



## OPEN ACCESS

## EDITED BY

Pragya Dhruv Yadav,  
ICMR-National Institute of Virology, India

## REVIEWED BY

Juan Carlos Hurtado,  
Hospital Clinic of Barcelona, Spain  
Alexei Galatenko,  
Lomonosov Moscow State University, Russia

## \*CORRESPONDENCE

Jose Arturo Molina-Mora  
✉ jose.molinamora@ucr.ac.cr  
Alfredo Herrera-Estrella  
✉ alfredo.herrera@cinvestav.mx

## SPECIALTY SECTION

This article was submitted to  
Infectious Diseases: Epidemiology and  
Prevention,  
a section of the journal  
Frontiers in Public Health

RECEIVED 11 November 2022

ACCEPTED 06 February 2023

PUBLISHED 02 March 2023

## CITATION

Molina-Mora JA, Reales-González J,  
Camacho E, Duarte-Martínez F, Tsukayama P,  
Soto-Garita C, Brenes H, Cordero-Laurent E,  
Ribeiro dos Santos A, Guedes Salgado C, Santos  
Silva C, Santana de Souza J, Nunes G, Negri T,  
Vidal A, Oliveira R, Oliveira G,  
Muñoz-Medina JE, Salas-Lais AG,  
Mireles-Rivera G, Sosa E, Turjanski A,  
Monzani MC, Carobene MG, Remes Lenicov F,  
Schottlender G, Fernández Do Porto DA,  
Kreuze JF, Sacristán L, Guevara-Suarez M,  
Cristancho M, Campos-Sánchez R and  
Herrera-Estrella A (2023) Overview of the  
SARS-CoV-2 genotypes circulating in Latin  
America during 2021.  
*Front. Public Health* 11:1095202.  
doi: 10.3389/fpubh.2023.1095202

## COPYRIGHT

© 2023 Molina-Mora, Reales-González,  
Camacho, Duarte-Martínez, Tsukayama,  
Soto-Garita, Brenes, Cordero-Laurent, Ribeiro  
dos Santos, Guedes Salgado, Santos Silva,  
Santana de Souza, Nunes, Negri, Vidal, Oliveira,  
Oliveira, Muñoz-Medina, Salas-Lais,  
Mireles-Rivera, Sosa, Turjanski, Monzani,  
Carobene, Remes Lenicov, Schottlender,  
Fernández Do Porto, Kreuze, Sacristán,  
Guevara-Suarez, Cristancho, Campos-Sánchez  
and Herrera-Estrella. This is an open-access  
article distributed under the terms of the  
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).  
The use, distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in this  
journal is cited, in accordance with accepted  
academic practice. No use, distribution or  
reproduction is permitted which does not  
comply with these terms.

# Overview of the SARS-CoV-2 genotypes circulating in Latin America during 2021

Jose Arturo Molina-Mora<sup>1\*</sup>, Jhonnatan Reales-González<sup>2</sup>,  
Erwin Camacho<sup>3</sup>, Francisco Duarte-Martínez<sup>4</sup>, Pablo Tsukayama<sup>5</sup>,  
Claudio Soto-Garita<sup>4</sup>, Hebleen Brenes<sup>4</sup>, Estela Cordero-Laurent<sup>4</sup>,  
Andrea Ribeiro dos Santos<sup>6</sup>, Cláudio Guedes Salgado<sup>6</sup>,  
Caio Santos Silva<sup>6</sup>, Jorge Santana de Souza<sup>7</sup>, Gisele Nunes<sup>8</sup>,  
Tatianne Negri<sup>8</sup>, Amanda Vidal<sup>8</sup>, Renato Oliveira<sup>8</sup>,  
Guilherme Oliveira<sup>8</sup>, José Esteban Muñoz-Medina<sup>9</sup>,  
Angel Gustavo Salas-Lais<sup>9</sup>, Guadalupe Mireles-Rivera<sup>10</sup>,  
Ezequiel Sosa<sup>11,12</sup>, Adrián Turjanski<sup>11,12</sup>, María Cecilia Monzani<sup>12,13</sup>,  
Mauricio G. Carobene<sup>12,13</sup>, Federico Remes Lenicov<sup>12,13</sup>,  
Gustavo Schottlender<sup>11</sup>, Darío A. Fernández Do Porto<sup>11</sup>,  
Jan Frederik Kreuze<sup>14</sup>, Luisa Sacristán<sup>15</sup>,  
Marcela Guevara-Suarez<sup>15</sup>, Marco Cristancho<sup>15</sup>,  
Rebeca Campos-Sánchez<sup>16</sup> and Alfredo Herrera-Estrella<sup>10\*</sup>

<sup>1</sup>Centro de investigación en Enfermedades Tropicales and Facultad de Microbiología, Universidad de Costa Rica, San José, Costa Rica, <sup>2</sup>Grupo de Genómica de Microorganismos Emergentes, Instituto Nacional de Salud, Bogotá, Colombia, <sup>3</sup>Investigaciones Biomédicas, Universidad de Sucre, Sincelejo, Colombia, <sup>4</sup>Laboratorio de Genómica y Biología Molecular, Instituto Costarricense de Investigación y Enseñanza en Nutrición y Salud, Tres Ríos, Cartago, Costa Rica, <sup>5</sup>Facultad de Ciencias y Filosofía, Universidad Peruana Cayetano Heredia, Lima, Peru, <sup>6</sup>Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém, Brazil, <sup>7</sup>Instituto Metrópole Digital, Universidade Federal do Rio Grande do Norte, Natal, Brazil, <sup>8</sup>Environmental Genomics, Vale Institute of Technology, Belém, Pará, Brazil, <sup>9</sup>Coordinación de Calidad de Insumos y Laboratorios Especializados, Instituto Mexicano del Seguro Social, Ciudad de México, Mexico, <sup>10</sup>Laboratorio Nacional de Genómica para la Biodiversidad-Unidad de Genómica Avanzada, Centro de Investigación y de Estudios Avanzados, Irapuato, Mexico, <sup>11</sup>Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina, <sup>12</sup>Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina, <sup>13</sup>Facultad de Medicina de la Universidad de Buenos Aires, Buenos Aires, Argentina, <sup>14</sup>International Potato Center, Lima, Peru, <sup>15</sup>Vicerrectoría de Investigación y Creación, Universidad de Los Andes, Bogotá, Colombia, <sup>16</sup>Centro de Investigación en Biología Celular y Molecular, Universidad de Costa Rica, San José, Costa Rica

Latin America is one of the regions in which the COVID-19 pandemic has a stronger impact, with more than 72 million reported infections and 1.6 million deaths until June 2022. Since this region is ecologically diverse and is affected by enormous social inequalities, efforts to identify genomic patterns of the circulating SARS-CoV-2 genotypes are necessary for the suitable management of the pandemic. To contribute to the genomic surveillance of the SARS-CoV-2 in Latin America, we extended the number of SARS-CoV-2 genomes available from the region by sequencing and analyzing the viral genome from COVID-19 patients from seven countries (Argentina, Brazil, Costa Rica, Colombia, Mexico, Bolivia, and Peru). Subsequently, we analyzed the genomes circulating mainly during 2021 including records from GISAID database from Latin America. A total of 1,534 genome sequences were generated from seven countries, demonstrating the laboratory and bioinformatics capabilities for genomic surveillance of pathogens that have been developed locally. For Latin America, patterns regarding several variants associated with multiple re-introductions, a relatively low percentage of sequenced samples, as well as an increment in the mutation frequency since the

beginning of the pandemic, are in line with worldwide data. Besides, some variants of concern (VOC) and variants of interest (VOI) such as Gamma, Mu and Lambda, and at least 83 other lineages have predominated locally with a country-specific enrichments. This work has contributed to the understanding of the dynamics of the pandemic in Latin America as part of the local and international efforts to achieve timely genomic surveillance of SARS-CoV-2.

#### KEYWORDS

COVID-19, Latin America, SARS-CoV-2, genomic surveillance, CABANA, coronavirus

## Introduction

In December 2019, several cases of a new respiratory illness were described in Wuhan, China. About a month later, it was confirmed that the illness COVID-19 (coronavirus disease 2019) was caused by a novel coronavirus which was subsequently named SARS-CoV-2 (1, 2). Until June 2022, the COVID-19 pandemic had impacted the world with >549 million confirmed cases of COVID-19, including >6.3 million deaths. Latin America was one of the most strongly impacted regions with more than 72 million reported infections and >1.6 million deaths during the same period.

SARS-CoV-2 genome sequences have been reported from many regions of the world and these data have been proven useful in tracking the global spread of the virus. Genomic epidemiology of SARS-CoV-2 has shed light on the origins of regional outbreaks, global dispersal, and epidemiological history of the virus (3, 4). Until April 2022, over 11.5 million genomes had been deposited in the GISAID database (<https://www.gisaid.org/>), out of which >376,000 were reported by Latin American countries.

Since its appearance, a large genetic diversity has been recognized for SARS-CoV-2 due to widespread transmission and geographical isolation (5). The emergence of new genotypes (lineages, clades, variants, etc.) is the product of a natural process that occurs when viruses replicate at high rates as it happens during a pandemic (4). The World Health Organization (WHO) has classified five divergent genotypes as variants of concern (VOC: Alpha, Beta, Gamma, Delta, Omicron), as well as some lineages into variants of interest (VOI: Lambda, Mu, Epsilon, Zeta, Theta, Iota, Eta, Kappa, and others) and variants under monitoring (VUM: B.1.640 and XD) (6). All reported variants and other lineages have been identified in Latin America (7), including genotypes that were first reported regionally, such as Gamma in Brazil, Mu in Colombia, and Lambda in Peru (6), as well as unique lineages in Costa Rica and Central America (8, 9). Those descriptions of locally enriched genotypes exemplify the opportunities that SARS-CoV-2 has found in Latin America for spreading and evolving. This scenario is in part explained by the complex environmental and human reality in this region, with huge ecological diversity and social inequalities (10, 11). Thus, efforts on revealing the behavior of SARS-CoV-2 are necessary to identify regionally emerging patterns for the suitable management of the pandemic, which cannot be inferred from North America, Europe, or Asia (11).

In this context, the CABANA initiative (Capacity building for Bioinformatics in Latin America, Global Challenges Research

Fund GCRF: [www.cabana.online](http://www.cabana.online)) supported the development of a regional project titled “The SARS-CoV-2 genome, its evolution and epidemiology in Latin America” during 2021. The project had the direct participation of seven institutions from Argentina, Brazil, Bolivia, Colombia, Costa Rica, Mexico, and Peru. Efforts of this project included not only the sequencing and genome assembly of the SARS-CoV-2 virus from a total of 1,534 COVID-19 cases in those countries, but also to bring a more complete overview of the SARS-CoV-2 genotypes circulating in Latin America during 2021 using public databases. Thus, this study aimed to contribute to the genomic surveillance of the SARS-CoV-2 to understand the dynamics of the pandemic in Latin America by providing genome sequences and analyzing circulating genotypes during the year 2021.

## Methods

### Samples and ethical considerations

Respiratory samples were obtained from public and private laboratories belonging to the national network of SARS-CoV-2 diagnostics in each country. Adequate transportation and storage conditions were guaranteed to preserve the samples. Every sample was anonymized to protect patients' identity. Being a notifiable disease, the metadata was collected from the forms that accompanied the samples, either in the national reference laboratories or in the ministries of health. See [Supplementary material](#) for IDs to access metadata in the GISAID database.

### Sample sequencing and genome analysis

To contribute with SARS-CoV-2 genome sequences from Latin America, seven participant countries (Argentina, Bolivia, Brazil, Costa Rica, Colombia, Mexico, and Peru) were involved in sample processing from COVID-19 patients. Diagnosis using RT-qPCR, genome sequencing and assembly, as well as genotyping, were implemented using the laboratory protocols and bioinformatic pipelines that are being locally used as part of the genomic surveillance efforts in each country as shown in [Table 1](#) and reported in (9, 12, 13). Genome sequences were uploaded to the GISAID database (<https://www.gisaid.org/>). Details regarding the number of processed samples (assembled genomes), laboratory

and bioinformatic protocols for each country are summarized in [Table 1](#). GISAID accession numbers (ID) for assembled genomes are presented in [Supplementary material](#).

## Analysis of circulating SARS-CoV-2 genotypes in Latin America

To gain insights into the SARS-CoV-2 genotypes circulating in Latin America during 2021, a general analysis was done using the genome sequences available at the GISAID database (<https://www.gisaid.org/>). Selection of countries, statistics of sequenced samples, and plots of circulating genomes and mutation frequency were obtained using the tools of the GISAID platform. The number of COVID-19 cases per country was retrieved from the daily reports of the Pan American Health Organization (14). All analyses were performed considering sequences collected until January 31st, 2022. PANGOLIN lineage database (15, 16) was used to analyze the frequency of lineages among countries.

## Results and discussion

Genomic surveillance has been a hallmark of the COVID-19 pandemic that, in contrast to other pandemics, achieves tracking of the virus evolution and spread worldwide almost in real-time (4).

In this work, we extended the repertoire of SARS-CoV-2 genome sequences with a total of 1,534 sequences from seven Latin American countries ([Table 1](#)). Whereas, this was a relatively modest contribution to the overall quantity of sequences produced in this period in Latin America for certain time-intervals and countries it provided important complementarity for the genomic surveillance of the virus. In Bolivia for example, our efforts represented 38% of all sequences produced over this time. To perform a more complete examination, we included all sequences from Latin America available at the GISAID database collected up to January 2022. A total of 221,228 genomes sequences, including the 1,534 provided by this work, were analyzed by genotype and the mutation profile.

According to the GISAID database records, the numbers of sequences is still small in comparison to the number of diagnosed cases in Latin America ([Table 2](#)). On average, only 0.39% of COVID-19 cases in Latin America had been sequenced, with Mexico and Chile having the highest rates with 0.98 and 0.92%, respectively. In the case of Nicaragua, in which the pandemic has been downplayed (17, 18), the reports of diagnosed patients and other statistics are considered unrealistic, including the 2.92% of sequenced samples. Thus, we did not conduct comparisons of Nicaragua among other countries due to the extremely biased data.

On the other extreme, Bolivia, Honduras and Venezuela have barely sequenced even 0.03% of samples derived from all patients diagnosed with the disease. There is no single Latin American country that has sequenced more samples, relative to the number

**TABLE 1** Sequencing strategy and bioinformatic pipelines used for the genomic surveillance of the SARS-CoV-2 in five Latin American countries, CABANA initiative.

Country	Number of samples	Participant institutions	Sequencing protocol	Bioinformatic pipeline
Argentina	220	Universidad de Buenos Aires (UBA) and Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)	Illumina platform: Nextera XT DNA library and MiSeq sequencer (Illumina, San Diego, CA, USA).	Genome assembly, variant calling, and genotyping: Custom protocol with BWA-MEM/Freebayes/SNPeff.
Bolivia	94	Laboratorio de diagnóstico e investigación BIOSCIENCE SRL and Laboratorio de Genómica Microbiana, Universidad Peruana Cayetano Heredia	Illumina platform: Illumina COVIDSeq Test and Illumina NextSeq 550 sequencer (Illumina, Inc., San Diego, CA, USA).	Genome assembly, variant calling, and genotyping: Illumina DRAGEN COVID Lineage v3.5.3 BaseSpace App.
Brazil	167	Vale Institute of Technology, Belém, PA, Brazil	Illumina platform, following a custom protocol and Illumina NextSeq sequencer.	Genome assembly, variant calling, and genotyping: PipeCoV pipeline described in (12).
Costa Rica	190	Universidad de Costa Rica (UCR) and Instituto Costarricense de Investigación y Enseñanza en Nutrición y Salud (INCIENSA)	Illumina platform, following the protocol described in (9, 13) with a Illumina DNA Prep Kit/Nextera DNA flex library and MiSeq sequencer (Illumina, San Diego, CA, USA).	Genome assembly, variant calling, genotyping, and phylogeny: Described in (9)
Colombia	147	Universidad de Los Andes.	Nanopore platform with a GridION sequencer.	Genome assembly: custom protocol for long-reads sequencing with Minimap2/Nanopolish. Variant calling, genotyping, and phylogeny: (9)
Mexico	472	Unidad de Genómica Avanzada del Centro de Investigación y de Estudios Avanzados (UGA-LANGEBIO, CINVESTAV)	Illumina platform: Illumina COVIDSeq test and Illumina MiSeq sequencer (Illumina, Inc., San Diego, CA, USA).	Genome assembly, variant calling, genotyping, and phylogeny: Described in (9)
Peru	244	Laboratorio de Genómica Microbiana, Universidad Peruana Cayetano Heredia	Illumina platform: Illumina COVIDSeq Test and Illumina NextSeq 550 sequencer (Illumina, Inc., San Diego, CA, USA).	Genome assembly, variant calling, and genotyping: Illumina DRAGEN COVID Lineage v3.5.3 BaseSpace App.

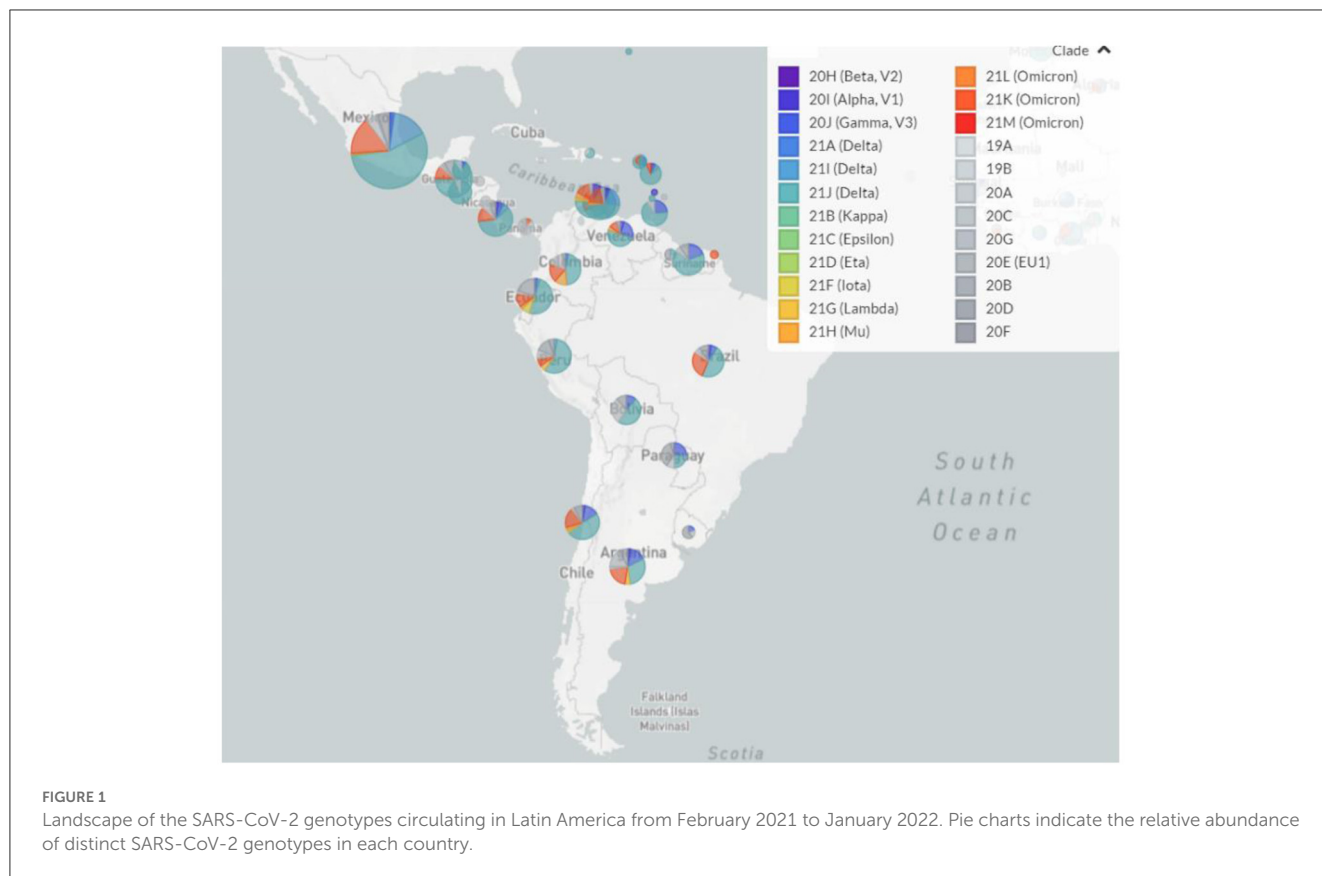
TABLE 2 Comparison of COVID-19 cases and sequenced samples among Latin American countries from the beginning of the pandemic to January 31th, 2022.

Country	Population	Total COVID-19 cases			Sequenced samples	
		Absolute number	Percentage of the population (%)	Deaths	Absolute number	Percentage of COVID-19 cases (%)
Mexico	131,026,542	4,942,590	3.77	306,091	48,329	0.98
Chile	19,369,864	2,190,561	11.31	39,733	20,086	0.92
Belize	408,778	50,487	12.35	625	376	0.74
Ecuador	18,057,002	732,038	4.05	34,533	3,770	0.52
Peru	33,681,601	3,239,538	9.62	205,834	14,790	0.46
Brazil	214,895,351	25,454,105	11.84	627,589	101,532	0.40
Costa Rica	5,166,024	694,865	13.45	7,575	2,719	0.39
El Salvador	6,536,807	135,109	2.07	3,899	308	0.23
Colombia	51,721,161	5,887,261	11.38	13,400	12,520	0.21
Panama	4,419,710	700,274	15.84	7,732	1,223	0.17
Guatemala	18,427,485	690,290	3.75	16,385	1,420	0.21
Paraguay	7,267,941	583,662	8.03	17,321	882	0.15
Uruguay	3,492,345	668,425	19.14	6,479	717	0.11
Argentina	45,836,859	8,378,656	18.28	121,273	11,553	0.14
Honduras	10,147,994	391,874	3.86	10,504	116	0.03
Venezuela	28,311,334	485,974	1.72	5,447	123	0.03
Bolivia	11,919,079	855,705	7.18	20,951	248	0.03
Nicaragua	6,746,365	17,650	0.26	216	516	2.92
Latin America	617,432,242	56,099,064	9.09	1,445,587	221,228	0.39
Worldwide	7,900,000,000	380,099,991	4.81	5,695,345	7,748,697	2.04

of cases reported, than the world average that corresponds to 2.04%, which is low too. The current scenario is congruent with a previous report with <0.5% of sequenced samples for Latin American countries (19). These findings represent not only part of the regional disparities in the SARS-CoV-2 genomic surveillance efforts in Latin America, but also that this geographic region needs to increase the effort to achieve the sequencing of at least 5% of positive samples to detect emerging viral lineages when their prevalence is <1% of all strains in a population, as suggested previously (20). In fact, globally, only 6.8% of 189 countries around the world reached this value (19). This situation is like that of other latitudes around the world in which only a very small portion of the countries has reached the recommended percentage, suggesting that sequencing at least 0.5% of the cases, with a time in days between sample collection and genome submission <21 days, could be a benchmark for SARS-CoV-2 genomic surveillance efforts for low- and middle-income countries (19) taking into account the high cost of sequencing reagents and equipment in these countries. In high income countries, around 25% of the genomes were submitted within 21 days, contrasting with the pattern observed in 5% of the genomes from low- and middle-income countries. Thus, the identification of patterns regarding the circulating genotypes in Latin America should be interpreted with cautions due the differences of SARS-CoV-2 surveillance systems,

including sequencing capacity and sampling strategies between countries in the region.

Regarding the circulating genotypes, the reports on the diversity of lineages are similar to other studies in Latin America (9, 21–23) and other distant geographic regions (24, 25). For divergent SARS-CoV-2 genomes, all VOCs have been reported in all Latin American countries, resulting in a large diversity of genotypes circulating in each country (Figure 1). This is in line with the expected pattern of multiple and independent re-introductions due to population mobility within Latin America, as well as to and from other countries and continents (26–28). Besides, some genotypes have been reported with an epicenter in Latin America. As presented in Figure 2, those country-specific variants were predominant in the first semester of 2021, such as the Gamma variant in Brazil until August 2021, the Mu variant in Colombia from April to September 2021, and Lambda in Peru during the period from March to June 2021 (29). Other remarkable genome versions were the case of the Gamma variant predominating between June and August 2021 in Argentina, as well as the more mixed pattern with distinct variants in Mexico, similar to the average for the entire Latin American region. In comparison to the rest of the world, Latin America reported similar transitions between the Alpha, Delta, and Omicron variants. Nonetheless, the increased reports of Mu and Lambda



in this region were minimal for the worldwide representation (Figure 2).

In Latin America recurrent dissemination of SARS-CoV-2 through shared borders between countries has been evidenced (30), allowing rapid entrance and dissemination of different lineages to the different countries (31). Territories with no restriction to international interchange are more likely to introduce multiple SARS-CoV-2 variants, including variants of concern and/or interest and even lineages with mutations of concern and emerging variants with different mutation patterns (32). These introductions of VOCs to Latin America were more evident during the second half of the year 2021, where the Delta variant displaced other variants in several countries and became predominant as shown in Figure 2, while during the first semester of the year lineage predominance varied among these countries.

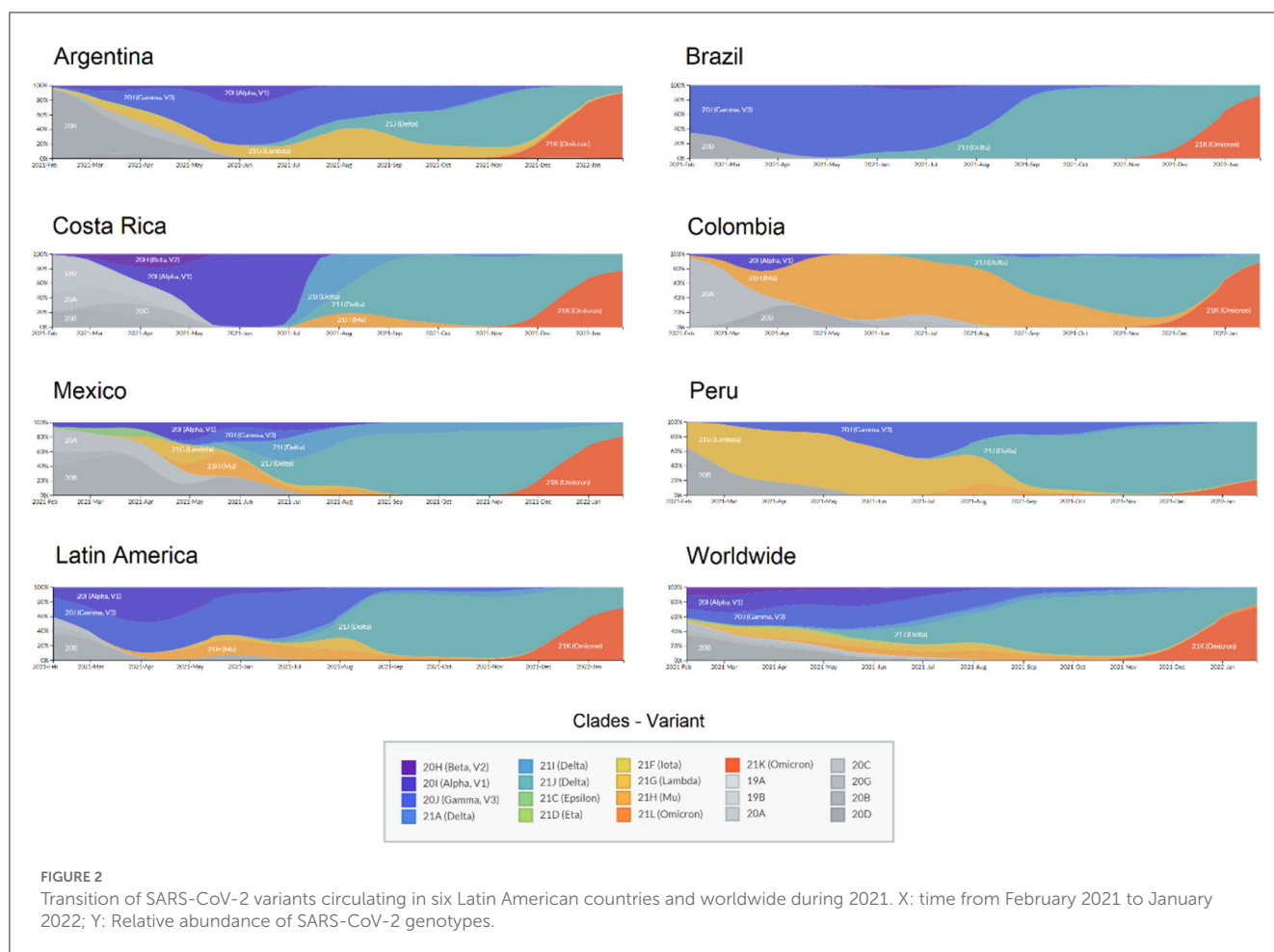
Although several epidemiological aspects can be associated with these patterns, the extensive opening of the borders during the middle of 2021 possibly favored the spread of new variants of concern in the region. Besides, the presence of multiple mutations that have been associated with increased infectivity and/or escape from immune response in variants such as Delta (33) helped this variant to displace other variants, as it occurred worldwide.

For other genotypes, at least 83 out of >1,500 PANGOLIN lineages have been reported with a high predominance in a Latin American country (Table 3). The full list of lineages is presented in the Supplementary material. As an example, lineage C.39 was predominant in Chile with 45.0% of all the sequences reported,

followed by France (15.0%), Peru (12.0%), Guinea (10.0%), and Germany (8.0%). From these lineages, at least 80% of the sequences from 51 lineages have been reported to come from a Latin American country (Tables 3, 4). In the distribution by country, Brazil, Peru, Chile, Costa Rica, and Mexico have more reports of lineages with a frequency >80% locally.

For instance, Peru had 97.0% of the sequences reported for lineage C.40 and 95% of lineage C.4. Also, Peru was the main country in which the AY.25.1.1 and AY.26.1 genotypes (Delta sub-lineages) were documented. Brazil reported 97% of the 22,815 cases of lineage AY.99.2 (firstly reported in Colombia), that was demonstrated to successfully disseminate among different locations in the country (34, 35). Lineages derived from the Gamma variant were also reported frequently in Brazil (e.g., P.1.4, P.1.7, P.4, and others). During 2020 the lineage B.1.1.389, which harbors the specific mutation spike:T1117, was reported as predominant in Costa Rica (86% of cases of this lineage were reported in this country) (9). Despite its dominance, few changes were predicted on the virus behavior (transmission, immune response, and other) and it was quickly replaced by the lineage Central America and subsequently by VOCs such as Alpha and Gamma (8). In the case of Mexico, two lineages (P.1.10.2 and B.1.243.2) were mainly found in this country (frequency >80%) but in a limited number. Lineage B.1.1.519 was a relevant genotype reported in Mexico, despite it was mainly reported in the United States. This version predominated in Mexico during the first quarter of 2021 while the Alpha variant (B.1.1.7) was also spreading. Interestingly, unlike other cases, the Alpha variant did not displace B.1.1.519 in this





**TABLE 3** Number of lineages in which a Latin American country is predominant by frequency.

Country	Lineages in which the country is predominant*	Lineages in which the country has frequency >80%**
Argentina	4	3
Brazil	30	21
Chile	11	6
Colombia	2	0
Costa Rica	6	3
Ecuador	1	0
Mexico	5	2
Panama	1	1
Peru	21	13
Uruguay	2	2
Total	83	51

\*Refers to those lineages in which the main source of sequences is from a Latin American country. See text for details.

\*\*Similar to the previous case, but lineages are only counted if the predominant country has a percentage >80%. See text for details.

country (36). B.1.1.519 was assigned as a VUM by WHO in 02-Jun-2021 and was degraded to a FMV (formerly monitoring variant) on 9-Nov-2021 (<https://www.who.int/activities/tracking-SARS-CoV-2-variants>).

Jointly, these results indicate that specific mutations and the subsequent consolidation into lineages were detected in Latin America and evidenced by genomic surveillance in the region. Interestingly, 17 of these lineages were first reported in a different country from where it was subsequently found to be predominant (>80%). This includes neighboring countries, such as the case of lineage P.1.7.1 which was enriched in Peru but was first reported in Bolivia. This pattern was more frequent for Brazil, with eight lineages that were first reported in other countries including from Europe and Asia, but that became dominant in this country.

Tracking of specific mutations into Latin American lineages that could be used as local markers, may help to identify transmission networks locally and globally, highlighting the need for each country and territory to strengthen the sequencing and bioinformatic capacities. These capacities can also be of use to locally study other scenarios such as clinical profiles for COVID-19 patients (37), immune escape (38), long-term COVID-19 (39), identification of co-infections (40) or identify

TABLE 4 Lineages reported with a frequency &gt;80% among Latin American countries.

Lineage	Most common countries	First detected—country	Total sequences worldwide
A.2.4	Panama 96.0%, Costa_Rica 2.0%, Northern_Mariana_Islands 1.0%, United States of America 1.0%, United Kingdom 0.0%	Panama	291
C.4	Peru 95.0%, United States of America 2.0%, Turkey 2.0%, Republic of Serbia 1.0%, Switzerland 1.0%	Peru	130
C.11	Chile 97.0%, Peru 2.0%, Argentina 1.0%	Chile	95
C.13	Peru 84.0%, United States of America 14.0%, Japan 2.0%	Peru	45
C.14	Peru 92.0%, United States of America 3.0%, Japan 2.0%, Democratic Republic of the Congo 1.0%, Brazil 1.0%	Peru	253
C.25	Peru 100.0%	Peru	8
C.29	Chile 89.0%, Russia 6.0%, Denmark 6.0%	Chile	18
C.32	Peru 95.0%, Turkey 2.0%, Denmark 2.0%, United Kingdom 2.0%	Peru	57
C.33	Peru 90.0%, Japan 5.0%, United States of America 5.0%	Peru	20
C.40	Peru 97.0%, Chile 3.0%	Peru	37
P.1.3	Brazil 100.0%	Brazil	29
P.1.4	Brazil 98.0%, Costa_Rica 1.0%, Peru 1.0%, United States of America 0.0%, Colombia 0.0%	Brazil	1039
P.1.5	Brazil 100.0%	Brazil	14
P.1.6	Brazil 99.0%, Sweden 0.0%, French_Guiana 0.0%	Brazil	562
P.1.7	Brazil 91.0%, United States of America 5.0%, Spain 2.0%, Peru 0.0%, Mexico 0.0%	Japan	3666
P.1.7.1	Peru 94.0%, United States of America 2.0%, Brazil 1.0%, Chile 1.0%, Puerto_Rico 0.0%	Bolivia	773
P.1.8	Brazil 94.0%, Sweden 1.0%, France 1.0%, United Kingdom 1.0%, Spain 1.0%	Brazil	257
P.1.9	Brazil 97.0%, United States of America 1.0%, Mexico 1.0%, Belgium 1.0%	Brazil	236
P.1.10.2	Mexico 100.0%	Mexico	23
P.1.11	Brazil 96.0%, Ecuador 2.0%, Spain 1.0%, United States of America 1.0%	Brazil	102
P.1.12.1	Peru 91.0%, United States of America 4.0%, Italy 2.0%, Chile 1.0%, Switzerland 1.0%	Italy	161
P.4	Brazil 100.0%, United States of America 0.0%	Brazil	234
P.5	Brazil 95.0%, Philippines 3.0%, United States of America 2.0%	Brazil	44
P.6	Uruguay 95.0%, Philippines 2.0%, United States of America 1.0%, Norway 0.0%, Spain 0.0%	Uruguay	298
B.1.1.33	Brazil 83.0%, United States of America 5.0%, Chile 3.0%, Argentina 2.0%, Paraguay 1.0%	United States	2130
N.3	Argentina 97.0%, Bolivia 2.0%, Hong_Kong 1.0%, Chile 1.0%	Argentina	124
N.4	Chile 92.0%, Brazil 4.0%, United States of America 1.0%, Peru 1.0%, New_Zealand 0.0%	Canada	232
N.5	Argentina 87.0%, United States of America 7.0%, Spain 1.0%, Italy 1.0%, United Kingdom 1.0%	India	365
N.6	Chile 97.0%, Brazil 1.0%, Japan 1.0%, Paraguay 1.0%	Chile	138
N.7	Uruguay 100.0%	Uruguay	33
N.9	Brazil 96.0%, Ireland 1.0%, Argentina 1.0%, Japan 1.0%, Chile 1.0%	Brazil	138
N.10	Brazil 97.0%, Germany 3.0%	Brazil	22
B.1.1.110	Peru 92.0%, Finland 8.0%	Finland	11
B.1.1.324	Chile 100.0%	Chile	17
B.1.1.332	Brazil 100.0%	Brazil	29
B.1.1.389	Costa_Rica 86.0%, United States of America 6.0%, Spain 5.0%, Finland 2.0%, Australia 1.0%	Costa Rica	139
B.1.1.442	Argentina 80.0%, Turkey 7.0%, United States of America 5.0%, Germany 5.0%, Austria 2.0%	Argentina	47
B.1.1.516	Costa_Rica 100.0%	Nicaragua	21

(Continued)

TABLE 4 (Continued)

Lineage	Most common countries	First detected—country	Total sequences worldwide
B.1.110.1	Chile 82.0%, United States of America 18.0%	Chile	12
B.1.205	Peru 98.0%, Israel 2.0%	Peru	40
B.1.243.2	Mexico 81.0%, United States of America 19.0%	Mexico	60
B.1.291	Costa_Rica 94.0%, United States of America 4.0%, Australia 2.0%	Nicaragua	67
AY.25.1.1	Peru 82.0%, United States of America 14.0%, Chile 2.0%, Colombia 1.0%, Switzerland 1.0%	USA	172
AY.26.1	Peru 81.0%, United States of America 12.0%, Mexico 2.0%, Israel 1.0%, Chile 1.0%	Peru	154
AY.43.1	Brazil 96.0%, United Kingdom 1.0%, United States of America 1.0%, France 1.0%, Chile 1.0%	Poland	1036
AY.43.2	Brazil 99.0%, France 0.0%, Japan 0.0%, Belgium 0.0%, Switzerland 0.0%	India	1295
AY.43.7	Brazil 98.0%, Netherlands 1.0%, France 1.0%, Israel 1.0%	France	184
AY.46.3	Brazil 95.0%, United States of America 1.0%, Turkey 1.0%, India 0.0%, Czech_Republic 0.0%	India	1565
AY.99.1	Brazil 85.0%, United Kingdom 10.0%, United States of America 4.0%, India 1.0%, Spain 0.0%	India	1372
AY.99.2	Brazil 97.0%, United States of America 1.0%, Chile 0.0%, France 0.0%, Portugal 0.0%	Brazil	22815
AY.101	Brazil 81.0%, Chile 10.0%, Colombia 4.0%, Peru 2.0%, United States of America 1.0%	Colombia	4300

recombinant genomes (a recognized mechanism of viral diversity in coronaviruses).

Despite the reports of differences in the enriched genotypes in the first half of 2021, the emergence of new variants of the viral genome in Latin America was consistent with the rest of the world inferred from the mutation frequency (Figure 3A). During 2020, the mutation frequency for the S1 region of the spike gene was estimated at around 2-3 mutations per month. At the beginning of 2021, this frequency increased to 8.32 and subsequently to around 12 with the predominance of Delta. However, with the arrival of the Omicron variant, the frequency at the very end of 2021 and the first month of 2022 reached values of 28 mutations, in both, Latin America and the world. Thus, this accumulative divergence has impacted the mutation rate over the pandemic, which until January 2022 was estimated to be around  $8.74 \times 10^{-4}$  substitutions per site per year (Figure 3B). This mutation frequency and rate values are consistent with other local and global studies during the pandemic (41–43), including the rate of  $0.8 - 2.38 \times 10^{-3}$  substitutions per site per year described by Banerjee et al. (44). Following the gradual reopening of borders and worldwide travels, the frequency of infections and the appearance of mutations and new genotypes are expected to increase (45). Thus, more genome sequencing studies, including robust metadata collection, and more financial support are needed to continue with the surveillance of the pandemic in Latin America.

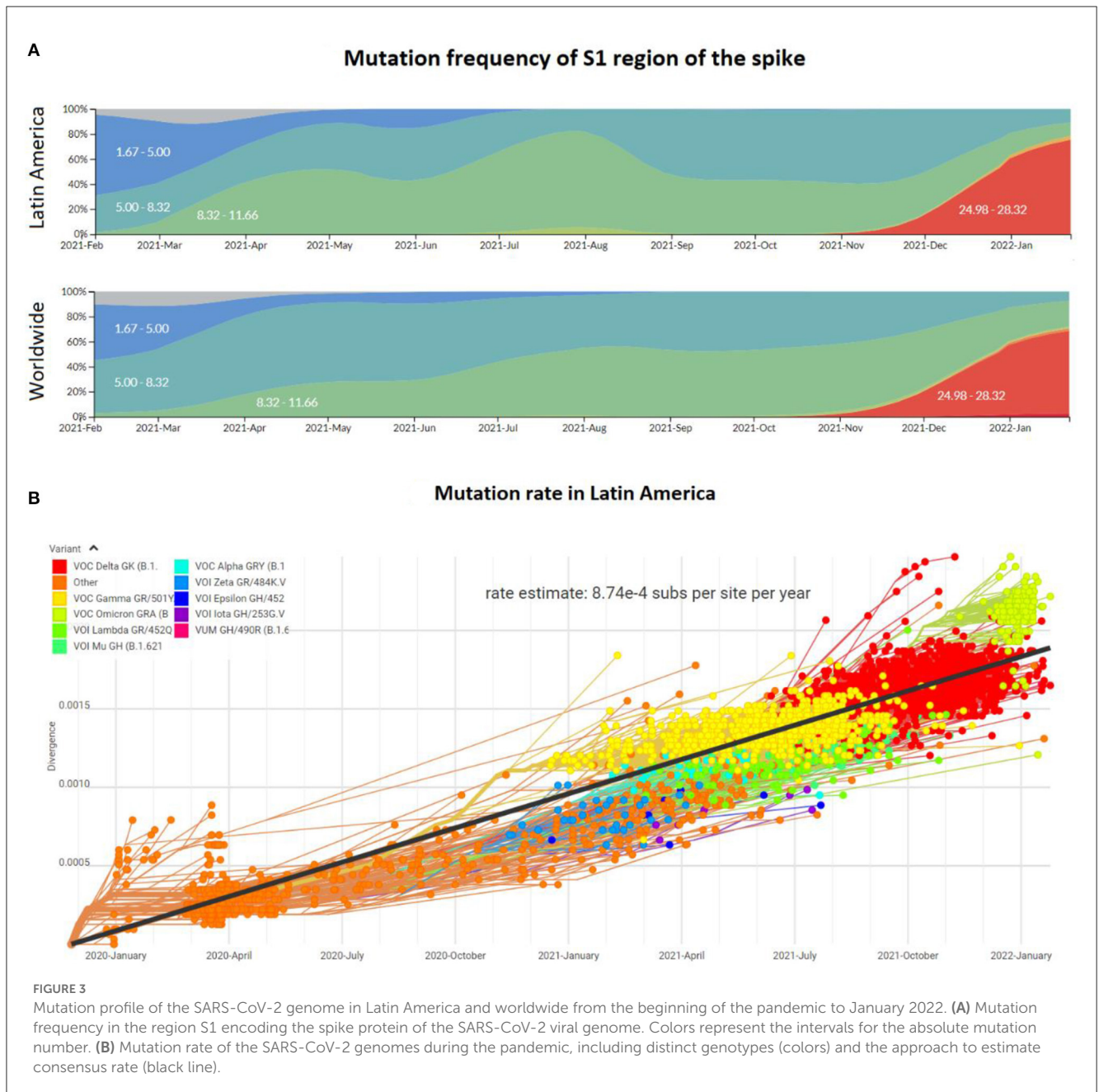
Finally, since most countries in this region are considered low- and middle-income countries, the impact of the COVID-19 pandemic on society has been devastating socially and economically (46). Genomic surveillance is pivotal as a powerful tool for decision-makers regarding the management of the pandemic in the Latin American context concerning social and economic measures, as well as practical decisions in terms of the diagnostic tools, treatments, and vaccines (4). On the other hand,

local and prompt reports of emerging genotypes demonstrated the laboratory and bioinformatic capabilities in Latin American countries. These capabilities were developed locally in the last years for the surveillance of pathogens and other applications. Jointly, the local and international efforts to achieve the genomic surveillance of SARS-CoV-2 have contributed to the understanding of the dynamics of the pandemic in Latin America, which is an ongoing process.

In addition, the infrastructure related to molecular diagnostic techniques experimented a relevant advance due to the pandemic. Before the pandemic outbreak, these techniques were only available in advanced clinical laboratories but now an expanded availability and a cost-effective implementation are found in most clinical laboratories toward-becoming routine tests to study other pathogens and diseases (47).

Regarding limitations, the main drawback of this study is that we assumed that all sequences were comparable, with no segregation by experimental or bioinformatic conditions. The GISAID platform accepts a variety of conditions to upload genome sequences without restriction associated with the sample processing strategy, sequencing technology, genome assembler, variant callers and others, which were not considered here to assess their impact on the genotyping. Although previous reports have found differences in the used pipelines (48), we made the analysis using the whole set of available sequences as performed in other studies (19, 49, 50). Also, as an infectious disease, the clinical outcome of COVID-19 depends on the epidemiological triad: (i) environmental conditions (social behavior, restriction measurements, management of cases, others), (ii) host factors (ethnicity, risk factors, genetic profile of HLA or ACE-II alleles, others), and (iii) the virus (genotype and mutations that impact transmission, immune response, others). Here we have only considered the SARS-CoV-2 genotypes in the period but other





data associated with the epidemiological triad and change of the circulating versions of the virus is relevant to include in further analyses.

## Conclusions

In conclusion, with this study we have contributed to the genomic surveillance of the SARS-CoV-2 in Latin America by providing 1,534 genome sequences from seven countries and the subsequent global analysis of circulating genomes mainly during 2021. For Latin America, patterns regarding several variants associated with multiple re-introductions, a relatively low proportion of sequenced samples, as well as an increase

in the mutation frequency, are in line with worldwide data. Additionally, some genotypes such as Gamma, Mu and Lambda variants and 83 lineages have emerged locally with a subsequent country-specific predominance. Regional efforts demonstrate the laboratory and bioinformatics capabilities for the genomic surveillance of pathogens that have been developed in Latin America, and which is expected to continue during the current COVID-19 pandemic.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories

and accession number(s) can be found in the article/[Supplementary material](#).

## Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

## Author contributions

JM-Mo, PT, GO, JK, MC, RC-S, and AH-E participated in the conception and design of this study. FD-M, CS-G, HB, EC-L, JM-Me, AV, AR, CG, CS, JS, AS-L, LS, MG-S, and GM-R were involved in sample processing. JM-Mo standardized the bioinformatics pipelines and drafted the manuscript. JM-Mo, GN, RO, TN, JR-G, and EC were involved in data analysis. JM-Mo, FD-M, GO, JK, MC, RC-S, and AH-E participated in the interpretation of results. All authors reviewed and approved the final manuscript.

## Funding

This work was supported by the CABANA project, which was funded by the BBSRC under the Global Challenges Research Fund (GCRF) Growing Research Capability call, Contract Number BB/P027849/1 (<http://www.ukri.org/research/global-challenges-research-fund/funded-projects/>), through the CABANA Innovation Fund to AH-E.

## References

- Ashour HM, Elkhatib WF, Rahman MM, Elshabrawy HA. Insights into the recent 2019. Novel coronavirus (SARS-CoV-2) in light of past human coronavirus outbreaks. *Pathog.* (2020) 9:186. doi: 10.3390/pathogens9030186
- World Health Organization. *WHO Director-General's Opening Remarks at the Media Briefing on COVID-19 - 11 March 2020*. (2020). Available online at: <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19--11-march-2020> (accessed February 2, 2022).
- Young BE, Fong SW, Chan YH, Mak TM, Ang LW, Anderson DE, et al. Effects of a major deletion in the SARS-CoV-2 genome on the severity of infection and the inflammatory response: an observational cohort study. *Lancet.* (2020) 396:603–11. doi: 10.1016/S0140-6736(20)31757-8
- Gouvêa dos Santos W. Co-infection, re-infection and genetic evolution of SARS-CoV-2: Implications for the COVID-19 pandemic control. *Comment dos Santos.* (2021) 2:56–61. doi: 10.46439/cancerbiology.2.025
- Zhou H-Y, Cheng Y-X, Xu L, Li J-Y, Tao C-Y, Ji C-Y, et al. Genomic evidence for divergent co-infections of SARS-CoV-2 lineages. *bioRxiv [Preprints]*. (2021). doi: 10.1101/2021.09.03.458951
- World Health Organization. *Tracking SARS-CoV-2 Variants*. (2022). Available online at: <https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/> (accessed January 20, 2022).
- Pan American Health Organization. *Epidemiological Update: SARS-CoV-2 Variants in the Region of the Americas - 1 December 2021*. (2021). Available online at: <https://www.paho.org/en/documents/epidemiological-update-sars-cov-2-variants-region-americas-1-december-2021> (accessed January 20, 2022).
- Molina-Mora JA. Insights into the mutation T1117I in the spike and the lineage B.1.1.389 of SARS-CoV-2 circulating in Costa Rica. *Gene Reports.* (2022) 27:1–9. doi: 10.1016/j.genrep.2022.101554
- Molina-Mora JA, Cordero-Laurent E, Godínez A, Calderón-Osorno M, Brenes H, Soto-Garita C, et al. SARS-CoV-2 genomic surveillance in Costa Rica: Evidence of a divergent population and an increased detection of a spike T1117I mutation. *Infect Genet Evol.* (2021) 92:104872. doi: 10.1016/j.meegid.2021.104872
- The Lancet. COVID-19 in Latin America: a humanitarian crisis. *Lancet.* (2020) 396:1463. doi: 10.1016/S0140-6736(20)32328-X
- Callejas D, Echevarría JM, Carrero Y, Rodríguez-Morales AJ, Moreira R. The SARS-CoV-2 pandemic in Latin America: the need for multidisciplinary approaches. *Curr Trop Med reports.* (2020) 7:120–5. doi: 10.1007/s40475-020-00219-w
- Oliveira RRM, Negri TC, Nunes G, Medeiros I, Araújo G, de Oliveira Silva F, et al. PipeCoV: a pipeline for SARS-CoV-2 genome assembly, annotation and variant identification. *PeerJ.* (2022) 10:13300. doi: 10.7717/peerj.13300
- Resende PC, Motta FC, Roy S, Appolinario LR, Fabri A, Xavier J, et al. SARS-CoV-2 genomes recovered by long amplicon tiling multiplex approach using nanopore sequencing and applicable to other sequencing platforms. *bioRxiv [Preprints]*. (2020). doi: 10.1101/2020.04.30.069039
- Pan American Health Organization. *PAHO Daily COVID-19 Update- 20 January 2022*. PAHO/WHO | Pan American Health Organization (2022). Available online at: <https://www.paho.org/en/documents/paho-daily-covid-19-update-20-january-2022> (accessed January 21, 2022).

## Acknowledgments

We thank public and private clinical laboratories for the samples of confirmed cases of COVID-19 from the participant countries.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpubh.2023.1095202/full#supplementary-material>

15. Rambaut A, Holmes EC, O'Toole Á, Hill V, McCrone JT, Ruis C, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol.* (2020) 5:1403–7. doi: 10.1038/s41564-020-0770-5
16. O'Toole Á, Scher E, Underwood A, Jackson B, Hill V, McCrone JT, et al. Assignment of epidemiological lineages in an emerging pandemic using the pangolin tool. *Virus Evol.* (2021) 7:64. doi: 10.1093/ve/veab064
17. Pearson AA, Prado AM, Colburn FD. Nicaragua's surprising response to COVID-19. *J Global Health.* (2020) 10:371. doi: 10.7189/jogh.10.010371
18. Salazar Mather TP, Gallo Marin B, Medina Perez G, Christophers B, Paiva ML, Oliva R, et al. Love in the time of COVID-19: negligence in the Nicaraguan response. *Lancet Glob Heal.* (2020) 8:e773. doi: 10.1016/S2214-109X(20)30131-5
19. Brito AF, Semenova E, Dudas G, Hassler GW, Kalinich CC, Kraemer MUG, et al. Global disparities in SARS-CoV-2 genomic surveillance. *medRxiv Prepr Serv Heal Sci [Preprints].* (2022). doi: 10.1101/2021.08.21.21262393
20. Vavrek D, Speroni L, Curnow KJ, Oberholzer M, Moeder V, Febbo PG. Genomic surveillance at scale is required to detect newly emerging strains at an early timepoint. *medRxiv [Preprints].* (2021). doi: 10.1101/2021.01.12.21249613
21. Romero PE, Dávila-Barclay A, Salvatierra G, González L, Cuicapuzo D, Solís L, et al. The Emergence of SARS-CoV-2 Variant Lambda (C.37) in South America. *Microbiol Spectr.* (2021) 9:21. doi: 10.1128/Spectrum.00789-21
22. Vargas-Herrera N, Araujo-Castillo RV, Mestanza O, Galarza M, Rojas-Serrano N, Solari-Zerpa L. SARS-CoV-2 Lambda and Gamma variants competition in Peru, a country with high seroprevalence. *Lancet Reg Heal Am.* (2022) 6:100112. doi: 10.1016/j.lana.2021.100112
23. Laiton-Donato K, Franco-Muñoz C, Álvarez-Díaz DA, Ruiz-Moreno HA, Usme-Ciro JA, Prada DA, et al. Characterization of the emerging B.1.621 variant of interest of SARS-CoV-2. *Infect Genet Evol.* (2021) 95:105038. doi: 10.1016/j.meegid.2021.105038
24. Wang R, Chen J, Gao K, Hozumi Y, Yin C, Wei GW. Analysis of SARS-CoV-2 mutations in the United States suggests presence of four substrains and novel variants. *Commun Biol.* (2021) 4:1–14. doi: 10.1038/s42003-021-01754-6
25. Alizon S, Haim-Boukobza S, Foulongne V, Verdurme L, Trombert-Paolantoni S, Lecorche E, et al. Rapid spread of the SARS-CoV-2 Delta variant in some French regions, June 2021. *Euro Surveill.* (2021) 26:2100573. doi: 10.2807/1560-7917.ES.2021.26.28.2100573
26. Ramírez JD, Florez C, Muñoz M, Hernández C, Castillo A, Gomez S, et al. The arrival and spread of SARS-CoV-2 in Colombia. *J Med Virol.* (2021) 93:1158–63. doi: 10.1002/jmv.26393
27. da Silva Filipe A, Shepherd JG, Williams T, Hughes J, Aranday-Cortes E, Asamaphan P, et al. Genomic epidemiology reveals multiple introductions of SARS-CoV-2 from mainland Europe into Scotland. *Nat Microbiol.* (2020) 6:112–22. doi: 10.1038/s41564-020-00838-z
28. Spanakis N, Kassela K, Dovrolis N, Bampali M, Gatzidou E, Kafasi A, et al. A main event and multiple introductions of SARS-CoV-2 initiated the COVID-19 epidemic in Greece. *J Med Virol.* (2021) 93:2899–907. doi: 10.1002/jmv.26778
29. GISAID. *GISAID - Clade and Lineage Nomenclature Aids in Genomic Epidemiology of Active hCoV-19 Viruses.* GISAID (2021). Available online at: <https://www.gisaid.org/references/statements-clarifications/clade-and-lineage-nomenclature-aids-in-genomic-epidemiology-of-active-hcov-19-viruses/> (accessed November 18, 2020).
30. Paniz-Mondolfi A, Muñoz M, Florez C, Gomez S, Rico A, Pardo L, et al. SARS-CoV-2 spread across the Colombian-Venezuelan border. *Infect Genet Evol.* (2020) 86:616. doi: 10.1016/j.meegid.2020.104616
31. Mir D, Rego N, Resende PC, Tort F, Fernández-Calero T, Noya V, et al. Recurrent dissemination of SARS-CoV-2 through the Uruguayan–Brazilian border. *Front Microbiol.* (2021) 12:1018. doi: 10.3389/fmicb.2021.653986
32. Williams GH, Llewelyn A, Brandao R, Chowdhary K, Hardisty KM, Loddo M. SARS-CoV-2 testing and sequencing for international arrivals reveals significant cross border transmission of high risk variants into the United Kingdom. *eClinicalMedicine.* (2021) 38:101021. doi: 10.1016/j.eclinm.2021.101021
33. Shieh-zadegan S, Alaghemand N, Fox M, Venketaraman V. Analysis of the delta variant B16172 COVID-19. *Clin Pract.* (2021) 11:778–84. doi: 10.3390/clinpract11040093
34. Gualarte JS, da Silva MS, Mosena ACS, Demoliner M, Hansen AW, Filippi M, et al. Early introduction, dispersal and evolution of Delta SARS-CoV-2 in Southern Brazil, late predominance of AY.99.2 and AY.101 related lineages. *Virus Res.* (2022) 311:198702. doi: 10.1016/j.virusres.2022.198702
35. de Souza UJB, Dos Santos RN, de Melo FL, Belmok A, Galvão JD, de Rezende TCV, et al. Genomic epidemiology of SARS-CoV-2 in Tocantins state and the diffusion of P17 and AY992 Lineages in Brazil. *Viruses.* (2022) 14:659. doi: 10.3390/v14040659
36. Zárate S, Taboada B, Muñoz-Medina JE, Iša P, Sanchez-Flores A, Boukaidia C, et al. The alpha variant (B.1.1.7) of SARS-CoV-2 failed to become dominant in Mexico. *Microbiol Spectr.* (2022) 10:21. doi: 10.1128/spectrum.02240-21
37. Molina-Mora JA, González A, Jiménez-Morgan S, Cordero-Laurent E, Brenes H, Soto-Garita C, et al. Clinical profiles at the time of diagnosis of SARS-CoV-2 infection in Costa Rica during the pre-vaccination period using a machine learning approach. *Phenomics.* (2022) 1:1–11. doi: 10.1007/s43657-022-00058-x
38. Acevedo ML, Gaete-Argel A, Alonso-Palomares L, de Oca MM, Bustamante A, Gaggero A, et al. Differential neutralizing antibody responses elicited by CoronaVac and BNT162b2 against SARS-CoV-2 Lambda in Chile. *Nat Microbiol.* (2022) 7:524–9. doi: 10.1038/s41564-022-01092-1
39. Reales Gonzalez J, Prada Cardozo D, Corchuelo S, Zabaleta G, Alarcón Z, Sepulveda MTH, et al. Prolonged SARS-CoV-2 nucleic acid conversion time in military personnel outbreaks with presence of specific IgG antibodies. *J Med Microbiol.* (2022) 71:1–8. doi: 10.1099/jmm.0.001498
40. Molina-Mora JA, Cordero-Laurent E, Calderón-Osorno M, Chacón-Ramírez E, Duarte-Martínez F. Metagenomic pipeline for identifying co-infections among distinct SARS-CoV-2 variants of concern: study cases from Alpha to Omicron. *Sci Reports.* (2022) 12:1–10. doi: 10.1038/s41598-022-13113-4
41. Corey L, Beyrer C, Cohen MS, Michael NL, Bedford T, Rolland M. SARS-CoV-2 Variants in Patients with Immunosuppression. *N Engl J Med.* (2021) 385:562–6. doi: 10.1056/NEJMs2104756
42. Desai S, Rashmi S, Rane A, Dharavath B, Sawant A, Dutt A. An integrated approach to determine the abundance, mutation rate and phylogeny of the SARS-CoV-2 genome. *Brief Bioinform.* (2021) 22:1065–75. doi: 10.1093/bib/bbaa437
43. Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, et al. BEAST 25: an advanced software platform for Bayesian evolutionary analysis. *PLoS Comput Biol.* (2019) 15:e1006650. doi: 10.1371/journal.pcbi.1006650
44. Banerjee A, Mossman K, Grandvaux N. Molecular determinants of SARS-CoV-2 variants. *Trends Microbiol.* (2021) 29:871–3. doi: 10.1016/j.tim.2021.07.002
45. Liu R, Wu P, Ogrodzki P, Mahmoud S, Liang K, Liu P, et al. Genomic epidemiology of SARS-CoV-2 in the UAE reveals novel virus mutation, patterns of co-infection and tissue specific host immune response. *Sci Reports.* (2021) 11:1–14. doi: 10.1038/s41598-021-92851-3
46. Bontan N, Hoffmann B, Vera-Cossio D. The unequal impact of the coronavirus pandemic: evidence from seventeen developing countries. *PLoS ONE.* (2020) 15:e0239797. doi: 10.1371/journal.pone.0239797
47. Alamri AM, Alkhilawi FA, Ullah Khan N. Era of molecular diagnostics techniques before and after the COVID-19 pandemic. *Curr Issues Mol Biol.* (2022) 44:4769–89. doi: 10.3390/cimb44100325
48. Liu T, Chen Z, Chen W, Chen X, Hosseini M, Yang Z, et al. A benchmarking study of SARS-CoV-2 whole-genome sequencing protocols using COVID-19 patient samples. *iScience.* (2021) 24:102892. doi: 10.1016/j.isci.2021.102892
49. Ohlsen EC, Hawksworth AW, Wong K, Guagliardo SAJ, Fuller JA, Sloan ML, et al. Determining gaps in publicly shared SARS-CoV-2 genomic surveillance data by analysis of global submissions. *Emerg Infect Dis.* (2022) 28:85–92. doi: 10.3201/eid2813.220780
50. Bhattacharya M, Sharma AR, Dhama K, Agoramoorthy G, Chakraborty C. Omicron variant (B11529) of SARS-CoV-2: understanding mutations in the genome, S-glycoprotein, and antibody-binding regions. *GeroScience.* (2022) 44:619–37. doi: 10.1007/s11357-022-00532-4