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.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Emergence and spreading of the largest SARS-CoV-2 deletion in the Delta AY.20 lineage from Uruguay

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

a Sección Genética Evolutiva, Departamento de Biología Animal, Instituto de Biología, Facultad de Ciencias, Universidad de la República, Iguazú 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

A R T I C L E  I N F O
Edited by: Christine Thien
Keywords:
SARS-CoV-2
Largest deletion
Accessory ORFs
Recombination hotspot

A B S T R A C T
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 coronavirus, 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, interspersed 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., 2021a, 2021b; 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 (Narayan 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 downstream ORFs (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 because they seem dispensable for replication (Grubaugh et al., 2020; Young et al., 2020). The same deletion in Southern Poland, an 872-nucleotide deletion (∆872) 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-Panasuik 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 (∆872) 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™ 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 (12 μl) 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 ∆872 (90 Left and 93 Right) as previously described (Mazur-Panasuik 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 GISAID 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 support to internal nodes, maximum-likelihood trees were inferred in Geneious using FastTree (Price et al., 2010) 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–December 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 (Fig. 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 L52300), 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 ORF6a, 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 1315 bp in delta strains (Fig. 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 (Fig. 1), the deletion has 872-nt (Δ872) compared with the reference Wuhan strain (Figs. 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 (Fig. 2). The deletion breakpoint was contiguous to the downstream TRS (the upstream TRS located at the very end of the deletion) (Fig. 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 (Fig. 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 subclasses. One comprises Uruguayan sequences, and the other clade clustered variants from Uruguay and South America (Fig. 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 (Alkhattib 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 in terms of the immune escape capability, and risk of reinfection (Alkhattib et al., 2021).
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 according to their classification. They are more related to wild-type strains 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.

Deletions are associated with RTC's homology-assisted mechanism to switch template strands during subgenomic and full-genome synthesis. RTC may dissociate from the template strand during the nascent strand synthesis and reassociate to the same template strand at different positions and had different phylogenetic relations with the same length that Poland and Japan variants. Remarkably, all AY.20 strains, suggesting a common origin (Fig. 3).

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-2, the truncation of ORF8 led to gradual virus attenuation in vitro (Muth et al., 2018). In SARS-CoV-2, the 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-2, the truncation of ORF8 led to gradual virus attenuation in vitro (Muth et al., 2018). In SARS-CoV-2, the 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-2, the truncation of ORF8 led to gradual virus attenuation in vitro (Muth et al., 2018). In SARS-CoV-2, the 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-2, the truncation of ORF8 led to gradual virus attenuation in vitro (Muth et al., 2018). In SARS-CoV-2, the accessory genes in host's immune response and viral pathogenicity.

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 (Fig. 3).

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.
Fig. 1. A) Deletion differences between Delta variant and the SARS-CoV-2 reference NC_045512/Wuhan. Delta strains have two small deletions (6 and 1 nt) 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.
Fig. 2. A: The 872-nucleotide deletion in the lineage 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.
protein encoded by ORF8 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 pathophysiological 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 (Mishra 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 (Mishra et al., 2021a). 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 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.

CRediT authorship contribution statement

Yanina Panzera: Conceptualization, Methodology, Validation, and Supervision. Maria Noel Cortinas: Data Curation and Formal analysis, Ana Marandino, Luciana Calleros, Victoria Bormida, Natalia Goti, Claudia Tetchera, Sofia Grecco, Joaquín Williman, Viviana Ramas and Leticia Coppola: Formal analysis. 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 competing interest

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.
Mishra, S., Mindermann, S., Sharma, M., Whittaker, C., Mellan, T.A., Wilton, T., Mazur-Panasiuk, N., Rabalski, L., Gromowski, T., Nowicki, G., Kowalski, M., Kemp, S.A., Meng, B., Ferriera, I.A.T.M., Datir, R., Harvey, W.T., Collier, D.A., Lytras, S., Lau, S.-Y., Wang, P., Mok, B.W.-Y., Zhang, A.J., Chu, H., Lee, A.C.-Y., Deng, S., Chen, P., Gorbalenya, A., Baker, S., Baric, R., de Groot, R., Drosten, C., Gulyaeva, A., Foster, C.S.P., Rawlinson, W.D., 2021. Rapid spread of a SARS-CoV-2 Delta variant with a Y.

Panzeria et al., 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.

Muth, D., Corman, V.M., Roth, B., Binger, T., Dijkman, R., Gottula, L.T., Glosa-Rausch, F., Balboni, A., Battilini, M., Rittaric, 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-CoV-2 genome 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, 113–121. https://doi.org/10.1016/j.virusres.2007.10.005.

Nemudryi, A., Nemudraia, A., Wiegand, T., Nichols, J., Snyder, D.T., Hedges, J.F., Ashenafi, N., Ramas, V., Coppola, L., Chiparelli, H., Arbiza, J., Dalfiero, A., Perez, R., 2022. Assignment of epidemiological lineages in an emerging pandemic using the pangolin tool. Virus Evol. 8 https://doi.org/10.1093/ve/veab064.

Pancer, K., Milewska, A., Owczarek, C., Babrowska, A., Kowalski, M., Labaj, 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.

Panzeria, Y., Ramos, N., Calleros, L., Marandino, A., Tomas, G., Tehera, G., Grecco, S., Frabasile, S., Cuquè, E., Coppola, L., Goti, N., Ramas, V., Sorhouet, C., Bormida, V., Burgueño, A., Brasesco, M., Garland, M.R., Molinari, S., Perez, M.T., Sorro, C., Somor, S., Morel, M.N., Mogdasy, C., Chiparelli, H., Arbiza, J., Dalfiero, A., Perez, R., 2021a. Translating the rate of COVID-19 cases from Uruguay: emergence and spreading of a novel SARS-CoV-2 ORF6 deletion. Mem. Inst. Oswaldo Cruz 116, 879–884. https://doi.org/10.1590/0074-027620210275.

Panzeria, Y., Ramos, N., Frabasile, S., Calleros, L., Marandino, A., Tomas, G., Tehera, G., Grecco, S., Frabasile, S., Cuquè, E., Coppola, L., Goti, N., Ramas, V., Sorhouet, C., Burgueño, A., Brasesco, M., Garland, M.R., Molinari, S., Perez, M.T., Sorro, C., Somor, S., Morel, M.N., Mogdasy, C., Chiparelli, H., Arbiza, J., Dalfiero, A., Perez, R., 2021b. A deletion in SARS-CoV-2 ORF7 identified in COVID-19 outbreak in Uruguay. Trends. Emerg. Dis. 68, 3075–3082. https://doi.org/10.1016/j.tred.2021.03.011.

Panzeria, Y., Calleros, L., G., Marandino, A., Tehera, G., Grecco, S., Ramos, N., Frabasile, S., Tomas, G., Condón, E., Cortinas, M., Ramas, V., Coppola, L., Sorhouet, C., Mogdasy, C., Chiparelli, H., Arbiza, J., Dalfiero, A., Perez, R., 2022. Concurrent 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.

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.

Perez, M.N., Dhali, P.S., Arkin, A.P., Liu, D.X., Fung, T.S., Chong, K.K., Shukla, A., Milenkovic, R., Foger, M., Bortner, R., Zehnder, J., Nguyen, Y., Dujardin, J., Cole, L., Niton-piédrosa, M., Lecoq, I., Meier, B.H., Alexanderens, S., Chaming, A., Bhutta, T.R., 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol. Biol. Evol. 26, 1401–1406. https://doi.org/10.1093/molev/bcp061.

Perez, M.N., Dhali, P.S., Arkin, A.P., Liu, D.X., Fung, T.S., Chong, K.K., Shukla, A., Milenkovic, R., Foger, M., Bortner, R., Zehnder, J., Nguyen, Y., Dujardin, J., Cole, L., Niton-piédrosa, M., Lecoq, I., Meier, B.H., Alexanderens, S., Chaming, A., Bhutta, T.R., 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol. Biol. Evol. 26, 1401–1406. https://doi.org/10.1093/molev/bcp061.

Pyke, A.T., Nair, N., van den Hurk, A.F., Burtonclay, P., Nguyen, S., Barcelon, J., Cucunuba, Z.M., Castro, M.C., de Souza Santos, A.A., Nascimento, V.H., Bhatta, T.R., 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.
Tegally, H., Wilkinson, E., Giovanetti, M., Iranzadeh, A., Fonseca, V., Giandhari, J., 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., Su, Y.C.F., Anderson, D.E., Young, B.E., Linster, M., Zhu, F., Jayakumar, J., Zhuang, Y., Stadler, K., Masignani, V., Eickmann, M., Becker, S., Abrignani, S., Klenk, H.-D., Singh, J., Rahman, S.A., Ehtesham, N.Z., Hira, S., Hasnain, S.E., 2021. SARS-CoV-2 from patients in Hong Kong. Microbiol. Resour. Announc. 10, 7

2021b. Genome sequences of three SARS-CoV-2 ORF7a deletion variants obtained during the early evolution of SARS-CoV-2. mBio 11. https://doi.org/10.1128/mbio.01610-20.

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.

Tegally, H., Wilkinson, E., Giovanetti, M., Iranzadeh, A., Fonseca, V., Giandhari, J., Doolabh, D., Pilla, S., San, E.J., Moumi, N., Miliana, K., von Gottberg, A., Walaza, S., Allam, M., Ismail, A., Mzhou, T., Glass, A.J., Engelbrecht, S., van Zyl, G., Reiser, W., Petruccione, F., Sigal, A., Hardie, D., Marais, G., Huialo, N.-Y., Korsman, S., Davies, M.-A., Tyers, I., Mudau, I., York, D., Maslo, C., Goedhals, D., Abraham, S., Laguda-Akingba, O., Alizoboud-Delordri, A., Godzik, A., Wibmer, C., Sewell, B.T., Lourenço, J., Alcantara, L.C.J., Kosakovsky Pond, S.L., Weaver, S., Martin, D., Lassells, R.J., Ihihman, 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, A., Hertzig, T., Schelle, B., Bayer, S., Weilbircher, 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/cia1898.

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.00253-21.

Volz, E., Mishra, S., Chand, M., Barrett, J.C., Johnson, R., Geidelberg, L., Hinsley, W.R., Laydon, D.J., Dobrara, G., O’Toole, A., Amato, R., Ragonnet-Cronin, M., Harrison, L., Jackson, B., Ariani, C.V., Boyd, D., Loman, N.J., McClure, J.T., Gonzalez, 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., Ratman, 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, 323–331. https://doi.org/10.1002/jmv.20994.

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, L.-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, 2542–2547. 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., Muro-Maurer-Stroh, S., Le, Y.-S., 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. 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 ORF8 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. 54, 251–255. https://doi.org/10.1016/j.jmb.2020.10.045.