

Genomic monitoring unveils a high prevalence of SARS-CoV-2 Omicron variant in vaccine breakthrough cases

Gúbio Soares Campos¹, Marta Giovanetti^{2,3,4}, Laise de Moraes^{5,6}, Helena Souza da Hora¹, Keila Oliveira Motta de Alcantara^{1,7}, Silvia Ines Sardi^{1*}

1- Laboratory of Virology, Health Institute of Science, Federal University of Bahia, Salvador, Bahia, Brazil;

2- Laboratório de Flavivírus, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, Rio de Janeiro, Brazil;

3- Department of Science and Technology for Humans and the Environment, University of Campus Bio-Medico di Roma, Rome, Italy;

4- Federal University of Minas Gerais; Belo Horizonte, Minas Gerais, Brazil;

5- Laboratório de Enfermidades Infecciosas Transmitidas por Vetores, Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Bahia, Brazil;

6- Programa de Pós-graduação em Ciências da Saúde, Faculdade de Medicina da Bahia, , Federal University of Bahia, Salvador, Bahia, Brazil.

7-Programa de Pos-graduação em Microbiologia, Instituto de Biologia, Federal University of Bahia, Salvador, Bahia, Brazil

Keywords: COVID-19, SARS-CoV-2, Omicron Variant, vaccine breakthrough

* Corresponding author E-mail: sissardi@yahoo.com.br

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

Abstract

Genome sequencing proved to be an excellent tool to monitor the molecular epidemiology of the disease caused by SARS-CoV-2, i.e., coronavirus disease (COVID-19). Some reports of infected, vaccinated individuals have aroused great interest because they are primarily being infected with circulating variants of concern (VOCs). To investigate the cases of infected, vaccinated individuals in Salvador, Bahia, Brazil, we performed genomic monitoring to estimate the magnitude of the different VOCs in these cases. Nasopharyngeal swabs from infected (symptomatic and asymptomatic), fully vaccinated individuals (n=29) who were of varying age and had RT-qPCR Ct values of ≤ 30 were subjected to viral sequencing using Nanopore technology. Our analysis revealed that the Omicron variant was found in 99% of cases and that only one case was due to the Delta variant. Infected, fully vaccinated patients have a favorable clinical prognosis; however, within the community, they become viral carriers with the aggravating factor of viral dissemination of VOCs not neutralized by the vaccines.

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which causes COVID-19, a disease that emerged in December 2019 in Wuhan, China, fueled worldwide efforts to develop vaccines to control the rapid spread of infection.

Currently, there are several vaccines against SARS-CoV-2 approved by the World Health Organization (WHO)¹. However, the viral replication of this RNA virus has general characteristics common to other RNA viruses; the high mutation rate allows it to generate new viral variants to improve its chances of survival in the host and escape to immune detection².

It has been shown that SARS-CoV-2 vaccines do not protect against viral infection, although they significantly reduce morbidity and mortality in infected patients. Moreover, the neutralizing antibodies persist for no more than 4 or 5 months, even with a full two-dose regimen³. These factors may contribute to the fact that fully vaccinated individuals, as well as those with booster doses, may acquire the infection. Although the clinical situation of the infected patient who is fully vaccinated has a very favorable prognosis, they become a possible viral carrier and can trigger viral dissemination within the community^{4,5}.

Recent studies of infected, vaccinated individuals have generated notable interest because they show that these individuals can primarily be infected with the circulating

variants of concern (VOCs), such as Omicron (B.1.529). This variant is more aggressive, with greater transmissibility and infectivity compared to the Delta variant⁶.

Here, we conducted a genomic study to estimate the magnitude and range of SARS-CoV-2 VOCs in cases of infected (symptomatic and asymptomatic), fully vaccinated individuals.

Materials and Methods

Molecular detection of SARS-CoV2

Patients (n=29) of any gender or age group, symptomatic or not, after consenting, clinical information and oropharyngeal and nasal swab were collected at the Laboratory of Virology at Federal University of Bahia. Detection of SARS-CoV-2 was carried out using an RT-qPCR (GoTaq® Probe 1-Step RT-qPCR System; Promega, USA) assay following the CDC 2019 Novel Coronavirus (2019-nCoV) Real-Time Reverse Transcriptase (RT)-PCR Diagnostic Panel⁷.

Ethics statement

This research was reviewed and approved by the Ethical Committee of the Federal University of Bahia (CAAE 30687320.9.0000.5662), and informed consent of all participating subjects or their legal guardians have been obtained.

cDNA synthesis and whole genome sequencing

Samples (n=29) were selected for sequencing based on the Ct value (≤ 30) and availability of epidemiological metadata (sex, age, residence in Salvador, symptoms, etc.) (**Table 1**). The preparation of SARS-CoV-2 genomic libraries was performed using the nanopore sequencing technology⁸. The SuperScript IV Reverse Transcriptase kit (Invitrogen, USA) was initially used for cDNA synthesis, following the manufacturer's instructions. The cDNA generated was subjected to multiplex PCR sequencing using the Q5 High Fidelity Hot-Start DNA Polymerase (New England Biolabs, UK) and a set of specific primers designed by the ARTIC Network for sequencing the complete SARS-CoV-2 genome (version 4)⁹. All experiments were performed in a biosafety level-2 cabinet. Amplicons were purified using 1x AMPure XP Beads (Beckman Coulter, USA) and quantified on a Qubit 3.0 fluorimeter (ThermoFisher Scientific, USA) using Qubit™ dsDNA HS Assay Kit (ThermoFisher Scientific, USA). DNA library preparation was performed using the Ligation Sequencing Kit SQK-LSK109 (Oxford Nanopore Technologies, UK) and the Native Barcoding Kit (EXP-NBD104 and EXP-NBD114,

Oxford Nanopore Technologies, UK). Sequencing libraries were loaded into an R9.4 flow cell (Oxford Nanopore Technologies, UK). In each sequencing run, we used negative controls to prevent and check for possible contamination with less than 2% mean coverage.

Generation of consensus sequences from nanopore

The fast5 files generated during sequencing were basecalled under the high-accuracy model performed by Guppy v.6.0.1 (Oxford Nanopore Technologies, UK). The basecalled fastQ files with a minimum Q score of 7 were selected for subsequent trim adaptor and demultiplex processes performed by Guppy v.6.0.1. Furthermore, the fastQ files were submitted to the ARTIC Network's field bioinformatics pipeline v.1.2.1⁹. Briefly, a length filtering was performed to remove additional chimeras using artic guppyplex. The assembly was performed by Minimap2 v.2.17¹⁰ using Genbank accession no. MN908947.3 as genome reference. The primer sequences were trimmed with align_trim.py, the assembly was then polished, and the variant calling was performed by Medaka v.1.0.3 (Oxford Nanopore Technologies, UK) and evaluated by LongShot v.0.4.1¹¹. The consensus sequences were then masked with "N" at regions with coverage depth <20, and the variant candidates were incorporated into the consensus genome using BCFtools v.1.10.2¹². The alignment statistics were calculated with SAMtools v.1.10 (using htlib 1.10.2)¹², exonerate v.2.4.0 (using glib version v.2.68.0)¹³ and Seqtk v.1.3-r106¹⁴. This entire workflow is available at <https://github.com/khourious/vgapONT>.

Phylogenetic inference

Lineage assignment was performed using the Pangolin lineage classification software tool¹⁵. The newly identified isolates were compared to a diverse pool of genome sequences (n= 3,441) sampled worldwide collected up to October 28th, 2021. All sequences were aligned using the ViralMSA tool v.1.1.20^{16,17}, and IQ-TREE2 v.2.2.0¹⁸ was used for phylogenetic analysis using the maximum likelihood approach. TreeTime v.0.8.5¹⁹ was used to transform this ML tree topology into a dated tree using a constant mean rate of 8.0×10^{-4} nucleotide substitutions per site per year, after the exclusion of outlier sequences.

Results

In this study, all individuals (n=29) were vaccinated with a complete two-dose regimen, except for four 7- and 11-year-olds who were not vaccinated (Table 1). The symptoms self-referred by participants were mild or asymptomatic, without the need for hospital care. The RT-qPCR Ct values varied between 14.6 and 27.2, even in individuals who declared themselves as asymptomatic (Fig. 1). From those vaccinated individuals (n=25) 24 confirm the presence of Omicron variant and only one Delta variant. From unvaccinated individuals (n=4) confirmed an Omicron Variant's infection. Our genomic surveillance analysis revealed that in the state of Bahia, three VOCs primarily dominated the epidemiology of the state. The Gamma variant (P.1) that was prominent during the second epidemic wave persisted until August 2021, when it was replaced by Delta (B.1.617.2) (Fig. 2a), which in turn was replaced by the emerging Omicron variant in December 2021. The Alpha variant was also detected in the state of Bahia at the end of December 2021, but it remained at a very low frequency (less than 1%). To explore the relationship between the sequenced genomes, we also constructed a phylogenetic tree with these VOCs and those from other parts of the world. Our time-stamped phylogeny revealed that the Omicron isolates in the Bahia state are scattered throughout the phylogeny, suggesting that multiple independent introductions have occurred through time (Fig. 2b).

Discussion

Since the start of the COVID-19 pandemic in December 2019, efforts have been directed to the development of vaccines against SARS-CoV-2. Given this global situation, WHO has authorized the use of several types of vaccines utilizing mRNA technology, adenoviruses, and inactivated viruses¹.

Recent reports confirm that vaccines currently in use do not protect against infection, but they do greatly reduce the severity of cases. Indeed, countries that have achieved vaccination coverage in at least 50% of their population managed to decrease the mortality rate. However, the current scenario shows the tendency of vaccinated individuals becoming infected with emerging VOCs that cannot be neutralized by the immunity generated from these vaccines². The new variants have emerged as a result of increased viral circulation, primarily in countries where vaccine coverage is low, which quickly spread worldwide. Examples of these variant leaks include Delta in India and Omicron in South Africa.

Our results reinforce the notion that virus migration generally follows national and international patterns of human mobility, facilitating the spread of emerging VOCs not only within countries, but also globally. The identification of SARS-CoV-2 suspected cases via genome-wide sequencing in the state of Bahia also revealed this, and as of 2020, co-circulation of three different VOCs such as Gamma (P.1), Delta, and Omicron appears to have dominated the epidemiological history in the state. The Alpha variant currently remains at a very low frequency (less than 1%) in the state, consistent with the effectiveness of SARS-CoV-2 vaccines against it while being less effective against new VOCs⁵

Initially, it was considered in the literature that reaching herd immunity ($\geq 70\%$) would prevent the circulation of the virus in the population; this fact is now being questioned. The countries that have reached or surpassed this percentage (England, Israel, and Brazil) have high infection rates since the emergence of Omicron^{20,21}. In our state, 74.9 % of the vaccination coverage has currently been reached with two doses, but our work demonstrated breakthrough infection by Omicron with the high viral load even in double or triple vaccinated individuals. Moreover, asymptomatic vaccinated individuals can still spread the virus to other people.

One limitation of our work is not knowing the level of protective immunity of those who are currently vaccinated; however, even with three doses, there have been cases of viral infection. Dose reinforcements in countries such as England and Israel with a four-dose regimen raise doubts as to whether they will solve the situation of viral recrudescence for new VOCs in circulation^{22,23}.

It is important to acknowledge the limitations of these vaccines and to urgently develop new vaccines to emergent VOCs, as is the case of the vaccine for Influenza; going through new doses of the same vaccines is “more of the same”.

Measures to control the use of face masks, social distancing, and crowd avoidance continue to be effective, but these need to be combined with new immunogens that prevent viral infection (an elementary concept for the approval of vaccines for widespread use).

Acknowledgments

This work was supported by Fundação de Apoio a Extensão e Pesquisa -Bahia-Brazil (FAPESB) GRANT 2020 /COVID 19; MG is supported by Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) (E-26/202.248/2018(238504) and by

the CRP- ICGEB RESEARCH GRANT 2020 Project CRP/BRA20-03, Contract CRP/20/03.

Data availability statement

Newly generated SARS-CoV-2 sequence have been deposited in GISAID under Accession Numbers: EPI_ISL_9265729, EPI_ISL_9265745, EPI_ISL_9265746, EPI_ISL_9265743, EPI_ISL_9265744, EPI_ISL_9265749, EPI_ISL_9265728, EPI_ISL_9265747, EPI_ISL_9265748, EPI_ISL_9265730, EPI_ISL_9265752, EPI_ISL_9265731, EPI_ISL_9265753, EPI_ISL_9265750, EPI_ISL_9265751, EPI_ISL_9265734, EPI_ISL_9265756, EPI_ISL_9265735, EPI_ISL_9265732, EPI_ISL_9265754, EPI_ISL_9265733, EPI_ISL_9265755, EPI_ISL_9265738, EPI_ISL_9265739, EPI_ISL_9265736, EPI_ISL_9265737, EPI_ISL_9265741, EPI_ISL_9265742, EPI_ISL_9265740.

References

1. World Health Organization. WHO Coronavirus Disease (COVID-19) <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/covid-19-vaccines>
2. Mengist HM, Kombe Kombe AJ, Mekonnen D, Abebaw A, Getachew M, Jin T. Mutations of SARS-CoV-2 spike protein: Implications on immune evasion and vaccine-induced immunity. *Semin Immunol.* 2021 Jun;55:101533. Doi: 10.1016/j.smim.2021.101533.
3. Mizrahi B, Lotan R, Kalkstein N, Peretz A, Perez G, Ben-Tov A, Chodick G, Gazit S, Patalon T. Correlation of SARS-CoV-2-breakthrough infections to time-from-vaccine. *Nat Commun.* 2021 Nov 4;12(1):6379. Doi: 10.1038/s41467-021-26672-3.
4. Servellita V, Morris MK, Sotomayor-Gonzalez A, Gliwa AS, Torres E, Brazer N, Zhou A, Hernandez KT, Sankaran M, Wang B, Wong D, Wang C, Zhang Y, Reyes KR, Glasner D, Deng X, Streithorst J, Miller S, Frias E, Rodgers M, Cloherty G, Hackett J Jr, Hanson C, Wadford D, Philip S, Topper S, Sachdev D, Chiu CY. Predominance of antibody-resistant SARS-CoV-2 variants in vaccine breakthrough cases from the San Francisco Bay Area, California. *Nat Microbiol.* 2022 Feb;7(2):277-288. Doi: 10.1038/s41564-021-01041-4.

5. Hacısuleyman E, Hale C, Saito Y, Blachere NE, Bergh M, Conlon EG, Schaefer-Babajew DJ, DaSilva J, Muecksch F, Gaebler C, Lifton R, Nussenzweig MC, Hatzioannou T, Bieniasz PD, Darnell RB. Vaccine Breakthrough Infections with SARS-CoV-2 Variants. *N Engl J Med*. 2021 Jun 10;384(23):2212-2218. Doi: 10.1056/NEJMoa2105000.
6. Chen J, Wang R, Gilby NB, Wei GW. Omicron (B.1.1.529): Infectivity, vaccine breakthrough, and antibody resistance. *J Chem Inf Model*. 2022 Jan 24;62(2):412-422.
7. National Center for Immunization and Respiratory Diseases (NCIRD) D of VD. Novel Coronavirus (2019-nCoV) Real-time RT-PCR Primers and Probes. <https://www.cdc.gov/Coronavirus/2019-Ncov/Lab/Rt-Pcr-Panel-Primer-Probes.Html>.
8. Giovanetti M. et al. Genomic epidemiology reveals how restriction measures shaped the SARS-CoV-2 epidemic in Brazil. *Medrxiv*. <https://doi.org/10.1101/2021.10.07.21264644>
9. ARTICnetwork. ARTIC field bioinformatics pipeline [Internet]. GitHub. Available from: <https://github.com/artic-network/fieldbioinformatics>
10. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 2018;34(18):3094–100.
11. Edge P, Bansal V. Longshot enables accurate variant calling in diploid genomes from single-molecule long read sequencing. *Nat Commun*. 2019 Dec 11;10(1):4660.
12. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. *Gigascience*. 2021 Jan 29;10(2):1–4.
13. Slater GSC, Birney E. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics*. 2005 Feb 15; 6:31.
14. Li H. Seqtk [Internet]. GitHub. Available from: <https://github.com/lh3/seqtk>
15. Rambaut A, Holmes EC, O’Toole Á, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nature Microbiology*. Published online July 15, 2020. Doi:10.1038/s41564-020-0770-5.
16. Moshiri N. ViralMSA: massively scalable reference-guided multiple sequence alignment of viral genomes. *Bioinformatics*. 2021;37(5):714-716. Doi:10.1093/bioinformatics/btaa743.
17. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 2018;34(18):3094-3100. Doi:10.1093/bioinformatics/bty191.

18. Minh BQ, Schmidt HA, Chernomor O, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*. 2020;37(5):1530-1534. Doi:10.1093/molbev/msaa015
19. Sagulenko P, Puller V, Neher RA. TreeTime: Maximum-likelihood phylodynamic analysis. *Virus Evol*. 2018;4(1). Doi:10.1093/ve/vex042.
20. Goldberg Y, Mandel M, Bar-On YM, Bodenheimer O, Freedman L, Haas EJ, Milo R, Alroy-Preis S, Ash N, Huppert A. Waning Immunity after the BNT162b2 Vaccine in Israel. *N Engl J Med*. 2021 Dec 9;385(24):e85. Doi: 10.1056/NEJMoa2114228.
21. Rössler A, Riepler L, Bante D, von Laer D, Kimpel J. SARS-CoV-2 Omicron Variant Neutralization in Serum from Vaccinated and Convalescent Persons. *N Engl J Med*. 2022 Jan 12;NEJMc2119236. Doi: 10.1056/NEJMc2119236.
22. World Health Organization. WHO Coronavirus Disease (COVID-19) <https://www.who.int/countries/isr/> ; <https://www.who.int/countries/gbr/>.
23. Eggink D, Andeweg SP, Vennema H, van Maarseveen N, Vermaas K, Vlaemyneck B, Schepers R, van Gageldonk-Lafeber AB, van den Hof S, Reusken CB, Knol MJ. Increased risk of infection with SARS-CoV-2 Omicron BA.1 compared with Delta in vaccinated and previously infected individuals, the Netherlands, 22 November 2021 to 19 January 2022. *Euro Surveill*. 2022 Jan;27(4). Doi: 10.2807/1560-7917.ES.2022.27.4.2101196.

Table 1. Demographic and Clinical characteristics in vaccinated or unvaccinated study group.

<i>Characteristic</i>	<i>Groups</i>	<i>Number/Total</i>
<i>Age</i>	6-19	5/29
	20-30	8/29
	31-45	7/29
	46-60	5/29
	> 61	4/29
<i>Sex</i>	F	15/29
	M	14/29
<i>Immunization</i>	2 doses	18/29
	3 doses	6/29
	Not Immunized	4/29
<i>Symptomatic</i>	Fever-Cough- Sore Throat- Headache	15/29
<i>Asymptomatic</i>		14/29

Figure 1

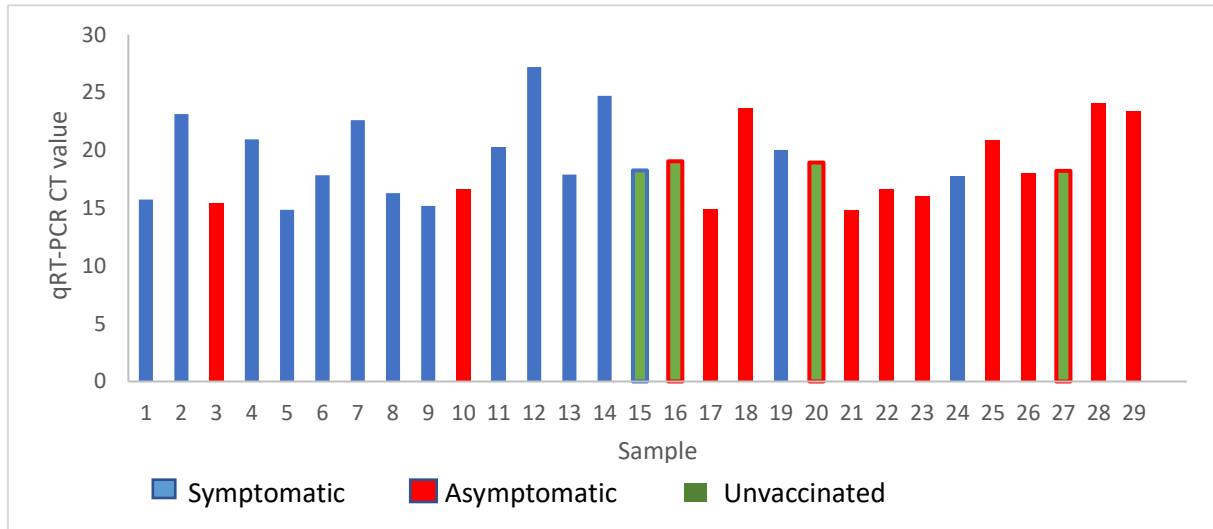


Figure 2

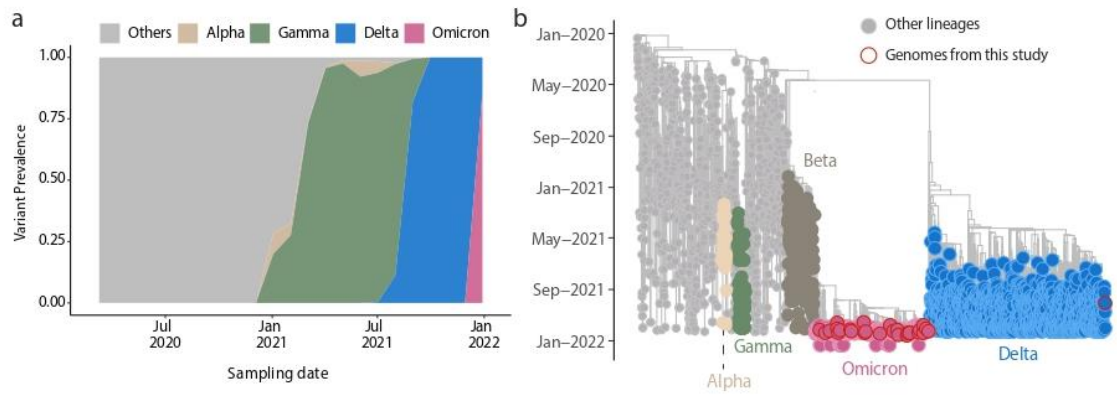


Figure Legends

Figure 1. Distribution of Ct values in the clinical samples from vaccinated or unvaccinated infected by Omicron or Delta variant.

Figure 2. Genomic epidemiology of the SARS-CoV-2. Omicron variant in Salvador, Bahia, Northeast Brazil. A) Dynamics of the SARS-CoV-2 epidemic in Bahia showing the progression in the proportion of circulating variants in the state over time, with the rapid replacement of the Delta by the Omicron variant; B) Time-resolved maximum likelihood phylogenetic tree including the newly (n=29) Omicron isolates obtained in this study plus n = 4,249 representative SARS-CoV-2 genomes collected up to January 16th, 2021. Alpha, Beta, Gamma, Delta and Omicron VOCs are highlighted in the tree. Newly Omicron and Delta genomes obtained in this study are highlighted with a red border. Genomes of other lineages are showed in grey.