



OPEN ACCESS

EDITED BY

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

REVIEWED BY

Dorothee Von Laer,
Innsbruck Medical University, Austria
Jose Arturo Molina Mora,
University of Costa Rica, Costa Rica

*CORRESPONDENCE

Pedro Alves
pedro.alves@fiocruz.br

SPECIALTY SECTION

This article was submitted to
Infectious Diseases – Surveillance,
Prevention and Treatment,
a section of the journal
Frontiers in Medicine

RECEIVED 01 August 2022

ACCEPTED 02 September 2022

PUBLISHED 28 September 2022

CITATION

Silva TdS, Salvato RS, Gregianini TS,
Gomes IA, Pereira EC, de Oliveira E,
de Menezes AL, Barcellos RB,
Godinho FM, Riediger I, Debur MdC,
de Oliveira CM, Ribeiro-Rodrigues R,
Miyajima F, Dias FS, Abbud A,
Monte-Neto Rd, Calzavara-Silva CE,
Siqueira MM, Wallau GL, Resende PC,
Fernandes GdR and Alves P (2022)
Molecular characterization of a new
SARS-CoV-2 recombinant cluster XAG
identified in Brazil.
Front. Med. 9:1008600.
doi: 10.3389/fmed.2022.1008600

COPYRIGHT

© 2022 Silva, Salvato, Gregianini,
Gomes, Pereira, de Oliveira, de
Menezes, Barcellos, Godinho, Riediger,
Debur, de Oliveira, Ribeiro-Rodrigues,
Miyajima, Dias, Abbud, Monte-Neto,
Calzavara-Silva, Siqueira, Wallau,
Resende, Fernandes and Alves. This is
an open-access article distributed
under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).
The use, distribution or reproduction in
other forums is permitted, provided
the original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Molecular characterization of a new SARS-CoV-2 recombinant cluster XAG identified in Brazil

Thais de Souza Silva¹, Richard Steiner Salvato²,
Tatiana Schäffer Gregianini², Ighor Arantes Gomes³,
Elisa Cavalcante Pereira³, Eneida de Oliveira⁴,
André Luiz de Menezes⁴, Regina Bones Barcellos⁵,
Fernanda Marques Godinho⁵, Irina Riediger⁶,
Maria do Carmo Debur⁶, Cristina Mendes de Oliveira⁷,
Rodrigo Ribeiro-Rodrigues⁸, Fabio Miyajima⁹,
Fernando Stehling Dias⁹, Adriano Abbud¹⁰,
Rubens do Monte-Neto¹, Carlos Eduardo Calzavara-Silva¹,
Marilda Mendonça Siqueira³, Gabriel Luz Wallau¹¹,
Paola Cristina Resende³, Gabriel da Rocha Fernandes¹ and
Pedro Alves^{1*}

¹Instituto René Rachou, Fundação Oswaldo Cruz, Belo Horizonte, Brazil, ²Laboratório Central de Saúde Pública do Estado do Rio Grande do Sul, Porto Alegre, Brazil, ³Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil, ⁴Laboratório Municipal de Referência, Setor de Biologia Molecular, Belo Horizonte, Brazil, ⁵Centro de Desenvolvimento Científico e Tecnológico, Porto Alegre, Brazil, ⁶Laboratório Central de Saúde Pública do Estado do Paraná, Curitiba, Brazil, ⁷Dasa, Barueri, Brazil, ⁸Laboratório Central de Saúde Pública do Estado do Espírito Santo, Vitória, Brazil, ⁹Fiocruz Ceará, Fundação Oswaldo Cruz, Eusébio, Brazil, ¹⁰Instituto Adolfo Lutz, São Paulo, Brazil, ¹¹Instituto Aggeu Magalhães, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

Recombination events have been described in the Coronaviridae family. Since the beginning of the SARS-CoV-2 pandemic, a variable degree of selection pressure has acted upon the virus, generating new strains with increased fitness in terms of viral transmission and antibody scape. Most of the SC2 variants of concern (VOC) detected so far carry a combination of key amino acid changes and indels. Recombination may also reshuffle existing genetic profiles of distinct strains, potentially giving origin to recombinant strains with altered phenotypes. However, co-infection and recombination events are challenging to detect and require in-depth curation of assembled genomes and sequencing reads. Here, we present the molecular characterization of a new SARS-CoV-2 recombinant between BA.1.1 and BA.2.23 Omicron lineages identified in Brazil. We characterized four mutations that had not been previously described in any of the recombinants already identified worldwide and described the likely breaking points. Moreover, through phylogenetic analysis, we showed that the newly named XAG lineage groups in a highly supported monophyletic clade confirmed its common evolutionary history from parental Omicron lineages and other recombinants already

described. These observations were only possible thanks to the joint effort of bioinformatics tools auxiliary in genomic surveillance and the manual curation of experienced personnel, demonstrating the importance of genetic, and bioinformatic knowledge in genomics.

KEYWORDS

SARS-CoV-2, recombinant, genomic surveillance, XAG, variants, omicron

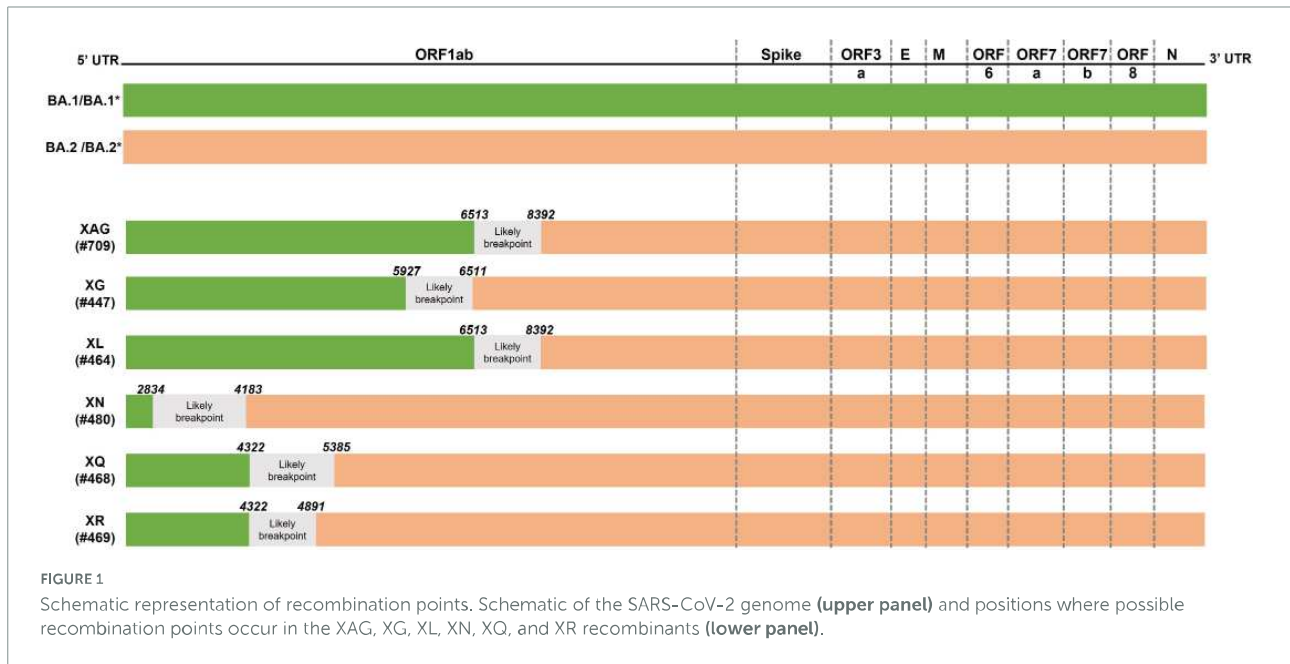
Introduction

Since the first detection of SARS-CoV-2 in Wuhan, China (1), many new viral strains/lineages emerged, carrying some differences from the initial virus. Over time, the evolution and fixation of mutations, especially in variants of concern (VOC), demonstrates that new strains have a high chance of continuing to emerge due to the changing selective pressure on the virus lineages associated with extensive transmission worldwide (2). Epidemiological surveillance is essential for the control of COVID-19. However, genomic surveillance is just as important (3). Genome analyzes can provide additional information for epidemiological surveillance, demonstrating outbreak dynamics in space and time, characterizing transmission, and allowing the identification of mutations that can lead to the emergence of new variants with the potential to impact public health and the epidemiology of COVID-19 (4). The effectiveness of molecular surveillance as a tool for monitoring pandemics is dependent on continuous and consistent sampling through time and space, rapid virus genome sequencing, and rapid reporting (5). Enhancing genomic surveillance and sequencing efforts across the globe is a valuable tool to detect and understand emerging variants (6), and genomics-based SARS-CoV-2 surveillance is a helpful tool for monitoring the current and future phases of the COVID-19 pandemic (7).

Since the beginning of the COVID-19 pandemic, five lineages have been considered VOCs by the WHO: Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (B.1.617.2), and Omicron (B.1.1.529). The Omicron variant, initially identified in South Africa and Botswana on November 21, 2021, was considered by the WHO to be a VOC on November 26, 2021 (8). Shortly after its identification, Omicron showed great potential for dissemination, with a significant increase over the Delta variant, which, since its identification, was the variant with the highest frequency worldwide. Omicron has been classified into five sublineages: BA.1/BA.1.*, BA.2/BA.2.*, BA.3, BA.4, and BA.5. BA.2 is replacing BA.1 as the dominant subvariant in more countries over time (9). To date (June 2022), the circulation of the BA.2/BA.2.* sublineage currently represents 61.5% of the genomes sequenced in Brazil (10). Omicron has potential

for dissemination and should be closely monitored due to the high number of mutations present in the genome (at least 32) that can increase infectivity and immune escape compared with the early wild-type lineage and the other VOCs (11–13).

With the circulation of different variants in the same place and at the same time, co-infections become possible, potentially leading to the emergence, and rise of new variants through viral recombination. Unfortunately, the number of co-infections is challenging to determine, mainly because genomic surveillance is suboptimal in most countries (14). Some studies in different countries have found values ranging from 0.06 to 4.0% of co-infection and seem to be underestimated (15–20). Some tools based on metagenomics and bioinformatics have been proposed to identify and evaluate co-infections (21–25). It is unclear whether co-infections can result in more severe disease. However, the dominance of one strain over another in a co-infection has already been observed in the same patient, which can be explained by the higher virulence of the dominant strain (22). Genome recombination is an important evolutionary mechanism for the emergence and re-emergence of human pathogens and a significant source of viral evolution (9). SARS-CoV-2 originated from recombination may have advantages for viral dispersion, immune evasion potential, and decrease in vaccine effectiveness, but little is known about it and, consequently, it highlights the importance of studying the recombinants (26). Coronaviruses (CoVs) are highly recombinogenic, unlike other viruses that have emerged in the past two decades (27). Recombination occurs when genetic material from two circulating lineages is combined within the host, giving rise to a viable descent lineage (28). While some SARS-CoV-2 lineages disappear, others can become dominant through the fixation of key mutations in the genome that allow improved adaptation of these viruses regarding transmissibility and faster dispersion (6). The first recombinant to be identified, named XA, was detected in the UK in samples collected between Oct 2020 and Jan 2021 and resulted from recombination involving the Alpha variant (29). An increase in the detection of recombination events occurred between Delta and Omicron variants (20, 23, 30, 31) that co-circulated between November 2021 and February 2022 (10)



and between Omicron variant sublineages such as BA.1 and BA.2 (32).

The gold standard technique for identifying and classifying SARS-CoV-2 lineages is based primarily on partial or whole genome sequencing through New Generation Sequencing (NGS) (6). Reverse transcription Polymerase chain reaction (RT-PCR) assays have been used to identify specific characteristics that are unique in specific variants, such as S-gene target failure (SGTF) in the Alpha and Omicron variant, and genotyping assays have also been applied for this purpose (33). Sanger sequencing is also a methodology that allows monitoring SARS-CoV-2 variants with a rapid response (34–36). Recombination can be challenging to detect by classification methods, as the recombinant sequences have high similarity to their shared ancestor (28, 32). Bioinformatics tools are used to identify and classify variants from the results obtained by NGS. However, these tools must be used with caution for potential new recombinant lineage classification once more in-depth human intervention is usually required for correct recombinant identification. Otherwise, recombinant lineages may be unreported until a high prevalence is reached (37, 38).

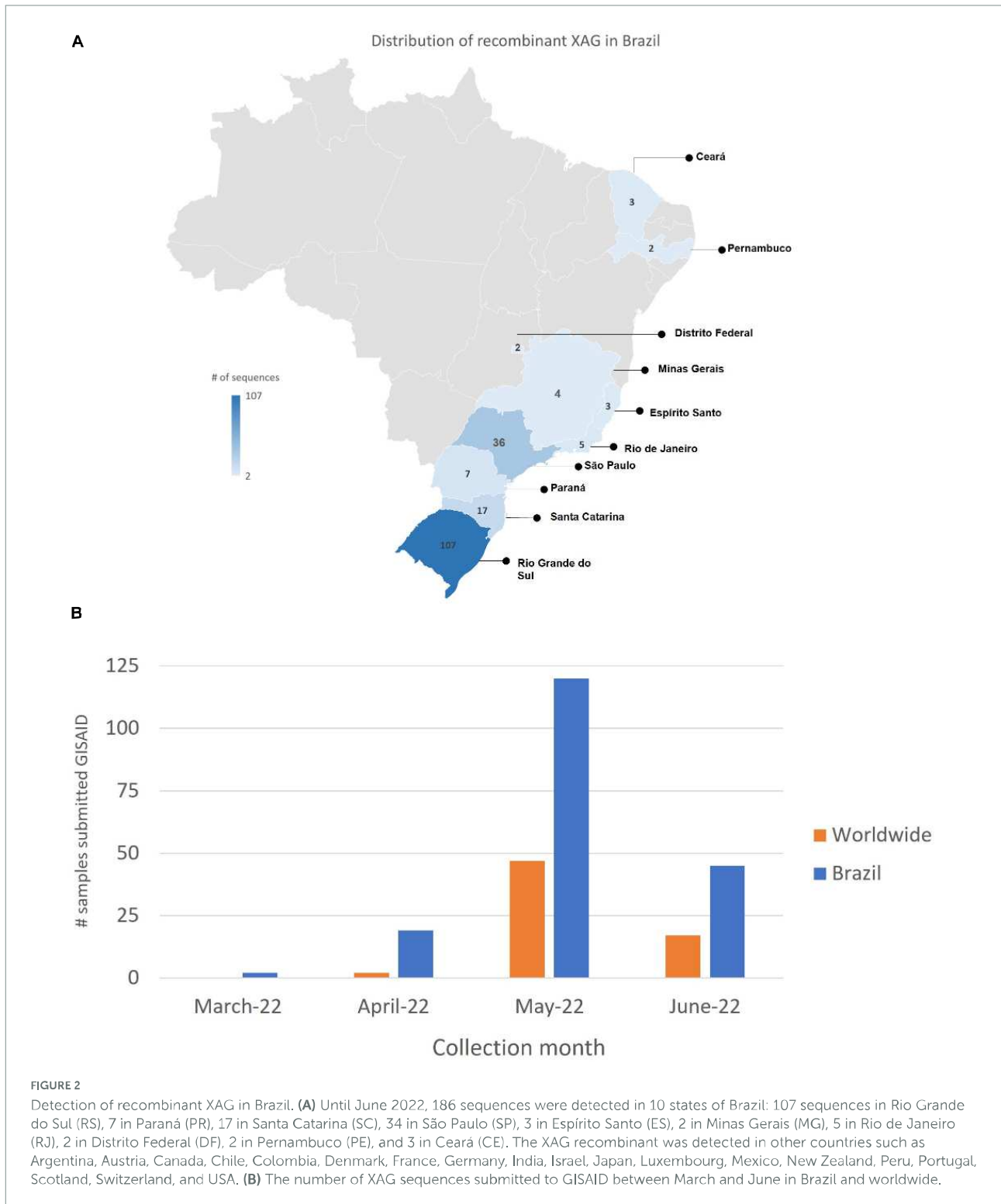
The aim of this study was the molecular characterization of a new recombinant lineage from samples collected in Brazil between April and May 2022. We demonstrate that this lineage was not initially classified by the available tools, such as Pangolin (v4.0.6 at that time). We performed several complementary analyzes and showed that this new recombinant, now named XAG, is the result of the recombination between two sublineages of the Omicron VOC, BA.1.1 and BA.2.23.

Materials and methods

Sequencing, variant calling, characterization, and phylogeny

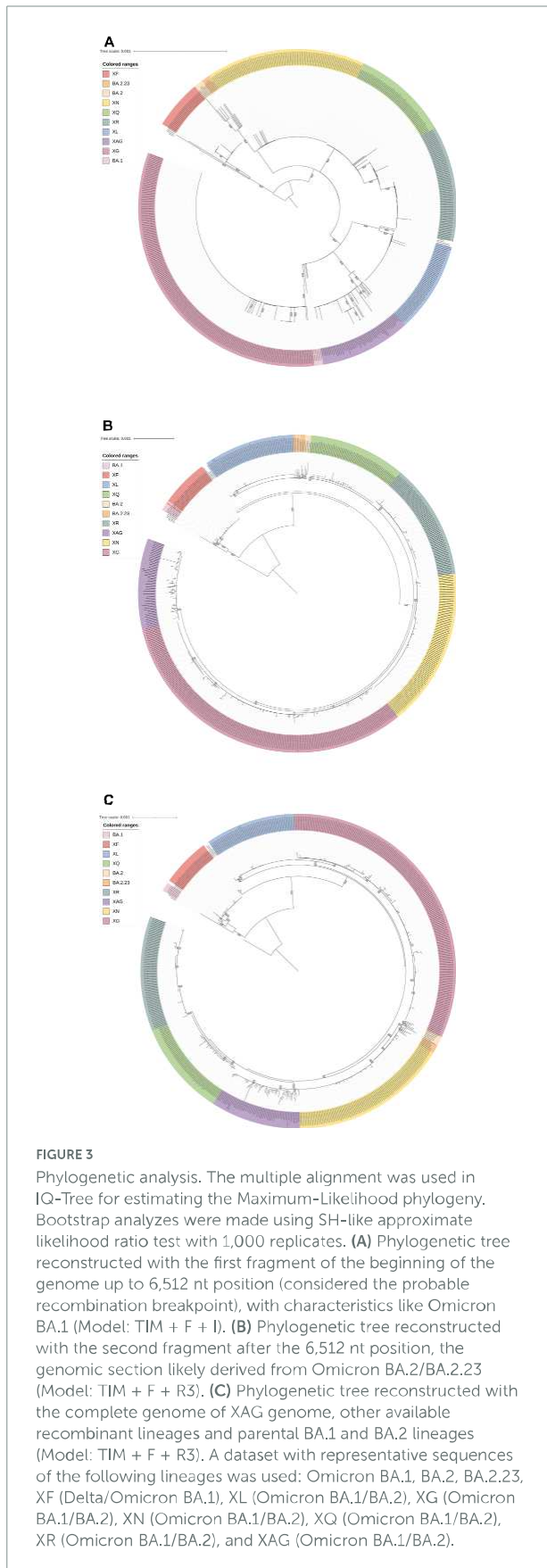
The whole genome sequencing of the samples was performed using the COVIDSeq Illumina test protocol adapted by the Fiocruz Genomic Network.¹ The assembly and variant calling was done through the ViralFlow (39) workflow that performs the assembly of the genome according to the reference sequence and additional analysis. The molecular characterization was performed by aligning the sequences with a dataset of sequences (Supplementary Data Sheet 1) from previously detected recombinant lineages submitted to the EpiCoV database at GISAID (40) using MAFFT (41) and visualization of the mutations profiles found through the AliView program (42). The multiple alignments were used in IQ-Tree (43) for estimating the Maximum-Likelihood phylogeny. Three different phylogenetic trees were reconstructed: I—Using the complete genome of the XAG recombinant, other available recombinant lineages, and parental BA.1 and BA.2 lineages; II—Using two fragments of the XAG recombinant genome splitted at the likely breaking point. The first fragment consists of the beginning of the genome up to 6,515 nt position (considered the probable recombination breakpoint), with characteristics like Omicron BA.1.* and the second fragment after this position, the genomic section likely

¹ <https://www.genomahcov.fiocruz.br/>



derived from Omicron BA.2*. A dataset with representative sequences of the following lineages was used: Omicron BA.1, BA.2, BA.2.23, XF (Delta/Omicron BA.1), XL (Omicron BA.1/BA.2), XG (Omicron BA.1/BA.2), XN (Omicron

BA.1/BA.2), XQ (Omicron BA.1/BA.2), XR (Omicron BA.1/BA.2), and XAG (Omicron BA.1/BA.2). This study was reviewed and approved by Research Ethics Committee involving human beings at Instituto René Rachou, Fundação



Oswaldo Cruz, under license protocol number: 4,084,902 and CAAE (certificate of presentation for ethical appreciation): 31984720300005091.

Results

Identification and classification of sequences collected in Brazil with initial variant identification failure

In May 2022, the first sequence (EPI_ISL_13019803) that failed to be identified by Pangolin was deposited and submitted to GISAID, identified in the Rio Grande do Sul, Brazil. The genetic profile present in this sample was detected in other submitted sequences collected later in Brazil. Through analysis of the mutations present in the sequences, a specific genetic signature of both the BA.1 and BA.2 lineages of the Omicron variant was identified, raising the hypothesis of recombination between these two lineages. In the genome analysis, a possible recombination point was detected between positions 6,512 and 8,395, at the same likely breaking point described for the XL recombinant. Up to position 6,512, the samples present a genetic signature characteristic of the BA.1.1 variant (including the deletion at position 6,512 found only in BA.1), while after position 8,392, the genetic signature resembles that of the BA.2.23 lineage (**Figure 1** and **Supplementary Table 1**).

A new recombinant cluster was detected through molecular and phylogenetic analyses, later named XAG.² The first sequence of the XAG variant was collected on March 10, 2022, in the city of Caxias do Sul, Rio Grande do Sul/Brazil, from a 17-year-old male patient. Currently, 186 sequences were detected in 10 Brazilian states (**Figure 2**) and have already been detected in other countries such as Argentina, Austria, Canada, Chile, Colombia, Denmark, France, Germany, India, Israel, Japan, Luxembourg, Mexico, New Zealand, Peru, Portugal, Scotland, Switzerland, and the USA. On July 7, 2022, 252 sequences belonging to the recombinant XAG cluster were deposited in GISAID (**Supplementary Data Sheet 1**).

Phylogenetic analyzes and molecular characterization

In the phylogenetic analyses, it is possible to observe differences in the distances between BA.1 and BA.2 clusters when using the different genome portions. The proximity of the BA.1 branches to the BA.1-corresponding portion of the XAG genome is observed (**Figure 3A**), as well as the greater proximity of the BA.2/BA.2.23 branches to the portion corresponding to

² <https://github.com/cov-lineages/pango-designation/issues/709>

BA.2 of the XAG genome (Figure 3B). In the complete genome analysis of XAG, it is possible to verify that it is a new cluster, both regarding the recombinants of Omicron BA.1/BA.2 and XF (Delta/Omicron BA.1) (Figure 3C). Nevertheless, in less detailed analyses, it is possible to confuse the XAG recombinants with other Omicron BA.1/BA.2 recombinants such as XG and XQ.

The XAG recombinant has four distinct markers: C2857T (synonymous), C5585A (L1774I), A12334G, and C17502T (synonymous), all present in the region of the ORF1ab gene. Up to position 6512, XAG has features of the BA.1 sublineage and, after that, shows a genetic signature like the BA.2 sublineage (Figure 4).

Discussion

In nature, mutations, recombination, and reassortment are critical evolutionary processes that generate genetic diversity (44). The CoVs, like most other viruses, have developed a variety of genetic mechanisms, among which recombination and generation of defective-interfering (DI) RNA that, as a side effect, generate diversity (45, 46). SARS-CoV-2, despite being genetically distinct from the viruses that cause SARS epidemic in 2003 and Middle East respiratory syndrome coronavirus (MERS-CoV), shows high levels of genetic similarity with a strain from bats and strains obtained from pangolins (47–49), demonstrating that recombination can help to develop transmission strategies between species by establishing more susceptible hosts (50). Frequent genomes recombination, large genetic diversity, and high human-animal interface enable CoVs to emerge from time to time in humans due to occasional spillover and recurrent cross-species infectious events (44). Recombination and reassortment are essential processes that allow new antigenic combinations and altered phenotypes in emerging viruses that might aid the course of cross-species diffusion (44). RNA recombination is required during normal replication. The mechanisms and determinants of CoVs recombination are not known (51). Recombination of MERS-CoVs was already described in camels, leading to human outbreaks in 2015 (52). The high number of accumulated mutations in the Omicron variant may be due to recombination events. However, there is still no scientific evidence to support this hypothesis (53).

It was confirmed that co-infections by Omicron and Delta variants have already occurred in specific populations (9). Currently, the Omicron variant has the largest circulation in Brazil (approximately 99% of the genomes sequenced in June 2022), represented by 0.7% of the BA.1.* sublineage, 64% of the BA.2.*, 14.7% of the BA.4 and 20.5% of BA.5.* (10). This variant overlapped the Delta variant, but they co-circulated between November 2021 and January

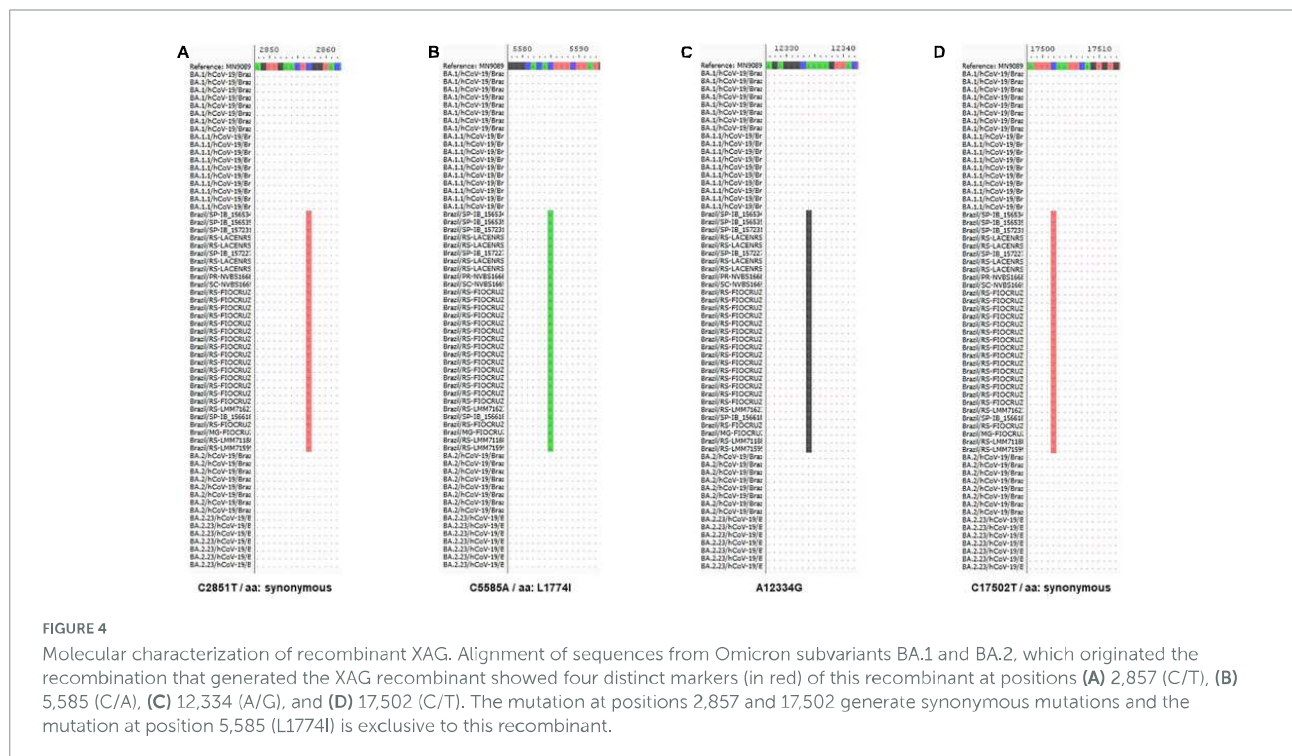
2022, which contributed to co-infections and, consequently, allowed recombination between these two strains (9, 23). The first recombination identified was named XA and is the result of a recombination of the Alpha variant.³ After that, recombinant strains B.1.634 and B.1.631 were detected (XB), and so far, 32 recombinants have been identified in several countries,⁴ most of them occurred between Omicron variant sublineages recombination.

The XAG recombinant described in this work has four unique mutations. Two mutations, C2857T and C17502, are synonymous. Synonymous nucleotide changes may be related to virus adaptation and more efficient use of host tRNA profile, but it also may impact virus genome hairpins and 3D RNA structure (54). In addition, the mutation at position 5,585 (L17714I) is unique to the XAG recombinant and had already been observed in circulating strains previously, but at a low frequency, disappearing over time (55). SARS-CoV-2 and other CoVs have moderate genetic variability because they have an RNA-dependent RNA-polymerase (RdRP) with correction activity during viral genome replication and transcription (56). Even with this mechanism, errors can occur and become fixed if they present adaptive advantages. Host-related factors can also induce mutations, such as the antiviral mechanism mediated by APOBEC (Apolipoprotein B mRNA Editing Catalytic Polypeptide-like) proteins. APOBEC-like directional C→U transitions of genomic plus-strand RNA are overrepresented in SARS-CoV-2 genome sequences of variants emerging during the COVID-19 pandemic (57) and may affect the identification of co-infection events (15). De Maio et al. demonstrated that two mutation rates, C→U and G→U, are similar and much higher than all other mutation rates, leading to extremely frequent homoplasies (58). Sequence changes in SARS-CoV-2 and other coronavirus genomes may be partially or restricted to several mutational hot spots that promote convergent changes between otherwise genetically unlinked strains (59).

Recombination can be challenging to detect, mainly because they have similar characteristics to other circulating lineages at a higher frequency (9, 32). In addition to monitoring circulating lineages, the purpose of genomic surveillance is the rapid identification of new emerging lineages (5). Genomic surveillance also plays a significant role in studies on developing prophylactic measures and vaccines. Monitoring SARS-CoV-2 genetic changes, especially at the epitopes with implications for immune escape, is crucial (60). Since the first identification of a recombinant in Rio Grande do Sul, surveillance has been intensified, providing identification of more XAG recombinant sequences (see text footnote 1). This recombinant community transmission was observed in several neighboring and distant Brazilian states. This fact demonstrates the importance of

³ <https://virological.org/t/recombinant-sars-cov-2-genomes-involving-lineage-b-1-1-7-in-the-uk/658>

⁴ <https://github.com/cov-lineages/pango-designation>



genomic surveillance associated with an epidemiological link to validate genetic findings further.

As SARS-CoV-2 circulates worldwide, new lineages emerge and are tracked using the Pango dynamic hierarchical nomenclature system (28). Bioinformatics tools are essential and greatly help genomic surveillance, especially in pandemic scenarios, where surveillance needs to be assertive and fast (37). However, these tools have some limitations. In our work, it was possible to identify a flaw in the classification of the SARS-CoV-2 lineage by Pangolin since our sequences have mutations shared with the ancestral lineages and, in addition, they have synonymous mutations, which could only be detected through manual analyzes of all genome positions. To minimize this effect, Pangolin developers regularly train the tool with the latest designated sequencing (38). Despite this, viruses, especially RNA viruses, have high mutation rates that lead to an eminent environmental adaptation with rapid evolution, contrasting with the identification tools update time (44). We strongly suggest that classification tools are used allied with manual curation, especially in cases such as the one described in this study, since minority mutations may go unnoticed and help to tease apart co-infection and recombination events of epidemiological importance. The main limitation of our study is the focus only on the agent, based on the epidemiological triad (Agent, Host, and Environment), which demonstrates the need for further studies on recombinants (61, 62). We could not perform viral isolation, and it was also not possible to collect serial samples from each patient to assess the impact

of XAG infection. The impact must be evaluated considering factors such as vaccination, social distancing measures, and recombination events that can occur in animals and viruses jumping back to humans (26, 61).

Conclusion

In conclusion, it was possible to identify the emergence of a new SARS-CoV-2 recombinant, a result of the recombination between two sublineages of the VOC Omicron, applying bioinformatics tools for the identification of variants together with manual analyzes for the characterization of unique unlabeled mutations present in the new cluster, called XAG. In addition, new markers were identified in the XAG recombinant that had not been found in other previously described recombinants, demonstrating the potential for these viruses to evolve through recombination. This study demonstrates the need for continuous genomic surveillance of SARS-CoV-2, in which recombination appears essential in its evolution. The real impact of the recombinants needs to be further studied, considering the possibility of the occurrence of these events in animal-human interfaces and the emergence of new lineages.

Data availability statement

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

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

Ethics statement

The studies involving human participants were reviewed and approved by Research Ethics Committee involving human beings at Instituto René Rachou, Fundação Oswaldo Cruz, under license protocol number: 4084902 and CAAE (certificate of presentation for ethical appreciation): 31984720300005091. The ethics approval was issued on June 12, 2020. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

TS, PA, and GF: conceptualization, experimental design, and, writing—original draft. TS, PA, GF, RS, TG, IG, EP, RB, FG, IR, MD, CO, RR-R, FD, and AA: investigation and perform experiments. TS, PA, GF, GW, RS, TG, IG, EP, CO, FM, FD, AA, and PR: data analyses. RM-N, GF, PA, RS, TG, IG, EP, MS, and PR: reagents, materials, analysis, and tools. MS, GF, and PA: supervision. TS, PA, GF, RM-N, CC-S, RS, TG, IG, EP, MS, PR, EO, AM, RB, FG, IR, MD, CO, RR-R, FM, FD, AA, and GW: writing—review and editing. All authors discussed the results and contributed to the final manuscript.

Funding

This work was funded by the Departamento de Ciência e Tecnologia/Secretaria de Ciência, Tecnologia, Inovação e Insumos Estratégicos em Saúde/Ministério da Saúde (Grant nos. 402457/2020-9 and 403276/2020-9) and Innovative Products to face COVID-19 pandemics (Grant no. VPPIS-005-FIO-20-2-45) to PA. TS received a Doctoral's fellowship from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, CAPES. RM-N was CNPq Research Fellow (Grant no. #310640/2017-2).

References

- Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with Pneumonia in China, 2019. *N Engl J Med.* (2020) 382:727–33.
- Campbell F, Archer B, Laurenson-Schafer H, Jinnai Y, Konings F, Batra N, et al. Increased transmissibility and global spread of SARS-CoV-2 variants of concern as at June 2021. *Euro Surveill.* (2021) 26:2100509.
- Robishaw JD, Alter SM, Solano JJ, Shih RD, DeMets DL, Maki DG, et al. Genomic surveillance to combat COVID-19: challenges and opportunities. *Lancet Microbe.* (2021) 2:e481–4.
- World Health Organization. *Genomic Sequencing of SARS-CoV-2: A Guide to Implementation for Maximum Impact on Public Health.* Geneva: WHO (2021).

Acknowledgments

We thank all technicians from Laboratório Municipal de Referência, Setor de Biologia Molecular, Belo Horizonte, Brazil, for all support in sample selection. We thank Anna Christina de M. Salim, Flávio M. G. de Araújo, and Caroline P. Rocha from Plataforma de Sequenciamento NGS (P01-007) from Instituto René Rachou, Fundação Oswaldo Cruz, Belo Horizonte, Brazil by NGS. We thank all researchers, students, and technicians in the Imunologia de Doenças Virais (IDV) group for their help and expertise. TS, RM-N, GF, and PA were part of Fiocruz COVID-19 Genomic Surveillance Network (<http://www.genomahcov.fiocruz.br>) which we are also extremely grateful for the possibility of doing the NGS sequencing. We thank all authors who have kindly deposited and shared genome data on GISAID.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmed.2022.1008600/full#supplementary-material>

5. Wilkinson E, Giovanetti M, Tegally H, San JE, Lessells R, Cuadros D, et al. A year of genomic surveillance reveals how the SARS-CoV-2 pandemic unfolded in Africa. *Science*. (2021) 374:423–31. doi: 10.1126/science.abc4336
6. Chen Z, Azman AS, Chen X, Zou J, Tian Y, Sun R, et al. Global landscape of SARS-CoV-2 genomic surveillance and data sharing. *Nat Genet*. (2022) 54:499–507.
7. Goswami C, Sheldon M, Bixby C, Keddache M, Bogdanowicz A, Wang Y, et al. Identification of SARS-CoV-2 variants using viral sequencing for the Centers for Disease Control and Prevention genomic surveillance program. *BMC Infect Dis*. (2022) 22:404. doi: 10.1186/s12879-022-07374-7
8. World Health Organization. *Classification of Omicron (B.1.1.529): SARS-CoV-2 Variant of Concern*. Geneva: WHO (2021).
9. Ou J, Lan W, Wu X, Zhao T, Duan B, Yang P, et al. Tracking SARS-CoV-2 Omicron diverse spike gene mutations identifies multiple inter-variant recombination events. *Signal Transduct Target Ther*. (2022) 7:138. doi: 10.1038/s41392-022-00992-2
10. Rede Genômica Fiocruz. *Dashboard Rede Genômica*. (2021). Available online at: <https://www.genomahcov.fiocruz.br/dashboard-pt/> (accessed November 5, 2021).
11. Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, et al. SARS-CoV-2 variants, spike mutations and immune escape. *Nat Rev Microbiol*. (2021) 19:409–24.
12. Khan A, Zia T, Suleman M, Khan T, Ali SS, Abbasi AA, et al. Higher infectivity of the SARS-CoV-2 new variants is associated with K417N/T, E484K, and N501Y mutants: an insight from structural data. *J Cell Physiol*. (2021) 236:7045–57. doi: 10.1002/jcp.30367
13. Tian D, Sun Y, Xu H, Ye Q. The emergence and epidemic characteristics of the highly mutated SARS-CoV-2 Omicron variant. *J Med Virol*. (2022) 94:2376–83.
14. de Oliveira CM, Romano CM, Sussuchi L, Cota BDCV, Levi JE. SARS-CoV-2 BA.1 and BA.2 coinfection detected by genomic surveillance in Brazil, January 2022. *Arch Virol*. (2022) [Epub ahead of print]. doi: 10.1007/s00705-022-05532-5
15. Dezordi FZ, Resende PC, Naveca FG, do Nascimento VA, de Souza VC, Paixão ACD, et al. Unusual SARS-CoV-2 intrahost diversity reveals lineage superinfection. *Microb Genom*. (2022) 8:000751. doi: 10.1099/mgen.0.000751
16. Zhou HY, Cheng YX, Xu L, Li JY, Tao CY, Ji CY, et al. Genomic evidence for divergent co-infections of SARS-CoV-2 lineages. *bioRxiv* [Preprint]. (2021) doi: 10.1101/2021.09.03.458951
17. Tonkin-Hill G, Martincorena I, Amato R, Lawson AR, Gerstung M, Johnston I, et al. Patterns of within-host genetic diversity in SARS-COV-2. *eLife*. (2021) 10:e66857.
18. Lythgoe KA, Hall M, Ferretti L, de Cesare M, MacIntyre-Cockett G, Trebes A, et al. SARS-CoV-2 within-host diversity and transmission. *Science*. (2021) 372:eabg0821.
19. Bal A, Simon B, Destras G, Chalvignac R, Semanas Q, Oblette A, et al. Detection and prevalence of SARS-CoV-2 co-infections during the Omicron variant circulation, France, December 2021 – February 2022. *medRxiv* [Preprint]. (2022) doi: 10.1101/2022.03.24.22272871
20. Bolze A, Basler T, White S, Rossi AD, Wyman D, Roychoudhury P, et al. Evidence for SARS-CoV-2 Delta and Omicron co-infections and recombination. *medRxiv* [Preprint]. (2022) doi: 10.1101/2022.03.09.22272113
21. Molina-Mora JA, Cordero-Laurent E, Calderón-Osorno M, Chacón-Ramírez E, Duarte-Martínez F. Metagenomic pipeline for identifying co-infections among distinct SARS-CoV-2 variants of concern: study cases from Alpha to Omicron. *Sci Rep*. (2022) 12:9377. doi: 10.1038/s41598-022-13113-4
22. Samoilov AE, Kaptelova VV, Bukharina AY, Shipulina OY, Korneenko EV, Saenko SS, et al. Case report: change of dominant strain during dual SARS-CoV-2 infection. *BMC Infect Dis*. (2021) 21:959. doi: 10.1186/s12879-021-06664-w
23. Rockett RJ, Draper J, Gall M, Sim EM, Arnott A, Agius JE, et al. Co-infection with SARS-CoV-2 Omicron and Delta variants revealed by genomic surveillance. *Nat Commun*. (2022) 13:2745.
24. van Tan L, Thi Thu Hong N, My Ngoc N, Tan Thanh T, Thanh Lam V, Anh Nguyet L, et al. SARS-CoV-2 and co-infections detection in nasopharyngeal throat swabs of COVID-19 patients by metagenomics. *J Infect*. (2020) 81:e175–7.
25. Li Y, Jiang Y, Li Z, Yu Y, Chen J, Jia W, et al. Both simulation and sequencing data reveal coinfections with multiple SARS-CoV-2 variants in the COVID-19 pandemic. *Comput Struct Biotechnol J*. (2022) 20:1389–401. doi: 10.1016/j.csbj.2022.03.011
26. Mohapatra RK, Kandi V, Tuli HS, Chakraborty C, Dhama K. The recombinant variants of SARS-CoV-2: concerns continues amid COVID-19 pandemic. *J Med Virol*. (2022) 94:3506–8. doi: 10.1002/jmv.27780
27. Boni MF, Lemey P, Jiang X, Lam TTY, Perry BW, Castoe TA, et al. Evolutionary origins of the SARS-CoV-2 sarbecovirus lineage responsible for the COVID-19 pandemic. *Nat Microbiol*. (2020) 5:1408–17. doi: 10.1038/s41564-020-0771-4
28. Gutierrez B, Castelan Sánchez HG, da Silva Candido D, Jackson B, Fleishon S, Houzet R, et al. Emergence and widespread circulation of a recombinant SARS-CoV-2 lineage in North America. *Cell Host Microbe*. (2022) 30: 1112–23.e3. doi: 10.1016/j.chom.2022.06.010
29. Jackson B, Boni MF, Bull MJ, Collier A, Colquhoun RM, Darby AC, et al. Generation and transmission of interlineage recombinants in the SARS-CoV-2 pandemic. *Cell*. (2021) 184:5179–88.e8. doi: 10.1016/j.cell.2021.08.014
30. Lacek KA, Rambo-Martin BL, Batra D, Zheng XY, Sakaguchi H, Peacock T, et al. Identification of a novel SARS-CoV-2 delta-omicron recombinant virus in the United States. *bioRxiv* [Preprint]. (2022) doi: 10.1101/2022.03.19.484981
31. Duerr R, Dimartino D, Marier C, Zappile P, Wang G, Plitnick J, et al. Delta-Omicron recombinant SARS-CoV-2 in a transplant patient treated with Sotrovimab. *bioRxiv* [Preprint]. (2022) doi: 10.1101/2022.04.06.487325
32. Focosi D, Maggi F. Recombination in coronaviruses, with a Focus on SARS-CoV-2. *Viruses*. (2022) 14:1239.
33. Vogels CBF, Breban MI, Ott IM, Alpert T, Petrone ME, Watkins AE, et al. Multiplex qPCR discriminates variants of concern to enhance global surveillance of SARS-CoV-2. *PLoS Biol*. (2021) 19:e3001236. doi: 10.1371/journal.pbio.3001236
34. Bezerra MF, Machado LC, de Carvalho V, do CV, Docena C, Brandão-Filho SP, et al. A Sanger-based approach for scaling up screening of SARS-CoV-2 variants of interest and concern. *Infect Genet Evol*. (2021) 92:104910. doi: 10.1016/j.meegid.2021.104910
35. Lim HJ, Park MY, Jung HS, Kwon Y, Kim I, Kim DK, et al. Development of an efficient Sanger sequencing-based assay for detecting SARS-CoV-2 spike mutations. *PLoS One*. (2021) 16:e0260850. doi: 10.1371/journal.pone.0260850
36. Dorlass EG, Lourenço KL, Magalhães RDM, Sato H, Fiorini A, Peixoto R, et al. Survey of SARS-CoV-2 genetic diversity in two major Brazilian cities using a fast and affordable Sanger sequencing strategy. *Genomics*. (2021) 113:4109–15. doi: 10.1016/j.ygeno.2021.10.015
37. Hu T, Li J, Zhou H, Li C, Holmes EC, Shi W. Bioinformatics resources for SARS-CoV-2 discovery and surveillance. *Brief Bioinform*. (2021) 22:631–41.
38. O'Toole Á, Scher E, Underwood A, Jackson B, Hill V, McCrone JT, et al. Assignment of epidemiological lineages in an emerging pandemic using the pangolin tool. *Virus Evol*. (2021) 7:veab064. doi: 10.1093/ve/veab064
39. Dezordi FZ, Neto AM, da S, Campos T, de L, Jeronimo PMC, et al. ViralFlow: a versatile automated workflow for SARS-CoV-2 genome assembly, lineage assignment, mutations and intrahost variant detection. *Viruses*. (2022) 14:217. doi: 10.3390/v14020217
40. Khare S, Gurry C, Freitas LB, Schultz M, Bach G, Diallo A, et al. GISAID's role in pandemic response. *China CDC Wkly*. (2021) 3:1049–51.
41. Katoh K, Standley DM. MAFFT, multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. (2013) 30:772–80. doi: 10.1093/molbev/mst010
42. Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics*. (2014) 30:3276–8. doi: 10.1093/bioinformatics/btu531
43. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. (2015) 32:268–74. doi: 10.1093/molbev/msu300
44. Rehman SU, Shafique L, Ihsan A, Liu Q. Evolutionary trajectory for the emergence of novel coronavirus SARS-CoV-2. *Pathogens*. (2020) 9:240.
45. Lai MMC, Cavanagh D. The molecular biology of coronaviruses. *Adv Virus Res*. (1997) 48:1–100.
46. Pérez-Losada M, Arenas M, Galán JC, Palero F, González-Candelas F. Recombination in viruses: mechanisms, methods of study, and evolutionary consequences. *Infect Genet Evol*. (2015) 30:296–307.
47. Liu P, Chen W, Chen JP. Viral metagenomics revealed sendai virus and coronavirus infection of malayan pangolins (*Manis javanica*). *Viruses*. (2019) 11:979. doi: 10.3390/v11110979
48. Lam TTY, Jia N, Zhang YW, Shum MHH, Jiang JF, Zhu HC, et al. Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. *Nature*. (2020) 583: 282–5.
49. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A Pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. (2020) 579:270–3.
50. Li X, Giorgi EE, Marichannegowda MH, Foley B, Xiao C, Kong XP, et al. Emergence of SARS-CoV-2 through recombination and strong purifying selection. *Sci Adv*. (2020) 6:eabb9153.

51. Gribble J, Stevens LJ, Agostini ML, Anderson-Daniels J, Chappell JD, Lu X, et al. The coronavirus proofreading exoribonuclease mediates extensive viral recombination. *PLoS Pathog.* (2021) 17:e1009226. doi: 10.1371/journal.ppat.1009226
52. Sabir JSM, Lam TTY, Ahmed MMM, Li L, Shen YEM, Abo-Aba S, et al. Co-circulation of three camel coronavirus species and recombination of MERS-CoVs in Saudi Arabia. *Science.* (2016) 351:81–4. doi: 10.1126/science.aac8608
53. Callaway E. Heavily mutated Omicron variant puts scientists on alert. *Nature.* (2021) 600:21. doi: 10.1038/d41586-021-03552-w
54. Kim SJ, Nguyen VG, Park YH, Park BK, Chung HCA. Novel synonymous mutation of SARS-CoV-2: Is this possible to affect their antigenicity and immunogenicity? *Vaccines.* (2020) 8:220. doi: 10.3390/vaccines8020220
55. Mullen JL, Tsueng G, Latif AA, Alkuzweny M, Cano M, Haag E, et al. *Outbreak.Info.* (2020). Available online at: <https://outbreak.info/> (accessed November 5, 2021).
56. Kim K, Calabrese P, Wang S, Qin C, Rao Y, Feng P, et al. The roles of APOBEC-mediated RNA editing in SARS-CoV-2 mutations, replication and fitness. *bioRxiv* [Preprint]. (2022) doi: 10.1101/2021.12.18.473309
57. Ratcliff J, Simmonds P. Potential APOBEC-mediated RNA editing of the genomes of SARS-CoV-2 and other coronaviruses and its impact on their longer term evolution. *Virology.* (2021) 556:62–72. doi: 10.1016/j.virol.2020.12.018
58. De Maio N, Walker CR, Turakhia Y, Lanfear R, Corbett-Detig R, Goldman N. Mutation rates and selection on synonymous mutations in SARS-CoV-2. *Genome Biol Evol.* (2021) 13:evab087.
59. Simmonds P. Rampant C→U Hypermutation in the genomes of SARS-CoV-2 and other coronaviruses: causes and consequences for their short- and long-term evolutionary trajectories. *mSphere.* (2020) 5:e408–20. doi: 10.1128/mSphere.00408-20
60. Bugembe DL, Phan MVT, Ssewanyana I, Semanda P, Nansumba H, Dhaala B, et al. Emergence and spread of a SARS-CoV-2 lineage A variant (A.23.1) with altered spike protein in Uganda. *Nat Microbiol.* (2021) 6:1094–101. doi: 10.1038/s41564-021-00933-9
61. Tsui BCH, Deng A, Pan S. COVID-19: epidemiological factors during aerosol-generating medical procedures. *Anesth Analg.* (2020) 131:e175–8.
62. Molina-Mora JA, González A, Jiménez-Morgan S, Cordero-Laurent E, Brenes H, Soto-Garita C, et al. Clinical profiles at the time of diagnosis of SARS-CoV-2 infection in costa rica during the pre-vaccination period using a machine learning approach. *Phenomics* (2022) [Epub ahead of print]. doi: 10.1007/s43657-022-00058-x