Whole genome stability of SARS-CoV-2: abridged secondary publication

MH Wang *, MKC Chong, ELY Wong, PKS Chan, Z Chen, CKC Lai, RCW Chan

KEY MESSAGES

1. Understanding virus evolution is important for pandemic preparedness and response. Characterisation of genome stability of SARS-CoV-2 enables evaluation of the virus's evolutionary rate and the potential risk of COVID-19 vaccine breakthrough.
2. The receptor-binding domain region of the spike protein was associated with a significant risk of COVID-19 vaccine breakthrough.

Hong Kong Med J 2025;31(Suppl 4):S22-3
HMRF project number: COVID190103

[1] **MH Wang,** [1] **MKC Chong,** [1] **ELY Wong,** [2] **PKS Chan,** [2] **Z Chen,** [3] **CKC Lai,** [3] **RCW Chan**

[1] *JC School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong SAR, China*
[2] *Department of Microbiology, The Chinese University of Hong Kong, Hong Kong SAR, China*
[3] *Department of Health, Hong Kong SAR, China*

* Principal applicant and corresponding author: maggiew@cuhk.edu.hk

Introduction

The continuous evolution of SARS-CoV-2 has generated successive epidemic waves, posing a major threat to COVID-19 prevention and control.[1] Evaluations of viral evolutionary traits help predict future trajectories and support assessment of the health burden and disease risks posed by emerging variants.[2] The substitution rate, or evolutionary rate, is a key metric for understanding the evolutionary characteristics of a species.[3] At the genomic level, the substitution rate of SARS-CoV-2 aggregates information across its constituent proteins.[4] However, various genes and genomic regions are subject to different degrees of immune pressure. For example, surface proteins or epitope sites of the influenza virus tend to have more mutations that lead to antigenic escape or positive selection.[5] Therefore, analyses of segment-specific substitution rates may better reveal the evolutionary characteristics of SARS-CoV-2. Such information is valuable for monitoring and assessing the risks posed by circulating variants. This study aimed to evaluate substitution rates across specific genome segments of SARS-CoV-2.

Methods

The dataset used included 73 062 SARS-CoV-2 sequences retrieved from the Global Initiative on Sharing All Influenza Data[6] and 110 viral genomes sequenced at the Prince of Wales Hospital, Hong Kong SAR between March 2020 and March 2022. Substitution rates were estimated using BEAST v1.10 for Bayesian phylogenetic analysis.[7]

We developed a method to evaluate the effect of mutations in major immunodominant regions on the reduction of vaccine effectiveness at the population level.[8] Cumulative genetic distance was calculated between the antigen sequences of the COVID-19 of circulating viruses. Real-world vaccine effectiveness data from observational studies were collected. A model was then constructed to capture the relationship between genetic mutations and vaccine protection.[8]

Results

During the early phase of the pandemic, cumulative mutation activity across the SARS-CoV-2 genome was profiled over time. Increasing trends in mutation activity were observed, particularly in the S protein, ORF1b, and the N protein.[9] The basic framework for profiling mutation activity was based on average genetic distance from the ancestral Wuhan strain.

The estimated genomic-level substitution rate was 0.86 (95% confidence interval [CI]=0.76-0.96) $\times 10^{-3}$ substitutions per nucleotide site per year (s/n/y), consistent with values reported in the literature.[4,10] The highest substitution rates were observed in the N and ORF8 proteins at 2.27 and 2.19 $\times 10^{-3}$ s/n/y, respectively—approximately 2.5-fold higher than the genomic average, followed by the S protein at 1.64 $\times 10^{-3}$ s/n/y—approximately 1.9-fold higher than the genomic average, and then the E and ORF10 proteins at 0.17 and 0.35 $\times 10^{-3}$ s/n/y, respectively.

Based on the vaccine effectiveness–genetic distance model, as of March 2022, a single mutation in the receptor-binding domain region corresponded to an absolute reduction in vaccine effectiveness of 5.2% (95% CI=2.4-8.0) for mRNA vaccines and 15.8% (95% CI=12.4-19.3) for inactivated vaccines.

Discussion

Genome stability of SARS-CoV-2 was characterised on a gene segment basis. The N, ORF8, and S proteins were identified to be the fastest-evolving among all evaluated genomic segments. The receptor-binding domain region of the spike protein was associated with a significant risk of COVID-19 vaccine breakthrough. All results presented should be interpreted with regard to the time of data collection.

Conclusions

Substitution rates in the SARS-CoV-2 genome vary. Conserved genomic segments containing epitopes may serve as potential vaccine antigen candidates. Regular reporting on viral genome stability may assist health authorities in understanding local viral genetic diversity and transmission patterns. Regular estimation of vaccine protection levels for both COVID-19 and influenza may support timely issuance of health alerts and allocation of hospital resources.

Funding

This study was supported by the Health and Medical Research Fund Commissioned Research on the Novel Coronavirus Disease, Health Bureau, Hong Kong SAR Government (#COVID190103). The full report is available from the Health and Medical Research Fund website (https://rfs1.healthbureau.gov.hk).

Disclosure

The results of this research have been previously published in:

1. Cao L, Lou J, Chan SY, et al. Rapid evaluation of COVID-19 vaccine effectiveness against symptomatic infection with SARS-CoV-2 variants by analysis of genetic distance. Nat Med 2022;28:1715-22.
2. Lou J, Zhao S, Cao L, et al. Temporal patterns in the evolutionary genetic distance of SARS-CoV-2 during the COVID-19 pandemic. Public Health Genomics 2022. doi:10.1159/000520837
3. Lou J, Zheng H, Zhao S, et al. Quantifying the effect of government interventions and virus mutations on transmission advantage during COVID-19 pandemic. J Infect Public Health 2022;15:338-42.
4. Zhao S, Lou J, Cao L, et al. Quantifying the transmission advantage associated with N501Y substitution of SARS-CoV-2 in the UK: an early data-driven analysis. J Travel Med 2021;28:taab011.
5. Zhao S, Lou J, Chong MKC, et al. Inferring the association between the risk of COVID-19 case fatality and N501Y substitution in SARS-CoV-2. Viruses 2021;13:638.
6. Chan RWY, Chan KCC, Lui GCY, et al. Mucosal antibody response to SARS-CoV-2 in paediatric and adult patients: a longitudinal study. Pathogens 2022;11:397.
7. Zhao S, Lou J, Cao L, et al. Differences in the case fatality risks associated with SARS-CoV-2 Delta and non-Delta variants in relation to vaccine coverage: an early ecological study in the United Kingdom. Infect Genet Evol 2022;97:105162.

References

1. Tasakis RN, Samaras G, Jamison A, et al. SARS-CoV-2 variant evolution in the United States: high accumulation of viral mutations over time likely through serial Founder Events and mutational bursts. PLoS One 2021;16:e0255169.
2. Markov PV, Ghafari M, Beer M, et al. The evolution of SARS-CoV-2. Nat Rev Microbiol 2023;21:361-79.
3. Duffy S, Shackelton LA, Holmes EC. Rates of evolutionary change in viruses: patterns and determinants. Nat Rev Genet 2008;9:267-76.
4. Duchene S, Featherstone L, Haritopoulou-Sinanidou M, Rambaut A, Lemey P, Baele G. Temporal signal and the phylodynamic threshold of SARS-CoV-2. Virus Evol 2020;6:veaa061.
5. Koel BF, Burke DF, Bestebroer TM, et al. Substitutions near the receptor binding site determine major antigenic change during influenza virus evolution. Science 2013;342:976-9.
6. Shu Y, McCauley J. GISAID: global initiative on sharing all influenza data - from vision to reality. Euro Surveill 2017;22:30494.
7. Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol 2018;4:vey016.
8. Cao L, Lou J, Chan SY, et al. Rapid evaluation of COVID-19 vaccine effectiveness against symptomatic infection with SARS-CoV-2 variants by analysis of genetic distance. Nat Med 2022;28:1715-22.
9. Lou J, Zhao S, Cao L, et al. Temporal patterns in the evolutionary genetic distance of SARS-CoV-2 during the COVID-19 pandemic. Public Health Genomics 2022. doi:10.1159/000520837
10. Ghafari M, du Plessis L, Raghwani J, et al. Purifying selection determines the short-term time dependency of evolutionary rates in SARS-CoV-2 and pH1N1 influenza. Mol Biol Evol 2022;39:msac009.