A CRISPR-Cas12a-based specific enhancer for more sensitive detection of SARS-CoV-2 infection

EBioMedicine · 2020 · DOI: 10.1016/j.ebiom.2020.103036

Background: Real-time reverse transcription-PCR (rRT-PCR) has been the most effective and widely implemented diagnostic technology since the beginning of the COVID-19 pandemic. However, fuzzy rRT-PCR readouts with high Ct values are frequently encountered, resulting in uncertainty in diagnosis.

Methods: A Specific Enhancer for PCR-amplified Nucleic Acid (SENA) was developed based on the Cas12a trans-cleavage activity, which is specifically triggered by the rRT-PCR amplicons of the SARS-CoV-2 Orf1ab (O) and N fragments. SENA was first characterized to determine its sensitivity and specificity, using a systematic titration experiment with pure SARS-CoV-2 RNA standards, and was then verified in several hospitals, employing a couple of commercial rRT-PCR kits and testing various clinical specimens under different scenarios.

Findings: The ratio (10 min/5 min) of fluorescence change (FC) with mixed SENA reaction (mix-FCratio) was defined for quantitative analysis of target O and N genes, and the Limit of Detection (LoD) of mix-FCratio with 95% confidence interval was 1.2-1.6-2.1. Totally, 295 clinical specimens were analyzed, among which 21 uncertain rRT-PCR cases as well as 4 false negative and 2 false positive samples were characterized by SENA and further verified by next-generation sequencing (NGS). The cut-off values for mix-FCratio were determined as 1.145 for positive and 1.068 for negative.

Interpretation: SENA increases both the sensitivity and the specificity of rRT-PCR, solving the uncertainty problem in COVID-19 diagnosis and thus providing a simple and low-cost companion diagnosis for combating the pandemic.

Keywords: COVID-19; SARS-CoV-2; rRT-PCR; CRISPR diagnosis; Cas12a; SENA

1. Introduction

Real-time reverse transcription-PCR (rRT-PCR) has been the most effective and widely implemented diagnostic technology since the beginning of the COVID-19 pandemic. However, fuzzy rRT-PCR readouts with high Ct values are frequently encountered, resulting in uncertainty in diagnosis. Since December 2019, the outbreak of COVID-19, caused by the infection of the SARS-CoV-2, has rapidly spread throughout the world, and is now a global pandemic. One of the greatest public health concerns in combating the pandemic is a prompt response to the urgent demand for rapid and accurate diagnosis of the virus. Currently, nucleic acid amplification-based molecular diagnostics (MDx) is the most accurate, fast and affordable and thus the preferred method for diagnosis of SARS-CoV-2 infection. However, since its clinical application, the diagnostic performance of rRT-PCR for SARS-CoV-2 has brought some urgent challenges, particularly the uncertain negative or positive readouts associated with the frequently encountered high Ct-value "grey zones". Besides of "human error" factors such as misconducted sampling, unqualified reagents and uncalibrated diagnostic equipment, inefficient RT reaction and PCR amplification of clinical samples with very low viral loads are likely the major intrinsic causative factors for the fuzzy rRT-PCR readouts and uncertain diagnosis.

With the characterization of non-specific trans-cleavage activities against single-stranded nucleic acids in several CRISPR-associated (Cas) proteins, e.g., Cas12, Cas13 and Cas14, Clustered Regularly Interspaced Short Palindromic Repeats Diagnostics (CRISPR-Dx) technology was established and has been developing rapidly. The underlying mechanism for CRISPR-Dx, as illustrated by the Cas12a-based HOLMES system for example, is based on the efficient trans-cleavage activity against a fluorophore quencher (FQ)-labeled single-stranded DNA reporter by Cas12a triggered upon target DNA recognition, which is guided by a specific CRISPR RNA (crRNA), generating exponentially increasing fluorescence signal within several minutes. With this mechanism, here, we design a Specific Enhancer for detection of PCR-amplified Nucleic Acids (SENA) to improve both the detection sensitivity and specificity against the pre-amplified targeted SARS-CoV-2 genomic fragments. Briefly, the COVID-19 clinical samples are firstly analyzed by the

well-established rRT-PCR assays and amplicons with uncertain readouts are then verified by SENA in a physically isolated space, avoiding contamination of the PCR laboratory during pipetting.

2. Methods

2.1 rRT-PCR reactions

Reactions were conducted in a 25 ul reaction mixture following the instructions offered by the commercial suppliers of the reaction kits. Usually, the reaction cycle parameters were set as reverse transcription at 50 C for 10 min, denaturation at 95 C for 5 min, then followed by 45 cycles of amplification, i.e., 95 C for 10 s and 55 C for 40 s.

2.2 SENA reagents and detection

Candidate crRNA guide sequences with the "TTN" PAM sequences were designed, and the crRNAs were prepared following the procedures previously described. Except the target DNA, the 2x SENA reagent comprises of 2x NEB buffer 3.1, 500 nM LbCas12a (Tolo Biotech.), 1 uM synthesized crRNAs, 1 uM FQ-reporter, and 1 U/ul RNase inhibitor. To avoid the aerosol contamination of the MDx laboratory, after the rRT-PCR reaction, their products must be transferred to a physically isolated room to perform SENA detection. To prepare a 20 ul SENA reaction system, with corresponding positive and negative controls, 2 ul PCR products and 8 ul RNase-free H2O were mixed with 10 ul 2x SENA reagent, and the mixture was then measured on an appropriate fluorescence reader with FAM fluorescence collected following the programs: 48 C 30 s per cycle, 20 cycles.

2.3 Next generation sequencing (NGS)

The rRT-PCR products were purified by AMPure XP beads, and libraries were then constructed following the procedures of end repair, dA-tailing and adaptor ligation. After PCR amplification, samples were sequenced on Illumina Miniseq to produce 2 x 150 bp paired-end reads. After adaptor trimming and quality trimming, the clean reads were mapped to the reference genome of SARS-CoV-2 (MN908947.3) using Bowtie2.

2.4 Systematic titration and regression analyses

The SARS-CoV-2 RNA standards were serially diluted, generating RNA templates ranging from 0.025 to 25 copies per reaction (Rx). To overcome the Poisson sampling ambiguity problem, we performed 9 replicas for groups with 1 and 0.5 RNA template copies/Rx while 6 replicas for each of the rest concentrations. Three readout parameters were compared: the slope, the FC (fold of change of fluorescence), and the FCratio (the ratio between FC of SENA detection at 10 min versus 5 min after the beginning of the fluorescence reading). The FCratio was defined as the standard readout, and mix-SENA (crRNAs targeting both O and N sequences) was chosen as the standard reagent for clinical application.

2.5 SENA detection of clinical samples

Nucleic acids are extracted from the total 295 clinical specimens, such as the pharyngeal swabs and nasopharyngeal swabs of the suspects of SARS-CoV-2 infection, which collected in Ruijin Hospital, Shenzhen Second People's Hospital, and Jinan Center for Disease Control and Prevention, and then subject to rRT-PCR analysis followed by SENA detection and NGS.

3. Results

3.1 Development and characterization of SENA

To prepare appropriate crRNAs for SENA detection, we firstly determined the amplicon sequences from several commercial rRT-PCR kits used in China and then designed specific crRNAs corresponding to each of the distinct amplicons. The apparently most appropriate crRNAs were chosen for the final formulation of the SENA assay reagents. In general, four formulations of reagents were prepared employing crRNAs against Orf1ab (O), E or N genes, individually, and the mixture of O and N (mix).

The performance of SENA was quantitatively characterized via a systematic titration upon rRT-PCR amplicons employing pure SARS-CoV-2 RNA standards. Consistent with the theoretical analysis, along with the decrease of the RNA templates to less than 3 copies/Rx, the rRT-PCR Ct values in some replicas passed 38 but were less than 40, entering the "grey zone". Employing Ct=38 as the cut-off for "positive" detection, we estimated the LoD for O and N genes with 95% confidence interval of this set of rRT-PCR assay as 3.3-4.0-6.1 and 4.0-4.1-4.4, respectively.

The mix-FCratio was demonstrated as the most sure-proof index for rRT-PCR confirmation, and we empirically estimated that mix-FCratio >= 1.145 for positive cut-off, and mix-FCratio <= 1.020 for negative cut-off. As expected, the N-SENA LoD (3.7-4.3-4.8) is very close to that of the N-Ct of rRT-PCR, while the LoD of O-SENA (1.1-1.3-1.7) is significantly lower than that of O-Ct. The LoD of mix-SENA (1.2-1.6-2.1) is slightly higher than that of O-SENA, but it is chosen for clinical applications because of its capability of confirming ambiguous amplicons in the extremely low concentration cases.

3.2 Verification of SENA in different clinical scenarios

SENA was further verified in a few hospitals, testing various clinical specimens under different scenarios. Totally 295 clinical samples collected from 282 individuals were tested by rRT-PCR followed by SENA detection. Samples from 139 patients of Ruijin Hospital were assayed, revealing not only the 12 suspected as negative but also identifying one more positive among the original 123 negative individuals, clearly a case of false negative diagnosis. In Shenzhen Second People's Hospital, SENA correctly resolved uncertain rRT-PCR readouts, and all amplicons were further confirmed by NGS analysis, obtaining consistent results with those of SENA. SENA also assisted in providing evidence of viral clearance for COVID-19 recovering patients. The cut-off value for SENA mix-FCratio remains unchanged as 1.145 for positive diagnosis while slightly increased to 1.068 for negative.

4. Discussion

Instead of developing a closed CRISPR-Dx system, SENA was created to match the commercially available and widely applied rRT-PCR kits, and to solve the uncertainty challenge of the rRT-PCR "grey zone" in COVID-19 diagnosis. SENA specifically increases both the detection sensitivity (with the LoD of 1.6) and the specificity (with false positive and false negative samples detected) in COVID-19 diagnosis. In a SENA detection system, the amplicons of rRT-PCR are taken as the target nucleic acids for the Cas12a-based trans-cleavage reaction. The crRNAs were carefully designed to avoid targeting either the primers or the probes, so the remaining probes will not trigger the trans-cleavage activities of Cas12a. Considering the fact that detection of FAM is supported by most quantitative PCR machines, we simply chose FAM for labelling the SENA FQ-reporter. Since the outbreak of COVID-19 pandemic, dozens of rRT-PCR diagnosis kits that use distinct PCR primer pairs and probes have been successfully developed and are commercially available. The SENA cut-off values were shown to be consistent among the three tested rRT-PCR kits in this work. When a mix-FCratio falls between 1.068 and 1.145, which is the "grey zone" of SENA detection at present, either NGS or clinical symptoms should be used to help diagnosis. Taken together, considering the fact that rRT-PCR is the most popular MDx system and SENA is simple to operate, SENA has the potential to be widely used in various scenarios to solve the uncertainty problems of rRT-PCR as well as other nucleic acid amplification-based MDx.

Data sharing

All data is available in the main text or the Supplementary materials.

Funding sources

J.W. (Jin Wang) is supported by the National Natural Science Foundation of China (31922046); W.H. is supported by the National Key R&D Program of China (2019YFA0906000), NSFC (81772737), and Shenzhen High-level Hospital Construction Fund; G.Z. is supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB19040200).

Declarations of Interests

G.Z. and J.W. are cofounders of Tolo Biotechnology Co., Ltd. W.H., G.Z. and J.W. have filed patent applications relating to the work in this manuscript.