Review

Xiao Ge, Huizi Zhou, Fangyuan Shen, Guimao Yang, Yubo Zhang, Xiaoyu Zhang and Heng Li*

SARS-CoV-2 subgenomic RNA: formation process and rapid molecular diagnostic methods

https://doi.org/10.1515/cclm-2023-0846
Received August 4, 2023; accepted November 13, 2023; published online November 27, 2023

Abstract: Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) which caused coronavirus disease-2019 (COVID-19) is spreading worldwide and posing enormous losses to human health and socio-economic. Due to the limitations of medical and health conditions, it is still a huge challenge to develop appropriate discharge standards for patients with COVID-19 and to use medical resources in a timely and effective manner. Similar to other coronaviruses, SARS-CoV-2 has a very complex discontinuous transcription process to generate subgenomic RNA (sgRNA). Some studies support that sgRNA of SARS-CoV-2 can only exist when the virus is active and is an indicator of virus replication. The results of sgRNA detection in patients can be used to evaluate the condition of hospitalized patients, which is expected to save medical resources, especially personal protective equipment. There have been numerous investigations using different methods, especially molecular methods to detect sgRNA. Here, we introduce the process of SARS-CoV-2 sgRNA formation and the commonly used molecular diagnostic methods to bring a new idea for clinical detection in the future.

Keywords: COVID-19; detection; molecular diagnostic techniques; subgenomic RNA

Introduction

COVID-19 caused by SARS-CoV-2 is spreading globally and has been declared a public health emergency worldwide by the World Health Organization (WHO). Patients with COVID-19 are initially prone to “flu” symptoms with fever and subsequently severe cases leading to respiratory distress, pneumonia, renal failure, and even death [1, 2]. Vaccination can achieve rapid herd immunity, and could effectively reduce the death resulting from SARS-CoV-2 around the world [3]. However, some poor countries or regions still have not been vaccinated. On the other hand, even after vaccination, COVID-19 may still spread to other countries or regions. And, patients who have recovered or vaccinated are still at risk of reinfection. Previous studies indicated that severity of the first infection with COVID-19 was comparable to that of the second infection [4]. So, achieving rapid, accurate, and inexpensive detection of SARS-CoV-2 remains an indispensable component of the containment of COVID-19.

SARS-CoV-2 is an enveloped virus with ~30 kb of coronavirus genome, contains 14 open reading frames (ORFs), and encodes 29 viral proteins [2, 5]. ORF1a and ORF1b are located in the 5’ terminal region of the genome, which account for approximately two-thirds of the total genome length, and encode 16 nonstructural proteins (nsps) [2]. The 3’ untranslated region (3’ UTR) encodes four structural proteins including spike (S), envelope (E), membrane (M), and nucleocapsid (N) proteins participating in the assembly of virions. Besides, there are some genes encoding accessory proteins (f such as ORF3a, ORF3b, ORF6, ORF7a, ORF7b, etc.) [2] (Figure 1A and B). These proteins are translated by newly synthesized viral subgenomic RNA (sgRNA) [8]. The SARS-CoV-2 genome carries out an RNA synthesis process including two parts: replication of genomic RNA (gRNA) and transcription of sgRNA. Replication of the genome is performed for continuous RNA synthesis, while transcription follows the principle that RNA viruses perform discontinuous processes and eventually form a nested set of sgRNA. Viral sgRNA serves as template for the translation of the structural proteins and other nsps.

Although the measurement of SARS-CoV-2 has been widely used clinically using molecular methods, such as reverse transcription-quantitative polymerase chain reaction (RT-qPCR), with primers and probes designed in the core coding region, cannot distinguish the presence of non-
Figure 1: Schematic presentation of the SARS-CoV-2 virion structure, lifecycle, genome organization, subgenome biogenesis, and nine canonical sgRNA.

(A) Virion structure of SARS-CoV-2. SARS-CoV-2 is an enveloped virus. S, E, M, N proteins are incorporated into the virion. N protein is packed inside the virion. S, E, and M are incorporated in the virion membrane (S, blue; E, green; M, pink; N, gold). (B) Genome organization. ORF1a and ORF1b of SARS-CoV-2 encode 16 nsps (nsp1–nsp16). The structural genes encode the structural proteins S, E, M, and N, respectively; and some genes encoding accessory proteins are located among these structural genes. The figure is adapted from previous studies [6, 7]. (C) Lifecycle of SARS-CoV-2. SARS-CoV-2 attaches to the host cell membrane and fuses with it. (I) Viral entry: the virus releases the viral gRNA into the host cell. (II) Primary translation: +gRNA hijacks host ribosomes and produces the two viral replicase polyproteins. 16 nsps are processed and part of them assembled into RTC. (III) (IV) gRNA and sgRNA synthesis: RTC binds to the +gRNA 3′ end to generate full-length −gRNA and a set of −sgRNA which can be used as the templates to transcribe offspring.
infectious residual virions from infectious (replicative) virus particles [10, 11]. Therefore, the results driven by RT-qPCR are not a very accurate assessment of the duration of viral shedding and transmission capacity of SARS-CoV-2 in patients [12, 13]. Isolation and culture of the virus can be a better indication for determining viral replication activity. Yet, it is difficult, labor-intensive, require biosafety level III laboratories. It is not feasible to utilize viral culture as a routine clinical diagnostic method for identifying viable viral infections [13–15]. In recent years, nucleic acid detection methods targeting sgRNA have emerged. For example, Wölfl et al. [16] designed special primers/probe sets to detect E sgRNA in their experiments and used this method to distinguish between actively replicating virus and the viral RNA load. This is owing to the fact that only actively replicating SARS-CoV-2 viruses can initiate RNA synthesis, which leads to the replication of gRNA and transcription of sgRNA [12]. What’s more, sgRNA is not believed to be commonly packaged into new virions [17]. There are numerous studies supporting the utility of sgRNA in routine diagnostics to identify viral replication [18, 19]. Previous studies showed the duration of detection of the sgRNA was similar to that of the culturable virus, and suggested that sgRNA can infer the duration and prevalence of viable virus shedding [11, 14].

In this article, we summarized the general process of SARS-CoV-2 entering host cells to produce genome and subgenomic and introduced common sgRNA detection methods. It may provide a new idea for the rapid diagnosis of SARS-CoV-2 and another possible coronavirus emerging in the future.

Formation process of SARS-CoV-2 sgRNA

SARS-CoV-2 attaches to and fuses with the host cell membrane. After fusion, the viral gRNA is release and began the process of replication and transcription [20–22]. SARS-CoV-2 takes about 6 h to hijack the host transcriptional machinery and transcribe its sgRNA [23]. By hijacking host ribosomes, ORF1a and ORF1b are translated to produce peptides. ORF1a produces the polypeptide pp1a. The frameshift mutation of 1 upstream of the ORF1a stop codon allows the ribosome to continue translating ORF1b and eventually produce the polypeptide pp1ab. As a result, 16 nsps are co- and post translationally processed by pp1a and pp1ab [6, 24]. Most of these nsps are assembled into the replication-translation complex (RTC) with RNA-dependent RNA polymerase (RdRP) activity to drive the replication and transcription of viral RNA [24]. The synthesis process of viral gRNA and sgRNA occurs within double-membrane vesicles (DMVs) and is mediated via RTC consisting of nsp7, nsp8, nsp9, nsp12, nsp13, and short RNA primers [8]. The core of RTC consists of RdRP (nsp12) and three accessory subunits: one nsp7 and two copies of nsp8 [25].

SARS-CoV-2 virus infection of host cells undergoes a complex RNA synthesis process. The positive-strand gRNA (+gRNA) released to act as the template to generate intermediates negative-strand gRNA (−gRNA) and negative-strand sgRNA (−sgRNA). The −gRNA and −sgRNA act as templates to participate in the synthesis process of offspring +gRNA and positive-strand sgRNA (+sgRNA), respectively [6]. Like other coronaviruses, SARS-CoV-2 generates gRNA and nine canonical sgRNA (S, E, M, N, 3a, 6, 7a, 7b, 8). The highly structured gRNA 3′ end of the virus contains binding sites for RTC. The replication process of SARS-CoV-2 gRNA is continuous. After +gRNA released, RTC binds to the +gRNA 3′ end to synthesize the full-length −gRNA intermediates which can transcribe viral +gRNA [8]. Meanwhile, the +gRNA synthesizes a set of −sgRNA which can be templates for newly synthesized +sgRNA to encode structural and accessory proteins of the virus [24]. Normally, each +sgRNA translates a protein from the initial ORF in the +sgRNA, and the data from previous studies show no evidence for ORF10 expression [5, 6, 26]. The nascent +gRNA can be encapsulated by structural protein N and then covered by the viral envelope and eventually released from infected host cells [8] (Figure 1C).

Coronavirus +gRNA is first transcribed into −sgRNA, which are then used as templates to transcribe +sgRNA. +sgRNA is thought to encode structural and nonstructural proteins mentioned above [6, 23]. In the studies of other coronaviruses, the existing transcriptional model is Leader-
to-body fusion, which appears in the minus-strand synthesis of transcriptional regulatory sequences (TRSs) immediately adjacent to ORFs [5]. TRSs contain a conserved 6–7 nt core sequence (CS) surrounded by variable sequences. And the CS of SARS-CoV-2 is ACGAAC [8, 9]. Since the CS is identical for the 5′ Leader of the genome (CS-L) and the viral genome body (CS-B), the complementary sequence of CS-B (cCS-B) can be paired with CS-L. In the process of −sgRNA synthesis, RdRP pauses when passing through TRS upstream of each ORF (TRS-B) in vivo and switches the template into TRS in the precursor (TRS-L). The cCS-B pairs with CS-L, allowing the leader and body to fuse (Figure 1D). The abundance of subgenome is different and N sgRNA is the most abundantly expressed transcript [6, 9]. One reasonable explanation is that the process begins from 3′ end of +gRNA, when RTC initiates transcription form 3′ UTR of the gRNA, it first encounters TRS-B before the ORF of the N gene, switches the template to the TRS-L to generate N sgRNA. If leaky scanning or read-through occurs, the RTC continues scanning to the further TRS-B to produce another kind of sgRNA [8]. This process skips many internal genes, producing a series of nested −sgRNA. According to the fused antisense intermediates above, +sgRNA are transcribed [6, 8, 27]. As a result, a set of sgRNA containing common ~70 nt 5′ common leader sequences, TRS sequence, truncated ORF sequences of different lengths, and 3′ polyadenylated UTR were synthesized [28]. It is notable that ORF10 sgRNA reads were barely detected in previous studies, which does not show obvious homology to any known proteins [5, 6] (Figure 1E). Additional research is required to investigate how ORF10 is translated and what functional role this protein plays.

The formation process of sgRNA is generated through TRS-dependent template switching, however, it has been found that part of sgRNA is formed independently of TRS, resulting in some noncanonical transcripts. Kim et al. [6] found that in addition to the traditional TRS-mediated discontinuous transcription, there were several other atypical minor junction sites that were supported by a follow-up study [26]. There are three major types of such fusion events: the leader combined with the body at unexpected 3′ sites in the middle of ORFs or UTR; a long-distance fusion between sequences that do not have similarity to the leader; local fusion, which leads to smaller deletions. What’s more, it has been described that the sgRNA can give rise to shorter sgRNA, through additional RdRP pause-and-jump events [9]. The molecular mechanism by which SARS-CoV-2 undergoes discontinuous synthesis remains to be explored.

Detection methods of SARS-CoV-2 sgRNA

Presently, there is a lot of research utilizing various molecular methods to examine SARS-CoV-2 sgRNA. The main methods are reverse transcription-quantitative polymerase chain reaction (RT-qPCR), reverse transcription-droplet digital polymerase chain reaction (RT-ddPCR), next-generation sequence (NGS), clustered regularly interspaced short palindromic repeats/associated nuclease systems (CRISPR/Cas systems), and so on (Figure 2). Here, we briefly introduce these main methods for sgRNA detection in previous studies, as well as their benefits and limitations (Table 1).
RT-qPCR

RT-qPCR is a PCR method developed specifically for RNA detection, capable of directly detecting the genomic portion of the virus [29]. It is recognized the current gold standard for the detection of COVID-19 and has been extensively employed for the detection of RNA of SARS-CoV-2 [30]. This method uses an RNA extraction kit to extract RNA from the sample, adds the extract RNA to the mastermix containing the required components for the reaction, and obtains the fluorescence signals for the determination of positive or negative samples.

Most of the kits are designed with primers targeting the coding region of SARS-CoV-2, mainly ORF1ab, N, and E genes [30], which resulted in a test that only indicated the detection of the gRNA of the virus and didn’t determine the replication activity of the virus [10, 11]. Several laboratories have used RT-qPCR to detect the sgRNA which are transcribed only in infected cells and considered to be markers of viral replication [12, 18, 19, 31, 32]. The detection of sgRNA is the same as the detection of gRNA, differing only in the choice of primer positions. Current methods for detecting the SARS-CoV-2 gRNA design primers/probe sets in the body sequence of the gene. Whereas in the process of detecting sgRNA, forward primers are uniquely designed on the leader sequence, reverse primers and probes are designed in the coding region of the gene. When using this unique primers/probe sets for RT-qPCR, gRNA cannot be amplified because the amplicon is too long (about 30 kb). In contrast, sgRNA can be amplified because of their nested structures formed by TRS-mediated template conversion [33, 34] (Figure 1B and E). Therefore, using this special primers/probe sets, only the presence of alive viruses with sgRNA can be detected during amplification. Non-infectious samples that may carry dead virus will not be amplified despite the presence of residual gRNA. In other words, the unique design of the primers/probe sets allows RT-qPCR to detect only actively replicating virus.

There have been several studies utilizing RT-qPCR approaches for subgenome detection. They compared the results of RT-qPCR for sgRNA in clinical samples with the results of viral culture, which is considered the gold standard for active viral replication. Perera et al. [13] carried out RT-qPCR to detect N sgRNA and viral cultures on 33 clinical specimens of various types which had viral loads >5.0 log_{10} virus genome copies/mL. Both tests showed positive results for 12 (36.4 %) specimens, both tests showed negative results for 12 (36.4 %) specimens, positive culture and negative sgRNA results in 2 (6.1 %) specimens, and positive sgRNA and negative culture results in 7 (21.2 %) specimens. Tariq et al. [35] evaluated the interrater agreement between two types of antigen tests and sgRNA detection with viral cell culture technique. The kappa statistic indicated fair agreement between sgRNA detection and cell culture methods (kappa=0.35). A subsequent study compared cell culture results with RT-qPCR results for the detection of sgRNA from 155 clinical samples. The sensitivities of N sgRNA and S sgRNA were equivalent at 100 %, and the specificities of N sgRNA and S sgRNA are 65 and 68 %, respectively [11].

RT-qPCR with high sensitivity, specificity, rapidity, and efficiency plays a crucial role in epidemic prevention and control. Prevent studies have indicated that RT-qPCR also has good sensitivity and specificity in detecting SARS-CoV-2 sgRNA. However, this approach has limitations especially in laboratories in poor countries or regions, such as high prices, complex instruments, and well-trained laboratory personnel. Additionally, some reports indicated that RT-qPCR was less sensitive in diagnosing low viral load samples and was susceptible to contaminants and interferents in the samples [36].

RT-ddPCR

RT-ddPCR is an emerging technology as a third-generation PCR technique that has been used commercially in 2011 [37].
In ddPCR, the PCR reaction is divided into thousands of individual reaction vessels before amplification, and typical PCR cycles are performed in each reaction well [38, 39]. It is a powerful tool to provide more accurate and reproducible data from very low amounts of target molecules even in the contaminated samples [40–43]. Nowadays, ddPCR is applied in different clinical fields and is one of the sensitive and accurate methods to diagnose numerous pathologies, including COVID-19 diagnosis [44–46]. Compared to RT-qPCR, ddPCR technique showed higher sensitivity and specificity in diagnosing COVID-19 in false-negative nasopharyngeal swab samples [39, 45]. Under the same primers/probe sets and template conditions, the sensitivity of ddPCR for SARS-CoV-2 detection is approximately 500 times (maximum) than that of RT-PCR in low-level analyte [47].

Recent studies have demonstrated that the subgenomic transcript abundance of SARS-CoV-2 is low [5, 6, 16]. The combination of low abundance targets and the potential presence of inhibitors during operation and other factors prompted the development of ddPCR techniques to quantify SARS-CoV-2 subgenome [38, 48, 49]. A previously published study [50] applied RT-ddPCR to carry out the detection of SARS-CoV-2 sgRNA. The results indicate that their method was highly specific and reproducible due to low false-positive rates (FPR) and low intra- and inter-assay coefficient of variation (CV%). The limits of detection (LODs) of DNA-template based assay were 2.3 copies/reaction for N sgRNA and 5.2 copies/reaction for S sgRNA. Telwatte et al. [48] designed primer/probe sets for seven typical sgRNA of SARS-CoV-2. The results showed that for all assays, the assay efficiencies were >63 %, the demonstrated linearity was over 4–5 orders of magnitude, and all the assays could detect 1–10 copies of plasmids. Hwang et al. [33] developed a multiplex ddPCR method for the simultaneous quantification of SARS-CoV-2's sgRNA and gRNA in 144 clinical samples. The area under the curve (AUC) was 0.864 and 0.876 for sgRNA and gRNA (p=0.33). To confirm specificity, they performed this method for N sgRNA and gRNA on nasal swabs from five healthy volunteers, and no N sgRNA and gRNA was not detected.

Compared with RT-qPCR, ddPCR provides absolute quantification, tends to tolerate more sequence mismatches in primer/probe sequence, and may be more precise at low copies, while providing similar sensitivity and reproducibility [51–53]. However, ddPCR has a complex instrumentation and high technical requirements for the operators [39]. The majority of current studies using ddPCR to detect SARS-CoV-2 sgRNA are designed to quantify subgenome to investigate the abundance of different types of subgenome and the relationship with total RNA and to study the relationship between transcriptome and clinical disease severity. It requires about 15 % more time and 5–10 % more cost than the RT-qPCR, but its higher sensitivity tips the balance of the cost/benefit scale in favor of ddPCR in some instances. As a result, ddPCR can be a complementary or even alternative tool in identifying COVID-19 positive patients with low viral loads [45].

NGS

NGS can perform massively parallel sequencing with millions of nucleic acid fragments sequenced simultaneously in each run [54]. Furthermore, it presents an unbiased method for detecting pathogens that do not need culture or require clinical assumptions. Most SARS-CoV-2 genome sequencings are done with NGS sequencing technology. Presently, the main purpose of sequencing SARS-CoV-2 utilizing NGS is to understand the genomic structure and related mechanisms of SARS-CoV-2 to provide theoretical basis and data support for subsequent diagnosis as well as treatment. Via this approach, researchers obtained the architecture of SARS-CoV-2 transcriptome that contain the nested sgRNA, ORFs, and TRSs [6, 55, 56].

However, NGS still has certain limitations, such as the requirement of specialized personnel, expensive equipment, a certain level of data storage capacity and security, and lack of adequate research literature and standardized clinical practice, which hinder the possibility of NGS for SARS-CoV-2 in routine clinical testing [57]. Because of the wealth of data provided by NGS, it serves more as a tool to help understand the characteristics of the SARS-CoV-2 genome and transcriptome as well as its physical and chemical properties. In the future, as research continues to deepen and technology continues to develop, NGS may be used as part of the routine detection of pathogens.

CRISPR-Cas systems

At present, CRISPR/Cas systems have already been used not only for genome engineering and RNA editing but for nucleic acid detection of pathogens [58–62]. In the CRISPR-Cas systems, there are various Cas proteins that can be utilized for detection including Cas12a, Cas13a and so on. After form the Cas12a/crRNA/target DNA ternary complex, Cas12a can be activated to carry out cleavage of target DNA (cis-cleavage) and adjacent DNA reporters (trans-cleavage) [63]. In parallel,
Cas13a is an RNA-guide RNase, which can cleavage RNA targets and bystander RNA molecules termed ‘collateral activity’. Utilizing the unique multiple cleavage activities of different Cas enzymes, many laboratories have combined them with isothermal amplification techniques to develop detection tools for different purposes, such as DETECTR and SHERLOCK [59, 64].

Liu et al. [10] developed a detection method (CRISPR-actCoV) based on reverse transcription recombinase polymerase amplification (RT-RPA) combined with a CRISPR Cas12-assisted fluorescence reporter system for the detection of SARS-CoV-2 sgRNA. They performed this approach in 30 clinical specimens, and 21 specimens were detected to be sgRNA positive. This assay was found to be highly sensitive and specific. It was able to detect 10 copies of sgRNA within 35 min, and the results could be observed by the naked eye.

The combined isothermal amplification techniques such as RT-RPA and reverse transcription-loop-mediated isothermal amplification (RT-LAMP) with CRISPR-Cas systems can complete amplification and detection at a constant temperature without the need for temperature cyclers. They can be combined with test strips or handheld devices for subsequent interpretation of the results, which has the advantages of simplicity and convenience, rapid efficiency, and low price, and has great prospects. Yet such methods still have some limitations. For example, isothermal amplification of RNA in the pre-amplification process is prone to aerosol contamination and results in false positives [65, 66]. Although some methods have been developed to detect low copies without pre-amplification, such as combining biosensors, screening for crRNAs, and adding multiple crRNAs. These approaches can also yield promising results, but obviously, the cost raises.

**Significance and limitation of detecting sgRNA of SRAR-CoV-2**

Previous studies have assumed that intermediates required by viruses to express assembly proteins would only be transcribed in actively infected cells, unpackaged into viral particles, and would only be produced during viral replication [6, 67]. This is supported by several subsequent studies. A study of African green monkeys inoculated with infectious SARS-CoV-2 and replication-inactive γ-irradiated SARS-CoV-2 respectively, discovered that gRNA was highly stable, but sgRNA appeared to be rapidly degraded in the absence of replication-active viral infection [68]. Another research found that SARS-CoV-2 and sgRNA were seldom detectable beyond 8 days after the onset of symptoms, yet gRNA could remain detectable by RT-qPCR for several weeks [13]. A subsequent study supported this finding, after comparing RT-qPCR results of gRNA and sgRNA with cell culture results, Kim et al. [11] found that the mean durations of positive culture and sgRNA (11.39 ± 10.34 days and 13.75 ± 11.22 days after symptom onset, respectively) were notable similar. In addition, compared with cell cultures, the detection of sgRNA had comparable sensitivity to gRNA (100 %) and higher specificity than that of gRNA for alive viruses (N, 65 % and S, 68 % of sgRNA, N 23 %, and S, 17 % of gRNA, respectively). All of the above studies have demonstrated that subgenome is better detection targets than genome, enabling the differentiation between actively replicating virus and the viral RNA load.

In contrast, the outcomes of some other research did not corroborate the above conclusions. For example, sgRNA still detects positive results despite negative viral cultures in some studies. Van et al. found that the detection of sgRNA outlasted that of infectious viruses and proposed two explanations, and one is that sgRNA can continue to be detected after replication has ceased [69]. This interpretation is consistent with another study [5]. They proposed sgRNA is stable and is associated with cellular membranes which can protect them just like gRNA from nuclease. In addition, some studies have pointed out that the ratio of sgRNA:gRNA is highly correlated, and sgRNA is not able to provide additional detection value [31, 33]. Moreover, since sgRNA is a series of nested RNAs with various abundances, the time required for different sgRNA to be rendered undetectable is different which makes the selection of the assay target an aspect worth considering [26, 31, 70]. In summary, there are two possible interpretations for this, one is that the viral load is too low to be obtained by culture, but can be obtained by sgRNA, which is more able to demonstrate the high sensitivity of sgRNA. The second is that there is no longer any actively replicating virus in the sample at this time, but the sgRNA is consistently detected because it is protected by the cell membrane and not degraded by nuclease [71].

**Discussion**

WHO states, “The best way to prevent and slow down transmission is to be well informed about SARS-CoV-2, the disease it causes and how it spreads.” Isolation and culture of
the virus is the most accurate method. However, they are labor-intensive, require biosafety level III laboratories, and are difficult to perform on a large scale [13–15]. At present, the gold standard is nucleic acid detection in clinical sets because it has a shorter window period than immunoassays based on antigen-antibody binding, making it easier to make an early diagnosis. But it still has two disadvantages. For one thing, RT-qPCR requires certain technical requirements, and for another, the gRNA of the virus doesn’t indicate the viral replication [10, 11], which prolongs the hospitalization time of patients and causes the waste of medical resources. Because of uncertainty about whether Ct values are representative of infected viruses or simply detect whole gRNA and methodological differences between laboratories, Michalakis et al. [72] did not recommend Ct values as an indicator for determining viral replication.

Viral sgRNA of coronaviruses are produced during viral replication and can be used as markers to assist in determining viral replication instead of secondary biomarkers such as antigens or antibodies, and are gradually being used by more experimenters [10, 11, 14, 16, 29]. The significance of detecting the presence of sgRNA varies in different published articles. For example, Bhosle et al. [49] agreed with the statement that RT-qPCR detection of sgRNA differentiates an actively replicating virus from the viral RNA load. Liu et al. [10] thought that the detection of sgRNAs can distinguish replicable and intact virus from viral RNA. Wölfel et al. [16] supposed that sgRNA can indicate the presence of actively infected cells in samples. A study also pointed out that viable virus can be inferred due to the presence of sgRNA [34]. The detection of sgRNA has been used as a marker to determine the replication of other coronaviruses in previous studies [73]. Many RT-qPCR-based approaches were sensitive and specific and demonstrated a certain concordance with the results of virus culture [11, 13, 14, 16, 26, 70]. It can accurately and rapidly detect sgRNA with low abundance and is rarely contaminated by other components [33, 48]. NGS with its high throughput is widely used for its ability to obtain more information at the maximum level to explore the transcriptional pattern of SARS-CoV-2 [6, 9, 26, 55, 74, 75]. Meanwhile, the approach that RT-RPA combined with the CRISPR-Cas system has been developed for diagnosing sgRNA [10]. It is combined with lateral flow strips so that the results can be read out by the naked eye, which allows it to be easily and quickly accessible and can be carried out on a large scale in poor areas. Researchers can choose various assays according to different research objectives and laboratory conditions.

In conclusion, although the results of some studies do not support the idea that sgRNA can be used as markers for determining the replication of SARS-CoV-2 activity, more studies support their roles as indicators for determining viral replication [5, 11, 69]. Currently, the studies on SARS-CoV-2 sgRNA are focused more on the regulatory features [9], the transcriptional patterns of coronavirus [6, 55, 56], the effects of drug antiviral therapy [49, 76, 77] and so on. On the other hand, there are relatively few studies and few detection methods, and some well-defined commercial kits that have not been applied in the detection of sgRNA for the time being. The reliability of the methods needs to be verified by more subsequent experimental studies. Although COVID-19 is now in a more controllable stage, it still needs to be studied in depth to prevent the recurrence of SARS-CoV-2 on the one hand, and to be able to be used for rapid diagnosis of infectiousness of other coronaviruses that may be new in the future on the other hand. There is still a long way to go.

Research ethics: Not applicable.
Informed consent: Not applicable.
Author contributions: The authors have accepted responsibility for the entire content of this manuscript and approved its submission.
Competing interests: The authors state no conflict of interest.
Research funding: National Natural Science Foundation of China (grant no. 81902170).
Data availability: Not applicable.

References


