In vitro evaluation of the effect of mutations in primer binding sites on detection of SARS-CoV-2 by RT-qPCR

Fee Zimmermann, Maria Urban, Christian Krüger, Mathias Walter, Roman Wölfel, Katrin Zwirglmaier*

*Corresponding author katrinzwirglmaier@bundeswehr.org

Bundeswehr Institute of Microbiology, Neuherbergstr, 11, 80937 Munich, Germany

Phone +49 80 9926923933

Abstract

A number of RT-qPCR assays for the detection of SARS-CoV-2 have been published and are listed by the WHO as recommended assays. Furthermore, numerous commercial assays with undisclosed primer and probe sequences are on the market. As the SARS-CoV-2 pandemic progresses, the virus accrues mutations, which in some cases – as seen with the B.1.1.7 variant – can outperform and push back other strains of SARS-CoV-2. If mutations occur in primer or probe binding sites, this can impact RT-qPCR results and impede SARS-CoV-2 diagnostics. Here we tested the effect of primer mismatches on RT-qPCR performance in vitro using synthetic mismatch in vitro transcripts. The effects of the mismatches ranged from a shift in ct values from -0.13 to +7.61. Crucially, we found that a mismatch in the forward primer has a more detrimental effect for PCR performance than a mismatch in the reverse primer. Furthermore, we compared the performance of the original Charité RdRP primer set, which has several ambiguities, with a primer version without ambiguities and found that without ambiguities the ct values are ca. 3 ct lower. Finally, we investigated the shift in ct values observed with the Seegene Allplex kit with the B.1.1.7 SARS-CoV-2 variant and found a three-nucleotide mismatch in the forward primer of the N target.

Running Title

Effect of primer mismatches on SARS-CoV-2 RT-PCR

Keywords

SARS-CoV-2, RT-PCR, primer mismatch, diagnostics

1. Introduction

As part of the SARS-CoV-2 response the international scientific community rapidly published a number of RT-PCR diagnostic assays for SARS-CoV-2 detection. 15 RT-PCR assays of national reference laboratories are listed by the WHO for SARS-CoV-2 detection in patient samples (WHO, 2020). Although SARS-CoV-2 has a slower mutation rate than, for example, influenza (Manzanares-Meza and Medina-Contreras, 2020), it nevertheless has accrued mutations in most of the primer and probe binding sites of globally used PCR assays within a few months of the start of the pandemic. Aligning 177 genomes collected up to July 2020 from across Brazil with primer sequences of 15 of the WHO-listed SARS-CoV-2 assays revealed that only 3 assays (NIID_2019-nCoV-N, nCoV-IP4 and CN-CDC-E) had a perfect match to all 177 genomes (Santos et al., 2020). A similar study using sequence data from 375 genomes collected around the world up to April 2020 showed that only 2 out of 12 WHO-listed assays had a perfect match to all analyzed genomes (CDC-N1 and Charité RdRP) (Toms et al., 2020). Depending on location and nature of the mismatches this could have significant effects on assay performance and negatively impact sensitivity. The possible effect of primer mismatches in silico has been examined in a number of publications (Bru et al., 2008; Huang et al., 1992; Kwok et al., 1990; Lefever et al., 2013; Stadhouders et al., 2010). However, in vitro results don't always agree with in silico analyses, since various confounding factors can influence the outcome (Lefever et al., 2013; Mendelman et al., 1990; Stadhouders et al., 2010). Therefore, the aim of this study was to check the impact of in silico described primer mismatches on the in vitro performance of several of the WHO-listed SARS-CoV-2 RT-PCR assays. Furthermore, approaching the issue from the other side, we wanted to investigate the in silico basis of an in vitro observed dropout of a PCR target of a commercial SARS-CoV-2 assay.

2. Material and Methods

2.1. SARS-CoV-2 strains used in this study

We used extracted RNA from SARS-CoV-2 strain mucIMB-1 to generate the in vitro transcripts of the mismatch and perfect match templates. mucIMB-1 (EPI_ISL_406862) was isolated in January 2020 from a patient in Bavaria, Germany, and is phylogenetically close to the first published SARS-CoV-2 sequence, Wuhan-Hu1. For investigating the performance of several commercial SARS-CoV-2 PCR kits with SARS-CoV-2 variant B.1.1.7, we used strain mucIMB-CB (EPI_ISL-755639) in comparison with mucIMB-1. mucIMB-CB was isolated from a patient with a history of travel from the UK in December 2020.

2.2. Selection of primer mismatches

The European Centre for Disease Control (ECDC) provides an online tool "PrimerScan" that regularly screens all available SARS-CoV-2 genomes published on GISAID for mutations in the primer and probe binding sites of WHO recommended RT-PCR assays (https://primerscan.ecdc.europa.eu/). The website also plots the frequency of these mutations over time. Based on this information we chose several of the most frequent mutations between March and September 2020 for this study. Several of these mutations have since disappeared again.

2.3. Construction of mismatch templates

In order to generate RNA templates with the required mismatches in the primer binding sites, we used primers containing the mismatch (or complement base of the mismatch for reverse primers) (Supplementary table 1) and performed an RT-PCR with SARS-CoV-2 mucIMB-1 strain (which has no mismatch to any of the PCR assays tested here). For several mismatches we had to relax the

2

annealing temperature in order to be able to generate a PCR product (Table 1). Both forward and reverse primers contained a short nonsense sequence 5' of the primer sequence so as to generate a PCR product that is slightly longer than the template for the subsequent experiments. The purpose is to avoid artefacts due to primers binding right at the end of a template. The forward primers carried the T7 promotor at the 5' end. We then used in vitro transcription to produce the mismatch templates. Correct (i.e. no mismatch) templates were produced in the same way.

The detailed protocol is described below.

To generate the DNA templates needed for in vitro transcription PCR was performed on a Thermo™ Scientific Piko™ Thermal Cycler (Thermo Fisher Scientific, Waltham, MA, USA) using the QIAGEN One-Step RT-PCR Kit (QIAGEN, Hilden, Germany) in 50 µl reactions containing 10 µl 5x QIAGEN OneStep RT-PCR Buffer, 1.4 µl dNTP Mix (10 mM of each dNTP), 0.6 µl dNTP Mix dUTP Solution (premix of 10 mM dATP, dCTP, dGTP and 20 mM dUTP; Jena Bioscience GmbH, Jena, Germany), 1.5 µl Antarctic UDG (New England Biolabs, Ipswich, MA, USA) and 2 µl QIAGEN OneStep RT-PCR Enzyme Mix. 5 µl (2 x 10e4 copies/µl) SARS-CoV-2 RNA were used as template. The amplification conditions were 50 °C for 30 min, 95°C for 15 min, followed by 45 cycles of 94 °C, 15 sec and assay specific annealing temperature according to Table 2 for 30 sec and a final elongation at 72 °C for 30 sec. PCR products were analyzed by 2 % agarose gel electrophoresis. Bands corresponding to the correct fragment size were cut out and the DNA extracted using the QIAquick Gel Extraction Kit (QIAGEN, Hilden, Germany). The sequence was confirmed by Sanger sequencing.
RNA templates were produced by in vitro transcription using the MEGAscript T7 Kit (Thermo Fisher Scientific, Waltham, MA, USA) following the manufacturer's protocol including a TURBO DNase treatment step (15 min at 37°C). The in vitro transcripts (IVTs) were cleaned up using the NucleoSpin RNA Clean-Up Kit (Macherey-Nagel, Düren, Germany) according to manufacturer's instructions and subsequently quantified by using a Qubit RNA BR Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA). The copy number was calculated using the online NEBio Calculator (https://nebiocalculator.neb.com/#!/ssrnaamt).

2.4. RT-PCR with correct and mismatch IVTs

Dilution series (triplicates) of in vitro transcripts of the correct and mismatch templates were amplified by RT-qPCR using the published primers and following the published protocols (ref WHO and references in Supplementary Table 1). All experiments were done with Qiagen OneStep RT-PCR reagents (Qiagen, Hilden, Germany) on a MIC thermocycler (BMS, Brisbane, Australia).

2.5. Detecting SARS-CoV-2 variant B.1.1.7 with several commercial PCR kits

Extracted RNA of strains mucIMB-1 (wild type SARS-CoV-2) and mucIMB-CB (B.1.1.7 variant) was used with i) the Seegene Allplex SARS-CoV-2/FluA/FluB/RSV assay (Seegene, Seoul, Korea) and run on a Biorad CFX96DX cycler (Hercules, CA, USA), ii) the VitaPCR SARS-CoV-2 assay (Credo Diagnostics, Singapore) and iii) the GeneXpert Xpress SARS-CoV-2 assay (Cepheid, Sunnyvale, CA, USA). Each assay was performed according to the respective manufacturer's instructions.

2.6. Datamining

We downloaded a total of 600,443 sequences and their corresponding metadata from GISAID at 2021-02-24 10:05. Then we used the search_oligodb command from usearch (Edgar, 2010) to search for the non-degenerated primer sequences (see tables 2a, 3a, 4a) in every SARS-CoV-2 genome sequence with a maximum divergence of three (parameter: -maxdiffs 3). The result sets

were enriched with the corresponding pangolin lineage (Rambaut et al., 2020) obtained from the metadata.tsv and statistics were generated using csvtk (Shen, 2021). For some strains, no pangolin lineage was provided with the metadata.tsv. In such case, we classified the strains using pangolin client v2.3 (O'Toole et al., 2021).

3. Results

3.1. Mismatches in WHO-listed primers

Based on the ECDC PrimerScan tool (https://primerscan.ecdc.europa.eu) we chose 14 mismatches in primer binding sites of six different WHO-listed RT- qPCR assays. We generated in vitro transcripts of the mismatch templates and then performed RT-qPCR with the published primers, comparing the ct values obtained with the "correct" templates and the "mismatch" templates (Table 1 and Figure 1). The effect of the mismatch on the ct values ranged from -0.13 to +7.61. Within a dilution series ranging from 10e5 copies/reaction to 10e0 copies/reaction the observed shift in ct values between the correct and the mismatch template was consistent across all dilutions, i.e. independent of the template concentration. Regarding the distance of the mismatch from the 3' end of the primer and the magnitude of the ct shift, there was only a weak correlation (Pearson -0.320) for the entire dataset. However, when analyzing forward and reverse primer separately, we found a strong negative correlation for mismatches in the forward primer (Pearson -0.822) and no correlation for the reverse primer (Pearson 0.004). Analyzing the data from a recent study on primer mismatches in SARS-CoV-2 RT-qPCR assays (Nakabayashi et al., 2021) under this aspect, we found the same trend, i.e. a strong negative correlation between the distance of the mismatch from the 3' end and the magnitude of the ct shift for the forward primer (Pearson -0.576) and only a weak correlation for the reverse primer (Pearson -0.159).

With regard to the mismatch type, there was no consistent pattern. A G:T mismatch is thought to be generally well tolerated in PCR (Kwok et al., 1990; Rejali et al., 2018; Stadhouders et al., 2010). However, in our study a G:T mismatch resulted in a large delta ct of 6.79 at a distance of 10 bases from the 3' end (CDC-N1-MM R15) in one instance and a delta ct of only 2.3 at a distance of 2 bases from the 3' end (IP2-MM R17) in another instance.

3.2. Charité RdRP primer set

The original Charité primer set for RdRP (Corman et al., 2020) carries three ambiguous positions, one in the forward primer (position 5) and two in the reverse primer (positions 3 and 12 (Supplementary table 1). We conducted an in silico analysis of the primer binding sites using > 600,000 SARS-CoV-2 genomes published on GISAID (as of February 2021). 99.40 % of genomes carry an A at position 5 in the forward primer and 99.97 % carry a T (i.e. an A in the primer) at both positions 3 and 12 of the reverse primer (Tables 2a and 3a). We therefore used primers with the dominant sequence (designated RdRP-SARSr-F2.new and RdRP-SARSr.R1.new) and compared their performance to primers containing the ambiguities. Both the correct template as well as a template including a mismatch in the binding site of the forward primer, RdRP MM F8, yielded PCR products with lower ct values (-3.16 for correct template, -3.60 for MM F8) with the .new primers compared to the original Charité primers (Fig. 1a). In both primer sets (Charité and .new) the difference in ct values between the correct template and mismatch template was negligible (Table 1), indicating that the mismatch has no pronounced effect on amplification.

4

Further in silico analysis showed that for the reverse primer sequence only 23 respectively 15 out of 600,443 SARS-CoV-2 genomes carry a mismatch to the .new primer sequence in positions 3 respectively 12. The mismatches are spread over eight (position 3) and nine (position 12) lineages with no clear cluster in any one lineage (Table 3b). No genome carries both mismatches at the same time. For the forward primer 44 genomes carry a C and 36 a G at position 5 of the primer. The C mismatch only appears in lineage B.1.177 (0.052 % of B.1.177 genomes), while the G mismatch is spread over 9 lineages. A more dominant mismatch is the mismatch in position 8 of the forward primer, mentioned above. This mismatch was investigated in vitro, as 2,050 genomes carried a T, instead of a G (0.34 % of all genomes). This appears in 41 different lineages and is dominant (i.e. present in > 98 % of genomes) in 5 lineages (Table 2b).

3.3. SARS-CoV-2 variant B.1.1.7 amplified with several commercial PCR assays

We amplified RNA from wild type SARS-CoV-2 and variant B.1.1.7 with the GeneXpert Xpress SARS-CoV-2 assay, Vita PCR SARS-CoV-2 assay and Seegene Allplex SARS-CoV-2/FluA/FLuB/RSV assay. According to the manufacturers, GeneXpert targets the N and E gene, VitaPCR has two N gene targets and Seegene has targets in the RdRP, S and N gene. Neither GeneXpert nor VitaPCR showed any difference in ct values between wild type and B.1.1.7 (data not shown). The Seegene assay, however, showed a drop in ct values of 7.2 ct in the N target, but not in the S or RdRP target. The drop in ct values was consistent across template concentrations between 10 and 10e5 copies/reaction and led to a loss of sensitivity of 2 log for the N target (Fig 2). We cloned and sequenced the PCR product to identify the mismatch causing the drop in PCR performance. Aligning the PCR amplicon to SARS-CoV-2 WuhanHu1 and B.1.1.7 revealed a triple mismatch in the forward primer (GAT => CTA), corresponding to SARS-CoV-2 genome positions 28,280-28,282 at the start of the N gene (Fig. 3). This mismatch currently (as of February 2021) has an overall frequency of 16.55 % in 600,000 genomes on gisaid.org and appears in 98.74 % of all B.1.1.7 genomes (UK variant) and 57.23 % of all B.1.214 genomes (DRC variant) as well as three further lineages with < 1 % frequency (Table 4).

4. Discussion

4.1. Effect of primer mismatches

Several factors can influence the impact of a mismatch on PCR amplification. One intuitive and well documented factor is the proximity of the mismatch to the 3' end of the primer (Bru et al., 2008; Kwok et al., 1990; Stadhouders et al., 2010). Interestingly, in our study this correlation was not obvious when looking at the whole dataset, but was only true for mismatches in the forward primers. SARS-CoV-2 is a single-stranded positive-sense RNA virus, which means that in an RT-PCR reaction, the first step, the reverse transcription, involves the binding of the reverse primer to the sense RNA strand. The forward primer is only used during the subsequent PCR step. Studies from the early days of PCR noted that reverse transcriptases are more tolerant towards mismatches than DNA polymerases (Huang et al., 1992; Mendelman et al., 1990). Therefore, in an RT-PCR reaction, the reverse transcriptase is more likely to be able to create a cDNA strand from a mismatched primer, correcting the mismatch in the process, while the Taq polymerase is more discriminant against mismatches, resulting in decreased amplification. Thus, in the case of SARS-CoV-2, a mismatch in the forward primer should have a stronger negative impact than in the reverse primer. Our data confirms this, albeit with a limited number of data points. For SARS-CoV-2 RT-PCR assays, this hypothesis is also supported by data from Nakabayashi et al. (Nakabayashi et al., 2021). In an earlier study using HIV-RT-PCR assays, Stadhouders et al. (Stadhouders et al., 2010) made a similar observation. They also noted that when using rTth polymerase (an enzyme that has both reverse transcription and DNA polymerase functionality) instead of a combination of MMLV reverse transcriptase and Taq polymerase for the RT-PCR reaction, mismatches in the forward primer had almost no effect, but mismatches in the reverse primer had a detrimental effect. The authors hypothesize that the different incubation temperatures of MMLV and rTth of 60 °C and 48 °C as well as differences in ionic strength of the reaction buffers contribute to this phenomenon. The varying influence of different PCR reagents on mismatches was also observed by Levefer et al. (Lefever et al., 2013).

Another factor impacting PCR amplification is the mismatch type. A transversion, i.e. resulting in a mispairing of two pyrimidines (C:C, T:T, C:T, T:C) or two purines (A:A, G:G, A:G, GA) is thought to have a greater impact than a transition, i.e. resulting in a mispairing of a pyrimidine and a purine (A:C, C:A, T:G, G:T), mainly due to sterical reasons (Huang et al., 1992; Rejali et al., 2018; Stadhouders et al., 2010). This trend is not obvious in our dataset, particularly illustrated by the example of a G:T mismatch, which had a small effect in close proximity to the 3' end but a rather large effect further away. A recent study on primer mismatches in SARS-CoV-2 RT-PCR assays has shown the effect from a transition mismatch type in the ultimate 3' position of a primer to result in a shift in ct values ranging from 0.51 to 6.29 up to a complete PCR failure (Nakabayashi et al., 2021). Taken together, this indicates that not only the mismatch type and position, but also the surrounding sequence of the template and primer influence the outcome, presumably due to secondary structures and interactions based on these (Mendelman et al., 1990; Rejali et al., 2018). Consequently, this serves as a warning that in silico analysis of the effect of primer mismatches can only give an estimate on the impact on PCR performance and should always be tested in vitro.

4.2. RdRP primers

The Charité RdRP assay (Corman et al., 2020) was published on 23th Jan 2020, when only the original Wuhan-Hu1 sequence of SARS-CoV-2 was available as a reference. Based on alignments with human SARS-CoV-1 and bat-related SARS-CoV several ambiguities were introduced in the primer sequences. Particularly the base at position 12 from the 5' end of the reverse primer has sparked some

discussion since. Corman et al. inserted an S (= C or G) at this position. Pillonel (Pillonel et al., 2020), after noticing a poorer performance of the RdRP assay compared to the E assay and comparing 1623 SARS-CoV-2 genome sequences available by May 2020, suggested using an R (= A or G), while the ECDC PrimerScan website lists an ECDC-corrected version of the primer with a W (= A or T). Vogels (Vogels, 2020) also noted the poor performance of the RdRP assay and stated that 990 of 992 genomes available at the time carried a T at position 12 of the reverse primer. More than a year into the pandemic and with over 600,000 SARS-CoV-2 genomes sequenced it turns out that 99.97 % carry a T at position 12 as well as position 3, where the Charité reverse primer has the second ambiguity. Since the use of ambiguous primer sequences essentially reduces the amount of "correct" primer for the template (with two ambiguities in the reverse primer, R and S, this results in four primer variants, i.e. only 25 % of the reverse primers in the PCR reaction have the correct sequence) this would account for the poor performance of the primer. Using the primer without ambiguities, with A (to complement the T on the template) in positions 3 and 12 improves the sensitivity of the assay by about 1 log.

For the forward primer the situation is more complex. Only a minute fraction of genomes (0.01 %) carry a G instead of an A at position 5, which would be covered by the ambiguity in the Charité primer sequence. However, we discovered another mismatch in position 8 which, although only present in 0.34 % of the total genome dataset, is the dominant variant in several smaller lineages. This highlights the problem that primer mismatches in localized outbreaks may go unnoticed, while still causing problems with diagnostic assays locally.

4.3. Commercial SARS-CoV-2 RT-PCR assays

The 7.2 ct loss in the N target of the Seegene Allplex assay we observed with the B.1.1.7 variant exemplifies the problem with commercial SARS-CoV-2 PCR assays. With proprietary primer and probe sequences the users cannot (or only with considerable effort) check for themselves whether new SARS-CoV-2 variants affect the performance of the assay. They have to rely on communication of manufacturers or serendipitous findings. Other examples of PCR failures with commercial assays were described by Ziegler et al (Ziegler et al., 2020), who found a drop out of the N target in GeneXpert (Cepheid) in one patient sample. Analysis of the sequence revealed a point mutation in the N gene and comparison with GISAID data showed a frequency of 0.2 % for that SNP. Furthermore, Artesi et al. (Artesi et al., 2020) noted a drop out of the E gene target in the Cobas system (Roche) in 0.2 % of SARS-CoV-2 positive samples from Belgian healthcare workers in March/April 2020. The identified SNP at position 26,340 in the E gene was present in just 0.09 % of global SARS-CoV-2 genomes on GISAID. This reiterates the problem pointed out in the last paragraph – not all mutations are as successful as B.1.1.7 and many appear only within a limited geographical region. Another example is a cluster of SNPs in the N gene found in Madeira County, California, USA, that resulted in a loss of 5.4 ct in the Japan N assay NIID_2019-nCoV (Vanaerschot et al., 2021). These regionally restricted mutations may also disappear over time, but while they exist, they can impede diagnostics. This is particularly problematic in countries that do not have the capacity to retest samples with multiple assays or follow up suspicious samples with whole genome sequencing.

On the other hand, a drop out of a PCR target can also be used to screen for a certain mutation, provided that at least a second target is used to confirm SARS-CoV-2. An example is the drop out of the S gene target in the Thermo Fisher TaqPath assay, which is widely used across the UK. It guided and supported the genome sequencing effort that documented the fast spread of variant B.1.1.7 in Britain (Chand et al., 2020).

4.5. Conclusions

This in vitro study provides data for the effect of primer mismatches on PCR performance. It highlights the fact that experimental data does not necessarily follow the theoretical predictions, particularly with regard to the magnitude of the ct shift with mismatches close to the 3' end. This emphasizes the importance of using more than one target in a diagnostic SARS-CoV-2 RT-PCR (or any other diagnostic PCR for that matter) to counteract PCR dropouts due to mutations. With several targets a PCR drop out or relative shift in ct values compared to the other targets used can actually serve as an indication for a mutation, prompting further investigation of the sample by genome sequencing. Furthermore, it emphasizes the importance to remain vigilant with regard to new mutations and their potential effect on PCR performance. In this context it is also important to keep in mind the underrepresentation of genome data from some geographical regions (e.g. Africa), which may lead to mutations from these regions being overlooked, despite causing problems in diagnostics in these regions.

5. Acknowledgements

6. Author contributions

Fee Zimmermann – conceptualization, methodology, formal analysis, visualization, writing: review and editing, supervision , Maria Urban – investigation, validation, writing: review and editing, Christian Krüger - investigation, Mathias Walter - software, writing: review and editing, Roman Wölfel – formal analysis, writing: review and editing, Katrin Zwirglmaier – conceptualization, methodology, formal analysis, writing: original draft, supervision

7. References

Artesi, M., Bontems, S., Göbbels, P., Franckh, M., Maes, P., Boreux, R., Meex, C., Melin, P., Hayette, M.-P., Bours, V., Durkin, K., 2020. A Recurrent Mutation at Position 26340 of SARS-CoV-2 Is Associated with Failure of the E Gene Quantitative Reverse Transcription-PCR Utilized in a Commercial Dual-Target Diagnostic Assay. J Clin Microbiol 58, e01598-20, /jcm/58/10/JCM.01598-20.atom. https://doi.org/10.1128/JCM.01598-20

Bru, D., Martin-Laurent, F., Philippot, L., 2008. Quantification of the Detrimental Effect of a Single Primer-Template Mismatch by Real-Time PCR Using the 16S rRNA Gene as an Example. AEM 74, 1660–1663. https://doi.org/10.1128/AEM.02403-07

Chand, M., Hopkins, S., Dabrera, G., Achison, C., Barclay, W., Ferguson, N., Volz, E., Loman, N., Rambaut, A., Barrett, J., 2020. Investigation of novel SARS-COV_2 variant VOC2020-12-01, PHE Technical Briefing. Public Health England.

Corman, V.M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D.K., Bleicker, T., Brünink, S., Schneider, J., Schmidt, M.L., Mulders, D.G., Haagmans, B.L., van der Veer, B., van den Brink, S., Wijsman, L., Goderski, G., Romette, J.-L., Ellis, J., Zambon, M., Peiris, M., Goossens, H.,

Reusken, C., Koopmans, M.P., Drosten, C., 2020. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Eurosurveillance 25. https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045

Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26, 2460–2461. https://doi.org/10.1093/bioinformatics/btq461

Huang, M.-M., Arnheim, N., Goodman, M.F., 1992. Extension of base mispairs by *Taq* DNA polymerase: implications for single nucleotide discrimination in PCR. Nucl Acids Res 20, 4567–4573. https://doi.org/10.1093/nar/20.17.4567

Kwok, S., Kellogg, D.E., McKinney, N., Spasic, D., Goda, L., Levenson, C., Sninsky, J.J., 1990. Effects of primer-template mismatches on the polymerase chain reaction: Human immunodeficiency virus type 1 model studies. Nucl Acids Res 18, 999–1005. https://doi.org/10.1093/nar/18.4.999

Lefever, S., Pattyn, F., Hellemans, J., Vandesompele, J., 2013. Single-Nucleotide Polymorphisms and Other Mismatches Reduce Performance of Quantitative PCR Assays 11.

Manzanares-Meza, L.D., Medina-Contreras, O., 2020. SARS-CoV-2 and influenza: a comparative overview and treatment implications. BMHIM 77, 4621. https://doi.org/10.24875/BMHIM.20000183

Mendelman, L.V., Petruska, J., Goodman, M.F., 1990. Base mispair extension kinetics. Comparison of DNA polymerase alpha and reverse transcriptase. Journal of Biological Chemistry 265, 2338–2346. https://doi.org/10.1016/S0021-9258(19)39981-8

Nakabayashi, T., Kawasaki, Y., Murashima, K., Omi, K., Yuhara, S., 2021. Evaluation of the effects of SARS-CoV-2 genetic mutations on diagnostic RT-PCR assays (preprint). Molecular Biology. https://doi.org/10.1101/2021.01.19.426622

Nao, N., Shirato, K., Katano, H., Matsuyama, S., Takeda, M., 2020. Detection of second case of 2019-nCoV infection in Japan (corrected version). NIID - National Institute of Infectious Diseases Japan.

O'Toole, À., McCrone, J., Scher, E., 2021. Pangolin – A Software Package for Assigning SARS-CoV-2 Genome Sequences to Global Lineages.

Pillonel, T., Scherz, V., Jaton, K., Greub, G., Bertelli, C., 2020. Letter to the editor: SARS-CoV-2 detection by real-time RT-PCR. Eurosurveillance 25. https://doi.org/10.2807/1560-7917.ES.2020.25.21.2000880

Rambaut, A., Holmes, E.C., O'Toole, Á., Hill, V., McCrone, J.T., Ruis, C., du Plessis, L., Pybus, O.G., 2020. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nat Microbiol 5, 1403–1407. https://doi.org/10.1038/s41564-020-0770-5

Rejali, N.A., Moric, E., Wittwer, C.T., 2018. The Effect of Single Mismatches on Primer Extension. Clinical Chemistry 64, 801–809. https://doi.org/10.1373/clinchem.2017.282285

Santos, R. da S., Bret, R.S.C., Moreira, A.C. de O.M., Campos, A.R., Silva, A.R.A. e, Lima, D.M., Tavares, K.C.S., 2020. In silico analysis of mismatches in RT-qPCR assays of 177 SARS-CoV-2 sequences from Brazil. Rev. Soc. Bras. Med. Trop. 53, e20200657. https://doi.org/10.1590/0037-8682-0657-2020

Shen, W., 2021. Csvtk – A Cross-Platform, Efficient and Practical CSV/TSV Toolkit in Golang.

Stadhouders, R., Pas, S.D., Anber, J., Voermans, J., Mes, T.H.M., Schutten, M., 2010. The Effect of Primer-Template Mismatches on the Detection and Quantification of Nucleic Acids Using the 5 Nuclease Assay. The Journal of Molecular Diagnostics 12, 109–117. https://doi.org/10.2353/jmoldx.2010.090035

Toms, D., Li, J., Cai, H.Y., 2020. Evaluation of WHO listed COVID-19 qPCR primers and probe in silico with 375 SERS-CoV-2 full genome sequences (preprint). Infectious Diseases (except HIV/AIDS). https://doi.org/10.1101/2020.04.22.20075697

US-CDC, C. for D.C. and P., 2020. 2019-Novel Coronavirus (2019-nCoV) Real-time rRT-PCR Panel Primers and Probes. URL https://www.cdc.gov/coronavirus/2019-ncov/lab/rt-pcr-panel-primer-probes.html

Vanaerschot, M., Mann, S.A., Webber, J.T., Kamm, J., Bell, S.M., Bell, J., Hong, S.N., Nguyen, M.P., Chan, L.Y., Bhatt, K.D., Tan, M., Detweiler, A.M., Espinosa, A., Wu, W., Batson, J., Dynerman, D., Consortium, C., Wadford, D.A., Puschnik, A.S., Neff, N., Ahyong, V., Miller, S., Ayscue, P., Tato, C.M., Paul, S., Kistler, A.L., DeRisi, J.L., Crawford, E.D., 2021. Identification of a Polymorphism in the N Gene of SARS-CoV-2 That Adversely Impacts Detection by Reverse Transcription-PCR. Journal of Clinical Microbiology 59, 4.

Vogels, C.B.F., 2020. Analytical sensitivity and efficiency comparisons of SARS-CoV-2 RT–qPCR primer–probe sets. Nature Microbiology 5, 13.

WHO, 2020. Molecular assays to diagnose COVID-19: Summary table of available protocols. URL https://www.who.int/publications/m/item/molecular-assays-to-diagnose-covid-19-summary-table-of-available-protocols

Ziegler, K., Steininger, P., Ziegler, R., Steinmann, J., Korn, K., Ensser, A., 2020. SARS-CoV-2 samples may escape detection because of a single point mutation in the N gene. Eurosurveillance 25. https://doi.org/10.2807/1560-7917.ES.2020.25.39.2001650

Figures and Tables

Figures

Fig. 1. PCR results using the published primers with the correct templates and mismatched templates. Dilution series were done in triplicate. Size of the symbols represents one (small), two (medium sized) or three (large) positive replicates.

Fig. 2. PCR results using the Seegene Allplex assay with SARS-CoV-2 wild type and variant B.1.1.7. Dilution series were done in triplicate. Size of the symbols represents one (small), two (medium sized) or three (large) positive replicates.

Fig. 3. Alignment of the 5' end of the amplicon of the N-assay of the Seegene Allplex against WuhanHu1 and B1.1.7

Tables

Tab. 1. Effect of mismatches in template on PCR. * RdRP primer RdRP_SARSr-F2/RdRP_SARSr-R1.corrECDC (with ambiguities), § RdRP primers RdRP_SARSr-F2.new/RdRP_SARSr-R1.new (without ambiguities)

Table 2. Mismatches of Charité RdRP forward primer in 600 000 published SARS-CoV-2 genomes.

Table 2a. Number of sequences with mismatches in specific positions

Table 2b. Lineage analysis of the mismatches. Lineage information retrieved from Pangolin lineages (https://cov-lineages.org/) on 15.03.2021. Only lineages where > 10 sequences carry the mutation are shown

Table 3. Mismatches of Charité RdRP reverse primer in 600 000 published SARS-CoV-2 genomes.

Table 3a. Number of sequences with mismatches in specific positions

Table 3b. Lineage analysis of the mismatches. Lineage information retrieved from Pangolin lineages (https://cov-lineages.org/) on 15.03.2021. Only lineages where > 3 sequences carry the mutation are shown

Table 4. Triple mismatch in the Seegene Allplex N-assay in genome position 28280 -28282

Table 4a. Number of sequences with the triple mismatch

Table 4b. Lineage analysis of the mismatches. Lineage information retrieved from Pangolin lineages (https://cov-lineages.org/) on 15.03.2021. Only lineages where > 10 sequences carry the mutation are shown

Supplementary information

Supplementary Table. 1. Primer and probe sequences used in this study.
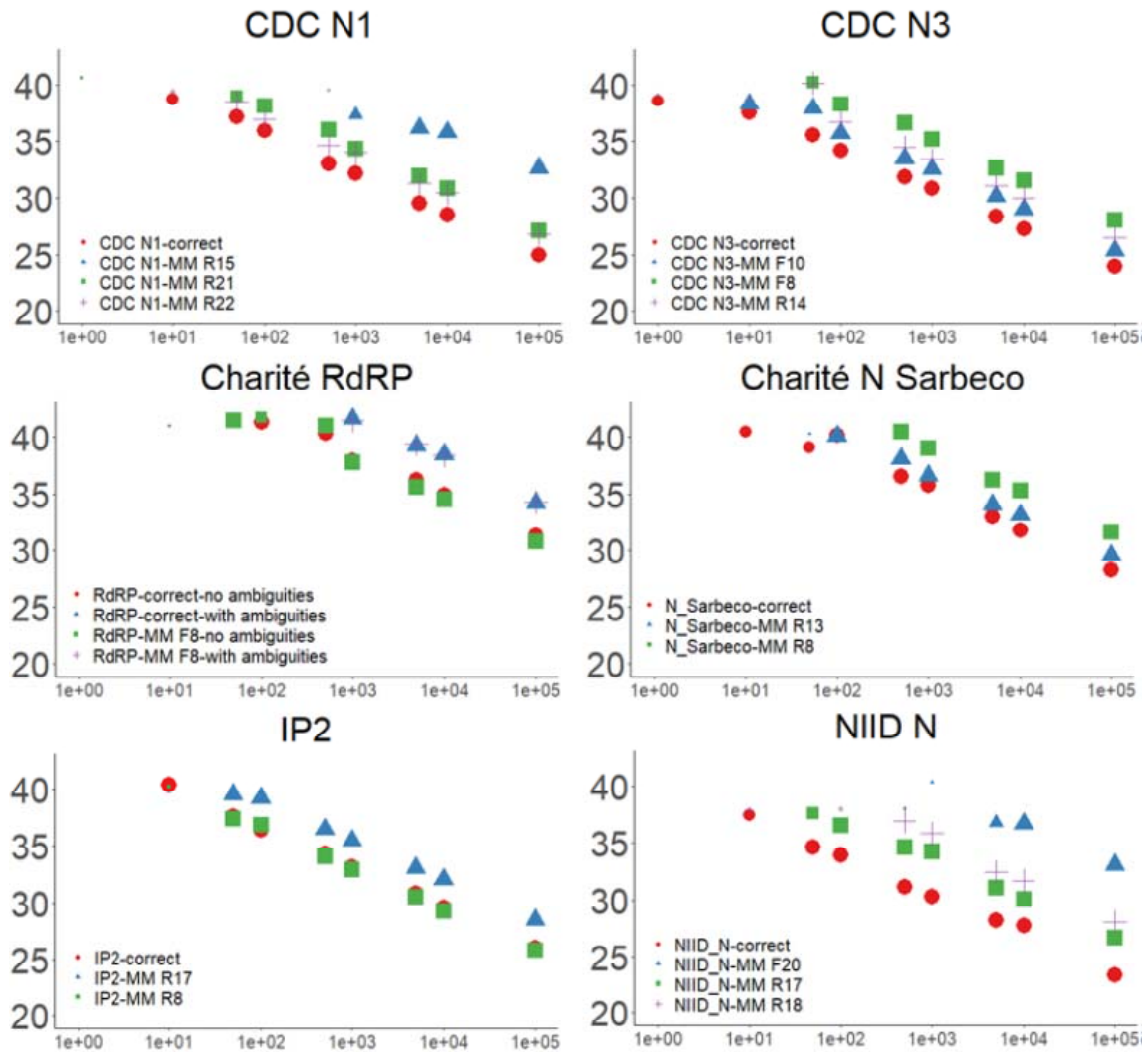
Fig. 1

PCR results using the published primers with the correct templates and mismatched templates. Dilution series were done in triplicate. Size of the symbols represents one (small), two (medium sized) or three (large) positive replicates.

Fig. 2. PCR results using the Seegene Allplex assay with SARS-CoV-2 wild type and variant B.1.1.7. Dilution series were done in triplicate. Size of the symbols represents one (small), two (medium sized) or three (large) positive replicates.

13
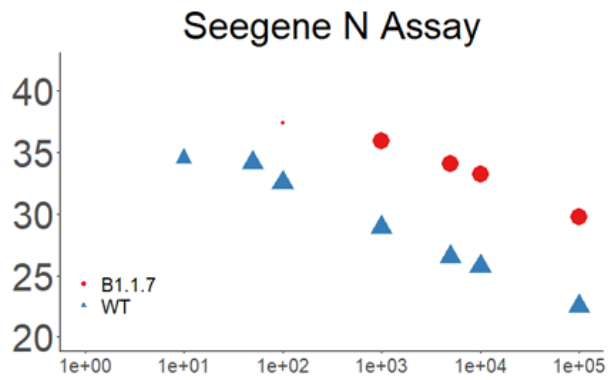
Fig. 3. Alignment of the 5' end of the amplicon of the N-assay of the Seegene Allplex against SARS-CoV-2 WuhanHu1 and B1.1.7

```
                        28270     28280     28290     28300     28310     28320     28330
                        .|....|....|....|....|....|....|....|....|....|....|....|....|....|....
Seegene N amplicon      ------TGTCTGATAATGGACCCCAAAATCAGCGAAATGCACCCCGCATTACGTTTGGTGGACCCTCAGAT
MN908947.3(WuhanHu1)    CTAAAATGTCTGATAATGGACCCCAAAATCAGCGAAATGCACCCCGCATTACGTTTGGTGGACCCTCAGAT
EPI_ISL_755639(B1.1.7)  CTAA-ATGTCTCTAAATGGACCCCAAAATCAGCGAAATGCACCCCGCATTACGTTTGGTGGACCCTCAGAT
```

14

Tab. 1. Effect of mismatches in template on PCR. * RdRP primer RdRP_SARSr-F2/RdRP_SARSr-R1.corrECDC (with ambiguities), § RdRP primers RdRP_SARSr-F2.new/RdRP_SARSr-R1.new (without ambiguities)

Mismatch	max fequency % /date	Mutation	Mismatch type	distance from 3' end	delta ct compared to wildtype	Annealing temp used to generate mismatch template	
	data from ECDC PrimerScan		primer : mismatch IVT			normal PCR	to generate mismatch template
RdRP MM F8	0.46 /06-2020	G => T	G : A	15	-0.01*/+0.16 §	58	58
N_Sarbeco MM R8	1.37 /06-2020	C => A	G : A	13	3.48	58	56
N_Sarbeco MM R13	2.40 /06-2020	C => A	G : A	8	1.05	58	56
CDC N1 MM R15	0.17 /03-2020	C => T	G : T	10	6.79	55	55
CDC N1 MM R21	0.50 /08-2020	T => G	A : G	4	2.31	55	55
CDC N1 MM R22	n.d.	G => T	C : T	3	1.54	55	55
CDC N3 MM F8	4.18 /02-2020	T => C	T : G	15	4.37	55	55
CDC N3 MM R14	0.17 /02-2020	G => T	C : T	8	2.86	55	55
CDC N3 MM F10	0.39 /06-2020	G => T	C : T	13	1.6	55	55
Japan N MM F20	0.27 /07-2020	C => T	C : T	1	7.61	60	60
Japan N-neu MM R17	0.50 /05-2020	G => C	C : C	4	3.44	60	51
Japan N-neu MM R18	0.64 /01-2020	T => G	A : G	3	5.25	60	51
IP2 MM R8	n.d.	A => G	T : G	11	-0.13	58	56
IP2 MM R17	0.45 /06-2020	C => T	G : T	2	2.3	58	58

Table 2. Mismatches of Charité RdRP forward primer in 600 000 published SARS-CoV-2 genomes.

Table 2a. Number of sequences with mismatches in specific positions

5' – 3'		
GTGA**R**ATGGTCATGTGTGGCGG	RdRP_SARSr-F2	
GTGA**A**ATGGTCATGTGTGGCGG	RdRP_SARSr-F2.new	
GTGAAATGGTCATGTGTGGCGG	SARS-CoV-2 ref seq.	
	Number of sequences	%
	600443	100
GTGAAATGGTCATGTGTGGCGG		
....	596848	99.401
....C.................	44	0.007
....G.................	36	0.006
.......T.............	2050	0.341

Table 2b. Lineage analysis of the mismatches. Lineage information retrieved from Pangolin lineages (https://cov-lineages.org/) on 15.03.2021. Only lineages where > 10 sequences carry the mutation are shown

mismatch	sequences	% of lineage	Pangolin lineage	Most common countries	Earliest date	Number designated	Number assigned	Description
....C.................	44	0.055	B.1.177	United Kingdom 68.0%, Denmark 8.0%, Switzerland 4.0%, Spain 4.0%, Netherlands 3.0%	14.02.2020	25431	84265	Spain to Europe
....G.................	18	0.022	B.1.177	United Kingdom 68.0%, Denmark 8.0%, Switzerland 4.0%, Spain 4.0%, Netherlands 3.0%	14.02.2020	25431	84265	Spain to Europe
.......T.............	598	100.000	B.1.36.18	Canada 98.0%, New Zealand 1.0%, United Kingdom 0.0%, India 0.0%, United States of America 0.0%	30.08.2020	42	339	Canadian lineage
.......T.............	561	98.768	B.1.1.50	Israel 80.0%, United Kingdom 8.0%,	29.03.2020	119	689	UK/ Israel/ Palestine lineage

				Palestine 8.0%, Denmark 1.0%, United States of America 1.0%				
......T..............	252	98.438	B.1.223	United Kingdom 81.0%, Spain 10.0%, Ecuador 3.0%, Poland 2.0%, Netherlands 2.0%	16.03.2020	238	254	Previously B.1.5.6, Spain/ UK lineage
......T..............	166	0.300	B.1	United States of America 48.0%, United Kingdom 12.0%, Canada 4.0%, Spain 4.0%, France 3.0%	24.01.2020	19976	55060	A large European lineage that corresponds to the Italian outbreak.
......T..............	157	62.800	B.1.146	United Kingdom 65.0%, Switzerland 26.0%, Luxembourg 2.0%, Austria 2.0%, Germany 1.0%	23.06.2020	134	248	Swiss lineage
......T..............	103	100.000	B.1.411	Sri Lanka 94.0%, Australia 2.0%, United Kingdom 2.0%, Israel 1.0%, Singapore 1.0%	03.10.2020	12	101	Sri Lanka lineage
......T..............	39	0.110	B.1.2	United States of America 97.0%, Canada 2.0%, Denmark 0.0%, United Kingdom 0.0%, Australia 0.0%	28.02.2020	3875	42269	USA lineage
.......T.............	31	0.637	B.1.243	United States of America 98.0%, Mexico 1.0%, Canada 0.0%, Switzerland 0.0%, South Korea 0.0%	23.03.2020	1090	5286	USA lineage
......T..............	31	0.037	B.1.177	United Kingdom 68.0%, Denmark 8.0%, Switzerland 4.0%, Spain 4.0%, Netherlands 3.0%	14.02.2020	25431	84265	Spain to Europe
......T.............	23	1.139	B.1.36.17	United Kingdom 97.0%, Canada 1.0%, Switzerland 1.0%, United States of America 0.0%, Denmark 0.0%	16.07.2020	1485	2007	UK lineage
......T..............	18	0.511	B.1.1	United States of America 26.0%, United Kingdom 21.0%, Bangladesh 6.0%, Canada 5.0%, Germany 5.0%	24.02.2020	280	3495	Lineage reassigned. Reassigned in the current tree. European lineage with 3 clear SNPs `28881GA`,`28882GA`,`28883GC`
.......T..............	11	100.000	F.1	New Zealand 64.0%, Hong Kong 18.0%, Canada 9.0%, China 9.0%	06.09.2020	6	11	Alias of B.1.36.17.1, New Zealand lineage

17

Table 3. Mismatches of Charité RdRP reverse primer in 600 000 published SARS-CoV-2 genomes.

Table 3a. Number of sequences with mismatches in specific positions

5' - 3'		
CA**R**ATGTTAAA**S**ACACTATTAGCATA	RdRp_SARSr-R	
CA**R**ATGTTAAA**W**ACACTATTAGCATA	RdRP_SARSr-R1.corrECDC	
CA**A**ATGTTAAA**A**ACACTATTAGCATA	RdRP_SARSr-R1.new	
CAAATGTTAAAAACACTATTAGCATA	SARS-CoV-2 ref seq. (rev. comp.)	
	Number of sequences	%
	600443	100
....	600272	99.972
..G......................	20	0.003
..C......................	2	0.000
..T......................	1	0.000
..........G..............	10	0.002
..........C..............	5	0.001

Table 3b. Lineage analysis of the mismatches. Lineage information retrieved from Pangolin lineages (https://cov-lineages.org/) on 15.03.2021. Only lineages where > 3 sequences carry the mutation are shown

mismatch	Sequences	% of lineage	Pangolin lineage	Most common countries	Earliest date	Number designated	Number assigned	Description
..G......................	15	0.015	B.1.1.7	UK 68.0%, Denmark 8.0%, Switzerland 4.0%, Spain 4.0%, Netherlands 3.0%	14.02.2020	25431	84265	Spain to Europe

Table 4. Triple mismatch in the Seegene Allplex N-assay in genome position 28280 -28282

Table 4a. Number of sequences with the triple mismatch

	Number of sequences	%
	594926	100
......GAT........	494713	83.16
......CTA........	98488	16.55

Table 4b. Lineage analysis of the mismatches. Lineage information retrieved from Pangolin lineages (https://cov-lineages.org/) on 15.03.2021. Only lineages where > 10 sequences carry the mutation are shown

mismatch	sequences	% of lineage	Pangolin lineage	Most common countries	Earliest date	Number designated	Number assigned	Description
......CTA........	98296	98.74	B.1.1.7	United Kingdom 68.0%, Denmark 8.0%, Switzerland 4.0%, Spain 4.0%, Netherlands 3.0%	14.02.2020	25431	84265	Spain to Europe
......CTA........	91	57.23	B.1.214	Belgium 31.0%, Denmark 24.0%, Democratic Republic of the Congo 15.0%, France 10.0%, United Kingdom 5.0%	18.04.2020	42	241	DRC lineage
......CTA........	35	0.17	B.1.1.29	United Kingdom 46.0%, United States of America 14.0%, Denmark 4.0%, Portugal 4.0%, Russia 3.0%	16.02.2020	110	20096	(expanded a bit) UK and Ireland
......CTA........	23	0.65	B.1.1	United States of America 26.0%, United Kingdom 21.0%, Bangladesh 6.0%, Canada 5.0%, Germany 5.0%	24.02.2020	280	3495	Lineage reassigned. Reassigned in the current tree. European lineage with 3 clear SNPs `28881GA`,`28882GA`,`28883GC`
......CTA........	17	0.03	B.1	United States of America 48.0%, United Kingdom 12.0%, Canada 4.0%, Spain 4.0%, France 3.0%	24.01.2020	19976	55060	A large European lineage that corresponds to the Italian outbreak.

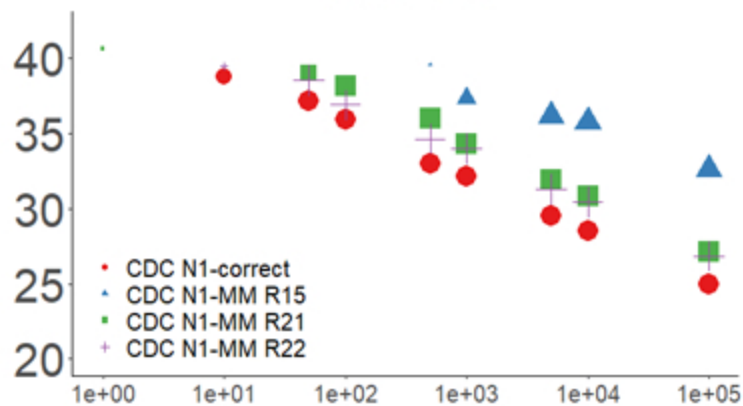Supplementary Table. 1. Primer and probe sequences used in this study.

Underlined = T7 promotor, italicized = nonsense sequence to extend the amplicon, bold and red = mismatch

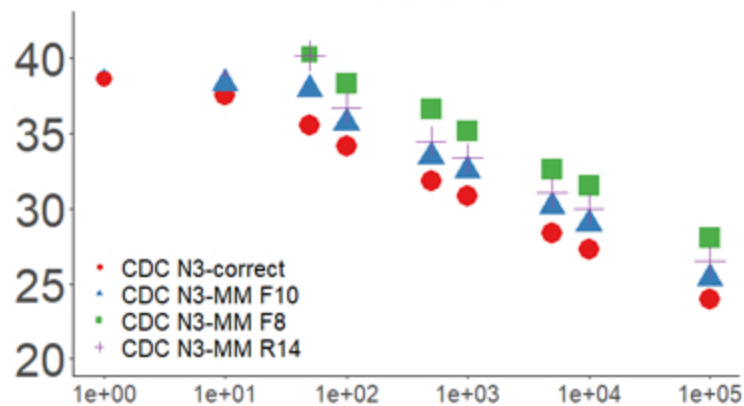| Assay	Primer name	Sequence 5' - 3'	Description and Reference
Charité-RdRP	RdRP_SARSr-F2	GTGARATGGTCATGTGTGGCGG	(Corman et al., 2020)
	RdRP_SARSr-R1	CARATGTTAAASACACTATTAGCATA	(Corman et al., 2020)
	RdRP_SARSr-P2	FAM-CAGGTGGAACCTCATCAGGAGATGC-BHQ1	(Corman et al., 2020)
	RdRP long F	<u>TAATACGACTCACTATAGGGAGA</u> *ACGTTGGCCA* GTGAAATGGTCATGTGTGGCGG	used to generate correct template
	RdRP long R	*ACGTGCATAATTGGCTACGTCGTGA* CAAATGTTAAAAACACTATTAGCATA	used to generate correct and mismatch template
	RdRP MM F8	<u>TAATACGACTCACTATAGGGAGA</u> *ACGTTGGCCA* GTGAAAT**T**GTCATGTGTGGCGG	used to generate mismatch template mismatch in position 8 from 5' end
	RdRP_SARSr-R1.corrECDC	CARATGTTAAAWACACTATTAGCATA	ECDC corrected version of reverse primer
	RdRP_SARSr-F2.new	GTGA**A**ATGGTCATGTGTGGCGG	Charité primer without ambiguities
	RdRP_SARSr-R1.new	CA**A**ATGTTAAA**A**ACACTATTAGCATA	Charité primer without ambiguities
Charité-N	N_Sarbeco_F1	CACATTGGCACCCGCAATC	(Corman et al., 2020)
	N_Sarbeco_R1	GAGGAACGAGAAGAGGCTTG	(Corman et al., 2020)
	N_Sarbeco_P1	FAM-ACTTCCTCAAGGAACAACATTGCCA-BHQ1	(Corman et al., 2020)
	N_Sarbeco long F	<u>TAATACGACTCACTATAGGGAGA</u> *ACGTTGGCCA* CACATTGGCACCCGCAATC	used to generate correct and mismatch template
	N_Sarbeco long R	*ACGTGCATAATTGGCTACGTCGTGA* GAGGAACGAGAAGAGGCTTG	used to generate correct template
	N_Sarbeco MM R8	*ACGTGCATAATTGGCTACGTCGTGA* GAGGAAC**A**AGAAGAGGCTTG	used to generate mismatch template, mismatch in position 8 from 5' end
	N_Sarbeco MM R13	*ACGTGCATAATTGGCTACGTCGTGA* GAGGAACGAGAA**T**AGGCTTG	used to generate mismatch template, mismatch in position 13 from 5' end

US-CDC N1	2019-nCoV_N1-F	GACCCCAAAATCAGCGAAAT	(US-CDC, 2020)
	2019-nCoV_N1-R	TCTGGTTACTGCCAGTTGAATCTG	(US-CDC, 2020)
	2019-nCoV_N1-P	Joe-ACCCCGCATTACGTTTGGTGGACC-BHQ1	(US-CDC, 2020)
	CDC N1 long F	*TAATACGACTCACTATAGGGAGA ACGTTGGCCA* GACCCCAAAATCAGCGAAAT	used to generate correct and mismatch template
	CDC N1 long R	*ACGTGCATAATTGGCTACGTCGTGA* TCTGGTTACTGCCAGTTGAATCTG	used to generate correct template
	CDC N1 MM R15	*ACGTGCATAATTGGCTACGTCGTGA* TCTGGTTACTGCCA**A**TTGAATCTG	used to generate mismatch template, mismatch in position 15 from 5' end
	CDC N1 MM R21	*ACGTGCATAATTGGCTACGTCGTGA* TCTGGTTACTGCCAGTTGAA**C**CTG	used to generate mismatch template, mismatch in position 21 from 5' end
	CDC N1 MM R22	*ACGTGCATAATTGGCTACGTCGTGA* TCTGGTTACTGCCAGTTGAAT**A**TG	used to generate mismatch template, mismatch in position 22 from 5' end
US-CDC N3	2019-nCoV_N3-F	GGGAGCCTTGAATACACCAAAA	(US-CDC, 2020)
	2019-nCoV_N3-R	TGTAGCACGATTGCAGCATTG	(US-CDC, 2020)
	2019-nCoV_N3-P	FAM-AYCACATTGGCACCCGCAATCCTG-BHQ1	(US-CDC, 2020)
	CDC N3 long F	*TAATACGACTCACTATAGGGAGA ACGTTGGCCA* GGGAGCCTTGAATACACCAAAA	used to generate correct and mismatch template
	CDC N3 long R	*ACGTGCATAATTGGCTACGTCGTGA* TGTAGCACGATTGCAGCATTG	used to generate correct and mismatch template
	CDC N3 MM F8	*TAATACGACTCACTATAGGGAGA ACGTTGGCCA* GGGAGCC**C**TGAATACACCAAAA	used to generate mismatch template, mismatch in position 8 from 5' end
	CDC N3 MM R14	*ACGTGCATAATTGGCTACGTCGTGA* TGTAGCACGATTG**A**AGCATTG	used to generate mismatch template, mismatch in position 14 from 5' end
	CDC N3 MM F10	*TAATACGACTCACTATAGGGAGA ACGTTGGCCA* GGGAGCCTT**T**AATACACCAAAA	used to generate mismatch template, mismatch in position 10 from 5' end
NIID Japan N	NIID_2019-nCOV_N_F2	AAATTTTGGGGACCAGGAAC	(Nao et al., 2020)
	NIID_2019-nCOV_N_R2.new	TGGCACCTGTGTAGGTCAAC	(Nao et al., 2020)
	NIID_2019-nCOV_N_P2	FAM-ATGTCGCGCATTGGCATGGA-BHQ1	(Nao et al., 2020)
	Japan N long F	*TAATACGACTCACTATAGGGAGA ACGTTGGCCA* AAATTTTGGGGACCAGGAAC	used to generate correct and mismatch template
	Japan N long R	*ACGTGCATAATTGGCTACGTCGTGA*	used to generate correct and mismatch

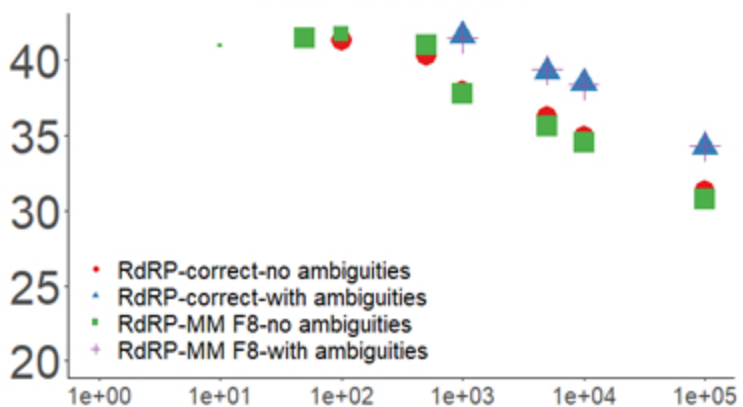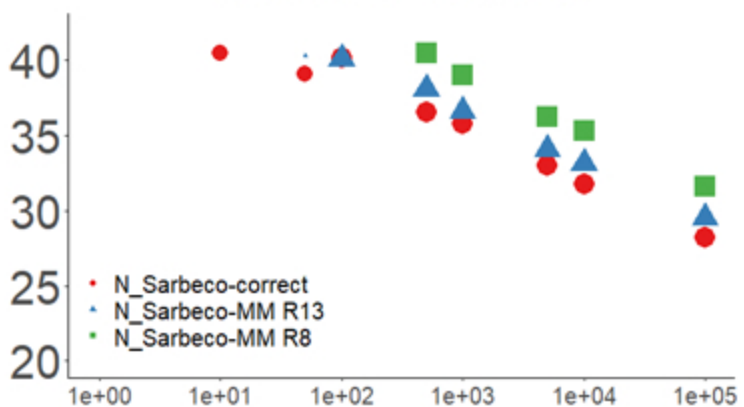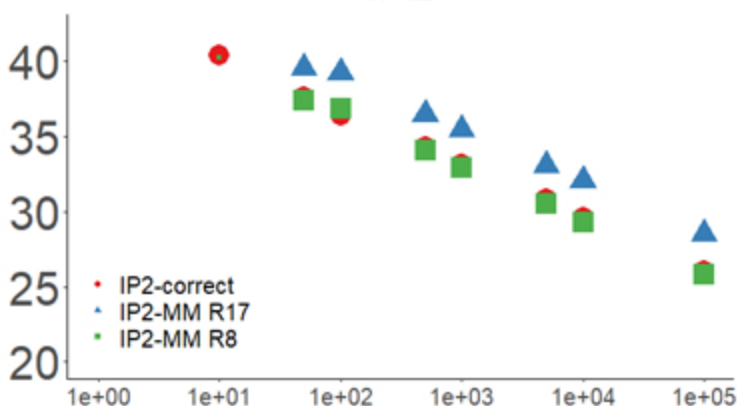|	neu	TGGCACCTGTGTAGGTCAAC	template
	Japan N MM F20	TAATACGACTCACTATAGGGAGA *ACGTTGGCCA* AAATTTTGGGGACCAGGAA**T**	used to generate mismatch template, mismatch in position 20 from 5' end
	Japan N-new MM R17	*ACGTGCATAATTGGCTACGTCGTGA* TGGCACCTGTGTAGGT**G**AAC	used to generate mismatch template, mismatch in position 17 from 5' end
	Japan N-new MM R18	*ACGTGCATAATTGGCTACGTCGTGA* TGGCACCTGTGTAGGTC**C**AC	used to generate mismatch template, mismatch in position 18 from 5' end
France RdRP IP2	nCoV_IP2-12669Fw	ATGAGCTTAGTCCTGTTG	(WHO, 2020)
	n CoV_IP2-12759Rv	CTCCCTTTGTTGTGTTGT	(WHO, 2020)
	nCoV_IP2-12696bProbe(+)	AGATGTCTTGTGCTGCCGGTA	(WHO, 2020)
	IP2 long F	TAATACGACTCACTATAGGGAGA *ACGTTGGCCA* ATGAGCTTAGTCCTGTTG	used to generate correct and mismatch template
	IP2 long R	*ACGTGCATAATTGGCTACGTCGTGA* CTCCCTTTGTTGTGTTGT	used to generate correct and mismatch template
	IP2 MM R8	*ACGTGCATAATTGGCTACGTCGTGA* CTCCCTT**C**GTTGTGTTGT	used to generate mismatch template, mismatch in position 8 from 5' end
	IP2 MM R17	*ACGTGCATAATTGGCTACGTCGTGA* CTCCCTTTGTTGTGTT**A**T	used to generate mismatch template, mismatch in position 17 from 5' end

CDC N1

- CDC N1-correct
- CDC N1-MM R15
- CDC N1-MM R21
- CDC N1-MM R22

CDC N3

- CDC N3-correct
- CDC N3-MM F10
- CDC N3-MM F8
- CDC N3-MM R14

Charité RdRP

- RdRP-correct-no ambiguities
- RdRP-correct-with ambiguities
- RdRP-MM F8-no ambiguities
- RdRP-MM F8-with ambiguities

Charité N Sarbeco

- N_Sarbeco-correct
- N_Sarbeco-MM R13
- N_Sarbeco-MM R8

IP2

- IP2-correct
- IP2-MM R17
- IP2-MM R8

NIID N

- NIID_N-correct
- NIID_N-MM F20
- NIID_N-MM R17
- NIID_N-MM R18

Seegene N Assay

Legend:
- B1.1.7 (red circle)
- WT (blue triangle)
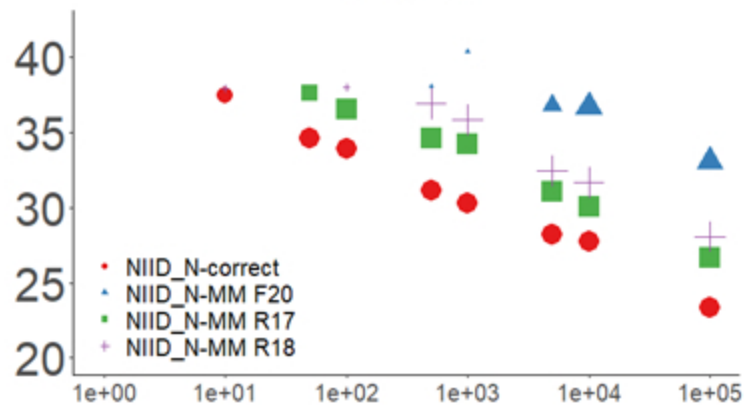
```
                            28270     28280     28290     28300     28310     28320     28330
                            .|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....
    Seegene N amplicon      ------TGTCTGATAATGGACCCCAAAATCAGCGAAATGCACCCCGCATTACGTTTGGTGGACCCTCAGAT
    MN908947.3(WuhanHu1)    CTAAAATGTCTGATAATGGACCCCAAAATCAGCGAAATGCACCCCGCATTACGTTTGGTGGACCCTCAGAT
    EPI_ISL_755639(B1.1.7)  CTAA-ATGTCTCTAAATGGACCCCAAAATCAGCGAAATGCACCCCGCATTACGTTTGGTGGACCCTCAGAT
```