

Article

Comparative Dynamics of Delta and Omicron SARS-CoV-2 Variants across and between California and Mexico

Sanjay R. Mehta ^{1,2} , Davey M. Smith ^{1,2}, Celia Boukadida ³  and Antoine Chaillon ^{1,*} 

¹ Department of Medicine, University of California, San Diego, CA 92093, USA; srmehtha@health.ucsd.edu (S.R.M.); d13smith@ucsd.edu (D.M.S.)

² Veterans Affairs Health System, San Diego, CA 92093, USA

³ Centro de Investigación en Enfermedades Infecciosas, Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas, Ciudad de México 14080, Mexico; celia.boukadida@cieni.org.mx

* Correspondence: achaillon@health.ucsd.edu

Abstract: Evolutionary analysis using viral sequence data can elucidate the epidemiology of transmission. Using publicly available SARS-CoV-2 sequence and epidemiological data, we developed discrete phylogeographic models to interrogate the emergence and dispersal of the Delta and Omicron variants in 2021 between and across California and Mexico. External introductions of Delta and Omicron in the region peaked in early July (2021-07-10 [95% CI: 2021-04-20, 2021-11-01]) and mid-December (2021-12-15 [95% CI: 2021-11-14, 2022-01-09]), respectively, 3 months and 2 weeks after first detection. These repeated introductions coincided with domestic migration events with no evidence of a unique transmission hub. The spread of Omicron was most consistent with gravity centric patterns within Mexico. While cross-border events accounted for only 5.1% [95% CI: 4.3–6] of all Delta migration events, they accounted for 20.6% [95% CI: 12.4–29] of Omicron movements, paralleling the increase in international travel observed in late 2021. Our investigations of the Delta and Omicron epidemics in the California/Mexico region illustrate the complex interplay and the multiplicity of viral and structural factors that need to be considered to limit viral spread, even as vaccination is reducing disease burden. Understanding viral transmission patterns may help intra-governmental responses to viral epidemics.

Keywords: COVID-19; SARS-CoV-2; variants; Delta; Omicron; phylogeography



Citation: Mehta, S.R.; Smith, D.M.; Boukadida, C.; Chaillon, A.

Comparative Dynamics of Delta and Omicron SARS-CoV-2 Variants across and between California and Mexico. *Viruses* **2022**, *14*, 1494. <https://doi.org/10.3390/v14071494>

Academic Editor: Manuela Sironi

Received: 7 June 2022

Accepted: 4 July 2022

Published: 8 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Identifying and tracking new SARS-CoV-2 variants is critical for public health efforts. Previous studies have shown that in the early phase of epidemic spread, intensive surveillance and interventions reduced transmission rates and even prevented the establishment of SARS-CoV-2 in some locations [1,2]. However, as the infectiousness of SARS-CoV-2 has increased with new variants, patterns of transmission have evolved, requiring a reconsideration of interventions. To help in this effort, scientists from around the world have curated and made publicly available over 8 million complete viral sequences since the start of the COVID-19 pandemic. As these genomic data are being collected and published in near real-time, this also enables the monitoring of the evolution of the virus across the world at an unprecedented speed and scale [3]. As a result, the Delta variant, first detected in India in October 2020, was classified as a variant of concern (VOC) by the World Health Organization (WHO) in May 2021 while the Omicron variant, first detected in South Africa and Botswana in November 2021, was designated as a VOC that same month. In addition, the Centers for Disease Control (CDC) used these genomic data to track the rapid rise of the Delta variant of concern (VOC) in June–July 2021 (>50% circulating SARS-CoV-2 in the USA at the end of June, and >95% at the end of July). These data were used to inform a coordinated approach to combat the pandemic, including temporary travel restriction and the extension of US–Mexico land border closure to non-essential travel [4] in July 2021 until

8 November 2021, when travel resumed for fully vaccinated foreign travelers to the USA. These data also identified the emergence of the Omicron VOC (first detected on 1 December in the USA) and its rapid rise in the following weeks (>50% of viruses in the USA for the week ending 25 December and >99% at the week ending 22 January 2022, while the relative proportion of Delta dropped to <1% of cases) [5].

Publicly available genomic data can also be used to estimate epidemiologic characteristics and the spatiotemporal dynamics of viral spread [6]. Furthermore, since the evolution of a single VOC is relatively slow and steady ($\sim 9 \times 10^{-4}$ mutations per site per year [7]), Bayesian phylogeographic inferences from sequences that have been annotated with geographic origin and time can provide a framework to gain insight into the epidemiology of these SARS-CoV-2 variants [8–10]. The spatial heterogeneity in these transmission patterns can also impact the intensity of cases and population level immunity [7,11].

In this study, we use all publicly annotated full-length SARS-CoV-2 sequences available on GISAID (Global Initiative on Sharing All Influenza Data, GISAID public database [12]) as of 31 January 2022 to understand differences in the dispersal of Delta and Omicron across and between Mexico and CA, USA. We focus on these areas because Mexico and the US state of California are tightly linked culturally and commercially. For example, the San Diego–Tijuana land border alone has over 45 million crossings annually [13]. Additionally, as the most populated state in the USA with a robust economy, California also has cultural and economic ties to much of the rest of the world. For this reason, when new pandemics enter the USA (e.g., HIV, H1N1 influenza), California is usually one of the earliest USA states to be affected. Similarly, Mexico City, as an economic and cultural hub with ties across the world, is often impacted early in pandemics.

2. Materials and Methods

2.1. Compilation of Datasets and Data Preparation for Phylogenetic Inference

We analyzed full-length SARS-CoV-2 genomes with known dates of sampling collected on or before 31 January 2022 and made available on the GISAID public database [12]. From this dataset, we extracted Delta and Omicron variants based on the variant assignment provided on GISAID. Sequences with missing or incomplete dates of sampling or geospatial information were excluded. The remaining sequences were curated and aligned using Nextalign implemented in Nextclade [14]. To allow the simultaneous viewing of sampling dates and geographic information, data were uploaded to the MicroReact platform [15] under <https://microreact.org/project/rqnwd9nymcvdkvq1dktbjf> and <https://microreact.org/project/d5mblhkwgxtofu6wdtxt> for Delta and Omicron, respectively.

2.2. Subsampling

The massive amount of available viral genomic data provides a unique opportunity to gain valuable real-time insight into dispersal dynamics, but it also leads to computational challenges and is not necessarily representative, as testing and sequencing efforts vary through time and between locations. Hence, in Supplementary Table S1 and Figures S1 and S2, we show a higher number of infections sequenced in San Diego and Los Angeles Counties compared to other counties in California and across Mexico, highlighting the lack of correlation between cases and sampling depth. These structural factors (e.g., testing efforts, sequencing capacity) may bias attempts to accurately model these epidemics. To tackle these limitations and optimize the phylogenetic inferences described below, we applied a subsampling strategy to enhance the computational feasibility of our phylogeographic analyses. This strategy was informed by the COVID-19 epidemic in the California–Mexico region. First, we gathered the number of confirmed COVID-19 cases in Californian counties [16] and Mexican states [17] since January 2021. Next, we normalized reported cases per geospatial unit to the population size (i.e., number of cases per 100,000 individuals, see Supplementary Figure S3). To obtain datasets that reflected the local dynamics of the past year and limit sampling bias (i.e., the over-representation of sequences from a particular location due to increased sequencing effort), we applied

a random subsampling strategy informed by the number of cases/100k individuals per month (for Delta sequences) or week (for Omicron sequences).

2.3. Lineage Identification

We used the Phylogenetic Assignment of Named Global Outbreak Lineages (Pangolin) (version 1.1.14) command-line tool and PANGO nomenclature system [18] to determine the lineage of all SARS-CoV2 genomes. The distribution lineages by Californian counties and Mexican states are provided in Figure 1 and Figure S4 using the MicroReact projects noted.

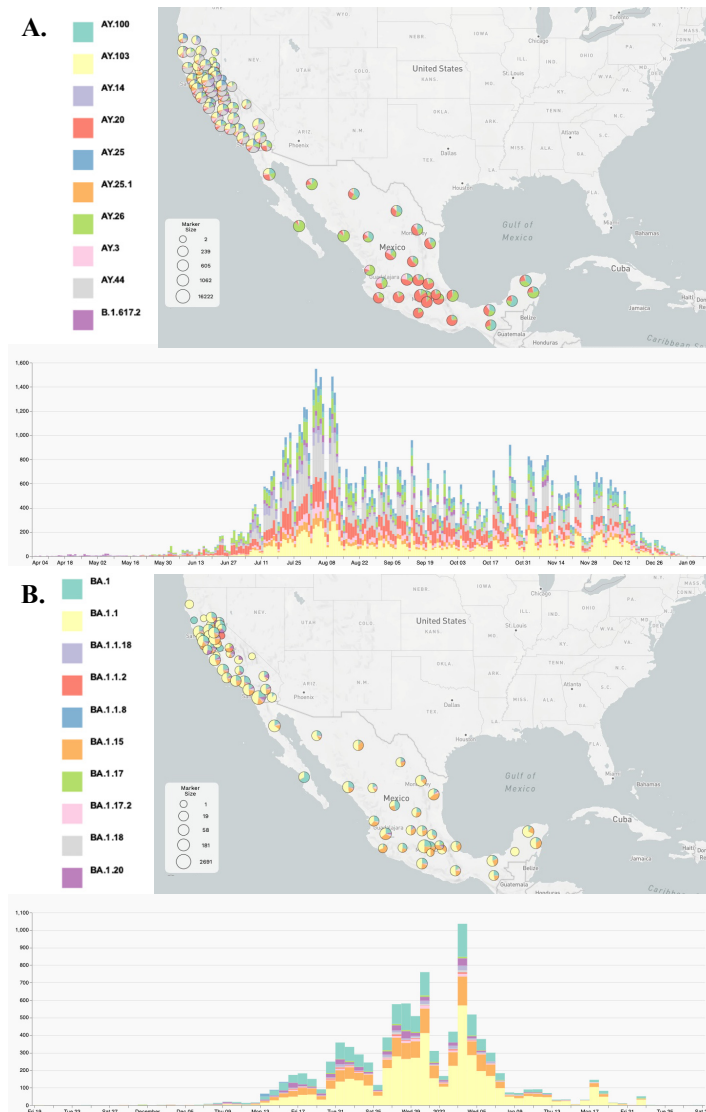


Figure 1. Visualization of the ten most prevalent Delta (A) and Omicron (B) genomes using MicroReact. MicroReact [15] display of the geolocation of the ten most prevalent genomes with associated lineage inferred by the Pangolin assignment tool [18]. See also Supplementary Figure S4 and MicroReact projects <https://microreact.org/project/rqnwd9nymcvdkvq1dktbjf> and <https://microreact.org/project/d5mblhkwgxtofu6wdtxt> for all circulating Delta and Omicron sub-variants, respectively.

2.4. Identification of Representative Phylogenetic Clades

We applied the following step-by-step analytical approach [19]: using the final alignments for each subset of Delta and Omicron sequences combined with their respective background described above, a maximum likelihood phylogenetic tree was inferred using FastTree 2 [20] under a GTR+ Γ evolutionary model. All clades within the trees were analyzed to delineate the multiplicity and timing of external introductions of COVID-19

in the region for each VOC. Of those, clades of size ≥ 5 including sequences originating from ≥ 2 distinct locations (i.e., Californian counties and Mexican states) were retained for further evaluating the dispersal of both epidemics within California and Mexico and across the California–Mexico border.

2.5. Timing of External Introductions

For each clade, the phylogeny was rescaled into units of time with *treedater* [21], assuming a strict molecular clock with the rate of SARS-CoV-2 genome evolution (of a single VOC) drawn from an externally estimated distribution [7]. Specifically, a normal distribution was specified with mean 9.41×10^{-4} nucleotide substitutions per site per year and a standard deviation of 4.99×10^{-5} [7]. This allowed us to infer the time to the most recent common ancestor (tMRCA) of each clade, which reflects the timing of that particular introduction. To incorporate uncertainty in the estimated clock rate, molecular clock estimation was replicated 100 times and the average timing of introduction for each clade was inferred.

2.6. Phylogeographic Inference

Phylogeographic inference was performed using the asymmetric discrete trait diffusion model and a GTR+ Γ substitution model [22] implemented in the BEAST 1.10.5 software package [23]. Briefly, we employed a Metropolis–Hastings algorithm for MCMC chains that required 3 parallel runs for Omicron and 10 parallel runs for Delta variants that were run for sufficiently long (from 1.9×10^8 to 3.5×10^9 iterations) to obtain enough samples from the posterior distribution in order to reach adequate ESS values as estimated by the program Tracer v1.7 [24]. Trees were sampled every 250,000 iterations and 10% of the sampled trees were discarded as burn-in. Estimates of the posterior probability of the expected number of variant introduction and migration events between all pairs of locations (Markov jumps) were computed through stochastic mapping techniques [25]. To identify the subset of transition rates that was most informative to reconstruct the dispersal history and assess its significance, we applied a measure of statistical support associated with a lower false-positive rate by incorporating information on the relative abundance of samples from each geospatial location in the data set, which was coined as the adjusted Bayes factor (BF_{adj} ; see our publication [26] for the detailed procedure). In our analyses, we only considered variant migration events associated with a BF_{adj} support ≥ 3 as evidence of positive support [27].

2.7. Mobility Data

To estimate human mobility patterns across the California–Mexico border, we analyzed travel pattern edges between the United States and Mexico provided by Facebook [28]. These data can be used to provide insight into factors impacting viral migration, as shown in the UK [29]. These data are based on the number of Facebook users moving over large distances, such as air or train travel. Counts of international travel events are updated daily based only on users who have opted to share precise location data from their device with the Facebook mobile app through location services.

3. Results

3.1. Clade Identification and Variant External Introduction in the Border Region

We analyzed 110,378 and 9610 SARS-CoV-2 complete genome sequences identified as Delta and Omicron VOCs sampled in California ($n = 86,688$ and 5656 , respectively) and Mexico ($n = 23,690$ and 3954 , respectively) in combination with a set of representative background sequences sampled outside of the California–Mexico region ($n = 88,826$ and $19,531$ Delta and Omicron complete genomes available). See Supplementary Figure S1 and Table S1 and MicroReact projects <https://microreact.org/project/rqnwd9nymcvdkvq1dktbj> and <https://microreact.org/project/d5mblhkwgxtofu6wdtxt> for Delta and Omicron, respectively. To accommodate the computational challenges with the large dataset,

we generated ten random subsets of Delta and Omicron sequences from California/Mexico (Supplementary Table S2). These random subsets were combined with a background set of sequences from outside the region with a maximum of ten sequences per country per month for Delta and per week for Omicron, and were used to perform discrete phylogeographic analyses to identify the ancestral internal nodes of distinct descending clades of infection in the California–Mexico region.

Delta variant. A mean of 747 [range: 744–753] sequences (79 locations, i.e., Californian counties or Mexican states) were included in each subsampled phylogeny along with a representative background dataset of external sequences, and we identified a mean of 101 clades of size ≥ 2 across 69 locations. The peak time of external introductions of Delta into California/Mexico was estimated to be in early July 2021 (2021-07-10 [95% CI: 2021-04-20, 2021-11-01]) for all down-sampled subsets (i.e., tMRCA, see Figures 2 and S5).

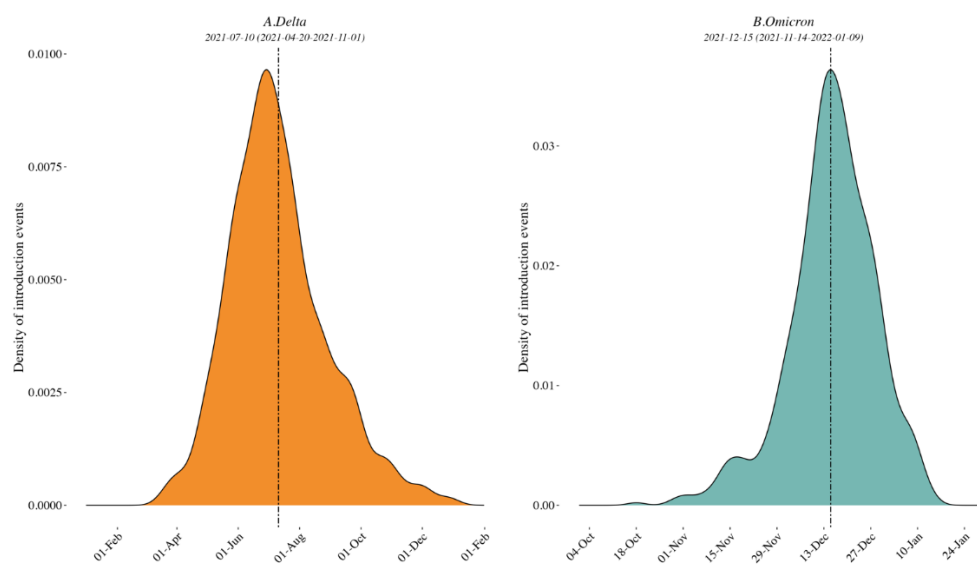


Figure 2. Density plots of external introduction of the identified clades for Delta (A) and Omicron (B) variants. A normal distribution was specified for the evolutionary rate with a mean of 9.41×10^{-4} nucleotide substitutions per site per year and a standard deviation of 4.99×10^{-5} [7]. All 10 replicates are combined to show the timing of the introductions of the clades in the region. See also Supplementary Figure S5 for the results of each of the 10 individual replicates.

Omicron variant. A mean of 642 [range: 631–661] sequences (64 locations, i.e., Californian county or Mexican state) were included in each subsampled phylogeny, and we identified a mean of 70 clades of size ≥ 2 across 25 locations. The peak time of introduction in the California–Mexico border region was estimated to be in mid-December 2021 across all data replicates (2021-12-15 [95% CI: 2021-11-14, 2022-01-09], see Figures 2 and S5). Characteristics of each replicate dataset for each VOC are summarized in Supplementary Table S2.

3.2. Lineage Distribution of the Delta and Omicron Variants

Next, we analyzed the geographical distribution of Delta and Omicron lineages. After they emerged, the parental B.1.617.2 lineage (Delta variant) and B.1.1.529 lineage (Omicron variant) diversified into 245 and 111 lineages designated as AY and BA in the Pango nomenclature system, respectively. In general, California and Mexico exhibited notable differences in lineage distribution patterns (Figure 1 and Figure S4 and related MicroReact projects (links above) for Delta and Omicron, respectively). For Delta, California was dominated by the AY.44 and AY.103 lineages, with relative frequencies of 22.7% and 20.1%, respectively, while the AY.20 (45.15%) and AY.26 (23.2%) lineages were most prevalent in Mexico. Interestingly, the epidemiology in the state of Baja California (Mexico) was more similar to southern California (US), with a higher prevalence of AY.103 and a lower relative frequency of AY.20 and AY.26 compared to other Mexican states. The distribution of

Omicron lineages was more similar between California and Mexico, with a predominance of BA1, BA1.1, and BA1.15.

3.3. Viral Dynamics within California and Mexico

To interrogate viral migration patterns in the California–Mexico border region, we used a mean of 32 clades of size ≥ 5 including ≥ 2 locations within our study region for Delta, and 10 clades of size ≥ 5 including ≥ 2 locations for Omicron.

Delta variant. Of the ten replicated discrete phylogeographic analyses, four showed the western state of Hidalgo and central states of Mexico City and Jalisco as a source of variant dispersal within Mexico. There was no clear predominant hub of dispersal observed in California (Figures 3A and 4A; see also Supplementary Figures S6A and S7A for the results of each replicate). The variation across the ten replicates was likely due to the limited number of sequences included in each subset analysis, and is further discussed below. To analyze the viral diffusion processes more finely, we interrogated the density of variant migrations over time (Figures 5A and S8A). Delta variant domestic migration events peaked in mid-summer 2021 and then progressively declined through the end of 2021 when Omicron was first identified. This coincided with the intensity of external introduction events.

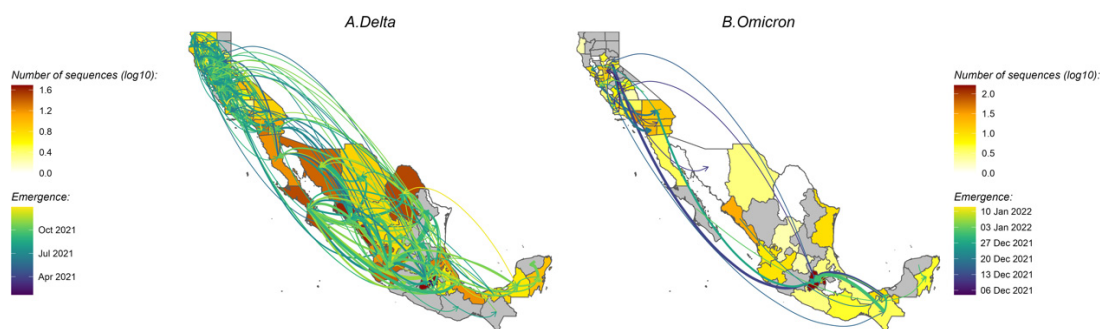


Figure 3. Map of migration events between locations for Delta (A) and Omicron (B) variants. The thickness of the arrows reflects the average number of inferred transmission events between locations and the colors of the arrows indicate the timing of the transmission event. California counties and Mexico states are colored according to the number of sequences. All 10 replicates are combined to show the timing of the introductions of the clades in the region. See also Supplementary Figure S6 for the results of each of the 10 replicates. Only the migration events with strong support ($BF \geq 20$) are shown.

Omicron variant. We again found that the city and state of Mexico (center south region) were important hubs of dispersal within Mexico (Figures 3B and 4B; see also Supplementary Figures S6B and S7B). Similar to what we observed for Delta, the timing and intensity of Omicron local migrations and external variant introductions coincided and contributed simultaneously to the spatial dispersion of Omicron in the California–Mexico region (Figures 5B and S8B) in December 2021.

3.4. Variant Migration across the California–Mexico Border

To evaluate the contribution of cross-border variant migration events (i.e., migration from California toward Mexico or from Mexico toward California) to the global dispersal of the Delta and Omicron epidemics, we evaluated the relative contribution of domestic and cross-border viral migrations.

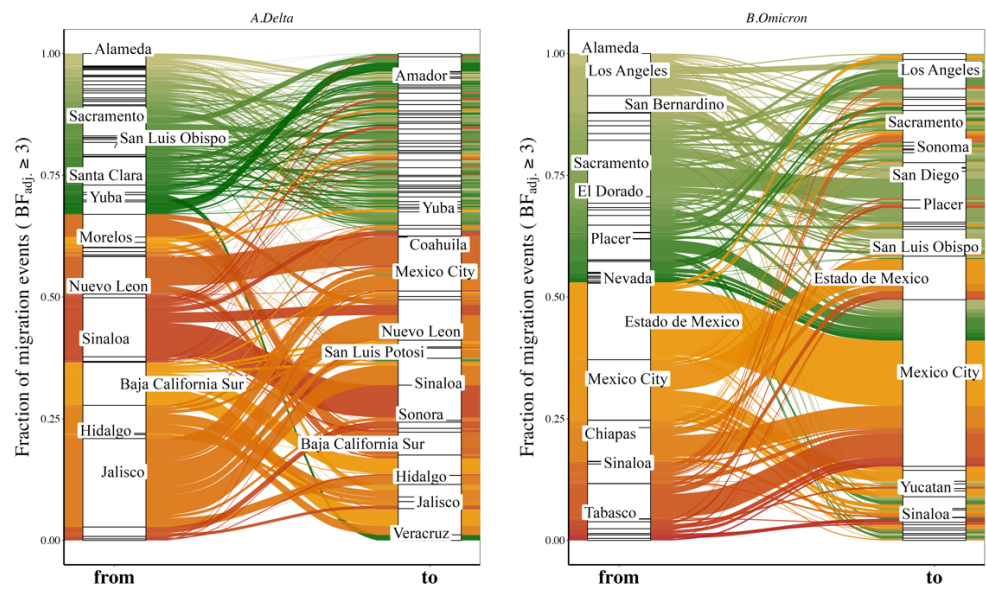


Figure 4. Proportion of viral migration events from each source location (“from”) toward the recipient location (“to”) for Delta (A) and Omicron (B) variants. The Sankey plot shows the proportion of viral migration events from each source location (Californian county or Mexican state) toward the recipient location. Left side of the plot shows the source location (“from”) and right side of the plot shows the destination location (“to”). We only report migration events associated with an adjusted Bayes factor ($BF_{adj} \geq 3$). See also Supplementary Figure S7 for results of each of the 10 individual replicates. Locations accounting for $\geq 2\%$ of all events are labeled.

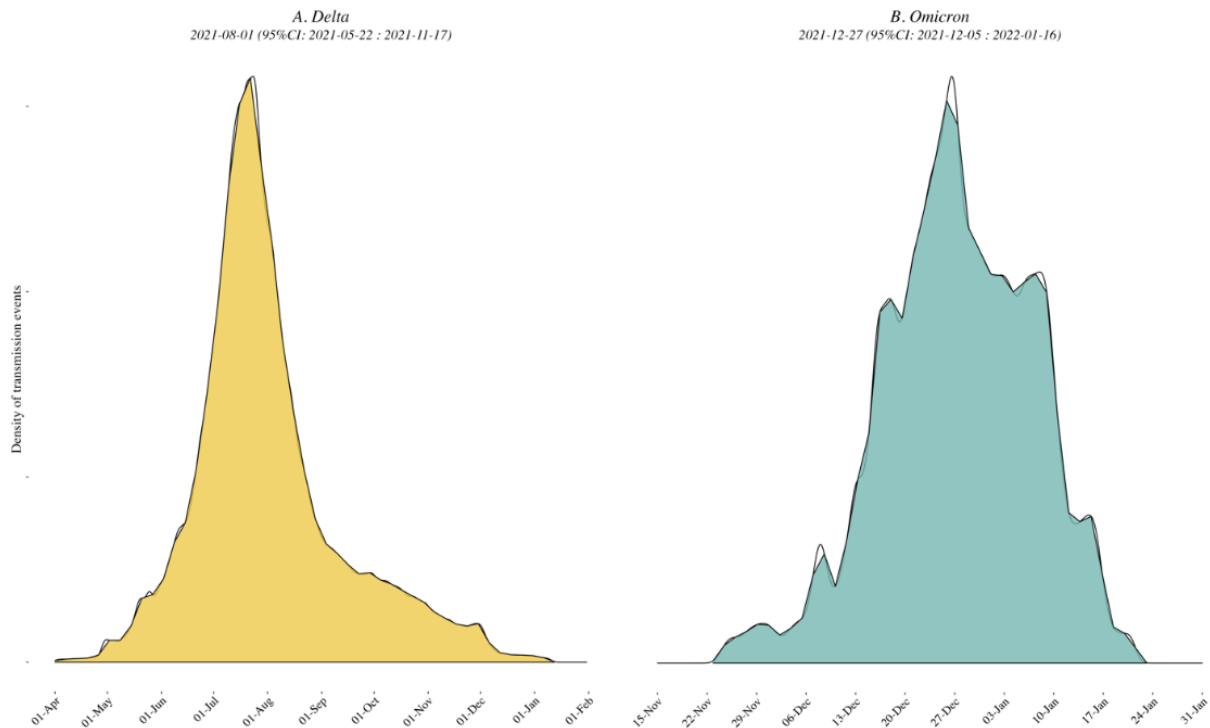


Figure 5. Probability density of variant migration events between locations for Delta (A) and Omicron (B) variants. Only migration events supported with an adjusted Bayes Factor ≥ 3 are considered. See also Supplementary Figure S8 for results of each of the 10 individual replicates.

Delta variant. During the Delta epidemic in 2021, the majority of viral movement was domestic, with a small amount of international migration in the border region with cross-border events accounting for 5.1% [95% CI: 4.31–6] of all inferred Delta variant migration events (mean of 0.92% [95% CI: 0.59–1] from California toward Mexico and 4.18% [95% CI: 3.58–5] from Mexico toward California, Figure 6A). Longitudinally, we showed that these cross-border transmissions were detected as early as May 2021 but remained limited with no peak through the study period (Supplementary Figure S9A).

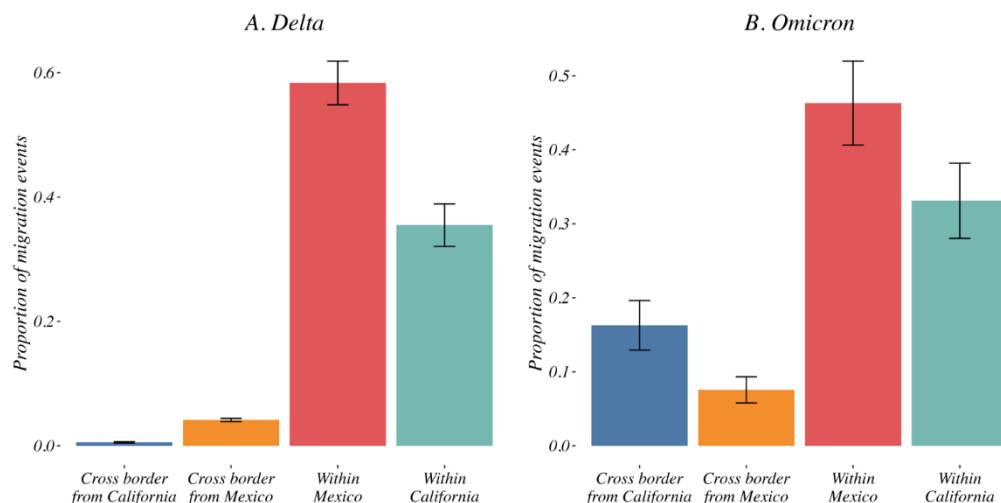


Figure 6. Relative contribution of national vs. international transmission events to the Delta (A) and Omicron (B) epidemics. The variant migration events across the border (from California toward Mexico and from Mexico toward California in blue and orange respectively) accounted for 3.4% and 23% of all inferred migration events for Delta and Omicron, respectively. Migration events within California and within Mexico are shown in red and green. See Supplementary Figure S9 for a description of these contributions over time.

Omicron variant. Distinct from what we observed for the Delta variant, we identified a high proportion of cross-border events for Omicron (from California toward Mexico and from Mexico toward California). We observed within the first weeks of December a sudden increase in cross-border viral migration accounting for 20.58% [95% CI: 12.38–29] of all inferred events (mean of 16.28% [95% CI: 8.37–24] from California toward Mexico and 7.56% [95% CI: 3.55–12] from Mexico toward California, Figure 6B). These cross-border migration events peaked in the last weeks of December 2021 (Supplementary Figure S9B).

3.5. Human Mobility and Variant Dispersal

To interrogate the role of human mobility, we examined the relationship between geospatial distance and migration events. As expected, we observed a negative correlation between distance and the intensity of transmission events (i.e., more events between close locations), and this association between proximity and dispersal was observed for both the Delta and Omicron variants ($p = 0.01$ and 0.06 , respectively). Interestingly, while overall long-distance dispersal events were relatively infrequent (Supplementary Figure S10), they were more frequent for Omicron than for Delta, consistent with the inferences from our Bayesian models that showed a higher proportion of cross-border migration events for Omicron (Figure 6).

To assess the association between human mobility and cross-border variant movement, we analyzed Facebook international travel mobility data [28]. The density of variant migration events from California toward Mexico (green line) and from Mexico toward California (red line) increased as international travel increased in June–July 2021 and December 2021 (Supplementary Figure S11). While the increase in cross-border travel observed in June–July 2021 was not associated with significant cross-border variant movement for the

Delta epidemic, the sharp increase in cross-border Omicron movement in late 2021 occurred when the number of international travel events peaked in December.

4. Discussion

Our phylogenetic investigations of the Delta and Omicron epidemics in the California/Mexico region illustrate a complex interplay between viral and structural forces for viral spread. The objective of our first phylogeographic analysis was to delineate clades corresponding to distinct viral introductions of SARS-CoV-2 Delta and Omicron lineages into the California/Mexico region. We found evidence of numerous external viral introductions of Delta and Omicron clades in the region throughout the study period. Of note, the probability density distribution of introduction events was highly negatively skewed from the peak for Omicron (-1.90) as compared to Delta (0.64 ; see Supplementary Table S3) with a high amount of kurtosis (13.8816) as compared to Delta (3.57), suggesting that the initial introduction of Omicron may have occurred much earlier than the estimated date of introduction, or similarly that lesser (limited propagation) introductions may have been occurring prior to the estimated peak introduction date. Discrete phylogeographic analyses of these clades demonstrated their contribution to the ongoing local circulation of Delta and Omicron in the region. Our results suggest that transmission chains circulating in California and Mexico were not established by a restricted number of isolated infectious cases but rather a combination of multiple external viral introductions, with cross-border and domestic variant movements occurring at the same time (Figures 2 and 5). This observation is in line with other reports analyzing the first waves of the epidemic where no predominant lineage was identified [8,30]. However, this is distinct from what was seen by Kraemer et al. in the UK, where London acted as a dominant source of dispersal of B.1.1.7 VOC [10], or in the USA, where New York was considered the location of emergence and expansion of VOC B.1.526 [8,31]. These differences likely reflect the fact that these two variants were first detected in London and New York and these cities acted as hubs. These results are consistent with a classic gravity model framework where large metropolitan areas act as the main transmission hub in an epidemic. Similarly, our analyses of the Omicron variant within Mexico showed that the city and state of Mexico (center south region) were important hubs of dispersal within Mexico. Our results were somewhat different for the Delta epidemic; while Mexico City and state were important, they were not singular predominant sources of viral dispersal. Rather, the epidemic was driven by a combination of repeated viral introductions from outside the region together with the circulation of multiple clades across Mexico.

We found that in the Mexican state of Baja California, the distribution of Delta lineages was more similar to southern California than central Mexico, highlighting the importance of proximity in the diffusion of these SARS-CoV-2 variants and suggesting cross-border transmission. Concerning Omicron, we noted a lower diversity and a more heterogeneous distribution of lineages across California and Mexico, with BA.1.1, BA.1 and BA.1.15 being the most prevalent. This may reflect the more recent emergence and introduction of Omicron compared to Delta.

To untangle the factors leading to this complex epidemiology, we interrogated the contribution of domestic and international variant migration events. We found that cross-border viral movement contributed marginally to the Delta epidemic in mid-2021 (5.1% [95% CI: 4.31–6] of inferred migration events), while nearly a quarter of all inferred variant migration events identified for Omicron in late 2021 were cross-border. A small positive skewness was observed for Delta, possibly suggesting an ongoing trickle of migration during the Delta wave. The observed differences in the spatial invasion of the two major VOCs may be explained by a combination of viral and structural/behavioral factors. The shorter duration of the Omicron wave and time to peak migration may be a function of the increased transmissibility of the Omicron variant compared to the Delta variant, which may be a function of the increased affinity of the Omicron spike to the ACE2 receptor or overall hydrophobicity of the Omicron spike [32–34]. In addition to differences in

variant characteristics (such as per-contact transmissibility, duration of infectiousness, and immune evasion [35,36]), local circulation and variant replacement dynamics are likely affected by human mobility but also by spatiotemporal heterogeneity in various structural factors (e.g., access and coverage of vaccinations) that need to be considered. Previous studies have shown that mobility data can be used to predict the spread of SARS-CoV2 [29,37,38] and other viruses [39]. In early phases of the COVID-19 pandemic, travel bans were effective in preventing COVID-19 importation [40,41]. Here, we interrogated the association between international viral migration and international travel using Facebook travel mobility data [28]. The increase in travel across the USA/Mexico border during the summer (Supplementary Figure S10) was not reflected in the cross-border migration of Delta. On the other hand, the emergence of the highly transmissible Omicron occurred in late 2021 when non-essential travel resumed for fully vaccinated foreign travelers to the US (effective as of 8 November 2021 [42], while no restrictions were in place for entry into Mexico). This timing also corresponded with the holiday travel season, and the Facebook data reflected these differences. In the case of Omicron, this increase in cross-border mobility was associated with significant cross-border viral migration.

Other structural factors, such as the heterogeneity of vaccination coverage between locations in 2021 [43] could have also impacted viral dynamics. As of 1 June 2021, when we observed a rise of Delta variant introductions and variant migrations between sampled locations, 55% of the population in California received at least one dose (46.9% fully vaccinated) while only 16.9% in Mexico (9.7% fully vaccinated) had received one dose [44]. At the time of the upsurge of Omicron cases in early December, vaccination coverage was 71.8% (64.4% fully vaccinated) in California and 59.1% (50% fully vaccinated) in Mexico. While this may appear counterintuitive, the intrinsic characteristics of Omicron [35,36] may be responsible for the rapid spread of Omicron in late 2021 despite higher rates of vaccination coverage.

An important limitation of this observational study is that these analyses are highly dependent on the data available. While the volume of genomic data available presents a unique opportunity to gain valuable real-time insight into viral dynamics, it also raises major computational challenges and potential biases. Failing to correct for ecologic and sampling bias can potentially lead to artifactual outcomes in phylogeographic reconstructions [45,46]. Here, we proposed a subsampling strategy informed by epidemiological data [8] to mitigate sampling bias. The lack of full congruence among the ten replicates that we observed was expected and is likely due to the relatively limited number of samples considered in each replicated analysis, as shown in other settings [8]. Random subsampling selection can also lead to different dispersal inferences, illustrating the impact of the sampled phylogenetic diversity in shaping the inferred dispersal history, as previously shown by Dellicour et al. in New York [8]. However, consistently across all replicates and for both Delta and Omicron, we showed a complex interplay of external introductions and domestic viral dispersal. With regard to our cross-border analyses, we relied on Facebook data which only note travel between countries and may not be totally representative of travel specifically between California and Mexico. Furthermore, these mobility data may not be representative, as Facebook use may vary geographically and demographically.

Our investigations of the Delta and Omicron epidemics in the California/Mexico region illustrate the complex interplay and the multiplicity of viral and structural factors that need to be considered to limit viral spread, even as vaccination is reducing disease burden. While our phylodynamic analysis provided insight into patterns of viral introductions into the region and migration across the border, future analysis using compartmental ordinary differential equation models could provide additional insight on how changes in transmission rates, vaccination rates and flow across the border could impact these and future SARS-CoV-2 variant epidemics in the border region [47–55]. Given the rapid observed diffusion of Delta, and then Omicron, as the infectiousness of successive variants increases, containment strategies would need to be increasingly extreme to reduce spread. This recent history suggests that containment approaches are unlikely to be successful for

any further waves of the SARS-CoV-2 pandemic. Benefitting from our recent experience with SARS-CoV-2, we are undoubtedly better prepared to face the emergence of potential new viral pandemics, but untangling the driving forces of future epidemics and balancing the societal risks and benefits of containment vs. mitigation strategies will always remain a challenge.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/v14071494/s1>, Figure S1: Map of publicly available SARS-CoV-2 sequences; Figure S2: Number of COVID-19 confirmed cases and full-length genome sequences by location; Figure S3: Map of COVID-19 cases per 100,000 individuals up to 31 January 2022 [56]; Figure S4: Visualization of Delta (A) and Omicron (B) genomes using MicroReact; Figure S5: Density plots of external introduction of the identified clades for delta (Panel A) and omicron (Panel B) variants for the 10 replicates; Figure S6: Map of migration events between location for Delta (Panel A) and Omicron (Panel B) variants; Figure S7: Proportion of viral migration events from each source location ('from') toward the recipient location ('to') for Delta (Panel A) and Omicron (Panel B) variants; Figure S8: Probability Density of variant migration events between locations for Delta (Panel A) and Omicron (Panel B) variants; Figure S9: Relative contribution of national vs international transmission events to the Delta (Panel A) and Omicron (Panel B) epidemics overtime; Figure S10: Distribution of the geographic distances of phylogenetic lineage movement events for Delta (Panel A) and Omicron (Panel B) variants; Figure S11: Density of viral migration events across the USA Mexico border and international travels; Table S1: Distribution of full-length SARS-CoV2 genomes by county sampled in California and by state sampled in Mexico in 2021–2022. Table S2: Sampling overview; Table S3: Skewness and kurtosis of the distribution of the external introduction events and viral migrations for Omicron and Delta variants.

Author Contributions: A.C. and S.R.M. designed the study, performed the analyses, and wrote the manuscript. D.M.S. and C.B. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by grants from the NIH (San Diego Center for AIDS Research, CFAR, AI306214 and AI100665), the Department of Veterans Affairs, the John and Mary Tu Foundation, and the James B. Pendleton Charitable Trust. A.C. was also supported by NIH Grants DA049644, AI145555, MH128153 (R01), AI106039 and DP2 CA051915. S.R.M. was also supported by NIH Grants AI106039, MH124590, AI135992, DA049644, and the PIMSA foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data Availability Statement: Data can be visualized with MicroReact projects <https://microreact.org/project/rqnwd9nymcvdkvq1dktbfj> and <https://microreact.org/project/d5mbllhkwgxtofu6wdtxt>.

Acknowledgments: We gratefully acknowledge the authors and the originating and submitting laboratories of the genetic sequence and metadata made available through GISAID on which this research is based. A full acknowledgement table can be found with the following EPI_SET-IDs, EPI_SET_20220517tr, EPI_SET_20220517xz, EPI_SET_20220517au and EPI_SET_20220517zh in Data Acknowledgement Locator under GISAID resources (<https://www.gisaid.org/> accessed on 12 February 2022). Facebook mobility data are not publicly available owing to stringent licensing agreements [28]. Facebook mobility data are made available through the Data for Good programme (<https://dataforgood.fb.com> accessed on 5 March 2022) under the terms of a data license agreement that defines the allowed terms of use by partners (contact: disastermaps@fb.com). We are also grateful to Santiago Avila Rios for his assistance in writing the manuscript.

Conflicts of Interest: The authors have no financial interest to report. D.M.S. has consulted for Linear Therapies, Model Medicines, Vx Biosciences, Bayer, Kiadis Pharmaceuticals, Signant Health and FluxErgy. S.R.M. has consulted for MedialEarlySign.

References

1. Worobey, M.; Pekar, J.; Larsen, B.B.; Nelson, M.I.; Hill, V.; Joy, J.B.; Rambaut, A.; Suchard, M.A.; Wertheim, J.O.; Lemey, P. The emergence of SARS-CoV-2 in Europe and North America. *Science* **2020**, *370*, 564–570. [[CrossRef](#)] [[PubMed](#)]
2. Lu, J.; du Plessis, L.; Liu, Z.; Hill, V.; Kang, M.; Lin, H.; Sun, J.; François, S.; Kraemer, M.U.G.; Faria, N.R.; et al. Genomic Epidemiology of SARS-CoV-2 in Guangdong Province, China. *Cell* **2020**, *181*, 997–1003.e9. [[CrossRef](#)] [[PubMed](#)]

3. Cyranoski, D. Alarming COVID variants show vital role of genomic surveillance. *Nature* **2021**, *589*, 337–338. [[CrossRef](#)]
4. Shephardson, D. U.S. Extends Travel Curbs at Canada, Mexico Land Borders through 21 August. Available online: <https://www.reuters.com/world/us/us-extends-travel-restrictions-canada-mexico-land-borders-through-aug-21-2021-07-21/> (accessed on 15 February 2022).
5. Lambrou, A.S.; Shirk, P.; Steele, M.K.; Paul, P.; Paden, C.R.; Cadwell, B.; Reese, H.E.; Aoki, Y.; Hassell, N.; Zheng, X.Y.; et al. Genomic Surveillance for SARS-CoV-2 Variants: Predominance of the Delta (B.1.617.2) and Omicron (B.1.1.529) Variants—United States, June 2021–January 2022. *MMWR Morb. Mortal. Wkly. Rep.* **2022**, *71*, 206–211. [[CrossRef](#)]
6. Nishiura, H.; Ito, K.; Anzai, A.; Kobayashi, T.; Piantham, C.; Rodriguez-Morales, A.J. Relative Reproduction Number of SARS-CoV-2 Omicron (B.1.1.529) Compared with Delta Variant in South Africa. *J. Clin. Med.* **2021**, *11*, 30. [[CrossRef](#)] [[PubMed](#)]
7. Du Plessis, L.; McCrone, J.T.; Zarebski, A.E.; Hill, V.; Ruis, C.; Gutierrez, B.; Raghwan, J.; Ashworth, J.; Colquhoun, R.; Connor, T.R.; et al. Establishment and lineage dynamics of the SARS-CoV-2 epidemic in the UK. *Science* **2021**, *371*, 708–712. [[CrossRef](#)]
8. Dellicour, S.; Hong, S.L.; Vrancken, B.; Chaillon, A.; Gill, M.S.; Maurano, M.T.; Ramaswami, S.; Zappile, P.; Marier, C.; Harkins, G.W.; et al. Dispersal dynamics of SARS-CoV-2 lineages during the first epidemic wave in New York City. *PLoS Pathog.* **2021**, *17*, e1009571. [[CrossRef](#)]
9. Dellicour, S.; Durkin, K.; Hong, S.L.; Vanmechelen, B.; Martí-Carreras, J.; Gill, M.S.; Meex, C.; Bontems, S.; André, E.; Gilbert, M.; et al. A Phylodynamic Workflow to Rapidly Gain Insights into the Dispersal History and Dynamics of SARS-CoV-2 Lineages. *Mol. Biol. Evol.* **2021**, *38*, 1608–1613. [[CrossRef](#)]
10. Kraemer, M.U.G.; Hill, V.; Ruis, C.; Dellicour, S.; Bajaj, S.; McCrone, J.T.; Baele, G.; Parag, K.V.; Battle, A.L.; Gutierrez, B.; et al. Spatiotemporal invasion dynamics of SARS-CoV-2 lineage B.1.1.7 emergence. *Science* **2021**, *373*, 889–895. [[CrossRef](#)]
11. Rambaut, A.; Pybus, O.G.; Nelson, M.I.; Viboud, C.; Taubenberger, J.K.; Holmes, E.C. The genomic and epidemiological dynamics of human influenza A virus. *Nature* **2008**, *453*, 615–619. [[CrossRef](#)]
12. Shu, Y.; McCauley, J. GISAID: Global initiative on sharing all influenza data—From vision to reality. *Eur. Surveill. Eur. Commun. Dis. Bull.* **2017**, *22*, 30494. [[CrossRef](#)] [[PubMed](#)]
13. United States Department of Transportation Bureau of Transportation Statistics (BTS). Available online: <https://www.bts.gov/topics/national-transportation-statistics> (accessed on 1 March 2022).
14. Hadfield, J.; Megill, C.; Bell, S.M.; Huddleston, J.; Potter, B.; Callender, C.; Sagulenko, P.; Bedford, T.; Neher, R.A. Nextstrain: Real-time tracking of pathogen evolution. *Bioinformatics* **2018**, *34*, 4121–4123. [[CrossRef](#)] [[PubMed](#)]
15. Argimón, S.; Abudahab, K.; Goater, R.J.E.; Fedosejev, A.; Bhai, J.; Glasner, C.; Feil, E.J.; Holden, M.T.G.; Yeats, C.A.; Grundmann, H.; et al. Microreact: Visualizing and sharing data for genomic epidemiology and phylogeography. *Microb. Genom.* **2016**, *2*, e000093. [[CrossRef](#)] [[PubMed](#)]
16. USAFacts Understanding the COVID-19 Pandemic. Available online: <https://usafacts.org/issues/coronavirus/> (accessed on 13 February 2022).
17. Government of Mexico COVID-19 Data for Mexico. Available online: <https://datos.covid-19.conacyt.mx/#DOView> (accessed on 13 February 2022).
18. Rambaut, A.; Holmes, E.C.; O’Toole, Á.; Hill, V.; McCrone, J.T.; Ruis, C.; du Plessis, L.; Pybus, O.G. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat. Microbiol.* **2020**, *5*, 1403–1407. [[CrossRef](#)] [[PubMed](#)]
19. Cuypers, L.; Vrancken, B.; Fabeni, L.; Marascio, N.; Cento, V.; Di Maio, V.C.; Aragri, M.; Pineda-Pena, A.C.; Schrooten, Y.; Van Laethem, K.; et al. Implications of hepatitis C virus subtype 1a migration patterns for virus genetic sequencing policies in Italy. *BMC Evol. Biol.* **2017**, *17*, 70. [[CrossRef](#)] [[PubMed](#)]
20. Price, M.N.; Dehal, P.S.; Arkin, A.P. FastTree 2—Approximately maximum-likelihood trees for large alignments. *PLoS ONE* **2010**, *5*, e9490. [[CrossRef](#)] [[PubMed](#)]
21. Volz, E.M.; Frost, S.D.W. Scalable relaxed clock phylogenetic dating. *Virus Evol.* **2017**, *3*, vex025. [[CrossRef](#)]
22. Lemey, P.; Rambaut, A.; Drummond, A.J.; Suchard, M.A. Bayesian phylogeography finds its roots. *PLoS Comput. Biol.* **2009**, *5*, e1000520. [[CrossRef](#)]
23. Suchard, M.A.; Lemey, P.; Baele, G.; Ayres, D.L.; Drummond, A.J.; Rambaut, A. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **2018**, *4*, vey016. [[CrossRef](#)]
24. Rambaut, A.; Drummond, A.J.; Xie, D.; Baele, G.; Suchard, M.A. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst. Biol.* **2018**, *67*, 901–904. [[CrossRef](#)]
25. Minin, V.N.; Bloomquist, E.W.; Suchard, M.A. Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. *Mol. Biol. Evol.* **2008**, *25*, 1459–1471. [[CrossRef](#)] [[PubMed](#)]
26. Chaillon, A.; Gianella, S.; Dellicour, S.; Rawlings, S.A.; Schlub, T.E.; Faria De Oliveira, M.; Ignacio, C.; Porrachia, M.; Vrancken, B.; Smith, D.M. HIV persists throughout deep tissues with repopulation from multiple anatomical sources. *J. Clin. Investig.* **2020**, *130*, 1699–1712. [[CrossRef](#)] [[PubMed](#)]
27. Kass, R.E.; Raftery, A.E. Bayes Factors. *J. Am. Stat. Assoc.* **1995**, *90*, 773–795. [[CrossRef](#)]
28. Maas, P. Facebook Disaster Maps: Aggregate Insights for Crisis Response & Recovery. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA, 4–8 August 2019; p. 3173.
29. Lemey, P.; Ruktanonchai, N.; Hong, S.L.; Colizza, V.; Poletto, C.; Van den Broeck, F.; Gill, M.S.; Ji, X.; Lévassieur, A.; Oude Munnink, B.B.; et al. Untangling introductions and persistence in COVID-19 resurgence in Europe. *Nature* **2021**, *595*, 713–717. [[CrossRef](#)]

30. Deng, X.; Gu, W.; Federman, S.; du Plessis, L.; Pybus, O.G.; Faria, N.; Wang, C.; Yu, G.; Pan, C.-Y.; Guevara, H.; et al. A Genomic Survey of SARS-CoV-2 Reveals Multiple Introductions into Northern California without a Predominant Lineage. *Science* **2020**, *369*, 582–587. [[CrossRef](#)]
31. Annavajhala, M.K.; Mohri, H.; Wang, P.; Nair, M.; Zucker, J.E.; Sheng, Z.; Gomez-Simmonds, A.; Kelley, A.L.; Tagliavia, M.; Huang, Y.; et al. Emergence and expansion of SARS-CoV-2 B.1.526 after identification in New York. *Nature* **2021**, *597*, 703–708. [[CrossRef](#)]
32. Kumar, S.; Thambiraja, T.S.; Karuppanan, K.; Subramaniam, G. Omicron and Delta variant of SARS-CoV-2: A comparative computational study of spike protein. *J. Med. Virol.* **2022**, *94*, 1641–1649. [[CrossRef](#)]
33. McCallum, M.; Czudnochowski, N.; Rosen, L.E.; Zepeda, S.K.; Bowen, J.E.; Walls, A.C.; Hauser, K.; Joshi, A.; Stewart, C.; Dillen, J.R.; et al. Structural basis of SARS-CoV-2 Omicron immune evasion and receptor engagement. *Science* **2022**, *375*, 864–868. [[CrossRef](#)]
34. Shishir, T.A.; Jannat, T.; Naser, I.B. An in-silico study of the mutation-associated effects on the spike protein of SARS-CoV-2, Omicron variant. *PLoS ONE* **2022**, *17*, e0266844. [[CrossRef](#)]
35. Medigeshi, G.; Batra, G.; Murugesan, D.R.; Thiruvengadam, R.; Chattopadhyay, S.; Das, B.; Gosain, M.; Ayushi; Singh, J.; Anbalagan, A.; et al. Sub-optimal Neutralisation of Omicron (B.1.1.529) Variant by Antibodies induced by Vaccine alone or SARS-CoV-2 Infection plus Vaccine (Hybrid Immunity) post 6-months. *eBioMedicine* **2022**, *78*, 103938. [[CrossRef](#)]
36. Ai, J.; Zhang, H.; Zhang, Y.; Lin, K.; Zhang, Y.; Wu, J.; Wan, Y.; Huang, Y.; Song, J.; Fu, Z.; et al. Omicron variant showed lower neutralizing sensitivity than other SARS-CoV-2 variants to immune sera elicited by vaccines after boost. *Emerg. Microbes Infect.* **2022**, *11*, 337–343. [[CrossRef](#)] [[PubMed](#)]
37. Kraemer, M.U.G.; Yang, C.H.; Gutierrez, B.; Wu, C.H.; Klein, B.; Pigott, D.M.; du Plessis, L.; Faria, N.R.; Li, R.; Hanage, W.P.; et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* **2020**, *368*, 493–497. [[CrossRef](#)] [[PubMed](#)]
38. Hodcroft, E.B.; Zuber, M.; Nadeau, S.; Vaughan, T.G.; Crawford, K.H.D.; Althaus, C.L.; Reichmuth, M.L.; Bowen, J.E.; Walls, A.C.; Corti, D.; et al. Emergence and spread of a SARS-CoV-2 variant through Europe in the summer of 2020. *Nature* **2021**, *595*, 707–712. [[CrossRef](#)] [[PubMed](#)]
39. Lemey, P.; Rambaut, A.; Bedford, T.; Faria, N.; Bielejec, F.; Baele, G.; Russell, C.A.; Smith, D.J.; Pybus, O.G.; Brockmann, D. Unifying viral genetics and human transportation data to predict the global transmission dynamics of human influenza H3N2. *PLoS Pathog.* **2014**, *10*, e1003932. [[CrossRef](#)] [[PubMed](#)]
40. Gwee, S.X.W.; Chua, P.E.Y.; Wang, M.X.; Pang, J. Impact of travel ban implementation on COVID-19 spread in Singapore, Taiwan, Hong Kong and South Korea during the early phase of the pandemic: A comparative study. *BMC Infect. Dis.* **2021**, *21*, 799. [[CrossRef](#)]
41. Burns, J.; Movsisyan, A.; Stratil, J.M.; Biallas, R.L.; Coenen, M.; Emmert-Fees, K.M.; Geffert, K.; Hoffmann, S.; Horstick, O.; Laxy, M.; et al. International travel-related control measures to contain the COVID-19 pandemic: A rapid review. *Cochrane Database Syst. Rev.* **2021**, *3*, CD013717. [[PubMed](#)]
42. US Homeland Security Guidance for Travelers to Enter the U.S. at Land Ports of Entry and Ferry Terminals. Available online: <https://www.dhs.gov/news/2021/10/29/fact-sheet-guidance-travelers-enter-us-land-ports-entry-and-ferry-terminals> (accessed on 13 February 2022).
43. Diesel, J.; Sterrett, N.; Dasgupta, S.; Kriss, J.L.; Barry, V.; Vanden Esschert, K.; Whiteman, A.; Cadwell, B.L.; Weller, D.; Qualters, J.R.; et al. COVID-19 Vaccination Coverage among Adults—United States, 14 December 2020–22 May 2021. *MMWR Morb. Mortal. Wkly. Rep.* **2021**, *70*, 922–927. [[CrossRef](#)]
44. Reuters.com COVID-19 Vaccination Tracker. Available online: <https://graphics.reuters.com/world-coronavirus-tracker-and-maps/vaccination-rollout-and-access/> (accessed on 19 February 2022).
45. De Maio, N.; Wu, C.-H.; O’Reilly, K.M.; Wilson, D. New Routes to Phylogeography: A Bayesian Structured Coalescent Approximation. *PLoS Genet.* **2015**, *11*, e1005421. [[CrossRef](#)]
46. Kalkauskas, A.; Perron, U.; Sun, Y.; Goldman, N.; Baele, G.; Guindon, S.; De Maio, N. Sampling bias and model choice in continuous phylogeography: Getting lost on a random walk. *PLoS Comput. Biol.* **2021**, *17*, e1008561. [[CrossRef](#)]
47. Spiliotis, K.; Koutsoumaris, C.C.; Reppas, A.I.; Papaxenopoulou, L.A.; Starke, J.; Hatzikirou, H. Optimal vaccine roll-out strategies including social distancing for pandemics. *iScience* **2022**, *25*, 104575. [[CrossRef](#)]
48. Beira, M.J.; Sebastião, P.J. A differential equations model-fitting analysis of COVID-19 epidemiological data to explain multi-wave dynamics. *Sci. Rep.* **2021**, *11*, 16312. [[CrossRef](#)] [[PubMed](#)]
49. El Koufi, A.; El Koufi, N. Stochastic differential equation model of Covid-19: Case study of Pakistan. *Results Phys.* **2022**, *34*, 105218. [[CrossRef](#)] [[PubMed](#)]
50. Ahmad, Z.; El-Kafrawy, S.A.; Alandijany, T.A.; Giannino, F.; Mirza, A.A.; El-Daly, M.M.; Faizo, A.A.; Bajrai, L.H.; Kamal, M.A.; Azhar, E.I. A global report on the dynamics of COVID-19 with quarantine and hospitalization: A fractional order model with non-local kernel. *Comput. Biol. Chem.* **2022**, *98*, 107645. [[CrossRef](#)] [[PubMed](#)]
51. Sintunavarat, W.; Turab, A. Mathematical analysis of an extended SEIR model of COVID-19 using the ABC-fractional operator. *Math. Comput. Simul.* **2022**, *198*, 65–84. [[CrossRef](#)]
52. Majee, S.; Adak, S.; Jana, S.; Mandal, M.; Kar, T.K. Complex dynamics of a fractional-order SIR system in the context of COVID-19. *J. Appl. Math. Comput.* **2022**, 1–24. [[CrossRef](#)]

53. Boonpatcharanon, S.; Heffernan, J.M.; Jankowski, H. Estimating the basic reproduction number at the beginning of an outbreak. *PLoS ONE* **2022**, *17*, e0269306. [[CrossRef](#)]
54. Mwalili, S.; Kimathi, M.; Ojiambo, V.; Gathungu, D.; Mbogo, R. SEIR model for COVID-19 dynamics incorporating the environment and social distancing. *BMC Res. Notes* **2020**, *13*, 352. [[CrossRef](#)]
55. Kucharski, A.J.; Klepac, P.; Conlan, A.J.K.; Kissler, S.M.; Tang, M.L.; Fry, H.; Gog, J.R.; Edmunds, W.J. Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: A mathematical modelling study. *Lancet Infect. Dis.* **2020**, *20*, 1151–1160. [[CrossRef](#)]
56. Dong, E.; Du, H.; Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **2020**, *20*, 533–534. [[CrossRef](#)]