Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Emergence and spreading of the largest SARS-CoV-2 deletion in the Delta AY.20 lineage from Uruguay

Yanina Panzera, Maria Noel Cortinas, Ana Marandino, Lucía Calleros, Victoria Bormida, Natalia Goñi, Claudia Techera, Sofía Grecco, Joaquín Williman, Viviana Ramas, Leticia Coppola, Cristina Mogdasy, Héctor Chiparelli, Ruben Pérez

PII: S2452-0144(22)00211-4

DOI: https://doi.org/10.1016/j.genrep.2022.101703

Reference: GENREP 101703

To appear in: *Gene Reports*

Received date: 21 September 2022

Revised date: 24 October 2022

Accepted date: 26 October 2022

Please cite this article as: Y. Panzera, M.N. Cortinas, A. Marandino, et al., Emergence and spreading of the largest SARS-CoV-2 deletion in the Delta AY.20 lineage from Uruguay, *Gene Reports* (2022), https://doi.org/10.1016/j.genrep.2022.101703

Emergence and spreading of the largest SARS-CoV-2 deletion in the Delta AY.20 lineage from Uruguay

Yanina Panzera [a], Maria Noel Cortinas [b], Ana Marandino [a], Lucía Calleros [a], Victoria Bormida [b], Natalia Goñi [c], Claudia Techera [a], Sofía Grecco [a], Joaquín Williman [a], Viviana Ramas [c], Leticia Coppola [c], Cristina Mogdasy [c], Héctor Chiparelli [c], Ruben Pérez [a*]

*Corresponding author

E-mails: rperez@fcien.edu.uy (RP)

a) Sección Genética Evolutiva, Departamento de Biología Animal, Instituto de Biología, Facultad de Ciencias, Universidad de la República, Iguá 4225, 11400, Montevideo, Uruguay.

b) Genómica. Departamento de Laboratorios de Salud Pública, Ministerio de Salud Pública, Alfredo Navarro 3051 (entrada N), 11600, Montevideo, Uruguay.

c) Centro Nacional de Referencia de Influenza y otros Virus Respiratorios, Departamento de Laboratorios de Salud Pública, Ministerio de Salud Pública, Alfredo Navarro 3051 (entrada N), 11600, Montevideo, Uruguay.

Keywords: SARS-CoV-2, largest deletion, accessory ORFs, recombination hotspot.

Abstract

The genetic variability of SARS-CoV-2 (genus *Betacoronavirus*, family *Coronaviridae*) has been scrutinized since its first detection in December 2019. Although the role of structural variants, particularly deletions, in virus evolution is little explored, these genome changes are extremely frequent. They are associated with relevant processes, including immune escape and attenuation. Deletions commonly occur in accessory ORFs and might even lead to the complete loss of one or more ORFs. This scenario poses an interesting question about the origin and spreading of extreme structural rearrangements that persist without compromising virus viability. Here, we analyze the genome of SARS-CoV-2 in late 2021 in Uruguay and identify a Delta lineage (AY.20) that experienced a large deletion (872 nucleotides according to the reference Wuhan strain) that removes the 7a, 7b, and 8 ORFs. Deleted viruses coexist with wild-type (without deletion) AY.20 and AY.43 strains. The Uruguayan deletion is like those identified in Delta strains from Poland and Japan but occurs in a different Delta clade. Besides providing proof of the circulation of this large deletion in America, we infer that the 872-deletion arises by the consecutive occurrence of a 6-nucleotide deletion, characteristic of delta strains, and an 866-nucleotide deletion that arose independently in the AY.20 Uruguayan lineage. The largest deletion occurs adjacent to transcription regulatory sequences needed to synthesize the nested set of subgenomic mRNAs that serve as templates for transcription. Our findings support the role of transcription sequences as a hotspot for copy-choice recombination and highlight the remarkable dynamic of SARS-CoV-2 genomes.

1. Introduction

The pandemic coronavirus disease 2019 (COVID-19) is one of the most concerning and challenging global health threats. Massive research has focused on its causative agent, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Gorbalenya et al., 2020). Like other coronaviruses, SARS-CoV-2 is an enveloped virus with a positive polarity single-stranded RNA genome of ~30,000 nucleotides with high mutation and recombination rates (Duchene et al., 2020).

The multicistronic SARS-CoV-2 genome is organized in open reading frames (ORFs) that code for structural and non-structural proteins. The largest 1a/1ab ORF constitutes the 5'-proximal two-thirds of the genome and gives rise to the non-structural proteins associated with the replication process. In addition, some ORFs encoding the structural proteins spike, envelope, membrane, and nucleocapsid are found downstream, interspaced with ORFs encoding accessory proteins 3a, 6, 7a, 7b, and 8 (Kim et al., 2020; Pancer et al., 2020). The large 1a/1ab ORF is directly expressed by the replication–transcription complex (RTC) from genomic RNA. Still, other viral proteins' expression is achieved using subgenomic mRNAs produced during replication (Lai and Cavanagh, 1997). Subgenomic mRNAs arise by discontinuous transcription produced by the RTC template switch from the transcription regulatory sequences (TRS), located upstream of most ORFs, to the leader TRS located about 70 nt from the 5´end of the genome.

Millions of reported SARS-CoV-2 sequences allow us to analyze the viral microevolution on a scale not previously attempted for any pathogen. Numerous changes within SARS-CoV-2 genomes are continuously reported, from minor single-nucleotide polymorphisms (SNPs) to insertions and deletions (indels). Regardless of its

size, all variations may impair the protein structure and function and cause changes in the virus phenotype, infectivity, virulence, and host immune response.

Most studies initially focused on SNPs that produce non-synonymous substitutions in relevant genome regions. The spike (S) protein has been the most studied because of its essential role in virus cell entry and infection (Moore et al., 2020). However, genome structural variations are becoming more relevant as the virus expands in the human population. Indels in coronaviruses are enhanced by the discontinuous RNA synthesis of the polymerase machinery and remain uncorrected by the proofreading activity of nsp14-exoribonuclease (Chen et al., 2020). Although most indels likely negatively affect viral fitness (Grubaugh et al., 2020), a small number emerge and spread in viral populations, suggesting a positive effect on fitness and adaptive evolution (Foster and Rawlinson, 2021; Kemp et al., 2021; Lau et al., 2020; McCarthy et al., 2021; Panzera et al., 2021b, 2021a; Su et al., 2020; Young et al., 2020). Thus, analyzing these indels may reveal evolutionary trends and provide new insight into the surprising variability and rapidly spreading capability that SARS-CoV-2 has shown since its recent emergence. Deletions are particularly interesting because they can help us understand the selective pressures on different genome regions and how the virus loses genetic material to evolve in the host population. Deletions could also be related to the virus's attenuation in the human population (Lau et al., 2020; Young et al., 2020). The modification and inactivation of some genes could indicate that the virus adapts to rapid spread in the population but produces a less harmful effect.

One usual target of deletions is the five accessory ORFs that occupy almost 7% of the genome and have been implicated in modulating cellular innate immune responses

(Narayanan et al., 2008; Pancer et al., 2020; Stadler et al., 2003). Accessory ORFs may be prone to deletion because they seem dispensable for replication *in vitro* and *vivo*.

Some of these deletions in accessory genes relate to coronavirus adaptation to the host cell environment. For example, in SARS-CoV-1, a 45-nucleotide deletion in the ORF7b emerged following passage in Vero E6 cells, suggesting attenuation (Thiel et al., 2003). The same deletion occurs in other cultured strains (Yan et al., 2004; Yeh et al., 2004). In SARS-CoV-2, three different deletions in ORF7a truncate the C-terminal half of the protein; two of them result in growth defects and failures to suppress the immune response, providing potential advantages for the virus adaptation in humans (Joonlasak et al., 2021; Nemudryi et al., 2021; Pyke et al., 2021; Zinzula, 2020).

Deletions can even remove complete ORFs and are expected to alter the structure of the whole genome. For example, two large and phylogenetically unrelated deletions (392 and 227 nucleotides long) fuse the ORF7a with downstream ORFs (Addetia et al., 2020). The 392-nucleotide deletion lacks the ORF7b and creates a new ORF from the N-terminus of ORF7a and the ORF8. The 227-nucleotide deletion results in a new ORF by combining the N-terminus of ORF7a with the ORF7b. These deletions have become extinct or appear as sporadic or unique variants (Addetia et al., 2020; Tse et al., 2021b). On the other hand, a 382-nucleotide deletion that removes most of the ORF8 was a circulating form hypothesized to lead to an attenuated phenotype of SARS-CoV-2 (Su et al., 2020; Young et al., 2020).

Remarkably, certain deletions could be so extensive that they eliminate contiguous accessory ORFs. For example, in Southern Poland, an 872-nucleotide deletion (Δ872) in the AY.4 lineage of the Delta (also known as B.1.617.2 or AY) variants of concern (VOC) caused the complete loss of three consecutive accessory ORFs (7a, 7b, and 8)

(Mazur-Panasiuk et al., 2021). In addition, a deletion in the B.1.36.27 lineage from Hong Kong also lacks these three ORFs and has the last 12 nucleotides of the ORF6 replaced by ~60 nucleotides from the 5´- UTR (Tse et al., 2021a). These deletions are the largest genomic rearrangements that occurred naturally in SARS-CoV-2 and pose interesting questions about the origin and spreading of extreme structural rearrangements that persist without apparently affecting virus infectivity.

Here, we present genomic evidence supporting recent increases in the Delta VOC number and proportion in Uruguay and identify a similar Polish deletion (Δ872) in the AY.20 lineage. Besides providing proof of the circulation of this large deletion outside Europe, we reveal that this deletion arose independently in Uruguay and might have originated by RTC intramolecular translocation at transcription-regulating sequences (TRS).

2. Materials and Methods

2.1. Samples and SARS-CoV-2 diagnosis

Combined nasopharyngeal and oropharyngeal swab samples from Uruguayan patients were collected from September to December 2021 by the Reference Center for Influenza and other Respiratory Viruses, National Institute of Health Laboratories (DLSP-MSP). The collection and analysis of samples were performed according to the Declaration of Helsinki; no specific authorization was required because the activities were conducted as part of routine virological surveillance (anonymously, without identification of patients) by the Uruguayan official Institution (DLSP-MSP). The SARS-CoV-2 diagnosis was performed by RNA extraction with the Qiamp Viral RNA Minikit (Qiagen USA) followed by real-time reverse transcription-polymerase chain reaction (RT-qPCR) using the protocol recommended by the Panamerican Health Organization (PAHO-WHO) (Corman et al., 2020).

2.2. Full-length genome amplification using COVIDSeq assay

The SARS-CoV-2 genome of 37 samples was processed with the COVIDSeq assay from Illumina (USA), which uses a modified version of the ARTIC multiplex PCR (Bhoyar et al., 2021). The reverse transcription was performed from 8.5 µL of RNA, and the library preparation employed dual indexing (IDT for Illumina-PCR Indexes Set 3). Whole-genome sequencing was performed on an Illumina MiniSeq platform using MiniSeq^{TM} Mid Output Reagent Cartridge (300-cycles, paired-end reads). The 37 libraries were sequenced in three runs. Samples Adapter/quality trimming and filtering of raw data were performed with BBDuk, and clean reads were mapped to the consensus genome using Minimap2 in Geneious Prime 2020.1.2

(https://www.geneious.com). The consensus sequences were also obtained using the DRAGEN genome pipeline from Illumina BaseSpace (https://basespace.illumina.com).

2.3. PCR, Sanger sequencing, and capillary electrophoresis

An aliquot of RNA (12µL) was used for cDNA synthesis with Superscript II ® reverse transcriptase (Thermo Fisher, USA) and random primers. cDNA was amplified by PCR with a single pair of primers surrounding the Δ872 (90_Left and 93_Right) as previously described (Mazur-Panasiuk et al., 2021). The deleted amplicon has 450 bp, and the wild type 1323 bp in the reference sequence NC_045512, or 1316 bp in Delta sequences with two characteristic deletions of 6 and 1 nucleotides concerning NC_045512. Amplicons were analyzed by capillary electrophoresis on a Fragment Analyzer 5200 system using the High Sensitivity NGS Analysis Kit (Agilent Technologies, USA). Amplicons were also subjected to Sanger sequencing in Macrogen (Korea).

2.4. Deletion characterization and comparison

The CoV-GLUE and outbreak.info (Lineage|Mutation) trackers were used to identify markers (SNPs and indels) in SARS-CoV-2 lineages (Hughes et al., 2022; Singer et al., 2020).

Sequence comparison and identification of strains with similar deletions were performed using GISAID Audacity Instant and BLAST searches.

2.5. Secondary structure in the transcription-regulating sequence (TRS)

We used RNAfold prediction software with the default parameters (Lorenz et al., 2011)to estimate the secondary structure of the reference SARS-CoV-2 and the Delta strain RNA genome; the Geneious RNA fold utility was used for visualization.

2.6. SARS-CoV-2 lineage assignment and phylogenetic analysis

The lineage of the strains was assigned according to the nomenclature system proposed by Rambaut et al. (Rambaut et al., 2020) using the Pangolin tool (O'Toole et al., 2021).

The phylogenetic analysis includes the newly obtained Uruguayan sequences and closely related South American and global sequences. Sequences and associated metadata were retrieved from the GISAD EpiCoV database using the search and blast tools (Shu and McCauley, 2017). The dataset was filtered and reduced to generate a comprehensive genome dataset.

DNA alignments were performed with MAFFT (Katoh and Standley, 2013) and manually edited to exclude varied-length 5′- and 3′-untranslated regions (UTRs) based on the annotation of NC_045512. With 1000-replicates bootstrap to support internal nodes, maximum-likelihood trees were inferred in Geneious using FastTree (Price et al., 2009) and visualized with the ggTree package in R.

3. Results

3.1. Whole-genome sequencing

We sequenced 37 SARS-CoV-2 genomes from Uruguay in September-November 2021. Genomes were obtained with an average coverage of 1700× (range: 300× to 4200×) and submitted to GenBank (Table 1). All strains were classified as belonging to the Delta lineage (Table 1).

All strains from Uruguay have characteristic Delta (B.1.617.2) changes in the spike (19R, Δ156-157, 158G, 452R, 478K, 614G, 681R, and 950N). Other common B.1.617.2 residues occur in the ORF1a (1306S, 2046L, 2287S, 2930L, 3255I, 3646A, 4715L), ORF3a (26L), N (63G, 203M, 215C, 377Y), M (82T), ORF 7a (82A and 120I) and ORF 7b (40I). Delta strains also have a deletion of 6 bases in the ORF8 that removes two residues (119-120) concerning Wuhan reference strains. In the intergenic region between ORF8 and N, there is an additional 1-nt deletion. These two small deletions surround the TRS needed to generate the subgenomic mRNA coding for the N protein. The secondary structure in the TRS-containing region differs between the reference and Delta strains. The reference sequence TRS folds in a stem with an internal loop; the Delta strain TRS folds in a hairpin base and a bulge loop (Figure 1).

The Delta strains were further classified in the AY.20 (n=6) and AY.43 (n=31) lineages (Table 1). AY.20 and AY.43 lineages show residue changes in ORF1ab (V599A and L5230I), S (Y69H), and N (Q9L)

3.2. Deletion in the Delta AY.20 lineage

The consensus obtained on the Illumina DRAGEN application from five of the six AY.20 sequences had a long stretch of Ns between nucleotide positions 27,385 and 28,256, corresponding to the region of the ORF7a, ORF7b, and ORF8. Mapping with the reference SARS-CoV-2 genome (GenBank NC_045512) revealed an absence of NGS reads in this part of the genome and single reads that map on both sides (ORF6 and N) of the putative deletion, suggesting a large genomic loss. This deletion does not occur in one of the early collected AY.20 sample p2262 (Table 1).

The deletion was confirmed by amplifying the surrounding region using flanking primers. Capillary electrophoresis showed that the amplification products were 450 bp instead of the expected size of 1,315 bp in delta strains (Figure 2). The amplicons were subjected to Sanger sequencing to confirm the deletion within the AY.20 lineage. The deletion spans 866-nt (Δ866) concerning the wild-type AY.20 strain. Still, as Delta strains have a previous 6-nt deletion in the same region (Figure 1), the deletion has 872-nt (Δ872) compared with the reference Wuhan strain (Figures 1 and 2). The Δ872 produces the loss of the entire ORF7a, ORF7b, and ORF8 without affecting the upstream ORF6 and the downstream N gene (Figure 2). The deletion breakpoint was contiguous to the downstream TRS (the upstream TRS located at the very end of the deletion) (Figure 2).

3.3. Phylogenetic analysis

The newly sequenced genomes from the AY.20 and AY.43 lineages were added to related genomes retrieved from the GISAID. This dataset included all the good-quality sequences from America (Argentina, Aruba, Brazil, Canada, Chile, Colombia, Ecuador, Guyana, Paraguay, Peru, Mexico, and the USA), Europe, and Asia. Sequences of both lineages appeared separated in the phylogenetic tree, as expected by their corresponding pangolin classification.

The AY.20 lineage samples from Uruguay fell in a clade with undeleted Mexican, Chilean, USA, and Canadian AY.20 sequences (Figure 3).

The Δ872 variant from Poland (AY.4) fell in a different position in the phylogenetic tree associated with Poland strains that lack the deletion. Some variants from Japan with

the same deletion belonged to a third Delta lineage (AY.90); a closely related Japanese lineage (AY.99) lacks Δ872 but has a smaller deletion in the same position.

The AY.43 lineage samples appeared distributed in two sister clades with well-supported subclades. One comprises Uruguayan sequences, and the other clade clustered variants from Uruguay and South America (Figure 3).

4. Discussion

SARS-CoV-2 genetics studies have been fundamental to understanding the emergence and spreading of VOCs, including B.1.1.7 (Alpha), B.1.351 (Beta), P.1 (Gamma), and B.1.617.2 (Delta) that were first reported in England (Volz et al., 2021), South Africa (Tegally et al., 2021), Brazil (Faria et al., 2021) and India (Singh et al., 2021), respectively. These VOCs have been associated with extensive transmission following emergence, leading to substantial infection and mortality rates even in populations with high seroprevalence (Faria et al., 2021; Sabino et al., 2021). VOCs are more transmissible than ancestral SARS-CoV-2 lineages and carry mutations that contribute to partial immune escape (E484K and T478K).

We observed a dynamic shift in the composition of SARS-CoV-2 lineages driving transmission across Uruguay in late 2021. The predominant Gamma (P1) VOC was replaced by the Delta VOC during the last month of 2021. Similar behavior occurred in other countries in early 2021, and the Delta VOC has become dominant worldwide since its first detection in India in late 2020 (Mazur-Panasiuk et al., 2021; Mishra et al., 2021). This variant raises concern due to its high transmissibility, immune escape capability, and risk of reinfection (Alkhatib et al., 2021).

Delta strains from Uruguay belong mostly to the AY.43 and a lesser extent, to AY.20 lineages. However, both lineages are quite divergent and appear separated in the phylogenetic tree (Figure 3).

The AY.20 lineage is polymorphic for an 866-nt deletion (Δ866) that emerged in a Delta genome background with a 6-nt deletion (Δ6) in the same region (Figures 1 and 2). Accordingly, the Δ872 arouses by two consecutive deletions, the former Δ6 and the latter Δ866 (Δ872 = Δ866 + Δ6). The existence of an AY.20 wild type (without Δ866) in Uruguay might indicate that the Δ866 emerged locally. Moreover, the deleted variants were collected from September to October, indicating that it is a circulating form that undergoes local transmission (Table 1). Uruguayan AY.20 is similar to North American and South American (Chile) strains, suggesting a common origin (Figure 3).

The Uruguayan Δ872 occurs in the same genomic position and has the same length that Poland and Japan variants. Remarkably, all AY.20 strains with Δ872 fell into different phylogenetic tree branches according to their classification. They are more related to wild-type strains of the same origin. The genomes from the three countries differed in several nucleotide positions and had different phylogenetic relationships. Therefore, detecting the same Δ872 in phylogenetically unrelated samples supports that Δ872 arose by the occurrence of Δ866 in different Delta clades with the ancestral Δ6 variant.

Deletions are associated with RTC's homology-assisted mechanism to switch template strands during subgenomic and full-genome synthesis. RTC may disassociate from the template strand during the nascent strand synthesis and reassociate to the same template strand at different loci, producing indels (Chrisman et al., 2021; Lai, 1992; Simon-Loriere and Holmes, 2011). The Δ872 occurs adjacent to TRS needed to synthesize the nested set of minus-strand subgenomic-RNAs that serve as templates for transcription

(Sawicki and Sawicki, 1995). These short TRS sequences precede the ORFs that encode structural and accessory proteins. SARS-CoV-2 TRS comprises a conserved core sequence (5′-ACGAAC-3′) flanked by motives of variable length that may also contribute to the base-pairing interaction with the leader TRS (Finkel et al., 2021; Kim et al., 2020; Su et al., 2021) (Figure 1). If the nascent RNA "jumps" and reassociates to an upstream TRS, a complete ORF deletion may emerge in that position. For this reason, TRS are considered hot spots for copy-choice recombination in the genome (Chrisman et al., 2021). The large deletion observed in Hong Kong supports this hypothesis. Adjacent to this deletion, there is a ~60-nucleotide translocated sequence identical to the upstream sequence of the leader TRS located at the 5'-end of the plus-strand full-length genome, suggesting that it originated from a "jump" or translocation of the RTC between the TRS preceding the ORF N and the leader TRS (Tse et al., 2021a).

Consecutive deletions in Delta lineages could also be associated with TRS and RTC translocation. The first deletion (Δ6) upstream of the TRS in all Delta strains and the 1-nt downstream deletion might affect the secondary RNA genomic structure and favor the transfer of the nascent RNA to the next TRS, producing complete gene deletion (Figure 1). We observed that the secondary structure in the TRS and surrounding regions are altered in the Delta strain compared with the Wuhan reference strain (Figure 1), supporting a role of RNA fold in RTC translocation.

Massive deletions like those described here are restricted to accessory genes that exhibit extensive plasticity and resilience to undergo structural changes (Zinzula, 2020). Deletions in these accessory ORFs have interesting outcomes and potential effects on virus evolution (Gamage et al., 2020; Michel et al., 2020; Narayanan et al., 2008;

Panzera et al., 2022; Pereira, 2020). The removal of 7a, 7b, and 8 ORFs in the Δ872 variant leaves only two accessory genes (3a and 6) in the SARS-CoV-2 genome and represent almost 3% of the viral genetic material, offering new opportunities to understand the role of accessory genes in host's immune response and viral pathogenicity.

ORF8 has high variability and multiple alterations, including SNPs, short indels causing frameshifts, and partial or complete gene deletions. In recombinant SARS-CoV-1, the truncation of ORF8 led to gradual virus attenuation *in vitro* (Muth et al., 2018). In SARS-CoV-2, the accessory protein encoded by ORF3 has a function related to evasion of the host adaptive immune response via downregulation of the major histocompatibility complex-1 (Zhang et al., 2021). Moreover, ORF8 modulates the host's interferon-mediated antiviral response (Li et al., 2020). Human angiotensin-converting enzyme 2-transgenic mice infected with ΔORF8 recombinants had similar pathological lesions and mortality to normal strain, suggesting that ORF8 does not contribute to virus pathogenicity (Silvas et al., 2021). However, individuals infected with SARS-CoV-2 lacking a functional ORF8 gene have fewer probabilities of developing hypoxia (Young et al., 2020).

The ORF7a and ORF7b are two contiguous genes with a short overlap of four nucleotides, which undergo deletions that could strongly affect protein structure (Addetia et al., 2020; Joonlasak et al., 2021; Tse et al., 2021a). Both ORFs encode transmembrane proteins localized in the endoplasmic reticulum and Golgi network (Liu et al., 2014). ORF7a activates the nuclear factor kappa-light-chain-enhancer of the activated B cell pathway and induces proinflammatory cytokine expression (Su et al., 2021; Zhou et al., 2021). In addition, protein 7a activates the type-I interferon signaling

pathway and promotes the expression of IFN-beta, interleukin-6, and the pro-apoptotic tumor necrosis factor alpha (Yang et al., 2021). Based on studies in animal models, both proteins have a minor impact on pathological lesions and disease outcomes (Silvas et al., 2021).

We did not observe any distinctive signs in the patients with the Δ872 variant (Table 1) that suggest a particular disease course. Patients did not require hospitalization for COVID-19, which is quite common in Uruguay due to the high vaccination rate among the population (more than 80% of the population has two doses) (Table 1). The deleted AY.20 variant was detected in symptomatic and asymptomatic patients with and without vaccines; the non-vaccinated individual had been infected 9 months ago with another SARS-CoV-2 variant (Table 1). The small number of deleted viruses identified avoids hypothesizing about the pathogenic effect of the deleted variant.

Our findings underscore the remarkable variability in the number of accessory ORFs in circulating SARS-CoV-2 strains. Therefore, the re-analysis of raw sequencing data or Sanger sequencing is needed to detect and confirm deletions in SARS-CoV-2 sequences to provide real-time information on the highly dynamic genome. Such information is crucial to understanding SARS-CoV-2 evolutionary trends and the emergence of novel SARS-CoV-2 variants with new biological properties.

Acknowledgments

We thank BIKO S.A. (Uruguay) for Illumina's technical support and supplies.

5. References

Addetia, A., Xie, H., Roychoudhury, P., Shrestha, L., Loprieno, M., Huang, M.-L., Jerome, K.R., Greninger, A.L., 2020. Identification of multiple large deletions in ORF7a resulting in in-frame gene fusions in clinical SARS-CoV-2 isolates. J Clin Virol 129, 104523. https://doi.org/10.1016/j.jcv.2020.104523

Alkhatib, M., Svicher, V., Salpini, R., Ambrosio, F.A., Bellocchi, M.C., Carioti, L., Piermatteo, L., Scutari, R., Costa, G., Artese, A., Alcaro, S., Shafer, R., Ceccherini-Silberstein, F., 2021. SARS-CoV-2 Variants and Their Relevant Mutational Profiles: Update Summer 2021. Microbiol Spectr 9. https://doi.org/10.1128/Spectrum.01096-21

Bhoyar, R.C., Jain, A., Sehgal, P., Divakar, M.K., Sharma, D., Imran, M., Jolly, B., Ranjan, G., Rophina, M., Sharma, S., Siwach, S., Pandhare, K., Sahoo, S., Sahoo, M., Nayak, A., Mohanty, J.N., Das, J., Bhandari, S., Mathur, S.K., Kumar, A., Sahlot, R., Rojarani, P., Lakshmi, J.V., Surekha, A., Sekhar, P.C., Mahajan, S., Masih, S., Singh, P., Kumar, V., Jose, B., Mahajan, V., Gupta, V., Gupta, R., Arumugam, P., Singh, A., Nandy, A., P. V., R., Jha, R.M., Kumari, A., Gandotra, S., Rao, V., Faruq, M., Kumar, S., Reshma C., B., Varma G., N., Roy, S.S., Sengupta, A., Chattopadhyay, S., Singhal, K., Pradhan, S., Jha, D., Naushin, S., Wadhwa, S., Tyagi, N., Poojary, M., Scaria, V., Sivasubbu, S., 2021. High throughput detection and genetic epidemiology of SARS-CoV-2 using COVIDSeq next-generation sequencing. PLoS One 16, e0247115. https://doi.org/10.1371/journal.pone.0247115

Chen, Y., Liu, Q., Guo, D., 2020. Emerging coronaviruses: Genome structure, replication, and pathogenesis. J Med Virol 92, 418–423. https://doi.org/10.1002/jmv.25681

Chrisman, B.S., Paskov, K., Stockham, N., Tabatabaei, K., Jung, J.-Y., Washington, P., Varma, M., Sun, M.W., Maleki, S., Wall, D.P., 2021. Indels in SARS-CoV-2 occur at template-switching hotspots. BioData Min 14, 20. https://doi.org/10.1186/s13040-021-00251-0

Corman, V.M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D.K.W., Bleicker, T., Brünink, S., Schneider, J., Schmidt, M.L., Mulders, D.G.J.C., Haagmans, B.L., van der Veer, B., van den Brink, S., Wijsman, L., Goderski, G., Romette, J.L., Ellis, J., Zambon, M., Peiris, M., Goossens, H., Reusken, C., Koopmans, M.P.G., Drosten, C., 2020. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Eurosurveillance. https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045

Duchene, S., Featherstone, L., Haritopoulou-Sinanidou, M., Rambaut, A., Lemey, P., Baele, G., 2020. Temporal signal and the phylodynamic threshold of SARS-CoV-2. Virus Evol 6, veaa061. https://doi.org/10.1093/ve/veaa061

Faria, N.R., Mellan, T.A., Whittaker, C., Claro, I.M., Candido, D. da S., Mishra, S., Crispim, M.A.E., Sales, F.C.S., Hawryluk, I., McCrone, J.T., Hulswit, R.J.G., Franco, L.A.M., Ramundo, M.S., de Jesus, J.G., Andrade, P.S., Coletti, T.M.,

Ferreira, G.M., Silva, C.A.M., Manuli, E.R., Pereira, R.H.M., Peixoto, P.S., Kraemer, M.U.G., Gaburo, N., Camilo, C. da C., Hoeltgebaum, H., Souza, W.M., Rocha, E.C., de Souza, L.M., de Pinho, M.C., Araujo, L.J.T., Malta, F.S. v., de Lima, A.B., Silva, J. do P., Zauli, D.A.G., Ferreira, A.C. de S., Schnekenberg, R.P., Laydon, D.J., Walker, P.G.T., Schlüter, H.M., dos Santos, A.L.P., Vidal, M.S., del Caro, V.S., Filho, R.M.F., dos Santos, H.M., Aguiar, R.S., Proença-Modena, J.L., Nelson, B., Hay, J.A., Monod, M., Miscouridou, X., Coupland, H., Sonabend, R., Vollmer, M., Gandy, A., Prete, C.A., Nascimento, V.H., Suchard, M.A., Bowden, T.A., Pond, S.L.K., Wu, C.-H., Ratmann, O., Ferguson, N.M., Dye, C., Loman, N.J., Lemey, P., Rambaut, A., Fraiji, N.A., Carvalho, M. do P.S.S., Pybus, O.G., Flaxman, S., Bhatt, S., Sabino, E.C., 2021. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. Science (1979) 372, 815–821. https://doi.org/10.1126/science.abh2644

Finkel, Y., Mizrahi, O., Nachshon, A., Weingarten-Gabbay, S., Morgenstern, D., Yahalom-Ronen, Y., Tamir, H., Achdout, H., Stein, D., Israeli, O., Beth-Din, A., Melamed, S., Weiss, S., Israely, T., Paran, N., Schwartz, M., Stern-Ginossar, N., 2021. The coding capacity of SARS-CoV-2. Nature 589, 125–130. https://doi.org/10.1038/s41586-020-2739-1

Foster, C.S.P., Rawlinson, W.D., 2021. Rapid spread of a SARS-CoV-2 Delta variant with a frameshift deletion in ORF7a. medRxiv 2021.08.18.21262089. https://doi.org/10.1101/2021.08.18.21262089

Gamage, A.M., Tan, K. sen, Chan, W.O.Y., Liu, J., Tan, C.W., Ong, Y.K., Thong, M., Andiappan, A.K., Anderson, D.E., Wang, D.Y., Wang, L.-F., 2020. Infection of human Nasal Epithelial Cells with SARS-CoV-2 and a 382-nt deletion isolate lacking ORF8 reveals similar viral kinetics and host transcriptional profiles. PLoS Pathog 16, e1009130. https://doi.org/10.1371/journal.ppat.1009130

Gorbalenya, A., Baker, S., Baric, R., de Groot, R., Drosten, C., Gulyaeva, A., Haagmans, B., Lauber, C., Leontovich, A., Neuman, B., 2020. Coronaviridae Study Group of the International Committee on Taxonomy of Viruses. The species severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nat Microbiol 2020.

Grubaugh, N.D., Petrone, M.E., Holmes, E.C., 2020. We shouldn't worry when a virus mutates during disease outbreaks. Nat Microbiol 7–8. https://doi.org/10.1038/s41564-020-0690-4

Hughes, L., Gangavarapu, K., Latif, A.A., Mullen, J., Alkuzweny, M., Hufbauer, E., Tsueng, G., Haag, E., Zeller, M., Aceves, C., Zaiets, K., Cano, M., Zhou, J., Qian, Z., Sattler, R., Matteson, N., Levy, J., Lee, R., Freitas, L., Maurer-Stroh, S., Suchard, M., Wu, C., Su, A., Andersen, K., 2022. Outbreak.info genomic reports: scalable and dynamic surveillance of SARS-CoV-2 variants and mutations. Res Sq. https://doi.org/10.21203/rs.3.rs-1723829/v1

Joonlasak, K., Batty, E.M., Kochakarn, T., Panthan, B., Kümpornsin, K., Jiaranai, P., Wangwiwatsin, A., Huang, A., Kotanan, N., Jaru-Ampornpan, P., Manasatienkij, W., Watthanachockchai, T., Rakmanee, K., Jones, A.R., Fernandez, S., Sensorn, I.,

Sungkanuparph, S., Pasomsub, E., Klungthong, C., Chookajorn, T., Chantratita, W., 2021. Genomic surveillance of SARS-CoV-2 in Thailand reveals mixed imported populations, a local lineage expansion and a virus with truncated ORF7a. Virus Res 292, 198233. https://doi.org/10.1016/j.virusres.2020.198233

Katoh, K., Standley, D.M., 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol Biol Evol 30, 772–780. https://doi.org/10.1093/molbev/mst010

Kemp, S.A., Meng, B., Ferriera, I.A.T.M., Datir, R., Harvey, W.T., Collier, D.A., Lytras, S., Papa, G., Carabelli, A., Kenyon, J., Lever, A., James, L.C., Robertson, D., Gupta, R.K., 2021. Recurrent Emergence and Transmission of a SARS-CoV-2 Spike Deletion H69/V70. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3780277

Kim, D., Lee, J.-Y., Yang, J.-S., Kim, J.W., Kim, V.N., Chang, H., 2020. The Architecture of SARS-CoV-2 Transcriptome. Cell 181, 914-921.e10. https://doi.org/10.1016/j.cell.2020.04.011

Lai, M.M., 1992. RNA recombination in animal and plant viruses. Microbiol Rev 56, 61–79.

Lai, M.M., Cavanagh, D., 1997. The molecular biology of coronaviruses. Adv Virus Res 48, 1–100.

Lau, S.-Y., Wang, P., Mok, B.W.-Y., Zhang, A.J., Chu, H., Lee, A.C.-Y., Deng, S., Chen, P., Chan, K.-H., Song, W., Chen, Z., To, K.K.-W., Chan, J.F.-W., Yuen, K.-Y., Chen, H., 2020. Attenuated SARS-CoV-2 variants with deletions at the S1/S2 junction. Emerg Microbes Infect 9, 837–842. https://doi.org/10.1080/22221751.2020.1756700

Li, J.Y., Liao, C.H., Wang, Q., Tan, Y.J., Luo, R., Qiu, Y., Ge, X.Y., 2020. The ORF6, ORF8 and nucleocapsid proteins of SARS-CoV-2 inhibit type I interferon signaling pathway. Virus Res 286, 198074. https://doi.org/10.1016/j.virusres.2020.198074

Liu, D.X., Fung, T.S., Chong, K.K., Shukla, A., Hilgenfeld, R., 2014. Accessory proteins of SARS-CoV and other coronaviruses. Antiviral Res 109, 97–109. https://doi.org/10.1016/j.antiviral.2014.06.013

Lorenz, R., Bernhart, S.H., Höner zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F., Hofacker, I.L., 2011. ViennaRNA Package 2.0. Algorithms for Molecular Biology 6, 26. https://doi.org/10.1186/1748-7188-6-26

Mazur-Panasiuk, N., Rabalski, L., Gromowski, T., Nowicki, G., Kowalski, M., Wydmanski, W., Szulc, P., Kosinski, M., Gackowska, K., Drweska-Matelska, N., Grabowski, J., Piotrowska-Mietelska, A., Szewczyk, B., Bienkowska-Szewczyk, K., Swadzba, J., Labaj, P., Grzybek, M., Pyrc, K., 2021. Expansion of a SARS-CoV-2 Delta variant with an 872 nt deletion encompassing ORF7a, ORF7b and ORF8, Poland, July to August 2021. Euro Surveill 26, 1–6. https://doi.org/10.2807/1560-7917.ES.2021.26.39.2100902

McCarthy, K.R., Rennick, L.J., Nambulli, S., Robinson-McCarthy, L.R., Bain, W.G., Haidar, G., Duprex, W.P., 2021. Recurrent deletions in the SARS-CoV-2 spike glycoprotein drive antibody escape. Science 371, 1139–1142. https://doi.org/10.1126/science.abf6950

Michel, C.J., Mayer, C., Poch, O., Thompson, J.D., 2020. Characterization of accessory genes in coronavirus genomes. Virol J 17, 131. https://doi.org/10.1186/s12985-020-01402-1

Mishra, S., Mindermann, S., Sharma, M., Whittaker, C., Mellan, T.A., Wilton, T., Klapsa, D., Mate, R., Fritzsche, M., Zambon, M., Ahuja, J., Howes, A., Miscouridou, X., Nason, G.P., Ratmann, O., Semenova, E., Leech, G., Sandkühler, J.F., Rogers-Smith, C., Vollmer, M., Unwin, H.J.T., Gal, Y., Chand, M., Gandy, A., Martin, J., Volz, E., Ferguson, N.M., Bhatt, S., Brauner, J.M., Flaxman, S., 2021. Changing composition of SARS-CoV-2 lineages and rise of Delta variant in England. EClinicalMedicine 39, 101064. https://doi.org/10.1016/j.eclinm.2021.101064

Moore, S.C., Penrice-Randal, R., Alruwaili, M., Randle, N., Armstrong, S., Hartley, C., Haldenby, S., Dong, X., Alrezaihi, A., Almsaud, M., Bentley, E., Clark, J., García-Dorival, I., Gilmore, P., Han, X., Jones, B., Lau, L., Sharma, P., Shawli, G., Sun, Y., Zhao, Q., Pullan, S.T., Carter, D.P., Bewley, K., Dunning, J., Zhou, E., Solomon, T., Beadsworth, M., Cruise, J., Crook, D.W., Matthews, D.A., Davidson, A.D., Mahmood, Z., Aljabr, W., Druce, J., Vipond, R., Ng, L., Renia, L., Openshaw, P.J.M., Baillie, J.K., Carroll, M.W., Stewart, J., Darby, A., Semple, M., Turtle, L., Hiscox, J.A., 2020. Amplicon-Based Detection and Sequencing of SARS-CoV-2 in Nasopharyngeal Swabs from Patients With COVID-19 and Identification of Deletions in the Viral Genome That Encode Proteins Involved in Interferon Antagonism. Viruses 12, 1164. https://doi.org/10.3390/v12101164

Muth, D., Corman, V.M., Roth, H., Binger, T., Dijkman, R., Gottula, L.T., Gloza-Rausch, F., Balboni, A., Battilani, M., Rihtarič, D., Toplak, I., Ameneiros, R.S., Pfeifer, A., Thiel, V., Drexler, J.F., Müller, M.A., Drosten, C., 2018. Attenuation of replication by a 29 nucleotide deletion in SARS-coronavirus acquired during the early stages of human-to-human transmission. Sci Rep 8, 15177. https://doi.org/10.1038/s41598-018-33487-8

Narayanan, K., Huang, C., Makino, S., 2008. SARS coronavirus accessory proteins. Virus Res 133, 113–21. https://doi.org/10.1016/j.virusres.2007.10.009

Nemudryi, A., Nemudraia, A., Wiegand, T., Nichols, J., Snyder, D.T., Hedges, J.F., Cicha, C., Lee, H., Vanderwood, K.K., Bimczok, D., Jutila, M.A., Wiedenheft, B., 2021. SARS-CoV-2 genomic surveillance identifies naturally occurring truncation of ORF7a that limits immune suppression. Cell Rep 35, 109197. https://doi.org/10.1016/j.celrep.2021.109197

O'Toole, Á., Scher, E., Underwood, A., Jackson, B., Hill, V., McCrone, J.T., Colquhoun, R., Ruis, C., Abu-Dahab, K., Taylor, B., Yeats, C., du Plessis, L., Maloney, D., Medd, N., Attwood, S.W., Aanensen, D.M., Holmes, E.C., Pybus, O.G., Rambaut, A., 2021. Assignment of Epidemiological Lineages in an

Emerging Pandemic Using the Pangolin Tool. Virus Evol. https://doi.org/10.1093/ve/veab064

Pancer, K., Milewska, A., Owczarek, K., Dabrowska, A., Kowalski, M., Łabaj, P.P., Branicki, W., Sanak, M., Pyrc, K., 2020. The SARS-CoV-2 ORF10 is not essential in vitro or in vivo in humans. PLoS Pathog 16, e1008959. https://doi.org/10.1371/journal.ppat.1008959

Panzera, Y., Calleros, L., Goñi, N., Marandino, A., Techera, C., Grecco, S., Ramos, N., Frabasile, S., Tomás, G., Condon, E., Cortinas, M.N., Ramas, V., Coppola, L., Sorhouet, C., Mogdasy, C., Chiparelli, H., Arbiza, J., Delfraro, A., Pérez, R., 2022. Consecutive deletions in a unique Uruguayan SARS-CoV-2 lineage evidence the genetic variability potential of accessory genes. PLoS One 17, e0263563. https://doi.org/10.1371/journal.pone.0263563

Panzera, Y., Ramos, N., Calleros, L., Marandino, A., Tomás, G., Techera, C., Grecco, S., Frabasile, S., Fuques, E., Coppola, L., Goñi, N., Ramas, V., Sorhouet, C., Bormida, V., Burgueño, A., Brasesco, M., Garland, M.R., Molinari, S., Perez, M.T., Somma, R., Somma, S., Morel, M.N., Mogdasy, C., Chiparelli, H., Arbiza, J., Delfraro, A., Pérez, R., 2021a. Transmission cluster of COVID-19 cases from Uruguay: emergence and spreading of a novel SARS-CoV-2 ORF6 deletion. Mem Inst Oswaldo Cruz 116. https://doi.org/10.1590/0074-02760210275

Panzera, Y., Ramos, N., Frabasile, S., Calleros, L., Marandino, A., Tomás, G., Techera, C., Grecco, S., Fuques, E., Goñi, N., Ramas, V., Coppola, L., Chiparelli, H., Sorhouet, C., Mogdasy, C., Arbiza, J., Delfraro, A., Pérez, R., 2021b. A deletion in SARS-CoV-2 ORF7 identified in COVID-19 outbreak in Uruguay. Transbound Emerg Dis 68, 3075–3082. https://doi.org/10.1111/tbed.14002

Pereira, F., 2020. Evolutionary dynamics of the SARS-CoV-2 ORF8 accessory gene. Infect Genet Evol 85, 104525. https://doi.org/10.1016/j.meegid.2020.104525

Price, M.N., Dehal, P.S., Arkin, A.P., Liu, D.X., Fung, T.S., Chong, K.K., Shukla, A., Hilgenfeld, R., Fogerson, M., Montserret, R., Zehnder, J., Nguyen, M., Dujardin, M., Cole, L., Minor Pedrosa, M., Lecoq, L., Meier, B.H., Alexandersen, S., Chamings, A., Bhatta, T.R., 2009. FastTree: Computing Large Minimum Evolution Trees with Profiles instead of a Distance Matrix. Mol Biol Evol 26, 1641–1650. https://doi.org/10.1093/molbev/msp077

Pyke, A.T., Nair, N., van den Hurk, A.F., Burtonclay, P., Nguyen, S., Barcelon, J., Kistler, C., Schlebusch, S., McMahon, J., Moore, F., 2021. Replication Kinetics of B.1.351 and B.1.1.7 SARS-CoV-2 Variants of Concern Including Assessment of a B.1.1.7 Mutant Carrying a Defective ORF7a Gene. Viruses 13, 1087. https://doi.org/10.3390/v13061087

Rambaut, A., Holmes, E.C., O'Toole, Á., Hill, V., McCrone, J.T., Ruis, C., du Plessis, L., Pybus, O.G., 2020. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nat Microbiol 5, 1403–1407. https://doi.org/10.1038/s41564-020-0770-5

Sabino, E.C., Buss, L.F., Carvalho, M.P.S.S., Prete, C.A., Crispim, M.A.E.E., Fraiji, N.A., Pereira, R.H.M., Parag, K. v., da Silva Peixoto, P., Kraemer, M.U.G.G., Oikawa, M.K., Salomon, T., Cucunuba, Z.M., Castro, M.C., de Souza Santos, A.A., Nascimento, V.H., Pereira, H.S., Ferguson, N.M., Pybus, O.G., Kucharski, A., Busch, M.P., Dye, C., Faria, N.R., 2021. Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. Lancet 397, 452–455. https://doi.org/10.1016/S0140-6736(21)00183-5

Sawicki, S.G., Sawicki, D.L., 1995. Coronaviruses use discontinuous extension for synthesis of subgenome-length negative strands. Adv Exp Med Biol 380, 499–506. https://doi.org/10.1007/978-1-4615-1899-0_79

Shu, Y., McCauley, J., 2017. GISAID: Global initiative on sharing all influenza data – from vision to reality. Eurosurveillance 22, 30494. https://doi.org/10.2807/1560-7917.ES.2017.22.13.30494

Silvas, J.A., Vasquez, D.M., Park, J.-G., Chiem, K., Allué Guardia, A., Garcia-Vilanova, A., Platt, R.N., Miorin, L., Kehrer, T., Cupic, A., Gonzalez-Reiche, A.S., Bakel, H. van, García-Sastre, A., Anderson, T., Torrelles, J.B., Ye, C., Martinez-Sobrido, L., 2021. Contribution of SARS-CoV-2 Accessory Proteins to Viral Pathogenicity in K18 Human ACE2 Transgenic Mice. J Virol 95, e0040221. https://doi.org/10.1128/JVI.00402-21

Simon-Loriere, E., Holmes, E.C., 2011. Why do RNA viruses recombine? Nat Rev Microbiol 9, 617–626. https://doi.org/10.1038/nrmicro2614

Singer, J., Gifford, R., Cotten, M., Robertson, D., 2020. CoV-GLUE: A Web Application for Tracking SARS-CoV-2 Genomic Variation. Preprint. https://doi.org/10.20944/preprints202006.0225.v1

Singh, J., Rahman, S.A., Ehtesham, N.Z., Hira, S., Hasnain, S.E., 2021. SARS-CoV-2 variants of concern are emerging in India. Nat Med 27, 1131–1133. https://doi.org/10.1038/s41591-021-01397-4

Stadler, K., Masignani, V., Eickmann, M., Becker, S., Abrignani, S., Klenk, H.-D., Rappuoli, R., 2003. SARS--beginning to understand a new virus. Nat Rev Microbiol 1, 209–18. https://doi.org/10.1038/nrmicro775

Su, C.-M., Wang, L., Yoo, D., 2021. Activation of NF-κB and induction of proinflammatory cytokine expressions mediated by ORF7a protein of SARS-CoV-2. Sci Rep 11, 13464. https://doi.org/10.1038/s41598-021-92941-2

Su, Y.C.F., Anderson, D.E., Young, B.E., Linster, M., Zhu, F., Jayakumar, J., Zhuang, Y., Kalimuddin, S., Low, J.G.H., Tan, C.W., Chia, W.N., Mak, T.M., Octavia, S., Chavatte, J.-M., Lee, R.T.C., Pada, S., Tan, S.Y., Sun, L., Yan, G.Z., Maurer-Stroh, S., Mendenhall, I.H., Leo, Y.-S., Lye, D.C., Wang, L.-F., Smith, G.J.D., 2020. Discovery and Genomic Characterization of a 382-Nucleotide Deletion in ORF7b and ORF8 during the Early Evolution of SARS-CoV-2. mBio 11. https://doi.org/10.1128/mBio.01610-20

Tegally, H., Wilkinson, E., Giovanetti, M., Iranzadeh, A., Fonseca, V., Giandhari, J., Doolabh, D., Pillay, S., San, E.J., Msomi, N., Mlisana, K., von Gottberg, A., Walaza, S., Allam, M., Ismail, A., Mohale, T., Glass, A.J., Engelbrecht, S., van Zyl, G., Preiser, W., Petruccione, F., Sigal, A., Hardie, D., Marais, G., Hsiao, N.-Y., Korsman, S., Davies, M.-A., Tyers, L., Mudau, I., York, D., Maslo, C., Goedhals, D., Abrahams, S., Laguda-Akingba, O., Alisoltani-Dehkordi, A., Godzik, A., Wibmer, C.K., Sewell, B.T., Lourenço, J., Alcantara, L.C.J., Kosakovsky Pond, S.L., Weaver, S., Martin, D., Lessells, R.J., Bhiman, J.N., Williamson, C., de Oliveira, T., 2021. Detection of a SARS-CoV-2 variant of concern in South Africa. Nature 592, 438–443. https://doi.org/10.1038/s41586-021-03402-9

Thiel, V., Ivanov, K.A., Putics, Á., Hertzig, T., Schelle, B., Bayer, S., Weißbrich, B., Snijder, E.J., Rabenau, H., Doerr, H.W., Gorbalenya, A.E., Ziebuhr, J., 2003. Mechanisms and enzymes involved in SARS coronavirus genome expression. J Gen Virol 84, 2305–2315. https://doi.org/10.1099/vir.0.19424-0

Tse, H., Lung, D.C., Wong, S.C.-Y., Ip, K.-F., Wu, T.-C., To, K.K.-W., Kok, K.-H., Yuen, K.-Y., Choi, G.K.-Y., 2021a. Emergence of a Severe Acute Respiratory Syndrome Coronavirus 2 Virus Variant With Novel Genomic Architecture in Hong Kong. Clin Infect Dis 73, 1696–1699. https://doi.org/10.1093/cid/ciab198

Tse, H., Wong, S.C.-Y., Ip, K.-F., Cheng, V.C.-C., To, K.K.-W., Lung, D.C., Choi, G.K.-Y., 2021b. Genome Sequences of Three SARS-CoV-2 ORF7a Deletion Variants Obtained from Patients in Hong Kong. Microbiol Resour Announc 10, 7–9. https://doi.org/10.1128/MRA.00251-21

Volz, E., Mishra, S., Chand, M., Barrett, J.C., Johnson, R., Geidelberg, L., Hinsley, W.R., Laydon, D.J., Dabrera, G., O'Toole, Á., Amato, R., Ragonnet-Cronin, M., Harrison, I., Jackson, B., Ariani, C. v, Boyd, O., Loman, N.J., McCrone, J.T., Gonçalves, S., Jorgensen, D., Myers, R., Hill, V., Jackson, D.K., Gaythorpe, K., Groves, N., Sillitoe, J., Kwiatkowski, D.P., COVID-19 Genomics UK (COG-UK) consortium, Flaxman, S., Ratmann, O., Bhatt, S., Hopkins, S., Gandy, A., Rambaut, A., Ferguson, N.M., 2021. Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. Nature 593, 266–269. https://doi.org/10.1038/s41586-021-03470-x

Yan, H., Xiao, G., Zhang, J., Hu, Y., Yuan, F., Cole, D.K., Zheng, C., Gao, G.F., 2004. SARS coronavirus induces apoptosis in Vero E6 cells. J Med Virol 73, 323–31. https://doi.org/10.1002/jmv.20094

Yang, R., Zhao, Q., Rao, J., Zeng, F., Yuan, S., Ji, M., Sun, X., Li, J., Yang, J., Cui, J., Jin, Z., Liu, L., Liu, Z., 2021. SARS-CoV-2 Accessory Protein ORF7b Mediates Tumor Necrosis Factor-α-Induced Apoptosis in Cells. Front Microbiol 12, 654709. https://doi.org/10.3389/fmicb.2021.654709

Yeh, S.-H., Wang, H.-Y., Tsai, C.-Y., Kao, C.-L., Yang, J.-Y., Liu, H.-W., Su, I.-J., Tsai, S.-F., Chen, D.-S., Chen, P.-J., National Taiwan University SARS Research Team, 2004. Characterization of severe acute respiratory syndrome coronavirus

genomes in Taiwan: molecular epidemiology and genome evolution. Proc Natl Acad Sci U S A 101, 2542–7. https://doi.org/10.1073/pnas.0307904100

Young, B.E., Fong, S.-W., Chan, Y.-H., Mak, T.-M., Ang, L.W., Anderson, D.E., Lee, C.Y.-P., Amrun, S.N., Lee, B., Goh, Y.S., Su, Y.C.F., Wei, W.E., Kalimuddin, S., Chai, L.Y.A., Pada, S., Tan, S.Y., Sun, L., Parthasarathy, P., Chen, Y.Y.C., Barkham, T., Lin, R.T.P., Maurer-Stroh, S., Leo, Y.-S., Wang, L.-F., Renia, L., Lee, V.J., Smith, G.J.D., Lye, D.C., Ng, L.F.P., 2020. Effects of a major deletion in the SARS-CoV-2 genome on the severity of infection and the inflammatory response: an observational cohort study. The Lancet 396, 603–611. https://doi.org/10.1016/s0140-6736(20)31757-8

Zhang, Y., Chen, Y., Li, Y., Huang, F., Luo, B., Yuan, Y., Xia, B., Ma, X., Yang, T., Yu, F., Liu, J., Liu, B., Song, Z., Chen, J., Yan, S., Wu, L., Pan, T., Zhang, X., Li, R., Huang, W., He, X., Xiao, F., Zhang, J., Zhang, H., 2021. The ORF8 protein of SARS-CoV-2 mediates immune evasion through down-regulating MHC-I. Proc Natl Acad Sci U S A 118. https://doi.org/10.1073/pnas.2024202118

Zhou, Ziliang, Huang, C., Zhou, Zhechong, Huang, Z., Su, L., Kang, S., Chen, X., Chen, Q., He, S., Rong, X., Xiao, F., Chen, J., Chen, S., 2021. Structural insight reveals SARS-CoV-2 ORF7a as an immunomodulating factor for human CD14+ monocytes. iScience 24, 102187. https://doi.org/10.1016/j.isci.2021.102187

Zinzula, L., 2020. Lost in deletion: The enigmatic ORF8 protein of SARS-CoV-2. Biochem Biophys Res Commun 64, 151–5. https://doi.org/10.1016/j.bbrc.2020.10.045

Tables

Table 1. Epidemiological data of the Delta (B.1.617.2) samples from Uruguay. Accession numbers are indicated.

Figure captions

Figure 1. A) Deletion differences between Delta variant and the SARS-CoV-2 reference NC_045512/Wuhan. Delta strains have two small deletions (6 and 1nt) upstream of the N gene according to the reference strain (NC045512). These small deletions are close to transcription-regulating sequences (TRS) (zoomed diagram below). B) The deletion surrounding the TRS (6 + 1-nt deletions) alters the RNA

folding and might promote the translocation of the replication–transcription complex (RTC) to other TRS. This translocation could be related to the subsequent occurrence of a large deletion (866 nucleotides) that removes three consecutive ORFs: 7a, 7b, and 8.

Figure 2. A: The 872-nucleotide deletion in the linage AY.20 of the Delta variant of concern. The deletion removes three consecutive ORFs: 7a, 7b, and 8. B: zoomed diagram details the transcription-regulating sequence (TRS) at the boundary of the 872-nucleotide deletion and Sanger sequencing chromatograms. The curved arrow represents the ability of the replicative machinery to "jump" between TRS and favor deletions. C: chromatogram peaks amplicon by capillary electrophoresis analysis, wildtype, and Δ872 variant.

Figure 3. Maximum-likelihood tree based on different lineages of the Delta (AY) SARS-CoV-2 variant of concern. The 872-nucleotide deletion of the AY.20, AY.90, and AY.4 from Uruguay, Japan, and Poland are highlighted in magenta.

Table 1. Epidemiological data of the Delta (B.1.617.2) samples from Uruguay.

Sample	Lineage	Date	Origin	Sex	Age	Clinical Data	Immunization	Comorbidities	Accession
p2262	AY.20	3-9-21	Montevideo	F	64	Symptomatic	2 (2 Sinovac)	NA	ON653593
p2272	AY.20	10-9-21	Montevideo	F	43	Symptomatic	2	No	OM971724
p2277	AY.20	13-9-21	Montevideo	M	33	NA	NA	NA	ON653595
p2278	AY.20	13-9-21	Montevideo	M	49	Asymptomatic	not vaccinated	already infected in Jan 2021	OM971728
6553	AY.20	6-12-21	Montevideo	M	34	NA	NA	No	OM971721
p2339	AY.20	6-12-21	Montevideo	F	NA	Symptomatic	NA	NA	ON653594
p2269	AY.43	10-9-21	Montevideo	F	55	Symptomatic	3 (2 Sinovac + 1 Pfizer)	No	OM974322
p2270	AY.43	10-9-21	Montevideo	M	53	Symptomatic	2 (Pfizer)	No	OM971722
*p2271	AY.43	10-9-21	Montevideo	F	47	Asymptomatic	2	In hospital for	OM971723

								surgery	
p2273	AY.43	10-9-21	Uruguay	F	53	NA	NA	NA	OM971725
p2274	AY.43	10-9-21	Canelones	F	67	Mild symptoms, only cough	NA	No	OM971726
p2275	AY.43	10-9-21	Montevideo	M	78	NA	NA	No	OM971727
p2283	AY.43	20-9-21	Canelones	M	61	Asymptomatic	3 (3 Sinovac + 1 Pfizer)	No	OM971729
p2284	AY.43	20-9-21	Montevideo	M	49	Symptomatic	2 (Sinovac)	No	OM971730
p2285	AY.43	20-9-21	Montevideo	M	16	Symptomatic	not vaccinated	No	OM971731
5827	AY.43	22-9-21	Montevideo	F	33	Symptomatic	2 (Sinovac)	Diabetes	OM971713
5842	AY.43	24-9-21	Montevideo	M	5	Symptomatic	2 (Sinovac)	No	OM971714
p2287	AY.43	24-9-21	Montevideo	F	83	Symptomatic	NA	No	OM971732
p2288	AY.43	27-9-21	Uruguay	M	53	NA	NA	No	OM971733
p2289	AY.43	27-9-21	Uruguay	F	58	NA	NA	No	OM971734
p2290	AY.43	27-Sep-21	Montevideo	M	26	Mild Symptoms	2 (Sinovac)	No	OM971735
5868	AY.43	28-9-21	Montevideo	M	26	Symptomatic	3	No	OM971715
5900	AY.43	4-10-21	Canelones	M	44	Symptomatic	3	No	OM971716
p2292	AY.43	5-10-21	Canelones	F	40	Symptomatic	2 (Pfizer)	No	OM971736
6041	AY.43	14-10-21	Montevideo	F	83	Mild symptoms, only cough	2 (Pfizer)	No	OM971717
p2294	AY.43	15-10-21	Canelones	M	21	Symptomatic	3 (2 Sinovac + 1 Pfizer)	No	OM971737
6126	AY.43	20-10-21	Montevideo	F	9	NA	Not vaccinated	No	OM971718
p2299	AY.43	21-10-21	Montevideo	F	36	Odynophagia	3 (Pfizer)	No	OM971739
p2300	AY.43	21-10-21	Montevideo	F	35	NA	NA	NA	OM971740
p2301	AY.43	25-10-21	Uruguay	F	59	NA	NA	NA	OM971741
p2302	AY.43	25-10-21	Treinta y Tres	M	25	NA	NA	NA	OM971742
p2306	AY.43	3-11-21	Canelones	F	31	NA	NA	NA	OM971743
p2308	AY.43	3-11-21	Montevideo	M	37	NA	NA	NA	OM971744
p2311	AY.43	4-Nov-21	Montevideo	F		Symptomatic	2 (Sinovac)	No	OM971745
p2314	AY.43	8-Nov-21	Montevideo	M	37	Mild Symptoms	3 (2 Sinovac + 1 Pfizer)	smoker	OM971746
6226	AY.43	9-Nov-21	Montevideo	F	31	Symptomatic	NA	Recent smoker	OM971719
6533	AY.43	3-Dec-21	Montevideo	F	91	Mild Symptoms	2 (Pfizer)	No	OM971720

*patient p2271 required hospitalization for surgery (none of the other patients required hospitalization). NA: no available data. There is no record of an epidemiological link among the patients.

Abbreviation list

Δ	deletion
ORF	Open Reading Frame
RTC	replication–transcription complex
TRS	transcription-regulating sequence
VOC	variants of concern

Author contributions

Yanina Panzera: Conceptualization, Methodology, Validation, and Supervision. Maria Noel Cortinas: Data Curation and Formal analysis, Ana Marandino, Lucía Calleros, Victoria Bormida, Natalia Goñi, Claudia Techera, Sofía Grecco, Joaquín Williman, Viviana Ramas and Leticia Coppola: Formal análysis. Cristina Mogdasy and Héctor Chiparelli: Resources and Funding acquisition. Ruben Pérez. Conceptualization, Visualization, Methodology, and Writing (original draft). All authors revised and approved the manuscript.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Highlights

- An extreme SARS-CoV-2 deletion emerged in a Delta strain from Uruguay
- The deletion spans 872 nucleotides and removes the 7a, 7b, and 8 ORFs
- This largest deletion occurs adjacent to transcription regulatory sequences (TRS)
- TRS might promote deletions by intramolecular translocation of the replication–transcription complex

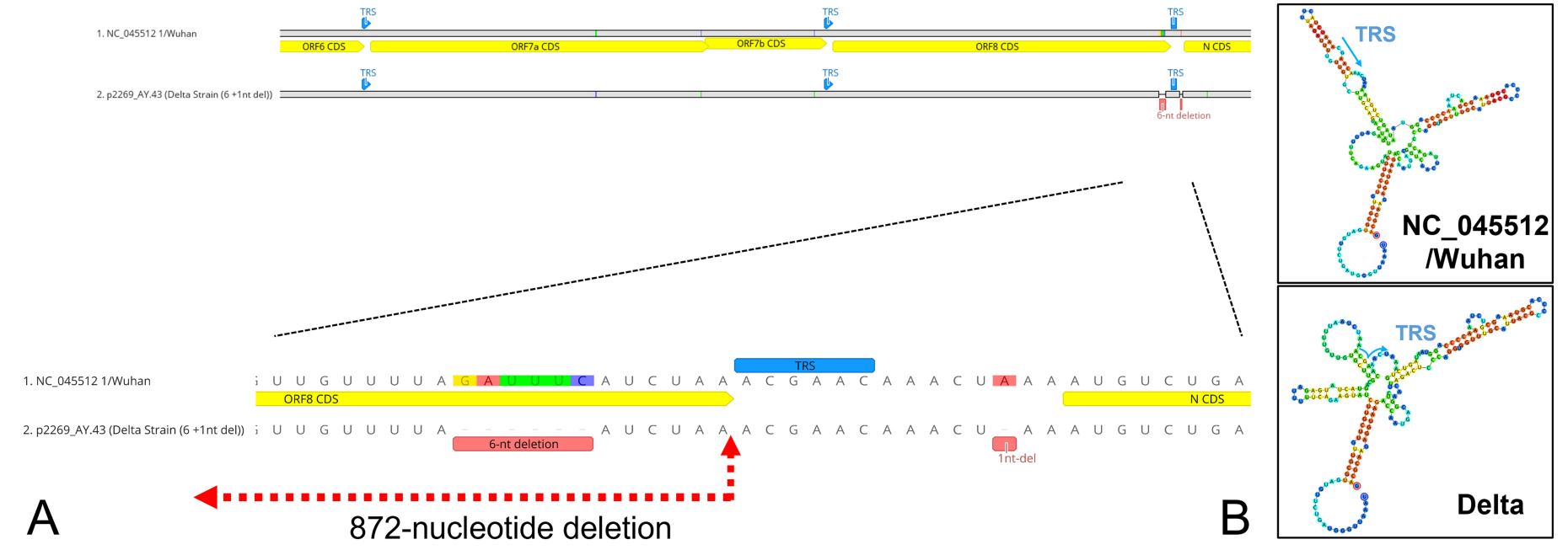
1. NC_045512 1/Wuhan
TRS
ORF6 CDS
ORF7a CDS
ORF7b CDS
ORF8 CDS
N CDS
2. p2269_AY.43 (Delta Strain (6 +1nt del))
6-nt deletion
1. NC_045512 1/Wuhan
; U U G U U U U A G A U U U C A U C U A A A C G A A C A A A A C U A A A A U G U C U G A
ORF8 CDS
2. p2269_AY.43 (Delta Strain (6 +1nt del))
; U U G U U U U A A U C U A A A C G A A C A A A A C U A A A U G U C U G A
6-nt deletion
1nt-del
872-nucleotide deletion
A
B
TRS
NC_045512
/Wuhan
Delta

Figure 1


A
ORF6 CDS
ORF7a CDS
ORF7b CDS
ORF8 CDS
N CDS
872 nt deletion
B
27,380
27,390
27,400
28,260
28,270
28,280
TRS
AGAUUGAUUAAACGAACAUGAAAAUU
AGAUUGAUUAA
CAUCUAAACGAACAAACUAAAAUGUCUGAU
ACGAACAAACU-AAAUGUCUGAU
872 nt deletion
27,358
27,368
27,378
27,388
27,398
27,408
27,418
GAAGAGCAACCAAUGGAGAUUGAUUAAACGAACAAACUAAAAUGUCUGAUAAUGGACCCCAAAAUCAGCGA
REV Amplif.Delta AY.20
FWD Amplif.Delta AY.20
C
1323
450
RFU
Size (bp)

Figure 2

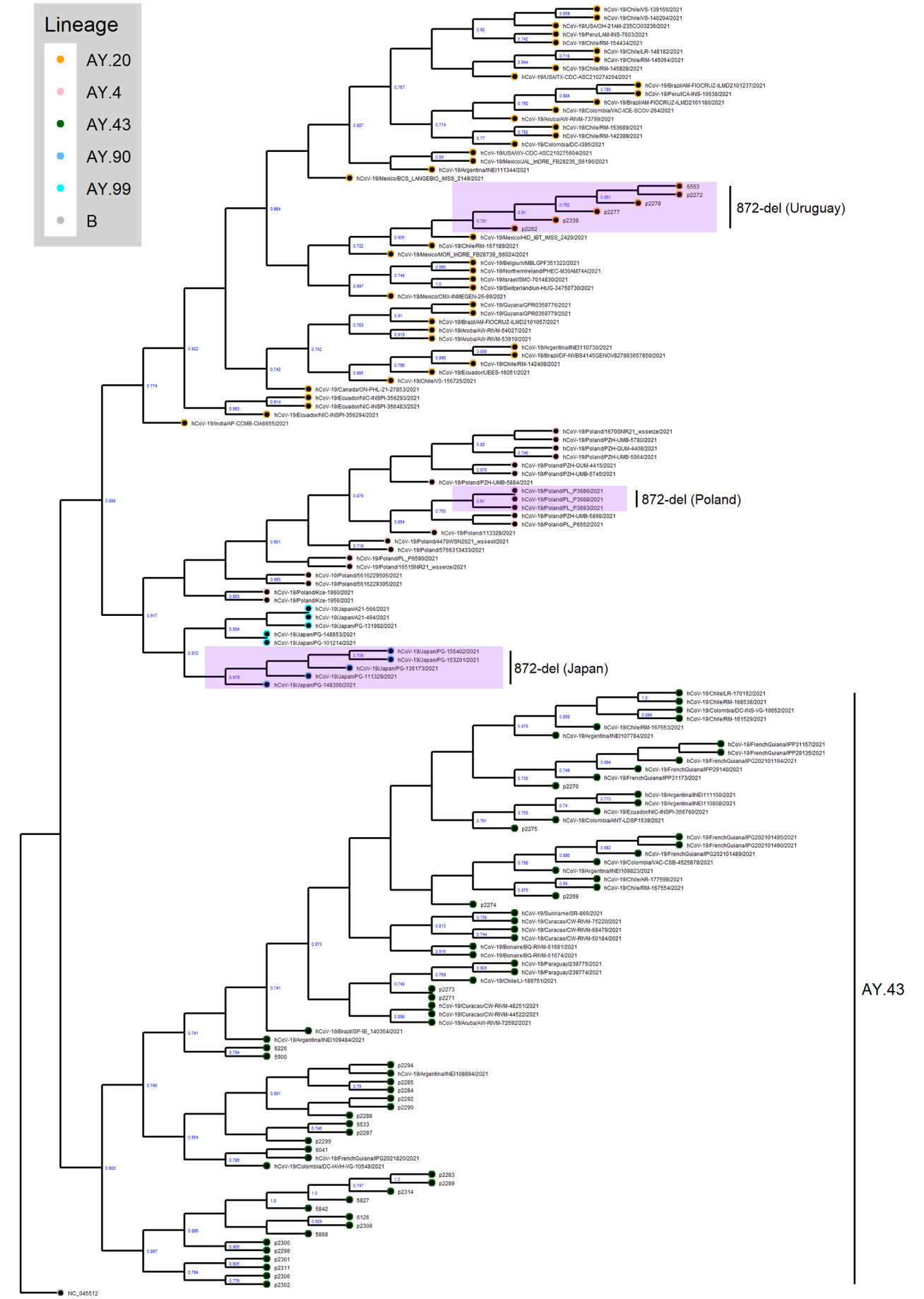
Lineage
AY.20
AY.4
AY.43
AY.90
AY.99
B
872-del (Uruguay)
872-del (Poland)
872-del (Japan)
AY.43
NC_045512

Figure 3