Design and *in silico* validation of polymerase chain reaction primers to detect severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)

**Daniel C. F. Lanza** (danielclanza@gmail.com)
Applied Molecular Biology Lab - LAPLIC, Department of Biochemistry, Biosciences Center, Federal University of Rio Grande do Norte, Natal, RN, Brazil

**João P. M. S. Lima**
Laboratory of Metabolic Systems and Bioinformatics - LASIS, Department of Biochemistry, Biosciences Center, Federal University of Rio Grande do Norte, Natal, RN, Brazil

**Selma M. B. Jerônimo**
Institute of Tropical Medicine of Rio Grande do Norte (IMT), Federal University of Rio Grande do Norte, Natal, RN, Brazil

**Research Article**

**Keywords:** COVID-19, universal PCR primers, coronavirus, blood, real-time PCR, qPCR, environmental samples

**Posted Date:** April 30th, 2020

**DOI:** https://doi.org/10.21203/rs.3.rs-26306/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

**Version of Record:** A version of this preprint was published at Scientific Reports on June 15th, 2021. See the published version at https://doi.org/10.1038/s41598-021-91817-9.
Abstract

Accurate designing of polymerase chain reaction (PCR) primers targeting conserved segments in viral genomes is desirable for preventing false negative results and decreasing the need for standardization across different PCR protocols. In this work, we designed and described a set of primers and probes targeting conserved regions identified from a multiple sequence alignment of 2341 Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) genomes from the Global Initiative on Sharing All Inuenza Data (GISAID). Those primers and probes were subsequently validated in 3067 SARS-CoV-2 whole-genome sequences. From these analyses, we obtained nine systems (forward primer + reverse primer + probe) that potentially anneal to highly conserved regions of the virus genome. In silico predictions also demonstrated that those primers do not bind to nonspecific targets for human, bacterial, fungal, or apicomplexan sequences. The availability of these primer and probe sequences will make it possible to accelerate the beginning of in vitro tests in order to validate more efficient protocols for the identication of SARS-CoV-2.

Introduction

The identication of viral genetic material using the PCR technique is considered the gold standard for determining SARS-CoV-2 in nasal swab samples from symptomatic patients. Since the outbreak started, WHO released some SARS-CoV-2 polymerase chain reaction (PCR) protocol assays produced by different reference institutions in the world [1]. In addition to these initial protocols, there is an increasing number of works and commercial kits suggesting new alternatives to the identication of SARS-CoV-2 by molecular or immunological approaches [2-4].

Concomitant with those advances, many polymorphic regions in SARS-CoV-2 have been identied as a result the rapid increase in the number of available SARS- CoV-2 genome sequences from different localities. It is expected that some of the proposed primers may have targeted these polymorphic regions, which can compromise accurate identication of some viral variants and increase the number of false negatives or inconclusive results.

In this context, the design of primers that target conserved regions in the genome in order to detect a large number of viral variants possible is imperative. In this work, we identied 26 conserved segments (CS) in the SARS-CoV-2 genome based on an alignment of 2341 full genome sequences and used these regions as a target for the design of universal primers and probes.

Results

At the end of the analysis, we elected nine candidate systems (forward primer + reverse primer + probe) that met all requirements (Table 1). In general, in silico analyzes revealed that the primers pairs proposed in this study (UFRN_primers) (Table 1) are more compatible with each other (evaluated by lower differences between forward and reverse primers' Tm), have lower self-complementarity (both overall and 3'), and higher specicity than the previously described primers (PD_primers) (Table 2). Regarding the proposed probes, only the probes UFRN_3_P and UFRN_4_P did not reach a Tm higher than that of their respective primer pairs.

By comparing the number of SARS-CoV-2 sequences that anneal without mismatches ("No mis" in Tables 1 and 2), using in silico PCR methodologies, it is safe to assume that the set of UFRN_primers targets less polymorphic sites in the viral genome than the PD_primers set. Among the 3067 SARS-CoV-2 genomic sequences used as targets, the primers from UFRN_primers set anneal with 100% identity with at least 3060. Also, the newly-designed primers are still capable of annealing with at least 3065 sequences from the tested database (Table 1) when performing the same analysis allowing a maximum of 10% of mismatches.

Concerning the specicity, both primers set performed well. Tests allowing 20% mismatch against Apicomplexa targets revealed that the 2019-nCoV_N2-F / 2019- nCoV_N2-R and UFRN_8_F / UFRN_8_R primer pairs could generate 746 bp and 755 bp amplicons with Toxoplasma gondii sequences from accession codes XTG08368.2 and XM_002364674.2, respectively. The other pairs of primers did not present non- specific amplicons allowing values between 0 and 20% of mismatches.

Discussion

Early detection of pathogens is a crucial step to disease prevention [5] and containment, especially during epidemic outbreaks [6]. PCR is one of the reliable and relatively accessible molecular methods that directly recognize pathogen-derived
material from patients samples [7]. However, the optimization of PCR protocols is strongly dependent on primers specificity and efficiency [8]. This reason, combined with the increasing number of SARS-CoV-2 sequences available and its crescent polymorphism, led us to design a set of new primers that can address very conserved regions of the virus genomes.

Therefore, in order to aid PCR optimization, the UFRN_primers were designed to present Tm values that were as close as possible. It is expected that these configurations will allow the use of at least two systems using the same thermal cycling parameters. In this way, it would be possible to perform the PCR test identifying different regions of the viral genome at the same time, according to the protocols already described for the PD_primers. In this context, possibly the systems UFRN_3 and UFRN_4, will have different thermal cycling parameters comparing to the other systems since, in this case, the probe Tm is similar to the primers (Table 1). Probably these systems will depend on more annealing time in order to ensure that the probe has interacted in DNA template before the amplification starts.

The higher specificity of UFRN_primers confirmed by in silico analysis is mainly due to the availability of 2,341 genome sequences, which made it possible to identify the conserved regions with greater accuracy from the alignment. Quite possibly, at the time of publication of this work, a considerably larger number of additional sequences will be available, which may reveal new polymorphic sites in the target regions of UFRN_primers and PD_primers. In this way, our research group will continue this bioinformatics work, and whenever relevant, new updates on the primer sequences or new primers set will be reported.

Another critical point is that primers presented here were tested against the updated databases of RNA sequences from bacteria, fungi, and protozoa, and did not generate non-specific amplicons in any case. Although executed through in silico analyses, this negative prediction increases the potential for applying these primers to different samples such as blood, feces, or even in environmental samples. Currently, the most suitable sample for detecting SARS-CoV-2 is the human nasal swab; however, there are already studies that have shown digestive symptoms (e.g. diarrhea and vomiting) [9,10] and other less frequent symptoms (e.g. conjunctivitis) in patients who tested positive for SARS-CoV-2 [11-13]. This diversity of symptoms makes clinical diagnosis difficult, and testing new types of samples may be needed in a short time. The application of UFRN_primers to detect SARS-CoV-2 in blood or fecal samples is likely efficient since these primers should not interact non-specifically with RNAs of the main protozoa and bacteria that cause health problems in humans.

The use of universal primers makes it possible to identify several variants of the virus using the same PCR protocol. UFRN_primers are strong candidates to help simplify the procedures and supply chain for detecting SARS-CoV-2, allowing, for example, the mass production of primers and kits that could be applied in different parts of the world with equivalent efficiency. However, the primers presented here still depend on in vitro validation. The availability of these sequences at this time will be crucial so that these new protocols can be validated promptly to assist in the control of the SARS-CoV-2 pandemic.

Methods

Whole-genome sequences of SARS-CoV-2 from human isolates were retrieved from the Global Initiative on Sharing All Influenza Data (GISAIID - gisaid.org) [14] and Virus Variation from the National Center for Biotechnology Information (NCBI - https://www.ncbi.nlm.nih.gov/genome/viruses/variation/) [15] databases, every three days, between Mar 30 and Apr 13, 2020. To minimize sequencing errors and artifacts, we activated the filters “complete (>29,000 bp),” “high coverage only” and “low coverage excl” at sequence retrieval in GISAID database and the filter “Complete” under the option “Nucleotide completeness” from the Virus Variation database. The full list of authors and laboratories of GISAID submissions and the Virus Variation sequences accessions are available in Supplementary Material 1.

Complete fasta sequences were then aligned using Clustal-Omega, version 1.2.4 [16], with standard parameters, using a supercomputer. To avoid excessive misaligned gaps and to better identify conserved polymorphic sites, we trimmed the multiple sequence alignments (MSAs) using the trimAltool, version 1.2 [17], with the “-automated” option. We used the sequence from a Wuhan seafood market pneumonia virus (GenBank Accession code MN908947) [18] as a reference for all alignments to identify site and region positions.

The CSs were submitted to online Primer-BLAST [19] to design primer pairs adopting the following criteria: PCR product size = 90 – 150 nt; primer melting temperatures (ºC) minimum = 55, optimum = 58, maximum = 63 and maximum melting temperature (Tm) difference = 2 ºC. The specificity check was performed using the complete Refseq RNA databases for Homo sapiens (taxid: 9606), Bactéria (taxid: 2), Fungi (taxid:4751), Apicomplexa (taxid:5794). The primer specificity stringency was set so that the primer must have at least 3 total mismatches to unintended targets, including at least 2 mismatches within the last 5 bps at the 3’ ignoring targets that have 5 or more
mismatches to the primer. The other Primer-BLAST parameters have been kept in the default configuration to confirm the newly-designed primers pairs features.

From all the primers generated by the Primer-BLAST, we selected 124 primers pairs that presented low self-complementarity for total annealing (max 5 nt) and also for annealing in the 3' region (max 3 nt). After individual evaluation using the Geneious suite (version 9.1.8, 2017), we elected 9 primer pairs that target regions with 100% identity among all of the 2143 genomes. These primers comprise ORF1a, ORF1b, and S regions of the SARS-CoV-2 genome. The TaqMan® probes for each primer pair were also designed considering the same alignment and prioritizing conserved regions inside each of the predicted amplicons.

To compare and assess the annealing specificity of the already used and newly-designed primers and probes, we used three different tools: PrimerSearch from the Emboss package [20], the stand-alone BLAST+ [21] and the on-line Primer-BLAST. For the first two tools, we used four different custom databases: (1) SARS-CoV-2 sequences from GISAID (3067 genome sequences retrieved on April 6, 2020 at 9:00 am); (2) SARS-CoV-2 sequences from Virus Variation; (3) RefSeq RNAs from Apicomplexa taxon, retrieved from GenBank on Mar 30, 2020; and (4) RefSeq RNAs from Toxoplasma taxon, also from GenBank (Mar 30, 2020). The first step of the specificity test was to search all 5' and 3' primers pairs sequences to verify amplicon possibilities using PrimerSearch, against each of the databases mentioned above. At this screening, we used three different mismatch allowance percentages (0, 10 and 20%). For probes similarity searches, we evaluated the number of hits subject sequences from stand-alone BLAST+, the aligned start and end regions, and the number of mismatches for each alignment.

Declarations

Acknowledgments

We are grateful for the authors and laboratories that contributed the SARS-CoV-2 sequences to GISAID. The authors would also like to thank the "Núcleo de Processamento de Alto Desempenho - NPAD" from Digital Metropolis Institute (IMD) at the Federal University of Rio Grande do Norte (UFRN), for the use of the supercomputer for multiple sequence alignments.

Author Contributions

D.C.F.L, Conceived the experiments, identified the conserved regions, established the parameters for primer design, designed the primers, analysed the results and wrote the manuscript. J.P.M.S.L Collected and selected the sequence data, performed the alignments, performed the specificity tests, analysed the results and wrote the manuscript. S.M.B.J Conceived the experiments and assisted in the writing of the manuscript.

Competing interests

The author(s) declare no competing interests.

Funding

This research received financial support from the Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq, and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)

References

1. WHO. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/laboratory-guidance (2020).
2. Chan, JF et al. Improved molecular diagnosis of COVID-19 by nthe novel, highly sensitive and specific COVID-19-RdRp/Hel real-time reverse transcription- polymerase chain reaction assay validated in vitro and with clinical specimens. J Clin Microbiol. 2020 Mar pii: JCM.00310-20. doi: 10.1128/JCM.00310-20.
3. Wang Y, Kang H, Liu X, Tong Z.Y. Combination of RT qPCR Testing and clinical features for diagnosis of COVID-19 facilitates management of SARS-CoV-2 J Med Virol. 2020 Feb 25. doi: 10.1002/jmv.25721
4. Udugama B, Kadhiresan P, Kozlowski HN, Malekjahani A, Osborne M, Li VYC Chen H, Mubareka S, Gubbay JB, Chan WCW Diagnosing COVID-19: The Disease and Tools for ACS Nano. 2020 Apr 9. doi: 10.1021/acsnano.
Table 1 – Primers designed in this study.
| Primer name | Sequence 5'-> 3' | Length | Tm  | GC(%) | SC | SC 3' | Target | Size | No mis | 10% mis |
|-------------|------------------|--------|-----|-------|----|-------|--------|------|--------|---------|
| UFRN_1_F    | GGGCATAACACTGCCTATGTC | 20     | 58,22 | 55     | 4  | 3     | ORF1a  | 101  | 3064   | 3067    |
| UFRN_1_R    | GCATGAAGCTTACCAGCAC | 20     | 57,73 | 50     | 6  | 0     | ORF1a  | 137  | 3061   | 3065    |
| UFRN_1_P    | TCTGTGGCCCTGATGGCTACCCT | 23     | 67,22 | 60,87  | 7  | 2     | ORF1a  | 101  | 3064   | 3067    |
| UFRN_2_F    | GCCTACTAAACTGCACTGC | 20     | 57,22 | 50     | 5  | 2     | ORF1a  | 101  | 3064   | 3067    |
| UFRN_2_R    | TAACATTGGGCCAGCAAGCA | 20     | 58,02 | 45     | 4  | 1     | ORF1a  | 101  | 3064   | 3067    |
| UFRN_2_P    | GGGTGGTATAGTGGTTTAAAGCGG | 24     | 62,33 | 50     | 4  | 1     | ORF1a  | 101  | 3064   | 3067    |
| UFRN_3_F    | TTCACTGTTGGCTCGCCCAAT | 20     | 58,37 | 45     | 2  | 1     | ORF1a  | 101  | 3064   | 3067    |
| UFRN_3_R    | TGGTGAATTAGTAACTTCTGT | 20     | 57,1  | 50     | 4  | 2     | ORF1a  | 101  | 3064   | 3067    |
| UFRN_3_P    | GTGCAGCTAATGTCATGGCACTGTTTATA | 28     | 58,01 | 32,14  | 8  | 4     | ORF1a  | 101  | 3064   | 3067    |
| UFRN_4_F    | AGGCACACTGACAACAGAA | 20     | 58,27 | 50     | 4  | 0     | ORF1b  | 105  | 3063   | 3066    |
| UFRN_4_R    | CAATTCAGCAGGACAGAAGCG | 20     | 58,31 | 50     | 4  | 2     | ORF1b  | 105  | 3063   | 3066    |
| UFRN_4_P    | GTGCCAGCATGTTCCTCCGGA | 24     | 64,18 | 54,17  | 8  | 6     | ORF1b  | 105  | 3063   | 3066    |
| UFRN_5_F    | TCTTCACGACATTGGTAACCC | 21     | 57,95 | 47,62  | 5  | 3     | ORF1b  | 105  | 3063   | 3066    |
| UFRN_5_R    | GGTACCTGAGGCTGTGACATC | 21     | 57,9  | 50     | 4  | 2     | ORF1b  | 105  | 3063   | 3066    |
| UFRN_5_P    | TACCTCAAGCTGGAATGGGAAG | 26     | 60,41 | 42,31  | 8  | 0     | ORF1b  | 105  | 3063   | 3066    |
| UFRN_6_F    | CTTGTCAGATTGGTAACCT | 21     | 57,95 | 47,62  | 5  | 3     | ORF1b  | 105  | 3063   | 3066    |
| UFRN_6_R    | CTTCACGACATTGGTAACCT | 21     | 57,95 | 47,62  | 5  | 3     | ORF1b  | 105  | 3063   | 3066    |
| UFRN_7_F    | GGTACCTGAGGCTGTGACATC | 21     | 57,9  | 50     | 4  | 2     | ORF1b  | 105  | 3063   | 3066    |
| UFRN_7_R    | GTGTACCTCAAGCAGAATGG | 26     | 61,4  | 46,15  | 8  | 0     | ORF1b  | 105  | 3063   | 3066    |
| UFRN_7_P    | GGTACCTCAAGCAGAATGG | 26     | 61,4  | 46,15  | 8  | 0     | ORF1b  | 105  | 3063   | 3066    |
| UFRN_8_F    | GGCACAGGTGGTTCTTTACTGA | 20     | 57,46 | 50     | 4  | 1     | S      | 107  | 3062   | 3066    |
| UFRN_8_R    | TCAAGTGCTGATGTAATGGGAAG | 20     | 57,56 | 50     | 4  | 2     | S      | 107  | 3062   | 3066    |
| UFRN_8_P    | CCAACAATTGGCCAGACATTGC | 24     | 61,62 | 45,83  | 5  | 3     | S      | 107  | 3062   | 3066    |
| UFRN_9_F    | AGGCACAGGTGTTCTTTACTG | 20     | 57,46 | 50     | 4  | 1     | S      | 107  | 3062   | 3066    |
| UFRN_9_R    | TCACGGACAGCATGATG | 20     | 58,45 | 50     | 3  | 2     | S      | 93   | 3063   | 3066    |
| UFRN_9_P    | TCAACAATTTGGCCAGACATTGC | 25     | 62,75 | 44     | 5  | 3     | S      | 93   | 3063   | 3066    |

F = forward primer; R= reverse primer; P = probe; Tm = melting temperature; GC% = G + C percentage; SC = self complementarity; SC 3’ = self 3’-complementarity; No mis = number of sequences that anneal to the primer without mismatches; 10% mis = number of sequences that anneal to the primer allowing 10% mismatches.

Table 2 – Primers released by WHO to detect SARS-CoV-2 using polymerase chain reaction.
| Location/Primer or probe name | Sequence (5’>3’) | Length | Tm  | GC(%) | SC | SC’ | Target | Size | No mis | 10% mis |
|-------------------------------|------------------|--------|-----|-------|----|-----|--------|------|--------|---------|
| **Germany (17 January 2020)** |                  |        |     |       |    |     |        |      |        |         |
| RdRP_SARSr-F2                | GTGARATGGTACATGTTGGCGG | 22     | 63,25 | 57,14 | 5,5 | 1   | RdRp   |      |        |          |
| RdRP_SARSr-R1                | CARATGTTAASACACATTAGCATA | 26     | 54,25 | 25    | 5   | 4   | RdRp   | 100  | 1      | 3067     |
| RdRP_SARSr-P2                | FAM-CAGTGGACCTCATCAGGAGATGC-BBQ | 25     | 64,89 | 56    | 6   | 5   | RdRp   |      |        |          |
| E_Sarbeco_F1                 | ACAGGTACGGTTAATAGTTAATAGGCT | 26     | 58,29 | 34,62 | 8   | 8   | E      |      |        |          |
| E_Sarbeco_R2                 | ATATTGCAGACAGTCACACA | 22     | 60,39 | 45,45 | 7   | 1   | E      | 113  | 3063   | 3064     |
| E_Sarbeco_P1                 | FAM-ACACTAGCCATCTTTACTGCGCTTGC-BBQ | 26     | 66,78 | 53,85 | 4   | 2   | E      |      |        |          |
| N_Sarbeco_F1                 | CACATTGGCAACCGCAATC | 19     | 60,15 | 57,89 | 4   | 0   | N      |      |        |          |
| N_Sarbeco_R1                 | GAGGAAGGAAAGGCTTTC | 20     | 58    | 55    | 3   | 1   | N      | 128  | 3048   | 3063     |
| N_Sarbeco_P1                 | AM-ACTTCCTCAAGGAACACATGGCCA-BBQ | 25     | 63,15 | 44    | 8   | 3   | N      |      |        |          |
| **Hong Kong (23 January 2020)** |                  |        |     |       |    |     |        |      |        |         |
| HKU-ORF1b-nsp14F             | TGGGYYTTACRGGTAACCT | 20     | 47,07 | 50    | 7,5 | 4,5 | ORF1b  |      |        |          |
| HKU-ORF1b-nsp14R             | AACRCGCTTTAACAAGCACTC | 21     | 53,44 | 45    | 4   | 0   | ORF1b  | 132  | 3060   | 3067     |
| HKU-ORF1b-nsp141P            | FAM-TAGTTGTGATGWCATGACTAG-TAMRA | 24     | 54,86 | 39,13 | 10,5 | 6,5 | ORF1b  |      |        |          |
| HKU-NF                       | TAATCAGACAAAGGAACGTATTA | 22     | 52,27 | 31,82 | 7   | 7   | N      |      |        |          |
| HKU-NR                       | CGAAGGTGTGACTTCCATG | 19     | 55,95 | 52,63 | 4   | 4   | N      | 110  | 3049   | 3067     |
| HKU-NP                       | FAM-GCAAAATTGTGCATTTGGCGG-TAMRA | 20     | 58,05 | 45    | 14  | 6   | N      |      |        |          |
|                     |                  |      |      |      |      |      |      |          |          |          |
|---------------------|------------------|------|------|------|------|------|------|----------|----------|----------|
| **ORF1ab_F**        | CCCTGTGGGTTTTTACCTTAA | 21   | 55,7 | 42,86 | 4   | 4   | ORF1ab |
| **ORF1ab_R**        | ACGATTGTGATCATGCTGTA | 19   | 57,46| 47,37 | 8   | 8   | ORF1ab | 119      | 3062     | 3064     |
| **ORF1ab_P**        | FAM-CCGCTCTCGGTATGAGGTTATGG-BHQ1 | 28   | 67,24| 53,57 | 3   | 0   | ORF1ab |
| **N_F**             | GGGGAACCTTCTCTGCTGATTAA | 22   | 59,23| 50    | 7   | 2   | N      |
| **N_R**             | CAGACATTTTTGCTCTCAAGCTG | 22   | 58,18| 45,45 | 4   | 2   | N      | 99       | 2580     | 2954     |
| **N_P**             | FAM-TTGCTGCTGCTTTGACAGATT-TAMRA | 20   | 58,39| 45    | 4   | 1   | N      |

**Japan (24 January 2020)**

|                     |                  |      |      |      |      |      |      |          |          |          |
|---------------------|------------------|------|------|------|------|------|------|----------|----------|----------|
| **NIID_WH-1_F501-F** | TTCGATCTCAGAAGGCAACC | 21   | 63,27| 57,14 | 4   | 0   | ORF1a  | 413      | 3016     | 3059     |
| **NIID_WH-1_R913-R** | CTTTACGACGTGCTGAGG | 23   | 61,47| 52,17 | 10  | 10  | ORF1a  |
| **NIID_WH-1_F509-F** | CTGAAGCTGAACCTGCTGGA | 19   | 58,24| 57,89 | 4   | 2   | ORF1a  | 346      | 3045     | 3059     |
| **NIID_WH-1_R854-R** | CAGAAGTTTATAAGGCTGACATAGC | 22   | 55,05| 40,91 | 4   | 3   | ORF1a  |
| **NIID_WH-1_Seq_F519** | ACCTCATGTCAGTGTATGGA | 20   | 54,79| 45    | 6   | 1   | ORF1a  | 322      | 3047     | 3057     |
| **NIID_WH-1_Seq_R840** | GACATAGCGAGTGTAGGCA | 19   | 55,61| 52,63 | 4   | 3   | ORF1a  |
| **WuhanCoV-spk1-f** | TTGCAAAATTCAAGACTCACTT | 24   | 58,02| 33,33 | 5   | 3   | S      | 547      | 3054     | 3067     |
| **WuhanCoV-spk2-r** | TGTGGTTCTAAAATTTCTTTTGTG | 25   | 56,98| 32    | 4   | 3   | S      |
| **NIID_WH-1_F24381** | TCAAGACTCACTTCTCTCAC | 21   | 55,48| 42,86 | 4   | 0   | S      | 493      | 3044     | 3066     |
| **NIID_WH-1_R24873** | ATTTGAAACAAAGACACCTTAC | 23   | 56,13| 34,78 | 5   | 0   | S      |
| Sequence ID         | Primer Sequence                     | Length | Tm  | GC% | Amplicon Size | Sense/Reverse | V3/V4 | RNAse P |
|---------------------|--------------------------------------|--------|-----|-----|---------------|---------------|-------|---------|
| NIID_WH-1_Seq_F24383 | AAGACTCACCTTCTCACACAG               | 21     | 55.47 | 42.86 | 4 1 S         | 483           | 3045  | 3065    |
| NIID_WH-1_Seq_R24865 | CAAAGACACCTTCAGG                   | 19     | 55.88 | 52.63 | 3 2 S         |               |       |         |
| NIID_2019-nCOV_N_F2  | AAATTTGGGGACAGGAC                 | 20     | 56.09 | 45      | 6 1 N         |               |       |         |
| NIID_2019-nCOV_N_R2  | TGGCAGCTGGTAGGTCACAC               | 20     | 60.25 | 55      | 6 2          | 108 0        | 3065  |
| NIID_2019-nCOV_N_P2  | FAM-ATGTCGCGCATGGCATGGA-BHQ       | 20     | 63.5  | 55      | 6 0          |               |       |         |
| **Thailand (23 January 2020)** |                                     |        |       |       |               |               |       |         |
| WH-NIC N-F           | CGTTTGTGTTGGACCTCAGAT              | 20     | 59.68 | 55      | 4 2 N         |               |       |         |
| WH-NIC N-R           | CCCACTGCGTTCTCATT                  | 19     | 60    | 57.89 | 3 1 N         | 57 3048       | 3067  |
| WH-NIC N-P           | FAM-CAACTGGCAGTAACCA-BQH1          | 16     | 50.27 | 50      | 7 1 N         |               |       |         |
| **USA (24 January 2020)** |                                     |        |       |       |               |               |       |         |
| 2019-nCoV_N1-F       | GACCCCCAAATCAGCGAAAT              | 20     | 56.67 | 45      | 2 2 N         |               |       |         |
| 2019-nCoV_N1-R       | TCTGTATTACTGCCCAGTTTAGCTG         | 24     | 60.8  | 45.83  | 7 5 N         | 72 3036       | 3067  |
| 2019-nCoV_N1-P       | FAM-ACCCCCGATTACGGTTGGGACCG-BHQ1  | 24     | 67.48 | 58.33  | 4 4 N         |               |       |         |
| 2019-nCoV_N2-F       | TTACAAACATTGGGCGCAAA              | 20     | 57.11 | 40      | 5 5 N         |               |       |         |
| 2019-nCoV_N2-R       | GCGCGACATTCCGAAAGAA              | 18     | 58.53 | 55.56  | 5 2 N         | 67 3058       | 3067  |
| 2019-nCoV_N2-P       | FAM-ACAATTGGCCCCAGGGCTCGAC-BHQ1   | 23     | 66.45 | 56.52  | 6 2 N         |               |       |         |
| 2019-nCoV_N3-F       | GGGAGCCTTGAAATACACAAA             | 22     | 58.84 | 45.45  | 4 0 N         |               |       |         |
| 2019-nCoV_N3-R       | TGTAGCAGATTGCAAGCTTGG            | 21     | 59.87 | 47.62  | 5 3 N         | 72 2985       | 3067  |
| 2019-nCoV_N3-P       | FAM-AYCACATTGGCACCAGGACATCCTG-BHQ1| 24     | 65.21 | 56.52  | 4 1 N         |               |       |         |
| RP-F                 | AGATTGCGACCTGCGACG                | 19     | 60.45 | 57.89  | 3 2 RNAse P   |               |       |         |
| F = forward primer; R = reverse primer; P = probe; Tm = melting temperature; GC% = G + C percentage; SC = self complementarity; SC 3’ = self 3’- complementarity; No mis = number of sequences that anneal to the primer without mismatches; 10% mis = number of sequences that anneal to the primer allowing 10% mismatches. |

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

- SupplementaryMaterialbbb5ffcab9e45ef84950ac4ebf7584f.xlsx