Full vaccination is imperative to suppress SARS-CoV-2 delta variant mutation frequency

Ting-Yu Yeh and Gregory P. Contreras

Auxergen Inc., Columbus Center, Institute of Marine and Environmental Technology,
University of Maryland, 701 E Pratt Street, Baltimore, MD 21202, USA

*Ting-Yu Yeh is Corresponding author

E-mail: yehty@auxergen.com
Tel: 1-443-762-1974

Key words: COVID-19, SARS-CoV-2, mutation frequency, Tajima's *D* test
Short running title: Suppression of SARS-CoV-2 mutation by vaccination
Word count: 977 words.

Abstract

Public distrust has undermined COVID-19 vaccine acceptance and has become a major public health issue in the battle against SARS-CoV-2 transmission globally. Here we present the first evidence that the vaccination coverage rate is inversely correlated to the mutation frequency of the SARS-CoV-2 delta variant in 16 countries (R^2=0.878), strongly indicating that full vaccination against COVID-19 is critical to suppress emergent mutations. We also present a promising tool to forecast new COVID-19 outbreaks. The Tajima's *D* test, an evolutionary algorithm, with a threshold value of -2.50 is shown to be an accurate predictor of new outbreaks. We recommend that universal vaccination, as well as mitigation strategies, and genomic surveillance continue to be employed to prevent further viral transmission.

Introduction

Legally required vaccination against various infectious diseases is essential to public health policy in many countries. However, resistance to voluntary COVID-19 vaccination has emerged worldwide, including in the United States. In mid-June 2021, 33% of Americans said they were not willing to be vaccinated [1]. Public distrust has undermined COVID-19 vaccine acceptance, in association with a belief that the vaccine is ineffective, dangerous and compromises individual freedom [2]. Therefore, overcoming COVID-19 vaccination resistance has become a major challenge to prevent global SARS-CoV-2 transmission.

SARS-CoV-2 is characterized by an average mutation rate of 7.23 mutations per viral sample [3]. Mutations drive genome variability, generating many different SARS-CoV-2 variants as the virus evolves to escape vaccine-mediated immunity and thereby, develop drug or vaccine resistance [4]. The delta (lineage B1.617.2) variants were first documented in October 2020 in India and were designated variants of concern by the World Health Organization due to its increased transmissibility and virulence [5]. All SARS-CoV-2 mutations result from two major mechanisms: (1) spontaneous substitution/deletion of nucleotides, and (2) RNA recombination [3,4,6]. However, it remains unclear exactly how human interventions (vaccinations, lockdowns, etc.) affect viral mutations.

Findings

To explore this question, we first analyzed the correlation between the rates of full vaccination and the point mutation frequency (*Mf*) of COVID-19 delta variants' genome in 20 countries. Complete SARS-CoV-2 genome sequences with high coverage from June 20 to July 3 2021 in 20 countries: Australia (N=121), France (N=788), Germany (N=955), Indonesia (N=97), India (N=171), Ireland (N=617), Israel (N=333), Italy (N=642), Japan (N=105), Mexico (N=368), Netherland (N=456), Norway (N=142), Portugal (N=782), Singapore (N=131), Spain (N=689), Switzerland (N=131), Sweden (N=786), Turkey (N=428), United States (N=537), and United Kingdom (UK, N=3534) were collected from the Global Initiative on Sharing All Influenza Data (GISAID)[7]. *Mf* was calculated as described previously [3]. We found that *Mf* was logarithmically reduced as the full vaccination rate increased in 16 of the 20 countries (R^2=0.878, Figure 1A). To our knowledge, this is the first evidence that vaccinations successfully suppressed viral mutations. With 10.8% vaccinated rate, *Mf* was exceptionally low in Australia, likely as a result of the national lockdown restrictions. In contrast, *Mf* is higher in Japan, Switzerland, and United States, suggesting that their mitigation strategies have been less successful.

Previously we reported that *de novo* recombination of viral genome and positive selection pressure led to prevalent SARS-CoV-2 mutations among passengers in quarantine on the Diamond Princess cruise ship between January and March 2020 [6]. It is still unknown whether COVID-19 variants are evolving randomly or under the selection pressure generated by mitigation efforts. We chronologically analyzed SARS-CoV-2 delta variant sequences in the UK (N=27,344), India (N=4,451) and

Australia (N=305) from 13 February to 3 July 2021, respectively in the GISAID [7]. Genome alignments and analysis using the Tajima D test, a neutrality statistic widely used in population genetics, were carried out as previously [6,8,9,10]. The Tajima D test (D') compares pairwise nucleotide diversity (π, the average number of nucleotide differences per site between two sequences) and total polymorphism to infer selection and demographic events [6, 9, 10].

We observed that π was significantly lower in the UK ($\pi < 1.5 \times 10^{-4}$) than others which emerged after April 24, 2021 ($\pi > 2.0 \times 10^{-4}$; $p<0.01$) (Figure 1B), suggesting that the UK delta variants emerged with rapid clonal expansion, while others gained more singleton mutations. D' value decreased drastically in both the UK and India (Figure 1C). All variants eventually remained between -2.68 to -2.84 ($p<0.01$), which was similar to the D' value calculated from the alpha variant sequences in UK (-2.58 to -2.78; N=15,396) during the period of study. Negative D' signifies stronger demographic expansion or purifying (positive) selection, witnessed in both the UK and Indian variants[10]. In Australia, D' decreased in June 2021, suggesting positive selection exerted by lockdowns [6].

Together with available epidemiological reports [5], we found that one to three weeks after D' was reduced below -2.50, new COVID-19 outbreaks emerged in the UK and India (Figure 1C). Although drivers of positive selection may vary (vaccinations, lockdowns, etc.), we propose that a D' score with a threshold of -2.50 is a promising forecasting tool for COVID-19 outbreaks. It is worth noting that D' of the Australian delta variants reached -2.57 on 3 July 2021, suggesting that a further extension in restrictions was required to contain outbreaks due to low vaccination rates nationwide.

Conclusion and recommendation

More virulent strains of SARS-CoV-2 have emerged with enhanced transmissibility and immune evasion properties. For example, multiple variants have escaped neutralizing antibodies developed to target the spike protein receptor-binding or N-terminal domain [11]. What's more, the case numbers of breakthrough infection caused by the delta variant have increased drastically worldwide [12].

Our observations showed that countries with higher vaccination rates generated fewer mutations versus low vaccinated countries. This suggests that there is less of a chance for the virus to gain more virulent mutations in high vaccinated countries. It must be noted that the virus becomes more contagious as it is screened through the vaccinated population, eventually to become the dominant strain to infect the entire population. This is what we are seeing in the delta variant in the United States, the UK and many countries [12]. A negative effect of COVID-19 vaccination resistance could result in a large increase of virulent mutations and failure to reach herd immunity nationwide or globally.

Thus, we recommend that: 1) universal vaccination should be administered as soon as possible to suppress the generation of deadly mutations; 2) mitigations strategies such as personal protection equipment, social distancing, etc., should be re-implemented with new urgency until full vaccination is reached in an effort to prevent viral transmission until the virus is eradicated; 3) genomic surveillance should be undertaken to monitor for new mutations; and 4) The Tajima *D'* threshold test further provides a tool by which we can forecast future outbreaks.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Author contributions

All authors contributed to study concept, rationale, and initial manuscript drafts, interpretation of data, and final manuscript preparation. All authors have read and approved the final version of the manuscript.

Funding

No funding

Ethical approval

None declared.

References

1. Kirzinger A, Sparks G, Brodie M, KFF COVID-19 vaccine monitor-Rural America. https://www.kff.org/coronavirus-covid-19/poll-finding/kff-covid-19-vaccine-monitor-in-their-own-words-six-months-later/ Accessed 13 July 2021.
2. Schmelz K, Bowles S, Overcoming COVID-19 vaccination resistance when alternative policies affect the dynamics of conformism, social norms, and crowding out. Proc Natl Acad Sci U S A 2021;118(25):e2104912118. PMID: 34099578
3. Mercatelli D, Giorgi FM, Geographic and Genomic Distribution of SARS-CoV-2 Mutations. Front Microbiol 2020;11:1800. PMID: 32793182

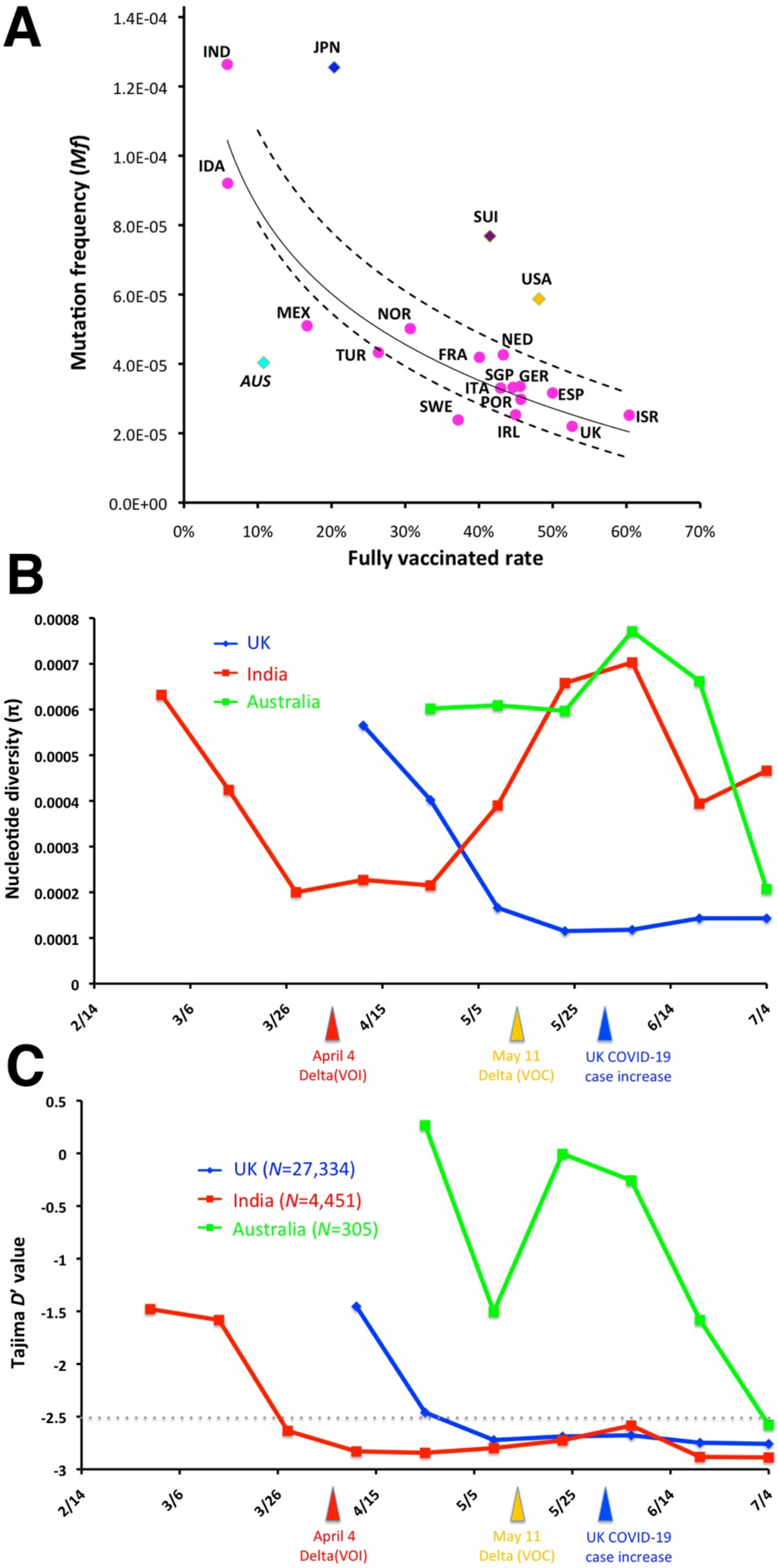
4. Roy C, Mandal, SM, Mondal SK, Mukherjee S, Mapder T, et al., Trends of mutation accumulation across global SARS-CoV-2 genomes: Implications for the evolution of the novel coronavirus. Genomics 2020;112(6):5331-5342. PMID: 33161087
5. The World Health organization. Tracking SARS-CoV-2 variants. https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/
6. Yeh, TY, Contreras, GP, Viral transmission and evolution dynamics of SARS-CoV-2 in shipboard quarantine. Bull World Health Organ. 2021; 99(7): 486–495. PMID: 34248221
7. Global Initiative on Sharing All Influenza Data. Available from: https://www.gisaid.org/
8. Katoh, K, Standley, DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30(4):772–80. PMID: 23329690
9. Kumar S, Stecher G, Tamura K. MEGA7: Molecular evolutionary genetics analysis Version 7.0 for bigger datasets. Mol Biol Evol. 2016;33(7):1870–4. PMID: 27004904
10. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989;123(3):585–595. PMID: 2513255
11. Harvey WT, Carabelli AM, Jackson B, et al. SARS-CoV-2 variants, spike mutations and immune escape. Nat Rev Microbiol. 2021;19(7):409-424. PMID: 34075212

12. Farinholt T, Doddapaneni H, Qin X, Menon V, Meng Q, et al. Transmission event of SARS-CoV-2 Delta variant reveals multiple vaccine breakthrough infections. medRxiv 2021 Jul 12;2021.06.28.21258780. doi: 10.1101/2021.06.28.21258780. PMID: 34268529

13. Coronavirus (COVID-19) Vaccinations. Our World in Data. 2021 July. https://ourworldindata.org/covid-vaccinations

Figure Legend

Figure 1. (A) Correlation between full vaccinated rate [13] and mutation frequency (*Mf*) from June 20 to July 3 2021 in 20 countries: Australia (AUS), France (FRA), Germany (GER), Indonesia (IDA), India (IND), Ireland (IRL), Israel (ISR), Italy (ITA), Japan (JPN), Mexico (MEX), Netherland (NED), Norway (NOR), Portugal (POR), Singapore (SGP), Spain (ESP), Switzerland (SUI), Sweden (SWE), Turkey (TUR), United States (USA), and UK. Logarithmic regression (solid) line was draw based on 16 countries (pink dots) with a calculated 95% confidence interval (dashed lines). Japan, Switzerland, USA, and Australia are labeled in different colors as outliers. (B and C) Chronology of nucleotide diversity (π) (B) and Tajima D' value (C) of SARS-CoV-2 delta variants in UK (N=27,344, blue), Indian (N=4,451, red), and Australian (N=305, green). Data were plotted every two weeks, and the data only represent the effective population size with more than 3 high quality sequences. The arrows label the epidemiological events of COVID-19 delta variants announced by the World Health Organization (WHO). WHO classified the delta variant as a global variant of interest (VOI) on 4 April 2021, and variants of concern (VOC) on 11 May 2021 [5]. The dashed line in (C) labels the cut-off threshold -2.50 in Tajima D' test.


A
Mutation frequency (Mf)
1.4E-04
1.2E-04
1.0E-04
8.0E-05
6.0E-05
4.0E-05
2.0E-05
0.0E+00
0%
10%
20%
30%
40%
50%
60%
70%
Fully vaccinated rate
IND
JPN
IDA
SUI
USA
MEX
NOR
NED
TUR
FRA
SGP
GER
ITA
ESP
POR
SWE
ISR
IRL
UK
AUS
B
Nucleotide diversity (π)
0.0008
0.0007
0.0006
0.0005
0.0004
0.0003
0.0002
0.0001
0
UK
India
Australia
2/14
3/6
3/26
4/15
5/5
5/25
6/14
7/4
April 4
Delta(VOI)
May 11
Delta (VOC)
UK COVID-19
case increase
C
Tajima D' value
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3
UK (N=27,334)
India (N=4,451)
Australia (N=305)
2/14
3/6
3/26
4/15
5/5
5/25
6/14
7/4
April 4
Delta(VOI)
May 11
Delta (VOC)
UK COVID-19
case increase