



Diversidade genética do novo coronavírus SARS-CoV-2 (COVID-19) em Portugal

Mais informações em <https://insaflu.insa.pt/covid19/>

Relatório de situação

5 de Fevereiro de 2021

O Instituto Nacional de Saúde Doutor Ricardo Jorge, I.P. (INSA) analisou até à data **3261 sequências do genoma do novo coronavírus SARS-CoV-2**, obtidas de amostras colhidas em 71 laboratórios/hospitais/instituições representando 235 concelhos. 1054 e 44 sequências foram geradas nas instituições colaboradoras **Instituto Gulbenkian de Ciência (IGC)** e **Institute of Biomedicine - iBiMED (Universidade de Aveiro)**, respectivamente.

Desde o último relatório de 12 de Janeiro de 2021, **foram analisadas mais 919 sequências (Figura 1), incluindo:**

- i) **532 sequências obtidas de amostras colhidas entre 10 e 19 de Janeiro de 2021 por laboratórios distribuídos pelas duas Regiões Autónomas (Madeira e Açores) e por todos os distritos de Portugal continental (no total de 138 concelhos)**, no âmbito da vigilância de periodicidade mensal com amostragem nacional que o INSA está a coordenar.
- ii) **199 amostras obtidas no âmbito da solicitação do Ministério da Saúde para que fossem enviadas para o INSA amostras suspeitas da presença das variantes 501Y.V1 (VOC-202012/01; linhagem B.1.1.7, associada ao Reino Unido), 501Y.V2 (linhagem B.1.351, associada à África do Sul) e 501Y.V3 (linhagem P.1, associada ao Brasil) ou de outros estudos específicos (ex., surtos hospitalares).**
- iii) 188 sequências obtidas de amostras colhidas entre 23 e 28 de Dezembro de 2020, no âmbito de uma colaboração específica entre o Instituto Gulbenkian de Ciência (IGC) e o laboratório Germano de Sousa.

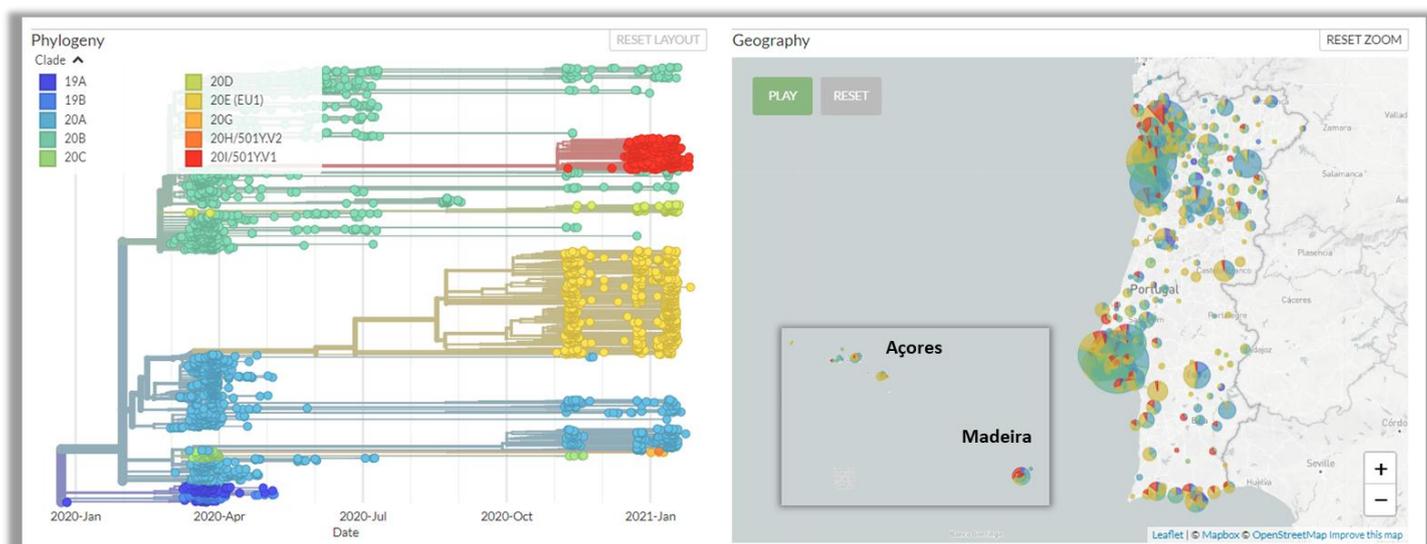


Figura 1. Visão global da diversidade genética e dispersão geotemporal do vírus SARS-CoV-2 em Portugal. Os diferentes genomas (representados por círculos no painel à esquerda) estão coloridos de acordo com a classificação de Clade e com a mesma tonalidade no mapa – o tamanho dos círculos no mapa é proporcional ao número de genomas sequenciados por concelho (consultar o site <https://insaflu.insa.pt/covid19/> para mais detalhes).



Diversidade genética do novo coronavírus SARS-CoV-2 (COVID-19) em Portugal



A Tabela seguinte apresenta a **frequência relativa das principais variantes genéticas do SARS-CoV-2 detectadas na amostragem nacional de 10 a 19 Janeiro de 2021 (n=532)**, bem como o número total de sequências dessas variantes detectadas até à data (n = 3261).

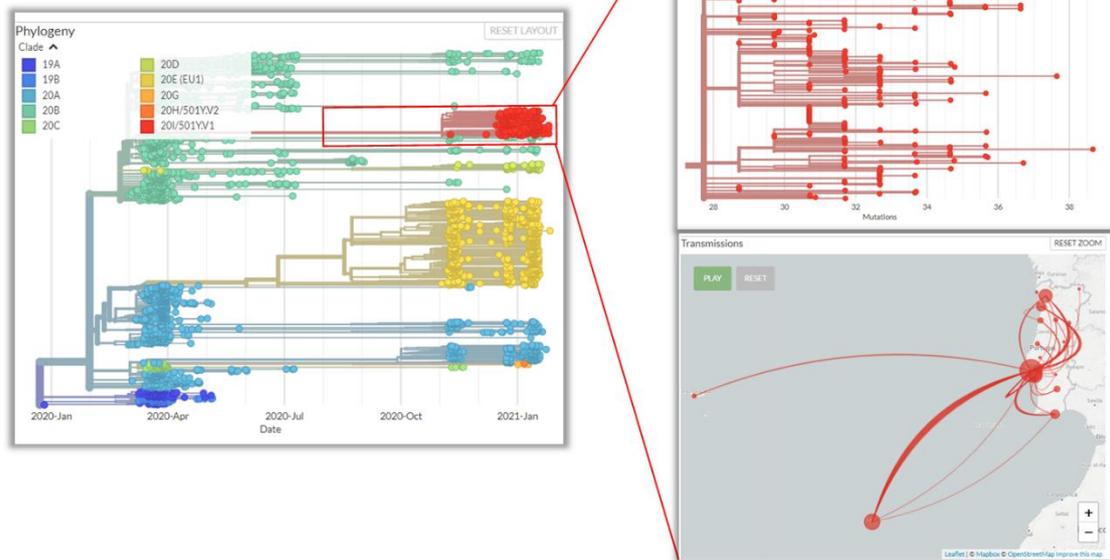
Variante / linhagem	Frequência relativa na amostragem nacional, 10-19 Janeiro, 2021 (n=532) ^a	Total de sequências até à data (n = 3261)	CONTEXTO
S: D614G+A222V (clade "20E / EU1"; linhagem e sub-linhagens B.1.177)	54,70%	782	- Este cluster (sub-clade 20A.EU1) terá tido origem em Espanha (em meados de Junho), tendo revelado uma elevada disseminação a nível global. Hodcroft et, al, medRxiv. https://www.medrxiv.org/content/10.1101/2020.10.25.20219063v2
S: D614G+N501Y+H69del/V70del+ + (clade 20I; VOC-202012/01; 501Y.V1; linhagem B.1.1.7)	16,0%	235	- A mutação A222V na Spike altera um potencial local antigénico (epítipo). Mateus et al, 2020, Science. https://science.sciencemag.org/content/370/6512/89 - Uma variante com esta mutação A222V este na origem de um caso de re-infecção descrito em Hong Kong. To et al, 2020, CID. https://academic.oup.com/cid/advance-article/doi/10.1093/cid/ciaa1275/5897019
S: D614G+S477N (sub-clade "20A.EU2"; linhagem B.1.160)	13,2%	162	- Esta variante apresenta múltiplas mutações na proteína Spike, potencialmente mediadoras de uma maior capacidade de transmissão. Foi detectada em Setembro no Reino Unido e, desde então, em múltiplos países. https://www.ecdc.europa.eu/sites/default/files/documents/SARS-CoV-2-variant-multiple-spike-protein-mutations-United-Kingdom.pdf
S: D614G+L452R+ + (sub-clade 20D)	6,8%	57	- Esta variante apresenta uma deleção no gene S, a qual está associada à falha na detecção deste gene em alguns testes de RT-PCR. https://www.ecdc.europa.eu/sites/default/files/documents/SARS-CoV-2-variant-multiple-spike-protein-mutations-United-Kingdom.pdf
S: D614G+N439Y+H69del/V70del (sub-clade 20A; sub-linhagens B.1.258)	1,7%	20	- Este cluster (sub-clade 20A.EU2) apresenta uma considerável disseminação internacional. Hodcroft et, al, medRxiv. https://www.medrxiv.org/content/10.1101/2020.10.25.20219063v2 - A mutação S477N na Spike altera o domínio de ligação ao receptor ("RBD") ACE2, podendo aumentar a afinidade da ligação da Spike às células do hospedeiro. Starr et al, 2020, Cell. https://www.sciencedirect.com/science/article/pii/S0092867420310035
S: D614G+S98F (sub-clade 20A; linhagem B.1.221)	1,1%	29	- Esta variante apresenta a mutação L452R, a qual se prevê afectar o domínio de ligação da proteína Spike ao receptor ("RBD") ACE2, podendo mediar resistência a anticorpos neutralizantes. Li et al, 2020, Cell. https://www.sciencedirect.com/science/article/abs/pii/S0092867420308771
S: D614G+E484K+ + (clade 20B; linhagem P.2)	0,6%	5	- A mutação N439Y na Spike altera o domínio de ligação ao receptor ("RBD") ACE2, podendo aumentar a afinidade da ligação da Spike às células do hospedeiro e alterar o reconhecimento por anticorpos. Weisblum et al, 2020, Elife. https://elifesciences.org/articles/61312 ; Barnes et al, 2020, Nature https://www.nature.com/articles/s41586-020-2852-1 ; Thomson et al, 2020, https://www.biorxiv.org/content/10.1101/2020.11.04.355842v1
S: D614G+N501Y+E484K+ + (clade 20H; 501Y.V2; linhagem B.1.351)	0%	2	- Esta variante apresenta uma deleção no gene S, a qual está associada à falha na detecção deste gene em alguns testes de RT-PCR. https://www.ecdc.europa.eu/sites/default/files/documents/SARS-CoV-2-variant-multiple-spike-protein-mutations-United-Kingdom.pdf
OUTRAS	5,9%	1969	- Esta variante apresentou maior frequência na Bélgica e Holanda, tendo já sido identificada em múltiplos países. Hodcroft et, al, medRxiv. https://www.medrxiv.org/content/10.1101/2020.10.25.20219063v2 ; https://github.com/emmahodcroft/cluster_scripts/blob/master/README.md#sn501 Esta variante apresenta uma mutação no domínio de ligação ao receptor ("RBD") ACE2 da proteína Spike, a qual é potencialmente mediadora de resistência a anticorpos neutralizantes (S: E484K). Foi recentemente associada a casos de re-infecção no Brasil Resende et al, 2021, virological. https://virological.org/t/spike-e484k-mutation-in-the-first-sars-cov-2-reinfection-case-confirmed-in-brazil-2020/584 - Esta variante apresenta múltiplas mutações na proteína Spike, potencialmente mediadoras de uma maior capacidade de transmissão (ex, S: N501Y) e de resistência a anticorpos neutralizantes (ex, S: E484K). Foi detectada pela primeira vez na África do Sul e, desde então, em múltiplos países. Tegally et al, 2020, medRxiv. https://www.medrxiv.org/content/10.1101/2020.12.21.20248640v1



Principais destaques:

- A variante **501Y.V1** (VOC-202012/01; linhagem B.1.1.7), associada ao **Reino Unido**, foi **detetada por sequenciação com uma frequência de 16% na amostragem nacional** de 10 e 19 Janeiro, 2021, num total de 532 sequências analisadas. Esta **frequência relativa é concordante com a que foi estimada a partir dos dados de falha na detecção do gene S por RT-PCR para a mesma semana**, no âmbito do estudo de monitorização contínua desta variante em colaboração com a Direção Geral da Saúde e o laboratório Unilabs, Portugal (<https://virological.org/t/tracking-sars-cov-2-voc-202012-01-lineage-b-1-1-7-dissemination-in-portugal-insights-from-nationwide-rt-pcr-spike-gene-drop-out-data/600>). Os resultados obtidos pelas duas estratégias de vigilância indicam que esta variante está **amplamente dispersa por todo o território nacional (Figura 2)**.
- Foram detetados, até à data, **dois casos da variante 501Y.V2** (linhagem B.1.351), associada à África do Sul, sendo que não se detetou nenhum caso desta variante na amostragem nacional do período 10-19 de Janeiro, 2021. Esta observação sugere que **a circulação desta variante é ainda limitada em Portugal**.
- Não foi detetado **nenhum caso associado à variante 501Y.V3 (P.1)**, primeiramente detetada no **Brasil**, nomeadamente na região de Manaus, Amazónia. Contudo, foram detetados **cinco casos da variante P.2**, também detetada inicialmente no Brasil e associada a casos de reinfeção.
- Destaca-se ainda a emergência de uma **variante com a mutação de interesse L452R na proteína Spike**, tendo sido detetada com uma **frequência relativa de 6.8%** na amostragem nacional do período 10-19 de Janeiro, 2021. Esta variante foi associada a três casos na amostragem nacional de Novembro, tendo agora um total de **57 sequências detetadas em 32 concelhos, abrangendo 9 distritos de Portugal continental e a Região Autónoma dos Açores**, indicando alguma dispersão a nível nacional (**Figura 2**). Dentro deste cluster foram detetados dois sub-clusters com as mutações adicionais **P681H** (comum também à variante **501Y.V1**) e **Q675H**, ambas próximas de uma região da proteína Spike com elevada relevância funcional (*furin cleavage site*).
- **As actividades de vigilância laboratorial do SARS-CoV-2 continuarão em articulação com as autoridades de Saúde, mantendo especial foco na detecção de novas introduções e monitorização de variantes a suscitar particular interesse pela comunidade científica e autoridades de Saúde.**
- **Neste âmbito, destaca-se a recente publicação do Diário da República (Despacho n.º 331/2021 - Diário da República n.º 6/2021, Série II de 2021-01-11), a qual determina o reforço da vigilância laboratorial genética e antigénica do vírus SARS-CoV-2, sob coordenação do INSA.**

Variante 501Y.V1



Variante com mutação L452R

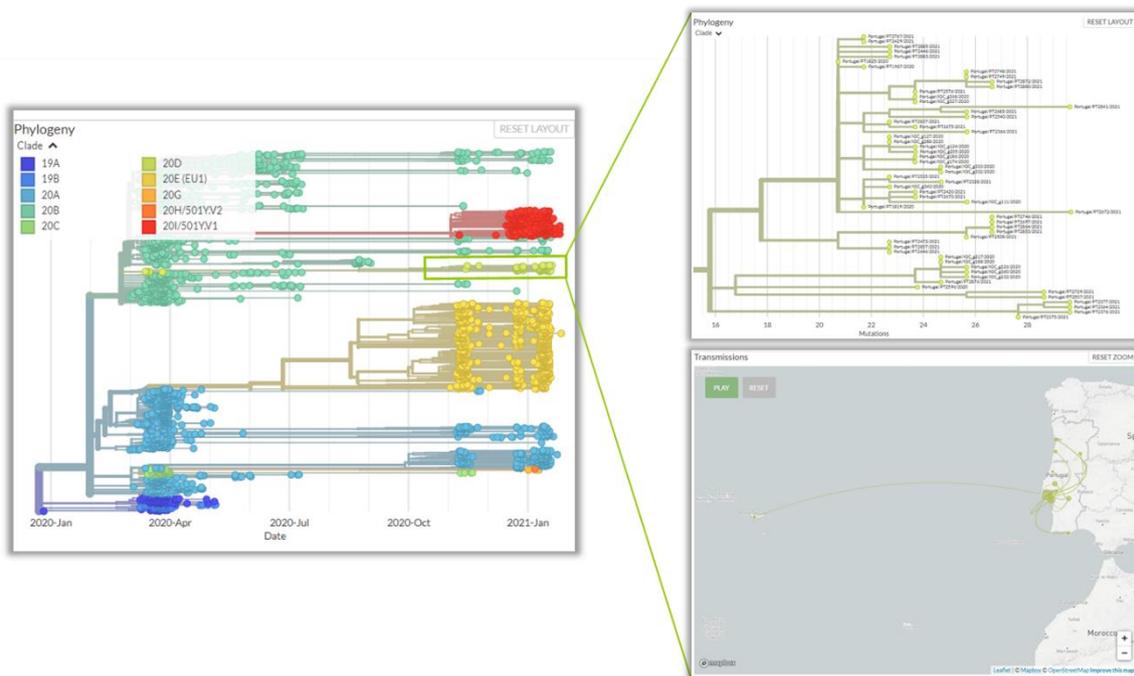


Figura 2. Enquadramento da variante 501Y.V1 (painel superior), associada ao Reino Unido (VOC-202012/01; linhagem B.1.1.7), e da variante emergente com a mutação L452R (painel inferior) no contexto da diversidade genética de SARS-CoV-2 em Portugal até à data. As árvores filogenéticas e os mapas detalham a diversidade genética destas variantes detectada em Portugal e a sua distribuição geográfica (<https://insaflu.insa.pt/covid19/>).



NOTAS ADICIONAIS (também disponível em <https://insaflu.insa.pt/covid19/>)

Objectivos gerais do estudo:

- Determinação dos **padrões de disseminação do vírus nas diferentes regiões de Portugal** e em diferentes grupos populacionais.
- Determinação dos perfis mutacionais do SARS-CoV-2 para **identificação e monitorização de cadeias de transmissão**, bem como **identificação de novas introduções do vírus** em Portugal.
- Prever o início da transmissão na comunidade e aferir o impacto das medidas de contenção, avaliando a **contribuição da transmissão local versus importações do vírus**.
- Determinação do **grau de variabilidade genética de antígenos ou alvos de fármacos antivirais** com possível impacto no desenvolvimento / eficiência de medidas profiláticas (vacinas) e terapêuticas, bom como **monitorização das mutações em alvos genéticos de testes de diagnóstico**.
- Determinação de possíveis **associações entre perfis genéticos (mutacionais)** do SARS-CoV-2 e **determinadas manifestações clínicas** (ex. diferentes graus severidade da COVID-19) ou diferente **capacidade de transmissão** do vírus.
- Estudar os **mecanismos evolutivos do vírus** e a sua relação com os perfis de disseminação em diferentes regiões de Portugal e em diferentes grupos populacionais.
- **Contribuir para a avaliação da relevância funcional e fenotípica** de mutações particulares.

Métodos

- **Procedimento Pre-NGS:** adaptado da Artic Network (<https://artic.network/ncov-2019>, <https://www.protocols.io/view/ncov-2019-sequencing-protocol-bbmuik6w>)
- **Procedimento NGS:** Nextera XT e MiSeq (Illumina)
- **Dos "reads" às sequências do genoma:** [INSaFLU](#)
- **Das sequências do genoma à "filogeografia*":** [Nextstrain](#) (mais detalhes sobre o método em <https://nextstrain.org/ncov> e <https://github.com/nextstrain/ncov>) e [Microreact](#).

** O posicionamento geográfico reflecte o local (Distrito – "Admin Division" ou Concelho – "Location") de residência ou, caso não exista informação, local de ocorrência ou da entidade que enviou a amostra.*

Agradecimentos

- A todos os laboratórios nacionais que enviam amostras clínicas (suspeitas ou positivas para SARS-CoV-2) para o Laboratório Nacional de Referência da Gripe e outros vírus respiratórios do INSA.
- À Sara Hill e Nuno Faria (Universidade de Oxford) e Joshua Quick e Nick Loman (Universidade de Birmingham) por nos terem gentilmente cedido os primers usados no início deste estudo.
- Às equipas dos projectos [Nextstrain](#) e [Microreact](#) pela libertação de algumas ferramentas de bioinformática usadas neste estudo.
- Ao Miguel Pinheiro (iBiMED / Universidade de Aveiro) pelo seu trabalho na atualização da plataforma [INSaFLU](#) para o novo coronavírus SARS-CoV-2.
- À [Infraestrutura Nacional de Computação Distribuída \(INCD\)](#), por ter fornecido recursos computacionais para testar a plataforma INSaFLU. O INCD foi financiado pela FCT e FEDER sob o projeto 22153-01 / SAICT / 2016.

Este estudo é co-financiado pela Fundação para a Ciência e Tecnologia e Agência de Investigação Clínica e Inovação Biomédica no âmbito da primeira e segunda "calls" do programa Research 4 COVID-19 (projectos 234 e 657, respectivamente).