# *In silico* assessment of the impact of 2019 novel coronavirus genomic variation on the efficiency of published real-time quantitative polymerase chain reaction detection assays

**Hang Fan, Xiang-Li-Lan Zhang, Ya-Wei Zhang, Yong Huang, Yue Teng, Yan Guo, Zhi-Qiang Mi, Rui-Fu Yang, Ya-Jun Song, Yu-Jun Cui**

State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, Beijing 100071, China.

*To the Editor:* In December 2019, coronavirus disease 2019 outbreak caused by the 2019 novel coronavirus (2019-nCoV) happened in Wuhan, China. Now, it has posed a worldwide public health threat. Real-time quantitative polymerase chain reaction (RT-qPCR) was recommended as an effective pathogen detection method and has played an important role in prevention and control of the current outbreak. Many research institutions have released their primer sets for RT-qPCR. If the variant sites were located in the primer regions, the efficiency of RT-qPCR would be reduced, thus possibly causing false negative results, and leading to unpredictable impact on the diagnosis of patients and the control of this outbreak. Therefore, a comprehensive investigation on 2019-nCoV genome variation is necessary to evaluate the effectiveness of current released RT-qPCR methods.

Here, we analyzed 77 public full-length genome sequences of 2019-nCoV from the GISAID website [Supplementary Table 1, http://links.lww.com/CM9/A222]. All the sequences were aligned by using MAFFT v7.450.[1] A total of 85 variant sites were found, all of which are single nucleotide variants. Among the 85 variant sites, seven were shared by two 2019-nCoV sequences and nine were found in three or more sequences [Supplementary Table 2, http://links.lww.com/CM9/A223].

We investigated the published 2019-nCoV RT-qPCR detection assays and found a total number of 13 RT-qPCR primer sets designed by eight institutions [Table 1, Supplementary Table 3, http://links.lww.com/CM9/A224]. These primers were designed to amplify genes of ORF1ab, Spike (S), Envelope (E), and Nucleocapsid (N). The reverse primers of primer sets 3 and 6 had one mismatch against all of the released 2019-nCoV sequences. Among all the observed variation sites, three (positions on the reference genome IVDC-HB-01: 28291, 28688, and 29200) are separately located on the region of forward primers of the primer sets 7 and 9, and the probe of the primer set 8. It is noted that the above three variations was found in the sequences BetaCoV/Shenzhen/SZTH-003/2020, BetaCoV/Shandong/IVDC-SD-001/2020|EPI_ISL_408482, and BetaCoV/Chongqing/YC01/2020|EPI_ISL_408478, respectively. These variants may affect the RT-qPCR detection efficiency. In particular, variations on the probe region of the primer set 8 may have largely negative effects on detection efficiency according to the previous research.[2]

In conclusion, using any of the five RT-qPCR primer sets mentioned above to detect 2019-nCoV may potentially cause false-negative results. Among the five, two have mismatches and three contain some 2019-nCoV genome variants which occurred during the outbreak. It is worth noting that the three primer sets containing variants are all located on the N gene. Therefore, it is suggested that conservative regions, such as nsp12 (*RdRp*) gene, would be preferable primer targets. Although the multiple-targets designation of RT-qPCR protocol would reduce the false-negative results caused by genome variation, more careful performance evaluation of the currently used primers is needed. Moreover, it is necessary to keep continuous surveillance on the genome variants and their effects on the RT-qPCR assays during the whole outbreak.

## Acknowledgements

We gratefully acknowledge the Authors, the Originating and Submitting Laboratories for their sequences and metadata shared through GISAID, on which this research

Access this article online

Quick Response Code:

Website: www.cmj.org

DOI: 10.1097/CM9.0000000000000817

**Correspondence to:** Dr. Yu-Jun Cui, State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, Beijing 100071, China
E-Mail: cuiyujun.new@gmail.com



Chinese Medical Journal 2020;133(13)

**Received:** 26-02-2020 **Edited by:** Pei-Fang Wei

**Table 1: The detailed information of publicly released 2019-nCoV RT-qPCR primers.**

| Primer No. | Target gene | Primer type | Primer name | Primer sequence (5′-3′) | Pos. on Ref. | Variation site |
|---|---|---|---|---|---|---|
| 1 | ORF1ab | Forward | HKU-ORF1b-nsp14F | TGGGGYTTTACRGGTAACCT | 18778..18797 | |
| | | Reverse | HKU- ORF1b-nsp14R | AACRCGCTTAACAAAGCACTC | 18889..18909 | |
| | | Probe | HKU-ORF1b-nsp141P | FAM-TAGTTGTGATGCWATCATGACTAG-TAMRA | 18849..18872 | |
| 2 | N | Forward | HKU-NF | TAATCAGACAAGGAACTGATTA | 29145..29166 | |
| | | Reverse | HKU-NR | CGAAGGTGTGACTTCCATG | 29236..29254 | |
| | | Probe | HKU-NP | FAM-GCAAATTGTGCAATTTGCGG-TAMRA | 29177..29196 | |
| 3 | ORF1ab | Forward | RdRP_SARSr-F2 | GTGARATGGTCATGTGTGGCGG | 15431..15452 | |
| | | Reverse | RdRP_SARSr-R1 | CARATGTTAAASACACTATTAGCATA | 15505..15530 | 15519 |
| | | Probe | RdRP_SARSr-P2 | FAM-CAGGTGGAACCTCATCAGGAGATGC-BBQ | 15470..15494 | |
| 4 | E | Forward | E_Sarbeco_F1 | ACAGGTACGTTAATAGTTAATAGCGT | 26269..26294 | |
| | | Reverse | E_Sarbeco_R2 | ATATTGCAGCAGTACGCACACA | 26360..26381 | |
| | | Probe | E_Sarbeco_P1 | FAM-ACACTAGCCATCCTTACTGCGCTTCG-BBQ | 26332..26357 | |
| 5 | N | Forward | NbatCoV_F1 | TTTGGTGGACCCTCAGATTC | 28322..28341 | |
| | | Reverse | NbatCoV_R1 | GGGTGCCAATGTGATCTTTT | 28699..28718 | |
| 6 | N | Forward | NIID_2019-nCOV_N_F2 | AAATTTTGGGGACCAGGAAC | 29125..29144 | |
| | | Reverse | NIID_2019-nCOV_N_R2 | TGGCAGCTGTGTAGGTCAAC | 29263..29282 | 29277 |
| | | Probe | NIID_2019-nCOV_N_P2 | FAM-ATGTCGCGCATTGGCATGGA-BHQ | 29222..29241 | |
| 7 | N | Forward | 2019-nCoV_N1-F | GACCCCAAAATCAGCGAAAT | 28287..28306 | 28291 |
| | | Reverse | 2019-nCoV_N1-R | TCTGGTTACTGCCAGTTGAATCTG | 28335..28358 | |
| | | Probe | 2019-nCoV_N1-P | FAM-ACCCCGCATTACGTTTGGTGGACC-BHQ1 | 28309..28332 | |
| 8 | N | Forward | 2019-nCoV_N2-F | TTACAAACATTGGCCGCAAA | 29164..29183 | |
| | | Reverse | 2019-nCoV_N2-R | GCGCGACATTCCGAAGAA | 29213..29230 | |
| | | Probe | 2019-nCoV_N2-P | FAM-ACAATTTGCCCCCAGCGCTTCAG-BHQ1 | 29188..29210 | 29200 |
| 9 | N | Forward | 2019-nCoV_N3-F | GGGAGCCTTGAATACACCAAAA | 28681..28702 | 28688 |
| | | Reverse | 2019-nCoV_N3-R | TGTAGCACGATTGCAGCATTG | 28732..28752 | |
| | | Probe | 2019-nCoV_N3-P | FAM-AYCACATTGGCACCCGCAATCCTG-BHQ1 | 28704..28727 | |
| 10 | ORF1ab | Forward | F | CCCTGTGGGTTTTACACTTAA | 13342..13362 | |
| | | Reverse | R | ACGATTGTGCATCAGCTGA | 13442..13460 | |
| | | Probe | P | FAM-CCGTCTGCGGTATGTGGAAAGGTTATGG-BHQ1 | 13377..13404 | |
| 11 | N | Forward | F | GGGGAACTTCTCCTGCTAGAAT | 28881..28902 | |
| | | Reverse | R | CAGACATTTTGCTCTCAAGCTG | 28958..28979 | |
| | | Probe | P | FAM-TTGCTGCTGCTTGACAGATT-TAMRA | 28934..28953 | |
| 12 | E | Forward | F | ACTTCTTTTTCTTGCTTTCGTGGT | 26295..26318 | |
| | | Reverse | R | GCAGCAGTACGCACACAATC | 26357..26376 | |
| | | Probe | P | CY5-CTAGTTACACTAGCCATCCTTACTGC-BHQ1 | 26326..26351 | |
| 13 | N | Forward | WH-NIC N-F | CGTTTGGTGGACCCTCAGAT | 28320..28339 | |
| | | Reverse | WH-NIC N-R | CCCCACTGCGTTCTCCATT | 28358..28376 | |
| | | Probe | WH-NIC N-P | FAM-CAACTGGCAGTAACCA- BQH1 | 28341..28356 | |

The reference genome is IVDC-HB-01. The underlined letters in the column of primer sequence represent the variant sites. 2019-nCoV: 2019 Novel coronavirus; RT-qPCR: Real-time quantitative polymerase chain reaction; Pos: Position; Ref: Reference genome.

is based. All submitters of data may be contacted directly via https://www.gisaid.org.


## Funding

This work was supported by the National Key Research & Development Program of China (No. 2020YFC0840900) and the Beijing Municipal Science and Technology Project (No. Z201100001020004).


## Conflicts of interest

None.

## References


1. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 2013;30:772–780. doi: 10.1093/molbev/mst010.
2. Süss B, Flekna G, Wagner M, Hein I. Studying the effect of single mismatches in primer and probe binding regions on amplification curves and quantification in real-time PCR. J Microbiol Methods 2009;76:316–319. doi: 10.1016/j.mimet.2008.12.003.



**How to cite this article:** Fan H, Zhang XLL, Zhang YW, Huang Y, Teng Y, Guo Y, Mi ZQ, Yang RF, Song YJ, Cui YJ. *In silico* assessment of the impact of 2019 novel coronavirus genomic variation on the efficiency of published real-time quantitative polymerase chain reaction detection assays. Chin Med J 2020;133:1612–1613. doi: 10.1097/CM9.0000000000000817