medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Genome evolution and early introductions of the SARS-CoV-2 Omicron Variant in Mexico

Hugo G. Castelán-Sánchez[1], León P. Martínez-Castilla[1], Gustavo Sganzerla-Martínez[2,3], Jesús Torres-Flores[1], Gamaliel López-Leal[1†]

[1] Programa de Investigadoras e Investigadores por México. Grupo de Genómica y Dinámica Evolutiva de Microorganismos Emergentes. Consejo Nacional de Ciencia y Tecnología. Av. Insurgentes Sur 1582, Crédito Constructor, Benito Juárez, CP 03940, Ciudad de México, México.

[2] Laboratory of Immunity, Shantou University Medical College, Shantou, People's Republic of China

[3] The Department of Microbiology and Immunology Dalhousie University, Halifax, NS, Canada

Hugo G Castelán-Sánchez hugo.castelan@conacyt.mx

León P. Martinez-Castilla leon.martinez@conacyt.mx

Gustavo Sganzerla Martínez gustavo.sganzerla@dal.ca

Jesús Torres-Flores jesus.torres@conacyt.mx

Gamaliel Lopez-Leal gamaliel.lopez@conacyt.mx

† These authors contributed equally to this work

* Correspondence to: Hugo G. Castelán Sánchez hugo.castelan@conacyt.mx

Keywords

Omicron, introductions, natural selection, human mobility

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

ABSTRACT

A new variant of SARS-CoV-2 Omicron (Pango lineage designation B.1.1.529), was first reported to the World Health Organization (WHO) by South African health authorities on November 24, 2021. The Omicron variant possesses numerous mutations associated with increased transmissibility and immune escape properties. In November 2021, Mexican authorities reported Omicron's presence in the country. In this study, we infer the first introductory events of Omicron and the impact that human mobility can have on the spread of the virus. We also evaluated the adaptive evolutionary processes in Mexican SARS-CoV-2 genomes during the first month of circulation of Omicron.

We infer 173 introduction events of Omicron in Mexico in the first two months of detection; subsequently, of the introductions, there was an increase in the prevalence for January.

This higher prevalence of the novel variant results in a peak of cases reported, on average, six weeks after a higher mobility trend was reported. The peak of cases reported is due to the BA.1.1 Omicron sub-lineage dominated, followed by BA.1 and BA.15 sub-lineages in the country from January to February 2022.

Additionally, we identified the presence of diversifying natural selection in the genomes of Omicron and found mainly five non-synonymous mutations in the RDB domain of the Spike protein, all of them related to evasion of the immune response. In contrast, the other proteins in the genome are highly conserved—however, there are homoplasies mutations in non-structural proteins, indicating a parallel evolution.

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

54

55 1. INTRODUCTION

56 The new coronavirus SARS-CoV-2, the causative agent of coronavirus disease 2019
57 (COVID-19), emerged at the end of 2019 in the province of Wuhan, China, and rapidly
58 spread to other countries, causing more than 5.2 million deaths worldwide as of
59 November 18, 2022 (Platto, Xue, y Carafoli 2020; Zhu et al. 2020). As new variants of
60 SARS-CoV-2 emerged in different regions of the world, the World Health Organization
61 (WHO) designated new variants as variants of interest (VOI), variants under
62 monitoring (VUM), and variants of concern (VOC), depending on their epidemiological
63 and clinical behavior (Aleem, Akbar Samad, y Slenker 2022). The emergence of VOCs
64 has been of particular interest because they are associated with remarkable increases
65 in the incidence of COVID-19 worldwide (Choi y Smith 2021). For example, the Alpha
66 (B.1.1.7), Gamma (P.1.X), and Delta (B.1.617.2) variants were associated with drastic
67 increases in the number of infections and COVID-19-related deaths in the United
68 Kingdom, Brazil, and India, respectively (Tao et al. 2021).
69 On November 24, 2021, health authorities in Botswana and South Africa reported
70 circulating a new SARS-CoV-2 variant that was later designated Omicron and
71 classified as a VOC by the WHO (Viana et al. 2022). This new variant displayed
72 several mutations across the genome, some previously found in other VOCs (Thakur
73 y Ratho 2022). One of the most critical traits of Omicron is the presence of a high
74 number of mutations (32 mutations) in the spike protein compared to other VOCs,
75 which might confer this variant an increased affinity towards the angiotensin-
76 converting enzyme 2(ACE2 receptor), enhancing entry into the host cell, and the ability
77 to escape neutralization by antibodies induced by either natural infection or

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

vaccination (Thakur y Ratho 2022). Moreover, recent *in vitro* studies reported that the replication rates for Omicron are almost 70 times higher in bronchial tissue than in lung tissue in contrast to the Delta variant, which might explain the high transmissibility of Omicron (Chi-wai MC, Nicholls J, Pui-yan KH, Peiris M, Wah-Ching T, Lit-man LP s/f) Omicron has already spread across 77 countries, and their genomes are available in GISAID (Thakur y Ratho 2022). The Omicron variant was first detected in Mexico in November 2021 and has rapidly spread since, probably causing the recent increase in symptomatic COVID-19 cases observed in the country. Due to the high infectivity of Omicron in contrast to other VOCs, the study of the evolution and transmission patterns of this variant in different countries is crucial for the implementation of effective strategies to reduce its impact in a clinical context. In this work, we focus on identifying the first events of Omicron introduction to understand how the virus spreads and spreads in Mexico, with which it is possible to identify more than 100 multiple independent introductions. In addition, we analyzed adaptation signals throughout the genome and found only evidence of natural selection in protein Spike and homoplasic mutations in nonstructural proteins.

2. MATERIALS AND METHODS

2.1 Collection of genomes

Two hundred fifty-four complete Omicron genomes reported in Mexico were downloaded between November 28 to December 27, 2021, from the Global Initiative on Sharing All Influenza Data GISAID platform (https://www.gisaid.org/). Additionally, the first 20 genomes of Omicron reported per country were downloaded. However, there were fewer genomes available for some because there is no comprehensive

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

sequencing coverage. Finally, 2,148 Omicron SARS-COV-2 genomes sequenced worldwide were analyzed (including 187 of other lineages to guide the phylogeny). Additionally, to analyze the spread and diversification of Omicron in Mexico, genomes of this GISAID variant were downloaded from November 2021 to April 2022 with the parameters of inclusion, complete genome, high coverage, and low coverage exclusion.

2.2 Phylogenomics and phylogeographic analysis

Genomes were aligned using MAFFTv7 software (Katoh, Rozewicki, y Yamada 2019), trimming the 5' and 3' ends, and classified according to their viral lineage assignment under the Pango nomenclature system using Pangolin v3.1.7, version (Rambaut et al. 2020).

To estimate the phylogenetic relationships and phylogeography between SARS-COV-2 genomes, a maximum likelihood (ML) phylogenetic tree was constructed using the IQ-Tree (Nguyen et al. 2015) under the GTR + G model of nucleotide substitution with empirical base frequencies and four free site rate categories and statistical support (-alrt 1000). Sequences from lineages other than Omicron were added to this phylogeny to establish the monophyly of Omicron subtrees, while the sequences from Wuhan were used to root the tree. The tree was evaluated for a temporal signal using TempEst (Rambaut et al. 2016) (the outliers were eliminated).

A time-scaled phylogenetic tree was built with TreeTime v0.7.4 (Sagulenko, Puller, y Neher 2018), which specified a clock rate of $8x10^{-4}$ substitutions per site per year (s/s/y) according to the Nextstrain workflow (Hadfield et al. 2018).

The phylogeographic analysis detected and quantified Omicron lineage introduction events into Mexico. This analysis was reconstructed from the time-scaled tree

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

generated previously with TreeTime v0.7.4. We use two different approaches based on several models and inference methods. To identify independent introduction events of Omicron in Mexico, we used a discrete diffusion model implemented in the software package BEAST v1.10.4 (Drummond et al. 2012). The Bayesian analysis through Markov chain Monte Carlo (MCMC) was run on 106 generations and sampled every 1,000 generations using the BEAST v1.10.4 tool. The time-scaled phylogeny previously built was a fixed empirical tree and considered two possible ancestral locations: "Mexico" and "other location," according to the described pipeline Dellicour et al., 2021 (Dellicour, Durkin, et al. 2021, 2).

The MCMC convergence and mixing properties were inspected using Tracer v1.72 (Rambaut, A., Suchard, M.A., Xie, D. and Drummond, A.J. 2014, 6), and the effective sample size (ESS) values above 200 were achieved for all parameters. The phylogenies were annotated with a 10% burn-in by Tree Annotator. Subsequently, the "Seraphim" package (Dellicour et al. 2016) extracted the Spatio-temporal information from the data and visualized the phylogeographic reconstructions.

Second, we use PASTML (Ishikawa et al. 2019) briefly reconstructs the ancestral character states and their changes along with the trees, with maximum likelihood marginal posterior probabilities approximation (MPPA) and Felsenstein 1981 (F81) model options. For the parameters in PASTML, we used MPPA as a prediction method (standard settings). We added the character predicted by the joint reconstruction, even if it was not selected by the Brier score (option forced_joint).

On the other hand, the Omicron sequences from Mexico that were downloaded from November to April were built in to reconstruct a maximum likelihood phylogeny using iqtree v.2.1.1, with the GTR+F+R3 substitution model and statistical support (-alrt 1000) and the phylogenetic tree was visualized in iTOL v5 (Letunic y Bork 2021).

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

2.3 Daily mobility and cases report

The mobility data used in this study was obtained from the Google Mobility Reports service (https://www.google.com/covid19/mobility/) for all Mexican states. The data is found in the following categories: *i*) grocery and pharmacy; *ii*) retail and recreation; *iii*) parks; *iv*) workplaces; *v*) residential; *vi*) transit stations. The service has been updated daily since 26-02-2020. The last date of collection was 17-04-2022. The data were treated in weekly variations. Finally, we added up all six instances of mobility to come up with a single unitary value representing the mobility of a given week.

The country was divided into seven regions to analyze the mobility as follows: Northeast (NE; Coahuila, Nuevo León, and Tamaulipas), Northwest (NW; Baja California, Baja California Sur, Chihuahua, Durango, Sonora, and Sinaloa) Central North (CN; Aguascalientes, Guanajuato, Querétaro, San Luis Potosí, and Zacatecas), Central South (CS; Mexico City, Estado de México, Morelos, Hidalgo, Puebla, and Tlaxcala), West (W; Colima, Jalisco, Michoacán, and Nayarit), Southeast (SE; Guerrero, Oaxaca, Chiapas, Veracruz, and Tabasco) and South (S; Campeche, Yucatán, and Quintana Roo).

Moreover, daily cases of the COVID-19 transmission in Mexico were retrieved from the service https://datos.covid-19.conacyt.mx/. The regionalized (please see above) daily cases were analyzed in a 7-day window in order to eliminate data granularity. Next, the data was converted into cases/100k inhabitants. All data analysis was conducted through in-house R scripts.

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

2.4 Natural Selection Analysis

Natural selection can manifest in directional, diversifying, purifying, and episodic selection. We evaluated these types of natural selection in Omicron genomes from Mexico. The Omicron SARS-COV-2 genome was annotated using RCoV19 (https://ngdc.cncb.ac.cn/ncov/?lang=en) (Gong et al. 2020). The directional selection search was performed with FEL (fixed effects likelihood), SLAC (single likelihood ancestor counting) (Delport et al. 2010), and Fast Unconstrained Bayesian AppRoximation (FUBAR), assuming that the selection pressure is constant across the phylogeny for each site (Murrell et al. 2013). Furthermore, the episodic diversifying selection was evaluated with the mixed-effects model (MEME) (Murrell et al. 2012).

Homoplasy test

We evaluated the sequences of Omicron (from Mexico) that have arisen independently several times (homoplasies). Homoplasies are likely candidates for adaptation due to convergent evolution, recombination, or errors during sequence data processing. Homoplasies were detected on phylogeny using HomoplasyFinder (Crispell, Balaz, y Gordon 2019). We collected all the Omicron sequences from Mexico and two sequences from Bostwana (NICD-N22418 and NICD-N22397) as reference genomes. Additionally, we used Snippy (https://github.com/tseemann/snippy) to detect single-nucleotide variations (SNVs) using the Omicron genome NICD-N22418 as reference genomes.

3. RESULTS

3.1. Phylogenetic evidence for multiple introductions Omicron and the human mobility facilitates its spread

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

On December 3, 2021, the first WGS-confirmed case of symptomatic COVID-19 caused by Omicron (B.1.1.519) was reported by the Instituto Nacional de Diagnóstico y Referencia Epidemiológica (InDRE), corresponding to a patient from South Africa that arrived at Mexico City on 16 November 2021. Since then and up to December 27, 2021, 254 genomes were sequenced in Mexico, corresponding to the states of Baja California (*n=1*), Chiapas (*n=1*), Guerrero (*n=1*), Hidalgo (*n=1*), Mexico City (*n =159*), Oaxaca (*n=1*), Puebla (*n=4*), Quintana Roo (*n =18*), Sinaloa (*n=2*), State of Mexico (*n=35*), Tabasco (*n=12*), Tamaulipas (*n=5*), Veracruz (*n=1*) and Yucatán (*n=13*). Interestingly, most of the Omicron sequences come from Mexico City and the State of Mexico, probably due to most international flights arriving in Mexico through the country's capital. However, other states with international airports as well as important tourist destinations such as Baja California Sur, Quintana Roo, Yucatán, and Sinaloa also reported Omicron sequences and increases in cases in the first days of December (**Figure S1**).

To determine Omicron introductions to Mexico, we constructed a phylogenetic tree using 244 Omicron genome sequences from Mexico and 1903 foreign sequences deposited in GISAID. First, we inferred phylogeny for said sequences, to trace the movement of Omicron in Mexico through time and space. First, we note that the sequences of the Mexican genomes do not form a single clade but rather are scattered throughout the tree, consistent with multiple introductions to the country; In Figure A, the green branches correspond to the Mexican sequences, while the gray branches correspond to international sequences.

Through a discrete phylogeographic analysis following the Dellicour pipeline, introductions were evaluated (Dellicour, Durkin, et al. 2021). For this purpose, we used a time-scaled phylogenetic tree, and subsequently, introduction events were identified

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

in the MCC tree with an associated HPD interval by reporting the number of introduction events. The analysis detected a minimum number of 173 introduction events of the Omicron lineage (95% HPD interval = [162.925-176.025]) identified from the phylogenetic analysis of 244 samples collected in Mexico (**Figure 1A**).

The discrete phylogeographic analysis in BEAST show that the majority of introductions were represented by singletons (80%) and formed three small monophyletic clades composed of at least ten sequences from Mexico, corresponding to the center of Mexico, where the virus possibly initially spread throughout the country, driven by the mobility of people. However, the highest peak of Omicron in Mexico was reached by mid-January 2022, which may have had a more significant number of introductions.

The PastML analysis predicted Botswana as the location of the tree root of Omicron, suggesting that the Omicron variant origin might have been in southern African.

Four main clusters emerged from the Botswana cluster; the oldest clusters were located in Hong Kong, India, Australia, and South Africa (**Figure S2**), which could indicate that the first transmissions and spreads of the virus were possibly in those locations. After the Botswana cluster, there is the cluster of South Africa, which is linked to one of India, from which a small cluster of Mexico emerges; this result suggests that the first independent introductions of Omicron to Mexico (four sequences) could have come from South Africa or India.

On the other hand, another small cluster from Mexico is connected to the Singapore cluster, which comes from a large group from Hong Kong (**Figure S2**). These results indicate that the introduction of the virus to Mexico could have been through different locations in Africa, India, and Asia. This analysis showed that the sequences of these

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

regions share similarities with the first sequences of Mexico, and possibly, this was the spread of Omicron in the world towards Mexico.

As mentioned above, the peak of omicron was reached in mid-January 2022. A maximum-likelihood phylogenetic tree was built on understanding the distribution of Omicron during the increase in cases in the context of the first circulating genomes in Mexico. We found multiple introductions of different Omicron sublineages, which are marked with a red star in the phylogeny **(Figure 1B).**

The first genomes that circulated during the months of November and December belong to the Omicron sub-lineages, a total of 244 genomes. **BA.1** 23.3% (n=57), **BA.1.1** 39% (n=97), **BA.1.1. 10** 0.4% (n=1), **BA.1.13** 0.4% (n=1), **BA.1.15** 29.91% (n=73), **BA.1.17** 1.6% (n=4), **BA.1.17.2** 0.4% (n=1), **BA.1.18** 0.81% (n=2), and **BA.1.7** 4.9% (n=12).

These lineages were identified for the first time in Mexico City. For example, **BA.1** on 16/11/2021 imported from South Africa; BA**.1.1** on 19/11/2021; BA**.1.1.10** on 21/12/2021. **BA.1.17** on 18/12/2021, BA**.1.17.2** on 16/12/2021, and **BA.1.18** on 10/12/2021. While BA**.1.15** was first identified on 03/12/2021 in Tamaulipas. Therefore, most of the introductions of the lineages are from the center of the country, and these introductions spread to the rest of Mexico.

In the tree, the first genomes are distributed within the clades of the omicron sub-lineages, indicating that these first introductions gave rise to the spread of the variant throughout the country. However, not all introductory lineages reached a high frequency, such as BA.1.7, BA.18, BA.1.17.2. Otherwise, lineages BA.1, BA.1.1, and BA.1.15 were the introductions that dominated and increased their frequency during the Omicron peak, which correspond to the largest clades in the phylogenetic tree (**Figure 1B**).

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Interestingly, lineage BA.1.1 is subdivided into two large clades (clades in color Blue), followed by lineage BA.1.5 distributed in three clades (Green), followed by BA.1 (red), the most ancestral lineage in the tree. In contrast, the other lineages are found in a lesser proportion (**Figure 1B**).

We also investigated whether the introduction and spread of the Omicron variant could be due, in part, to trends of increased population mobility (**Figure 2**). We obtained human mobility data and divided the country into seven regions NE, NW, CN, CS, W, SE, and S. First, we found a mobility peak in December 2021 (**Figure 2-A**), involving all Mexican regions. The mobility increase observed in December 2021 is the highest observed in the Omicron selection period. Second, in **Figure 2-B** shows the increase in reported cases in January 2022. Next, in **Figure 2-C**, we analyze the correlation between increased mobility resulting in future cases. On average, it took six weeks for the correlation between the two variables to start showing a stronger correlation (except in the Northwest region). To explain the odd behavior of the Northwest region, in **Figure S3** the mobility of the states of Baja California, Chihuahua, Durango, and Sonora can be seen to increase earlier. Therefore, the earliest of these trends caused the correlation of this region to be different than the others. In **Figure 2-D,** The Omicron lineage peaked in mid-January 2022, where the B.1.1 lineage predominated throughout the country, followed by the BA.1.15 and BA.1 lineages. Therefore, the increase in cases is mainly associated with the BA.1.1 lineage. In addition, different sub-lineages of Omicron have been established, which have not increased in frequency, and perhaps many have become extinct.

As mentioned above, the peak of mobility increased in December 2021 in all regions of Mexico; therefore, we also evaluated the prevalence and spread of omicron sub-lineages over time in the seven regions (**Figure 3**). Initially, in November, the

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

prevalence of the Delta variant throughout the country was 99%. However, in December, there was an increase in mobility, and most of the cases this month in the Central and South region seem to be related to lineages BA.1 (~10.80%), BA.1.1 (~74.3%), and BA.1.15 (~11.25%). Whereas some states are located on the US-Mexico border (Baja California, Chihuahua, Coahuila). On December NE and NW had a higher prevalence of Delta (70%) than the Central-South and South states that presented a 60% prevalence of Omicron sublineages. Subsequently, in January in all regions, Omicron was already dominant with the three sub-lineages BA.1, BA.1.1, and BA.1.1.15. Additionally, there was already a diversity of Omicron sub-lineages such as BA.2 (first time detected on 9/01/2022 in Jalisco), BA.1.1.10, BA.1.17, BA.1.17.2, BA.1.18 and BA.1.15. By February, BA.1.1 was already dominant over the other sub-lineages in all regions, while BA.2 in March began to increase in frequency, especially in the Central South and South regions.

3.2 Analysis of natural selection of Omicron

We then searched for evidence of positive selection in Omicron genomes isolated from Mexico with different algorithms focused on a codon-based phylogenetic framework by calculating the ratio of non-synonymous (dN) to synonymous (dS) substitutions per coding sequence site (dN/dS). The evaluation of selection throughout the genome revealed that sites are highly conserved and show no evidence of episodic and directional selection. There was no evidence of an excess of non-synonymous mutations (dN/dS>1), representing potential purifying selection pressure, except for the gene encoding Spike. For the gene that codes for Spike protein, positive diversifying selection was detected on five codon sites, and non-synonymous changes in only five (dN/dS>1) through the FUBAR method, which assumes that the selection

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

pressure for each site is constant along the entire phylogeny with a Bayesian algorithm to infer rates (dN/dS) (Murrell et al. 2013). Thus, genes were subjected to a distinct natural selection strength.

The sites under positive natural selection found in the Spike gene were; L371S, R346K, N417K, E484A, S496G, and N1098D (**Figure 4**). The L371S mutation (on Spike) is within the receptor-binding domain (RBD, residues numbered 319–541). The other observed mutations, R346K, N417K, E484A, and S496G are also found in the RBD. Mutation R346K described in the Mu variant can alter the interactions with monoclonal antibodies (Fratev 2022). Mutation E484A has been shown to confer up to 10-fold greater resistance to vaccinee sera (Wang et al. 2021, 1). The mutations L371S and S496G by neutral polar amino acid probably produce no difference in the receptor's binding affinity (Rath, Padhi, y Mandal 2022). Finally, the N1098D mutation occurs in the spike protein's S2 subunit (amino acid residues 686–1273). These changes produced in the Spike protein by charged polar amino acids could have significant repercussions, favoring a better interaction with the receptor. Interestingly, these mutations may be local adaptations of the virus to the Mexican population.

3.3 Detection of Homoplasies

Homoplasies can be produced by base-pair substitutions or by recombination and can emerge naturally in the context of neutral evolution or positive natural selection. In order to detect homoplasies in the Omicron sequences in Mexico, we used HomoplasyFinder [24]. We used 244 Omicron (Mexico) sequences to construct a phylogeny, adding two South African genomes, NICD-N22418 and NICD-N22397, as the reference genomes. We identified homoplasies at 25 different positions (**Figure S4**), corresponding to 24.5% of the 102 polymorphic sites identified. Of these, 18 were

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

in the ORF1a nsp11 gene. In addition, the homoplasy at position 15854 was in RNA polymerase, and the rest of the homoplasies (at positions 1694 to 9870) were in the non-structural proteins (nsp). The homoplasies detected tend to change for thymine or cytosine (**Supplementary table S1**).

On the other hand, using SNIPPY (using NICD-N22418 and NICD-N22397 as reference genomes), we found that the homoplasies at positions 1694, 2205, 5250, 5465 can be directly observed from the Botswana genomes, suggesting that the rest of the homoplasies could be generated from the descendants of the internal nodes.

4. DISCUSSION

In this study, we performed a phylogenetic and phylogeographic analysis to infer Omicron introductions in Mexico and how mobility might affect virus spread following introductions. Subsequently, we analyzed the adaptive evolutionary processes in local genomes during the first month of circulation of Omicron in Mexico. According to the data obtained from the genomes collected, the earliest detected genomes come mainly from Mexico City, the State of Mexico, Quintana Roo, and Yucatán, of which the first of the reported cases was imported from South Africa. However, most of the introduction cases in Mexico were singletons, which indicates multiple independent introductions (173 independent introductions). These results are like those reported during the first epidemic wave in New York City by Dellicour et al., 2021; discrete phylogeographic analysis found 116 independent introductions in the city and relatively small clades. These results show the heterogeneity of the virus to establish local transmission successfully (Dellicour, Hong, et al. 2021). The multiple introductions of variants to Mexico have already been explored in other lineages, such as B.1.1.519 (Taboada et al. 2021) and Alpha (Zárate et al. 2022). For example, in the

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Alpha lineage, the introductions probably came from the United States, contributing to the spread of the Alpha lineage from the north to the south of the country (Zárate et al. 2022)

However, in the case of Omicron, the increase in cases was in Mexico City, probably due to the arrival of international flights. As inferred from the PastML phylogeographic analysis, the Mexican clusters came from South Africa and India. The other introductions came from Hong Kong or Singapore, and this lineage arose from South Africa; Therefore, we argue that human mobility is a fundamental factor in introducing variants to a country. South Africa *per se* is an important airport hub in the continent (Newfarmer, Page, y Tarp 2018). An example of that is mobility data that shows that there have been approximately 93 thousand trips from South Africa to other 30 non-African regions by November 2021. Moreover, by the time the WHO had declared Omicron as a VoC, the probability of importation of the new variant exceeded 50% in North America, Europe, Middle East, and Oceania (Bai et al. 2021). Therefore, people's inbound through international mobility is likely to be the introductory event of new variants and viruses.

From the moment a novel variant arrives in a country, internal mobility is among the factors that will facilitate its local transmission (Fakir y Bharati 2021; Oka, Wei, y Zhu 2021). We show that by observing a mobility increase and correlating it with the number of cases. In the regions where the first introductions of Omicron likely happened, i.e., Southwest, Northwest, and Central South, we observed that higher mobility was associated with higher cases six weeks after. On top of that, we argue that the mobility trends of a population follow a seasonal trend, as reported in (Hoogeveen, Kroes, y Hoogeveen 2022), since our data depicts a movement peak in the end-of-year season. Next, we show that the increase in cases reported in early

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

January 2022 was due to the variants BA.1.1, BA.1, and BA.15, and it seems that there is no regionalization of the sub-lineages. And, the local transmission chains could reflect the higher internal mobility observed in the Southwest and Central South of the country; this has been observed in the Delta variant in Mexico [...].

Once knowing the introductions of omicron to Mexico, we evaluated the adaptive changes in the genomes of Mexico. Since natural selection is one of the main evolutionary forces that generate diversity and show signs of adaptation of an organism, we evaluate the diversifying directional selection. We use FUBAR, which shows better performance on SLAC and FEL (Delport et al. 2010; Murrell et al. 2013). The results revealed six sites under positive selection on Spike protein (dN/dS>1). The residues under positive selection are located in RBD on spike protein surrounding the receptor-binding site (R346K, L371S, K417N, E484A, S496G, and N1098D) associated manly with evasion of antibodies. For example, the R346K amino acid change, previously detected in lineage Mu (B.1.621) (Laiton-Donato et al. 2021, 2) is associated with increased resistance to convalescent plasma therapy, and less neutralization of antibodies(Liu et al. 2022). The substitution R346K was previously reported in genomes from Mexico City with a prevalence of 40% in the capital and according to phylogenetic analyses, the sequences from Mexico City form a monophyletic clade (Cedro-Tanda et al. 2022). Genomes with the R346K mutation now designated as BA.1.1 sub-lineage (Mohandas et al. 2022) which predominated in Mexico and spread throughout the country.

The amino acid substitution L371S has been associated with a significantly increased binding affinity to human ACE2 (Kimura et al. 2022). The amino acid substitution K417N has been previously reported in the Alpha and Gamma lineages, which facilitates the escape of monoclonal antibodies (bamlanivimab/LY-CoV555) (Starr et

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

al. 2021, 016) and the escape of vaccine sera such as BNT162b2 (Hoffmann et al. 2021). The amino acid substitution E484A is under selection in genomes in Mexico; this substitution was previously E484K in the Beta, Gamma, and Mu variants related to the escape to neutralization of vaccines. Currently, mutation E484A is present in Omicron, weakening the mAbs binding to escape the immune response (Shah y Woo 2022). The S496G amino acid change is related to evasion of mAbs binding (Kannan et al. 2022).

In general, the sites under selection in the protein S of the Omicron variants in Mexico are associated with the evasion of antibodies. In addition, several recent studies have proven that natural selection is evident in the Spike a glycoprotein (Chaw et al. 2020; Tang et al. 2020). Changes in the amino acid sequence of RBD can dramatically impact the binding affinity of Spike for the ACE2 receptor, in addition to causing infectivity and transmission of SARS-CoV-2.

The methods used to detect natural selection determined sites under selection in Spike and not in other Omicron genome regions. However, ORF1ab, ORF3a, and ORF8 proteins have been reported to have sites that evolve under positive natural selection (Velazquez-Salinas et al. 2020). Many of the mutations can occur repeatedly and independently throughout the genome several times (homoplasies), which produce non-synonymous changes at the protein level, are candidates for possible adaptation due to positive natural selection. Therefore, we evaluated the presence of homoplasies, which are recurrent mutations in the genomes of Mexico; we found that most of the changes were associated with the ORF1a nsp11 gene, RNA polymerase, and non-structural proteins (nsp). The recurrent mutation signal in Orf1a, which encodes the nonstructural proteins Nsp11 and Nsp13, has been previously reported

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

(van Dorp et al. 2020, 2). In our analysis, the Nsp11 was the recurrent gene with homoplasies.

5. CONCLUSIONS

The analyses presented here show multiple introductions of omicron that took place in different regions of central-southern Mexico, then spreading to the rest of the country during the first months of its circulation. We have also gathered evidence to show that higher levels of mobility may have facilitated the spread of the variant to the outer regions of the country. Here, we show that the increased mobility at the end of the year resulted in a peak of Omicron cases; on average, it started being observed six weeks later. The identification of introductions is essential to determine the spread of the virus in a country. It also highlights the presence of diversifying natural selection in the Spike protein and the presence of homoplasies mutations in non-structural proteins, which may be possible candidates for adaptation of the virus to the Mexican population. It is crucial to identify possible adaptation signatures in SARS-CoV-2 for the continued development of vaccines and treatments.

Acknowledgments

We thank to Consorcio Mexicano de Vigilancia Genómica (CoviGen).

We thank the submitting laboratories for generating the genetic sequence and metadata and sharing it through the GISAID initiative, on which this research is based.

Authors contributions

HGCS conceived the initial study. HGCS, GLL, JTF and LMC designed the experiments. HGCS, GLL, and GSM performed analysis. HGCS, GLL, GMS, JTF, and LMC have contributed to the interpretation and discussion of the results. HCGS and

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

GLL wrote the first manuscript draft with editing from GMS, JTF, and LMC. The final manuscript was approved by all authors.

Reference

Aleem, Abdul, Abdul Bari Akbar Samad, y Amy K. Slenker. 2022. "Emerging Variants of SARS-CoV-2 And Novel Therapeutics Against Coronavirus (COVID-19)". En *StatPearls*. Treasure Island (FL): StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK570580/.

Bai, Yuan, Zhanwei Du, Mingda Xu, Lin Wang, Peng Wu, Eric H. Y. Lau, Benjamin J. Cowling, y Lauren Ancel Meyers. 2021. "International risk of SARS-CoV-2 Omicron variant importations originating in South Africa". *medRxiv*, diciembre, 2021.12.07.21267410. https://doi.org/10.1101/2021.12.07.21267410.

Cedro-Tanda, Alberto, Laura Gómez-Romero, Guillermo de Anda-Jauregui, Dora Garnica-López, Yair Alfaro-Mora, Sonia Sánchez-Xochipa, Eulices F. García-García, et al. 2022. "Early Genomic, Epidemiological, and Clinical Description of the SARS-CoV-2 Omicron Variant in Mexico City". *Viruses* 14 (3): 545. https://doi.org/10.3390/v14030545.

Chaw, Shu-Miaw, Jui-Hung Tai, Shi-Lun Chen, Chia-Hung Hsieh, Sui-Yuan Chang, Shiou-Hwei Yeh, Wei-Shiung Yang, Pei-Jer Chen, y Hurng-Yi Wang. 2020. "The Origin and Underlying Driving Forces of the SARS-CoV-2 Outbreak". bioRxiv. https://www.biorxiv.org/content/10.1101/2020.04.12.038554v1.

Chi-wai MC, Nicholls J, Pui-yan KH, Peiris M, Wah-Ching T, Lit-man LP. s/f. "HKUMed Finds Omicron SARS-CoV-2 Can Infect Faster and Better than Delta in Human Bronchus but with Less Severe Infection in Lung". Consultado el 12 de abril de 2022. https://www.med.hku.hk/en/news/press/20211215-omicron-sars-cov-2-infection.

Choi, Jun Yong, y Davey M. Smith. 2021. "SARS-CoV-2 Variants of Concern". *Yonsei Medical Journal* 62 (11): 961–68. https://doi.org/10.3349/ymj.2021.62.11.961.

Crispell, Joseph, Daniel Balaz, y Stephen Vincent Gordon. 2019. "HomoplasyFinder: a simple tool to identify homoplasies on a phylogeny". *Microbial Genomics* 5 (1): e000245. https://doi.org/10.1099/mgen.0.000245.

Dellicour, Simon, Keith Durkin, Samuel L. Hong, Bert Vanmechelen, Joan Martí-Carreras, Mandev S. Gill, Cécile Meex, et al. 2021. "A Phylodynamic Workflow to Rapidly Gain Insights into the Dispersal History and Dynamics of SARS-CoV-2 Lineages". *Molecular Biology and Evolution* 38 (4): 1608–13. https://doi.org/10.1093/molbev/msaa284.

Dellicour, Simon, Samuel L. Hong, Bram Vrancken, Antoine Chaillon, Mandev S. Gill, Matthew T. Maurano, Sitharam Ramaswami, et al. 2021. "Dispersal Dynamics of SARS-CoV-2 Lineages during the First Epidemic Wave in New York City". *PLOS Pathogens* 17 (5): e1009571. https://doi.org/10.1371/journal.ppat.1009571.

Dellicour, Simon, Rebecca Rose, Nuno R. Faria, Philippe Lemey, y Oliver G. Pybus. 2016. "SERAPHIM: Studying Environmental Rasters and Phylogenetically Informed Movements". *Bioinformatics* 32 (20): 3204–6. https://doi.org/10.1093/bioinformatics/btw384.

Delport, Wayne, Art F. Y. Poon, Simon D. W. Frost, y Sergei L. Kosakovsky Pond. 2010. "Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology". *Bioinformatics* 26 (19): 2455–57. https://doi.org/10.1093/bioinformatics/btq429.

Dorp, Lucy van, Mislav Acman, Damien Richard, Liam P. Shaw, Charlotte E. Ford, Louise Ormond, Christopher J. Owen, et al. 2020. "Emergence of Genomic Diversity and Recurrent Mutations in SARS-CoV-2". *Infection, Genetics and Evolution* 83 (septiembre): 104351. https://doi.org/10.1016/j.meegid.2020.104351.

Drummond, Alexei J., Marc A. Suchard, Dong Xie, y Andrew Rambaut. 2012. "Bayesian Phylogenetics with BEAUti and the BEAST 1.7". *Molecular Biology and Evolution* 29 (8): 1969–73. https://doi.org/10.1093/molbev/mss075.

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Fakir, Adnan M. S., y Tushar Bharati. 2021. "Pandemic Catch-22: The Role of Mobility Restrictions and Institutional Inequalities in Halting the Spread of COVID-19". *PLOS ONE* 16 (6): e0253348. https://doi.org/10.1371/journal.pone.0253348.

Fratev, Filip. 2022. "R346K Mutation in the Mu Variant of SARS-CoV-2 Alters the Interactions with Monoclonal Antibodies from Class 2: A Free Energy Perturbation Study". *Journal of Chemical Information and Modeling* 62 (3): 627–31. https://doi.org/10.1021/acs.jcim.1c01243.

Gong, Zheng, Jun-Wei Zhu, Cui-Ping Li, Shuai Jiang, Li-Na Ma, Bi-Xia Tang, Dong Zou, et al. 2020. "An Online Coronavirus Analysis Platform from the National Genomics Data Center". *Zoological Research* 41 (6): 705–8. https://doi.org/10.24272/j.issn.2095-8137.2020.065.

Hadfield, James, Colin Megill, Sidney M. Bell, John Huddleston, Barney Potter, Charlton Callender, Pavel Sagulenko, Trevor Bedford, y Richard A. Neher. 2018. "Nextstrain: real-time tracking of pathogen evolution". *Bioinformatics* 34 (23): 4121–23. https://doi.org/10.1093/bioinformatics/bty407.

Hoffmann, Markus, Prerna Arora, Rüdiger Groß, Alina Seidel, Bojan F. Hörnich, Alexander S. Hahn, Nadine Krüger, et al. 2021. "SARS-CoV-2 Variants B.1.351 and P.1 Escape from Neutralizing Antibodies". *Cell* 184 (9): 2384-2393.e12. https://doi.org/10.1016/j.cell.2021.03.036.

Hoogeveen, Martijn J., Aloys C. M. Kroes, y Ellen K. Hoogeveen. 2022. "Environmental Factors and Mobility Predict COVID-19 Seasonality in the Netherlands". *Environmental Research* 211 (agosto): 113030. https://doi.org/10.1016/j.envres.2022.113030.

Ishikawa, Sohta A., Anna Zhukova, Wataru Iwasaki, y Olivier Gascuel. 2019. "A Fast Likelihood Method to Reconstruct and Visualize Ancestral Scenarios". *Molecular Biology and Evolution* 36 (9): 2069–85. https://doi.org/10.1093/molbev/msz131.

Kannan, Saathvik R., Austin N. Spratt, Kalicharan Sharma, Anders Sönnerborg, Subbu Apparsundaram, Christian Lorson, Siddappa N. Byrareddy, y Kamal Singh. 2022. "Complex Mutation Pattern of Omicron BA.2: Evading Antibodies without Losing Receptor Interactions", abril. https://doi.org/10.20944/preprints202204.0120.v1.

Katoh, Kazutaka, John Rozewicki, y Kazunori D. Yamada. 2019. "MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization". *Briefings in Bioinformatics* 20 (4): 1160–66. https://doi.org/10.1093/bib/bbx108.

Kimura, Izumi, Daichi Yamasoba, Hesham Nasser, Jiri Zahradnik, Yusuke Kosugi, Jiaqi Wu, Kayoko Nagata, et al. 2022. "SARS-CoV-2 Spike S375F Mutation Characterizes the Omicron BA.1 Variant". Preprint. Microbiology. http://biorxiv.org/lookup/doi/10.1101/2022.04.03.486864.

Laiton-Donato, Katherine, Carlos Franco-Muñoz, Diego A. Álvarez-Díaz, Hector Alejandro Ruiz-Moreno, José A. Usme-Ciro, Diego Andrés Prada, Jhonnatan Reales-González, et al. 2021. "Characterization of the emerging B.1.621 variant of interest of SARS-CoV-2". *Infection, Genetics and Evolution* 95 (noviembre): 105038. https://doi.org/10.1016/j.meegid.2021.105038.

Letunic, Ivica, y Peer Bork. 2021. "Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation". *Nucleic Acids Research* 49 (W1): W293–96. https://doi.org/10.1093/nar/gkab301.

Liu, Haolin, Pengcheng Wei, John W. Kappler, Philippa Marrack, y Gongyi Zhang. 2022. "SARS-CoV-2 Variants of Concern and Variants of Interest Receptor Binding Domain Mutations and Virus Infectivity". *Frontiers in Immunology* 13. https://www.frontiersin.org/article/10.3389/fimmu.2022.825256.

Mohandas, Sreelekshmy, Pragya D. Yadav, Gajanan Sapkal, Anita M. Shete, Gururaj Deshpande, Dimpal A. Nyayanit, Deepak Patil, et al. 2022. "Pathogenicity of SARS-CoV-2 Omicron (R346K) Variant in Syrian Hamsters and Its Cross-Neutralization with Different Variants of Concern". *EBioMedicine* 79 (mayo). https://doi.org/10.1016/j.ebiom.2022.103997.

Murrell, Ben, Sasha Moola, Amandla Mabona, Thomas Weighill, Daniel Sheward, Sergei L. Kosakovsky Pond, y Konrad Scheffler. 2013. "FUBAR: A Fast, Unconstrained

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Bayesian Approximation for Inferring Selection". *Molecular Biology and Evolution* 30 (5): 1196–1205. https://doi.org/10.1093/molbev/mst030.

Murrell, Ben, Joel O. Wertheim, Sasha Moola, Thomas Weighill, Konrad Scheffler, y Sergei L. Kosakovsky Pond. 2012. "Detecting Individual Sites Subject to Episodic Diversifying Selection". *PLoS Genetics* 8 (7): e1002764. https://doi.org/10.1371/journal.pgen.1002764.

Newfarmer, Richard, John Page, y Finn Tarp, eds. 2018. *Industries without Smokestacks: Industrialization in Africa Reconsidered*. WIDER Studies in Development Economics. Oxford: Oxford University Press. https://oxford.universitypressscholarship.com/10.1093/oso/9780198821885.001.0001/oso-9780198821885.

Nguyen, Lam-Tung, Heiko A. Schmidt, Arndt von Haeseler, y Bui Quang Minh. 2015. "IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies". *Molecular Biology and Evolution* 32 (1): 268–74. https://doi.org/10.1093/molbev/msu300.

Oka, Tatsushi, Wei Wei, y Dan Zhu. 2021. "The Effect of Human Mobility Restrictions on the COVID-19 Transmission Network in China". *PLOS ONE* 16 (7): e0254403. https://doi.org/10.1371/journal.pone.0254403.

Platto, Sara, Tongtong Xue, y Ernesto Carafoli. 2020. "COVID19: An Announced Pandemic". *Cell Death & Disease* 11 (9): 1–13. https://doi.org/10.1038/s41419-020-02995-9.

Rambaut, A., Suchard, M.A., Xie, D. and Drummond, A.J. 2014. "Tracer v1.6".

Rambaut, Andrew, Edward C. Holmes, Áine O'Toole, Verity Hill, John T. McCrone, Christopher Ruis, Louis du Plessis, y Oliver G. Pybus. 2020. "A Dynamic Nomenclature Proposal for SARS-CoV-2 Lineages to Assist Genomic Epidemiology". *Nature Microbiology* 5 (11): 1403–7. https://doi.org/10.1038/s41564-020-0770-5.

Rambaut, Andrew, Tommy T. Lam, Luiz Max Carvalho, y Oliver G. Pybus. 2016. "Exploring the Temporal Structure of Heterochronous Sequences Using TempEst (Formerly Path-O-Gen)". *Virus Evolution* 2 (1): vew007. https://doi.org/10.1093/ve/vew007.

Rath, Soumya Lipsa, Aditya K. Padhi, y Nabanita Mandal. 2022. "Scanning the RBD-ACE2 molecular interactions in Omicron variant". *Biochemical and Biophysical Research Communications* 592 (febrero): 18–23. https://doi.org/10.1016/j.bbrc.2022.01.006.

Sagulenko, Pavel, Vadim Puller, y Richard A. Neher. 2018. "TreeTime: Maximum-Likelihood Phylodynamic Analysis". *Virus Evolution* 4 (1): vex042. https://doi.org/10.1093/ve/vex042.

Shah, Masaud, y Hyun Goo Woo. 2022. "Omicron: A Heavily Mutated SARS-CoV-2 Variant Exhibits Stronger Binding to ACE2 and Potently Escapes Approved COVID-19 Therapeutic Antibodies". *Frontiers in Immunology* 12. https://www.frontiersin.org/article/10.3389/fimmu.2021.830527.

Starr, Tyler N., Allison J. Greaney, Adam S. Dingens, y Jesse D. Bloom. 2021. "Complete map of SARS-CoV-2 RBD mutations that escape the monoclonal antibody LY-CoV555 and its cocktail with LY-CoV016". *Cell Reports Medicine* 2 (4): 100255. https://doi.org/10.1016/j.xcrm.2021.100255.

Taboada, Blanca, Selene Zárate, Pavel Iša, Celia Boukadida, Joel Armando Vazquez-Perez, José Esteban Muñoz-Medina, José Ernesto Ramírez-González, et al. 2021. "Genetic Analysis of SARS-CoV-2 Variants in Mexico during the First Year of the COVID-19 Pandemic". *Viruses* 13 (11): 2161. https://doi.org/10.3390/v13112161.

Tang, Xiaolu, Changcheng Wu, Xiang Li, Yuhe Song, Xinmin Yao, Xinkai Wu, Yuange Duan, et al. 2020. "On the origin and continuing evolution of SARS-CoV-2". *National Science Review* 7 (6): 1012–23. https://doi.org/10.1093/nsr/nwaa036.

Tao, Kaiming, Philip L. Tzou, Janin Nouhin, Ravindra K. Gupta, Tulio de Oliveira, Sergei L. Kosakovsky Pond, Daniela Fera, y Robert W. Shafer. 2021. "The Biological and Clinical Significance of Emerging SARS-CoV-2 Variants". *Nature Reviews. Genetics* 22 (12): 757–73. https://doi.org/10.1038/s41576-021-00408-x.

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Thakur, Vikram, y Radha Kanta Ratho. 2022. “OMICRON (B.1.1.529): A New SARS-CoV-2 Variant of Concern Mounting Worldwide Fear”. *Journal of Medical Virology* 94 (5): 1821–24. https://doi.org/10.1002/jmv.27541.

Velazquez-Salinas, Lauro, Selene Zarate, Samantha Eberl, Douglas P. Gladue, Isabel Novella, y Manuel V. Borca. 2020. “Positive Selection of ORF1ab, ORF3a, and ORF8 Genes Drives the Early Evolutionary Trends of SARS-CoV-2 During the 2020 COVID-19 Pandemic”. *Frontiers in Microbiology* 11 (octubre): 550674. https://doi.org/10.3389/fmicb.2020.550674.

Viana, Raquel, Sikhulile Moyo, Daniel G. Amoako, Houriiyah Tegally, Cathrine Scheepers, Christian L. Althaus, Ugochukwu J. Anyaneji, et al. 2022. “Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa”. *Nature* 603 (7902): 679–86. https://doi.org/10.1038/s41586-022-04411-y.

Wang, Pengfei, Manoj S. Nair, Lihong Liu, Sho Iketani, Yang Luo, Yicheng Guo, Maple Wang, et al. 2021. “Antibody Resistance of SARS-CoV-2 Variants B.1.351 and B.1.1.7”. *Nature* 593 (7857): 130–35. https://doi.org/10.1038/s41586-021-03398-2.

Zárate, Selene, Blanca Taboada, José Esteban Muñoz-Medina, Pavel Iša, Alejandro Sanchez-Flores, Celia Boukadida, Alfredo Herrera-Estrella, et al. 2022. “The Alpha Variant (B.1.1.7) of SARS-CoV-2 Failed to Become Dominant in Mexico”. *Microbiology Spectrum* 10 (2): e0224021. https://doi.org/10.1128/spectrum.02240-21.

Zhu, Na, Dingyu Zhang, Wenling Wang, Xingwang Li, Bo Yang, Jingdong Song, Xiang Zhao, et al. 2020. “A Novel Coronavirus from Patients with Pneumonia in China, 2019”. *New England Journal of Medicine* 382 (8): 727–33. https://doi.org/10.1056/NEJMoa2001017.

Figure legends

Figure 1 - Phylogeography and Phylogenomics analysis Omicron. A) Time-scaled analysis of Omicron showing the 173 introduction events to Mexico, shown at the nodes in red color, and branches color in green represent genomes sampled from Mexico. In contrast, the branches in gray correspond to samples from other places. B) Maximum likelihood phylogeny of Omicron Mexican lineages, tips are colored by lineage.

Figure 2 - Mobility. A) Google mobility trends (retail, recreation, grocery, pharmacy, parks, public transport, workplace, and residential) of each Mexican region are computed from 2021-09-18 to 2022-03-12. Mobility levels above 0 indicate a trend above the baseline. The baseline was a period of observation from before the pandemic (the baseline day is the median value from the 5 week period Jan 3 – Feb 6, 2020. B) Daily COVID cases are computed in each Mexican region from 2021-09-18 to 2022-03-12. C) Pearson and Spearman correlations between COVID cases and mobility are computed. In order to analyze how long it takes for a spike in mobility to result in more cases, the cases were assessed in *n* weeks in the future (after the mobility spike). D) Number of SARS-CoV-2 genomes

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

obtained at the national level. The BA.1.1 lineage was the most abundant in January and February. BA.2 in March, it seems its frequency begins to increase.

Figure 3 - Distribution of Omicron lineages in Mexico between November 2021 and April 2022. (A), Northeast (B), Central North (C), Central South (D), West (E), Southeast (F), and South (G). The lineages B.1.1, B.1, and B.1.15 were predominant in all regions in January.

Figure 4 - Spike protein model of Omicron SARS-COV. The sites that will evolve under positive selection according to FUBAR method are shown in red color.

Figure Supplementary

Figure S1. Number of cases per 100,000 population.

The number of cases per 1000 inhabitants in the first months of the introduction of omicron is shown in each state of Mexico. The states of Baja California Sur, Mexico City, State of Mexico, Quintana Roo, and Yucatan.

Figure S2. Ancestral-state reconstruction using maximum likelihood inferred with PastML. Colors correspond to different countries. Numbers inside (or next to) the circles indicate the number of strains assigned to the specific node, and the size of the circles is proportional to the number of tips.

Figure S3. Mobility of Northwest region.

The states composing the Northwest region were isolated in terms of their mobility trends in order to explain the odd behavior of the mobility *vs.* cases correlation reported in the Northwest region by Figure 2-C. The mobility index reported is the result of the mobility categories Google displays (i.e., retail, recreation, grocery, pharmacy, parks, public transport, workplace, and residential). The baseline indicates how mobility used to be in the timespan of Jan 3 – Feb 6, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Figure S4. Homoplasies in Omicron from Mexico.

A phylogenetic tree was reconstructed using 244 Omicron genome sequences. The colored right columns represent the 25 homoplasies identified. The nucleotides associated with these positions in each sequence are plotted and colored according to their type (Adenine=red, Cytosine=blue, Guanine=cyan, Thymine=orange, and no information for the nucleotide (Ns) = white). The tree's base is in the NICD-N22418 and NICD-N22397 as reference genomes. Homoplasies are indicated in the internal nodes (Mexican sequences).

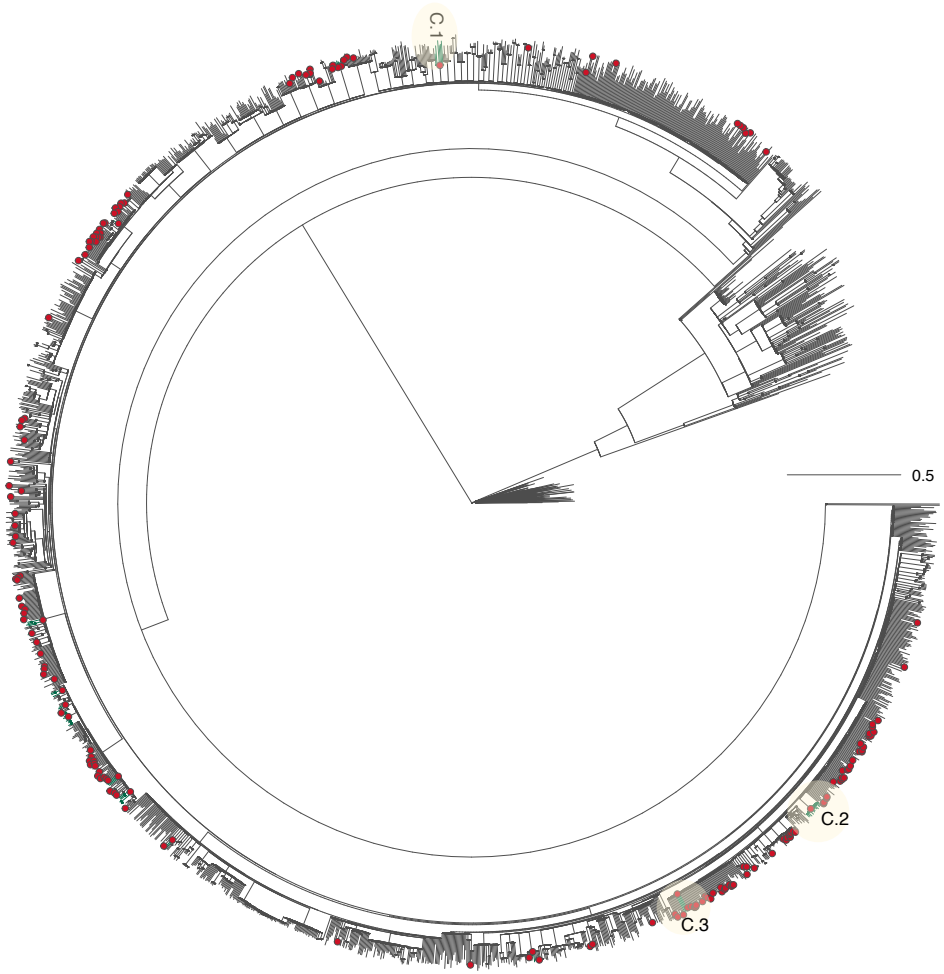
Supplementary Table S1

Supplementary Table 1. The number of changes observed (multiple alleles present) for each of sites on the phylogeny.

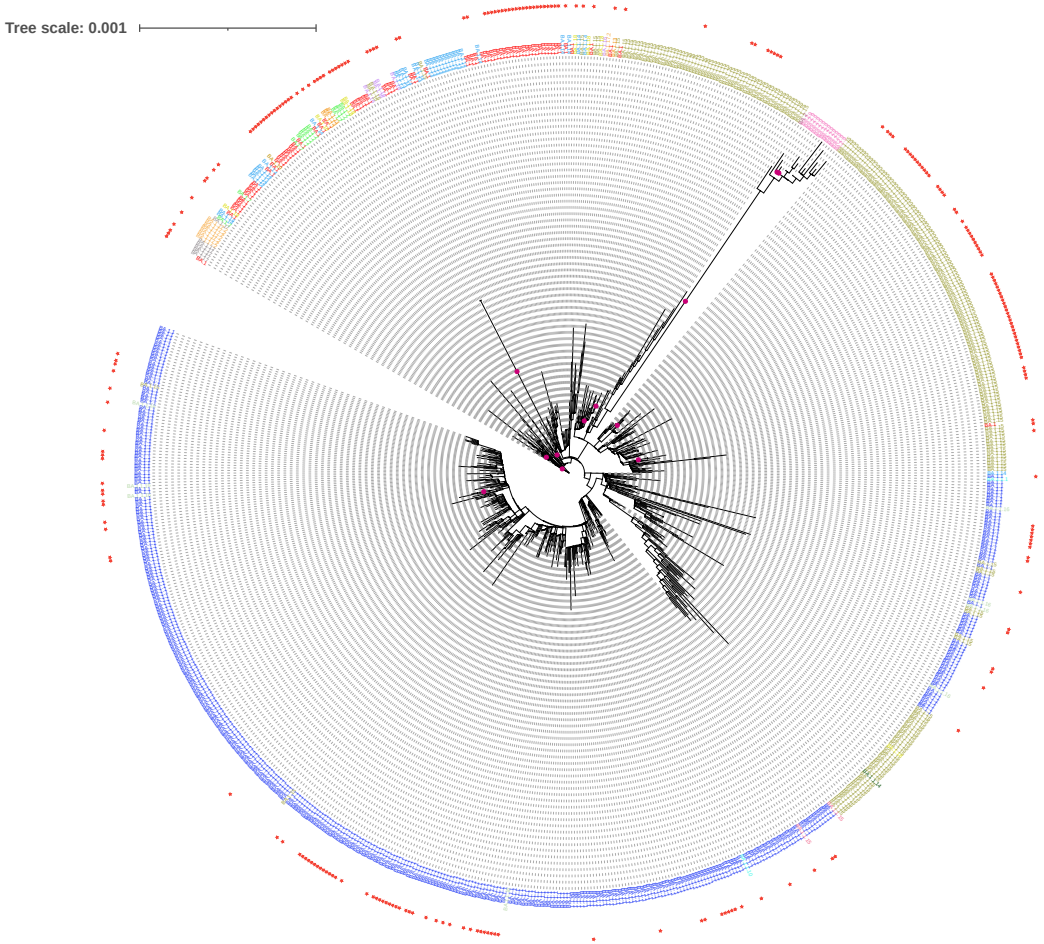
medRxiv preprint doi: https://doi.org/10.1101/2022.07.13.22277603; this version posted July 15, 2022. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

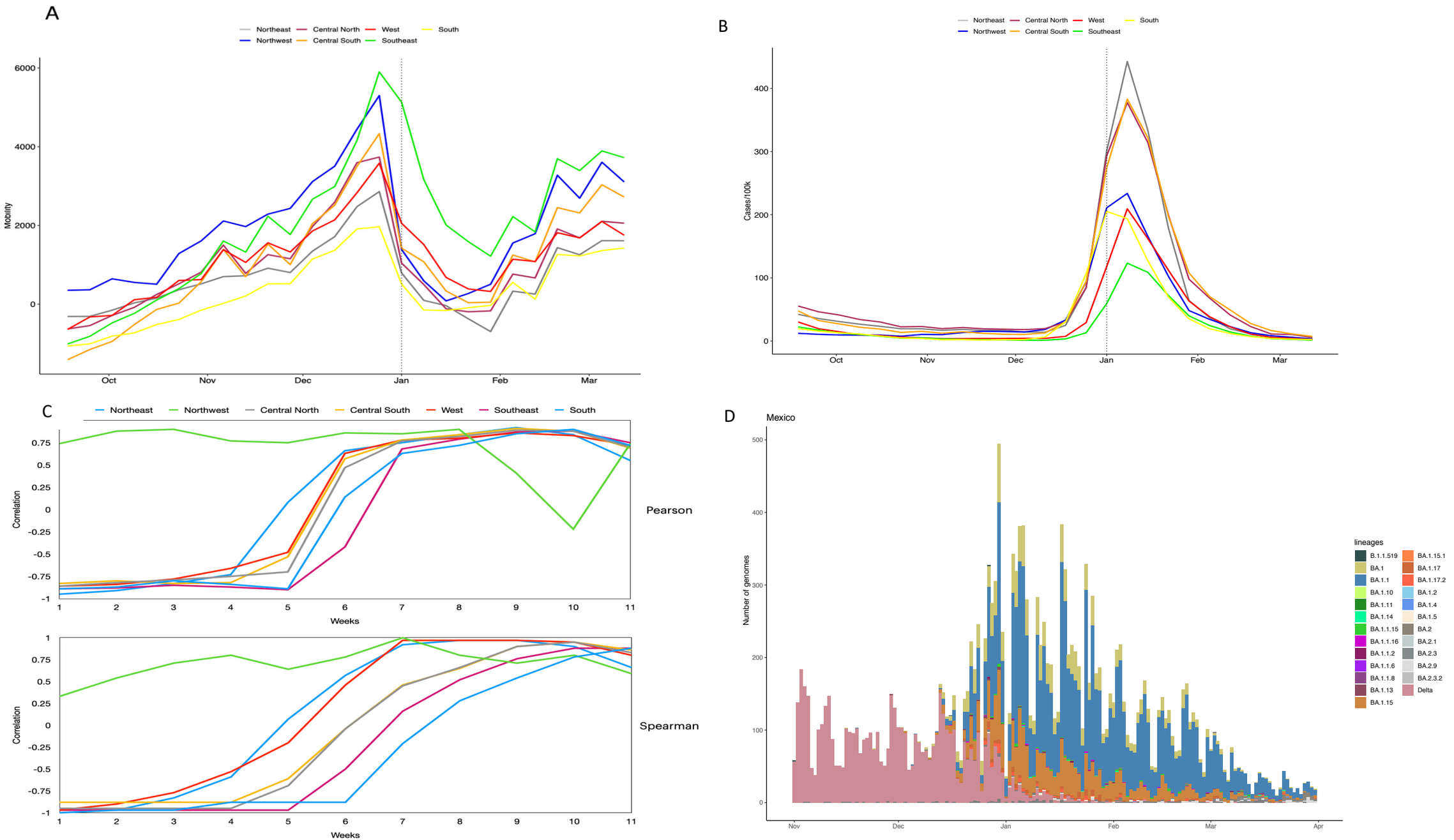
723

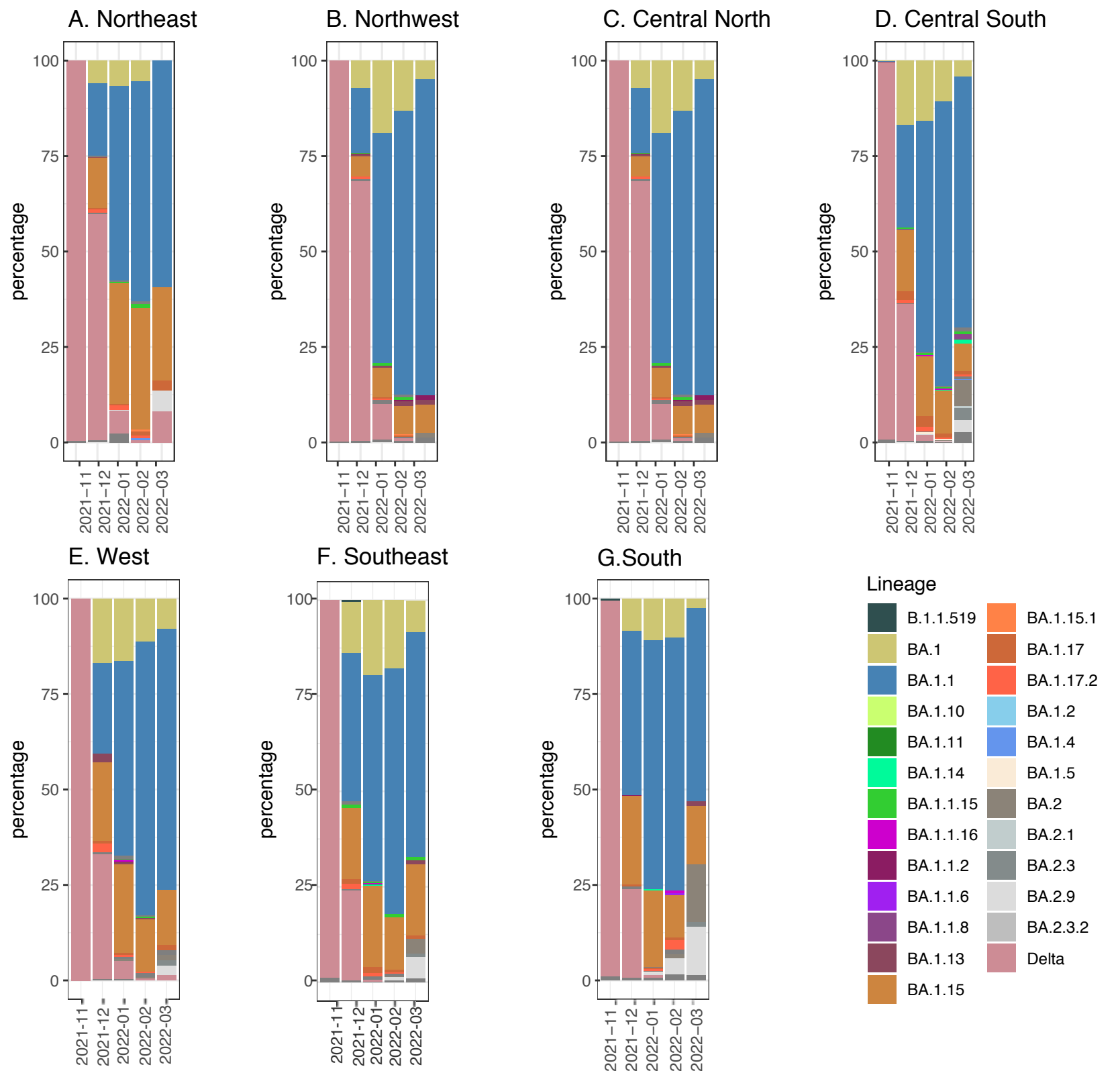
A

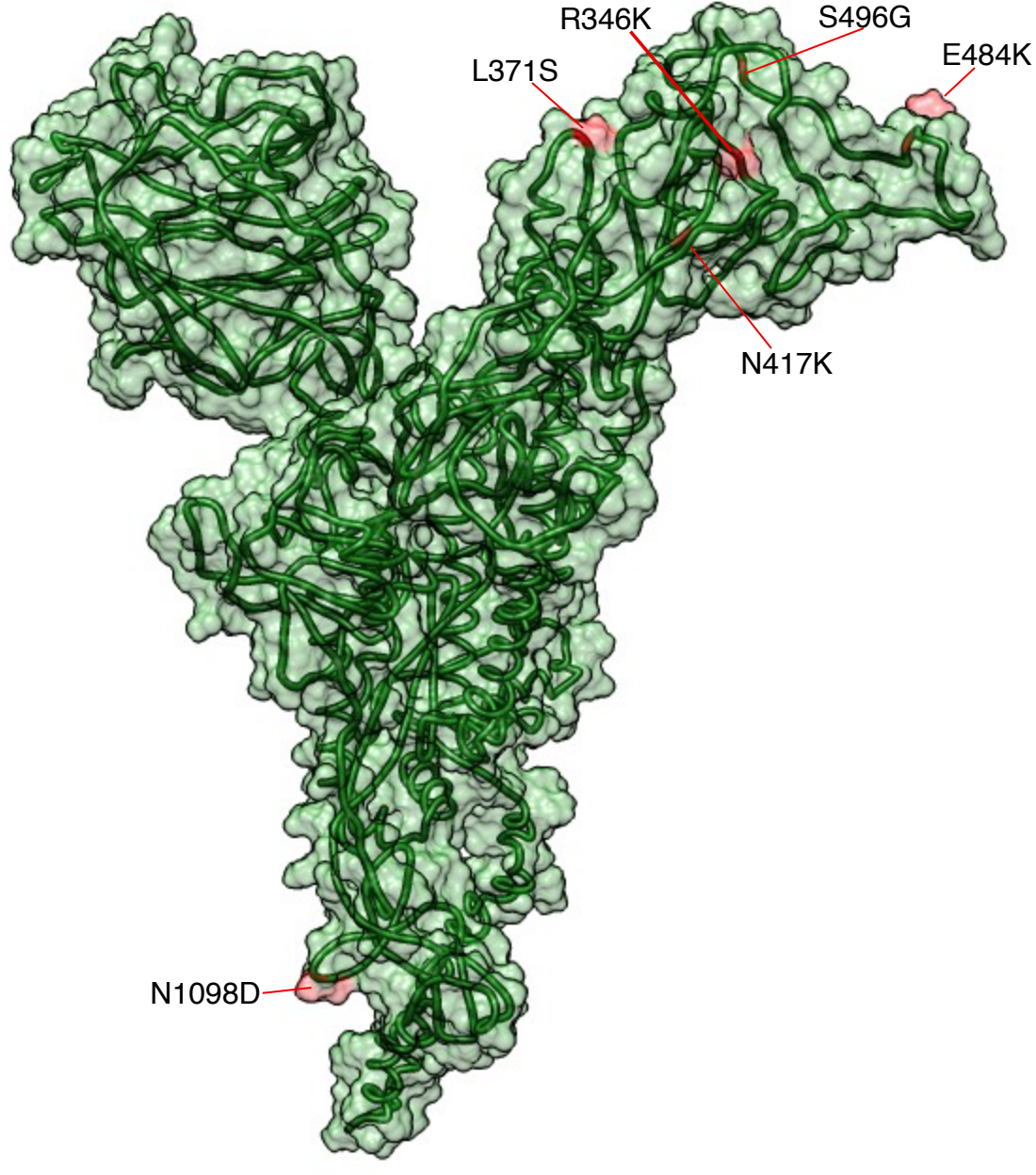


B




A
Northeast
Central North
West
South
Northwest
Central South
Southeast
Mobility
6000
4000
2000
0
Oct
Nov
Dec
Jan
Feb
Mar
B
Cases/100k
400
300
200
100
0
C
Correlation
0.75
0.5
0.25
-0.25
-0.5
-0.75
-1
1
2
3
4
5
6
7
8
9
10
11
Weeks
Pearson
Spearman
D
Mexico
Number of genomes
500
400
300
200
100
Nov
Dec
Jan
Feb
Mar
Apr
lineages
B.1.1.519
BA.1
BA.1.1
BA.1.10
BA.1.11
BA.1.14
BA.1.15
BA.1.1.16
BA.1.1.2
BA.1.1.6
BA.1.1.8
BA.1.13
BA.1.15
BA.1.15.1
BA.1.17
BA.1.17.2
BA.1.2
BA.1.4
BA.1.5
BA.2
BA.2.1
BA.2.3
BA.2.9
BA.2.3.2
Delta


A. Northeast
B. Northwest
C. Central North
D. Central South
E. West
F. Southeast
G.South
percentage
100
75
50
25
0
2021−11
2021−12
2022−01
2022−02
2022−03
Lineage
B.1.1.519
BA.1
BA.1.1
BA.1.10
BA.1.11
BA.1.14
BA.1.1.15
BA.1.1.16
BA.1.1.2
BA.1.1.6
BA.1.1.8
BA.1.13
BA.1.15
BA.1.15.1
BA.1.17
BA.1.17.2
BA.1.2
BA.1.4
BA.1.5
BA.2
BA.2.1
BA.2.3
BA.2.9
BA.2.3.2
Delta


R346K
S496G
E484K
L371S
N417K
N1098D