PCR-Sequencing Approaches to Assess Informative Mutations in SARS-Cov-2 Spike (S) and ORF7, ORF8 and N Genes Characterizing Variants of Concern and Variants of Interest

Glauco Akelinghton Freire Vitiello
Institute for Cancer Research of Guarapuava
Isabela Medeiros Oliveira
Institute for Cancer Research of Guarapuava
Fernanda Ivanski
UNICENTRO: Universidade Estadual do Centro-Oeste
Bárbara Luisa Fermino
UNICENTRO: Universidade Estadual do Centro-Oeste
Emanuele Cristina Gustani-Buss
UNICENTRO: Universidade Estadual do Centro-Oeste
Rafael dos Santos Bezerra
USP Ribeirao Preto: Universidade de Sao Paulo Campus de Ribeirao Preto
Carlos Eduardo Buss
UNICENTRO: Universidade Estadual do Centro-Oeste
João Paulo Bianchi Ximenez
USP Ribeirao Preto: Universidade de Sao Paulo Campus de Ribeirao Preto
Bárbara Mendes Paz
UNICENTRO: Universidade Estadual do Centro-Oeste
Katiuscia de Oliveira Francisco Gabriel
UNICENTRO: Universidade Estadual do Centro-Oeste
Carlos Alberto Oliveira Biaggi Jr
USP Ribeirao Preto: Universidade de Sao Paulo Campus de Ribeirao Preto
Omics working group NAPI Genomics Consortium
Institute for Cancer Research (IPEC) Guarapuava
Daiane Priscila Simão Silva
Institute for Cancer Research (IPEC) - Guarapuava-PR
Emerson Carraro
UNICENTRO: Universidade Estadual do Centro-Oeste
Kamila Chagas Peronni
Institute for Cancer Research (IPEC) - Guarapuava-PR

David Livingstone Alves Figueiredo
UNICENTRO: Universidade Estadual do Centro-Oeste
Wilson Araujo Silva Jr (wilsonjr@usp.br)
Ribeirão Preto Medical School, University of Sao Paulo https://orcid.org/0000-0001-9364-2886

Short Report

Keywords: SARS-CoV-2, COVID-19, Sanger-sequencing, lineages, variants of concern, variants of interest.

Posted Date: November 19th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1072564/v1

Abstract

Background: The high infectivity rates of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and the prolonged duration of coronavirus disease 2019 (COVID-19) pandemics have contributed to the emergence of viral variants endowed with evolutionary advantages, leading to enhanced infectivity. The tracking of these lineages is urgent. However, the need to sequence whole-viral genomes through next-generation sequencing (NGS) represents a barrier hampering the massive identification of these variants. Therefore, in the current study, we developed Sanger-sequencing approaches targeting regions of interest containing vast lineage-defining mutations in the SARS-CoV-2 S gene and ORF8 region, allowing for unambiguous identification of all SARS-CoV-2 variants of concern (VOCs) and of interest (VOIs).

Methods and results: Primers were designed for polymerase chain reaction (PCR) and nested-PCR to amplify and sequence samples with a low-viral burden. The primers' annealing sites conservancy were checked in a large group of sequences. Amplification protocols were standardized, and sequencing reactions were performed in a cohort of samples for validation. The primers were highly efficient and sequencing of the targeted regions matched those generated by NGS in the same samples. The sequencing results allowed for the unambiguous identification of B.1.1.7, P.1 and P.2 samples, and would also allow the identification of B.1.617.2, B.1.351 and B.1.427/B.1.429 lineages, which were absent in our cohort.

Conclusion: Implementing Sanger-sequencing-based approaches to identify SARS-CoV-2 lineages may represent an alternative to tracking these variants by more laboratories around the world and providing valuable molecular and epidemiologic information to inform health systems.

Introduction

The coronavirus disease 2019 (COVID-19), caused by the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), was firstly notified in Wuhan (China) at the end of 2019 [1], and has rapidly spread around the world. In March 2020, with more than 118 thousand cases in 114 countries, the World Health Organization (WHO) declared the outbreak a pandemic [2]. There are more than 200 million cases and around 4,5 million deaths worldwide from the disease [3].

The prolonged duration of the pandemic and the high infection rates by the virus contributed to the emergence of new variants with potential evolutionary advantages, resulting in increased infectivity, which represents a challenge for COVID-19 control, even in highly vaccinated countries [4-6].

Variants of concern (VOCs) that emerged so far include the B.1.1.7 (α) lineage, firstly detected in the United Kingdom [7,8], the B.1.351 (β) lineage, detected for the first time in South Africa [9], the P.1 (γ) variant, firstly reported in Brazil [10] and the B.1.617.2 (δ), firstly identified in India [11]. Also, variants of interest (VOIs) were described, including the P.2 (ζ) lineage, which was firstly described in Brazil [12] and B.1.621 (μ), firstly detected in Colombia [13]. Other variants that harbor mutations suspected to impact disease course are classified as variants under monitoring (VUM) by the WHO; these currently include the

B.1.617.1 (κ), B.1.427/B.1.429 (ε), B.1.526 (ι), B.1.525 (η) (https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/)

The efficient identification of these variants is crucial to implement measures aiming to control their spreading [14]. However, economic and technological barriers in sequencing complete viral genomes through next-generation sequencing (NGS) hampers its practical implementation in many countries and communities as a surveillance method [15-17].

In this context, Sanger sequencing of regions of interest (ROIs) in the viral genome that include lineage-defining mutations emerges as a suitable strategy that may allow more laboratories around the world to track these mutations in SARS-CoV-2 genomes of diagnosed individuals, contributing with relevant molecular and epidemiological information [15,18, 19].

Therefore, the present study describes the development and validation of Sanger-sequencing methods allowing the identification of VOCs, VOIs and VUMs by targeting ROIs in the Spike (S) and ORF7/ORF8/N genes, enabling the unambiguous characterization of B.1.1.7 (α), B.1.351 (β), B.1.617.2 (δ), P.1 (γ), P.2 (ζ), B.1.621 (μ), B.1.617.1 (κ), B.1.427/B.1.429 (ε), B.1.526 (ι) and B.1.525 (η) lineages.

Methods

Samples

Samples for the present work were selected from an epidemiological genomic surveillance project of SARS-CoV-2 lineages in Paraná state, Brazil. Nasopharyngeal swab samples were collected from patients with active SARS-CoV-2 infection and stored in viral transport media (VTM) under -80ºC. RNA was extracted from VTM using the QIAmp Viral RNA mini kit (QIAGEN®, USA) in QIAcube equipment and quantified in a NanoDrop® One spectrophotometer (ThermoFisher Scientific®) and a Qubit™ Fluorometer (Invitrogen™) using the Qubit RNA HS Assay kit (Invitrogen™).

Libraries for viral genome sequencing were constructed using the AmpliSeq™ for Illumina SARS-CoV-2 research panel (Illumina®, Cat. 20020496), and sequencing was performed in a NovaSeq™ 6000 system (Illumina, USA). The PANGOLin lineage assignment tool (https://cov-lineages.org/pangolin.html) was used for lineage classification.

Samples belonging to P.1, P.2, and B.1.1.7 lineages carrying the mutations of interest in S, ORF7, ORF8, and N genes were selected to validate the technique developed in the current study. Samples belonging to the B.1.1, B.1.1.28, and B.1.1.33 not carrying the mutations of interest were also selected as wild-type samples for the regions of interest, serving as negative controls. For all selected samples, cDNA was newly synthesized from -80ºC-stored RNA samples using the High-Capacity cDNA reverse transcription kit (Applied Biosystems™, USA, Cat. 4368814) following the manufacturer's instructions.

Primer design

Polymerase chain reaction (PCR) primers were designed to target ROIs in S and ORF7a/ORF8/N genes using the primer-BLAST tool. Nested-PCR protocols were planned with two primer pairs designed for each ROI to amplify and sequence samples with low viral burden. The first (outer) amplifies a fragment of 600-800 bp, and the second (inner) amplifies a fragment of 400-600 bp inside the first amplicon.

Conservation of annealing sites was checked in the NCBI nucleotide bank using the BLAST tool and 159 SARS-CoV-2 newly-generated genomes aligned in MEGA 8.0 software. Primer pairs not targeting regions containing any mutation in the analyzed sequences were tested *in silico* for potential secondary structures using the OligoAnalyzer™ tool (IDT).

For the S gene, selected primer pairs were predicted to amplify a 737 bp fragment spanning from nucleotide 22563 to 23299 of SARS-CoV-2 reference genome in the first PCR round and a fragment of 544 bp from nucleotide 22670 to 23213 in the nested-PCR round (Supplementary table S1). In this setting, both primer pairs flanked a large portion of the S-protein receptor-binding domain (RBD) and included several mutations of interest in the S gene, which combined could differentiate all VOCs and VOIs (Supplementary table S2).

Another primer set was designed to target mutations around the ORF8 region. For the first PCR round, a primer pair amplifying a fragment of 784 bp between nucleotides 27593 and 28377 of the SARS-CoV-2 reference genome was selected. This fragment spanned from codon 67 of ORF7a to codon 31 of N gene, flanking the entire coding regions of ORF7b and ORF8 genes.

The nested-PCR primers amplified a 432 bp fragment from nucleotides 27915 to 28347 of SARS-CoV-2 genome (supplementary table S1), spanning from codon 8 of ORF8 gene to codon 25 of N gene. The regions flanked by both inner and outer primers included several lineage-defining mutations, allowing for unambiguous identification of all VOCs and VOIs (Supplementary table S3).

Polymerase chain reaction

For PCR standardization, three viral samples were randomly selected. Amplification protocol was established by testing two $MgCl_2$ concentrations (1mM and 1.5mM) and four annealing temperatures (56°C, 58°C, 60°C, and 62°C) in a Veriti™ 96 well thermal cycler (Applied biosystems™). All PCR reagents were purchased from Invitrogen®, including the recombinant *Taq* DNA polymerase (5U/µL) as well as the accompanying 10X PCR Buffer and 50mM $MgCl_2$ solutions (Cat. 10342020), the 100mM dNTP set (Cat. 10297018), and all the primers. Amplified fragments were analyzed after electrophoresis in 2% agarose gels stained with SYBR Safe® (Invitrogen™), run with Tris-Acetate-EDTA (TAE) buffer. Gels were visualized using an iBright system (Applied biosystems™).

Primers were highly specific, amplifying a single fragment with the predicted size for any tested $MgCl_2$ concentrations or annealing temperatures. The best amplification for the S and ORF7/ORF8/N fragments was achieved with 1.5mM $MgCl_2$ and at 62°C of annealing temperature.

Therefore, all PCRs were performed in 25µL of the reaction mix, setting the final reagent concentrations as follows: 1X PCR Buffer, 1.5mM $MgCl_2$, 0.1µM of each primer, 0.1mM of each dNTP, and 1U of *Taq* DNA polymerase. To the nested protocol, 2µL of cDNA was used in the first PCR round, and 1µL of the PCR product was added in the second PCR reaction. Ultrapure water was used to complete the final volume. Temperature cycling consisted of initial denaturation at 95°C for 5 minutes; 40 cycles of denaturation at 95°C for 30 seconds, annealing at 62°C for 30 seconds and extension at 72°C for 45 seconds, and a final extension at 72°C for 7 minutes.

Sanger sequencing

For Sanger sequencing, PCR fragments were amplified using the first primer pairs for both S and ORF7/ORF8/N regions of interest using Platinum™ *Taq* DNA High Fidelity (Invitrogen™), and amplicons were sliced and purified from agarose gels using the Wizard® SV Gel and PCR Clean-Up System (Promega™). Gel-purified PCR products were quantified in a NanoDrop spectrophotometer (ThermoFisher Scientific™) and diluted to 4ng/µL.

PCR products were sequenced for forward and reverse strands using the BigDye™ Terminator v3.1 Cycle Sequencing Kit (Applied biosystems™) with nested-PCR inner primers. The reaction was performed in a final volume of 10µL using 5µL of purified PCR product at 4ng/µL, 2µL of 2.5µM primer (forward or reverse), 1µL of 5X BigDye Sequencing Buffer, and 2µL BigDye™ Terminator v3.1 Ready Reaction Mix, following the manufacturer's recommendations.

The reactions took place in a Veriti™ thermal cycler using the universal BigDye™ protocol: 96ºC for 1 minute, followed by 35 cycles of 15 seconds at 96°C, 15 seconds at 50°C, and 4 minutes at 60°C; importantly, the decrease from 96°C to 50°C was done at a rate of 1°C per second, as recommended by the sequencing kit supplier (Applied biosystems™).

Sequencing amplicons were purified using the BigDye XTerminator™ Purification kit (Applied Biosystems™, Cat. 4376486) following the manufacturer's instructions and submitted to capillary electrophoresis in a SeqStudio Genetic analyzer (Applied Biosystems™) using the medium run protocol with default parameters. Chromatograms were inspected using BioEdit 7.2 software.

Results

S-gene region of interest sequencing

Ten samples from different lineages were sequenced for the S-gene region of interest. All samples were amplified using the outer primers for the sequencing reaction, and sequencing reactions were performed on amplified fragments using the nested PCR inner primers. The sequenced samples included two samples from P.1 (γ) lineage, two samples classified as P.2 (ζ), one sample classified as B.1.1.7 (α), one sample characterized as N.9 and four samples classified as B.1.1.28. Interestingly, one of the B.1.1.28

samples harbored the K417T (22812 A>C), E484K (23012 G>A) and N501Y (23063 A>T) mutations in S gene, which are characteristics of P.1, according to the FASTA sequence for the whole viral genome. This sample is referred to as P.1-like from here throughout the text.

All sequenced fragments were concordant with the whole genomes sequenced through NGS, and the fragments correctly displayed the mutations of interest characteristic of each lineage: The B.1.1.7 lineage sample showed an A>T transversion in nucleotide 372 from forwarding inner primer start corresponding to N501Y and a C>A transversion in nucleotide 580, corresponding to A570D mutation. P.2 samples showed a G>A transition in nucleotide 321 corresponding to the E484K mutation. Lastly, P.1 and P.1-like samples showed an A>C transversion in nucleotide 121 corresponding to K417T mutation in addition to E484K and N501Y mutations. Figure 1 shows a representative chromatogram for a P.1 sample, harboring the three mutations mentioned above.

No mutation was found in classical B.1.1.28 and N.9 lineages. We did not have any sample belonging to other VOCs, VOIs and VUMs in our cohort, however these would be characterized by mutations depicted in supplementary table S2 in corresponding positions of the sequenced fragments.

ORF8-gene region of interest sequencing

For the ORF8 region of interest, 23 samples were sequenced, including the following lineages: B.1.1.28 (7, including the P.1-like sample), B.1.1.33 (2), B.1.1 (1), N.9 (1), B.1.1.7 (α, 3), P.1 (γ, 3) and P.2 (ζ, 6). Notably, according to the FASTA for the whole-genome sequence, the P.1-like sample lacked the P.1-defining E92K mutation in ORF8.

Again, all Sanger-generated sequences matched those generated using NGS for this region, confirming the method's accuracy. For B.1.1.7 samples, the following mutations could be detected: a C>T transition in position 35 from inner forward primer start site corresponding to Q27* ORF8 mutation; a G>T transversion, corresponding to R52I ORF8 mutation; an A>G transition representing the Y73C mutation; an adenine deletion in position 334, in a noncoding region of SARS-CoV-2 genome, between ORF8 and N genes, and, finally a three-nucleotide mutation from GAT to CAT in position 343-345, corresponding to the D3L mutation in N gene (Figure 2).

P.1 samples showed a G>A transition in nucleotide 230 from inner primer, representing E92K mutation in ORF8, while P.2 samples showed a C>T transversion in nucleotide 316, characterizing the silent mutation in nucleotide 28253 of SARS-CoV-2 genome (F120-) (Figure 3). Importantly, the lack of E92K mutation in the P.1-like sample was confirmed in the current protocol. This fact argues against the hypothesis of this lacking being derived from a sequencing error in NGS and suggests that this sample might be an evolutionary intermediate between B.1.1.28 and P.1 lineages, harboring some but not all P.1-defining mutations. Supplemental table S3 depict mutations in the targeted region that characterize other VOCs, VOIs and VUMs that are not included in our sample.

Discussion

Since its' emergency in late 2019 in the province of Wuhan, China, COVID-19 has rapidly spread through the world, being declared a pandemic in less than six months from its first report. Widespread transmission and the prolonged period that the pandemic has lasted until now, new SARS-CoV-2 lineages have emerged, some of which harbor mutations conferring them evolutionary advantages, resulting in increased transmissibility and predominance in communities where they are found [4].

According to the level of evidence, the WHO classifies these lineages as VOIs and VOCs according to the level of evidence that they result in increased transmissibility, virulence or are associated with a decrease in the effectiveness of public health measurements, diagnostic tests, vaccines, or therapeutics (https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/, access on 7th November 2021). Given their epidemiological relevance, the early detection and tracking of these variants are crucial to inform health systems worldwide, allowing the adoption of adequate measurements to contain their spread [14]. The current tools for lineage-assignment, such as the PANGOlin web resource, requires whole-genome sequencing of infecting SARS-CoV-2. Despite the technological advances which decreased the time and costs of generating whole-genomes through NGS in recent years, the implementation of these technologies beyond big centers and reference laboratories is still a challenge due to the technological and economic barriers as well as to the need for specialized and well-trained staff for sample processing, data acquisition, and analysis. Therefore, studies suggested Sanger-sequencing of ROIs in the SARS-Cov-2 genome, mainly in the S gene, as a feasible method to identify VOCs and VOIs [15,18, 19].

In the current work, we provide an additional set of primers allowing the amplification and sequencing of regions of interest in S and ORF7/ORF8/N genes containing informative mutations that characterize all the currently listed VOIs and VOCs. Further, two additional primer pairs were designed and validated for each region for use in nested-PCR protocols, allowing the robust amplification of both regions of interest even for samples with a low-viral burden. Importantly, all primer pairs were checked for annealing site conservancy in a large number of samples, and the absence of mutations suggests that highly conserved regions were targeted, diminishing the possibility of amplification failing.

Since Sanger-sequencing is widespread and requires fewer resources than NGS, the approach described in the current study may assist the identification of VOCs and VOIs by a more significant number of diagnostic and research laboratories serving health systems with valuable information regarding the spread of these variants. Furthermore, Sanger-sequencing has advantages over other non-NGS-based techniques aiming to identify SARS-CoV-2 mutations and lineages, such as probe-based qRT-PCR methods [20], including the identifications of several mutations in the targeted region in a single run, allowing for SARS-CoV-2 haplotyping and lineage assignment, and the possibility of discovering new mutations in that region.

In conclusion, the present work provides alternative methods based on PCR-sequencing to identify informative mutations and classify SARS-CoV-2 VOIs and VOCs. The implementation of these methods by more laboratories may help to overcome the technical and economic bottlenecks involved in the identification of these variants through NGS and provide a more realistic picture of their emergence and

spreading throughout geographical regions that are currently underrepresented in SARS-CoV-2 whole-genomes databanks, allowing for more effective public policies for their contention.

Declarations

Authors contributions

G.A.F.V.: Conceptualization, Investigation, Methodology, Formal analyses, Data Curation, Visualization, Writing – Original Draft; **I.M.O:** Investigation, Methodology, Data Curation, Supervision, Project administration, Writing – Review & editing; **F.I.:** Investigation, Methodology, Writing – Original Draft; **B.L.F.:** Investigation, Methodology, Writing – Original Draft; **E.C.G.:** Data curation, Formal analyses, Software, Writing – Review & editing; **R.S.B.:** Data curation, Formal analyses, Software; **C.E.B.:** Data curation, Software, Formal Analyses; **J.P.B.X.:** Formal analyses, Software; **B.M.P.:** Investigation, Methodology; **K.O.F.G.:** Investigation, Methodology; **C.A.O.B.Jr:** Data curation Resources, Software; **Omics working group - NAPI Genomics:** Conceptualization, Resources, Funding acquisition, Writing – Review & Editing, Supervision; **D.P.S.S.:** Supervision, Funding acquisition, Project administration, Writing – Review & Editing; **E.C.:** Supervision, Funding acquisition, Project administration; **K.C.P.:** Methodology, Investigation, Supervision, Funding acquisition, Project administration; **D.L.A.F.:** Supervision, Funding acquisition, Resources, Project administration; **W.A.S.Jr:** Conceptualization, Methodology, Data curation, Supervision, Funding acquisition, Project administration, Resources, Project administration, Writing – Review & editing.

Acknowledgements

This project was supported by research grants from the Araucária Foundation for support of scientific and technological development of Paraná State (FAP-PR), the Paraná State Secretariat of Science, Technology and Higher Education (SETI-PR), the Guarapuava City Council, The National Council for Scientific and Technological Development (CNPq) and by the São Paulo Research Foundation (FAPESP). G.A.F.V., B.L.F, B.M.P and K.O.F.G. received research scholarships from the Araucária Foundation (FAP-PR). F.I. received research scholarship from the Brazilian Coordination of Superior Level Staff Improvement (CAPES). We would like to thank the Presidents of FAP-PR Prof. Dr. Ramiro Wahrhaftig and the SETI-PR, Prof. Dr. Aldo Nelson Bona, for institutional support. We also would like to thank the members of the NAPI-Genomics Consortium, who worked hard in the COVID-19 Genome Project of the Paraná State.

Funding

This work was financed by Fundação Araucária/Superintendência de Ciência, Tecnologia e Ensino Superior do Paraná (SETI), Grant 0035/2020; Guarapuava City Council, Grant No. 003/2020; The National Council for Scientific and Technological Development (CNPq), grant 465539/2014-9 and São Paulo Research Foundation (FAPESP), grants 2013/08135-2, 2017/23205-8, 2019/08528-0, and 2020/02939-6.

Conflicts of interest

None.

Availability of data and material

Consensus genomes generated in the current project used for primer design have been submitted to the GISAID database under accession numbers EPI_ISL_2758648-2758806. Other data are available from the corresponding author upon reasonable request.

Code availability

Not applicable.

Ethics approval

The study was approved by the ethics committee for research involving human beings of the Midwestern Paraná State University (CAAE: 37660820.8.1001.0106)

Consent to participate

Signed free-informed consent was obtained from all individual participants included in the study prior to biological material or clinical data collection.

Consent for publication

Not applicable.

References

1. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, Zhao X, Huang B, Shi W, Lu R, Niu P, Zhan F, Ma X, Wang D, Xu W, Wu G, Gao GF, Tan W (2020) A Novel Coronavirus from Patients with Pneumonia in China, 2019. New England Journal of Medicine 382 (8):727-733. doi:10.1056/NEJMoa2001017
2. Ciotti M, Ciccozzi M, Terrinoni A, Jiang W-C, Wang C-B, Bernardini S (2020) The COVID-19 pandemic. Critical Reviews in Clinical Laboratory Sciences 57 (6):365-388. doi:10.1080/10408363.2020.1783198
3. WHO (2021). "Weekly epidemiological update on COVID-19 - 2 November 2021.World Health Organization (WHO).https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19—2-november-2021. Accessed 07 November 2021.
4. Mascola JR, Graham BS, Fauci AS (2021) SARS-CoV-2 Viral Variants—Tackling a Moving Target. Jama 325 (13):1261. doi:10.1001/jama.2021.2088
5. van Oosterhout C, Hall N, Ly H, Tyler KM (2021) COVID-19 evolution during the pandemic – Implications of new SARS-CoV-2 variants on disease control and public health policies. Virulence 12 (1):507-508. doi:10.1080/21505594.2021.1877066

6. Wadman M (2021) Israel's grim warning: Delta can overwhelm shots. Science 373 (6557):838-839. doi:10.1126/science.373.6557.838
7. Frampton D, Rampling T, Cross A, Bailey H, Heaney J, Byott M, Scott R, Sconza R, Price J, Margaritis M, Bergstrom M, Spyer MJ, Miralhes PB, Grant P, Kirk S, Valerio C, Mangera Z, Prabhahar T, Moreno-Cuesta J, Arulkumaran N, Singer M, Shin GY, Sanchez E, Paraskevopoulou SM, Pillay D, McKendry RA, Mirfenderesky M, Houlihan CF, Nastouli E (2021) Genomic characteristics and clinical effect of the emergent SARS-CoV-2 B.1.1.7 lineage in London, UK: a whole-genome sequencing and hospital-based cohort study. The Lancet Infectious Diseases. doi:10.1016/s1473-3099(21)00170-5
8. Wise J (2020) Covid-19: New coronavirus variant is identified in UK. Bmj:m4857. doi:10.1136/bmj.m4857
9. Tegally H, Wilkinson E, Giovanetti M, Iranzadeh A, Fonseca V, Giandhari J, Doolabh D, Pillay S, San EJ, Msomi N, Mlisana K, von Gottberg A, Walaza S, Allam M, Ismail A, Mohale T, Glass AJ, Engelbrecht S, Van Zyl G, Preiser W, Petruccione F, Sigal A, Hardie D, Marais G, Hsiao N-y, Korsman S, Davies M-A, Tyers L, Mudau I, York D, Maslo C, Goedhals D, Abrahams S, Laguda-Akingba O, Alisoltani-Dehkordi A, Godzik A, Wibmer CK, Sewell BT, Lourenço J, Alcantara LCJ, Kosakovsky Pond SL, Weaver S, Martin D, Lessells RJ, Bhiman JN, Williamson C, de Oliveira T (2021) Detection of a SARS-CoV-2 variant of concern in South Africa. Nature 592 (7854):438-443. doi:10.1038/s41586-021-03402-9
10. Faria NR, Mellan TA, Whittaker C, Claro IM, Candido DdS, Mishra S, Crispim MAE, Sales FCS, Hawryluk I, McCrone JT, Hulswit RJG, Franco LAM, Ramundo MS, de Jesus JG, Andrade PS, Coletti TM, Ferreira GM, Silva CAM, Manuli ER, Pereira RHM, Peixoto PS, Kraemer MUG, Gaburo N, Camilo CdC, Hoeltgebaum H, Souza WM, Rocha EC, de Souza LM, de Pinho MC, Araujo LJT, Malta FSV, de Lima AB, Silva JdP, Zauli DAG, Ferreira ACdS, Schnekenberg RP, Laydon DJ, Walker PGT, Schlüter HM, dos Santos ALP, Vidal MS, Del Caro VS, Filho RMF, dos Santos HM, Aguiar RS, Proença-Modena JL, Nelson B, Hay JA, Monod M, Miscouridou X, Coupland H, Sonabend R, Vollmer M, Gandy A, Prete CA, Nascimento VH, Suchard MA, Bowden TA, Pond SLK, Wu C-H, Ratmann O, Ferguson NM, Dye C, Loman NJ, Lemey P, Rambaut A, Fraiji NA, Carvalho MdPSS, Pybus OG, Flaxman S, Bhatt S, Sabino EC (2021) Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. Science 372 (6544):815-821. doi:10.1126/science.abh2644
11. Dhar MS, Marwal R, Vs R, Ponnusamy K, Jolly B, Bhoyar RC, Sardana V, Naushin S, Rophina M, Mellan TA, Mishra S, Whittaker C, Fatihi S, Datta M, Singh P, Sharma U, Ujjainiya R, Bhatheja N, Divakar MK, Singh MK, Imran M, Senthivel V, Maurya R, Jha N, Mehta P, A V, Sharma P, Vr A, Chaudhary U, Soni N, Thukral L, Flaxman S, Bhatt S, Pandey R, Dash D, Faruq M, Lall H, Gogia H, Madan P, Kulkarni S, Chauhan H, Sengupta S, Kabra S, Gupta RK, Singh SK, Agrawal A, Rakshit P, Nandicoori V, Tallapaka KB, Sowpati Divya T, Thangaraj K, Bashyam MD, Dalal A, Sivasubbu S, Scaria V, Parida A, Raghav SK, Prasad P, Sarin A, Mayor S, Ramakrishnan U, Palakodeti D, Seshasayee ASN, Bhat M, Shouche Y, Pillai A, Dikid T, Das S, Maitra A, Chinnaswamy S, Biswas Nidhan K, Desai AS, Pattabiraman C, Manjunatha MV, Mani Reeta S, Arunachal Udupi G, Abraham P,

Atul Potdar V, Cherian SS (2021) Genomic characterization and epidemiology of an emerging SARS-CoV-2 variant in Delhi, India. Science. doi:10.1126/science.abj9932

12. Voloch CM, da Silva Francisco R, de Almeida LGP, Cardoso CC, Brustolini OJ, Gerber AL, Guimarães APdC, Mariani D, da Costa RM, Ferreira OC, Cavalcanti AC, Frauches TS, de Mello CMB, Leitão IdC, Galliez RM, Faffe DS, Castiñeiras TMPP, Tanuri A, de Vasconcelos ATR, Parrish CR (2021) Genomic Characterization of a Novel SARS-CoV-2 Lineage from Rio de Janeiro, Brazil. Journal of Virology 95 (10). doi:10.1128/jvi.00119-21
13. Laiton-Donato K, Franco-Muñoz C, Álvarez-Díaz DA, Ruiz-Moreno HA, Usme-Ciro JA, Prada DA, Reales-González J, Corchuelo S, Herrera-Sepúlveda MT, Naizaque J, Santamaría G, Rivera J, Rojas P, Ortiz JH, Cardona A, Malo D, Prieto-Alvarado F, Gómez FR, Wiesner M, Martínez MLO, Mercado-Reyes M (2021) Characterization of the emerging B.1.621 variant of interest of SARS-CoV-2. Infection, Genetics and Evolution 95:105038. doi:10.1016/j.meegid.2021.105038
14. Cyranoski D (2021) Alarming COVID variants show vital role of genomic surveillance. Nature 589 (7842):337-338. doi:10.1038/d41586-021-00065-4
15. Bezerra MF, Machado LC, De Carvalho VdCV, Docena C, Brandão-Filho SP, Ayres CFJ, Paiva MHS, Wallau GL (2021) A Sanger-based approach for scaling up screening of SARS-CoV-2 variants of interest and concern. Infection, Genetics and Evolution 92:104910. doi:10.1016/j.meegid.2021.104910
16. Brito AF, Semenova E, Dudas G, Hassler GW, Kalinich CC, Kraemer MUG, Hill SC, Danish Covid-19 Genome C, Sabino EC, Pybus OG, Dye C, Bhatt S, Flaxamn S, Suchard MA, Grubaugh ND, Baele G, Faria NR (2021) Global disparities in SARS-CoV-2 genomic surveillance. medRxiv. doi:10.1101/2021.08.21.21262393
17. Liu X, Liu C, Liu G, Luo W, Xia N (2020) COVID-19: Progress in diagnostics, therapy and vaccination. Theranostics 10 (17):7821-7835. doi:10.7150/thno.47987
18. Anaclerio F, Ferrante R, Mandatori D, Antonucci I, Capanna M, Damiani V, Tomo PD, Ferrante R, Ranaudo M, De Laurenzi V, Stuppia L, De Fabritiis S (2021) Different Strategies for the Identification of SARS-CoV-2 Variants in the Laboratory Practice. Genes 12 (9). doi:10.3390/genes12091428
19. Daniels RS, Harvey R, Ermetal B, Xiang Z, Galiano M, Adams L, McCauley JW (2021) A Sanger sequencing protocol for SARS-CoV-2 S-gene. Influenza and Other Respiratoy Viruses 15 (6):707-710. doi: 10.1111/irv.12892
20. Sugden B, Vogels CBF, Breban MI, Ott IM, Alpert T, Petrone ME, Watkins AE, Kalinich CC, Earnest R, Rothman JE, Goes de Jesus J, Morales Claro I, Magalhães Ferreira G, Crispim MAE, Singh L, Tegally H, Anyaneji UJ, Hodcroft EB, Mason CE, Khullar G, Metti J, Dudley JT, MacKay MJ, Nash M, Wang J, Liu C, Hui P, Murphy S, Neal C, Laszlo E, Landry ML, Muyombwe A, Downing R, Razeq J, de Oliveira T, Faria NR, Sabino EC, Neher RA, Fauver JR, Grubaugh ND (2021) Multiplex qPCR discriminates variants of concern to enhance global surveillance of SARS-CoV-2. PLOS Biology 19 (5):e3001236. doi:10.1371/journal.pbio.3001236

Figures

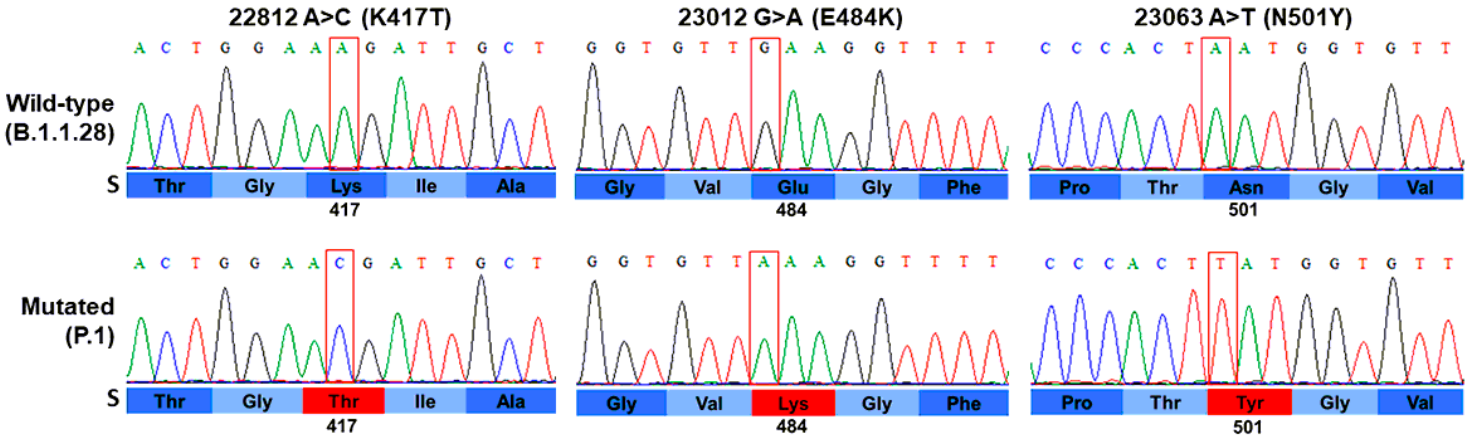

Figure 1

Identification of S gene mutations characterizing P.1 (γ) lineage samples. Lower panels show chromatograms corresponding to the P.1 lineage, identified by the presence of 22812 A>C (K417T), 23012 G>A (E484K) and 23063 (N501Y) mutations, while upper panels are from a B.1.1.28 lineage sample, which is representative for wild-type lineage samples.

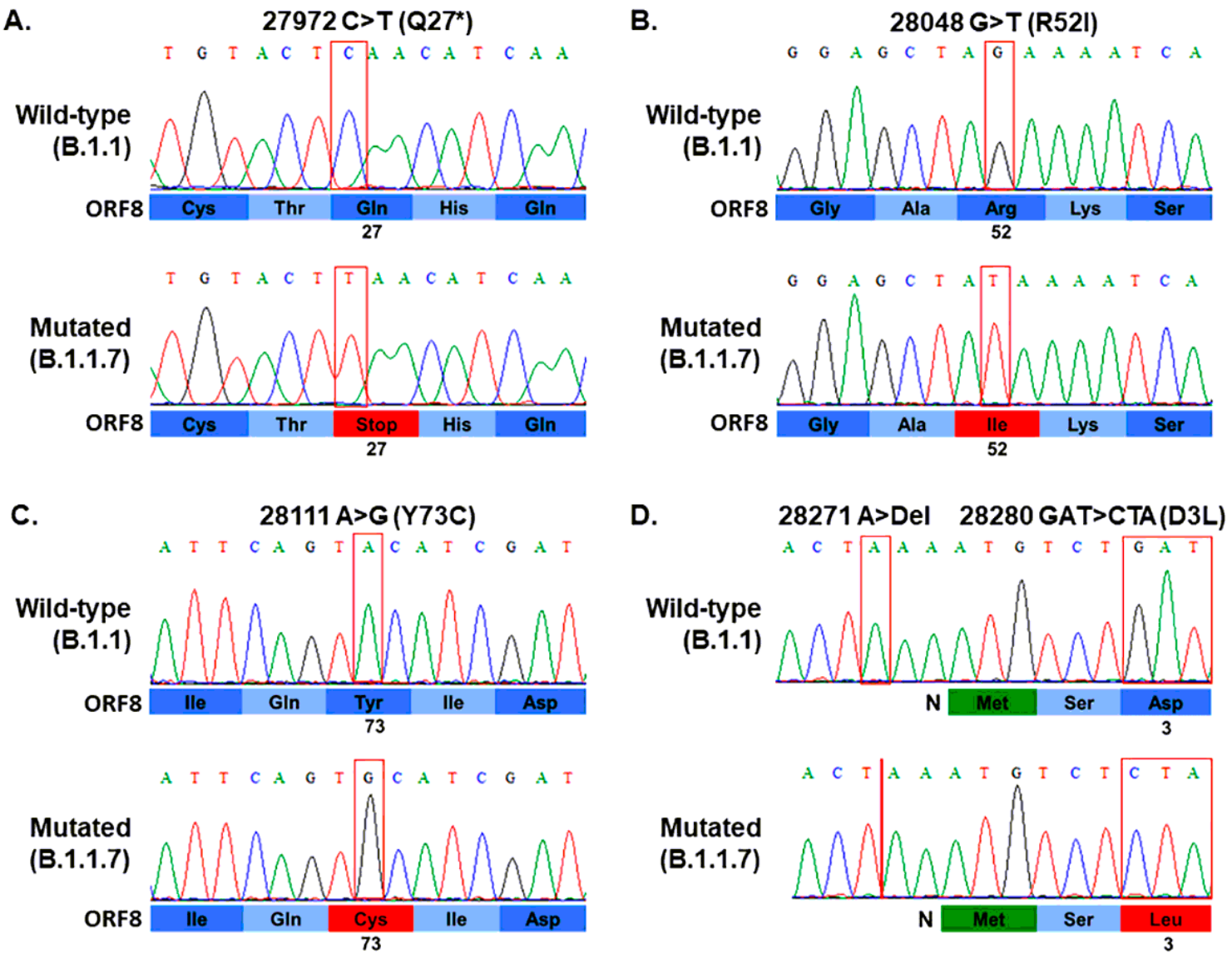

Figure 2

Identification of ORF8/N region mutations characterizing B.1.1.7 (α) lineage samples. A. Identification of 27972 C>T (Q27*) ORF8 mutation. B. Identification of 28048 G>T (R52) ORF8 mutation. C. Identification of 28111 A>G (Y73C) mutation. D. Identification of the 28271 A deletion and of 28280-28282 GAT>CTA (D3L) N gene mutation. For each panel (A, B, C and D), upper chromatograms are representative of a wild-type sample (B.1.1 lineage) and lower chromatograms are representative of B.1.1.7 lineages.

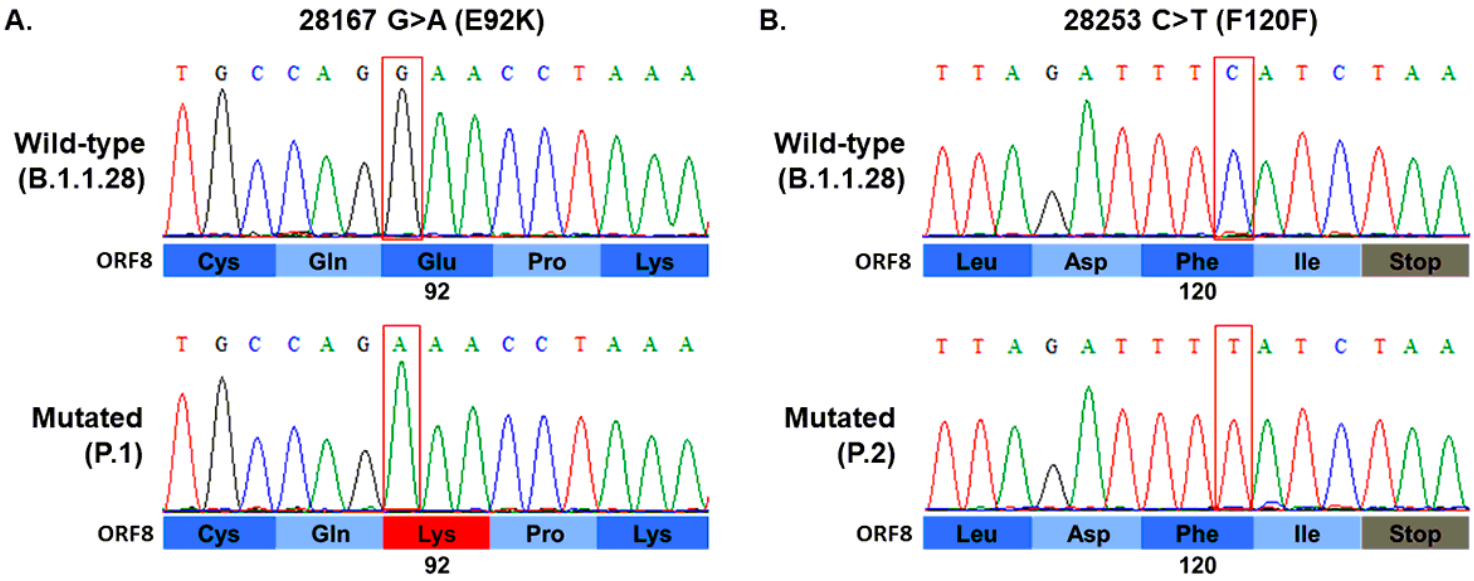

Figure 3

Identification of ORF8 mutations characterizing P.1 (γ) and P.2 (ζ) lineage samples. A. Identification of the 28167 G>A (E92K) ORF8 mutation, characteristic of P.1 (γ) lineage. B. Identification of the 28253 C>T ORF8 silent mutation, characteristic of P.2 (ζ) lineage. For each panel (A and B), upper chromatograms are representative of a wild-type sample (B.1.1.28 lineage) and lower chromatograms are representative of mutated lineages (P.1 and P.2, respectively).

Supplementary Files

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

- Supplementaryinformation.docx
- AppendixA.docx