Rapid spread of a SARS-CoV-2 Delta variant with a frameshift deletion in ORF7a

Dr. Charles S.P. Foster, PhD; Prof. William D. Rawlinson MBBS, PhD

1School of Medical Sciences, University of New South Wales Sydney, Sydney, NSW, Australia
2Serology and Virology Division, NSW Health Pathology, Prince of Wales Hospital, Sydney, NSW, Australia

Abstract

Australia is currently experiencing COVID-19 outbreaks from infection with SARS-CoV-2 Delta variants (B.1.617.2, AY.3). Analysis of the index case reveals a sub-consensus level of sequencing reads (~25%) that support a 17-nucleotide deletion in ORF7a (ORF7aΔ17del). ORF7aΔ17del induces a frameshift mutation in ORF7a, which truncates the peptide and potentially leads to reduced suppression of host restriction factor BST-2/CD317/Tetherin. Despite this, the mutation has rapidly become represented at the consensus level in subsequent cases: approximately 72% of SARS-CoV-2 genomes in the Australian outbreak possess ORF7aΔ17del, and 99.7% (1534/1538) of Delta genomes on GISAID with ORF7aΔ17del originate from the current Australian outbreak (5 August 2021). The global abundance of this mutation might be underestimated given the difficulty of variant calling software correctly calling insertion/deletions (indels), the common inability of phylogenetics software to take indels into account, and the tendency of GISAID to not release submissions that contain a frameshift mutation (unless specifically requested). Overall, the rapid increase of persistent ORF7aΔ17del variants is concerning, and suggests either a chance founder effect with a neutral mutation yet to be purged, or that the ORF7aΔ17del mutation provides a direct selective advantage.

Main text

SARS-CoV-2 encodes proteins that modulate antiviral responses (ORF3b, ORF6, ORF7a), including reducing type 1 interferon. The ORF7a of SARS-CoV-2 is an ortholog of the corresponding SARS-CoV antagonist of host restriction factor BST-2/CD317/Tetherin that induces apoptosis. Cells with reduced BST2 enhance SARS-CoV-2 replication thereby releasing more virus. Deletion mutations at the SARS-CoV-2 C-terminus of ORF7a are frequent. An 81-nucleotide in-frame deletion in ORF7a removes signal peptide and beta strand sequences, which was detected in March 2020. However, ORF7a deletions are generally transient. This may be due to reduced viral fitness, possibly as a consequence of reduced suppression of host innate immunity.

Australia is currently experiencing COVID-19 outbreaks (>5000 cases) from infection with SARS-CoV-2 Delta variants (B.1.617.2, AY.3). These are characterised by increased frequency of detection of Delta variants with a 17-nt frameshift deletion in ORF7a (ORF7aΔ17del) and minimal other changes in the gene. This deletion spans genome positions 27607–27623 and leads to a truncated (78 vs 121 amino acids) peptide (Fig. 1A,B). The first virus of the current outbreak was collected on June 16, and the consensus genome exhibited a 100% match to a USA sequence with complete ORF7a (lacking ORF7aΔ17del) that was the likely progenitor (Fig. 1A). Notably, the index case of current Australian COVID-19 outbreaks and close contacts possessed a sub-consensus fraction of sequencing reads (~25%) with ORF7aΔ17del. We sequenced the first Delta genome with consensus-level ORF7aΔ17del in Australia on June 18, and ~72% of SARS-CoV-2 genomes in the Australian outbreak possess ORF7aΔ17del (Fig. 1C). This is likely an underestimate as it includes returned travellers and incomplete genome sequences. Analysis of 392,734 whole-genome SARS-CoV-2 Delta variant sequences from GISAID (5 August 2021) showed 99.7% (1534/1538) originate from the current Australian outbreak.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
The occurrence of ORF7a deletions in viruses from multiple countries throughout the pandemic suggests many deletions do not impair fitness, although gene truncation is generally associated with production of proteins with different or no activity. If ORF7a reduces BST2 effects in coronaviruses and BST2 knockdown enhances virion production, the persistence of ORF7aΔ17del suggests it arose by either a chance founder effect or provides a direct selective advantage compensating for loss of BST2 suppression. The latter should be investigated with functional studies of viruses containing ORF7aΔ17del, such as replicon and in vitro organoid research into phenotypic change, as done with S-protein mutations. Given the many Delta variant COVID19 cases worldwide, the possible global spread of ORF7aΔ17del should be closely monitored.

References


Figure 1. Visualisation of a 17-nucleotide deletion in ORF7a of Delta variant SARS-CoV-2 sequences (ORF7aΔ17del) from Australia, 2021, the possible functional consequences at the peptide level, and the increasing frequency in NSW, Australia. Sequences chosen for visualisation include the index case of the current outbreaks (lacking ORF7aΔ17del), the first Australian case possessing ORF7aΔ17del, the only sample assigned to lineage AY.3 that possesses ORF7aΔ17del, the three USA samples (all lacking ORF7aΔ17del) that match the Australian index case, and the only four non-Australian samples to possess ORF7aΔ17del. Accession numbers for all samples analysed for this panel are in the supplementary appendix, available with the full text of this letter online. Panel A shows the sample dates of all chosen samples relative to the index case of the current outbreaks (16 June 2021), and all variants in the sample genomes relative to the SARS-CoV-2 reference genome (created in part using snipit (https://github.com/aineniamh/snippet)). Panel B illustrates the consequences of ORF7aΔ17del on the amino acid translation of ORF7a, leading to a premature stop codon after amino acid 78 and a truncated peptide sequence (~64% complete). Amino acids 1–44 are identical to the SARS-CoV-2 reference genome in all samples and are not presented. Panel C demonstrates the number of genomes of all lineages from Australia from 01 January 2021 until 11 August 2021 with ORF7aΔ17del, ORF7aΔ17del plus other deletion(s), other deletions in ORF7a only, or no deletions in ORF7a. The purple dashed line indicates the beginning of the current outbreaks in Australia, with shading also indicating pre-outbreak and outbreak samples.
### Sample Table

<table>
<thead>
<tr>
<th>Sample</th>
<th>Accession</th>
<th>Collection Date</th>
<th>Day of Outbreak</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia: ORF7aΔ17del</td>
<td>EPI_ISL_3200832</td>
<td>2021-07-31</td>
<td>Day 46</td>
</tr>
<tr>
<td>Australia: ORF7aΔ17del</td>
<td>EPI_ISL_2695795</td>
<td>2021-06-18</td>
<td>Day 3</td>
</tr>
<tr>
<td>Australia: Index</td>
<td>EPI_ISL_2695793</td>
<td>2021-06-16</td>
<td>Day 1</td>
</tr>
<tr>
<td>USA 1</td>
<td>EPI_ISL_2398023</td>
<td>2021-05-20</td>
<td>Day -27</td>
</tr>
<tr>
<td>USA 2</td>
<td>EPI_ISL_2398087</td>
<td>2021-05-13</td>
<td>Day -34</td>
</tr>
<tr>
<td>USA 3</td>
<td>EPI_ISL_2306457</td>
<td>2021-05-13</td>
<td>Day -34</td>
</tr>
<tr>
<td>England: ORF7aΔ17del</td>
<td>EPI_ISL_2551963</td>
<td>2021-06-08</td>
<td>Day -8</td>
</tr>
<tr>
<td>England: ORF7aΔ17del</td>
<td>EPI_ISL_2520482</td>
<td>2021-06-05</td>
<td>Day -11</td>
</tr>
<tr>
<td>Russia: ORF7aΔ17del</td>
<td>EPI_ISL_3122970</td>
<td>2021-06-29</td>
<td>Day 14</td>
</tr>
<tr>
<td>Reference</td>
<td>NC_045512.2</td>
<td>2021-12-26</td>
<td>Day -551</td>
</tr>
</tbody>
</table>

* Pango lineage for sample is AY.3

### Diagrams

#### A

- A color-coded diagram showing the ORF7a position (base) for each sample.
- Sample and accession numbers are indicated.
- The color of the base represents the deletion status:
  - C: No deletion
  - T: ORF7aΔ17del

#### B

- A sequence alignment showing amino acid changes.
- Samples are compared against the reference sequence.
- The alignment includes positions 45 to 120 of the ORF7a gene.
- Deletion types are indicated by color coding:
  - Orange: ORF7aΔ17del
  - Blue: No deletion
  - Yellow: Other deletion
  - Green: Other deletion and ORF7aΔ17del

#### C

- A bar graph showing the number of genomes over time.
- The x-axis represents the time period (Dec 21 to Aug 21).
- The y-axis represents the number of genomes.
- Deletion types are indicated by different colors:
  - Pre-outbreak (gray)
  - Outbreak (red)
  - No Deletion (blue)
  - Other Deletion (yellow)
  - Other Deletion and ORF7aΔ17del (green)

---

CC-BY-NC-ND 4.0 International license

It is made available under a perpetual license that waives the right to make commercial use of this work and its derivatives.

The copyright holder for this preprint is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

medRxiv preprint doi: https://doi.org/10.1101/2021.08.18.21262089; this version posted August 21, 2021. © 2021 The copyright holder for this preprint is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.