 doi:10.20944/preprints202205.0131.v1

Article

Dominance of three sublineages of the SARS-CoV-2 Delta variant in Mexico

Blanca Taboada[*#1], Selene Zárate[*#2], Rodrigo García-López[1], José Esteban Muñoz-Medina[3], Alejandro Sanchez-Flores[4], Alfredo Herrera-Estrella[5], Celia Boukadida[6], Bruno Gómez-Gil[7], Nelly Selem Mojica[8], Mauricio Rosales-Rivera[1], Angel Gustavo Salas-Lais[9], Rosa María Gutiérrez-Ríos[10], Antonio Loza[1], Xaira Rivera-Gutierrez[1], Joel Armando Vazquez-Perez[11], Margarita Matías-Florentino[6], Marissa Pérez-García[6], Santiago Ávila-Ríos[6], Juan Manuel Hurtado[12], Carla Herrera-Nájera[13], José Núñez-Contreras[14], Brenda Sarquiz-Martínez[15], Víctor García-Arias[16], María Santiago-Mauricio[17], Bernardo Martínez-Miguel[18], Julissa Enciso-Ibarra[7], Cristóbal Cháidez-Quiróz[19], Pavel Isa[1], Rosa María Wong-Chew[20], María-Eugenia Jiménez-Corona[21], Susana López[1], Carlos F. Arias[*1]

1 Departamento de Genética del Desarrollo y Fisiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico.
2 Posgrado en Ciencias Genómicas, Universidad Autónoma de la Ciudad de México, Ciudad de México, Mexico.
3 Coordinación de Calidad de Insumos y Laboratorios Especializados, Instituto Mexicano del Seguro Social, Ciudad de México, Mexico.
4 Unidad Universitaria de Secuenciación Masiva y Bioinformática, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico.
5 Laboratorio Nacional de Genómica para la Biodiversidad-Unidad de Genómica Avanzada, Centro de Investigación y de Estudios Avanzados del IPN, Irapuato 36821, Mexico.
6 Centro de Investigación en Enfermedades Infecciosas, Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas, Ciudad de México, Mexico.
7 Centro de Investigación en Alimentación y Desarrollo AC, Mazatlán, Sinaloa, Mexico.
8 Centro de Ciencias Matemáticas, Universidad Nacional Autónoma de México, Morelia, Michoacan, Mexico.
9 Estancia Posdoctoral (CONACYT), Laboratorio Central de Epidemiología, Instituto Mexicano del Seguro Social, Ciudad de México, Mexico.
10 Departamento de Microbiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico.
11 Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas, Ciudad de México, Mexico.
12 Unidad de Computo, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico.
13 Unidad de Investigación Médica Yucatán, Instituto Mexicano del Seguro Social, Mérida, Yucatán, Mexico/
14 Unidad de Investigación Biomédica de Zacatecas, Instituto Mexicano del Seguro Social, Zacatecas, Zacatecas, Mexico.
15 Laboratorio Central de Epidemiología, Instituto Mexicano del Seguro Social, Ciudad de México, Mexico.
16 Centro de Investigación Biomédica de Occidente, Instituto Mexicano del Seguro Social, Guadalajara, Jalisco, Mexico.
17 Centro de Investigación Biomédica del Noreste, Instituto Mexicano del Seguro Social, Monterrey, Nuevo León, Mexico.
18 División de Laboratorios Especializados, Instituto Mexicano del Seguro Social, Ciudad de México, México.
19 Centro de Investigación en Alimentación y Desarrollo AC, Culiacán, Sinaloa, Mexico.
20 Facultad de Medicina, Laboratorio de Investigación en Enfermedades infecciosas, División de investigación, Universidad Nacional Autónoma de Mexico.
21 Departamento de Epidemiología, Instituto Nacional de Cardiología Ignacio Chávez, Ciudad de México, Mexico.
* Correspondence: btaboada@ibt.unam.mx; Tel +52-77-7329-1612 (B.T.); selene.zarate@uacm.edu.mx; Tel +52-55-5488-6661, ext. 15318 (S.Z.); carlos.arias@ibt.unam.mx; Tel +52-77-7311-4701 (C.F.A.)
# B.T. and S.Z. contributed equally to this work.

Abstract: In this study, we analyzed sequences of SARS-CoV-2 isolates of the Delta variant in Mexico, which completely replaced other previously circulating variants in the country due to its transmission advantage. Among Delta sublineages detected, 81.5 % were classified as AY.20, AY.26, and AY.100. According to publicly available data, these sublineages only reached a world prevalence of less than 1%, suggesting a possible Mexican origin. The signature mutations of these sublineages are described, and phylogenetic analyses and haplotype networks were used to track their spread across the country. Other frequently detected sublineages include AY.3, AY.62, AY.103, and AY.113.

 doi:10.20944/preprints202205.0131.v1

Over time, the principal sublineages showed different geographical distributions, with AY.20 predominant in Central Mexico, AY.26 in the North, and AY.100 in the Northwest and South/Southeast. This work describes the circulation, from May to November 2021, of the primary sublineages of the Delta variants associated to the third wave of the COVID-19 pandemic in Mexico and reinforces the importance of SARS-CoV-2 genomic surveillance for timely identification of emerging variants that may impact public health.

Keywords: SARS-CoV-2; Variants of Concern; Delta Variant; genomic surveillance.

1. Introduction

Lineage B.1.617.2 of SARS-CoV-2 was first detected in India in October 2020. The World Health Organization (WHO) first classified it as a variant of interest (VOI) on April 4, 2021, and later as a variant of concern (VOC) on May 11, 2021, receiving the denomination of Delta [1]. By May 2021, the global spread of this variant had accelerated and quickly displaced all other circulating lineages throughout the world [2], accounting for 99.7% of SARS-CoV-2 infections by mid-November 2021, until the emergence of Omicron, the VOC that eventually displaced it.

As the Delta lineage diversified, a dynamic classification into sublineages was established, using the nomenclature "AY.X", giving rise to over 129 sublineages within the Pango classification (https://cov-lineages.org/) [3], distributed across different geographical regions. Therefore, the Delta variant is a group of distinct sublineages (AY.1 – AY.133) that accumulated specific mutations while circulating regionally or locally. The core mutations that confer Delta its transmissibility advantage in the original B.1.617.2 lineage [4] consists of 18 nonsynonymous substitutions and two deletions, with seven mutations and one deletion located in the spike protein (S). Some of these aa changes have been characterized more thoroughly, such as S:T478K, which confers higher receptor binding affinity [5]; S:L452R, which can increase transmissibility and immune evasion [6]; and S:P681R, which has been shown to improve fusogenicity [7]. Also, sublineages such as AY.4.2 presented higher transmissibility than the parental lineage [8], while some specific mutations, like S:A222V, may be associated with slightly higher viral titers thought to result in higher transmissibility in linage B.1.177 [9], or S:E484K, which can enhance immune evasion [10]. Therefore, genomic surveillance is critical for tracking the evolution and spread of SARS-CoV-2 lineages to understand their transmission and dispersion, and detect specific mutations linked to changes in virulence and immune evasion, especially in a population with increasing vaccination levels.

In Mexico, SARS-CoV-2 genomic surveillance has enabled the study of the expansion, and eventual succession, of increasingly-transmissible variants, such as B.1.1.519, which became dominant in late 2020 during the second pandemic wave observed in the country, peaking at 81% in March 2021 [11,12]. Later, the variants of concern (VOC) Alpha (B.1.1.7) and Gamma (P.1 and its sublineages) circulated in Mexico, albeit with a different geographical distribution, reached their highest prevalence in May (20%) and June (41%) 2021, respectively, during the period between the second and third waves. The Delta variant was associated with Mexico's third wave of infections, starting its spread in May 2021. Delta reached over a 90% prevalence in August 2021 and 99% in September, and its dominance was maintained until the Omicron variant arrived in December of the same year.

In this study, we characterized the dynamic distribution of the most prevalent sublineages of Delta in Mexico: AY.3, AY.20, AY.26, AY.62, AY.100, AY.103, and AY.113, from May through November 2021. These sublineages showed different geographic distribution patterns, with some of them dominating particular regions but not the whole country. Moreover, AY.20, AY.26, and AY.100, which accounted for 81.5 % of the infections, were practically exclusive to, and probably originated in Mexico; their dispersion and mutations are thoroughly studied here.

 doi:10.20944/preprints202205.0131.v1

2. Materials and Methods

2.1 Samples

Oropharyngeal or nasopharyngeal samples were obtained from all 32 Mexican states as part of the national governmental epidemiological surveillance program. Most of the samples were collected in public clinics and hospitals of the Ministry of Health of Mexico [Instituto Mexicano del Seguro Social (IMSS) and Instituto Nacional de Enfermedades Respiratorias (INER)]. RNA from 14,529 samples, positive for SARS-CoV-2 by real-time reverse transcription-polymerase chain reaction (RT-qPCR), anonymized before the start of the study, were analyzed by the CoViGen-Mex consortium under the Mexican Official Norm NOM-017-SSA2-2012 [13] between February 2020 and November of 2021, as previously described [14]. Sequencing was done using the Illumina COVIDSeq kit or Nextera XT (Illumina, Inc., San Diego, CA), in accordance with the manufacturer's instructions, using different Illumina platforms and 2x150 cycles of paired-end runs. FASTQ reads were generated by the Illumina pipeline at BaseSpace (https://basespace.illumina.com). Adapters, low-quality bases, and duplicate reads were removed from the data of each sample using a customized pipeline described previously[15], and the remaining reads were mapped with Bowtie2 v2.3.4.3 [16] against the Wuhan-Hu-1 (NC_045512) reference genome. Afterward, consensus sequences were generated using iVar (v1.3.1)[17] with the following settings: base score Q>20, 10X as minimum read coverage to call a base, and a threshold of at least 55% under majority rule for Single Nucleotide Polymorphism (SNP) calling. Lineage assignment was performed using the Pangolin lineage classification software tool [3].

2.2 Mexican Delta sequences dataset

The CoViGen-Mex consortium has deposited 13,600 sequences with genome coverage >90%, together with accompanying metadata, in the Global Initiative on Sharing All Influenza Data (GISAID) platform (https://www.gisaid.org/)[18] by December 14, 2021. Of these sequences, 6,696 were assigned as Delta (B.1.617.2+AY.x) and deposited in GenBank (Supplementary Table S1). An additional dataset of 13,072 Mexican Delta sequences and their metadata was retrieved from GISAID on the same date, resulting in a total of 19,768 Delta variant genomes from Mexico, which were used in this study (Supplementary Table S1).

2.3 Geographical circulation of Delta sublineages

To analyze the variation of the most prevalent Delta sublineages in Mexico, sequences from AY.3 (n=1,003), AY.20 (n= 9,749), AY.26 (n= 4,673), AY.62 (n=224), AY.100 (n= 1,650), AY.103 (n=215), and AY.113 (n=572) were included. Moreover, to inspect their geographical distribution and spread, the country was divided into seven regions (Fig. S1) as previously reported[14], each including the following states: Northwest (NW): Baja California, Baja California Sur, Chihuahua, Durango, Sonora, and Sinaloa; Northeast (NE): Coahuila, Nuevo Leon, and Tamaulipas; Central North (CN): Aguascalientes, Guanajuato, Querétaro, San Luis Potosi, and Zacatecas; Central South (CS): Mexico City, State of Mexico, Morelos, Hidalgo, Puebla, and Tlaxcala; West (W): Colima, Jalisco, Michoacán, and Nayarit; South (S): Guerrero, Oaxaca, Chiapas, Veracruz, and Tabasco; Southeast (SE): Campeche, Yucatan, and Quintana Roo. A stack density plot was built for each region using ggplot2 v.3.3.5[19] in R v.4.1.0. Also, correlation analyses were performed to assess the differences in lineage distributions among regions.

2.4 Mutation description

Only sequences from the three most prevalent Delta sublineages in Mexico, AY.20, AY.26, and AY.100, accounting for 81.5% were included in the analysis of their mutations and spread over time. Synonymous/nonsynonymous substitutions were identified using the Nextclade single nucleotide variant (SNV) calling system. Only SNVs present in more than 1% of the genomes were considered. A heatmap of mutation frequency was built using ComplexHeatmap v.2.8.0 [20] in R.

2.5 Phylogeny and haplotype network

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2022 doi:10.20944/preprints202205.0131.v1

A randomized subset of 1,190 sequences, representative of AY.20, AY.26, and AY.100, was selected, reflecting the sublineages' monthly circulation proportions observed in Mexico and evenly drawn from all the Mexican states. Additionally, 591 genomic sequences from other countries were included: 295 from the rest of North America, 88 from Europe and South America each, 60 from Asia, and 30 from Africa and Oceania each. These sequences were aligned against the reference Wuhan-Hu-1 isolate sequence (NC_045512) using MAFFTv7 with the option -addfragments [21]. The reference sequence and the UTR regions were manually eliminated from the alignment. A maximum-likelihood phylogeny was built using iqtree2 [22], with the GTR+R+F3 substitution model [23], and time-scaled using LSD2 and collection dates [24]. Finally, ggtree v.3.0.2 [25] and treeio v.1.16.1 [26] packages were used to display the tree using R. The alignment was converted into Nexus format and imported into PopArt v.1.7 software [27] to generate a Templeton, Crandall, and Sing (TCS) network [28], masking sites with more than 5% undefined states. Each sample's lineage and location information were used to estimate and visualize the geographical and lineage relationships of haplotype groups. SNVs were associated with a particular haplotype group.

2.6 Statistical analysis

Independence χ^2 tests were carried out in R with 10,000 Monte Carlo replicates to evaluate the association between sublineages and categorical variables (gender, patient status, cycle threshold [Ct, evaluating different binning sizes 1, 2, 5, and 10], region, and age). Results were evaluated with pairwise Student's t-tests with 10,000 resamples, using the lowest sublineage depth to overcome sampling bias. Normality was tested with the Shapiro-Wilk test. False discovery rate (FDR) correction was carried out for multiple sublineage testing. Finally, the odds ratio between Nextclade clades and categorical variables, at 95% of confidence, were calculated using binary logistic regression models in SPSS v22.

3. Results

3.1. Spread and epidemiology of Delta sublineages

In Mexico, as of March 2022, the pandemic had presented four distinct nationwide epidemiological waves of SARS-CoV-2 (Fig. 1A shows the first three waves), with corresponding peaks of cases in June 2020, January 2021, August 2021, and January 2022, respectively. The swift expansion of the Delta VOC ensued the third epidemiological wave in Mexico, lasting from late June to November 2021 (Fig. 1A). Delta overtook the previously prevalent Alpha and Gamma VOC and the dominant lineage B.1.1.519 (Fig. S2). Delta showed an exponential growth starting in May (5.5%), becoming the most prevalent variant in the country by June (38.6%), and reaching over 97% prevalence in August 2021. Onward, our study will focus on Delta genomes from samples collected from May to November 2021.

In Mexico, 89 sublineages of Delta were detected in the analyzed period (Supplementary Table S1), dominated by AY.20, AY.26, and AY.100, which accounted for about 49.4%, 23.7%, and 8.4%, respectively, of all genomes reported in the study period (Fig. 1B). Of interest, all three lineages seem to have originated in Mexico since their prevalence was extremely low in the rest of the world, with a percentage of 0.6%, 1.0%, and 0.7%, respectively (Fig. S3). Other less abundant sublineages in the country include AY.3 (Mexico(M): 5.1%; World(W): 3.6%), AY.103 (M: 1.1%; W: 6.4%), AY.113 (M: 2.9%; W: 0.3%), and AY.62 (M: 1.1%; W: 0.1%). These seven Delta sublineages constitute the majority (91.6%; n=18,576) of the cases during Delta circulation in Mexico.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2022 doi:10.20944/preprints202205.0131.v1

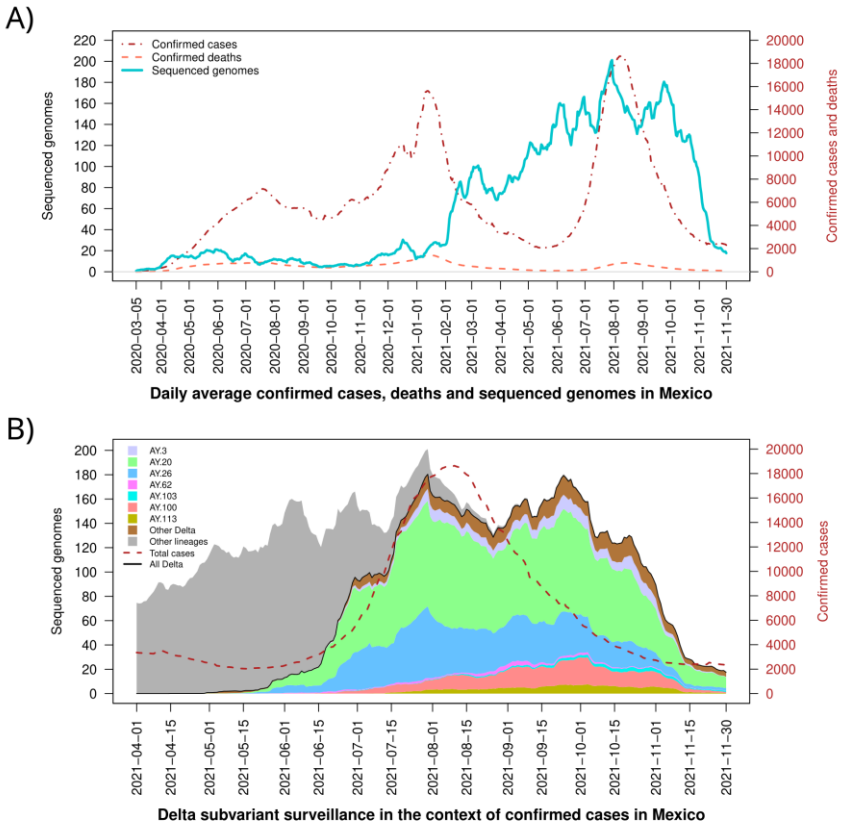

Figure 1. SARS-CoV-2-cases and diversity of Delta variant viruses in Mexico. (**A**) Confirmed cases and deaths in the country and number of Delta genomes sequenced between March 1, 2020, and November 30, 2021. (**B**) Stacked area plot showing the lineage diversity from April through November 2021. Vertically, lineages are stacked on top of one. The black line corresponds to the total Delta genomes reported and the dashed red line to the number of confirmed cases for context.

Sublineage AY.20 remained the most prevalent since May, reaching 60% in June and decreasing to 46% in November 2021 (Fig. 1B and Fig. S2), while AY.26 had a maximum prevalence of 33% during June and dropped to 9% in November. On the other hand, AY.3, AY.100, AY.103, and AY.113 were initially detected at low levels (<1.5%, each) but then increased their prevalence, with a peak in November (7.5%, 12.5%, 3.9%, and 6.3%, respectively), when the Omicron variant reached Mexico. Finally, AY.62 had a very low prevalence at the national level, but it was detected almost exclusively in the State of Yucatan in the SE region, with a frequency higher than 10% throughout.

A significant association was found between Delta sublineages and their monthly distribution (at α=0.05; p-value=0.0001), as well as with the state of sampling (p-value=0.0001). On the contrary, no significant associations with cycle threshold values (Ct values), gender, patient status, or age group were observed (at α=0.05; p-values>0.05 in all cases).

3.2 Delta sublineages geographical dispersion

The country was divided into seven geographical regions to analyze the spread of the Delta sublineages throughout Mexico, as seen in Figure S1. Lineage AY.20 was the most prevalent in the country, and Figure 2 shows that it reached at least 75% of all cases in the CS region (Fig. 2D), however, it was identified in less than 25% of the cases in the NW (Fig. 2A) and SE regions (Fig. 2G). Moreover, AY.20 was identified in the CS region as early as May 2021 and spread to other regions later. Contrastingly, the most abundant lineage in the NW was AY.26 (>65%), appearing before May 2021. This lineage was also detected throughout the country, although at a lower percentage (8%-33%), and at a later date. Likewise, AY.103 was present almost exclusively in the NW region and extended to the NE and CN but did not spread further. In the SE, lineages AY.20, AY.26, and AY.100

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2022 doi:10.20944/preprints202205.0131.v1

circulated with similar frequencies, while lineage AY.62 was only detected in this region (Fig. 2G), suggesting a local transmission chain. In the S region (Fig. 2F), AY.20, AY.26, AY.100, and AY.113 lineages were also frequently identified. Regarding the NE, CN, and W regions, AY.3 (15%-25%), AY.20 (45%-53%), and AY.26 (19%-25%) were the most abundant sublineages (Fig. 2B, 2C, and 2E). In these regions, the Pearson correlation coefficient, calculated between their sublineages' monthly distributions, was greater than 0.94 ($p < 0.05$), suggesting similar spread patterns of these sublineages over time; thus, they were considered a single region (NE+CN+W) in subsequent analyses.

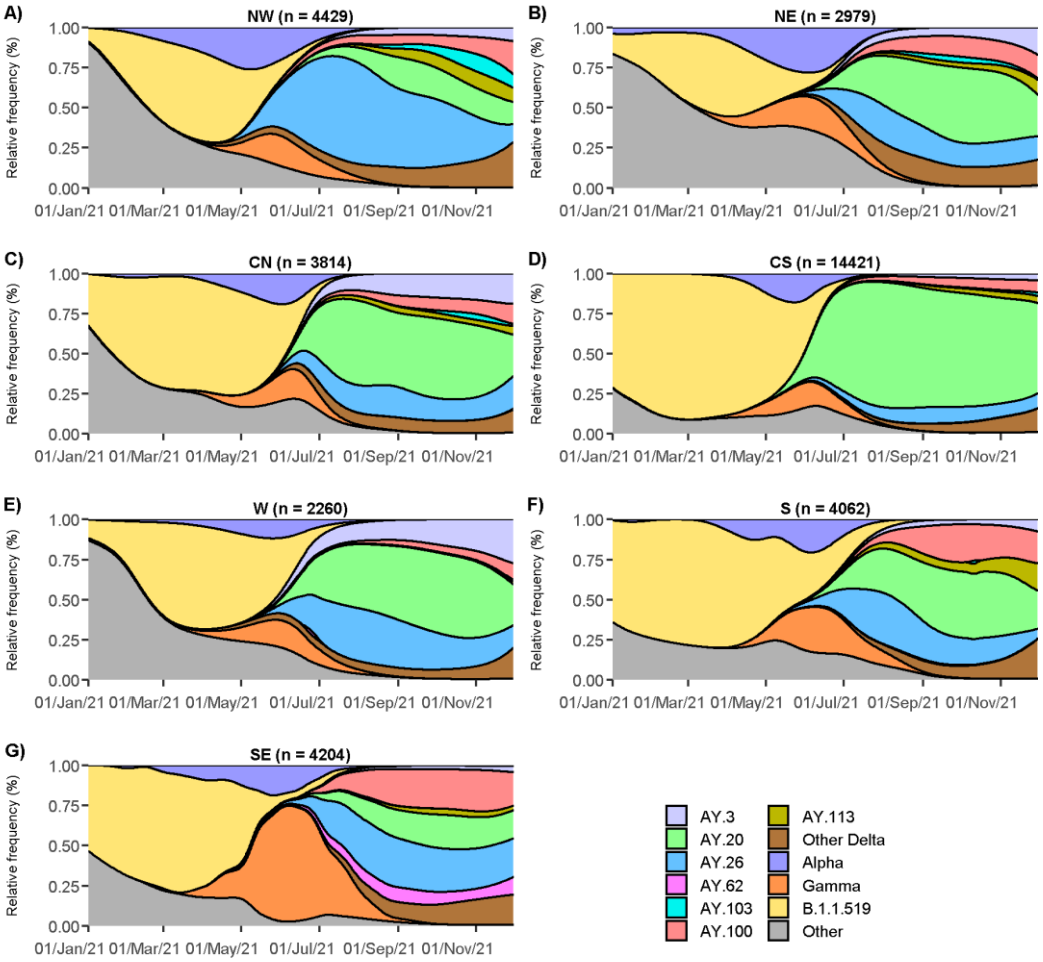

Figure 2. Density plots of relative lineage distribution by geographical region, considering sequences from January to November 2021. **(A)** Northwest, **(B)** Northeast, **(C)** Central-north, **(D)** Central-south, **(E)** West, **(F)** South, and **(G)** Southeast.

3.3 Mutations in Delta AY.20, AY.26 and AY.100 sublineages

To further characterize the most prevalent sublineages in Mexico, AY.20, AY.26, and AY.100, their defining mutations were examined. All available Mexican sequences of these sublineages were analyzed, resulting in 16,093 genomes (Supplementary Table S2). In total, 14,541 nucleotide changes were identified. Of these, 57.8% (n=8,429) were nonsynonymous SNPs, and the rest were synonymous SNPs (n=6,112). Among the nonsynonymous SNPs, only 66 showed a prevalence higher than 1%. When analyzing the accumulation of mutation patterns by sublineages, it was seen that AY.20 and AY.100 had, on average, more substitutions (40.6 and 41.6, respectively) than AY.26 (35.1). Interestingly, AY.20 and AY.100 belong to the 21J Nextclade Delta subclade, the most widespread, whereas AY.26 falls within subclade 21I, which started decreasing its frequency even before Omicron appeared.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2022 doi:10.20944/preprints202205.0131.v1

In Figure 3, all mutations with a prevalence higher than 5% in at least one sublineage are shown. The 18 defining changes of the Delta variant (B.1.1.617.2), shared by all its sublineages, can be observed. Seven substitutions are found in Spike (S), three in ORF1b and N, two in ORF7a, and one in ORF3a, M, and ORF9b, with none in ORF1a, E, ORF6, ORF7b, and ORF8. In addition, all Delta sequences harbor two amino acid deletions (S:157/158 and ORF8:119/120) and three stop codons. Sublineages AY.20, AY.26, and AY.100 harbor 13, 9, and 13 additional nonsynonymous changes, respectively, compared to the signature mutations of the Delta variant, having (Fig. 3).

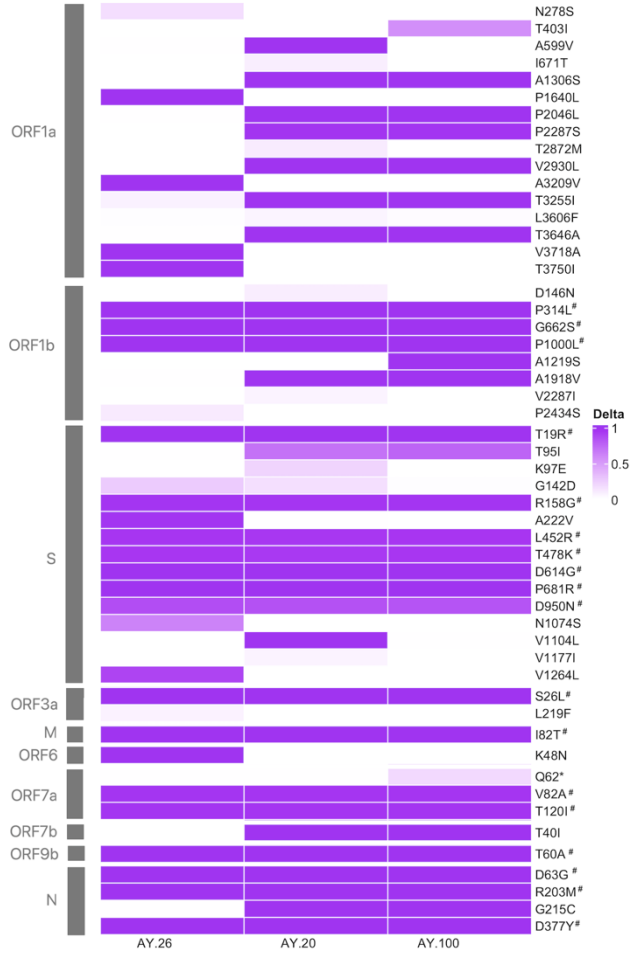

Figure 3. Frequency of aa changes in SARS-CoV-2 genome sequences of Mexico compared against Wuhan reference genome (NC_045512). Only changes above 5% prevalence in at least one of the sublineages are considered. The Delta defining substitutions are marked with #.

Some sublineage-specific substitutions are present in prevalences ranging between 86% and 99% of their genomic sequences analyzed; eight in lineage AY.26 (four in ORF1a, three in S, and one in ORF6, see Fig. 3), two present only in AY.20 (ORF1a:A599V and S:V1104L), and one in AY.100 (ORF1b:A1219S). This, in addition to the ten high-frequency substitutions shared by these lineages (six in ORF1a, and one in each ORF1b, S, ORF7b, and N), as expected given that both are placed in clade 21J. Moreover, other substitutions occurred at a prevalence of <50%, suggesting that these three sublineages may have distinct local transmission chains that lead to local viral subpopulations. This observation was further explored using a haplotype network and a phylogenetic reconstruction.

3.4 Delta haplotype network

A time-scaled phylogenetic tree and a haplotype network were built to explore the spread of Delta sublineages at Mexico's regional level. The NE+CN+W regions (Fig. S1) were collated into a single region, given the high correlation of circulating sublineages in these areas over time. The network in Figure 4 reveals a clear separation of the sublineages

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2022 doi:10.20944/preprints202205.0131.v1

and their association with specific regions in which they circulated; for lineage AY.20, three large (marked as A, B, and C) and two minor (D and E) subclusters can be clearly identified, indicating that different haplotypes of AY.20 circulated together. In the periphery of the main haplotypes (corresponding to the largest circles) of A and C subclusters, many sequences from the CS (purple) region were present, but also from the NE+CN+W (orange) regions. Correspondingly, in the phylogeny (Fig. S4), AY.20 sequences from the CS region appeared earlier than those from the NE+CN+W region, suggesting that the spread of this sublineage started there.

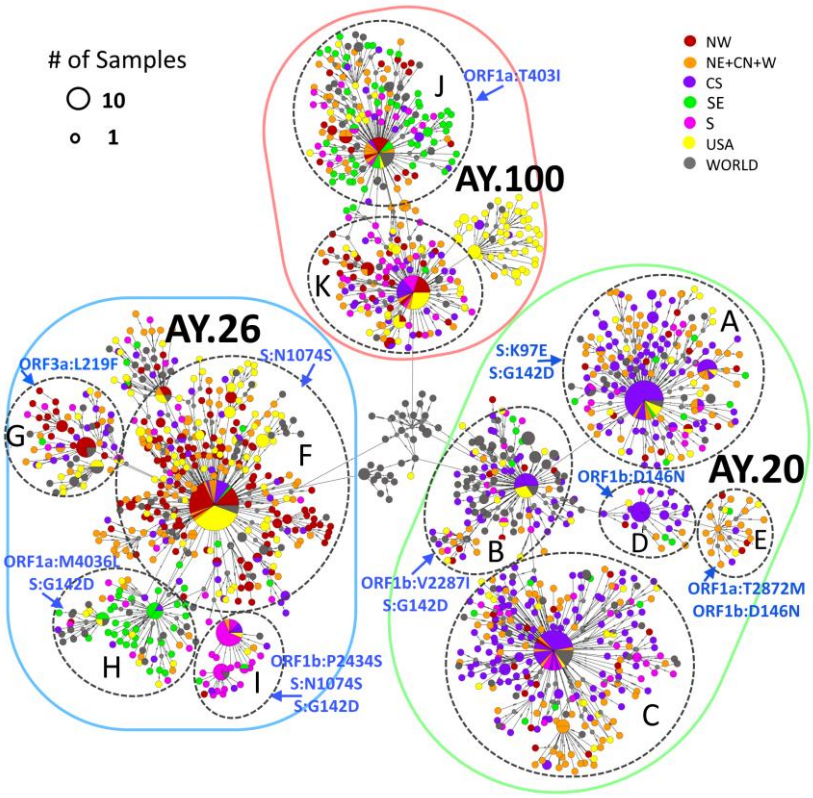

Figure 4. Haplotype network using mutations of AY.20, AY.26, and AY.100 sublineages from Mexico and worldwide. Colors indicate the regions in Mexico where sequences were isolated, with yellow and gray circles representing the USA and the rest of the world. The size of the circles indicates the number of samples within the same haplotype (scale is provided). Main subclusters within each lineage are marked with black discontinuous circles, and their specific mutations are indicated in blue.

Interestingly, sequences within subcluster B, which had the highest number of international samples (grey), harbor two additional nonsynonymous substitutions (ORF1b:V2287I and S:G142D) (Fig. 4). This is in contrast to subcluster C, composed mainly of Mexican sequences, which carry only the sublineage-specific mutations. These observations also support the idea that the original AY.20 started its circulation in the CS region and, as it spread, it acquired more mutations. The spatiotemporal distribution of AY.20 (Fig. 5A) showed that the earliest identification was observed in the CS region, particularly in Mexico City (28% in May-June), and in the South region, with the highest prevalence in Guerrero (14% in May-June), following dispersion to other regions, especially NE+CN+W.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2022 doi:10.20944/preprints202205.0131.v1

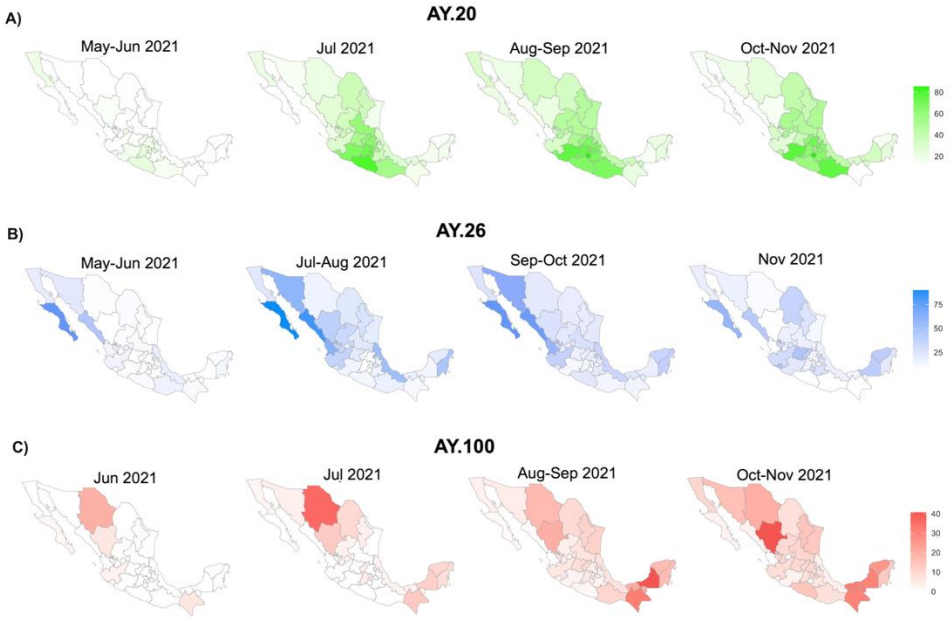

Figure 5. Map series showing the spatiotemporal distribution of the most prevalent Delta sublineages in Mexico. **(A)** Sublineage AY.20. (**B)** Sublineage AY.26. (**C)** Sublineage AY.100. The prevalence was calculated for each State over time.

For sublineage AY.26, the network revealed the presence of different haplotypes with distinct geographical distributions. The largest subcluster F, as well as G, were dominated by sequences obtained from samples collected in the NW region (red), NE+CN+W regions (orange), and USA (yellow). Also, subcluster I from the SE region (green) and H from the SE region (green) derived from F suggest different haplotypes with local transmission chains. These observations are consistent with the phylogeny (Fig. S4), where samples from the NW region dominated, especially on earlier dates, along with a few sequences from the SE region. Later, sequences from NE+CN+W regions appeared, followed by sequences from the rest of the country, but with distinct clades for the S and SE regions. Moreover, the data shown in Figure 5B confirms that AY.26 spread initially in the northern states of Mexico, probably corresponding to haplotypes F and G, and at a lower frequency in the SE region with haplotypes H and I; then, this lineage spread to NE+CN+W and S regions.

Finally, sublineage AY.100 had two main subclusters, J and K (Fig. 4). Subcluster K included sequences mainly from the USA (yellow) and all regions of Mexico except the SE region (green). In subcluster J, samples from the SE region (green) were observed along with the S region (pink) and at a lower prevalence of CS and NE-CN+W region, suggesting two separate transmission chains for these haplotypes, one started spreading from the North and other from the South. These patterns match those observed in the phylogeny (Fig. S4) and the spatiotemporal distribution of AY.100 (Fig. 5C).

Of note, some of the less-frequent mutations (present in <85% of total Mexican Delta sequences) found in the heatmap (Fig. 3) corresponded to a particular subcluster in the network, supporting the idea of different local transmission chains of Delta sublineages. For instance, mutation S:K97E, present only in 21.6% of AY.20 genomes, was present in subcluster A (Fig. 4). Also, mutation ORF1a:T403I, identified in 55.7% of AY.100 sequences, was only found in subcluster J. On the other hand, some less-frequent mutations, such as S:G142D, were shared by sublineages. This mutation was present in 25.8% of AY.20 sequences corresponding to subclusters A and B and in 9.8% of AY.26 sequences belonging to subclusters H and I. In some cases, the subclusters with additional mutations are restricted to a geographical region, as can be observed in subclusters H and I of AY.26, which share mutations S:G142D (25.8% of prevalence) and S:N1074S (61.4% of

 doi:10.20944/preprints202205.0131.v1

prevalence), plus ORF1b:P2434S (10.2% of prevalence) in H, and ORF1a:N278S (16.4% of prevalence) in I.

4. Discussion

The Delta variant of SARS-CoV-2 spread worldwide during 2021 and replaced most previously circulating lineages due to a transmission advantage and its ability to partially escape the pre-existing immunity [28] [29]. The ample circulation of Delta led to a regional diversification that resulted in 133 sublineages. In Mexico, Delta was principally responsible for the third COVID-19 wave of the pandemic, accounting for more than of 99% infections by November 2021.

Contrary to other countries in which the Delta variant was associated with increased severity of COVID-19 in unvaccinated patients (higher number of hospitalizations and severe symptoms) [30,31], in Mexico the proportion of severe illness and case fatality rates remained lower than in the previous waves [12]. Despite having low vaccination rates when Delta appeared (<9% with a complete initial vaccination protocol and <12% partly vaccinated) and during its peak surge (<43% and <24%, respectively) [32], a decrease in hospitalization and mortality in Mexico was observed in comparison with the second wave. Various factors may have contributed to lessening the impact of the Delta wave, such as the decision of health authorities to prioritize the vaccination in older people, a better knowledge of the medical treatment of the disease, and seroprevalence of 23.4% before the entry of Delta (Muñoz-Medina JE et at, unpublish data). However, according to the same group, by the end of the third wave, seroprevalence reached 86.7%, probably due to Delta and the progress in vaccination.

This work describes the presence and geographical distribution of Delta sublineages AY.3, AY.20, AY.26, AY.62, AY.100, AY.113, which were the most prevalent in Mexico, with a particular focus on the three most abundant (AY.20, AY.26, and AY.100). It seems that migration patterns could have influenced the sublineage distribution chains. For instance, sublineages AY.100 and AY.113 were highly prevalent in the S and SE regions, correlating with their detection frequencies in Central American countries, such as Costa Rica (3.4 and 21%, respectively) and Guatemala (30 and 68%). This observation may reflect the continuous influx of migrants from Central America to Mexico. However, AY.100 was also detected early in the NW, suggesting at least two possible spread chains of distinct haplotypes in each region. By contrast, sublineage AY.3, which was equally prevalent in the USA and Mexico, exhibited a higher prevalence only in the NE+CN+W regions (15%-25%). A very interesting case is sublineage AY.62, which in Mexico was almost exclusively found in the state of Yucatán in the SE region, where it reached 20% prevalence from July to September. This lineage was also detected in the USA, particularly in Texas and Kentucky; however, it only accounted for 0.75% of genomes in these locations during that period, suggesting that it entered Mexico through the tourism ports of the region and stayed there.

Of interest, sublineages AY.20 and AY.26 seem to be almost exclusive to Mexico, representing nearly 74% of Delta infections in the country, while it was detected at a much lower prevalence (<1.6%) elsewhere (Fig. S3). Regarding sublineage AY.20, all subclusters in the network were dominated by samples from the CS region, particularly Mexico City, which also corresponded to the earliest samples in the phylogeny and on the map (Fig. S4 and Fig 5). Therefore, this sublineage either entered or originated in Mexico City and later spread efficiently from there to other regions. The haplotype network and mutation analysis showed limited regional clustering of AY.20 populations, but different spread chains could be observed. Interestingly, AY.26 sublineage was detected early in the north of Mexico, particularly in the NW region, although bordering USA states did not exhibit high frequencies of this sublineage (California and Arizona reported a frequency of AY.26 of 5 % vs. 50% of NW Mexico, according to GISAID data), indicating that the circulation of AY.26 through the Mexico-US border was limited in contrast to what was reported for the Alpha variant [14]. This data suggests that this sublineage originated in the NW of Mexico. Interestingly, the prevalence of AY.26 decreased by November 2021 (<10%), as did the

 doi:10.20944/preprints202205.0131.v1

majority of worldwide Delta sublineages belonging to clade 21I (of Nextclade), while AY.20 remained the most prevalent (>44%) until the arrival of the Omicron variant. Overall, AY.20 started to spread from central and southern Mexico, where it reached its highest prevalence, whereas AY.26 and AY.100 seem to have spread simultaneously from the NW and the S/SE regions towards the center of the country and the NE+CN+W regions. In general, the region NE+CN+W seems to be a transition area where the three main sublineages co-circulated but were introduced at different timepoints.

The mutation analysis was carried out only for the three most prevalent sublineages: AY.20, AY.26, and AY.100. Notably, in addition to the defining Delta substitutions in the Spike protein, namely L452R, T478K, D614G, and P681R, that increase transmissibility and immune escape [5–7,29], other substitutions were identified in these sublineages, which may have also played a role in Delta enhanced phenotypes. For instance, several substitutions in the N-terminal domain of S were present, including T19R and R158G, which were common to the three sublineages; T95I was shared by AY.20 and AY.100; and G142D was shared by AY.20 and AY.26. These substitutions have been identified in other Delta sublineages, such as AY.4 and some of them in Omicron (T95I and G142D). The role of each specific mutation has not been completely elucidated, but they all have been implicated in immune escape [33–35]. In addition, mutation S:A222V, present in AY.26, slightly increases human cell receptor (ACE2) binding [36], while mutations N:R203M, common to the three sublineages, and N:G215C, present in AY.20 and AY.100, have also been shown to improve viral replication and assembly [37,38]. Even though Delta-defining mutations do not include changes in ORF1a, many Delta sublineages harbor specific mutations in this ORF. Here, we identified 16 aa changes in ORF1a, six of which were shared by AY.20 and AY.100, probably due to their placement in Nextclade subclade 21J. Despite the abundance of mutations in ORF1ab in all SARS-CoV-2 genomes, most of them remain primarily uncharacterized.

5. Conclusions

This study describes the evolution of the most prevalent sublineages of the Delta variant in Mexico, contributing to a better understanding of their dynamics and spread. This knowledge is relevant since new SARS-CoV-2 lineages have the potential to have a considerable impact on public health systems as they can potentially lead to an increase in infectivity, transmissibility, and/or disease severity.

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/s1, Figure S1: Map of Mexico with the seven geographical regions indicated; Figure S2: Stacked density plot of relative frequency of SARS-CoV-2 lineage circulation in Mexico; Figure S3: Frequency of the most prevalent Delta sublineages in Mexico; Figure S4: Time-scaled, maximum likelihood phylogeny; Table S1: Patient demographic and clinical characteristics of Mexican sequences classified as Delta and reported as of 14 December,2021.

Author Contributions: Conceptualization, B.T., S.Z. and C.F.A.; methodology, B.T., S.Z., J.E.M.-M., A.S.-F., A.H.-E., C.B., and B.G.-G.; software, B.T., S.Z., R.G.-L., N.S.M., M.R.-R, and J.M.H.; validation, B.T., S.Z., A.S.-F., C.B., and J.M.H.; formal analysis, B.T., S.Z., and R.G.-L.; investigation, J.E.M.-M, A.S.-F., A.H.-E., C.B., B.G.-G. A.G.S.-L., J.A.V.-P., M.M.-F., M.P.-G., S.A.-R., C.H.-N., J.N.-C., B.S.-N., V.G.-A., M.S.-M., B.M.-M., J.E.-I., and C.C.-Q.; resources, J.E.M.-M., A.S.-F., A.H.-E., C.B., B.G.-G., A.G.S.-L.,C.H.-N., J.N.-C., B.S.-N., V.G.-A., M.S.-M., B.M.-M., J.E.-I., C.C.-Q., and C.F.A.; data curation, B.T., S.Z., M.R.-R., R.M.G.-R., A.L., X.R.-G., P.I., R.M.W.-C., M.E,J.-C., S.L.; writing—original draft preparation, B.T., S.Z. and C.F.A.; writing—review and editing, B.T., S.Z., R.G.L., J.E.M.-M., A.H.-E., C.B., B.G.-G, M.R.-R, A.G.S.-L., R.M.G.-R, A.L., J.A.V.-P., S.A.-R., X.R.-G, P.I., R.M.W.-C., M.E.J.-C., S.L., and C.F.A.; visualization, B.T., S.Z., R.G.-L., and M.R.-R.; supervision, B.T., S.Z., J.E.M.-M., A.S.-F., A.H.-E., B.G.-G, S.A.-R., C.F.A.; project administration, C.F.A.; funding acquisition, J.A.V.-P., and C.F.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially supported by grants "Vigilancia Genómica del Virus SARS-CoV-2 en México-2022" (PP-F003; to CFA), "Caracterización de la diversidad viral y bacteriana" (FORDECYT to JAVP)) from the National Council for Science and Technology-México (CONACyT), grant 057 from the "Ministry of Education, Science, Technology and Innovation (SECTEI) of Mexico

City" (to CFA), and grant "Genomic surveillance for SARS-CoV-2 variants in Mexico" to CFA) from the AHF Global Public Health Institute at the University of Miami, as well as by the national epidemiological surveillance system. R.G.-L. (ProNacEs #I1000/023/2021; C-08/2021) and A.G. S.-L (408350) are recipients of postdoctoral fellowships from CONACyT.

Informed Consent Statement: It was waived because all samples used in this study were unlinked from any personal identifiers prior to commencement of the study. Table S1

Data Availability Statement: The generated sequences of SARS-CoV-2 used in this study have been publicly shared through the Global Initiative on Sharing All Influenza Data (GISAID) repository and have also been deposited in the Genbank NCBI database. Accession numbers are listed in Supplementary Table S1.

Acknowledgments: We gratefully acknowledge authors from the originating laboratories re-sponsible for obtaining the SARS-CoV-2 specimens and the submitting laboratories from which genetic sequence data were generated and shared via the GISAID Initiative. We appreciate the computer assistance provided by Jerome Verleyen, Roberto Bahena and Veronica Jacinto. The project LANCAD-UNAM-DGTIC-350 of the Dirección General de Cómputo y Tecnologías de la Información (DGTIC-UNAM) provided Supercomputing resources in MIZTLI).

Conflicts of Interest: There are no conflicts of interest declared by the authors. The funding institutions were not involved in the study's design, data collection, analysis, interpretation, manuscript writing, nor the decision to publish the findings.

References

1. Tracking SARS-CoV-2 Variants Available online: https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/ (accessed on 23 February 2022).
2. Campbell, F.; Archer, B.; Laurenson-Schafer, H.; Jinnai, Y.; Konings, F.; Batra, N.; Pavlin, B.; Vandemaele, K.; Van Kerkhove, M.D.; Jombart, T.; et al. Increased Transmissibility and Global Spread of SARSCoV- 2 Variants of Concern as at June 2021. *Euro Surveill.* **2021**, *26*, 2100509, doi:10.2807/1560-7917.ES.2021.26.24.2100509.
3. 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. *Nature Microbiology* **2020**, *5*, 1403–1407, doi:10.1038/s41564-020-0770-5.
4. Chowdhury, S.; Bappy, M.H.; Chowdhury, S.; Chowdhury, Md.S.; Chowdhury, N.S. Current Review of Delta Variant of SARS-CoV-2. *European Journal of Medical and Health Sciences* **2021**, *3*, 23–29, doi:10.24018/ejmed.2021.3.6.1120.
5. Wang, Y.; Liu, C.; Zhang, C.; Wang, Y.; Hong, Q.; Xu, S.; Li, Z.; Yang, Y.; Huang, Z.; Cong, Y. Structural Basis for SARS-CoV-2 Delta Variant Recognition of ACE2 Receptor and Broadly Neutralizing Antibodies. *Nature Communications* **2022**, *13*, 1–12, doi:10.1038/s41467-022-28528-w.
6. Liu, C.; Ginn, H.M.; Dejnirattisai, W.; Supasa, P.; Wang, B.; Tuekprakhon, A.; Nutalai, R.; Zhou, D.; Mentzer, A.J.; Zhao, Y.; et al. Reduced Neutralization of SARS-CoV-2 B.1.617 by Vaccine and Convalescent Serum. *Cell* **2021**, *184*, 4220-4236.e13, doi:10.1016/J.CELL.2021.06.020.
7. Saito, A.; Irie, T.; Suzuki, R.; Maemura, T.; Nasser, H.; Uriu, K.; Kosugi, Y.; Shirakawa, K.; Sadamasu, K.; Kimura, I.; et al. Enhanced Fusogenicity and Pathogenicity of SARS-CoV-2 Delta P681R Mutation. *Nature* **2021**, *602*, 300–306, doi:10.1038/s41586-021-04266-9.
8. Eales, O.; Page, A.J.; Martins, L. de O.; Wang, H.; Bodinier, B.; Haw, D.; Jonnerby, J.; Atchison, C.; Consortium, T.C.-19 G.U. (COG-U.; Ashby, D.; et al. SARS-CoV-2 Lineage Dynamics in England from September to November 2021: High Diversity of Delta Sub-Lineages and Increased Transmissibility of AY.4.2. *medRxiv* **2021**, 2021.12.17.21267925, doi:10.1101/2021.12.17.21267925.

9. 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. Spread of a SARS-CoV-2 Variant through Europe in the Summer of 2020. *Nature* **2021**, *595*, 707–712, doi:10.1038/s41586-021-03677-y.
10. Lassaunière, R.; Polacek, C.; Fonager, J.; Bennedbæk, M.; Boding, L.; Rasmussen, M.; Fomsgaard, A. Neutralisation of the SARS-CoV-2 Delta Variant Sub-Lineages AY.4.2 and B.1.617.2 with the Mutation E484K by Comirnaty (BNT162b2 MRNA) Vaccine-Elicited Sera, Denmark, 1 to 26 November 2021. *Euro Surveill.* **2021**, *26*, doi:10.2807/1560-7917.ES.2021.26.49.2101059.
11. Rodríguez-Maldonado, A.P.; Vázquez-Pérez, J.A.; Cedro-Tanda, A.; Taboada, B.; Boukadida, C.; Wong-Arámbula, C.; Nuñez-García, T.E.; Cruz-Ortiz, N.; Barrera-Badillo, G.; Hernández-Rivas, L.; et al. Emergence and Spread of the Potential Variant of Interest (VOI) B.1.1.519 of SARS-CoV-2 Predominantly Present in Mexico. *Archives of Virology* **2021**, *166*, 3173–3177, doi:10.1007/S00705-021-05208-6/FIGURES/2.
12. Taboada-Ramírez, B.; Zárate, S.; Iša, P.; Boukadida, C.; Vázquez-Pérez, J.A.; Muñoz-Medina, J.E.; Ramírez-González, J.E.; Comas-García, A.; Grajales-Muñiz, C.; Rincón-Rubio, A.; et al. Genetic Analysis of SARS-CoV-2 Variants in Mexico during the First Year of the COVID-19 Pandemic. *Viruses* **2021**, *13*, 2161, doi:10.3390/V13112161.
13. NORMA Oficial Mexicana NOM-017-SSA2-2012, Para La Vigilancia Epidemiológica. Available online: http://dof.gob.mx/nota_detalle.php?codigo=5288225&fecha=19/02/2013 (accessed on 15 October 2021).
14. Zárate, S.; Taboada, B.; Muñoz-Medina, J.E.; Iša, P.; Sanchez-Flores, A.; Boukadida, C.; Herrera-Estrella, A.; Mojica, N.S.; Rosales-Rivera, M.; Gómez-Gil, B.; et al. The Alpha Variant (B.1.1.7) of SARS-CoV-2 Failed to Become Dominant in Mexico. *Microbiology Spectrum* **2022**, doi:10.1128/SPECTRUM.02240-21.
15. Taboada, B.; Vazquez-Perez, J.A.; Muñoz-Medina, J.E.; Ramos-Cervantes, P.; Escalera-Zamudio, M.; Boukadida, C.; Sanchez-Flores, A.; Isa, P.; Mendieta-Condado, E.; Martínez-Orozco, J.A.; et al. Genomic Analysis of Early SARS-CoV-2 Variants Introduced in Mexico. *Journal of Virology* **2020**, *94*, doi:10.1128/JVI.01056-20.
16. Langmead, B.; Salzberg, S.L. Fast Gapped-Read Alignment with Bowtie 2. *Nat Methods* **2012**, *9*, 357–359, doi:10.1038/NMETH.1923.
17. Grubaugh, N.D.; Gangavarapu, K.; Quick, J.; Matteson, N.L.; Jesus, J.G. De; Main, B.J.; Tan, A.L.; Paul, L.M.; Brackney, D.E.; Grewal, S.; et al. An Amplicon-Based Sequencing Framework for Accurately Measuring Intrahost Virus Diversity Using PrimalSeq and IVar. *Genome Biology* **2019**, *20*, 1–19, doi:10.1186/S13059-018-1618-7.
18. Shu, Y.; McCauley, J. GISAID: Global Initiative on Sharing All Influenza Data – from Vision to Reality. *Euro Surveill.* **2017**, *22*, 30494, doi:10.2807/1560-7917.ES.2017.22.13.30494.
19. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis*; Use R!; 2nd ed.; Springer International Publishing: Cham, 2016; ISBN 978-3-319-24275-0.
20. Gu, Z.; Eils, R.; Schlesner, M. Complex Heatmaps Reveal Patterns and Correlations in Multidimensional Genomic Data. *Bioinformatics* **2016**, *32*, 2847–2849, doi:10.1093/BIOINFORMATICS/BTW313.
21. Katoh, K.; Standley, D.M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* **2013**, *30*, 772–780, doi:10.1093/MOLBEV/MST010.
22. Minh, B.Q.; Schmidt, H.A.; Chernomor, O.; Schrempf, D.; Woodhams, M.D.; von Haeseler, A.; Lanfear, R. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution* **2020**, *37*, 1530–1534, doi:10.1093/MOLBEV/MSAA015.
23. Kalyaanamoorthy, S.; Minh, B.Q.; Wong, T.K.F.; Haeseler, A. von; Jermiin, L.S. ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates. *Nature Methods* **2017**, *14*, 587–589, doi:10.1038/nmeth.4285.
24. To, T.-H.; Jung, M.; Lycett, S.; Gascuel, O. Fast Dating Using Least-Squares Criteria and Algorithms. *Systematic Biology* **2016**, *65*, 82–97, doi:10.1093/SYSBIO/SYV068.

25. Yu, G.; Smith, D.K.; Zhu, H.; Guan, Y.; Lam, T.T.-Y. Ggtree: An r Package for Visualization and Annotation of Phylogenetic Trees with Their Covariates and Other Associated Data. *Methods in Ecology and Evolution* **2017**, *8*, 28–36, doi:10.1111/2041-210X.12628.
26. Wang, L.-G.; Lam, T.T.-Y.; Xu, S.; Dai, Z.; Zhou, L.; Feng, T.; Guo, P.; Dunn, C.W.; Jones, B.R.; Bradley, T.; et al. Treeio: An R Package for Phylogenetic Tree Input and Output with Richly Annotated and Associated Data. *Molecular Biology and Evolution* **2020**, *37*, 599–603, doi:10.1093/MOLBEV/MSZ240.
27. Leigh, J.W.; Bryant, D. Popart: Full-Feature Software for Haplotype Network Construction. *Methods in Ecology and Evolution* **2015**, *6*, 1110–1116, doi:10.1111/2041-210X.12410.
28. Clement, M.; Snell, Q.; Walke, P.; Posada, D.; Crandall, K. TCS: Estimating Gene Genealogies. *Proceedings - International Parallel and Distributed Processing Symposium, IPDPS 2002* **2002**, 184, doi:10.1109/IPDPS.2002.1016585.
29. Mlcochova, P.; Kemp, S.; Dhar, M.S.; Papa, G.; Meng, B.; Ferreira, I.A.T.M.; Datir, R.; Collier, D.A.; Albecka, A.; Singh, S.; et al. SARS-CoV-2 B.1.617.2 Delta Variant Replication and Immune Evasion. *Nature* **2021**, *599*, 114–119, doi:10.1038/s41586-021-03944-y.
30. Twohig, K.A.; Nyberg, T.; Zaidi, A.; Thelwall, S.; Sinnathamby, M.A.; Aliabadi, S.; Seaman, S.R.; Harris, R.J.; Hope, R.; Lopez-Bernal, J.; et al. Hospital Admission and Emergency Care Attendance Risk for SARS-CoV-2 Delta (B.1.617.2) Compared with Alpha (B.1.1.7) Variants of Concern: A Cohort Study. *The Lancet Infectious Diseases* **2022**, *22*, 35–42, doi:10.1016/S1473-3099(21)00475-8.
31. Ong, S.W.X.; Chiew, C.J.; Ang, L.W.; Mak, T.-M.; Cui, L.; Toh, M.P.H.; Lim, Y.D.; Lee, P.H.; Lee, T.H.; Chia, P.Y.; et al. Clinical and Virological Features of SARS-CoV-2 Variants of Concern: A Retrospective Cohort Study Comparing B.1.1.7 (Alpha), B.1.315 (Beta), and B.1.617.2 (Delta). *Clinical Infectious Diseases* **2021**, doi:10.1093/cid/ciab721.
32. Coronavirus (COVID-19) Vaccinations - Our World in Data Available online: https://ourworldindata.org/covid-vaccinations (accessed on 12 April 2022).
33. Saunders, N.; Planas, D.; Bolland, W.; Rodriguez, C.; Fourati, S.; Buchrieser, J.; Planchais, C.; Prot, M.; Staropoli, I.; Guivel-Benhassine, F.; et al. Fusogenicity and Neutralization Sensitivity of the SARS-CoV-2 Delta Sublineage AY.4.2. *bioRxiv* **2022**, 2022.01.07.475248, doi:10.1101/2022.01.07.475248.
34. Singh, P.; Sharma, K.; Singh, P.; Bhargava, A.; Negi, S.S.; Sharma, P.; Bhise, M.; Tripathi, M.K.; Jindal, A.; Nagarkar, N.M. Genomic Characterization Unravelling the Causative Role of SARS-CoV-2 Delta Variant of Lineage B.1.617.2 in 2nd Wave of COVID-19 Pandemic in Chhattisgarh, India. *Microbial Pathogenesis* **2022**, *164*, 105404, doi:10.1016/J.MICPATH.2022.105404.
35. Planas, D.; Veyer, D.; Baidaliuk, A.; Staropoli, I.; Guivel-Benhassine, F.; Rajah, M.M.; Planchais, C.; Porrot, F.; Robillard, N.; Puech, J.; et al. Reduced Sensitivity of SARS-CoV-2 Variant Delta to Antibody Neutralization. *Nature* **2021**, *596*, 276–280, doi:10.1038/s41586-021-03777-9.
36. Ginex, T.; Marco-Marín, C.; Wieczór, M.; Mata, C.P.; Krieger, J.; López-Redondo, M.L.; Francés-Gómez, C.; Ruiz-Rodriguez, P.; Melero, R.; Sánchez-Sorzano, C.Ó.; et al. The Structural Role of SARS-CoV-2 Genetic Background in the Emergence and Success of Spike Mutations: The Case of the Spike A222V Mutation. *bioRxiv* **2021**, 2021.12.05.471263, doi:10.1101/2021.12.05.471263.
37. Syed, A.M.; Taha, T.Y.; Tabata, T.; Chen, I.P.; Ciling, A.; Khalid, M.M.; Sreekumar, B.; Chen, P.Y.; Hayashi, J.M.; Soczek, K.M.; et al. Rapid Assessment of SARS-CoV-2–Evolved Variants Using Virus-like Particles. *Science (1979)* **2021**, *374*, 1626–1632, doi:10.1126/SCIENCE.ABL6184/SUPPL_FILE/SCIENCE.ABL6184_MDAR_REPRODUCIBILITY_CHECKLIST.PDF.

 doi:10.20944/preprints202205.0131.v1

38. Zhao, H.; Nguyen, A.; Wu, D.; Li, Y.; Hassan, S.A.; Chen, J.; Shroff, H.; Piszczek, G.; Schuck, P. Plasticity in Structure and Assembly of SARS-CoV-2 Nucleocapsid Protein. *bioRxiv* **2022**, 2022.02.08.479556, doi:10.1101/2022.02.08.479556.