The Medical Journal of Australia – **Preprint – 15 May 2020**

Emerging viral mutants in Australia suggest RNA recombination event in the SARS-CoV-2 (COVID-19) genome

Ting-Yu Yeh
Principle Investigator
Institute of Marine and Environmental Technology
Auxergen Inc., Columbus Center
Baltimore, Maryland
United States

Institute of Marine and Environmental Technology
GreenWave Bioscience, Columbus Center
Baltimore, Maryland
United States

Gregory P Contreras
Institute of Marine and Environmental Technology
Auxergen Inc., Columbus Center
Baltimore, Maryland
United States

The Medical Journal of Australia – **Preprint – 15 May 2020**

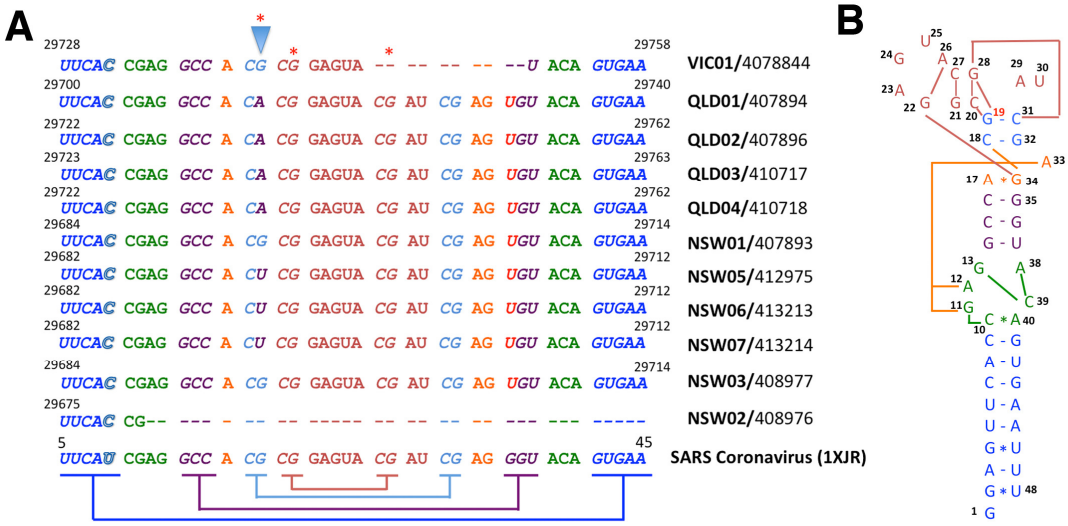
Letter to Editor:

To the Editor: The 2019 novel coronavirus disease (COVID-19) outbreak has become a public health emergency globally.[1,2] Until April 29 2020, there were 6,746 confirmed cases reported in Australia (https://coronavirus.jhu.edu/map.html) However, SARS-CoV-2 specimens independently isolated in Australia (Sydney, Gold Coast, and Melbourne)[3] exhibited very unusual mutations, which have not been identified in other countries (Figure 1A).

Currently 1,319 sequences of the Australian SARS-CoV-2 isolates are available in GISAID.[3] Except for the NSW03 and NSW01 isolates, viral mutations are located at the stem loop-II (s2m) motif, an extremely conserved RNA element in the 3' untranslated region (Figure 1A). The NSW02 and VIC01 isolates have deletion of 41 and 10 nucleotides, respectively. All Queensland cases have single G to A substitution (nucleotide 29714/QLD01, 29736/QLD02 and QLD04, 29737/QLD03). Moreover, patients with NSW05, NSW06, NSW07, NSW15, NSW18, NSW19, NSW21, NSW24, NSW26, NSW28, NSW31 (nucleotide 29696) have single G to U substitution at the same nucleotide. This substitution is only present in Australian patients, not in other SARS-CoV-2 isolates from other countries.

Phylogenetic analysis showed that SARS with 30 other coronaviruses and astroviruses all possess the s2m motif, suggesting that this motif is conserved in both nucleotide sequence and secondary structure folding during evolution in an otherwise rapidly mutable RNA genome.[3,4] The three-dimensional crystal structure of the s2m RNA element of SARS virus shows guanosine 19, which is mutated in Australian isolates, is critical for tertiary contacts to form an RNA base quartet involving two adjacent G–C pairs (G19, C20, G28, and C31)[5] (Figure 1B). Because s2m plays an essential role for the viral RNA to substitute host protein synthesis, we hypothesize the disruption of s2m could alter the viral viability or infectivity dramatically.

The s2m sequence of coronaviruses is highly conserved, and spontaneous mutations in this motif were not expected to have occurred during the apparent short period of time that SARS-CoV-2 has been present, therefore it is highly likely that the changes are due to recombination.[4] Because a high frequency of recombination events in coronaviruses occur, RNA recombination could lead to either enhance the adaptation process to its new host like human or cause unpredictable changes in virulence during infection.



(A) Deletions and mutations in the primary, secondary, and tertiary structures of the COVID-19 s2m RNA motif based on the three-dimensional crystal structure of SARS virus. Conventional RNA helical base-pairing are indicated in italics. Sequence complements are indicated using color-coded brackets. G19 mutation (arrowhead) of Australian SARS-CoV-2 is shown with purple color. Asterisks label the RNA recombination breakpoints based on analysis of 1319 Australia SARS-CoV-2 sequences using Recco algorithm (http:/bioinf.mpi-inf.mpg.de/recco/). The sequence p value < 0.002. (B) Schematic representation of the s2m RNA secondary structure of SARS virus, with tertiary structural interactions indicated as long range contacts.

References

1.Lu R, Zhao X, Li J, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* 2020; 395: 565-574.

2. COVID-19, Australia: Epidemiology report 13: Reporting week ending 23:59 AEST 26 April 2020. https://doi.org/10.33321/cdi.2020.44.35.

3. Global Initiative on Sharing All Influenza Data (GISAID). EpiCoV™ : Pandemic coronavirus causing COVID-19. https://gisaid.org (viewed 29 April, 2020)

4. Tengs T, Jonassen C. Distribution and evolutionary history of the mobile genetic element s2m in coronaviruses. *Diseases* 2016; 4: 27.

5. Robertson MP, Igel H, Baertsch R, et al. The structure of a rigorously conserved RNA element within the SARS virus genome. *PLoS Biol* 2005; 3: e5.