Endonuclease-based genotyping of the RBM as a method to track the emergence or evolution of SARS-CoV-2 variants

Eva Lopez, Margot Barthélémy, Cécile Baronti, Shirley Masse, Alessandra Falchi, Fabien Durbesson, Renaud Vincentelli, Xavier de Lamballerie, Rémi Charrel, Bruno Coutard

Endonuclease-based genotyping of the RBM as a method to track the emergence or evolution of SARS-CoV-2 variants

Eva Lopez[1], Margot Barthélémy[1], Cécile Baronti[1], Shirley Masse[2], Alessandra Falchi[2], Fabien Durbesson[3], Renaud Vincentelli[3], Xavier de Lamballerie[1], Rémi Charrel[1,4], Bruno Coutard[1]

1 Unité des Virus Émergents (UVE: Aix-Marseille Univ-IRD 190-Inserm 1207), Marseille, France.

2 UR7310, Laboratoire de Virologie, Université de Corse-Inserm, 20250 Corte, France.

3 Unité Mixte de Recherche (UMR) 7257, Centre National de la Recherche Scientifique (CNRS) Aix-Marseille Université, Architecture et Fonction des Macromolécules Biologiques (AFMB), Marseille, France.

4 Comité de Lutte contre les Infections Nosocomiales, Hôpitaux Universitaires de Marseille, AP-HM, Marseille, France

Corresponding author and lead contact: Bruno Coutard (bruno.coutard@univ-amu.fr)

Abstract

Since the beginning of the Covid-19 pandemics, variants have emerged. Some of them display increased transmissibility and/or resistance to immune response. Most of the mutations involved in the functional adaptation are found in the Receptor Binding Module (RBM), close to the interface with the receptor ACE2. We thus developed a fast molecular assay to detect mutations in the RBM coding sequence. After amplification, the amplicon is heat-denatured and hybridized with an amplicon of reference. The presence of a mutation can be detected using a mismatch-specific endonuclease and the cleavage pattern is analysed by capillary electrophoresis. The method was validated on RNA of SARS-CoV-2 variants produced *in vitro* before being implemented for clinical samples. The assay showed 97.8% sensitivity and 97.8% specificity. The procedure can be set up for high throughput identification of the presence of mutations and serve as a first-line screening to select the samples for full genome sequencing.

Keywords

Variants of Concern, SARS-CoV-2, molecular assay, mutations, Spike protein.

Introduction

In December 2019, individuals with pneumonia of unknown aetiology were recorded in the city of Wuhan, China. The number of cases increased steadily in the following weeks including in the countries surrounding China (Taiwan, Thailand, Malaysia, etc.), and then throughout the world via aerial transport, triggering the WHO to announce a Public Health Emergency of International Concern (PHEIC) on the 1st of February 2020. Shortly after, the disease was named "COVID-19", for "Coronavirus disease 2019". The coronavirus in question, SARS-CoV-2, is an enveloped, positive single stranded RNA virus. The viral particle exposes at its surface the envelope glycoprotein Spike (S). The S protein is a multi-domain protein (Figure 1A) involved in host cell recognition, in particular *via* its receptor binding domain (RBD) which specifically binds the Human Angiotensin-2 Converting Enzyme (hACE2). The S protein is considered a major determinant of viral infectivity and antigenicity (Li et al., 2020; Walls et al., 2020), and mutations in the coding sequence of the S protein are susceptible to affect the biology of SARS-CoV-2.

Since the emergence of SARS-CoV-2 several non-synonymous mutations have been reported in the coding sequence of the S protein. Some of these non-synonymous mutations are particularly monitored because they are suspected to affect functions of the protein and thereby impact the biology of the virus. One of these first mutations led to D614G, a substitution contributing to the enhancement of viral loads in the upper respiratory tract with possible increased transmission (Plante et al., 2021). Latter, several variants defined as Variant of Concern (VoC) have emerged and are disseminating. Those variants may demonstrate increased transmissibility or severity of the disease, reduction of sero-neutralization by antibodies induced by previous infection or vaccination, or resistance to therapeutic treatments. Among the VoCs, the first one - 501Y.V1 or B.1.1.7 lineage - was identified in the UK and showed enhanced human-to-human transmission and increased disease severity (Davies et al., 2021a) (Davies et al., 2021b). Then, variants of B1.351 (501Y.V2), P.1 (501Y.V3) and B.617 lineages were isolated and characterized in South Africa, Brazil/Japan and India, respectively. Both B1.351 (501Y.V2), P.1 (501Y.V3) variants show increased resistance to antibody neutralization (Davies et al., 2021; Y. Liu et al., 2021; P. Wang et al., 2021). VoCs have in common to present at least one non-synonymous mutation in the spike receptor binding motif (RBM). RBM is the sub-domain of the RBD containing most of the hACE2-contacting residues and is also characterized by the presence of epitopes for neutralising antibodies (Figure 1B). Some of the characterized mutations are N501Y for 501Y.V1, E484K/N501Y for 501Y.V2 and

501Y.V3. By themselves, these mutations can affect the binding of the S protein to hACE2 and/or the potency of neutralizing antibodies (Greaney et al., 2021; Z. Liu et al., 2021; Shang et al., 2020; P. Wang et al., 2021; Zahradník et al., 2021). The genetic evolution of this region is specifically scrutinized to identify possible new VoCs. The rapid detection of VoCs is thus pivotal for mitigating transmission in hospital settings and for adjusting therapies to avoid lowering efficacy.

The detection of VoCs and surveillance of the evolution of SARS-CoV-2 population are currently surveyed by different approaches. Recently, researchers have developed methods such as RT-LAMP (LAMP SARS-CoV-2 Variant detection panel, LaCAR), CRISPR-Cas9 or CRISPR-Cas13-based methods to specifically detect VoCs (Kumar et al., 2021.; Y. Wang et al., 2021). However the two most commonly used methods are the real time RT-PCR for the search of mutations at given positions and massive campaigns of New Generation Sequencing (NGS), both contributing to the public-health decision making (Oude Munnink et al., 2020) . On one side, real time RT-PCR is fast and operational on site, but it can detect only known mutations and does not address the newly emerging ones. On the other side, NGS can detect any mutation along the genome but the results is obtained in days rather than hours, delaying information required for medical decision to be taken upon sequence identification. In addition, all the biological samples cannot be sequenced and upstream sampling is mandatory for the selection of the most relevant biological samples to characterize. Here we present the proof-of-concept for an alternative method allowing the surveillance of the genetic drift of SARS-CoV-2 in the RBM region where mutations are susceptible to affect the dissemination, pathogenicity or antibody-resistance of the virus. The technique, relying on the amplification of the RBM coding sequence followed by an assay using a mismatch-specific endonuclease, has been validated on biological samples demonstrating its feasibility.

Results and discussion

Principle of the method

Some of the non-synonymous mutations occurring in the RBM coding sequence of SARS-CoV-2 are pivotal because they are susceptible to change the phenotype with possible influence on transmission pattern, increased pathogenesis or immune escape; the latter can result in iterative infections, reduced vaccine efficacy or resistance to Mab-based therapeutics. Accordingly, RBM genetic evolution has to be monitored for the surveillance of emerging variants. To detect mutations in the RBM region, we evaluated the SURVEYOR® Nuclease S, an endonuclease

cleaving double strand DNA where mismatches exist. The enzyme had already been used for genotyping for several purposes including the detection of mutations in *brca1* and *brca2* genes in the case of hereditary breast cancer (Davies et al., 2006; Pilato et al., 2012). The principle of the technique relies on the creation of heteroduplexes between an amplified DNA from a reference nucleic acid (*Ref*) and an amplified DNA from a sample to be evaluated (*Sample*). Both amplicons are mixed (Figure 1C). After denaturation/hybridization, a mixture of homoduplexes (*Ref*/*Ref* and *Sample*/*Sample*) and heteroduplexes (*Ref*/*Sample*) are produced. If the sequence from the S*ample* has mutations compared to the *Ref* sequence, mismatches happen in the heteroduplex. The latter is cleaved, resulting in cleavage products that can be evidenced by capillary electrophoresis, providing information on the number of Single Nucleotide Polymorphism (SNPs) and their approximate location(s). In contrast, if the sequences of *Ref* and *Sample* are identical no cleavage is observed and only the full-length PCR products are visible by electrophoresis.

Sensitivity of the RT-PCR assay targeting the RBM coding sequence

Within the RBD region of the Spike protein, the RBM contains amino-acids subjected to mutations (AA) positions – *eg* 452, 484 and 501 - with functional relevance and observed in VoCs (Figure 1A). Its coding sequence was thus well-suited for a genotyping assay. Prerequisite was also that the targeted region is centred on the positions of interest and short enough so that it excludes the identification of mutations with low to no functional effect, *ie* synonymous or non-synonymous with functional consequence. The size of the PCR product was thus set to 315-nt, allowing the detection of mutations in the region of the Spike protein spanning AA positions 431 to 524 of the S protein.

To set up the assay, we first evaluated the sensitivity of the RT-PCR system in the range 1-10^6 of RNA copies/µL. The evaluation was done either by real-time RT-PCR with SYBR green (Figure 2A, 2B), or by end-point analysis of the PCR products on agarose gel electrophoresis (Figure 2C). The amplification is linear from 10 to 10^5 copies of RNA/µL, with correlation coefficient R^2=0.9986 (Figure 2A), and has a limit of detection on agarose gel up to 10 copies/µL (Figure 2C). However, given the quantity of material needed for the nuclease assay (>25 ng/µL), we arbitrarily applied the threshold at 10^3/10^4 copies/µL, which corresponds to samples with Ct values between 28 to 30 in reference detection systems (Pezzi et al., 2020). It should be noted that the SYBR green inhibits the endonuclease used for the detection of

mismatches, likely by altering the structure of the DNA helix, thus rendering the resulting PCR product not suitable for subsequent capillary electrophoresis analysis (data not shown).

Detection and identification of mutations revealing in the RBM

Detection of mutations in reference material derived from SARS-CoV-2 isolates

To establish the proof of concept, first experiments were conducted using viral RNA derived from cell cultures infected by three well-characterized variants: SARS-CoV-2 BavPat1, 501Y.V1 and 501Y.V2; (i) 501Y.V1 has N501Y (nt A1501T) mutation, (ii) and 501Y.V2 has E484K and N501Y (nt G1450A and A1501T, respectively) mutations by reference to the BavPat1 respectively. The theoretical cleavage profiles for pairwise combinations of the three variants are presented in Table 1. In practice PCR products were mixed in pairs in approximate equimolar quantities before denaturation/hybridization prior to the mismatch-specific endonuclease assay. The results are presented in Figure 2D. The electropherogram (blue curve) corresponding to the hybridization of the BavPat1 amplicon with itself, with no mismatch expected, resulted in a unique 330-nt long fragment. When mixed with the one of 501Y.V1 (Fig 2D, orange curve), three DNA fragments were observed at 330, 232 and 93 bp, close to the anticipated profile (Table 1). For the identification of the 501Y.V2 (grey curve), five DNA fragments can be detected, corresponding to the products from both complete and partial cleavages of the mismatches at the expected positions. Finally, when mixing amplicons from the 501Y.V1 and 501Y.V2 samples (yellow curve), three DNA fragments were observed, where the two smaller being indicative of the difference at position 484 in the Spike protein sequence of the two variants.

Application to clinical samples containing SARS-CoV-2 RNA as diagnosed by routine real-time RT-PCR assay

To further assess the mutation detection method, 8 samples were tested after they were diagnosed as SARS-CoV-2 RNA positive using the routine diagnostic assay TaqPath real-time RT-PCR™ (ThermoFisher Scientific), known to discriminate the 501Y.V1 variant on the deletion observed in the S coding sequence, leading to the amplification of only two targets out the three of the test (Kidd et al., 2021). In order to make the analysis easy and rapid, it is advocated to test each sample against itself and against the selected reference with no prior quantification. In our case, the reference was the European lineage of SARS-CoV-2 (BavPat1 strain) since it was the dominant one at the time of the study. Clinical samples were identified

as either SARS-CoV-2 positive for the three targets of the TaqPath assay™ (Figure 3, samples 1, 3, 4, 5, 7 and 8) or SARS-CoV-2 501Y.V1 positive (Figure 3 samples 2 and 6) with Ct ranging from 18 to 34. For samples 5 and 8, no fragment corresponding to the expected amplicon was visible in the self-hybridization assay (Figure 3, lanes 13 and 22), in line with the high Ct values obtained from initial real time RT-PCR assays. By contrast, the amplification of the target region for the other sample was satisfactory (Figure 3, lanes 1, 4, 7, 10, 16 and 19). Overall, SARS-CoV-2 positive and 501Y.V1-putative samples (2 and 6) showed profiles matching with the expectations, *ie* one mismatch with BavPat1 reference amplicon and no mismatch with 501Y.V1 reference amplicon (Figure 3, lanes 5, 6 and 17, 18 respectively).

For the other samples (Figure 3, samples 1, 3, 4 and 7), described as non-501Y.V1 according to the TaqPath assay, no cleavage was expected when compared to the BavPat1 reference amplicon, and two cleaved products with the 501Y.V1. However, at least one additional fragment was observed against both reference samples, of about 135 bp length (Figure 3, lanes 2,3,8,9,11,12,20 and 21). The PCR products were submitted to Sanger sequencing and a non-synonymous mutation yielding S477N substitution was detected, in agreement with the size of the cleaved products. This mutation had already been reported in viral populations circulating in Europe (https://www.gisaid.org/), and the corresponding variant has been shown to slightly increase the receptor binding domain's affinity for hACE2 (Hodcroft et al., 2020).

We next blindly evaluated the assay on 92 SARS-CoV-2 positive samples, for which Ct values were below 28 and NGS data available (Supplemental Figure S1). The sensitivity, defined as the ability to generate an amplicon with yields compatible for the nuclease assay, is 97.83% as 2 samples out of 92 were not properly amplified. The specificity was defined as the ability to detect in the RT-PCR-positive samples a mutation compared to the BavPat1 strain or detect the lack of mutation for sequences identical to the reference. Compared to the sequence data, 2 samples out of 90 were inadequately identified, with selectivity of 97.78%. Out of the 88 remaining profiles, 82 had the same profile as in Figure 2D. Yet unmet patterns were observed and confirmed by the sequence analysis with E471D mutation detected in two samples, E484K in one sample, L452R in one sample and both L452R / N501Y in 2 others.

Altogether, these results demonstrate that the mismatch-specific nuclease assay coupled to capillary electrophoresis is a suitable assay for the detection in clinical samples of already reported mutations. The identification of atypical profiles, confirmed in this study by sequencing, also demonstrates that it is relevant for the discovery of yet unmet variants. With the incremental characterization of variants, the set of reference cleavage patterns can be updated to adapt the assay to circulating strains. As this technique can be dimensioned for

96/384 well devices and requires less than 4 hours from the extracted RNA to production of individual results, it can be used to filter SARS-CoV-2 positive clinical samples and identify those for which virus isolation and complete genome sequencing is justified for surveillance purpose.

In conclusion, we developed a molecular assay dedicated to the surveillance of SARS-CoV-2 variants, specifically targeting the RBM coding sequence known to be involved in the functional adaptation of the virus. The assay is suitable to screen biological samples and identify the presence new or emerging mutations.

Limitations of study

The technique we developed is based on RT-PCR amplification of the RBM coding sequence followed by mismatch-specific nuclease assay and detection by DNA capillary electrophoresis and is made possible only for samples with RNA titers yielding Ct values better than 28 to produce enough amplified DNA material. The proof of concept has been validated but the method remains to be deployed on a large cohort to assess if it is amenable to high throughput genotyping or to individual surveillance of virus evolution during a chronic infection by SARS-CoV-2, a situation favourable for virus adaptation (Kemp et al., 2021).

Author contribution

B.C conceived the idea and supervised the work. E.L, M.B, C.B, F.D and R.V carried out the experiments. S.M and A.L provided material on related information. E.L, R.C and B.C wrote the article. All the authors discussed the results and commented on the manuscript.

Acknowledgements

We thank Professor Bruno Lina and the Centre National de Référence des virus des infections respiratoires for sharing extracted RNA of clinical samples and sequencing information. This study was funded by (i) the “European Virus Archive Global” (EVA-GLOBAL) project H2020-INFRAIA-2019 program, Project No 871029, (ii) the “Advanced Nanosensing platforms for Point of care global diagnostics and surveillance” (CONVAT), H2020, Project No 101003544. Eva Lopez is the recipient a DGA fellowship (Direction Générale de l’Armement).

Declaration of interests

The authors declare no competing interests

Figure titles and legend

Figure 1: **RBM is a key region for viral for viral entry and seroneutralization.** A: Primary structure of the SARS-CoV-2 Spike. NTD: N-terminal domain; RBD: Receptor binding domain; RBM: Receptor binding motif; SD1: Subdomain 1; SD2: Subdomain 2; FP: Fusion peptide; HR1: Heptad repeat 1; HR2: Heptad repeat 2; TM: Transmembrane region; CT: Cytoplasmic region. Amino acid positions of the RBD and RBM boundaries are presented in black and blue numbers, respectively. Functional mutations in the RBM are shown in orange (E484K) and red (N501Y). Purple arrows and numbers correspond to the RT-PCR amplified region. B: Structure representation of the interaction between SARS-CoV-2 RBD core (green) and hACE2 receptor (white) (PDB 6vw1). RBM is shown in blue. E484 and N501 positions are highlighted in orange and red, respectively. C: Detection of mutation(s) in a sample compared to a reference sequence. Results of the digestion by a mismatch-specific endonuclease of an auto-control (amplicon of a sample alone) and mix (amplicons of reference and sample) are analysed by electrophoresis and fragment patterns are detected.

Figure 2: **Evaluation of the target amplification and mismatch-specific nuclease assay on reference material.** A and B: Linearity of the real time RT-PCR assay targeting the RBM coding sequence from 10^5 (blue curve), 10^4 (red curve), 10^3 (green curve), 10^2 (yellow curve) and 10^1 (pink curve) RNA copy/µL. The linear regression (dashed line) in Panel A has a correlation coefficient R^2=0.9986. C: End-point analysis of PCR fragments on gel electrophoresis. D: Electropherograms of the digestion products. Values above the peaks correspond to the size of the fragment. Peaks on the right (330bp) correspond to the homoduplexes created during the hybridization.

Figure 3: Validation of the assay on SARS-CoV-2 positive clinical samples. The First row indicates number of the samples, from 1 to 8. Ct obtained from TaqPath COVID-19 kit™ (ThermoFisher) real time RT-PCR for ORF1ab, S and N genes are presented in the three next rows, respectively (nd: not determined, Ct>35), leading to the following classification: (+) refers to the presence of the deletion indicative of a 501Y.V1 variant, and (-) to a non-501Y.V1

sample. Below the table, the results of the mismatch-specific assay followed by capillary electrophoresis are presented in a "gel-like" format for the eight samples. The molecular weight ladder is indicated on the left. For each sample (1 to 8), the experiment was conducted in three conditions: The first lane is the self-hybridization (sample/sample; Lane A), the second one is sample/BavPat1 reference (lane B) and the third sample/501Y.V1 reference (Lane C). The fragment above 330 bp corresponds to uncleaved homoduplexes.

Table 1: Theoretical cleavage profiles for pairwise combinations of amplicons from BavPat1, 501Y.1 and 501Y.V2 variants. The amino-acid residue and corresponding position in the sequence of the Spike protein for each variant are in italic. Values in the table correspond the expected size of the fragments after treatment with the mismatch-specific nuclease. Underlined: fragments resulting from the complete cleavage of the heteroduplexes; Bold: fragment resulting from an incomplete cleavage.

	0 **BavPat1**	*N501Y* **501.V1**	*E848K ; N501Y* **501Y.V2**
0 **BavPat1**	315	84 / 231 / 315	51 / 84 / **135** / 180 / **231** / 315
N501Y **501.V1**	-	315	135 / 180 / 315
E848K ; N501Y **501Y.V2**	-	-	315

STAR Methods

Resource availability

Lead contact

Requests for further information should be directed to Bruno Coutard (bruno.coutard@univ-amu.fr).

Materials availability

This study did not generate new unique material or reagents.

Data and code availability

Data

All data reported in this paper will be shared by the lead contact upon request.

Code

This paper does not report original code.

Other items

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Experimental model and subject detail

Purified viral RNA from infected cell cultures

VeroE6/TMPRSS2+ (CFAR#100978) cells were grown in minimal essential medium (MEM) (Thermo Fisher Scientific) with 7.5 % heat-inactivated foetal calf serum (FCS; Thermo Fisher Scientific), with 1% penicillin/streptomycin (PS, 5000 U.mL−1 and 5000 µg.mL−1 respectively; Thermo Fisher Scientific), supplemented with 1% non-essential amino acids (Thermo Fisher Scientific) and L-Glutamine (Thermo Fisher Scientific), at 37 °C with 5% CO2. SARS-CoV-2 strain BavPat1 was obtained from Pr. C. Drosten through EVA GLOBAL (https://www.european-virus-archive.com/). SARS-CoV-2 strain 2021/FR/7b-ex UK (EVA-G ref: 001V-04044) belonging to lineage 501Y.V1 and strain 2021/FR/1299-ex SA (EVA-G ref: 001V-04067) belonging to 501Y.V2 were isolated from Human nasopharyngeal swabs. The virus stocks were produced on VeroE6/TMPRSS2+ cells. Briefly, a 25 cm² flask of sub-confluent cells was inoculated with each strain at MOI=0.001. Cells were incubated at 37 °C overnight, after which the medium was changed, and incubation was continued for 24 hours. The supernatant was then collected, clarified by spinning at 1500 g for 10 min., supplemented with 25 mM HEPES (Sigma-Aldrich) and stored at -80 °C in several aliquots. All infectious experiments were conducted in a biosafety level 3 laboratory (BSL3).

RNA from clinical samples

For the first set of 8 clinical samples, we collected anonymised residual nasopharyngeal (NP) samples of COVID-19 patients confirmed by real time RT-PCR from clinical laboratories based in Corsica in mid-January 2021. No nominative nor sensitive data on participant people have been collected. This study falls within the scope of the French Reference Methodology MR-004 according to 2016–41 law dated 26 January 2016 on the modernization of the French health

system. The ethics committee of University of Corsica Pascal Paoli (IRB UCPP 2020-01) approved this study.

The collection of 92 SARS-CoV-2 positive clinical samples used in the second evaluation phase was kindly provided by the National Reference Center for Respiratory Viral Infections (CNR Lyon, France). The SARS-CoV-2 genomes from the 92 samples were fully sequenced by the CNR for the surveillance of the variants.

Method details

Extraction of the RNA from infected cell cultures

RNA extraction was performed using the QIAamp 96 DNA kit and the Qiacube HT plasticware kit on the Qiacube HT automate (Qiagen). Viral RNA obtained was quantified by real-time RT-PCR (EXPRESS One-Step Superscript qRT-PCR Kit (Thermo Fisher Scientific) using 3.5 µL of RNA and 6.5 µL of qRT-PCR mix and standard fast cycling parameters, i.e., 10 min at 50 °C, 2 min at 95 °C, and 40 amplification cycles (95 °C for 3 sec. followed by 30 sec. at 60 °C) (Touret et al., 2020). The quantification was provided by four log 2 serial dilutions of an appropriate T7-generated synthetic RNA standard of known quantities (102 to 108 copies). Real-time RT-PCR reactions were performed on QuantStudio 12K Flex Real-Time PCR System (APPLIED BIOSYSTEMS, Waltham, USA) and analysed using QuantStudio 12 K Flex Applied Biosystems software v1.2.3. Primers and probe targeting SARS-CoV-2 N gene, are: Forward: 5'-GGCCGCAAATTGCACAAT-3'; Reverse: 5'-CCAATGCGCGACATTCC-3'; Probe: FAM-CCCCCAGCGCTTCAGCGTTCT-BHQ1.

Extraction of the RNA from clinical samples

Total nucleic acids were extracted from 200 µL of NP samples using the QIACUBE processing system with the QIAamp 96 Virus QIAcube HT Kit (Qiagen) and eluted to 100 µL of total nucleic acid. Real-time RT-PCR was performed for each sample with TaqPath COVID-19 KitTM (Thermo Fisher Scientific). The TaqPath COVID-19 KitTM is a multiplex real time RT-PCR diagnostic assay targeting three regions of the SARS-CoV-2 genome (N, S, and ORF1ab) which was approved by the French National Center of respiratory diseases for the detection of SARS-CoV-2. TaqPath RT-PCRTM assay is a useful tool enabling a rapid screening of SARS-CoV-2 as a S-gene target failure was observed in variants with a deletion at positions 69-70 (ΔHV69-70) whereas ORF1ab and N targets are correctly amplified (Bal et al., 2021).

Choice of the target sequence for amplification

The target of the SARS-CoV-2 Spike coding sequence ranges from nt position 1273 to 1587, including the primers, and is 315 nucleotide-long. The corresponding amino acid sequence (amino acid 425 to 529) encompasses the RBM module (Figure 1A).

RT-PCR protocol of the targeted sequence

RT-PCR was performed with SuperScript II One-Step Platinium Kit (#10928042, Thermo Fisher Scientific), on a Biometra T3000 thermal cycler. Primers were synthesized and provided by Thermo Fisher Scientific. The forward and reverse primer sequences are: 5'-TTACCAGATGATTTTACAGGC-3' and 5'-AGACTTTTTAGGTCCACAAAC-3', respectively. The cycling conditions were defined as follows : 50 °C for 30 min. ; 94 °C for 2 min. ; 40 cycles of 94 °C for 15 sec. ; 55 °C for 20 sec. and 68°C for 20 sec. The final elongation was performed at 72°C for 2 min.

Evaluation of the linearity in the designed system

To determine the linearity of the designed system, real time RT-PCR was performed with QuantiTect SYBR Green RT-PCR Kit (#204243, Qiagen), on a BioRad CFX96™ thermal cycler, software version 3.0 (Bio-Rad Laboratories). The cycling conditions were : 50 °C for 30 min. ; 95 °C for 15 min. ; 40 cycles of 95 °C for 15 sec. ; 50 °C for 30 sec. and 72°C for 45 sec. PCR products were then loaded on a 2% agarose gel. The size of the fragment size is estimated using a molecular size ladder (DNA 1kb Plus #10787026, ThermoFisher Scientific).

Production of the reference PCR products

SARS-CoV-2 BavPat1 strain was selected as the reference for this study. A large scale production of amplicons was done for this strain, following the RT-PCR protocol defined above.

Production of the PCR products from biological samples

Each sample used in this study has been amplified following the RT-PCR protocol described above.

Mismatch-specific endocnuclease assay and detection of cleavage

The presence of a mutation in the amplified target sequence was detected using the Surveyor® Mutation Detection Kit (#706021, Integrated DNA Technologies). Two PCR products (a reference product and a sample to be analysed) were mixed in a 10 µL final volume and the

endonuclease mismatch specific cleavage was performed following the manufacturer recommendations. The mixture was then loaded on capillary electrophoresis system (Fragment Analyser 5200, Agilent or GXII, Perkin Elmer) prior to analysis of the cleavage profile. When needed for confirmation of the presence of mutations, the PCR products were sequenced using the Sanger method with the forward and reverse primers used for the RT-PCR (Genewiz).

References

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., Group, the C.-D.H.S., 2021. 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. Eurosurveillance 26, 2100008. https://doi.org/10.2807/1560-7917.ES.2021.26.3.2100008

Davies, H., Dicks, E., Stephens, P., Cox, C., Teague, J., Greenman, C., Bignell, G., O'Meara, S., Edkins, S., Parker, A., Stevens, C., Menzies, A., Blow, M., Bottomley, B., Dronsfield, M., Futreal, P.A., Stratton, M.R., Wooster, R., 2006. High throughput DNA sequence variant detection by conformation sensitive capillary electrophoresis and automated peak comparison. Genomics 87, 427–432. https://doi.org/10.1016/j.ygeno.2005.11.008

Davies, N.G., Abbott, S., Barnard, R.C., Jarvis, C.I., Kucharski, A.J., Munday, J.D., Pearson, C.A.B., Russell, T.W., Tully, D.C., Washburne, A.D., Wenseleers, T., Gimma, A., Waites, W., Wong, K.L.M., van Zandvoort, K., Silverman, J.D., CMMID COVID-19 Working Group, COVID-19 Genomics UK (COG-UK) Consortium, Diaz-Ordaz, K., Keogh, R., Eggo, R.M., Funk, S., Jit, M., Atkins, K.E., Edmunds, W.J., 2021a. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. Science. https://doi.org/10.1126/science.abg3055

Davies, N.G., Jarvis, C.I., CMMID COVID-19 Working Group, Edmunds, W.J., Jewell, N.P., Diaz-Ordaz, K., Keogh, R.H., 2021b. Increased mortality in community-tested cases of SARS-CoV-2 lineage B.1.1.7. Nature. https://doi.org/10.1038/s41586-021-03426-1

Greaney, A.J., Loes, A.N., Crawford, K.H.D., Starr, T.N., Malone, K.D., Chu, H.Y., Bloom, J.D., 2021. Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. Cell Host Microbe 29, 463-476.e6. https://doi.org/10.1016/j.chom.2021.02.003

Hodcroft, E.B., Zuber, M., Nadeau, S., Crawford, K.H.D., Bloom, J.D., Veesler, D., Vaughan, T.G., Comas, I., Candelas, F.G., Stadler, T., Neher, R.A., 2020. Emergence and spread of a SARS-CoV-2 variant through Europe in the summer of 2020. medRxiv. https://doi.org/10.1101/2020.10.25.20219063

Kemp, S.A., Collier, D.A., Datir, R.P., Ferreira, I.A.T.M., Gayed, S., Jahun, A., Hosmillo, M., Rees-Spear, C., Mlcochova, P., Lumb, I.U., Roberts, D.J., Chandra, A., Temperton, N., CITIID-NIHR BioResource COVID-19 Collaboration, COVID-19 Genomics UK (COG-UK) Consortium, Sharrocks, K., Blane, E., Modis, Y., Leigh, K.E., Briggs, J.A.G., van Gils, M.J., Smith, K.G.C., Bradley, J.R., Smith, C., Doffinger, R., Ceron-Gutierrez, L., Barcenas-Morales, G., Pollock, D.D., Goldstein, R.A., Smielewska, A., Skittrall, J.P., Gouliouris, T., Goodfellow, I.G., Gkrania-Klotsas, E., Illingworth, C.J.R.,

McCoy, L.E., Gupta, R.K., 2021. SARS-CoV-2 evolution during treatment of chronic infection. Nature 592, 277–282. https://doi.org/10.1038/s41586-021-03291-y

Kidd, M., Richter, A., Best, A., Cumley, N., Mirza, J., Percival, B., Mayhew, M., Megram, O., Ashford, F., White, T., Moles-Garcia, E., Crawford, L., Bosworth, A., Atabani, S.F., Plant, T., McNally, A., 2021. S-variant SARS-CoV-2 lineage B1.1.7 is associated with significantly higher viral loads in samples tested by ThermoFisher TaqPath RT-qPCR. J. Infect. Dis. https://doi.org/10.1093/infdis/jiab082

Kumar, M., Gulati, S., Ansari, A.H., Phutela, R., Acharya, S., Azhar, M., Murthy, J., Kathpalia, P., Kanakan, A., Maurya, R., Vasudevan, J.S., S, A., Pandey, R., Maiti, S., Chakraborty, D., n.d. FnCas9-based CRISPR diagnostic for rapid and accurate detection of major SARS-CoV-2 variants on a paper strip. eLife 10, e67130. https://doi.org/10.7554/eLife.67130

Li, Q., Wu, J., Nie, J., Zhang, Li, Hao, H., Liu, S., Zhao, C., Zhang, Q., Liu, H., Nie, L., Qin, H., Wang, M., Lu, Q., Li, Xiaoyu, Sun, Q., Liu, J., Zhang, Linqi, Li, Xuguang, Huang, W., Wang, Y., 2020. The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity. Cell 182, 1284-1294.e9. https://doi.org/10.1016/j.cell.2020.07.012

Liu, Y., Liu, J., Xia, H., Zhang, X., Fontes-Garfias, C.R., Swanson, K.A., Cai, H., Sarkar, R., Chen, W., Cutler, M., Cooper, D., Weaver, S.C., Muik, A., Sahin, U., Jansen, K.U., Xie, X., Dormitzer, P.R., Shi, P.-Y., 2021. Neutralizing Activity of BNT162b2-Elicited Serum. N. Engl. J. Med. https://doi.org/10.1056/NEJMc2102017

Liu, Z., VanBlargan, L.A., Bloyet, L.-M., Rothlauf, P.W., Chen, R.E., Stumpf, S., Zhao, H., Errico, J.M., Theel, E.S., Liebeskind, M.J., Alford, B., Buchser, W.J., Ellebedy, A.H., Fremont, D.H., Diamond, M.S., Whelan, S.P.J., 2021. Identification of SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. Cell Host Microbe 29, 477-488.e4. https://doi.org/10.1016/j.chom.2021.01.014

Oude Munnink, B.B., Nieuwenhuijse, D.F., Stein, M., O'Toole, Á., Haverkate, M., Mollers, M., Kamga, S.K., Schapendonk, C., Pronk, M., Lexmond, P., van der Linden, A., Bestebroer, T., Chestakova, I., Overmars, R.J., van Nieuwkoop, S., Molenkamp, R., van der Eijk, A.A., GeurtsvanKessel, C., Vennema, H., Meijer, A., Rambaut, A., van Dissel, J., Sikkema, R.S., Timen, A., Koopmans, M., Dutch-Covid-19 response team, 2020. Rapid SARS-CoV-2 whole-genome sequencing and analysis for informed public health decision-making in the Netherlands. Nat. Med. 26, 1405–1410. https://doi.org/10.1038/s41591-020-0997-y

Pezzi, L., Charrel, R.N., Ninove, L., Nougairede, A., Molle, G., Coutard, B., Durand, G., Leparc-Goffart, I., de Lamballerie, X., Thirion, L., 2020. Development and Evaluation of a duo SARS-CoV-2 RT-qPCR Assay Combining Two Assays Approved by the World Health Organization Targeting the Envelope and the RNA-Dependant RNA Polymerase (RdRp) Coding Regions. Viruses 12. https://doi.org/10.3390/v12060686

Pilato, B., De Summa, S., Danza, K., Papadimitriou, S., Zaccagna, P., Paradiso, A., Tommasi, S., 2012. DHPLC/SURVEYOR Nuclease: A Sensitive, Rapid and Affordable Method to Analyze BRCA1 and BRCA2 Mutations in Breast Cancer Families. Mol. Biotechnol. 52, 8–15. https://doi.org/10.1007/s12033-011-9468-5

Plante, J.A., Liu, Y., Liu, J., Xia, H., Johnson, B.A., Lokugamage, K.G., Zhang, X., Muruato, A.E., Zou, J., Fontes-Garfias, C.R., Mirchandani, D., Scharton, D., Bilello, J.P., Ku, Z., An, Z., Kalveram, B., Freiberg, A.N., Menachery, V.D., Xie, X., Plante, K.S., Weaver, S.C., Shi, P.-Y., 2021. Spike mutation D614G alters SARS-CoV-2 fitness. Nature 592, 116–121. https://doi.org/10.1038/s41586-020-2895-3

Shang, J., Ye, G., Shi, K., Wan, Y., Luo, C., Aihara, H., Geng, Q., Auerbach, A., Li, F., 2020. Structural basis of receptor recognition by SARS-CoV-2. Nature 581, 221–224. https://doi.org/10.1038/s41586-020-2179-y

Touret, F., Gilles, M., Barral, K., Nougairède, A., van Helden, J., Decroly, E., de Lamballerie, X., Coutard, B., 2020. In vitro screening of a FDA approved chemical library reveals potential inhibitors of SARS-CoV-2 replication. Sci. Rep. 10, 13093. https://doi.org/10.1038/s41598-020-70143-6

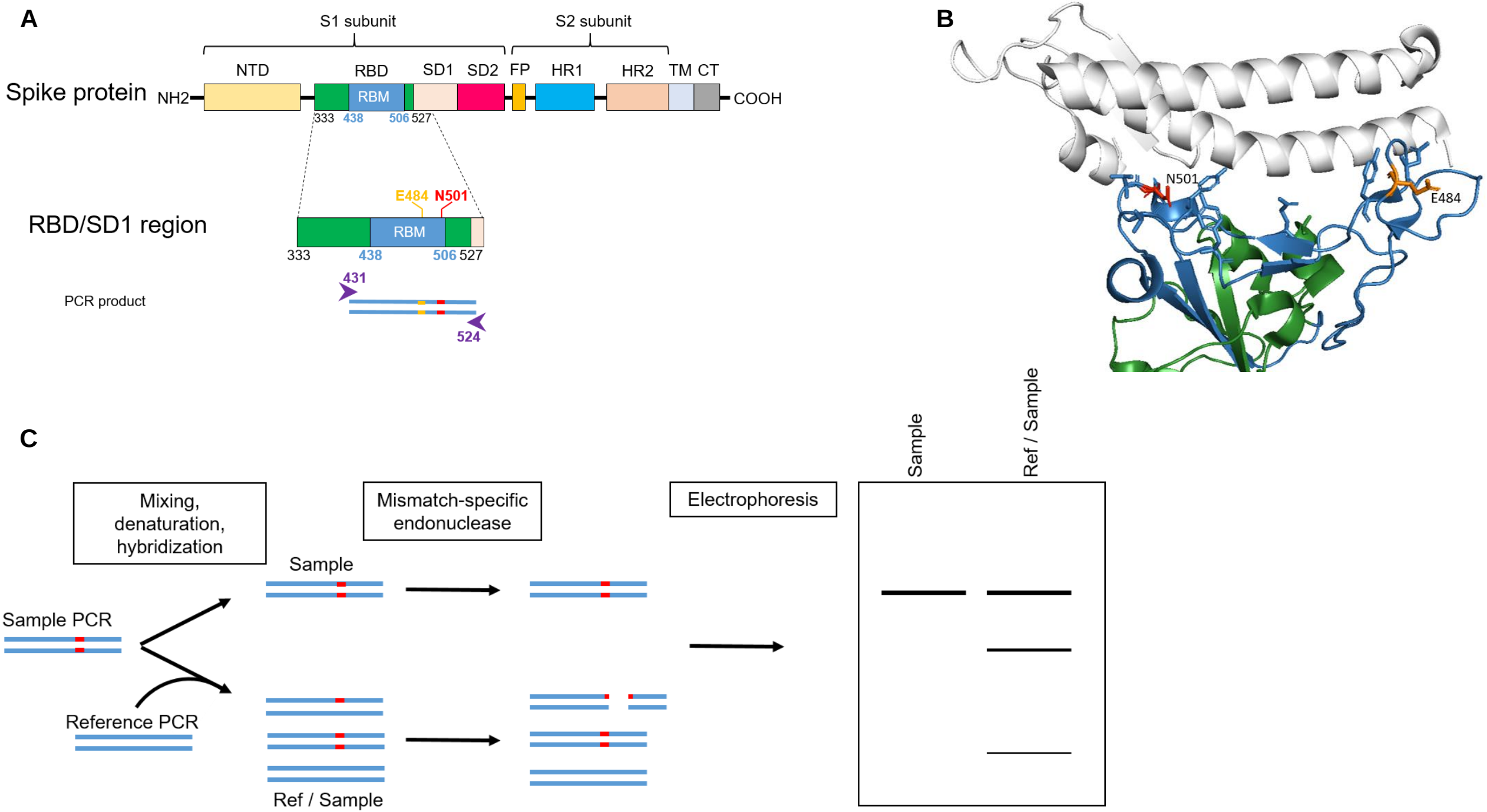
Walls, A.C., Park, Y.-J., Tortorici, M.A., Wall, A., McGuire, A.T., Veesler, D., 2020. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. Cell 181, 281-292.e6. https://doi.org/10.1016/j.cell.2020.02.058

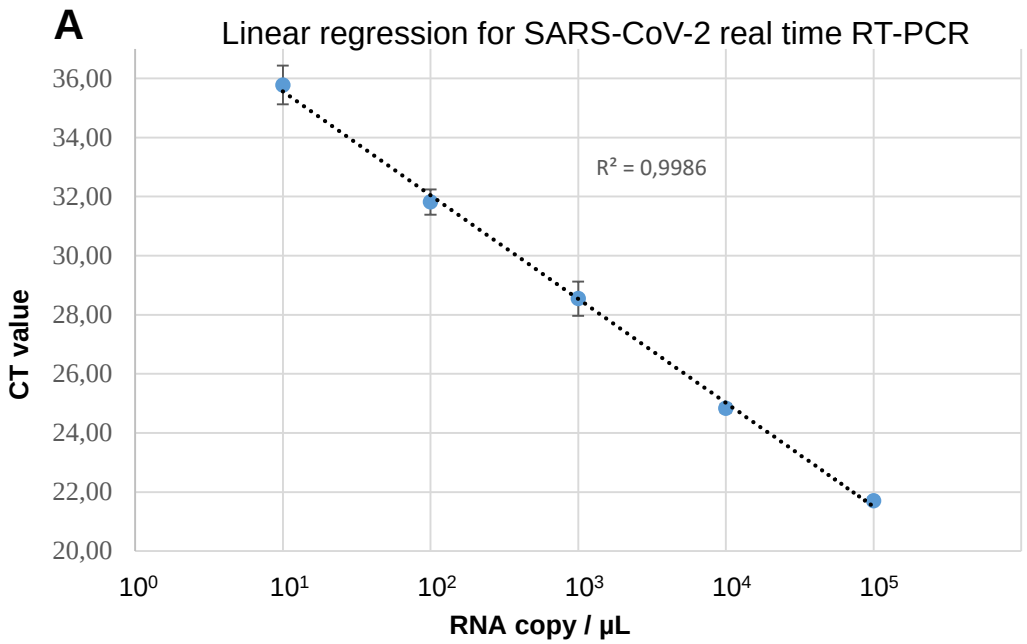
Wang, P., Nair, M.S., Liu, L., Iketani, S., Luo, Y., Guo, Y., Wang, M., Yu, J., Zhang, B., Kwong, P.D., Graham, B.S., Mascola, J.R., Chang, J.Y., Yin, M.T., Sobieszczyk, M., Kyratsous, C.A., Shapiro, L., Sheng, Z., Huang, Y., Ho, D.D., 2021a. Increased Resistance of SARS-CoV-2 Variants B.1.351 and B.1.1.7 to Antibody Neutralization. BioRxiv Prepr. Serv. Biol. https://doi.org/10.1101/2021.01.25.428137

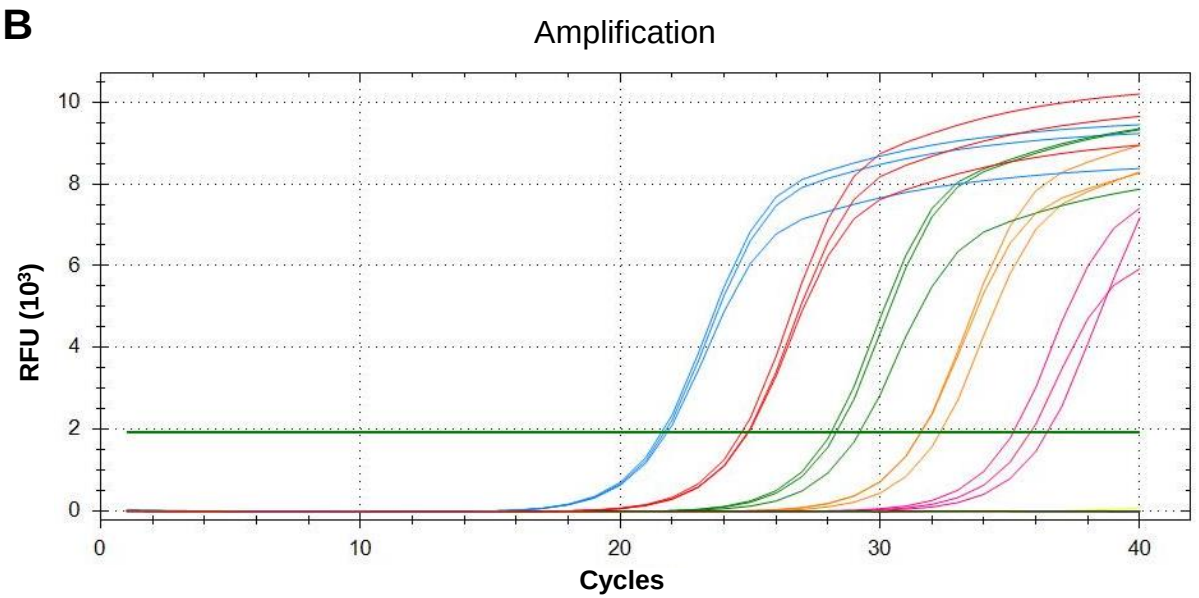
Wang, P., Nair, M.S., Liu, L., Iketani, S., Luo, Y., Guo, Y., Wang, M., Yu, J., Zhang, B., Kwong, P.D., Graham, B.S., Mascola, J.R., Chang, J.Y., Yin, M.T., Sobieszczyk, M., Kyratsous, C.A., Shapiro, L., Sheng, Z., Huang, Y., Ho, D.D., 2021b. Antibody Resistance of SARS-CoV-2 Variants B.1.351 and B.1.1.7. bioRxiv 2021.01.25.428137. https://doi.org/10.1101/2021.01.25.428137

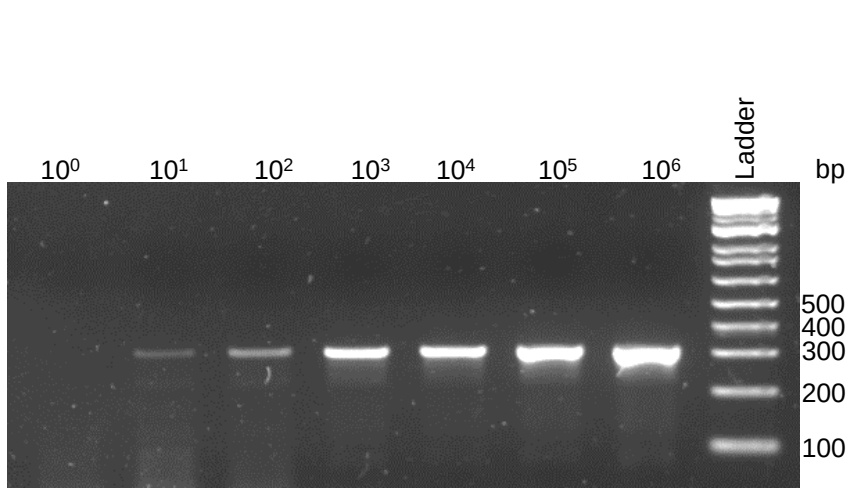
Wang, Y., Zhang, Y., Chen, J., Wang, M., Zhang, T., Luo, W., Li, Y., Wu, Y., Zeng, B., Zhang, K., Deng, R., Li, W., 2021. Detection of SARS-CoV-2 and Its Mutated Variants via CRISPR-Cas13-Based Transcription Amplification. Anal. Chem. acs.analchem.0c04303. https://doi.org/10.1021/acs.analchem.0c04303

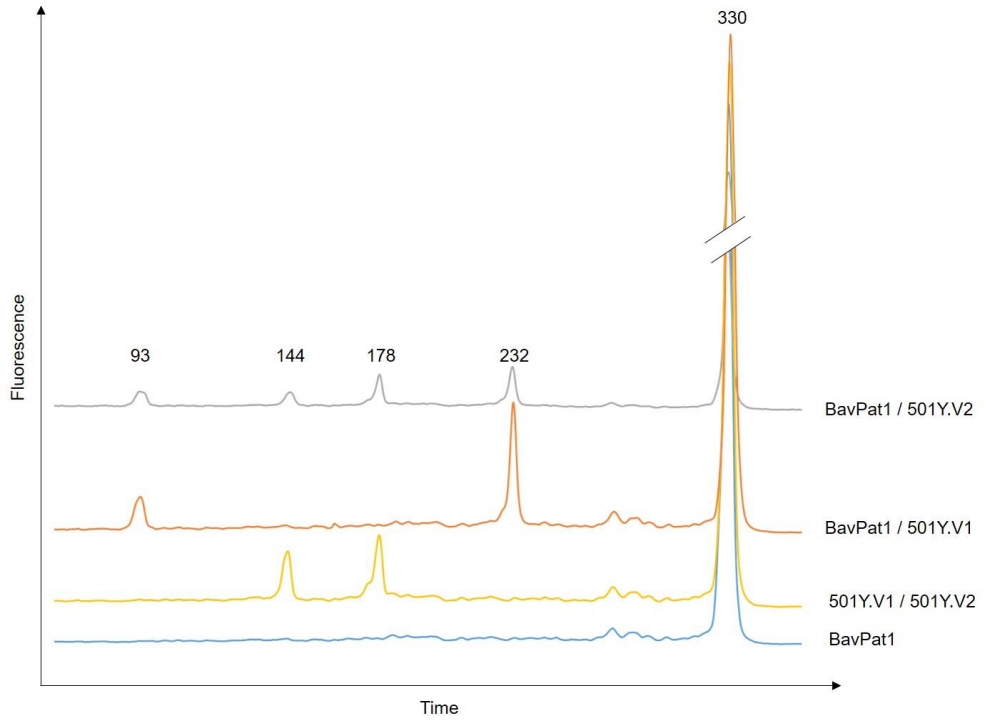
Zahradník, J., Marciano, S., Shemesh, M., Zoler, E., Chiaravalli, J., Meyer, B., Rudich, Y., Dym, O., Elad, N., Schreiber, G., 2021. SARS-CoV-2 RBD in vitro evolution follows contagious mutation spread, yet generates an able infection inhibitor. bioRxiv 2021.01.06.425392. https://doi.org/10.1101/2021.01.06.425392


A
S1 subunit
S2 subunit
NTD
RBD
SD1
SD2
FP
HR1
HR2
TM
CT
Spike protein
NH2
COOH
RBM
333
438
506
527
RBD/SD1 region
E484
N501
PCR product
431
524
B
N501
E484
C
Mixing, denaturation, hybridization
Mismatch-specific endonuclease
Electrophoresis
Sample
Sample PCR
Reference PCR
Ref / Sample
Sample
Ref / Sample


A
Linear regression for SARS-CoV-2 real time RT-PCR
36,00
34,00
32,00
30,00
28,00
26,00
24,00
22,00
20,00
CT value
R² = 0,9986
10⁰
10¹
10²
10³
10⁴
10⁵
RNA copy / µL


B
Amplification
10
8
6
4
2
0
RFU (10³)
0
10
20
30
40
Cycles


C
10⁰
10¹
10²
10³
10⁴
10⁵
10⁶
Ladder
bp
500
400
300
200
100


D
330
93
144
178
232
Fluorescence
BavPat1 / 501Y.V2
BavPat1 / 501Y.V1
501Y.V1 / 501Y.V2
BavPat1
Time

Samples	1	2	3	4	5	6	7	8
TaqPath COVID-19 kit™ , Ct values for each gene								
ORF1ab	24	21	25	22	30	23	27	34
S	24	nd	25	22	30	nd	27	34
N	23	18	24	22	30	22	27	34
Presence (+) or absence (-) of the deletion indicative of 501Y.V1								
501Y.V1	(-)	(+)	(-)	(-)	(-)	(+)	(-)	(-)