Development and Evaluation of Novel Real-Time Reverse Transcription-PCR Assays with Locked Nucleic Acid Probes Targeting Leader Sequences of Human-Pathogenic Coronaviruses


State Key Laboratory of Emerging Infectious Diseases, Department of Microbiology, Research Centre of Infection and Immunology, and Carol Yu Centre for Infection, The University of Hong Kong, and Department of Pathology, Hong Kong Sanatorium and Hospital, Hong Kong Special Administrative Region, China

Based on findings in small RNA-sequencing (Seq) data analysis, we developed highly sensitive and specific real-time reverse transcription (RT)-PCR assays with locked nucleic acid probes targeting the abundantly expressed leader sequences of Middle East respiratory syndrome coronavirus (MERS-CoV) and other human coronaviruses. Analytical and clinical evaluations showed their noninferiority to a commercial multiplex PCR test for the detection of these coronaviruses.

Coronaviruses (CoVs) have repeatedly crossed species barriers, and some have emerged as important human pathogens (1, 2). Human coronavirus (HCoV)-229E, HCoV-OC43, HCoV-NL63, and HCoV-HKU1 predominantly cause mild upper respiratory tract infections, while severe acute respiratory syndrome CoV (SARS-CoV) and Middle East respiratory syndrome CoV (MERS-CoV) frequently cause severe pneumonia with extrapolmonary manifestations (3–6). Highly sensitive and specific laboratory diagnostic tests are essential for the control of emerging CoV outbreaks (7). The gold standard for the laboratory diagnosis of CoV infection is the isolation of infectious virus from the respiratory tract and/or other clinical specimens. However, most CoVs are either difficult or dangerous to culture in cell lines (8, 9). The need for convalescent-phase samples and potential false-positive results due to cross-reactivity with other CoVs limit the use of serum antibody detection assays in the acute setting (10). The overall sensitivity of antigen detection assays is inferior to that of molecular assays such as reverse transcription (RT)-PCR (11, 12). With the increasing availability of molecular diagnostic facilities and expertise in clinical microbiology laboratories worldwide, RT-PCR has become the test of choice for diagnosing CoV infections (7, 13–15).

Traditionally, the preferred targets of RT-PCR assays are genes that are conserved and/or abundantly expressed from the viral genome (16). For CoVs, the most commonly employed targets include the structural nucleocapsid (N) and spike (S) genes, and the nonstructural RNA-dependent RNA polymerase (RdRp) and replicase ORF1a/b genes (4, 7). Recently, other unique noncoding genome regions not present in related CoVs have also been utilized to develop an RT-PCR for the emerging MERS-CoV (7, 13–15). The World Health Organization (WHO) recommends using the upE assay (regions upstream of the envelope [E] gene) for laboratory screening of suspected MERS cases, followed by confirmation with either the ORF1a or ORF1b assay (7). Notably, a number of single nucleotide mismatches at different positions included in the upE assay forward primer and probe have been detected in recent strains of MERS-CoV and may affect the sensitivity of this assay (17). We hypothesize that additional gene targets may be suitable for design of RT-PCR assays for CoVs and would increase the options of molecular diagnosis for circulating and emerging CoV infections. In this study, we designed and evaluated novel real-time RT-PCR assays with locked nucleic acid (LNA) probes for clinically important CoVs based on the identification of the abundantly expressed leader sequence in the 5′-untranslated region (UTR) in small RNA-sequencing (Seq) data analysis.

We included MERS-CoV (strain HCoV-EMC/2012, passage 8, provided by Ron Foucheir, Erasmus Medical Center), HCoV-229E, HCoV-OC43, HCoV-NL63, and HCoV-HKU1 in the study. SARS-CoV was not included as there has not been any human case since 2005. The MERS-CoV isolate was amplified by one additional passage in Vero cells to make working stocks of the virus (5.62 × 10^{5} 50% tissue culture infective doses [TCID_{50}] /ml) as previously described (18). All experimental protocols involving live MERS-CoV followed the approved standard operating procedures of the biosafety level 3 facility at the Department of Microbiology, The University of Hong Kong, as previously described (19). High-titer stocks of HCoV-229E, HCoV-OC43, and other respiratory viruses were prepared, and their TCID_{50} values were determined using standard methods as previously described (20, 21–23). Attempts to culture HCoV-NL63 and HCoV-HKU1 were unsuccessful because of their difficulty in growing in the cell lines available in our laboratories. Virus-positive clinical specimens (n = 14) and laboratory strains (n = 13) were used for evaluating cross-
reactivities with other respiratory viruses in the novel assays were obtained from nasopharyngeal aspirates archived at the clinical microbiology laboratory at Queen Mary Hospital, Hong Kong. Total nucleic acid extractions of clinical specimens and laboratory cell culture with virus strains were performed on 200 μl of sample using an EZ1 virus minikit v2.0 (Qiagen) according to the manufacturer’s instructions. The elution volume was 60 μl. Extracts were stored at −70°C or below until use. Total nucleic acid extracts of ResPlex-II HCoV-positive (n = 49) and -negative (n = 180) respiratory clinical specimens prepared by using the QIAamp MinElute virus spin kit were provided by the Hong Kong Sanatorium and Hospital. A total of 229 fresh or frozen nasopharyngeal aspirates (NPAs) collected between 1 January 2012 and 31 October 2014 from 229 pediatric and adult patients, including 128 males and 101 females, aged 1 to 97 years, who were managed in Queen Mary Hospital and Hong Kong Sanatorium and Hospital for upper and/or lower respiratory tract symptoms were included in the study. The study was approved by the Institutional Review Board of The University of Hong Kong/Hospital Authority Hong Kong West Cluster.

The most abundantly expressed sequence in the MERS-CoV genome was determined by small RNA-Seq data analysis (see the supplemental material). Approximately 2.6% of the trimmed reads could be mapped onto the MERS-CoV genome. Among the mapped sequences, the mapping analysis revealed that most of these small RNA sequence reads, accounting for >6,000 sequences (6.3%), matched the 67-nucleotide leader sequence at the 5′-untranslated region of the genome (Fig. 1). In contrast, the other peaks at the ORF1a, S, and N gene regions accounted for <3.0% of the mapped small RNA sequence reads. Our mapping analysis also showed that the percentages of mapped small RNA sequence reads at the gene regions targeted by the previously described upE, ORF1a, ORF1b, N2, N3, NSeq, and RdRpSeq assays, which had longer sequences than the 67-nucleotide MERS-CoV leader sequence, were only 0.2%, 0.1%, 0.1%, 0.2%, 0.1%, 2.8%, and 0.2%, respectively (13, 14, 24). HCoV-229E, HCoV-OC43, HCoV-NL63, and HCoV-HKU1 similarly possess 70- to 72-nucleotide leader sequences at the same region in their respective genomes (Fig. 1) (25–28). Although leader sequences of around 60 to 90 nucleotides in length are found at the 5′-UTR upstream to the transcription regulatory sequence in the genomes and at the subgenomic RNAs of all CoVs, the function of these leader sequences remains poorly understood (29–31). In view of the abundance of the leader sequences and since infected cells are known to contain large amounts of viral subgenomic RNA at which the leader sequences are abundantly found (32, 33), we hypothesized that the leader sequence might be a valuable diagnostic target not only for MERS-CoV but also for other currently circulating HCoVs (Fig. 1).

To overcome the relatively short length of the leader sequences, we employed LNA probes (Exiqon, Copenhagen, Denmark) to develop novel real-time RT-PCR assays for these human-pathogenic CoVs. LNA is a nucleic acid analogue with an extra bridge connecting the 2′ oxygen and 4′ carbon that has exceptionally high hybridization affinity toward cDNA and RNA and efficient mismatch discrimination (34). These properties are associated with an increased melting temperature of the oligonucleotides, which allows the application of shorter probes when LNA rather than DNA nucleotides are used in the nucleic acid amplification assays (34). In recent years, LNA probes have been increasingly

![Schematic diagram of the MERS-CoV genome, with the leader sequence at the 5′-untranslated region enlarged to illustrate the abundance of the small RNA sequences. The percentages of mapped small RNA sequence reads at the leader sequence, ORF1a, S, and N gene regions are quantified and shown. Leader sequences of 70 to 72 nucleotides in length are also present in other human coronaviruses (HCoV-229E, HCoV-OC43, HCoV-NL63, and HCoV-HKU1).](image-url)
used in the design of real-time PCR assays for other respiratory infections such as those caused by avian influenza A/H5N1 virus, rhinovirus, enteroviruses, respiratory syncytial virus, and *Myco-plasma pneumoniae* (35–38). Using LNA probes, we developed five novel real-time RT-PCR assays (named MERS-CoV-LS, HCoV-229E-LS, HCoV-OC43-LS, HCoV-NL63-LS, and HCoV-HKU1-LS) targeting the short leader sequences of these CoVs (Table 1).

The analytical sensitivities and specificities of the assays were excellent. The limits of detection with *in vitro* RNA transcripts for MERS-CoV-LS, HCoV-229E-LS, and HCoV-OC43-LS were 10 RNA copies/reaction and those for HCoV-NL63-LS and HCoV-HKU1-LS were 5 RNA copies/reaction. Linear amplification was achieved over an 8-log dynamic range, from $10^1$ to $10^8$ RNA copies/reaction for all five assays, with calculated linear correlation coefficients ($R^2$) of 0.99 to 1.00 and amplification efficiency values of 1.93 to 2.27. The limits of detection with viral RNA were approximately $5.62 \times 10^{-2}$ TCID$_{50}$/ml, $5.00 \times 10^{-2}$ TCID$_{50}$/ml, and $3.16 \times 10^{-3}$ TCID$_{50}$/ml for MERS-CoV-LS, HCoV-229E-LS, and HCoV-OC43-LS, respectively (see Tables S1 and S2 in the supplemental material). The limit of detection for the MERS-CoV-LS assay was about 1 log TCID$_{50}$/ml higher than that for the MERS-CoV upE assay in parallel runs and was comparable with those for the other assays currently recommended for screening MERS by the WHO, including the ORF1a, ORF1b, RdRpSeq, and NSseq assays (see Table S1 in the supplemental material) (7, 14). Comparatively, the ORF1b assay for MERS-CoV has the least optimal limit of detection of 64 RNA copies/reaction (13, 14). Our assays showed no cross-reactivity among the individual CoVs and with other common respiratory viruses, including adenovirus, influenza A and B viruses, parainfluenza virus types 1 to 4, rhinovirus, respiratory syncytial virus, and human metapneumovirus (see Table S3 in the supplemental material).

Additionally, we assessed the diagnostic performance of our assays and compared it with that of ResPlex-II in in-use evaluation using 229 NPAs. ResPlex-II is a commercially available multiplex PCR assay which detects 18 respiratory viruses, including HCoV-229E, HCoV-OC43, HCoV-NL63, and HCoV-HKU1 in a single run. It is commonly employed for the laboratory diagnosis of viral respiratory tract infections in many clinical laboratories worldwide (39, 40). Forty-nine NPAs which tested positive for HCoVs by ResPlex-II and another 180 NPAs which tested negative for respiratory viruses by ResPlex-II (Table 2) were tested in an operator-blinded manner. All 49 (100%) ResPlex-II HCoV-positive NPAs tested positive for the corresponding HCoVs in our assays with viral loads of $1.37 \times 10^3$ to $3.86 \times 10^8$ RNA copies/reaction (Table 2) ($P = 1.00$; Fisher’s exact test). Moreover, our assays detected HCoVs in an additional 4/180 (2.2%) NPAs which initially tested negative by ResPlex-II, probably because of the low viral loads of $2.29 \times 10^4$ to $2.40 \times 10^8$ RNA copies/reaction ($P = 0.12$; Fisher’s exact test). Sequencing analysis and two-step confirmatory real-time RT-PCR assays using specific primers targeting the N gene of HCoV-OC43 and HCoV-NL63 (see the supplemental material) confirmed that the results for these four ResPlex-II HCoV-negative specimens concurred with their CoV real-time RT-PCR assay results (two were positive for HCoV-OC43 and two were positive for HCoV-NL63). Overall, these results suggest that our assays are highly sensitive and specific and not inferior to ResPlex-II for the detection of HCoVs *in vitro* and in clinical samples. It is important to note that while ResPlex-II and other multiplex PCR assays have the advantage of being able to detect multiple viruses simultaneously, the sensitivity may be <50% and inferior to that of monoplex PCR assays for HCoVs and other respiratory viruses such as influenza A viruses (39, 41). This relatively poorer sensitivity would especially limit the application of these multiplex PCR assays for the detection of future emerging CoVs and avian influenza A viruses.

### Table 1: Primer and probe sequences of CoV real-time RT-PCR assays with LNA probes in the present study

<table>
<thead>
<tr>
<th>Assay</th>
<th>Genome target</th>
<th>Nucleotide position in genome</th>
<th>Primer/probe</th>
<th>Sequence (5’ to 3’)</th>
<th>GenBank accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MERS-CoV-LS</td>
<td>Leader sequence</td>
<td>14–32 Forward</td>
<td>AGCTTGCCATCTCACCCTTC</td>
<td>JX869059.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>47–69 Reverse</td>
<td>AGTTCGTTAAAATCAAGTCTCG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>34–47 Probe</td>
<td>C+CT+CGT+T+C+CT+TG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HCoV-229E-LS</td>
<td>Leader sequence</td>
<td>20–41 Forward</td>
<td>CTACAGATAGAAGATGTCGTTT</td>
<td>NC_002645.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>57–75 Reverse</td>
<td>ggTCGGTAGTGGAGAAAGT</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>44–59 Probe</td>
<td>AGACT+T+TG+T+CT+A+CT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HCoV-OC43-LS</td>
<td>Leader sequence</td>
<td>17–28 Forward</td>
<td>aaaaCGTGGTCGATC</td>
<td>NC_005147.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>43–66 Reverse</td>
<td>AGATTACAAAAGATCTAAAGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>32–48 Probe</td>
<td>C+TTCA+CTG+ATCT+C+T+TGT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HCoV-NL63-LS</td>
<td>Leader sequence</td>
<td>23–46 Forward</td>
<td>ggAGATAGAGAATTTCATTAGA</td>
<td>NC_005831.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>60–77 Reverse</td>
<td>gTTTGTTAGTTGAGAAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>50–66 Probe</td>
<td>TGTTG+C+TAC+T+C+TTCT+CA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HCoV-HKU1-LS</td>
<td>Leader sequence</td>
<td>21–37 Forward</td>
<td>CGTACCTGTCATAGAGCT</td>
<td>NC_006577.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>48–71 Reverse</td>
<td>GTTTAGATTTAAGATGAGTCTGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>39–52 Probe</td>
<td>ACGA+T+CT+C+T+TG+T+CA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Probes were labeled at the 5' end with the reporter molecule 6-carboxyfluorescein (6-FAM) and at the 3' end with Iowa Black FQ (Integrated DNA Technologies, Inc.). Lowercase letters represent the additional bases added which are not from the original genome sequence. The letters following "a" represent LNA bases which are modified with an extra bridge connecting the 2’ oxygen and 4’ carbon. The bridge "locks" the ribose in the 3’-endo (north) conformations and significantly increases the hybridization properties of the probe.*

*HCoV, human coronavirus; MERS-CoV, Middle East respiratory syndrome coronavirus.*
which are potential pandemic agents that might have a significant public health impact if a case was misdiagnosed.

Our study has demonstrated the previously unknown diagnostic value of the CoV leader sequence and the usefulness of small RNA-Seq data analysis in the selection of optimal gene targets for the development of molecular diagnostic assays. The application of LNA probes allowed the use of relatively short sequences such as the leader sequence of CoV genomes as a diagnostic target in RT-PCR assays. The same approach may be applied to identify and design real-time RT-PCR assays for other emerging viruses, including novel CoVs that are likely to emerge in the future, once their genomic data become available. As for any other gene targets used in RT-PCR assays, particular attention should be paid to the presence of polymorphisms in the leader sequences, which may affect the sensitivity of the assays. The novel CoV real-time RT-PCR assays with LNA probes described in the present study should be further evaluated in large-scale in-field evaluations. The development of these assays into multiplex assays with comparable sensitivity and specificity and additional detection of other novel or reemerging CoVs may further enhance their clinical utility.

ACKNOWLEDGMENTS

We thank Johnson Lau and Manson Fok for their facilitation of the study. This work was supported in part by the donations of Respiratory Viral Research Foundation Limited, Larry Chi-Kin Yung, and the Hui Hoy and Chow Sin Lan Charity Fund Limited and by funding from the Consultancy Service for Enhancing Laboratory Surveillance of Emerging Infectious Disease of the Department of Health, Hong Kong Special Administrative Region, China, the University Development Fund, and the Committee for Research and Conference Grant, The University of Hong Kong.

We declare no conflicts of interest.

REFERENCES


