

Haplotypic Distribution of SARS-CoV-2 Variants in Cases of Intradomiciliary Infection in the State of Rondônia, Western Amazon

Bioinformatics and Biology Insights
Volume 18: 1–8
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/11779322241266354



Karolaine Santos Teixeira^{1,2,3}, Márlon Grégori Flores Custódio¹,
Gabiella Sgorlon^{1,2,4,5}, Tárco Peixoto Roca^{1,2,6},
Jackson Alves da Silva Queiroz^{1,2,4,5}, Ana Maisa Passos-Silva^{1,2,4},
Jessiane Rodrigues Ribeiro^{1,2} and Deusilene Vieira^{1,2,4,5}

¹Laboratório de Virologia Molecular, Fundação Oswaldo Cruz Rondônia—FIOCRUZ/RO, Porto Velho, Brazil. ²Instituto Nacional de Epidemiologia da Amazônia Ocidental—INCT EpiAmO, Porto Velho, Brazil. ³Programa Institucional de Bolsas de Iniciação em Desenvolvimento Tecnológico e Inovação—PIBITI/FIOCRUZ, Rio de Janeiro, Brazil. ⁴Programa de Pós-Graduação em Biologia Experimental, Universidade Federal de Rondônia – UNIR/FIOCRUZ/RO, Porto Velho, Brazil. ⁵Rede Genômica FIOCRUZ, Rio de Janeiro, Brazil. ⁶Laboratório de Hepatites Virais, Instituto Oswaldo Cruz/IOC, FIOCRUZ, Rio de Janeiro, Brazil.

ABSTRACT: Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has a high transmissibility profile which favors the accumulation of mutations along its genome, providing the emergence of new variants. In this context, haplotype studies have allowed mapping specific regions and combining approaches and tracking phylogenetic changes. During the COVID-19 pandemic, it was notorious that home environments favored the circulation of SARS-CoV-2, in this study we evaluated 1,407 individuals positive for SARS-CoV-2, in which we located 53 families in the period from June 2021 to February 2023. The epidemiological data were collected in E-SUS notifica and SIVEP-gripe. Then, the genetic material was extracted using the commercial kit and the viral load was evaluated and the viral genomes were sequenced using the Illumina MiSeq methodology. In addition, the circulation of 3 variants and their respective subvariants was detected. The delta variant represented the highest number of cases with 45%, the Omicron variant 43% and the lowest number with 11% of cases the Gamma variants. There were cases of families infected by different subvariants, thus showing different sources of infection. The haplotype network showed a distribution divided into 6 large clusters that were established according to the genetic characteristics observed by the algorithm and 224 Parsimony informative sites were found. In addition, 92% of subjects were symptomatic and 8% asymptomatic. The secondary attack rate of this study was 8.32%. Therefore, we can infer that the home environment favors the spread of SARS-CoV-2, so it is of paramount importance to carry out genomic surveillance in specific groups such as intradomiciliary ones.

KEYWORDS: Household, household infection cases, haplotypes, SARS-cov-2, COVID-19

RECEIVED: August 9, 2023. **ACCEPTED:** June 9, 2024.

TYPE: Original Research Article

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was funded by Fundação Oswaldo Cruz de Rondônia—FIOCRUZ/RO (PROEP 2021 Process: VPGDI-008-FIO-21-2-17), Departamento de Ciência e Tecnologia (DECIT), Fundação para o Desenvolvimento da Ação Científica e Tecnológica e à Pesquisa do Estado de Rondônia—FAPERO (Process: 01133100038-0000.72 / 2016; Public bid invitation: 012/2016 PRO-RONDÔNIA and PPSUS 001/2021 Process: 350.095.442.048.526.000.000)

and by Instituto Nacional de Epidemiologia da Amazônia Ocidental—INCT EpiAmO. FGN is a CNPq fellow. Departamento de Ciência e Tecnologia (DECIT) of the Brazilian MoH, US/CDC and OPAS, Brazilian office.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Deusilene Vieira, Laboratório de Virologia Molecular, Fundação Oswaldo Cruz Rondônia—FIOCRUZ/RO, Rua da Beira, 7176, Bairro Lagoa, 76812-245 Porto Velho, Rondônia, Brazil. Email: deusilene.vieira@fiocruz.br

Introduction

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection has become a public health issue, with more than 767 million cases worldwide.¹ In evolutionary terms, SARS-CoV-2 presents a high transmissibility rate that favors the accumulation of mutations that impact the number of variations at specific sites of the genome.^{2,3}

In addition, the spread of SARS-CoV-2 has shown that most cases of infection with this virus occur in the home environment, characterized as a place that allows intense and prolonged contact between residents.^{4,5}

In this context, the study of haplotypes has allowed the mapping of specific regions of the genome.⁶ Haplotype determination, coupled with approaches and tracking of phylogenetic changes has become increasingly relevant in classifying

new lineages and understanding the dynamics of infections among close populations, such as in the household environment. It is characterized as a technique that favors the visualization of the behavior of certain groups of organisms because it uses metrics different from traditional phylogenies to establish distances.^{5,7}

Studies related to intrahousehold transmission have shown the emergence of new haplotypes in the household environment even with very close infections.^{8,9} In this study, we relate the main haplotypic markers and dynamics of infections involving the Omicron, Delta, and Gamma variants in individuals who share the same family environment, through the method of determining haplotypes associated with phylogeny, using data from complete viral sequences from the far west of the Amazon.



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

Methods

Place of study

This study was carried out at Laboratório de Virologia Molecular da Fundação Oswaldo Cruz de Rondônia-FIOCRUZ/RO in partnership with Rede Genômica da FIOCRUZ.

Ethical aspects

The study was approved by the Research Ethics Committee of the Centro de Pesquisa em Medicina Tropical de Rondônia-CEP/CEPEM with opinion No.4.637.465.

Study population and epidemiological data

A cohort of 1405 individuals positive for SARS-CoV-2 was selected from June 2021 to February 2023. Biological samples were collected at reference centers in different municipalities in the state of Rondônia. Epidemiological data were collected from SIVEP-Gripe and E-SUS.

Nucleic acid extraction and RT-qPCR

Biological samples from nasopharyngeal swab were forwarded to the Laboratório de Virologia Molecular (FIOCRUZ/RO) where viral RNA was extracted using the kit QIAamp® viral RNA Mini Kit (QIAGEN, Alemanha), according to the manufacturer's instructions. Viral RNA was subjected to the quantitative RT-qPCR test to determine the viral load of SARS-CoV-2, developed by Queiroz et al.¹⁰ Biological samples with Ct values <25 for the target were selected for genomic sequencing.

Complete genome sequencing of SARS-CoV-2

Genomic sequencing was performed using Illumina MiSeq and NextSeq platforms (Illumina, San Diego, CA, USA) and COVIDSEQ kit as previously reported.¹¹

Mutation analysis and lineage determination

Severe acute respiratory syndrome coronavirus 2 genomes were classified into lineages using the available software Pangolin COVID-19 Lineage Assigner v2.1.7¹² and mutations were analyzed with Nextclade v.1.5.4.¹³ Nextclade implements several quality controls (QCs) to flag potentially problematic sequences due to errors during sequencing or assembly. Sequences that produced mediocre and poor metrics were excluded from mutation analyses.

Phylogenetic and statistical analysis

High quality (>29 kb) whole reference genomes (<1% of N) of variants of concern (VOCs) Gamma, Delta, and Omicron

sampled in Brazil (n = 138) were downloaded from the GISAID EpiCoV database on September 11, 2022. The sequences generated in this the study (n=95) and the retrieved sequences were aligned using MAFFT v.7.487.¹⁴ The best model of nucleotide substitution was measured (GTR + G + I) using ModelFinder¹⁵ and the phylogenetic tree was reconstructed using the maximum likelihood method in the program IQ-TREE v.2.1.3.¹⁶ Branch support values were obtained using Ultrafast Bootstrap with 1000 replicates. The tree was visualized and edited with FigTree v.1.4.4.¹⁷ Statistical analysis and graphics were performed using R v4.0.3 software.

Analysis of the secondary attack rate among households

The secondary attack rate was estimated from the number of infected household contacts with the overall value of cases per household. Age group, vaccination, and other criteria were not considered.¹⁸

Haplotypes analysis

For this study, the Templeton, Crandal e Sing (TCS) method, named after its creators Templeton, Crandal and Sing, was used,¹⁹ and defaults to building a distance matrix between individuals (identifying haplotypes) and then identifying the shortest paths between groups and thus establishing network relationships. Because of the particularities of the statistical model used, this methodology can infer microevolutionary events.²⁰

The application of the TCS method for the inference of the haplotype network was conducted using the software PopART,²¹ a tool that allows the construction, visualization, and editing of haplotype networks through statistical inference models, enabling the comparison of different network conformations from the same dataset.

Results

Epidemiological data and secondary attack rate

A total of 1405 SARS-CoV-2-positive individuals were analyzed between June 24, 2021 and February 17, 2023, and 53 families were identified, which are formed between 2 and 4 individuals, totaling a number of 117 study participants (117/1405). The individuals were aged between 1 and 82 years, with a mean of 38 years, 55% (64/117) were male and 45% (53/117) female.

The cases of intradomiciliary infection were distributed in 20 municipalities of the state, covering 38% (20/52) of the entire territory. The capital Porto Velho accounted for the largest number of households characterized with 32% (17/53), followed by the municipalities of Ariquemes, Alvorada do Oeste, Chupinguaia, Teixeiraópolis, Jaru, Ji-Paraná, which had 15% (3/53) for each municipality. The municipalities of Cacoal,

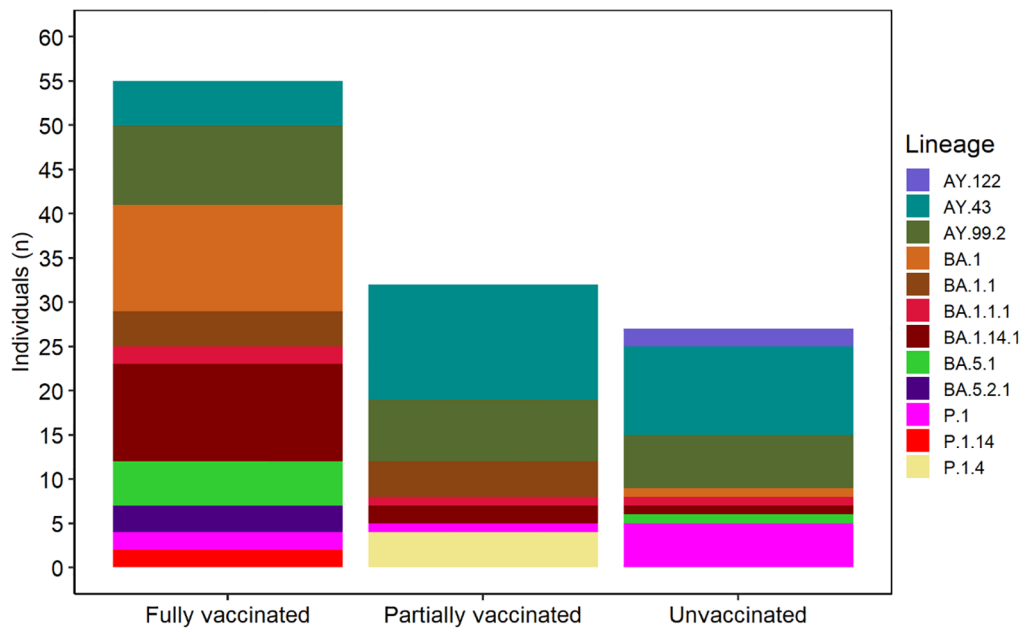


Figure 1. Bar plot showing lineages and vaccination data. Family 42 was removed for not containing vaccination data.

Cerejeiras, Machadinho do Oeste, Seringueiras, Urupá had 10% (2/53) of families for each municipality, while Castanheiras, Guajará-Mirim, Itapuã do Oeste, Presidente Médici, São Miguel do Guaporé, Vilhena, Candeias do Jamari, and Novo Horizonte do Oeste corresponded to 5% (1/53) of families for each municipality. The secondary attack rate of the study population was 8.32% (117/1405).

The most prevalent symptoms reported by individuals were cough 63% (74/117), headache 56% (66/114), fever 55% (64/117), runny nose 44% (51/117) and sore throat 32% (37/117), followed by less prevalent olfactory disorders 9% (11/117), taste disorders 9% (11/117), and dyspnea 4% (5/117). In addition, 8% (9/117) of cases were asymptomatic.

A total of 12% (14/117) of the individuals had comorbidities. Reported diseases include long-term respiratory diseases, heart disease, advanced stage chronic kidney disease, obesity, diabetes, and hypertension. Regarding the vaccination profile of the study population, 3% completed the vaccination schedule with 4 doses, 12% (14/117) were vaccinated with 3 doses, 35% (41/117) had 2 doses, 27% (32/117) were classified as partially vaccinated, as they had received only one dose and 23% (27/117) of the individuals were not vaccinated. Vaccination data of the study population compared with identified subvariants are shown in Figure 1.

Phylogenetic analysis

Three variants of SARS-CoV-2 were traced among the families. Through genomic sequencing analyses, the variants Gamma, Delta, and Omicron were confirmed, being the only

variants characterized in this study. For the construction of the phylogenetic tree and haplotype network, 95 high-quality sequences were included.

Delta variant accounted for the largest number of families representing 45% (24/53), presenting 3 subvariants, with AY.43 accounting for 50% (12/24) of the families, followed by AY.99.2 with 46% (11/24) and AY.122 with 4% (1/24). The Omicron variant presented 43% (23/53) of the families and was the variant that obtained the highest number of subvariants in intradomiciliary infections, being BA.1 with 26% (6/23), BA.1.1 with 13% (3/23), BA.1.1.1 with 9% (2/23), and BA.1.14.1 with 22% (5/23) of the families.

However, the Gamma variant was responsible for the lowest number of cases, presenting 11% (6/53) of the intradomiciliary cases, with the following subvariants: P.1 with 50% (3/6), P.1.14 with 17% (1/6), and P.1.4 with 33% (2/6) of the families (Figure 2).

The subvariants BA.2.36, BA.5.2.1, and BA.5.1 were identified in only one family member, as 11.32% (6/53) of the families showed infection with different subvariants.

Haplotype network

The analysis of SARS-CoV-2 sequences in individuals of the same family through a haplotype network were carried out with the intention of analyzing the possible different sources of infections and the virus evolution considering the geographical context and the dates of detection of the virus in individuals. The haplotype network showed a distribution divided into 6 major clusters that were established according to the genetic

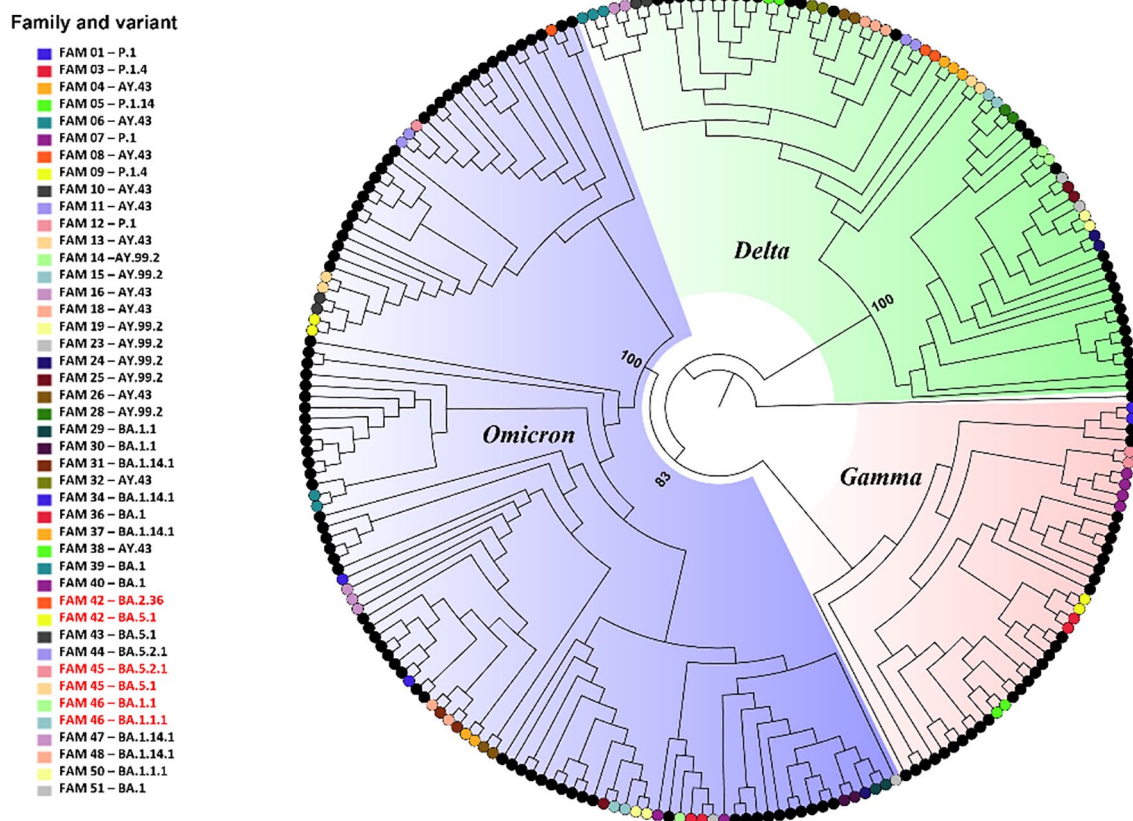


Figure 2. Maximum likelihood (ML) phylogenetic tree including isolated representatives of the Gamma, Delta, and Omicron variants. Maximum likelihood phylogenetic tree showing 95 sequences with “good” quality in NEXTCLADE referring to 44 families obtained in this study and 138 reference genomes retrieved from GISAID. The tree was rooted with the most ancestral sequence (EPI_ISL_402123). Gamma, Omicron, and Delta are shaded in red, blue, and green, respectively.

characteristics observed by the algorithm, and which are also in agreement with the classification inferred by Nextclade. As expected, individuals from the same family who had infection of the same SARS-CoV-2 strain were grouped in a single cluster; however, some individuals from the same family were grouped in different halogroups and others were not grouped in any cluster (Figure 3).

Parsimony informative sites contains at least 2 types of nucleotides, which occur at the same site at least twice. In this study, 224 Parsimony informative sites were found, which are plotted in green in Figure 4:

Discussion

The COVID-19 pandemic resulting from the high spread of SARS-CoV-2 has been relevant in recent years, showing a worrying scenario in Brazil, mainly in the Amazon due to successive introductions of variants and viral strains.²²⁻²⁴

The secondary attack rate obtained among the sample of this study was 8.32%, differing from the study carried out in 2022, in the State of Rondônia,²⁵ in which the secondary attack rate was higher, this decrease may be related to the

decrease in testing for SARS-CoV-2 and the advance of vaccination.

The symptomatology presented by the study population follows the clinical pattern of COVID-19,²⁶ it is important to emphasize the low frequency of manifestations such as olfactory and taste disorders and dyspnea, which at the beginning of the pandemic were more prevalent in infected individuals.²⁷

The different comorbidities related in this study population are already characterized as factors contributing to the development of severe cases of the disease.²⁸⁻³⁰ However, most individuals were vaccinated, being an important protective factor against severe SARS-CoV-2 infections and decreased risk of mortality.³¹

Regarding the profile of variants associated with infections in vaccinated individuals, the predominance of the BA.1 lineage and other lineages that present important mutations that confer characteristics of higher transmission, immunological and vaccine escape is notable.³² Which can be observed by the increase in cases in the State of Rondônia between January and March 2022, where the circulation of the Omicron variant was predominant.³³ Gamma, Delta, and Omicron variant clades

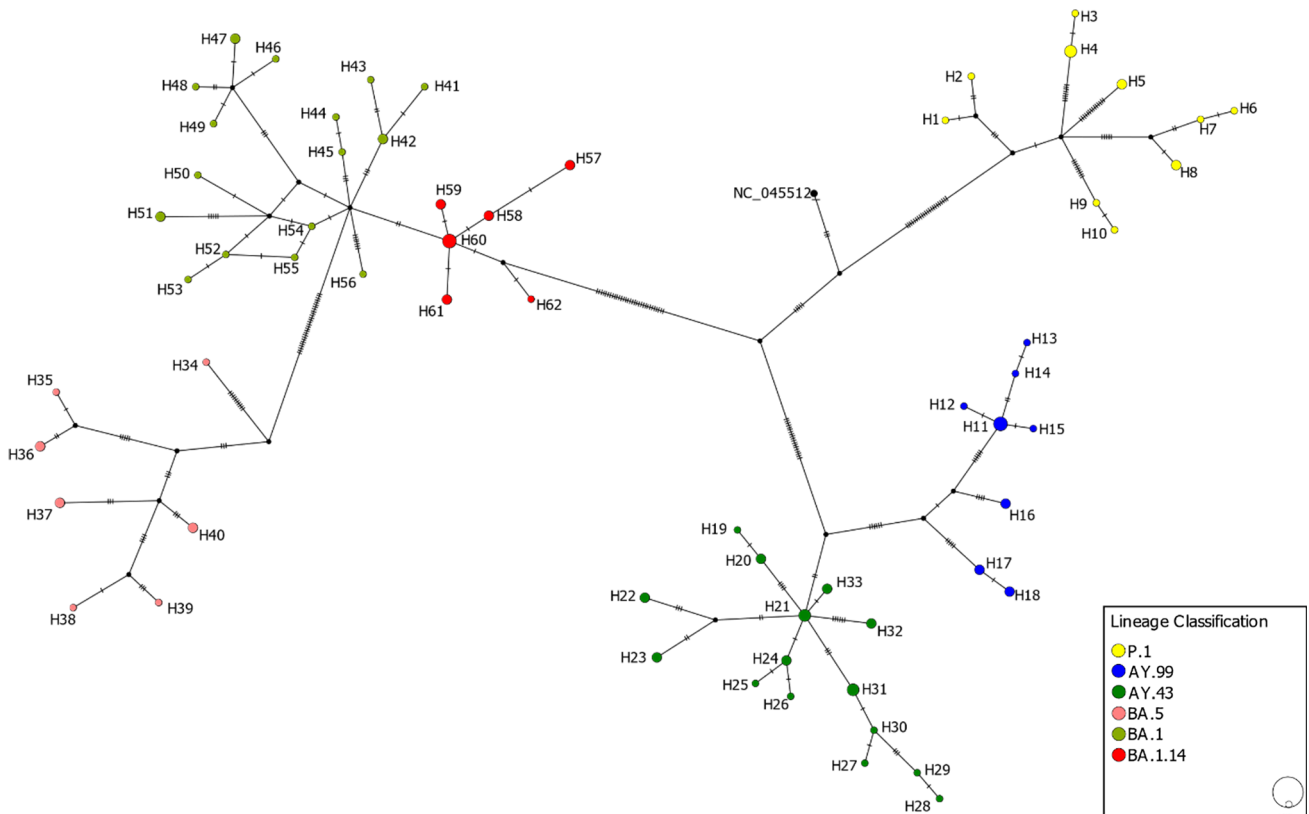


Figure 3. Haplotype network containing the Gamma, Delta, and Omicron variant clades and their respective subvariants. Haplotype network containing the strains of the Gamma, Delta, and Omicron variants that circulated in the State of Rondônia in the period corresponding to the study, each represented by different colors.

had wide distribution of cases,³⁴ demonstrating the ability of these variants that were in circulation to cause intra-household infections, as shown in some studies.^{35,36} It is noteworthy that few studies have correlated intra-household infections with the Gamma, Delta, and Omicron clades as demonstrated here.

In general, the construction of haplotype networks for SARS-CoV-2 has demonstrated that the technique is able to recognize clusters in large-sample groups.^{37,38} In this case, having the ability to consider phylogeography, as well as in small groups of samples.³⁹

In this study, it was evidenced that the haplogroups formed were not derived from a predominant haplogroup as already seen in other studies,³⁹ where from one event, there is dispersion into sub-clusters of haplotypes. Although there are modulated clusters according to lineages, our results show that there are no determinant haplotypes that could concentrate many individuals. Instead, the clusters were more discrete and limited to family profiles, which demonstrates great adaptive capacity of the virus, already evidenced in other studies.^{9,40} As expected, individuals from the same family who had infection of the same SARS-CoV-2 strain were grouped in a single cluster; however, some individuals from the same family grouped in different halogroups and others did not group in any cluster.

This can be explained by the presence of point mutations in the genome of these variants that were determinant for required halogroup identity.

Although SARS-CoV-2 has a revision mechanism in the viral replication process, mutations can occur in the virus genome, increasing its transmissibility and the ability to cause escape in the host immune system.⁴¹ It is also important to note that the process of sequencing and assembling genomes has a crucial influence on the process of inferring haplotype networks, as well as on phylogeny inferences.

We observed that individuals from 6 families (fam42, fam44, fam45, fam46, fam49 and fam52) became infected with different strains of SARS-CoV-2, possibly showing that these individuals from the same family became infected from different sources of infection. One of the factors that could elucidate the occurrence of infections from different sources consists of behavioral habits combined with the relaxation of restriction measures. It is noteworthy that these infections occurred only among the Omicron sublineages (BA. *). Compared to the Gamma and Delta variants, Omicron is characterized by a high number of accumulated mutations, high transmissibility and subsequently an increase in the number of cases of the disease.^{42,43}

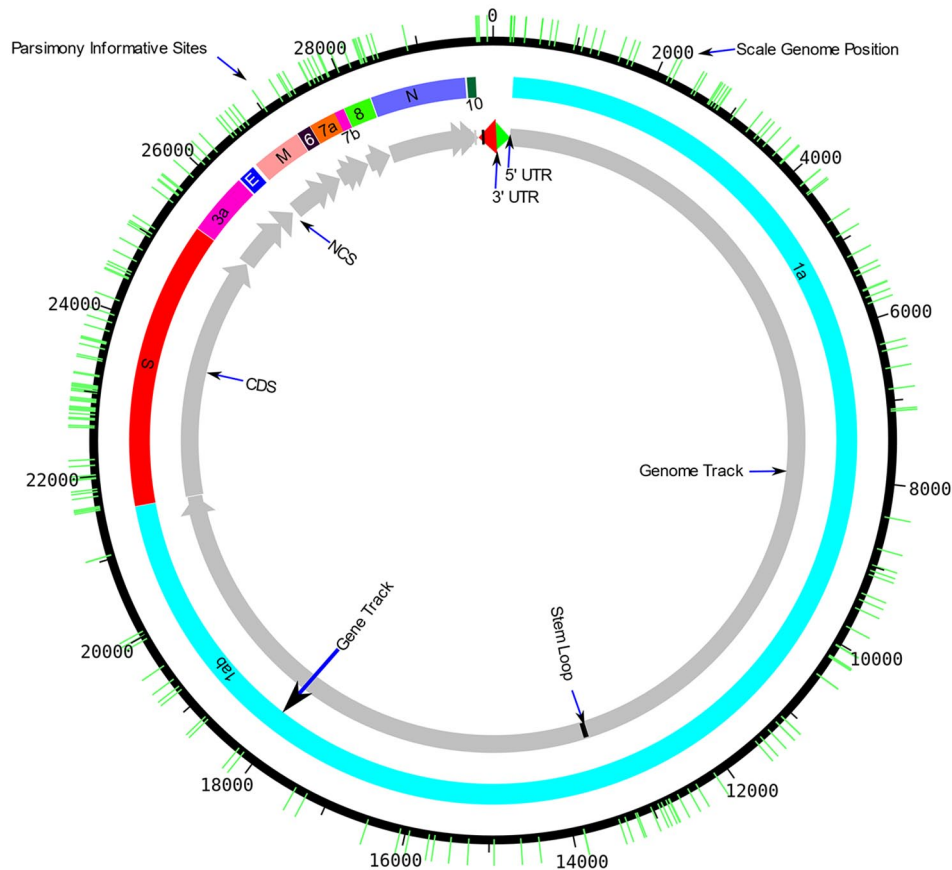


Figure 4. Distribution of parsimony informative sites along the SARS-CoV-2 virus genome. The parsimony informative sites are represented by the green color, in the position they were located in the SARS-CoV-2 virus genome.

Conclusion

We can infer that the intradomiciliary environment is a favorable place for the spread of the SARS-CoV-2, since it is a place that allows intense and prolonged contact between those who share the same domestic environment. In view of this, it is important to carry out genomic surveillance in groups that are more vulnerable, as is the case with intradomiciliary.

Acknowledgements

This study was developed by a group of researchers from Laboratório de Virologia Molecular da Fundação Oswaldo Cruz, in Rondônia, with financial support from the Fiocruz Genomic Network, Departamento de Ciência e Tecnologia (DECIT), Fundação para o Desenvolvimento das Ações Científicas e Tecnológicas da Pesquisa do Estado de Rondônia—FAPERÓ, Programa de Pesquisa para o SUS (PPSUS), as well as Instituto Nacional de Ciência e Tecnologia de Epidemiologia da Amazônia Ocidental—INCT- EpiAmo who have been important contributors to scientific development in the Amazon region. Collaboration from Coordenação de Aperfeiçoamento Pessoal de Nível Superior—CAPES, from whom some authors received financial aid (scholarships) during the production of this study, the Vice president of

Vigilância em Saúde e Laboratórios de Referências of Fiocruz, Instituto de Biologia Molecular do Paraná (IBMP) and Laboratório Central de Saúde Pública de Rondônia (LACEN/RO) were essential for the development of the study.

Author Contributions

Conceptualization: KST, MGFC, GOS, and DV; Data curation: KST, MGFC, and GSO; Formal analysis: KST, GOS, MGFC, TPR, AMSP, JASQ, and DV; Funding Acquisition: DV; Investigation: Methodology: KST, MGFC, GSO, TPR, and DV; Project Administration: DV; Supervision: DV; Writing—original draft: KST, MGFC, GSO, TPR, JR, AMSP, JASQ, ALFS, and DV; Writing—review & editing: MGFC and DV; All authors have read and agreed to the published version of the manuscript.

Data Availability

All the SARS-CoV-2 genomes generated and analyzed in this study are available in the EpiCov database in GISAID under the following ID numbers: EPI_ISL_11112665-11112679, EPI_ISL_11112681-11112700, EPI_ISL_11112702-11112704, EPI_ISL_11112706-11112719, EPI_ISL_11112721-11112725, EPI_ISL_11112727-11112746, EPI_ISL_11112748-11112760, EPI_ISL_11112762, EPI_ISL_11112764-11112768, EPI_ISL_11112770, EPI_ISL_

11112772-11112776, EPI_ISL_11112778-11112789, EPI_ISL_11112791-11112798, EPI_ISL_11622642-11622700, EPI_ISL_11622702-11622725, EPI_ISL_5030021, EPI_ISL_6575689-6575706, EPI_ISL_6575708, EPI_ISL_6575710-6575739, EPI_ISL_8623163-8623164, EPI_ISL_8623166-8623256, EPI_ISL_8623258-8623269, EPI_ISL_9414682-9414748, EPI_ISL_9414750-9414774, EPI_ISL_9636793-9636797, EPI_ISL_9636798-9636804 and EPI_ISL_9636805-9636878. The list of accession IDs may be found in the attached file in the supplemental materials.

SUPPLEMENTAL MATERIAL

Supplemental material for this article is available online.

REFERENCES

- WHO coronavirus (COVID-19) dashboard WHO coronavirus (COVID-19) dashboard with vaccination data. Accessed August 1, 2023. <https://covid19.who.int/>
- Potdar V, Vipat V, Ramdasi A, et al. Phylogenetic classification of the whole-genome sequences of SARS-CoV-2 from India & evolutionary trends. *Indian J Med Res.* 2021;153:166-174. Accessed August 2, 2023. https://journals.lww.com/ijmr/Fulltext/2021/01000/Phylogenetic_classification_of_the_whole_genome.14.aspx
- Flores-Vega VR, Monroy-Molina JV, Jiménez-Hernández LE, Torres AG, Santos-Preciado JI, Rosales-Reyes R. SARS-CoV-2: evolution and emergence of new viral variants. *Viruses.* 2022;14:653. Accessed August 2, 2023. <https://www.mdpi.com/1999-4915/14/4/653/htm>
- Madewell ZJ, Yang Y, Longini IM, Halloran ME, Dean NE. Household transmission of SARS-CoV-2: a systematic review and meta-analysis. *JAMA Netw Open.* 2020;3:e2031756. Accessed August 2, 2023. <https://pubmed.ncbi.nlm.nih.gov/33315116/>
- De Marco C, Marascio N, Veneziano C, et al. Whole-genome analysis of SARS-CoV-2 in a 2020 infection cluster in a nursing home of Southern Italy. *Infect Genet Evol.* 2022;99:105253.
- Lindley RA, Steele EJ. Analysis of SARS-CoV-2 haplotypes and genomic sequences during 2020 in Victoria, Australia, in the context of putative deficits in innate immune deaminase anti-viral responses. *Scand J Immunol.* 2021;94:e13100. Accessed August 2, 2023. <https://onlinelibrary.wiley.com/doi/full/10.1111/sji.13100>
- Clement M, Posada D, Crandall KA. TCS: a computer program to estimate gene genealogies. *Mol Ecol.* 2000;9:1657-1659. Accessed August 2, 2023. <https://onlinelibrary.wiley.com/doi/full/10.1046/j.1365-294x.2000.01020.x>
- Manuto L, Grazioli M, Spitaleri A, et al. Rapid SARS-CoV-2 intra-host and within-household emergence of novel haplotypes. *Viruses.* 2022;14:399. Accessed August 2, 2023. <https://www.mdpi.com/1999-4915/14/2/399/htm>
- Wang Y, Wang D, Zhang L, et al. Intra-host variation and evolutionary dynamics of SARS-CoV-2 populations in COVID-19 patients. *Genome Med.* 2021;13:1-13. Accessed August 2, 2023. <https://genomemedicine.biomedcentral.com/articles/10.1186/s13073-021-00847-5>
- Queiroz JADS, Rampazzo RCP, Filho EBDS, et al. Development of a quantitative one-step multiplex RT-qPCR assay for the detection of SARS-CoV-2 in a biological matrix. *Int J Infect Dis.* 2021;104:373-378. Accessed August 2, 2023. <http://www.ijidonline.com/article/S1201971221000023/fulltext>
- Naveca FG, Nascimento V, Souza V, et al. Spread of gamma (P.1) sub-lineages carrying spike mutations close to the furin cleavage site and deletions in the N-terminal domain drives ongoing transmission of SARS-CoV-2 in Amazonas, Brazil. *Microbiol Spectr.* 2022;10:e02366-21. Accessed August 2, 2023. <https://journals.asm.org/doi/10.1128/spectrum.02366-21>
- COG-UK. Accessed August 2, 2023. <https://pangolin.cog-uk.io/>
- Nextclade. Accessed August 2, 2023. <https://clades.nextstrain.org/>
- Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform.* 2019;20:1160-1166. doi:10.1093/bib/bbx108
- Kalyanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods.* 2017;14:587-589. Accessed August 2, 2023. <https://www.nature.com/articles/nmeth.4285>
- Trifunopoulos J, Nguyen LT, von Haeseler A, Minh BQ. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* 2016;44:W232-W235. doi:10.1093/nar/gkw256
- FigTree. Accessed August 2, 2023. <http://tree.bio.ed.ac.uk/software/figtree/>
- Cauchemez S, Donnelly CA, Reed C, et al. Household transmission of 2009 pandemic influenza A (H1N1) virus in the United States. *N Engl J Med.* 2009;361:2619-2627. Accessed August 3, 2023. <https://pubmed.ncbi.nlm.nih.gov/20042753/>
- Hudson RR, Slatkin M, Maddison WP. Estimation of levels of gene flow from DNA sequence data. *Genetics.* 1992;132:583-589. doi:10.1093/genetics/132.2.583
- Posada D, Crandall KA. Intraspecific gene genealogies: trees grafting into networks. *Trends Ecol Evol.* 2001;16:37-45. Accessed August 2, 2023. <http://www.cell.com/article/S0169534700020267/fulltext>
- Leigh JW, Bryant D. Popart: full-feature software for haplotype network construction. *Methods Ecol Evol.* 2015;6:1110-1116. Accessed August 2, 2023. <https://onlinelibrary.wiley.com/doi/full/10.1111/2041-210X.12410>
- Botelho-Souza LF, Nogueira-Lima FS, Roca TP, et al. SARS-CoV-2 genomic surveillance in Rondônia, Brazilian Western Amazon. *Sci Rep.* 2021;11:1-12. Accessed August 2, 2023. <https://www.nature.com/articles/s41598-021-83203-2>
- Sgorlon G, Queiroz JADS, Roca TP, et al. Clinical and epidemiological aspects of Delta and Gamma SARS-CoV-2 variant of concern from the western Brazilian Amazon. *Mem Inst Oswaldo Cruz.* 2023;117:e220155. Accessed August 2, 2023. <https://www.scielo.br/j/mioc/a/Cv9wxmpsfXpqb7ntsyH96Zw/?lang=en>
- Naveca FG, Nascimento V, de Souza VC, et al. COVID-19 in Amazonas, Brazil, was driven by the persistence of endemic lineages and P.1 emergence. *Nat Med.* 2021;27:1230-1238. Accessed April 18, 2022. <https://www.nature.com/articles/s41591-021-01378-7>
- Genomic surveillance of SARS-CoV-2: study of cases intradomestic infection in Brazilian Amazonia. Accessed October 12, 2022. <https://austinpublishing-group.com/infectious-diseases/fulltext/ajid-v9-id1065.php>
- Zhang JJ, Dong X, Cao YY, et al. Clinical characteristics of 140 patients infected with SARS-CoV-2 in Wuhan, China. *Allergy.* 2020;75:1730-1741. Accessed August 3, 2023. <https://onlinelibrary.wiley.com/doi/full/10.1111/all.14238>
- Qiu C, Cui C, Hautefort C, et al. Olfactory and gustatory dysfunction as an early identifier of COVID-19 in adults and children: an international multicenter study. *Otolaryngol Head Neck Surg.* 2020;163:714-721. Accessed August 3, 2023. <https://onlinelibrary.wiley.com/doi/full/10.1177/0194599820934376>
- Singh MK, Mobeen A, Chandra A, Joshi S, Ramachandran S. A meta-analysis of comorbidities in COVID-19: which diseases increase the susceptibility of SARS-CoV-2 infection? *Comput Biol Med.* 2021;130:104219
- Guan WJ, Liang WH, Zhao Y, et al. Comorbidity and its impact on 1590 patients with COVID-19 in China: a nationwide analysis. *Eur Respir J.* 2020;55:2000547. Accessed August 3, 2023. <https://erj.ersjournals.com/content/55/5/2000547>
- Chen N, Zhou M, Dong X, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet.* 2020;395:507-513. Accessed August 3, 2023. <http://www.thelancet.com/article/S0140673620302117/fulltext>
- Zhang J-J, Dong X, Liu G-H, Gao Y-D. Risk and protective factors for COVID-19 morbidity, severity, and mortality. *Clin Rev Allergy Immunol.* 2022;64:90-107. Accessed August 3, 2023. <https://link.springer.com/article/10.1007/s12016-022-08921-5>
- Shrestha LB, Foster C, Rawlinson W, Tedla N, Bull RA. Evolution of the SARS-CoV-2 omicron variants BA.1 to BA.5: implications for immune escape and transmission. *Rev Med Virol.* 2022;32:e2381. Accessed August 3, 2023. <https://onlinelibrary.wiley.com/doi/full/10.1002/rmv.2381>
- Dashboard—en—Genomahcov—Fiocruz. Accessed August 3, 2023. <https://www.genomahcov.fiocruz.br/dashboard-en/>
- Sgorlon G, Roca TP, Passos-Silva AM, et al. SARS-CoV-2 spike protein mutations in different variants: a comparison between vaccinated and unvaccinated population in Western Amazonia. *Bioinform Biol Insights.* 2023;17:11779322231186477. Accessed August 7, 2023. <https://journals.sagepub.com/doi/10.1177/11779322231186477>
- Lyngse FP, Mortensen LH, Denwood MJ, et al. Household transmission of the SARS-CoV-2 Omicron variant in Denmark. *Nat Commun.* 2022;13:1-7. Accessed August 3, 2023. <https://www.nature.com/articles/s41467-022-33328-3>
- McLean HQ, Grijalva CG, Hanson KE, et al. Household transmission and clinical features of SARS-CoV-2 infections. *Pediatrics.* 2022;149:e2021054178. Accessed August 3, 2023. <https://publications.aap.org/pediatrics/article/149/3/e2021054178/184868/Household-Transmission-and-Clinical-Features-of>
- Wu Y, Zheng S, Liu T, et al. Viral haplotypes in COVID-19 patients associated with prolonged viral shedding. *Front Cell Infect Microbiol.* 2021;11:715143.
- Kumar P, Pandey R, Sharma P, et al. Integrated genomic view of SARS-CoV-2 in India. *Wellcome Open Res.* 2020;5:184. Accessed August 3, 2023. <https://pubmed.ncbi.nlm.nih.gov/32995557/>
- Sekizuka T, Itokawa K, Kageyama T, et al. Haplotype networks of SARS-CoV-2 infections in the diamond princess cruise ship outbreak. *Proc Natl Acad Sci USA.* 2020;117:20198-20201. Accessed August 3, 2023. <https://www.pnas.org/doi/abs/10.1073/pnas.2006824117>

40. Tomaszewski T, Ali MA, Caetano-Anollés K, Caetano-Anollés G. Seasonal effects decouple SARS-CoV-2 haplotypes worldwide. *F1000Res*. 2023;12:267. Accessed August 3, 2023. <https://f1000research.com/articles/12-267>
41. Peng Q, Zhou R, Liu N, et al. Naturally occurring spike mutations influence the infectivity and immunogenicity of SARS-CoV-2. *Cell Mol Immunol*. 2022;19:1302-1310. Accessed August 3, 2023. <https://www.nature.com/articles/s41423-022-00924-8>
42. Andrews N, Stowe J, Kirsebom F, et al. Effectiveness of COVID-19 vaccines against the Omicron (B.1.1.529) variant of concern. *medRxiv*. 2021. Accessed August 3, 2023. <https://www.medrxiv.org/content/10.1101/2021.12.14.21267615v1>
43. Torjesen I. Covid-19: omicron may be more transmissible than other variants and partly resistant to existing vaccines, scientists fear. *BMJ*. 2021;375:n2943. Accessed August 3, 2023. <https://www.bmj.com/content/375/bmj.n2943>