Emergence of within-host SARS-CoV-2 recombinant genome after coinfection by Gamma and Delta variants

Ronaldo da Silva Francisco Junior

Laboratorio Nacional de Computacao Cientifica

Luiz G P de Almeida

Laboratorio Nacional de Computacao Cientifica

Alessandra P Lamarca

Laboratorio Nacional de Computacao Cientifica

Liliane Cavalcante

Laboratorio Nacional de Computacao Cientifica

Yasmmin Martins

Laboratorio Nacional de Computacao Cientifica

Alexandra L Gerber

Laboratorio Nacional de Computacao Cientifica

Ana Paula de C Guimarães

Laboratorio Nacional de Computacao Cientifica

Ricardo Barbosa Salviano

Departamento de Genetica, Instituto de Biologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

Fernanda Leitão dos Santos

Departamento de Genetica, Instituto de Biologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

Thiago Henrique de Oliveira

Departamento de Genetica, Instituto de Biologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

Isabelle Vasconcellos de Souza

Unidades de Apoio ao Diagnostico da Covid-19, Rio de Janeiro, Brazil

Erika Martins de Carvalho

Unidades de Apoio ao Diagnostico da Covid-19, Rio de Janeiro, Brazil

Mario Sergio Ribeiro

Secretaria Estadual de Saude do Rio de Janeiro, Rio de Janeiro, Brazil

Silvia Carvalho

Secretaria Estadual de Saude do Rio de Janeiro, Rio de Janeiro, Brazil

Flávio Dias da Silva

Secretaria Municipal de Saude Rio de Janeiro, Rio de Janeiro, Brazil

Marcio Henrique de Oliveira Garcia

Secretaria Municipal de Saude Rio de Janeiro, Rio de Janeiro, Brazil

Leandro Magalhães de Souza

Laboratorio Central de Saude Publica Noel Nutels, Rio de Janeiro, Brazil

Cristiane Gomes Da Silva

Laboratorio Central de Saude Publica Noel Nutels, Rio de Janeiro, Brazil

Caio Luiz Pereira Ribeiro

Secretaria Municipal de Saude Rio de Janeiro, Rio de Janeiro, Brazil

Andréa Cony Cavalcanti

Laboratorio Central de Saude Publica Noel Nutels, Rio de Janeiro, Brazil

Claudia Maria Braga de Mello

Secretaria Estadual de Saude do Rio de Janeiro, Rio de Janeiro, Brazil

Amilcar Tanuri

Departamento de Genetica, Instituto de Biologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

Ana Tereza R Vasconcelos (✉ atrv@lncc.br)

Laboratorio Nacional de Computacao Cientifica

Research Article

Keywords: Recombination, SARS-CoV-2, coinfection, Gamma, Delta

Posted Date: December 20th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1159806/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Since the first reports of patients coinfecting by two genetically-distinct lineages of SARS-CoV-2, the scientific community raised concerns about the recombination of intra-host viral RNA sequences as a possible mechanism underlying the emergence of novel variants. Indeed, this phenomenon occurs at a relatively high frequency among betacoronaviruses. Nevertheless, the few existing studies about recombination between genetically-distinct lineages of SARS-CoV-2 are restricted to detect the inter-host dissemination of genomes post-recombination events. However, the high genomic similarity between the current co-circulating lineages challenges the identification of these events. Here, we report the first case of intra-host SARS-CoV-2 recombination during a coinfection by the variants of concern (VOC) AY.33 (Delta) and P.1 (Gamma) supported by sequencing reads harboring a mosaic of lineage-defining mutations. By using next-generation sequencing reads intersecting regions that simultaneously overlap lineage-defining mutations from Gamma and Delta, we were able to identify a total of six recombinant regions across the SARS-CoV-2 genome within a sample. Four of them mapped in the spike gene and two in the nucleocapsid gene. We detected mosaic reads harboring a combination of lineage-defining mutations from each VOC. To our knowledge, this is the first report of intra-host RNA-RNA recombination between two lineages of SARS-CoV-2, which can represent a threat to public health management during the COVID-19 pandemic due to the possibility of the emergence of viruses with recombinant phenotypes.

Full Text

In late July, a 32-year-old male with mild flu-like symptoms of fever, headache, cough, fatigue, and sore throat was referred to the Santo Antônio de Pádua municipal health department in the northeastern region of the state of Rio de Janeiro, Brazil. After two days from the onset of symptoms, the diagnosis of COVID-19 was confirmed by SARS-CoV-2 reverse transcriptase polymerase chain reaction (RT-PCR) assay of a nasopharyngeal swab. Given the mild symptoms, no hospitalization was required and the patient was quarantined at home. The patient did not report any comorbidities and had not been vaccinated by the time the sample was collected.

Whole-genome sequencing analysis detected 73 intra-host single nucleotide variants (iSNVs) with allele frequency > 5% and depth > 100 reads. Of these, 26 were lineage-defining mutations exclusively found in the variant of concern (VOC) Gamma (P.1) and 32 were from Delta (AY.33; Table S1). By using a hypergeometric distribution approach¹, we estimate an overall haplotype frequency of approximately 16% and 82% for Delta and Gamma, respectively. Nevertheless, two iSNVs characteristic of VOC Delta (ORF1ab: I1091V - AF = 94%; and ORF7b: T40I - AF = 70%) were found with a high frequency in our analysis, suggesting that some genomic copies from Gamma haplotype could also harbor the same mutations (Table S1).

We then sought to investigate the occurrence of within-host recombination events between the two lineages using next-generation sequencing. In theory, mutations characteristic of Gamma should be kept in different reads from those defining mutations of Delta. We divided the SARS-CoV-2 genome into

windows with at least 100 nucleotides, which is smaller than the read length of the Illumina COVIDSeq Test (Illumina) sequencing kit used in this study. Next, we only selected windows that intersected at least two iSNVs, being one lineage-defining mutation from Gamma and the other from Delta. We aimed to detect unfragmented reads that covered the entire window harboring the iSNVs selected. A total of six recombinant candidate regions across the SARS-CoV-2 genome were found (see Supplementary Material), four mapped in the spike gene and two in the nucleocapsid gene (Table 1). We detected an average of two recombinant haplotypes supported by reads harboring a combination of the discriminating mutations from Gamma and Delta (Table 1; Figures S1-S6). In addition, the within-host recombinant sequences were formed by a single breakpoint event with a minimal and maximum interval of 15 and 75 nucleotides between two discriminate mutations where an event of recombination was observed, respectively.

Although the number of reads intersecting the discriminating iSNVs in the recombinant haplotypes seems to be small, we also observed a huge number of reads covering each mutation separately (Table S1; median read depth: 1947). Thus, the read coverage and minor allele frequency in each site suggest that these mutations were not caused by low mapping quality or miscalling variant issues. The hypothesis that the recombination might be the product of laboratory contamination is highly discouraged due to the protocol used by us. All viral particles are inactivated before RNA extraction, which stops replication and, consequently, impairs recombination from happening due to sample contamination. In addition, the high efficiency of ARTIC pool of primers utilized in our library construction rule out the possibility of the recombinant fragment being made by the initial PCR amplification step. Moreover, the 100 nucleotide window used in our genome scan and the similarity of sequence breakpoints analyzed in different genome regions make it difficult to imagine that it can be produced by the PCR step used in our library construction.

Recent studies reported the detection of SARS-CoV-2 recombinant lineages circulating at a low frequency²⁻⁶. Nevertheless, these studies were restricted to detecting the inter-host dissemination of genomes post-recombination events, which could be biased by spurious mutations generated due to library preparation, laboratory contamination, and unavailability of raw sequenced data. Here, we report the first case of intra-host SARS-CoV-2 recombination during a coinfection by the VOC Delta (AY.33) and Gamma (P.1) supported by sequencing reads harboring a mosaic of lineage-defining mutations as well as putative sequences breakpoints. We did not detect the circulation of recombinant sequences in the state, so far. Finally, Gamma and Delta variants were notable by causing an elevated number of Covid-19 cases and deaths in Brazil and worldwide, respectively. Both VOC are known to enhance SARS-CoV-2 transmissibility and induce an immune-scape response. Therefore, recombination between both sequences could represent the emergence of novel variants of concern with new phenotypic combinations.

Declarations

The authors declare that this study was approved by the Ethics Committee (30161620.0.1001.5257 and 34025020.0.0000.5257). The research protocol was approved without informed consent in accordance with Brazilian National Health Council's Resolution 510/2016. All samples were residual COVID-19 clinical diagnostic samples de-identified before receipt by the researchers.

Acknowledgment

We would like to thank the patient for taking part in this study. We also acknowledge the authors and administrators of the GISAID database, which allowed this study of genomic epidemiology to be conducted properly. The authors declare that funding sources had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' contributions

Conceptualisation: ATRV; DNA extraction and sequencing: LC, ALG, APCG, RBS, FLS, and THO; project administration: IVS, EMC, MSR, SC, FDS, MHOG, LMS, CGS, CLPR, ACC, and CMBM; formal analysis: RSFJ, LGPA, APL, and YM; data Curation: RSFJ and LGPA; funding: ATRV; writing – original draft: RSFJ, APL, LGPA, ATRV; writing – review & editing: AT and ATRV.

Conflicts 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.

Funding

This work was developed in the frameworks of Corona-ômica-RJ (FAPERJ = E-26/210.179/2020). A.T.R.V. is supported by CNPq (303170/2017-4) and FAPERJ (E-26/202.903/20); A.T. by FAPERJ E-26/010.002434/2019 and E-26/210.178/2020. R.S.F.J is a recipient of a graduate fellowship from CNPq, A.P.L is granted a post-doctoral scholarship (DTI-A) from CNPq. We acknowledge the support from the Rede Corona-ômica BR MCTI/FINEP affiliated to RedeVírus/MCTI (FINEP 01.20.0029.000462/20, CNPq 404096/2020-4).

References

1. Zhou H-Y, Cheng Y-X, Xu L, et al. Genomic evidence for divergent co-infections of SARS-CoV-2 lineages. *bioRxiv*. 2021;: 2021.09.03.458951.
2. Jackson B, Boni MF, Bull MJ, et al. Generation and transmission of interlineage recombinants in the SARS-CoV-2 pandemic. *Cell* 2021; 184: 5179–88.e8.
3. VanInsberghe D, Neish AS, Lowen AC, Koelle K. Recombinant SARS-CoV-2 Genomes Circulated at Low Levels Over The First Year of The Pandemic. *Virus Evolution*. 2021. DOI:10.1093/ve/veab059.

4. Varabyou A, Pockrandt C, Salzberg SL, Pertea M. Rapid detection of inter-clade recombination in SARS-CoV-2 with Bolotie. *Genetics* 2021; published online May 13. DOI:10.1093/genetics/iyab074.
5. Yi H. 2019 Novel Coronavirus Is Undergoing Active Recombination. *Clin Infect Dis* 2020; 71: 884–7.
6. Ignatieva A, Hein J, Jenkins PA. Ongoing recombination in SARS-CoV-2 revealed through genealogical reconstruction. DOI:10.1101/2021.01.21.427579.

Table 1

Table 1 is available in the Supplementary Files section.

Figures

Figure 1

Intra-host single nucleotide variant frequencies evidence coinfection by two genetically distinct lineages. A) Identification of iSNVs across the different genes of the SARS-CoV-2 genome within our sample. In blue and red, we showed the lineage-defining mutations from the variant Gamma and Delta, respectively. iSNVs represented in black are mutations commonly found in both lineages that descend from their ancestor B.1. B) Zoom in gene S showing the location of each genomic window (I-IV) where recombination events were investigated. C) Zoom in gene N showing the distribution of both selected regions (V and VI) across the gene.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterial.docx](#)
- [floatimage2.png](#)