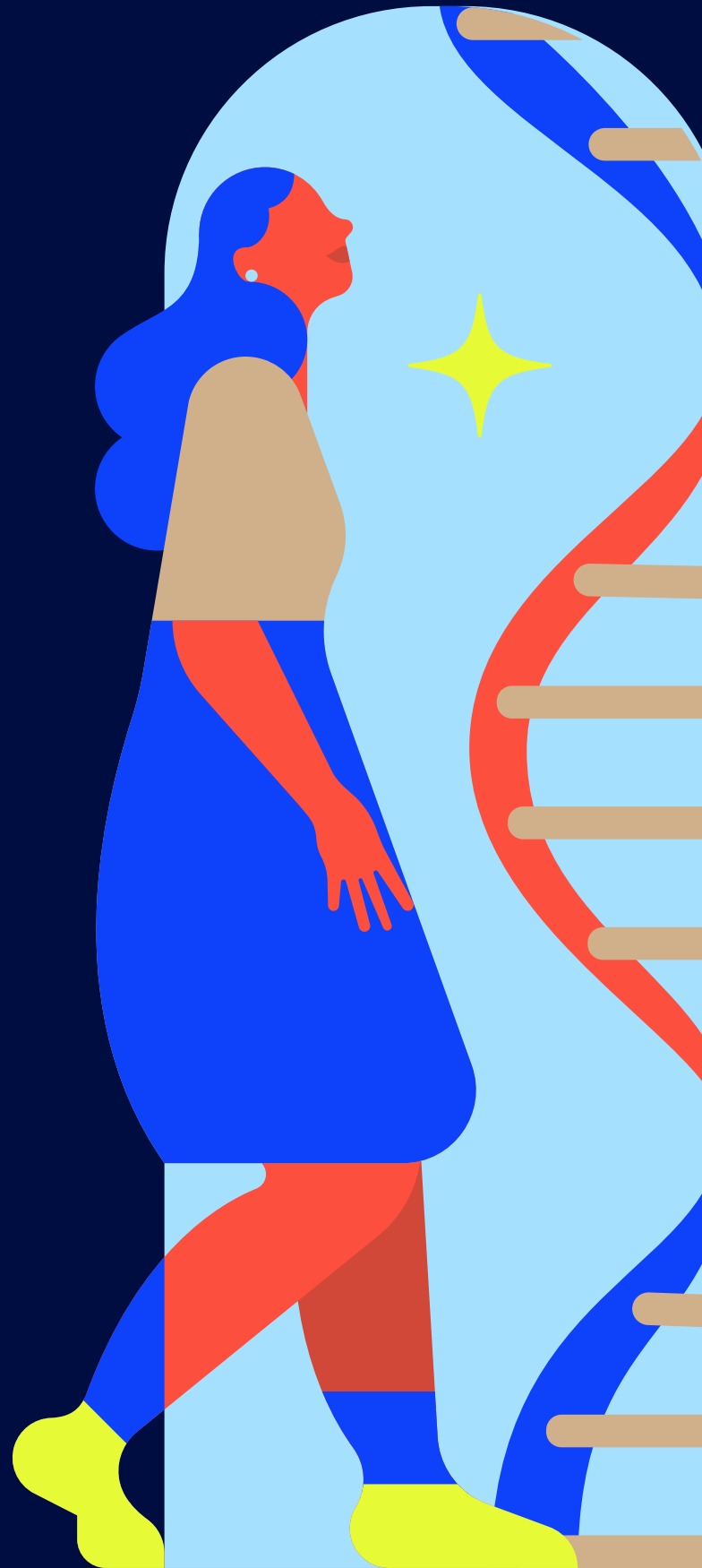


Relatório Genov

15/07/2021



Sumário

introdução.....	3
Total de amostras.....	4
Principais variantes do SARS-CoV-2 e a nomenclatura conforme classificação PANGO	5
Questões Éticas	5
Seleção das amostras	6
Processamento das amostras	6
Sequenciamento dos genomas virais.....	6
Classificação de variantes	7
Resultados	7
Referências	10



Introdução

Neste primeiro relatório trazemos os dados do sequenciamento completo de 502 amostras colhidas na 1ª quinzena de Maio de 2021, providas de pacientes da Dasa com RT-PCR positivo para SARS-CoV-2. Este é um número ainda pequeno mas importante na soma de esforços de vigilância genômica com outras redes e instituições brasileiras. Embora não se saiba qual é a taxa ideal de sequenciamentos/casos positivos, o Brasil hoje, com cerca de 22.000 genomas sequenciados tem uma taxa aproximada de 0.1%, considerando-se 19.2 milhões de casos (15/7/21), enquanto os Estados Unidos da América, país que tem o maior número de casos e mortes, sequenciou cerca de 650.000 genomas de seus 33 milhões de casos ($\cong 2\%$, 15/7/21). Nossa proposta é de sequenciar ao longo do projeto Genov 30.000 genomas, numa taxa de 3.000/mês.

Ainda assim, este relatório já traz muitas informações de importância epidemiológica, como por exemplo a predominância absoluta da variante de preocupação Gama (P.1) em todas regiões do Brasil, respondendo por 98% do total. Este achado reforça nossa percepção de que embora não se deva minimizar o risco representado por variantes que possam eventualmente serem importadas para o nosso país, como a Delta, o maior risco reside na evolução local da variante Gama, acumulando novas mutações, como aqui verificado na aquisição da mutação P681H/R, aliás também presente na variante Delta (P681R).

Também foi identificada uma amostra contendo duas mutações localizadas no gene N exatamente nas regiões onde estão descritos os primers e probes do ensaio de RT-PCR proposto pelo CDC-EUA e largamente utilizado, sugerindo potencial impacto no diagnóstico molecular.

Como aqui apresentado, o sequenciamento genômico produz informações que vão muito além da frequência de variantes. A análise aprofundada dos genomas em conjunto com informações epidemiológicas e dados clínicos e laboratoriais gera novos conhecimentos, mostrando-se fundamental no enfrentamento à Covid-19.

Dr. José Eduardo Levi – em nome da equipe do Genov
CRBio 23.407/01-D



Total de amostras = 502*

1. Região Norte

n=19

- » Idade mediana: 36 anos (1 – 84),
média: 37
- » Sexo: Masculino: 47,4%

Linhagem detectada: P.1 100% (19/19)

2. Região Nordeste

n=22

- » Idade mediana: 31.5 anos (10 –
78), média: 35,55
- » Sexo: Masculino: 59,1%

Linhagem detectada: P.1 81,8% (18/22);
P.1.1 9,1% (2/22); P.1.2 9,1% (2/22)

3. Região Centro-Oeste

n=128

- » Idade mediana: 38 anos (0 –
83), média: 39,41
- » Sexo: Masculino: 48%

Linhagem detectada: P.1 92,1% (118/128);
P.1.1 5,5% (7/128); P.1.2 0,8% (1/128)

4. Região Sudeste só ES+MG

n=9

- » Idade mediana: 43 anos (15 – 70),
média: 44,33
- » Sexo: Masculino: 33,3%

Linhagem detectada: P.1 77,8% (7/9);
B.1.1.7 22,2% (2/9)

5. Estado São Paulo

n=52

- » Idade mediana: 35 anos (3 – 67),
média: 53,8%
- » Sexo: Masculino: 47,4%

Linhagem detectada: P.1 94,2% (48/52);
P.1.1 3,8% (3/52)

6. Estado Rio de Janeiro

n=195

- » Idade mediana: 40 anos (1-82),
média: 40,96
- » Sexo: Masculino: 50,3%

Linhagem detectada: P.1 94,4% (184/195)
P.1.1 2,6% (5/195); P.1.2 1,0% (2/195); B.1.1.7
2,0% (4/195)

7. Região Sul

n=74

- » Idade mediana: 36 anos (6-74),
média: 36,59
- » Sexo: Masculino: 47,9%

Linhagem detectada: P.1 91,8% (69/74);
P.1.1 6,8% (4/74)



Tabela – Principais variantes do SARS-CoV-2 e a nomenclatura conforme classificação PANGO

(<https://cov-lineages.org/>) e nova classificação pela Organização Mundial da Saúde (OMS).

PANGO	OMS	Local da 1ª identificação
VARIANTES DE ATENÇÃO (VOCs)		
B.1.1.7	Alfa	Reino Unido
B.1.351	Beta	África do Sul
P.1	Gama	Brasil
B.617.2	Delta	Índia
VARIANTES DE INTERESSE (VOIs)		
B.1.525	Eta	Vários países
B.1.526	Iota	Estados Unidos da América
B.1.617.1	Kappa	Índia
C.37	Lambda	Peru

Disponível em: <https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/>

12 JULHO 21



Questões Éticas:

Apenas os genomas virais das amostras são sequenciados e associados somente a dados básicos dos pacientes como sexo, idade e local de coleta, mantendo-se o anonimato. Desta forma, o Comitê de Ética em Pesquisa do Hospital 9 de Julho, afiliado à CONEP, aprovou o estudo e a dispensa do Termo de Consentimento Livre e Esclarecido (CAAE: 45540421.0.0000.5455).



Seleção das amostras:

Foi realizado o sequenciamento de 502 genomas completos de SARS-CoV-2 amostrados em diferentes regiões do país na primeira quinzena de Maio, 2021. A escolha das amostras objetivou representar todas regiões do país, ao mesmo tempo refletindo a prevalência do SARS-CoV-2 no período. Por questões técnicas, foram selecionadas apenas amostras com RT-PCR positivo e valor de Ct < 30 (Cycle Treshold) correspondendo a cargas virais que permitem o sequenciamento do genoma completo com qualidade. O número mensal de amostras sequenciadas irá aumentar progressivamente até atingir 3.000/mês que é o limite de capacidade logístico-operacional do GENOV.

Processamento das amostras

As amostras são submetidas ao processo de extração de RNA no equipamento Chemagic, parte do sistema automatizado de pipetagem das amostras e extração (Janus, Perkin-Elmer) utilizando tecnologia de “beads” magnéticas com o kit Viral DNA/RNA 960 (Perkin Elmer), em grupos de 96 amostras. O RNA extraído é eluído em placas de 96 e submetido à síntese de cDNA e amplificação do genoma viral por PCR em fragmentos sobrepostos de aproximadamente 300 pares de bases, com 2 pools de primers já fornecidos no kit o kit Illumina COVIDSeq Test (Illumina, CA, EUA).



Sequenciamento dos genomas virais:

As sequências dos genomas virais são obtidas por Sequenciamento de Nova Geração (NGS), utilizando reagentes do kit Illumina COVIDSeq Test e Mid Flowcell em equipamento NextSeq 550 (Illumina, CA, USA).

O processamento dos dados brutos é feito usando o pipeline COVID Lineage na plataforma de análise Dragen. Após controle de qualidade, são feitas: montagem de sequência consenso, mapeamento e identificação de sítios variantes com relação a cepa referência (NC_045512.2).



Classificação de variantes:

A classificação das linhagens seguiu o sistema dinâmico de classificação proposto por Rambaut e colaboradores (2020), através do software Phylogenetic Assignment of Named Global Outbreak LINages, disponível em:

<https://github.com/cov-lineages/pangolin> e também pelo NextClade. A classificação definitiva das linhagens foi feita após confirmação por análise filogenética contendo sequencias representativas das principais linhagens circulantes.

Resultados:

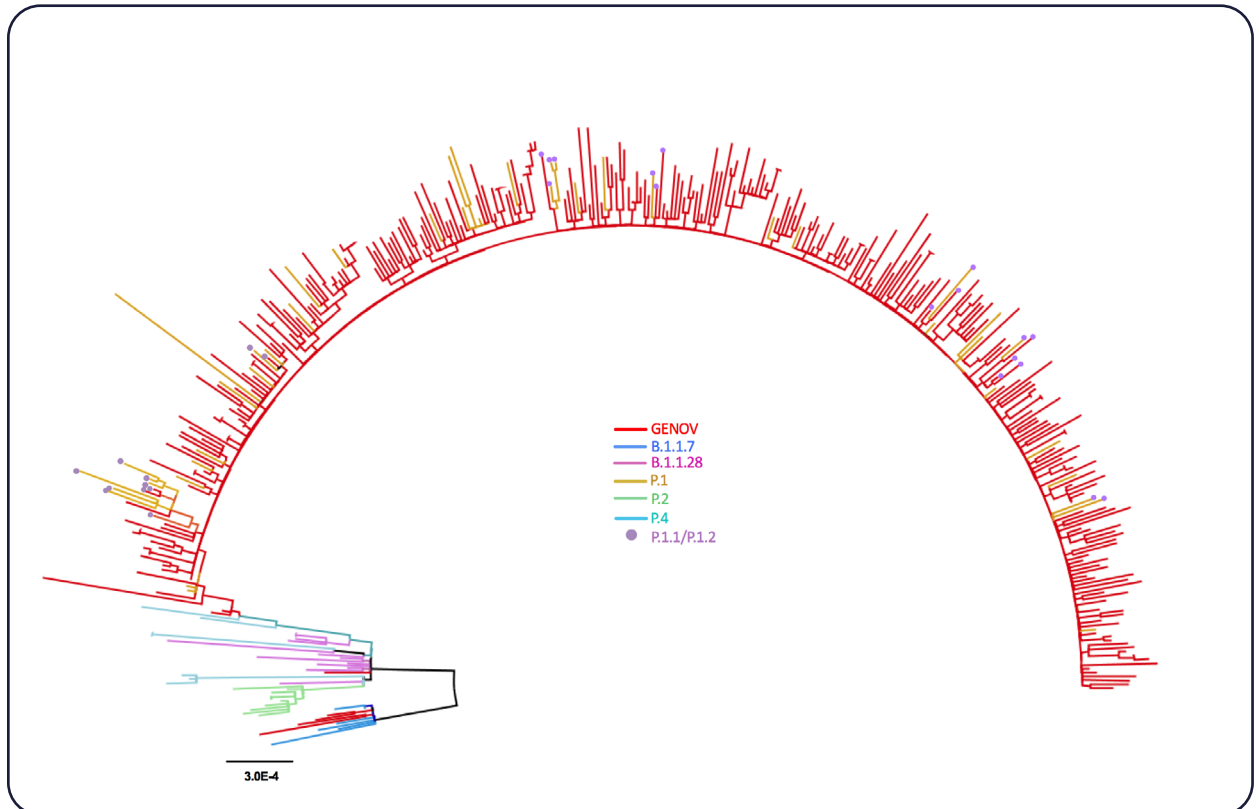
Observou-se a presença das linhagens de atenção Alfa (B.1.1.7) e Gama (P.1 e linhagens derivadas), além das linhagens B.1. e P.4, esta última, restrita ao Estado de São Paulo. A linhagem Gama (P.1) teve origem no Brasil em meados de Outubro/Novembro de 2020 (Faria et al., 2021), e rapidamente se espalhou pelo país e pelo mundo. De fato, a Gama foi aqui a mais frequentemente detectada, respondendo por 97% das amostras sequenciadas, sendo ainda a responsável por 100% das amostras da região Norte e Nordeste. A linhagem Alfa (B.1.1.7), de origem no Reino Unido, foi inicialmente detectada no país em Dezembro de 2020 (Claro et al., 2021), e, embora em baixa prevalência, está atualmente presente em todas as regiões do país, porém não estava presente em nenhuma amostra das regiões Norte e Nordeste.

Também não foi encontrada nenhuma amostra representativa da linhagem de preocupação Delta (B.1.617.2), sugerindo que de fato, a mesma não havia sido introduzida no país até a primeira quinzena de Maio.

A árvore filogenética (ver figura 1) contextualiza a distribuição das amostras sequenciadas neste projeto em relação as linhagens de referência. As recém classificadas P.1.1 e P.1.2, descendentes da linhagem P.1, foram encontradas em todas as regiões amostradas, porém aparecem dispersas na árvore filogenética. Até o momento, sugerimos que a interpretação da classificação de amostras pertencentes as linhagens P.1.1 e P.1.2 deve ser feita com cautela. Foi identificada uma amostra da linhagem P.4 proveniente da cidade de Registro-SP, no sul do estado de São Paulo. Esta linhagem foi descrita inicialmente no noroeste de São Paulo (Bittar C et al., 2021) e já havia sido identificada, entre outras, na cidade de Capão Bonito, na região centro-sul do estado de SP, já próxima à Registro, sugerindo sua expansão geográfica.



Figura 1 – Árvore filogenética de 502 genomas completos de SARS-CoV-2 do Genov.

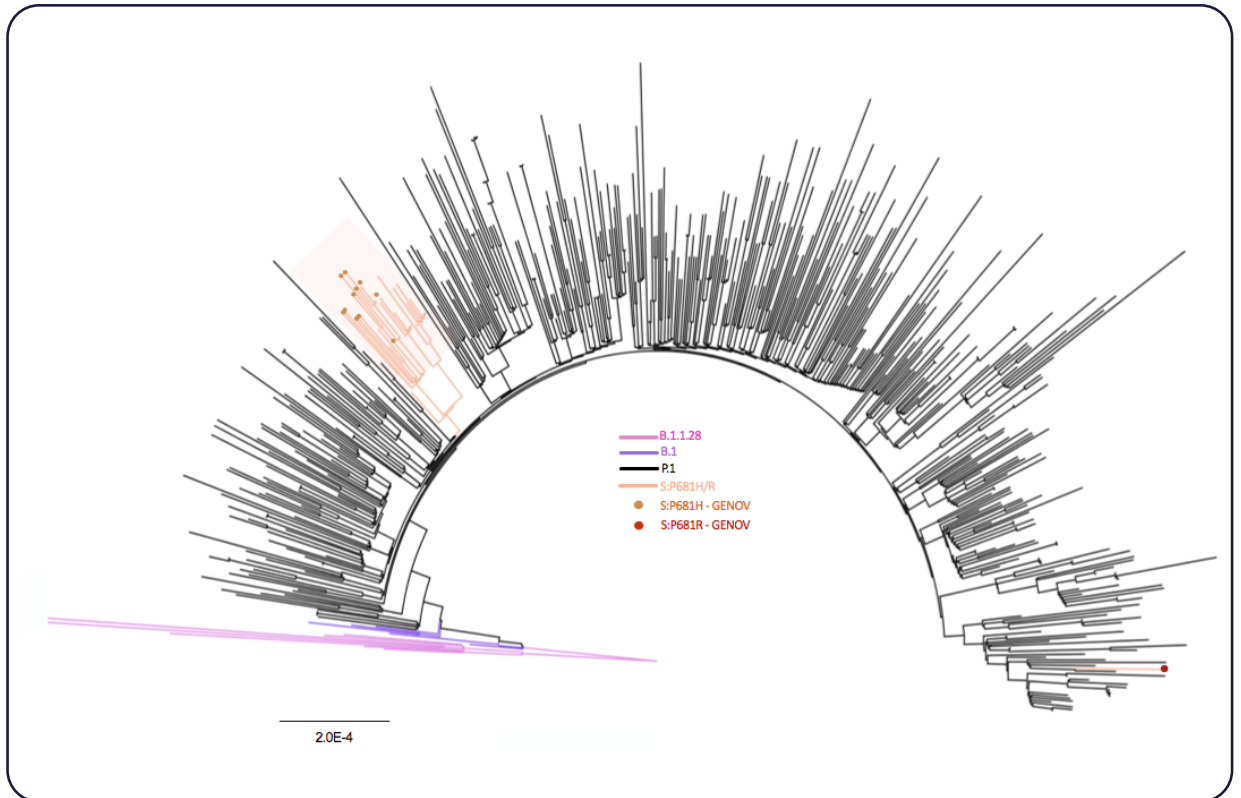


Legenda. A filogenia inclui amostras sequenciadas pelo GENOV (ramos vermelhos) e amostras de SARS-CoV-2 representativas das principais linhagens circulantes no Brasil (ramos de outras cores – vide legenda para identificação). Amostras classificadas pelo sistema Pangolin como P.1.1 e P.1.2 não formam um grupo monofilético e estão representadas pelo símbolo (•) nos respectivos ramos.

Dentre as sequências amostradas pelo GENOV e classificadas como Gama (n= 489), foi encontrado um clado monofilético (n=11) de vírus que possuem a mutação P681H (prolina para histidina) na proteína spike. Essa mutação ocorre no sítio de clivagem pela enzima furina (681-P-R-R-A-R|S-686). A mutação nesse sítio já foi descrita em todas as variantes de preocupação (VOC) e de maneira convergente em diferentes regiões do mundo. É preocupante, no entanto, que embora esta não fosse uma mutação característica da linhagem Gama, a prevalência de vírus desta linhagem que possuem o aminoácido histidina está aumentando rapidamente no país, conforme recém descrito no estado do Amazonas, sendo denominada “Gama-plus” (Naveca et al., 2021) A árvore filogenética abaixo destaca o clado monofilético de 11 amostras Gama contendo a mutação P681H, além de uma amostra Gama também do GENOV com mutação neste mesmo sítio, porém para o aminoácido arginina (P681R). Destas 11 amostras, 5 são do estado de Goiás, 2 do Tocantins, 1 do Mato Grosso, 1 do Ceará, 1 de Santa Catarina e 1 do Paraná.



Figura 2 - Árvore filogenética contendo amostras P.1 geradas pelo GENOV e de referência.



Legenda. O clado em destaque compreende vírus contendo a mutação P681H. As amostras com mutação nesse sítio (P681H/R) geradas no GENOV estão ressaltadas com pontos coloridos nos respectivos ramos.



Referências

Rambaut, A., Holmes, E.C., O’Toole, Á. et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol* 2020; 5:1403–1407. <https://doi.org/10.1038/s41564-020-0770-5>

Faria NR, Mellan TA, Whittaker C, Claro IM, Candido DDS, Mishra S, et al. Genomics and epidemiology of a novel SARS-CoV-2 lineage in Manaus, Brazil. *Science* 2021;372(6544):815-821. doi: 10.1126/science.abh2644. Epub 2021 Apr 14.

Claro IM, da Silva Sales FC, Ramundo MS, Candido DS, Silva CAM, de Jesus JG, Manuli ER, de Oliveira CM, Scarpelli L, Campana G, Pybus OG, Sabino EC, Faria NR, Levi JE. Local Transmission of SARS-CoV-2 Lineage B.1.1.7, Brazil, December 2020. *Emerg Infect Dis.* 2021;27(3):970-972. doi: 10.3201/eid2703.210038. Epub 2021 Jan 26. PMID: 33496249; PMCID: PMC7920684.

Bittar C, Possebon FS, Ullmann LS, Almeida LGP, Banho CA, Campos G et al. Potential new B.1.1.28 sub-lineage with L452R in Brazil. Disponível em: <https://github.com/cov-lineages/pango-designation/issues/68>

Naveca F, Nascimento V, Souza V, Corado A, Nascimento F. Emergence and spread of SARS-CoV-2 P.1 (Gamma) lineage variants carrying Spike mutations 141-144, N679K or P681H during persistent viral circulation in Amazonas, Brazil. Disponível em: <https://virological.org>