Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

Summarize of the available molecular methods for detection of SARS-CoV-2 during the ongoing pandemic

Fabio Arena[1,2,*], Simona Pollini [3,4], Gian Maria Rossolini [3,4] and Maurizio Margaglione [1]

Affiliations:

1. Department of Clinical and Experimental Medicine, University of Foggia, Foggia, Italy.
2. IRCCS Don Carlo Gnocchi Foundation, Florence, Italy.
3. Department of Experimental and Clinical Medicine, University of Florence, Florence, Italy.
4. Clinical Microbiology and Virology Unit, Florence Careggi University Hospital, Florence, Italy.

*Correspondence: FA, fabio.arena@unifg.it; Tel. 0881/588064, Dipartimento di Medicina Clinica e Sperimentale, Università degli Studi di Foggia, Centro di Ricerche Biomediche "E. Altomare" Via Luigi Pinto – 71122, Foggia.

Abstract:

Starting from early 2020, the COVID-19 pandemic has caused a morbidity and mortality excess worldwide. Containment strategies rely firstly on rapid and sensitive laboratory diagnosis with molecular detection of the viral genome in respiratory samples being the gold standard. Reliability of diagnostic protocols could be affected by SARS-CoV-2 genetic variability. In fact, mutations occurring during SARS-CoV-2 genomic evolution can involve the regions targeted by the diagnostic probes. Following a review of the literature and an *in silico* analysis of the most recently described virus variants (including the UK B 1.1.7 and the South Africa 501Y.V2 variants), we conclude that the described genetic variability should have minimal or no effect on the sensitivity of existing diagnostic protocols for SARS-CoV-2 genome detection. However, given the continuous emergence

© 2021 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

of new variants, the situation should be monitored in the future, and protocols including multiple targets should be preferred.

Keywords: molecular methods, virus, pandemic, Coronaviridae, COVID-19, diagnosis, UK variant, B.1.1.7 variant, 501Y.V2 variant.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

Background

In March 2020, the World Health Organization declared the severe respiratory disease caused by a new coronavirus (initially named 2019nCoV), named COVID-19, a global pandemic (https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020). At the same time, the Coronaviridae Study Group (CSG) of the International Committee on Taxonomy of Viruses designated the virus responsible for COVID-19 as SARS-CoV-2 [1].

Thenceforth, COVID-19 has continued expanding globally causing almost 80 millions infections and claiming more than 1.5 millions lives from January 2020, as reported to the World Health Organization (situation updated as of 29 December 2020).

SARS-CoV-2 genomic features and variability

SARS-CoV-2 belongs to the Coronaviridae family, subfamily Orthocoronavirinae. This subfamily is further divided into four genera, namely Alphacoronavirus, Betacoronavirus, Gammacoronavirus and Deltacoronavirus. The majority of clinically relevant Coronaviridae belong to the Alphacoronavirus and Betacoronavirus [2]. The Alphacoronavirus and Betacoronavirus genera are currently divided into twelve and five subgenera, respectively, able to cause infections in a wide range of animal hosts (mainly bats but also cows, dogs, horses, pigs and dromedaries). Coronaviridae infecting humans belong to the following subgenera: Duvinacovirus and Setracovirus for Alphacoronavirus, and Embecovirus, Sarbecovirus, Merbecovirus for Betacoronavirus (https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?mode=Undef&id=694002&lvl=3&p=has_linkout&p=blast_url&p=genome_blast&keep=1&srchmode=1&unlock). Subgenus Merbecovirus comprises the Middle East respiratory syndrome (MERS)-related coronaviruses. SARS-CoV-2, together with SARS-CoV (responsible for the 2002-03 SARS outbreak) are currently classified within the subgenus Sarbecovirus.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

The Betacoronaviruses, like all other members of the Coronaviridae family, have relatively large RNA genomes of around 30 kb in size. The genomes have short untranslated regions (UTR) at both ends, with a 5' methylated cap and a 3' polyadenylated tail. Typically, Coronaviridae genomes contain 9-12 open reading frames (ORF), six of which are conserved and follow the same order: the replicase/transcriptase polyproteins and the spike (S), envelope (E), membrane (M) and nucleocapsid (N) structural proteins. Replicase/transcriptase is organized in two overlapping ORFs, called ORF1a (11-13 kb) and ORF1b (7-8 kb), that occupy nearly two thirds of the genome. These ORFs are translated into two polyproteins that later cleave themselves to form several nonstructural proteins, most of which involved in genome replication and translation [3].

The first whole genome sequence of SARS-CoV-2 (strain Wuhan-HU-1) was deposited in the NCBI Genbank on January 5th 2020 [4]. Since then, the number of available genomes has increased dramatically during the pandemic, with thousands of SARS-CoV-2 whole-genome sequences available on the rapid data sharing service hosted by the Global Initiative on Sharing All Influenza Data (GISAID; https://www.epicov.org). Soon after the beginning of the pandemic, it appeared evident that SARS-CoV-2 is a recombinant virus between the bat coronavirus and a coronavirus of unknown origin [5].

The virus was first reported in the city of Wuhan, China [4-6], where an intermediate host, with high probability an animal sold at the seafood market in Wuhan, has likely facilitated the emergence of the virus in humans [7,8]. The early phases of dissemination of the virus outside China were linked to intercontinental travel originating to multiple introductions of different subclones in various geographic regions [9-11].

Even if Coronaviridae have the capacity of proofreading during genome replication, thanks to the presence of a non-structural exonuclease able to excise erroneous nucleotides inserted by the RNA polymerase [12], the SARS-CoV-2 global population has accumulated considerable genetic diversity at this stage of the COVID-19 pandemic [13]. Available data suggest that the SARS-CoV-2 genome accumulates variability at a rate of approximately 9.8×10^{-4} substitutions per site per year [13-17].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

Mutations are generally rapidly purged from the viral population if highly deleterious. By contrast, neutral and advantageous mutations can reach higher frequencies. Some mutations may facilitate SARS-CoV-2 adaptation to the human host (decreasing virulence, increasing transmissibility or escaping immune responses) and could emerge repeatedly and independently under natural selection. A series of small deletions across the whole genome and single nucleotide polymorphisms (SNPs) occurring with high frequency have been identified and are summarized in figure 1 [14].

Genomic variability allows the classification of several SARS-CoV-2 lineages. The two major classification efforts have been produced by GISAID (https://www.gisaid.org/references/statements-clarifications/clade-and-lineage-nomenclature-aids-in-genomic-epidemiology-of-active-hcov-19-viruses/) and Nextstrain (https://nextstrain.org/ncov) initiatives, respectively. Nextstrain assigns nomenclature through designation of SARS-CoV-2 clades to label well-defined clades that reached geographic spread with significant frequency [18]. GISAID clade definitions are informed by the statistical distribution of genome distances in phylogenetic clusters followed by merging of smaller lineages into major clades based on shared marker variants [19]. The two systems produce largely overlapping phylogenetic trees [20].

The first isolates that appeared in Wuhan in December 2019 belonged to the L clade (GISAID classification). Its first variant, the S clade, appeared at the beginning of 2020, while, since mid-January 2020, the V and G variants became prevalent. To date, clade G is the most widespread, and has evolved in three subclades, namely GR and GH (which appeared at the end of February 2020) and GV, appeared later. Today, GR, GH and GV are by far the most widespread in Europe. In North America, the most widespread is GH, while in South America GR seems prevalent. In Asia, where the Wuhan L strain initially appeared, the spread of G, GH and GR is increasing. Globally, G, GH and GR are constantly increasing while S, L and V strains are gradually disappearing [21 and https://www.gisaid.org/epiflu-applications/phylodynamics/].

More recently a distinct phylogenetic cluster derived from the SARS-CoV-2 GR clade (named lineage B.1.1.7) has spread rapidly starting from early December 2020 in UK locations. The

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

emergence of this variant is cause of concerns since it seems to be associated with increased transmissibility and an unusually large number of genetic changes, particularly in the spike protein (https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563). However, some of the modifications occurring in the S protein of the B.1.1.7 lineage (e.g. the N501Y substitution and the deletion of 6 bases at positions 69 and 70, respectively, in the viral S gene) have been circulating, globally for many months previously [20,22]. Another emerging lineage named 501Y.V2, characterized by some lineage-defining mutations in the spike protein, has spread rapidly, becoming within weeks the dominant lineage in the Eastern Cape and Western Cape Provinces (South Africa) [23].

Epidemiological investigations aiming at assessing new virus variants and their spreading are useful to prioritize relevant mutations and unravel their potential impact on molecular diagnostics.
The information on genomic variability should be taken into account when a new diagnostic assay is released and to monitor the reliability of already released methods. Ideally, diagnostics should target relatively invariant, strongly constrained regions of the SARS-CoV-2 genome, while multiple targets are preferred to increase detection sensitivity.

Diagnostic tests for SARS-CoV-2 infection

Diagnostic tests for SARS-CoV-2 infection belong to three categories including: i) nucleic acid amplification tests (NAATs) detecting the presence of viral RNA by RT-PCR or other amplification methods; ii) tests detecting the presence of viral antigens; and iii) tests detecting the presence of antibodies against SARS-CoV-2 antigens (Table 1).
NAATs detecting viral RNA in respiratory specimens remain the reference test for diagnosis of SARS-CoV-2 infection in patients with COVID-19-like symptoms and for patient triage and isolation in healthcare facilities, including long-term care facilities, outbreak investigations, and contact tracing activities. Testing at pre-determined time intervals can also be adopted as screening for infection in

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

certain high-risk groups, like healthcare workers and essential services workers as part of surveillance programs.

Soon after the emergence in China, in January 2020 WHO announced several RT-PCR-based diagnostic schemes for SARS-CoV-2, based on the amplification of different viral targets (details available at https://www.who.int/docs/default-source/coronaviruse/whoinhouseassays.pdf) including some specific for SARS-CoV-2 (i. e. those targeting the viral RNA-dependent RNA polymerase-encoding RdRp gene and the viral nucleocapsid N gene) and one common to members of subgenus Sarbecovirus (i. e. the envelope E gene) (Table 2). The latter could also be used as a screening test followed by the detection of SARS-CoV-2 specific targets [24-26]. The different viral targets were associated with different specificity and sensitivity, with the E gene being reported as the target with the highest sensitivity and the RdRp as the most specific [25].

More recently, additional assays based on the isothermal amplification of viral nucleic acids, also in combination with CRISPR-based detection methods, have been developed; these methods, that do not require thermal cycling, are generally more rapid that RT-PCR, declare good sensitivity and specificity and are also considered suitable as point-of-care tests for the detection of SARS-CoV-2 [27-29].

Since the development of the first in-house diagnostic tests, several manufacturers have quickly developed commercial kits for molecular detection of SARS-CoV-2, based on existing diagnostic platforms. As a result, the number of commercial RT-PCR-based and isothermal nucleic acid amplification assays is at present considerably high and novel tests are continuously increasing the repertoire of available in vitro diagnostic assays (IVDs). Since the pandemic beginning, numerous tests have received the CE-IVD mark or the Food and Drug Administration (FDA) emergency use authorization (EUA) that is required to be placed in the market [30] (EUA assays are available at https://www.fda.gov/emergency-preparedness-and-response/mcm-legal-regulatory-and-policy-framework/emergency-use-authorization#covidinvitrodev). An online tool for existing SARS-CoV-2 assays performance comparation is available at https://www.finddx.org/covid-19/sarscov2-eval/.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

Furthermore, a meta-analysis focused on diagnostic accuracy of point-of-care antigen and molecular-based methods for COVID-19 diagnosis has recently been published. The review underscored a high variability in sensitivity of rapid test across available studies (especially for antigen tests) [31]. The vast majority of these assays are RT-PCR schemes that require a separated viral RNA extraction step. Most of them target multiple viral genes (in most cases the N and Orf1ab/RdRp genes), with a minority only being able to detect a single gene. The detection of multiplex viral targets has the potential advantage of improving test sensitivity, particularly in case of low viral load in the initial specimen or RNA degradation during specimen handling, or in the event of viral genome mutations affecting one of the targeted regions [32]. In fact, test sensitivity is an important issue in the present scenario, where many assays, that may differ in their capability of viral genome detection, are proposed for laboratory diagnosis. Available commercial assays declare limits of detection in a quite large range (from less than 1 copy/PFU to up to 1000 copies/PFU per mL) (https://www.fda.gov/emergency-preparedness-and-response/mcm-legal-regulatory-and-policy-framework/emergency-use-authorization#covidinvitrodev), possibly resulting in detection differences (particularly for low viral loads) among laboratories that use different diagnostic tests or in centers running side-by-side multiple assays on routine samples. However, the significance of extremely low viral loads remains to be ascertained, since in these cases, the SARS-CoV-2 quantity is apparently below the threshold at which replication-competent virus can be isolated by culture methods [33,34].

Although WHO diagnostic schemes have been deemed the gold standard at the beginning of the pandemic, they required specialized reagents, equipment, and personnel training. In the pandemic scenario, the possibility of rapidly scaling-up the number of tests and of automation are crucial points in helping to face the ever-increasing number of required tests. Many commercially available tests can be automated by using robotic platforms able to separately extract viral RNA and prepare PCR assays, for high-throughput batch processing of clinical specimens (Table 3). This approach may still require some significant expertise, dedicated equipment, and relatively long turn-around-times

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

(TAT). At present, there are also a number of assays proposed as sample-to-result platforms (Table 3). Some of them (i.e. Hologic Panther/Aptima SARS-CoV-2 assay, Roche Cobas 6800/8800/cobas SARS-CoV-2, and Abbott Alinity m System/Alinity m SARS-CoV-2 assay) are high-throughput methods with a turnaround-time of approx. 2 to 3.5, while others can perform a smaller volume of tests with similar or reduced TAT (i.e. ELITech InGenius platform/SARS-CoV-2 ELITe MGB® Kit and Simplexa COVID-19 Direct assay), or are rapid single-test assays that give results in as fast as 13 minutes (i.e. Abbott ID Now COVID-19 assay) and up to 40-50 minutes (e.g. Cepheid Xpert® Xpress SARS-CoV-2, Bosch Vivalytic VRI test and Menarini VitaPCR™ SARS-CoV-2 assay) (Table 3). Some RT-PCR assays (e.g. BIOFIRE® Respiratory Panel 2.1 and QIAstat-Dx Respiratory SARS-CoV-2 Panel) are also developed as rapid syndromic panels; such tests are usually single-sample assays able to give results in up to 1 hour and may help healthcare providers to rapidly discriminate between common respiratory pathogens (e.g. flu and other viral pathogens) and SARS-CoV-2 in patients with COVID-19-like symptoms. Regarding the rapid assays, the availability of such tests appears to be of particular importance in managing suspect SARS-CoV-2-positive patients, mostly for fast patients' triage and correct isolation procedures in the emergency departments.

Influence of SARS-CoV-2 genetic variability on molecular diagnostic protocols

Several studies have previously evaluated *in silico* the potential effect of mutations occurring in the target regions of published assays listed by the WHO and other Agencies (Table 4).

In June 2020, Khan and Cheung published an exhaustive evaluation of the sequence variability within the primer/probe target regions of the viral genome using more than 17000 viral sequences from around the world (35). Overall, the authors found moderate mutation rate in the SARS-CoV-2 genome regions of interest. However, they reported a mismatch with all the viral sequences in the region of the Charité-ORF1b primer. Furthermore, they found a relatively high frequency of mutation in the region of the forward N gene primer released by the China CDC.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

Three independent works confirmed, on a global and local scale, the overall high inclusivity of publicly available sets of primers and probes, with the exception of the forward N China CDC primer (occurrence of the so called "AAC" variant) and set N3 of the US CDC, subsequently withdrawn [35-39].

In the largest bioinformatic project on this topic, Peñarrubia L. et al. analyzed 9 different publicly available primers/probe sets with more than 30.000 genomes. The authors found a relatively high frequency of mutations in the regions of interest of various primers (approx. 34% of included genomes) and concluded that adopting multiple targets approaches may mitigate the risk of loss of sensitivity [32] (Table 4). However, all these authors found only a small proportion of mutations involving the more problematic 3' end of primers annealing regions, therefore, they concluded that, from a practical point of view, the impact of genetic variability on primers reliability should be minimal.

To expand the analysis to more recent sequences we downloaded the selection of genomes present in the primary global analysis of the GISAID interactive database on December 9, 2020 (https://www.gisaid.org/epiflu-applications/phylodynamics/) and analyzed the variability in the WHO RT-PCR primers/probes regions of interest (Table 5). The GISAID system automatically subsamples 120 genomes per admin division (geographical area), per month, to obtain a more geographically representative subset. We further customized the analysis including only genomes obtained from human source and sequences uploaded from November 9th to December 9th.

A total of 1251 sequences were therefore included in the study and were available for further analysis (Table S1). The majority of the sequences included in this study originated from Europe (27.6%), North America (26.6%) and Africa (21.1%), while there was only a relatively small number sequences from Asia, Oceania and South America (cumulatively 24.7%). For the analysis of presence of mismatches between PCR assays primers and probes and the selected SARS-CoV-2 genomes we used a previously validated bioinformatic software pipeline [35]. Only the sequences variants with occurrence >=1%, among included genomes, were further considered and are shown in Table 5. The

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

Wuhan-Hu-1/2019, 29,903 bp long, genome (access. MN908947.3) was obtained from NCBI GenBank and used as reference for site numbering and identification of primers and probes regions of interest.

Interestingly, most primer/probe binding regions showed no mutations or mutations/mismatches with a few sequences (one or two genomes, frequency <1%) representing probably extremely low prevalent variants or sequencing errors. In seven cases only, the frequency of occurrence of single mutations reached 1%. The most relevant phenomenon was that of the circulation of the known "AAC" variant, causing mismatch in the region of interest of the China CDC N forward primer. The variant was mainly found among recently sequenced genomes from Europe and North America, belonging to the GR and GH clades. Issues with this set of primers/probe were also linked to the presence of a mutation (G28975C) found in approx. 4% of genomes, in the reverse primer region. This mutation also was found among GH clade genomes from Europe.

Adjunctively we expanded the analysis to more recent variants. UK B.1.1.7 variant has an unusually large number of genetic changes in the spike protein. Assays targeting the S-gene are not widely used for viral detection. Furthermore, relying only on one target for detection of SARS-CoV-2 infection using RT-PCR is not recommended. However, the B.1.1.7 lineage shows higher rate of molecular evolution, compared to other SARS-CoV-2 lineages, also outside the S gene. We screened a representative genome belonging to B.1.1.7 variant (GISAID EPI_ISL_744131) for the presence of mismatches in regions of interest of the WHO RT-PCR primers/probes sets. Using the online NIH Basic Local Alignment Search Tool, we found a perfect match between primers/probes released by WHO (Table 2) and the genome sequence of the variant. As expected, the only mismatches found were in the first three position of the China CDC N forward primer ("AAC variant"). The same analysis was conducted with the EPI_ISL_660190 genome, representative of the South Africa 501Y.V2 variant, and we found only two mismatches involving the central parts of China CDC N forward primer and the Japan National Institute of Infectious Diseases N reverse primer.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

Fewer information are available on detection capabilities of commercial tests, based on proprietary primers and probes. Many suppliers declare the ability of their test to in silico detect the viral variants that were circulating and included in freely available repositories (e.g. NCBI Genbank and GISAID) at the time of test approval; anyway, proprietary PCR primers sequences included in the assays are unavailable, preventing assessment by the users of the test ability for detection of new viral variants. A number of suppliers (e.g. Hologic, Roche, ELITEch, Diasorin, Seegene) have declared the ability of their tests to detect the B.1.1.7 SARS-CoV-2 variant.

Conclusions

Overall, considering previously published data, the bioinformatic analysis performed in this work and the information provided by companies producing commercial diagnostic systems, we can conclude that, currently, the known variability occurring in the SARS-CoV-2 population have minimal or no effect on the sensitivity of existing molecular systems for viral detection. Furthermore, the majority of mismatches observed were not near the 3′ end and should be tolerated.

The only exception that should be mentioned is the three nucleotide substitutions (GGG→AAC) which occur frequently in the three first positions of the China CDC-N forward primer binding site. The so called “AAC variant” is particularly frequent among the GR and GH recent clades.

Based on available sequencing data, also the new UK B.1.1.7 and the South Africa 501Y.V2 variants should be reliably recognized by most widely used commercial and in-house protocols.

However, the continuous description of novel genomic variants represents an important diagnostic issue that need to be monitored in the future, while a multiple target strategy is suggested to minimize the effect SARS-CoV-2 genetic variability on assays sensitivity.

Given the increasing importance of genome sequencing data availability (to monitor the variability of the viral population), the development of rapid, inexpensive and standardized laboratory/informatic pipelines for SARS-CoV-2 genome sequencing is urgently needed.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

289 **Supplementary Materials:** Table S1: GISAID "Strains metadata" of genomes included in
290 this work and full listing of all originating and submitting laboratories.

291 **Funding:** this research received no external funding

292 **Acknowledgments:** we are grateful to Mr. Antonio Bratti for his support with the
293 bioinformatic analysis.

294 **Author Contributions:** Conceptualization, F.A.; Software, F.A.; Writing – Original Draft
295 Preparation, F.A., S.P. G.M.R. and M.M.; Writing – Review & Editing, F.A., S.P. G.M.R.
296 and M.M.

297 **Institutional Review Board Statement:** Not applicable

298 **Data Availability Statement:** Figure 1 was obtained from an open access and interactive web
299 resource at: https://macman123.shinyapps.io/ugi-scov2-alignment-screen/. SARS-CoV-2
300 genome data for screening of mismatch with primers/probes sequences were downloaded
301 from GISAID for SARS-CoV-2 online repository available at: https://www.epicov.org. The
302 online tool for comparison of SARS-CoV-2 diagnostic tests performance is available at:
303 https://www.finddx.org/covid-19/sarscov2-eval/. The list of emergency use authorization
304 (EUA) diagnostics is available at: https://www.fda.gov/emergency-preparedness-and-
305 response/mcm-legal-regulatory-and-policy-framework/emergency-use-
306 authorization#covidinvitrodev. The genome sequence EPI_ISL_744131 was sequenced at
307 Wales Specialist Virology Centre Sequencing lab and submitted to GISAID by Catherine
308 Moore C. et al.. The genome sequence EPI_ISL_660190 was sequenced at KRISP, KZN
309 Research Innovation and Sequencing Platform, and submitted to GISAID by Giandhari J. et
310 al A

311 **Conflicts of Interest:** The authors declare no conflict of interest relevant to this work.

312

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

References:

1) Coronaviridae Study Group of the International Committee on Taxonomy of Viruses. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nat Microbiol. 2020 Apr;5(4):536-544. doi: 10.1038/s41564-020-0695-z.
2) Cui J, Li F, Shi ZL. Origin and evolution of pathogenic coronaviruses. Nat Rev Microbiol. 2019 Mar;17(3):181-192. doi: 10.1038/s41579-018-0118-9.
3) Sola I, Almazán F, Zúñiga S, Enjuanes L. Continuous and Discontinuous RNA Synthesis in Coronaviruses. Annu Rev Virol. 2015;2(1):265-88. doi: 10.1146/annurev-virology-100114-055218.
4) Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, Hu Y, Tao ZW, Tian JH, Pei YY, Yuan ML, Zhang YL, Dai FH, Liu Y, Wang QM, Zheng JJ, Xu L, Holmes EC, Zhang YZ. A new coronavirus associated with human respiratory disease in China. Nature. 2020;579(7798):265-269. doi: 10.1038/s41586-020-2008-3.
5) Ji W, Wang W, Zhao X, Zai J, Li X. Cross-species transmission of the newly identified coronavirus 2019-nCoV. J Med Virol. 2020;92(4):433-440. doi: 10.1002/jmv.25682.
6) Benvenuto D, Giovanetti M, Salemi M, Prosperi M, De Flora C, Junior Alcantara LC, Angeletti S, Ciccozzi M. The global spread of 2019-nCoV: a molecular evolutionary analysis. Pathog Glob Health. 2020;114(2):64-67. doi: 10.1080/20477724.2020.1725339.
7) Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, Wang W, Song H, Huang B, Zhu N, Bi Y, Ma X, Zhan F, Wang L, Hu T, Zhou H, Hu Z, Zhou W, Zhao L, Chen J, Meng Y, Wang J, Lin Y, Yuan J, Xie Z, Ma J, Liu WJ, Wang D, Xu W, Holmes EC, Gao GF, Wu G, Chen W, Shi W, Tan W. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet. 2020 Feb 22;395(10224):565-574. doi: 10.1016/S0140-6736(20)30251-8.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

8) Lam TT, Jia N, Zhang YW, Shum MH, Jiang JF, Zhu HC, Tong YG, Shi YX, Ni XB, Liao YS, Li WJ, Jiang BG, Wei W, Yuan TT, Zheng K, Cui XM, Li J, Pei GQ, Qiang X, Cheung WY, Li LF, Sun FF, Qin S, Huang JC, Leung GM, Holmes EC, Hu YL, Guan Y, Cao WC. Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. Nature. 2020;583(7815):282-285. doi: 10.1038/s41586-020-2169-0.

9) Giovanetti M, Angeletti S, Benvenuto D, Ciccozzi M. A doubt of multiple introduction of SARS-CoV-2 in Italy: A preliminary overview. J Med Virol. 2020;92(9):1634-1636. doi: 10.1002/jmv.25773.

10) Gonzalez-Reiche AS, Hernandez MM, Sullivan MJ, Ciferri B, Alshammary H, Obla A, Fabre S, Kleiner G, Polanco J, Khan Z, Alburquerque B, van de Guchte A, Dutta J, Francoeur N, Melo BS, Oussenko I, Deikus G, Soto J, Sridhar SH, Wang YC, Twyman K, Kasarskis A, Altman DR, Smith M, Sebra R, Aberg J, Krammer F, García-Sastre A, Luksza M, Patel G, Paniz-Mondolfi A, Gitman M, Sordillo EM, Simon V, van Bakel H. Introductions and early spread of SARS-CoV-2 in the New York City area. Science. 2020;369(6501):297-301. doi: 10.1126/science.abc1917.

11) Stefanelli P, Faggioni G, Lo Presti A, Fiore S, Marchi A, Benedetti E, Fabiani C, Anselmo A, Ciammaruconi A, Fortunato A, De Santis R, Fillo S, Capobianchi MR, Gismondo MR, Ciervo A, Rezza G, Castrucci MR, Lista F, On Behalf Of Iss Covid-Study Group. Whole genome and phylogenetic analysis of two SARS-CoV-2 strains isolated in Italy in January and February 2020: additional clues on multiple introductions and further circulation in Europe. Euro Surveill. 2020;25(13):2000305. doi: 10.2807/1560-7917.ES.2020.25.13.2000305.

12) Snijder EJ, Bredenbeek PJ, Dobbe JC, Thiel V, Ziebuhr J, Poon LL, Guan Y, Rozanov M, Spaan WJ, Gorbalenya AE. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J Mol Biol. 2003;331(5):991-1004. doi: 10.1016/s0022-2836(03)00865-9.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

13) van Dorp L, Acmana M, Richardb D, Shawd LP, Forda CE, Ormonda L, Owena CJ, Panga J, Tana CCS, BoshiereFAT, Ortiza AT, Ballouxa F. Emergence of genomic diversity and recurrent mutations in SARS-CoV-2. Infection, Genetics and Evolution. 2020; 83-10435. doi.org/10.1111/tbed.13931.

14) van Dorp L, Richard D, Tan CCS, Shaw LP, Acman M, and Balloux F. No evidence for increased transmissibility from recurrent mutations in SARS-CoV-2. Nat Commun. 2020; 11: 5986. doi: 10.1038/s41467-020-19818-2

15) Li XG, et al. Transmission dynamics and evolutionary history of 2019-nCoV. J. Med. Virol. 2020;92:501–511. doi: 10.1002/jmv.25701.

16) Giovanetti M, et al. The first two cases of 2019-nCoV in Italy: where they come from? J. Med. Virol. 2020;92:518–521. doi: 10.1002/jmv.25699.

17) Duchene S,Featherstone L,Haritopoulou-Sinanidou M,Rambaut A,Lemey P, and Baele G. "Temporal Signal and the Phylodynamic Threshold of SARS-CoV-2." Virus Evol. 2020 Aug 19;6(2):veaa061. doi: 10.1093/ve/veaa061.

18) Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, Sagulenko P, Bedford T, Neher RA Nextstrain: real-time tracking of pathogen evolution. Bioinformatics. 2018;34(23):4121-4123. doi: 10.1093/bioinformatics/bty407.

19) Han AX, Parker E, Scholer F, Maurer-Stroh S, Russell CA. Phylogenetic Clustering by Linear Integer Programming (PhyCLIP). Mol Biol Evol. 2019;36(7):1580-1595. doi: 10.1093/molbev/msz053.

20) Alm E, Broberg EK, Connor T, Hodcroft EB, Komissarov AB, Maurer-Stroh S, Melidou A, Neher RA, O'Toole Á, Pereyaslov D; WHO European Region sequencing laboratories and GISAID EpiCoV group; WHO European Region sequencing laboratories and GISAID EpiCoV group*. Geographical and temporal distribution of SARS-CoV-2 clades in the WHO European Region, January to June 2020. Euro Surveill. 2020;25(32):2001410. doi: 10.2807/1560-7917.ES.2020.25.32.2001410.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

21) Mercatelli D, Giorgi FM. Geographic and Genomic Distribution of SARS-CoV-2 Mutations. Front Microbiol. 2020;11:1800. doi: 10.3389/fmicb.2020.01800.

22) Kemp SA, Datir RP, Collier DA, et al. Recurrent emergence and transmission of a SARS-CoV-2 Spike deletion DeltaH69/DeltaV70. bioRxiv 2020.12.14.422555; doi: 10.1101/2020.12.14.422555

23) Tegally H,Wilkinson E,Giovanetti M et al. Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. medRxiv 2020.12.21.20248640; doi:https://doi.org/10.1101/2020.12.21.20248640

24) World Health Organization. Coronavirus Disease (COVID-19) Technical Guidance: Laboratory Testing for 2019-nCoV in Humans (19 March 2020). Available online at: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance-publications.

25) Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DK, Bleicker T, Brünink S, Schneider J, Schmidt ML, Mulders DG, Haagmans BL, van der Veer B, van den Brink S, Wijsman L, Goderski G, Romette JL, Ellis J, Zambon M, Peiris M, Goossens H, Reusken C, Koopmans MP, Drosten C. 2020. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro Surveill 25(3):pii=2000045. doi:10.2807/1560-7917.ES.2020.25.3.2000045.

26) Holshue ML, DeBolt C, Lindquist S, Lofy KH, Wiesman J, Bruce H, Spitters C, Ericson K, Wilkerson S, Tural A, Diaz G, Cohn A, Fox L, Patel A, Gerber SI, Kim L, Tong S, Lu X, Lindstrom S, Pallansch MA, Weldon WC, Biggs HM, Uyeki TM, Pillai SK, Washington State 2019-nCoV Case Investigation Team. 2020. First case of 2019 novel coronavirus in the United States. N Engl J Med 382:929 –936. https://doi.org/10.1056/NEJMoa2001191.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

27) Esbin MN, Whitney ON, Chong S, Maurer A, Darzacq X, Tjian R. Overcoming the bottleneck to widespread testing: a rapid review of nucleic acid testing approaches for COVID-19 detection. RNA. 2020;26(7):771-783. doi: 10.1261/rna.076232.120.

28) James AS, Alwneh JI. COVID-19 Infection Diagnosis: Potential Impact of Isothermal Amplification Technology to Reduce Community Transmission of SARS-CoV-2. Diagnostics (Basel). 2020;10(6):399. doi: 10.3390/diagnostics10060399.

29) Shen M, Zhou Y, Ye J, Abdullah Al-Maskri AA, Kang Y, Zeng S, Cai S. Recent advances and perspectives of nucleic acid detection for coronavirus. J Pharm Anal. 2020;10(2):97-101. doi: 10.1016/j.jpha.2020.02.010.

30) World Health Organization. Diagnostic testing for SARS-CoV-2, Interim guidance (11 September 2020). Available online at: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance-publications.

31) Dinnes J, Deeks JJ, Adriano A, Berhane S, Davenport C, Dittrich S, Emperador D, Takwoingi Y, Cunningham J, Beese S, Dretzke J, Ferrante di Ruffano L, Harris IM, Price MJ, Taylor-Phillips S, Hooft L, Leeflang MM, Spijker R, Van den Bruel A; Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection. Cochrane COVID-19 Diagnostic Test Accuracy Group.Cochrane Database Syst Rev. 2020 Aug 26;8:CD013705. doi: 10.1002/14651858.CD013705.

32) Peñarrubia L, Ruiz M, Porco R, Rao SN, Juanola-Falgarona M, Manissero D, López-Fontanals M, Pareja J. Multiple assays in a real-time RT-PCR SARS-CoV-2 panel can mitigate the risk of loss of sensitivity by new genomic variants during the COVID-19 outbreak. Int J Infect Dis. 2020;97:225-229. doi: 10.1016/j.ijid.2020.06.027.

33) Wolfel, R, Corman, VM, Guggemos W, Seilmaier M, Zange S, Müller MA, Niemeyer D, Jones TC, Vollmar P, Rothe C, et al. Virological assessment of hospitalized patients with COVID-19. Nature 2020, 581, 465–469. doi: 10.1038/s41586-020-2196-x.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

34) La Scola B, Le Bideau M, Andreani J, Hoang VT, Grimaldier C, Colson P, Colson P, Raoult D. Viral RNA load as determined by cell culture as a management tool for discharge of SARS-CoV-2 from infectious disease wards. Eur. J. Clin. Microbiol. Infect. Dis. 2020, 39, 1059–1061. doi: 10.1007/s10096-020-03913-9.

35) Khan KA, Cheung P. Presence of mismatches between diagnostic PCR assays and coronavirus SARS-CoV-2 genome. R Soc Open Sci. 2020;7(6):200636. doi: 10.1098/rsos.200636.

36) Vogels CBF, Brito AF, Wyllie AL, Fauver JR, Ott IM, Kalinich CC, Petrone ME, Casanovas-Massana A, Catherine Muenker M, Moore AJ, Klein J, Lu P, Lu-Culligan A, Jiang X, Kim DJ, Kudo E, Mao T, Moriyama M, Oh JE, Park A, Silva J, Song E, Takahashi T, Taura M, Tokuyama M, Venkataraman A, Weizman OE, Wong P, Yang Y, Cheemarla NR, White EB, Lapidus S, Earnest R, Geng B, Vijayakumar P, Odio C, Fournier J, Bermejo S, Farhadian S, Dela Cruz CS, Iwasaki A, Ko AI, Landry ML, Foxman EF, Grubaugh ND. Analytical sensitivity and efficiency comparisons of SARS-CoV-2 RT-qPCR primer-probe sets. Nat Microbiol. 2020;5(10):1299-1305. doi: 10.1038/s41564-020-0761-6.

37) Gand M, Vanneste K, Thomas I, Van Gucht S, Capron A, Herman P, Roosens NHC, De Keersmaecker SCJ. Use of Whole Genome Sequencing Data for a First in Silico Specificity Evaluation of the RT-qPCR Assays Used for SARS-CoV-2 Detection. Int J Mol Sci. 2020;21(15):5585. doi: 10.3390/ijms21155585.

38) Álvarez-Díaz DA, Franco-Muñoz C, Laiton-Donato K, Usme-Ciro JA, Franco-Sierra ND, Flórez-Sánchez AC, Gómez-Rangel S, Rodríguez-Calderon LD, Barbosa-Ramirez J, Ospitia-Baez E, Walteros DM, Ospina-Martinez ML, Mercado-Reyes M. Molecular analysis of several in-house rRT-PCR protocols for SARS-CoV-2 detection in the context of genetic variability of the virus in Colombia. Infect Genet Evol. 2020;84:104390. doi: 10.1016/j.meegid.2020.104390.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

464 39) Kuchinski KS, Jassem AN, Prystajecky NA. Assessing oligonucleotide designs from early
465 lab developed PCR diagnostic tests for SARS-CoV-2 using the PCR_strainer pipeline. J Clin
466 Virol. 2020 Oct;131:104581. doi: 10.1016/j.jcv.2020.104581.
467

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

468 **Table 1. Summary of main features of existing tests for COVID-19 diagnosis**

Assay type	Principle of the assay	Intended use
Nucleic acid tests	detect the presence of viral RNA, generally by RT-PCR	decision making for clinical, infection control or public health management (screening close contacts, outbreak investigations or surveillance programmes)
Antigen tests	detect the presence of a viral antigen, typically part of a surface protein, by lateral-flow assays or chemiluminescence immunoassays	decision making for clinical, infection control or public health management (screening close contacts, outbreak investigations or surveillance programmes)
Antibody tests	detect the presence of antibodies generated against SARS-CoV-2. The three most used assays are enzyme-linked immunosorbent assays, chemoluminescence assays and lateral flow assays	sero-epidemiological surveys and studies; complement to the virus detection tests

469

470

471 **Table 2. Summary of primers/probes sets released by WHO for in-house RT-PCR detection of**

472 **SARS-CoV-2**

473

Source	Primer/probe name	Target gene	Sequence	Lenght	Genomic region*
ChinaCDC,China	正向引物（F）	ORF1ab	CCCTGTGGGTTTTACACTTAA	21	13342-13362
ChinaCDC,China	反向引物（R）	ORF1ab	ACGATTGTGCATCAGCTGA	19	13442-13460
ChinaCDC,China	荧光探针（P）	ORF1ab	CCGTCTGCGGTATGTGGAAAGGTTATGG	28	13377-13404
ChinaCDC,China	正向引物（F)	N	GGGGAACTTCTCCTGCTAGAAT	22	28881-28902
ChinaCDC,China	反向引物（R）	N	CAGACATTTTGCTCTCAAGCTG	22	28958-28979
ChinaCDC,China	荧光探针（P）	N	TTGCTGCTGCTTGACAGATT	20	28934-28953
Institut Pasteur	nCoV_IP2-12669Fw	RdRp	ATGAGCTTAGTCCTGTTG	18	12690-12707
Institut Pasteur	nCoV_IP2-12759Rv	RdRp	CTCCCTTTGTTGTGTTGT	18	12780-12797
Institut Pasteur	nCoV_IP2-12696bProbe(+)	RdRp	ATGTCTTGTGCTGCCGGTA	19	12719-12737
Institut Pasteur	nCoV_IP4-14059Fw	RdRp	GGTAACTGGTATGATTTCG	19	14080-14098
Institut Pasteur	nCoV_IP4-14146Rv	RdRp	CTGGTCAAGGTTAATATAGG	20	14167-14186

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

Institut Pasteur	nCoV_IP4-14084Probe(+)	RdRp	TCATACAAACCACGCCAGG	19	14105-14123
Institut Pasteur	E_Sarbeco_F1	E	ACAGGTACGTTAATAGTTAATAGCGT	26	26269-26294
Institut Pasteur	E_Sarbeco_R2	E	ATATTGCAGCAGTACGCACACA	22	26360-26381
Institut Pasteur	E_Sarbeco_P1	E	ACACTAGCCATCCTTACTGCGCTTCG	26	26332-26357
USCDC	2019-nCoV_N1-F	ORF9b	GACCCCAAAATCAGCGAAAT	20	28287-28306
USCDC	2019-nCoV_N1-R	ORF9b	TCTGGTTACTGCCAGTTGAATCTG	24	28335-28358
USCDC	2019-nCoV_N1-P	ORF9b	ACCCCGCATTACGTTTGGTGGACC	24	28309-28332
USCDC	2019-nCoV_N2-F	ORF9b	TTACAAACATTGGCCGCAAA	20	29164-29183
USCDC	2019-nCoV_N2-R	ORF9b	GCGCGACATTCCGAAGAA	18	29213-29230
USCDC	2019-nCoV_N2-P	ORF9b	ACAATTTGCCCCCAGCGCTTCAG	23	29188-29210
USCDC	2019-nCoV_N3-F	ORF9b	GGGAGCCTTGAATACACCAAAA	22	28681-28702
USCDC	2019-nCoV_N3-R	ORF9b	TGTAGCACGATTGCAGCATTG	21	28732-28752
USCDC	2019-nCoV_N3-P	ORF9b	ATCACATTGGCACCCGCAATCCTG	24	28704-28727
National Institute of Infectious Diseases, Japan	NIID_2019-nCOV_N_F2	N	AAATTTTGGGGACCAGGAAC	20	29142-29161
National Institute of Infectious Diseases, Japan	NIID_2019-nCOV_N_R2	N	TGGCAGCTGTGTAGGTCAAC	20	29280-29299
National Institute of Infectious Diseases, Japan	NIID_2019-nCOV_N_P2	N	ATGTCGCGCATTGGCATGGA	20	29239-29258
Charité,Germany	RdRP_SARSr-F2	RdRp	GTGAAATGGTCATGTGTGGCGG	22	15431-15452
Charité,Germany	RdRP_SARSr-R1	RdRp	CAAATGTTAAAAACACTATTAGCATA	26	15505-15530
Charité,Germany	RdRP_SARSr-P2	RdRp	CAGGTGGAACCTCATCAGGAGATGC	25	15470-15494
Charité,Germany	E_Sarbeco_F1	E	ACAGGTACGTTAATAGTTAATAGCGT	26	26269-26294
Charité,Germany	E_Sarbeco_R2	E	ATATTGCAGCAGTACGCACACA	22	26360-26381
Charité,Germany	E_Sarbeco_P1	E	ACACTAGCCATCCTTACTGCGCTTCG	26	26332-26357
HKU,HongKongSAR	HKU-ORF1b-nsp14F	ORF1b	TGGGGTTTTACAGGTAACCT	20	18778-18797
HKU,HongKongSAR	HKU-ORF1b-nsp14R	ORF1b	AACACGCTTAACAAAGCACTC	21	18889-18909
HKU,HongKongSAR	HKU-ORF1b-nsp141P	ORF1b	TAGTTGTGATGCAATCATGACTAG	24	18849-18872
HKU,HongKongSAR	HKU-NF	N	TAATCAGACAAGGAACTGATTA	22	29145-29166
HKU,HongKongSAR	HKU-NR	N	CGAAGGTGTGACTTCCATG	19	29236-29254

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

HKU,HongKongSAR	HKU-NP	N	GCAAATTGTGCAATTTGCGG	20	29177-29196
National Institute of Health, Thailand	WH-NICN-F	ORF9b	CGTTTGGTGGACCCTCAGAT	20	28320-28339
National Institute of Health, Thailand	WH-NICN-R	ORF9b	CCCCACTGCGTTCTCCATT	19	28358-28376
National Institute of Health, Thailand	WH-NICN-P	ORF9b	CAACTGGCAGTAACCA	16	28341-28356

474 *Site numbering uses Wuhan-Hu-1/2019 as reference (access.MN908947.3)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

475 **Table 3. Main CE-IVD and/or EUA-labelled integrated extraction/amplification platforms and**

476 **sample-to-result assays for the detection of SARS-CoV-2.**

Assay	Manufacturer	Viral genes	Assay/equipment type	Approx. time-to-result
Xpert® Xpress SARS-CoV-2	Cepheid	N, E	RT-PCR/single test, sample-to-result	45 min.
Vivalytic analyzer/Vivalytic VRI test	BOSCH	na[a]	RT-PCR/single test, sample-to-result	39 min.
VitaPCR™ platform/ VitaPCR™ SARS-CoV-2 assay	Menarini	N	RT-PCR/single test, sample-to-result	20 min
GenMark ePlex instrument/ePlex® SARS-CoV-2 Test	GenMark	N	RT-PCR/single test, sample-to-result	90 min.
ARIES® SARS-CoV-2 Assay	Luminex Corporation	Orf1ab, N	RT-PCR/single test, sample-to-result	2 h
ID Now COVID-19	Abbott	RdRp	Isothermal amplification/single test, sample-to-result	13 min.
Simplexa COVID-19 Direct assay	DiaSorin	orf1ab, S	RT-PCR/batch testing, sample-to-result	80 min.
ELITech InGenius platform/SARS-CoV-2 ELITe MGB® Kit	ELITech	RdRp, Orf8	RT-PCR/batch testing, sample-to-result	2 h 30 min.
Cobas 6800/8800/cobas SARS-CoV-2	Roche	orf1ab, E	RT-PCR/batch testing, sample-to-result	3 h 30 min.
Alinity m System/Alinity m SARS-CoV-2 assay	Abbott	RdRp, N	RT-PCR/batch testing, sample-to-result	2 h
NeoMoDx™ molecular system/ NeuMoDx™ SARS-CoV-2 Assay	QIAGEN	Nsp2, N	RT-PCR/batch testing, sample-to-result	80 min.
BD MAX™ System/BD SARS-CoV-2 Reagents	Becton Dickinson	N	RT-PCR/batch testing, sample-to-result	3 h
Panther/Aptima SARS-CoV-2 assay	Hologic	orf1ab	Isothermal amplification/batch testing, sample-to-result	3 h 30 min.
Seegene NIMBUS/STARlet/ Maelstrom 9600/Allplex™ SARS-CoV-2 Assay	Seegene	RdRp, N, S, E	RT-PCR/batch testing, integrated equipment for extraction and amplification[b]	From 3 h 20 min. to 4 h 40 min.
KingFisher Flex Purification system/TaqPath™ COVID-19 RT-PCR Kit	Life Technologies Corporation	orf1ab, N, S	RT-PCR/batch testing, integrated equipment for extraction and amplification	na
BIOFIRE® Respiratory Panel 2.1	Biomérieux	S, M	RT-PCR/syndromic panel, sample-to-result	45 min.
QIAstat-Dx Respiratory SARS-CoV-2 Panel	QIAGEN	RdRp, E	RT-PCR/syndromic panel, sample-to-result	60 min.

477 [a] not declared by manufacturer.

478 [b] Main assays using proprietary equipment for both extraction and amplification steps are reported.

479

480

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

481 **Table 4. Summary of previously published papers analyzing the presence of mismatches**
482 **between publicly available RT-PCR primers/probes and SARS-CoV-2 genomes.**

N. of genomes	N. of primers/probes set evaluated	Relevant findings	Source	Period	Reference
17027	27	100% of mutation frequency in the Charité-ORF1b and 18% in the forward primer of CN-CDC-N	GISAID	Genomes sequenced before 7 May 2020	35
992	10	mutations in the first 5' three positions of the China CDC N forward primer, frequency 13%	GISAID	Genomes sequenced before 22 Mrch 2020	36
2569	30	mutations in the first 5' three positions of the China CDC N forward primer, frequency 14%	GISAID	Genomes sequenced before 7 April 2020	37
30	13	mutations in the China CDC N forward primer, frequency 16%	Locally sequenced genomes from Colombia	Period March 6th–24th 2020	38
15001	15	A single mismatch in the Charité group's RdRP gene assay and the Japan NIID's N gene assay; "AAC" variant at the 5' end of the China CDC N forward primer, frequency 18,8%	GISAID	Genomes sequenced before 8 June 2020	39
33819	9	"AAC" variant at the 5' end of the China CDC N forward primer, frequency 24%	GISAID and GenBank	Genomes sequenced before June 2020	32

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

483 **Table 5. Summary of mismatches between publicly available RT-PCR primers/probes and**

484 **SARS-CoV-2 genomes analyzed in this work.**

Source	Primer/probe name	Target gene	Sequence	Lenght	Genomic region*	Mutation	Frequency (%)	Clade Nextstrain	Clade GISAID	Country
ChinaCDC, China	正向引物 (F)	N	GGGGAACTTCT CCTGCTAGAAT	22	28881-28902	G28881A	37,1	20A, 20B	G, GH, GR	Worldwide
						G28882A	36,9	20A, 20B	GH, GR	Worldwide
						G28883C	36,9	20A, 20B	GH, GR	Worldwide
						C28887T	2,9	19A, 20A, 20B, 20C	G, GH, GR, O	SriLanka
ChinaCDC, China	反向引物 (R)	N	CAGACATTTTGC TCTCAAGCTG	22	28958-28979	G28975C	4,6	20A	GH	Europe
USCDC	2019-nCoV_N3-P	ORF9b	ATCACATTGGCACC CGCAATCCTG	24	28704-28727	A28715T	2,0	20A, 20B	GH, GR	Japan
HKU HongKong SAR	HKU-NR	N	CGAAGGTGTGACT TCCATG	19	29236-29254	G29254A	1,0	20A, 20B, 20C	GH, GR	Latvia

485

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2021 doi:10.20944/preprints202101.0473.v1

Figure 1. Graphical representation of SARS-CoV-2 genomic diversity derived from alignment of 46723 genomes obtained from different locations worldwide (from https://macman123.shinyapps.io/ugi-scov2-alignment-screen/.) Vertical bars indicate the number of SNPs found for each genome position.

Figure 2. Subdivision in GISAID clades of genomes included in this work.