



# Evidence for Recombination as an Evolutionary Mechanism in Coronaviruses: Is SARS-CoV-2 an Exception?

Zisis Kozlakidis\*

*International Agency for Research on Cancer, World Health Organization, Lyon, France*

**Keywords:** COVID-19, recombination, evolution, viral recombination, coronavirus, SARS

## INTRODUCTION

The ability of RNA viruses to exhibit high rates of mutation and replication has been proven for over half a century and has been well documented with the advent of -omics technologies during the last two decades. These high rates of mutation relative to their hosts, allow them to evolve through the genomic evolutionary space, to broaden their variability and in some cases may afford them to acquire advantageous phenotypes in response to environmental pressures, e.g., anti-viral treatments, the latter changes can then become established in the particular evolutionary lineage of the virus (1, 2). Two additional, distinct but not mutually exclusive types of genetic exchange operate in RNA viruses, as a mechanism to acquire advantageous genomic changes, as well to be able to purge accumulated deleterious mutations. These are: firstly re-assortment, for viruses with segmented viral genomes such as Influenza, where antigenic shift in Influenza A is a well documented occurrence (3, 4). The second mechanism is recombination, which can occur both in segmented (5, 6) and non-segmented viruses, when such a mechanism exists; effectively when a “donor” sequence is introduced into a single contiguous genome to produce a new recombinant one. There is much excellent, recent literature summarizing the current knowledge and characterization of recombination for different RNA viruses at a population level (7–10).

Specifically in the Coronavirus family, recombination has been observed previously on a number of genomic studies. For example, recombination was reported in the MERS-CoV species (11, 12), while further phylogenetic analysis of the MERS-CoV full-genome sequences revealed recombination signatures that defined at least five major phylogenetically stable lineages, all of which contained human and camel MERS-CoV sequences (13). Similarly, for SARS-CoV there has been evidence for potential recombination events during its evolution (14, 15), as has also been suggested for human coronavirus HCoV-NL63, the latter exhibiting signs of having arisen from multiple recombination events from its nearest relative over its evolution (16, 17). As such, it is often reported that recombination is a normal consequence of coronavirus replication, required for the generation of the sub-genomic mRNAs and is also implicated in novel strain emergence (18–20).

## THE CASE OF SARS-COV-2

In light of the above, it is of interest to consider the current evidence of recombination observed in the case of SARS-CoV-2. The SARS-CoV-2 virus was hypothesized to have emerged as a result of a recombination event between strains of beta-coronaviruses endemic to certain species of bats and pangolins (21), however this theory has invited intense debate as regards convincingly proving the proximal origin of the virus (22, 23). Specifically for the SARS-CoV-2 origins hypotheses,

## OPEN ACCESS

### Edited by:

Aleksandra Barac,  
University of Belgrade, Serbia

### Reviewed by:

Vladimir Makarek,  
Université du Québec à  
Montréal, Canada  
Paolo Bonilauri,  
Experimental Zooprophyllactic  
Institute of Lombardy and Emilia  
Romagna (IZSLER), Italy

### \*Correspondence:

Zisis Kozlakidis  
kozlakidis@iarc.fr

### Specialty section:

This article was submitted to  
Infectious Diseases - Surveillance,  
Prevention and Treatment,  
a section of the journal  
Frontiers in Public Health

**Received:** 21 January 2022

**Accepted:** 16 February 2022

**Published:** 17 March 2022

### Citation:

Kozlakidis Z (2022) Evidence for  
Recombination as an Evolutionary  
Mechanism in Coronaviruses: Is  
SARS-CoV-2 an Exception?  
Front. Public Health 10:859900.  
doi: 10.3389/fpubh.2022.859900

several authors provided arguments supporting the possibility that the SARS-CoV-2 genome is a chimera of the RaTG13 and Guangdong Pangolin coronavirus (i.e., a virus found in dead Malayan pangolins in the Guangdong province of China) (24–26) or in the place of the latter of close relatives of the bat CoV ZC45 and ZXC21 strains (24). Similarly, according to current hypotheses, evidence was presented that SARS-CoV-2 might be the result of recombination into RaTG13 from some unknown CoV strains (27). Such recombination events remain likely hypotheses at present, especially as in previous outbreaks intermediate hosts were implicated in the  $\beta$ -CoV transmission (e.g., civets for SARS-CoV and camels for MERS-CoV) (20), suggesting that SARS-CoV-2 may have co-circulated with other coronaviruses in the wild in the same intermediary hosts, and also may have been transmitted to humans in this way.

To date, the SARS-CoV-2 genetic diversity increases slowly compared to other RNA viruses: given the many millions of infections globally and hundreds of thousands of genomes deposited in public databases (e.g., in the GISAID database) (28), there are only 7–8 major circulating clades observed, being identified based on multiple variants common to large numbers of isolates. It is this relative genomic stability of the circulating viral forms that allowed for the rapid development of effective vaccines and therapeutics, as well as supporting the deciphering of the SARS-CoV-2 pathology. However, inter- and intra-host recombination events in coronaviruses are well studied and evidenced to occur frequently (29, 30). As such the question arises on the lack of recombination events reported for circulating SARS-CoV-2 viruses. There have been a limited number of publications reporting any such recombination events (31–34).

It is becoming evident that while homologous recombination exists, recombinants seem to circulate at low levels for SARS-CoV-2 (31, 35, 36) with current estimates that at most 5% of circulating strains in the United Kingdom and USA are recombinants (36), or 16 recombinant sequences from the whole UK dataset of 279,000 sequences up to March 7, 2021 (31). On the other hand, it is also technically challenging to demonstrate homologous recombination when the genomic lineage evolution is driven by a limited number of single nucleotide polymorphisms. Furthermore, in order for homologous recombination to occur, the same cells within an individual need to be co-infected by genetically distinct viruses. Such co-infection of an individual requires that multiple viral lineages co-circulate within a population and, given the short duration of most SARS-CoV-2 infections, is most likely to be observed when virus prevalence is high in the population. Thus, the potential window of opportunity for the currently circulating SARS-CoV-2 variants is of limited time. To date, no heterologous recombination events have been reported, for example between SARS-CoV-2 and other co-circulating seasonal coronaviruses.

Having said that, the existing literature has demonstrated that the coronavirus proofreading exoribonuclease (nsp14-ExoN) is required to maintain the rates and loci of recombination generated during infection, and strongly supports that recombination mechanisms have been conserved across different evolutionary trajectories and host species specificity (18). Specifically, group 2a (MHV), 2b (SARS-CoV-2), and 2c

(MERS-CoV)  $\beta$ -CoVs demonstrated many strong similarities in their patterns of recombination junctions across the genomes and in the types of recombined RNAs produced (18). Furthermore, during mixed infections of cell cultures with murine coronaviruses, at least 10% of progeny viruses were recombinants showing multiple independent recombination breakpoints (37). While such events appear unbiased in culture, in nature, events of recombination will be guided by natural selection pressures in regions with roles in host interactions. Among coronaviruses such areas of interest are centered in spike proteins (38). In light of the above, a particular case in SARS-CoV-2 can be hypothesized, with an upper ceiling of homologous recombination frequency (as evidenced by the experiments in culture) and potential recombination hotspots (spike protein) where the identification of such events would be most likely.

## DISCUSSION

The coronavirus family is inclusive of many members, and the relative levels of recombination might be very different between different viruses even of the same family. In particular, the exact molecular mechanisms and determinants of RNA recombination in coronaviruses are only now becoming understood in greater detail, through the scaled-up surveillance and whole genome sequencing analyses (39), though the exact mechanisms and determinants of CoV recombination are not known (18). Additionally, for some outbreaks, there is little reason to suspect recombination, e.g., negative sense single-stranded RNA viruses are thought to recombine over evolutionary, not population-level, time scales (40). As more mutations and lineages of SARS-CoV-2 get fixed in the population and sequenced over a number of consecutive SARS-CoV-2 waves, a recombination event caused by a co-infection of a single patient with genetically distinct clades may lead to emergence of novel lineages, posing risks to the efficacy of future treatments.

Therefore, the following actions need to be considered: (i) a rapid and consistent surveillance of the sequenced SARS-CoV-2 genomes both for novel mutations and recombinations; (ii) a unified collection of genomic, epidemiological and clinical data; and (iii) further developed bioinformatics pipelines that allow for such recombination events to be detected within the limitations of the SARS-CoV-2 low genomic variation. While the first point is largely in place, the latter two points vary greatly between different geographic locations between and within countries. The UK presents a useful example in this respect, as the high rate of genomic surveillance and unified collection of genomic, epidemiological, and geographic data provide multiple lines of evidence for evaluating the identification of recombinant viruses. Establishing and operating such an integrated approach to viral surveillance on a consistent basis, remains critical to the ongoing identification of recombinants.

## AUTHOR CONTRIBUTIONS

ZK conceived and wrote the Opinion manuscript.

## REFERENCES

1. de Béthune MP. Non-nucleoside reverse transcriptase inhibitors (NNRTIs), their discovery, development, and use in the treatment of HIV-1 infection: a review of the last 20 years (1989–2009). *Antiviral Res.* (2010) 85:75–90. doi: 10.1016/j.antiviral.2009.09.008
2. Johnson VA, Calvez V, Günthard HF, Paredes R, Pillay D, Shafer RW, et al. Update of the drug resistance mutations in HIV-1: March 2013. *Topics Antiviral Med.* (2013) 21:6.
3. Steel J, Lowen AC. Influenza A virus reassortment. *Influenza Pathogenesis and Control-Vol.* (2014) 14:377–401. doi: 10.1007/82\_2014\_395
4. Lowen AC. Constraints, drivers, and implications of influenza A virus reassortment. *Ann Rev Virol.* (2017) 4:105–21. doi: 10.1146/annurev-virology-101416-041726
5. Bonnet J, Fraile A, Sacristán S, Malpica JM, García-Arenal F. Role of recombination in the evolution of natural populations of Cucumber mosaic virus, a tripartite RNA plant virus. *Virology.* (2005) 332:359–68. doi: 10.1016/j.virol.2004.11.017
6. Varsani A, Lefevre P, Roumagnac P, Martin D. Notes on recombination and reassortment in multipartite/segmented viruses. *Curr Opin Virol.* (2018) 33:156–66. doi: 10.1016/j.coviro.2018.08.013
7. González-Candelas F, López-Labrador FX, Bracho MA. Recombination in hepatitis C virus. *Viruses.* (2011) 3:2006–24. doi: 10.3390/v3102006
8. Yebra G, Frampton D, Gallo Cassarino T, Raffle J, Hubb J, Ferns RB, et al. A high HIV-1 strain variability in London, UK, revealed by full-genome analysis: Results from the ICONIC project. *Plos one.* (2018) 13:e0192081. doi: 10.1371/journal.pone.0192081
9. Bbosa N, Kaleebu P, Ssemwanga D. HIV subtype diversity worldwide. *Curr Opin HIV and AIDS.* (2019) 14:153–60. doi: 10.1097/COH.0000000000000534
10. Bentley K, Evans DJ. Mechanisms and consequences of positive-strand RNA virus recombination. *J General Virol.* (2018) 99:1345–56. doi: 10.1099/jgv.0.001142
11. Corman VM, Ithete NL, Richards LR, Schoeman MC, Preiser W, Drosten C, et al. Rooting the phylogenetic tree of middle East respiratory syndrome coronavirus by characterization of a conspecific virus from an African bat. *J Virol.* (2014) 88:11297–303. doi: 10.1128/JVI.01498-14
12. Dudas G, Rambaut A. MERS-CoV recombination: implications about the reservoir and potential for adaptation. *Virus Evol.* (2016) 2:1. doi: 10.1093/ve/vev023
13. Sabir JS, Lam TT, Ahmed MM, Li L, Shen Y, EM Abo-Aba S, et al. Co-circulation of three camel coronavirus species and recombination of MERS-CoVs in Saudi Arabia. *Science.* (2016) 351:81–4. doi: 10.1126/science.aac8608
14. Holmes EC, Rambaut A. Viral evolution and the emergence of SARS coronavirus. *Philosophic Transact Royal Soc London. Series B: Biologic Sci.* (2004) 359:1059–65. doi: 10.1098/rstb.2004.1478
15. Zhang XW, Yap YL, Danchin A. Testing the hypothesis of a recombinant origin of the SARS-associated coronavirus. *Archiv Virol.* (2005) 150:1–20. doi: 10.1007/s00705-004-0413-9
16. Pyrc K, Dijkman R, Deng L, Jebbink ME, Ross HA, Berkhout B, et al. Mosaic structure of human coronavirus NL63, one thousand years of evolution. *J Mol Biol.* (2006) 364:964–73. doi: 10.1016/j.jmb.2006.09.074
17. Abdul-Rasool S, Fielding BC. Understanding human coronavirus HCoV-NL63. *Open Virol J.* (2010) 4:76. doi: 10.2174/1874357901004010076
18. Gribble J, Stevens LJ, Agostini ML, Anderson-Daniels J, Chappell JD, Lu X, et al. The coronavirus proofreading exoribonuclease mediates extensive viral recombination. *PLoS Pathog.* (2021) 17:e1009226. doi: 10.1371/journal.ppat.1009226
19. Lai MM, Cavanagh D. The molecular biology of coronaviruses. *Adv Virus Res.* (1997) 48:1–00. doi: 10.1016/S0065-3527(08)60286-9
20. Graham RL, Baric RS. Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission. *J Virol.* (2010) 84:3134–46. doi: 10.1128/JVI.01394-09
21. Zhang T, Wu Q, Zhang Z. Probable pangolin origin of SARS-CoV-2 associated with the COVID-19 outbreak. *Curr Biol.* (2020) 30:1346–51. doi: 10.1016/j.cub.2020.03.022
22. Seyran M, Pizzol D, Adadi P, El-Aziz TM, Hassan SS, Soares A, et al. Questions concerning the proximal origin of SARS-CoV-2. *J Med Virol.* (2021) 93:1204. doi: 10.1002/jmv.26478
23. Zhu Z, Meng K, Meng G. Genomic recombination events may reveal the evolution of coronavirus and the origin of SARS-CoV-2. *Scientific Rep.* (2020) 10:1–0. doi: 10.1038/s41598-020-78703-6
24. Makarenkov V, Mazouze B, Rabusseau G, Legendre P. Horizontal gene transfer and recombination analysis of SARS-CoV-2 genes helps discover its close relatives and shed light on its origin. *BMC Ecol Evol.* (2021) 21:1–8. doi: 10.1186/s12862-020-01732-2
25. Li X, Giorgi EE, Marichannelowda MH, Foley B, Xiao C, Kong XP, et al. Emergence of SARS-CoV-2 through recombination and strong purifying selection. *Sci Adv.* (2020) 6:eabb9153. doi: 10.1126/sciadv.abb9153
26. Xiao K, Zhai J, Feng Y, Zhou N, Zhang X, Zou JJ, et al. Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins. *Nature.* (2020) 583:286–9. doi: 10.1038/s41586-020-2313-x
27. Boni MF, Lemey P, Jiang X, Lam TT, Perry BW, Castoe TA, et al. Evolutionary origins of the SARS-CoV-2 sarbecovirus lineage responsible for the COVID-19 pandemic. *Nature microbiology.* (2020) 5:1408–17. doi: 10.1038/s41564-020-0771-4
28. Elbe S, Buckland-Merrett G. Data, disease and diplomacy: GISAID's innovative contribution to global health. *Glob Challe.* (2017) 1:33–46. doi: 10.1002/gch2.1018
29. Amoutzias GD, Nikolaidis M, Tryfonopoulou E, Chlichlia K, Markoulatos P, Oliver SG. The remarkable evolutionary plasticity of coronaviruses by mutation and recombination: insights for the COVID-19 pandemic and the future evolutionary paths of SARS-CoV-2. *Viruses.* (2022) 14:78. doi: 10.3390/v14010078
30. Su S, Wong G, Shi W, Liu J, Lai AC, Zhou J, et al. Epidemiology, genetic recombination, and pathogenesis of coronaviruses. *Trends Microbiol.* (2016) 24:490–502. doi: 10.1016/j.tim.2016.03.003
31. Jackson B, Boni MF, Bull MJ, Collier A, Colquhoun RM, Darby AC, et al. Generation and transmission of interlineage recombinants in the SARS-CoV-2 pandemic. *Cell.* (2021) 184:5179–88. doi: 10.1016/j.cell.2021.08.014
32. Varabyou A, Pockrandt C, Salzberg SL, Pertea M. Rapid detection of inter-clade recombination in SARS-CoV-2 with Bolotie. *Genetics.* (2021) 218:iyab074. doi: 10.1093/genetics/iyab074
33. Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, et al. Spike mutation pipeline reveals the emergence of a more transmissible form of SARS-CoV-2. *BioRxiv.* (2020). doi: 10.1101/2020.04.29.069054
34. Ignatieva A, Hein J, Jenkins PA. Evidence of ongoing recombination in SARS-CoV-2 through genealogical reconstruction. *BioRxiv.* (2021) 21:615. doi: 10.1101/2021.01.21.427579
35. Nie Q, Li X, Chen W, Liu D, Chen Y, Li H, et al. Phylogenetic and phylodynamic analyses of SARS-CoV-2. *Virus Res.* (2020) 287:198098. doi: 10.1016/j.virusres.2020.198098
36. VanInsberghe D, Neish AS, Lowen AC, Koelle K. Recombinant SARS-CoV-2 genomes are currently circulating at low levels. *bioRxiv.* (2021) 1:2002–08. doi: 10.1101/2020.08.05.238386
37. Makino S, Keck JG, Stohman SA, Lai MMC. High-frequency RNA recombination of murine coronaviruses. *J Virol.* (1986) 57:729–37. doi: 10.1128/jvi.57.3.729-737.1986
38. Bobay LM, O'Donnell AC, Ochman H. Recombination events are concentrated in the spike protein region of Betacoronaviruses. *PLoS Genetics.* (2020) 16:e1009272. doi: 10.1371/journal.pgen.1009272
39. Pollett S, Conte MA, Sanborn M, Jarman RG, Lidl GM, Modjarrad K, et al. A comparative recombination analysis of human coronaviruses and implications for the SARS-CoV-2 pandemic. *Scien Rep.* (2021) 11:1. doi: 10.1038/s41598-021-96626-8
40. Chare ER, Gould EA, Holmes EC. Phylogenetic analysis reveals a low rate of homologous recombination in negative-sense RNA viruses. *J General Virol.* (2003) 84:2691–703. doi: 10.1099/vir.0.19277-0

\* References that are pre-prints, have been denoted within the text in italics.

**Author Disclaimer:** Where authors are identified as personnel of the International Agency for Research on Cancer/WHO, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer/WHO.

**Conflict of Interest:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may

be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2022 Kozlakidis. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*