Nucleotide Sequence Sharing between the Human Genome and Primers for Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Detection

Darja Kanduc

1 Department of Biosciences, Biotechnologies, and Biopharmaceutics, University of Bari, Bari, Italy

Address for correspondence Darja Kanduc, PhD, Department of Biosciences, Biotechnologies, and Biopharmaceutics, University of Bari, Via Orabona 4, Bari 70125, Italy (e-mail: d.kanduc@gmail.com).

Abstract

This study shows that oligonucleotide sequences are shared between the human genome and primers that have been proposed/used for SARS-CoV-2 detection by polymerase chain reaction (PCR). The high level of sharing (namely, up to 19mer with a maximum number of gaps equal to 2) might bear implications for the diagnostic validity of SARS-CoV-2 detection by PCR.

Keywords

► PCR primers
► SARS-CoV-2 detection
► false positives

Introduction

Defining the relationship(s) between infectious agents and the human host is a crucial topic in immunology, microbiology, and infectious medicine. Although it has been proposed that genetic factors might play a role,1,2 the exact mechanisms of chronic infections and occasional (re)activation of pathogens in the human host are largely misunderstood and poorly studied. The issue became even more relevant in light of the recent Ebola virus, Dengue virus, and SARS outbreaks associated with high morbidity and mortality.3–5 In this context, there is a need not only for knowing the molecular basis of infections to define effective and safe preventive and therapeutic interventions but also for sensitive and specific diagnostic tools. Indeed, accurate screening of asymptomatic, presymptomatic, and symptomatic subjects might be key to effective epidemiological measures during pandemics. However, especially in analyzing SARS-CoV-2 as a paradigmatic example, contrasting data have been reported on the analytical performance of SARS-CoV-2 detection methods and claims about the rates of false negatives and false positives have been published.6–11

On the basis of all these, this study focused on the possible genetic basis of potential false polymerase chain reaction (PCR) results by comparing the nucleotide sequence of proposed/used SARS-CoV-2 primers versus the human genome. The scientific rationale is that—given the high level of amino acid sequence sharing between SARS-CoV-2 proteins and the human proteome12–15—parallel sequence matching at the nucleotide level might exist between the SARS-CoV-2 primer sequences and the human genome, in this way possibly explaining the generation of false-positive SARS-CoV-2 detection results. Data are reported here that confirm the likelihood of the research hypothesis.

De facto, using the nucleotide Basic Local Alignment Search Tool (BLASTn) program from NCBI (http://blast.ncbi.nlm.nih.gov,16,17 a sample of 12 primers retrieved from literature,18,19 proposed/used even by government health institutions19 to detect SARS-CoV-2, and described here in ►Table 1, was analyzed for nucleotide sequence sharing with the human genome. BLASTn analyses documented a relevant viral versus human oligonucleotide overlap, with shared primer sequences repeatedly present in the human genome, disseminated among different chromosomes, and located in plus strands, minus strands, mRNAs, pseudogenes, etc. Due to space constraints, an in extenso description of the complete nucleotide sequence sharing is practically not possible, and only a synthetic snapshot is shown in ►Table 2.

In conclusion, this communication highlights the likelihood that viral versus human nucleotide sequence overlap...
### Table 1  Nucleotide sequence of primers used/proposed for PCR detection of SARS-CoV-2<sup>a</sup>

<table>
<thead>
<tr>
<th>Primer no.</th>
<th>Target gene&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Primer direction</th>
<th>Primer nucleotide sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>S 2</td>
<td>F</td>
<td>CCACTAGTCCTCAGTCAGTGAATAT</td>
</tr>
<tr>
<td>2</td>
<td>S 2</td>
<td>R</td>
<td>AAACGGAATCTGAAACTTGTC</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>F</td>
<td>GGAGCTAGAAAATCAGCACCTT</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>R</td>
<td>TCGATGTACTGAAATGCAGTTAG</td>
</tr>
<tr>
<td>5</td>
<td>E</td>
<td>F</td>
<td>ACAGGTACGTATATAGGTTAGCT</td>
</tr>
<tr>
<td>6</td>
<td>E</td>
<td>R</td>
<td>ATATTGCAGCAGTACGACAGA</td>
</tr>
<tr>
<td>7</td>
<td>N</td>
<td>F</td>
<td>GACCCAAAATCAGCGAAT</td>
</tr>
<tr>
<td>8</td>
<td>N</td>
<td>R</td>
<td>TCTGTTACTGCAGTTGAATCTG</td>
</tr>
<tr>
<td>9</td>
<td>N</td>
<td>F</td>
<td>GGGGAACCTCTCCTGGTACTGAAT</td>
</tr>
<tr>
<td>10</td>
<td>N</td>
<td>R</td>
<td>CAGACATTTGGCTCTAAAAGCTG</td>
</tr>
<tr>
<td>11</td>
<td>N</td>
<td>R</td>
<td>TAATCAAGAAGAAGACTGATTA</td>
</tr>
<tr>
<td>12</td>
<td>N</td>
<td>F</td>
<td>TGCCAGCTGTGGTAGTCAAC</td>
</tr>
</tbody>
</table>

Abbreviations: F, forward; PCR, polymerase chain reaction; R, reverse.

<sup>a</sup>Primers retrieved from Gadkar et al<sup>18</sup> and Qasem et al,<sup>19</sup> and further details and references therein.

<sup>b</sup>Gene names given according to Uniprot.<sup>20</sup>

### Table 2  Oligonucleotide sharing between the human genome and polymerase chain reaction (PCR) primers proposed/used to detect SARS-CoV-2: a few examples<sup>4</sup>

1. CCACATGCCTAGTCAGTGAATAT  
   Glypican 5 (GPC5), Chromosome 13, Strand: Plus/Plus  
   864805 TCTAGTCAGTGAATAT  864822

2. AAACGGAATCTGAAACTTGTC  
   DEF domain containing 5, Chromosome 22, Strand: Plus/Minus  
   132374 CTGAGCTAGTAAACTT  132396

3. GGAGCTAGAAAATCAGCACCTTAA  
   DNA damage regulated autophagy modulator 2 (DRAM2), Chromosome 1, Strand: Plus/Plus  
   3702 AGAACATCGACCTTTTAA  3720

4. TCGATGTACTGAAATGGGTATTAG  
   Isolate CM13 chromosome 17, Strand: Plus/Plus  
   5169199 GATGACTGAAAGGCTGATTTA  5169220

5. ACAGGTACGTAGATAGGCTGCT  
   Chromosome 18, SeqID: AF023478.1, Strand: Plus/Minus  
   34259565 GTACGTTAAATGACTGAAATA  34259548

6. ATATTCACGACATGACCACAA  
   Hemicentin 1, HMCN1, Chromosome 1, Strand: Plus/Plus  
   379167 ATATTCACGATACAGCACAG  379185

7. GACCCAAAATCAGCGAAT  
   SLAM family member 8, SLAM8, transcript variant 2, mRNA, SeqID: NM_001330741.2, Strand: Plus/Plus  
   161 CCCCCACATCGCAGAAT  178

8. TCGGGTTACTGCGATTGATGCT  
   Sciatic injury induced RNA upregulator of SOX11, long non-coding RNA, SeqID: NR_026832.1, Strand: Plus/Minus  
   9779 TGGTTACTCCCAAGTGAAT  9761
can interfere with nucleic acid amplification testing and generate PCR false-positive results in SARS-CoV-2 detection, in this way affecting medical diagnoses.

Funding
None.

Conflict of Interest
None declared.

References