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

Mais informações em **https://insaflu.insa.pt/covid19/**

Relatório de situação — 12 de Janeiro de 2021

O **Instituto Nacional de Saúde Doutor Ricardo Jorge, I.P. (INSA)** analisou até à data **2342 sequências do genoma do novo coronavírus SARS-CoV-2**, obtidas de amostras colhidas em 69 laboratórios/hospitais/instituições representando 199 concelhos. 593 sequências foram geradas no **Instituto Gulbenkian de Ciência (IGC)**, instituição colaboradora do INSA neste estudo.

Nesta actualização do site foram inseridas **mais 55 sequências** obtidas no âmbito da solicitação do Ministério da Saúde para que fosse pesquisada a presença da nova variante recentemente identificada no Reino Unido (VOC-202012/01; N501Y.V1; linhagem B.1.1.7) em amostras suspeitas, incluindo amostras em que foi verificada a falha na detecção do gene S por RT-PCR e amostras associadas a casos positivos de COVID-19 com historial de viagem.

Entre as novas sequências, destacam-se **38 novas sequências da variante recentemente identificada no Reino Unido (VOC-202012/01; N501Y.V1; linhagem B.1.1.7)* (Figura 1),** as quais são referentes a amostras colhidas nos aeroportos de Lisboa (n=8) e Porto (n=11) e a amostras provenientes de todas as regiões de Saúde do país (com excepção da Região Autónoma da Madeira).

Balanço total - variante (VOC-202012/01; N501Y.V1; linhagem B.1.1.7):

Foram detectadas no território português até à data um total de **72 casos de infecção associados a esta nova variante**, distribuindo-se pelas duas Regiões Autónomas (Madeia e Açores) e por 10 distritos de Portugal continental, num total de 28 concelhos.

A diversidade genética e dispersão geográfica desta variante (Figura 1) é concordante com a ocorrência de múltiplas introduções independentes e aponta para a existência de transmissão comunitária. Estes resultados foram reportados às entidades de Saúde Pública para que sejam monitorizados potenciais contactos e cadeias de transmissão.

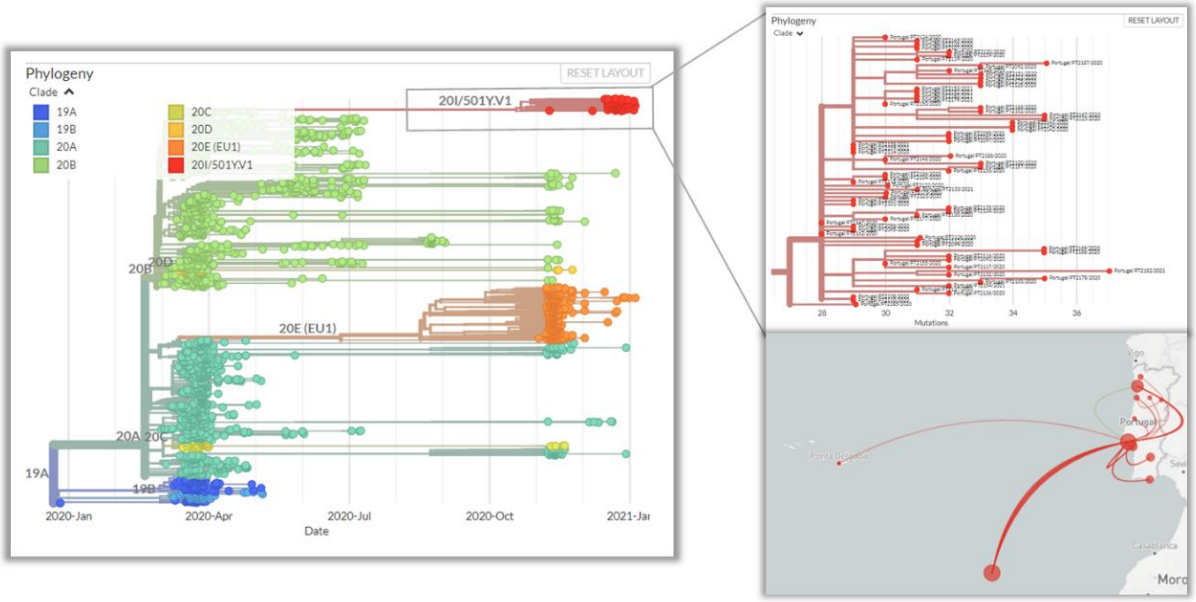

Figura 1. **Enquadramento da nova variante recentemente identificada no Reino Unido (VOC-202012/01; N501Y.V1; linhagem B.1.1.7) no contexto da diversidade genética de SARS-CoV-2 em Portugal até à data** (a árvore filogenética e o mapa à direita detalham a diversidade genética desta variante detectada em Portugal e a sua distribuição geográfica). Até à data, foram analisadas 75 sequências desta variante (associadas a 72 casos de infecção). *Para duas destas 75 amostras foi apenas obtida a sequência parcial do genoma viral, pelo que não foram integradas na árvore filogenética (https://insaflu.insa.pt/covid19/).

Núcleo de Bioinformática | Departamento de Doenças Infecciosas
INSTITUTO NACIONAL DE SAÚDE DOUTOR RICARDO JORGE
Avenida Padre Cruz, 1649-016 Lisboa, PORTUGAL

A Tabela seguinte e a Figura 2 apresentam a distribuição e frequência actualizadas das principais variantes genéticas do SARS-CoV-2 detectadas em Portugal desde o ínicio de Novembro (n=512).

Variante	PROPORÇÃO (n=512) [a]	CONTEXTO
S: D614G+**A222V** **(clade "20E / EU1"; linhagem B.1.177)**	**307 (60,0%)**	- **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
		- A mutacão A222V na Spike altera um potencial local antigénico (epítopo). 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)**	**39 (7,6%)**	- **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 mutacão S477N na Spike altera o domínio de ligação ao receptor ("RBD") ACE2, pondendo 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+**N501Y+H69del/V70del++** (clade 20I; **VOC-202012/01**; N501Y.V1; **linhagem B.1.1.7**)	**75 (72 casos) (15,2%)[a]**	**- 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
		- Esta variante apresenta uma delecção 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+**S98F** **(sub-clade 20A; linhagem B.1.221)**	**20 (3,9%)**	**- Esta variante é encontrada com 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
S: D614G+**N439Y+H69del/V70del** **(sub-clade 20A; linhagem B.1.258)**	**7 (1,4%)[b]**	**- A mutacão N439Y na Spike altera o domínio de ligação ao receptor ("RBD") ACE2, pondendo aumentar a afinidade da ligação da Spike às células do hospedeiro e alterar a 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
		- Esta variante apresenta uma delecção 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+**D80Y** **(sub-clade 20C; linhagem B.1.367)**	**4 (0,8%)**	**- Esta variante é encontrada com maior frequência na França, Noruega 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
OUTRAS	**60 (11,7%)**	----

[a] Até à data, foram analisadas 75 sequências da variante VOC-202012/01 / linhagem B.1.1.7 (associadas a 72 casos de infecção). Para duas destas 75 amostras (e, portanto, duas das 512 amostras analisadas desde o ínicio de Novembro), foi apenas obtida a sequência parcial do genoma viral, pelo que não foram integradas na árvore filogenética (https://insaflu.insa.pt/covid19/).

[b] Para 3 destas 7 sequências apenas foi possível obter parcialmente a sequência do gene S. Contudo, a partilha de outras mutações aponta que o posicionamento filogético e classificação dessa sequência como "S: D614G+N439Y+H69del/V70del (linhagem B.1.258) " é correcto.

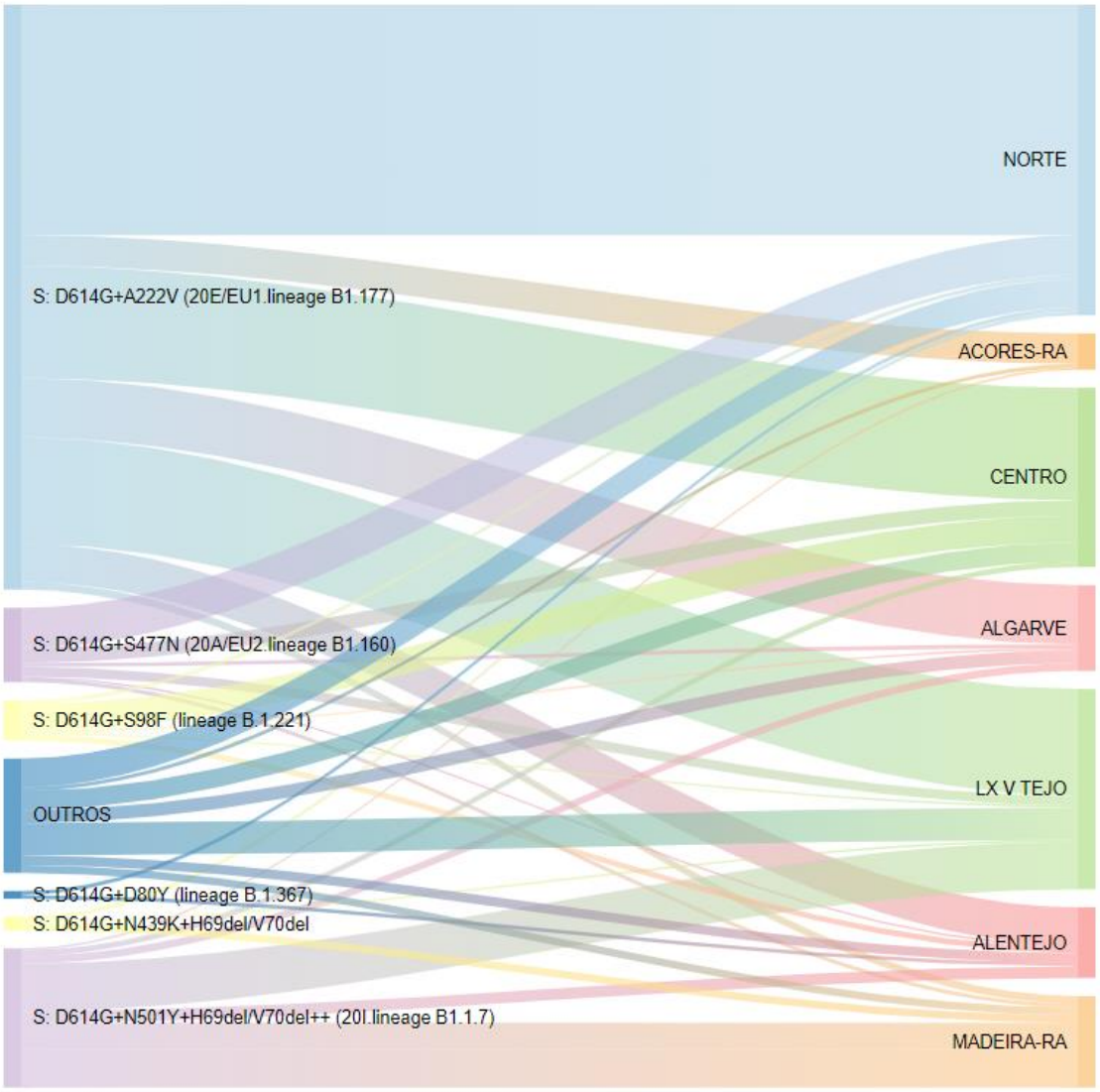

Figura 2. **Frequência das principais variantes genéticas do SARS-CoV-2 detectadas em Portugal desde o ínicio de Novembro (à esquerda) e sua distribuição por região (à direita).**

- **O INSA, em colaboração com o IGC, prosseguirá com as actividades de vigilância laboratorial do SARS-CoV-2 em articulação com as autoridades de Saúde, mantendo especial foco na detecção de novas introduções e monitorização da circulação da variante VOC-202012/01 (N501Y.V1; linhagem B.1.1.7), bem como de outras 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.**

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

Objectivos gerais do estudo:

- Determinação dos **padrões de disseminação do vírus nas dife rentes 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 antigénios 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 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 ínicio 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).

Núcleo de Bioinformática | Departamento de Doenças Infecciosas
INSTITUTO NACIONAL DE SAÚDE DOUTOR RICARDO JORGE
Avenida Padre Cruz, 1649-016 Lisboa, PORTUGAL