Characterization of genetic variants of SARS-CoV-2 isolated from cadavers: comparative and validation study between RT-PCR and NGS methodologies

Magda Isabel Cerdeira Franco

Mestrado em Genética Forense

Departamento de Biologia

2022

Orientadora

Laura Cainé, MD, PhD, Faculdade de Medicina da Universidade do Porto (FMUP), Instituto Nacional de Medicina Legal e Ciências Forenses, I.P. (INMLCF)

Coorientador

António Amorim, PhD, Faculdade de Ciências da Universidade de Lisboa (FCUL), Instituto Nacional de Medicina Legal e Ciências Forenses, I.P. (INMLCF)

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.
O Presidente do Júri,

Porto, _____/_____/_____

Declaração de Honra

Eu, Magda Isabel Cerdeira Franco, inscrita no Mestrado em Genética Forense da Faculdade de Ciências da Universidade do Porto declaro, nos termos do disposto na alínea a) do artigo 14.º do Código Ético de Conduta Académica da U.Porto, que o conteúdo da presente dissertação reflete as perspetivas, o trabalho de investigação e as minhas interpretações no momento da sua entrega.

Ao entregar esta dissertação, declaro, ainda, que a mesma é resultado do meu próprio trabalho de investigação e contém contributos que não foram utilizados previamente noutros trabalhos apresentados a esta ou outra instituição.

Mais declaro que todas as referências a outros autores respeitam escrupulosamente as regras da atribuição, encontrando-se devidamente citadas no corpo do texto e identificadas na secção de referências bibliográficas. Não são divulgados na presente dissertação quaisquer conteúdos cuja reprodução esteja vedada por direitos de autor.

Tenho consciência de que a prática de plágio e auto-plágio constitui um ilícito académico.

Magda Isabel Cerdeira Franco

Lisboa, 30 de setembro de 2022

Agradecimentos

Em primeiro lugar quero agradecer à minha mãe e ao meu pai por estarem sempre presentes na minha vida, por me motivarem e apoiarem em todo o meu percurso académico. Por me terem dado sempre tudo e por me ajudarem sempre no que toca a seguir os meus sonhos. Nada disto teria sido possível sem esse apoio e, por isso, estou extremamente grata. À minha irmã, que é a pessoa que mais me compreende neste mundo, obrigada por estares sempre presente, mesmo estando longe. Obrigada por aturares todos os meus momentos de stress, por me dizeres que vai ficar tudo bem mesmo quando tudo parece que vai colapsar. Obrigada por me motivares a não desistir de acabar este Mestrado, porque foram muitos os momentos em que o quis fazer.

Quero agradecer à minha orientadora, Professora Doutora Laura Cainé, por ter aceite orientar-me na realização desta dissertação. Por ser tão querida e estar sempre disposta a ajudar-me a ultrapassar os problemas que foram surgindo ao longo de todo o processo. Obrigada por me ter proporcionado esta oportunidade incrível de estagiar no Instituto de Medicina Legal. Algo que eu nunca pensei que fosse possível num futuro próximo, e por isso, estou-lhe eternamente grata.

Ao meu coorientador, Professor Doutor António Amorim, por ter disponibilizado o seu tempo para me orientar e guiar na realização desta dissertação. Obrigada por ter aceite a proposta da Professora Doutora Laura Cainé e me ter acolhido na Delegação do Sul do INMLCF. Sem si, a realização deste sonho também não teria sido possível.

À Joana e à Vânia por me terem ensinado todo o processo laboratorial e por me terem acompanhado no INMLCF. Obrigada por todos os momentos que passei com vocês.

A todos os meus colegas do LGM, que se foram tornando mais como uma família, um obrigada enorme por todos os momentos de riso, de descontração, de apoio emocional. Obrigada por tornarem os momentos de stress mais fáceis de lidar. Quero agradecer em especial à Dani e ao João, por serem os melhores amigos de sempre e se terem tornado pessoas tão próximas e que vou levar sempre no coração, independentemente de onde a vida nos levar. Por todas as conversas que certamente me ajudaram a ver as coisas de uma perspetiva diferente. Um obrigada especial também à Olena e ao Rui por serem as pessoas mais compreensivas que conheço. Sinto que não tenho palavras suficientes para vos agradecer tudo o que fizeram por mim.

Por último, quero agradecer ao Ivan, que seguramente é a pessoa que mais me aturou, que mais me ouviu a reclamar da vida, que apesar de todos os dramas, esteve sempre lá para me acalmar. Estou para sempre grata por te ter conhecido.

Estes últimos dois anos foram anos de mudanças, de novas etapas, nem tudo foi fácil e por isso, quero agradecer do fundo do coração a todas as pessoas que passaram na minha vida. Sei que todos, de uma maneira ou de outra, me ajudaram a evoluir e a tornar-me na pessoa que sou hoje.

Abstract

SARS-CoV-2 is a recent coronavirus that appeared in the end of 2019. The World Health Organization named the infection, caused by this new coronavirus, Coronavirus Disease 2019 (COVID-19). This disease was considered a pandemic on March 11th 2020 by the same organization.

Coronaviruses rapidly acquire new mutations and, consequently, new variants keep emerging. There are some variants with an associated risk in the increase of the transmissibility and that cause more severe disease and reduction in the neutralization by antibodies and, for this reason, are called variants of concern.

The aim of this project was to identify the main VOCs and VOIs circulating in the region of Lisbon, applying the methodology of real time RT-PCR in cadavers that tested positive for SARS-CoV-2. To meet this goal, we used three assays: Allplex™ SARS-CoV-2 Variants I, Allplex™ SARS-CoV-2 Variants II and Allplex™ SARS-CoV-2 Variants V. The first one detects defining mutations of the Alpha, Beta and Gamma variants (N501Y, E454K and HV69/70del), the second assay detects mutations present in the Delta, Beta, Gamma and California variants (L452R, K417T, K417N and W152C) and the third one detects defining mutations of the Delta and Lambda variants (L452R, P681R, L452Q and F690S).

In addition, we also wanted to understand if these RT-PCR assays were efficient to correctly identify these variants, comparing the results obtained to the reference methods for the determination of variants – Next Genration Sequencing (NGS). Two different NGS methodologies were used to compare the results – NGS-ONT (Oxford Nanopore Technologies) and NGS-Sanger based.

We concluded that 30% of the samples belonged to the Alpha variant and 70% belonged to the Delta variant and that, in general, the three RT-PCR assays applied in this study were efficient in correctly identifying these two variants, based on the comparison to the NGS results.

Keywords: COVID-19, Cadavers, Allplex™ SARS-CoV-2 Variants I, Allplex™ SARS-CoV-2 Variants II, Allplex™ SARS-CoV-2 Variants V, Alpha varitant, Delta variant, Real Time PCR

Resumo

SARS-CoV-2 é um coronavírus novo que apareceu no final de 2019. A Organização Mundial de Saúde nomeou a infeção causada por este vírus como COVID-19, tendo sido considerada uma pandemia no dia 11 de março de 2020.

Os coronavírus adquirem novas mutações muito rapidamente e, consequentemente, novas variantes vão emergindo. Existem algumas variantes com um risco associado ao aumento da transmissão e que causam doença mais severa e redução na neutralização por anticorpos, sendo, por esta razão, chamadas de variantes de preocupação.

O objetivo deste estudo foi identificar as principais VOCs e VOIs na região de Lisboa, aplicando a metodologia de RT-PCR em amostras de cadáveres que testaram positivo para SARS-CoV-2. Para tal, foram utilizados três kits de RT-PCR: Allplex™ SARS-CoV-2 Variants I, Allplex™ SARS-CoV-2 Variants II e Allplex™ SARS-CoV-2 Variants V. O primeiro deteta mutações presentes nas variantes Alpha, Beta e Gamma (N501Y, E454K e HV69/70del), o segundo deteta mutações presentes nas variantes Delta, Beta, Gamma e California (L452R, K417T, K417N e W152C) e o terceiro deteta mutações presentes nas variantes Delta e Lambda (L452R, P681R, L452Q e F690S).

Posteriormente, também queríamos perceber se estes kits eram eficientes na correta identificação destas variantes. Para isso, comparamos os resultados obtidos por RT-PCR com dois métodos diferentes de sequenciação: NGS-ONT (Oxford Nanopore Technologies) e NGS-Sanger based.

Concluímos, portanto, que 30% das amostras pertenciam à variante Alpha e 70% à variante Delta. Em geral, os três kits de RT-PCR aplicados neste estudo foram eficientes na correta definição destas duas variantes, baseando esta conclusão na comparação com os resultados obtidos por sequenciação.

Palavras-chave: COVID-19, Cadáveres, Allplex™ SARS-CoV-2 Variants I, Allplex™ SARS-CoV-2 Variants II, Allplex™ SARS-CoV-2 Variants V, Variante Alpha, Variante Delta, Real Time PCR

Table of contents

List of tables

List of figures

List of abbreviations

ACDP	Advisory Committee on Dangerous Pathogens
ACE2	Angiotensin-converting enzyme 2
CatB/L	Cathepsin B and L
CDC	Center for Disease Control and Prevention
cDNA	Complementary DNA
COVID-19	Coronavirus Disease 2019
CTD	C-terminal domain
Ct	Cycle threshold
del	Deletion
DNA	Deoxyribonucleic acid
DAD	Diffuse alveolar damage
DPP4	Dipeptidyl peptidase-4
ER	Endoplasmic reticulum
E	Envelope
ELISA	Enzyme-linked immunosorbent assay
et al.	et alli/et aliae
HG	Hazard groups
IgG	Immunoglobulin G
IgM	Immunoglobulin M
INMLCF	Instituto Nacional de Medicina Legal e Ciências Forenses
INSA	Instituto Nacional de Saúde Doutor Ricardo Jorge
IC	Internal control
LFIA	Lateral flow immunoassay
M	Matrix
mRNA	Messenger RNA
MERS-CoV	Middle East respiratory syndrome coronavirus
nm	Nanometer
NTD	N-terminal domain
N	Nucleocapsid
ORF	Open reading frame
PPE	Personal protective equipment
PCR	Polymerase chain reaction
RAT	Rapid antigen test

RBD	Receptor binding domain
Rpm	Revolutions per minute
RTC	Replication-transcription complex
RT-PCR	Reverse transcription polymerase chain reaction
RNA	Ribonucleic acid
RdRp	RNA-dependent RNA polymerase
SARS-CoV-2	Severe acute respiratory syndrome coronavirus
SARS-CoV	Severe acute respiratory syndrome-associated virus
S	Spike
S1	Spike protein subunit 1
S2	Spike protein subunit 2
TMPRSS2	Transmembrane serine protease
UK	United Kingdom
USA	United States of America
UTR	Untranslated region
VOC	Variant of concern
VOI	Variant of interest
n	Number of samples
WHO	World Health Organization
n	Number of samples
NGS	Next Generation Sequencing
NT	Nucleotide
SB	Sanger Based
ONT	Oxford Nanopore Technologies
AA	Aminoacid
WGS	Whole Genome Sequencing

1. Introduction

1.1. Epidemiology

A novel coronavirus appeared in 2019 and it was termed severe acute respiratory syndrome coronavirus (SARS-CoV-2). The World Health Organization (WHO) named the infection, caused by this new coronavirus, Coronavirus Disease 2019 (COVID-19). On March 11th, this disease was considered a pandemic by the WHO [1].

This outbreak started in December of 2019, in a province in China, more specifically, in Wuhan. SARS-CoV-2 was identified and isolated from patients with pneumonia, by the Chinese Center for Disease Control and Prevention (China CDC) and the genomic sequence was released on January 11th [1]. Initially there were 3 main variants of the virus, depending on the part of the globe. Variant A – the original variant with 96.2% sequence similarity to the bat coronavirus – was predominantly found in USA and Australia and in Americans who used to live in Wuhan. Variant B differs from variant A by two mutations and was the predominant variant in Wuhan and East Asia. Variant C was mostly found in Europe [2,3]. However, new variants keep emerging, and with them an associated risk in the increase of the transmissibility [4,5].

Different coronaviruses have been identified years before. Severe acute respiratory syndrome-associated virus (SARS-CoV) was responsible for the first outbreak of an infectious disease in the XXI century, more concretely from 2002 to 2003 [6]. Another different strain was isolated from a patient with severe respiratory infection, in Saudi Arabia, in 2012. It was later labelled Middle East respiratory syndrome coronavirus (MERS-CoV) [7].

RNA viruses evolve rapidly, which allows them to accumulate mutations that might improve their transmissibility capabilities. A specific mutation on the spike protein of SARS-CoV-2 – D614G substitution (from aspartic acid to glycine on the position 614) – has allowed an increase of the virus infectivity [8–10]. Interestingly this mutation is not located in the receptor binding domain (RBD) region – the site where the virus binds to the human angiotensin-converting enzyme 2 (ACE2) receptor [2,3,11].

SARS-CoV and MERS-CoV were previously known to infect only animals. However, with the increasing contact between humans and wildlife, and the growth of agriculture and urban ecosystems, they acquired new mutations that made it possible to enter the human population [3,12,13]. The natural reservoir of coronaviruses (including SARS-CoV-2) are some bat species. It seems that intermediate hosts have an enormous importance in the genetic modification and adaptation of coronaviruses to cross species

barriers. In the case of MERS-CoV and SARS-CoV, those might have been dromedary camels and civet cats (*Paguma larvata*), respectively [1,3]. It was thought that pangolins (*Manis javanica*) might have played that role in SARS-CoV-2 [3,14] but recent findings have excluded this hypothesis, suggesting that the transmission might occur directly from bats to humans [2,3]. Mutations in the RBD of SARS-CoV have contributed for the ability of infecting human cells [11]. The spike protein of SARS-CoV has a few amino acid differences from the SARS-CoV-2 spike, which might explain the higher transmissibility rate of the latter [15].

SARS-CoV-2 is transmitted via infected droplets, direct or indirect contacts, as well as SARS-CoV and MERS-CoV [1,15,16]. The virus particles can remain in the air for great periods of time (up to three hours as described by De Ungria, 2020) and can be transferred to other individuals at close range [18,19]. The virus may persist on surfaces on indoor environments for days so the droplets can be transferred from an individual's hands to their own mucosa of mouth, nose or eyes [20–22]. Airborne transmission has only been described when aerosol-generating procedures are conducted (mainly during autopsies) [17].

The incubation period of SARS-CoV-2 is between 2–14 days with a median of 5.1 days [1,23]. COVID-19 commences with symptoms on the upper respiratory tract and later might turn to a lower respiratory tract infection [24]. Symptoms may include: fever, loss of taste and smell, headache, cough, sore throat, fatigue, dyspnea, among others [20,25–27]. Asymptomatic individuals are able to transmit the virus but it seems to happen at a lower degree compared to individuals who have symptoms throughout the disease [20,28]. The estimated percentage of asymptomatic cases is around 20% [25,28,29]. The period before symptoms onset (1-5 days) seems to be critical in spreading the infection, being that 25-40% of transmissions are fruit of the pre-symptomatic period [25,30]. Therefore, it can be inferred that pre-symptomatic transmissions play a greater role in the disease spreading than asymptomatic transmissions [20,28,29,31].

In some patients who have been infected, it was described the presence of SARS-CoV-2 in fecal samples that persisted for about 5 weeks after the upper respiratory samples have turned negative [32,33]. An autopsy study has described the presence of mild to high viral RNA loads in the small and large intestine [34].

Studies on viral isolation of SARS-CoV-2 from RT-PCR positive nasopharyngeal swabs have demonstrated that 8-9 days after symptoms onset the virus isolates are not obtained anymore, even though the viral load might remain high [20,35,36]. It was also observed that the culture of the virus is more successful for Ct values of 25 or less, and Ct values higher than 35 show a low probability of obtaining a virus isolate [34]. Contact tracing investigations corroborate the viral culture data, demonstrating that no infections

occurred after 5 days of symptoms onset [28]. Regarding stool samples, isolates were never obtained despite the high viral load [35].

In the analysis of 9 patients, Wölfel et al., 2020 showed that seroconversion (the period when a certain antibody is produced and detected in the blood) of immunoglobulin M (IgM) occurred 7 days after the symptoms onset in half of the individuals and until day 14 in every patient and the viral load remained about the same. Iyer et al., 2020 observed seroconversion of IgM and IgG against the RBD of the spike protein within approximately 12 days since the emergence of the symptoms in 343 patients [35,38]. The median for the decay of IgM was 49 days. However, IgG levels were barely reduced in 90 days and seroreversion (the period when a specific antibody cannot be detected in the blood anymore) has only occurred in a 3 of the 343 patients. The study of Li et al., 2020 is also consistent with this information, being that IgG was detected 35 days after the symptoms onset in almost every patient. It is interesting to note that IgM levels decrease earlier than IgG which might help in the distinction between older and more recent SARS-CoV-2 infections. The detection of antibodies after a long period of time might be more useful in severe cases as it seems that in mild or asymptomatic individuals seroreversion occurs faster and the IgG levels remain higher in the severe cases after several weeks [37,39]. This possibly indicates that there is a correlation between the severity of symptoms and antibody production [37,40]. In addition, Li et al., 2020 showed that in the first 3 weeks the antibody response was bigger in the mild cases than in the severe ones, suggesting that lower antibody production during the first weeks might be correlated with the progression of the disease. Epidemiological studies have shown that herd immunity cannot be achieved without a vaccine, since some patients who had the disease were able to get infected again some months later, suggesting that immunological memory is not effective by natural infection [41]. In spite of this evidence, there is still a limitation on the knowledge concerning the duration of antibody responses, the protection against possible re-infections and cross-reactivity with other respiratory viruses [37].

The main concern related to this pandemic is to decrease the transmission of COVID-19. Social distancing of 2 meters and a temporary lockdown to prevent the saturation of hospitals were the measures taken to reduce the dissemination of the coronavirus [42].

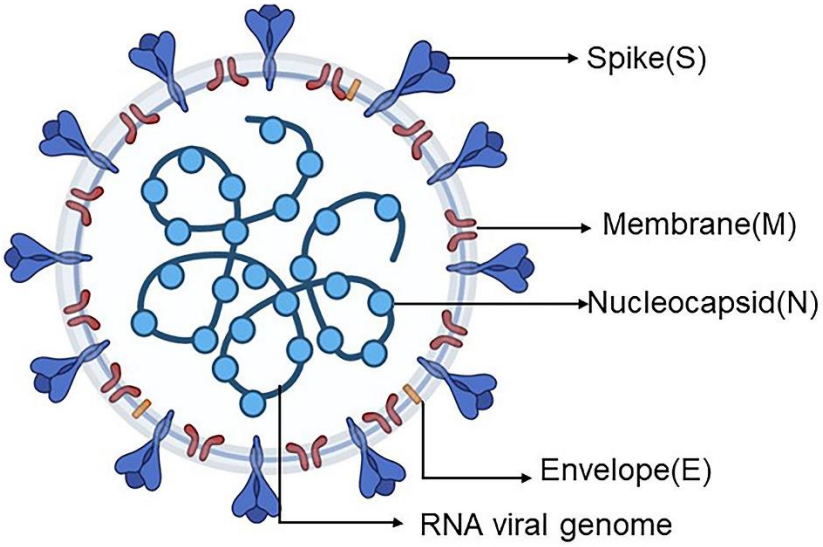
According to the Advisory Committee on Dangerous Pathogens (ACDP), pathogens are classified in hazard groups (HG1-4) regarding their risk to infect humans, to spread and whether there is a treatment available or not. SARS-CoV-2 belongs to the HG3, meaning that this microorganism should be manipulated in a containment level 3 [43–45].

1.2. Zoonotic coronaviruses

Zoonotic coronaviruses were detected in the 1960s. They are viruses that usually exist in vertebrate animals but can also jump to humans and cause disease in both animals and humans [6,46,47]. They are single and positive stranded RNA viruses with a size between 80-200 nm in diameter with spikes of 12-24 nm that bring to mind a crown shape, hence the name “corona” [12,48–50]. Interestingly, they have the greatest genome within the known RNA viruses, being that SARS-CoV-2 is approximately 30,000 nucleotides long with a 5’ – methylated cap and a 3’ – poly A tail and encodes about 9860 amino acids [2,48,51–53]. They belong to the order Nidovirales, family *Coronaviridae* and subfamily *Coronavirinae* [1]. There are four coronavirus genera – alpha, beta, gama and deltacoronavirus – being that SARS-CoV, MERS-CoV and SARS-CoV-2 belong to the betacoronavirus genera [46,54].

The RNA of these viruses is wrapped by a nucleocapsid protein (N), which forms a coiled tubular structure. This structure is surrounded by the viral envelope (E) which has the matrix protein (M) and the spike structural protein (S) linked to it [1] (Figure 1).

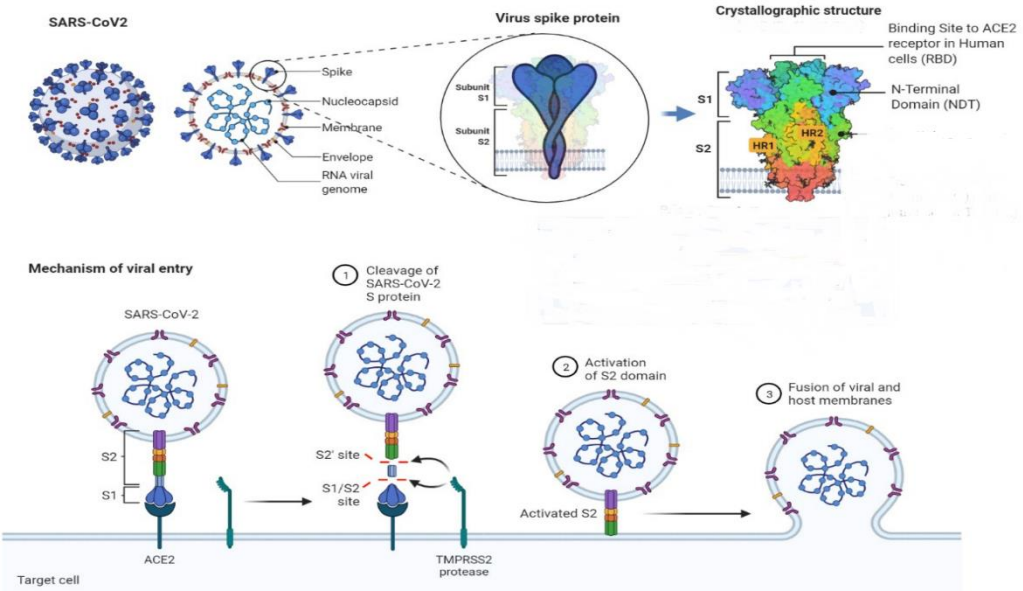
Figure 1. Coronavirus structure (Adapted from Petrovszki et al., 2021)[55].



Spike protein mediates the binding to the host cell receptor – ACE2 – through the RBD [3,10,39,56–58]. The host cell proteases, namely type II transmembrane serine protease (TMPRSS2) and cathepsin B and L (CatB/L), cleave the S protein into S1 and S2 subunits which allows the virus fusion with the host cell membrane (through the exposal of a peptide in the C-terminal S2 domain), enabling the entrance in the host [3,9,14,15,46,48,59–62] (Figure 2).

After cleavage of the S protein, the viral RNA genome is released into the cytoplasm and the cell recognizes the viral RNA as its own, which allows the virus to control the cellular machinery. The positive RNA strand of the virus is translated by the host ribosome into two replicase polyproteins (pp1a and pp1ab), which will be involved in the replication and transcription. During this step of translation, a ribosome frameshift occurs and originates pp1ab, which is a hybrid of pp1a and pp1b. The two replicase polyproteins are proteolyzed into smaller proteins, which form a replication-transcription complex (RTC). The RTC binds to the positive RNA strand of the virus and is replicated into a negative RNA strand. This negative RNA strand might originate more positive stranded RNA or might be transcribed by a process called discontinuous transcription. During this process, the RdRp can bind to different sites of the negative RNA and, therefore, initiate transcription at different points. This will generate several mRNAs (subgenomic RNAs) of different lengths, which encode for different structural proteins. These structural proteins are translated in the rough endoplasmic reticulum (ER) and then sent to the Golgi apparatus, together with the positive RNA strand, to be packaged into vesicles, assembled into virions and exit the cell to infect and spread to other cells [15,50,52,54,59,63,64].

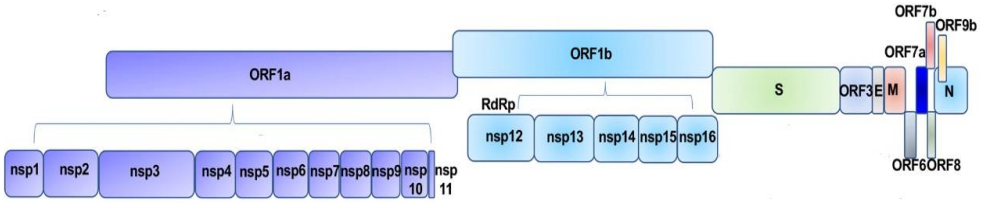
Figure 2. Structure of SARS-CoV-2 spike protein and mechanism of entrance in the host cell (Adapted from Petrovszki et al., 2021).



The M protein determines the shape of the viral envelope. The E and N proteins are associated with the viral assembly and budding, being that the latter is the only one that binds to the RNA genome [46,65].

Their genome has 14 open reading frames (ORFs) which encode for 4 structural proteins (S, E, M and N), 16 nonstructural proteins (Nsp 1-16), including RdRp, and 9 accessory proteins (Orf3a, Orf3b, Orf6, Orf7a, Orf7b, Orf8, Orf9b, Orf9c, and Orf10) [52,59,66] (Figure 3). They also contain a 5' – untranslated region (UTR) and a 3' – UTR [2].

Figure 3. Representation of the viral genome of SARS-CoV-2 (Adapted from Zhang, Q., Xiang, R., Huo, S. et al., 2021)[67].



SARS-CoV-2 shares a genome sequence similarity of 51.8% and 79% with MERS-CoV and SARS-CoV, respectively [1]. N, E and M proteins are more conserved than the S protein when comparing different coronavirus [48].The host cell receptor of SARS-CoV and SARS-CoV-2 is ACE2 while the receptor of MERS-CoV is the dipeptidyl peptidase-4 (DPP4) [68].

Seven human coronavirus are identified to this date and can be divided in two groups according to the level of pathogenicity: the ones that cause mild diseases (HKU1, OC43, NL63 and 229E) and the ones that cause more severe symptoms (MERS-CoV, SARS-CoV and SARS-CoV-2) [46,69].

Coronaviruses infect the organism by entering through the respiratory tract. They replicate in the upper respiratory tract (in the epithelial cells) and then disseminate to the lower respiratory tract [12,58].

1.3. SARS-CoV-2 variants and important mutations

Lineage A is considered the root of the pandemic. It appeared in China and spread to Asia, Australia, the USA and Europe [70].

Lineage B.1, which was originated from lineage B, is an European lineage that was responsible for the Italian outbreak in the beginning of 2020 [71].

The spike protein of SARS-CoV-2 is composed by two main parts: S1 – which contains the N-terminal domain (NTD) and the receptor binding domain (RBD) – mediates the binding to the human ACE2 receptor; and S2 – which contains the C-terminal domain (CTD) – allows the membrane fusion and entry in the host cell. This protein is approximately 1200 amino acids long and is under a huge selective pressure

to increase its binding to the human ACE2 receptor and to escape from the immune system, so that the virus can continue to spread. The RBD is the site that mediates the virus interaction with the host cell, thus mutations occurring in those amino acids are concerning to the human population as it might affect antibody neutralization and vaccines efficacy [62,72].

A new strain with the D614G substitution in the spike protein started circulating in Europe in January 2020 and rapidly spread to other parts of the world. Compared with the Wuhan seafood market pneumonia virus reference sequence NC_045512.3, amino acid changes D614G in the spike protein and P323L in the RdRp might be associated with disease severity and it is known that the mutation that originates the amino acid change D614G is responsible for an increased rate of transmissibility. It was verified that a higher mutation rate of the virus is associated with mutations in the *RdRp* gene. P323L amino acid change comes from a mutation in *ORF1ab* gene that originates P4715L on the ORF1ab polyprotein [64,73,74].

An amino acid deletion at positions 69 and 70 in the N-terminal domain of the S1 subunit of the spike protein (HV69/70del) was observed in the cluster-5 variant which was found in humans and minks in Denmark. This deletion has been detected together with the RBD mutation Y453F (in the cluster-5 variant) and with RBD mutations N439K (in Europe) and N501Y (in the UK) [75–77]. N439K and Y453F mutations have been associated with increased affinity to ACE2 receptor and higher viral loads when compared to the wild-type but don't seem to decrease plasma antibody neutralization [78,79].

E484K mutation has emerged independently in different places in the world [80] but has only been defined as a lineage-defining mutation in B.1.351 (Beta variant) and P.1 (Gamma variant). It has also been observed in some sequences from isolates of the Alpha variant. It was demonstrated that this amino acid change is associated with resistance to neutralizing antibodies, both from human sera or plasma and monoclonal and polyclonal antibodies but doesn't seem to affect the affinity of the virus to the ACE2 [79,81–84].

N501Y appeared independently in the three variants of concern (VOCs) – Alpha, Beta and Gamma variants – and is associated with higher affinity between the spike of the virus and the human ACE2 receptor, which might contribute to higher transmissibility and virulence but does not seem to affect neutralization by most human plasma [54,79,84–87].

1.3.1. Variants of Concern (VOC)

A Variant of Concern is defined by the Centers for Disease Control and Prevention (CDC) as *a variant for which there is evidence of an increase in transmissibility, more severe disease (e.g., increased hospitalizations or deaths), significant reduction in neutralization by antibodies generated during previous infection or vaccination, reduced effectiveness of treatments or vaccines, or diagnostic detection failures.*

1.3.1.1. Alpha variant (B.1.1.7)

501N lineage was circulating in the UK before two new lineages appeared with the amino acid substitution N501Y in the RBD of the spike protein. The N501Y lineage, that also contained the deletion HV69/70 in the N-terminal domain of the spike protein, was found to be 75% more transmissible than the 501N lineage. The N501Y lineage (also called B.1.1.7 or VOC-202012/01 or Alpha) quickly became the dominant variant circulating in the UK in December 2020 and became a VOC characterized by many different mutations (Table 3) [88,89]. Volz et al., 2021 concluded that this strain has a higher reproduction number (R) compared to preceding lineages. The N501Y amino acid change and the Y145del are in host receptor and antibody binding sites which makes them relevant mutations. *ORF8* deletions have been detected in some sequences of SARS-CoV-2 and are associated with the Q27 early stop codon which is also observed in the Alpha variant [90]. P681H is also relevant as it is associated with the furin cleavage site in the spike protein, which is a crucial site responsible for infection [91,92].

It is important to know if people who were infected by one strain acquire immunity to different strains [88]. According to Public Health England, 2021, individuals who were infected with non-B.1.1.7 strains show neutralizing responses against the B.1.1.7 lineage and vice-versa, thus it seems the neutralizing activity remains unchanged when compared to the D614G strain [81]. This evidence is in concordance with the study performed by C. Liu et al., 2021 which shows that individuals infected by B.1.1.7 variant seem to develop protection against all VOCs (B.1.351, P.1 and B.1.617.2).

This variant is mostly found in the UK (25%), USA (19%), Germany (10%), Sweden (6%) and Denmark (6%) [94].

1.3.1.2. Beta variant (B.1.351)

Another different variant emerged in October, the South African variant (501Y.V2 or B.1.351 or Beta), which also has the N501Y amino acid change but not the deletion HV69/70 [88,95].

K417N is an important substitution present in this variant, as it affects neutralization by some monoclonal antibodies. However, this mutation alone doesn't seem to affect the binding of RBD of the spike protein to the human ACE2 receptor [79].

The combination of the two mutations mentioned above together with E484K has been shown to reduce neutralization, obtained by previous infection or vaccination, more than those mutations separated [79,81]. In addition, the combination of N501Y, E484K and K417N has been associated with higher affinity to ACE2 and increased transmissibility [6296].

This variant also contains a mutation close to the furin cleavage site – A701V [92].

The list of all mutations present in this lineage is shown in Table 3.

This lineage is predominantly found in South Africa (18%), Sweden (9%), USA (8%), Germany (8%) and France (7%) [97].

1.3.1.3. Gamma variant (P.1)

Brazil/Japan variant (also called P.1 or 501Y.V3 or Gamma) is characterized as a VOC due to a set of defining mutations in RBD of the spike protein – E484K, N501Y and K417T. It also contains mutations in the N-terminal domain (L18F, T20N, P26S, D138Y and R190S) and in the furin cleavage site (H655Y) (Table 3) [98]. N501Y and E484K have been shown to be associated with a higher transmissibility rate and escape from neutralizing antibodies, respectively [99]. Wang, Casner, et al., 2021 showed that neutralization from convalescent plasma and vaccines is lower compared to the D614G strain but not as low as in the B.1.351 strain.

P.1 was first identified in Brazil in individuals who arrived from Japan [99]. It arose from B.1.1.28 lineage, which was first detected in Brazil in February 2020, together with P.2, which is considered a Variant of Interest (VOI) and also contains the E484K mutation [83,100]. It is more frequently found in Brazil (30%), USA (29%), Canada (17%), Chile (5%) and Mexico (4%) [101].

1.3.1.4. Delta variant (B.1.617.2)

B.1.617.2, also known as the Indian variant or Delta, is defined by L452R, T478K and P681R mutations. It also contains D614G, D950N; and T19R, 156/157del, R158G, G142D and A222V in the N-terminal domain of the spike protein [62,102–104]. L452R and P681R occurring in the RBD are associated with an increased transmissibility and are also present in the B.1.617.1 lineage, which is a VOI [105]. P681R occurs specifically at the S1/S2 cleavage site of the spike protein and has been shown to affect replication, originating higher viral loads, hence the increased transmission [104].

The mutations, present in this lineage, occurring in the NTD have been shown to escape completely some monoclonal antibodies. Studies show that L452R and E484Q block the activity of some neutralizing RBD antibodies but T478K doesn't seem to affect any neutralization [62].

Convalescent sera from individuals previously infected by B.1.351 and P.1 doesn't seem to provide protection against B.1.617.2 strain [62]. Studies on vaccine induced neutralization showed that its effectiveness doesn't seem to be reduced significantly compared to the B.1.1.7 variant [72].

Recently, this variant has become the predominant variant in various countries, including Portugal (July 2021), replacing the Alpha variant [62]. It is more frequently found in USA (18%), India (18%), Turkey (16%), UK (12%) and Germany (5%)[106].

1.3.1.5. B.1.617.2 AY.1

This is a sub-lineage of B.1.617.2 that appeared in March 2021 and it is more prevalent in the USA (64%), Japan (6%), Portugal (6%), UK (5%) and Switzerland (4%) [107]. It is characterized by the following mutations: A1306S, P2046L, P2287S, V2930L, T3255I and T3646A in ORF1a, P314L, G662S, P1000L and A1918V in ORF1b, T19R, E156G, del157/158,W258L, K417N, L452R, T478K, D614G, P681R and D950N in spike, S26L in ORF3a, I82T in the matrix, V82A and T120I in ORF7a, T40I in ORF7b, D119I and del120/121 in ORF8, and D63G, R203M, G215C and D377Y in the nucleocapsid [108].

1.3.1.6. Omicron variant (B.1.1.529)

The lineage B.1.1.529 was named by the WHO as Omicron and is considered a VOC. It is characterized by a high number of spike mutations compared to the other

VOCs [109] (Table 1), which suggests that it evolved in a different environment that might be an animal host or an immunocompromised patient [110]. This variant was first identified in South Africa and Botswana in November 2021. It is more frequently found in the USA (58%), Turkey (18%), India (4%), Germany (3%) and Italy (3%) [111].

Table 1. List of defining amino acid changes in the Omicron variant [112].

Gene	Amino acid change
ORF1a	A2710T
ORF1a	T3255I
ORF1a	P3395H
ORF1a	I3758V
ORF1b	P314L
ORF1b	I1566V
S	T95I
S	G142D
S	D614G
S	H655Y
S	N679K
S	P681H
S	Q954H
S	N969K
E	T9I
M	Q19E
M	A63T
ORF8	S84L
N	P13L
N	31/33del
N	R203K
N	G204R

1.3.2. Variants of Interest (VOI)

A Variant of Interest is defined by the Centers for Disease Control and Prevention (CDC) as *a variant with specific genetic markers that have been associated with changes to receptor binding, reduced neutralization by antibodies generated against previous*

infection or vaccination, reduced efficacy of treatments, potential diagnostic impact, or predicted increase in transmissibility or disease severity.

1.3.2.1. Lineage A.23.1

This is an international lineage, detected in Uganda for the first time in October 2020 and is mostly found in Canada (37%), UK (16%), Uganda (14%), Rwanda (11%) and USA (7%) [113]. It is characterized by the following mutations: M3655I, L3667F and M3752I in ORF1a, F157L, V367F, Q613H and P681R in spike, L84S and E92K in ORF8 and S202N in the nucleocapsid [114].

1.3.2.2. Lineage B.1.429

This lineage was first identified in the USA in September 2020 [115] and is mostly found in the USA (97%), Canada (2%) and Mexico (1%) [116]. It contains the following concerning mutations: L452R in the RBD and W152C in the N-terminal domain of spike protein. These two sites are targets of neutralizing antibodies but studies have shown that vaccines are still effective against this variant [117]. It also contains the following mutations: T265I and I4205V in ORF1a, P314L and D1183Y in ORF1b, S13I, W152C, L452R and D614G in the spike, Q57H in ORF3a and T205I in the nucleocapsid [118].

1.3.2.3. Lineage B.1.525

This is an international lineage and is found mostly in Canada (21%), USA (16%), Germany (10%), France (9%) and Denmark (8%) [119]. It was first identified in the UK and Nigeria in December 2020 [102,103,115,120,121]. It is characterized by the following mutations: T2007I, S3675K and del3676/3678 in ORF1a, P314F in ORF1b, Q52R, A67V, del69/70, del144/145, E484K, D614G, Q677H and F888L in spike, L21F in the envelope protein, I82T in the matrix protein, F2H and del3/3 in ORF6, and S2Y, del3/3, A12G and T205I in the nucleocapsid [122].

1.3.2.4. Lineage B.1.526

B.1.526 appeared in New York in November 2020 and is predominantly found in USA (97%) and Ecuador (1%) [84]. It harbors the following mutations: T265I, L3201P, S3675K and del3676/3678 in ORF1a, P314L and Q1011H in ORF1b, T95I, D253G and

D614G in spike, P42L and Q57H in ORF3a and T11I in ORF8 [123]. In some sequences but not all, it was detected E484K, L5F, S477N and A701V [103].

1.3.2.5. Lineage B.1.526.1

This lineage has the following mutations in the spike protein: D80G, F157S, L452R, D614G and D950H. T961I and T859N have been detected in some sequences and has been associated with resistance to neutralizing antibodies[124]. It was identified in New York in October 2020 [103].

1.3.2.6. Lineage B.1.616

This lineage only exists in France and was first detected in February 2021 [125]. It is characterized by the following mutations: T265I, N1324S, T1638I, S2261Y, Y316OH and L3606F in ORF1a, P314L, Q813R, L1681F, T2537I and K2674R in ORF1b, H66D, G142V, del144/145, D215G, V483A, D614G, H655Y, G669S, Q949R and N1187D in spike, Q57H in ORF3a, F20L and T30I in the envelope, H125Y in the matrix, del23/32 in ORF6, T325I in the nucleocapsid and A8V in ORF10 [126].

1.3.2.7. Lineage B.1.617

This lineage only exists in India and it was first detected in February 2021 [103]. It is characterized by the following mutations: del17/17, del1698/1698 and A1812D in ORF1a, P314L, del1130/1129, del1353/1353, del1546/1545, A2306V and H2583R in ORF1b, L452R, E484Q, D614G, del682/681 and del1072/1071 in spike, S26L and del96/96 in ORF3a, I82S in the matrix, P93S, E106Q and del122/121 in ORF8, and del204/204 and del215/215 in the nucleocapsid [127].

1.3.2.8. Lineage B.1.617.1

This lineage harbors the following substitutions: T1567I and T3646A in ORF1a, P314L, M1352I and K2310R in ORF1b, L452R, E484Q, D614G, P681R and Q1071H in spike, S26L in ORF3a, V82A in ORF7a, and R203M and D377Y in the nucleocapsid [128]. It was first identified in India in December 2020 and now it is mostly found in India (70%), UK (7%), Canada (5%), USA (5%) and Ireland (3%) [103].

1.3.2.9. Lineage B.1.617.3

This lineage was first detected in India in October 2020 and is predominantly found in India (90%), UK (4%), Malawi (2%), USA (2%) and Russia (1%) [129]. It contains the following mutations: A2344V, T2648I, A3457S and A3686V in ORF1a, P314L in ORF1b, T19R, L452R, E484Q, D614G and P681R in spike, V82A in ORF7a, T26I in ORF8, and P67S, R203M and D377Y in the nucleocapsid [130].

1.3.2.10. Lineage B.1.620

It appeared in February 2021 and was first detected in Cameroon. It is found mostly in South Korea (39%), Lithuania (18%), France (12%), Germany (5%) and Switzerland (4%) [131]. It contains the following mutations: T403I, V1991I, S3675K and del3676/3678 in ORF1a, P314L and A1215S in ORF1b, P26S, del69/70, V126A, del144/145, del241/243, H245Y, S477N, E484K, D614G, P681H, T1027I and D1118H in spike, and L14A and del15/15 in ORF7b [132].

1.3.2.11. Lineage B.1.621

It was first detected in Colombia in January 2021 and is spread mostly in the USA (45%), Colombia (13%), Mexico (9%), Spain (9%) and Chile (4%) [133]. It is characterized by the following mutations: T1055A, T1538I, T3255I and Q3729R in ORF1a, P314L and P1342S in ORF1b, T95I, Y144S, Y145N, R346K, E484K, N501Y, D614G, P681H and D950N in spike, Q57H and del257/257 in ORF3a, T11K, P38S and S67F in ORF8, and T205I in the nucleocapsid [134].

1.3.2.12. Lineage B.1.626

This lineage was first detected in April 2021 and is mostly found in the USA (70%), Portugal (9%), Luxembourg (8%), Netherlands (4%) and France (3%) [135]. It contains the following mutations: T265I, R550H, S3384L, A3548V, V3653F and T4249I in ORF1a, P314L and P1165L in ORF1b, L18F, del66/68, G257S, E484K, F490S, N501T, D614G and T859N in the spike, S40L, Q57H and T64I in ORF3a, I82S in the matrix, T14I in ORF7a, D3Y and T205I in the nucleocapsid and R24C in ORF10 [136]. Some sequences have, additionally, the substitution K417N [102].

1.3.2.13. Lineage P.2

P.2 appeared in Brazil, in Rio de Janeiro, in April 2020 and is currently found in Brazil (55%), USA (26%), Canada (4%), Suriname (2%) and Argentina (2%) [137]. It harbors the following mutations: L3468V and L3930F in ORF1a, P314L in ORF1b, E484K, D614G and V1176F in spike, and A119S, R203K, G204R and M234I in the nucleocapsid [138]. F565L was detected in some sequences but not all [99,103].

1.3.2.14. Lineage P.3

This lineage was detected in the Philippines in January 2021 and is more frequent in Philippines (65%), USA (6%), Germany (4%), Malaysia (4%) and UK (3%) [139]. It is characterized by the following mutations: D1554G, S2625F, D2980N, L3201P, D3681E and L3930F in ORF1a, P314L, L1203F and A1291V in ORF1b, E484K, N501Y, D614G, P681H, E1092K, H1101Y and V1176F in spike, K2Q in ORF8, and R203K and G204R in the nucleocapsid [140].

1.3.2.15. Lambda variant (C.37)

This variant appeared in Peru in November 2020 and in June 2021 was classified as a VOI by the WHO and named Lambda [141142]. It contains the following mutations: T1246I, P2287S, F2387V, L3201P, T3255I, G3278S and del 3675/3677 in ORF1a, P314L in ORF1b, G75V, T76I, R246N, del 247/253, **L452Q**, **F490S**, D614G and T859N in the spike, S84L in ORF8 and P13L, R203K, G204R and G214C in the nucleocapsid [143]. It is potentially more transmissible and resistant to neutralizing antibodies due to some spike substitutions, such as, L452Q which is similar to L452R present in the Delta variant, and T859N also present in the B.1.526.1 [124,141]. It is mostly found in Peru (43%), Chile (24%), USA (14%), Argentina (6%) and Ecuador (3%) [142].

1.3.3. Other lineages

1.3.3.1. Lineage A.21

This lineage appeared in April 2020 in the USA. It is mostly found in France (31%), Burkina Faso (22%), USA (13%), Portugal (5%) and Spain (4%). It contains D614N spike mutation, V3718F in ORF1a, A99S in ORF3a, L84S in ORF8 and S202N in the nucleocapsid [144,145].

1.3.3.2. Lineage B.1

This lineage is mostly spread in the USA (43%), Turkey (10%), UK (9%), France (4%) and Canada (3%). It appeared in January 2020 during the Italian outbreak [146]. It contains the D614G spike mutation and P314L substitution in ORF1b [147].

1.3.3.3. Lineage B.1.1.216

This lineage is mostly found in India (78%), following by the UK (5%), Indonesia (3%), USA (2%) and Canada (2%) and it appeared in India in May 2020 [148]. It contains two mutations in the nucleocapsid (R203K and G204R), D614G in the spike protein, P314L in ORF1b and T223I in ORF3a [149].

1.3.3.4. Lineage B.1.1.160

This lineage is mostly found in Japan (73%) and the UK (27%). It appeared in May 2020 in Northern Ireland [150]. It contains two substitutions in ORF1a (S1361P and P3371S), two in ORF1b (P314L and A414V), one in spike (D614G) and three in the nucleocapsid (P151L, R203K and G204R) [151].

1.3.3.5. Lineage B.1.177

This lineage is from Europe and was first identified in February 2020. It is more frequently found in the UK (62%), Spain (12%), Germany (5%), Switzerland (4%) and Italy (4%) [152]. It is characterized by the following mutations: P314L in ORF1b, A222V and D614G in spike, A220V in the nucleocapsid and V30L in ORF10 [153].

1.3.3.6. Lineage B.1.177.32

This lineage was first identified in Switzerland and Portugal in September 2020. It is mostly common in Spain (47%), Portugal (32%), Switzerland (8%), Luxembourg (3%) and France (3%) [154]. It is characterized by the following mutations: A1812V in ORF1a, P314L in ORF1b, A222V and D614G in spike, A220V in the nucleocapsid and V30L in ORF10 [155].

1.3.3.7. Lineage B.1.177.52

This lineage appeared in Europe in August 2020 and it is mostly found in Portugal (36%), UK (25%), Germany (18%), Luxembourg (5%) and the Netherlands (3%) [156]. It is characterized by the following mutations: M1586I in ORF1a, P314L in ORF1b, A222V, D614G and P1162R in spike, A220V in the nucleocapsid and V30L in ORF10 [157].

1.3.3.8. Lineage B.1.177.72

This lineage appeared in Portugal in November 2020. It is mostly found in Portugal (72%), France (9%), Switzerland (6%), Luxembourg (6%) and Spain (3%) [146]. It is characterized by the following mutations: P314L and L820F in ORF1b, A222V and D614G in spike, H125Y in the matrix protein, A220V in the nucleocapsid and V30L in ORF10 [158].

1.3.3.9. Lineage B.1.177.85

This lineage was first identified in Portugal in October 2020. It is mostly frequent in Portugal (60%), Switzerland (19%), Luxembourg (14%), France (5%) and UK (1%) [159]. It is characterized by the following mutations: A1082T in ORF1a, P314L in ORF1b, A222V and D614G in spike, A220V in the nucleocapsid and V30L in ORF10 [160].

1.3.3.10. Lineage B.1.214

This lineage was identified in the Democratic Republic of the Congo (DRC) in April 2020. It is commonly found in DRC (56%), Uganda (10%), Belgium (8%), Republic of the Congo (8%) and Angola (6%) [161]. It is characterized by the following mutations: I1398V, T1881I and A4016V in ORF1a, P314L in ORF1b and D614G in spike [162].

1.3.3.11. Lineage B.1.258

This lineage appeared in the UK in March 2020 and it is spread in Germany (19%), UK (17%), Denmark (14%), Sweden (8%) and Switzerland (5%) [163]. It is

characterized by the following mutations: I2501T in ORF1a, P314L and H1213Y in ORF1b and N439K and D614G in spike [164].

1.3.3.12. Lineage B.1.258.17

This lineage appeared in Europe in August 2020 and is commonly found in Slovenia (72%), Switzerland (8%), Germany (7%), Sweden (6%) and Austria (2%) [165]. It is characterized by the following mutations: I2501T and M4241I in ORF1a, P314L, V711I, P976L, H1213Y, A1521S, P1570L and E1977D in ORF1b, del69/70, L189F, N439K, D614G and V772I in spike, and Q185H in ORF3a [166].

1.3.3.13. Lineage B.1.258.20

This lineage appeared in Luxembourg in September 2020 and is mostly found in Luxembourg (48%), Portugal (22%), Belgium (18%), India (5%) and UK (2%) [167]. It is characterized by the following mutations: I2501T and M4241I in ORF1a, P314L, V711I, H1213Y, A1521S and T2432I in ORF1b, del 69/7, N439K and D614G in spike, and K67N and G172D in ORF3a [168].

1.3.3.14. Lineage B.1.575

This lineage was identified in the USA and Aruba in April 2020 and is mostly found in the USA (89%), Spain (3%), Finland (2%) and Dominican Republic (2%) [169]. It is characterized by the following mutations: T265I in ORF1a, P314L in ORF1b, D614G in spike, Q57H in ORF3a, I82T in the matrix, and T205I in the nucleocapsid [170].

1.3.3.15. Lineage C.16

This lineage appeared in Portugal in June 2020 and is mostly found in France (17%), Angola (15%), Portugal (14%), USA (13%) and Switzerland (9%). It is characterized by the following mutations: E1013K, T1246I, A3143V, G3278S and T4175I in ORF1a, P314L in ORF1b, L452R and D614G in spike, and R203K and G204R in nucleocapsid [171].

1.3.3.16. Lineage C.36

This lineage was identified in Egypt in March 2020. It is predominantly found in Egypt (34%), Germany (11%), UK (11%), USA (8%) and Denmark (7%) [172]. It is characterized by the following mutations: T1246I, G3278S and T4090I in ORF1a, P314L in ORF1b, D614G and Q677H in spike, R203K, G204R and G212V in the nucleocapsid [173].

1.3.3.17. Lineage B.1.160

This lineage is a European lineage and was first detected in February 2020. It is predominantly found in France (20%), Denmark (13%), Switzerland (8%), Canada (8%) and UK (7%) [174]. It is characterized by the following mutations: M3087I in ORFa; A176S, P314L, V767L, K1141R and E1184D; S477N and D614G in the spike; Q57H in ORF3a; S84L in ORF8; M234I and A376T in the nucleocapsid [151].

1.3.3.18. Lineage B.1.367

This is a European lineage, and it was first identified in July 2020. It is mostly found in the UK (28%), France (23%), Norway (12%), Switzerland (10%) and Italy (7%) [175]. It is characterized by the following mutations: T265I, T945I, T1567I, A2994V, Q3346K, V3475F and M3862I in ORF1a; P255T and P314L in ORF1b; D80Y and D614G in the spike; Q57H in ORF3a; R80I in ORF7a; S84L in ORF8; S186Y and D377Y in the nucleocapsid [176].

1.3.3.19. Lineage B.1.177.58

This lineage was first identified in the UK in September 2020 and is mostly found in the UK (92%), Ireland (7%) and Switzerland (1%) [177]. It is characterized by the following mutations: P2260L in ORF1a; P314L in ORF1b; L18F, A222V and D614G in the spike; T64I in ORF3a; S84L in ORF8; A220V in the nucleocapsid and V30L in ORF10 [178].

1.3.4. Frequency of the lineages circulating in Portugal

According to the reports of Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA), in August 2020 (Table 2), the clades circulating in Portugal were the following:

19A (6,4%), 19B (2,4%), 20A (36,6%), 20B (52%) and 20C (2,6%). Being that clades 20A, 20B and 20C contain the spike mutation D614G [179]. During November 2020 [102], the most frequent lineages circulating in Portugal were B.1.777 (17,7%) together with its sub-lineages B.1.177.52 (13,2%), B.1.177.72 (12,4%), B.1.777.85 (7,7%) and lineage B.1.160 (8,2%). During January 2021, the most frequent variant was still B.1.177 (16,9%). The Alpha variant (B.1.1.7) started circulating during this month with a relative frequency of 16%, becoming the second most frequent variant during this month. B.1.177.72, B.1.160, B.1.177.32, C.16, B.1.177.52 and B.1.177.85 circulated with a frequency of 12%, 11,7%, 10,7%, 6,6%, 4,7% and 3,8%, respectively. In February 2021, the Alpha variant became the most frequent variant circulating in Portugal with a relative frequency of 58,2% while B.1.177 decreased to 6,6%. B.1.160 (5,8%), B.1.177.72 (5,1%), C.16 (5,1%), B.1.177.85 (3,7%) and B.1.177.52 (2,8%) were still part of the 8 more frequent lineages circulating in Portugal during this period. The Gamma variant also appeared during this month with a relative frequency of 0,3%. In March 2021, the relative frequency of the Alpha variant increased to 82,9%. B.1.177 decreased to 2.9%. The Beta variant (B.1.351) started circulating in Portugal, being the third more frequent variant this month with only 2,5%. B.1.160, B.1.177.72 and C.16 had a relative frequency of 1,9%, 1,3% and 0,9%, respectively. The Gamma variant (P.1) had a relative frequency of 0,4%. In April 2021, the Alpha variant was still the predominant variant with 91,2%. The Gamma variant became the second most frequent variant with 4,3% and the Beta variant the third most frequent variant with 1,3%. During May 2021, the Alpha variant was still the most frequent variant but the relative frequency decreased to 88,4%. The Delta variant started circulating in Portugal during this period and became the second most frequent variant with 3,7%. The relative frequency of the Gamma variant decreased to 2,3%, becoming the third most frequent variant this month. The Beta variant was the fourth most frequent variant with a relative frequency of 1,8%. B.1.525 started circulating in March 2021 but during May its relative frequency increased to 1,1%, being the fifth most frequent variant circulating this month. In June 2021, the relative frequency of the Alpha variant decreased to 40,2%, being that the Delta variant became the most frequent variant with 53.3%. The Gamma variant decreased to 2%. B.1.621 circulated for the first time in Portugal with a relative frequency of 1%. The Beta variant decreased to 0,1%. In the end of July 2021, the Delta variant became dominant with a relative frequency of 97,9%. The relative frequency of the Alpha variant decreased to 0,9%, together with the Gamma and Beta variants which had a frequency of 0,4% each. The sub-lineage B.1.617.2 AY.1 also had a relative frequency of 0,4%. From August to November 2021, the Delta variant maintained its relative frequency of 100%.

Table 2. Relative frequency (%) of the lineages circulating in Portugal from November 2020 to September 2021 [102].

	Relative Frequency (%)									
Lineage	Month									
	November	January	February	March	April	May	June	July	August	September
A.21	0	0	0,6	0,5	0	0	0	0	0	0
B.1	3,5	1,1	0,7	0	0,1	0	0,3	0	0	0
B.1.1.7	0	16	58,2	82,9	91,2	88,4	40,2	0,9	0	0
B.1.1.216	0	0,8	0	0	0	0	0	0	0	0
B.1.160	8,2	11,7	5,8	1,9	0,1	0,2	0	0	0	0
B.1.177	17,7	16,9	6,6	2,9	0	0	0	0	0	0
B.1.177.32	11,7	10,7	3,7	0,9	0	0	0	0	0	0
B.1.177.52	13,2	4,7	2,1	0,6	0	0	0	0	0	0
B.1.177.72	12,4	12	5.1	1.3	0.8	0	0	0	0	0
B.1.177.85	7,7	3,8	2,8	0,7	0,1	0,1	0	0	0	0
B.1.214	0	0,6	0,2	0,1	0,3	0,1	0	0	0	0
B.1.221	4,7	1,1	0,5	0	0	0	0	0	0	0
B.1.258	0	0,4	0	0	0,1	0	0	0	0	0
B.1.258.17	0	0,4	0	0	0	0	0	0	0	0
B.1.258.20	0	0,9	0,1	0,1	0	0	0	0	0	0
B.1.351	0	0	0,1	2,5	1,3	1,8	0,1	0	0	0
B.1.525	0	0	0	0,3	0,2	1,1	0,4	0,4	0	0
B.1.526	0	0	0	0,2	0	0		0	0	0
B.1.575	0	0	0,1	0	0,1	0	0	0	0	0
B.1.617.1	0	0	0	0	0,4	0,1	0	0	0	0
B.1.617.2	0	0	0	0	0	3,7	53,3	97,9	100	99,7
B.1.617.2 AY.1	0	0	0	0	0	0	0	0,4	0	0
B.1.620	0	0	0	0	0,1	0,2	0	0	0	0
B.1.621	0	0	0	0	0	0	1	0	0	0
C.16	0,7	6,6	5,1	0,9	0	0,1	0	0	0	0
C.36	0	0	0	0	0,2	0,6	0	0	0	0
P.1	0	0	0,3	0,4	4,3	2,3	2	0,4	0	0,3
P.2	0	0,6	0,5	0,1	0	0	0	0	0	0

Table 3. List of defining amino acid changes in the four main VOCs – B.1.1.7, B.1.351, P. 1 and B.1.617.2 [88,90,121].

Gene	**Amino acid change**			
	Alpha (B.1.1.7)	**Beta (B.1.351)**	**Gamma (P.1)**	**Delta (B.1.617.2)**
ORF1a	T1001I	T265I	S1188L	P2046L
	A1708D	K1655N	K1795Q	V2930L
	I2230T	K3353R	S3675K	
	S3675K	S3675K	del3676/3678	
	del3676/3678	del3676/3678		
ORF1b	P314L	P314L	P314L	P314L

			E1264D	G662S
				P1000L
				A1918V
S	del69/70	D80A	L18F	T19R
	del144/145	D215G	T20N	E156G
	N501Y	del241/243	P26S	del157/158
	A570D	K417N	D138Y	L452R
	D614G	E484K	R190S	T478K
	P681H	N501Y	K417T	D614G
	T716I	D614G	E484K	P681R
	S982A	A701V	N501Y	D950N
	D1118H		D614G	
			H655Y	
			T1027I	
			V1176F	
ORF3a		Q57H	S253P	S26L
		S171L		
M				I82T
ORF7a				V82A
				T120I
ORF8	Q27*		E92K	D119I
	R52I			del120/121
	Y73C			
E		P71L		
N	D3L	T205I	P80R	D63G
	R203K		R203K	R203M
	G204R		G204R	G215C
	S235F			D377Y

1.4. Molecular and serological methods for SARS-CoV-2 diagnostic

1.4.1. RT-PCR

Evidence shows that RT-PCR is highly sensitive and specific for COVID-19 diagnostic, mainly in the first week following the onset symptoms, which matches the time where the highest viral load is observed [27,35,37,180,181].

This technique uses an enzyme – reverse transcriptase – to transform the virus RNA into a complementary DNA (cDNA) strand, which can be used as a template for the PCR reaction. The amplification of the DNA is monitored in real time through the binding of fluorescent dyes to the target sequences [23]. The cycle threshold (Ct) value is defined

as the number of cycles required for the fluorescent signal to cross the background level of fluorescence [182] and is used to quantify the initial concentration of RNA present in the sample. The lower the Ct values, the higher the viral RNA loads [23].

RT-PCR is capable of detecting viral RNA for a mean of 17 days [20]. Lower Ct values, which indicate higher viral RNA loads, have been observed in patients who show more severe symptoms [183]. In addition, higher rates of antibody production have been observed in more severely ill patients than in mild cases (Chen et al., 2020). In spite of the advantages, it can still generate false-positive or false-negative results [184].

Usually the methods used for collecting the samples for testing are nasopharyngeal, throat or saliva swabs [185]. Gilsenan-Reed et al., 2020 showed that swabbing is a better method than tissue sampling for PCR detection of viral infections. Throat swabs have been shown to be greatly sensitive during the first week of the disease. In individuals who displayed symptoms related to the lungs, it is observed a higher viral RNA concentration in sputum samples [35]. The composition of the swabs seems to interfere with the sensitivity of the tests. In comparison to polyester, rayon and cotton, nylon swabs seem to be the best option as they are able to retrieve a higher volume of viral particles [184].

The assays usually target the *E*, *N*, *S*, *RdRp* and *ORF1* genes [183,187].

1.4.2. Serological tests

Serological tests detect the presence of antibodies or antigens of SARS-CoV-2 proteins [23]. IgG and IgM antibodies for spike and nucleocapsid proteins are usually used as targets of serological diagnosis [20,23,187]. Nucleocapsid protein is the most abundant protein which makes the assays that detect its specific antibodies the most sensitive. The most specific are the ones that target the antibodies of RBD of the spike protein as it is the protein that binds to the host cells [183].

Although the identification of antibodies against SARS-CoV-2 proteins is not very useful for early detection of the infection [37], it can be very helpful for monitoring the progress of the disease, to identify past infections and to study the epidemiology of the virus [188–190].

Enzyme-linked immunosorbent assay (ELISA) and lateral flow immunoassay (LFIA) are the most common used serological tests for SARS-CoV-2 detection [188].

ELISA can detect antigens or antibodies, by direct or indirect methods. When there is an antigen detection, the color of the substrate changes and it is quantified according to the antigen concentration. In the direct method, there is a primary antibody linked to an enzyme and if the sample has the virus, the antibody binds to the antigen,

and it can be detected directly. In the in indirect approach, there are two types of antibodies: the primary and the secondary. The primary antibody is specific to the antigen we want to detect. If the antigen is present in the sample, it will bind to the primary antibody. After a washing step, only the antibodies that got linked to the target antigen stay in the plate. The secondary enzyme-labeled antibodies, which are specific to the primary antibodies, are added and only the primary antibodies that got attached to the antigens will bind to the secondary antibodies and be detected [23].

Rapid antigen tests (RATs) are a type of lateral flow immunoassay and rely on the binding kinetics of the monoclonal antibodies to the viral antigen [23,185]. This is a qualitative method, so it can detect the presence of the antigens with the naked eye, but cannot quantify them [23].

These types of tests are less sensitive than RT-PCR but are important to get a fast result and are easy to use [191]. They seem to work better with high viral loads (Ct<25). Their specificity is generally high and can be compared to the specificity of RT-PCR tests [184].

The most common target protein present in these tests is the nucleocapsid protein as it is more conserved than the other SARS-CoV-2 proteins [184].

1.5. Importance of autopsies during the pandemic

Forensic medicine should play an important role in identifying “death due to COVID-19” and “death with COVID-19” leading this way to a more accurate mortality statistics, since the person might have the virus but the cause of death could be another co-existing disease [45,192–194]. Edler et al. performed 80 autopsies on individuals infected with SARS-CoV-2 in Germany and concluded that the clinical cause of death and the forensic results were vastly different.

It is also of great importance to understand if SARS-CoV-2 infection would accelerate death in individuals with underlying conditions [195]. Taking tissue samples to perform wider analysis, such as genetic, serological and histopathologic analysis, is important to understand the pathological mechanisms of the virus which can help in prevention and treatment of the disease [192,196,197]. It is well known that SARS-CoV-2 affects the throat and lungs but little is known about its ability to infect or replicate in other organs [24,193].

The virus has been detected until 90-175 hours *post-mortem* but there is no evidence whether it is still infectious or not [198].

It is crucial to understand how long the virus can survive in a body after death to evaluate the likelihood of transmission to the personnel performing autopsies [192].

1.6. Autopsy procedures

It has not been described any case of COVID-19 transmission from a deceased body to an individual to date (19/01/2021) but all the corpses should be handled with care [199,200]. They also should be tested for COVID-19, as described in Santurro et al., 2020. The study performed by Aquila et al., 2021 showed that deceased bodies remain COVID-19 positive for bronchial, nasal and oropharyngeal swabs at least 24 hours after death which means that the virus can survive most of the *post-mortem* biological modifications [203].

During the moving and storage of the corpse, it should be inside a closed bag to reduce the contact with body fluids and all the body opening cavities should be filled with cotton balls soaked with disinfectant. The face and the mouth should be covered with a mask. The professional should use gloves. The ventilation of the morgue should be different from the building one to avoid airborne transmission to other rooms. The bag should be disinfected before reused or going to waste [44,204].

Inside the autopsy room only the professionals needed should be there [204] and they must use a full personal protective equipment (PPE), which includes gloves, another pair of long-sleeved rubber gloves, a long-sleeved water proof gown, FFP2, FFP3 or N95 masks, protective glasses or face shield, a disposable surgical cap and waterproof shoes [44,197,199].

Organs manipulation should be avoided unless it is absolutely necessary and minimally invasive autopsy supported by *post-mortem* imaging should be preferred [204]. Craniotomy box technique was tested by Hasmi et al., 2020 as an alternative to normal skull and brain removal procedures and its use is described in the respective article. Only round-ended scissors and PM40 blades should be used and only one body cavity should be managed at a time [44].

1.7. *Post-mortem* findings

The histopathological findings for COVID-19 described in the literature are in agreement with those described for SARS-CoV and MERS-CoV [68,205].

Diffuse alveolar damage (DAD) is frequently reported in autopsy investigations related to COVID-19 with deposition of hyaline membranes which hampers the gas exchange in the alveoli [20,47,206–210]. In addition, the lungs are presented with vascular occlusion provoked by thrombosis of pulmonary arteries, which leads to a decrease of oxygen levels in the blood (hypoxemia) and consequently to respiratory failure [50,197,211].

It is observed a boost in type-2 pneumocytes which habitually occurs when there is lung damage [47,60,205,208,209]. It has been shown that the co-expression of ACE2 and TMPRSS2 is related to higher risk of SARS-CoV-2 infection [212] and, interestingly, the expression of ACE2 and TMPRSS2 proteins occur in alveolar type-1 and 2 pneumocytes in the areas of DAD [14,60]. It is also known that ACE2 is widely expressed in endothelial cells [34,49].

Studies on COVID-19 related autopsies found that SARS-CoV-2 infection induces the release of megakaryocytes from the bone marrow, which clogs the lung capillaries, and can cause micro or macrothrombi [47,49,210,213,214]. Megakaryocytes are located in the bone marrow and are responsible for the fabrication of thrombocytes (platelets) which are fundamental in the blood clotting process [215]. These cells were also found in the heart, kidneys and skin [215–217]. Interestingly, *ACE2* and *TMPRSS2* are found expressed in platelets and exposure to the virus leads to platelet aggregation, spreading and clot retraction [8].

Normally when there is a pathogen invasion, pro-inflammatory cytokines are released by monocytes and macrophages and B and T cells are brought to the place of the infection so that the virus can be eliminated [16,218]. However, coronaviruses are able to induce a “cytokine cascade/storm” which is generated by an excessive release of inflammatory cytokines, caused by rapid viral replication [50], that ultimately culminate in organ damage due to the maladaptive response of the immune system to the virus infection [8,50,209,210,218,219]. This overproduction of inflammatory responses might cause various histopathologies, such as mucus plugs with fibrinous exudate, desquamation, lung fibrosis (causes shortness of breath), pulmonary oedema (accumulation of fluids in the lungs) and apoptosis of epithelial cells [8,20,208,219–221]. This high expression of cytokines has been associated with *ORF3a* gene in SARS-CoV and MERS-CoV and it is suggested to be similar in SARS-CoV-2 [11].

Ducloyer et al., 2020 described the presence of SARS-CoV-2 RNA in the upper and lower respiratory airways, as well as in feces and blood more than 48h after death. There are cases where COVID-19 is only detected *post-mortem* [210].

Some of the comorbidities found within patients who died with COVID-19 were diabetes mellitus, hypertension, obesity, ischemic cardiac disease, history of stroke, cancer, chronic kidney disease, among others [49,50,197,210,221–223].

Genomic and subgenomic RNA of SARS-CoV-2 were observed in the cornea of deceased bodies but it was not concluded if the virus could still be infectious or not [224]. Alexandersen et al., 2020 claimed that the presence of subgenomic RNAs is not a good indicator of active replication, since these mRNAs can be detected for a long time after the active infection.

1.8. Aims of the study

The aim of this project is to identify the main VOCs and VOIs circulating in the region of Lisbon, applying the methodology of real time RT-PCR in cadavers that tested positive for SARS-CoV-2. To meet this goal, we used three assays: Allplex™ SARS-CoV-2 Variants I, Allplex™ SARS-CoV-2 Variants II and Allplex™ SARS-CoV-2 Variants V. The first one detects defining mutations of the Alpha, Beta and Gamma variants (N501Y, E454K and HV69/70del), the second assay detects mutations present in the Delta, Beta, Gamma and California variants (L452R, K417T, K417N and W152C) and the third one detects defining mutations of the Delta and Lambda variants (L452R, P681R, L452Q and F690S).

In addition, we also want to understand if these RT-PCR assays are efficient to correctly identify these variants, comparing the results obtained to the reference methods for the determination of variants – Next Generation Sequencing (NGS). Two different NGS methodologies were used to compare the results – NGS-ONT (Oxford Nanopore Technologies) and NGS-Sanger based.

2. Methodology

The National Institute of Legal Medicine and Forensic Sciences, South Delegation (Lisbon) receives deceased bodies from Lisbon, Barreiro, Cascais and Torres Vedras. Before the autopsy, as routine due to the current pandemic, they need to be tested for the eventual presence of SARS-CoV-2 by the Laboratory of Virology and Infectious Diseases from the Institute. Nasopharyngeal swabs are collected from the corpses and placed in 1.5 mL of NaCl 0,9% in a microtube. The positive samples are stored in the refrigerator at -23ºC. Allplex™ SARS-CoV-2 and 3plex™ SARS-CoV-2 RT-PCR assays from SeeGene are used as a laboratory routine to detect the presence of SARS-CoV-2 RNA. To identity the genetic variants, Allplex™ SARS-CoV-2 Variants I, Allplex™ SARS-CoV-2 Variants II and Allplex™ SARS-CoV-2 Variants V were used. The samples are inactivated by heat at 56ºC during 30 minutes before the RNA extraction.

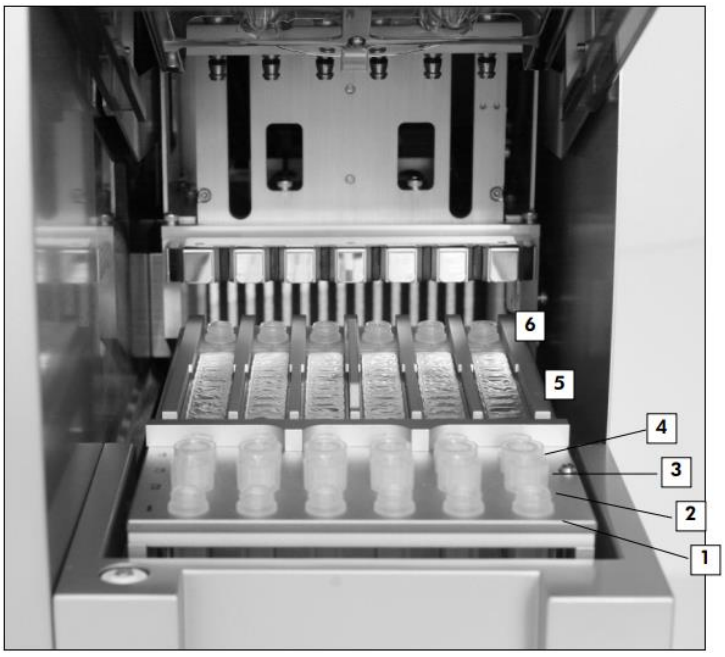
2.1. RNA extraction

Each tube was placed in the vortex for approximately 15-20 seconds at 3000 rpm to free the nucleic acids. RNA was extracted using an automatic extractor, the BioRobot EZ1 DSP. A preprogrammed EZ1 DSP Card for purification of viral nucleic acids was placed in the device. Only six samples can be extracted simultaneously. Elution tubes of

1.5 mL were loaded into the first row (Figure 4 (1)). Disposable tip holders and disposable filter-tips were placed in the second row (Figure 4 (2)). 6 tubes of 1.5 mL containing 60 µL of a mix of carrier RNA and the elution buffer (AVE) were placed in the third row (Figure 4 (3)). The mix contains 25.2 µL of carrier RNA and 394.8 µL of AVE. Carrier RNA is important in the purification step because it helps the binding of viral RNA to the silica surface of the magnetic particles and prevents the RNA degradation if the RNases were not correctly denatured by the chaotropic salts and detergent present in the lysis buffer. 200 µL of each sample was pipetted to the corresponding sample tube of 2 mL to be loaded to the fourth row (Figure 4 (4)). In the fifth row, the reagent cartridges, which contain every necessary reagent for viral RNA extraction except carrier RNA, were placed, after being inverted 5-10 times to mix the reagents (Figure 4 (5)). In the sixth row, tubes of 2 mL were loaded for lysis (Figure 4 (6)).

Then the protocol for virus RNA extraction was set. The option of “200 µL” was selected under the operation “Sample Volume”. We selected a final volume of 60 µL for the RNA elution. When the protocol is finished, the elution tubes of the first row are kept to be used in the RT-PCR and the remaining is discarded.

Figure 4. The image represents the worktable of the BioRobot EZ1 DSP with the respective steps indicated in the section 2.1 (Courtesy of QIAGEN EZ1® DSP Virus Kit Handbook guidelines).



2.2. Pre-PCR

2.2.1. SeeGene Allplex™ SARS-CoV-2 Variants I

The master mix was prepared using the following reagents:

- SC2V1 MOM – a mix of oligonucleotides for amplification and detection of the target genes (primers)
- EM5 – contains the reverse transcriptase, the DNA polymerase, Uracil-DNA glycosylase (to prevent mutagenesis) and a buffer containing deoxyribonucleotides triphosphate (dNTPs)
- EM5 Buffer – buffer for Real Time PCR (Bovine Serum Albumin – BSA – and glycerol)

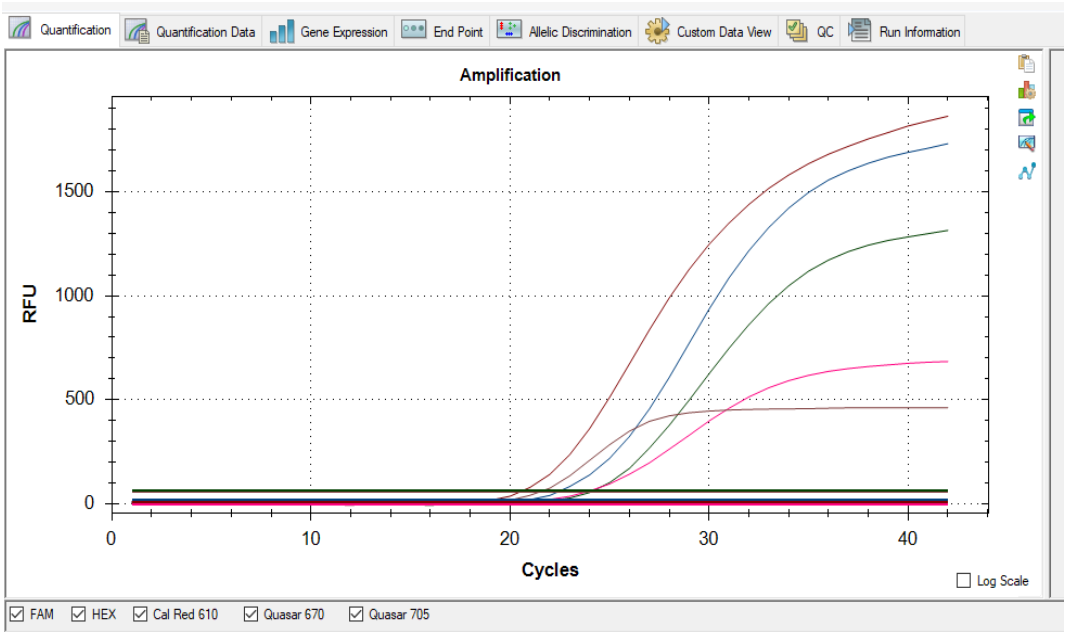
For each reaction, 5 µL of SC2V1 MOM, 5 µL of EM5 and 5 µL of EM5 Buffer were added in a microtube to reach a total volume of 15 µL.

15 µL of the master mix were pipetted in each well of the RT-PCR plate.

Afterwards, 5 µL of each sample (including the positive and negative controls*) were added to each well to reach a total volume of reaction of 20 µL. Then the plate was covered with a plastic strip and ready for the real time RT-PCR.

*The positive control is composed by a mix of the target genes and an internal control (IC) (Figure 5) and the negative control is composed by RNase-free water.

Figure 5. Graphic showing the curves for the positive control on SeeGene Allplex™ SARS-CoV-2 Variants I.



2.2.2. SeeGene Allplex™ SARS-CoV-2 Variants II

The master mix was prepared using the following reagents:

- SC2V2 MOM – a mix of oligonucleotides for amplification and detection of the target genes (primers)
- EM5 – contains the reverse transcriptase, the DNA polymerase, Uracil-DNA glycosylase (to prevent mutagenesis) and a buffer containing deoxyribonucleotides triphosphate (dNTPs)
- EM5 Buffer – buffer for Real Time PCR (Bovine Serum Albumin – BSA – and glycerol)

For each reaction, 5 µL of SC2V2 MOM, 5 µL of EM5 and 5 µL of EM5 Buffer were added in a microtube to reach a total volume of 15 µL.

15 µL of the master mix were pipetted in each well of the RT-PCR plate.

Afterwards, 5 µL of each sample (including the positive and negative controls) were added to each well to reach a total volume of reaction of 20 µL. Then the plate was covered with a plastic strip and ready for the real time RT-PCR.

2.2.3. SeeGene Allplex™ SARS-CoV-2 Variants V

The master mix was prepared using the following reagents:

- SC2V5 MOM – a mix of oligonucleotides for amplification and detection of the target genes (primers)
- EM5 – contains the reverse transcriptase, the DNA polymerase, Uracil-DNA glycosylase (to prevent mutagenesis) and a buffer containing deoxyribonucleotides triphosphate (dNTPs)
- EM5 Buffer – buffer for Real Time PCR (Bovine Serum Albumin – BSA – and glycerol)

For each reaction, 5 µL of SC2V5 MOM, 5 µL of EM5 and 5 µL of EM5 Buffer were added in a microtube to reach a total volume of 15 µL.

15 µL of the master mix were pipetted in each well of the RT-PCR plate.

Afterwards, 5 µL of each sample (including the positive and negative controls) were added to each well to reach a total volume of reaction of 20 µL. Then the plate was covered with a plastic strip and ready for the real time RT-PCR.

2.3. RT-PCR

The extracted RNA was amplified using CFX96™ Real-time PCR Detection System (Bio-Rad). 72 samples were run on RT-PCR with SeeGene Allplex™ SARS-CoV-2 Variants I, 33 samples with SeeGene Allplex™ SARS-CoV-2 Variants II and 14 samples with SeeGene Allplex™ SARS-CoV-2 Variants V assay.

SeeGene Allplex™ SARS-CoV-2 Variants I is an assay that amplifies and detects mutations on the *spike* gene that originate the amino acid changes E484K, N501Y and HV69/70del, and it also detects the *RdRp* gene of SARS-CoV-2, while SeeGene Allplex™ SARS-CoV-2 Variants II Assay amplifies and detects the *spike* mutations that originate the amino acid changes L452R, W152C, K417T and K417N, through multiplex RT-PCR. The first assay can detect mutations present in the Alpha variant (N501Y and HV69/70del), in the Beta variant (E484K and N501Y) and in the Gamma variant (E484K and N501Y). This assay alone cannot differentiate between the Beta and the Gamma variants but when used together with SeeGene Allplex™ SARS-CoV-2 Variants II, it can differentiate between the two, since it detects K417N (present in the Beta variant) and K417T (present in the Gamma variant). This second assay can also detect L452R, present in the Delta variant and in the California variant (B.1.427 and B.1.429 lineages), and W152C present exclusively in the California variant.

SeeGene Allplex™ SARS-CoV-2 Variants V Assay was developed later and detects the following spike mutations: L452Q, L452R, F490S and P681R. L452R and P681R are present in the Delta variant while L452Q and F490S are found within a new variant called Lambda (C.37 lineage).

The experiment was set up according to SeeGene Allplex™ SARS-CoV-2 Variants I, SeeGene Allplex™ SARS-CoV-2 Variants II and SeeGene Allplex™ SARS-CoV-2 Variants V assay recommendations (which are the same), following three main steps: “Protocol Setup”, “Plate Setup” and “Start Run”.

Under the operation “Protocol setup”, in the main menu, it was selected “File”, following by “New” and finally “Protocol” which led to the next operation – Protocol Editor. Here we defined the number of cycles, temperatures and duration for each step of the thermal cycler, as shown in Table 4. Step 1 was set at 50ºC during 20 minutes and step 2 at 95ºC during 15 minutes. In the “number of cycles” option, we selected “1”. Step 3 was set at 95ºC during 10 seconds; step 4 at 60ºC during 40 seconds and step 5 at 72ºC during 20 seconds. For these 3 steps, under the “number of cycles” option, we selected “3”. On step 6, we selected “GOTO step 3, twice”. Step 7 was set at 95ºC during 10 seconds; step 8 at 60ºC during 15 seconds and step 9 at 72ºC during 10 seconds. For

these steps, “42” was selected under the option “number of cycles”. On step 10, we selected “GOTO step 7, 41 times”. The fluorescence is detected at 60ºC and 72ºC during steps 8 and 9.

Under the “Sample Volume” option, we introduced “20 µL” and saved the protocol.

Under the operation “Experiment Setup”, we selected “Plate”, followed by “Create New” which led to the “Plate Editor”. Under the option “Select Fluorophores”, we selected the fluorophores “FAM”, “HEX”, “Cal Red 610”, “Quasar 670” and “Quasar 705” (Table 5, Table 6 and Table 7).

In the main menu of “Plate Editor”, under the operation “Settings”, we selected “Plate Size (96 wells)” and “Plate Type (BR White)”.

Then, under the operation “Start Run” in the “Experiment Setup” window, we selected “Close Lid” and saved the assays.

Table 4. Number of cycles, temperatures and duration for each step of the thermal cycler.

Step	Number of cycles	Temperature	Duration
1	1	50ºC	20 min
2		95ºC	15 min
3	3	95ºC	10 sec
4		60ºC	40 sec
5		72ºC	20 sec
6	GOTO Step 3, 2 more times		
7	42	95ºC	10 sec
8		60ºC	15 sec
9		72ºC	10 sec
10	GOTO Step 7, 41 more times		

Table 5. Mutations/genes detected by each fluorophore and respective colors with SeeGene Allplex™ SARS-CoV-2 Variants I Assay.

FAM	E484K
Cal Red 610	N501Y
Quasar 705	HV69/70 del

HEX	*RdRp* gene
Quasar 670	IC

Table 6. Mutations detected by each fluorophore and respective colors with SeeGene Allplex™ SARS-CoV-2 Variants II Assay.

FAM	L452R
Cal Red 610	K417T
Quasar 705	K417N
HEX	W152C
Quasar 670	IC

Table 7. Mutations detected by each fluorophore and respective colors with SeeGene Allplex™ SARS-CoV-2 Variants V Assay.

FAM	L452Q
Cal Red 610	P681R
Quasar 705	L452R
HEX	F490S
Quasar 670	IC

2.4. Sequencing

To validate the RT-PCR results, each sample was sequenced by two methods of Next Generation Sequencing (NGS): NGS-Oxford Nanopore Technologies (NGS-ONT) and NGS-Sanger Based. NGS-ONT was performed by my laboratory colleague Joana Rodrigues and NGS-SB by INSA.

NGS-Sanger Based works based on the principles of Sanger sequencing, being that fluorescent-labeled nucleotides are added by DNA polymerase and the signal emitted by the fluorescence is used to decode each base. For sample preparation, the DNA has to be divided in fragments of the same size and then sequencing adaptors are added so that the DNA fragments can bind to the flow cell. Each fragment is amplified by PCR, generating a cluster. DNA polymerase can only add one base at a time so for each cluster, after each base is added, there is a washing step so that the next base can be added and this process is repeated over and over until a complete sequence of that fragment is achieved. Then the computer compares all the reads and generates a contiguous sequence [226].

Nanopore sequencing works with a nanopore protein to detect the RNA sequence. When the RNA molecule passes through the nanopore, it measures the difference in the electric current. This change in the current produces a signal that is then decoded to define the RNA sequence in real time[227].

2.5. Data analysis

72 samples were tested with both methodologies (RT-PCR and NGS) but only 60 could be used for comparison purposes. In 12 samples it was not possible to obtain results with NGS. RT-PCR results were compared to at least one of the sequencing methods in 39 samples and compared to both methods in 21 samples (Section 3.3).

A comparison between the common mutations detected by RT-PCR and NGS was performed (Section 3.4).

A comparison between the variants attributed by RT-PCR and NGS was performed to see if they matched in all samples (Section 3.5).

For each assay individually, after the assignment of the variants, we checked if in all of them the defining mutations of the respective variant were detected (to confirm if they are efficient methods to detect these variants) (Section 3.6).

3. Results and discussion

3.1. RT-PCR

The Ct value is defined as the number of cycles required for the fluorescent signal to cross the background level of fluorescence [182]. The results were considered valid when the positive control showed a Ct value inferior to 42 for each mutation/gene amplified and the negative control showed no signs of amplification. If a target gene and the Internal Control (IC) show fluorescence, it means that the nucleic acid was detected. If a target gene was amplified but not the IC, it can still be considered valid and positive for that target (high viral RNA loads might hamper the detection of the IC). Results that did not show any signal for a target gene nor for the IC were considered invalid and therefore were excluded from the experiment.

3.1.1. SeeGene Allplex™ SARS-CoV-2 Variants I Assay

With SeeGene Allplex™ SARS-CoV-2 Variants I Assay, 72 samples were tested.

The nucleotide substitution 1501A>T (amino acid change N501Y) was detected in 28 samples and 204_209delACATGT (amino acid deletion HV69/70) was detected in 59 samples (Table 8).

23012G>T (amino acid change E484K) was not detected in any sample.

HV69/70del alone was detected in 31 samples (Table 8).

The combination of N501Y and HV69/70del occurred in 28 samples (Table 8).

Table 8. RT-PCR results obtained with SeeGene Allplex™ SARS-CoV-2 Variants I Assay. The "X" indicates the presence of N501Y and/or HV69/70del in that sample.

Sample	c.1501A>T (p.N501Y)	c.204_209delACATGT (p.HV69/70del)
LX196	X	X
CV6825	X	X
CV6546	X	X
CV6864	X	X
CV6561	X	X
CV6919	X	X
CV6595	X	X
CV6900	X	X
CV6609	X	X
CF1688	X	X
LX271	X	X
LX181	X	X
LX197	X	X
LX225	X	X
LX270	X	X
LX320	X	X
LX209	X	X
CC73	X	X
LX525	X	X
TV53	X	X
BRR411	X	X
LX145	X	X
LX380	X	X
LX516	X	X
TV60	X	X
LX508	X	X
LX384	X	X

Sample	c.1501A>T (p.N501Y)	c.204_209delACATGT (p.HV69/70del)
408		X
LX1035		X
LX1039		X
LX1082		X
LX1119		X
CC212		X
LX1144		X
LX1155		X
CF1824		X
CF1830		X
CV8057		X
CV7996	X	X
CV8064		X
CV8068		
CV8072		
CV8074		
LX2992		
LX3009		
LX3010		
LX2993		
CV9713		X
CV9705		
CV9704		
LX3019		
CF1980		
CV9738		X
LGM179		X
LX3028		
LX3021		X
CV9575		X
LX3034		X
CDA2237		X
LX3147		
LX3152		X
CF2012		X
LX3155		X

Sample	c.1501A>T (p.N501Y)	c.204_209delACATGT (p.HV69/70del)
CF2018		X
CV10172		X
CV10191		X
CV10192		X
LX3162		X
LX3166		X
CV10205		X
CF2025		X
326		X

3.1.2. SeeGene Allplex™ SARS-CoV-2 Variants II Assay

With SeeGene Allplex™ SARS-CoV-2 Variants II Assay, 33 samples were tested.

The nucleotide substitution 1355T>G (amino acid change L452R) was detected in 30 samples (Table 9).

The combination of L452R and 22812G>T (amino acid change K417N) occurred in 1 sample (Table 9).

The nucleotide substitution 22812A>C (amino acid change K417T) and 22018G>T (amino acid change W152C) were not detected in any sample.

Table 9. RT-PCR results obtained with SeeGene Allplex™ SARS-CoV-2 Variants II Assay. The "X" indicates the presence of L452R and/or K417N in that sample.

Sample	c.1355T>G (L452R)	22812G>T (K417N)
408	X	
LX1035	X	
LX1039	X	
LX1082	X	
LX1119	X	
CC212	X	
LX1144	X	
LX1155	X	
CV8057	X	
CV7996		
CV8064	X	
CV8068	X	
CV8072	X	X

Sample	c.1355T>G (L452R)	22812G>T (K417N)
CV8074	X	
LX2992	X	
LX3009	X	
LX3010	X	
LX2993		
CV9713	X	
CV9705	X	
CV9704	X	
LX3019	X	
CF1980	X	
CV9738	X	
LGM179	X	
LX3028		
LX3021	X	
CV9575	X	
LX3034	X	
CDA2237	X	
LX3147	X	
LX3152	X	
CF2012	X	

3.1.3. SeeGene Allplex™ SARS-CoV-2 Variants V Assay

With SeeGene Allplex™ SARS-CoV-2 Variants V Assay, 14 samples were tested.

The nucleotide substitution 1355T>G (amino acid change L452R) and 2042C>G (amino acid change P681R) were both detected in all samples (n=14) (Table 10).

L452Q and F490S were not detected in any sample.

Table 10. RT-PCR results obtained with SeeGene Allplex™ SARS-CoV-2 Variants V Assay. The "X" indicates the presence of P681R and/or L452R in that sample.

Sample	c.2042C>G (P681R)	c.1355T>G (L452R)
CF1824	X	X
CF1830	X	X
LX2993	X	X
LX3028	X	X

Sample	c.2042C>G (P681R)	c.1355T>G (L452R)
LX3155	X	X
CF2018	X	X
CV10172	X	X
CV10191	X	X
CV10192	X	X
LX3162	X	X
LX3166	X	X
CV10205	X	X
CF2025	X	X
326	X	X

3.1.4. Overall results

In the samples where SeeGene Allplex™ SARS-CoV-2 Variants I Assay was applied, when N501Y and HV69/70del were amplified, the sequence was considered to belong to the Alpha variant.

In the samples where SeeGene Allplex™ SARS-CoV-2 Variants II Assay was applied, when L452R was detected, the sequence was assigned to the Delta variant.

In the samples where SeeGene Allplex™ SARS-CoV-2 Variants V Assay was applied, when L452R and P681R were detected, the sequence was considered to belong to the Delta variant.

In the samples where both assays SeeGene Allplex™ SARS-CoV-2 Variants I and SeeGene Allplex™ SARS-CoV-2 Variants II were applied, when L452R and HV69/70del were amplified, the sequence was considered to belong to the Delta variant.

In the samples where SeeGene Allplex™ SARS-CoV-2 Variants II was applied, when L452R and K417N were detected, the sequence was considered to belong to the Delta variant.

With the three assays together:

- The combination of N501Y and HV69/70del occurred in 28 samples. The lineage attributed to these samples was B.1.1.7 (Alpha variant) (Table 11).
- The combination of L452R and HV69/70del was found in 18 samples. The lineage attributed to these samples was B.1.617.2 (Delta variant) (Table 11).
- The combination of L452R and P681R was found in 2 samples. The lineage attributed to these samples was B.1.617.2 (Delta variant) (Table 11).

- The combination of L452R, P681R and HV69/70del was found in 12 samples. The lineage attributed to these samples was B.1.617.2 (Delta variant) (Table 11).
- The combination of L452R and K417N occurred in 1 sample, to which it was attributed the lineage B.1.617.2 (Delta variant) (Table 11).
- L452R alone occurred in 11 samples. The lineage attributed to these samples was B.1.617.2 (Delta variant) (Table 11).

Overall, based on RT-PCR results, 38.89% of the samples (n=28) belong to the lineage B.1.1.7 (Alpha variant) and 61.12% (n=44) belong to B.1.617.2 (Delta variant).

Table 11. Comparison of the overall RT-PCR results of the three assays together (SeeGene Allplex™ SARS-CoV-2 Variants I, SeeGene Allplex™ SARS-CoV-2 Variants II and SeeGene Allplex™ SARS-CoV-2 Variants V). The table shows the mutations detected in each sample and the respective variant attributed to them by RT-PCR.

Samples	Variant	c.1501A>T (p.N501Y)	c.204_209delACATGT (p.HV69/70del)	c.1355T>G (p.L452R)	c.2042C>G (p.P681R)	c.22812G>T (p.K417N)
LX196	B.1.1.7	X	X			
CV6825	B.1.1.7	X	X			
CV6546	B.1.1.7	X	X			
CV6864	B.1.1.7	X	X			
CV6561	B.1.1.7	X	X			
CV6919	B.1.1.7	X	X			
CV6595	B.1.1.7	X	X			
CV6900	B.1.1.7	X	X			
CV6609	B.1.1.7	X	X			
CF1688	B.1.1.7	X	X			
LX271	B.1.1.7	X	X			
LX181	B.1.1.7	X	X			
LX197	B.1.1.7	X	X			
LX225	B.1.1.7	X	X			
LX270	B.1.1.7	X	X			
LX320	B.1.1.7	X	X			
LX209	B.1.1.7	X	X			
CC73	B.1.1.7	X	X			
LX525	B.1.1.7	X	X			
TV53	B.1.1.7	X	X			
BRR411	B.1.1.7	X	X			
LX145	B.1.1.7	X	X			
LX380	B.1.1.7	X	X			
LX516	B.1.1.7	X	X			

Samples	Variant	c.1501A>T (p.N501Y)	c.204_209delACATGT (p.HV69/70del)	c.1355T>G (p.L452R)	c.2042C>G (p.P681R)	c.22812G>T (p.K417N)
TV60	B.1.1.7	X	X			
LX508	B.1.1.7	X	X			
LX384	B.1.1.7	X	X			
408	B.1.617.2		X	X		
LX1035	B.1.617.2		X	X		
LX1039	B.1.617.2		X	X		
LX1082	B.1.617.2		X	X		
LX1119	B.1.617.2		X	X		
CC212	B.1.617.2		X	X		
LX1144	B.1.617.2		X	X		
LX1155	B.1.617.2		X	X		
CF1824	B.1.617.2		X	X	X	
CF1830	B.1.617.2		X	X	X	
CV8057	B.1.617.2		X	X		
CV7996	B.1.1.7	X	X			
CV8064	B.1.617.2		X	X		
CV8068	B.1.617.2			X		
CV8072	B.1.617.2			X		X
CV8074	B.1.617.2			X		
LX2992	B.1.617.2			X		
LX3009	B.1.617.2			X		
LX3010	B.1.617.2			X		
LX2993	B.1.617.2			X	X	
CV9713	B.1.617.2		X	X		
CV9705	B.1.617.2			X		
CV9704	B.1.617.2			X		
LX3019	B.1.617.2			X		
CF1980	B.1.617.2			X		
CV9738	B.1.617.2			X		
LGM179	B.1.617.2		X	X		
LX3028	B.1.617.2			X	X	
LX3021	B.1.617.2		X	X		
CV9575	B.1.617.2		X	X		
LX3034	B.1.617.2		X	X		
CDA2237	B.1.617.2		X	X		
LX3147	B.1.617.2			X		
LX3152	B.1.617.2		X	X		
CF2012	B.1.617.2		X	X		
LX3155	B.1.617.2		X	X	X	

Samples	Variant	c.1501A>T (p.N501Y)	c.204_209delACATGT (p.HV69/70del)	c.1355T>G (p.L452R)	c.2042C>G (p.P681R)	c.22812G>T (p.K417N)
CF2018	B.1.617.2		X	X	X	
CV10172	B.1.617.2		X	X	X	
CV10191	B.1.617.2		X	X	X	
CV10192	B.1.617.2		X	X	X	
LX3162	B.1.617.2		X	X	X	
LX3166	B.1.617.2		X	X	X	
CV10205	B.1.617.2		X	X	X	
CF2025	B.1.617.2		X	X	X	
326	B.1.617.2		X	X	X	

3.2. Next Generation Sequencing (NGS)

3.2.1. NGS-Oxford Nanopore Technologies (NGS-ONT)

With this sequencing method, results were obtained for 33 samples.

The deletion 204_209ACATGT (amino acid deletion **HV69/70**) was detected in 1 samples, 432_434TTA (amino acid deletion Y145) in 2 samples, 665C>T, 1055C>T (amino acid change A352V) in 1 sample, 1355T>G (amino acid change **L452R**) in 23 samples, 1501A>T (amino acid change **N501Y**) in 5 samples, 1709C>A (amino acid change A570D) in 8 samples, 1841A>G (amino acid change D614G) in 35 samples, 2042C>A (amino acid change P681H) in 5 samples, 2042C>G (amino acid change **P681R**) in 21 samples, 2147C>T (amino acid change T716I) in 5 samples, 2944T>G (amino acid change S982A) in 6 samples, 3352G>C (amino acid change D1118H) in 8 samples, 4752C>T (amino acid change T678I) in 1 sample, 21618C>G (amino acid change T19R) in 9 samples, 22971C>T (amino acid change T470I) in 1 sample, 22995C>A (amino acid change T478K) in 23 samples, 24051A>G (amino acid change D830G) in 1 sample, 24124G>T (amino acid change K854N) in 2 samples, 24197G>T (amino acid change A879S) in 1 sample, 24410G>A (amino acid change D950N) in 8 samples, 24872G>T (amino acid change V1104L) in 1 sample, 24883G>T (amino acid change R1107S) in 1 sample.

3.2.2. NGS-Sanger Based (NGS-SB)

With this sequencing method, results were obtained for 48 samples.

The deletion 204_209ACATGT (amino acid deletion **HV69/70**) was detected in 15 samples, 432_434TTA (amino acid deletion Y145) in 12 samples, 1355T>G (amino

acid change **L452R**) in 31 samples, 1501A>T (amino acid change **N501Y**) in 17 samples, 1709C>A (amino acid change A570D) in 14 samples, 1841A>G (amino acid change D614G) in 49 samples, 2042C>A (amino acid change P681H) in 17 samples, 2042C>G (amino acid change **P681R**) in 31 samples, 2147C>T (amino acid change T716I) in 17 samples, 2944T>G (amino acid change S982A) in 3 samples, 3352G>C (amino acid change D1118H) in 17 samples, 21618C>G (amino acid change T19R) in 33 samples, 21846 C>T (amino acid change T95I) in 11 samples, 21995T>C (amino acid change Y145H) in 1 sample, 22151A>G (amino acid I197V) in 1 sample, 22216G>T (amino acid change Q218H) in 1 sample, 22314C>T (amino acid change P251L) in 2 samples, 22806C>A (amino acid change T415N) in 1 sample, 22971C>T (amino acid change T470I) in 1 sample, 22995C>A (amino acid change T478K) in 31 samples, 23262G>T (amino acid change R567I) in 1 sample, 23484A>G (amino acid change N641S) in 3 samples, 23593G>C (amino acid change Q677H) in 1 sample, 24124G>T (amino acid change K854N) in 1 sample, 24410G>A (amino acid change D950N) in 31 samples, 24872G>T (amino acid change V1104L) in 1 sample, 25151C>T (amino acid change (L1197F) in 1 sample.

3.2.3. Overall results

With NGS-ONT alone, 1 sample was assigned to the lineage B.1.1.7 (Alpha variant) and 11 samples to the lineage B.1.617.2 (Delta variant).

With NGS-SB alone, 11 samples were attributed to the lineage B.1.1.7 (Alpha variant) and 16 samples to the lineage B.1.617.2 (Delta variant).

With the two NGS methods together, 6 samples were attributed to the lineage B.1.1.7 (Alpha variant) and 15 samples to the lineage B.1.617.2 (Delta variant).

3.3. RT-PCR and NGS results in each sample

The 12 samples where it was not possible to obtain results with NGS and, therefore, a comparison with RT-PCR results could not be made were the following: LX196, CV6825, CV6864, LX197, LX270, TV53, LX380, TV60, LX384, LX1082, CV7996 and CV8074.

3.3.1. Sample CV6546

With RT-PCR and NGS-SB, the lineage attributed to this sample was B.1.1.7.

The two defining mutations of the lineage B.1.1.7 (N501Y and HV69/70del) were detected by RT-PCR and NGS-SB.

The other spike mutations belonging to B.1.1.7 were detected by NGS-SB.

3.3.2. Sample CV6561

With RT-PCR and NGS-SB, the lineage attributed to this sample was B.1.1.7.

The two defining mutations of the lineage B.1.1.7 (N501Y and HV69/70del) were detected by RT-PCR and NGS-SB.

The other spike mutations belonging to B.1.1.7, except S982A, were detected by NGS-SB.

3.3.3. Sample CV6919

The lineage attributed to this sample according to RT-PCR and NGS-ONT results was B.1.1.7.

Both methods detected N501Y. HV69/70del was detected by RT-PCR but not NGS-ONT.

The rest of the spike mutations of this variant were detected by NGS-ONT, except Y145del.

The following mutations were also detected by NGS-ONT: T1001I, P314L and P403P in orf1ab; L15F in ORF3a; Y73C in ORF8; and D3L in the nucleocapsid.

3.3.4. Sample CV6595

With RT-PCR and NGS-SB, the lineage assigned to this sample was B.1.1.7.

The two defining mutations of the lineage B.1.1.7 (N501Y and HV69/70del) were detected by RT-PCR and NGS-SB.

The other spike mutations belonging to B.1.1.7 (except S982A) were detected by NGS-SB.

3.3.5. Sample CV6900

With RT-PCR and NGS-SB, the lineage attributed to this sample was B.1.1.7.

The two defining mutations of the lineage B.1.1.7 (N501Y and HV69/70del) were detected by RT-PCR and NGS-SB.

The other spike mutations present in B.1.1.7, except S982A, were detected by NGS-SB.

3.3.6. Sample CV6609

With RT-PCR and NGS-SB, the lineage attributed to this sample was B.1.1.7.

The two defining mutations of the lineage B.1.1.7 (N501Y and HV69/70del) were detected by RT-PCR and NGS-SB.

The other spike mutations belonging to B.1.1.7, except S982A and Y145del, were detected by NGS-SB.

3.3.7. Sample CF1688

The lineage attributed to this sample was the same with the three methods: B.1.1.7.

The two defining mutations of B.1.1.7 (N501Y and HV69/70del) were detected by RT-PCR and NGS-SB, but not with NGS-ONT.

A570D, D614G, S982A and D1118H were detected by both NGS-SB and NGS-ONT.

Y145del, P681H, T716I were detected by NGS-SB but not with NGS-ONT.

The following mutations were also detected by NGS-ONT: V337I, T1001I, A1708D, V3875G and P314L in orf1ab; Y73C in ORF8; and D3L in the nucleocapsid.

3.3.8. Sample LX271

The lineage attributed to this sample was the same with the three methods: B.1.1.7.

N501Y was detected by the three methods, while HV69/70del was detected by RT-PCR and NGS-SB but not with NGS-ONT.

A570D, D614G, P681H, T716I and D1118H were detected by both NGS-SB and NGS-ONT.

Y145del was detected by NGS-SB but not with NGS-ONT.

S982A was detected by NGS-ONT but not with NGS-SB.

The following mutations were also detected by NGS-ONT: T1001I, A1708D, 3675/3677del, P314L, K1383R in orf1ab; D155Y in ORF3a; Y73C in ORF8 and D3L in the nucleocapsid.

3.3.9. Sample LX181

The lineage attributed to this sample was the same with the three methods: B.1.1.7.

N501Y was detected by the three methods, while HV69/70del was detected by RT-PCR and NGS-ONT but not with NGS-SB.

A570D, D614G, P681H, T716I and D1118H were detected by both NGS-SB and NGS-ONT.

Y145del and S982A were detected by NGS-ONT but not with NGS-SB.

The following mutations were also detected by NGS-ONT: T1001I, G1125C, A1708D, 3675/3677del and P314L in orf1ab; Q27, R52I, Y73C, D119E and I121fs in ORF8; and D3L, RG203KR and S235F in the nucleocapsid.

3.3.10. Sample LX225

With RT-PCR and NGS-SB, the lineage attributed to this sample was B.1.1.7.

The two defining mutations of the lineage B.1.1.7 (N501Y and HV69/70del) were detected by RT-PCR and NGS-SB.

The other spike mutations present in B.1.1.7, except A570D, S982A and Y145del, were detected by NGS-SB.

3.3.11. Sample LX320

With RT-PCR and NGS-SB, the lineage attributed to this sample was B.1.1.7.

The two defining mutations of the lineage B.1.1.7 (N501Y and HV69/70del) were detected by RT-PCR and NGS-SB.

The other spike mutations present in B.1.1.7, except A570D, S982A and Y145del, were detected by NGS-SB.

3.3.12. Sample LX209

With RT-PCR and NGS-SB, the lineage assigned to this sequence was the same: B.1.1.7.

Both methods detected N501Y and HV69/70del.

The rest of the spike mutations present in this variant were detected by NGS-SB, except S982A.

3.3.13. Sample CC73

The lineage attributed to this sample was the same with the three methods: B.1.1.7.

N501Y was detected by the three methods, while HV69/70del was detected by RT-PCR and NGS-SB but not with NGS-ONT.

Y145del, A570D, D614G, P681H, T716I and D1118H were detected by both NGS-SB and NGS-ONT.

S982A was not detected by any of the sequencing methods.

The following mutations were also detected by NGS-ONT: R24C, T1001I, A1708D, Q3966R, P314L, K1383R and T1621I in orf1ab; Y73C in ORF8; and D3L in the nucleocapsid.

3.3.14. Sample LX525

With RT-PCR and NGS-SB, the lineage assigned to this sequence was the same: B.1.1.7.

Both methods detected N501Y and HV69/70del.

The rest of the spike mutations of this variant were detected by NGS-SB, except Y145del and S982A.

3.3.15. Sample BRR411

The lineage attributed to this sample was the same with the three methods: B.1.1.7.

N501Y and HV69/70del were detected by RT-PCR and NGS-SB but not with NGS-ONT.

A570D, D614G and D1118H were detected by both NGS-SB and NGS-ONT.

Y145del, T716I and P681H were detected by NGS-SB but not with NGS-ONT.

S982A was detected by NGS-ONT, but not with NGS-SB.

The following mutations were also detected by NGS-ONT: P309L, T1001I, A1708D and P314L in orf1ab; Y73C in ORF8; and D3L, RG203KR and S235F in the nucleocapsid.

3.3.16. Sample LX145

With RT-PCR and NGS-SB, the lineage attributed to this sample was B.1.1.7.

N501Y was detected by the two methods, while HV69/70del by RT-PCR but not with NGS-SB.

A570D, D614G, P681H, T716I and D1118H were detected by NGS-SB.

Y145del and S982A were not detected by NGS-SB.

3.3.17. Sample LX516

The lineage attributed to this sample was the same with the three methods: B.1.1.7.

N501Y was detected by the three methods, while HV69/70del was detected by RT-PCR and NGS-SB but not with NGS-ONT.

A570D, D614G, T716I, P681H and D1118H were detected by both NGS-SB and NGS-ONT.

Y145del was detected by NGS-SB but not with NGS-ONT.

The following mutations were also detected by NGS-ONT: T1001I, A1708D and P314L in orf1ab; Y73C and I121fs in ORF8; and D3L in the nucleocapsid.

3.3.18. Sample LX508

With RT-PCR and NGS-SB, the lineage assigned to this sequence was the same: B.1.1.7.

Both methods detected N501Y and HV69/70del.

The rest of the spike mutations present in this variant were detected by NGS-SB, except Y145del, S982A and A570D.

3.3.19. Sample 408

With RT-PCR and NGS-SB, the lineage attributed to this sample was the same: B.1.617.2.

L452R was detected by both RT-PCR and NGS-SB, while P681R was detected by NGS-SB but not RT-PCR.

The rest of the spike mutations usually present in this variant (T19R, T478K, D614G and D950N) were detected by NSG-SB.

3.3.20. Sample LX1035

The lineage attributed to this sample was the same with the three methods: B.1.617.2.

L452R was detected by the three methods, while P681R was detected by both NGS methods but not RT-PCR.

The rest of the spike mutations present in this variant (T19R, T478K, D614G and D950N) were detected by NSG-SB and NGS-ONT.

3.3.21. Sample LX1039

With RT-PCR and NGS-ONT, the lineage assigned to this sequence was the same: B.1.617.2.

L452R was detected by the two methods, while P681R was detected by NGS-ONT but not RT-PCR.

T478K and D614G were detected by NGS-ONT. However, T19R and D950N were not detected.

3.3.22. Sample LX1119

The lineage attributed to this sample was the same with the three methods: B.1.617.2.

L452R was detected by the three methods, while P681R was detected by NGS-SB but not with RT-PCR and NGS-ONT.

T19R was detected by NSG-SB but not NGS-ONT.

T478K, D614G and D950N were detected by both NGS methods.

3.3.23. Sample CC212

The lineage attributed to this sample was the same with the three methods: B.1.617.2.

L452R was detected by the three methods, while P681R was detected by both NGS methods but not RT-PCR.

T19R was detected by NSG-SB but not NGS-ONT.

T478K, D614G and D950N were detected by both NGS methods.

3.3.24. Sample LX1144

The lineage attributed to this sample was the same with the three methods: B.1.617.2.

L452R was detected by the three methods, while P681R was detected by both NGS methods but not RT-PCR.

T19R was detected by NSG-SB but not NGS-ONT.

T478K and D614G were detected by both NGS methods.

D950N was detected by NGS-SB but not NGS-ONT.

3.3.25. Sample LX1155

With RT-PCR and NGS-ONT, the lineage assigned to this sequence was the same: B.1.617.2.

L452R was detected by the two methods, while P681R was detected by NGS-ONT but not RT-PCR.

T478K, T19R, D950N and D614G were detected by NGS-ONT.

3.3.26. Sample CF1824

The lineage attributed to this sample was the same with the three methods: B.1.617.2.

L452R and P681R were detected by the three methods.

T19R, T478K and D614G were detected by both NGS methods.

D950N was detected by NGS-SB but not NGS-ONT.

3.3.27. Sample CF1830

With RT-PCR and NGS-ONT, the lineage assigned to this sequence was the same: B.1.617.2.

L452R and P681R were detected by both RT-PCR and NGS-ONT.

T478K, T19R and D614G were detected by NGS-ONT.

D950N was not detected by NGS-ONT.

3.3.28. Sample CV8057

With RT-PCR and NGS-SB, the lineage attributed to this sample was the same: B.1.617.2.

L452R was detected by both RT-PCR and NGS-SB, while P681R was detected by NGS-SB but not RT-PCR.

The rest of the spike mutations usually present in this variant (T19R, T478K, D614G and D950N) were detected by NSG-SB.

3.3.29. Sample CV8064

With RT-PCR and NGS-SB, the lineage attributed to this sample was the same: B.1.617.2.

L452R was detected by both RT-PCR and NGS-SB, while P681R was detected by NGS-SB but not RT-PCR.

The rest of the spike mutations present in this variant (T19R, T478K, D614G and D950N) were detected by NSG-SB.

3.3.30. Sample CV8068

The lineage attributed to this sample was the same with the three methods: B.1.617.2.

L452R was detected by RT-PCR and NGS-SB, but not with NGS-ONT, while P681R was detected only by NGS-SB.

D614G was detected by both NGS methods.

T19R, T478K and D950N were detected by NGS-SB but not NGS-ONT.

3.3.31. Sample CV8072

With RT-PCR and NGS-SB, the lineage attributed to this sample was the same: B.1.617.2.

L452R was detected by both RT-PCR and NGS-SB, while P681R was detected by NGS-SB but not RT-PCR.

The rest of the spike mutations present in this variant (T19R, T478K, D614G and D950N) were detected by NSG-SB.

3.3.32. Sample LX2992

With RT-PCR and NGS-ONT, the lineage assigned to this sequence was the same: B.1.617.2.

L452R was detected by both RT-PCR and NGS-ONT, while P681R was detected by NGS-ONT but not RT-PCR.

T478K and D614G were detected by NGS-ONT.

T19R and D950N were not detected by NGS-ONT.

3.3.33. Sample LX3009

The lineage attributed to this sample was the same with the three methods: B.1.617.2.

L452R was detected with the three methods, while P681R was detected by both NGS methods but not RT-PCR.

D614G and T478K were detected by both NGS methods.

T19R and D950N were detected by NGS-SB but not NGS-ONT.

3.3.34. Sample LX3010

With RT-PCR and NGS-SB, the lineage assigned to this sample was B.1.617.2.

L452R was detected by RT-PCR and NGS-SB, while P681R was detected by NGS-SB but not with RT-PCR.

The other spike mutations present in B.1.617.2 (T478K, T19R, D950N, D614G) were detected by NGS-SB.

3.3.35. Sample LX2993

With RT-PCR and NGS-ONT, the lineage assigned to this sequence was the same: B.1.617.2.

L452R and P681R were detected by both RT-PCR and NGS-ONT.

T478K and D614G were detected by NGS-ONT.

T19R and D950N were not detected by NGS-ONT.

3.3.36. Sample CV9713

With RT-PCR and NGS-SB, the lineage attributed to this sample was the same: B.1.617.2.

L452R was detected by both RT-PCR and NGS-SB, while P681R was detected by NGS-SB but not RT-PCR.

The rest of the spike mutations usually present in this variant (T19R, T478K, D614G and D950N) were detected by NSG-SB.

3.3.37. Sample CV9705

With RT-PCR and NGS-SB, the lineage attributed to this sample was the same: B.1.617.2.

L452R was detected by both RT-PCR and NGS-SB, while P681R was detected by NGS-SB but not RT-PCR.

The rest of the spike mutations usually present in this variant (T19R, T478K, D614G and D950N) were detected by NSG-SB.

3.3.38. Sample CV9704

With RT-PCR and NGS-ONT, the lineage assigned to this sequence was the same: B.1.617.2.

L452R was detected by both RT-PCR and NGS-ONT, while P681R was detected by NGS-ONT but not RT-PCR.

T478K and D614G were detected by NGS-ONT.

T19R and D950N were not detected by NGS-ONT.

3.3.39. Sample LX3019

With RT-PCR and NGS-SB, the lineage attributed to this sample was the same: B.1.617.2.

L452R was detected by both RT-PCR and NGS-SB, while P681R was detected by NGS-SB but not RT-PCR.

The rest of the spike mutations present in this variant (T19R, T478K, D614G and D950N) were detected by NSG-SB.

3.3.40. Sample CF1980

With RT-PCR and NGS-SB, the lineage attributed to this sample was the same: B.1.617.2.

L452R was detected by both RT-PCR and NGS-SB, while P681R was detected by NGS-SB but not RT-PCR.

The rest of the spike mutations present in this variant (T19R, T478K, D614G and D950N) were detected by NSG-SB.

3.3.41. Sample CV9738

With RT-PCR and NGS-SB, the lineage attributed to this sample was the same: B.1.617.2.

L452R was detected by both RT-PCR and NGS-SB, while P681R was detected by NGS-SB but not RT-PCR.

The rest of the spike mutations present in this variant (T19R, T478K, D614G and D950N) were detected by NSG-SB.

3.3.42. Sample LGM179

The lineage attributed to this sample was the same with the three methods: B.1.617.2.

L452R was detected by the three methods, while P681R was detected by both NGS methods but not RT-PCR.

D614G, T19R and T478K were detected by both NGS methods.

D950N was detected by NGS-SB but not NGS-ONT.

3.3.43. Sample LX3028

With RT-PCR and NGS-ONT, the lineage assigned to this sequence was the same: B.1.617.2.

L452R was detected by both RT-PCR and NGS-ONT, while P681R was detected by RT-PCR but not NGS-ONT.

T478K and D614G were detected by NGS-ONT.

T19R and D950N were not detected by NGS-ONT.

3.3.44. Sample LX3021

The lineage attributed to this sample was the same with the three methods: B.1.617.2.

L452R was detected with RT-PCR and NGS-SB, but not with NGS-ONT, while P681R was detected only by NGS-SB.

D614G was detected by both NGS methods.

T19R, D950N and T478K were detected by NGS-SB but not NGS-ONT.

3.3.45. Sample CV9575

With RT-PCR and NGS-ONT, the lineage assigned to this sequence was the same: B.1.617.2.

L452R was detected by both RT-PCR and NGS-ONT, while P681R was detected by NGS-ONT but not RT-PCR.

T478K, T19R, D950N and D614G were detected by NGS-ONT.

3.3.46. Sample LX3034

With RT-PCR and NGS-SB, the lineage attributed to this sample was the same: B.1.617.2.

L452R was detected by both RT-PCR and NGS-SB, while P681R was detected by NGS-SB but not RT-PCR.

The rest of the spike mutations present in this variant (T19R, T478K, D614G and D950N) were detected by NSG-SB.

3.3.47. Sample CDA2237

With RT-PCR and NGS-SB, the lineage attributed to this sample was the same: B.1.617.2.

L452R was detected by both RT-PCR and NGS-SB, while P681R was detected by NGS-SB but not RT-PCR.

The rest of the spike mutations usually present in this variant (T19R, T478K, D614G and D950N) were detected by NSG-SB.

3.3.48. Sample LX3147

With RT-PCR and NGS-SB, the lineage assigned to this sample was B.1.617.2.

L452R was detected by RT-PCR and NGS-SB, while P681R was detected by NGS-SB but not with RT-PCR.

The other spike mutations present in B.1.617.2 (T478K, T19R, D950N, D614G) were detected by NGS-SB.

3.3.49. Sample LX3152

With RT-PCR and NGS-SB, the lineage attributed to this sample was the same: B.1.617.2.

L452R was detected by both RT-PCR and NGS-SB, while P681R was detected by NGS-SB but not RT-PCR.

The rest of the spike mutations usually present in this variant (T19R, T478K, D614G and D950N) were detected by NSG-SB.

3.3.50. Sample CF2012

With RT-PCR and NGS-SB, the lineage attributed to this sample was the same: B.1.617.2.

L452R was detected by both RT-PCR and NGS-SB, while P681R was detected by NGS-SB but not RT-PCR.

The rest of the spike mutations usually present in this variant (T19R, T478K, D614G and D950N) were detected by NSG-SB.

3.3.51. Sample LX3155

The lineage attributed to this sample was the same with the three methods: B.1.617.2.

L452R and P681R were detected with the three methods.

T19R, D950N, D614G and T478K were detected by both NGS methods.

3.3.52. Sample CF2018

The lineage attributed to this sample was the same with the three methods: B.1.617.2.

L452R and P681R were detected with the three methods.

D614G and T478K were detected by both NGS methods.

T19R and D950N were detected by NGS-SB but not NGS-ONT.

3.3.53. Sample CV10172

With RT-PCR and NGS-ONT, the lineage assigned to this sequence was the same: B.1.617.2.

L452R and P681R were detected by RT-PCR but not with NGS-ONT.

D614G was detected by NGS-ONT, while T478K, T19R and D950N were not detected by NGS-ONT.

3.3.54. Sample CV10191

The lineage attributed to this sample was the same with the three methods: B.1.617.2.

L452R and P681R were detected with the three methods.

D614G and T478K were detected by both NGS methods.

T19R and D950N were detected by NGS-SB but not NGS-ONT.

3.3.55. Sample CV10192

The lineage attributed to this sample was the same with the three methods: B.1.617.2.

L452R and P681R were detected by the three methods.

T19R, D614G and T478K were detected by both NGS methods.

D950N was detected by NGS-SB but not NGS-ONT.

3.3.56. Sample LX3162

The lineage attributed to this sample was the same with the three methods: B.1.617.2.

L452R and P681R were detected with the three methods.

D950N, D614G and T478K were detected by both NGS methods.

T19R was detected by NGS-SB but not NGS-ONT.

3.3.57. Sample LX3166

With RT-PCR and NGS-SB, the lineage attributed to this sample was the same: B.1.617.2.

L452R and P681R were detected by both RT-PCR and NGS-SB.

The rest of the spike mutations usually present in this variant (T19R, T478K, D614G and D950N) were detected by NSG-SB.

3.3.58. Sample CV10205

With RT-PCR and NGS-ONT, the lineage assigned to this sequence was the same: B.1.617.2.

L452R and P681R were detected by both RT-PCR and NGS-ONT.

D614G and T478K were detected by NGS-ONT, while T19R and D950N were not detected.

3.3.59. Sample CF2025

The lineage attributed to this sample was the same with the three methods: B.1.617.2.

L452R and P681R were detected with the three methods.

T19R, D614G and T478K were detected by both NGS methods.

D950N was detected by NGS-SB but not NGS-ONT.

3.3.60. Sample 326

With RT-PCR and NGS-ONT, the lineage assigned to this sequence was the same: B.1.617.2.

L452R and P681R were detected by both RT-PCR and NGS-ONT.

D614G and T478K were detected by NGS-ONT, while T19R and D950N were not detected.

3.4. Comparison of the detected mutations between RT-PCR and NGS

The only mutations that can be compared between RT-PCR and NGS results are N501Y, HV69/70del, L452R and P681R.

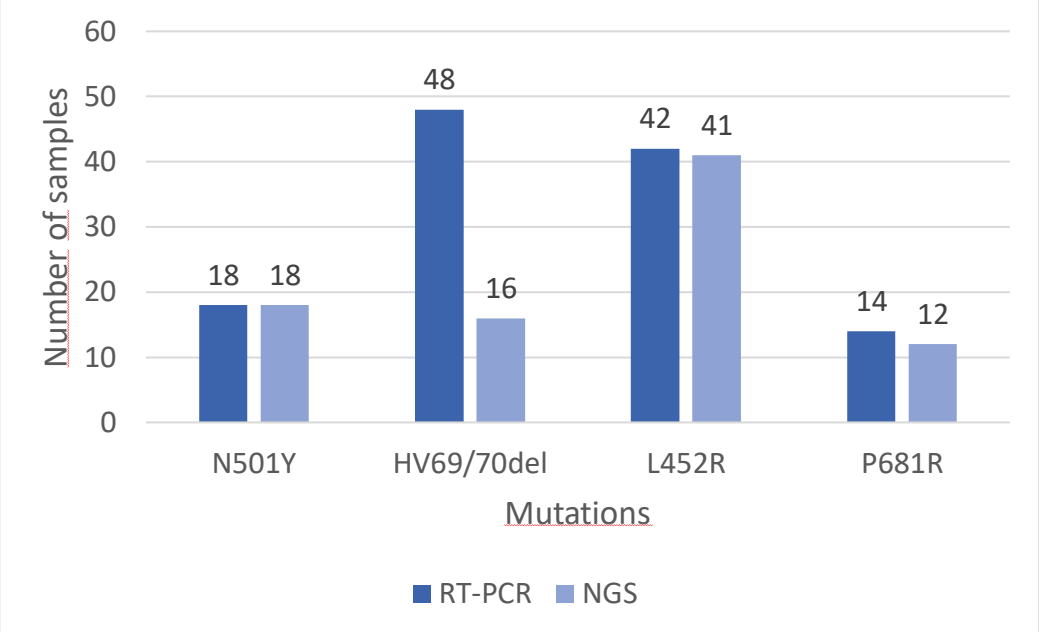
With RT-PCR, N501Y was detected in 18 samples. These results matched at least one of the NGS methods in all samples (n=18) (Figure 6).

With RT-PCR, HV69/70del was detected in 48 samples. However, these results only matched NGS results in 16 samples (Figure 6). Usually, this deletion is only present in the Alpha variant but according to the reports of INSA[228] there are some sequences of the Delta variant that contain the respective deletion. As it did not affect the definition of the variant, I did not consider this discrepancy relevant. Nevertheless, the low coverage of some parts of the genome might have affected the identification of this mutation. It is interesting to note that it was only detected by NGS-ONT in one sample. These results are supported by the literature as this technology detects longer reads than NGS-SB which makes it difficult to sequence insertions/deletions[229].

With RT-PCR, L452R was detected in 42 samples. These results matched at least one of the NGS methods in 41 samples (Figure 6). In the only sample where a match was not obtained, this mutation could not be detected by NGS-ONT and the reason might be due to the bad quality of the sample and low coverage of that part of the genome.

With RT-PCR, P681R was detected in 14 samples. These results matched at least one of the NGS methods in 12 samples (Figure 6). In the two samples where a match was not obtained, this mutation could not be detected by NGS-ONT due to the same reason mentioned above.

Figure 6. Comparison of the mutations detected by the different methodologies.



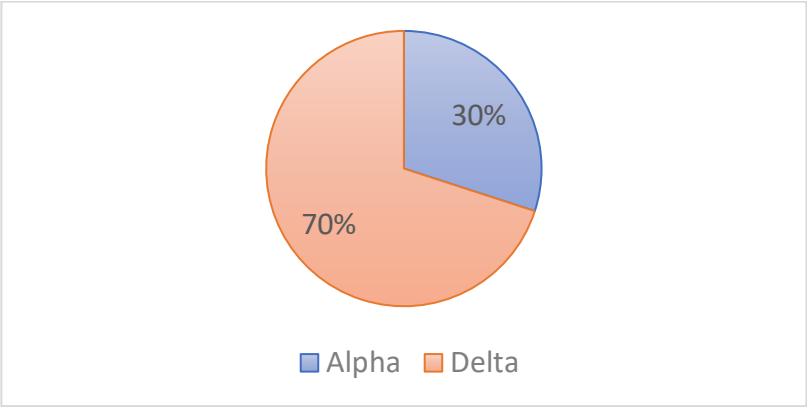
3.5. Comparison of the assigned variants

The variants assigned by RT-PCR were coincident with the variants attributed by at least one sequencing method in 100% of the samples (n=60) (Table S1).

When it was possible to obtain results with the three methodologies, the assigned variants were consistent between the three methods in 100% of the samples (n=20) (Table S1).

Based on RT-PCR and at least one of the NGS methods results, B.1.1.7 (Alpha variant) was assigned to 30% of the samples (n=18) and B.1.617.2 (Delta variant) was assigned to 70% (n=42) (Figure 7).

Figure 7. Graphic showing the percentage of samples belonging to the Alpha and Delta variants.



3.6. Validation of the assays

In the samples tested with Allplex™ SARS-CoV-2 Variants I Assay that were assigned to B.1.1.7 and were confirmed, subsequently, by at least one of the NGS methods (n=18), in 100% of them (n=18), it was detected N501Y and HV69/70del.

In the samples tested with Allplex™ SARS-CoV-2 Variants II Assay that were assigned to B.1.617.2 and were confirmed, subsequently, by at least one of the NGS methods (n=32), in 93.75% of them (n=30), it was detected L452R.

In the samples tested with Allplex™ SARS-CoV-2 Variants V Assay that were assigned to B.1.617.2 and were confirmed, subsequently, by at least one of the NGS methods (n=14), in 100% of them (n=14), both L452R and P681R were detected.

4. Conclusion

Based on RT-PCR and at least one of the NGS methods results, B.1.1.7 (Alpha variant) was assigned to 30% of the samples (n=18) and B.1.617.2 (Delta variant) was assigned to 70% (n=42). The samples that were assigned to the Alpha variant were collected between February and May, which corresponds to the period when the Alpha variant was the predominant lineage circulating in Portugal. The samples assigned to the Delta variant were collected between July and September, corresponding to the period when this variant was the predominant lineage circulating in Portugal.

In the samples tested with Allplex™ SARS-CoV-2 Variants I Assay that were assigned to B.1.1.7 and were confirmed, subsequently, by at least one of the NGS methods, in 100% of them (n=18), it was detected N501Y and HV69/70del. Since these are defining mutations of the Alpha variant, it confirms the hypothesis that this is a good assay to correctly identify this variant.

In the samples tested with Allplex™ SARS-CoV-2 Variants II Assay that were assigned to B.1.617.2 and were confirmed, subsequently, by at least one of the NGS methods, in 93.75% of them (n=30), it was detected L452R. Since this is a defining mutation of the Delta variant, it helps to validate the hypothesis that this is a good assay to correctly identify this variant. The purpose of this assay was also to identify the presence of K417N and K417T, defining mutations of the Beta and Gamma variants, respectively. However, none of the samples tested in this experiment belonged to these variants, so it could not be assessed the efficacy of this assay to identify these lineages. This assay can also potentially identify the California variant (B.1.427/B.1.429), as it detects both L452R and W152C. However, we could not detect any of these mutations in this study so the correct definition of this variant by this assay could not be evaluated.

In the samples tested with Allplex™ SARS-CoV-2 Variants V Assay that were assigned to B.1.617.2 and were confirmed, subsequently, by at least one of the NGS methods, in 100% of them (n=14), both L452R and P681R were detected. Since these are defining mutations of the Delta variant, it confirms the hypothesis that this is a good assay to correctly identify this lineage. This assay can also potentially identify the Lambda variant (C.37), as it detects both L452Q and F490S. However, we could not detect any of these mutations in this study so the correct definition of this variant by this assay could not be evaluated.

Between Allplex™ SARS-CoV-2 Variants II Assay and Allplex™ SARS-CoV-2 Variants V Assay, the latter would be more accurate to identify the Delta variant, as it

can detect two of the mutations present in this lineage. Moreover, it was able to detect L452R in the 2 samples where Allplex™ SARS-CoV-2 Variants II could not.

Allplex™ SARS-CoV-2 Variants II Assay should be tested with more samples, especially samples that would potentially belong to the Beta and Gamma variants, so that it could be confirmed if it is a good assay to define these variants. To get even more robust results, this assay should be used together with Allplex™ SARS-CoV-2 Variants I Assay, as together they can detect three defining mutations of each variant (K417T+E484K+N501Y and K417N+E484K+N501Y).

In conclusion, these three assays have been shown to correctly define the Alpha and Delta variants. To evaluate the correctly definition of the remaining variants (Beta, Gamma, Lambda and California variants), more tests should be performed.

5. References

1. Hong, K. H. *et al.* Guidelines for Laboratory Diagnosis of Coronavirus Disease 2019 (COVID-19) in Korea. *Ann. Lab. Med.* **40**, 351–360 (2020).
2. Giovanetti, M. *et al.* Evolution patterns of SARS-CoV-2: Snapshot on its genome variants. *Biochem. Biophys. Res. Commun.* (2020) doi:10.1016/j.bbrc.2020.10.102.
3. Platto, S., Wang, Y., Zhou, J. & Carafoli, E. History of the COVID-19 pandemic: Origin, explosion, worldwide spreading. *Biochem. Biophys. Res. Commun.* (2020) doi:10.1016/j.bbrc.2020.10.087.
4. Mahase, E. Covid-19 : What new variants are emerging and how are they being investigated ? 18–19 (2021) doi:10.1136/bmj.n86.
5. S., A. Genetic Variants of SARS-CoV-2 — What Do They Mean ? **2**, 1–3 (2021).
6. Du, L. *et al.* The spike protein of SARS-CoV - A target for vaccine and therapeutic development. *Nat. Rev. Microbiol.* **7**, 226–236 (2009).
7. Chafekar, A. & Fielding, B. C. MERS-CoV: Understanding the latest human coronavirus threat. *Viruses* **10**, (2018).
8. Schultze, J. L. & Aschenbrenner, A. C. COVID-19 and the human innate immune system. *Cell* (2021) doi:10.1016/j.cell.2021.02.029.
9. Hoffmann, M., Arora, P., Mu, J., Po, S. & Hahn, A. S. SARS-CoV-2 variants B . 1 . 351 and P . 1 escape from neutralizing antibodies Article escape from neutralizing antibodies. 1–10 (2021) doi:10.1016/j.cell.2021.03.036.
10. Mansbach, R. A. *et al.* The SARS-CoV-2 Spike variant D614G favors an open conformational state. *Sci. Adv.* **7**, 1–10 (2021).

11. Zhang, M., Li, L., Luo, M. & Id, B. L. Genomic characterization and evolution of SARS-CoV-2 of a Canadian population. 1–14 (2021) doi:10.1371/journal.pone.0247799.
12. Korsman, S. N. J., Zyl, G. U. van, Nutt, L., Andersson, M. I. & Preiser, W. *Virology*. (2012).
13. Alsharif, K. F. *et al.* The prevalence of MERS-CoV among military personnel and their families: A single center study. *J Med Virol.* 0–2 (2020) doi:10.1002/jmv.26642.
14. Calabrese, F. *et al.* Pulmonary pathology and COVID-19 : lessons from autopsy . The experience of European Pulmonary Pathologists. 359–372 (2020).
15. Giordo, R., Paliogiannis, P., Mangoni, A. A. & Pintus, G. SARS-CoV-2 and endothelial cell interaction in COVID-19: molecular perspectives. *Vasc. Biol.* **3**, R15–R23 (2021).
16. Quek, E. Treatment of COVID-19 : a review of current and prospective pharmacotherapies. 1–9 (2021).
17. De Ungria, M. C. A. Forensic DNA testing during the SARS-CoV-2 pandemic. *Forensic Sci. Int. Genet.* **48**, 102346 (2020).
18. Rani, S. A review of the management and safe handling of bodies in cases involving COVID-19. *Med. Sci. Law* (2020) doi:10.1177/0025802420949044.
19. Chirizzi, D. *et al.* SARS-CoV-2 concentrations and virus-laden aerosol size distributions in outdoor air in north and south of Italy. *Environ. Int.* **146**, 106255 (2020).
20. Cevik, M., Kuppalli, K., Kindrachuk, J. & Peiris, M. Virology , transmission , and pathogenesis of SARS-CoV-2. **2019**, 1–6 (2020).
21. Bueckert, M., Gupta, R., Gupta, A., Garg, M. & Mazumder, A. Infectivity of SARS-CoV-2 and Other Coronaviruses on Dry Surfaces : Potential for Indirect Transmission. **2**, 1–16 (2020).
22. Shih, K. C. *et al.* Tropism , replication competence , and innate immune responses of the coronavirus SARS-CoV-2 in human respiratory tract and conjunctiva : an analysis in ex-vivo and in-vitro cultures. *Lancet Respir.* **8**, 687–695 (2020).
23. Sheikhzadeh, E., Eissa, S., Ismail, A. & Zourob, M. Diagnostic techniques for COVID-19 and new developments. (2020).
24. Adachi, T. *et al.* Clinicopathologic and immunohistochemical findings from autopsy of patient with COVID-19, Japan. *Emerg. Infect. Dis.* **26**, 2157–2161 (2020).
25. Meyerowitz, E. A., Richterman, A., Bogoch, I. I., Low, N. & Cevik, M. Towards an

Accurate and Systematic Characterization of Persistently Asymptomatic Infection with SARS-CoV-2. 1–24 (2020).

26. Asadi-pooya; & Ali, A. Central nervous system manifestations of COVID-19: A systematic review. (2020).
27. Sessa, F. *et al.* Clinical-Forensic Autopsy Findings to Defeat COVID-19 Disease: A Literature Review. *J. Clin. Med.* **9**, 2026 (2020).
28. Qiu, X. *et al.* Defining the role of asymptomatic and pre-symptomatic SARS-CoV-2 transmission – a living systematic review 3. (2020).
29. Buitrago-Garcia, D. *et al.* Occurrence and transmission potential of asymptomatic and presymptomatic SARSCoV-2 infections: A living systematic review and meta-analysis. *PLoS Med.* **17**, 1–25 (2020).
30. Liu, Z., Chu, R., Gong, L., Su, B. & Wu, J. The assessment of transmission efficiency and latent infection period in asymptomatic carriers of SARS-CoV-2 infection. *Int. J. Infect. Dis.* **99**, 325–327 (2020).
31. Ashcroft, P. *et al.* COVID-19 infectivity profile correction. *Swiss Med. Wkly.* **150**, 3–5 (2020).
32. Falahi, S. & Kenarkoohi, A. COVID-19 re-infection: prolonged shedding or true reinfection. *New Microbes New Infect.* 100812 (2020) doi:10.1016/j.nmni.2020.100812.
33. Zhang, J., Wang, S. & Xue, Y. Fecal specimen diagnosis 2019 novel coronavirus – infected pneumonia. 680–682 (2020) doi:10.1002/jmv.25742.
34. Deinhardt-Emmer, S. *et al.* Early postmortem mapping of SARS-CoV-2 RNA in patients with COVID-19 and the correlation with tissue damage. *Elife* **10**, 1–22 (2021).
35. Wölfel, R. *et al.* Virological assessment of hospitalized patients with COVID-2019. *Nature* **581**, 465–469 (2020).
36. Huang, C. *et al.* Culture-Based Virus Isolation To Evaluate Potential Infectivity of Clinical Specimens Tested for COVID-19. **58**, 1–8 (2020).
37. Iyer, A. S. *et al.* Persistence and decay of human antibody responses to the receptor binding domain of SARS-CoV-2 spike protein in COVID-19 patients. **0367**, 1–13 (2020).
38. Id, B. K. *et al.* Prevalence of IgG antibodies against SARS- CoV-2 among healthcare workers in a tertiary pediatric hospital in Poland. **2**, 1–11 (2021).
39. Li, K. *et al.* Dynamic changes in anti-SARS-CoV-2 antibodies during SARS-CoV-2 infection and recovery from COVID-19. (2020) doi:10.1038/s41467-020-19943-y.
40. Chen, X. *et al.* Disease severity dictates SARS-CoV-2-speci fi c neutralizing

antibody responses in COVID-19. (2020) doi:10.1038/s41392-020-00301-9.

41. Chen, J. & Lu, H. New challenges to fighting COVID-19: Virus variants, potential vaccines, and development of antivirals. 1–3 (2021) doi:10.5582/bst.2021.01092.
42. Sakelliadis, E. I., Katsos, K. D., Zouzia, E. I., Spiliopoulou, C. A. & Tsiodras, S. Impact of Covid-19 lockdown on characteristics of autopsy cases in Greece. Comparison between 2019 and 2020. *Forensic Sci. Int.* **313**, 19–22 (2020).
43. Hanley, B., Lucas, S. B., Youd, E., Swift, B. & Osborn, M. Autopsy in suspected COVID-19 cases. *J Clin Pathol* **73**, 239–242 (2020).
44. Baj, J., Ciesielka, M., Buszewicz, G., Maciejewski, R. & Budzyńska, B. COVID-19 in the autopsy room – requirements, safety, recommendations and pathological findings. *Forensic Sci. Med. Pathol.* (2020) doi:10.1007/s12024-020-00341-1.
45. Moretti, M. *et al.* The roles of medical examiners in the COVID-19 era: a comparison between the United States and Italy. *Forensic Sci. Med. Pathol.* (2021) doi:10.1007/s12024-021-00358-0.
46. Malik, Y. A. Properties of Coronavirus and SARS-CoV-2. *Malays. J. Pathol.* **42**, 3–11 (2020).
47. Elsoukkary, S. S. *et al.* Autopsy Findings in 32 Patients with COVID-19: A Single-Institution Experience. *Pathobiology* (2020) doi:10.1159/000511325.
48. Udugama, B. *et al.* Diagnosing COVID-19: The Disease and Tools for Detection. *ACS Nano* (2020) doi:10.1021/acsnano.0c02624.
49. Sekhawat, V., Green, A. & Mahadeva, U. COVID-19 autopsies: conclusions from international studies. *Diagnostic Histopathol.* 1–5 (2020) doi:10.1016/j.mpdhp.2020.11.008.
50. Pandey, P. & Agarwal, S. Lung Pathology in COVID-19: A Systematic Review. **2019**, (2020).
51. Chan, J. F.-W. *et al.* Improved Molecular Diagnosis of COVID-19 by the Novel, Highly Sensitive and Specific COVID-19-RdRp/Hel Real-Time Reverse Transcription-PCR Assay Validated In Vitro and with Clinical Specimens. *J Clin Microbiol* **58**, 10–20 (2020).
52. Wit, E. De, Doremalen, N. Van, Falzarano, D. & Munster, V. J. SARS and MERS : recent insights into emerging coronaviruses. *Nat. Publ. Gr.* (2016) doi:10.1038/nrmicro.2016.81.
53. Khan, K. A. & Cheung, P. Evaluation of the Sequence Variability within the PCR Primer/Probe Target Regions of the SARS-CoV-2 Genome. **10**, (2021).
54. Peacock, T. P., Randal, R. P.-, Hiscox, J. A. & Barclay, W. S. SARS-CoV-2 one

year on: evidence for ongoing viral adaptation. (2021) doi:10.1099/jgv.0.001584.

55. Petrovszki, D. *et al.* Penetration of the SARS-CoV-2 Spike Protein across the Blood – Brain Barrier , as Revealed by a Combination of a Human Cell Culture Model System and Optical Biosensing. (2022).
56. Pan, B. *et al.* Chinese herbal compounds against SARS-CoV-2: puerarin and quercetin impair the binding of viral S-protein to ACE2 receptor. *Comput. Struct. Biotechnol. J.* (2020) doi:10.1016/j.csbj.2020.11.010.
57. Shah, M., Ahmad, B., Choi, S. & Goo, H. Mutations in the SARS-CoV-2 spike RBD are responsible for stronger ACE2 binding and poor anti-SARS-CoV mAbs cross-neutralization. *Comput. Struct. Biotechnol. J.* **18**, 3402–3414 (2020).
58. Ackermann, M. *et al.* Pulmonary vascular endothelialitis, thrombosis, and angiogenesis in Covid-19. *N. Engl. J. Med.* **383**, 120–128 (2020).
59. Haas, P., Muralidharan, M., Krogan, N. J., Kaake, R. M. & Hu, R. Proteomic Approaches to Study SARS-CoV-2 Biology and COVID-19 Pathology. (2020) doi:10.1021/acs.jproteome.0c00764.
60. Damiani, S. *et al.* Pathological Post Mortem Findings in Lungs Infected With Sars-Cov 2. *J. Pathol.* (2020) doi:10.1002/path.5549.
61. Antonela, M., Rivero, M., Ramiro, H., Beatriz, V. & Patricia, V. Virtual screening of plant-derived compounds against SARS-CoV-2 viral proteins using computational tools. *Sci. Total Environ.* **781**, 146400 (2021).
62. Liu, C. *et al.* Reduced neutralization of SARS-CoV-2 B.1.617 by vaccine and convalescent serum. 1–17 (2021) doi:10.1016/j.cell.2021.06.020.
63. Marle, G. Van *et al.* Arterivirus discontinuous mRNA transcription is guided by base pairing between sense and antisense transcription-regulating sequences. (1999).
64. Pachetti, M. *et al.* Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant. *J. Transl. Med.* **18**, 1–9 (2020).
65. Acheson, N. H., Witt, K., Palumbo, M., Morris, L. & Soo, J. *FUNDAMENTALS OF MOLECULAR VIROLOGY*. (2011).
66. De Luca, C. *et al.* Evaluation of a fully closed real time PCR platform for the detection of SARS-CoV-2 in nasopharyngeal swabs: A pilot study. *J. Clin. Pathol.* 1–4 (2021) doi:10.1136/jclinpath-2021-207516.
67. Zhang, Q., Xiang, R., Huo, S., Zhou, Y. & Jiang, S. Molecular mechanism of interaction between SARS-CoV-2 and host cells and interventional therapy. *Signal Transduct. Target. Ther.* (2021) doi:10.1038/s41392-021-00653-w.
68. Colling, R. An overview of COVID-19 for diagnostic pathologists: clinicopathological correlation and diagnostic techniques. *Diagnostic Histopathol.*

26, 529–536 (2020).

69. Chu, D. K. W. *et al.* Molecular Diagnosis of a Novel Coronavirus (2019-nCoV) Causing an Outbreak of Pneumonia. **555**, 549–555 (2020).
70. PANGO lineages. Lineage Description List, accessed 20 June 2021. https://cov-lineages.org/lineage_description_list.html (2021).
71. PANGO lineages. Lineage B, accessed 20 June 2021. https://cov-lineages.org/lineages/lineage_B.html (2021).
72. Liu, J. *et al.* BNT162b2-elicited neutralization of B.1.617 and other SARS-CoV-2 variants. *Nature* (2021) doi:10.1038/s41586-021-03693-y.
73. Biswas, S. K. & Mudi, S. R. Spike protein D614G and RdRp P323L: the SARS-CoV-2 mutations associated with severity of COVID-19. 0–6 (2020).
74. Zhang, L. *et al.* The D614G mutation in the SARS-CoV-2 spike protein reduces S1 shedding and increases infectivity. *bioRxiv Prepr. Serv. Biol.* (2020) doi:10.1101/2020.06.12.148726.
75. Bal, A. *et al.* Two-step strategy for the identification of SARS-CoV-2 variant of concern 202012/01 and other variants with spike deletion H69–V70, France, August to December 2020. *Eurosurveillance* **26**, 1–5 (2021).
76. 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* **2**, (2020).
77. Bayarri-Olmos, R. *et al.* The SARS-CoV-2 Y453F mink variant displays a pronounced increase in ACE-2 affinity but does not challenge antibody neutralization. *J. Biol. Chem.* **296**, 100536 (2021).
78. Thomson, E. C. *et al.* Circulating SARS-CoV-2 spike N439K variants maintain fitness while evading antibody-mediated immunity. *Cell* **184**, 1171-1187.e20 (2021).
79. Greaney, A. J. *et al.* Comprehensive mapping of mutations in the SARS- CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal . 463–476 (2021) doi:10.1016/j.chom.2021.02.003.
80. Laiton-Donato, K. *et al.* Novel Highly Divergent SARS-CoV-2 Lineage With the Spike Substitutions L249S and E484K. *Front. Med.* **8**, 1–6 (2021).
81. Planas, D. *et al.* Sensitivity of infectious SARS-CoV-2 B.1.1.7 and B.1.351 variants to neutralizing antibodies. *Nat. Med.* **27**, 917–924 (2021).
82. Francisco, R. da S. *et al.* Pervasive transmission of E484K and emergence of VUI-NP13L with evidence of SARS-CoV-2 co-infection events by two different

lineages in Rio Grande do Sul, Brazil. *Virus Res.* **296**, (2021).

83. Resende, P. C. *et al.* A Potential SARS-CoV-2 Variant of Interest (VOI) Harboring Mutation E484K in the Spike Protein Was Identified within Lineage B.1.1.33 Circulating in Brazil. 1–7 (2021).
84. Zhao, Y. *et al.* A novel diagnostic test to screen SARS-CoV-2 variants containing E484K and N501Y mutations. *Emerg. Microbes Infect.* **10**, 994–997 (2021).
85. Ibba, G., Sau, R., Angioj, F., Abbondio, M. & Rubino, S. A straightforward molecular strategy to retrospectively investigate the spread of SARS-CoV-2 VOC202012/01 B.1.1.7 variant. **2**, (2021).
86. Li, M., Lou, F. & Fan, H. SARS-CoV-2 variants: a new challenge to convalescent serum and mRNA vaccine neutralization efficiency. *Signal Transduct. Target. Ther.* **6**, 1–3 (2021).
87. Xie, X. *et al.* Neutralization of SARS-CoV-2 spike 69/70 deletion, E484K and N501Y variants by BNT162b2 vaccine-elicited sera. *Nat. Med.* **27**, (2021).
88. Leung, K., Shum, M. H. H., Leung, G. M., Lam, T. T. Y. & Wu, J. T. Early transmissibility assessment of the N501Y mutant strains of SARS-CoV-2 in the United Kingdom, October to November 2020. *Eurosurveillance* **26**, (2020).
89. Volz, E. *et al.* Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. (2021) doi:10.1038/s41586-021-03470-x.
90. GISAID. UK reports new variant, termed VUI 202012/01, accessed 26 May 2021. https://www.gisaid.org/references/gisaid-in-the-news/uk-reports-new-variant-termed-vui-20201201/ (2021).
91. Davies, N. G. *et al.* Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. (2021) doi:10.1126/science.abg3055.
92. Wang, P. *et al.* Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. *Nature* **2021**, (2021).
93. Public Health England. Investigation of novel SARS-CoV-2 variant. Variant of Concern 202012/01 Technical briefing 4 Nomenclature of variants in the UK. *Investig. SARS-CoV-2 Var. concern Engl.* (2021).
94. CoV lineages. CoV lineages, acessed 15 September 2021. https://cov-lineages.org/lineage.html?lineage=B.1.1.7.
95. Giallonardo, F. Di *et al.* Emergence and Spread of SARS-CoV-2 Lineages B.1.1.7 and P.1 in Italy. (2021).
96. Yadav, P. D. *et al.* Imported SARS-CoV-2 V501Y.V2 variant (B.1.351) detected in travelers from South Africa and Tanzania to India. *Travel Med. Infect. Dis.* **41**, (2021).
97. CoV lineages. Lineage B.1.351, accessed 15 September 2021. https://cov-

lineages.org/lineage.html?lineage=B.1.351 (2021).

98. Wang, P., Casner, R. G., Shapiro, L. & Ho, D. D. Increased resistance of SARS-CoV-2 variant P.1 to antibody neutralization Increased resistance of SARS-CoV-2 variant P.1 to antibody neutralization. *Cell Host Microbe* **29**, 747-751.e4 (2021).
99. Soares, M. *et al.* Early detection of SARS-CoV-2 P.1 variant in Southern Brazil and reinfection of the same patient by P.2. **2**, 1–8 (2021).
100. Ferrareze, P. A. G. *et al.* E484K as an innovative phylogenetic event for viral evolution: Genomic analysis of the E484K spike mutation in SARS-CoV-2 lineages from Brazil. **93**, (2021).
101. CoV lineages. Lineage P.1, accessed 15 September 2021. https://cov-lineages.org/lineage.html?lineage=P.1 (2021).
102. INSA. Diversidade genética do novo coronavírus SARS-CoV-2 (COVID-19) em Portugal, accessed 30 June 2021. https://insaflu.insa.pt/covid19/ (2021).
103. CDC. SARS-CoV-2 Variant Classifications and Definitions, accessed 27 June 2021. https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html#Interest (2021).
104. Gower, C. *et al.* Effectiveness of Covid-19 Vaccines against the B.1.617.2 (Delta) Variant. 1–10 (2021) doi:10.1056/NEJMoa2108891.
105. Lustig, Y. *et al.* Neutralising capacity against Delta (B.1.617.2) and other variants of concern following Comirnaty (BNT162b2, BioNTech/Pfizer) vaccination in health care workers, Israel. *Eurosurveillance* **26**, 1–5 (2021).
106. PANGO lineages. Lineage B.1.617.2, accessed 14 December 2022. (2022) doi:https://cov-lineages.org/lineage.html?lineage=B.1.617.2.
107. CoV lineages. Lineage AY.1, accessed 16 September 2021. https://cov-lineages.org/lineage.html?lineage=AY.1 (2021).
108. outbreak.info. AY.1 Lineage Report, accessed 16 September 2021. https://outbreak.info/situation-reports?pango=AY.1 (2021).
109. PANGO lineages. B.1.1.529, accessed 27 January 2022. https://cov-lineages.org/global_report_B.1.1.529.html (2022).
110. Sun, Y., Lin, W., Dong, W. & Xu, J. Origin and evolutionary analysis of the SARS-CoV-2 Omicron variant. *J. Biosaf. Biosecurity* **4**, 33–37 (2022).
111. PANGO lineages. Lineage B.1.1.529, accessed 28 January 2022. https://cov-lineages.org/lineage.html?lineage=B.1.1.529 (2022).
112. Outbreak.info. B.1.1.529 Lineage Report, accessed 28 January 2022. https://outbreak.info/situation-reports?pango=B.1.1.529 (2022).
113. CoV lineages. Lineage A.23.1, accessed 15 September 2021. https://cov-

lineages.org/lineage.html?lineage=A.23.1 (2021).

114. Outbreak.info. A.23.1 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=A.23.1 (2021).
115. ECDC. SARS-CoV-2 variants of concern as of 24 June 2021, accessed 29 June 2021. https://www.ecdc.europa.eu/en/covid-19/variants-concern (2021).
116. CoV lineages. Lineage B.1.429, accessed 15 September 2021. https://cov-lineages.org/lineage.html?lineage=B.1.429 (2021).
117. Zhao, Q. *et al.* Neutralization of SARS-CoV-2 Variants B.1.429 and B.1.351. *J. Med. Virol.* **69**, 2016–2017 (2020).
118. Outbreak.info. B.1.429 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=B.1.429 (2021).
119. CoV lineages. Lineage B.1.525, accessed 15 September 2021. https://cov-lineages.org/lineage.html?lineage=B.1.525 (2021).
120. CoVariants. Shared mutations, accessed 20 June 2021. https://covariants.org/shared-mutations (2021).
121. PANGO lineages. Global report investigating novel coronavirus haplotypes, acessed 26 May 2021. https://cov-lineages.org/global_report_B.1.1.7.html (2021).
122. Outbreak.info. B.1.525 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=B.1.525 (2021).
123. Outbreak.info. B.1.526 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=B.1.526 (2021).
124. Padilla-Rojas, C. *et al.* Genomic analysis reveals a rapid spread and predominance of lambda (C.37) SARS-COV-2 lineage in Peru despite circulation of variants of concern. *J. Med. Virol.* (2021) doi:10.1002/jmv.27261.
125. CoV lineages. Lineage B.1.616, accessed 15 September 2021. https://cov-lineages.org/lineage.html?lineage=B.1.616 (2021).
126. Outbreak.info. B.1.616 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=B.1.616 (2021).
127. Outbreak.info. B.1.617 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=B.1.617 (2021).
128. Outbreak.info. B.1.617.1 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=B.1.617.1 (2021).
129. CoV lineages. Lineage B.1.617.3, accessed 15 September 2021. https://cov-lineages.org/lineage.html?lineage=B.1.617.3 (2021).
130. Outbreak.info. B.1.617.3 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=B.1.617.3 (2021).

131. CoV lineages. Lineage B.1.620, accessed 15 September 2021. https://cov-lineages.org/lineage.html?lineage=B.1.620 (2021).
132. Outbreak.info. B.1.620 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=B.1.620 (2021).
133. CoV lineages. Lineage B.1.621, accessed 15 September 2021. https://cov-lineages.org/lineage.html?lineage=B.1.621 (2021).
134. Outbreak.info. B.1.621 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=B.1.621 (2021).
135. CoV lineages. Lineage B.1.626, accessed 10 October 2021. https://cov-lineages.org/lineage.html?lineage=B.1.626 (2021).
136. Outbreak.info. B.1.626 Lineage Report, accessed 10 October 2021. https://outbreak.info/situation-reports?pango=B.1.626 (2021).
137. CoV lineages. Lineage P.2, accessed 15 September 2021. (2021).
138. Outbreak.info. P.2 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=P.2 (2021).
139. CoV lineages. Lineage P.3, accessed 15 September 2021. https://cov-lineages.org/lineage.html?lineage=P.3 (2021).
140. Outbreak.info. P.3 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=P.3 (2021).
141. Darvishi, M., Rahimi, F. & Abadi, A. T. B. SARS-CoV-2 Lambda (C.37): An emerging variant of concern? *Gene Reports* **25**, 101378 (2021).
142. CoV lineages. Lineage C.37, accessed 16 October 2021. https://cov-lineages.org/lineage.html?lineage=C.37 (2021).
143. Outbreak.info. C.37 Lineage Report, accessed 16 October 2021. https://outbreak.info/situation-reports?pango=C.37 (2021).
144. Outbreak.info. A.21 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=A.21 (2021).
145. CoV lineages. Lineage A.21, accessed 15 September 2021. https://cov-lineages.org/lineage.html?lineage=A.21 (2021).
146. CoV lineages. Lineage List, accessed 15 September 2021. https://cov-lineages.org/lineage_list.html (2021).
147. Outbreak.info. B.1 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=B.1 (2021).
148. CoV lineages. Lineage B.1.1.216, accessed 15 September 2021. https://cov-lineages.org/lineage.html?lineage=B.1.1.216 (2021).
149. Outbreak.info. B.1.1.216 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=B.1.1.216 (2021).

150. CoV lineages. Lineage B.1.1.160, accessed 15 September 2021. https://cov-lineages.org/lineage.html?lineage=B.1.1.160 (2021).
151. Outbreak.info. B.1.1.160 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=B.1.1.160 (2021).
152. CoV lineages. Lineage B.1.177, accessed 15 September 2021. https://cov-lineages.org/lineage.html?lineage=B.1.177 (2021).
153. Outbreak.info. B.1.177 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=B.1.177 (2021).
154. CoV lineages. Lineage B.1.177.32, accessed 15 September 2021. https://cov-lineages.org/lineage.html?lineage=B.1.177.32 (2021).
155. Outbreak.info. B.1.177.32 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=B.1.177.32 (2021).
156. CoV lineages. Lineage B.1.177.52, accessed 15 September 2021. https://cov-lineages.org/lineage.html?lineage=B.1.177.52 (2021).
157. Outbreak.info. B.1.177.52 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=B.1.177.52 (2021).
158. Outbreak.info. B.1.177.72 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=B.1.177.72 (2021).
159. CoV lineages. Lineage B.1.177.85, accessed 15 September 2021. https://cov-lineages.org/lineage.html?lineage=B.1.177.85 (2021).
160. Outbreak.info. B.1.177.85 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=B.1.177.85 (2021).
161. CoV lineages. Lineage B.1.214, accessed 15 September 2021.
162. Outbreak.info. B.1.214 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=B.1.214 (2021).
163. CoV lineages. Lineage B.1.258, accessed 15 September 2021. https://cov-lineages.org/lineage.html?lineage=B.1.258 (2021).
164. Outbreak.info. B.1.258 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=B.1.258 (2021).
165. CoV lineages. Lineage B.1.258.17, accessed 15 September 2021. https://cov-lineages.org/lineage.html?lineage=B.1.258.17 (2021).
166. Outbreak.info. B.1.258.17 Lineage Report, accessed 15 September 2021. (2021).
167. CoV lineages. Lineage B.1.258.20, accessed 15 September 2021. https://cov-lineages.org/lineage.html?lineage=B.1.258.20 (2021).
168. Outbreak.info. B.1.258.20 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=B.1.258.20 (2021).

169. CoV lineages. Lineage B.1.575, accessed 15 September 2021. https://cov-lineages.org/lineage.html?lineage=B.1.575 (2021).
170. Outbreak.info. B.1.575 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=B.1.575 (2021).
171. Outbreak.info. C.16 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=C.16 (2021).
172. CoV lineages. Lineage C.36, accessed 15 September 2021. (2021).
173. Outbreak.info. C.36 Lineage Report, accessed 15 September 2021. https://outbreak.info/situation-reports?pango=C.36 (2021).
174. CoV lineages. Lineage B.1.160, accessed 30 October 2021. https://cov-lineages.org/lineage.html?lineage=B.1.160 (2021).
175. CoV lineages. Lineage B.1.367, accessed 30 October 2021. https://cov-lineages.org/lineage.html?lineage=B.1.367 (2021).
176. Outbreak.info. B.1.367 Lineage Report, accessed 30 October 2021. https://outbreak.info/situation-reports?pango=B.1.367 (2021).
177. CoV lineages. Lineage B.1.177.58, accessed 30 October 2021. https://cov-lineages.org/lineage.html?lineage=B.1.177.58 (2021).
178. Outbreak.info. B.1.177.58 Lineage Report, accessed 30 October 2021. https://outbreak.info/situation-reports?pango=B.1.177.58 (2021).
179. INSA. Diversidade genética do novo coronavírus SARS-CoV-2 (COVID-19) em Portugal, accessed 28 September 2021. https://insaflu.insa.pt/covid19/relatorios/INSA_SARS_CoV_2_DIVERSIDADE_GENETICA_relatorio_situacao_2020-08-21.pdf (2020).
180. He, X. *et al.* Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat. Med.* **26**, 672–675 (2020).
181. Shen, M. *et al.* Recent advances and perspectives of nucleic acid detection for coronavirus. *J. Pharm. Anal.* **10**, 97–101 (2020).
182. QIAGEN. What is the threshold cycle or Ct value? https://www.qiagen.com/us/resources/faq?id=783d4566-9ad9-4fab-9936-182beda65617&lang=en (2021).
183. Sethuraman, N., Stanleyraj, S., City, Y., Ryo, A. & City, Y. Interpreting Diagnostic Tests for SARS-CoV-2. **2019**, 2019–2021 (2020).
184. Kyosei, Y. *et al.* Antigen tests for COVID-19. *Biophys. Physicobiology* **18**, 28–39 (2021).
185. Yamayoshi, S. *et al.* Comparison of rapid antigen tests for covid-19. *Viruses* **12**, 1–8 (2020).
186. Gilsenan-Reed, C., Higgins, G. & Langlois, N. Determining a sampling regime

for PCR detection of respiratory tract viral infection at coronial post-mortem examinations. *Forensic Sci. Med. Pathol.* **16**, 457–462 (2020).

187. von Stillfried, S. & Boor, P. Detection methods for SARS-CoV-2 in tissue. *Pathologe* (2021) doi:10.1007/s00292-021-00920-1.
188. Carter, L. J. *et al.* Assay Techniques and Test Development for COVID-19 Diagnosis. (2020) doi:10.1021/acscentsci.0c00501.
189. Michel, M. *et al.* Evaluating ELISA , Immunofluorescence , and Lateral Flow Assay for SARS-CoV-2 Serologic Assays. **11**, 1–8 (2020).
190. Lisboa Bastos, M. *et al.* Diagnostic accuracy of serological tests for covid-19: Systematic review and meta-analysis. *BMJ* **370**, (2020).
191. Osborn, M. J. *et al.* CRISPR/Cas9-Based Lateral Flow and Fluorescence Diagnostics. 1–20 (2021).
192. Bogdanović, M. *et al.* Is the role of forensic medicine in the covid-19 pandemic underestimated? *Forensic Sci. Med. Pathol.* (2020) doi:10.1007/s12024-020-00308-2.
193. Edler, C. *et al.* Dying with SARS-CoV-2 infection—an autopsy study of the first consecutive 80 cases in Hamburg, Germany. *Int. J. Legal Med.* **134**, 1977 (2020).
194. Pomara, C. *et al.* Safe management strategies in clinical forensic autopsies of confirmed COVID-19 cases. *Diagnostics* **11**, (2021).
195. Evert, K. *et al.* Autopsy findings after long-term treatment of COVID-19 patients with microbiological correlation. (2021).
196. Pomara, C., Li Volti, G. & Cappello, F. COVID-19 Deaths: Are We Sure It Is Pneumonia? Please, Autopsy, Autopsy, Autopsy! *J. Clin. Med.* **9**, 1259 (2020).
197. Skok, K. *et al.* COVID-19 autopsies: Procedure , technical aspects and cause of fatal course. Experiences from a single-center. *Pathol. - Res. Pract.* **217**, 153305 (2021).
198. Dijkhuizen, L. G. M., Gelderman, H. T. & Duijst, W. L. J. M. Review: The safe handling of a corpse (suspected) with COVID-19. *J. Forensic Leg. Med.* **73**, 101999 (2020).
199. Yaacoub, S. *et al.* Safe management of bodies of deceased persons with suspected or confirmed COVID-19: A rapid systematic review. *BMJ Glob. Heal.* **5**, 1–10 (2020).
200. Hasmi, A. H. *et al.* The craniotomy box: an innovative method of containing hazardous aerosols generated during skull saw use in autopsy on a COVID-19 body. *Forensic Sci. Med. Pathol.* **16**, 477–480 (2020).
201. Santurro, A., Scopetti, M., D'Errico, S. & Fineschi, V. A technical report from the

Italian SARS-CoV-2 outbreak. Postmortem sampling and autopsy investigation in cases of suspected or probable COVID-19. *Forensic Sci. Med. Pathol.* **16**, 471–476 (2020).

202. Aquila, I. *et al.* Analysis of the persistence time of the SARS-CoV-2 virus in the cadaver and the risk of passing infection to autopsy staff. (2021) doi:10.1177/0025817220980601.
203. Beltempo, P., Maria, S., Maserati, R., Gherardi, M. & Castelli, M. Persistence of SARS-CoV-2 RNA in post-mortem swab 35 days after death: A case report. *Forensic Sci. Int.* **319**, 110653 (2021).
204. Keten, D., Okdemir, E. & Keten, A. Precautions in postmortem examinations in Covid-19 - Related deaths: Recommendations from Germany. *J. Forensic Leg. Med.* **73**, 102000 (2020).
205. Ducloyer, M. *et al.* Complete post-mortem data in a fatal case of COVID-19: clinical, radiological and pathological correlations. *Int. J. Legal Med.* **2**, 2209–2214 (2020).
206. Attoh, S. A. *et al.* Postmortem diagnosis of COVID-19: Antemortem challenges of three cases at the 37 Military Hospital, Accra, Ghana. *Afr. J. Lab. Med.* **9**, 1–8 (2020).
207. Brook, O. R. *et al.* Feasibility and safety of ultrasound-guided minimally invasive autopsy in COVID-19 patients. *Abdom. Radiol.* (2020) doi:10.1007/s00261-020-02753-7.
208. Ozmen, S. A. & Demirag, F. The Review of Histopathological Pulmonary Findings of COVID-19 : What We Learned from Postmortem Biopsy and Autopsies ; Beyond the Horizon. **52**, 307–308 (2020).
209. Wichmann, D. *et al.* Autopsy Findings and Venous Thromboembolism in Patients With COVID-19. **25**, (2020).
210. Keresztesi, A., Perde, F., Ghita-nanu, A. & Radu, C. Post-Mortem Diagnosis and Autopsy Findings in SARS-CoV-2 Infection: Forensic Case Series. (2021).
211. Bruce-Brand, C. *et al.* Postmortem lung biopsies from four patients with COVID-19 at a tertiary hospital in Cape Town, South Africa. **110**, 1195–1200 (2020).
212. Trougakos, I. P., Stamatelopoulos, K., Terpos, E., Tsitsilonis, O. E. & Aivalioti, E. Insights to SARS-CoV-2 life cycle, pathophysiology, and rationalized treatments that target COVID-19 clinical complications. *J. Biomed. Sci.* 1–18 (2021) doi:10.1186/s12929-020-00703-5.
213. Griffin, K. J. Autopsy in the time of COVID. *Diagnostic Histopathol.* 10–13 (2021) doi:10.1016/j.mpdhp.2020.12.002.
214. Harris, C. K., Hung, Y. P., Nielsen, G. P., Stone, J. R. & Ferry, J. A. Bone

Marrow and Peripheral Blood Findings in Patients Infected by SARS-CoV-2. 1–11 (2021) doi:10.1093/AJCP/AQAA274.

215. Campbell, R. A., Boilard, E. & Rondina, M. T. Is there a role for the ACE2 receptor in SARS-CoV-2 interactions with platelets? *J. Thromb. Haemost.* (2020) doi:10.1111/jth.15156.
216. Tombolini, A. & Scendoni, R. SARS-CoV-2-related deaths in routine forensic autopsy practice: histopathological patterns. *Int. J. Legal Med.* 2205–2208 (2020) doi:10.1007/s00414-020-02354-5.
217. Burkhard-koren, N. M. *et al.* Higher prevalence of pulmonary macrothrombi in SARS-CoV-2 than in influenza A: autopsy results from 'Spanish flu' 1918/1919 in Switzerland to Coronavirus disease 2019. (2020) doi:10.1002/cjp2.189.
218. Meng, Q. F. *et al.* Capturing Cytokines with Advanced Materials: A Potential Strategy to Tackle COVID-19 Cytokine Storm. *Adv. Mater.* **33**, 1–12 (2021).
219. Zahedipour, F. *et al.* Potential effects of curcumin in the treatment of COVID-19 infection. 2911–2920 (2020) doi:10.1002/ptr.6738.
220. Xu, G. *et al.* Persistent viral activity , cytokine storm , and lung fibrosis in a case of severe COVID-19. 1–5 (2020) doi:10.1002/ctm2.224.
221. Mullaguri, N. *et al.* COVID-19 Disease and Hypercoagulability Leading to Acute Ischemic Stroke. **11**, 131–136 (2021).
222. Hooper, J. E., Padera, R. F., Dolhnikoff, M. & Silva, L. F. F. A Postmortem Portrait of the Coronavirus Disease 2019 (COVID-19) Pandemic: A Large Multi-institutional Autopsy Survey Study. (2021) doi:10.5858/arpa.
223. Malézieux-Picard, A. *et al.* Undetected Causes of Death in Hospitalized Elderly with COVID-19: Lessons from Autopsy. *J. Clin. Med.* **10**, 1337 (2021).
224. Casagrande, M. *et al.* Presence of SARS-CoV-2 RNA in the Cornea of Viremic Patients With COVID-19. 1–6 (2021) doi:10.1001/jamaophthalmol.2020.6339.
225. Alexandersen, S., Chamings, A. & Bhatta, T. R. SARS-CoV-2 genomic and subgenomic RNAs in diagnostic samples are not an indicator of active replication. *Nat. Commun.* **11**, (2020).
226. Bharagava, R. N. & Mulla, S. I. Illumina Sequencing. https://www.sciencedirect.com/topics/immunology-and-microbiology/illumina-dye-sequencing (2019).
227. Technologies, O. N. Nanopore DNA Sequencing. https://nanoporetech.com/applications/dna-nanopore-sequencing#:~:text=Nanopore sequencing is a unique,specific DNA or RNA sequence (2022).
228. INSA. Diversidade genética do novo coronavírus SARS-CoV-2 (COVID-19) em

Portugal, accessed 10 October 2021. https://insaflu.insa.pt/covid19/ (2021).

229. Clara, D. & Nicolas, J. Sequencing DNA with nanopores: Troubles and biases. (2021) doi:10.1371/journal.pone.0257521.

6. Supplementary tables

Table S1. Comparison of the mutations N501Y, HV69/70del, L452R and P681R detected by RT-PCR and NGS in the 60 samples. In the first columns of the table, it shows the assigned variants to each sample. In the second part of the table, it shows the mutations detected in each sample by the different methods.

Samples	**Variants**			**c.204_209delACATGT(p.H69_V70del)**			**c.1355T>G (p.L452R)**			**c.1501A>T (p.N501Y)**			**c.2042C>G (p.P681R)**		
	RT-PCR	NGS-ONT	NGS-SB	RT-PCR	NGS-ONT	NGS-SB	RT-PCR	NGS-ONT	NGS-SB	RT-PCR	NGS-ONT	NGS-SB	RT-PCR	NGS-ONT	NGS-SB
CV6546	B.1.1.7	NR	B.1.1.7	X		X				X		X			
CV6561	B.1.1.7	NR	B.1.1.7	X		X				X		X			
CV6919	B.1.1.7	B.1.1.7	NR	X						X	X				
CV6595	B.1.1.7	NR	B.1.1.7	X		X				X		X			
CV6900	B.1.1.7	NR	B.1.1.7	X		X				X		X			
CV6609	B.1.1.7	NR	B.1.1.7	X		X				X		X			
CF1688	B.1.1.7	B.1.1.7	B.1.1.7	X		X				X		X			
LX271	B.1.1.7	B.1.1.7	B.1.1.7	X		X				X	X	X			
LX181	B.1.1.7	B.1.1.7	B.1.1.7	X	X					X	X	X			
LX225	B.1.1.7	NR	B.1.1.7	X		X				X		X			
LX320	B.1.1.7	NR	B.1.1.7	X		X				X		X			
LX209	B.1.1.7	NR	B.1.1.7	X		X				X		X			
CC73	B.1.1.7	NR	B.1.1.7	X		X				X	X	X			
LX525	B.1.1.7	NR	B.1.1.7	X		X				X		X			
BRR411	B.1.1.7	B.1.1.7	B.1.1.7	X		X				X		X			
LX145	B.1.1.7	NR	B.1.1.7	X						X		X			
LX516	B.1.1.7	B.1.1.7	B.1.1.7	X		X				X	X	X			
LX508	B.1.1.7	NR	B.1.1.7	X		X				X		X			
408	B.1.617.2	NR	B.1.617.2	X			X		X						X
LX1035	B.1.617.2	B.1.617.2	B.1.617.2	X			X	X	X					X	X
LX1039	B.1.617.2	B.1.617.2	NR	X			X	X						X	
LX1119	B.1.617.2	B.1.617.2	B.1.617.2	X			X	X	X						X
CC212	B.1.617.2	B.1.617.2	B.1.617.2	X			X	X	X					X	X

Samples	Variants			c.204_209delACATGT(p.H69_V70del)			c.1355T>G (p.L452R)			c.1501A>T (p.N501Y)			c.2042C>G (p.P681R)		
LX1144	B.1.617.2	B.1.617.2	B.1.617.2	X			X	X	X					X	X
LX1155	B.1.617.2	B.1.617.2	NR	X			X	X						X	
CF1824	B.1.617.2	B.1.617.2	B.1.617.2	X			X	X	X				X	X	X
CF1830	B.1.617.2	B.1.617.2	NR	X			X	X					X	X	
CV8057	B.1.617.2	NR	B.1.617.2	X			X		X						X
CV8064	B.1.617.2	NR	B.1.617.2	X			X		X						X
CV8068	B.1.617.2	B.1.617.2	B.1.617.2				X		X						X
CV8072	B.1.617.2	NR	B.1.617.2				X		X						X
LX2992	B.1.617.2	B.1.617.2	NR				X	X						X	
LX3009	B.1.617.2	B.1.617.2	B.1.617.2				X	X	X					X	X
LX3010	B.1.617.2	NR	B.1.617.2				X		X						X
LX2993	B.1.617.2	B.1.617.2	NR				X	X					X	X	
CV9713	B.1.617.2	NR	B.1.617.2	X			X		X						X
CV9705	B.1.617.2	NR	B.1.617.2				X		X						X
CV9704	B.1.617.2	B.1.617.2	NR				X	X						X	
LX3019	B.1.617.2	NR	B.1.617.2				X		X						X
CF1980	B.1.617.2	NR	B.1.617.2				X		X						X
CV9738	B.1.617.2	NR	B.1.617.2	X			X		X						X
LGM179	B.1.617.2	B.1.617.2	B.1.617.2	X			X	X	X					X	X
LX3028	B.1.617.2	B.1.617.2	NR				X	X					X		
LX3021	B.1.617.2	B.1.617.2	B.1.617.2	X			X		X						X
CV9575	B.1.617.2	B.1.617.2	NR	X			X	X						X	
LX3034	B.1.617.2	NR	B.1.617.2	X			X		X						X
CDA2237	B.1.617.2	NR	B.1.617.2	X			X		X						X
LX3147	B.1.617.2	NR	B.1.617.2				X		X						X
LX3152	B.1.617.2	NR	B.1.617.2	X			X		X						X
CF2012	B.1.617.2	NR	B.1.617.2	X			X		X						X
LX3155	B.1.617.2	B.1.617.2	B.1.617.2	X			X	X	X				X	X	X
CF2018	B.1.617.2	B.1.617.2	B.1.617.2	X			X	X	X				X	X	X
CV10172	B.1.617.2	B.1.617.2	NR	X			X						X		
CV10191	B.1.617.2	B.1.617.2	B.1.617.2	X			X	X	X				X	X	X
CV10192	B.1.617.2	B.1.617.2	B.1.617.2	X			X	X	X				X	X	X
LX3162	B.1.617.2	B.1.617.2	B.1.617.2	X			X	X	X				X	X	X

Samples	Variants			c.204_209delACATGT(p.H69_V70del)			c.1355T>G (p.L452R)		c.1501A>T (p.N501Y)				c.2042C>G (p.P681R)		
LX3166	B.1.617.2	NR	B.1.617.2	X			X		X				X		X
CV10205	B.1.617.2	B.1.617.2	NR	X			X	X					X	X	
CF2025	B.1.617.2	B.1.617.2	B.1.617.2	X			X	X	X				X	X	X
326	B.1.617.2	B.1.617.2	NR	X			X	X					X	X	

Table S2. List of symbols and respective amino acids.

Symbol	Amino acid
A	Alanine
C	Cysteine
D	Aspartic acid
E	Glutamic acid
F	Phenylalanine
G	Glycine
H	Histidine
I	Isoleucine
K	Lysine
L	Leucine
M	Methionine
N	Asparagine
P	Proline
Q	Glutamine
R	Arginine
S	Serine
T	Threonine
V	Valine
W	Tryptophan
Y	Tyrosine