

## RELATÓRIO TÉCNICO 2021/05

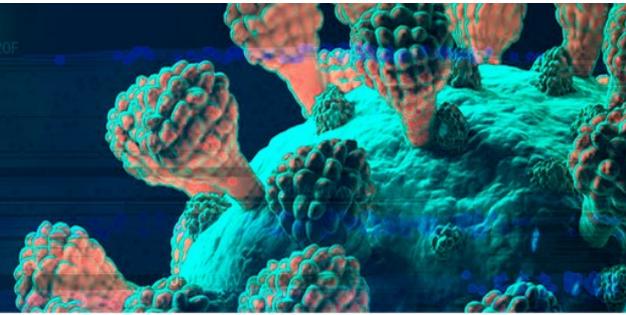
### REDE GENÔMICA FIOCRUZ / MINISTÉRIO DA SAÚDE

#### Identificação de linhagens relacionadas ao SARS-CoV-2 linhagem P.1 no Brasil

A emergência da variante de preocupação (VOC) de SARS-CoV-2 P.1, cujo o surgimento foi estimado em novembro de 2020 no estado do Amazonas <sup>1,2</sup>, teve uma rápida disseminação para outros estados brasileiros associada à segunda onda epidêmica que colapsou o sistema público de saúde no início de 2021. A VOC P.1 possui um grande número de mutações definidoras da linhagem, incluindo: 10 substituições não sinônimas na proteína *Spike* (S) (L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, H655Y, T1027I), cinco mutações não sinônimas distribuídas entre a NSP3 (S370L e K977Q), NSP13 (E341D); e as proteínas NS8 (E92K) e N (P80R). Mutações do tipo indel também são encontradas em NSP6 (S106del, G107del, F108del) e na região intergênica ORF8/N, com uma inserção de quatro nucleotídeos (ins28263) <sup>1,2</sup>.

Uma análise subsequente de um grande número de genomas do SARS-CoV-2 no Amazonas revelou quatro sequências semelhantes à P.1 de pacientes de Manaus, que se ramificam como um clado monofilético irmão em relação à linhagem P.1 <sup>1,3</sup>. O clado definido como P.1-like também acumulou um número excepcionalmente alto de mudanças genéticas em relação à linhagem B.1.1.28 parental, incluindo várias mutações definidoras de linhagem P.1 no S (L18F, P26S, D138Y, K417T, E484K, N501Y) e nas proteínas NSP3 (K977Q) e N (P80R), assim como mutações únicas nas proteínas NSP2 (K456R); NSP3 (T1189I); NSP6 (V149A); NSP13 (S74L); S (ins214 e D1139H) e NS8 (K2stop). Este achado sugere que mutações definidoras de linhagem P.1 não se acumularam em um evento único de diversificação genética em uma infecção de longo prazo, mas foram adquiridas em etapas sequenciais durante a evolução da linhagem B.1.1.28 no estado do Amazonas.

Neste relatório técnico, descrevemos uma segunda variante de SARS-CoV-2 relacionada à P.1 (P.1-like) que está se espalhando em vários estados das diferentes regiões brasileiras. A análise do perfil de mutações de amostras positivas para SARS-CoV-2 detectadas em diferentes estados brasileiros entre 12 de março de 2020 e 31 de março de 2021 revelou 44 sequências (EPI\_ISL\_2038926 ao EPI\_ISL\_2038968, EPI\_ISL\_2102018 e EPI\_ISL\_2102063) que abrigam 15 das 21 mutações definidoras da linhagem P.1, incluindo as três mutações de maior preocupação no domínio de



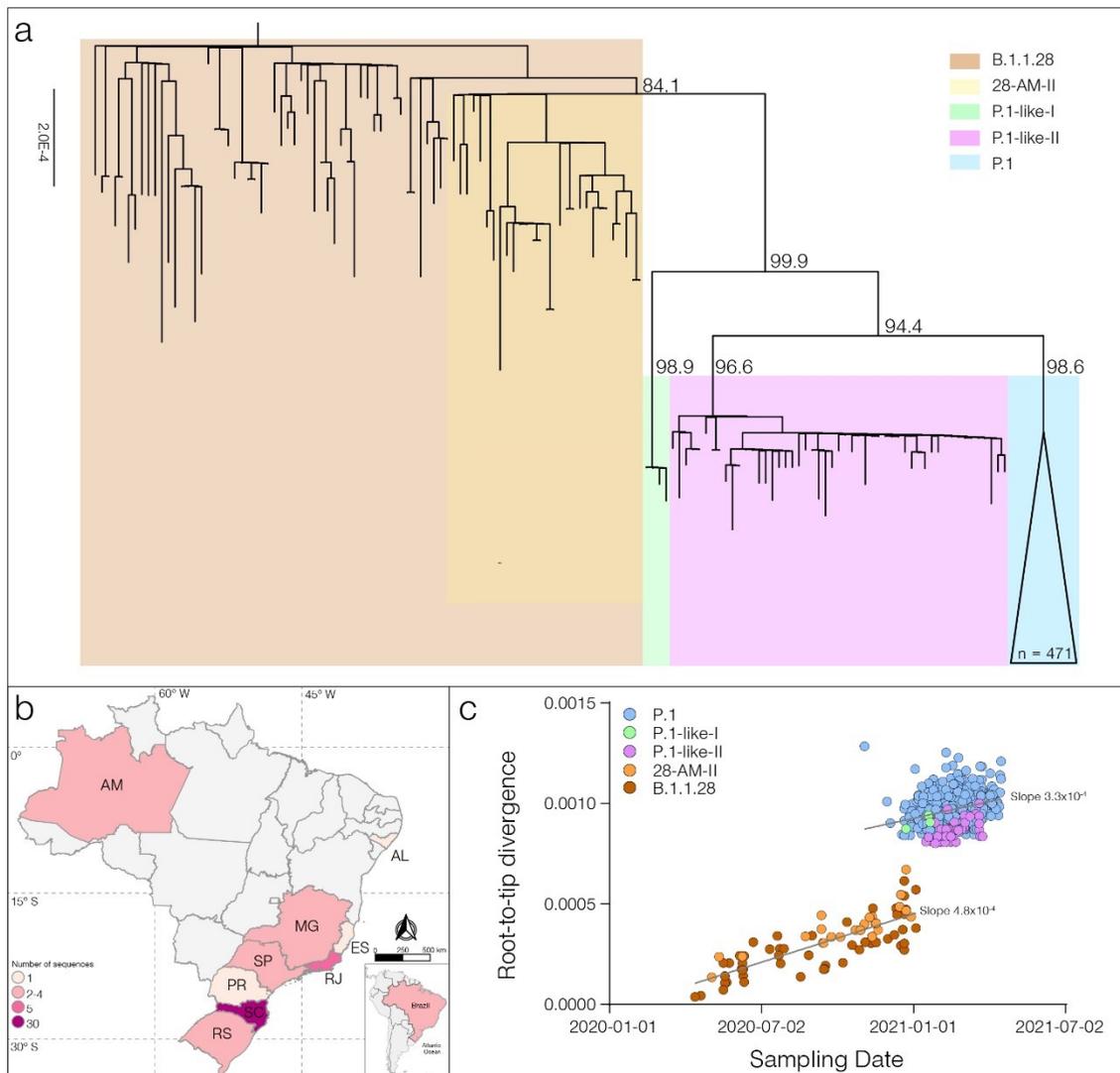
ligação do receptor (RBD) da proteína S (K417T, E484K e N501Y), a deleção na NSP6 (S106del, G107del, F108del) e a inserção de quatro nucleotídeos na região intergênica ORF8/N (ins28263). Estas sequências P.1-*like*, aqui designadas como P.1-*like*-II, não possuem algumas das mutações definidoras da linhagem P.1 em ORF1ab (C2749T, C12778T e C13860T); NSP13 (E341D); S (T20N) e NS8 (E92K) e exibem ainda seis substituições exclusivas na ORF1ab (C8905T, C16954T e A20931G); NSP4 (D217H); região intergênica E/M (A26492T) e N (P383L). As sequências P.1-*like*-II também compartilham nove mutações definidoras de linhagem P.1 com o clado P.1-*like* anteriormente caracterizado (agora designado como P.1-*like*-I).

A análise filogenética de máxima verossimilhança revelou que novas sequências P.1-*like*-II se ramificaram em um clado monofilético com alto suporte (SH-aLRT = 96,6%) junto com sete sequências obtidas do banco de dados EpiCoV (<https://www.gisaid.org/>) que exibiram o perfil de mutação P.1-*like*-II, mas estão atualmente classificadas como P.1 no mesmo banco de dados (**Figura 1a**). Os cladogramas P.1-*like*-I e P.1-*like*-II não estão agrupados dentro da diversidade da VOC P.1, portanto devem ser designados como novas linhagens “P.n” PANGO (a solicitação de designação de linhagem foi enviada em 10 de maio de 2021 <https://github.com/cov-lineages/pango-designation/issues/77>). Os genomas P.1-*like*-II foram amostrados em nove diferentes estados brasileiros, principalmente nas regiões Sul e Sudeste (**Figura 1b**), o mais antigo foi detectado no estado do Rio de Janeiro em 19 de janeiro de 2021 <sup>4</sup>, sendo o mais um recente identificado neste estudo no estado do Amazonas em 25 de março de 2021. O estado brasileiro que contém a maioria das sequências P.1-*like*-II identificadas até o momento foi Santa Catarina (59%), seguido por Rio de Janeiro (10%), Rio Grande do Sul (8%) e São Paulo (8%). A análise da estrutura temporal revelou que os cladogramas P.1 e P.1-*like* acumularam um maior número de mutações do que outras sequências B.1.1.28 (Figura 1c). Os diferentes modelos de relógio molecular usados traçaram consistentemente o tempo médio de surgimento do ancestral de todas as P.1 e P.1-*like* em meados de agosto de 2020 e o surgimento das linhagens P.1 e P.1-*like* no final de novembro e no final de dezembro de 2020, respectivamente.

Nossa vigilância genômica identificou uma nova linhagem irmã de P.1, derivada da diversidade amazônica B.1.1.28, e provisoriamente designada como linhagem P.1-*like*-II. Ela compartilha um ancestral comum e várias mutações definidoras de linhagem com a VOC P.1, incluindo aquelas no RBD da proteína S (K417T, E484K, N501Y). Essa nova linhagem está geograficamente dispersa no Brasil, principalmente nas regiões Sul e Sudeste do país. A P.1-*like*-II é a segunda linhagem relacionada a P.1 identificada por nosso grupo no Brasil, confirmando que a constelação completa de mutações que define o VOC P.1 não se acumulou em um único indivíduo, mas foram adquiridas em etapas sequenciais durante múltiplos eventos de transmissão. Nossos resultados sugerem ainda que um ancestral P.1 carregando metade das mutações definidoras da linhagem P.1, incluindo aquelas no RBD da proteína S, circulou silenciosamente no estado do Amazonas por vários meses antes do surgimento de P.1 e linhagens relacionadas a P.1. Digno de nota, é o fato que embora o ancestral de P.1 e as linhagens contemporâneas P.1-*like* possuam as principais mutações preocupantes na proteína S; nenhuma delas teve tanto sucesso em se espalhar como a VOC P.1. Isso sugere que outros fatores

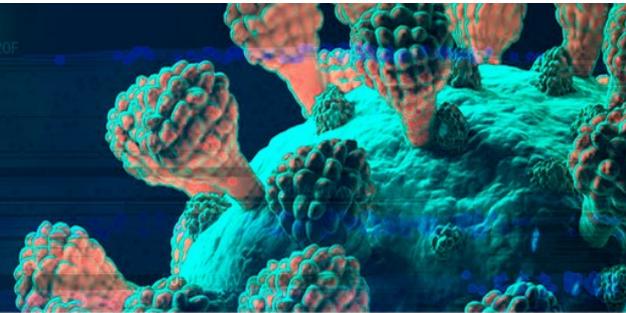
além das mutações virais, provavelmente relacionados ao comportamento humano, podem ter desempenhado um papel importante na disseminação notável da VOC P.1 no estado do Amazonas e depois em todo o Brasil.

**Figura 1.** Diversidade genética e distribuição das linhagens semelhantes a B.1.1.28, P.1 e P.1 no Brasil. **a)** Árvore filogenética de máxima verossimilhança (MV) das linhagens B.1.1.28, P.1 e P.1-*like* identificadas no Brasil. Cada linhagem foi destacada com cores conforme indicado na legenda. Os valores de suporte aLRT são indicados em ramificações principais e os comprimentos de ramificação são desenhados em escala com a barra lateral indicando substituições de nucleotídeos por local. **b)** Distribuição geográfica e frequência da linhagem P.1-*like*-II identificada no Brasil. Os nomes dos estados brasileiros seguem o padrão ISO 3166-2. O gradiente de cor representa o número de sequências identificadas neste estudo, seguindo a legenda. **c)** Correlação entre a data de amostragem de B.1.1.28, P.1, P.1-*like*-I e P.1-*like*-II e sua distância genética da raiz da árvore filogenética MV. Cada linhagem foi colorida de acordo com a legenda. A inclinação de cada regressão está indicada.



## Referências

- 1 Naveca, F. *et al.* COVID-19 epidemic in the Brazilian state of Amazonas was driven by long-term persistence of endemic SARS-CoV-2 lineages and the recent emergence of the new Variant of Concern P.1. *Research Square*, doi:10.21203/rs.3.rs-275494/v1 (2021).
- 2 Faria, N. R. *et al.* Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science*, doi:10.1126/science.abh2644 (2021).
- 3 Resende, P. C. *et al.* The ongoing evolution of variants of concern and interest of SARS-CoV-2 in Brazil revealed by convergent indels in the amino (N)-terminal domain of the Spike protein. *medRxiv*, doi:10.1101/2021.03.19.21253946 (2021).



- 4 Lamarca, A. *et al.* Genomic surveillance of SARS-CoV-2 tracks early interstate transmission of P.1 lineage and diversification within P.2 clade in Brazil. *medRxiv* doi:[10.1101/2021.03.21.21253418](https://doi.org/10.1101/2021.03.21.21253418) (2021).

Este relatório técnico foi elaborado por Tiago Gräf<sup>1A</sup>, Gonzalo Bello<sup>2A</sup>, Taina Moreira Martins Venas<sup>3</sup>, Elisa Cavalcante Pereira<sup>3</sup>, Anna Carolina Dias Paixão<sup>3</sup>; Luciana Reis Appolinario<sup>3</sup>; Renata Serrano Lopes<sup>3</sup>; Ana Carolina da Fonseca Mendonça<sup>3</sup>; Alice Sampaio Barreto da Rocha<sup>3</sup>; Fernando Couto Motta<sup>3</sup>, Tatiana Schäffer Gregianini<sup>4</sup>, Richard Salvato<sup>4</sup>, Sandra Bianchini Fernandes<sup>5</sup>, Darcita Buerger Rovaris<sup>5</sup>, Andrea Cony Cavalcanti<sup>6</sup>, Anderson Brandão Leite<sup>7</sup>, Irina Riediger<sup>8</sup>, Maria do Carmo Debur<sup>8</sup>, André Felipe Leal Bernardes<sup>9</sup>, Rodrigo Ribeiro-Rodrigues<sup>10</sup>, Beatriz Grinsztejn<sup>11</sup>, Filipe Zimmer Dezordi<sup>12,13</sup>, Gabriel Luz Wallau<sup>12,13B</sup>, Felipe Gomes Naveca<sup>14B</sup>, Edson Delatorre<sup>15B</sup>, Marilda Mendonça Siqueira<sup>3B</sup>, and Paola Cristina Resende<sup>3B</sup> on behalf of Fiocruz COVID-19 Genomic Surveillance Network.

<sup>A</sup> Theses authors contributed equally

<sup>B</sup> These authors share the senior authorship

1. Plataforma de Vigilância Molecular, Instituto Gonçalo Moniz, Fiocruz, Salvador, Bahia, Brazil
2. Laboratório de AIDS e Imunologia Molecular, Instituto Oswaldo Cruz, Fiocruz, Rio de Janeiro, Brazil
3. Laboratório de Vírus Respiratórios e do Sarampo (LVRS), Instituto Oswaldo Cruz, Fiocruz, Rio de Janeiro, Brazil
4. Laboratório Central de Saúde Pública do Estado do Rio Grande do Sul (LACEN-RS), Brazil
5. Laboratório Central de Saúde Pública do Estado de Santa Catarina (LACEN-SC)
6. Laboratório Central de Saúde Pública do Estado do Rio de Janeiro (LACEN-RJ)
7. Laboratório Central de Saúde Pública do Estado do Alagoas (LACEN-AL)
8. Laboratório Central de Saúde Pública do Estado do Paraná (LACEN-PR)
9. Laboratório Central de Saúde Pública do Estado de Minas Gerais (LACEN-MG)
10. Laboratório Central de Saúde Pública do Estado do Espírito Santo (LACEN-ES)
11. Instituto Nacional de Infectologia (INI), Fiocruz, Rio de Janeiro, Brazil
12. Departamento de Entomologia, Instituto Aggeu Magalhães, Fiocruz, Pernambuco, Brazil
13. Núcleo de Bioinformática (NBI), Instituto Aggeu Magalhães FIOCRUZ-Pernambuco, Brazil
14. Laboratório de Ecologia de Doenças Transmissíveis na Amazônia (EDTA), Instituto Leônidas e Maria Deane, FIOCRUZ-Amazonas, Brazil
15. Departamento de Biologia. Centro de Ciências Exatas, Naturais e da Saúde, Universidade Federal do Espírito Santo, Alegre, Brazil.

## **Agradecimentos**

Os autores agradecem a todos os profissionais de saúde e cientistas que trabalharam duro para lidar com esta ameaça de pandemia, a equipe GISAID e todos os colaboradores do banco de dados EpiCoV. Agradecemos também o apoio dos membros da Rede de Vigilância Genômica da Fiocruz COVID-19 (<http://www.genomahcov.fiocruz.br>), a Vigilância Genômica de Vírus Respiratórios. Coordenação Geral da Rede de Laboratórios (CGLab), Ministério da Saúde (MS), Laboratórios Centrais dos Estados (LACEN), Ministério da Saúde (MS) e ao Ministério da Ciência e Tecnologia (MCTI).

## **Financiamentos**

O apoio financeiro foi realizado pela Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM) (PCTI-EmergeSaude / AM chamada 005/2020 e Rede Genômica de Vigilância em Saúde-REGESAM); Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (bolsa 402457 / 2020-0); CNPq / Ministério da Ciência, Tecnologia, Inovações e Comunicações / Ministério da Saúde (MS / FNDCT / SCTIE / Decit) (processo 403276 / 2020-9); Inova Fiocruz / Fundação Oswaldo Cruz (Concessões VPPCB-007-FIO-18-2-30 e VPPCB-005-FIO-20-2-87), INCT-FCx (465259 / 2014-6) e Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) (26 / 210.196 / 2020). F.G.N, G.L.W, G.B e M.M.S são apoiados pelo CNPq por meio de suas bolsas de pesquisa de produtividade (306146 / 2017-7, 303902 / 2019-1, 302317 / 2017-1 e 313403 / 2018-0, respectivamente). G.B. também é financiado pela FAPERJ (Subsídio número E-26 / 202.896 / 2018).