Full vaccination suppresses SARS-CoV-2 delta variant mutation frequency

Ting-Yu Yeh, MD and Gregory P. Contreras, MBA

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 D test

Short running title: Suppression of SARS-CoV-2 mutation by vaccination

Word count: 600 words.

Key points

Question: It remains unclear how human interventions (vaccinations, lockdowns, etc.) affect viral mutation or generate selection pressure of SARS-CoV-2. It has also been obscure if there are differences in various geographic populations.

Findings: The vaccination coverage rate is inversely correlated to the mutation frequency of the SARS-CoV-2 delta variants in 16 countries of 20 countries studied. We also discovered delta variants evolved differently under the positive selection pressure in the United Kingdom and India.

Meaning: Full vaccination against COVID-19 is critical to suppress emergent mutations. Tajima's *D* test score, with a threshold value as -2.50, can provide a promising tool to forecast new COVID-19 outbreaks.

Abstract

This study presents the first evidence that full vaccination against COVID-19 suppresses emergent mutations of SARS-CoV-2 delta variants. An evolution algorithm, Tajima's *D* test with a threshold value as -2.50, can provide a promising tool to forecast new COVID-19 outbreaks.

Introduction

Mutations generate different SARS-CoV-2 variants to escape vaccine-mediated immunity and thereby, develop drug or vaccine resistance[1,2]. Until now, it remained unclear how human interventions (vaccinations, lockdowns, etc.) affected viral mutations.

Methods

Complete SARS-CoV-2 delta variant sequences from June 20 to July 3 2021: 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)[3]. Delta variant sequences of UK (*N*=27,344), India (*N*=4,451) and Australia (*N*=305) from 13 February to 3 July 2021, respectively, were downloaded from the GISAID. Genome alignments, point mutation frequency (*Mf*) and Tajima's *D* test were analyzed as described previously[1,4,5].

Results

We first analyzed the correlation between the full vaccination rates and *Mf* of COVID-19 delta variants' genome in 20 countries. We found that *Mf* was logarithmically reduced as the vaccination rate increased in 16 of the 20 countries

(R^2=0.878, Figure 1). 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. *Mf* was higher in Japan, Switzerland, and United States, suggesting that other mitigations may also facilitate suppression of viral mutations.

To explore whether mutations evolve randomly, we chronologically analyzed delta variant sequences in the UK, India and Australia with Tajima *D'* test, which compares pairwise nucleotide diversity (π) and total polymorphism to infer selection and demographic events[5]. π was significantly lower in UK ($\pi < 1.5 \times 10^{-4}$) than India or Australia ($\pi > 2.0 \times 10^{-4}$; $p<0.01$) (Figure 2A), suggesting that UK delta variants emerged with rapid clonal expansion, while others gained more singleton mutations. *D'* value decreased in UK and India prior to outbreaks, and remained between -2.68 to -2.84 (Figure 2B). Negative *D'* signifies stronger demographic expansion or purifying (positive) selection, witnessed there. In Australia, *D'* decreased in June 2021, suggesting positive selection exerted by lockdowns[4]. Excluding these 4 countries, our data indicates that delta variants evolved to become dominant in inverse correlation to vaccinations.

Together with epidemiological reports[2], we found that one to three weeks after *D*' reduced to less than -2.50, new COVID-19 outbreaks emerged in India and UK (Figure 2B). Although drivers of positive selection (vaccinations, lockdowns, etc.) vary, we propose that a *D*' score with a threshold of -2.50 is a promising forecasting

tool for COVID-19 outbreaks. *D'* of Australian delta variants reached -2.57 on July 3 2021, indicating that lockdowns in Australia were justified to contain outbreaks due to low vaccination rates nationwide.

Discussion

Our data shows that countries with higher vaccination rates generated fewer mutations, suggesting less of a chance for the virus to gain virulent mutations in high vaccinated countries. More virulent SARS-CoV-2 strains have emerged with enhanced transmissibility and immune evasion properties to cause breakthrough infections[6]. We conclude that the virus becomes more contagious as it is screened through the vaccinated population and the resultant strain becomes the dominant strain and able to infect the entire population.

We recommend that: 1) universal vaccination should be administered as soon as possible to suppress the generation of virulent mutations; 2) mitigations strategies such as personal protection equipment, social distancing, etc., should continue to be employed 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'* test provides a tool by which we can forecast future outbreaks.

Conflict of interest statement

No potential conflict of interest was reported by the authors.

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. Roy C, Mandal, SM, Mondal SK, 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. doi: 10.1016/j.ygeno.2020.11.003.
2. The World Health organization. Tracking SARS-CoV-2 variants. https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/
3. Global Initiative on Sharing All Influenza Data. Available from: https://www.gisaid.org/
4. 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. doi: 10.2471/BLT.20.255752.

5. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 1989;123(3):585–595.
6. Farinholt T, Doddapaneni H, Qin X, et al. Transmission event of SARS-CoV-2 Delta variant reveals multiple vaccine breakthrough infections. *medRxiv* 2021; doi: 10.1101/2021.06.28.21258780.

Figure Legend

Figure 1. Correlation between full vaccinated rate (https://ourworldindata.org/covid-vaccinations) 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 the United Kingdom (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.

Figure 2. Chronology of nucleotide diversity (π) (A) and Tajima *D'* value (B) of delta variant sequences of 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[2]. The dashed line in (B) labels the cut-off threshold -2.50 in Tajima *D'*.

Figure 1.

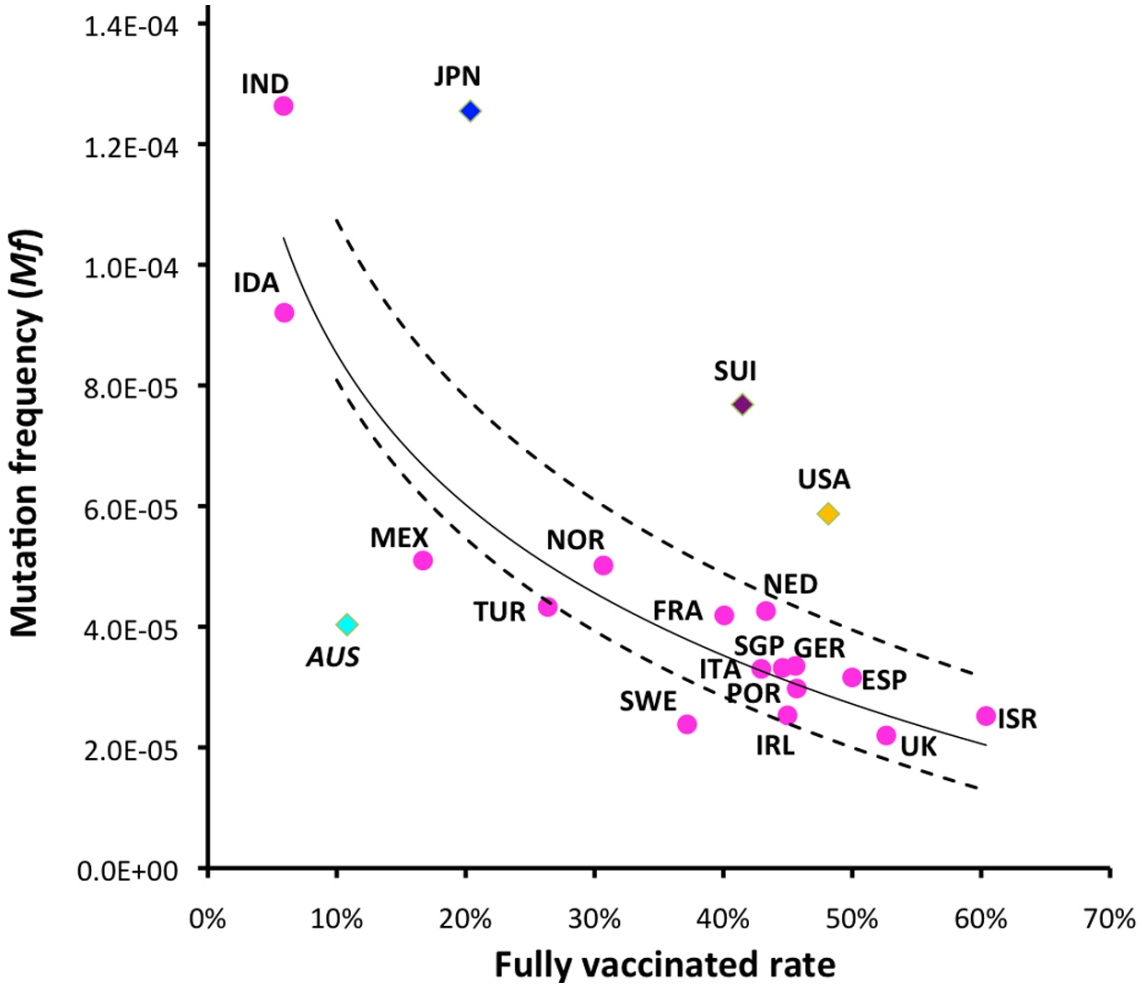
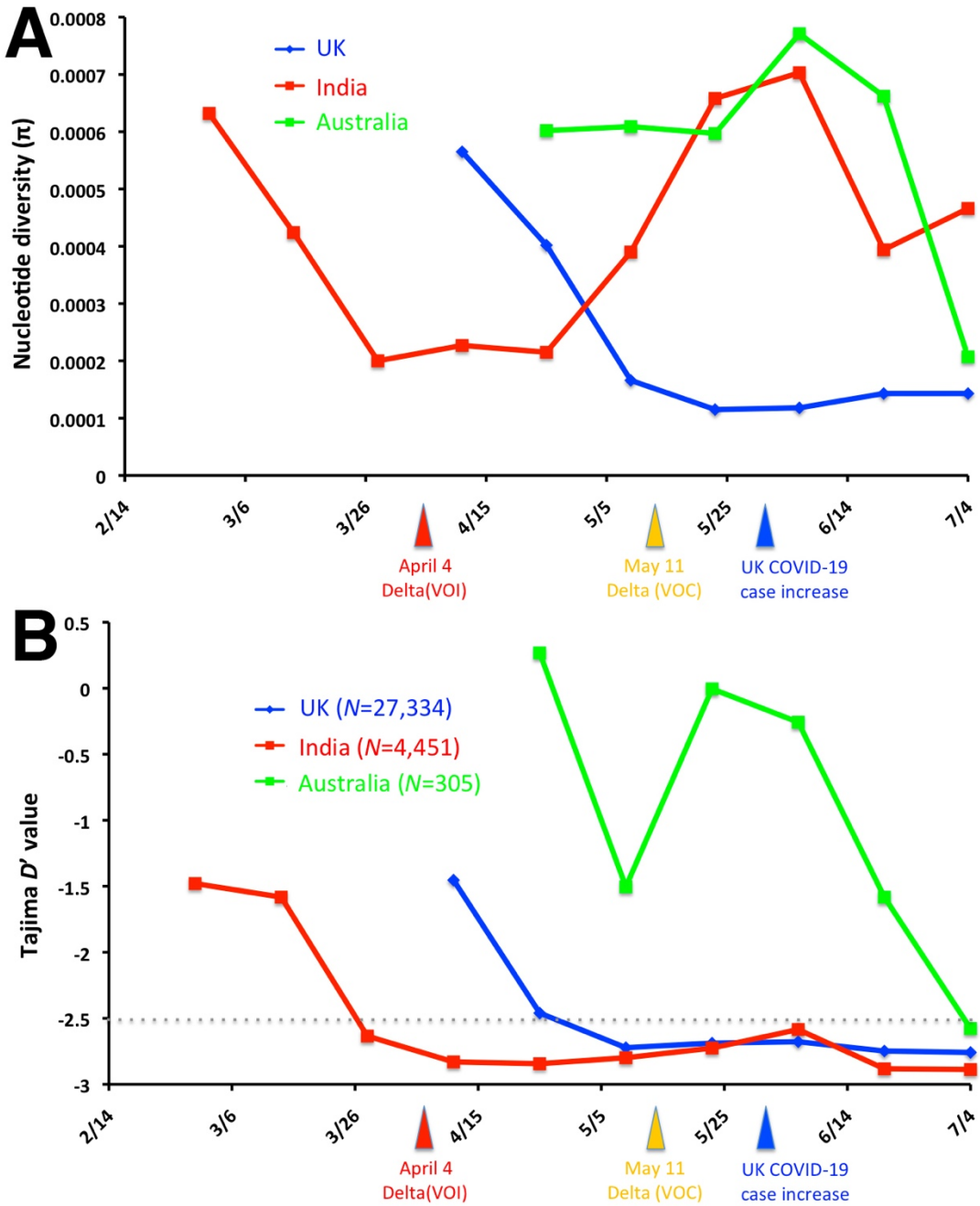

1.4E-04
1.2E-04
1.0E-04
8.0E-05
6.0E-05
4.0E-05
2.0E-05
0.0E+00
Mutation frequency (Mf)
0%
10%
20%
30%
40%
50%
60%
70%
Fully vaccinated rate
IND
JPN
IDA
SUI
USA
MEX
NOR
NED
TUR
FRA
AUS
SGP
GER
ITA
ESP
POR
SWE
ISR
IRL
UK

Figure 2.


A
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
B
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