CRISPR technology originated from repetitive nucleic acid sequences discovered in bacterial DNA. Among these sequences viral genome particles were found. These particles are used by CRISPR associated proteins to target the invading viral genome. Cas proteins have the ability of cleaving viral nucleic acids when activated by complementary base pairing of CRISPR RNA with the viral genome. Viruses are a threat to humanity as 65% of them are infectious and fatal. Thanks to CRISPR Cas technology, scientists can detect and eliminate viruses with high precision and accuracy. Among CRISPR Cas systems, scientists have taken the advantage of a natural tool called CRISPR Cas9, a powerful gene editing technology with an immense potential to treat human diseases and has wider applications. Cas 12 and Cas 13 proteins belonging to the Cas family, use their guide RNAs to search for the viral RNA sequence to introduce collateral cleavage of any RNA encountered. The technology is used as a diagnostic tool to detect vi
eral and bacterial pathogens. The CRISPR Cas gene-editing technique now plays a prominent role in diagnostic biology.[6] There are four sub-types of CRISPR RNA combined with Cas 13; namely, Cas 13 (V-A, VI-B, VI-C, and VI-D) and Cas 12 consisted of two subtypes; namely (V-A and V-B). [6] Nucleic acid samples go through amplification, which is an essential requirement for virus detection. The amplified fragment is made accessible by CRISPR-Cas proteins specially designed to get output from collateral cleavage, and its signal is detected by a flow sensing device. DETECTR (DNA Endonuclease Targeted CRISPR Trans Reporter) and SHERLOCK (Specific High Sensitivity Enzymatic Reporter UnLOCKing) are Cas12 and Cas 13-based detection techniques respectively. The natural RNase activity of Cas 13 and Cas 12 proteins is employed to target both RNA and non-specific ssDNA (single-stranded DNA) degradation respectively. The two systems, DETECTR and SHERLOCK generate a messenger signal to detect the presence of the target virus, SARS-CoV-2. The method follows the amplification of target DNA or RNA using RPA (recombinase polymerase amplification) or RT-RPA (reverse transcription recombinase polymerase amplification), respectively.[6] Cas 12a recognizes a 5-TTN-3 PAM site, whereas Cas 13 recognizes a 3 A/G/U PAM site. DETECTR and SHERLOCK systems can work in a similar way by targeting nucleotide-specific sequences and a fluorescent quencher reports the output after the nucleic acids have been degraded. The two detection systems were developed for rapid diagnosis, low cost and high-sensitivity that clearly distinguishing the close members of the coronavirus family, including SARS-CoV, MERS-CoV and others at at-tomolar levels within an hour.[6] The detection of this novel coronavirus, SARS-CoV-2 involves an incubation process in which the extracted nucleic acid is isothermally amplified with the addition of engineering the Cas protein.[6] Designing the sgRNA requires accurate target sites with the lowest off-target potential.[6] In this study, we suggest the in silico designed potential sgRNAs to detect the Spike protein gene of the COVID-19 virus from human nucleic acid samples.

Methods

Phylogenetic Tree

The S-gene sequences of SARS-CoV-2 from different geographical locations were selected from the NCBI server (https://www.ncbi.nlm.nih.gov/), and a phylogenetic tree was created according to the Maximum-Likelihood method using the open-source MEGA 7.0 version.[11] The conserved sequence of the S-gene (Gene ID: 43740568) was selected from the Wuhan reference genome (NC_045512.2) and compared to other closely related species encoding the spike glycoprotein.

Designing of CRISPR-Cas Targets

CHOPCHOP is an easy and straightforward tool to design sgRNA against any gene of interest and is available in both CRISPR and TALEN modes. CHOPCHOP algorithms search for potential off-targets and predict any frame-shift mutation frequencies. The selected conserved sequence of the S-gene was submitted to CHOPCHOP (https://chopchop.cbu.uib.no/). The genomic targets of 20 nucleotide gene sequences preceding 5-TTN-3 PAM were generated.[12] The output sequences were selected according to sequencing by minimum off-targets and were submitted to NCBI-Blast to further filter out the sequence showing the lowest sequence similarity to other viral genomes. We designed target-specific DNA oligos from the selected 20 nucleotide sequence using NEBioCalculator (https://nebiocalculator.neb.com/) and thus synthesized sgRNA. Hence, primer oligos were mainly designed for isothermal amplification of sgRNA followed by CRISPR-Cas12 based detection.

In Silico Vector Design

To design an expression vector construct for its introduction into the mammalian cells, we downloaded the pMBP-LbCas12a H759A vector from Addgene (www.addgene.org) which possesses a T7 promoter and other viral promoters encoding both the sgRNA and the Cas12 protein. We used SnapGene (https://www.snapgene.com) software to construct the expression vector inserted with the selected sgRNA sequence.[13] We used SnapGene software to introduce the 20 nucleotide S1-RBD target sequence into the exon region was generated by CHOPCHOP. Using the “Edit sequence” option, the selected 20-mer was inserted adjacent to the T7 promoter site.

Results

Phylogenetic Tree Analysis

The conserved sequence of the S-gene was identified using the Multiple sequence alignment (MSA) tool, where the sequences of the S1-RBD coding region (22553 to 23093) of SARS-CoV-2 from different geographical locations were compared with other viruses. The phylogenetic tree cluster was the best fit to showing maximum percent similarity under a common human clade SARS-CoV-2 and distantly related to other groups. The S protein sequence of 15 different corona virus isolates such as SARS-COV, Bat coronavirus, murine corona virus, murine hepatitis virus, and bovine coronavirus spike proteins were aligned using the muscle program in MEGA software and the sequence identities obtained for the full-length shown on the S protein diagram. A Maximum-Likelihood tree was generated based on the alignment (Fig. 1). In addition, the S-gene sequence was used to perform CHOPCHOP analysis.
**Target Sequence Identification**

We used CHOPCHOP (https://chopchop.rc.fas.harvard.edu) to identify the protospacer adjacent motif (PAM) sequence from the query RBD gene sequence of SARS-COV-2 in order to design sgRNA. CHOPCHOP generated 110 target sequences based on PAM provided within the exon region of the S-gene that was obtained from the reference genome (Wuhan coronavirus) and selected the 20 nucleotide target sequence (Fig. 2). The CHOPCHOP tool ranked the sequences by many in-built parameters, including GC content and off-targets. Besides, the CHOPCHOP tool provided target-specific primers for S-gene sequence amplification and restriction site analysis. The twenty nucleotide long sequence 5’ACCATACAGAGTAGTAGTAC3’ showed 100% similarity with SARS-CoV-2 and demonstrated <93% similarity with SARS-CoV and other closely related species (data not shown). The target sequence is (highlighted in green; Figure 3), transcription of T7 promoter site sequence (shown in red) and the overlapping sequence was added as the remaining sequence to design guide RNA shown in blue (Fig. 3).

**In silico Expression Vector Design**

Target-specific oligonucleotide sgRNA template was generated with an extra nucleotide ‘G’ added to the 5’ end using NEBiocalculator EnGen sgRNA Template Oligo design-
The sgRNA template was attached to the T7 promoter site for transcription. The NEBioCalculator also provided the overlapping sequence attached to the designed guide RNA (Fig. 3). The generated sequence \textit{in silico} has been used to synthesize the ssDNA oligonucleotide primers to develop the \textit{in vitro} transcription kit (Fig. 4).

**Discussion**

CRISPR genome editing consists of various CRISPR-Cas complex systems with a variety of proteins and enzymes. The computational \textit{in silico} method of designing guide RNA serves as one of the most important steps in targeting the novel coronavirus. The results showed that the conserved sequence was within the region of the RBD domain and the selected target sequence has been used to design CRISPR-Cas detection-based kits for the SARS-CoV-2. Once
the virus is detected by the designed sgRNA construct, the target region would be subjected to cleavage by the Cas 12 enzyme expressed by the pMBP-LbCas12a H759A plasmid construct and other collateral cleavage activity would be recognized using fluorescent quencher. The specificity of the designed guide RNA distinguishes the novel coronavirus from other closely related species, even with a single nucleotide difference. In vivo and in vitro validation methods should be used to test the accuracy of the designed novel coronavirus detection kit using in silico CRISPR-Cas system tools. The currently developed STOPCovid nucleic acid-based test kit could easily be performed by the user using a simple test strip. Currently, there is great competition and significant progress among several Bio-science companies in the development of home-use viral detection test kits. A simple cartridge microfluidic device-based CRISPR-Cas systems may further simplify the currently employed diagnostic method in the near future.

Disclosures
Acknowledgements: The authors acknowledge VFSTR (Deemed to be university) and DST-FIST (LSI-576/2013) for providing the networking facility that was crucial to carry out this work.
Ethics Committee Approval: Not applicable.
Peer-review: Externally peer-reviewed.
Conflict of Interest: There is no conflict of interest.
Authorship Contributions: Concept – A.P.K.; Design – A.P.K., V.T.C., V.S.A.; Supervision – V.T.C., K.S.; Data collection &/or processing – A.P.K., V.T.C., V.S.A.; Analysis and/or interpretation – A.P.K., V.T.C., V.S.A., I.M., K.S.; Literature search – A.P.K., R.R.A., A.G.; Writing – A.P.K., V.T.C., K.S., D.C.N., I.M., V.S.A., R.R.A., Y.N.B., A.G.; Critical review – D.C.N., Y.N.B.

References
1. Ishino Y, Krupovic M, Forterre P. History of CRISPR-Cas from Encounter with a Mysterious Repeated Sequence to Genome Editing Technology. J Bacteriol 2018;200:e00580–17.
2. Nasko DJ, Ferrell BD, Moore RM, Bhavsar JD, Polson SW, Wommack KE. CRISPR Spacers Indicate Preferential Matching of Specific Virolsporltton Genes. mBio 2019;10:e02651–18.
3. Jiang F, Doudna JA. CRISPR-Cas9 Structures and Mechanisms. Annu Rev Biophys 2017;46:505–29.
4. Zhu Y, Klompe SE, Vlot M, van der Oost J, Staals RHJ. Shooting the messenger: RNA-targetting CRISPR-Cas systems. Biosci Rep 2018;38:BSR20170788.
5. Strich JR, Chertow DS. CRISPR-Cas Biology and Its Application to Infectious Diseases. J Clin Microbiol 2019;57:e01307–18.
6. Tang Y, Fu Y. Class 2 CRISPR/Cas: an expanding biotechnology toolbox for and beyond genome editing. Cell Biosci 2018;8:59.
7. kellner MJ, Koob JG, Goottenberg JS, Abudayyeh OO, Zhang F. SHERLOCK: nucleic acid detection with CRISPR nucleases. Nat Protoc 2019;14:2986–3012.
8. Yuen KS, Ye ZW, Fung SY, Chan CP, Jin DY. SARS-CoV-2 and COVID-19: The most important research questions. Cell Biosci 2020;10:40.
9. Udagama B, Kadhiresan P, Kozlowski HN, Malekjahani A, Osborne M, Li YVC, et al. Diagnosing COVID-19: The Disease and Tools for Detection. ACS Nano 2020;14:3822–35.
10. Liu G, Zhang Y, Zhang T. Computational approaches for effective CRISPR guide RNA design and evaluation. Comput Struct Biotechnol J 2019;18:35–44.
11. Jaimes JA, André NM, Chappie JS, Millet JK, Whittaker GR. Phylogenetic Analysis and Structural Modeling of SARS-CoV-2 Spike Protein Reveals an Evolutionary Distinct and Proteolytically Sensitive Activation Loop. J Mol Biol 2020;432:3309–25.
12. Karlapudi AP, T C V, Tammineedi J, Sirama K, Kanumuri L, Prabhakar Kodali V. In silico sgRNA tool design for CRISPR control of quorum sensing in Acinetobacter species. Genes Dis 2018;5:123–9.
13. Abraham Peele K, Srihansa T, Krupanidhi S, Ayyagari VS, Venkateswarulu TC. Design of multi-epitope vaccine candidate against SARS-CoV-2: a in-silico study. J Biomol Struct Dyn. 2020 Jun 1;1-9. doi: 10.1080/07391102.2020.1770127. [Epub ahead of print].