

Um ano de pandemia da COVID-19: diversidade genética do SARS-CoV-2 no Brasil

A year of the COVID-19 pandemic: genetic diversity of SARS-COV-2 in Brazil

Carlos Henrique Camargo;^I Claudia Regina Gonçalves;^I Erica Valessa Ramos Gomes Pagnoca;^I Karoline Rodrigues Campos;^I Adriano Abbud;^{II} Adriana Bugno;^{III} Adele Caterino-de-Araujo;^{IV} Claudio Tavares Sacchi^I

^ILaboratório Estratégico. Centro de Respostas Rápidas. Instituto Adolfo Lutz. ^{II}Centro de Respostas Rápidas. Instituto Adolfo Lutz. ^{III}Diretoria Geral. Instituto Adolfo Lutz. ^{IV}Centro de Imunologia. Instituto Adolfo Lutz. Coordenadoria de Controle de Doenças. Secretaria de Estado da Saúde. São Paulo, Brasil.

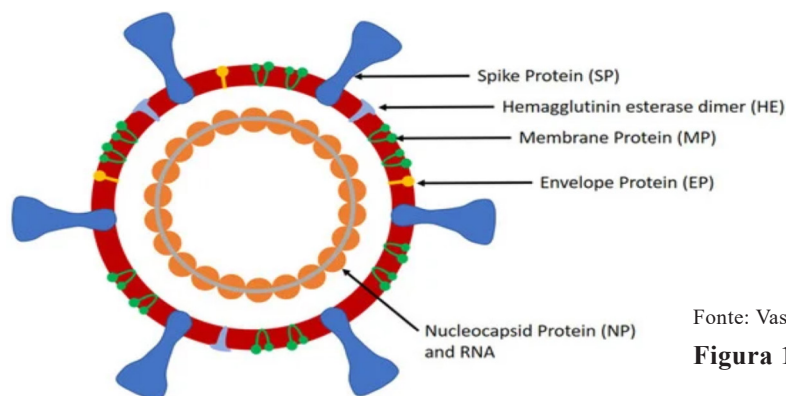
1. INTRODUÇÃO

A COVID-19 é uma doença infecciosa causada por um novo coronavírus anteriormente descrito como agente causal da *Severe Acute Respiratory Syndrome* (SARS) e atualmente denominado SARS-CoV-2. Este novo coronavírus foi identificado primeiramente em Wuhan na China, em dezembro de 2019¹⁻³, espalhando-se por todo o mundo e causando o que a Organização Mundial da Saúde considerou, em 11 de março de 2020, uma Emergência de Saúde Pública de Importância Internacional (ESPII), uma pandemia.^{4,5}

O SARS-CoV-2 compartilha algumas características com outros coronavírus que causaram síndromes respiratórias epidêmicas como a SARS-CoV em 2003 e a MERS-CoV em 2012.⁶ É altamente infectante, dissemina-se com rapidez e tem mostrado elevada

taxa de morbidade e mortalidade no Brasil. Diariamente, o Ministério da Saúde, por meio da Secretaria de Vigilância em Saúde (SVS/MS), divulga dados consolidados sobre COVID-19 e, até 28/03/2021, foram notificados 12.534.688 casos e 312.206 óbitos (<https://covid.saude.gov.br>).

O SARS-CoV-2 é um vírus RNA de fita simples, com polaridade positiva, esférico, envelopado e pleomórfico, cujo envelope consiste de uma camada lipídica onde estão ancoradas as proteínas estruturais de membrana (M), envelope (E) e as espículas (S). A glicoproteína S contém duas subunidades funcionais S1 e S2, sendo a S1 responsável pela ligação do vírus aos receptores celulares no hospedeiro e a S2 pela fusão. Há também uma proteína menor de superfície chamada hemaglutinina esterase (HE). No interior do vírus há um nucleocapsídeo (N) que protege seu RNA.^{7,8}



Fonte: Vashist, 2020⁸

Figura 1: Esquema do SARS-CoV-2.

A estrutura genômica do SARS-CoV-2 é de uma fita de RNA de cerca de 30 Kb, onde na região 5' se encontram as ORF1a/b que codificam poliproteínas (proteases, polimerase (RNA-dependent RNA polymerase, RdRp), helicase e endoribonuclease) e na região 3' os genes que codificam as proteínas estruturais S, E, M e N e outras ORFs que codificam as proteínas acessórias. O genoma do SARS-CoV-2 é flanqueado pelas terminações não codificantes repetitivas UTR 5' e UTR 3'.⁹

Atualmente, o diagnóstico de COVID-19 se baseia na investigação da história epidemiológica, de manifestações clínicas como febre, tosse seca, diarreia, perda de olfato e paladar, falta de ar, entre outras e, pelos exames de radiologia torácica e tomografia computadorizada dos pulmões, além dos testes laboratoriais de diagnóstico sorológico e molecular.^{8,10-13}

O diagnóstico laboratorial específico para COVID-19, realizado a partir de coleta de secreção naso-faríngea, inclui a detecção e o sequenciamento do genoma viral. Até o momento, mais de um milhão de amostras foram analisadas apenas pela plataforma de laboratórios coordenada pelo Governo do Estado de São Paulo.

Embora o diagnóstico imediato e o isolamento sejam importantes para o controle inicial do COVID-19, o desenvolvimento e a implantação de novas técnicas em biologia molecular possibilitam determinar com precisão a composição dos genomas dos agentes infecciosos e traz novas perspectivas para estudos epidemiológicos, ao permitir a caracterização molecular de amostras virais circulantes, estimativas de variabilidade

genética, assim como o conhecimento de sua distribuição geográfica¹⁴⁻¹⁶. Tais informações podem ser utilizadas para a predição de vacinas e medicamentos, para as ações de vigilância em saúde^{14,17-19}, seja no sentido de prever possíveis surtos ou epidemias, ou ainda contê-los, assim como estabelecer uma correlação entre a virulência das amostras e o impacto desses agentes na população.¹⁴⁻²⁰

Apesar de o sequenciamento completo do genoma do vírus não ser um método diagnóstico, desempenha papel relevante na luta contra epidemias, entre as quais a pandemia do COVID-19, pois a análise dos genomas permite quantificar e qualificar a diversidade genética viral, reconstruir origens da epidemia, estimar taxas de transmissão, fornecer informações para o desenvolvimento de vacinas e novos medicamentos, além de permitir o melhoramento de diagnósticos sorológicos e moleculares.^{15-17,20-22} Exemplos de aplicações práticas do sequenciamento do genoma do vírus SARS-CoV-2 incluem o desenvolvimento de testes laboratoriais capazes de detectar o vírus em amostras biológicas,^{14,21} após autoridades chinesas liberarem informações do sequenciamento de um novo tipo de vírus (SARS-CoV-2) e análises filogenéticas que demonstram a variabilidade genética do vírus SARS-CoV-2^{14,21,22} e a distribuição geográfica.^{22,23}

Neste sentido, torna-se indispensável a atuação do Laboratório Estratégico do Instituto Adolfo Lutz – LEIAL, que possui como atribuição proceder à resposta rápida aos agravos emergentes de interesse à saúde pública, com o uso de tecnologias avançadas de biologia molecular, dentre elas o Sequenciamento de Nova Geração –

NGS, cuja realização no Instituto se dá neste laboratório.

2. OBJETIVOS

Nosso objetivo é monitorar a diversidade de SARS-CoV-2, por meio do sequenciamento de genoma completo do vírus, para conhecer o perfil genômico do vírus, monitorar a evolução da pandemia de COVID-19, fornecer dados para estudos de filogenia e evolução geográfica da doença, identificar mutações e seu impacto na situação epidemiológica e avaliar possível correlação entre dados genômicos, laboratoriais, epidemiológicos e clínicos de pacientes com COVID-19.

3. METODOLOGIAS

3.1. Sequenciamento

O sequenciamento do genoma completo do SARS-CoV-2 é realizado pelo LEIAL através da plataforma de NGS *Ion Torrent S5* (ThermoFisher Scientific Inc., USA). Antes do sequenciamento, o RNA de todas as amostras é re-extraído utilizando o sistema automatizado *Loccus Extrata 32* e kit *Extracta Fast*, Cod. MVXA-P016 Fast (Loccus, SP, Brasil) e retestado por RT-qPCR utilizando-se o kit *Allplex 2019-nCoV Assay* (Seegene, Korea). Em seguida o cDNA é imediatamente produzido utilizando a *SuperScript VILO cDNA Syntesis* (Invitrogen, USA) e estocado a -80°C até seu uso. As bibliotecas são feitas com o uso do kit *AmpliSeq SARS-CoV-2* (ThermoFisher Scientific Inc., USA), de acordo com orientações do fabricante. O sequenciamento é feito com os Kits OT2 e ES, sendo que o pool de bibliotecas, depois de dosado e adequadamente equilibrado

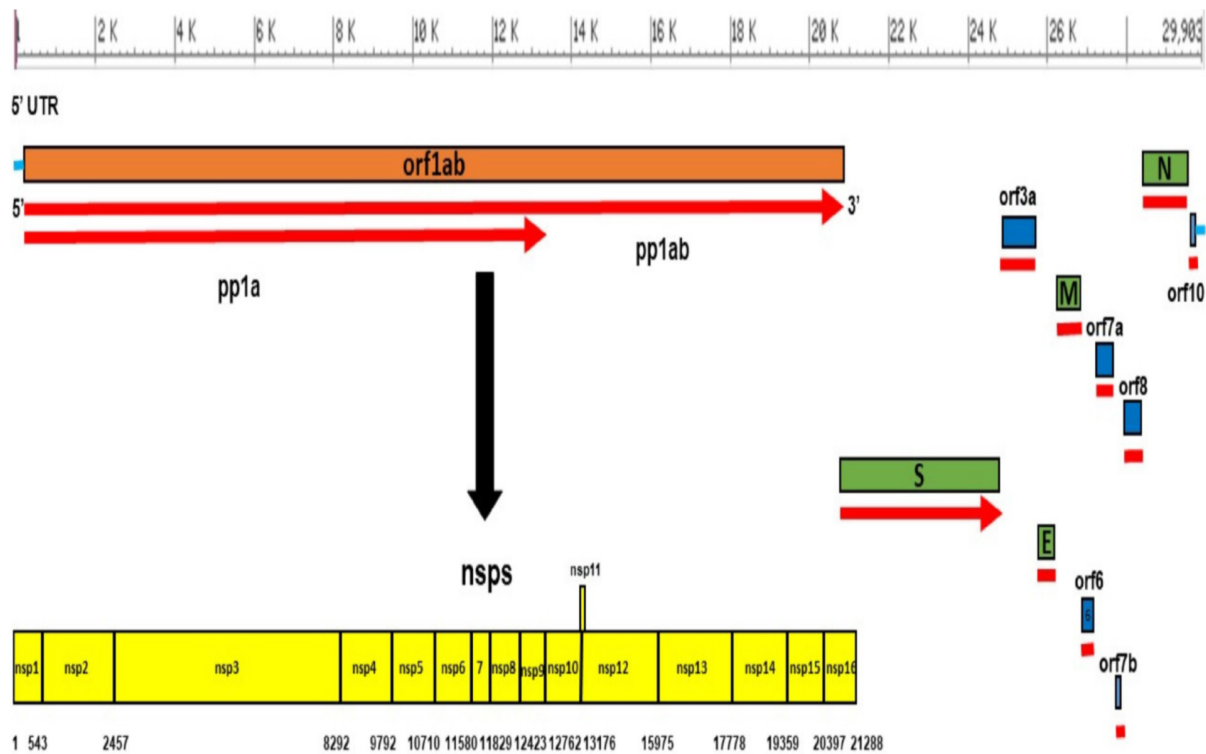
em molaridade, é aplicado em Chips 530 e inserido no sequenciador.

As sequências obtidas são processadas, avaliadas e alinhadas no software *IRMA*, disponibilizado pelo Centers for Disease Control and Prevention – CDC de Atlanta, USA e presente como um plugin no mesmo equipamento. Análise referente à qualidade das bases e variantes é feita através do plugin *Variant Caller* presente no *Ion Torrent S5*.

3.2. Determinação das Linhagens

O genoma do SARS-COV-2 possui 11 ORFs: ORF1ab, ORF2 (S), ORF3a, ORF4 (E), ORF5 (M), ORF6, ORF7a, ORF7b, ORF8, ORF9 (N) e ORF10. A primeira ORF (ORF1ab) codifica 2 poliproteínas (pp1a e pp1ab) que, após a clivagem, resulta em 16 proteínas não estruturais (nsp1 a nsp16), totalizando 26 sequências de codificação de proteínas (Figura 2). Nossas análises de similaridade foram realizadas com o software *BioNumerics v.8.0* (Applied Maths, BioMeriëux, Sint-Martens-Latem, Bélgica) utilizando o plugin *SARS-COV-2*. Para isso, o software analisa as 26 subsequências de codificação de proteínas. Nesta análise, sequências podem ser descartadas quando houver algum tipo de má qualidade ou ausência das sequências em quaisquer das ORFs analisadas. Esta verificação da qualidade das sequências é realizada a cada análise.

Para as análises filogenéticas apenas amostras que apresentam todas as 26 ORFs são consideradas e na identificação das linhagens, é utilizada a nomenclatura proposta pelo GISAID (*Global Initiative on Sharing All Influenza Data*).



Fonte: Helmy, 2020²⁴

Figura 2: Representação esquemática das Open Reading Frames (ORFs) do genoma de SARS-CoV-2

3.3. Minimum Spanning TREE (MST)

As análises e representações de similaridade são apresentadas no formato *Minimum Spanning Tree* (com indicação da UF e/ou linhagem a que pertencem), que leva em consideração as diferenças genéticas entre as sequências em relação à sequência de referência, Wuhan-Hu1 (número de acesso GenBank NC_045512).

4. AMOSTRAS

O GISAID (*Global Initiative on Sharing All Influenza Data*) promove o rápido compartilhamento de dados de todos os vírus Influenza e do coronavírus causador da COVID-19. Isso inclui a sequência genética, dados clínicos e epidemiológicos

relacionados e associados a vírus humanos, e dados geográficos e específicos de espécies associados a vírus aviários e de outros animais, com vistas a auxiliar os pesquisadores a compreenderem como os vírus evoluem e se espalham durante epidemias e pandemias (<https://www.gisaid.org>), sendo que todas as sequências feitas pelo LEIAL, após estarem prontas, são depositadas e disponibilizadas no GISAID.

Todas as sequências utilizadas neste estudo foram extraídas do GISAID, em 29/03/2021, tendo como controle externo, a sequência de referência Wuhan-Hu1 (número de acesso GenBank NC_045512).

Para adequada comparação filogenética, foram utilizadas apenas as sequências

que atenderam aos seguintes critérios de qualidade, que permitem classificá-las como tendo Mínima Qualidade para Análise (MQA): possuir genoma aproximadamente completo (sequência > 29.000 pb); e apresentar sequências com menos de 1% de Ns, mutações únicas (nunca vistas antes) em menos de 0,05% e nenhuma deleção ou inserção, a menos que verificada/confirmada, e serem passagem original (excluídas as passagens por células VERO ou outras). Sequências que não atendem tais critérios, classificadas, portanto, como Sem Qualidade para Análise – SQA, não foram consideradas neste estudo.

5. RESULTADOS

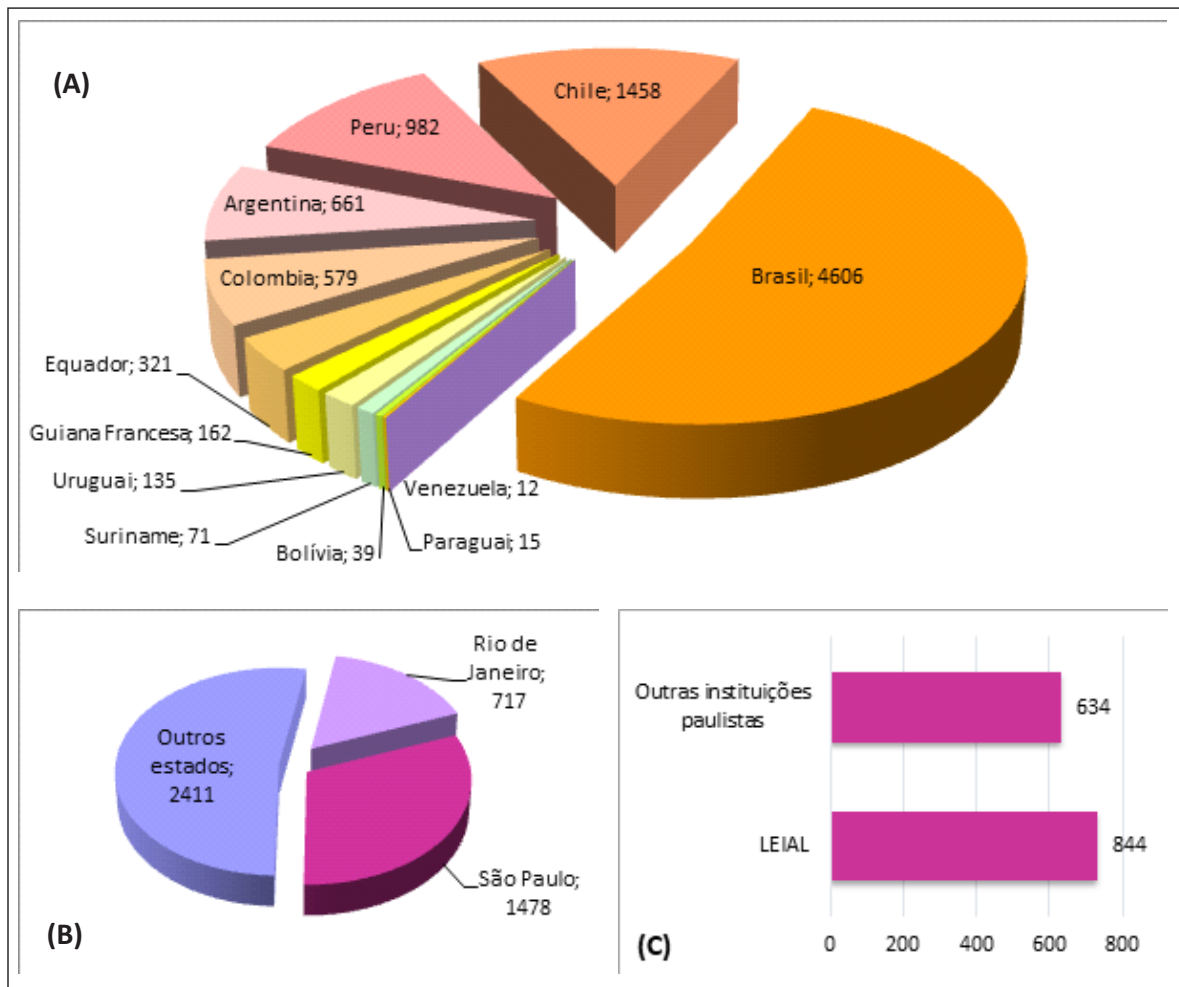
O Brasil tem se destacado na produção de sequências do novo coronavírus, considerando sua capacidade instalada, acesso a insumos e financiamentos em Ciência e Tecnologia. De acordo com Yuki Furuse²⁵, em artigo publicado este ano sobre os esforços dos países no sequenciamento genômico do SARS-CoV-2, o Brasil consta em 11º lugar na classificação mundial considerando indicadores estabelecidos no estudo (Figura 3).

Até 29/03/2021, pouco mais de um ano depois de decretada a pandemia de COVID-19, o Brasil depositou 4606 sequências no GISAID, sendo responsável por 51% dos dados de sequenciamento produzidos na América do Sul. O estado de São Paulo, que representa a principal força nacional na produção desses dados, depositou 32% das sequências produzidas no país, das quais o Laboratório Estratégico é responsável por 60% (Figura 4A, 4B e 4C).

País	Fração	Velocidade	Acesso
EUA	>1000	82	8
Islandia	581	343	12
Holanda	340	114	7
Reino Unido	>1000	33	12
Australia	736	25	15
RDC	231	101	21
China	201	28	23
Portugal	415	2	18
Canada	178	17	19
Singapura	291	24	34
Brasil	271	18	38
Japão	296	13	35
França	83	10	15
Senegal	66	23	19
Dinamarca	335	3	43
Bélgica	132	1	15
Espanha	337	0	24
Letônia	52	18	23
Suécia	273	1	42
Índia	133	2	34
Luxemburgo	127	1	21
Suíça	138	14	55
Nova Zelândia	224	8	64
Tailândia	176	2	48
Alemanha	52	3	31
Arábia Saudita	81	3	42
Chile	132	7	57
Irlanda	13	5	24
Rússia	79	1	26
Colômbia	120	2	48
Itália	41	2	26
Panamá	217	1	119
Austria	172	0	54,5
Finlândia	45	13	77
Quênia	102	0	46
Grécia	102	0	54
Coreia do Sul	104	2	131
Israel	75	0	46
Bangladesh	2	0	16
Noruega	26	0	28
Egito	2	0	21
Polônia	15	1	49
México	34	1	77
África do Sul	24	1	63
Turquia	41	0	53
Omã	66	0	79
UAE	44	0	97,5
Malásia	31	0	115
Romênia	11	0	63

Fonte: Furuse, 2021²⁵

Figura 3: Classificação dos países com maior esforço no sequenciamento genômico do SARS-CoV-2



Fonte: GISAID, 2021

Figura 4: Sequências de SARS-CoV-2 depositadas no GISAID por países da América do Sul (A), por São Paulo (B) e pelo LEIAL (C), considerando números absolutos.

Considerando a qualidade das sequências, verifica-se que 3380 (73,4%) sequências depositadas pelo Brasil no GISAID, até 29/03/2021, atenderam aos requisitos mínimos para serem consideradas como MQA e desta forma, apresentam informações confiáveis para respostas em vigilância. No mesmo período, o Estado de São Paulo foi responsável por depositar 1075 sequências com MQA (que correspondem a 31,8% das sequências com MQA brasileiras e a 72,7% das sequências com MQA paulistas), das quais 844 (78,5%) foram depositadas pelo Laboratório Estratégico do IAL.

5.1. Monitoramento das Variantes de Atenção (VOC)

Certas linhagens do SARS-CoV-2 são denominadas de variantes de atenção (*variants of concern* - VOC) por apresentarem aumento expressivo em sua incidência em curto espaço de tempo. As linhagens atualmente denominadas de VOC são B.1.1.7 (originária no Reino Unido), B.1.351 (originária da África do Sul) e P.1 (originária de Manaus, AM, Brasil). Entretanto, a linhagem B.1.351 não foi reportada no Brasil até a presente data (Tabela 1).

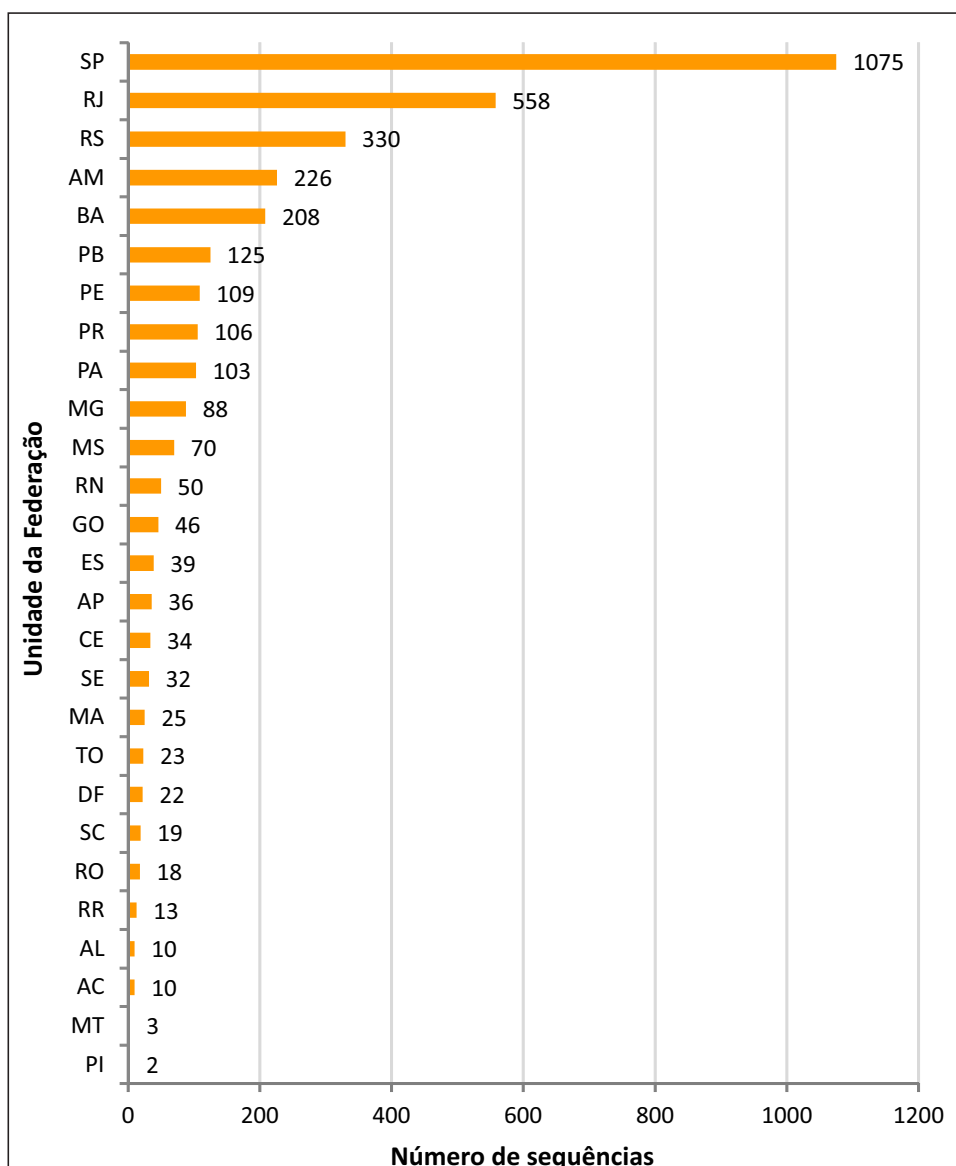
Tabela 1. Distribuição das variantes de preocupação predominantes circulantes no Brasil e no estado de São Paulo

Variante de preocupação	Brasil (n)	%	Estado de SP	%
P.1 (Manaus)	225	7,4	75	7,9
B.1.1.7 (Reino Unido)	44	1,5	22	2,3
B.1.351 (África do Sul)	0	0	0	0
Total de seqüências	3028	100	953	100

5.2. Circulação de SARS-CoV-2 no Brasil

A Figura 5 representa a distribuição geográfica das sequências brasileiras com

MQA, depositadas no GISAID, em função da Unidade da Federação na qual foi realizada a coleta da amostra.

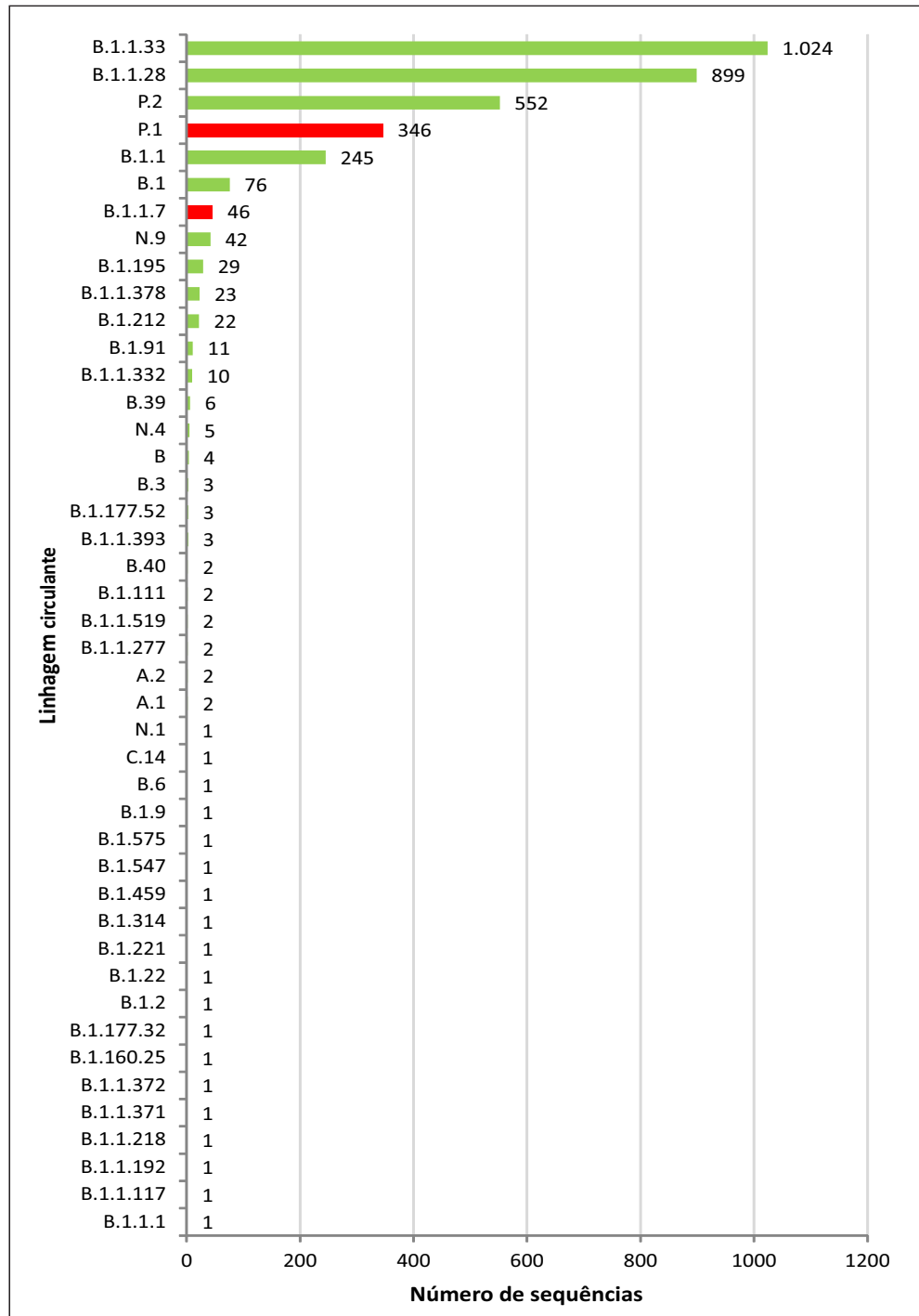


Fonte: GISAID, 2021

Figura 5. Distribuição das seqüências com MQA em função da Unidade da Federação de origem da amostra

Quarenta e quatro diferentes linhagens circularam no país desde o início da pandemia (Figura 6), sendo a distribuição das mais frequentes apresentada na Tabela 2.

Os dados indicam que duas linhagens são predominantes, a linhagem B.1.1.33 (n=1.024; 30,3%) e a B.1.1.28 (n=899; 26,6%), que juntas totalizam 56,9% das sequências realizadas.



Fonte: GISAID, 2021

Figura 6: Distribuição das linhagens de SARS-CoV-2 circulantes no Brasil desde o início da pandemia de COVID-19

Tabela 2. Distribuição das linhagens mais frequentes por Unidade da Federação

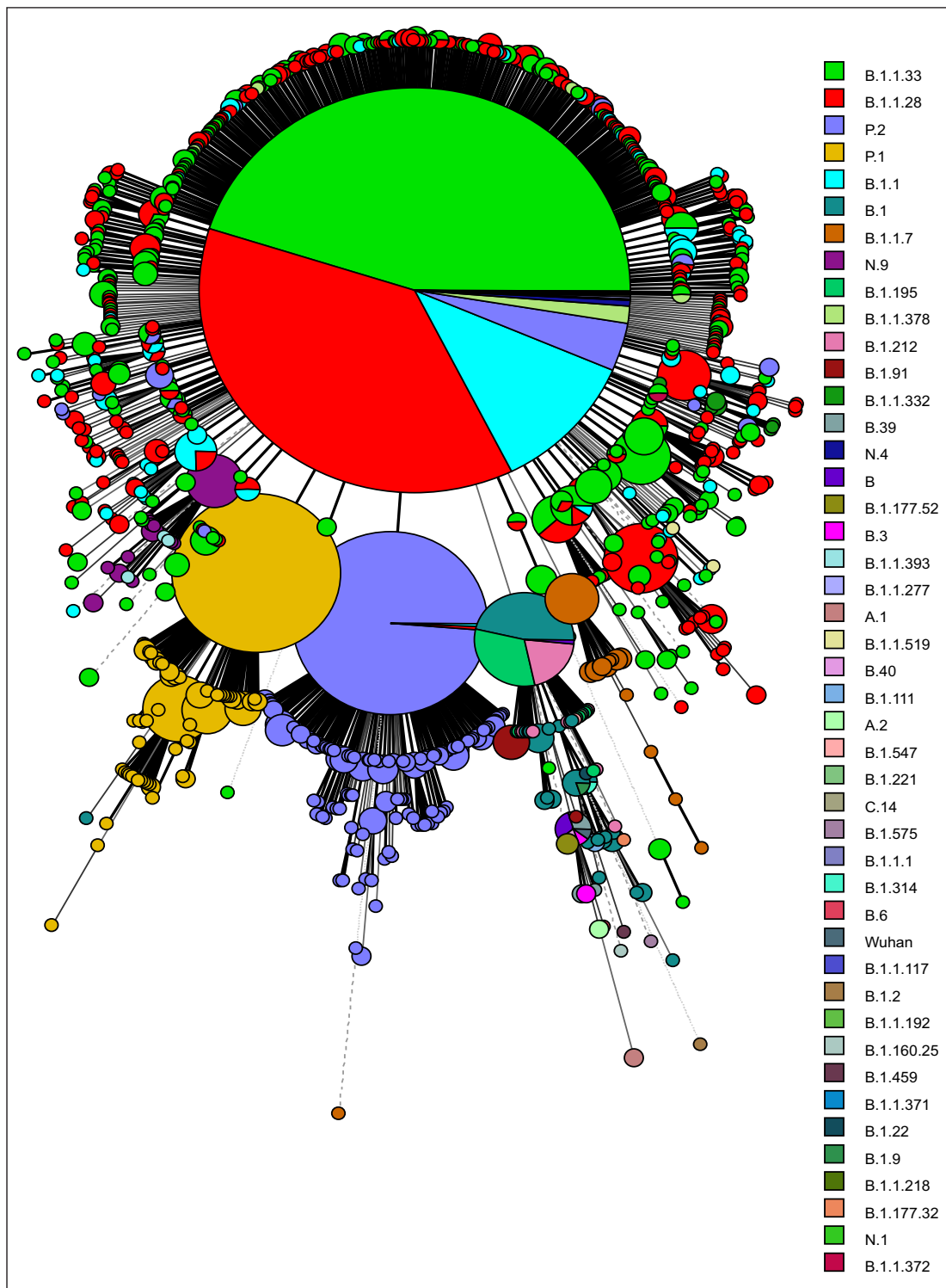
Região	UF	B.1.1.33	B.1.1.28	P.2	P.1	B.1.1.29	B.1.1.143	B.1.1.7	Outras 66 linhagens	Total Geral
Sudeste	SP	150	451	121	75	3	66	22	65	953
	RJ	408	39	41	-	2	1	2	77	570
	MG	19	7	25	3	-	-	11	15	80
	ES	11	-	-	-	-	-	1	1	13
Sul	RS	83	96	68	30	13	2	-	39	331
	PR	56	18	23	-	1	-	1	7	106
	SC	12	3	-	-	1	-	-	3	19
Norte	AM	24	53	12	89	-	4	-	35	217
	PA	19	42	28	2	1	1	-	9	102
	AP	17	-	12	-	-	-	-	2	31
	RR	4	-	-	7	-	-	-	1	12
	AC	7	-	-	-	-	-	-	3	10
	RO	5	-	-	-	-	-	-	2	7
	TO	2	9	2	1	1	-	-	1	16
Nordeste	BA	77	24	2	11	-	1	1	34	150
	PE	9	7	-	-	79	-	-	13	108
	PB	22	2	5	1	48	2	-	6	86
	SE	11	2	-	-	4	-	1	14	32
	CE	14	1	9	-	-	-	-	7	31
	MA	21	1	2	-	-	-	-	-	24
	AL	4	2	1	-	-	-	-	2	9
	RN	2	2	2	-	-	-	-	-	6
	PI	2	-	-	-	-	-	-	-	2
Centro-Oeste	MS	28	21	-	-	-	-	-	5	54
	GO	21	3	1	6	-	-	2	4	37
	DF	6	4	5	-	1	-	2	2	20
	MT	1	-	-	-	-	-	1	-	2
		1035	787	359	225	154	77	44	347	3028

Fonte: GISAID, 2021

A região Sudeste é representada por 1.760 sequências (52,1%) enquanto as sequências das demais regiões do país somam 1.620 (47,9%) sequências em condições de análise (Tabela 2).

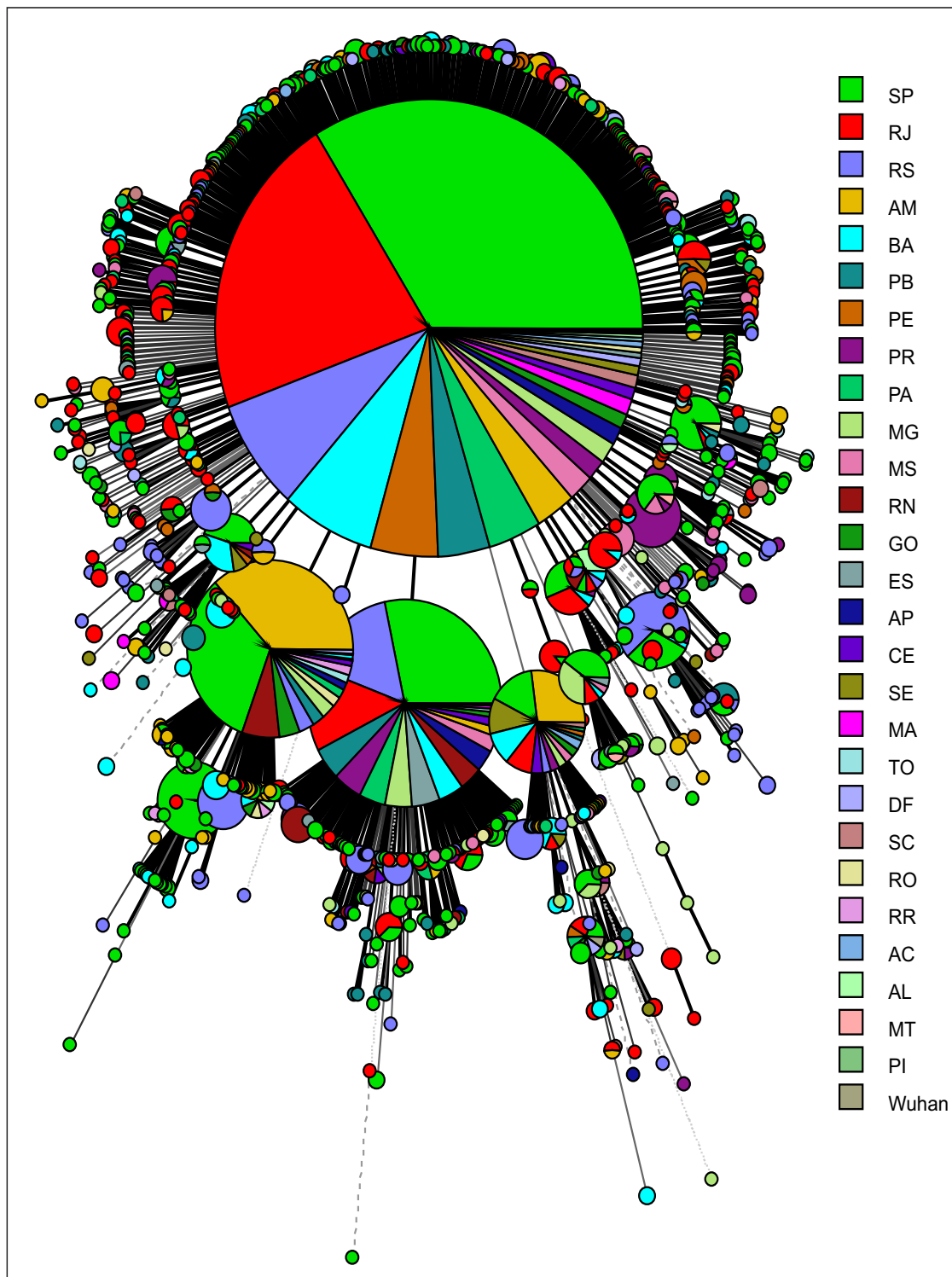
A representação esquemática destes dados é mostrada nas figuras 7 e 8.

Os dados, apresentados na Figura 9, indicam o predomínio da linhagem B.1.1.28 no estado de São Paulo (n=541) e da linhagem B.1.1.33 no estado do Rio de Janeiro (n=416) ($p < 0,001$, teste de qui-quadrado)



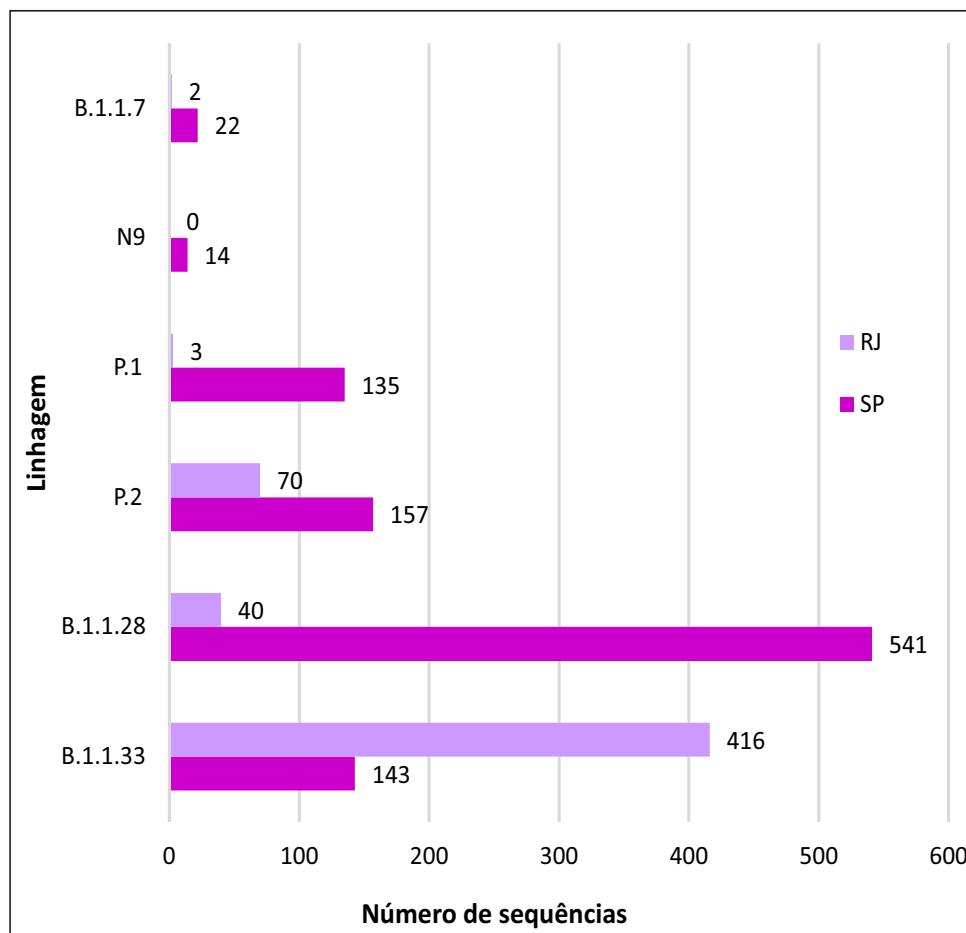
Fonte: GISAID, 2021

Figura 7. Representação esquemática (*Minimum Spanning Tree*) da semelhança entre as sequências de SARS-CoV-2 isolados no Brasil ($n=3.380$). O agrupamento é feito de acordo com as semelhanças entre elas (em relação aos SNPs nas 25 ORFs do vírus): cada sequência é representada por um círculo, o qual é colorido de acordo com a linhagem determinada pela similaridade com a cepa ancestral, sendo que o tamanho de cada círculo é proporcional ao número de sequências.



Fonte: GISAID, 2021

Figura 8. Representação esquemática (*Minimum Spanning Tree*) da semelhança entre as sequências de SARS-CoV-2 isolados no Brasil ($n=3.380$). O agrupamento é feito de acordo com as semelhanças entre elas (em relação aos SNPs nas 25 ORFs do vírus): cada sequência é representada por um círculo, o qual é colorido de acordo com a UF de isolamento, sendo que o tamanho de cada círculo é proporcional ao número de sequências.



Fonte: GISAID, 2021

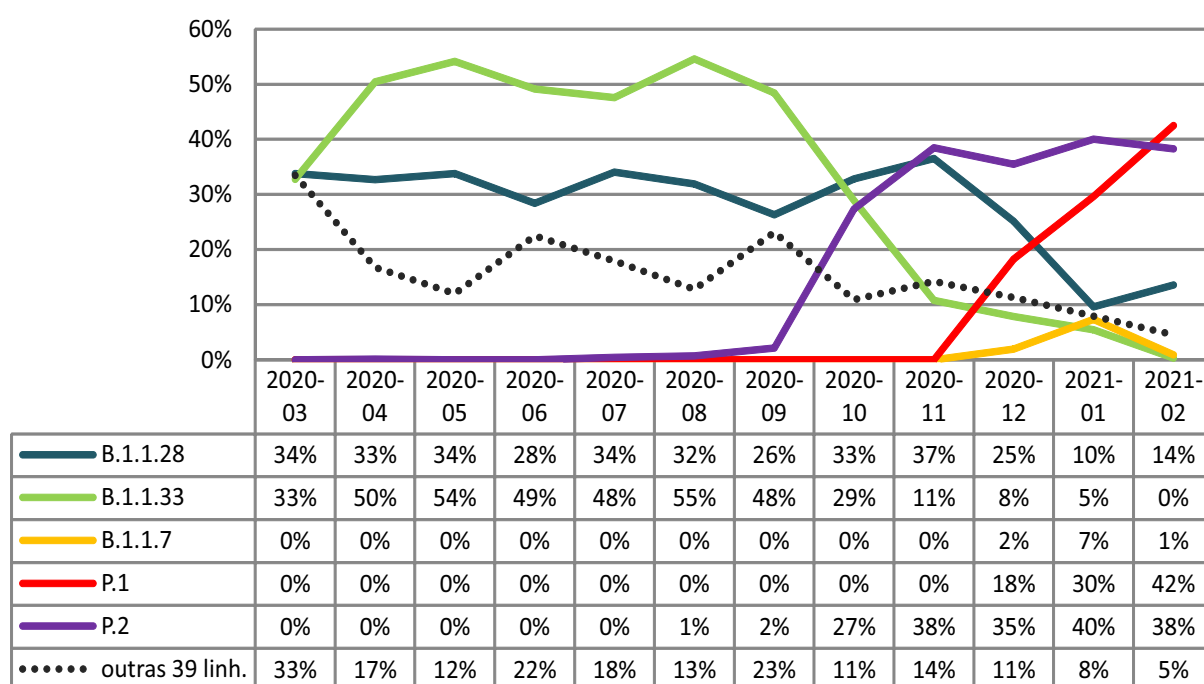
Figura 9. Distribuição das linhagens mais frequentes nos estados de SP e RJ.

De acordo com os dados do portal CoV-lineages (<https://cov-lineages.org/>), repositório mundial para as sequências de SARS-CoV-2, até 29/03/2021, a linhagem B.1.1.33 circula em 25 países dos cinco continentes, embora predominantemente no Brasil (83% de todas as sequências depositadas são do Brasil). Esta linhagem originou outras nove, das quais somente N.1, N.4 e N.9 foram reportadas no Brasil, até o momento.

A linhagem B.1.1.28 originou as linhagens P.1 (variante brasileira), P.2, frequentemente identificadas no Brasil, e P.3, ainda não reportada no país. Até o presente momento, a linha-

gem B.1.1.28 já foi encontrada em 46 países, a variante P.1, em 27, a P.2, em 28 países e a P.3 foi reportada apenas nas Filipinas.

Considerando as amostras brasileiras com informações disponíveis no GISAID (acesso em 29/03/2021, n=3.358) isoladas de março/2020 a fevereiro/2021, verifica-se a predominância das linhagens B.1.1.28 e B.1.1.33 ao longo do primeiro semestre da pandemia (Figura 10). A partir de agosto, entretanto, verifica-se um declínio proporcional da frequência destas linhagens com consequente aumento de outras, em particular das linhagens P.1, e P.2.



Fonte: GISAID, 2021

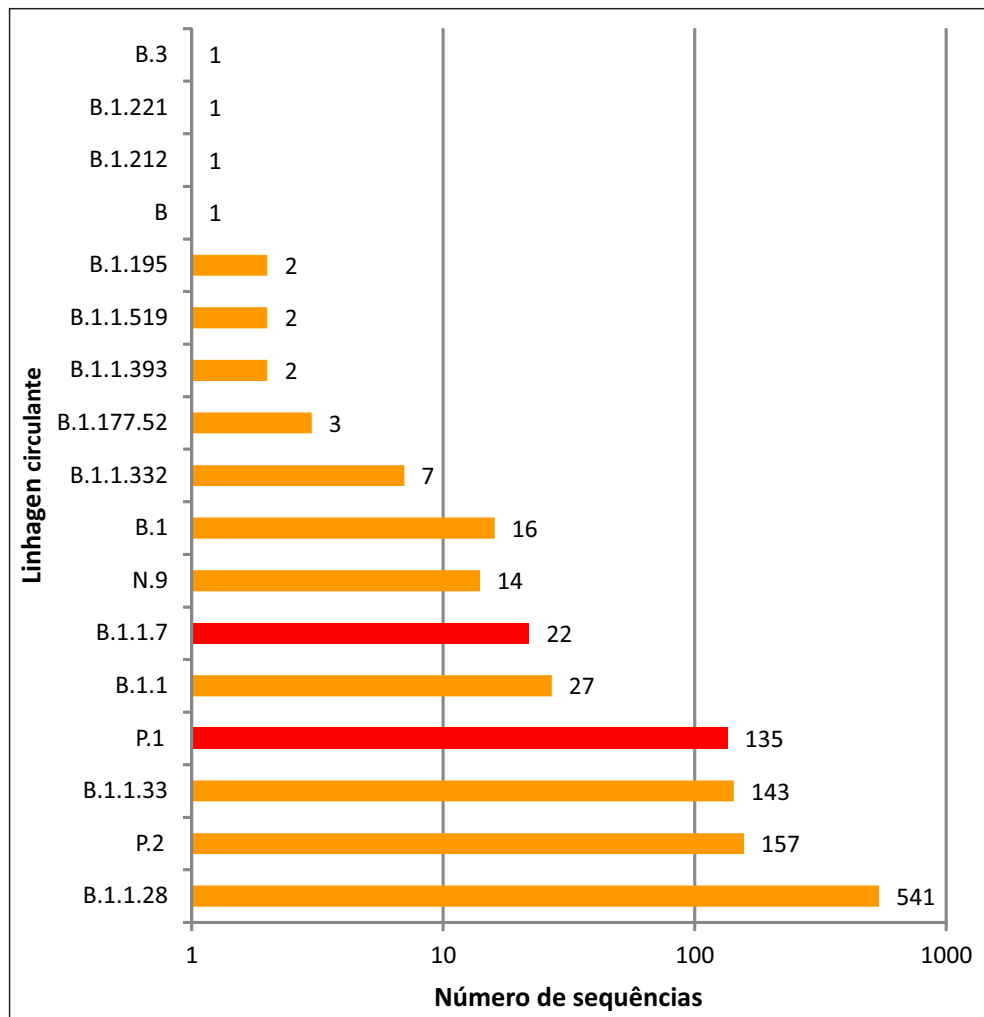
Figura 10. Distribuição temporal das principais linhagens circulantes no Brasil entre as sequências de SARS-CoV-2 depositadas no banco de dados GISAID (acesso em 29/03/2021)

5.3. Circulação de SARS-CoV-2 no Estado de São Paulo

Das 44 linhagens encontradas no país, dezessete (38,6%) diferentes linhagens circularam no estado de São Paulo desde o início da pandemia (Figura 11), sendo detectadas ao longo do ano, conforme apresentado na Figura 12.

A análise da distribuição geográfica das sequências paulistas com MQA em função do município de origem da amostra fica prejudicada, uma vez que o depósito por outras instituições paulistas muitas vezes não indicam em qual município do estado de São Paulo a amostra foi colhida. Entretanto, considerando que todas as sequências depositadas pelo LEIAL indicam o município de origem, a avaliação da distribuição geográfica no estado foi realizada prioritariamente com as amostras do LEIAL.

Devido à heterogeneidade do estado de São Paulo, informações geradas a partir de dados que compreendam todo o território tendem a gerar conclusões incorretas e por este motivo, as análises consideraram os Departamentos Regionais de Saúde (DRS), que correspondem a 17 divisões territoriais de natureza político-administrativa no estado de São Paulo, organizadas em função de particularidades como tradição de construção de políticas regionalizadas, dinâmica de movimento populacional, características geográficas, condições socioeconômicas e orçamentárias e participação nas diferentes Redes Regionais de Atenção à Saúde. A investigação completa para determinar a ocorrência das linhagens contribui para estratégias de vigilância, de maneira regionalizada, visando à tomada de medidas mais assertivas pelo poder público.

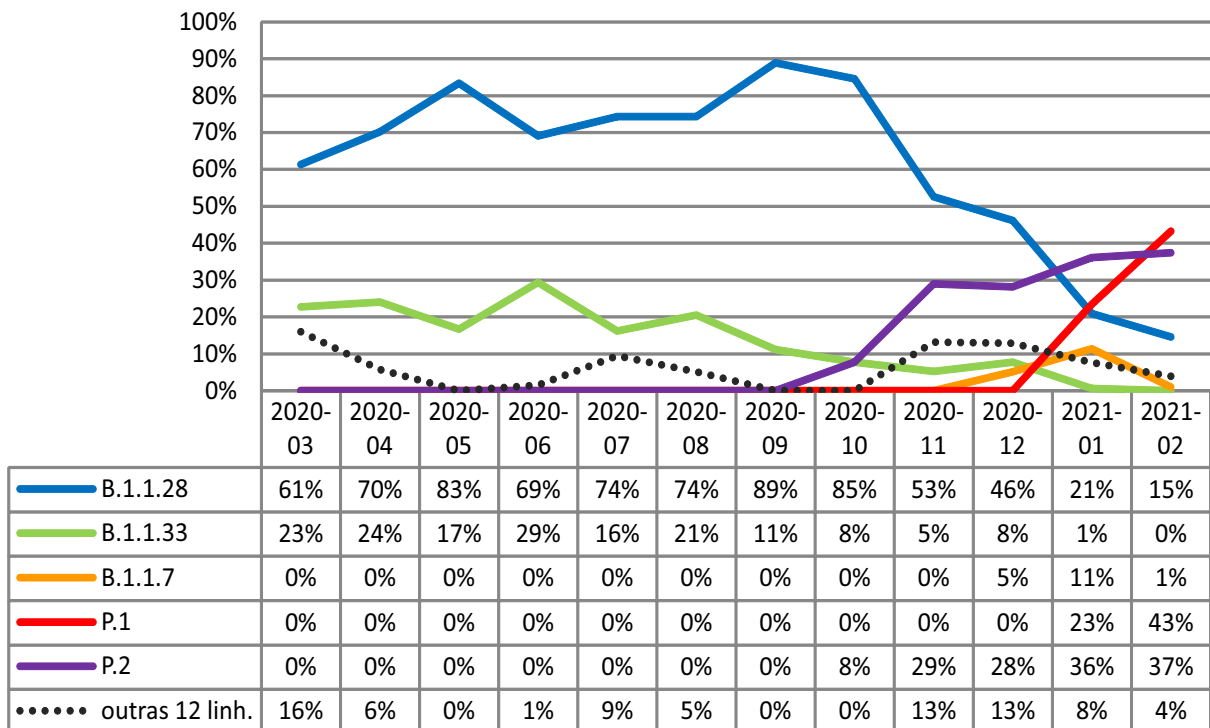


Fonte: GISAID, 2021

Figura 11. Distribuição das linhagens de SARS-CoV-2 circulantes no estado de São Paulo desde o início da pandemia de COVID-19

Inicialmente, um panorama de ocorrência das linhagens do SARS-CoV-2 nas diferentes Regiões de Saúde do estado (Figuras 13 a 21), está sendo construído pelos Grupos de Vigilância Epidemiológica, em conjunto com os Centros Regionais do IAL, que selecionam amostras positivas de relevância clínico-epidemiológica, com representatividade estatística, para que o Biobanco COVID-19 receba o material, proceda ao seu acondicionamento, realize sua separação e encaminhe ao Laboratório Estratégico do IAL,

local em que o processo de sequenciamento de genoma completo do vírus é realizado. A partir dos resultados obtidos, o Centro de Vigilância Epidemiológica do estado realiza as investigações de aspectos clínicos, históricos de viagem e rastreamento de contatos, para que os casos de ocorrência das Variantes de Atenção (VOC) possam ser confirmados, sendo que até o momento, foram identificados no Estado de SP 22 casos de infecção pela variante inglesa (B.1.1.7) e 75 pela variante brasileira (P.1).

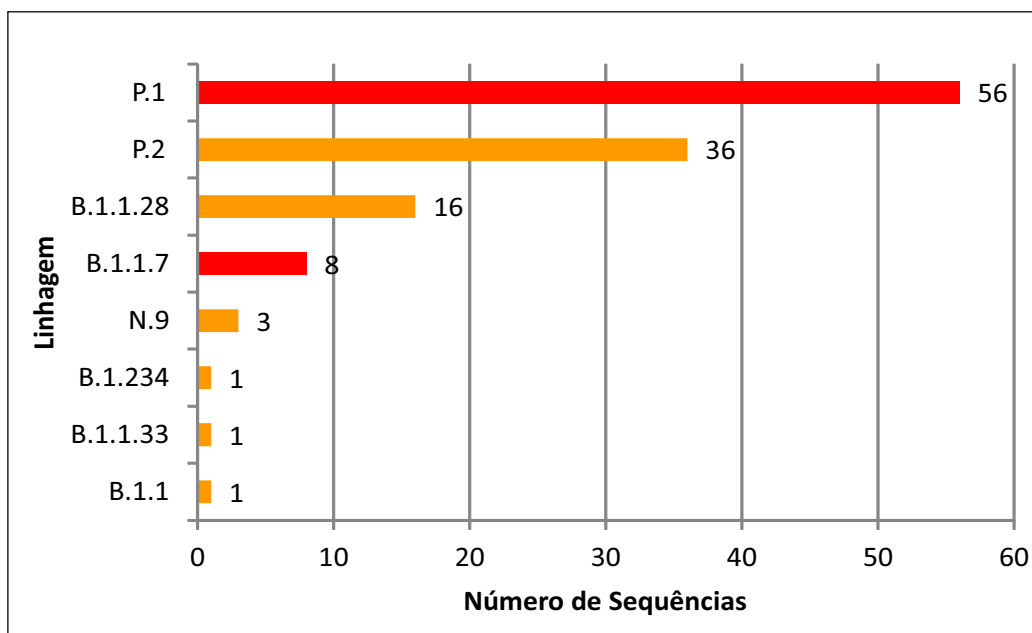


Fonte: GISAID, 2021

Figura 12. Distribuição temporal das principais linhagens circulantes no estado de São Paulo entre as sequências de SARS-CoV-2 depositadas no banco de dados GISAID (acesso em 11/03/2021)

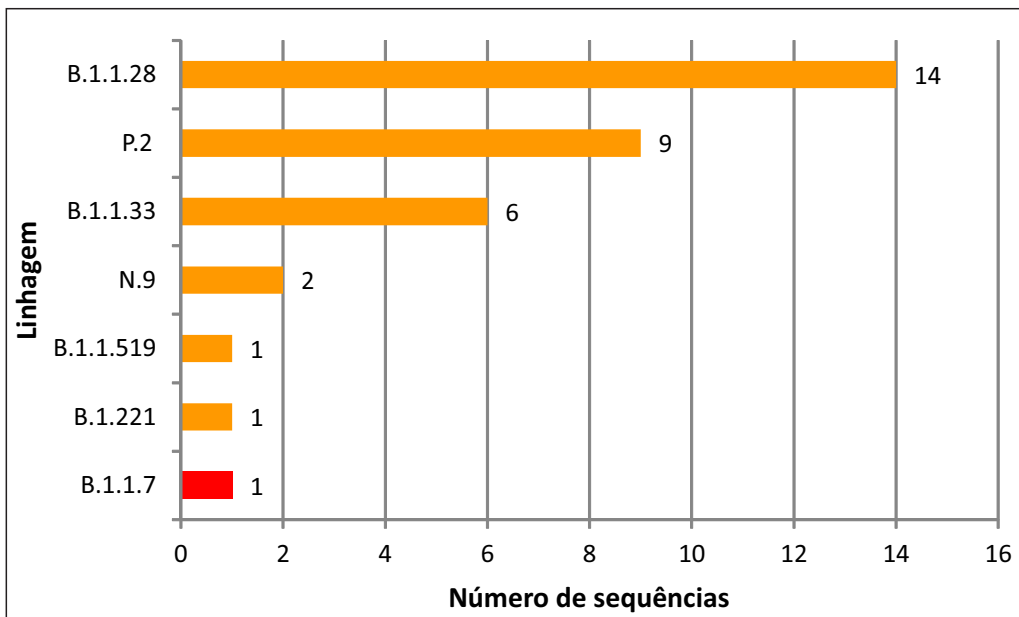
Do ponto de vista epidemiológico, as ações devem ser tomadas com base em dados abrangentes, e não em universos restritos,

como apenas a busca por Variantes de Atenção em detrimento do panorama genômico da COVID-19.



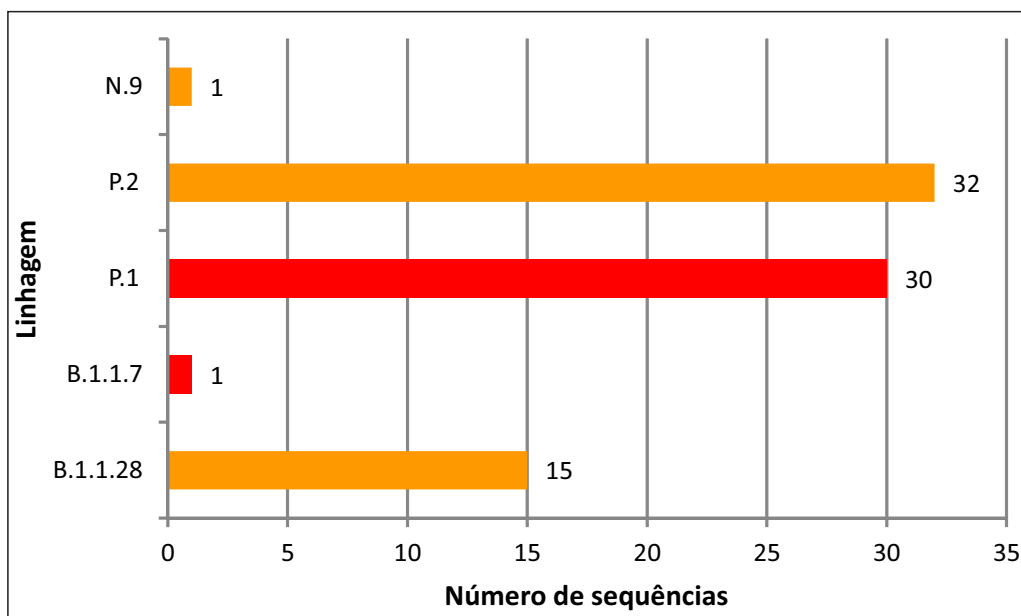
Fonte: GISAID, 2021

Figura 13: Distribuição das linhagens circulantes na DRS I – Grande São Paulo



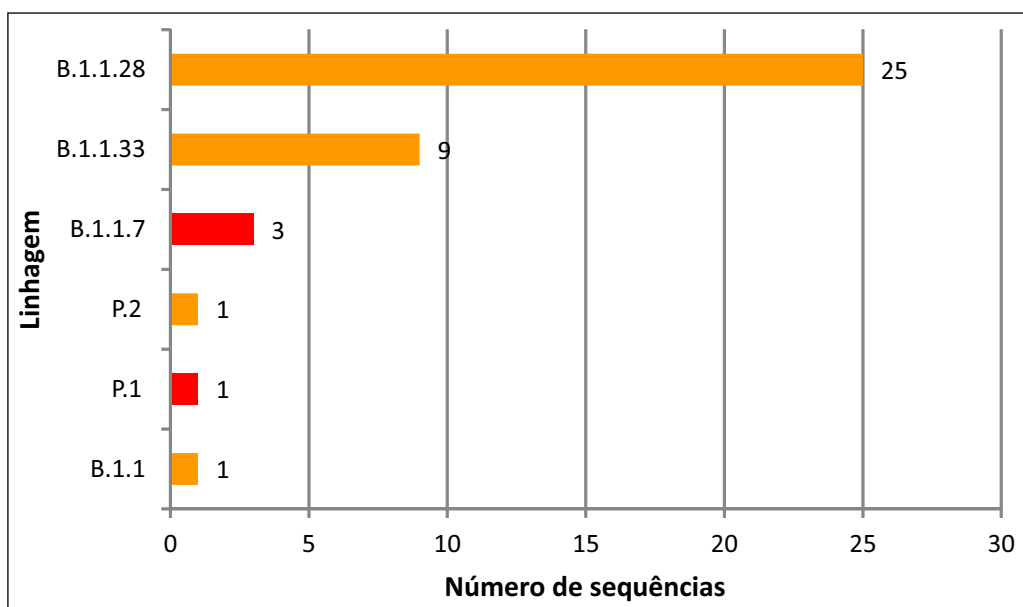
Fonte: GISAID, 2021

Figura 14: Distribuição das linhagens circulantes na DRS IV – Baixada Santista



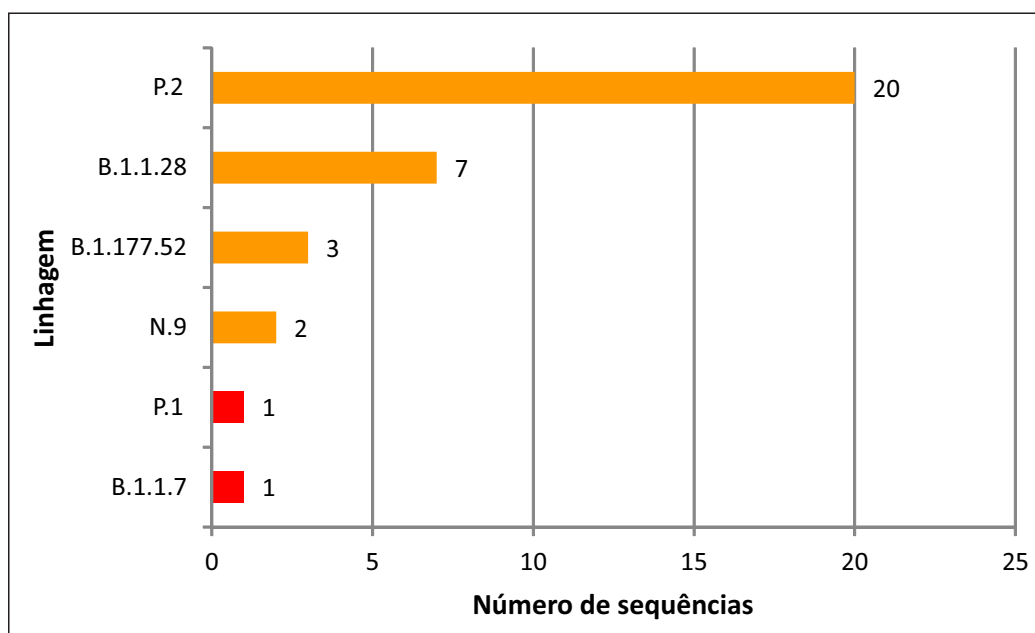
Fonte: GISAID, 2021

Figura 15: Distribuição das linhagens circulantes na DRS VI – Bauru



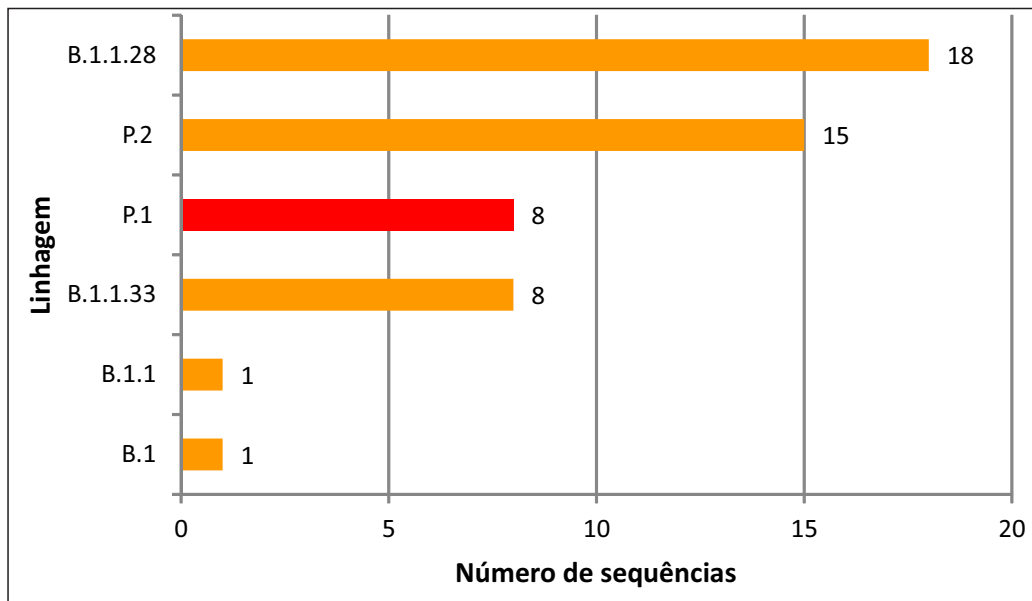
Fonte: GISAID, 2021

Figura 16: Distribuição das linhagens circulantes na DRS VII – Campinas



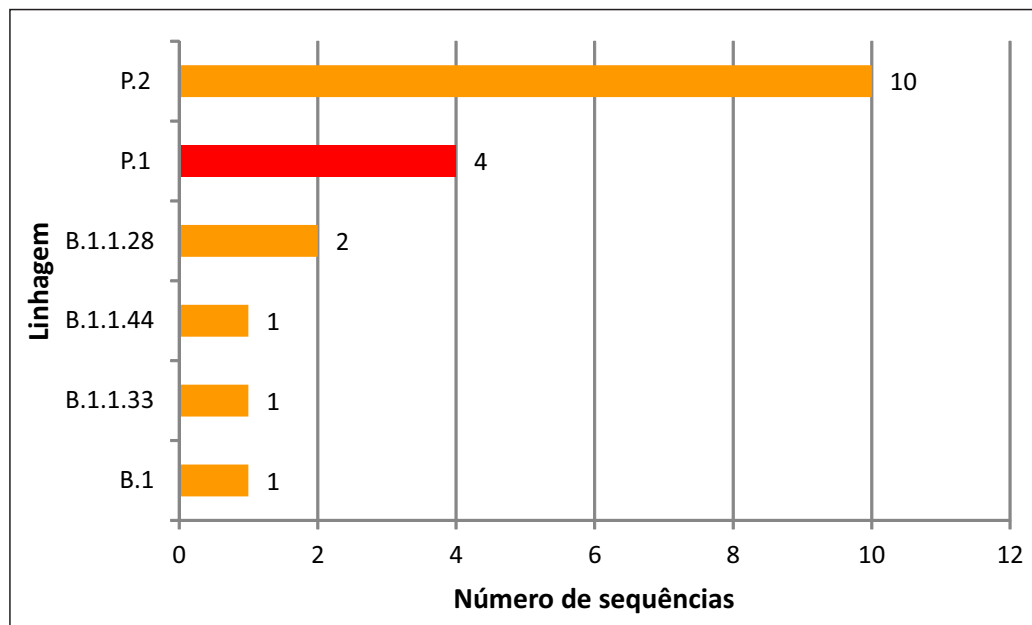
Fonte: GISAID, 2021

Figura 17: Distribuição das linhagens circulantes na DRS IX – Marília



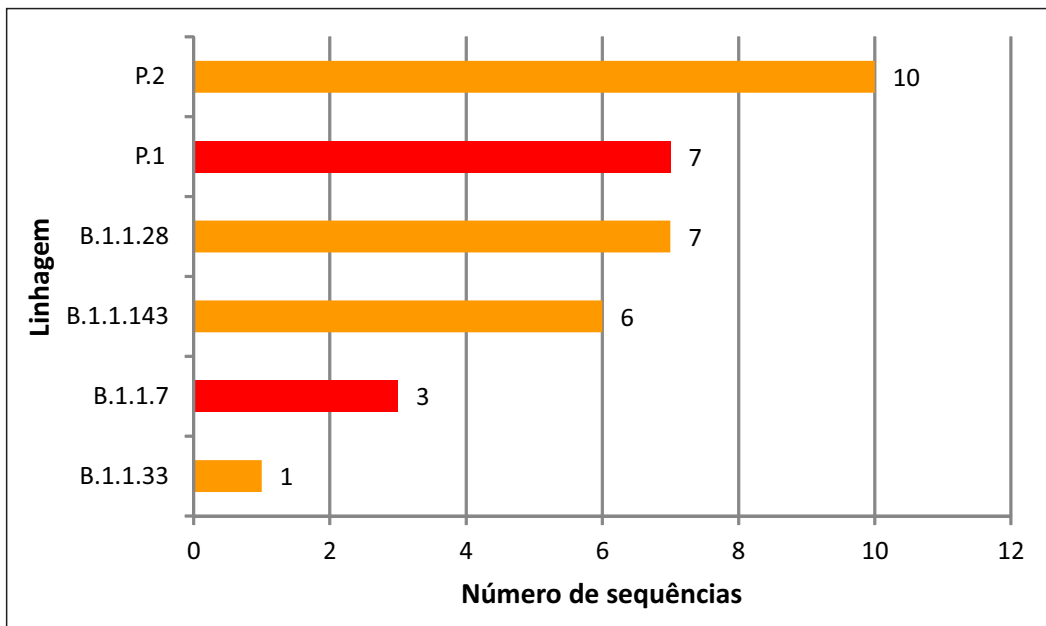
Fonte: GISAID, 2021

Figura 18. Distribuição das linhagens circulantes na DRS XI – Presidente Prudente



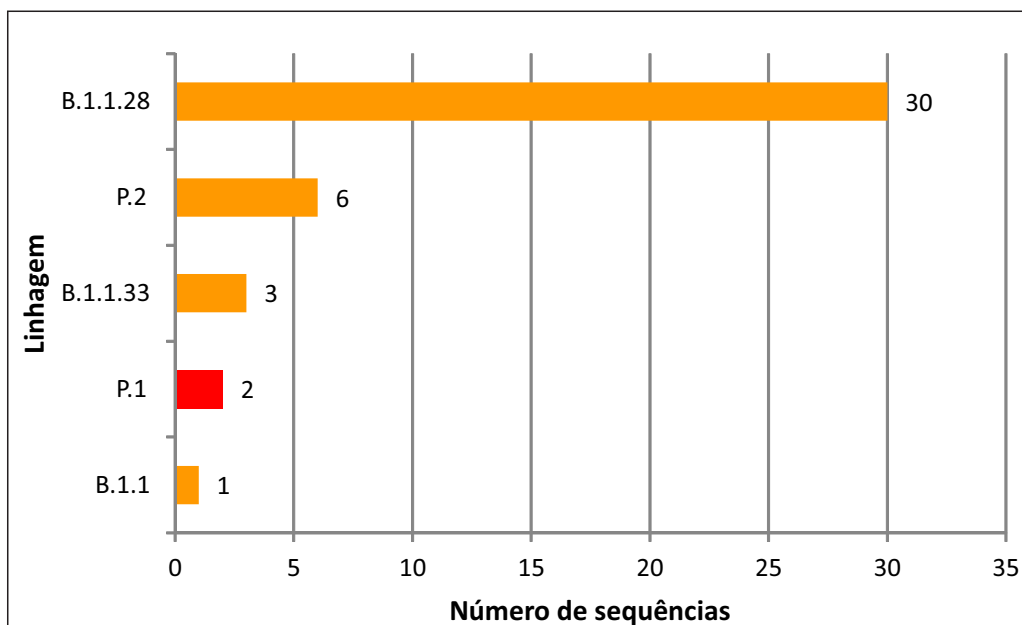
Fonte: GISAID, 2021

Figura 19. Distribuição das linhagens circulantes na DRS XV – São José do Rio Preto



Fonte: GISAID, 2021

Figura 20. Distribuição das linhagens circulantes na DRS XVI – Sorocaba



Fonte: GISAID, 2021

Figura 21. Distribuição das linhagens circulantes na DRS XVII – Taubaté

A partir da obtenção das linhagens circulantes do vírus em cada DRS, o monitoramento passará a ser realizado pela seleção das amostras por meio de análises prospectivas para identificação da ocorrência de casos e acompanhamento da disseminação da doença no espaço e no tempo, identificando municípios que apresentem um alto risco de casos e óbitos em comparação aos municípios vizinhos, para um monitoramento em tempo real.

6. CONSIDERAÇÕES

A análise das sequências com alta qualidade permitiu identificar que no Estado de São Paulo duas linhagens ainda são as mais frequentes, B.1.1.28 e B.1.1.33, as quais somam 63% do total desde o início da pandemia. Mais recentemente, as linhagens de atenção B.1.1.7 e P.1 passaram também a ser detectadas sem vínculo epidemiológico claro com os respectivos locais de origem (Reino Unido ou Manaus, respectivamente),

indicando a transmissão local. Os três primeiros casos da variante brasileira (P.1) sem vínculo epidemiológico óbvio foram identificados na segunda quinzena de janeiro de 2021, no município de Jaú, interior do Estado de São Paulo.

Na análise das sequências provenientes de todo o país, também há predomínio das linhagens B.1.1.28 e B.1.1.33 (que somam 57%), oscilando em frequência nos últimos meses, mas com um aumento recente de outras linhagens, a partir do mês de agosto. A VOC B.1.1.7 já foi detectada nos estados da Bahia, Espírito Santo, Goiás, Minas Gerais, Mato Grosso, Paraná, Rio de Janeiro, Sergipe, São Paulo e também Distrito Federal. Já a VOC P.1 é reportada nos estados de Amazonas, Bahia, Goiás, Minas Gerais, Pará, Paraíba, Roraima, Rio Grande do Sul, São Paulo e Tocantins. As duas VOC encontradas no Brasil apresentam incidência crescente nos meses mais atuais.

REFERÊNCIAS BIBLIOGRÁFICAS

1. Wuhan Municipal Health Committee. Available on: www.wuhan.gov.cn/front/web/showDetail/2019123108989 (accessed 7 March 2020)
2. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China. *N Engl J Med.* 2019; 382:727-33.
3. Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA.* 2020; 323:1061-9
4. Sohrabi C, Alsafi Z, O'Neill N, Khan M, Kerwan A, Al-Jabir A, et al. World health organization declares global emergency: a review of the 2019 novel coronavirus (COVID-19). *Int J Surg* 2020;76:71-6.
5. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. Available on: www.who.int/dg/speeches/detail/whodirector-general-s-opening-remarks-at-the-media-briefing-on-covid-19-11-march-2020 (accessed 11 March 2020).
6. de Wit E, van Doremalen N, Falzarano D, Munster VJ. SARS and MERS: recent

- insights into emerging coronaviruses. *Nature Rev Microbiol.* 2016; 14: 523-34. DOI:10.1038/nrmicro.2016.81
7. Infantino M, Damiani A, Li Gobbi F, Grossi V, Lari B, Macchia D, et al. Serological assays for SARS-Cov-2 infectious diseases: benefits, limitations and perspectives. *IMAJ.* 2020; 22:203-10.
 8. Vashist SK. In vitro diagnostic assays for COVID-19: recent advances and emerging trends. *Diagnostics.* 2020; 10(4):202. DOI: 10.3390/diagnostics10040202
 9. Khailany RA, Safdar M, Ozaslan M. Genomic characterization of a novel SARS-Cov-2. *Gene Reports.* 2020; 19:100682. <https://doi.org/10.1016/j.genrep.2020.100682>
 10. Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DKW, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill.* 2020; 25. DOI: 10.2807/1560-7917.ES.2020.25.3.2000045
 11. Lei J, Li J, Li X, Qi X. CT imaging of the 2019 novel coronavirus (2019-nCoV) pneumonia. *Radiology.* 2020; 295:18.
 12. Loeffelholz MJ, Tang Y-W. Laboratory diagnosis of emerging human coronavirus infections – the state of the art. *Emerg Microb & Infect.* 2020; 9(1):747-56. DOI: 10.1080/22221751.2020.1745095.
 13. Venter M, Richter K. Towards effective diagnostic assays for COVID-19: a review. *J Clin Pathol.* 2020; 73:370-377. DOI:10.1136/jclinpath-2020-206685
 14. Lu R, Zhao X, Li J, Yang B, Wu H, Wang W, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet.* 2020 Feb 22; 395(10224):565-574. DOI: 10.1016/S0140-6736(20)30251-8
 15. Cui J, Li F, Shi Z-L. Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol.* 2019; 17: 181-92.
 16. Benvenuto D, Angeletti S, Giovanetti M, Bianchi M, Pascarella S, Cauda R, et al. Evolutionary analysis of SARS-CoV-2: how mutation of Non-Structural Protein 6 (NSP6) could affect viral autophagy. *J Infect.* 2020; 81(1):e24-e27.
 17. Li Q, Wu J, Nie J, Zhang L, Hao H, Liu S, et al. The impact of mutations in SARS-CoV-2 spike on viral infectivity and antigenicity. *Cell.* 2020; 182:1284-1294.
 18. Gudbjartsson DF, Helgason A, Jonsson H, Magnusson OT, Melsted P, Norddahl GL, et al. Spread of SARS-CoV-2 in the Icelandic Population. *N Engl J Med.* 2020; 382:2302-15. DOI: 10.1056/NEJMoa2006100
 19. Chan JF-W, Yuan S, Kok K-H, To KK-W, Chu H, Yang J, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet.* 2020; 395:514-23. [https://doi.org/10.1016/S0140-6736\(20\)30154-9](https://doi.org/10.1016/S0140-6736(20)30154-9)
 20. Yadav PD, Potdar VA, Choudhary ML, Nyayanit DA, Agrawal M, Jadhav SM, et al. Full-genome sequences of the first two SARS-CoV-2 viruses from India. *Indian J Med Res.* 2020; 151:200-209. Available on: <https://www.ijmr.org.in/text.asp?2020/151/2/200/281471>
 21. Forster P, Forster L, Renfrew C, Forster M. Phylogenetic network analysis of SARS-CoV-genomes. *PNAS.* 2020; 117(17) 9241-3. <https://doi.org/10.1073/pnas.2004999117>
 22. Licastro D, Rajasekharan S, Dal Monego S, Segat L, D'Agaro P, Marcello A. Isolation and Full-Length Genome Characterization of SARS-CoV-2 from COVID-19 Cases in Northern Italy. *J Virol.* 2020; 94(11): e00543-20. DOI: 10.1128/JVI.00543-20

23. Giovanetti M, Angeletti S, Benvenuto D, Ciccozzi M. A doubt of multiple introduction of SARS-CoV-2 in Italy: A preliminary overview. *J Med Virol.* 2020; 92:1634-6. DOI: 10.1002/jmv.25773
24. Helmy YA, Fawzy M, Elasad A, Sobieh A, Kenney SP, Shebata AA. The COVID-19 Pandemic: a comprehensive review of taxonomy, genetics, epidemiology, diagnosis, treatment and control. *J. Clin. Med.* 2020,9(4), 1225; <https://doi.org/10.3390/jcm9041225>. A
25. Furuse Y. Genomic sequencing effort for SARS-CoV-2 by country during the pandemic, *International Journal of Infectious Diseases*, Volume 103, 2021 DOI: 10.1016/j.ijid.2020.12.034
-
-