes in
29 viral functionalities or pathogenic aspects have caused major concern. In December 2020 several
30 emerging variants of concern (VOCs) with increased transmission were reported from UK (B.1.1.7),
31 South Africa (B.1.351) and Brazil (P.1). To allow the timely adaptation of public health interventions it is
32 of utmost importance to monitor and precisely follow the ingress and distribution of these VOCs as
33 quickly as possible. They are each characterized by a set of mutations that can be used for variant
34 monitoring. Apart from the spike-protein mutation N501Y in all three currently circulating VOCs, B.1.351
35 and P.1 also share the mutations K417N/T and E484K, but are discriminated by the mutation V1176F
36 only present in P.1 [1-3]. The amino acid substitution E484K in the receptor-binding domain of the spike-
37 protein was reported to be associated with immune escape from neutralizing antibodies which is
38 especially worrisome regarding vaccination or reinfection [4].

39 The gold standard for the identification of virus variants is sequencing of full viral genomes. Currently,
40 RT-PCR melting curve analysis approaches followed by confirmation through viral genome sequencing
41 are widely used. However, the required time for NGS sequencing and the extra costs ask for faster and
42 more cost-effective test systems. Previously, we proposed a pooling approach for high-throughput RT-
43 PCR testing to identify SARS-CoV-2 infection in large cohorts of asymptomatic people [5]. Here we
44 extend this strategy to obtain a fast and accurate base-specific detection of VOCs by combining a
45 sequential RT-PCR based melting curve analysis as a pre-screening approach of SARS-CoV-2 positive
46 samples with a multiplexed IP-RP-HPLC-based single nucleotide primer extension approach (SIRPH,
47 6, 7]. This protocol allows a fast and comprehensive generation of “single base sequencing” signatures
48 for a relevant spectrum of VOCs to assign the variant lineages within the (local) pandemic situation.

49

50 **METHODS**

51 **SARS-CoV-2 (pool)-testing and N501Y RT-PCR melting curve analysis.** Respiratory samples were
52 tested individually (expected high positivity rate) or in pools (expected low positivity rate, 16 samples
53 per pool) performed with Biomek i5 Span8 (Beckman Coulter) by automated SARS-CoV-2 dual target
54 RT-PCR (E and ORF1) using Cobas 6800 (Roche Diagnostics). To identify VOC candidates positive
55 samples were extracted using RNAAdvance Viral and Biomek i5 MC96 (Beckman Coulter) and subjected
56 to N501Y RT-PCR melting curve analysis (VirSNiP SARS-CoV-2 Spike N501Y, TibMolBiol) using

57 LightCycler® 480 Instrument II (Roche Life Sciences). All analyses were performed according to the
58 manufacturers` instructions. The study was approved by the local ethics committee of the Saarland
59 University Medical Center at Saarland Ärztekammer.

60 **Reverse Transcription.** Isolated RNA was reverse transcribed using random hexamer primers (NEB)
61 and Maxima Reverse Transcriptase H- (ThermoFisher) following the manufacturer`s recommendations.
62 cDNA from the same preparation was submitted to PCRs for SNuPE and NGS Library generation.

63 **SNuPE and IP-RP-HPLC (SIRPH).** Two µl of cDNA were used as template in a 30 µl reaction in the
64 presence of 3 mM Tris–HCl (pH 8.8), 0.7 mM (NH₄)₂SO₄, 50 mM KCl, 2.5 mM MgCl₂, 0.06 mM of each
65 dNTP, 3 U HotFire DNA polymerase (Solis BioDyne) and 167 nM primers (Sup. Table 1). PCRs were
66 performed at 95° C for 15 min followed by 35 cycles 95°C/60 s, 54°C (58°C for A570D+D614G) /30 s,
67 72°C/30 s and a final extension 72°C/5 min. Five µl of PCR products were treated with 1U of ExoCIAP
68 (mixture of Exonuclease I (Jena Bioscience) and Calf Alkaline Intestine Phosphatase (Calbiochem)) for
69 30 min at 37°C. To inactivate the ExoCIP enzymes the reaction was incubated for 15 min at 80°C.
70 Afterwards, 14 µl primer extension mastermix (50 mM Tris–HCl, pH 9.5, 2.5 mM MgCl₂, 0.05 mM of
71 ddNTPs, 1.6 µM of each SNuPE primer, 2.5 U Termipol DNA polymerase (Solis BioDyne) was added.
72 Primer extension reactions were performed at 96°C for 2 min followed by 50 cycles 96°C/30 s, 50°C/30
73 s, 60°C/20 s. Separation of products was conducted on a XBridge BEH C18 2.5µm 4.6mmx50mm
74 column (Waters) at 0.9 ml/min and 50°C by continuously mixing buffer B (0.1 M TEAA, 25% acetonitril)
75 to buffer A (0.1 M TEAA, Sup. Table 2). Mutations were determined by relative retention times in
76 comparison to the non-mutated reference.

77 **Sanger sequencing.** Sanger sequencing of exemplary PCR products generated for SIRPH analysis
78 was performed at Macrogen (Amsterdam) using 40 ng purified PCR product and 10 pmol of
79 corresponding PCR primers.

80 **Viral genome sequencing.** NGS Libraries for viral genome sequencing were generated using the
81 primer design from the ARTIC v0.3 protocol [8] modified for Illumina sequencing. We added Illumina
82 compatible adapter sequences directly to the ARTIC Primers (Sup. Table 3). To obtain a more equal
83 coverage we introduced a third multiplex pool with 5 primer pairs covering difficult genomic regions (Sup.
84 Table 3). Multiplex PCRs were carried out as described in the ARTIC protocol. After purification PCR
85 products were submitted to an indexing PCR with TruSeq primers (Illumina) for 8 cycles. After final
86 purification and normalization samples were sequenced on a MiSeq (Illumina).

87 **NGS data processing.** Viral genome data were processed using CoVpipe
88 (https://gitlab.com/RKIBioinformaticsPipelines/ncov_minipipe). After trimming of adapter/primer
89 sequences and low quality reads (<Q30), reads were aligned to the SARS-CoV-2 reference
90 NC_045512.2. Reads mapping to the human genome were excluded. Variant calling was performed
91 with default parameter and consensus sequences were masked for lowly covered sites (<20X).
92 Visualization was achieved with Auspice and Nextstrain [9].

93

94 **RESULTS**

95 Single nucleotide primer extension (SNUPE) coupled to IP-RP-HPLC detection (SIRPH) is a powerful
96 and fast approach for point mutation analysis of known genomic variants. This assay is characterized
97 by primers that anneal with their 3'-end directly adjacent to potentially mutated sites (PMS). Using
98 ddNTPs the primer can only be extended by a single nucleotide complementary to the genomic
99 information on the annealed DNA strand. Due to the incorporation of different bases in wildtype and
100 mutant, extended primers differ in hydrophobicity which allows their separation by IP-RP-HPLC. We
101 adapted this approach for screening of VOCs in the SARS-CoV-2 genome in a multiplex assay that
102 allows simultaneous detection of several mutations.

103 We established the method on 80 SARS-CoV-2 positive samples collected in the south-west of Germany
104 (Saarland) from Nov/Dec 2020 for B.1.1.7, B.1.351 and P.1 (Figure 1A). We designed primers for our
105 SNUPE assay (Figure 1B/C, Sup. Table 1) to determine the nucleotide at the PMS D614G (A→G in all
106 clades except 19A), N501Y present in all 3 currently circulating VOCs (A→T) and the discriminatory
107 PMS A570D (C→A in B.1.1.7), P681H (C→A in B.1.1.7), T716I (C→T in B.1.1.7), E484K (G→A in
108 B.1.351 and P.1) as well as V11176T (G→T in P.1). Besides, we analyzed two local mutations, i.e. L9P
109 (T→C) and G172R (G→C) that we observed in a high frequency in NGS data of earlier samples (Sep/Oct
110 2020, Figure 1D) from Saarland to follow their further ingression and distribution.

111 Notably, mutations in close vicinity of each other that could be covered within the same PCR product
112 (Figure 1C, Sup. Figure 1) were assessed simultaneously in a multiplexed SNUPE reaction. After RNA
113 extraction and reverse transcription, cDNA was used as template for four different PCR reactions
114 yielding PCR products harbouring the PMS (Figure 1B). In a following one-tube reaction PCR products
115 were treated with exonuclease I and calf intestinal phosphatase (ExoCIP) to remove primers and
116 nucleotides, heat-inactivated and subjected to SNUPE followed by IP-RP-HPLC separation (Figure 1B)
117 in 96 sample formats. In these SNUPE reactions we either assessed two PMS at once (E484K, N501Y,

118 P681H, T716I, L9P, G172R) or one PMS (A570D, D614G) per reaction (Sup. Figure 1) resulting in a
119 total of five HPLC runs per sample.

120 All samples showed the mutation D614G which emerged early in the pandemic and is now predominant
121 in samples from across the globe [10]. The PMS characteristic for the three analyzed VOCs (E484K,
122 N501Y, P681H, T716I, A570D) were in concordance with the original SARS-CoV-2 sequence (Sup.
123 Table 4). Our local mutations revealed interesting patterns over time (Sup. Table 4). The spread of the
124 ORF9c mutation L9P dropped from 35.3% in Sep/Oct 2020 (n = 51) to 6.25 % in Nov/Dec 2020 (n = 80)
125 while the ORF3A mutation G172R remained more abundant (33.3% to 27.5%).

126 For validation we performed full viral genome sequencing on all 80 samples. For variant calling we only
127 considered sites with coverage of at least 20 (Sup. Table 4). Especially in the region harbouring our
128 local L9P variant we observed high coverage-related drop-out rates in the sequencing data. Thus, we
129 performed additional validation of exemplary samples by Sanger sequencing (Sup. Figure 2). All SNUPE
130 results were in full concordance with viral genome sequencing and Sanger sequencing-based variant
131 calling (Sup. Table 4, Sup. Figure 2) demonstrating the reliability and robustness of our SNUPE
132 approach.

133 Inspired by these results we developed a rapid screening strategy for prospective high-throughput
134 SARS-CoV-2 and subsequent VOC testing (Figure 2A). Asymptomatic individuals, i.e. health care
135 workers of the local University Hospital, were screened by SARS-CoV-2 pool testing in pools of 16, while
136 samples from symptomatic people or public health services (contact tracing) with expected positivity
137 rates above 2% were investigated by individual SARS-CoV-2 RT-PCR. Positive nucleic acid extracts
138 were immediately subjected to N501Y mutation specific RT-PCR melting curve analysis to identify
139 candidates of currently circulating VOCs. Subsequent SNUPE analysis was performed to clearly identify
140 and assign the respective VOC.

141 We applied this approach for a routine test on a large cohort to identify B.1.1.7, B.1.351 and P.1 variants
142 in Saarland generating reliable and highly accurate results. From 2059 pooled and 2987 individual
143 respiratory samples collected between January, 25th and 31st 2021, 542 samples were tested SARS-
144 CoV-2 positive by RT-PCR. Within these we identified 13 VOC candidates by RT-PCR melting curve
145 analysis based screening during this early phase of VOC introduction into Saarland. All candidates were
146 submitted to SIRPH analysis for confirmation and further VOC assignment revealing the presence of all
147 characteristic mutations 501Y, 570D, 681H and 716I in 12 out of 13 pre-defined samples clearly

148 classifying them as B.1.1.7, while 1 sample was identified as B.1.351 (Figure 2B, Sup. Table 6). The
149 whole process lasted about 24 hours from positive RT-PCR testing to PMS sequence validation.

150

151 **DISCUSSION**

152 Our presented strategy shows that a combination of pre-screening by (pooled) SARS-CoV-2 and
153 subsequent N501Y RT-PCR melting curve analysis together with a confirmatory SIRPH analysis
154 generates fast and reliable results to speed up monitoring of VOCs in larger cohorts. We demonstrate
155 how this strategy can quickly be implemented for VOC monitoring to rationalize local public health
156 interventions by supporting the tracing of VOCs. While whole-genome sequencing will remain the gold
157 standard for virus monitoring and the detection of new mutations, the proposed screening by
158 PCR/SIRPH has several advantages. First of all the method is robust, fast (24h) and very cost-effective
159 compared to NGS confirmation. Notably, it robustly calls regions with coverage drop-outs in sequencing
160 data (Sup. Figure 2). It can even be applied to samples with low viral load ($Ct \leq 33$) that typically fail or
161 result in low quality data during sequencing ($Ct > 27$). Furthermore the method is flexible and can be
162 implemented on a variety of existing RP-HPLC systems. In addition, we perform SNUPE assays as “one-
163 tube” reactions that can be multiplexed. New mutation specific assays can be rapidly established within
164 two working days to include newly emerging variants allowing a fast and precise population surveillance.
165 The combined pool PCR/SIRPH approach can be used to flank sequencing surveillance programs
166 where usually only up to 5% of randomly chosen positive cases will be full-genome sequenced. The
167 emergence and spreading of local variants asks for such fast and deep monitoring in the entire
168 population in a timely manner. Since time is crucial for surveillance, our method offers base-specific
169 results in one day as compared to NGS-based sequencing which usually require up to 5 working days.
170 Ongoing sequencing of full viral genomes in a proportion of samples will be used to screen for the
171 development of new SARS-CoV-2 variants. In an iterative process the information gained by sequencing
172 can be used to easily adapt the SIRPH approach to new VOCs that may emerge in the future course of
173 the pandemic.

174

175

176

177 **ACKNOWLEDGEMENT**

178 The authors wish to acknowledge the contributions of the technicians Katrin Thieser, Luisa Scheidhauer,
179 Helga Appel and Barbara Best.

180

181 **References**

- 182 1. Davies NG, Barnard RC, Jarvis CI, et al. Estimated transmissibility and severity of novel SARS-
183 CoV-2 Variant of Concern 202012/01 in England. *Epidemiology*; 2020 Dec. Available from:
184 <http://medrxiv.org/lookup/doi/10.1101/2020.12.24.20248822>
- 185 2. Tegally H, Wilkinson E, Giovanetti M, et al. Emergence and rapid spread of a new severe acute
186 respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike
187 mutations in South Africa [Internet]. *Epidemiology*; 2020 Dec. Available from:
188 <http://medrxiv.org/lookup/doi/10.1101/2020.12.21.20248640>
- 189 3. Voloch CM, Silva F R da, Almeida LGP de, et al. Genomic characterization of a novel SARS-
190 CoV-2 lineage from Rio de Janeiro, Brazil [Internet]. *Genetic and Genomic Medicine*; 2020 Dec.
191 Available from: <http://medrxiv.org/lookup/doi/10.1101/2020.12.23.20248598>
- 192 4. Greaney AJ, Starr TN, Gilchuk P, et al. Complete Mapping of Mutations to the SARS-CoV-2
193 Spike Receptor-Binding Domain that Escape Antibody Recognition. *Cell Host & Microbe*. **2021**;
194 29(1):44-57.e9.
- 195 5. Lohse S, Pfuhl T, Berkó-Göttel B, et al. Pooling of samples for testing for SARS-CoV-2 in
196 asymptomatic people. *The Lancet Infectious Diseases*. **2020**; 20(11):1231–1232.
- 197 6. Tierling S, Sers C, Lehmann A, Walter J. A fast, cost-efficient and sensitive approach for KRAS
198 mutation detection using multiplexed primer extension with IP/RP-HPLC separation.
199 *International Journal of Cancer*. **2012**; 130(3):567–574.
- 200 7. Hoogendoorn B, Owen MJ, Oefner PJ, Williams N, Austin J, O'Donovan MC. Genotyping
201 single nucleotide polymorphisms by primer extension and high performance liquid
202 chromatography. *Human Genetics*. **1999**; 104(1):89–93.
- 203 8. Tyson JR, James P, Stoddart D, et al. Improvements to the ARTIC multiplex PCR method for
204 SARS-CoV-2 genome sequencing using nanopore [Internet]. *Genomics*; 2020 Sep. Available
205 from: <http://biorxiv.org/lookup/doi/10.1101/2020.09.04.283077>

206 9. Hadfield J, Megill C, Bell SM, et al. Nextstrain: real-time tracking of pathogen evolution.
207 Bioinformatics. **2018**; 34(23):4121–4123.

208 10. Volz E, Hill V, McCrone JT, et al. Evaluating the Effects of SARS-CoV-2 Spike Mutation D614G
209 on Transmissibility and Pathogenicity. Cell [Internet]. **2020** [cited 2021 Jan 8]; . Available from:
210 <https://linkinghub.elsevier.com/retrieve/pii/S0092867420315373>

211

212

213 **Footnotes**

214 **Conflict of interest**

215 None declared.

216

217 **Funding**

218 This work was supported by the Staatskanzlei Saarland (SaarCoSeq) to Sigrun Smola and Jörn Walter
219 and the Ministry of Health Saarland (SaarCoMM) to Prof. Smola. Abdulrahman Salhab was supported
220 by the German Federal Ministry of Research and Education grant for de.NBI (031L0101D).

221

222 **Mention of any meeting(s) where the information has previously been presented**

223 Not applicable.

224

225 **Corresponding author contact information**

226 Prof. Dr. Jörn Walter – Tel: +49 (0) 681 302 4367; Fax: +49 (0) 681 302 2703; Email: [j.walter@mx.uni-](mailto:j.walter@mx.uni-saarland.de)
227 [saarland.de](mailto:j.walter@mx.uni-saarland.de).

228

229 **Data accession**

230 Fasta files of full viral genome sequencing can be accessed on GISAID (<https://www.gisaid.org/>) using
231 the identifiers listed in Sup. Table 6.

232

233

234 **FIGURE LEGENDS**

235 **Figure 1.** Multiplexed IP-RP-HPLC-based single nucleotide primer extension approach (SIRPH). A –
236 Phylogenetic tree of SARS-CoV-2 sequences derived from GISAID references (accessed on
237 2021/01/29) and sequencing data from Saarland (March – October 2020, available on GISAID using
238 identifiers listed in Sup. Table 6) colored by clade. Variants of concern are found in clades 20I/501Y.V1
239 (B.1.1.7), 20H/501Y.V2 (B.1.351) and 20J/501Y.V3 (P.1). B – SIRPH Workflow: Isolated RNA is reverse
240 transcribed, followed by a multiplexed PCR of regions of interest. After ExoCIP clean-up of residual
241 dNTPs and primers, multiplexed SNUPE reactions are performed that are separated by IP-RP-HPLC.
242 The whole workflow can be performed within 1 day. C – Exemplary SNUPE primer design (green) for
243 the PMS (blue) A570D and D614G within a single PCR product (yellow). D – Distribution of the local
244 mutations ORF9c:L9P (green) and ORF3A:G172R (yellow) in samples from Saarland (March – October
245 2020). E – Exemplary HPLC separation of SNUPE reactions at the PMS D614G showing a profile for
246 wildtype (Wt, A) and mutation (Mut, G).

247

248 **Figure 2.** High-throughput screening strategy for variant monitoring in a local population. A – Iterative
249 Workflow from Identification of SARS-CoV-2 positive samples (1) followed by variant candidate
250 identification by RT-PCR melting curve analysis (2) and final identification of specific VOCs as defined
251 by global sequencing efforts. Detailed proportions of VOCs or local variants of interest identified by
252 SIRPH analysis can then be used for selection of samples in a local sequencing projects. B – HPLC
253 separation profiles of the VOC characterizing PMS set for an exemplary non-VOC (ID65), a B.1.1.7 and
254 a B.1.351 variant as identified in the workflow described in A.

255

256

257

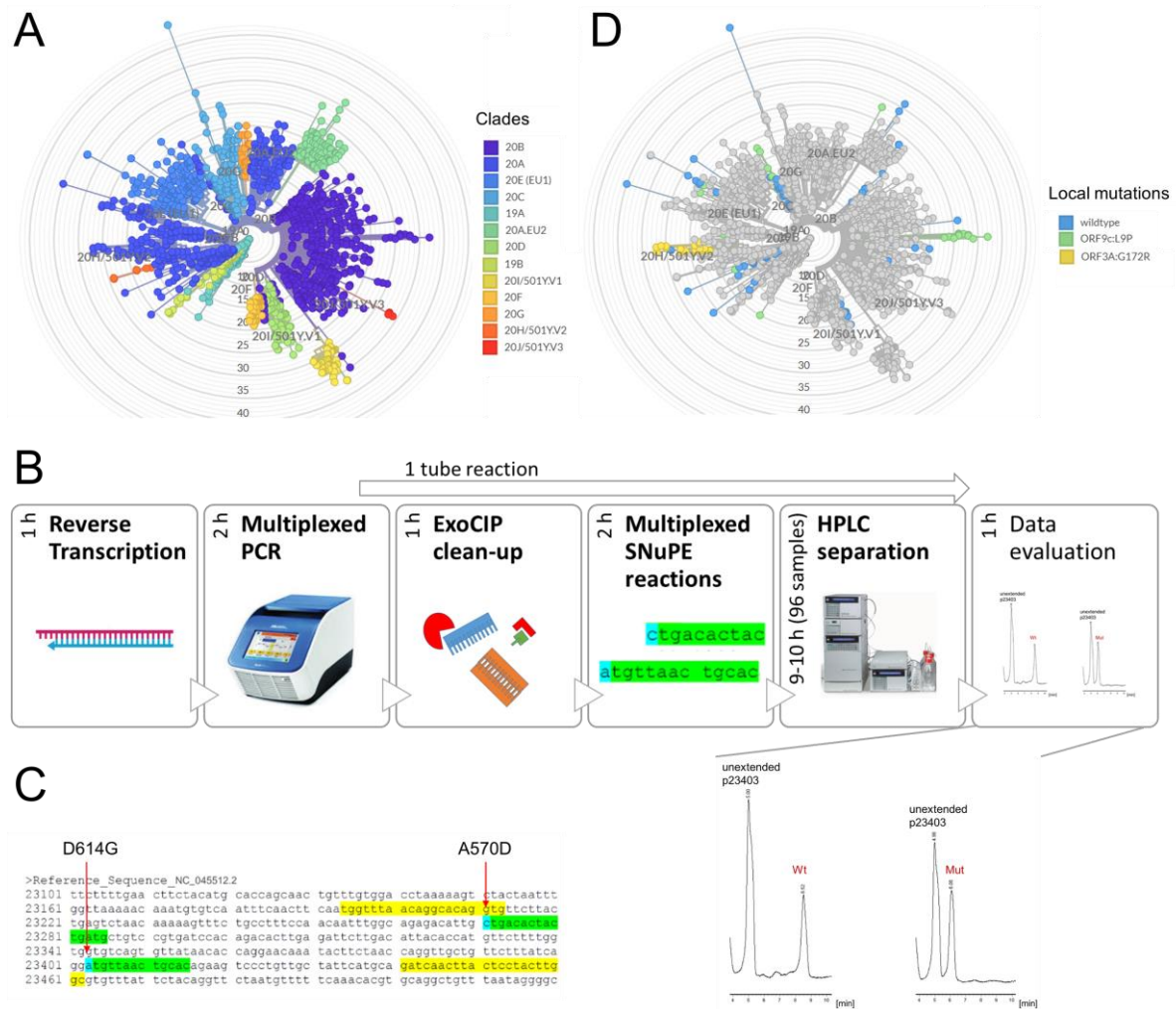
258

259

260

261

262 Figure 1



263

264

265

266

267

268

269

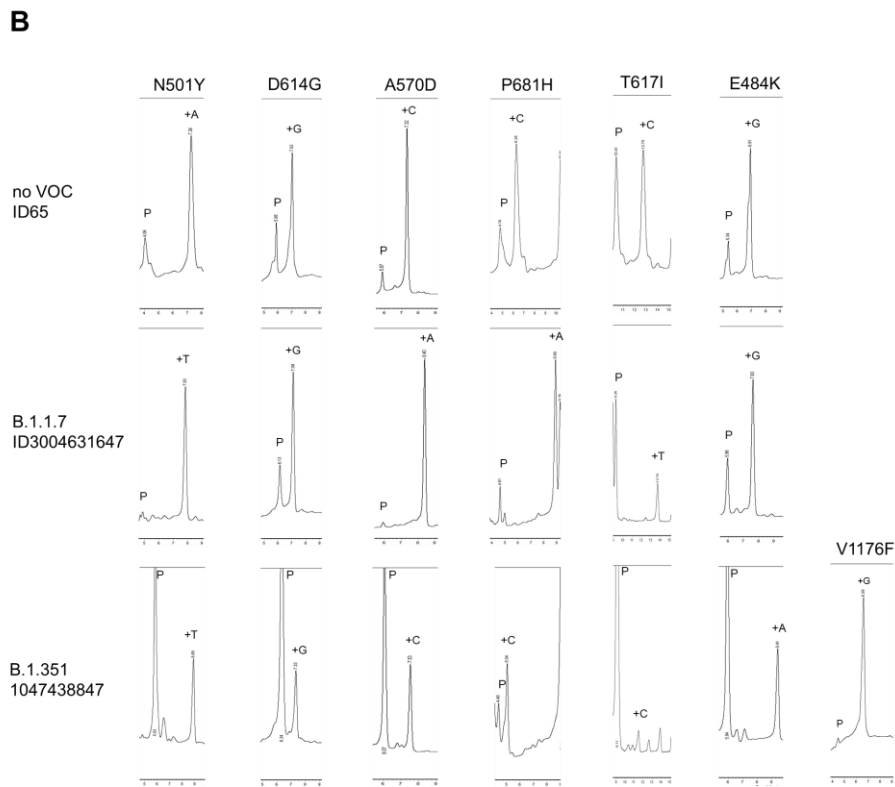
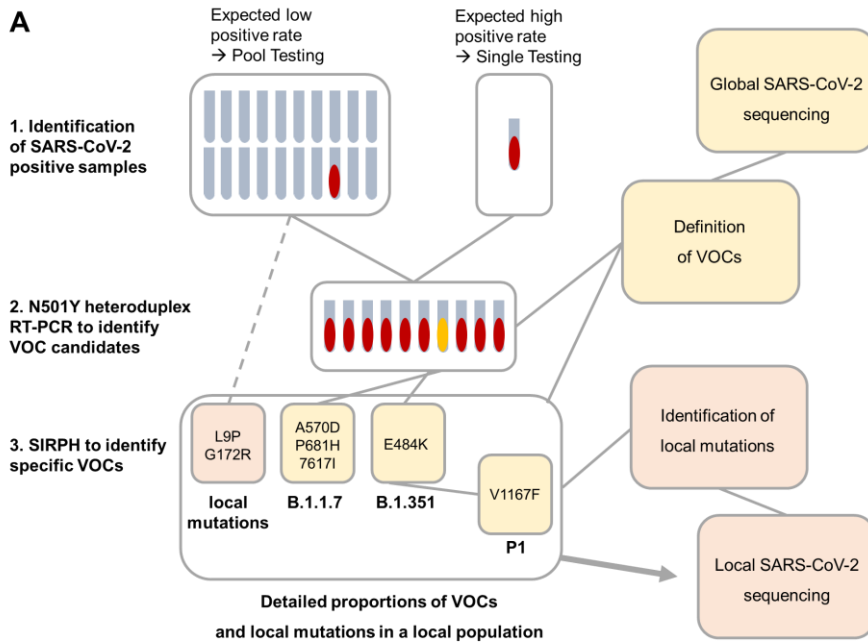
270

271

272

273

274 Figure 2



275