Cycle Threshold Probability Score for Immediate and Sensitive Detection of B.1.351 SARS-CoV-2 Lineage

Dieter De Smet
AZ Delta General Hospital
Merijn Vanhee
AZ Delta General Hospital
Brigitte Maes
Jessa Hospital
Koen Swaerts
AZ Delta General Hospital
Peter De Jaeger
AZ Delta General Hospital
Frederik Van Hoecke
AZ Delta General Hospital
Karel Maelegheer
AZ Sint-Lucas
Geert Antoine Martens (geert.martens@azdelta.be)
AZ Delta General Hospital

Research Article

Keywords: SARS-CoV-2, B.1.351 SARS-CoV-2 Lineage, Cycle Threshold, Probability Score, Immediate and Sensitive

Posted Date: June 14th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-558631/v1

Abstract

Background. Detection of SARS-CoV-2 variants of concern associated with immune escape is important to safeguard vaccination efficacy. We describe the potential of delayed N-gene amplification in the Allplex SARS-CoV-2 PCR assay for screening of the B.1.351 lineage.

Methods. In a study cohort of 397 consecutive PCR-positive samples genotyped by whole genome sequencing, amplification curves of E/N/S-RdRP targets indicated delayed N- versus E-gene amplification, characteristic for B.1.351. Logistic regression was used to calculate a VOC.V2 probability score that was evaluated as separate screening test in an independent validation cohort versus sequencing.

Results. B.1.351 showed a proportionally delayed amplification of N- versus E-gene. In logistic regression only N- and E-gene cycle thresholds independently contributed to B.1.351 prediction, allowing calculation of a VOC.V2 probability score with AUC=0.94. At an optimal dichotomous cut-off point of 0.12, VOC.V2 probability score achieved 98.7% sensitivity at 79.9% specificity resulting in 99.6% and 54.6% NPV and PPV respectively. Probability of B.1.351 increased with increasing VOC.V2 probability score, achieving likelihood ratio of 12.01 above 0.5. Near maximal NPV was confirmed in 153 consecutive validation samples.

Conclusion. Delayed N- versus E-gene amplification in the Allplex SARS-CoV-2 PCR assay can be used for fast and highly sensitive screening of B.1.351.

Introduction

In December 2020 an unexpected rise in SARS-CoV-2 infections in the UK was attributed to the emergence of a new SARS-CoV-2 variant of concern (VOC), lineage B.1.1.7 (20I/501.V1, VOC 202012/01), first detected by an efficient national scale genomic surveillance network. B.1.1.7 shows a 40–70% higher transmission rate [1,2], mainly explained by the N501Y amino acid substitution in the Spike-Receptor Binding Domain (S-RBD) that increases the affinity of S-RBD for the human ACE2 receptor 7-to 19-fold [3,4]. B.1.1.7 rapidly became the dominant strain in the UK and many European countries [5]. Its spread could be accurately monitored by an accidental interference by one of its defining mutations, the Spike 69/70 deletion (69.70del) in the widely used TaqPath PCR resulting in a characteristic signature with preserved amplification of N-gene and ORF1ab gene targets but S gene target failure (SGTF or S-dropout). Though recent data indicate that B.1.1.7 is overall 60% more deadly [6,7], it does not appear to display immune escape. B.1.1.7 appears refractory to N-terminal domain (NTD) targeting antibodies, but not to the S-RBD neutralizing antibodies that are induced by all currently approved vaccines [4,8], with generally preserved in vitro neutralization convalescent sera after wild type virus infection and by immune sera after the BNT162b2 (Pfizer/BioNTech) and AZD1222 (AstraZeneca/Oxford) [4,9,10] vaccines.

In an example of convergent evolution, the N501Y mutation independently arose in several other SARS-CoV-2 lineages, such as the B.1.351 (20H/501.V2, variant of concern 2) in South-Africa [11] and the P.1 (20J/501Y.V3, variant of concern 3) in Brazil [12]. Both lineages had higher infectivity than the wild type virus and rapidly achieved regional dominance. In another, more concerning, example of convergent evolution, the B.1.351 and P.1 lineages share the additional E484K Spike mutation that was independently confirmed as powerful driver of clinical immune escape. This partial immune escape explains the resurgence of COVID-19 in Manaus, Brazil by P.1 in a population with 76% seroprevalence [13] and the absence of protective immunity of prior wild type infection to B.1.351 in the placebo arms of the NVX-CoV2373 vaccine trial in South-Africa.

The impact of various SARS-CoV-2 mutations/variants on immune escape was mainly derived from several small-scale in vitro pseudovirus neutralization studies using polyclonal convalescent sera or monoclonal antibodies after wild type (non-variant of concern, non-VOC) infections and polyclonal sera after vaccination with the current first generation S-RBD targeted vaccines. Introduction of the E484K in a B.1.1.7 background leads to a 6-fold lower neutralization by convalescent sera [3,4]. Also, vaccine efficacy is reduced: P.1 shows a moderately (2- to 3-fold) lower neutralization by mRNA-1273 and BNT162b2

vaccine sera [8]. In case of B.1.351 the effect is more pronounced, likely due to additional effects of the NTD-mutations in this lineage, leading to 6 to 12-fold lower neutralization by both vaccines [8,14,15]. These effects are consistently confirmed in vivo. Recent data collected in South-Africa showed a drop of vaccine efficacy to prevent moderate to severe COVID-19 by the B.1.351 lineage as compared to non-VOC strains: from 95.6–50% for the NVX-CoV2373 (Novavax) vaccine, from 72–57% for Ad26.COV2.S (Janssen) and a complete loss of efficacy of the AZD1222 (AstraZeneca/Oxford) (66.7% for all variants to 10.6% for B.1.351) [16].

All data thus indicate that aggressively containing B.1.351 and P.1 variants in populations with low prevalence of these strains is crucial to safeguard the global vaccination strategy using the first-generation vaccines. This requires a combination of several test strategies: intensified genome surveillance by whole genome sequencing (WGS), the flexible introduction of targeted reflex PCR assays for lineage defining mutations in the spike protein (i.e. at position N501, D253, Q677) or mutations associated with potential immune escape (i.e. at position E484, L452, S477,…) depending on regional strain prevalences but also the intelligent use of subtle variations in standard SARS-CoV-2 PCR amplification curves. The latter could allow fast, high-throughput and low-cost screening of variants of concern, as illustrated by the impact of the SGFT/S-gene dropout for B.1.1.7 surveillance [1,17,18], rendering variant screening also available for health care systems with limited resources.

Here we describe a delayed amplification of the N gene-target, characteristic for the B.1.351 variant, in the Allplex Sars-CoV-2 Assay (Seegene, Korea) that is currently used in high volumes in over 70 countries worldwide. The N-gene delay was investigated in a study cohort of SARS-CoV-2 isolates containing a representative number of wild type lineages, B.1.1.7 and B.1.351. A probability score was calculated for presence of B.1.351 variant of concern 2 (VOC.V2 score) based on E/N/S-RdRP cycle threshold (Ct) values. We subsequently studied the diagnostic power of this new VOC.V2 probability score as separate test result in an independent validation cohort, as stand-alone test and in conjunction with N501Y and 69.70del mutation-specific PCRs.

Materials And Methods

Samples

The study was performed on consecutive samples of subjects (in- and outpatients) who underwent SARS-CoV-2 PCR and whole genome sequencing as part of routine diagnostic workup and baseline or active surveillance for variants of concern (VOC) divided in a study cohort and a validation cohort. The study cohort was consisted of 397 unique viral isolates, including 141 501Y.V1 clade (‘UK variant, hence B.1.1.7 lineage), 78 501Y.V2 clade (‘South African variant, hence B.1.351) and 178 samples with other SARS-CoV-2 clade not belonging to any of the three currently recognized VOC clades (hence non-VOC). The study cohort was representative for the regional prevalence as measured by unbiased surveillance by the Belgian National WGS consortium for which AZ Delta acts as a sentinel hub. Non-VOC Pangolin lineages included: B.1.0, B.1.162, B.1.214, B.1.221, B.1.258, B.1.160, B.1.177, B.1.1.222 and B.1.1.29. The validation cohort contained 153 unique viral isolates: 60 non-VOC, 57 B.1.1.7, 29 B.1.351 and 7 20J/501Y.V3 clade (‘Brazilian variant’, hence P.1). For detailed lineage composition of study and validation cohorts see Supplementary Table 1. The study was performed on analytical leftover samples and secondary use of data and samples obtained during standard of care diagnostic workup. It was performed according to the Helsinki declaration with full respect for individuals’ right to confidentiality and with a waiver of informed consent for COVID-19 related diagnostics and modeling, approved by the AZ Delta ethical committee (Clinical Trial Number IRB B1172020000008).

RNA extraction, RT-PCR for detection of SARS-CoV-2 and reflex PCR for N501Y and 69.70del

RNA was extracted from nasopharyngeal swabs using the STARMag 96x4 Viral DNA/RNA 200 C kit on Hamilton Starlet, followed by real-time PCR using the Allplex SARS-CoV-2 Assay (Seegene Inc, Korea). PCR amplification was run on a CFX96 real-time thermocycler (Biorad) and data were analyzed with the SARS-CoV-2 Viewer (Seegene). The assay simultaneously

detects 4 different SARS-CoV-2 genes, resulting in separate cycle threshold (Ct) values for the E and N gene, and one combined Ct value for the RdRp and S gene (RdRp/S). (raw data in Supplementary Table 1). For a subset of PCR-positive samples in the study and validation cohorts (raw data in Supplementary Table 2), two key mutations for VOC in the Spike protein were detected by reflex testing using Novaplex SARS-CoV-2/H-UK (N501Y & 69-70del) Assay (Seegene). Therefore, a new RNA extraction was performed on the original sample, using the STARMag 96x4 Universal kit, followed by manual PCR setup and amplification on a CFX96 (Biorad). Reflex PCR targets the N501Y amino acid substitution (A23063T) characteristic for VOC.V1-V3 and the HV 69–70 deletion (21765–21770 deletion, hence 69.70del) that, in presence of N501Y mutation, is highly suggestive for B.1.1.7. A conserved sequence in the RdRp gene is co-amplified as internal control, and negative results for N501Y/69.70del were only considered true negative if the RdRp internal control had Ct value < 33. For multicenter inter-laboratory validation, a secondary data analysis was performed on a second validation set of N = 308 consecutive samples analyzed by an independent laboratory (AZ Sint-Lucas Brugge Hospital) with Allplex SARS-CoV-2 Assay (Seegene Inc, Korea) using an alternative RNA extraction on Maelstrom 9600 (Taiwan Advanced Nanotech Inc, TANBead) followed by thermal cycling on CFX96 (Biorad).

SARS-CoV-2 whole genome sequencing and clade calling

As reference method for genomic fingerprinting of viral clades, WGS was performed using the Research Use Only AmpliSeq for Illumina SARS-CoV-2 Research Panel on Illumina MiSeq (40 samples on MicroV2 flow cell) according to the manufacturer's standard protocol: 7 µl RNA was reverse transcribed using Ampliseq cDNA synthesis for Illumina, followed by amplification of 237 virus specific amplicons covering > 99% of the 30kb reference genome aiming at a median coverage above 200x, minimal coverage for mutation calling of 10x and a minimum of 30,000 reads per sample and a maximum of 1kb bases below minimal coverage. A consensus sequence was constructed using an in-house pipeline containing Trimomatic for trimming, alignment by Burrows-Wheeler Aligner, mutation calling by Freebayes and inspection of sequence quality by IGV. Clade calling on the consensus FASTA was done by both Pangolin lineage assignment (v2.3.2, https://pangolin.cog-uk.io) and Nextclade (v0.14.1, https://clades.nextstrain.org) webtools and only concordant clade calling results were used for further analysis (Supplementary Table 1). All sequences were uploaded to GISAID.

Statistics

Statistical analysis was performed using Medcalc (software version 12.2.1, Mariakerke, Belgium). Data are expressed as medians (25th -75th percentiles) and Mann-Whitney U test was used to test statistical difference between groups. Confidence intervals for Likelihood Ratios (LR) were calculated with the method described by Simel et al.[7] Comparisons of Ct values within and between sample subsets were done using Spearman's rank correlation and Passing-Bablok regression analysis. Concordance analysis of integrated results for N501Y/69.70del reflex PCR and VOC.V2 probability scores with WGS was done by calculation of Cohen's kappa agreement coefficient (linear weights).

Different prediction models were trained using the Python Scikit-Learn library, i.e., Logistic Regression, Support Vector Machines, Random Forrest Trees, Bagging Classifier and Gaussian Naieve Bayes. Best accuracy and AUC was achieved with Logistic Regression (overview of accuracy and AUC of various tested models in Supplementary Table 3). The Logistic Regression classifier was validated and further adopted in the Medcalc statistical tool, which offers the benefit of interpretability. The cut-off for the Logistic Regression model was optimized to yield optimal recall for detecting the presence of the B.1.351 variant.

Results

Distinct E, N and S/RdRP cycle threshold profiles between SARS-CoV-2 viral lineages

Retrospective analysis of the study cohort (N = 397) indicated variations in the distribution profiles of E/N/S-RdRP Ct values between non-VOC, B.1.1.7 and B.1.351 (Fig. 1A, Table 1). In all clades except B.1.351, the N gene amplified earlier with left-

shifted Ct N versus Ct S/RdRP and Ct E (P < .05). B.1.351 was distinct from all other clades by showing a right-shifted (delayed) N gene amplification in the Allplex SARS-CoV-2 Assay. In the independent validation cohort (N = 153), which also contained 7 P.1 isolates, this delayed N gene amplification was confirmed as characteristic feature of the B.1.351 lineage (Fig. 1B, Table 1). Correlation analyses of Ct values indicated an identical amplification profile of E and S/RdRP in all strains (Fig. 1C, Table 2) and graphically illustrate that the delayed N amplification versus both S/RdRP (Fig. 1D) and E (Fig. 1E) in the B.1.351 lineage is clearly proportional (slope = 1.21, 95%CI 1.13–1.28) for N versus E (Table 2), indicating a gradually more pronounced N-gene amplification delay at lower viral loads.

Combined Use of Ct Values for detection of B.1.351 lineage: VOC.V2 probability score

Next, logistic regression analysis was done to investigate if Ct values could predict presence of the B.1.351 lineage in individual patient samples in the study cohort resulting in a fast and inexpensive screening tool. In a model including all three PCR targets, Ct E (P < .0001) and Ct N (P < .0001) but not S/RdRP (P = .5662) independently contributed to prediction of B.1.351 infection. Repeating regression analysis with only Ct E and Ct N indicated that a lower Ct E value and a higher Ct N value had strong diagnostic power to detect B.1.351 (AUC = 0.94, 95% CI 0.91–0.96). It allowed the calculation of a 'VOC.V2 probability score' (p) for each individual sample using the following formula: $\ln(p/[1-p]) = 0.9449 + 1.44091 \times Ct_{Ngene} - 1.51860 \times Ct_{Egene}$ (Supplementary Table 3). The distribution of this VOC.V2 probability score for all non-VOC clades, B.1.1.7 and B.1.351 is plotted in Fig. 2A. As expected, B.1.351 isolates had significantly higher VOC.V2 probability scores (P < .0001), although partially overlapping with other lineages.

The VOC.V2 probability score was further investigated as independent single test result and its diagnostic power for individual patient samples evaluated. ROC analysis (Fig. 2B) showed an AUC (95% CI) of 0.94 (0.91–0.96) with optimal cut-point of 0.12 (Table 3) for dichotomous test use resulting in a sensitivity of 98.7% (95% CI 93.1–100.0) and specificity of 79.9% (95% CI 75.1–84.2) for screening of B1.351 lineage. At a prevalence (pretest probability) of 19.6% of B.1.351 in our study cohort, this resulted in a very high negative predictive value (NPV) of 99.6% and a positive predictive value (PPV) of 54.6% (Table 3).

Calculation of LRs for different VOC.V2 probability score result intervals showed that the probability of SARS-CoV-2 B.1.351 variant increased with increasing score (Fig. 2C, Table 3). A score ≤ 0.12 almost completely ruled out B.1.351 with LR (95% CI) of 0.02 (0.00–0.11) while scores above 0.50 strongly increased probability with LRs (95% CI) of 12.01 (7.21–20.02). Relations of posttest and pretest probabilities for the proposed VOC.V2 probability score intervals are shown in Fig. 2D, illustrating the score's power to detect B.1.351 taking the prevailing prevalence of this lineage in the population as pretest probability. The clinical utility of the LR was summarized in a diagnostic algorithm in Fig. 2E.

Validation of VOC.V2 probability score in an independent cohort

The VOC.V2 probability score calculated from the study cohort was then tested in an independent validation cohort of N = 153 consecutive samples analyzed by WGS and PCR. This validation cohort contained a similar prevalence (Table 1) of B.1.1.7 (39.2%) and B.1.351 (19.0%) but also contained 7 (4.6%) P.1 isolates. Overall diagnostic performance was identical, with 100% NPV and 53.7% PPV using the optimal cut-point of 0.12, and increasing probability of B.1.351 screening with increasing VOC.V2 probability score, attaining a LR (95% CI) of 17.82 (8.05–39.41) at scores above 0.50.

Integration of VOC.V2 probability score with reflex PCR for N501Y and 69.70del for rapid cost-effective VOC detection

A subset (N = 172) of study and validation cohort samples was additionally analyzed by targeted PCR for N501Y and 69.70del mutations in the Spike protein (Supplementary Table 2). As shown in Fig. 3A, the combined use of the N amplification delay in the basic E/N/S-RdRP PCR assay with these two biologically relevant mutations allowed for a rapid

and cost-effective discrimination of the three variants of concern in our population (B.1.1.7, B.1.351 and P.1). All three share the N501Y mutation. Additional presence of 69.70del indicates B.1.1.7. Strains carrying N501Y but not 69.70del can be discerned as either B.1.351 or P.1 based on VOC.V2 probability score above or below 0.12, respectively. Figure 3B shows concordance of this simplified diagnostic approach versus the gold standard of WGS, with excellent agreement (Weighted Kappa (95% CI) = 0.987 (0.969-1.000)).

Multicenter validation

To evaluate the inter-laboratory robustness of the VOC.V2 probability score, an additional validation was performed on Allplex SARS-CoV-2 assay PCR Ct E/S-RdRP/N data obtained from an independent laboratory using an alternative RNA extraction setup. In a cohort containing N = 308 consecutive SARS-CoV-2 PCR-positive samples (N = 68 B.1.351, N = 73 non-VOC lineages and N = 167 B.1.1.7 isolates as measured by WGS) (Supplementary table 4), the VOC.V2 probability score achieved a sensitivity of 100% (95% CI 94.7–100%) at 60.8% (95% CI 54.4–67.1%) specificity using binary thresholding with single cut-off of 0.12. All B.1.351 isolates showed a VOC.V2 probability score above 0.5, a cut-off associated with a LR (95%CI) of 12.6 (8.2–19.5).

Discussion

Here we describe a VOC.V2 probability score calculated by regression modeling of the Ct values of the widely used Allplex (Seegene) SARS-CoV-2 PCR assay. As a standalone test, this derived score can screen for infection by the South-African B.1.351 (20H/501.V2) variant of concern, with a sensitivity approaching 100%, and thus selecting relevant samples for reflex-testing by mutation-specific PCRs or WGS.

The possible impact of our finding on pandemic control is significant: *(i)* the VOC.V2 probability score provides an immediate suggestion for suspected B.1.351 infection, based on the Ct results of the standard PCR (typically within 12-24h after sample reception), thus avoiding a delay of 24h to several days as compared to mutation-specific PCR and WGS, respectively. *(ii)* Fast detection of B.1.351 can improve infection containment which is crucial considering the proven immune escape of this strain towards immunity by prior infection or S-RBD-based vaccines. *(iii)* The VOC.V2 probability score can easily be implemented at no additional cost for labs using the widespread Allplex SARS-CoV-2 assay after simple regression modeling on their Ct data set. The underlying N-gene amplification delay in B.1.351 isolates is thus comparable to a similarly informative artefact, the SGFT in the TaqPath PCR assay that formed the basis of a cost-effective two-step screening approach for B.1.1.7 implemented in many countries and yielding dense epidemiological data [17,18]. In more general terms, our approach illustrates that additional diagnostic information can be extracted from raw PCR data and outlines an approach that can be iterated for other PCR assays, provided that a lab has access to a gold standard method for viral genotyping. The continuous emergence of new SARS-CoV-2 variants of concern urges labs all over the world to implement tools for rapid differentiation between circulating variants.

The strengths of our study are (i) its use of independent training and validation cohorts, both sufficiently sized and consisting of consecutive and representative samples, (ii) d the use of WGS as gold standard to unambiguously determine clinical truth (exact viral lineage) and (iii) the subsequent validation of our VOC.V2 probability score on PCR data obtained from an independent laboratory with a different RNA extraction setup, highlighting that the model is highly robust and can also be implemented on Allplex SARS-CoV-2 PCR assay Ct data from other labs, generated by other types of liquid handlers and/or thermal cyclers. The only limitation of our study is that its findings are directly transferable only to lab platforms using the Allplex SARS-CoV-2 assay, similarly to the limitation of the SGFT feature in the Taqpath assay. However, an overarching message of our study is that it demonstrates that intelligent analysis of subtle changes in PCR Ct-values can be of great use, additional to or even replacing strain-specific mutation PCR assays.

In conclusion, our study provides a fast and cost-effective screening tool for the concerning B.1.351 SARS-CoV-2 variant based on secondary use of data from a widely used SARS-CoV-2 PCR. More generally, it outlines a straightforward strategy to

create more diagnostic value from available data. This tool can be of help to monitor and hopefully contain the spread of the immune escape-prone B.1.351 strain.

Abbreviations

B.1.1.7 (20I/501.V1, variant of concern 1, VOC.V1, VOC 202012/01, UK SARS-CoV-2 variant), B.1.351 (20H/501.V2, variant of concern 2, VOC.V2, South-African SARS-CoV-2 variant); Ct: Cycle Threshold value; NPV: negative predictive value, LR: likelihood ratio; PCR: Polymerase Chain Reaction; P.1 (20J/501Y.V3, variant of concern 3, VOC.V3, Brazilian SARS-CoV-2 variant); PPV: positive predictive value; VOC: variant of concern; WGS: whole genome sequencing

Declarations

Disclosure statement: The authors declare no conflict of interest. This work was supported by a private donation by board members of Fagron (Nazareth, Belgium), a healthcare company, to RADar, the teaching and education initiative of AZ Delta General Hospital, to be used as unconditional research grant for data collection, collaborative collaboration and open access publication. The sponsor had no influence on the study design, data interpretation and drafting of the manuscript.

Trial registration: Clinical Trial Number IRB B1172020000009

References

1. Davies, N. G. *et al.* Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science*, doi:10.1126/science.abg3055 (2021).
2. Volz, E. *et al.* Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. *Nature*, doi:10.1038/s41586-021-03470-x (2021).
3. Collier, D. A. *et al.* Sensitivity of SARS-CoV-2 B.1.1.7 to mRNA vaccine-elicited antibodies. *Nature*, doi:10.1038/s41586-021-03412-7 (2021).
4. Supasa, P. *et al.* Reduced neutralization of SARS-CoV-2 B.1.1.7 variant by convalescent and vaccine sera. *Cell*, doi:10.1016/j.cell.2021.02.033 (2021).
5. Hodcroft., E. B. "CoVariants: SARS-CoV-2 Mutations and Variants of Interest. *https://covariants.org/* **Accesses March 30, 2021** (2021).
6. Davies, N. G. *et al.* Increased mortality in community-tested cases of SARS-CoV-2 lineage B.1.1.7. *Nature*, doi:10.1038/s41586-021-03426-1 (2021).
7. Challen, R. *et al.* Risk of mortality in patients infected with SARS-CoV-2 variant of concern 202012/1: matched cohort study. *BMJ* **372**, n579, doi:10.1136/bmj.n579 (2021).
8. Wang, Z. *et al.* mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. *Nature*, doi:10.1038/s41586-021-03324-6 (2021).
9. Liu, Y. *et al.* Neutralizing Activity of BNT162b2-Elicited Serum. *N Engl J Med*, doi:10.1056/NEJMc2102017 (2021).
10. Muik, A. *et al.* Neutralization of SARS-CoV-2 lineage B.1.1.7 pseudovirus by BNT162b2 vaccine-elicited human sera. *Science* **371**, 1152-1153, doi:10.1126/science.abg6105 (2021).
11. Tegally, H. *et al.* Emergence of a SARS-CoV-2 variant of concern with mutations in spike glycoprotein. *Nature*, doi:10.1038/s41586-021-03402-9 (2021).
12. Faria, N. R. *et al.* Genomics and epidemiology of a novel SARS-CoV-2 lineage in Manaus, Brazil. *medRxiv : the preprint server for health sciences*, doi:10.1101/2021.02.26.21252554 (2021).
13. Sabino, E. C. *et al.* Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. *Lancet* **397**, 452-455, doi:10.1016/S0140-6736(21)00183-5 (2021).

14. Wu, K. *et al.* mRNA-1273 vaccine induces neutralizing antibodies against spike mutants from global SARS-CoV-2 variants. *bioRxiv*, doi:10.1101/2021.01.25.427948 (2021).
15. Wibmer, C. K. *et al.* SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. *Nat Med*, doi:10.1038/s41591-021-01285-x (2021).
16. Madhi, S. A. *et al.* Efficacy of the ChAdOx1 nCoV-19 Covid-19 Vaccine against the B.1.351 Variant. *N Engl J Med*, doi:10.1056/NEJMoa2102214 (2021).
17. Bal, A. *et al.* 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 surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin* **26**, doi:10.2807/1560-7917.es.2021.26.3.2100008 (2021).
18. England, P. H. Investigation of novel SARS-CoV-2 variant: Varian of Concern 202012/01: technical briefing 1, accessed March 30, 2021. (2021).

Tables

Table 1. Cycle Threshold (Ct) characteristics of distinct SARS-CoV-2 lineages in Allplex SARS-CoV-2 PCR assay.

Patient group	Characteristic	Non-VOC	B.1.1.7 (20I/501Y.V1)	B.1.351 (20H/501Y.V2)	P.1 (20J/501Y.V3)
Study cohort (n = 397)					
	Number (%)	178 (44.8)	141 (35.5)	78 (19.6)	-
	Ct E gene, median (IQR)	18.7 (16.8-21.2)	20.6 (18.0-24.9) [a]	19.1 (16.8-21.9) [c]	-
	Ct S/RdRP gene, median (IQR)	18.4 (16.2-20.6)	20.7 (18.3-24.8) [a]	18.5 (16.5-21.3) [c]	-
	Ct N gene, median (IQR) [d]	17.3 (14.9-19.6) [d,e]	18.6 (15.9-23.1) [a,d,e]	20.1 (17.4-22.7) [b]	-
Validation cohort (n = 148)					
	Number (%)	60 (39.2)	57 (37.3)	29 (19.0)	7 (4.6)
	Ct E gene, median (IQR)	20.2 (16.8-23.4)	20.5 (18.1-24.2)	20.5 (18.2-25.0)	18.6 (18.2-23.3)
	Ct S/RdRP gene, median (IQR)	20.3 (16.6-23.3)	20.6 (18.5-25.0)	20.6 (17.5-25.0)	18.4 (17.6-22.2)
	Ct N gene, median (IQR)	19.0 (15.3-22.4)	18.4 (15.8-22.1) [d,e]	22.3 (19.4-26.5) [b,c,f]	17.0 (15.8-21.9)

[a] Indicates differences of VOC.V1 with non-VOC for which P values less than .05 were considered statistically significant.

[b] Indicates differences of VOC.V2 with non-VOC for which P values less than .05 were considered statistically significant.

[c] Indicates differences of VOC.V2 with VOC.V1 for which P values less than .05 were considered statistically significant.

[d] Indicates differences of Ct N gene with Ct E gene for which P values less than .05 were considered statistically significant.

[e] Indicates differences of Ct N gene with Ct S/RdRP gene for which P values less than .05 were considered statistically significant.

[f] Indicates differences of VOC.V2 with VOC.V3 for which P values less than .05 were considered statistically significant.

Table lists prevalences of the indicated SARS-CoV-2 clades in the Study cohort (upper panel) and Validation cohort (lower panel) and their median (interquartile range, IQR, 25^{th}-75^{th} percentiles) cycle threshold (Ct) values in the Allplex SARS-CoV-2 assay. From left to right: all non-variant of concern (non-VOC) lineages combined (detailed composition in Suppl. Table 1), the B.1.1.7 'UK variant' (20I/501Y.V1 clade, VOC.V1), the 'South-African' B.1.351 lineage (20H/501Y.V2 clade, VOC.V2) and the Brazilian P.1 (20J/501Y.V3, VOC.V3). Letters indicate corresponding statistical differences as calculated by Mann-Whitney U testing.

Table 2. Correlation and regression analysis of Ct values for E gene, S/RdRP and N gene within different SARS-CoV-2 lineages in the study cohort.

Variant strain	Ct E gene (Y) vs Ct S/RdRP gene (X)			Ct N gene (Y) vs Ct S/RdRP gene (X)			Ct N gene (Y) vs Ct E gene (X)		
	r_s (95%CI)	Slope (95%CI)	Intercept (95%CI)	r_s (95%CI)	Slope (95%CI)	Intercept (95%CI)	r_s (95%CI)	Slope (95%CI)	Intercept (95%CI)
Non-VOC (n=178)	0.99 (0.98 to 0.99)	0.98 (-1.08 to 1.08)	0.78 [b] (0.21 to 1.30)	0.88 (0.85 to 0.91)	1.08 (1.00 to 1.15)	-3.17 [b] (-4.54 to -1.71)	0.91 (0.88 to 0.93)	1.07 [a] (1.02 to 1.13)	-3.27 [b] (-4.41 to -2.19)
B.1.1.7 (20I/501Y.V1) (n=141)	0.98 (0.97 to 0.99)	0.96 [a] (0.93 to 0.99)	0.33 (-0.26 to 0.96)	0.97 (0.96 to 0.98)	1.00 (0.96 to 1.05)	-2.56 [b] (-3.66 to -1.63)	0.98 (0.98 to 0.99)	1.06 [a] (1.03 to 1.10)	-3.46 [b] (-4.29 to -2.74)
B.1.351 (20H/501Y.V2) (n=78)	0.99 (0.98 to 0.99)	0.99 (0.95 to 1.03)	0.52 (-0.35 to 1.28)	0.93 (0.90 to 0.96)	1.21 [a] (1.11 to 1.32)	-3.05 [b] (-4.78 to -1.11)	0.95 (0.92 to 0.97)	1.21 [a] (1.13 to 1.28)	-3.26 [b] (-4.53 to -1.77)

[a] Significant PROPORTIONAL difference.

[b] Significant SYSTEMATIC difference.

Correlation analysis of Ct E-gene (left panel) and N-gene (right panel) versus combined Ct of S/RdRP targets in the Study cohort (N=397) indicates that the N-gene amplification delay in B.1.351 lineage is proportional, with more pronounced N-gene amplification delay at higher Ct (lower viral load) (slope = 1.21 with confidence interval, CI, excluding 1).

Table 3. Diagnostic performance of VOC.V2 probability score in the Study and Validation cohorts using single cut-off or multiple result intervals.

VOC.V2 probability score	B.1.351	All other lineages combined	LR (95%CI)	Sn, % (95%CI)	Sp, % (95%CI)	PPV, % (95%CI)	Accuracy, % (95%CI)
Study cohort (n = 397)	n=78	n=319					
Single cut-off (dichotomous)							
< 0.12	1	255	0.02 (0.00 to 0.11)	98.7 (93.1 to 100.0)	79.9 (75.1 to 84.2)	54.6 (46.0 to 63.0)	83.6 (79.7 to 86.9)
≥ 0.12	77	64	4.92 (3.95 to 6.13)				
Multiple result intervals							
0.00 – 0.12	1	255	0.02 (0.00 to 0.11)	-	-	-	-
0.12 – 0.25	9	25	1.47 (0.72 to 3.03)				
0.25 – 0.50	21	23	3.73 (2.18 to 6.39)				
0.50 – 1.00	47	16	12.01 (7.21 to 20.02)				
Validation cohort (n = 153)	n=29	n=124					
Single cut-off (dichotomous)							
< 0.12	0	99	0.00 (0.00 to 0.34)	100.0 (88.1 to 100.0)	79.8 (71.7 to 86.5)	53.7 (39.6 to 67.4)	83.6 (77.0 to 88.7)
≥ 0.12	29	25	4.96 (3.49 to 7.04)				
Multiple result intervals							
0.00 – 0.12	0	99	0.00 (0.00 to 0.34)	-	-	-	-
0.12 – 0.25	1	11	0.39 (0.05 to 2.89)				
0.25 – 0.50	3	8	1.60 (0.45 to 5.68)				
0.50 – 1.00	25	6	17.82 (8.05 to 39.41)				

LR, likelihood ratio; Sn, sensitivity; Sp, specificity; PPV, positive predictive value; NPV, negative predictive value.

Diagnostic performance of the VOC.V2 probability score as a separate test result for detection of B.1.351 lineage as compared to all other lineages combined (non-VOC, B.1.1.7 and P.1) was assessed using a dichotomous single-cutoff approach or by using likelihood ratios (LR, 95% confidence intervals, CI) for the indicated multiple results intervals in the Study (upper panel) and Validation (lower panel) cohorts. For dichotomous use, from left to right: Sn: Sensitivity; Sp:

specificity; PPV: positive predictive value; NPV: negative predictive value each with 95% Confidence Interval (CI) with PPV and NPV based on the observed prevalences of B.1.351 in Study and Validation cohorts respectively.

Figures

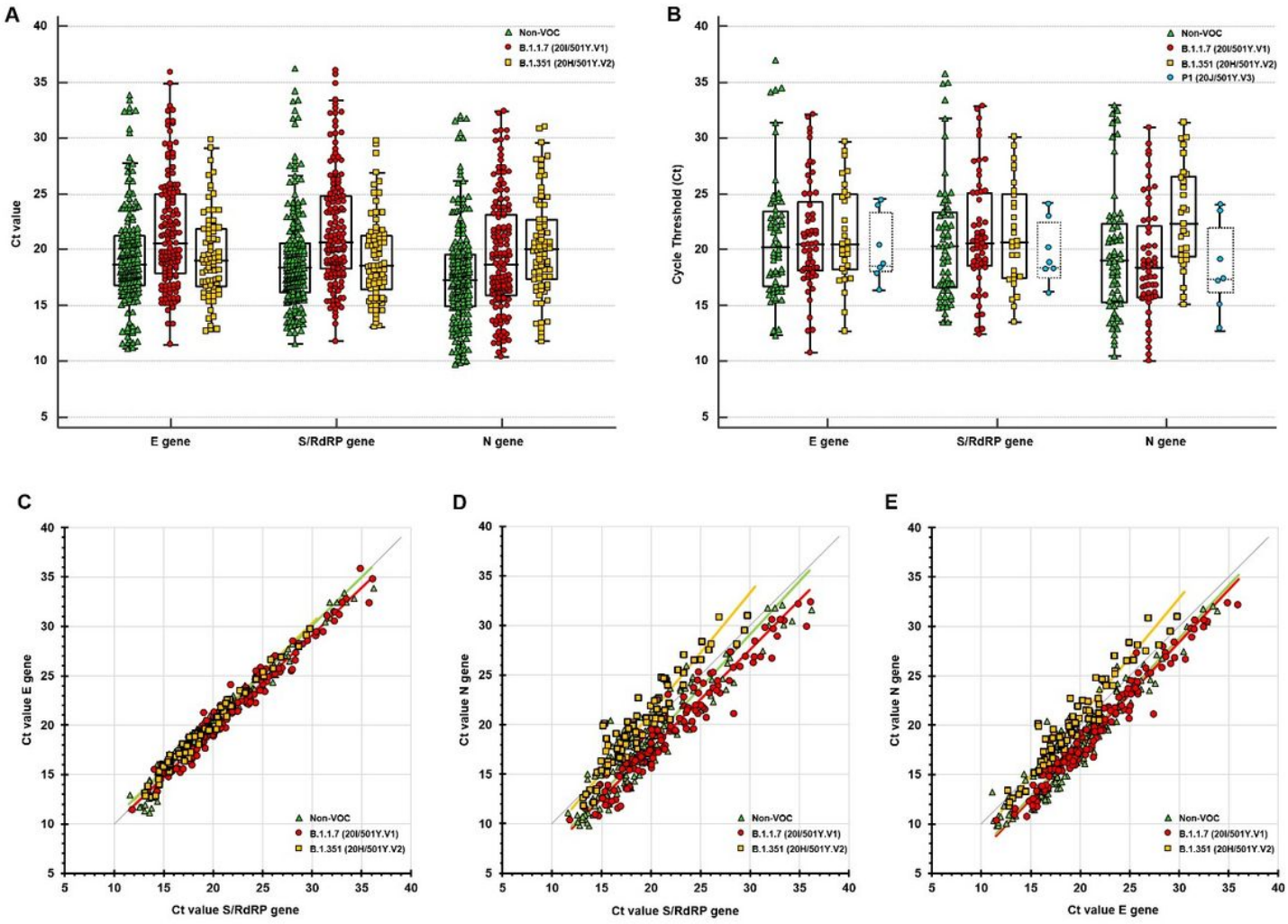

Figure 1

Distribution of Allplex SARS-CoV-2 PCR E/N/S-RdRP Ct values in B.1.1.7, B.1.351 and all non-variant of concern lineages combined. A-B: plotted distribution of raw Cycle threshold (Ct) values (X-axis) of the E-gene, S-RdRP and N-gene targets in the Study (A, N=397) and Validation (B, N=153) cohort in SARS-CoV-2 lineages genotyped by whole genome sequencing: B.1.1.7 (red circles), B.1.351 (yellow squares), P.1 (blue circles, only present in Validation cohort) and all other, non-variant of concern lineages combined (green triangles). C-E: Correlation analysis (see also Table 2) of Ct values of E versus S-RdRP channel (C), N versus S-RdRP channel (D) and N versus E channel (D) in the indicated SARS-CoV-2 lineages.

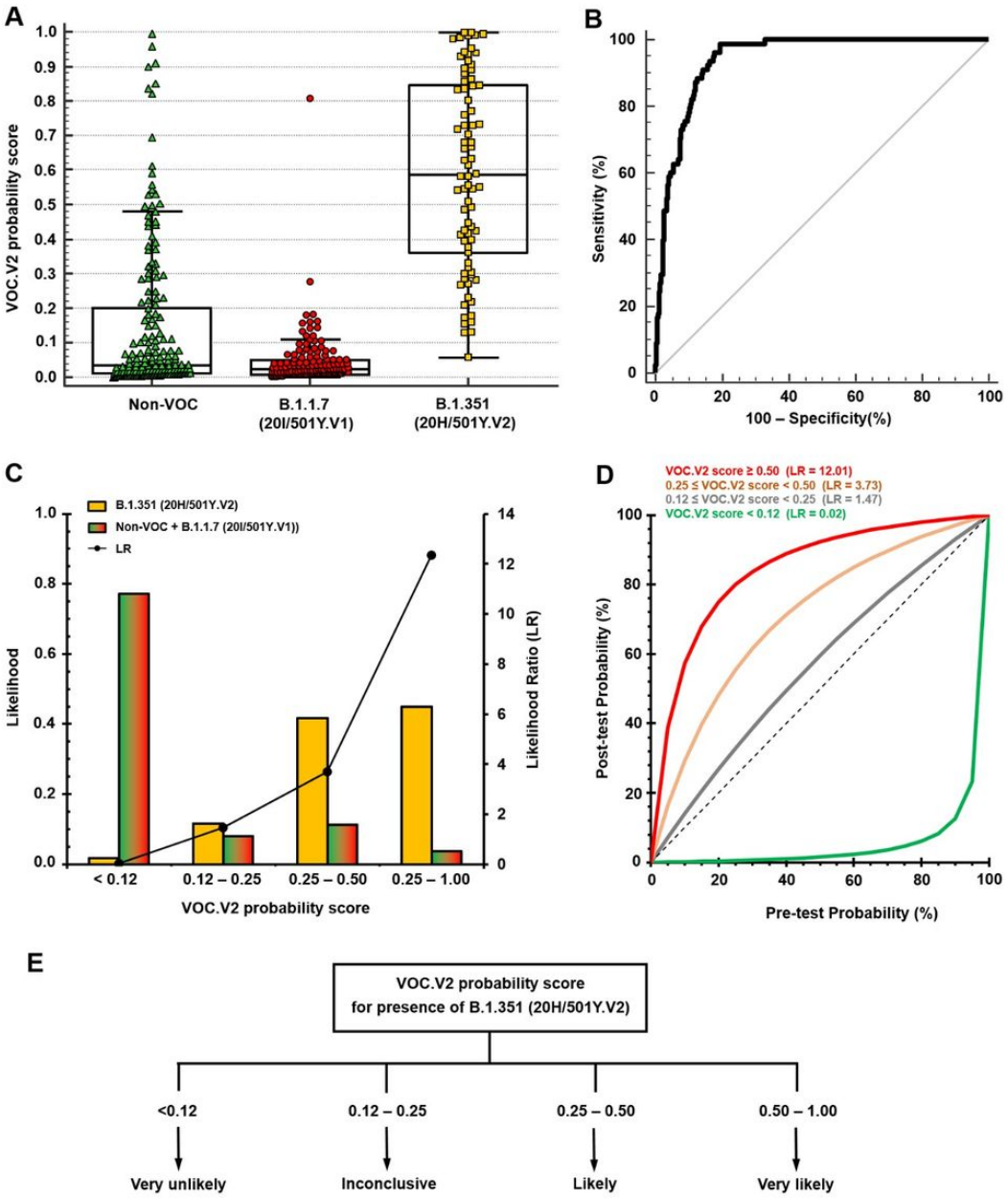

Figure 2

Diagnostic performance of VOC.V2 probability score for detection of B.1.351 variant versus all other lineages A. Distribution of calculated VOC.V2 probability scores of non-VOC samples combined (n=178, green squares), B.1.1.7 samples (n=141, red circles) and B.1.351 samples (n=78, yellow squares) in the Study cohort. B. Receiver operating characteristic curve of VOC.V2 probability score for detection of B.1.351 versus all other lineages combined (see also Table 3). C. Likelihood (left Y-axis) of the indicated result intervals of VOC.V2 probability scores given status B.1.351 (yellow bar) versus status non-VOC or B.1.1.7 combined (mixed green-red bar). Likelihood ratios are shown on the right Y-axis. D. Posttest probability for B.1.351 as a function of pretest probability and of the result interval of the VOC.V2 probability score. E. Summary of diagnostic algorithm of VOC.V2 probability score for screening of B.1.351.

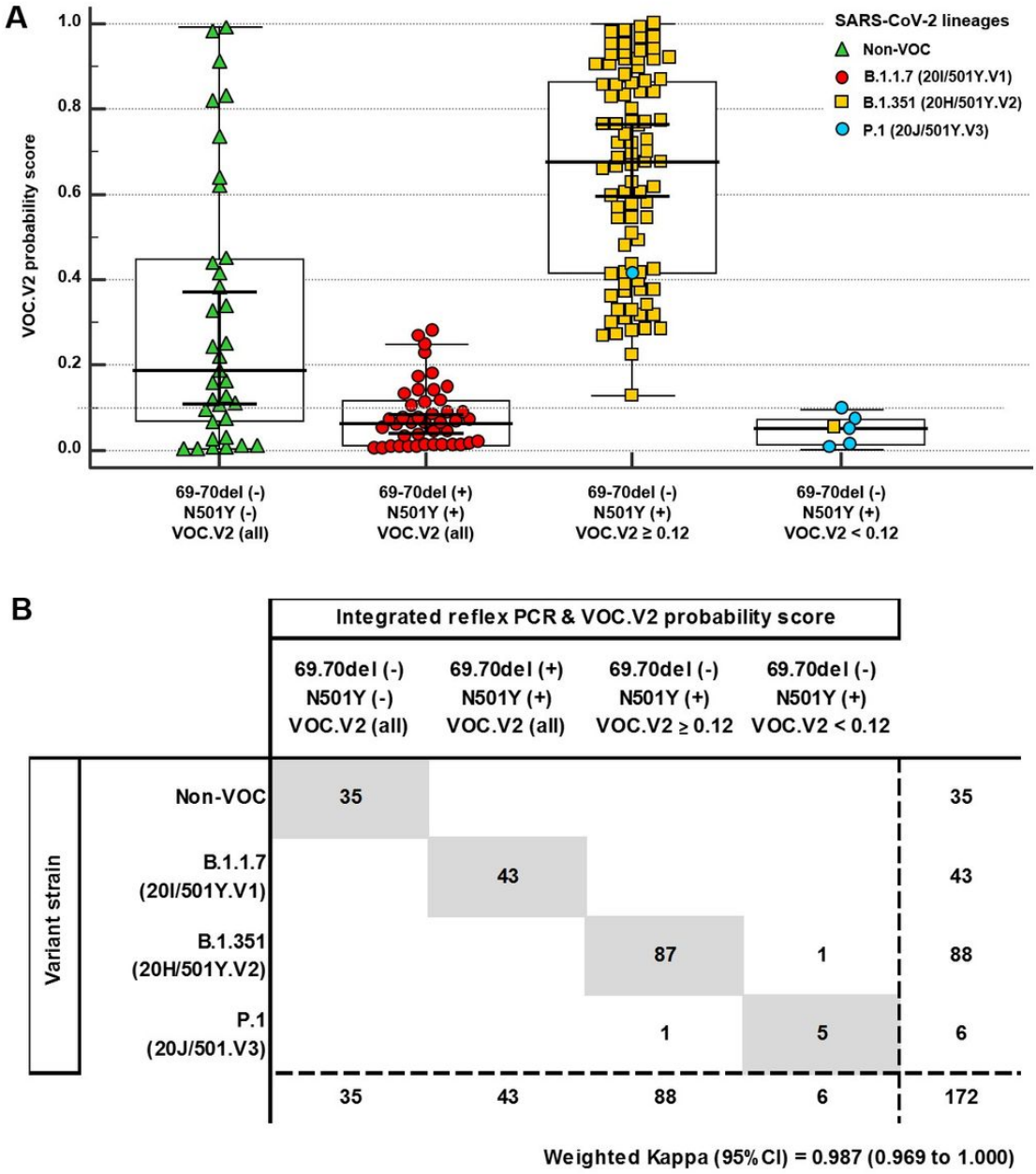

Figure 3

Integration of VOC.V2 probability score with N501Y and 69.70del mutation-specific reflex PCR for differentiation of SARS-CoV-2 variants of concern 1 to 3. A. Distribution of VOC.V2 probability score (Y-axis) in B.1.1.7 (red circles), B.1.351 (yellow squares), P.1 (blue circles, only present in Validation cohort) and all other, non-variant of concern lineages combined (green triangles) in a subset of samples for which mutation-specific reflex PCR was performed for the Spike N501Y and 69.70del mutations. B. Concordance analysis of an integrated PCR analysis of VOC.V2 probability score, N501Y and 69.70del mutations (horizontal) as compared to WGS genotyping (vertical) as gold standard in a subset of N=172 samples.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Suppl.Table1.xlsx

- Suppl.Table2.xlsx
- Suppl.Table3.xlsx
- Suppl.Table4.xlsx