

Affordable and time-effective high throughput screening of SARS-CoV-2 variants using Denaturing High-Performance Liquid Chromatography analysis

Maria Elena Turba¹, Domenico Mion⁵, Stavros Papadimitriou², Francesca Taddei³, Giorgio Dirani³, Vittorio Sambri^{3,4}, Fabio Gentilini^{5*}

¹Xenturion srl, Forlì, Italy

² ADS Biotec Inc, Omaha NE, USA

³Unit of Microbiology, The Great Romagna Hub Laboratory, Pievesestina, Italy

⁴Department of Experimental, Diagnostic and Specialty Medicine - DIMES, University of Bologna, Bologna, Italy

⁵Department of Veterinary Medical Sciences, University of Bologna, Ozzano dell'Emilia, Bologna, Italy

* Corresponding author

E-mail: fabio.gentilini@unibo.it (FG)

Abstract

Mutations in the receptor binding domain region of SARS-CoV-2 have been shown to impact the infectivity, pathogenicity and transmissibility of new variants. Even more worrisome, those mutations have the potential of causing immune escape, undermining the population immunity induced by ongoing mass vaccination programs.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

The massive parallel sequencing techniques have taken a lead role in the detection strategies of the new variants. Nevertheless, they are still cumbersome and labour-demanding. As a matter of fact, there is an urgent need for novel strategies and techniques aimed at the surveillance of the active emergence and spread of the variants of concern.

Two PCRs were designed to target the binding domain region, spanning residues N417 through N501 of the Spike protein. Furthermore, very rapid, affordable and straightforward Denaturing High-Performance Liquid Chromatography screening was set up. The screening consisted of mixing the unknown sample with a standard sample of a known sequence, denaturing at high temperature for two minutes, renaturing for 15 minutes followed by a 2-minute run using the WAVE® DHPLC system to detect the heteroduplexes which invariably originate whenever the unknown sample has a nucleotide difference with respect to the standard used. The workflow was able to readily detect new variants including the P.1 and the B.1.585 strains at a very affordable cost. This approach has the potential of greatly expediting surveillance of the SARS-CoV-2 variants.

Keyword: SARS-CoV-2, Denaturing High-Performance Liquid Chromatography, Variant of concern, surveillance, Region Binding Domain

Introduction

The SARS-CoV-2 genome is relatively stable due to the proofreading activity operated by a 3'-to-5' exoribonuclease (nsp14-ExoN) during replication which reduces the error rate of RNA polymerase 100–1000-fold as compared with other RNA viruses. This confers the capacity of maintaining its 30,000 nt genome to the virus without catastrophic mutational events hampering its integrity [Plante et al., 2021; Triggler et al., 2021]. Nonetheless, errors still occur in the SARS-CoV-2 genome at a higher rate than eukaryotic cells and, together with high replication rates, allow for the accumulation of mutations in the viral genome including amino acid changes, truncations or the loss of viral proteins [Plante et al., 2021]. Whenever these changes impact infectivity, pathogenicity and transmissibility, they could lead to higher fitness and undergo positive selection [Giovannetti et al., 2020; Li et al., 2020; Plante et al., 2021].

The D614G substitution of the Surface protein is the most investigated of the positively selected mutation. This mutation occurred early in the first months of 2020 and became rapidly prevalent worldwide due to the increased infectivity of the strains carrying it. The phenotypic advantage was first evident at the epidemiological level and it was then confirmed at the molecular level using *in-vitro* experiments [Groves et al., 2021; Jackson et al., 2021; Plante et al., 2021bis].

As the COVID-19 pandemic continues, and vaccination programs expand, however, a novel factor of positive selection has taken place represented by the rapid rise of acquired immunity in the population [Jung, 2021]. As the majority of vaccines exploit the immunogenicity of the Spike protein, due to its pivotal role in binding to the ACE2 cell receptor and entry into the host cell, the strict monitoring of mutations in the domains targeted by the neutralising antibodies should be implemented worldwide [Focosi and Maggi, 2021; Gomez et al., 2021; Dobey et al., 2021]. In particular, substitutions in the receptor binding domain (RBD) have emerged, and they are of particular concern due to the possibility of being responsible for immune escape. More specifically, the B.1.1.7 (originally identified in the United Kingdom), B.1.351 (Republic of South Africa; RSA), P1 (Brazil),

and B.429 (California, USA) variants carry mutations in the 417 and 484 codons which have strong implications for infectivity and immune evasion [Plante et al 2021].

In fact, some of the variants of concern which might escape vaccine immunity are being monitored. To this end, an unprecedented sequencing effort has been pushed forward by the implementation of next generation sequencing (NGS) programmes. However, NGS is still cumbersome and very labour intensive. As a consequence, the approach to sequencing any viral genome is still unaffordable and, likely, unwise. Alternatives have been proposed to be integrated with sequencing in order to more effectively monitor the variants of concern [Bal et al., 2021; Banada et al., 2021].

The aim of this study was to provide a very quick, cheap straightforward method; this aim was achieved using Denaturing High-Performance Liquid Chromatography (DHPLC) to screen the RBD of SARS-CoV-2 isolates.

Materials & methods

Positive samples diagnosed with Covid-19 by means of diagnostic reverse transcription polymerase chain reaction (RT-PCR) (Seegene) underwent DHPLC screening and entire genome sequencing using Oxford Nanopore. Furthermore, some samples from External Quality Assessment (EQA) for Molecular Diagnostics were also included in this proof-of-concept study. All the samples were retro-transcribed using an iScript cDNA Synthesis Kit (Bio-Rad). The overall workflow of the screening proposed is represented in Figure 1. All positive samples should undergo to the screening. The variants used as standard should be chosen according to the more prevalent variants identified in the lab.

The PCR reactions for amplifying two sub-regions within the RBD of the S gene were designed using the Primer Quest web design tool (<https://eu.idtdna.com/Primerquest/Home/Index>). The PCR reactions were designed keeping the amplicon approximately 200-300 bp long and by including the 417 codon in the RBD_1 PCR, and the 501 and 484 codons in the RBD_2 PCR. The PCR primers

were F_Var417_NCov19, 5' TAAGTGTATGGAGTGTCTCCTACTA 3', R_Var417_NCov19, 5' GGTTTGAGATTAGACTTCCTAAACAATC 3', F_Var484_501_NCov19, 5' TCTTGATTCTAAGGTTGGTGGTAA 3' and R_Var484_501_NCov19, 5' AGTTGCTGGTGCATGTAGAA 3'.

Briefly, the PCR reactions consisted of 1X Buffer (Phusion HF Buffer, Life Technologies), 0.2 mM MgCl₂, 350 nM each of the forward and reverse primers, deoxynucleotide triphosphates (dNTPs) 250 μM, 0.4 U of Phusion Taq DNA polymerase (ThermoFisher), 2 μL of cDNA template brought up to a final volume of 20 μL with molecular biology grade water. The cycling programme for both the conventional PCRs consisted of the following steps: 98°C for 30 s, 40 cycles at 98°C for 5 s, 63°C for 5 s and 72°C for 10 s, and a final elongation step at 72° C for 10 min.

DHPLC

Heteroduplexes were generated by denaturing the amplified fragments at 95°C for 10 min, and they were allowed to renature at room temperature for 15 min. Each fragment was run on an automated DHPLC apparatus (WAVE® DHPLC system, ADS Biotec) equipped with a proprietary column (DNASep®, ADS Biotec). The elutions were carried out with a mixture of Buffer A (0.1 mol/L triethylammonium acetate–TEAA) and Buffer B (0.1 mol/L TEAA, 25% acetonitrile).

The optimal oven temperatures for heteroduplex separation were determined using Navigator™ Software (ADS Biotec) which gave a computer-assisted determination of the melting profile and analytical conditions for each fragment which were then experimentally verified. A partially denaturing temperature of 55.5°C was then used for the DHPLC screening of the N417T/N mutation in RBD_1 amplicon and a partially denaturing temperature of 56.6° C for the E484K and N501Y mutations in the RBD_2 amplicon. The data analysis was carried out using Navigator™ Software (ADS Biotec). Following the DHPLC screening, the same amplicons were purified and sequenced using an ABI310 automated sequencer (ThermoFisher). All the variants used in the study had already

been fully sequenced using a modified ARTIC protocol (unpublished) involving Nanopore Sequencing (Oxford Nanopore Technologies).

Results and Discussion

When two amplicons containing nucleotide variations are denatured by heating and left to renature by cooling, in addition to homoduplexes, heteroduplexes due to cross-hybridisation between the mismatched strands are also formed; DHPLC exploits the different retention times of homoduplexes and heteroduplexes in order to detect their multiple retention peaks. To almost cover the entire RBD of the Spike protein, two PCRs were designed, and the respective amplicons were screened for mutations using DHPLC. A sample of a known sequence was used as a standard and was mixed with unknown samples. Whenever more than one retention peak was evident at DHPLC, it indicated that a mismatch, and hence a mutation, existed between the standard and the sample.

The setup of the assays was straightforward. To obtain sharp and easily interpretable peaks at DHPLC analysis, PCR reactions should avoid the formation of nonspecific amplicons and smears. The high-fidelity Phusion Taq polymerase is well-suited for this purpose; furthermore, the PCR reactions using Phusion Taq are very rapid in reducing the PCR step. It should also be pointed out that the Phusion buffer does not contain any detergent which would be detrimental for downstream DHPLC analysis.

In this study involving DHPLC analysis, the B.1 strain isolated in March 2020 carrying the K417, E484 and N501 residues in the RBD_1 and RBD_2 amplicons was first used. Using this strain, the prevalent Y501 [Kirby, 2021] could be identified (Figure 2) and used as a standard for subsequent analysis which allowed detecting the P.1 and B1.585 variants (Figure 3), as confirmed by Sanger Sequencing of the amplicons and whole genome sequencing using Oxford Nanopore (Oxford Nanopore Technologies).

Overall, the process was able to be completed in a few hours. Potentially, all new positives would be able to be screened on the same day of diagnosis using a dedicated thermal cycler and the WAVE® DHPLC system (Figure 1). Besides the cost of the additional PCR reaction, an additional reagent cost for each DHPLC analysis would be 1.40 Euros/amplicon. This amounts to 2.80 Euros/sample when considering both RBD_1 and RBD_2 targets.

The Spike protein has different hotspots of mutation and deletion; the most likely candidates for immune escape are those within the RBD, such as K417N/T, E484K, and N501Y; N439K, L452R, Y453F and S477N are also likely to have a negative impact on the effectiveness of the vaccination [Knoll et al., 2021; Kerim, 2021; Garcia et al., 2021; Groves et al., 2021; Plant et al., 2021]. Alarmingly, a convergent evolution has led to diverse groupings of those mutations among different clades [Focosi and Maggi, 2021]. Using known sequenced strains, DHPLC is able to rapidly screen a large number of samples, establishing whether they share the same combination of mutations in the RBD or eventually not. In the latter case, additional combinations may be rapidly ascertained in a few minutes by mixing and denaturing at high temperature, allowing renaturing and running in DHPLC with amplicons from known mutation combination. Other possible targets could also be added to the analysis whenever their role in vaccination escape in other domains emerges [McCarthy et al., 2020; Pereira, 2021]. The flexible, cost-effective streamlined approach herein described, which integrates DHPLC with Sanger sequencing and eventually NGS, should be considered to implement surveillance of the variant of concern.

Authors' contributions

Conceptualization, FG, MET; formal analysis, DM, FT, GD; data curation, FG, MET, DM, FT, GD; writing—original draft preparation, FG, MET; writing—review and editing, SP and VS; validation, DM, FT, GD, and VS; supervision, FG and SP; project administration, MET; funding acquisition, VS. All authors have read and agreed to the published version of the manuscript

Conflicts of interest

SP is an employee of ADS Biotec. ADS Biotec retains proprietary technologies to carry out DHPLC. SP supported the study by providing technical support. No financial support was provided by ADS Biotec which could have inappropriately influenced the conclusion of the study.

There are no patents products in development or marketed products to declare. All authors are committed to adhere to the Journal policies on sharing data and materials.

Funding information

Partial financial support was received from AUSL Romagna under the protocol code “COVdPCR”

Ethical approval

The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board of AUSL Romagna under the protocol code “COVdPCR of 07/02/2020. (include appropriate approvals or waivers)

Consent for publication

Not applicable. Informed Consent Statement: The samples included in this study were sent to the Unit of Micro-biology, Greater Romagna Area Hub Laboratory, Cesena, Italy, for routine diagnostic purposes and the laboratory data results were reported as answer to a clinical suspicion. As such, informed consent from patients was not required. Each sample included was preventively anonymized.

Acknowledgements

Some viral strain included in the study had been investigated within the interlaboratory “INSTAND e.V.” external proficiency testing program.

Reference

1. Plante JA, Liu Y, Liu J, Xia H, Johnson BA, Lokugamage KG, Zhang X, Muruato AE, Zou J, Fontes-Garfias CR, Mirchandani D, Scharton D, Bilello JP, Ku Z, An Z, Kalveram B, Freiberg AN, Menachery VD, Xie X, Plante KS, Weaver SC, Shi PY. Spike mutation D614G alters SARS-CoV-2 fitness. *Nature*. 2021 Apr;592(7852):116-121. doi: 10.1038/s41586-020-2895-3.
2. Gómez CE, Perdiguero B, Esteban M. Emerging SARS-CoV-2 Variants and Impact in Global Vaccination Programs against SARS-CoV-2/COVID-19. *Vaccines (Basel)*. 2021 Mar 11;9(3):243. doi: 10.3390/vaccines9030243.
3. Focosi D, Maggi F. Neutralising antibody escape of SARS-CoV-2 spike protein: Risk assessment for antibody-based Covid-19 therapeutics and vaccines. *Rev Med Virol*. 2021 Mar 16. doi: 10.1002/rmv.2231.
4. Plante JA, Mitchell BM, Plante KS, Debbink K, Weaver SC, Menachery VD. The variant gambit: COVID-19's next move. *Cell Host Microbe*. 2021 Mar 1:S1931-3128(21)00099-8. doi: 10.1016/j.chom.2021.02.020.
5. Groves DC, Rowland-Jones SL, Angyal A. The D614G mutations in the SARS-CoV-2 spike protein: Implications for viral infectivity, disease severity and vaccine design. *Biochem Biophys Res Commun*. 2021 Jan 29;538:104-107. doi: 10.1016/j.bbrc.2020.10.109. Epub 2020 Nov 5.
6. Jackson CB, Zhang L, Farzan M, Choe H. Functional importance of the D614G mutation in the SARS-CoV-2 spike protein. *Biochem Biophys Res Commun*. 2021 Jan 29;538:108-115. doi: 10.1016/j.bbrc.2020.11.026. Epub 2020 Nov 13.
7. Li Q, Wu J, Nie J, Zhang L, Hao H, Liu S, Zhao C, Zhang Q, Liu H, Nie L, Qin H, Wang M, Lu Q, Li X, Sun Q, Liu J, Zhang L, Li X, Huang W, Wang Y. The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity. *Cell*. 2020 Sep 3;182(5):1284-1294.e9. doi: 10.1016/j.cell.2020.07.012.

8. Pereira F. SARS-CoV-2 variants combining spike mutations and the absence of ORF8 may be more transmissible and require close monitoring. *Biochem Biophys Res Commun.* 2021 Apr 23;550:8-14. doi: 10.1016/j.bbrc.2021.02.080.
9. Giovanetti M, Benedetti F, Campisi G, Ciccozzi A, Fabris S, Ceccarelli G, Tambone V, Caruso A, Angeletti S, Zella D, Ciccozzi M. Evolution patterns of SARS-CoV-2: Snapshot on its genome variants. *Biochem Biophys Res Commun.* 2021 Jan 29;538:88-91. doi: 10.1016/j.bbrc.2020.10.102.
10. Kirby T. New variant of SARS-CoV-2 in UK causes surge of COVID-19. *Lancet Respir Med.* 2021 Feb;9(2):e20-e21. doi: 10.1016/S2213-2600(21)00005-9.
11. Jung J. Preparing for the Coronavirus Disease (COVID-19) Vaccination: Evidence, Plans, and Implications. *J Korean Med Sci.* 2021 Feb 22;36(7):e59. doi: 10.3346/jkms.2021.36.e59.
12. Knoll MD, Wonodi C. Oxford-AstraZeneca COVID-19 vaccine efficacy. *Lancet.* 2021 Jan 9;397(10269):72-74. doi: 10.1016/S0140-6736(20)32623-4.
13. Karim SSA. Vaccines and SARS-CoV-2 variants: the urgent need for a correlate of protection. *Lancet.* 2021 Apr 3;397(10281):1263-1264. doi: 10.1016/S0140-6736(21)00468-2.
14. Triggler CR, Bansal D, Ding H, Islam MM, Farag EABA, Hadi HA, Sultan AA. A Comprehensive Review of Viral Characteristics, Transmission, Pathophysiology, Immune Response, and Management of SARS-CoV-2 and COVID-19 as a Basis for Controlling the Pandemic. *Front Immunol.* 2021 Feb 26;12:631139. doi: 10.3389/fimmu.2021.631139. eCollection 2021.
15. Cobey S, Larremore DB, Grad YH, Lipsitch M. Concerns about SARS-CoV-2 evolution should not hold back efforts to expand vaccination. *Nat Rev Immunol.* 2021 Apr 1:1-6. doi: 10.1038/s41577-021-00544-9.
16. Garcia-Beltran WF, Lam EC, St Denis K, Nitido AD, Garcia ZH, Hauser BM, Feldman J, Pavlovic MN, Gregory DJ, Poznansky MC, Sigal A, Schmidt AG, Iafrate AJ, Naranbhai V, Balazs

AB. Multiple SARS-CoV-2 variants escape neutralization by vaccine-induced humoral immunity.

Cell. 2021 Mar 12:S0092-8674(21)00298-1. doi: 10.1016/j.cell.2021.03.013.

17. Banada P, Green R, Banik S, Chopoorian A, Streck D, Jones R, Chakravorty S, Alland D. A Simple RT-PCR Melting temperature Assay to Rapidly Screen for Widely Circulating SARS-CoV-2 Variants. medRxiv. 2021 Mar 8:2021.03.05.21252709. doi: 10.1101/2021.03.05.21252709. Preprint.

18. McCarthy KR, Rennick LJ, Nambulli S, Robinson-McCarthy LR, Bain WG, Haidar G, et al. Recurrent deletions in the SARS-CoV-2 spike glycoprotein drive antibody escape. bioRxiv. 2020:2020.11.19.389916. (Preprint)

19. Bal A, Destras G, Gaymard A, Stefic K, Marlet J, Eymieux S, Regue H, Semanas Q, d'Aubarede C, Billaud G, Laurent F, Gonzalez C, Mekki Y, Valette M, Bouscambert M, Gaudy-Graffin C, Lina B, Morfin F, Josset L; COVID-Diagnosis HCL Study Group. Two-step strategy for the identification of SARS-CoV-2 variant of concern 202012/01 and other variants with spike deletion H69-V70, France, August to December 2020. Euro Surveill. 2021 Jan;26(3):2100008. doi: 10.2807/1560-7917.ES.2021.26.3.2100008.

Figure 1: Workflow for the rapid screening of the variants of concern.

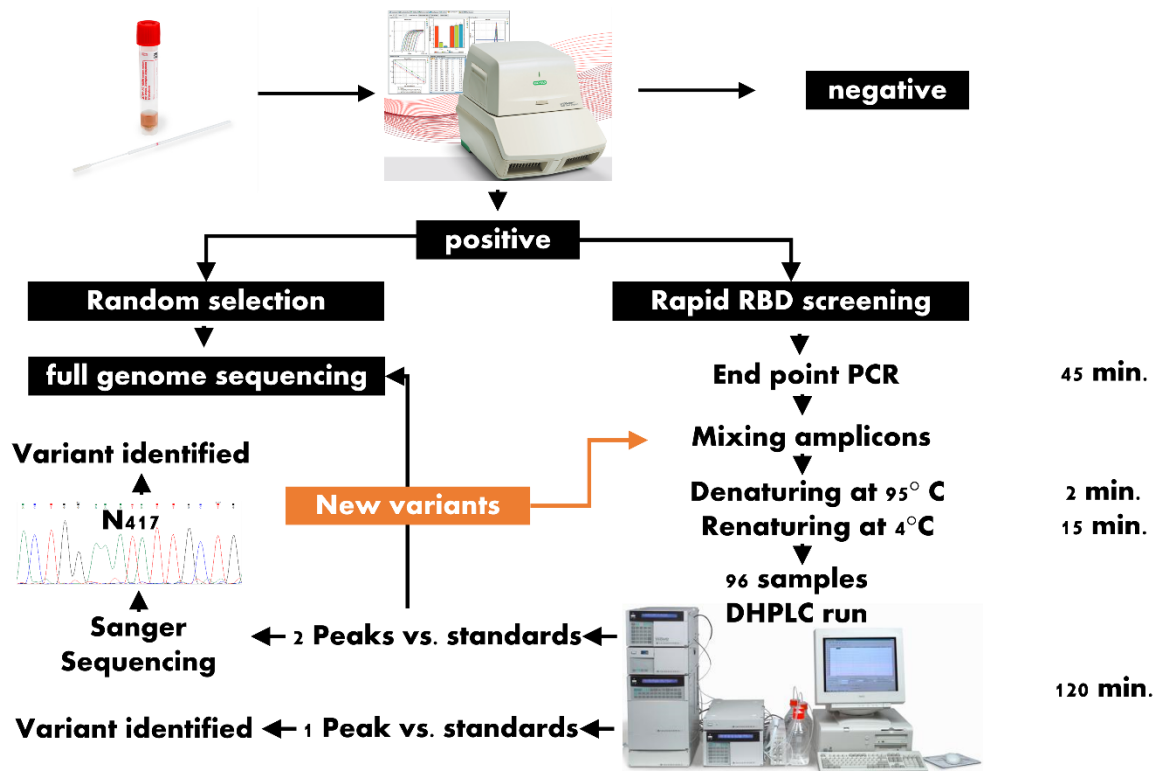


Figure 2: Detection of Y501 residue (B.1.1.7 variant) throughout heteroduplex detection between the region binding domain Target 2 (RBDN_2) amplified from a B.1 strain and a B.1.1.7 strain. The mixing, denaturing at 95°C and renaturing at room temperature between two mutations causes a heteroduplex formation detected by the Wave Denaturing High-Performance Liquid Chromatography system, such as peak splitting in E. A, B, C, D and F represent the typical appearance of homoduplexes.

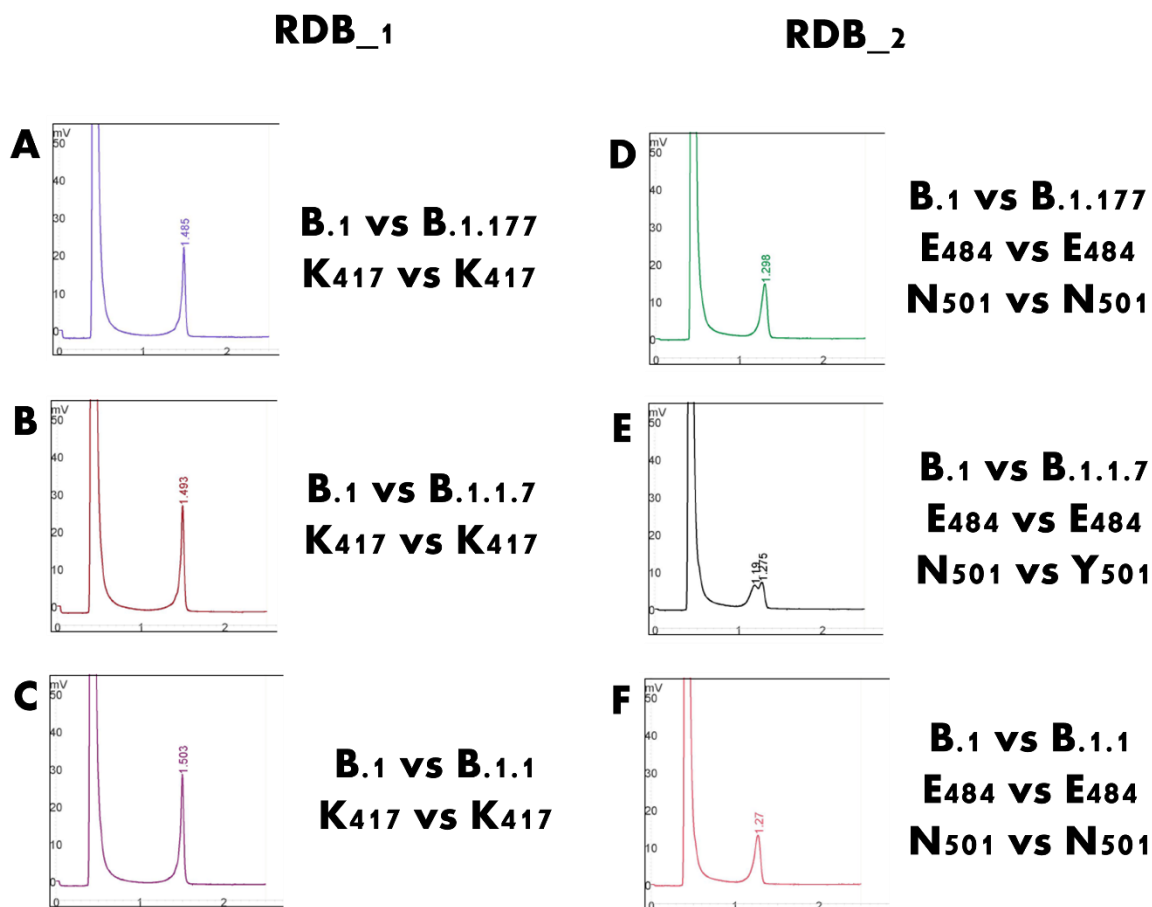
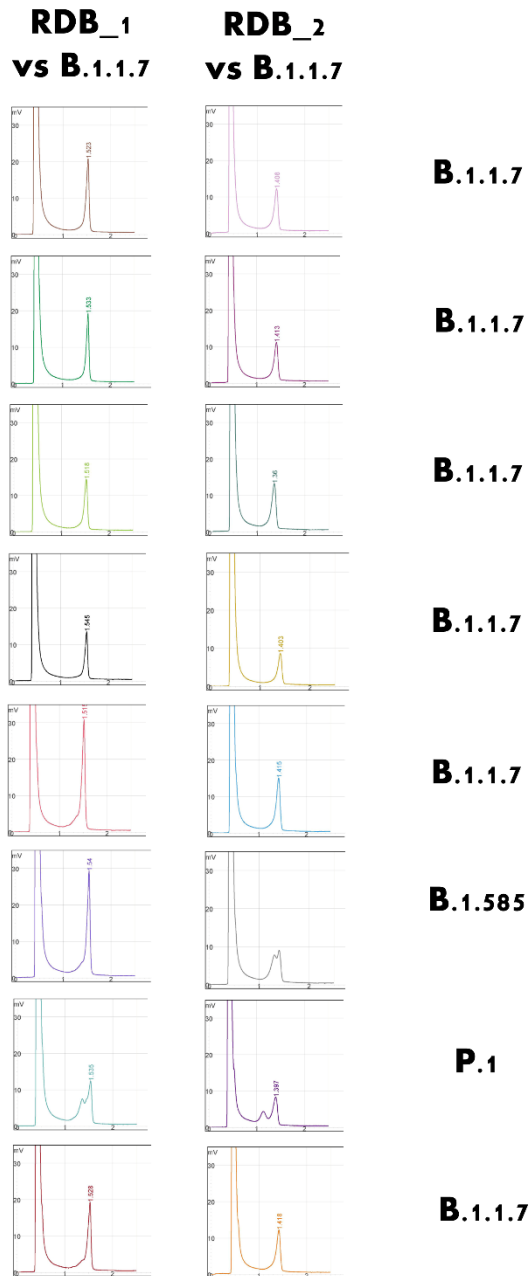


Figure 3: Selected example of the detection of rare variants of concern using the Denaturing High-Performance Liquid Chromatography system. Using the B.1.1.7 strain as a standard, the rarer B.1.585 (Nigerian) and P.1. (Brazilian) variants could be detected and confirmed.





Random selection

full genome sequencing

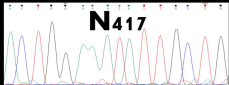


positive

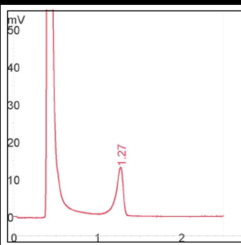
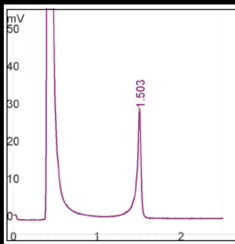
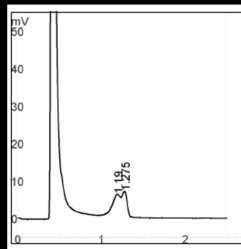
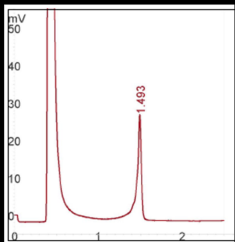
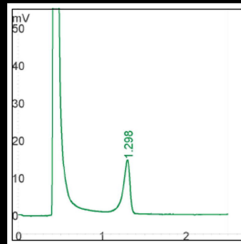
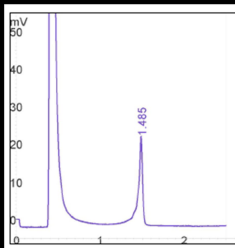
negative

Rapid RBD screening

New variants

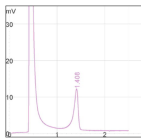
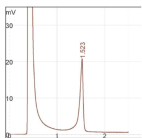


DHPLC run

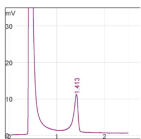
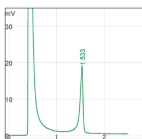


RDB_1 vs B.1.1.7

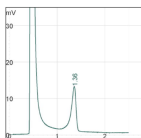
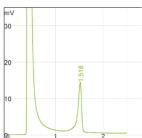
RDB_2 vs B.1.1.7



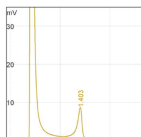
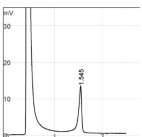
B.1.1.7



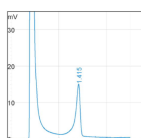
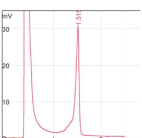
B.1.1.7



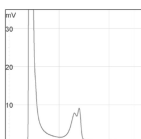
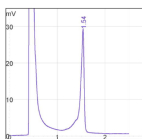
B.1.1.7



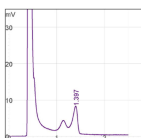
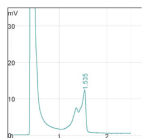
B.1.1.7



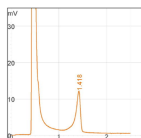
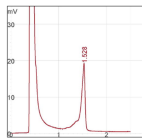
B.1.1.7



B.1.585



P.1



B.1.1.7