



NOTA TÉCNICA 2021/01 – REDE GENÔMICA FIOCRUZ / MINISTÉRIO DA SAÚDE

Instituto Oswaldo Cruz (IOC)

Laboratório de Vírus Respiratórios e do Sarampo (LVRS)

Referência Nacional para Coronavírus do Ministério da Saúde e Regional para Organização Mundial de Saúde.

Laboratório de Ecologia de Doenças Transmissíveis na Amazônia, Instituto Leônidas e Maria Deane, Manaus, Amazonas, Brasil.

Fundação de Vigilância em Saúde do Amazonas, Manaus, Amazonas, Brasil.

Laboratório Central de Saúde Pública do Amazonas, Manaus, Amazonas, Brasil.

Coordenação Geral de Laboratórios e Grupo Técnico em Influenza, COVID-19 e outros vírus respiratórios, Secretaria de Vigilância em Saúde, Ministério da Saúde (CGPNI/DEIDT/SVS/MS)

Relação filogenética de sequências SARS-CoV-2 do Amazonas com variantes emergentes brasileiras que abrigam mutações E484K e N501Y na proteína Spike

Felipe Naveca¹, Valdinete Nascimento¹, Victor Souza¹, André Corado¹, Fernanda Nascimento¹, George Silva¹, Ágatha Costa¹, Débora Duarte¹, Karina Pessoa¹, Luciana Gonçalves², Maria Júlia Brandão¹, Michele Jesus³, Cristiano Fernandes², Rosemary Pinto², Marineide Silva⁴, Tirza Mattos⁴, Gabriel Luz Wallau⁵, Marilda Mendonça Siqueira⁶, Paola Cristina Resende^{6*}, Edson Delatorre^{7*}, Tiago Gräf^{8*}, Gonzalo Bello^{9*}

**Esses autores contribuíram igualmente para o estudo*

¹ Laboratório de Ecologia de Doenças Transmissíveis na Amazônia, Instituto Leônidas e Maria Deane, Manaus, Amazonas, Brasil.

² Fundação de Vigilância em Saúde do Amazonas, Manaus, Amazonas, Brasil.

³ Laboratório de Diversidade Microbiana da Amazônia com Importância para a Saúde, Instituto Leônidas e Maria Deane, Manaus, Amazonas, Brasil.

⁴ Laboratório Central de Saúde Pública do Amazonas, Manaus, Amazonas, Brasil.

⁵ Instituto Aggeu Magalhães, Fundação Oswaldo Cruz, Recife, Pernambuco, Brasil.

⁶ Laboratório de Vírus Respiratórios e do Sarampo (LVRS), Instituto Oswaldo Cruz (IOC), Rio de Janeiro, Brasil.

⁷ Departamento de Biologia. Centro de Ciências Exatas, Naturais e da Saúde, Universidade Federal do Espírito Santo, Alegre, Brasil.

⁸ Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brasil.

⁹ Laboratório de AIDS e Imunologia Molecular, Instituto Oswaldo Cruz, FIOCRUZ, Rio de Janeiro, Brasil.

Resumo

Relatamos uma análise genômica preliminar da linhagem SARS-CoV-2 B.1.1.28 que circula na região amazônica brasileira e sua relação evolutiva com variantes emergentes e potenciais variantes emergentes do SARS-CoV-2 no Brasil que abrigam mutações no domínio de ligação com receptor da proteína Spike (S). A análise filogenética de 69 genomas da linhagem B.1.1.28 isoladas no estado do Amazonas revelou a existência de dois clados principais que evoluíram localmente sem mutações incomuns na proteína S de abril a novembro de 2020.

Os vírus B.1.1.28 abrigando mutações S: K417N, S: E484K e S: N501Y, recentemente detectados em viajantes japoneses retornando do Amazonas, ramificados dentro de um dos clados B.1.1.28 da Amazônia aqui identificados, sugerem que essas sequências poderiam ser representantes de um novo clado brasileiro emergente (não relatado), aqui designado B.1.1.28 (K417N / E484K / N501Y). Nossa análise também confirma que o novo clado putativo B.1.1.28 (K417N / E484K / N501Y) detectado em viajantes japoneses não evoluiu do clado B.1.1.28 (E484K) detectado recentemente no Rio de Janeiro e em outros estados brasileiros, mas ambas as variantes surgiram independentemente durante a evolução da linhagem B.1.1.28.

Introdução

O surgimento de novas variantes do SARS-CoV-2 que abrigam mutações na proteína Spike que podem impactar a aptidão viral e a transmissibilidade tem sido uma questão de grande preocupação, particularmente após a recente identificação de duas cepas emergentes independentes no Reino Unido e na África do Sul com um maior que número normal de mutações na proteína Spike que podem ter significado funcional. Ambas as linhagens B.1.1.7 emergentes de SARS-CoV-2 no Reino Unido (HV69-70del, Y144del, N501Y, A570D, P681H, T716I, S982A, D1118H)¹ e B.1.351 na África do Sul (L18F, D80A, D215G, R246I, K417N, E484K, N501Y e A701V)² adquiriu oito substituições de aminoácidos definidores de linhagem na proteína Spike, sendo a mutação N501Y no domínio de ligação ao receptor a única substituição de aminoácidos comum detectada em ambas as linhagens.

A epidemia de SARS-CoV-2 no Brasil foi dominada por duas linhagens designadas como B.1.1.28 e B.1.1.33 que provavelmente surgiram no país em fevereiro de 2020^{3,4}. Relatórios recentes chamam a atenção para o surgimento de novas variantes do SARS-CoV-2 B.1.1.28 no Brasil com mutações na proteína Spike comum às linhagens B.1.1.7 e B.1.351. Um estudo publicado em dezembro de 2020 descreveu o surgimento de um novo clado B.1.1.28 no estado do Rio de Janeiro que se distinguiu por cinco mutações definidoras de linhagem, incluindo uma na proteína Spike (E484K), também detectada no B.1.351 linhagem sul-africana⁵, mas de origem independente. Este clado, aqui designado como B.1.1.28 (E484K), foi posteriormente detectado

em outros estados brasileiros e foi ainda associado a dois casos de reinfecção em pacientes originalmente infectados pela linhagem B.1.1.33⁶⁻⁸.

Mais recentemente, o Ministério da Saúde japonês relatou a presença de uma nova variante SARS-CoV-2 B.1.1.28 em quatro viajantes que chegaram ao Japão retornando do estado do Amazonas em 2 de janeiro de 2021 que abriga 10 mutações sinapomórficas na Spike (L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, H655Y, T1027I), incluindo um (N501Y) detectado nas linhagens B.1.351 e B.1.1.7 e três também detectados na B.1.351 linhagem sul-africana apenas (L18F, K417N, E484K)⁹. A fim de compreender mais profundamente a origem da nova variante B.1.1.28 detectada no Japão, aqui designada como B.1.1.28 (K417N / E484K / N501Y), analisamos a diversidade genética de 148 genomas de SARS-CoV-2 circulando na Amazônia brasileira entre abril e novembro de 2020.

Resultados

Nosso levantamento de 148 sequências do genoma completo SARS-CoV-2 do estado do Amazonas, identificou 69 (47%) sequências B.1.1.28 amostradas de diferentes municípios entre 13 de abril e 13 de novembro de 2020, sendo a linhagem viral mais prevalente em neste estado brasileiro (Figura 1). Estas sequências B.1.1.28 foram alinhadas em seguida com 399 genomas da linhagem B.1.1.28 de alta qualidade (<10% de N e > 29.000 nucleotídeos de comprimento) de origem brasileira e as quatro sequências B.1.1.28 dos viajantes japoneses retornando do Amazonas que estavam disponíveis no banco de dados EpiCoV no GISAID em 10 de janeiro de 2021. Análise filogenética de máxima verossimilhança (ML) usando IQTree v2.1.2¹⁰ revelaram que as sequências B.1.1.28 brasileiras do Amazonas foram principalmente ramificadas em dois clados monofiléticos altamente suportados (teste da razão de verossimilhança aproximada [aLRT]> 80%) aqui designados como 28-AM-I (49%) e 28-AM- II (26%) (Figura 2). O clado monofilético 28-AM-I compreende 34 sequências isoladas em diferentes municípios amazônicos entre 20 de abril e 13 de novembro, agrupadas entre as sequências basais do estado de São Paulo. O clado monofilético 28-AM-II compreende 18 sequências isoladas em diferentes municípios amazônicos entre 20 de abril e 13 de novembro (EPI_ISL_801386 a EPI_ISL_801403) e também compreende as quatro sequências B.1.1.28 japonesas amostradas de viajantes retornando da região amazônica. O clado 28-AM-II foi agrupado dentro de um clado parafilético basal (aLRT = 77%) que compreende sequências amostradas de diferentes estados brasileiros entre 1º de abril e 11 de novembro de 2020 (Figura 2). A emergente linhagem B.1.1.28 (E484K) detectada no Brasil apareceu como um clado monofilético independente altamente suportado (aLRT = 98%) dentro da filogenia B.1.1.28 e compreende sequências de diferentes estados brasileiros, incluindo o único B.1.1.28 sequência do estado do Amazonas contendo a mutação E484K detectada em nosso estudo (Figura 2). Reconstrução bayesiana usando um modelo de relógio molecular estrito com uma taxa de substituição normal anterior ($8-10 \times 10^{-4}$ substituições

/ sítio / ano), conforme implementado no BEAST 1.10¹¹ estimou o surgimento do clado 28-AM-II em 27 de abril (95% de alta densidade posterior [HPD]: 30 de março - 28 de abril).

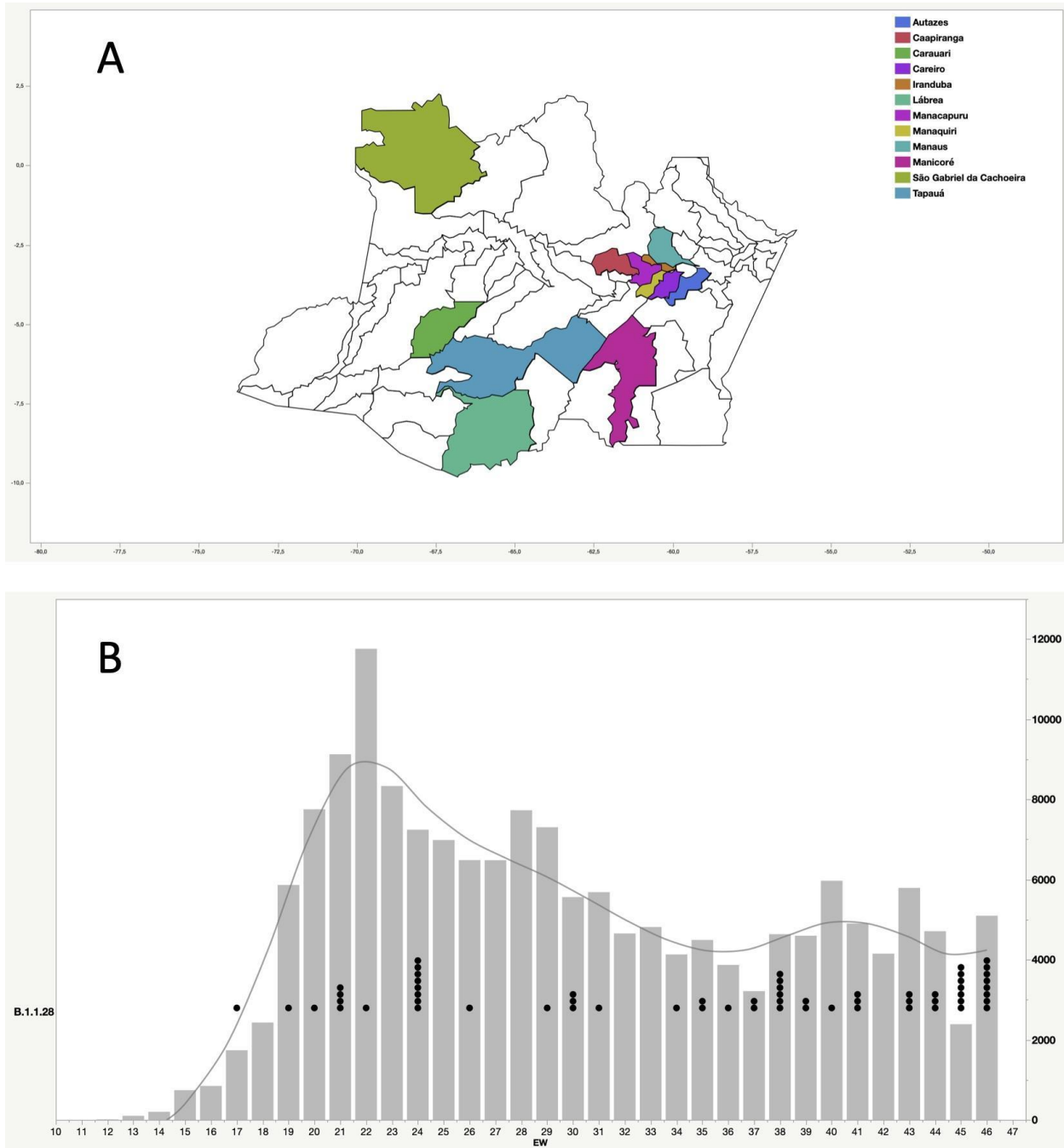


Figura 1. Distribuição geográfica e temporal das sequências SARS-CoV-2 B.1.1.28 amostradas no estado do Amazonas. A) Municípios do estado do Amazonas com amostras SARS-Cov-2 B.1.1.28 sequenciadas neste estudo. B) Sequências SARS-CoV-2 B.1.1.28 obtidas neste estudo (pontos pretos) e casos confirmados de SARS-CoV-2 (barras cinzas) no Amazonas por semana epidemiológica.

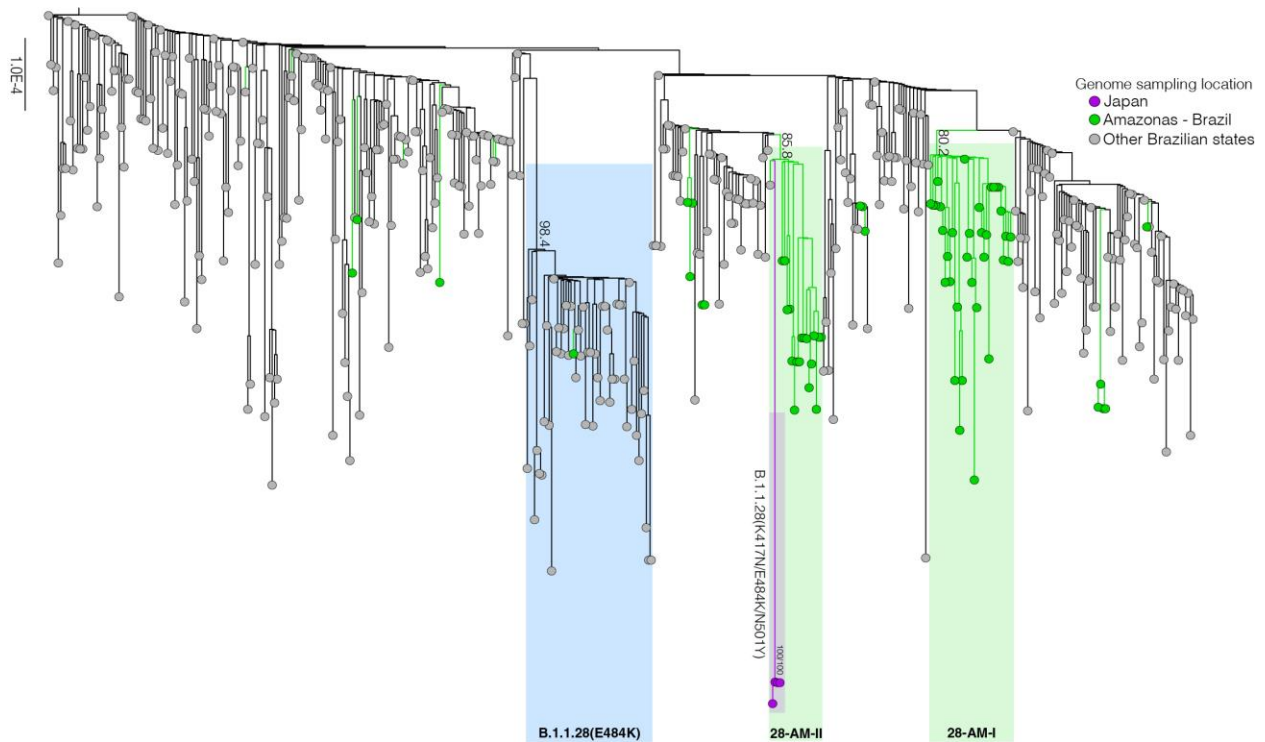


Figura 2. Árvore filogenética de máxima verossimilhança (MV) das sequências do genoma completo de B.1.1.28 do Brasil e do Japão. As sequências do estado do Amazonas são representadas por círculos verdes e as amostradas no Japão pelos roxos. O cluster B.1.1.28 (E484K) é destacado em azul, enquanto os dois clusters identificados no estado do Amazonas (28-AM-I e 28-AM-II) são destacados em verde. Os valores de suporte aLRT são indicados em nós-chave. A árvore foi enraizada na amostra mais antiga e os comprimentos dos ramos são desenhados em escala com a barra esquerda indicando substituições de nucleotídeos por local.

A maioria dos clados B.1.1.28 brasileiros identificados até agora foram caracterizados por um baixo número de mutações sinapomórficas que os distinguem de outras sequências B.1.1.28 detectados no Brasil. Identificamos uma mutação sinônima definidora de linhagem no clado 28-AM-I (C29284T), duas mutações no clado 28-AM-II e sequências basais (A6319G e T26149C [ORF3a: S253P]), uma mutação sinônima exclusiva do clado 28-AM-II (A6613G) e cinco mutações no clado B.1.1.28 (E484K) (C100U [5'UTR], T10667G [NSP5: L205V], C11824T [NSP6], G23012A [S: E484K] e G28628T [N: A119S]) (Tabela). O padrão é consistente com o acúmulo de mutações a uma taxa relativamente constante ao longo do tempo durante a evolução local e diversificação da linhagem B.1.1.28 no Brasil (Figura 3). O ramo que leva às sequências B.1.1.28 amostradas de viajantes japoneses, por outro lado, acumulou um número incomum de alterações genéticas (Figura 3), além daquelas na proteína Spike (Tabela), que se assemelham ao padrão observado em as linhagens B.1.1.7 e B.1.351 do Reino Unido e da África do Sul. Nenhuma das sequências

B.1.1.28 do Amazonas detectadas em nossa pesquisa genômica exibiu um padrão tão divergente de substituições de nucleotídeos e aminoácidos.

Tabela. Mutações definidoras de linhagem dos diferentes clados B.1.1.28 detectados no Brasil e no Japão. As caixas sombreadas destacam a mutação definidora de linhagem B.1.1.28 (cinza escuro) e aquelas comuns aos clados B.1.1.28-AM-II e B.1.1.28 (K417T / E484K / N501Y) (cinza claro).

Genomic Region	Nucleotide / Amino acid		
	B.1.1.28-AM-II	B.1.1.28(K417T/E484K/N501Y)	B.1.1.28(E484K)
ORF1a		T733C C2749T C3828T / ORF1a:S1188L A5648C / ORF1a:K1795Q	C100T
	A6319G A6613G	A6319G A6613G	
ORF1b		C12778T C13860T G17259T / ORF1b:E1264D	T10667G / ORF1a:L3468V C11824T C12053T / ORF1a:L3930F
		C21614T / S:L18F C21621A / S:T20N C21638T / S:P26S G21974T / S:D138Y G22132T / S:R190S A22812C / S:K417T G23012A / S:E484K A23063T / S:N501Y C23525T / S:H655Y C24642T / S:T1027I	G23012A / S:E484K
Spike	G25088T / S:V1176F	G25088T / S:V1176F	G25088T / S:V1176F
ORF3a	T26149C / ORF3a:S253P	T26149C / ORF3a:S253P G28167A / ORF8:E92K	C28253T
		C28512G / N:P80R, ORF9b:Q77E A28877T / N:R203K G28878C / N:R203K	G28628T / N:A119S
N/ORF9b			

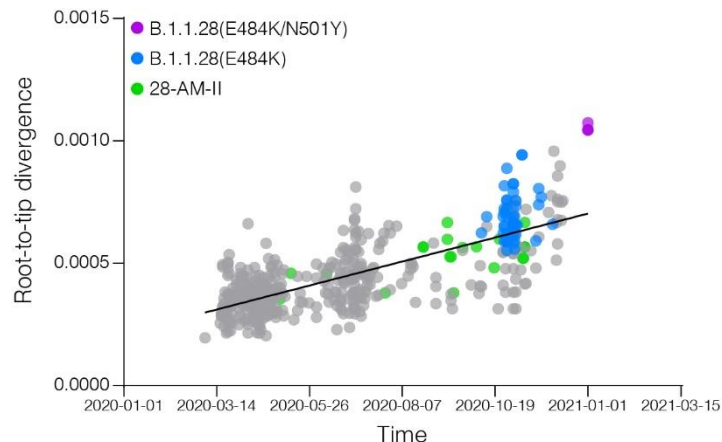


Figura 3. Correlação entre a data de amostragem das sequências B.1.1.28 e sua distância genética da raiz da árvore filogenética ML. As cores indicam o clado B.1.1.28 das sequências correspondentes de acordo com a legenda à esquerda.

Esses achados suportam que as cepas SARS-CoV-2 B.1.1.28 detectadas em viajantes japoneses retornando da região amazônica provavelmente evoluíram de uma linhagem viral que circula no estado do Amazonas desde abril de 2020, e podem ser representantes de um vírus potencialmente de uma linhagem emergente no Brasil, aqui designada como B.1.1.28 (K417N / E484K / N501Y). Local B.1.1.28 linhagens amazônicas parecem ter evoluído a uma taxa constante entre abril e novembro de 2020 e nenhuma das sequências aqui obtidas exibiu um número tão alto de mutações na Spike ou em qualquer outra região genômica. Esses achados suportam que a rápida taxa de mutação detectada na variante B.1.1.28 (K417N / E484K / N501Y) é um fenômeno recente que provavelmente ocorreu entre dezembro de 2020 - janeiro de 2021. Com o objetivo de detectar a circulação desta linhagem entre a Amazônia população, estamos atualmente conduzindo um levantamento genômico de indivíduos recentemente infectados com SARS-CoV-2 no Amazonas. Nossa análise preliminar também confirma que as linhagens brasileiras emergentes B.1.1.28 (E484K) e B.1.1.28 (K417N / E484K / N501Y) surgiram independentemente durante a diversificação da linhagem B.1.1.28 no Brasil. O surgimento simultâneo de diferentes linhagens B.1.1 virais que carregam mutações K417N, E484K e N501Y no domínio de ligação do receptor da proteína Spike em diferentes países ao redor do mundo durante a segunda metade de 2020 sugere mudanças seletivas convergentes na evolução de SARS-CoV-2 devido a similar pressão evolutiva durante o processo de infecção de milhões de pessoas. Se essas mutações conferem alguma vantagem seletiva para a transmissibilidade viral, devemos esperar um aumento da frequência dessas linhagens virais no Brasil e no mundo nos próximos meses.

Agradecimentos

Os autores desejam agradecer a todos os profissionais de saúde e cientistas que trabalharam duro para lidar com esta ameaça de pandemia, a equipe GISAID e todos os remetentes do banco de dados EpiCoV, em particular os membros do Instituto Nacional Japonês de Doenças Infecciosas (NIID), Dr. Tsuyoshi Sekizuka, Dr. Kentaro Itokawa, Rina Tanaka e Masanori Hashino para publicar os genomas. A tabela de reconhecimento GISAID contendo as sequências usadas neste estudo está anexada a este post (Tabela Suplementar 1). Gostaríamos também de agradecer ao Dr. Nuno Faria de <https://www.caddecentre.org> por compartilhar suas descobertas não publicadas sobre a linhagem SARS-CoV-2 B.1.1.28. Agradecemos também o apoio dos membros da Rede Fiocruz de Coronavírus Genômico e da Rede de Vigilância Genômica de Vírus Respiratórios da Coordenação Geral de Laboratórios (CGLab) do Ministério da Saúde (MS), dos Estados Laboratórios Centrais do Brasil (LACENS) e das equipes de vigilância do Amazonas para a parceria na vigilância viral no Brasil. Apoio financeiro FAPEAM (PCTI-EmergeSaude/AM chamada 005/2020 e Rede Genômica de Vigilância em Saúde - REGESAM); Conselho Nacional de Desenvolvimento Científico e Tecnológico (Grant 403276/2020-9); Inova Fiocruz/Fundação Oswaldo Cruz (Grant VPPCB-007-FIO-18-2-30 - Geração de conhecimento).

Referência

- 1 Rambaut, A. L., N.; Pybus, O.; Barclay, W.; Barrett, J.; Carabelli, A.; Connor, T.; Peacock, T.; Robertson, D.; Volz, E.; on behalf of COVID-19 Genomics Consortium UK (CoG-UK); . Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. *Virological.org* (2020). <<https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563>>.
- 2 Tegally, H. *et al.* Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. *medRxiv*, doi:10.1101/2020.12.21.20248640 (2020).
- 3 Candido, D. S. *et al.* Evolution and epidemic spread of SARS-CoV-2 in Brazil. *Science* **369**, 1255-1260, doi:10.1126/science.abd2161 (2020).
- 4 Resende P. C., D., E., Gräf T., Mir D., Motta F.C., Appolinario L., Paixão A. C., Mendonça A. C., Ogrzewalska M., Caetano B., Wallau G. L., Docena C., Santos M. C., Ferreira J., Sousa Junior E., Silva S., Fernandes S., Vianna L. A., Souza L., Ferro J. F, Nardy V., Santos C., Riediger I., Debur M., Croda J., Oliveira, W, Abreu A,, Bello G.. Siqueira M. M. Evolutionary dynamics and dissemination pattern of the SARS-CoV-2 lineage B.1.1.33 during the early pandemic phase in Brazil. *Frontier in Microbiology*, doi:10.3389/fmicb.2020.615280 (2020).
- 5 Voloch, C. M. *et al.* Genomic characterization of a novel SARS-CoV-2 lineage from Rio de Janeiro, Brazil. *medRxiv*, doi:10.1101/2020.12.23.20248598 (2020).
- 6 Resende, P. C. B., J.F.; Vasconcelos, R.H.T.; Arantes I.; Appolinario L.; Mendonça, A.C.; Paixao, A.C.; Rodrigues A.C.; Silva, T.; Rocha, A.S.; Pauvolid-Corrêa, A.; Motta, F.C.; Teixeira, D.L.F.T.; Carneiro, T.F.O.; Freire Neto, F.P.F.; Herbster, I.D.; Leite, A.B.; Riediger, I.N.; Debur, M.C.; Naveca, F.G.;

- Almeida, W.; Livorati, M.; Bello, G.; Siqueira, M.M. Spike E484K mutation in the first SARS-CoV-2 reinfection case confirmed in Brazil, 2020. *Virological.org* (2020). <<https://virological.org/t/spike-e484k-mutation-in-the-first-sars-cov-2-reinfection-case-confirmed-in-brazil-2020/584>>.
- 7 Vasques Nonaka, C. K. M. F., M.; Gräf, T.; Almeida Mendes, A.V.; Santana de Aguiar, R.; Giovanetti, M.; Solano de Freitas Souza, B. Genomic Evidence of a Sars-Cov-2 Reinfection Case With E484K Spike Mutation in Brazil. *Preprints*, doi:10.20944/preprints202101.0132.v1 (2021).
- 8 Brazilian Ministry of Health. *Ministério da Saúde confirma primeiro caso de reinfeção por Covid-19*, <<https://www.gov.br/saude/pt-br/assuntos/noticias/ministerio-da-saude-confirma-primeiro-caso-de-reinfeccao-por-covid-19>> (2020).
- 9 Japan., N. I. o. I. D. ブラジルからの帰国者から検出された新型コロナウイルスの新規変異株について. *Coronavirus disease (COVID-19)*, 4 (2021). <<https://www.niid.go.jp/niid/images/epi/corona/covid19-33-20210110.pdf>>.
- 10 Minh, B. Q. *et al.* IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol Biol Evol* **37**, 1530-1534, doi:10.1093/molbev/msaa015 (2020).
- 11 Suchard, M. A. *et al.* Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol* **4**, vey016, doi:10.1093/ve/vey016 (2018).