SARS-CoV-2 is the new strain of coronavirus identified in late 2019 in the context of the outbreak of pneumonia in China [1, 2, 3]. The virus causes COVID-19, a dangerous human respiratory coronavirus disease. Severe COVID-19 has pneumonia with acute respiratory failure as complications, which explains high mortality.

SARS-CoV-2 is a single-stranded RNA virus that belongs to the beta coronavirus genus, genetically close to SARS [4, 5, 6]. Today, beta coronaviruses OC43, HKU1, SARS, MERS, SARS-CoV-2 and 229E and NL63 alpha coronaviruses are considered to be of clinical importance [2, 5, 7].

The World Health Organization (WHO) declared COVID-19 a public health emergency of international concern (PHEIC); a pandemic was declared in March 2020 [6]. The spread of the disease in the world is regarded as very intense. The number of cases, including fatalities, and the number of affected countries are increasing steadily, and therefore the governments are taking unprecedented measures to prevent spread of the virus. As of June 30, 2020, there were 10360882 SARS-CoV-2 infection cases registered in the world, with 507014 of them ending in fatality. Considering the overall number of cases, Russia is third to USA and Brazil with 646929 confirmed infections and 9306 fatalities as of the mentioned June 30, 2020 [8].

Timely detection of the disease and prevention of its further spread on the territory of the Russian Federation necessitates urgent development of a highly sensitive and specific diagnostic system for detection of SARS-CoV-2 coronavirus RNA in biological samples. The goal of this research effort was to develop such a system.
MATERIALS AND METHODS

The development started on 19.01.2020; at that time, GISAID database contained nucleotide sequences of 8 full-length genomes of the SARS-CoV-2 (formerly CoV Wuhan), which had minor genetic differences: (BetaCoV/Nonthaburi/74/2020[EPI_ISL_403963-crop, BetaCoV/Nonthaburi/61/2020[EPI_ISL_403962, BetaCoV/Wuhan/VDC-HB-01/2019[EPI_ISL_402119, BetaCoV/Wuhan/VDC-HB-04/2020[EPI_ISL_402120, BetaCoV/Wuhan/VDC-HB-05/2019[EPI_ISL_402121, BetaCoV/Wuhan/IPBCAMS-WH-01/2019[EPI_ISL_402123, BetaCoV/Wuhan/WW04/2019[EPI_ISL_402124, BetaCoV/Wuhan-Hu-1/2019[EPI_ISL_402125 and короткий фрагмент BetaCoV/Kanagawa/1/2020[EPI_ISL_402126). We used Mega X software (Clustal W algorithm) to align genome of the new SARS-CoV-2 coronavirus and those of other coronaviruses. To select diagnostic primers and a probe, we identified RdRp, a genome region around the RNA-dependent RNA polymerase of the coronavirus, position 15643-15778 under the MN985325 sequence. At the time of development of the system, this region was conservative to all the known SARS-CoV-2 genomes. Moreover, it differed significantly (nucleotide differences) from the genome sequences of other closely related coronaviruses, including SARS-CoV.

We designed the primers and the probe in conformity with the standard oligonucleotide primers and TaqMan probes selection requirements [9, 10], relying on the Oligo Calc online resources: Oligonucleotide Properties Calculator [11] and OligoAnalyzer Tool [12]. Thermodynamic characteristics of fluorescent probes and their secondary structures were assessed with the help of the mfold Web Server online service [13]. 6-Carboxyfluorescein (ROX) with a black-hole quencher 1 (Hq1) and carboxyfluorescein (FAM) with BHQ1 were used as fluorophores for the probes. AO Genterra synthesized primers and probes.

AmpliTest SARS-CoV-2, the set of SARS-CoV-2 RNA detection reagents under development, covers all stages of testing: virus RNA extraction from samples, reverse transcription and PCR. The disease caused by the new SARS-CoV-2 virus mostly affects the respiratory tract, but in some cases patients suffered disorders in their intestines. Therefore, we examined three types of clinical material: nasopharynx and oropharynx mucosa swabs, sputum and feces.

Control samples were used to assess efficacy of the system at all stages of testing. The internal control sample (ICS) is an artificially synthesized recombinant RNA sequence, about 500 bp in length, enclosed in the ms2 phage envelope [14, 15]. ICS is added at the RNA extraction stage to all samples tested, which allows controlling the success of RNA extraction, reverse transcription and amplification. A positive control sample (PCS) is a recombinant RNA containing the SARS-CoV-2 genome target region measuring ~500 bp, in the ms2 bacteriophage envelope [14, 15]. PCS is introduced as a separate sample at the nucleic acid extraction stage. QX200 Droplet Digital PCR System (Bio-Rad Laboratories, USA) enabled measurement of ICS and PCS concentrations. Cleanliness of ICS and PCS of residual DNA was established through PCR without reverse transcription.

We followed the published clinical guidelines [16] in preparation of the samples of clinical material (smears and sputum). A slightly modified protocol was used in preparation of feces: the clarified extract was obtained through thorough resuspension of 0.1 g (0.1 ml) of the material in 0.9 ml of phosphate buffer, then it was centrifuged for 5 min at 7000 g (MiniSpin, Eppendorf), with subsequent collection of the upper phase. Treatment with guanidine isothiocyanate at 65 °C enabled extraction of nucleic acids, which was followed by the total DNA/RNA precipitation with isopropanol and glycogen as coprecipitation agent. The precipitate was washed to remove impurities and salts and then dissolved in a TE buffer with 0.02 mg/ml of potassium polyadenylate.

Reverse transcription and PCR were performed in one step. The volume of the reaction mixture was 50 μL. It contained the following components: 25 μL of RNA sample, 0.6 mM of each primer (AO Genterra, Russia), 0.3 mM of each probe (AO Genterra, Russia), 0.5 mM of each dNTP (Biosan, Russia), 1 μl of Taq® polymerase (AO Genterra, Russia), 0.5 μl of TM-revertase (MmT) (AO Genterra, Russia), random primers — 0.15 mM (AO Genterra, Russia), polyA — 0.01 mg/ml (AO Genterra, Russia), sodium azide 0.05% (Sigma-Aldrich, USA), tris-HCL buffer (pH 8.3) with 70 mM of tris (oxymethyl)-aminomethane (Sigma-Aldrich, USA), magnesium chloride — no more than 5 mM (Sigma-Aldrich, USA), potassium chloride — no more than 80 mM (Sigma-Aldrich, USA), enzyme stabilizer — no more than 0.2 mg/ml (AO Genterra, Russia), sterile H2O — up to 25 μl.

The format of RT PCR was multiplex, with the ICS fluorescence accumulation signal registered at the FAM fluorophore channel and the fluorescence accumulation signal associated with amplification of the target SARS-CoV-2 nucleic acid at the HEX fluorophore channel.

The amplification program included the following thermal cycling stages: 50 °C — 30 min; 95 °C — 15 min. The following stages were repeated for 45 cycles: 95 °C — 15 s, 60 °C — 30 s, 72 °C — 15 s. The temperature for detection at FAM/HEX fluorophore channels was 60 °C. Overall, the RT PCR process lasted about 2 hours. The result was evaluated with the help of the threshold method: Ct was determined by the intersection of the fluorescence curve and threshold line set in the middle of the fluorescence increase graph’s exponential section (logarithmic scale). The amplification results were interpreted as positive if fluorescence curve crossed threshold line set at the needed level.

The analytical specificity of RT PCR with the selected primers and probe was evaluated in the study of RNA strains of human coronavirus 229E (ATCC® RV-740TM), Betacoronavirus 1 OC43 (ATCC® VR-1558™), influenza A virus (H1N1) (ATCC® VR-1469), influenza A virus (H3N2) (ATCC® VR-776) and influenza B virus (Victoria Lineage) (ATCC® VR-1930) from the ATCC® collection (American Type Culture Collection, USA), HCoV 229E, HCoV OC43, HCoV Nl63, SARS-CoV HKU399489, MERS-CoV (European Virus Archive Global 011N-03868 — Coronavirus RNA specificity panel), DNA of Streptococcus pneumoniae strains (N° 131116), Streptococcus pyogenes (N° 130001), Haemophilus influenzae (N° 151221), Staphylococcus aureus (N° 201108), Klebsiella pneumoniae (N° 180129) from the State Collection of Pathogenic Microorganisms of Scientific Centre for Expert Evaluation of Medicinal Products at the concentration of at least 1x10^{4} genomic equivalents in 1 ml (GE/ml).

Analytical sensitivity (detection threshold) was assessed on model samples of biological material (opornaphy and nasopharynx mucosa swabs, sputum, feces) with the addition of dilutions of the standard sample — protected recombinant RNA containing the target region of SARS-CoV-2 coronavirus genome, in the ms2 bacteriophage envelope. The following dilutions were used: 1x10^{-6}, 5x10^{-6}, 2x10^{-6}, 1x10^{-5}, 5x10^{-5}, and 1x10^{-4} GE/ml. Each dilution was tested with 3 samples of each material, twice. The sensitivity threshold was set based on the minimum dilution detected in three takes.

We evaluated diagnostic indicators while analyzing all types of clinical material (opornaphy and nasopharynx mucosa swabs, sputum, feces) that was previously found to contain
or not contain SARS-CoV-2, as well as the material obtained from healthy people and patients with a different etiology of the disease, which was contaminated with the standard sample at a concentration of at least \(10^3\) GE/ml.

RESULTS AND DISCUSSION

The target we selected to enable SARS-CoV-2 RNA was RdRP, a gene of RNA-dependent RNA polymerase. The figure shows alignment of different sequences of SARS-CoV-2 and other coronaviruses at the site of primer and probe annealing. The nucleotide sequences in this region of SARS-CoV-2 genomes known at the time of development are completely identical; moreover, they have many differences with the genomes of other closely related coronaviruses, which ensures high specificity of the selected primers in terms of amplification of SARS-CoV-2 coronavirus RNA (Figure).

When we started developing this diagnostic system, there was no clinical material available in the Russian Federation (not a single case of COVID-19 was registered at the time). In this connection, we initially synthesized \(~500\) bp of RdRp gene region of the coronavirus genome. The target fragment of the coronavirus was cloned into a plasmid construct that allows obtaining recombinant RNA containing target region of genome of SARS-CoV-2 in the ms2 bacteriophage envelope [14, 15]. This recombinant RNA in the protein envelope served as PCS in the developed diagnostic system, which allowed evaluating effectiveness of all stages of testing.

To optimize the set of reagents part of the AmpliTest SARS-CoV-2 diagnostic system, we used the following PCR diagnostic devices registered in the Russian Federation as medical devices: Rotor-Gene Q (QIAGEN, Germany), CFX96 (Bio-Rad Laboratories, USA), Applied Biosystems QuantStudio 5 (Life Technologies Holdings Pte. Singapore), DTprime (DNK Tekhnologiya, Russia).

We used genetic material of other viruses and bacteria to assess specificity of AmpliTest SARS-CoV-2; these tests returned no cross-reactions, which confirmed 100% analytical specificity of the system. To control analytical sensitivity (detection threshold) of the reagents, we used model samples of biological material contaminated with the standard sample of ms2 recombinant bacteriophage containing a fragment of the SARS-CoV-2 genome. In case of nasopharynx and oropharynx membrane swabs, as well as sputum, the SARS-CoV-2 detection threshold was \(1\times10^3\) GE/ml, that for feces — \(5\times10^4\) GE/ml.

To assess diagnostic sensitivity and specificity, we used a sample of 115 model samples of various biological material (oropharynx and nasopharynx mucosa swabs, sputum, feces) contaminated with the standard sample to a concentration of at least \(10^3\) GE/ml, as well as 195 samples of biological material obtained from healthy people and patients suffering other pathologies. Later, with the spread of coronavirus infection in the Russian Federation, clinical trials were repeated. We examined 150 nasopharynx and oropharynx mucosa swabs, sputum and feces containing SARS-CoV-2 (samples obtained from patients with established COVID-19 infection), as well as 166 samples of biological material (same swabs, sputum, feces) that did not contain SARS-CoV-2 RNA (table 1). Diagnostic indicators (sensitivity and specificity) were at 100% (range from 94.2 to 100%, confidence level of 95%). Thus, we detected no false-positive and false-negative cases when assessing diagnostic sensitivity and specificity on samples from patients.

Coronaviruses are known to easily acquire new mutations [17]. Mutations in the regions of SARS-CoV-2 genome that are complementary to the primers and probe can translate into false-negative results or reduce sensitivity in detecting clinical isolates with nucleotide substitutions. To assess accumulation of mutations in the primer and probe regions, we compared them with the SARS-CoV-2 isolate sequences published in the GISAID database (multiple alignment of 50386 sequences).

Table 1. Results of repeated clinical trials (assessment of diagnostic sensitivity and specificity) of the testing system

| Sample type                  | Total samples tested | Device application results |
|------------------------------|----------------------|----------------------------|
|                              |                      | Positive | Negative |
| Nasopharynx and oropharynx swabs | 113                  | 50       | 0        |
|                              |                      | 0        | 63       |
| Sputum                       | 103                  | 50       | 0        |
|                              |                      | 0        | 53       |
| Feces                        | 100                  | 50       | 0        |
|                              |                      | 0        | 50       |
Table 2. Results of multiple alignment of 50386 known genome sequences of SARS-CoV-2 presented in the GISAID database, selected primers and probe regions. The columns indicate: 1 — sequence number of the nucleotide according to the MN985325 reference sequence, 2–5 identified polymorphisms among the analyzed sequences at this position, 6 — number of SARS-CoV-2 genome sequences in the GISAID database that do not differ at a given MN985325 position, 7 — number of sequences of SARS-CoV-2 GISAID genomes that differ in the given position, 8 — number of sequences of SARS-CoV-2 GISAID genomes that had no nucleotide established at the given position

| MN985325 nucleotide sequence | Polymorphisms | Number of GISAID seqs matching MN985325 | Number of GISAID seqs differing from MN985325 | Number of GISAID seqs for which there is no reliable reading |
|------------------------------|---------------|----------------------------------------|-----------------------------------------------|----------------------------------------------|
|                              | A  T  G  C    |                                        |                                               |                                              |
| 1                            | 2  3  4  5    | 6                                      | 7                                             | 8                                            |

Forward primer region

| MN985325 nucleotide sequence | Polymorphisms | Number of GISAID seqs matching MN985325 | Number of GISAID seqs differing from MN985325 | Number of GISAID seqs for which there is no reliable reading |
|------------------------------|---------------|----------------------------------------|-----------------------------------------------|----------------------------------------------|
|                              | A  T  G  C    |                                        |                                               |                                              |
| 15643                        | 50359         | 0                                      | 50359                                         | 0                                            |
| 15644                        | 50359         | 0                                      | 50359                                         | 0                                            |
| 15645                        | 50359         | 0                                      | 50359                                         | 0                                            |
| 15646                        | 50359         | 0                                      | 50359                                         | 0                                            |
| 15647                        | 50359         | 0                                      | 50359                                         | 0                                            |
| 15648                        | 50359         | 0                                      | 50359                                         | 0                                            |
| 15649                        | 50359         | 0                                      | 50359                                         | 0                                            |
| 15650                        | 50359         | 0                                      | 50359                                         | 0                                            |
| 15651                        | 50359         | 0                                      | 50359                                         | 0                                            |
| 15652                        | 50359         | 0                                      | 50359                                         | 0                                            |
| 15653                        | 50359         | 0                                      | 50359                                         | 0                                            |
| 15654                        | 50359         | 0                                      | 50359                                         | 0                                            |
| 15655                        | 50359         | 0                                      | 50359                                         | 0                                            |
| 15656                        | 50359         | 0                                      | 50359                                         | 0                                            |
| 15657                        | 50359         | 0                                      | 50359                                         | 0                                            |
| 15658                        | 50359         | 0                                      | 50359                                         | 0                                            |
| 15659                        | 50359         | 0                                      | 50359                                         | 0                                            |
| 15660                        | 50359         | 0                                      | 50359                                         | 0                                            |
| 15661                        | 50359         | 0                                      | 50359                                         | 0                                            |
| 15662                        | 50359         | 0                                      | 50359                                         | 0                                            |
| 15663                        | 50359         | 0                                      | 50359                                         | 0                                            |
| 15664                        | 50359         | 0                                      | 50359                                         | 0                                            |
| 15665                        | 50359         | 0                                      | 50359                                         | 0                                            |
| 15666                        | 50359         | 0                                      | 50359                                         | 0                                            |

Probe region

| MN985325 nucleotide sequence | Polymorphisms | Number of GISAID seqs matching MN985325 | Number of GISAID seqs differing from MN985325 | Number of GISAID seqs for which there is no reliable reading |
|------------------------------|---------------|----------------------------------------|-----------------------------------------------|----------------------------------------------|
|                              | A  T  G  C    |                                        |                                               |                                              |
| 1                            | 2  3  4  5    | 6                                      | 7                                             | 8                                            |


using MAFFT algorithm, available in the GISAID database as of June 30, 2020).

Table 2 shows presence of nucleotide polymorphisms in the primer and probe regions according to the alignment data for known SARS-CoV-2 genome sequences (as of June 30, 2020).

For the forward primer, 45 out of the known 50386 sequences were identified to have single polymorphisms (48 nucleotide differences; no more than 0.1% of the total). The substitutions are localized in the central region of the oligonucleotide (mainly one substitution per primer) and are not critical.

As for the reverse primer, there were 80 sequences (out of 50386) identified with one polymorphism in its region, of which 13 sequences have a G/A substitution at the 3’ end (the substitution leads to the formation of C/A effective noncanonical interaction [18]).

There were also 82 sequences with polymorphisms in the region of the probe identified (C/T substitutions).

All the polymorphisms identified belong to SARS-CoV-2 isolates found all over the world, mainly in the USA, Australia, England, the Netherlands, Switzerland, and China. The analysis of 237 known Russian isolates found in GISAID in the regions of forward and reverse primers revealed no nucleotide differences. One local isolate (hCoV-19 / Russia / StPetersburg-RII8955S / 2020 | EPI_ISL_450) has a G/A nucleotide difference in the region of the probe, but this substitution is not critical, since it allows a fairly stable noncanonical C/A interaction of the probe with the matrix [18], and lies close to the 3’ end of the probe. The results obtained indicate there are no nucleotide differences critical to PCR diagnostics in the regions of primers and probes peculiar to all known SARS-CoV-2 isolates (low prevalence of polymorphisms, maximum of one or two substitutions for an isolate, formation of stable noncanonical pairs).

Thus, the analysis of genomes of all known isolates of SARS-CoV-2 revealed current high reliability of the developed AmpliTest SARS-CoV-2 diagnostic PCR system, which detects RNA of coronavirus in the vast majority of cases. However, it should be noted that the high variability of coronavirus genomes suggests the need for constant monitoring of the accumulation of mutations in the primer and probe regions of new isolates, this monitoring allowing timely introduction of changes to the sequence of the oligonucleotides used ensuring high sensitivity of the system.

CONCLUSION

Specialists of Strategic Planning Center, Federal State Budgetary Institution under the Federal Medical-Biological Agency of Russia, developed AmpliTest SARS-CoV-2, a system (including a set of reagents) to detect RNA of SARS-CoV-2, a coronavirus causing severe acute respiratory syndrome (COVID-19). The set makes use of RT PCR in real time and enables control over all stages of testing. When the system
was developed, there were no sets of reagents detecting RNA of the new coronavirus registered in Russia. The system was registered as a medical device on 06.03.2020, registration certificate № РZN 2020/9765; on 30.06.2020 changes were made to the registration documents.

Technical and clinical laboratory investigation, as well as clinical practice of its use for testing purposes, confirmed high analytical and diagnostic sensitivity of the system, which makes it a promising device for timely detection of COVID-19. Today, the system is widely used in the Russian Federation.

References

1. Bogoch I. I. et al. Pneumonia of unknown aetiology in Wuhan, China: potential for international spread via commercial air travel // Journal of travel medicine. 2020. Vol. 27. №. 2. P.1–3.
2. Hul D. S. et al. The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health — The latest 2019 novel coronavirus outbreak in Wuhan, China // International Journal of Infectious Diseases. 2020. Vol. 91. P. 264–266.
3. Rothan H. A., Byrareddy S. N. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak // Journal of autoimmunity. 2020. №. 1–4.
4. Chan J. F. W. et al. Middle East respiratory syndrome coronavirus: another zoonotic betacoronavirus causing SARS-like disease // Clinical microbiology reviews. 2015. Vol. 28. №. 2. P. 465-522.
5. Ellify A. A., Mahdy S. M., Elshemy W. M. Quantitative structure activity relationship and molecular docking revealed a potency of anti-hepatitis C virus drugs against human corona viruses // Journal of medical virology. 2017. Vol. 89. №. 6. P. 1040-1047.
6. Ibrahim I. M. et al. COVID-19 spike-host cell receptor GRP78 binding site prediction // Journal of Infection. 2020. Vol. 80. №. 5. P. 554-562.
7. WHO: Middle East respiratory syndrome coronavirus (MERS-CoV) — The Kingdom of Saudi Arabia Retrieved. 24 February 2020. URL: https://www.who.int/csr/don/24-february-2020-mers-saudi-arabia/en/
8. John Hopkins University. Coronavirus Resource Center. [Электронный ресурс] URL: https://coronavirus.jhu.edu/map.html. Дата обращения: 30.06.2020.
9. Van Pelt-Verkuil E., van Belkum A., Hays J.P. Principles and Technical Aspects of PCR Amplification. — Springer Science & Business Media, 2008.
10. Yuryev A. Methods in Molecular Biology: PCR Primer Design. Totowa, New Jersey: Humana Press, 2007.
11. Kibbe W.A., OligoCalc: an online oligonucleotide properties calculator. Nucleic. Acids Res. 2007. URL: http://biotools.nubic.northwestern.edu/OligoCalc.html.
12. Integrated DNA Technologies. OligoAnalyzer Tool. URL: https://www.idtdna.com/pages/tools/oligoanalyzer.
13. The mfold Web Server (Hosted by The RNA Institute, College of Arts and Sciences). URL: http://unafold.mbi.albany.edu/?q=mfold-DNA-Folding-Form.
14. Cheng Y., Niu J., Zhang Y., Huang J., Li Q. Preparation of his-tagged armored RNA phage particles as a control for real-time reverse transcription-PCR detection of severe acute respiratory syndrome coronavirus // J Clin Microbiol. 2006; 44:3557–62.
15. Pasloske BL, Walkerpeach CR, Obermoeller RD, Winkler M., Du Bois D.S. Armored RNA technology for production of ribonuclease-resistant viral RNA controls and standards // J Clin Microbiol. 1998;36(12):3590–4.
16. Jatsyshina S. B. et al. Laboratory diagnosis of influenza and other acute respiratory viral infections by polymerase chain reaction // Laboratornaya sluzhba. 2017. Vol. 6. №. 3. P. 238–267.
17. Sanchez, C. M., F., Gebauer, C, Sune, A. et al. Genetic evolution and tropism of transmissible gastroenteritis coronaviruses // Virology. – 1992. P. 92–105.
18. Hatim T. Allawi and John SantaLucia, Jr. Nearest-Neighbor Thermodynamics of Internal A:C Mismatches in DNA: Sequence Dependence and pH Effects // Biochemistry. 1998. 37. 9435–9444.