



UNIVERSITÀ
DEGLI STUDI
FIRENZE

DOTTORATO DI RICERCA IN
SCIENZE CLINICHE

CICLO XXXVII

COORDINATORE Prof. Gian Maria Rossolini

COVID-19 and host genetic variability

Settore Scientifico Disciplinare MED/46

Dottorando

Dott. Capezzuoli Tommaso

Tutore

Prof. Gori Anna Maria

Coordinatore

Prof. Gian Maria Rossolini

1. INTRODUCTION	1
1.1 Covid-19 Study	1
NGS Study	1
Kringle IV Type 2 (KIV2) Repeats Evaluation.....	2
1.2. Vaccine-Induced Thrombotic Thrombocytopenia (VITT) Study	2
2. SARS-COV-2	3
2.1. Epidemiology	3
2.2. Clinical Aspects.....	5
Effects of Covid-19 on Cardiac Function	5
Effects of Covid-19 on Coagulation: clinical features and pathophysiology.....	6
Covid-19 and Lipoprotein (a).....	7
2.3. Treatments	11
Symptomatic Drugs.....	11
Drugs to be used only at Specific Stages of the Disease.....	11
2.4. Vaccines	13
mRNA Vaccines	13
Viral-Vector Vaccines.....	13
Protein-Subunit Vaccines	13
Inactivated-Virus Vaccines.....	14
Vaccine-Induced Thrombotic Thrombocytopenia (VITT)	14
2.5. Virus Classification and Taxonomy	16
2.6. Virus Structure	18
Spike Glicoprotein (S).....	18
The Membrane Glycoprotein (M).....	19
The Envelope Glycoprotein (E)	19
Nucleocapsid Protein (N).....	19
Haemagglutinin Esterase (HE).....	20
The Viral Genome	20

SARS-CoV-2 Variability	22
SARS-CoV-2 Life Cycle	24
3. GENETIC SUSCEPTIBILITY FACTORS: WHAT WE KNOW SO FAR	26
3.1. Genetic Susceptibility Factors: Virus Entry	26
ACE2.....	26
IL6.....	28
TMPRSS2	29
FURIN.....	30
HSPA5.....	31
CTSB and CTSL	31
3.2. Genetic Susceptibility Factors: Immune components, epigenetics and AB0 loci	32
4. MATERIALS AND METODHS.....	35
4.1 DNA manual extraction.....	35
Qiagen FlexiGene® DNA Kit.....	35
Qualitative and quantitative evaluation using Nanodrop	36
DNA quantification with Qubit® fluorometer	37
4.2. Gene panel design	38
4.3. Haloplex Hs Target Enrichment System	39
Library preparation.....	39
4.4. MiSeq Illumina Protocol	54
Dilution and denaturation of libraries for sequencing on MiSeq Illumina platform.....	54
Sequencing procedure on MiSeq Illumina platform	54
Bioinformatic analysis of MiSeq sequencing results	56
4.5. Kringle IV Type 2 (KIV2) Repeats Evaluation.....	57
4.6. SureSelectXT HS Target Enrichment System	59
Library preparation.....	59
4.7. NextSeq 500 Illumina protocol	75
Dilution and denaturation of libraries for sequencing on NextSeq Illumina Platform.....	75

Sequencing process on Illumina NextSeq 500	76
Bioinformatic analysis of NextSeq 500 sequencing results	78
5. RESULTS.....	79
5.1. Covid-19 Study	79
Study Population	79
Targeted Sequencing Results	80
Lp(a) and Covid-19	87
5.2. VITT Study	89
Study Population	89
WES Results	90
6. DISCUSSION	95
6.1. Covid-19 Study	95
Lp(a) and Covid-19	100
6.2. VITT Study	102
7. BIBLIOGRAFIA	108

1. INTRODUCTION

In this thesis, an attempt was made to study genetic profiles using high-throughput sequencing technologies (HTS) to identify potential susceptibility factors for COVID-19 (COronaVirus Disease 2019) and Vaccine-Induced Thrombotic Thrombocytopenia (VITT), a rare thrombotic condition associated with adenoviral-based vaccination against SARS-CoV-2. The use of HTS is a useful tool for determining the genetic background of individuals, and the information acquired at multiple loci could help define categories of individuals characterised by greater or lesser susceptibility to SARS-CoV-2 infection and VITT syndrome.

1.1 Covid-19 Study

NGS Study

The COVID-19 pandemic had an unprecedented global impact, making an in-depth understanding of the genetic dynamics of the SARS-CoV-2 virus and the human organism's responses to infection crucial. Genetic analysis of COVID-19 has proved crucial in identifying viral variants, monitoring their spread and understanding the mechanisms of individual resistance or susceptibility to the disease. Over 4000 variants have been identified in the SARS-CoV-2 genome¹, but none of which have been shown to have significant functional implications. Some mutations in the viral genome appear to increase the transmissibility and immune escape of the virus, but not the severity and pathogenicity of COVID-19. Thus, disease severity appears to be associated with host factors, rather than genetic variability of the virus².

This study aims to investigate the applications of genomic sequencing techniques to investigate host genetic predispositions. With this aim in mind, in our laboratory we analysed the genetic variability at the level of 7 genes (*ACE2*, *TMPRSS2*, *CTSL*, *CTSB*, *HSPA5*, *IL6*, *FURIN*) selected on the basis of their involvement in virus entry into the host cell and their association in the literature with COVID-19, in 80 patients with SARS-CoV-2 infection admitted to the Careggi University Hospital, in order to identify possible genetic predisposition profiles by means of High-Throughput Sequencing (HTS) technology.

A population of 80 subjects affected by SARS-CoV-2 infection: n=40 subjects with a more favourable clinical course admitted to the ordinary ward (non-ICU patients) and n=40 patients with a severe clinical course admitted to the intensive care unit (ICU/death patients) were investigated.

Kringle IV Type 2 (KIV2) Repeats Evaluation

Conflicting opinions in the literature³⁻⁵ on the possible role of the atherothrombotic risk factor lipoprotein (a) [Lp(a)] exist. Since more than 90% of the variance in Lp(a) levels is genetically determined, with the KIV2 repeat polymorphism being the main determinant^{6,7}, we decided to evaluate this CNV in our cohort using the ddPCR technique⁸ with the aim of identifying a possible relationship between Lp(a) levels and Covid-19 severity.

1.2. Vaccine-Induced Thrombotic Thrombocytopenia (VITT) Study

The widespread availability of COVID-19 vaccines has significantly changed the pandemic's trajectory. However, in February 2021, a prothrombotic syndrome, called Vaccine-Induced Thrombotic Thrombocytopenia (VITT), was observed in a small number of individuals who received adenoviral vector-based vaccines against SARS-CoV-2, specifically Vaxzevria (AstraZeneca) and Janssen (Johnson & Johnson). The condition primarily affected women under 55 years, particularly after the Vaxzevria and Janssen vaccines, with an incidence of about 1.02 cases per 100,000 doses for those over 60, rising to approximately 1 in 50,000-75,000 doses for younger individuals.⁸⁻¹⁰ VITT is characterized by the combination of thrombosis and thrombocytopenia, a phenomenon that is quite rare in medicine. It is particularly found in heparin-induced thrombocytopenia (HIT) in its autoimmune form (aHIT). In fact, the pathophysiological model of VITT is similar to that of aHIT, where platelet factor 4 (PF4) binds to negatively charged molecules on the endothelial surface through their FcγRIIA receptors. In the case of VITT, however, PF4 likely binds to components of the viral vector or the DNA contained within it.^{11,12}

The aim of this part of the PhD work was to analyze the genetic background of VITT subjects, focusing on rare variants in genes associated with platelet mechanisms, using a high-throughput whole exome sequencing approach. This investigation targeted six female patients of Caucasian origin who received either the VAXZEVRIA vaccine (AstraZeneca) or the JANSSEN vaccine (Johnson & Johnson) to explore potential genetic predispositions to VITT. The results of this thesis were published this year by our group¹³.

2. SARS-COV-2

2.1. Epidemiology

The emergence of the virus dates back to December 2019, when in Wuhan, China, a large number of people began presenting symptoms such as fever, malaise, dry cough and dyspnoea. The disease, initially diagnosed as viral pneumonia^{14,15}, was called ‘Wuhan pneumonia’ by the press, due to the geographical area affected by the symptoms. Later, whole-genome sequencing results showed that the aetiological agent was a member of the coronavirus family, specifically, the seventh member of the family capable of infecting humans¹⁶. On 12 January 2020, the World Health Organisation (WHO) temporarily named the new virus ‘2019 novel coronavirus’ (2019-nCoV) and on 12 February, officially catalogued the disease as ‘coronavirus disease 2019 (COVID-19)’. Subsequently, based on phylogeny and taxonomy studies, the pathogen was given the official name, SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2)¹⁷. The virus initially spread to China and several Asian countries before reaching Iran and Italy. Subsequently, despite initial attempts at containment, SARS-CoV-2 arrived throughout Europe and America in early 2020. Thus, on 11 March 2020, the WHO was forced to categorise COVID-19 as a pandemic disease, following the 1918 Spanish flu, the 1957 Asian flu, the 1968 Hong Kong flu and the 2009 Pandemic flu, which caused 50 million, 1.5 million, 1 million, 300 000 deaths respectively. Within a short time, SARS-CoV-2 reached virtually every corner of the World, infecting more than 140 million people and killing more than 3 million. Vaccines, as well as improving the prognosis of patients, has significantly lowered the number of infections. However, to date (at 10/17/2024), it has affected 776,471,644 people, including 7,068,677 deaths¹⁸. It should be noted that not all cases, especially asymptomatic ones, have been diagnosed and that, therefore, the number of infected is probably much higher. Italy was the first European nation to be affected by the pandemic. Although the first official case was diagnosed on 20 February, the potency of the outbreak suggests that the virus had already been circulating in Italy for weeks, even before 1 January¹⁹. It is difficult, if not impossible, to ascertain exactly when the virus began to circulate: during the influenza season many deaths from COVID-19 could easily have been mistaken for deaths from influenza and, among young people, the rapid spread could have gone unnoticed given the small number of symptoms in this category.

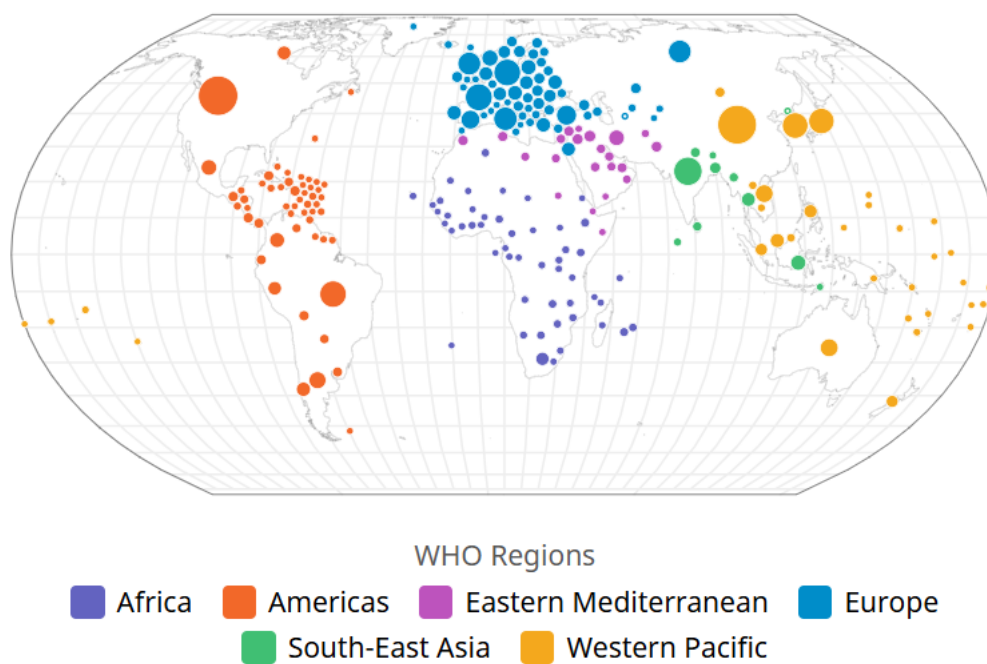


Figure 1. Number of COVID-19 cases reported to WHO (cumulative total).¹⁸

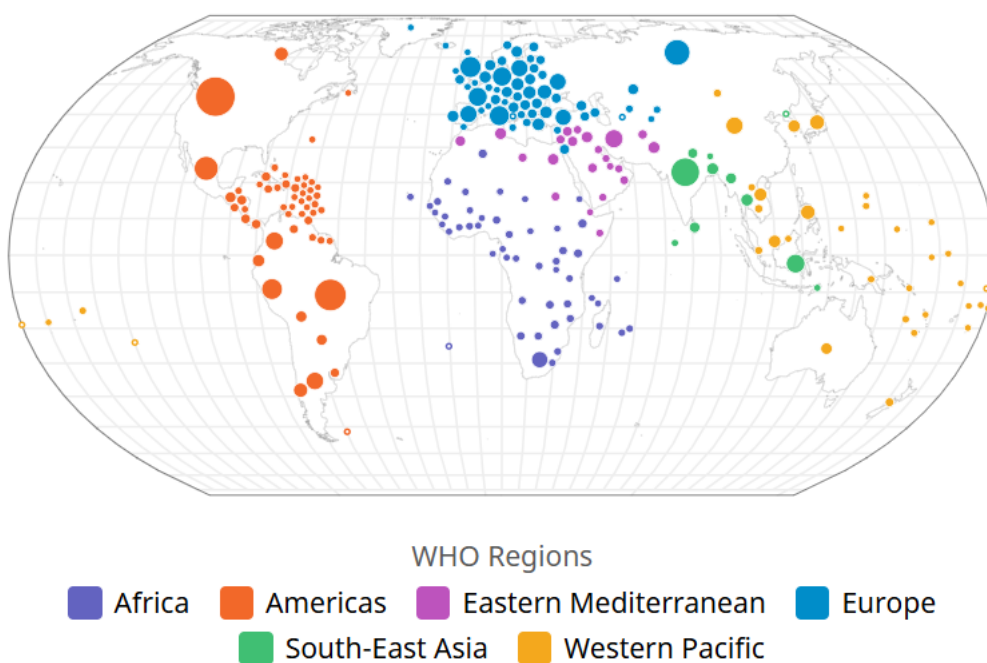


Figure 2. Number of COVID-19 deaths reported to WHO (cumulative total).¹⁸

2.2. Clinical Aspects

The clinical aspects are certainly milder now in the general population than in the early stages of the pandemic in the absence of vaccines. Moreover, they are partly dependent on the different virus variants identified so far. A typical COVID-19 infection has an incubation time of about 5 days. The prognosis is often unpredictable, especially in older patients with comorbidities. The clinical picture ranges from completely asymptomatic cases (about 10%) to rapidly devastating courses. The most common symptoms are fever (88%) and a dry cough (67.7%), shared with many other viral syndromes. Other frequent symptoms are headache, weakness and myalgia. A characteristic that seems to be strongly and independently associated with positivity is the possible alteration of smell and taste. Rhinorrhoea (4.8%) and gastrointestinal symptoms (diarrhoea 4% to 14%, nausea or vomiting 5%) appear to be infrequent. A significant majority of patients present mild symptoms (no pneumonia or mild pneumonia). Among patients with more consistent symptoms, 14% show severe symptoms (dyspnoea, respiratory rate ≥ 30 /min, oxygen saturation $\leq 93\%$, ratio of partial pressure of blood oxygen to the fraction of oxygen in inspired gases < 300 , or pulmonary infiltrates $> 50\%$ within 24-48 hours) that may lead to respiratory failure, septic shock or organ failure in 5% of cases²⁰. These require intensive care and mechanical ventilation. The fatality rate or case-fatality rate (CFR), i.e. the ratio of the number of deaths to the number of diagnosed cases, varies from country to country and shows a significant increase in elderly patients and those with comorbidities, especially in the presence of vascular disease, pulmonary disease, diabetes mellitus, hypertension and cancer.^{21,22}

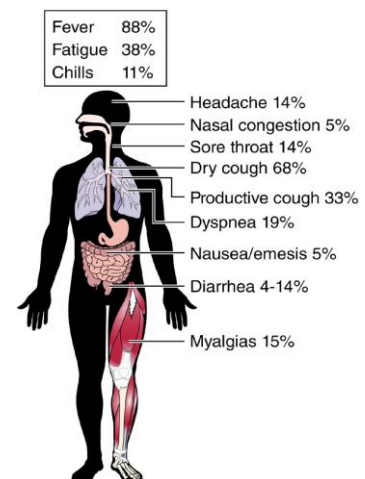


Figure 3. Common symptoms in Covid-19 patients ²³.

Effects of Covid-19 on Cardiac Function

There is growing evidence of direct and indirect effects of SARS-CoV-2 on the heart, especially in patients with pre-existing heart disease²⁴. SARS-CoV-2 has the potential to infect cardiomyocytes, pericytes and fibroblasts via ACE2-mediated mechanisms leading to direct myocardial injury, although the precise sequence of pathophysiological mechanisms is not yet

elucidated²⁵. Post-mortem in situ hybridisation analyses have suggested that the most likely localisation of SARS-CoV-2 is not in cardiomyocytes but in interstitial cells or macrophages invading myocardial tissue²⁶. It has also been shown that the ACE2 receptor is widely expressed on endothelial cells and that direct SARS-CoV-2 infection of the endothelial cell is possible, leading to widespread endothelial inflammation²⁷. Cases examined post-mortem indicate strong virus-induced vascular dysfunction²⁸. Clinically, COVID-19 can manifest as an acute cardiovascular syndrome (called 'ACovCS' for COVID-19 acute cardiovascular syndrome). Troponin is an important evaluation parameter in these patients: among 2,736 COVID-19 patients admitted to one of five New York City hospitals with troponin-I measurements within 24 hours of admission, 985 (36%) patients had elevated concentrations²⁹. In addition to ACovCS, a wide range of cardiovascular manifestations is described, including heart failure, cardiogenic shock, arrhythmia and myocarditis.

Effects of Covid-19 on Coagulation: clinical features and pathophysiology

Despite being a primary respiratory illness, it is commonly complicated by systemic involvement of the vasculature leading to arterial and venous thrombosis³⁰.

Coagulopathy is a commonly recognized feature of the inflammatory response in COVID-19, as 20–50% of hospitalized patients are noted to have abnormalities with elevated D-dimer, prolonged prothrombin time, thrombocytopenia, and abnormal fibrinogen levels. While some of these values resemble DIC (disseminated intravascular coagulation), there are clear differences in the patterns of laboratory abnormalities as well as the clinical complications³¹. The related thrombocytopenia and prolonged prothrombin time tend to be milder than that of DIC. Also, fibrinogen tends to be normal or elevated as an acute phase response rather than reduced in contrast to frank DIC³². The commonly elevated D-dimer in CAC may be one of the only similarities among laboratory findings of DIC and CAC. D-dimer also correlates to patient outcomes in COVID-19, where it has been noted that an elevated D-dimer increases risk of death fourfold^{33,34}. While DIC can have both thrombotic and bleeding phenotypes, the uncommonness of bleeding with CAC presents a key clinical difference between the two³⁵. Bleeding has been reported to occur in less than 5% of hospitalized patients with COVID-19. The lesser risk for bleeding complications with CAC may be related to the suppression of fibrinolysis, caused by two possible mechanisms: one is the high levels of thrombin activable fibrinolysis inhibitor (TAFI) and the release of PAI-1 from endothelial cells³⁶. A second mechanism is dysregulated RAAS due to SARS-CoV-2 mediated downregulation of ACE2, leading to unopposed Angiotensin II and increased levels of PAI-1³⁷.

10 to 30% of COVID-19 patients reported a thrombosis: venous thromboembolism (VTE) has

emerged as the most common form of COVID19 associated thrombosis³⁸. DVT (deep vein thrombosis) and PE (pulmonary embolism) were frequently reported complications of hospitalized COVID-19 patients, especially patients admitted to the ICU. Also arterial thrombotic events were frequently reported³⁹.

These complications arise from the inflammatory cascade triggered by the virus binding to the ACE2 (transmembrane angiotensin-converting enzyme 2) receptor. This cascade involves cytokines, white blood cells, and the complement system, ultimately leading to suppressed fibrinolysis and endothelial dysfunction³⁰. ACE2 is expressed ubiquitously and plays a critical role as a negative regulator of the renin-angiotensin-aldosterone system (RAAS). It does so by inactivating the pro-inflammatory angiotensin II by converting it to angiotensin 1-7.^{40,41} When SARS-CoV-2 binds the ACE2 receptor, it initiates a cascade that ultimately leads to the downregulation of the receptor⁴². Downregulation of ACE2 leads to unopposed angiotensin II effects, causing proinflammatory and prothrombotic changes⁴³. It stimulated the release of chemokines (CCL2, CCL3, CCL4, CXCL8, CXCL9, CXCL10) and cytokines (TNF α , IFN γ , IFN α , IL-12, IL-1 β , IL-6, IL-33, IL-18, TGF β) from monocytes, macrophages and neutrophils that activates the proinflammatory pathways mediated by NF-kB, MAPK, STAT3, and MyD88.^{14,44,45} Increased angiotensin II leads to vasoconstriction, but also oxidative stress damage, which is mediated by nicotinamide adenine dinucleotide phosphate (NADPH) activation and the production of reactive oxygen species (ROS)⁴⁶. The result is a massive hyperinflammatory response on the endothelium leading to prothrombotic changes⁴⁷. Indeed, 'cytokine storm' and oxidative stress lead to a downregulation of thrombomodulin (a glycoprotein that binds, blocking, thrombin) expressed on the surface of endothelial cells and an upregulation of tissue factor on macrophages and neutrophils, which activates the coagulation cascade⁴⁸. The von Willebrand factor also undergoes upregulation in this proinflammatory environment especially in severe patients⁴⁹.

Covid-19 and Lipoprotein (a)

Lipoprotein(a) [Lp(a)] is an atherogenic lipid particle composed of a polymorphic glycoprotein called apo(a), covalently bound to a low-density lipoprotein (LDL) portion containing apoB100. This structure includes a core of cholesterol esters and triacylglycerols, surrounded by a shell of phospholipids and free cholesterol⁵⁰. Apo(a) is characterised by the presence of ring structures, called kringles (K),^{51,52} which are also present in other molecules such as plasminogen, prothrombin, urokinase and tissue-type plasminogen activator (Figure 4). These features explain the involvement of Lp(a) in the atherosclerotic and pro-thrombotic process, supporting its

association with the pathogenesis of cardiovascular disease (CVD).^{6,7,53–56}

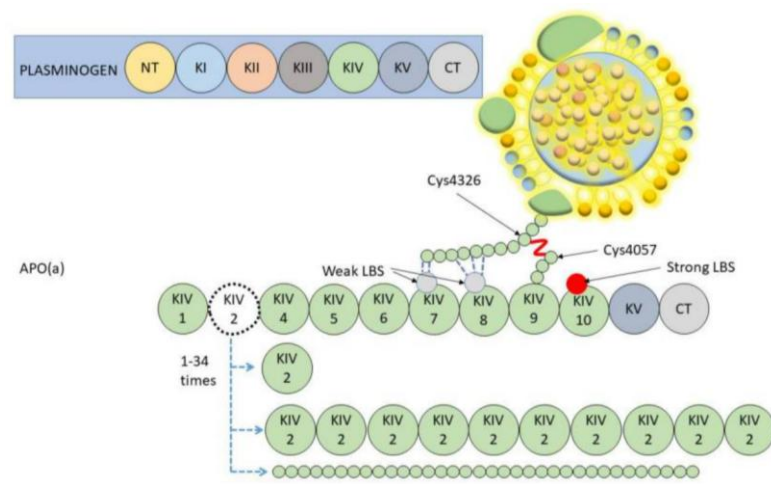


Figure 4. Schematic illustration of structural homology between plasminogen and apolipoprotein (a) [apo(a)]⁵⁷

The mechanisms underlying the influence of Lp(a) in atherogenesis are still not entirely clear. One proposed mechanism involves the direct deposition of lipoprotein on arterial walls. Since it is more susceptible to oxidation than LDL, this accelerates the uptake of Lp(a) by macrophages through scavenger receptors and results in their transformation into foam cells⁵⁸. In addition, the contribution of oxidised phospholipids (OxPLs) contained within the lipoprotein is also relevant in the expression of adhesion molecules (ICAM, VCAM, E-selectin) on the surface of endothelial cells and in facilitating their binding to monocytes by reducing their production of nitric oxide. Another possible proatherogenic mechanism of Lp(a) concerns the inverse relationship between lipoprotein levels and vascular reactivity. In this case, increased plasma levels induce endothelial dysfunction⁵⁹.

Lp(a) is synthesised in the liver under the control of the LPA gene, located on chromosome 6. Its plasma values vary from less than 1 mg/L to more than 3000 mg/L but, unlike other lipoproteins, its synthesis is little affected by environmental or dietary factors^{55,60,61}. In fact, between 70-90% of the variability in blood Lp(a) levels is due to genetic factors⁵⁰.

Although genome-wide association studies (GWAS) have identified more than 200 single nucleotide polymorphisms (SNPs), which could have a possible effect on the final concentration of Lp(a),⁶² a copy number variation (CNV) consisting of a variable number of a 5, 6 kb, comprising both exons 4 and 5 encoding the kringle IV type 2 (KIV2) protein domain, is its main genetic determinant, explaining approximately 70% of the variance in the Caucasian population. In fact, a variable number of copies per allele has been described, ranging from 3 to more than 40, thus generating >40 isoforms of apo(a) that have different final sizes and, in particular, small

isoforms with a low number of KIV2 repeats are associated with higher plasma concentrations of Lp(a).⁶¹

IS THERE A RELATIONSHIP BETWEEN ELEVATED LP(A) AND COURSE OF COVID-19?

Given its potential prothrombotic and proinflammatory properties, Lp(a) has been suggested as a possible contributor to a worse clinical outcome in patients with COVID-19.^{63,64} Indeed, due to its content of proinflammatory oxidized phospholipids, Lp(a) might directly cause endothelial injury and impair the antithrombotic function of the endothelium.^{63–66} Additionally, Lp(a) might act as an antifibrinolytic factor, promoting thrombus formation and progression due to its structural similarity to plasminogen.^{63–66} Although circulating Lp(a) levels are largely genetically determined, an increase in Lp(a) levels can occur in various inflammatory conditions due to the activation of the interleukin-6 (IL-6) response element on the LPA gene, which encodes apolipoprotein(a).^{63,64} As a result, the amplification of the thrombo-inflammatory cycle could occur in COVID-19 patients with elevated Lp(a) levels, potentially worsening their clinical outcomes³.

However, the data in the literature are discordant. In a study on 124 COVID-19 patients, patients with elevated Lp(a) level had significantly longer hospitalization time and higher oxygen demand on admission. Indeed, elevated Lp(a) was also significantly associated with intubation, ICU and Death rate⁴. Kaltoft and co-workers concluded that in COVID-19 patients, thrombotic activity marked by elevated D-dimer was associated with higher lipoprotein(a) while elevated inflammatory biomarkers of interleukin-6, C-reactive protein, and procalcitonin were associated with lower lipoprotein(a); however, elevated lipoprotein(a) was not associated with rate of hospital death or discharge⁵. In the opposite way, according to Bianconi and co-workers, Lp(a) does not influence biomarkers of plasma thrombotic activity and systemic inflammation nor has any impact on thrombotic events and adverse clinical outcomes in patients hospitalized for COVID-19³.

GENETIC CHARACTERIZATION OF LIPOPROTEIN (A) KRINGLE IV TYPE 2 REPEATS POLYMORPHISM

Numerous studies have confirmed the link between elevated plasma Lp(a) levels and increased cardiovascular disease (CVD) risk, highlighting the need to measure Lp(a) to support risk stratification of patients.^{54,55,67,68} However, the vast majority of available assays for measuring plasma Lp(a) levels utilise antibodies directed against the repetitive motif of the apo(a) protein, which is the unique component of Lp(a): due to the high degree of dimensional heterogeneity of apo(a), this causes measurement bias, making the development of accurate methods for measuring

plasma Lp(a) levels extremely complex.^{69,70}

To address this issue, since over 90% of the variance in Lp(a) levels is genetically determined, with the KIV2 repeat polymorphism being the main determinant, several groups have sought to demonstrate that characterizing this CNV can be used to indirectly assess Lp(a) levels and evaluate its role in modulating cardiovascular risk.^{6,7} Supporting this, Kamstrup and colleagues in 2009⁶⁷ also showed that genotyping could potentially be more effective in risk prediction compared to a single plasma measurement, which could be influenced by transient endogenous or exogenous factors.

The gold standard methods for evaluating KIV2 repeat polymorphism include Western blot analysis, considered the most informative as it analyzes the polymorphism at the protein level, separating the different apo(a) isoforms through SDS-PAGE electrophoresis; and Pulsed Field Gel Electrophoresis (PFGE), which provides the number of KIV2 repeats at the DNA level for each of the two apo(a) isoforms.⁷¹ These techniques are both highly accurate, but lengthy and methodologically complex, making them difficult to apply in routine diagnostic activities.

Consequently, a faster, high-throughput method such as quantitative real-time PCR (qPCR) has been developed, which provides the total number of KIV2 repeats across both alleles, but without offering genotype-specific information.^{71,72} Although this method is ideal for large-scale studies, it still has limitations in terms of stability and sensitivity.

An emerging method in various applications, including CNV analysis, is digital droplet PCR (ddPCR). ddPCR is rapidly gaining popularity over standard real-time PCR assays due to its relative ease of use, precise and accurate results, and ability to quantify nucleic acid concentrations without the need for standard samples.^{73,74} Barbieri et al. compared qPCR and ddPCR techniques on a population of 100 patients and data obtained support a better performance of ddPCR in the evaluation of KIV2 repeat polymorphism. In fact, A significant inverse correlation between Lp(a) levels and CNV values was confirmed for both techniques, but it was higher when evaluated by ddPCR than qPCR. When dividing subjects into two groups according to 500 mg/L Lp(a) cut-off value, a significantly lower number of KIV2 repeats emerged among subjects with greater Lp(a) levels, with stronger evidence in ddPCR than in qPCR⁷⁵. Indeed, it could be interesting to notice that ddPCR is actually a cost-effective technique, given the added value of a greater accuracy and stability, combined with a faster execution (thus no longer requiring duplicated analyses) leading to a reduction of the cost per man-hour. Moreover, no data normalization is required, so results are immediately available for clinical evaluation⁷⁵.

2.3. Treatments

The Italian Medicines Agency (AIFA) has drawn up guidelines that make explicit the therapeutic guidelines within which a controlled and safe use of the drugs used in this emergency can be envisaged⁷⁶.

Symptomatic Drugs

For the treatment of patients with mild symptoms (fever or joint or muscle pain), paracetamol or NSAIDs (non-steroidal anti-inflammatory drugs) are administered.

Drugs to be used only at Specific Stages of the Disease

Two **antivirals** (remdesivir, nirmatrelvir/ritonavir) are currently authorised by the EMA for the treatment of adults with COVID-19 who do not require supplemental oxygen therapy and who are at increased risk of progression to severe forms of COVID-19. The patient must not be hospitalised due to COVID-19, must present with a mild-to-moderate form and risk factors as Oncological/oncohaematological pathology in active phase, chronic renal failure and chronic obstructive pulmonary disease and/or other chronic chronic respiratory disease.

One is Remdesivir–Veklury ®, a propharmaceutical nucleotide analogue of adenosine that is metabolised in host cells to form the pharmacologically active nucleoside triphosphate metabolite. Remdesivir triphosphate acts as an adenosine triphosphate (ATP) analogue and competes with the natural ATP substrate for incorporation into nascent RNA chains by the SARS-CoV-2 dependent RNA-polymerase, causing delayed chain termination during viral RNA replication. The other one is Nirmatrelvir/ritonavir–Paxlovid ®, whose active ingredient is nirmatrelvir, which acts by reducing the ability of SARS-CoV-2 to multiply within the body through a mechanism of inhibition of certain viral proteases necessary for the virus to replicate. Ritonavir, on the other hand, inhibits the cytochrome P3A (CYP3A) metabolism of the active ingredient, so that the action of nirmatrelvir is prolonged.

The **monoclonal antibodies** available in Italy, authorised by EMA for the treatment of SARS-CoV-2 infection in individuals at high risk of progression to severe disease, are: the combination casirivimab/imdevimab, sotrovimab, and the combination tixagevimab/cilgavimab. All three bind the spike protein, preventing the virus from entering the host cell. The candidate population for therapy with the three treatments is represented by subjects aged 12 years or older (and at least 40 kg), SARS-CoV-2 positive, not hospitalised for COVID-19, not on oxygen therapy for COVID-19, with mild-to-moderate symptoms and who are at high risk of severe COVID-19. Indeed, COVID-19 must be of recent onset (no more than 7 days).

The use of **corticosteroids** (such as dexamethasone) is recommended in hospitalised individuals

with severe COVID-19 disease who require oxygen supplementation. This recommendation is based on the fact that there is currently evidence of a clinical benefit of these drugs only in this patient setting/disease phase.

In hospitalised patients with reduced mobility due to acute respiratory infection (including COVID-19), the use of **low molecular weight heparins** is recommended to prevent thromboembolic events for the duration of immobility. In non-hospitalised and non-bedridden patients, the routine use of heparins is not recommended as there is insufficient evidence of clinical benefit in this context.

2.4. Vaccines

Below are the vaccines currently approved by AIFA⁷⁶.

mRNA Vaccines

The European Medicines Agency (EMA) and AIFA have authorised two COVID-19 mRNA vaccines. They are the Pfizer mRNA BNT162b2 (**Comirnaty**) and COVID-19 Vaccine Moderna mRNA-1273 (**Spikevax**) vaccines. SARS-CoV-2 infects people using a surface protein called Spike, which acts like a key to enter cells and replicate. All current vaccines aim to block the Spike protein, preventing cell infection. The two approved mRNA COVID-19 vaccines use messenger RNA (mRNA) molecules, which instruct vaccinated cells to produce Spike proteins. These proteins stimulate the immune system to produce specific antibodies that, upon exposure to the virus, block the Spike protein and prevent it from entering cells. Vaccination also activates T cells, helping the immune system respond to future SARS-CoV-2 exposures. The vaccine doesn't introduce the actual virus but provides the genetic information for cells to produce Spike protein copies. If a vaccinated person encounters SARS-CoV-2 later, their immune system will recognize and fight the virus.

Viral-Vector Vaccines

EMA and AIFA have so far authorised two viral vector COVID-19 vaccines: **Vaxzevria** and **Jcovden** (formerly COVID-19 Vaccine Janssen). A viral vector vaccine uses a virus (usually an adenovirus incompetent to replicate) to carry the genetic code sequence coding for the spike protein into the cell. The immune system is activated against the protein and produces antibodies that, should the subject come into contact with the virus, will protect him or her from infection. Vaxzevria uses a non-replicating chimpanzee adenovirus (ChAdOx1) modified to deliver the genetic information needed to produce the SARS-CoV-2 Spike protein, while Jcovden uses a recombinant vector based on human adenovirus type 26 incompetent for replication. Both vaccines are designed to prevent COVID-19 in individuals aged 18, but the authorization for their commercialization has been withdrawn at the request of the manufacturer for commercial reasons.

Protein-Subunit Vaccines

The EMA and AIFA authorised two anti-COVID-19 protein subunit vaccines: **Nuvaxovid** and **VidPrevtyn Beta**. Protein subunit vaccines consist of "protein fragments" from the virus. In the production of this type of COVID-19 vaccine, a portion of DNA containing the instructions to produce the Spike protein is inserted into a baculovirus. The baculovirus then infects cells in vitro, releasing the genetic material needed to produce the Spike protein. This protein is then extracted, purified, and assembled into viral nanoparticles containing the Spike protein.

These nanoparticles, combined with an adjuvant molecule to further stimulate the immune system, are ready to be injected into the human body, prompting the production of antibodies against the SARS-CoV-2 Spike protein. Nuvaxovid is used in individuals over 12 years of age, while VidPrevtyn is indicated as a booster dose in adults who have previously received an anti-COVID-19 mRNA or adenoviral vector vaccine. The marketing authorisation for VidPrevtyn Beta vaccine was revoked at the holder's request on commercial grounds.

Inactivated-Virus Vaccines

EMA and AIFA also approved **Valneva**, the first inactivated viral vaccine, however, its marketing authorisation was revoked at the holder's request on commercial grounds. Inactivated virus vaccines are produced by culturing the SARS-CoV-2 virus in cell cultures and then chemically inactivating it. When the vaccine is administered, the immune system identifies the inactivated virus as foreign and produces antibodies and T-cells against it.

Vaccine-Induced Thrombotic Thrombocytopenia (VITT)

The widespread availability of COVID-19 vaccines has significantly changed the pandemic's trajectory. However, in February 2021, a prothrombotic syndrome was observed in a small number of individuals who received adenoviral vector-based vaccines against SARS-CoV-2, specifically Vaxzevria (AstraZeneca) and Janssen (Johnson & Johnson). The European Medicines Agency's Pharmacovigilance Risk Assessment Committee concluded that there is a causal relationship between Vaxzevria vaccination and these adverse events, known as vaccine-induced immune thrombotic thrombocytopenia (VITT).^{9,13}

Literature from 2021 outlined five criteria for diagnosing VITT: symptom onset 5 to 30 days post-vaccination, thrombosis presence, thrombocytopenia (platelet count $<150,000/\text{mm}^3$), elevated D-dimer levels, and detection of PF4 antibodies. The condition primarily affected women under 55 years, particularly after the Vaxzevria and Janssen vaccines, with an incidence of about 1.02 cases per 100,000 doses for those over 60, rising to approximately 1 in 50,000-75,000 doses for younger individuals.⁸⁻¹⁰

VITT is characterized by the combination of thrombosis and thrombocytopenia, a phenomenon that is quite rare in medicine. It is particularly found in heparin-induced thrombocytopenia (HIT) in its autoimmune form (aHIT). In fact, the pathophysiological model of VITT is similar to that of aHIT, where platelet factor 4 (PF4) binds to negatively charged molecules on the endothelial surface through their FcγRIIA receptors. In the case of VITT, however, PF4 likely binds to components of the viral vector or the DNA contained within it.^{11,12} Indeed, unlike mRNA vaccines, which release mRNA directly into the cytosol of host cells where it is translated into spike proteins without the need for splicing reactions, adenoviral vector-based vaccines require a

nuclear transcription process. This could lead to the formation of shorter spike protein variants, potentially responsible for the release of soluble forms capable of triggering an intense inflammatory process, thus promoting the development of VITT. Moreover, the presence of polysorbate as an adjuvant in ChAdOx1 and Ad26.CoV2.S vaccines, a non-anionic surfactant able to cross the blood-brain barrier and reach brain endothelial cells when complexed with nanoparticles^{45,77} might play a role in vaccine-induced complications.

Additional molecular mechanisms may also contribute to the activation of VITT; it has been suggested that genetically determined increased expression of the FcγRIIIa receptor, along with altered glycosylation of IgG produced in response to vaccination, could enhance reactivity towards the FcγRIIIa receptor on platelets, contributing to disease mechanisms^{78,79}. Furthermore, it is hypothesized that adenoviral vector leakage into the bloodstream and/or the presence of pre-existing antibodies cross-reactive with other coronaviruses may lead to the formation of immune complexes capable of activating platelets⁸⁰. In any case, this binding leads to the formation of a neoantigen that triggers a specific antibody response, creating a prothrombotic environment that is self-sustaining within the lumen of blood vessels and on cellular surfaces. Nevertheless, the exact sequence of events that leads to the syndrome has not yet been fully clarified^{11,12,81}. Ultimately, the result is a thrombocytopenic syndrome caused by the consumption of platelets and immunothrombosis.

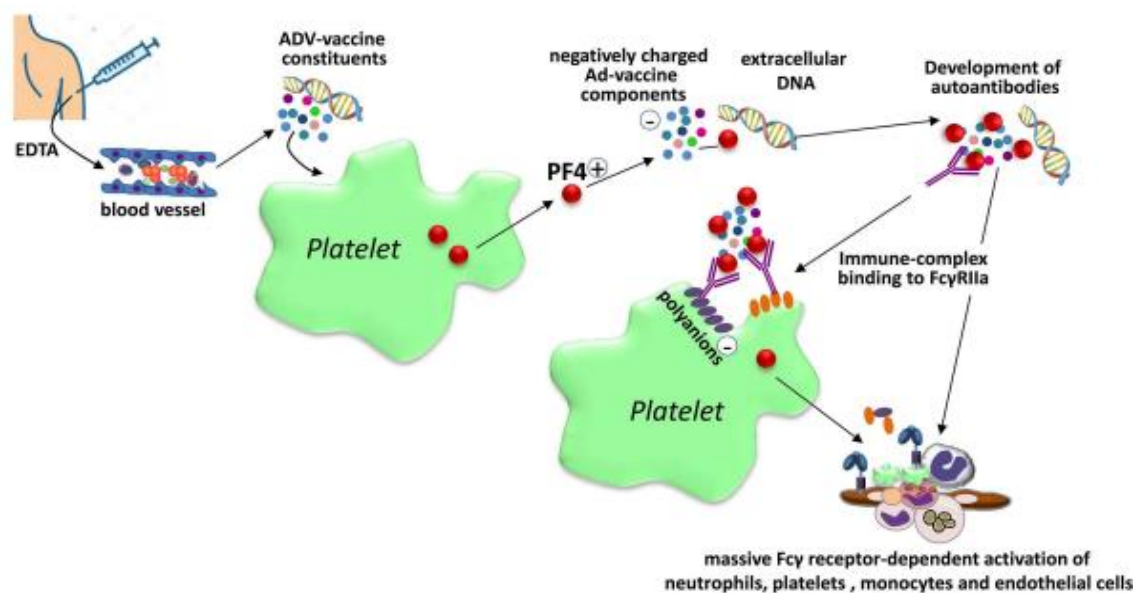


Figure 5. The autoimmune heparin-induced thrombocytopenia hypothesis.⁷⁸

2.5. Virus Classification and Taxonomy

The institution responsible for the classification and naming of Coronaviridae is the Coronaviridae Study Group (CSG) of the International Committee on Taxonomy of Viruses (ICTV). Its process involves assessing the genetic correlation between a new candidate virus and existing viruses within a taxon. A computational framework is then used to compare genomic sequences and identify related clusters and variations. A candidate virus is assigned to the appropriate cluster if it shares more than 46% nucleotide homology³⁷.

Under the current taxonomy, CoVs are classified as members of the Nidovirales order. The Coronaviridae family is further divided into two subfamilies: Coronavirinae and Torovirinae.

The Coronaviridae family is also divided into four genera: Alpha-, Beta-, Gammacoronavirus (previously corresponding to groups 1, 2, and 3, respectively), and the more recently identified Deltacoronavirus³⁷.

Members of the Coronaviridae family infect various species, including humans, other mammals, and birds, causing a wide spectrum of diseases. Among human CoVs, 229E and NL63 are classified within the Alphacoronavirus genus, while the others belong to Betacoronavirus. HKU1 and OC43 are part of the Embecovirus subgenus, SARS-CoV and SARS-CoV-2 are part of the Sarbecovirus genus, and MERS-CoV belongs to the Merbecovirus genus³⁷.

Numerous species of alpha and betacoronaviruses have been found in bats, making this mammalian order the most affected. For this reason, bats have been identified as reservoirs for incubation, evolution, and cross-species transmission⁸².

SARS-CoV-2 was classified by the CSG within the Coronaviridae family. Phylogenetic analysis conducted to determine the relationship between SARS-CoV-2 and other CoVs grouped it within the Betacoronavirus genus, specifically under the Sarbecovirus subgenus⁸³. Notably, there is a 94.4% homology with SARS-CoV in the seven conserved replicase domains in ORF1ab, forming a distinct clade within the species of coronaviruses related to severe acute respiratory syndrome (SARSr-CoV).

The SARSr-CoV species includes hundreds of known viruses, primarily isolated from humans and various bat species. Understandably, the reference to "SARS" can be misleading, as SARS-CoV-2, along with other SARSr-CoVs, does not cause SARS as a clinical disease. SARS-CoV was the prototype for a new viral species, which is why a unique name was assigned to the species according to established viral taxonomic practices. Therefore, virus nomenclature does not necessarily indicate a disease similar to SARS but refers to its phylogenetic grouping within the founding species^{37,84}.

ORDER	NIDOVIRALES		SPECIES	VIRUS
FAMILY	CORONAVIRIDAE			
SUBFAMILY	CORONAVIRINAE			
GENUS	ALPHA- (GROUP 1)			
			Bat CoV CDPHE15	Bat CoV CDPHE15
			Bat CoV HKU10	Bat CoV HKU10
			Human CoV 229E	HCov 229E
			Lucheng Bat rat CoV	Lucheng Bat rat CoV
			Mink CoV 1	Mink CoV 1
			Miniopterus bat CoV 1	Miniopterus bat CoV 1A
			Miniopterus bat CoV HKU8	Miniopterus bat CoV HKU8
			Porcine epidemic diarrhoea virus	Porcine epidemic diarrhoea virus
			Scotophilus bat CoV 312	Scotophilus bat CoV
			Rhinolophus bat CoV HKU2	Rhinolophus bat CoV HKU2
			Human CoV NL63	HCov NL63
			Alphacoronavirus 1	Transmissible gastroenteritis virus
GENUS	BETA- (GROUP 2)			
			Embecovirus	
			Betacoronavirus 1	Bovine CoV Mebus
			China Rattus CoV HKU24	China Rattus CoV HKU24
			HCov HKU1	HCov HKU1-A
			Marine CoV	Mouse hepatitis virus A59
			Myodes CoV J114	Myodes CoV J114
			HCov-OC45	HCov-OC45
			Sarbecovirus	
			SARS-related CoV	SARS-CoV
				SARS-CoV PC-4227
				SARS-CoV RA1013
				SARS-CoV-2
			Merbecovirus	
			Hedgehog CoV 1	Hedgehog CoV 1
			Tylonycteris bat CoV HKU4	Tylonycteris bat CoV HKU4-1
			Pipistrellus bat CoV HKU5	Pipistrellus bat CoV HKU5
			MERS-related CoV	MERS
			Norbecovirus	
			Eidolon bat CoV C704	Eidolon bat CoV C704
			Rousettus bat CoV GCDC1	Rousettus bat CoV GCDC1
			Rousettus bat CoV HKU9	Rousettus bat CoV HKU9-1
			Avian CoV	Infectious bronchitis virus
			Avian CoV 9203	Avian CoV 9203
			Duck CoV 2714	Duck CoV 2714
			Goose CoV CH17	Goose CoV CH17
			Beluga whale CoV SW1	Beluga whale CoV SW1
			Mamm CoV HKU19	Mamm CoV HKU19
			CoV HKU15	CoV HKU15
			Bubal CoV HKU11	Bubal CoV HKU11
DEL-TA-	GAMMA- (GROUP 3)			

Figure 6. Classification of Coronaviruses. The blue text highlights common respiratory infections circulating in humans. The orange text highlights zoonotic viruses with varying pathogenicity in humans³⁷.

2.6. Virus Structure

CoV virions have a spherical morphology, with an average diameter of 80–120 nm, and exhibit a moderate degree of pleomorphism. They are enveloped by a lipid bilayer and contain positive-sense single-stranded RNA. The Coronaviridae family has four key structural proteins: spike (S), membrane (M), and envelope (E), which form the viral envelope, and the nucleocapsid protein (N) housed within. Since SARS-CoV-2 is part of the CoV family, it retains these essential structural components⁸⁵.

Spike Glicoprotein (S)

The spike glycoproteins function as class I viral fusion proteins, mediating attachment to host cell receptors and facilitating fusion for cellular entry⁸⁶. Each spike assembles into a trimer, projecting about 20 nm from the viral envelope. These globular protrusions give CoVs their characteristic crown-like appearance when observed under an electron microscope.

Each S monomer consists of a long N-terminal ectodomain (extracellular), a transmembrane domain (anchored in the viral envelope), and a short C-terminal endodomain (intracellular). In most CoVs, the ectodomain is cleaved by host cell proteases into two functional polypeptide subunits, named S1 and S2. The S1 subunit corresponds to the globular head of the spike protein and contains the receptor-binding domain for the host cell, while the S2 subunit forms the narrow stalk and facilitates viral and cellular membrane fusion for entry.

Both SARS-CoV and SARS-CoV-2 utilize ACE2 (Angiotensin-Converting Enzyme 2, a metalloprotease normally involved in blood pressure regulation) as a receptor for host cell entry. Consequently, the interaction between the spike protein and ACE2 has been considered a promising therapeutic target for the development of vaccines,^{87,88} neutralizing antibodies,^{89,90} and antiviral compounds^{91,92} for SARS-CoV and SARS-CoV-2. The sequence homology between the spike protein of SARS-CoV and SARS-CoV-2 is approximately 76% for the entire protein, 73% for the receptor-binding domain, and 50% for the receptor-binding motif.⁹³

The spike protein of SARS-CoV-2 contains two distinctive features not present in the spike protein of SARS-CoV. First, SARS-CoV-2's spike protein retains a high binding affinity for human ACE2, even though five of the six residues within the receptor-binding motif of SARS-CoV, which are critical for binding human ACE2, have mutated in SARS-CoV-2 (L455, F486, Q493, S494, and N501).^{94,95} Cryo-electron microscopy studies have shown that SARS-CoV-2's spike protein can bind directly to human ACE2 with a similar or even higher affinity than that of SARS-CoV.^{93,96,97} This high binding affinity is likely related to the elevated transmissibility of SARS-CoV-2 and the severity of COVID-19.

Secondly, the spike protein of SARS-CoV-2 has an insertion of four amino acid residues (12 nucleotides) at the junction between the S1 and S2 subunits, introducing a novel furin cleavage site. This cleavage site is not found in SARS-CoV or other SARS-related coronaviruses originating from bats, and it appears to facilitate the cleavage of the spike protein at the S1-S2 boundary for priming by furin-like proteases, which are ubiquitously expressed. Although the function of this novel cleavage site remains unknown, similar cleavage sites have been described in highly pathogenic avian influenza viruses and Newcastle disease virus.^{94,95} This notable feature has been proposed to play a role in the expanded tissue and cell tropism of SARS-CoV-2, contributing to the multi-organ effects of COVID-19.

Given the spike glycoprotein's exposed surface location and its role in cell entry, it is considered the main target for neutralizing antibodies during infection. For this reason, the S monomers are heavily glycosylated with N-linked glycans that protrude from their surface to shield epitopes from neutralizing antibodies and aid immune evasion. Viral glycosylation also plays a critical role in the folding and stability of the spike glycoprotein, as well as modulating its accessibility to host proteases, which in turn affects viral tropism.⁹⁷

The Membrane Glycoprotein (M)

The membrane glycoprotein (M) is the most abundant of all CoV structural proteins. It exists as glycosylated dimeric molecules embedded in the viral envelope and can adopt either a compact or elongated form. The compact form is responsible for giving coronaviruses their spherical structure by bending the viral membrane. On the other hand, the elongated form plays a central role in viral assembly, promoting spike formation through interaction with the ribonucleocapsid. In addition to its structural function, M proteins also influence the organism's tropism and induce the release of interferon (IFN- β).

The Envelope Glycoprotein (E)

The envelope (E) protein is the smallest and least understood of the structural proteins in Coronaviridae. Although produced in large amounts during viral replication, only a small fraction is incorporated into the viral membrane, making it relatively scarce. The E protein can form multimers (dimers to pentamers) and consists of an N-terminal domain (NTD), a hydrophobic transmembrane domain, and a C-terminal domain (CTD). It plays a key role in viral assembly by interacting with the M protein through its CTD, which induces membrane curvature. Additionally, it facilitates the formation of pores (viroporins) across the membrane, crucial for effective virion trafficking through the secretory pathway.^{98,99}

Nucleocapsid Protein (N)

The N protein encapsulates the viral genome and is enclosed within the virion envelope. It forms

a helical structure with a diameter of 9-13 nm and a central channel of 3-4 nm, binding along the viral RNA in a pearl necklace-like arrangement. Phosphate groups on the N protein are thought to regulate its binding affinity for viral RNA. While the primary function of the N protein is to stabilize the viral RNA, it also plays a role in viral entry into the host cell and interacts with cellular processes post-fusion⁸⁵.

Haemagglutinin Esterase (HE)

This additional protein is present in a subset of beta-coronaviruses, such as murine CoVs and human coronavirus HKU1. The HE protein is located within the viral envelope as a dimer, forming short 5-10 nm projections. It acts as a cofactor for the S protein, aiding in viral attachment to host cells and assisting in transit through the extracellular mucosa. However, it is not present in SARS-CoV-2.

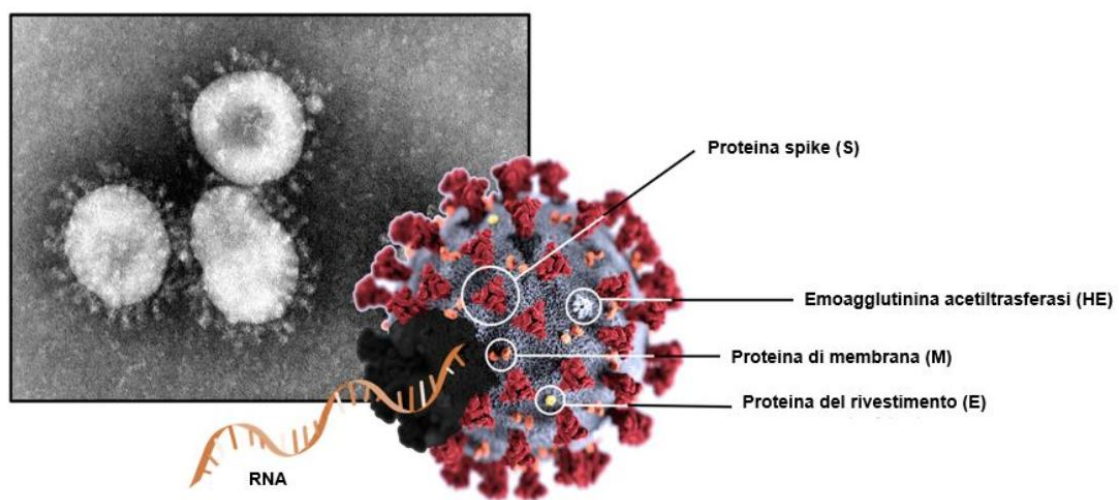


Figure 7. Structure of SARS-CoV-2 (CDC, Centres for Disease Control and Prevention)

The Viral Genome

The genome of the Coronaviridae family consists of single-stranded RNA, ranging from 26 to 32 kilobases in length. Unlike most RNA viruses, which typically have genomes under 10kb, CoV genomes are unusually large. Structurally, they share common features: a 5' cap structure, a polyadenylated 3' tail, untranslated regions (UTRs) at both ends, and a conserved set of structural and non-structural genes, along with strain-specific accessory genes. The genomes of all Coronaviridae members are organized in a 5' to 3' order: ORF1a/b (replicase gene), spike (S), envelope (E), membrane (M), and nucleocapsid (N), with accessory genes interspersed among the structural genes near the 3' end.

The replicase gene (ORF1a/b) encodes the pp1a and pp1b polyproteins, which are processed into

16 non-structural proteins (NSPs). Each NSP has a specific function in the replication of coronaviruses, although not all of these functions are fully understood. Their roles are summarized in the table below (Figure 9).

Downstream of the replicase gene, the genome contains ORFs coding for structural proteins (S, E, M, N) and accessory proteins. The accessory genes, which vary in number (ranging from one to eight), are interspersed among the structural genes. In some cases, these accessory ORFs overlap or are embedded within other genes, complicating bioinformatic analysis. The accessory proteins differ in size and position across viral subgroups, and while their exact functions remain unclear, they are thought to contribute to the pathogenicity in the virus's natural host. Studies involving knockouts or deletions of these accessory genes suggest that they are not essential for viral replication.

The SARS-CoV-2 genome follows the typical CoV organization, consisting of a 29,903-base, single-stranded, positive-sense RNA. In addition to the replicase and structural genes, it contains ten ORFs encoding accessory proteins. The genome sequence is as follows: 5'-cap, 5' UTR, ORF1a/b (replicase gene), spike (S), ORF3a, ORF3b, envelope (E), membrane (M), ORF6, ORF7a, ORF7b, ORF8, nucleocapsid (N), ORF9, ORF10, 3' UTR, and 3' polyadenylated tail (Figure 8).^{100,101}

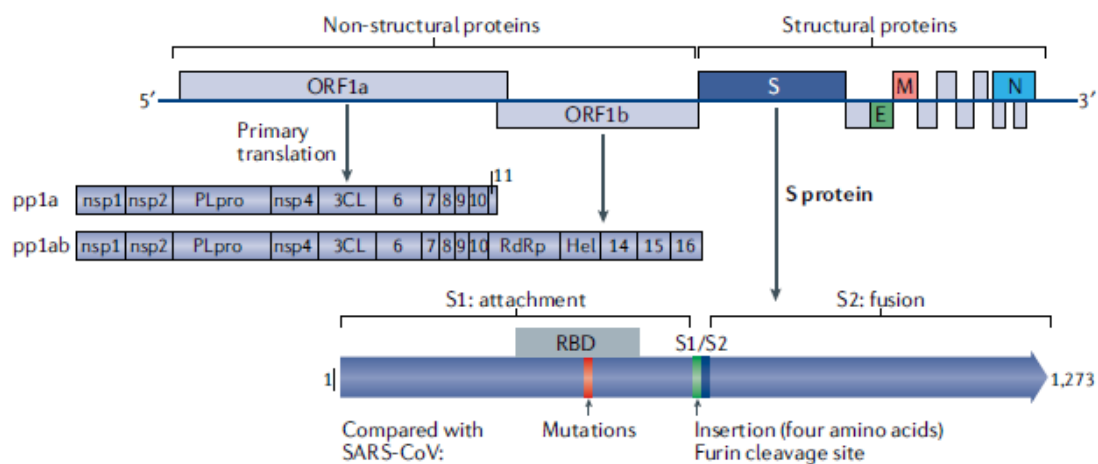


Figure 8. SARS-CoV-2 genome and structure of the S protein.¹⁰²

GENE	NON-STRUCTURAL PROTEIN (NSP)	FUNCTION
ORF 1A	NSP 1	- Inhibits interferon signaling - Inhibit host protein synthesis - Cellular mRNA degradation
	NSP 2	Unknown
	NSP 3	- Tethers genome to RTC allowing initiation of RNA synthesis - Papain-like protease (PLP) for polypeptide cleaving - Inhibit host innate immune response - Promotes cytokine expression
	NSP 4	- Transmembrane helices that anchor RTC to intracellular membranes - Double membrane vesicle formation
	NSP 5	- Main protease (Mpro) for polypeptide cleaving - Chymotrypsin-like protease (3CLpro) for polypeptide cleaving - Inhibits interferon signaling
	NSP 6	- Transmembrane helices that anchor RTC to intracellular membranes - Double membrane vesicle formation
	NSP 7	Essential small proteins: Hexadecameric complex (cofactor NSP8 and NSP12)
	NSP 8	Essential small proteins: - Hexadecameric complex, (cofactor NSP7 and NSP12) - Primase
	NSP 9	Essential small proteins: - RNA-binding protein - Dimerization
	NSP 10	Essential small proteins: - Zinc binding domain (ZBD) cofactor for 2'-O-methyltransferase (2'-O-MTase) - Scaffold protein for NSP14 and NSP16
ORF 1B	NSP 11	Unknown
	NSP 12	RNA-dependent polymerase (RdRP)
	NSP 13	- Zinc binding domain (ZBD) - RNA 5' triphosphatase - synthesis of 5' terminal cap structure of mRNA - RNA helicase - unwinds RNA duplexes with a 5'-3' polarity
	NSP 14	3'-5' exonuclease - some proof-reading activity that is unique to CoV - N7-methyltransferase
	NSP 15	Endoribonuclease
	NSP 16	2'-O-methyltransferase (2'-O-MTase) - Downregulating host innate immunity

Figure 9. The functions of the 16 non-structural proteins (NSPs). NSP-10 is translated by ORF1A, while NSP11-16 are translated by ORF1B⁸⁵.

SARS-CoV-2 Variability

RNA viruses, such as SARS-CoV-2, are known for their high mutation rates, primarily due to the lack of proofreading mechanisms. However, SARS-CoV-2 possesses the NSP14 enzyme, an exoribonuclease with proofreading capabilities, which helps maintain the stability of an unusually large genome. Although all sequenced genomes so far show a high similarity (over 99% sequence identity), there are regions of the genome with significant variability. The S1 subunit of the spike protein is particularly prone to mutations, especially in the receptor-binding motifs, which can lead to decreased protein stability.

The number of mutations in SARS-CoV-2, particularly in the receptor-binding region (RBD) of the spike protein, is steadily increasing, surpassing 4,000 mutations¹. These mutations can give rise to variants with diverse characteristics, such as increased transmissibility and the ability to evade the immune system. However, the severity of the disease appears to be more related to host

factors than to the genetic variability of the virus. While some mutations in the spike protein may enhance infectivity and immune escape, they do not seem to affect the severity and pathogenicity of COVID-19².

The WHO has identified variants of concern (VOCs), which exhibit increased transmissibility or virulence and negatively impact the epidemiology and clinical presentation of the disease (figure 10). The agency has implemented measures to address VOCs, including the sharing of complete genomic sequences and field investigations. Some mutations, such as E484K and N501Y, have been associated with a reduction in vaccine efficacy, necessitating updates to vaccination strategies and the development of booster shots to combat new variants¹⁰³.

The receptor-binding domain (RBD) is crucial for binding to ACE2 in host cells. Mutations like N439K and V483A in the RBM of the spike protein influence the binding affinity for ACE2, reducing the neutralizing activity of antibodies and promoting immune escape.^{104–106} Other mutations, such as D614G, also increase the virus's infectivity and enhance its stability.^{107,108} Mutations in non-structural proteins, such as Nsp1, can disrupt the host's innate immune responses, further contributing to immune evasion.¹⁰³

To date, vaccination has been the most effective medical intervention against COVID-19, but viral mutations can reduce the neutralization capacity of antibodies and compromise vaccine efficacy. Recent meta-analyses have highlighted the impact of VOCs on the effectiveness of widely used vaccines, such as Pfizer and Moderna (figure 11).¹⁰⁹ The decline in vaccine efficacy poses a significant public health challenge due to the continued persistence and recurrence of COVID-19 outbreaks. Comparative assessments of antibody neutralization in the sera of vaccinated individuals have been conducted to monitor vaccine efficacy against circulating variants.¹¹⁰

In conclusion, the continuous evolution of SARS-CoV-2 through mutations and the associated immune escape necessitate ongoing monitoring and the adaptation of vaccination strategies to address emerging challenges related to the pandemic.

Vaccine	Developers and Country	Vaccine Type	Doses	Injection Interval *	Vaccination Restrictions **
Comirnaty	Pfizer-BioNTech (New York City, NY, USA and Mainz, Germany)	mRNA	2	21 days	All individuals aged 6 months and above
Spikevax	Moderna, NIAID (Cambridge, MA, USA)	mRNA	2	28 days	All individuals aged 6 months and above
Johnson & Johnson vaccine	Janssen Pharmaceuticals Johnson & Johnson (New Brunswick, NJ, USA)	Viral vector	1	-	18 and above
Vaxzevria	AstraZeneca, University of Oxford (Cambridge, UK)	Viral vector	2	28 days	18 and above
CONVIDECIA	CanSino Biologics Inc. (Tianjing, China)	Viral vector	1	-	18 and above
CoronaVac	Sinovac (Beijing, China)	Inactivated (Vero Cells)	2	14 days	18 and above; Further assessment for pregnant women is needed
BBIBP-CorV (Vero Cells)	Sinopharm (Beijing, China)	Inactivated (Vero Cells)	2	21 days	18 and above; Further assessment for pregnant women is needed
BBV152/COVAXIN	Bharat Biotech (Hyderabad, India)	Inactivated	2	28 days	18 and above; recommended for pregnant women only when the benefits of vaccination outweigh the potential risks
VLA2001	Valneva SE (Saint-Herblain, France)	Inactivated	2	28 days	18 to 50 years
Nuvaxovid	Novavax (Gaithersburg, MD, USA)	Protein Subunit vaccine	2	21 days	12 and above; recommended for pregnant women only when the benefits of vaccination outweigh the potential risks

Figure 10. WHO-certified vaccines against COVID-19.¹⁰³

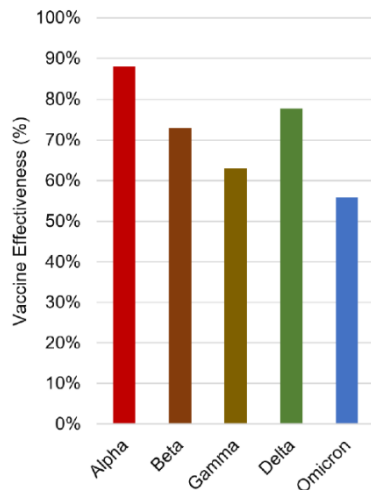


Figure 11. The impact of VOCs on vaccine effectiveness.¹⁰³

SARS-CoV-2 Life Cycle

SARS-CoV or SARS-CoV-2 infection involves the binding of the viral S protein to ACE2 on the host cell surface. The receptor binding domain on the S1 surface subunit of the S protein is responsible for the attachment of the virus to ACE2. After binding, the S protein is cleaved at the two cut-off sites S1/2 and S2' (in a process known as S protein priming) by the transmembrane serine protease TMPRSS2, which in turn facilitates the fusion of the viral membrane with the host cell membrane and the direct entry of the virus into the cytoplasm (Figure 12).^{94,111,112} Respiratory tract epithelial cells express both ACE2 and TMPRSS2 on their surface, and this direct or 'early' entry pathway appears to be the predominant mode of entry of SARS-CoV-2 in vivo and probably into respiratory tissue. Alternatively, SARS-CoV-2 may also utilise an endosomal entry pathway, whereby the ACE2-virus complex is translocated to endosomes and S-protein priming is performed by the endosomal cysteine proteases cathepsin B and cathepsin L, after which the virus is released from the endosome into the cytoplasm. This endosomal entry pathway, which can be blocked by lysosomotropic agents (such as hydroxychloroquine) or cathepsin inhibitors, may be the predominant entry pathway used by coronaviruses in the infection of cultured cells in vitro, but the importance of this pathway for in vivo infection remains unclear. After release of viral genomic RNA into the cytoplasm, the first ORF is translated into the polyproteins pp1a and pp1ab, which are then cleaved by viral proteases into small non-structural proteins such as RNA-dependent RNA polymerase (RdRP). The viral genomic RNA is then replicated using the viral RdRP and the four structural proteins (S, E, M and N) are translated through the endoplasmic reticulum and Golgi complex of the host cell. Finally, the genomic RNA and structural proteins are assembled into new viral particles, leading to their release via exocytosis (Figure 12).^{111,112} Each step of the viral life cycle described here is a potential therapeutic target, including S-protein priming by TMPRSS2 (target of the serine protease inhibitor camostat mesylate), membrane fusion and endocytosis (target of the antimalarial drug chloroquine and the anti-influenza drug

umifenovir) and RNA replication by RdRP (target of the antiviral agents favipiravir, remdesivir and ribavirin).^{113,114}

ALTERNATIVE GATEWAYS INTO THE HOST CELL

Another interesting feature of CoV-2-S is the presence, again in the RBD but upstream of the interaction domain with ACE2, of a specific amino acid sequence, called the RGD motif (R = arginine; G = glycine; D = aspartic acid) recognised by another class of receptors, the integrins¹¹⁵. The RGD motif is unique in CoV-2 among β -coronaviruses, whereas it is present in human adenoviruses and metapneumoviruses, which use integrins as entry pathways into cells. This sequence may have evolved as an alternative entry mechanism for the virus or to broaden its tropism. Furthermore, it has been shown that CD147, used by SARS-CoV as an alternative entry receptor, could also be used by CoV-2, as an antibody directed against CD147 prevents CoV-2 from entering cells¹¹⁶.

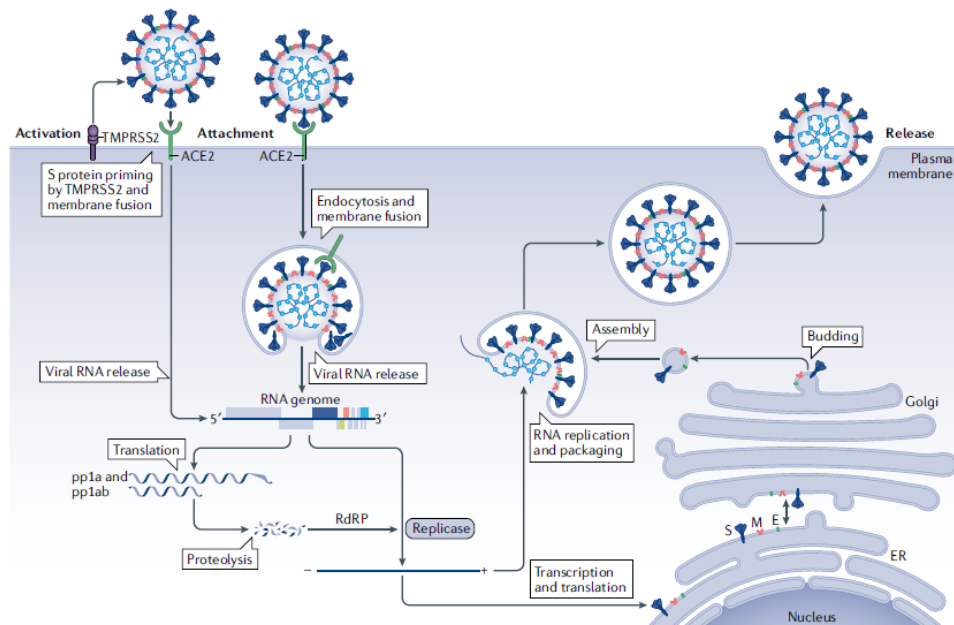


Figure 12. Life cycle of SARS-CoV-2¹⁰².

3. GENETIC SUSCEPTIBILITY FACTORS: WHAT WE KNOW SO FAR

Coronavirus disease 2019 (COVID-19) is associated with very heterogeneous clinical pictures, ranging from asymptomatic or paucisymptomatic infection to severe respiratory failure and death. Advanced age is a well-established risk factor, as are male gender and comorbidities such as hypertension and diabetes. As these risk factors only partially explain the broad spectrum of disease manifestations, the study of the genetic factors involved in the host response to SARS-CoV-2 is crucial for understanding the pathogenetic basis of the disease as well as the wide phenotypic variability associated with it.

3.1. Genetic Susceptibility Factors: Virus Entry

As mentioned above, as far as virus-associated genetic factors are concerned, the majority of mutations involve the S protein binding motif to the host cell receptor. More than 4000 mutations have been identified in the SARS-CoV-2 genome¹, none of which, however, have been shown to have significant functional implications¹¹⁷. Some mutations in the S-glycoprotein appear to increase transmissibility and immune escape of the virus but not the severity of clinical manifestations associated with infection. On the other hand, the most numerous and suggestive literature evidence on genetic factors of host response concerns critical genes involved in the mechanism of virus entry into the host cell together with those involved in the innate immune response. Variants in these genes have indeed been associated with different disease progression¹¹⁸. In fact, initial research focused on genetic variants encoding the proteins involved in viral entry into cells²⁵.

ACE2

The *ACE2* gene (Angiotensin Converting Enzyme 2, OMIM*300335) consists of 18 exons and is located on the X-chromosome (Xp22.2). The protein consists of three domains: an N-terminal peptidase domain (residues 19-615), a C-terminal collectrin-like domain [(CLD), residues 616-768], an end with a hydrophobic transmembrane region and an intracellular segment of 43 residues. Genetic variants of *ACE2* are described as a potentially influential factor in susceptibility to SARS-CoV-2.

A number of studies have examined the variability of *ACE2* in different populations. Studies, such as that of Hou et al., identified 63 potentially deleterious *ACE2* variants that could influence genetic susceptibility to COVID-19. These variants are more frequent in African/African

American populations (39%) and in non-Finnish European populations (54%), while they are less common in Asian and Latin American populations.¹¹⁹ Other studies such as the one conducted by Cao et al., however, found no significant differences that could influence virus binding to ACE2¹²⁰.

The rs2285666 variant, which could alter *ACE2* expression, is more frequent in European and African populations. The rs182366225 and rs2097723 variants have been associated with upregulation of *ACE2*, which is more common in East Asian populations, and could influence an increased susceptibility.¹²¹ Specific variants, such as p.(Lys26Arg), p.(Gly211Arg), and p.(Asn720Asp), were more frequently found in the Italian population than in the Asian population and could explain the higher mortality rates in Italy during the onset of the pandemic.¹²²

Several studies have examined whether these variants can influence the ability of the virus to bind to the ACE2 receptor and, thus, the infection and severity of the disease¹²³. Functional variants with an impact on the protein were detected. Some of the most relevant variants include: c.1517 T > C p.(Val506Ala), which has a major effect on protein stability, and c.631G > A p.(Gly211Arg) and c.77A > G p.(Lys26Arg), which could affect the virus' ability to bind to ACE2.¹²⁴ Fujikura and Uesaka identified eight variants, including S19P, E35K, E37K, that affect the contact residues between the SARS-CoV-2 spike protein and the human ACE2 receptor¹²⁵. Variants such as rs73635825 and rs143936283, which have a low binding affinity for the SARS-CoV-2 spike protein, could be associated with possible resistance to infection¹²⁶. Other variants such as p.(Leu584Ala), p.(Pro389His) and p.(Leu351Val) could interfere with the binding of the SARS-CoV-2 spike protein, potentially altering transmissibility and severity^{122,124}.

Some studies suggest that specific genetic variants, such as rs1476524 and rs2048683, may be associated with increased tissue expression of ACE2 and increased disease severity, including hospitalisation¹²⁷. In addition, the rs2106809 variant has been shown to be associated with circulating levels of ACE2, with the CC/CT genotype having higher levels than the TT genotype.¹²⁸

Some studies have examined gender differences, showing that differences in ACE2 expression between men and women are partly attributed to the gene's location on the X chromosome. This may explain why men tend to suffer from more severe forms of COVID-19, as women, having two X chromosomes, may be more protected due to heterozygosity.¹²⁹

Devaux and collaborators highlighted the role of ACE2 adaptation in SARS-CoV-2's evolution and the emergence of new variants of the virus. The emergence of key SARS-CoV-2 mutations (e.g., N501Y, P681H, K417N) is discussed in the context of human *ACE2* variants. In particular, the rare rs143936283 (E329G) ACE2 variant found in some European populations may confer

resistance to SARS-CoV-2 by weakening spike-ACE2 binding. This variant may have influenced the selection of viral mutations like N501Y. Molecular docking studies reveal that the E329G substitution in *ACE2* enhances interaction with the N501Y spike mutation, suggesting that *ACE2*-driven selection could have contributed to the emergence of variants like the Alpha (B.1.1.7) variant, especially in regions with this rare *ACE2* allele.¹³⁰

RELATIONSHIP BETWEEN GENETIC VARIANTS OF ACE2 AND IL-6 LEVELS

A strong correlation has been observed between interleukin-6 (IL-6) levels and the severity of COVID-19, and how this relationship is influenced by genetic variants of *ACE2*.¹³¹ IL-6 is a known biomarker of inflammation and the phenomenon of ‘cytokine storm’, an exaggerated immune reaction that can lead to systemic damage and worsening of clinical conditions. IL-6 levels were significantly correlated with disease progression, with patients with severe and critical cases having much higher IL-6 levels than mild or moderate cases. It was observed that patients carrying the CT or TT genotype of the rs2285666 *ACE2* polymorphism tended to have higher IL-6 levels than carriers of the CC (wild-type) genotype. Notably, the TT genotype was associated with increased disease severity and IL-6 overproduction, suggesting that these *ACE2* gene variants may indirectly influence the exacerbated inflammatory response.

These data provide clear evidence of the link between genetics and inflammation, suggesting that specific genetic variants may modulate the severity of COVID-19 through increased IL-6 production.

IL6

An association between genetic variants of the interleukin-6 (*IL6*) gene and the severity of COVID-19 has been reported in some works.

In particular, the polymorphism on the *IL6* promoter region rs1800795 (also known as -174G > C) was studied in relation to the severity of COVID-19. This polymorphism has been associated with an increased production of IL-6, which may contribute to the so-called ‘cytokine storm’.¹³² Previous studies had already shown that this variant is linked to severe outcomes of other viral infections such as influenza and respiratory syncytial virus (RSV)¹³³, and an analysis suggests that it may also influence the response to SARS-CoV-2¹³⁴. Another group studied the same polymorphism in a Moroccan population and concluded that individuals with the G/C genotype were less likely to develop severe forms of COVID-19 than those with the G/G or C/C genotypes. This suggests that certain genetic variants of *IL6* may also offer protection against progression to severe forms of the disease¹³⁵.

Furthermore, another study in paediatric patients showed an association between specific *IL6* polymorphisms and systemic inflammation, demonstrating that the severity of COVID-19 may also be related to these genetic variants in children. Children who expressed GC genotypes of

IL6 (-572G > C) polymorphism are at a considerably higher risk of developing a severe disease. This risk is significantly larger in the severe group of children than in children in critical condition who have GC genotypes of *IL6* (-174 G>C) polymorphism. While *IL6* (-597G >A) polymorphism has no role in COVID 19 severity in children¹³⁶.

However, research is still ongoing and other genetic and environmental factors contribute to the complexity of the response to the virus, including inflammatory markers such as C-reactive protein (CRP) and D-dimer, which have been associated with disease severity in several studies¹³⁵.

TMPRSS2

The *TMPRSS2* gene (Transmembrane Serine Protease 2, OMIM*602060) consists of 14 exons and is located at the level of the long arm of chromosome 21 (21q22.3). Its expression is regulated by androgens. *TMPRSS2* is mainly expressed on the luminal side of the prostate epithelium and is over-expressed in prostate cancer tissues¹³⁷. The protein has predominantly proteolytic activity but has also been shown to play an important role in facilitating the fusion of viral pathogens (influenza, CoV) to the host cell membrane. *TMPRSS2* thus represents a further factor involved in the S-glycoprotein priming of SARS-CoV-2.

The question was therefore raised as to whether the higher mortality rates observed in SARS-CoV-2 infected males were due to androgen-dependent *TMPRSS2* expression, but no differences in expression levels between the two sexes were observed in lung tissue¹³⁸. It was therefore hypothesised that underlying individual variability in susceptibility to SARS-CoV-2 infection might be genetic variability of *TMPRSS2* affecting expression levels, protein structure and function.

Several *TMPRSS2* polymorphism (rs12329760, rs17854725, rs12329760, rs2070788, rs2298659, rs456298, and rs462574) have been associated with COVID-19 severity.^{139–141} However, there are conflicting results and opinions on some of these variants. Iranian patients with severe COVID-19 appear to have the T allele of rs12329760 in *TMPRSS2* gene. This is contrary to most previous studies on this variant¹⁴². In another study of a Ukrainian population, the rs12329760 polymorphism in *TMPRSS2* was also associated with COVID-19 severity¹⁴³. The results of another study by Rokni et al.¹⁴⁰ in Iran suggested that the genotypes rs17854725 A > G (AA vs. AG and AA vs. GG), rs12329760C > T (CC vs. CT, CC vs. TT), and rs4303795 A > G (AA vs. AG) are associated with COVID-19 invasiveness. In contrast to previous studies, David et al.¹⁴⁴ in UK found *TMPRSS2* rs12329760 as a variant significantly associated with decreased severity of COVID-19.

TMPRSS2 polymorphisms rs2070788, rs383510, and rs12329760 were genotyped in 239 patients with SARS-CoV-2 and 253 patients SARS-CoV-2 negative. It was found, however, that neither the *TMPRSS2* rs2070788 nor the rs12329760 polymorphism were associated with COVID-19

infection risk. In an Indian population, instead, rs2070788 (homozygote mutated) was associated with high levels of the protein¹⁴⁵ and, in Brazilian hospitalized subjects increases the chance of death from COVID-19 4-fold in older adult patients¹⁴⁶. In a German cohort, the intron variant rs383510 in the gene *TMPRSS2* was associated with an increase in SARS-CoV-2 infection risk.¹⁴⁷

Indeed, in a study in Egypt, *TMPRSS2* rs12329760 and *ACE2* rs2285666 gene polymorphisms, in addition to routine cell and serum markers, provide a reliable prediction of COVID-19 disease severity, which accounts for 86 % of the severity owing to the increased expression of genotypes of *ACE2* rs2285666 and *TMPRSS2* rs1232976 (TT), CT + TT, and (T) alleles among severe COVID-19 groups compared to mild and control¹⁴⁸. *ACE2* rs2285666 was also associated with high circulant levels of IL6, as mentioned above (see “relationship between genetic variants of *ace2* and *il-6* levels”)

These variants, in combination with other genetic factors such as those in the *ACE2* gene, could significantly modulate the susceptibility and severity of COVID-19, influencing the inflammatory response and the clinical course of the disease.

FURIN

The SARS-CoV-2 spike protein contains a multibasic S1/S2 site at which the host cell's furin protease is able to exert its cleavage function⁹⁴. Bestle and co-workers (2020) demonstrated that *TMPRSS2* and furin are both essential for proteolytic activation of SARS-CoV-2, thus inhibition of these two proteases is indicated as a therapeutic approach for the treatment of Covid-19¹⁴⁹. Furthermore, it has been suggested that plasma furin levels are elevated in patients with diabetes mellitus, which could explain the susceptibility of diabetic patients to Covid-19¹⁵⁰. Six genetic variants of the *FURIN* gene (Furin, Paired Basic Amino Acid Cleaving Enzyme, OMIM*136950) have been identified in a small cohort of Covid-19, but none of them were associated with the disease¹⁵¹. Klaassen and co-workers studied the protease coding regions in a small Serbian cohort and identified two variants, p.(Gly146Ser) and p.(Thr33Ala) for which *in silico* analyses suggested an effect on the final structure and/or function of the protein¹⁵².

Intronic (c.372+5G>A) and missense variants (c.128C>T, p.Ala43Val; c.436G>A, p.Gly146Ser; c.893G>A, p.Arg298Gln; c.1906A>G, p.Ile636Val) of Furin have been identified in SARS-CoV-2-positive patients in Italy. The allele frequency of these variants was very low and a statistically significant difference was observed for the c.893G>A, p.(Arg298Gln) variant when compared to the overall allele frequency of the European population¹⁵³.

Some studies have suggested that polymorphisms in *FURIN* might influence the risk and severity of COVID-19 infection. For instance, variants such as rs6226 and rs4702 have been identified in genetic analyses as potential factors associated with a higher risk of developing severe forms of COVID-19. However, the evidence regarding the precise impact of these variants is still not

completely clear and further studies are needed to better understand their role in the disease^{154,155}. Dobrindt and colleagues also synthesised the mutated Furin-rs4702 construct using CRISPR (GG) in lung cells and neurons and then studied the effects of this genetic variant on SARS-CoV-2 infection¹⁵⁶. GG rs4702 lung cells actually showed a decrease in furin expression and SARS-CoV-2 infection. Both AA and GG neurons were then infected with SARS-CoV-2, showing a trend towards decreased SARS-CoV-2 infection in GG neurons. This study clearly shows that a single non-coding variant within the *FURIN* gene locus can influence SARS-CoV-2 infection.

HSPA5

Genetic variants of the HSPA5 gene (which codes for the protein known as GRP78, involved in the cellular stress response) have been studied in relation to the severity of COVID-19. However, unlike other genes such as ACE2 or TMPRSS2, direct evidence linking specific HSPA5 variants to the severity of SARS-CoV-2 infection is more limited. Some studies suggest that HSPA5 might play a role in the interaction between the virus and the host cell. In particular, it has been hypothesised that the GRP78 protein, which is highly expressed during cellular stress, may facilitate virus entry into cells. However, most research has focused on the functional role of the protein rather than its genetic variants. Studies indicate that *HSPA5* is expressed at significantly higher levels in various cancers, which correlates with increased susceptibility to SARS-CoV-2 infection. For instance, in lung cancer, *HSPA5* mRNA levels were found to be 253-fold higher than ACE2, suggesting that tissues with elevated HSPA5 expression are more vulnerable to viral invasion.¹⁵⁷ Furthermore, a systematic analysis revealed that COVID-19 patients with cancer exhibited a higher rate of severe outcomes (33.33%) compared to non-cancer patients (16.09%).¹⁵⁷ Additionally, genetic studies have identified GRP78 among proteins associated with COVID-19 severity, highlighting its potential as a therapeutic target.^{158,159}

However, it is essential to consider that while *HSPA5* may be a critical factor, other genetic and immune response variables also play significant roles in determining COVID-19 severity, indicating a complex interplay of factors influencing patient outcomes.

CTSB and CTSL

Genetic variants of the *CTSB* and *CTSL* genes, which code for cathepsins B and L respectively, appear to have a potential role in the severity of COVID-19, although research on these specific genes is not yet extensive. Cathepsins B and L are lysosomal proteases involved in the entry of the SARS-CoV-2 virus into human cells. Studies have shown that the virus can enter cells through proteolytic cleavage of the spike protein, which is facilitated not only by the TMPRSS2 protease, but also by cathepsins B and L. These proteins are crucial because they allow the virus to enter the cytoplasm of cells after fusion with the membrane.^{160,161}

However, there is currently little concrete evidence linking genetic variants in the *CTSB* and *CTSL* genes directly to the severity of COVID-19. A study that analysed the expression of these genes showed that cathepsin B is highly expressed in various tissues, including lymph nodes and bone marrow, and is particularly present in immune cells such as monocytes and macrophages. Cathepsin L, on the other hand, is expressed in various organs, with high levels in lung and liver tissues, which could be relevant to the way the virus affects these organs in severe patients.¹⁶¹

Despite the biological importance of *CTSB* and *CTSL* in the viral infection mechanism, there are still no large genetic association studies that have identified specific variants of these genes as strongly correlated with COVID-19 severity. Research in this field is ongoing, and a more detailed understanding of their role may emerge in the future.

3.2. Genetic Susceptibility Factors: Immune components, epigenetics and AB0 loci

Several GWAS have confirmed that genetic predisposition plays a role in COVID-19 susceptibility and severity^{162,163}. In addition to *ACE2* and the others gene involved in the virus entry route, significant associations were found with genes such as *SLC6A20*, *JAK1*, *IRF1*, *TLR7*, and *FOXP4*. A major European study identified associations with *LZTF1*, *ABO*, *TYK2*, *MAPT*, and other genes¹⁶⁴.

A more recent meta-analysis, based on over 200,000 cases and 3 million controls, identified 51 distinct genome-wide loci associated with COVID-19, expanding understanding of key biological pathways such as virus entry, mucus defense, and the role of interferons¹⁶⁵. Contrary to expectations, GWAS studies did not find significant associations with HLA alleles, despite their known role in viral infections by other coronaviruses, such as SARS-CoV-1 and MERS-CoV^{166,167}. However, a common variant, HLA-B*15:01, was found to be associated with asymptomatic SARS-CoV-2 infections, suggesting a protective effect in about 10% of individuals of European descent who carry this variant¹⁶⁸.

The candidate gene approach has also been widely used to confirm GWAS findings, providing valuable information on specific genes involved in disease pathogenesis^{169,170}. For example, impairment of the *CFTR* gene has been linked to COVID-19 severity and has been shown to interact synergistically with estrogens¹⁷⁰. Genes involved in innate immunity and the interferon pathway are now considered critical to SARS-CoV-2 susceptibility¹⁷¹. The recognition of viral pathogen-associated molecular patterns (PAMPs) activates signaling pathways that lead to interferon production and inflammatory responses¹⁷².

SARS-CoV-2 infection, as reported earlier, triggers a considerable production of pro-

inflammatory chemokines and cytokines. A review published in February 2021¹¹⁷ analysed the genetic component of IFN-alpha, interleukins and Toll-like receptors.

With regard to *IFN-alpha*, potentially relevant variants were identified in the cytokine itself as well as in the IFNAR receptor and IFN-inducible transmembrane proteins (IFTM). In addition, alterations in the coding regions of *OAS1* (p.Arg47Gln, p.Ile99Val and p.Arg130His) have been suggested as predictive markers of the inter-individual difference in the response to Covid-19¹⁵².

Loss-of-function variants of the *TLR7* gene (c.2129_2132del; p.Gln710Argfs*18; c.2383G>T; p.Val795Phe) have been reported in a case series of four young men from two unrelated families, admitted to the intensive care unit due to critical Covid-19¹¹⁷. Studies by the Covid Human Genetic Effort Consortium have identified rare variants of the *TLR7* and *TLR3* genes that confer susceptibility to severe COVID-19. Mutations in these genes, especially in homozygous form, increase the risk of hypoxemic pneumonia and reduce type I interferon production, a key element of the immune response^{173,174}. The link between *TLR7* variants and drugs like N-acetylcysteine (NAC) could open new therapeutic avenues, as NAC may inhibit NF-κB activation and reduce inflammation.

Interestingly, five of the 24 loci identified by GWAS as critical for COVID-19 susceptibility are linked to the type I interferon pathway, including *TYK2*, *IFNAR1*, and *OAS1*^{175,176}. However, genetic variants found through GWAS do not always overlap with rare variants identified through other approaches, as seen in breast cancer susceptibility research^{177,178}. Rare variants are increasingly recognized as important for understanding the mechanisms of complex diseases and represent promising targets for drug development and biomarkers for disease risk assessment^{179,180}.

Common genetic variants appear to have only modest effects, explaining only a small fraction of clinical variability¹⁸¹. Therefore, the search for rare variants that confer a stronger susceptibility to severe COVID-19 should be extended to population subgroups and individual families to better understand the genetic dynamics of the disease¹⁸².

In 2022, a group explored the relationship between aging, epigenetic age, telomere length, and COVID-19 severity. While aging isn't conclusively linked to COVID-19 susceptibility, patients with severe symptoms show a delayed acceleration in GrimAge (an epigenetic clock trained on mortality and respiratory-related factors), potentially due to genetic variations like rs35081325. Indeed, severe COVID-19 is associated with accelerated telomere shortening, aligning with studies that show telomere length decreases with age and severity of disease. However, these results are limited, and further research is needed to incorporate a wider range of aging biomarkers¹¹⁸.

Twenty-five susceptibility loci for SARS-CoV-2 infection and/or COVID-19 have been identified

through GWAS, with the *ABO* gene on chromosome 9q34.2 being the most frequently replicated finding. Contrary to the hypothesis that blood group A individuals have higher susceptibility due to SARS-CoV-2 RBD binding, studies favor the ABO-compatibility-dependence model, which shows that natural anti-AB antibodies may reduce SARS-CoV-2 transmission risk by up to 41% in blood group O individuals^{183–186}. Finally, the impact of blood groups on SARS-CoV-2 susceptibility regarding reinfections and newer variants remains to be fully understood. Future studies examining these factors may provide deeper insights into the relationship between ABO blood groups and COVID-19 severity¹⁸⁷.

4. MATERIALS AND METODHS

4.1 DNA manual extraction

The blood is taken in a Vacutainer® tube containing the anticoagulant EDTA or citrate. The vials must be stored until they are extracted at -20. It is essential, of course, to be very careful and precise in the various steps of the operation, since a good sample quality is essential to ensure reproducibility and reliability of the results of subsequent analyses.

Qiagen FlexiGene® DNA Kit

The FlexiGene" Kit ensures good yield and high DNA purity, free from contaminants and inhibitors for subsequent PCR and sequencing reactions. In addition, it provides a quick and simple method of purification of DNA from whole blood. The procedure is carried out starting from 0.5 ml of starting material and performed in a single Eppendorf cone, in order to minimize waste and the risk of sample exchange. Once extracted, the product can be stored at 2-8°C (for short periods) or at -20°C (for longer time). The lysis buffer is added to the sample, the cell nuclei and mitochondria are deposited to form pellets through the centrifugation process. The pellet is then resuspended and incubated in a denaturing buffer, which contains a caotropic salt (Guanidine hydrochloride) and QIAGEN protease; this step is useful for removing contaminants such as proteins. DNA is precipitated by the addition of isopropanol, recovered by centrifugation, and then washed in 70% entanol. The DNA is then dried and resuspended in hydration buffer (10mM tris.Cl, pH 8.5).

Protocol:

- Pipette 1.25 ml of Buffer FG1 into a 2 ml eppendorf cone and add 0.5 ml of whole blood; invert the vial at least 5 times. Centrifuge for 1 min at 12'000 x g.
- Discard the supernatant and leave the cone inverted on top of a piece of clean paper towel for 2 min, taking care that the pellet remains adhered to the bottom of the cone.
- In a 1.5-mL Eppendorf cone, prepare a mix consisting of the Protease enzyme (needed to break peptide bonds of proteins) and FG2 reaction buffer. Dispense 250 µl of the FG2/Protease solution into the vial, close and vortex vigorously until the pellet is homogenized and detached from the bottom.
- Invert the cone 3 times, place it inside the bath and incubate at 65°C for 5 min (the sample will change color from red to olive green, an indication of successful protein digestion).

- Add 250 µl of Isopropanol (100%) and mix by inversion until the DNA precipitates and becomes visible as a white filamentous tangle.
- Centrifuge for 3 min at 12'000 x g.
- Discard the supernatant and invert the conet over a piece of clean blotting paper, taking care that the pellet does not slide off.
- Add 250 µl of 70% ethanol to remove any contaminants and vortex for 5s.
- Centrifuge for 3 min at 12'000 x g.
- Discard the supernatant and leave the inverted cone on clean blotting paper for at least 5 min making sure that the pellet remains adhered to the bottom.
- Allow the pellet to dry by leaving the conet open to the air until all the liquid has evaporated (at least 5 min). Be careful not to over-dry the pellet, otherwise it may not dissolve completely with the addition of the elution Buffer (FG3).
- Add 200 µl of FG3 Buffer, vortex for 5s and dissolve the DNA by incubating it for 1 h in the thermostat bath at 65°C (if the DNA is not dissolved, it is advisable to leave the solution overnight at room temperature).

Qualitative and quantitative evaluation using Nanodrop

A good concentration and quality of the extracted sample are essential. The quickest and sufficiently reliable way to ascertain this is to use the *NanoDrop 1000 Spectrophotometer*, which through the ND1000 v.3.3.0 software analyzes DNA. The determination of concentration follows the Lambert-Beer law:

$$C = \frac{OD_{260} \varepsilon}{L}$$

Where:

OD₂₆₀ is the absorbance at 260 nm, defined as the log of the ratio of I₀ (input radiation intensity) to I (output radiation intensity).

ε is the wavelength-dependent molar extinction coefficient.

L is the optical path.

The instrument, for quality control, provides two parameters: the A₂₆₀/A₂₈₀ ratio, which is an indication of the purity of the nucleic acid (if the value is around 1.8, then the DNA is pure and free of proteins) and the A₂₆₀/A₂₃₀ ratio, which detects the possible presence of contaminants such as aromatic compounds, phenols, carbohydrates or peptides, the optimal value of which is about 2.2 (remember that the absorption wavelength of nucleic acids is 260nm while the absorption wavelength of proteins is 280nm). In order to hold the maximum absorbance peak of

nucleic acids at 260 nm, the spectrophotometer repeatedly scans the sample over a wavelength range of 220 nm to 350 nm. To use it initially, the instrument requires a step with distilled water to clean the pedestal on which the different materials (H₂O, buffer and samples) will be arranged; then the blank (the elution buffer FG3), must be supplied. The sample measurement is automatically calculated by deducting the blank value from the sample absorbance value.



Figure 13. NanoDrop ND-1000 Spectrophotometer

DNA quantification with Qubit® fluorometer

After the Nanodrop measurement, use Qubit® Fluorometric Method to accurately quantify the DNA starting material. Prepare the working solution by diluting the Qubit® dsDNA HS Reagent with a 1:200 ratio in the Qubit® dsDNA HS Buffer. Add 190 µL of the working solution into the cones used for standards, then add 10 µL of Qubit® standard into the respective cone, mix by vortexing for 2-3 seconds taking care not to form bubbles. Prepare the cones for the samples keeping in mind that the sum of the volumes between working buffer and sample should be exactly 200 µL; therefore, the sample should be between 1-20 µL and the corresponding volume of the working buffer between 180-199 µL; mix by vortexing for 2-3 seconds. Leave the cones to incubate for two minutes at room temperature, then go to the Qubit® fluorometer to read the standards first and then the samples.



Figure 14. The Invitrogen Qubit 4 Fluorometer

4.2. Gene panel design

The panel was constructed by selecting genes that, based on literature evidence and biological plausibility, appear to be involved in cellular mechanisms that regulate anchoring, entry and life cycle of SARS-CoV-2 virus within the host cell;

The chosen genes are shown in Table 1.

Table 1. Gene selected for the panel.

Gene	Protein	Cromosomal position hg19	Transcript	OMIM
ACE2	Angiotensin I converting enzyme 2	chrX:15,579,156-15,620,192	NM_021804	300335
TMPRSS2	Transmembrane protease, serine 2	chr21:42,836,478-42,879,992	NM_001135099	602060
CTSL	Homo sapiens cathepsin L	chr9:90,340,974-90,346,384	NM_001257972	116880
CTSB	Homo sapiens cathepsin B	chr8:11,700,034-11,725,646	NM_001908	116810
FURIN	Furin	chr15:91,411,885-91,426,687	NM_002569	136950
HSPA5 (GRP78)	Heat shock 70kDa protein 5	chr9:127,997,127-128,003,666	NM_005347	138120
IL6	Interleukine-6	chr7:22,766,766-22,771,621	NM_000600	147620

DNA libraries of the 80 COVID-19 patients were constructed using the Haloplex Hs Target Enrichment System and target sequencing was conducted through MiSeq Illumina Protocol.

4.3. Haloplex Hs Target Enrichment System

Library preparation

In every NGS protocol, preparation of the DNA library, which contains thousands of copies of the selected regions in the panel of interest, is required. This sequencing library was made through Illumina's MiSeq platform, using the HaloPlex HS Target Enrichment System kit from the Agilent company. The protocol consists of four basic steps (Figure 15).

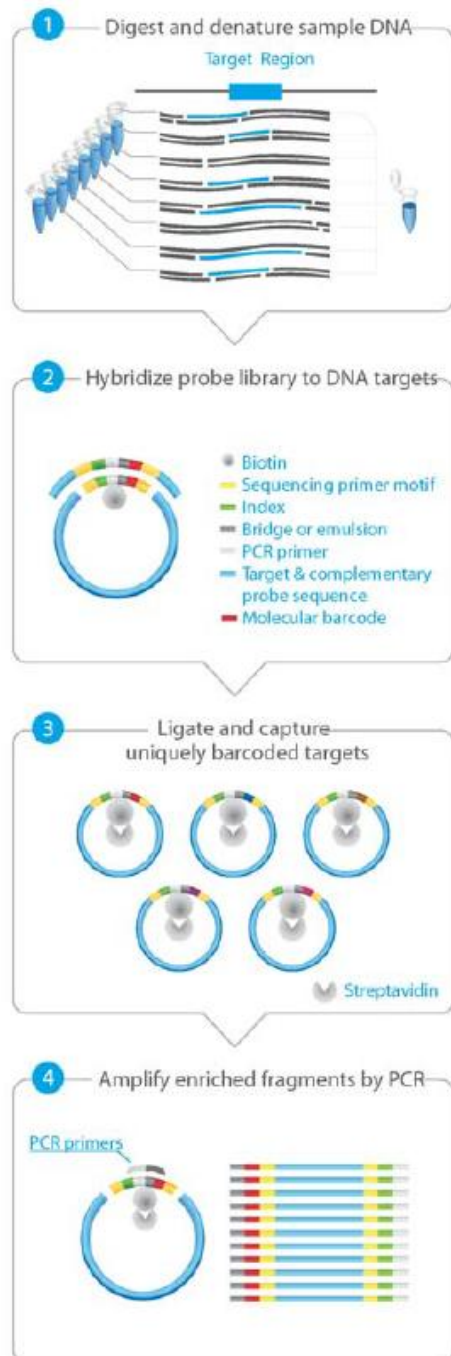


Figure 15. Overall HaloPlex HS target-enriched sequencing sample preparation workflow.

Step 1. Digest genomic DNA with restriction enzymes

In this step, gDNA samples are digested by 16 different restriction enzymes to create a library of gDNA restriction fragments.

1- Prepare the DNA samples for the run. For 12-reaction runs, prepare 11 gDNA samples and one Enrichment Control DNA sample (each sample is diluted to a concentration of 1.8 ng/μl in TrisHCl 10 mM buffer at pH 8.5).

2- Prepare restriction enzyme master mix: combine RE buffer and BSA solution in a 1.5ml cone as shown in the table. Then vortex.

Reagent	Volume for 1 reaction	Volume for 12 reactions (including excess)
RE buffer	24,6μL	344μL
BSA solution	0,64μL	9μL
Total Volume	25,2μL	353μL

Table 2. Restriction enzyme master mix preparation.

3- Prepare Restriction Enzyme Master Mix strip: DNA is digested in eight different reactions by combining eight pairs of reaction enzymes. The 16 enzymes are provided in 2 eight-well strips marked with a green indicator and a red indicator, respectively. The enzymes from both strips are combined into one strip with the master mix. Figure 16 illustrates how to prepare the Restriction Enzyme Master Mix strip for 12 samples.

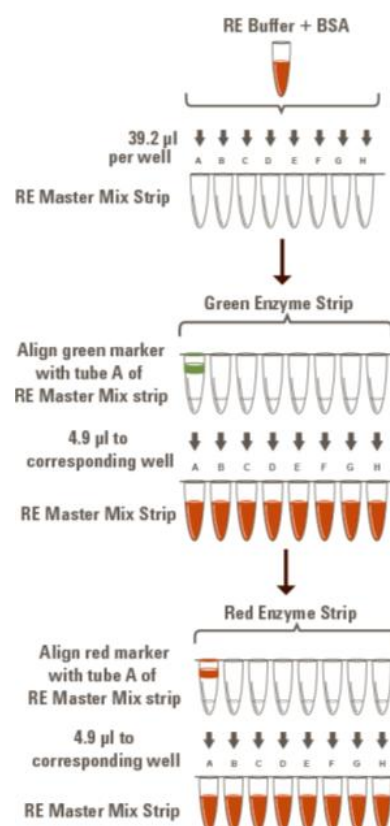


Figure 16. Restriction Enzyme Master Mix Strip preparation for 12 samples.

a) Dispense the volume of RE Master Mix shown in the table below into a new strip

Reagent	Volume for 1 reaction	Volume for 12 reactions (including excess)
RE buffer/BSA mixture	2,8 μ L	39,2 μ L

Table 3. RE Master Mix.

b) Use a multichannel pipette to add the appropriate volume of each enzyme from the green strip to the new strip with RE Master Mix. Repeat the same step for the red enzyme strip.

Reagent	Volume for 1 reaction	Volume for 12 reactions (including excess)
RE Enzymes from Green Enzyme Strip	0,35 μ L	4,9 μ L

Table 4. Required amounts of the green strip restriction enzymes for the digestion reaction.

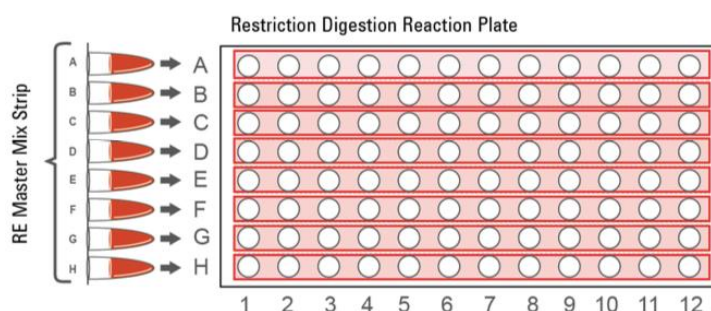
Reagent	Volume for 1 reaction	Volume for 12 reactions (including excess)
RE Enzymes from Red Enzyme Strip	0,35 μ L	4,9 μ L

Table 5. Required amounts of the red strip restriction enzymes for the digestion reaction.

c) Vortex and briefly spin the strip. Keep the strip on ice until it is used for the next step.

4- Aliquot the RE Master Mixes into a 96-well plate that will be used as the reaction plate for enzymatic digestion. Align the strip prepared in the previous steps vertically with respect to the 96-well PCR plate, as shown in figure 17.

Figure 17. Dispense the RE master mix into the reaction plate.



5- Carefully distribute 3.5µL of RE Master Mix into the wells of the plate using a multichannel pipette.

6- Aliquot 3.5µL of DNA prepared in step 1, into the wells of the reaction plate, arranging the strip horizontally with respect to the plate.

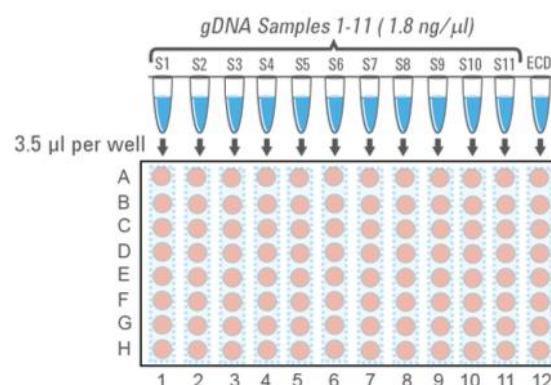


Figure 18. Arrangement of the 11 concentrated 1.8 ng/µL samples and ECD in the 96well reaction plate.

7- Cover the plate with adhesive protective film.

8- Vortex and briefly pin the plate.

9- Set the program given in the table in the thermal cycler.

Step	Temperature	Time
Step 1	37°C	30 minutes
Step 2	4°C	Hold

Table 6. Program to be set in the thermal cycler for the HaloPlex HS digestion reaction.

Meanwhile, it is advisable to take the Bioanalyzer reagents out of the refrigerator and let them equilibrate at room temperature.

10- Transfer 4µL of each ECD digestion reaction from the plate into 8 new 0.2mL PCR cones.

11- Incubate the freshly prepared ECD cones at 80°C for 5 minutes; this step is necessary to inactivate the restriction enzymes.

12- Freeze the plate of digested samples at -20°C until the experiment is continued.

13- Analyze the prepared samples using the Bioanalyzer 2100 evaluation method.

NOTE: ECD samples contain genomic DNA along with 800bp PCR products containing restriction sites for all enzymes used in digestion. When analyzing the results, the undigested control should have >2.5 kbp bands and the PCR products an 800bp band. Each of the 8 digested ECD samples must have a smear of DNA restriction fragments between 100 bp and 2500 bp, overlapping with the 3 predominant bands of 125, 225 and 450 bp. These 3 bands correspond to the PCR products derived from the restriction fragments.

VALIDATION WITH 2100 BIOANALYZER

The Bioanalyzer 2100 is an instrument equipped with platinum electrodes that enable qualitative and quantitative analysis of nucleic acids in sensitive applications such as NGS. The analysis involves reading samples with High Sensitivity DNA Kit and is based on capillary electrophoresis that takes place inside a particular chip. Each chip contains a set of tightly interconnected microchannels used for separation of nucleic acid fragments based on different molecular weight. At the end of the electrophoretic run, the samples are read according to their fluorescence, and this information is translated into electropherograms.



Figure 19. The Chip Priming Station and the BioAnalyzer 2100.

GEL DYE MIX PREPARATION:

- 1- equilibrate High Sensitivity DNA Dye concentrate and High Sensitivity DNA Gel Matrix reagents at room temperature for 30 min;
- 2- add 15 μ l of High Sensitivity DNA Dye concentrate to the vial of High Sensitivity DNA Gel Matrix;
- 3- vortex the solution well and centrifuge briefly; transfer to spin filter;
- 4- centrifuge at $2240g \pm 20\%$ for 10 min. Protect the solution from light and store it at 4°C .

“CHIP PRIMING STATION” SETTING:

- 1- set the syringe lock lever to the lowest position;
- 2- check that the base is in position “C”.

LOADING THE BIOANALYZER CHIP:

- 1- Transfer 9 μ l of Gel-Dye Mix into the Chip at the position marked by the black G (figure 20). You lower the syringe until it is wedged into the block. Start the timer for 60 seconds (time needed for the matrix to spread out on the Chip). Release the syringe block and let it rise spontaneously until it stops, at which point manually with unexaggerated speed you raise the syringe plunger to 1 ml;
- 2- load 9 μ l of Gel-Dye Mix into the other 3 G;
- 3- load 5 μ l of DNA marker into the other wells of the Chip;
- 4- load 1 μ l of DNA ladder into the indicated well;
- 5- load 1 μ l of sample into the remaining wells (row 1: samples 1-3, row 2: samples 3-6, row 3: samples 7-9, row 4: samples 10-11);
- 6- vortex the chip for 1 minute at 2400rpm;
- 7- check for bubbles in the Chip before loading it into the Bioanalyzer;
- 8- select from the computer dsDNA, High Sensitivity DNA assay, the number of samples loaded and their identifier. At this point select “start”
- 9- verify that the electropherogram shows a peak at the 245-325bp size range.

N.B. The Bioanalyzer is volume-sensitive so all wells must be loaded. If there are empty wells left because there are less than 11 samples, load an equal volume of markers in their place.

After analyzing the Bioanalyzer results, continue with the Haloplex HS protocol by resuming from step 2.



Figure 20. High Sensitivity DNA chip

STEP 2. HYBRIDIZATION OF DNA FRAGMENTS WITH HALOPLEX HS PROBES.

In this step, elements necessary for the identification of a specific DNA fragment are uniquely incorporated into the DNA fragments within the final pool: molecular barcodes, Illumina

sequencing motifs and also index sequences. Haloplex probes are designed to selectively hybridize fragments originating from target regions of the genome and to guide their circularization.

1- Prepare Hybridization Master Mix by combining Hybridization Solution and Haloplex HS Probe together in the manner shown in the table. Spin the individual reagents, then vortex and spin the resulting mix.

Reagents	Volume for 1 reaction	Volume for 12 reactions
Hybridization Solution	34µL	442µL
Haloplex HS Probe	5µL	65µL
Total Volume	39µL	507µL

Table 7. Hybridization Master Mix preparation.

2- In a new strip of 12 0.2-mL cones, distribute 39µL of Hybridization Master Mix into each cone.

3- Add 5µL of Haloplex HS Indexing Primer into the prepared cones. Indexers are provided by the kit in an orange 96-well plate.

4- Transfer the digested DNA samples from the 96well plate [step1.12] directly into the cones of the strip prepared in step 3. (After adding each digested DNA sample to the Hybridization Solution, mix by pipetting before adding the next one to ensure complete inactivation of the enzymes. Restriction enzymes, in fact, may still be active and catalyze inappropriate cleavages.) Spin the plate only if necessary. The final volume of each cone should be 100µL, so add about 56µL of sample. (For the ECD sample from the 12th cone instead, add 2.5µL of each digestion reaction and 36µL of nuclease-free water.)

5- Vortex the strip briefly and centrifuge. It is advisable to change the caps of the strip each time you go to incubate samples to avoid contamination.

6- Place the strip in the thermal cycler and start the hybridization reaction with the following program:

Step	Temperature	Time
Step 1	95°C	5 minutes
Step 2	58°C	2 hours

Table 8. Program to be set in the thermal cycler for the hybridization reaction of HaloPlex HS probes.

7- At least 30 minutes before the end of the Hybridization reaction, remove the reagents from cold storage to allow the solutions to reach the appropriate temperature:

a) from storage at -20°C, remove HS Hybridization Stop Solution, HS Ligation Solution, HS Capture Solution, HS Wash 1 Solution, HS Wash 2 Solution, and HS Elution Buffer and place

them at room temperature;

b) from storage at 4°C, remove Agencourt AMPure XP and Dynabeads MyOne Streptavidin T1 magnetic beads and place them at room temperature;

c) from storage at -20°C, remove the 10 mM rATP the HS DNA Ligase and place on ice.

STEP 3. REMOVAL OF THE HYBRIDIZATION BUFFER

In this step the hybridization buffer is removed, in preparation for the ligation step, using AMPure XP beads.

1- Prepare 400 µl of 70% ethanol per sample, plus excess, for use in [step 8](#).

2- Mix the AMPure XP bead suspension well, until the suspension appears homogeneous and consistent in color.

3- For each sample to be purified, prepare a bead mix by combining 20 µl of HS Hybridization Stop Solution and 80 µl of the homogeneous AMPure XP bead suspension. Mix well, until the bead suspension mixture appears homogeneous.

4- Add 100 µl of the homogeneous stop solution bead suspension prepared in step 3 to each 100-µl hybridized library sample. Mix by pipetting up and down 10 times using a 200-µl pipette set to 150 µl.

5- Incubate samples for 5 minutes at room temperature with continuous shaking at 1300 rpm.

Make sure the samples are properly mixing in the wells during the 5-minute incubation.

6- Spin briefly to collect the liquid, then place the tubes in the magnetic plate. Wait for the solution to clear (approximately 5 minutes).

7- Keep the tubes in the magnetic plate. Carefully remove and discard the cleared solution from each tube. Do not touch the beads while removing the solution.

8- Continue to keep the tubes in the magnetic plate while you add 200 µl of 70% ethanol into the tubes.

Use fresh 70% ethanol for optimal results.

9- Wait for 1 minute to allow any disturbed beads to settle, then remove and discard the ethanol.

10- Repeat step 8 and step 9 once for a total of two washes.

11- Spin the samples briefly, then place the tubes in the magnetic plate and remove any residual ethanol with a 10-µl volume pipette.

12- Air-dry the samples by keeping the tubes with open lids at room temperature until any residual

ethanol evaporates (approximately 1–5 minutes).

STEP 4. LIGATE THE CIRCULARIZED FRAGMENTS

In this step, DNA ligation reagents are added to the circularized hybridization products to close nicks in the probe-target DNA hybrids.

1- Prepare a 1 mM rATP solution by diluting the provided 10 mM rATP 1:10 with nuclease-free water. For 12-sample runs, combine 1 µl of the provided 10 mM rATP and 9 µl of nuclease-free water.

2- Prepare a DNA ligation master mix by combining the reagents in the following table. Mix the components thoroughly by gentle vortexing then spin the tube briefly.

Reagents	Volume for 1 reaction	Volume for 12 reactions
HS Ligation Solution	10µL	130µL
1mM rATP (step 1)	0,6µL	7,8µL
Nuclease-free water	39,4µL	512,2µL
Total volume	50µL	650µL

Table 9. Preparation of master mix for DNA ligation.

3- Remove the DNA sample tubes from the magnetic plate (see step 3.12), then add 50 µl of the DNA ligation master mix to each sample.

4- Resuspend the beads thoroughly by pipetting up and down 10 times using a 100-µl multichannel pipette set to 40 µl. Visually inspect the beads to verify complete resuspension; if needed, do additional rounds of pipetting up and down until the beads are in a homogeneous suspension.

5- Incubate the samples for 2 minutes at room temperature to elute the DNA samples from the beads.

6- Spin briefly to collect the liquid, then place the tubes in the magnetic plate. Wait for the solution to clear (approximately 5 minutes).

7- For each sample, transfer 47.5 µl of the cleared supernatant to a fresh tube. Pipette slowly to minimize transfer of beads with the supernatant.

8- Add 2.5 µl of HS DNA Ligase to each sample tube. Mix by gentle vortexing and then spin briefly to collect the liquid.

9- Incubate the tubes in a thermal cycler at 55°C for 10 minutes, using a heated lid.

During the 10-minute incubation, prepare the following components for later protocol steps: Dynabeads MyOne Streptavidin T1 magnetic beads and Wash 1 Mix.

STEP 5. CAPTURE THE TARGET DNA

In this step, the circularized target DNA-HaloPlex HS probe hybrids, containing biotin, are captured on streptavidin beads.

1- Vigorously resuspend the Dynabeads MyOne Streptavidin T1 magnetic beads on a vortex mixer. The magnetic beads settle during storage.

2- For each sample in the run prepare 40 µl (1 Volume) of Dynabeads MyOne Streptavidin T1 magnetic beads in HS Capture Solution, using the steps below:

a) Transfer the appropriate volume of bead suspension to a 1.5-ml tube.

Reagents	Volume for 1 reaction	Volume for 12 reactions (including excess)
Dynabeads MyOne Streptavidin T1	40µL	520µL

Table 10. Volume of the beads suspension to be transferred into the 1.5-mL cone.

b) Put the tube into a 1.5 ml tube-compatible magnetic rack until the solution has cleared (approximately 5 minutes).

c) Carefully remove and discard the cleared supernatant using a pipette.

d) Add an equivalent volume of HS Capture Solution to the beads and resuspend by pipetting up and down.

Reagents	Volume for 1 reaction	Volume for 12 reaction (including excess)
HS Capture Solution	40µL	520µL

Table 11. Volume of HS Capture Solution needed for each reaction.

3- Prepare Wash 1 Mix for use on step 5.8 using the following steps:

a) Prepare 10 µl per sample, plus excess, of fresh 1 M NaOH for use in step b. Prepare the 1 M NaOH solution from a 10 M NaOH stock solution.

b) Prepare Wash 1 Mix by combining the reagents in the following table.

Reagenti	Volume per 1 reazione	Volume per 12 reazioni
HS Wash 1 Solution	90µL	1170µL

NaOH 1 M	10µL	130µL
Volume Totale	100µL	1300µL

Table 12. Wash 1 mix preparation.

4- Remove the ligation reactions from the thermal cycler and immediately add 40 µl of the HS Capture Solution + streptavidin bead mixture (prepared in step 5.2) to each 50-µl ligation reaction.

5- Incubate samples for 15 minutes at room temperature with continuous shaking. Make sure the samples are properly mixing in the wells during the 15-minute incubation. During the 15-minute incubation, prepare the PCR Master Mix as described in step 6.

6- Briefly spin the tubes in a desktop centrifuge and then transfer the tubes to the magnetic plate.

7- Wait for the solution to clear, then remove and discard the supernatant using a pipette.

8- Wash the bead-bound samples:

a) Remove the capture reaction tubes from the magnetic plate and add 100 µl of the Wash 1 Mix prepared in step 5.3 to each tube.

b) Resuspend the beads thoroughly by pipetting up and down 10 times using a 100-µl multichannel pipette set to 80 µl then briefly spin the tubes in a desktop centrifuge.

c) Incubate the samples at room temperature for 1 minute.

d) Transfer the tubes to the magnetic plate. Wait for the solution to clear, then remove and discard the supernatant using a pipette.

e) Add 150 µl of the provided HS Wash Solution 2 to each tube.

f) Resuspend the beads thoroughly by pipetting up and down 10 times, then briefly spin the tubes in a desktop centrifuge.

g) Transfer the tubes to the magnetic plate. Wait for the solution to clear, then remove and discard the supernatant using a pipette.

STEP 6. PCR AMPLIFY THE CAPTURED TARGET LIBRARY

In this step, the captured target libraries are amplified by PCR.

1- Prepare the PCR master mix, on ice, by combining the reagents in the following table.

Reagents	Volume for 1 reaction	Volume for 12 reactions
Nuclease-free water	53.2µL	638.4µL
Herculase II Reaction Buffer	30µL	360µL

dNTPs (100mM, 25mM for each dNTP)	0,8μL	10,4μL
Primer 1	4μL	52μL
Primer 2	8μL	104μL
Herculase II Fusion DNA Polimerasi	4μL	52μL
Total volume	100μL	1300μL

Table 13. Preparation of PCR master mix

2- Mix the master mix components by gentle vortexing.

3- Remove the captured and washed DNA sample tubes from the magnetic plate (see step 5.8g), then transfer 100 μl of PCR Master Mix to each sample tube. Mix by pipetting up and down until the beads are in a homogeneous suspension.

4- Place the amplification reaction tubes in a thermal cycler and run the program in the following table, using a heated lid. The optimal amplification cycle number varies for each HaloPlex HS or ClearSeq HS probe design. For this panel, the recommended cycle number is 25.

Step	Number of cycles	Temperature	Time
1	1	98°C	2 minuti
2	25	98°C	30 seconds
		60°C	30 seconds
		72°C	1 minute
3	1	72°C	Hold
4	1	8°C	Hold

Table 14. HaloPlex HS post-capture DNA amplification PCR program

During amplification, remove the Agencourt AMPure XP magnetic beads from +4°C storage, and allow the beads to come to room temperature for use in step 5.7.

5- When the PCR program is complete, briefly spin the tubes in a desktop centrifuge and then transfer the tubes to the magnetic plate.

6- Wait for the solution to clear, then remove 40 μl of each PCR reaction sample to a fresh 0.2-ml tube for purification.

Store the remaining volume of each sample at –20°C for troubleshooting.

STEP 7. PURIFY THE AMPLIFIED TARGET LIBRARY

In this step, the amplified target DNA is purified using AMPure XP beads.

- 1- Verify that the Agencourt AMPure XP beads have been kept at room temperature for at least 30 minutes.
- 2- Prepare 400 µl of 70% ethanol per sample, plus excess, for use in step 7.9.
- 3- Mix the AMPure XP bead suspension well, until the suspension appears homogeneous and consistent in color.
- 4- For each sample to be purified, prepare a bead mix by combining 40 µl of nuclease-free water and 100 µl of the homogeneous AMPure XP bead suspension. Mix well, until the bead mix suspension appears homogeneous.
- 5- Add 140 µl of the homogeneous bead suspension prepared in step 7.4 to each 40-µl amplified library sample. Vortex thoroughly. Using this bead-to-sample volume ratio is imperative to ensure optimal purification results.
- 6- Incubate samples for 5 minutes at room temperature with continuous shaking. Make sure the samples are properly mixing in the wells during the 5-minute incubation.
- 7- Spin briefly to collect the liquid, then place the tubes in the magnetic plate. Wait for the solution to clear (approximately 5 minutes).
- 8- Keep the tubes in the magnetic plate. Carefully remove and discard the cleared solution from each tube using a 200-µl pipette set to 180 µl. Do not touch the beads while removing the solution.
- 9- Continue to keep the tubes in the magnetic plate while you add 200 µl of 70% ethanol into the tubes. Use fresh 70% ethanol for optimal results.
- 10- Wait for 1 minute to allow any disturbed beads to settle, then remove the ethanol using a 200-µl pipette set to 200 µl.
- 11- Repeat step 7.9 and step 7.10 once for a total of two washes.
- 12- Spin the samples briefly, then place the tubes in the magnetic plate and remove any residual ethanol with a 10-µl volume pipette.
- 13- Air-dry the samples by keeping the tubes with open lids at room temperature until any residual ethanol evaporates (approximately 1–5 minutes). Make sure all ethanol has evaporated before continuing.
- 14- Remove tubes from the magnetic plate and add 45 µl of HS Elution Buffer to each sample.
- 15- Mix thoroughly by pipetting up and down 10 times using a 100-µl pipette set to 30 µl.

- 16- Incubate the tubes at room temperature for 2 minutes to allow DNA elution.
- 17- Put the tubes in the magnetic plate and leave for 2 minutes or until the solution is clear.
- 18- Remove the cleared supernatant (approximately 40 μ l) to a fresh tube. You can discard the beads at this time.

STEP 8. VALIDATE ENRICHMENT AND QUANTIFY ENRICHED TARGET DNA

Prior to sample pooling and sequencing sample preparation, validate enrichment and quantify the enriched target DNA in each library sample by microfluidic analysis using the 2100 Bioanalyzer. Each amplicon in the prepared library contains one target insert surrounded by sequence motifs required for multiplexed sequencing using the Illumina platform, as shown in Figure 21.

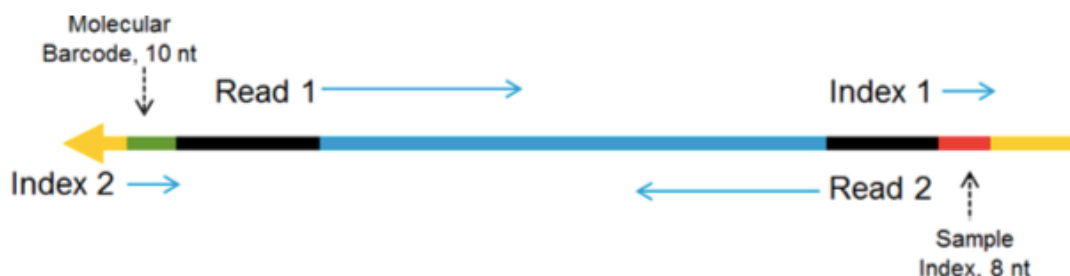


Figure 21. Content of HaloPlex HS-enriched target amplicons. Each amplicon contains one target insert (blue) surrounded by the Illumina paired-end sequencing elements (black), the sample index (red), the molecular barcode (green) and the library bridge PCR primers (yellow).

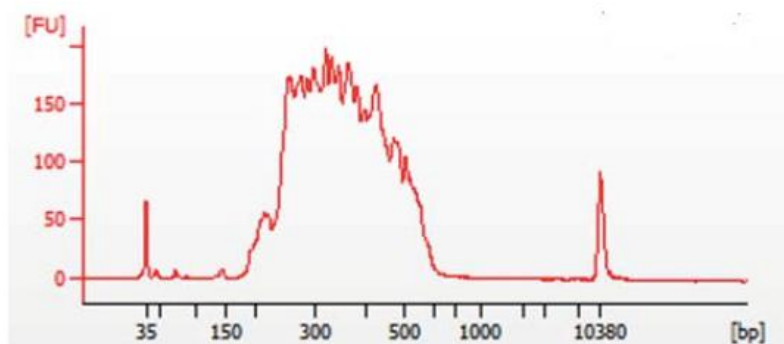


Figure 22. Validation of HaloPlex HS enrichment by 2100 Bioanalyzer analysis.

The amplicons should be approximately 190 to 640 bp in length, with the majority of products sized approximately 225 to 540 bp. Although the DNA fragment size distribution may vary for different DNA samples and different probe designs, use the constant size range of 175 to 625 bp for quantitation of the enriched target DNA in each sample. Any spurious DNA products outside of the 175 to 625 bp size range should be excluded from the target DNA quantitation results.

Some designs may generate a peak at about 140 bp. This peak is associated with an adaptor-dimer product which will cluster and generate sequence that does not map to the genome. If the molar fraction of the 140 bp peak is greater than 10%, do another round of AMPure purification after pooling samples. First, pool equimolar amounts of libraries to be multiplexed, using concentrations determined for the 175–625 bp peak of each sample. Using 40 µl of the pooled libraries, purify the DNA using AMPure XP beads according to the protocol step 7.

STEP 9. POOL SAMPLES FOR MULTIPLEXED SEQUENCING

Before proceeding with sequencing, it is necessary to prepare a pool of the indexed libraries with final concentration 20nM ($V_f=100\mu\text{l}$) such that each sample is present in equimolar amounts within the pool. The following formula can be used for calculation:

$$V_{\text{INDEXED SAMPLE}} = \frac{V_f \times C_f}{\# \times C_i}$$

V_f = Final desired volume of pool

C_f = Final desired concentration of pool

$\#$ = number of indexed samples

C_i = Initial concentration of each indexed sample

4.4. MiSeq Illumina Protocol

The MiSeq® Illumina system uses sequencing-by-synthesis (SBS) technology with a workflow that involves cluster generation, sequencing and data analysis.

Dilution and denaturation of libraries for sequencing on MiSeq Illumina platform

- 1- Dilute the 20nM pool to a final concentration of 4nM (Vf dilution=30µl).
- 2- Prepare a 0.2N NaOH solution, Vf=1ml (800µl distilled water + 200µl NaOH 1.0N).
- 4- Shake by inversion several times to homogenize the solution. The solution should be used within 12 hours after preparation.
- 5- Thaw the HT1 buffer at room temperature and store it at 2-8°C until use.
- 6- For denaturation, add 5 µl of the 4 nM library to 5 µl of 0.2N NaOH. Vortex briefly and centrifuge at 280x g for 1 min.
- 7- Incubate 5 min at room temperature.
- 8- Add 990 µl of pre-cooled HT1 buffer to the cone containing the denatured library. The result is 1ml of a 20pM denatured library.
- 9- Dilute the 20pM library with HT1 buffer in order to obtain the desired concentration, following the instructions in the table below:

Concentrazione	6pM	8pM	10pM	12pM	15pM	20pM
Libreria 20pM	180 µl	240 µl	300 µl	360 µl	450 µl	600 µl
HT1 pre-raffreddato	420 µl	360 µl	300 µl	240 µl	150 µl	0 µl

Table 15. Dilution of the library

(The 12pM concentration was used for the experiment).

Sequencing procedure on MiSeq Illumina platform

REAGENT CARTRIDGE THAWING

- 1- Reagent cartridge thawing is done using a water bath containing deionized water at room temperature (MiSeq v3 cartridges: approximately 60-90 minutes).
- 2- Remove the cartridge from the water bath and gently tap it on the bench to drain excess water from the cartridge base and dry the cartridge base.
- 3- Turn the cartridge upside down 10 times to mix the thawed reagents, then visually inspect all positions to make sure they are thawed.

4- Gently tap the cartridge on the bench to reduce air bubbles in the reagents.

5- Place the reagent cartridge on ice or store it at a temperature of 2-8 °C while waiting to set up the run.

RUN SETUP

Next, the “sample sheet” is prepared using the Illumina Experiment manager interface, entering index information into the worksheet. The MiSeq Control Software (MCS) interface provides commands to configure the instrument, set up and monitor runs, and perform maintenance procedures. RUN SETUP includes: loading the pre-filled reagent cartridge, buffer bottle and flow cell, without any other manual intervention.

The MiSeq Reagent Kit includes a flow cell and the reagents required for sequencing: the PR2 bottle and reagent cartridge, which use radio frequency identification (RFID) for compatibility and accurate tracking of consumables.

The MiSeq flow cell is a substrate on disposable glass on which clusters are generated and the sequencing reaction is performed. Reagents enter the interior through the inlet port, pass through the single-lane imaging area, and flow out the outlet port. Waste from the flow cell is collected in the waste bottle.

The libraries are loaded onto the reagent cartridge before the start of the run and then automatically transferred to the flow cell after the run is started. The MiSeq reagent cartridge is a disposable consumable consisting of small sealed reservoirs pre-filled with sufficient reagents for cluster generation and sequencing. Each reservoir in the cartridge is numbered. Sample libraries are loaded onto the cartridge at position 17, labeled Load Samples. After sequencing is complete, always perform an instrument wash.

CLEANING THE FLOW CELL

Gently rinse the flow cell with laboratory water until excess salts are completely removed from the glass and plastic cartridge. Dry carefully by blotting gently without leaving any residue.

LOADING THE FLOW CELL

Lift the door of the flow cell compartment, then press the release button to the right of the flow cell snap cover to open it. Holding the flow cell by its ends, place it on the cell tray. Gently press the flow cell snap cover to close it. Now close the flow cell compartment door.

LOADING THE LIBRARIES

Pierce the aluminum seal covering the reservoir marked Load Samples and transfer 600 µl of the prepared libraries into the sample loading reservoir (position 17 in figure 23).

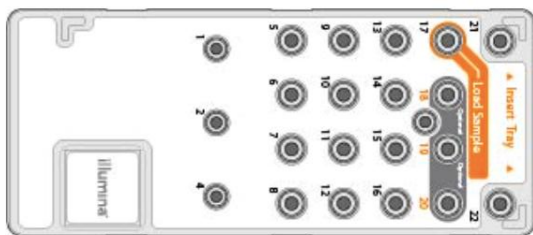


Figure 23. MiSeq reagent cartridge with reserved reservoirs.

LOADING REAGENTS

Open the reagent compartment door, lift the suction comb handle until it locks. Remove the wash bottle and load the PR2 bottle. Empty the contents of the waste bottle into the waste bottle container. Slowly lower the handle of the suction comb. Proceed with reagent cartridge loading: open the refrigerated reagent compartment door, grasp the reagent cartridge from the end with the Illumina label and slide it inside until it stops. Close the refrigerated compartment door and close the reagent compartment door.

Through the MCS interface, you can select the sample sheet of the experiment and check the settings related to the sequencing protocol before starting the run. The system checks all sequencing parameters. If there are positive results from the verification tests performed, you can select Start Run to start sequencing. When the run is complete, the Next button appears. Before proceeding, review the results on the Sequencing screen. Select Next to exit the Sequencing screen and proceed with post run washing.

PERFORMING THE POST-RUN WASH

At the end of a run, always perform a wash of the instrument. Follow the software instructions to load the wash components and perform the wash. Post-run flushing takes about 20 minutes and is recommended to be performed immediately after the completion of a run. An instrument wash is required before setting up a subsequent run.

In addition, a maintenance wash (wash duration approximately 90 minutes) using 0.5 percent Tween 20 wash solution is also performed monthly to ensure optimal instrument performance.

Bioinformatic analysis of MiSeq sequencing results

Concerning the Assembly and Variant Calling phases, reads were aligned with the human reference hg19 genome using the Burrows–Wheeler aligner¹⁸⁸. Fast and accurate short read alignment with Burrows–Wheeler transform16 was performed, and aligned reads were inspected with the IGV software (Broad Institute)¹⁸⁹. The variant calling for the identification of nucleotide variants was performed using the Genome Analysis Toolkit HaplotypeCaller Module¹⁹⁰. Coverage statistics and the average depth of coverage were computed using custom scripts in the R software environment [<https://www.r-project.org/>].

Rare SNVs and InDels were selected by removing all variants with no functional impact (intronic and missense) and with allelic frequency smaller than 0.0001.

Variant prioritization phase was carried out as follows: variants with a minor allele frequency (MAF) ≤ 0.01 in publicly available resources (dbSNP; in 1000 Genomes Project or Genome Aggregation Database) were considered. The pathogenic effect of the missense/splicing mutations was estimated using PROVEAN, PolyPhen-2, MutationTaster, FATHMM, NetGene2, ASSP, and SpliceAI tools. Attribution of pathogenicity was performed according to the American College of Medical Genetics (ACMG) guidelines¹⁹¹.

4.5. Kringle IV Type 2 (KIV2) Repeats Evaluation

The *LPA* KIV2 size polymorphism was genotyped through digital droplet polymerase chain reaction, using the QX200 Droplet Generator and reader system (Bio-Rad).

Genotyping resulted in an estimate of the total number (sum of repeats on both alleles) of KIV2 repeats. To improve precision, all samples were analysed in duplicates by the same molecular biologist, using the same calibrator and control samples.

Telomerase reverse transcriptase gene (*TERT*) was used as a single-copy reference gene; primers and probes for *TERT* were commercially available and well-validated (REF. 4401633, Applied Biosystems™, Waltham, MA, USA). Specific primers and probes targeting the exon 4 of the *LPA* gene were designed and generated for the technique (Applied Biosystems™, Waltham, MA, USA); Table 16 lists the sequences for primers and probes.

<i>LPA</i> exon 4	
Forward primer	5'-GTC AGG TGG GAG TAC TGC AA-3'
Reverse primer	5'-CGA CGG CAG TCC CTT CTG-3'
Probe	FAM-CCT GAC GCA ATG CTC A-MGBNFQ

Table 16. Primers and fluorogenic probe targeting exon 4 of *LPA* gene used in digital droplet PCR assay.

ddPCR: ANALYSIS AND DATA NORMALIZATION

ddPCR was performed using QX200™ Droplet Digital PCR (Bio-Rad Laboratories) according to manufacturer's instructions.

Accordingly, we prepared 22 μ L of a reaction mixture containing 10 ng of genomic DNA, 11 μ L

of 2X ddPCR Supermix for probes (no dUTP) (Bio-Rad, Cat#1863023), 900 nM and 400 nM of *LPA* exon 4 primers and probe, respectively, and 1.1 μ L of 20X TERT reference assay.

A 20 μ L volume of this mixture was then loaded into the sample wells of a DG8 cartridge (Bio-Rad, Cat#17005222), while 70 μ L of Droplet Generator Oil (Bio-Rad, Cat#1863005) was loaded into the oil wells. A total of 40 μ L of oil–water emulsion containing approximately 12000–20000 droplets was generated with the QX200 droplet generator and gently transferred into a separate well of a 96-well PCR plate, heat sealed with foil. PCR was performed under the following thermocycling conditions: enzyme activation at 95°C for 10 min (1 cycle), denaturation at 94 °C for 30 s and annealing/extension at 60°C for 1 min (40 cycles), enzyme deactivation at 98°C for 10 min (1 cycle), and a hold at 4°C. All steps had a ramp rate of 2 °C/s. Annealing at 60°C was selected after strong gradient tests and was proven optimal (data not shown). After amplification, the plate is then loaded to a droplet reader that singulates droplets prior to detection by a FAM/VIC two color fluorescence detector which samples at a rate of 100 kHz and identify where amplification has or has not occurred (positive and negatives, respectively), dividing the entire droplet population into four discrete clusters. These four populations are droplets containing either no target (F–/V–), one of the targets (F–/V+, F+, V–), or both targets (F+, V+). Setting a fluorescence threshold for each detection channel affords a digital method of droplet classification and computing the average number of copies per droplet based on the fraction of positive droplets and Poisson modelling.

three

Internal controls - C1, C2 and C3, derived from previous analyses in which calibrator and control samples were kindly supplied by Dr. Pia R. Kamstrup⁶ - were necessary to calculate CNV values of each sample and to make them comparable within and between different plates: C2 control sample, used as reference sample, was predicted to have an intermediate number of copies (estimated repeat number: N=50). As concerns C1 and C3 controls, they were predicted to have lower and higher number of repeats than C2, respectively (reported range values: 41<C1<47; 61<C3<68).



Figure 24. ddPCR instruments

4.6. SureSelectXT HS Target Enrichment System

With regard to the genetic susceptibility study in VITT subjects, DNA libraries were constructed using the SureSelectXT HS Target Enrichment System and WES was conducted through NextSeq 500 Illumina Protocol.

Library preparation

In every NGS protocol, preparation of the DNA library, which contains thousands of copies of the selected regions in the panel of interest, is required. This sequencing library was made through Illumina's NextSeq platform, using the SureselectXT HS Target Enrichment System kit and SureSelect XT HS Human All Exon V8 probes from the Agilent company. The SureSelectXT HS target enrichment workflow is summarized in Figure 25.

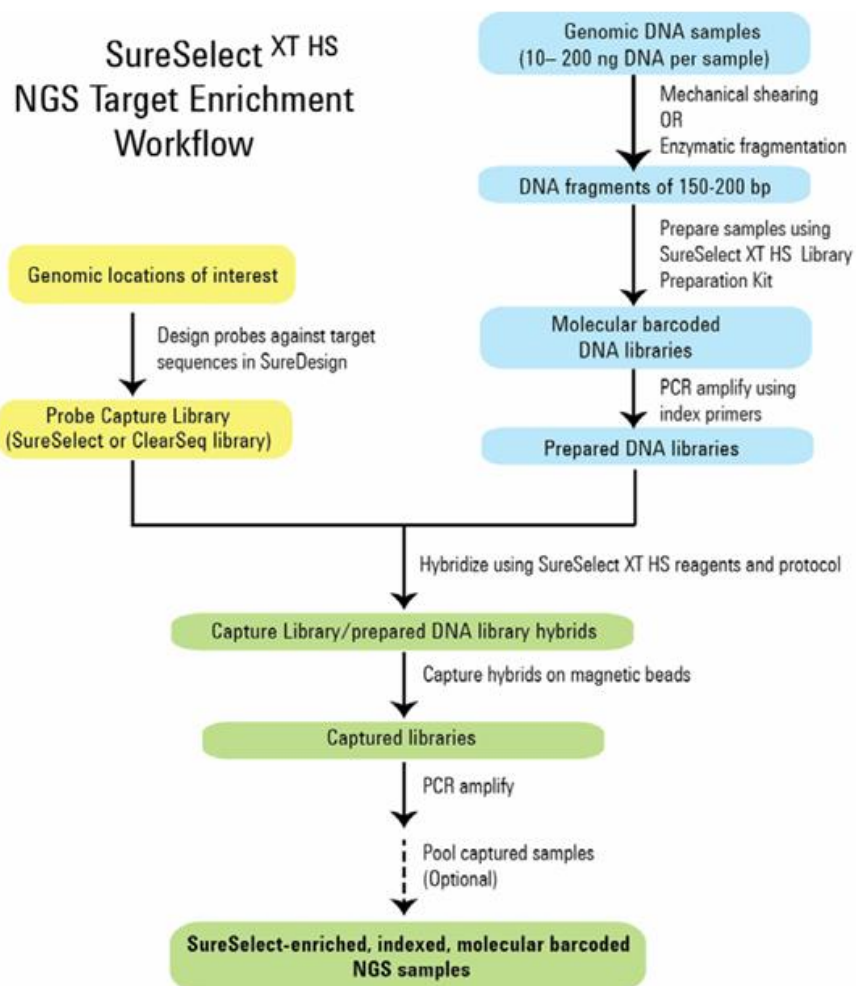


Figure 25. Overall SureSelectXT HS sequencing sample preparation workflow.

STEP 1. ENZYMATIC DNA FRAGMENTATION

In this step, gDNA samples are fragmented using Agilent's SureSelect Enzymatic Fragmentation Kit.

1- In wells of a thermal cycler-compatible strip tube or PCR plate, dilute 10 ng to 200 ng of each gDNA sample with nuclease-free water or 1X Low TE Buffer to a final volume of 7 μ l.

2- Thaw the vial of 5X SureSelect Fragmentation Buffer on ice, vortex, then keep on ice.

3- Preprogram a thermal cycler (with the heated lid ON) with the program in Table 17. Immediately pause the program, and keep paused until samples are loaded in step 7.

Step	Temperature	Time
Step 1	37°C	15 minutes for 150 to 200 bp target fragment size
		10 minutes for 180 to 250 bp target fragment size
Step 2	65°C	5 minutes
Step 3	4°C	Hold

Table 17. Thermal cycler program for enzymatic fragmentation.

4- Prepare the appropriate volume of Fragmentation master mix by combining the reagents in Table 18. Mix well by pipetting up and down 20 times or seal the tube and vortex at high speed for 5–10 seconds. Spin briefly to remove any bubbles and keep on ice.

Reagent	Volume for 1 reaction	Volume for 8 reactions (includes excess)	Volume for 24 reactions (includes excess)
5X SureSelect Fragmentation Buffer	2 μ l	18 μ l	50 μ l
SureSelect Fragmentation Enzyme	1 μ l	9 μ l	25 μ l
Total	3 μ l	27 μ l	75 μ l

Table 18. Preparation of Fragmentation master mix.

5- Add 3 μ l of the Fragmentation master mix to each sample well containing 7 μ l of input DNA.

6- Mix well by pipetting up and down 20 times or cap the wells and vortex at high speed for 5–10 seconds. Spin the samples briefly.

7- Immediately place the plate or strip tube in the thermal cycler and resume the thermal cycling program in Table 17.

8- When the program reaches the 4°C Hold step, remove the samples from the thermal cycler, add 40 µl of nuclease-free water to each sample, and place the samples on ice.

STEP 2. REPAIR AND DA-TAIL THE DNA ENDS

1- Before starting the end-repair protocol, prepare the Ligation master mix to allow equilibration to room temperature before use.

- a) Vortex the thawed vial of Ligation Buffer for 15 seconds at high speed to ensure homogeneity.
- b) Prepare the appropriate volume of Ligation master mix by combining the reagents in Table 18. Slowly pipette the Ligation Buffer into a 1.5-ml Eppendorf tube, ensuring that the full volume is dispensed. Slowly add the T4 DNA Ligase, rinsing the enzyme tip with buffer solution after addition. Mix well by slowly pipetting up and down 15–20 times or seal the tube and vortex at high speed for 10–20 seconds. Spin briefly to collect the liquid.

Keep at room temperature for 30–45 minutes before use on step 3.3.

Reagent	Volume for 1 reaction	Volume for 8 reactions (includes excess)	Volume for 24 reactions (includes excess)
Ligation Buffer (purple cap or bottle)	23 µl	207 µl	598 µl
T4 DNA Ligase (blue cap)	2 µl	18 µl	2 µl
Total	25 µl	225 µl	650 µl

Table 19. Preparation of Ligation master mix.

2- Preprogram a thermal cycler (with the heated lid ON) with the program in Table 17 for the End Repair and dA-Tailing steps. Immediately pause the program, and keep paused until samples are loaded in step 6.

Step	Temperature	Time
Step 1	20°C	15 minutes
Step 2	72°C	15 minutes
Step 3	4°C	Hold

Table 20. Thermal cycler program for End Repair/dA-Tailing

3- Vortex the thawed vial of End Repair-A Tailing Buffer for 15 seconds at high speed to ensure

homogeneity. Visually inspect the solution; if any solids are observed, continue vortexing until all solids are dissolved.

4- Prepare the appropriate volume of End Repair/dA-Tailing master mix, by combining the reagents in Table 21. Slowly pipette the End Repair-A Tailing Buffer into a 1.5-ml Eppendorf tube, ensuring that the full volume is dispensed. Slowly add the End Repair-A Tailing Enzyme Mix, rinsing the enzyme tip with buffer solution after addition. Mix well by pipetting up and down 15–20 times or seal the tube and vortex at high speed for 5–10 seconds. Spin briefly to collect the liquid and keep on ice.

Reagent	Volume for 1 reaction	Volume for 8 reactions (includes excess)	Volume for 24 reactions (includes excess)
End Repair-A Tailing Buffer (yellow cap or bottle)	16 µl	144 µl	416 µl
End Repair-A Tailing Enzyme Mix (orange cap)	4 µl	36 µl	104 µl
Total	20 µl	180 µl	520 µl

Table 21. Preparation of End Repair/dA-Tailing master mix.

5- Add 20 µl of the End Repair/dA-Tailing master mix to each sample well containing approximately 50 µl of fragmented DNA. Mix by pipetting up and down 15–20 times using a pipette set to 60 µl or cap the wells and vortex at high speed for 5–10 seconds.

6- Briefly spin the samples, then immediately place the plate or strip tube in the thermal cycler and resume the thermal cycling program in Table 20.

STEP 3. LIGATE THE MOLECULAR-BARCODED ADAPTOR

1- Once the thermal cycler reaches the 4°C Hold step, transfer the samples to ice while setting up this step.

2- Preprogram a thermal cycler (with the heated lid ON) for the Ligation step with the program in Table 22. Immediately pause the program, and keep paused until samples are loaded in step 5.

Step	Temperature	Time
Step 1	20°C	30 minutes
Step 2	4°C	Hold

Table 22. Thermal cycler program for Ligation.

3- To each end-repaired/dA-tailed DNA sample (approximately 70 µl), add 25 µl of the Ligation master mix that was prepared before and kept at room temperature. Mix by pipetting up and down at least 10 times using a pipette set to 85 µl or cap the wells and vortex at high speed for 5–10 seconds. Briefly spin the samples.

4- Add 5 µl of Adaptor Oligo Mix (white capped tube) to each sample. Mix by pipetting up and down 15–20 times using a pipette set to 85 µl or cap the wells and vortex at high speed for 5–10 seconds.

5- Briefly spin the samples, then immediately place the plate or strip tube in the thermal cycler and resume the thermal cycling program in Table 22.

STEP 4. PURIFY THE SAMPLE USING AMPURE XP BEADS

In this step, the DNA libraries are purified using AmpPure XP Beads. Critical purification protocol parameters are summarized for experienced users in Table 23.

Parameter	Value
Volume of RT AMPure XP bead suspension added to each sample well	80 µl
Final elution solvent and volume	35 µl nuclease-free water
Amount of eluted sample transferred to fresh well	Approximately 34.5 µl

Table 23 AMPure XP bead cleanup parameters after adaptor ligation

1- Verify that the AMPure XP beads were held at room temperature for at least 30 minutes before use.

2- Prepare 400 µl of 70% ethanol per sample, plus excess, for use in step 8.

3- Mix the AMPure XP bead suspension well so that the reagent appears homogeneous and consistent in color.

4- Add 80 µl of homogeneous AMPure XP beads to each DNA sample (approximately 100 µl) in the PCR plate or strip tube. Pipette up and down 15–20 times or cap the wells and vortex at high speed for 5–10 seconds to mix.

5- Incubate samples for 5 minutes at room temperature.

6- Put the plate or strip tube into a magnetic separation device. Wait for the solution to clear (approximately 5 to 10 minutes).

- 7- Keep the plate or strip tube in the magnetic stand. Carefully remove and discard the cleared solution from each well. Do not touch the beads while removing the solution.
- 8- Continue to keep the plate or strip tube in the magnetic stand while you dispense 200 µl of freshly-prepared 70% ethanol in each sample well.
- 9- Wait for 1 minute to allow any disturbed beads to settle, then remove the ethanol.
- 10- Repeat step 8 to step 9 once.
- 11- Seal the wells with strip caps, then briefly spin the samples to collect the residual ethanol. Return the plate or strip tube to the magnetic stand for 30 seconds. Remove the residual ethanol with a P20 pipette.
- 12- Dry the samples by placing the unsealed plate or strip tube on the thermal cycler, set to hold samples at 37°C, until the residual ethanol has just evaporated (typically 1–2 minutes).
- 13- Add 35 µl nuclease-free water to each sample well.
- 14- Seal the wells with strip caps, then mix well on a vortex mixer and briefly spin the plate or strip tube to collect the liquid.
- 15- Incubate for 2 minutes at room temperature.
- 16- Put the plate or strip tube in the magnetic stand and leave for approximately 5 minutes, until the solution is clear.
- 17- Remove the cleared supernatant (approximately 34.5 µl) to a fresh PCR plate or strip tube sample well and keep on ice. You can discard the beads at this time.

STEP 5. AMPLIFY THE ADAPTOR-LIGATED LIBRARY

- 1- Determine the appropriate index assignments for each sample. Use a different index for each sample to be sequenced in the same lane.
- 2- Preprogram a thermal cycler (with the heated lid ON) with the program in Table 24. Immediately pause the program, and keep paused until samples are loaded in step 6.

Segment	Number of Cycles	Temperature	Time
1	1	98°C	2 minutes
2	8 to 14, based on input DNA quality and quantity (see Table 24)	98°C	30 seconds
		60°C	30 seconds

		72°C	1 minute
3	1	72°C	5 minutes
4	1	4°C	Hold

Table 24. Pre-Capture PCR Thermal Cycler Program

Quality of Input DNA	Quantity of Input DNA	Cycles
Intact DNA from fresh sample	100 to 200 ng	8 cycles
	50 ng	9 cycles
	10 ng	11 cycles
FFPE sample DNA	100 to 200 ng	11 cycles
	50 ng	12 cycles
	10 ng	14 cycles

Table 25. Pre-capture PCR cycle number recommendations

3- Prepare the appropriate volume of pre-capture PCR reaction mix, as described in Table 26, on ice. Mix well on a vortex mixer.

Reagent	Volume for 1 reaction	Volume for 8 reactions (includes excess)	Volume for 24 reactions (includes excess)
5× Herculase II Reaction Buffer	10 µl	90 µl	250 µl
100 mM dNTP Mix	0.5 µl	4.5 µl	12.5 µl
Forward Primer	2 µl	18 µl	50 µl
Herculase II Fusion DNA Polymerase	1 µl	9 µl	25 µl
Total	13.5 µl	121.5 µl	337.5 µl

Table 26. Preparation of Pre-Capture PCR Reaction Mix

4- Add 13.5 µl of the PCR reaction mixture prepared in Table 23 to each purified DNA library sample (34.5 µl) in the PCR plate wells.

5- Add 2 µl of the appropriate SureSelect XT HS Index Primer to each reaction. Cap the wells then vortex at high speed for 5 seconds. Spin the plate or strip tube briefly to collect the liquid release any bubbles.

6- Before adding the samples to the thermal cycler, resume the program in Table 21 to bring the temperature of the thermal block to 98°C. Once the cycler has reached 98°C, immediately place the sample plate or strip tube in the thermal block and close the lid.

STEP 6. PURIFY THE AMPLIFIED LIBRARY WITH AMPURE XP BEADS

In this step, the amplified DNA libraries are purified using AmpPure XP Beads. Critical purification protocol parameters are summarized for experienced users in Table 27.

Parameter	Value
Volume of RT AMPure XP bead suspension added to each sample well	50 µl
Final elution solvent and volume	15 µl nuclease-free water
Amount of eluted sample transferred to fresh well	Approximately 15 µl

Table 27. AMPure XP bead cleanup parameters after pre-capture PCR

1- Verify that the AMPure XP beads were held at room temperature for at least 30 minutes before use.

2- Prepare 400 µl of 70% ethanol per sample, plus excess, for use in step 8.

3- Mix the AMPure XP bead suspension well so that the reagent appears homogeneous and consistent in color.

4- Add 50 µl of homogeneous AMPure XP beads to each 50-µl amplification reaction in the PCR plate or strip tube. Pipette up and down 15–20 times or cap the wells and vortex at high speed for 5–10 seconds to mix.

5- Incubate samples for 5 minutes at room temperature.

6- Put the plate or strip tube into a magnetic separation device. Wait for the solution to clear (approximately 5 minutes).

7- Keep the plate or strip tube in the magnetic stand. Carefully remove and discard the cleared solution from each well. Do not touch the beads while removing the solution.

8- Continue to keep the plate or strip tube in the magnetic stand while you dispense 200 µl of freshly-prepared 70% ethanol into each sample well.

- 9- Wait for 1 minute to allow any disturbed beads to settle, then remove the ethanol.
- 10- Repeat step 8 and step 9 step once.
- 11- Seal the wells with strip caps, then briefly spin the samples to collect the residual ethanol. Return the plate or strip tube to the magnetic stand for 30 seconds. Remove the residual ethanol with a P20 pipette.
- 12- Dry the samples by placing the unsealed plate or strip tube on the thermal cycler, set to hold samples at 37°C, until the residual ethanol has just evaporated (typically 1–2 minutes).
- 13- Add 15 µl nuclease-free water to each sample well.
- 14- Seal the wells with strip caps, then mix well on a vortex mixer and briefly spin the plate or strip tube to collect the liquid.
- 15- Incubate for 2 minutes at room temperature.
- 16- Put the plate or strip tube in the magnetic stand and leave for 2 to 3 minutes, until the solution is clear.
- 17- Remove the cleared supernatant (approximately 15 µl) to a fresh PCR plate or strip tube sample well and keep on ice. You can discard the beads at this time.

STEP 7. ASSESS QUALITY AND QUANTITY USING THE 2100 BIOANALYZER AND DNA 1000 ASSAY

Use a Bioanalyzer DNA 1000 chip and reagent kit. Perform the assay according to the Agilent DNA 1000 Kit Guide.

- 1- Set up the 2100 Bioanalyzer instrument as instructed in the reagent kit guide.
- 2- Prepare the chip, samples and ladder as instructed in the reagent kit guide, using 1 µl of each sample for the analysis. Load the prepared chip into the instrument and start the run within five minutes after preparation.
- 3- Verify that the electropherogram shows the peak of DNA fragment size positioned between 300 to 400 bp. Sample electropherograms are shown in Figure 26.

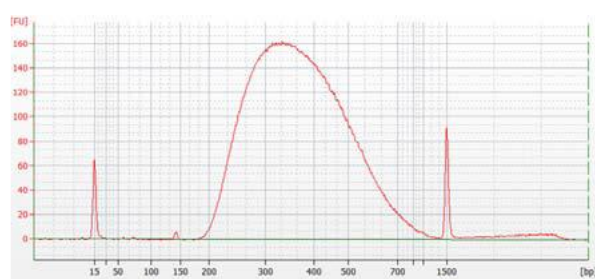


Figure 26. Pre-capture library prepared from gDNA sample analyzed using a DNA 1000 Bioanalyzer assay.

STEP 8. HYBRIDIZE DNA SAMPLES TO THE PROBE

In this step, the prepared gDNA libraries are hybridized to a target- specific probe. For each sample library prepared, do one hybridization and capture. Do not pool samples at this stage. The hybridization reaction requires 500–1000 ng of prepared DNA in a volume of 12 µl. Use the maximum amount of prepared DNA available within this range.

1- Preprogram a thermal cycler (with the heated lid ON) with the program in Table 28. Immediately pause the program, and keep paused until samples are loaded.

Segment #	Number of cycles	Temperature	Time
1	1	95°C	5 minutes
2	1	65°C	10 minutes
3	1	65°C	1 minute (Pause cycler here for reagent addition)
4	60	65°C	1 minute
		37°C	3 seconds
5	1	65°C	Hold briefly until ready to begin capture steps

Table 28 Pre-programmed thermal cycler program for Hybridization

2- Place 500–1000 ng of each prepared gDNA library sample into the hybridization plate or strip tube wells and then bring the final volume in each well to 12 µl using nuclease-free water. Use the maximum possible amount of each prepped DNA, within the 500–1000 ng range.

3- To each DNA library sample well, add 5 µl of SureSelect XT HS and XT Low Input Blocker Mix (blue cap). Seal the wells then vortex at high speed for 5 seconds. Spin the plate or strip tube briefly to collect the liquid and release any bubbles.

4- Transfer the sealed sample plate or strip to the thermal cycler and resume the thermal cycling program (Table 27), allowing the cycler to complete Segments 1 and 2 of the program. Important: The thermal cycler must be paused during Segment 3 to allow additional reagents to be added to the Hybridization wells in step 7. During Segments 1 and 2 of the thermal cycling program, begin preparing the additional hybridization reagents as described in step 5 and step 6 below. If needed, you can finish these preparation steps after pausing the thermal cycler in Segment 3.

5- Prepare a 25% solution of SureSelect RNase Block (1 part RNase Block to 3 parts water) according to Table 29. Prepare the amount required for the number of hybridization reactions in the run, plus excess. Mix well and keep on ice.

Reagent	Volume for 1 reaction	Volume for 8 reactions (includes excess)	Volume for 24 reactions (includes excess)
SureSelect RNase Block	0.5 µl	4.5 µl	12.5 µl
Nuclease-free water	1.5 µl	13.5 µl	37.5 µl
Total	2 µl	18 µl	50 µl

Table 29. Preparation of RNase Block solution

6- Prepare the Probe Hybridization Mix appropriate for your probe design size. Use Table 30 for probes <3 Mb. Combine the listed reagents at room temperature. Mix well by vortexing at high speed for 5 seconds then spin down briefly. Proceed immediately to step 7.

Reagent	Volume for 1 reaction	Volume for 8 reactions (includes excess)	Volume for 24 reactions (includes excess)
25% RNase Block solution (from step 5)	2 µl	18 µl	50 µl
Probe	2 µl	18 µl	50 µl
SureSelect Fast Hybridization Buffer	6 µl	54 µl	150 µl
Nuclease-free water	3 µl	27 µl	75 µl
Total	13 µl	117 µl	325 µl

Table 30. Preparation of Probe Hybridization Mix for probes <3 Mb

7- Once the thermal cycler starts Segment 3 of the program in Table 26 (1 minute at 65°C), pause the program. With the cycler paused, and while keeping the DNA + Blocker samples in the cycler, transfer 13 µl of the room-temperature Probe Hybridization Mix from step 6 to each sample well. Mix well by pipetting up and down slowly 8 to 10 times. The hybridization reaction wells now contain approximately 30 µl.

8- Seal the wells with fresh strip caps. Make sure that all wells are completely sealed. Vortex briefly, then spin the plate or strip tube briefly to remove any bubbles from the bottom of the wells. Immediately return the plate or strip tube to the thermal cycler.

9- Resume the thermal cycling program to allow hybridization of the prepared DNA samples to the probe.

STEP 9. PREPARE STREPTAVIDIN-COATED MAGNETIC BEADS

- 1- Vigorously resuspend the Dynabeads MyOne Streptavidin T1 magnetic beads on a vortex mixer. The magnetic beads settle during storage.
- 2- For each hybridization sample, add 50 µl of the resuspended beads to wells of a fresh PCR plate or a strip tube.
- 3- Wash the beads:
 - a) Add 200 µl of SureSelect Binding Buffer.
 - b) Mix by pipetting up and down 20 times or cap the wells and vortex at high speed for 5–10 seconds then spin down briefly.
 - c) Put the plate or strip tube into a magnetic separator device.
 - d) Wait at least 5 minutes or until the solution is clear, then remove and discard the supernatant.
 - e) Repeat step a through step d two more times for a total of 3 washes.
- 4- Resuspend the beads in 200 µl of SureSelect Binding Buffer.

STEP 10. CAPTURE THE HYBRIDIZED DNA USING STREPTAVIDIN-COATED BEADS

- 1- After all streptavidin bead preparation steps are complete, and with the hybridization thermal cycling program in the final hold segment (see Table 27), transfer the samples to room temperature.
- 2- Immediately transfer the entire volume (approximately 30 µl) of each hybridization mixture to wells containing 200 µl of washed streptavidin beads using a multichannel pipette. Pipette up and down 5–8 times to mix then seal the wells with fresh caps.
- 3- Incubate the capture plate or strip tube on a 96-well plate mixer, mixing vigorously (at 1400–1800 rpm), for 30 minutes at room temperature. Make sure the samples are properly mixing in the wells.
- 4- During the 30-minute incubation for capture, prewarm SureSelect Wash Buffer 2 at 70°C as described below.
 - a) Place 200-µl aliquots of Wash Buffer 2 in wells of a fresh 96-well plate or strip tubes. Aliquot 6 wells of buffer for each DNA sample in the run.
 - b) Cap the wells and then incubate in the thermal cycler, with heated lid ON, held at 70°C until used in step 9.
- 5- When the 30-minute incubation period initiated in step 3 is complete, spin the samples briefly

to collect the liquid.

6- Put the plate or strip tube in a magnetic separator to collect the beads. Wait until the solution is clear, then remove and discard all of the supernatant.

7- Resuspend the beads in 200 µl of SureSelect Wash Buffer 1. Mix by pipetting up and down 15–20 times, until beads are fully resuspended.

8- Put the plate or strip tube in the magnetic separator. Wait for the solution to clear (approximately 1 minute), then remove and discard all of the supernatant.

9- Remove the plate or strip tubes from the magnetic separator and transfer to a rack at room temperature. Wash the beads with Wash Buffer 2, using the protocol steps below.

a) Resuspend the beads in 200 µl of 70°C prewarmed Wash Buffer 2. Pipette up and down 15–20 times, until beads are fully resuspended.

b) Seal the wells with fresh caps and then vortex at high speed for 8 seconds. Spin the plate or strip tube briefly to collect the liquid without pelleting the beads. Make sure the beads are in suspension before proceeding.

c) Incubate the samples for 5 minutes at 70°C in the thermal cycler with the heated lid ON.

d) Put the plate or strip tube in the magnetic separator at room temperature.

e) Wait 1 minute for the solution to clear, then remove and discard the supernatant.

f) Repeat step a through step e five more times for a total of 6 washes.

10- After verifying that all wash buffer has been removed, add 25 µl of nuclease-free water to each sample well. Pipette up and down 8 times to resuspend the beads. Keep the samples on ice until they are used.

STEP 11. AMPLIFY THE CAPTURED LIBRARIES

In this step, the SureSelect- enriched DNA libraries are PCR amplified.

Prepare one amplification reaction for each DNA library.

1- Preprogram a thermal cycler (with the heated lid ON) with the program in Table 31. Immediately pause the program, and keep paused until samples are loaded in step 5.

Segment	Number of Cycles	Temperature	Time
1	1	98°C	2 minutes
2	12 (Probes 0.2–3 Mb)	98°C	30 seconds

		60°C	30 seconds
		72°C	1 minute
3	1	72°C	5 minutes
4	1	4°C	Hold

Table 31. Post-capture PCR thermal cyclor program

2- Prepare the appropriate volume of PCR reaction mix, as described in Table 32, on ice. Mix well on a vortex mixer.

Reagent	Volume for 1 reaction	Volume for 8 reactions (includes excess)	Volume for 24 reactions (includes excess)
Nuclease-free water	12.5 µl	112.5 µl	312.5 µl
5× Herculase II Reaction Buffer	10 µl	90 µl	250 µl
Herculase II Fusion DNA Polymerase	1 µl	9 µl	25 µl
100 mM dNTP Mix	0.5 µl	4.5 µl	12.5 µl
SureSelect Post-Capture Primer Mix	1 µl	9 µl	25 µl
Total	25 µl	225 µl	625 µl

Table 32. Preparation of post-capture PCR Reaction mix

3- Add 25 µl of the PCR reaction mix prepared in Table 31 to each sample well containing 25 µl of bead-bound target-enriched DNA.

4- Mix the PCR reactions well by pipetting up and down until the bead suspension is homogeneous. Avoid splashing samples onto well walls; do not spin the samples at this step.

5- Place the plate or strip tube in the thermal cyclor and resume the thermal cycling program in Table 30.

6- When the PCR amplification program is complete, spin the plate or strip tube briefly. Remove the streptavidin-coated beads by placing the plate or strip tube on the magnetic stand at room temperature. Wait 2 minutes for the solution to clear, then remove each supernatant (approximately 50 µl) to wells of a fresh plate or strip tube. The beads can be discarded at this

time.

STEP 12. PURIFY THE AMPLIFIED CAPTURED LIBRARIES USING AMPURE XP BEADS

In this step, the amplified enriched DNA libraries are purified using AmpPure XP Beads.

- 1- Let the AMPure XP beads come to room temperature for at least 30 minutes.
- 2- Prepare 400 µl of fresh 70% ethanol per sample, plus excess, for use in step 8.
- 3- Mix the AMPure XP bead suspension well so that the suspension appears homogeneous and consistent in color.
- 4- Add 50 µl of the homogeneous AMPure XP bead suspension to each amplified DNA sample (approximately 50 µl) in the PCR plate or strip tube. Mix well by pipetting up and down 15–20 times or cap the wells and vortex at high speed for 5–10 seconds. Check that the beads are in a homogeneous suspension in the sample wells. Each well should have a uniform color with no layers of beads or clear liquid present.
- 5- Incubate samples for 5 minutes at room temperature.
- 6- Put the plate or strip tube on the magnetic stand at room temperature. Wait for the solution to clear (approximately 3 to 5 minutes).
- 7- While keeping the plate or tubes in the magnetic stand, carefully remove and discard the cleared solution from each well. Do not disturb the beads while removing the solution.
- 8- Continue to keep the plate or tubes in the magnetic stand while you dispense 200 µl of freshly-prepared 70% ethanol in each sample well.
- 9- Wait for 1 minute to allow any disturbed beads to settle, then remove the ethanol.
- 10- Repeat step 8 and step 9 once for a total of two washes. Make sure to remove all of the ethanol at each wash step.
- 11- Seal the wells with strip caps, then briefly spin to collect the residual ethanol. Return the plate or strip tube to the magnetic stand for 30 seconds. Remove the residual ethanol with a P20 pipette.
- 12- Dry the samples by placing the unsealed plate or strip tube on the thermal cycler, set to hold samples at 37°C, until the residual ethanol has just evaporated (typically 1–2 minutes).
- 13- Add 25 µl of nuclease-free water to each sample well.
- 14- Seal the sample wells, then mix well on a vortex mixer and briefly spin to collect the liquid without pelleting the beads.

15- Incubate for 2 minutes at room temperature.

16- Put the plate or strip tube in the magnetic stand and leave for 2 minutes or until the solution is clear.

17- Remove the cleared supernatant (approximately 25 μ l) to a fresh well. You can discard the beads at this time.

STEP 13. ASSESS SEQUENCING LIBRARY DNA QUANTITY AND QUALITY USING THE AGILENT 2100 BIOANALYZER AND HIGH SENSITIVITY DNA ASSAY

Use the Bioanalyzer High Sensitivity DNA Assay to analyze the amplified captured DNA. Perform the assay according to the High Sensitivity DNA Kit Guide.

1- Set up the 2100 Bioanalyzer instrument as instructed in the reagent kit guide.

2- Prepare the chip, samples and ladder as instructed in the reagent kit guide, using 1 μ l of each sample for the analysis.

3- Load the prepared chip into the instrument and start the run within five minutes after preparation.

4- Verify that the electropherogram shows the peak of DNA fragment size positioned between 200 and 400 bp. Sample electropherograms are shown in Figure 27.

5- Measure the concentration of each library by integrating under the entire peak. For accurate quantification, make sure that the concentration falls within the linear range of the assay.

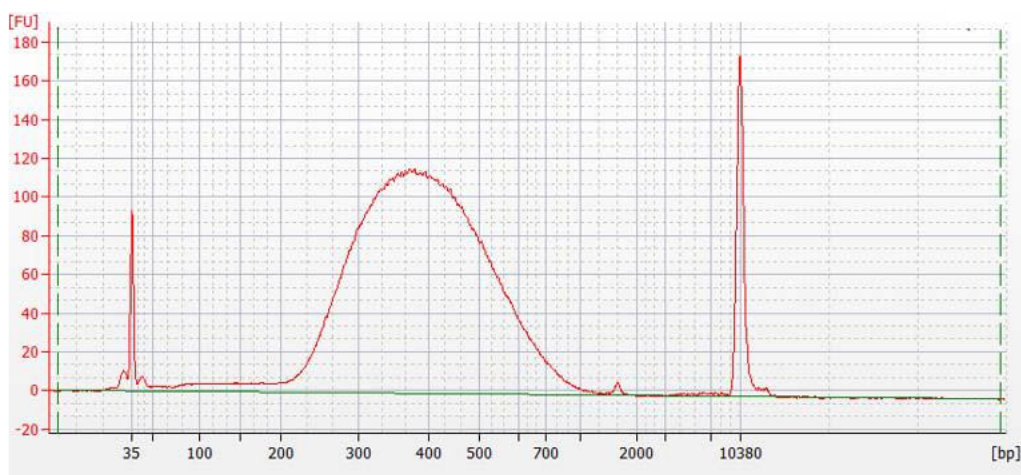


Figure 27. Post-capture library prepared from a high-quality gDNA sample analyzed using a Bioanalyzer system High Sensitivity DNA assay.

STEP 14. POOL SAMPLES FOR MULTIPLEXED SEQUENCING

The number of indexed libraries that may be multiplexed in a single sequencing lane is determined by the output specifications of the platform used, together with the amount of sequencing data required for your research design. Calculate the number of indexes that can be combined per lane, according to the capacity of your platform and the amount of sequencing data required per sample. Combine the libraries such that each index- tagged sample is present in equimolar amounts in the pool using the following method:

Starting with samples at different concentrations, add the appropriate volume of each sample to achieve equimolar concentration in the pool, then adjust the pool to the desired final volume using Low TE. The formula below is provided for determination of the amount of each indexed sample to add to the pool.

$$V_{INDEXED\ SAMPLE} = \frac{V_f \times C_f}{\# \times C_i}$$

V_f = Final desired volume of pool

C_f = Final desired concentration of pool (4 nM)

$\#$ = number of indexed samples

C_i = Initial concentration of each indexed sample

4.7. NextSeq 500 Illumina protocol

Dilution and denaturation of libraries for sequencing on NextSeq Illumina Platform

This procedure denatures and dilutes libraries to a final loading volume of 1.3 ml at a recommended concentration of 1.8 pM for high output kits and 1.5 pM for mid output kits. In practice, loading concentration can vary depending on library preparation and quantification methods.

Prepare a Fresh Dilution of NaOH

1- Combine the following volumes in a microcentrifuge tube:

- a) Laboratory-grade water (800 μ l)
- b) Stock 1.0 N NaOH (200 μ l)

The result is 1 ml of 0.2 N NaOH.

2- Invert the tube several times to mix.

NOTE: Use the fresh dilution within 12 hours.

Prepare HT1

- 1- Remove HT1 from -25°C to -15°C storage and thaw at room temperature.
- 2- Store at 2°C to 8°C until you are ready to dilute denatured libraries.

Prepare RSB

- 1- Remove the tube of RSB from -25°C to -15°C storage and thaw at room temperature.
- 2- Store thawed RSB at 2°C to 8°C until you are ready to dilute libraries.

NOTE: In place of RSB, you can use 10 mM Tris-HCl, pH 8.5 with 0.1% Tween 20

Denature Libraries

- 1- Combine 5 µl of library and 5 µl of freshly diluted 0.2 N NaOH in a microcentrifuge tube.
- 2- Vortex briefly and then centrifuge at $280 \times g$ for 1 minute.
- 3- Incubate at room temperature for 5 minutes.
- 4- Add 5 µl of 200 mM Tris-HCl, pH 7
- 5- Vortex briefly and then centrifuge at $280 \times g$ for 1 minute.

Dilute Denatured Libraries to 20 pM

- 1- Add the 985 µl of prechilled HT1 to the tube of denatured libraries. The result is a 20 pM denatured library.
- 2- Vortex briefly and then centrifuge at $280 \times g$ for 1 minute.
- 3- Place the 20 pM libraries on ice until you are ready to proceed to final dilution.

Dilute Libraries to Loading Concentration

- 1- Dilute the denatured 20 pM library solution to 1.5 pM as follows:

- a) Denatured library solution (97 µl)
- b) Prechilled HT1 (1203 µl)

The total volume is 1.3 ml at 1.5 pM.

- 2- Invert to mix and then pulse centrifuge.

Sequencing process on Illumina NextSeq 500

Prepare the Reagent Cartridge

- 1- Remove the reagent cartridge from -25°C to -15°C storage.
- 2- Thaw in a room temperature water bath until thawed (~60 minutes). Do not submerge the

cartridge.

3- Gently tap on the bench to dislodge water from the base, and then dry the base.

4- Invert the cartridge five times to mix reagents.

5- Inspect positions 29, 30, 31, and 32 to make sure that reagents are thawed.

6- Gently tap on the bench to reduce air bubbles.

Prepare the Flow Cell

1- Remove a new flow cell package from 2°C to 8°C storage.

2- Remove the flow cell from the foil package

3- Set the unwrapped flow cell package aside at room temperature for 30 minutes.

4- Open the clear plastic clamshell package and remove the flow cell.

5- Clean the glass surface of the flow cell with a lint-free alcohol wipe. Dry the glass with a low-lint lab tissue.

Load Libraries onto the Reagent Cartridge

1- Clean the foil seal covering reservoir #10 labeled Load Library Here using a low-lint tissue.

2- Pierce the seal with a clean 1 ml pipette tip.

3- Load 1.3 ml of prepared 1.8 pM libraries into reservoir #10. Avoid touching the foil seal as you dispense the libraries.

Set Up a Sequencing Run

1- From the Home screen, select Experiment.

2- On the Select Assay screen, select Sequence.

The Sequence command opens the imaging compartment door, releases consumables from a previous run, and opens the series of run setup screens. A short delay is normal.

When setting up a sequencing run, you select one of the following run modes to determine where to enter run information and how to analyze data.

Run Mode	Run Information	Data Analysis
Local Run Manager	Enter in Local Run Manager	The software saves data to the specified output folder for automatic analysis in Local Run Manager.

Manual	Enter in the NCS (manual run mode)	The software saves data to a specified output folder for later analysis off the instrument.
---------------	------------------------------------	---

Table 33. Run modes

Then, you can load the Flow Cell, the Buffer Cartridge and the Reagent Cartridge. After the run, there is an Automatic Post-Run Wash.

Bioinformatic analysis of NextSeq 500 sequencing results

Concerning the Assembly and Variant Calling phases, reads were aligned with the human reference hg19 genome using the Burrows–Wheeler aligner¹⁸⁸. Fast and accurate short read alignment with Burrows–Wheeler transform16 was performed, and aligned reads were inspected with the IGV software (Broad Institute)¹⁸⁹. The variant calling for the identification of nucleotide variants was performed using the Genome Analysis Toolkit HaplotypeCaller Module¹⁹⁰. Coverage statistics and the average depth of coverage were computed using custom scripts in the R software environment [<https://www.r-project.org/>].

Rare SNVs and InDels were selected by removing all variants with no functional impact (intronic and missense) and with allelic frequency smaller than 0.0001. For each patient, all genes affected by rare variants were analyzed for overrepresentation analysis by using the “gene list” analysis tool of the Reactome database (reactome.org). Reactome pathways with entities FDR smaller than 0.05 were considered statistically significant.

Genes for variants analysis have been selected according to gene ontology biological processes (blood coagulation [GO:0007596] and fibrinolysis [GO:0042730], platelet activation [GO:0030168] and aggregation [GO:0070527], integrin-mediated signaling pathway [GO:0007229], inflammatory response [GO:0006954], and Human Phenotype Ontology [autoimmune thrombocytopenia: HP:0001973]).

Variant prioritization phase was carried out as follows: nonsynonymous variants with a minor allele frequency (MAF) < 0.01 in publicly available resources (dbSNP; in 1000 Genomes Project or Genome Aggregation Database) were considered. The pathogenic effect of the missense/splicing mutations was estimated using PROVEAN, PolyPhen-2, MutationTaster, FATHMM, NetGene2, ASSP, and SpliceAI tools. Attribution of pathogenicity was performed according to the American College of Medical Genetics (ACMG) guidelines¹⁹¹.

5. RESULTS

5.1. Covid-19 Study

Study Population

In order to identify genetic profiles associated with the severity of the disease [death and admission in intensive care unit (ICU) vs early discharged patients] and thus potentially representing prognostic factors modulating the susceptibility to virus entry, in this thesis, 80 patients suffering from Covid-19, admitted at the Careggi University Hospital, were analysed. The blood samples were received at the Centre for Atherothrombotic Diseases of the Careggi University Hospital - Department of Experimental and Clinical Medicine, University of Florence as part of clinical-laboratory assessment/investigation paths. Patients' blood was collected from March 2020 to August 2021.

The population consisted of 52 males and 28 females, with a median age of 61.5 ± 12.6 . Forty subjects admitted to ordinary hospital wards ("non-ICU"), while 40 patients were admitted to intensive care or died ("ICU/death"). Table 34 reported clinical data of the eighty patients with Covid-19 analyzed.

Table 34. Clinical data of the eighty patients with Covid-19 analyzed.

	<i>Setting</i>			
	Total	ICU/Death	Non-ICU	p-value
<i>N of individuals (n, %)</i>	80 (100)	40 (50)	40 (50)	-
<i>Men (%)</i>	65	70	60	0.348
<i>Age ($\pm$SD; years)</i>	61.5 $\pm$ 12.6	60.0 $\pm$ 12.5	62.5 $\pm$ 12.9	0.424
<i>Smokers (%)</i>	16.3	10.0	22.5	0.130
<i>Hypertension (%)</i>	52.5	57.5	47.5	0.370
<i>Diabetes (%)</i>	10.0	12.5	7.5	0.456
<i>Cardiovascular disease (%)</i>	20.0	27.5	12.5	0.094
<i>Thromboembolic event (%)</i>	12.5	25.0	0	<0.001
<i>Dyslipidemia (%)</i>	12.5	5.0	20.0	0.043
<i>Pulmonary disease (%)</i>	8.8	10.0	7.5	0.692

Targeted Sequencing Results

For these patients, genetic-molecular characterisation by means of NGS technology was carried out at the Advanced Genetic-Molecular Laboratory (Reference: Prof. Betti Giusti) of the Centre for Atherothrombotic Diseases. The NGS genetic analysis available at the Centre involved the sequencing of a panel of 7 genes (*ACE2*, *TMPRSS2*, *CTSL*, *CTSB*, *HSPA5*, *IL6*, *FURIN*) selected on the basis of their involvement in the entry of SARS-CoV-2 into the host cell and their association in the literature with Covid-19 pathology.

Sequencing analysis of the 80 patients with Covid-19 resulted in a total number of genetic variants of 212, regarding all 7 genes included in the virus entry-related panel. Variant prioritization phase was carried out as follows: variants with a minor allele frequency (MAF) ≤ 0.01 in public available resources (dbSNP; in 1000 Genomes Project or Genome Aggregation Database) were considered. The deleterious effect of the missense/splicing mutations on proteins was estimated using PROVEAN, PolyPhen-2, MutationTaster, FATHMM, NetGene2, ASSP, and SpliceAI tools. Attribution of pathogenicity was performed according to the American College of Medical Genetics (ACMG) guidelines¹⁹¹ [pathogenic variants (P), likely pathogenic variants (LP), uncertain significance variants (VUS), likely benign variants (LB), benign variants (B)].

Forty-two out of 212 were rare variants, identified in 34/80 patients and reported in table 35.

In Figure 28 the variants by type in the whole population are reported. Most of the variants identified (48%) are missense (figure 28).

Observing the whole population of $n=80$ subjects, the most affected genes were *TMPRSS2* and *CTSB* (respectively 31% and 24% of total rare variants) (figure 29).

In Figure 30, variants by type, in each gene of the panel, in the whole population, are described: the gene most affected by missense variants is *TMPRSS2*. *TMPRSS2* is also the gene with the most VUS/LP variants (figure 31).

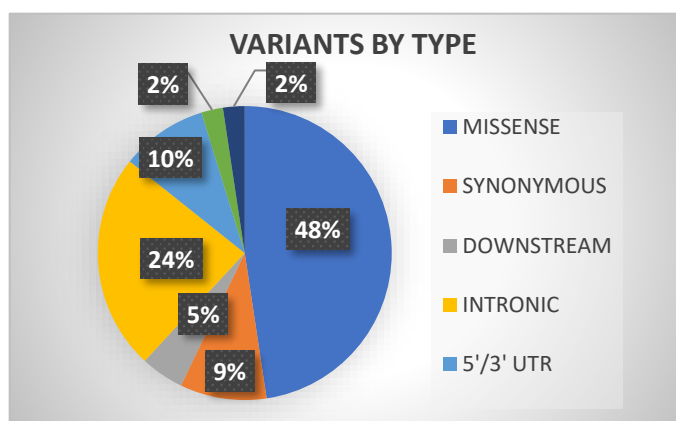


Figure 28. Variants by type in the whole population

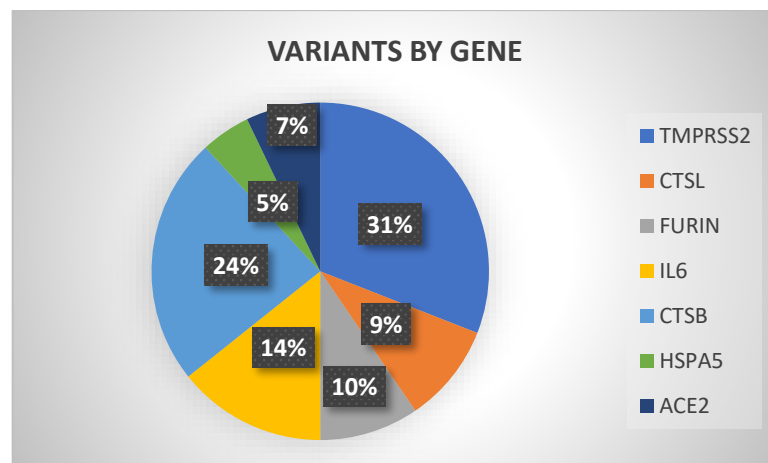


Figure 29. Variants by gene in the whole population

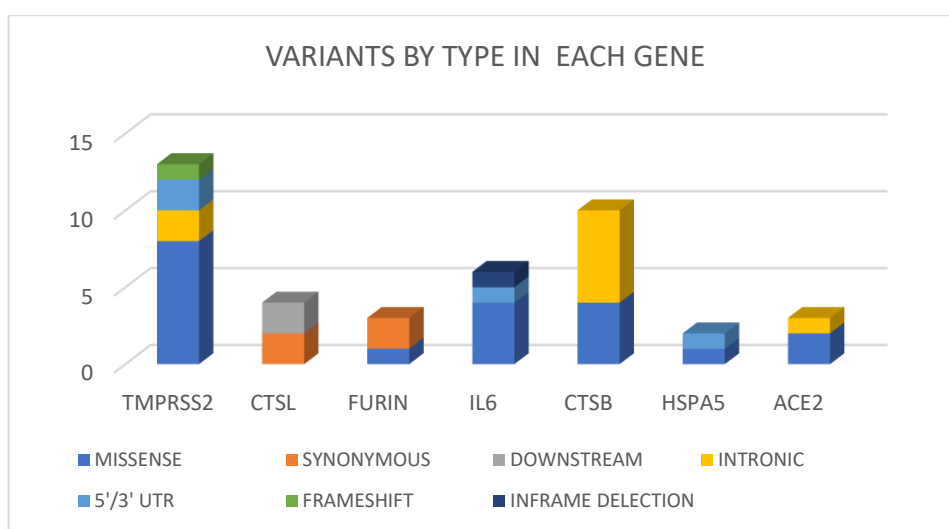


Figure 30. Variants by type, in each gene of the panel, in the whole population

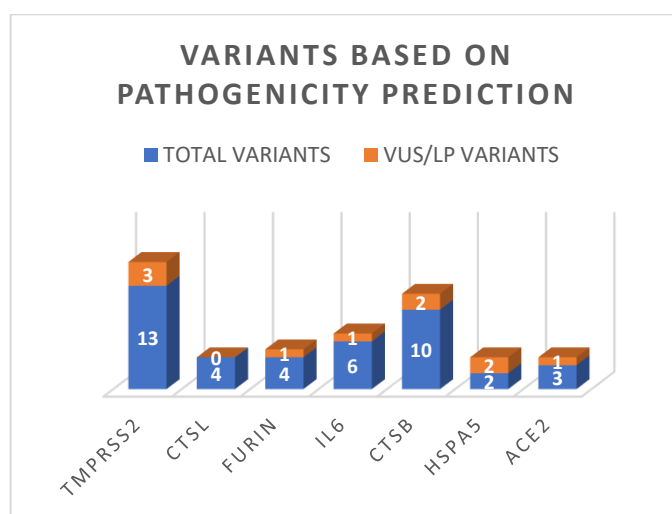


Figure 31. Variants by gene in the whole population, based on ACMG¹⁹¹ pathogenicity prediction.

We then analysed the two population subgroups divided according to the severity of the clinical course (non-ICU and ICU/death). It can be seen that *TMPRSS2*, the gene most affected by mutations in the entire cohort, has, compared to the other genes, a higher percentage of variants in the non-ICU group (40%) than in the ICU/death group (23%). In contrast, the *CTSB* gene follows an inverse trend, increasing from 15% in the non-ICU group to 32% in the ICU/death group (Figures 32-33). However, *TMPRSS2*, as can be seen from Figures 34-35, despite having fewer variants in the ICU/death group than in the non-ICU group, shows a higher number of VUS/LP variants in the group of patients with a worse clinical course (1/8 vs 2/5). The same figures (figures 34-35) also shows that the *HSPA5* gene has 2 VUS/LP variants in the ICU/death group and none in the non-ICU group.

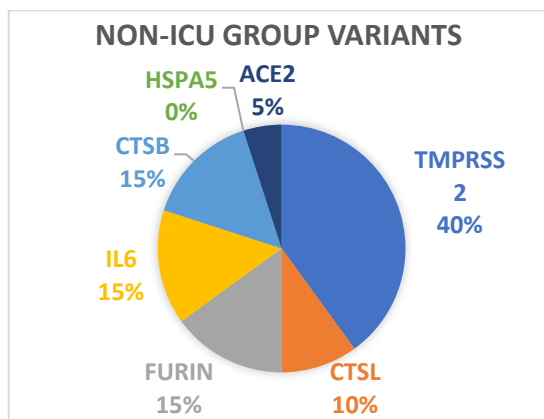


Figure 32. Variants by gene in non-ICU group

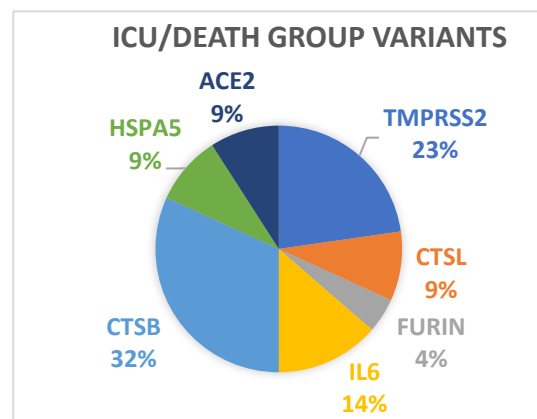


Figure 33. Variants by gene in ICU/death group

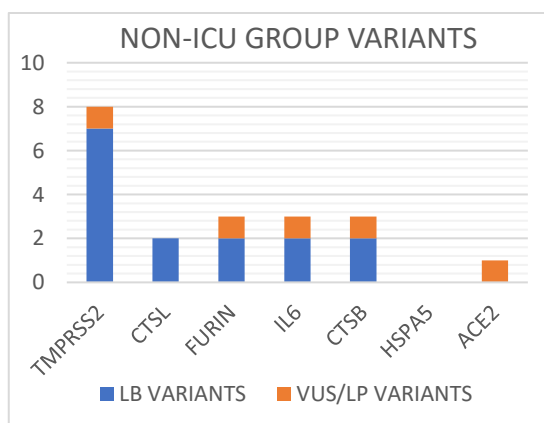


Figure 34. Variants by gene in non-ICU population, based on ACMG¹³⁴ pathogenicity prediction.

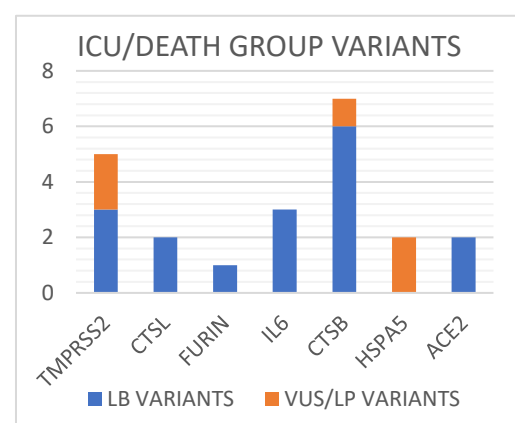


Figure 35. Variants by gene in ICU/death population, based on ACMG¹³⁴ pathogenicity prediction.

Table 35. Rare variants identified in Covid-19 patients in the selected genes.

Patients	Patients Setting	Genes	Position	Mutation cDNA	Mutation protein	SNP Id (rs)	MAF (Global)	MAF (Eur)	ACGM Variant Classification
COV_263	1	ACE2 NM_021804	chrX-15599395-T-C	c.1019A>G	p.Gln340Arg	N/A	N/A	N/A	VUS
COV_1124	2	ACE2 NM_021804	chrX-15589926-A-	c.1665-7del	-	rs774469453	0.00621	0.000493	LB
COV_452	2	ACE2 NM_021804	chrX-15618958-T-C (1/1)	c.77A>G (1/1)	p.Lys26Arg (1/1)	rs4646116	0.00397	0.00589	LB
COV_048	2	CTSB NM_001908	chr8-11703206-C-T	c.886G>A	p.Val296Ile	rs200851570	0.0000478	0.0000617	VUS
COV_061	2	CTSB NM_001908	chr8_11706546:G>C	c.446+9G>C	-	rs199582927	0.003076	0.004357	LB
COV_1121	2	CTSB NM_001908	chr8-11704650-C-T	c.704G>A	p.Ser235Asn	rs17573	0.0106	0.000941	LB
COV_1121	2	CTSB NM_001908	chr8-11705168-CCAGGCC	c.676+14_676+20del	-	rs138997275	0.0119	0.00081	LB
COV_1123	2	CTSB NM_001908	chr8-11706546-C-G	c.446+9G>C	-	rs199582927	0.00308	0.00504	LB
COV_602	2	CTSB NM_001908	chr8-11708421-G-A	c.281C>T	p.Pro94Leu	rs148117767	0.00291	0.00171	LB
COV_152	1	CTSB NM_001908	chr8-11708510-A-G	c.213-21T>C		rs1814495868	NA	NA	LB

COV_1034	1	CTSB NM_001908	chr8-11710956-T-G	c.8A>C	p.Gln3Pro	rs762550595	0.000168	0.000106	VUS
COV_434	2	CTSB NM_001908	chr8-11706546-C-G	c.446+9G>C	-	rs199582927			B
COV_711	1	CTSB NM_001908	chr8-11710096-G-A	c.212+23C>T	-	rs772212336			LB
COV_004	2	CTSL NM_001257971	chr9:90346400:C>T	DOWNSTREAM VARIANT	-	rs142739053	0.00449	0.00651	LB
COV_018	1	CTSL NM_001257971	chr9:90346013:C>T	c.993C>T	p.Pro331=	rs140683449	0.000649	0.00031	LB
COV_880	2	CTSL NM_001257971	chr9-90345360-C-T	c.849C>T	p.Tyr283=	rs190277186	0.0003	0.000159	LB
COV_747	1	CTSL NM_001257971	chr9-90346390-G-A	DOWNSTREAM VARIANT		rs45461199	0.00213	0.00376	LB
COV_011	1	FURIN NM_002569	chr15:91421411:G>T	c.717G>T	p.Ser239=	N/A	N/A	N/A	LB
COV_207	1	FURIN NM_002569	chr15:91420738:A>C	c.579-19A>C	-	rs138111870	0.00492	0.00804	LB
COV_1011	1	FURIN NM_002569	chr15-91419537-T-C	c.230T>C	p.Leu77Pro	rs143659625	7.97E-06	0.0000176	VUS
COV_708	2	FURIN NM_002569	chr15-91424029-C-T	c.1665C>T	p.Ser555=	rs142489043	0.00504	0.00292	B

COV_174	2	HSPA5 NM_005347	chr9-127999058-G-A	c.1778C>T	p.Thr593Ile	N/A	N/A	N/A	VUS
COV_708	2	HSPA5 NM_005347	chr9-127998680-C-	c.*191del	-	NA	NA	NA	VUS
COV_044	2	IL6 NM_000600	chr7_22771296:G>A	c.*104G>A	-	rs13306436	0.00127	0.00016	LB
COV_037	2	IL6 NM_000600	chr7_22771038:A>T	c.485A>T	p.Asp162Val	rs2069860	0.005967	0.007549	LB
COV_747	1	IL6 NM_000600	chr7-22766890-CTT-	c.11_13del	p.Phe4del	rs1034241131	3.98E-06	8.82E-06	VUS
COV_757 COV_880	757-1 880-2	IL6 NM_000600	chr7-22771039-T-A	c.486T>A	p.Asp162Glu	rs13306435	0.0287	0.0091	LB
COV_152	1	IL6 NM_000600	chr7-22771038-A-T	c.485A>T	p.Asp162Val	rs2069860	0.00597	0.00757	LB
COV_016	1	TMPRSS2 NM_001135099	chr21-42866278-C-A	c.349+5G>T	-	rs2146482909	N/A	N/A	VUS
COV_034 COV_385	34-2 385-1	TMPRSS2 NM_001135099	chr21-42866297-G-A	c.335C>T	p.Thr112Ile	rs61735793	0.007035	0.010218	LB
COV_048	2	TMPRSS2 NM_001135099	chr21:42860400::GGAGTCGC	c.470_477dupGCGACTCC	p.Ser160AlafsTer81	rs1167960590	0.000008	0.000015	LP
COV_174	2	TMPRSS2 NM_001135099	chr21-42861455-C-A	c.415G>T	p.Ala139Ser	rs781008294	0.000012	0	VUS

COV_207	1	TMPRSS2 NM_001135099	chr21-42866439-C-T	c.193G>A	p.Ala65Thr	rs61735791	0.00169	0.000139	LB
COV_309	1	TMPRSS2 NM_001135099	chr21-42879910-C-G	c.22G>C	p.Gly8Arg	rs200291871	0.00542	0.00962	LB
COV_400	1	TMPRSS2 NM_001135099	chr21-42837952-G-	c.*117del	-	rs149695119	0.0438	0.00239	LB
COV_848	2	TMPRSS2 NM_001135099	chr21-42838104-G-C	c.1468-24C>G	-	rs118028230	0.00708	0.000108	LB
COV_1016 COV_611 COV_806	1016-1 611-1 806-2	TMPRSS2 NM_001135099	chr21-42866297-G-A	c.335C>T	p.Thr112Ile	rs61735793	0.00704	0.0101	LB
COV_585	1	TMPRSS2 NM_001135099	chr21-42879987-C-T	c.-56G>A (5'UTR)	-	rs962605957	N/A	N/A	LB

Setting 1=non-ICU; Setting 2=ICU/death, CIOP=Conflicting Interpretation Of Pathogenicity; D=Damaging; DC=Disease Causing; LB=Likely Benign; LP=Likely Pathogenic; N=Neutral; NEG=Negative; P=Pathogenic; PD=Probably/Possibly Damaging; T=Tolerated; VUS=Variant of Uncertain Significance.

Lp(a) and Covid-19

Given the conflicting opinions in the literature.³⁻⁵, with the aim of detecting any relationship between genetic determinants of Lp(a) and Covid-19 severity, on the same COVID population the KIV2 repeat polymorphism was studied.

Genetic characterisation of the LPA KIV2 polymorphism produced an estimate of the total number (sum of repeats on both alleles) of KIV2 repeats.

No differences were observed for the number of KIV2 repeats according to disease severity (table 36-37-38):

-neither in the whole population, between ICU/death versus non-ICU patients (table 36),

-nor in the whole population, between subjects with (n=10) or without thrombosis (n=70) (table 37),

-nor when only ICU/death patients were considered (n=10 with thrombosis vs n=30 without thrombosis) (table 38).

Table 36. Association between the number of KIV2 repeats of the non-ICU population and that of the ICU/death population.

	non-ICU (40/80)	ICU/death (40/80)	U test Mann-Whitney
LPA KIV2 repeat	12.4 (9.7-22.1)	11.9 (10-21.1)	0.729

Median and interquartile range (25-75) are reported, as well as the significance value given by the Mann-Whitney test.

Table 37. Association between the number of KIV2 repeats of the thrombotic patients and that of the patients without thrombosis.

	Patients with thrombosis (10/80)	Patients without thrombosis (70/80)	U test Mann-Whitney
LPA KIV2 repeat	13.6 (10.4-24.5)	11.8 (9.7-20.1)	0.341

Median and interquartile range (25-75) are reported, as well as the significance value given by the Mann-Whitney test.

Table 38. Association between the number of KIV2 repeats of the thrombotic patients and that of the ICU patients without thrombosis, considering only ICU/death patients.

	Patients with thrombosis (10/40)	Patients ICU without thrombosis (30/40)	U test Mann-Whitney
LPA KIV2 repeat	13.6 (10.4-24.5)	10.8 (9.8-19.4)	0.396

Median and interquartile range (25-75) are reported, as well as the significance value given by the Mann-Whitney test.

5.2. VITT Study

Study Population

In order to study whether VITT genetic background might predispose to VITT¹³, n=6 patients referred to the Atherothrombotic Diseases Center (Department of Experimental and Clinical Medicine, Azienda Ospedaliero-Universitaria Careggi, Florence) with a diagnosis of definite VITT (based on criteria previously reported by Pavord S et al¹⁹²) underwent molecular characterization through Whole Exome Sequencing (WES) analysis. They are six female patients of Caucasian origin with a mean age of 64.2±13.8 yrs. Of these six patients, five (VITT02, VITT05, VITT06, VITT18, VITT21) received the VAXZEVRIA vaccine (AstraZeneca), and one (VITT03) the JANSSEN vaccine (Johnson & Johnson). Table 39 reported clinical data of the six patients with VITT analyzed. VITT05 (75 yrs) and VITT18 (42 yrs) patients exhibited the worse clinical outcome. VITT02 and VITT06 showed both venous and arterial thrombosis. Only 1 out of 6 patients, showed only arterial thrombosis (VITT18). The other 3 patients showed only venous thrombosis. No demographic and clinical differences were observed between the only subject (VITT03) receiving JANSSEN vaccine and the other five subjects receiving VAXZEVRIA vaccine.

Table 39. Demographic and clinical characteristics of the Study Population.

Patient	Sex	Age, yrs	Vaccine	Venous Thrombosis	Arterial Thrombosis	Bleeding	Clinical Outcome
VITT02	Female	57	Vaxzevria	PE, PVT	Abdominal aorta, lienal artery		No disability
VITT03	Female	78	Janssen	DVT			No disability
VITT05	Female	75	Vaxzevria	PVT, CSVT		SDH, IVH, IPH	Deceased
VITT06	Female	73	Vaxzevria	CSV, PE, RVT	Stroke, lower extremities		No disability
VITT18	Female	42	Vaxzevria		Stroke	IPH	Severe Disability
VITT21	Female	60	Vaxzevria	DVT, CSV, SVT, PE			No disability

PE=Pulmonary Embolism; PVT=Portal Venous Thrombosis; DVT=Deep Venous Thrombosis; CSV=Cerebral Sinuses Venous Thrombosis; RVT=Renal Venous Thrombosis; SVT=Splanchnic Venous Thrombosis; SDH=Subdural Hematoma; IVH=Intraventricular hemorrhage; IPH=Intraparenchymal Hemorrhage.

WES Results

Sequencing analysis of the 6 patients with VITT resulted in a total number of genetic variants of 140,563. The total number of variants per patient analyzed, splitted into rare (MAF $\leq$ 0.01) and non-rare variants is shown in Figure 36. The number (Figure 36) and the variant type distribution (Figure 37) is similar among each patient analyzed. Among rare variants, ranging from 1619 to 1774 among the six patients with VITT, the higher percentage is represented by intronic (ranging from 38.2% to 42.3%), missense (ranging from 28.4% to 30.6%) and synonymous variants (ranging from 18.5% to 21.3%), whereas splice sites (ranging from 3.6% to 5.3%), 5'/3'-UTR (ranging from 3.9% to 5.1%), nonsense (ranging from 0.4% to 0.7%) and frameshift (ranging from 0.2% to 0.8%) variants are less frequent (Figure 37).

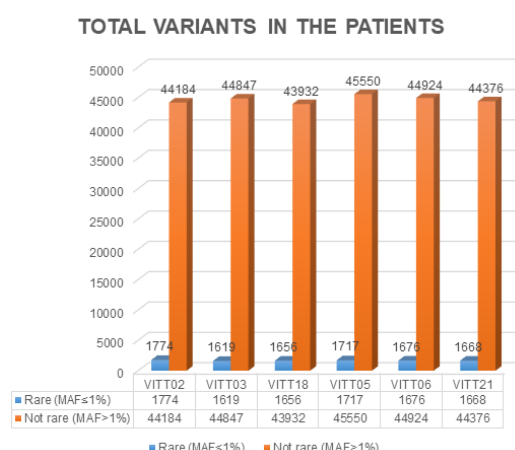


Figure 36. Total variants (rare and not rare) identified in the six patients with VITT.

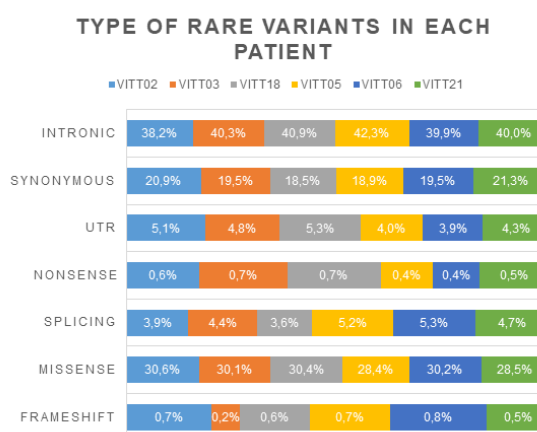


Figure 37. Rare variant type (intronic-missense-synonymous-UTR-splicing-frameshift-nonsense) distribution among patients with VITT

Due to the rare occurrence of VITT, we first of all focused the attention on very rare variants in the 6 patients. The global analysis of all high-quality rare variants evaluated by over-representation analysis by using the "gene list" analysis tool of the reactome database did not show a significant enrichment of mutated genes in biological/functional pathways common to the 6 different patients.

Next, we focused attention on rare variants identified in genes selected to be involved in biological/functional processes associated with blood coagulation (GO:0007596) and fibrinolysis (GO:0042730), platelet activation (GO:0030168) and aggregation (GO:0070527), integrin-mediated signaling pathway (GO:0007229), as well as autoimmune thrombocytopenia

(GO:0001973).

Altogether, in the six patients, n=112 rare variants were identified, all at the heterozygous state. As ACMG¹⁹¹ variant classification criteria were considered, n=31 variants (27.7%) have been classified as uncertain significance variants (VUS): 2 in VITT05, 2 in VITT21, 4 in VITT03, 7 in VITT06, 7 in VITT18, and 9 in VITT02 (Table 40). No variant has been classified as likely pathogenic/pathogenic variant. Even if different genetic variants, patient VITT02 and VITT06 carried one of their mutations in *FGA* gene, and VITT02 and VITT18 in *SERPINA12* gene.

Table 40. Rare uncertain significance variants identified in patients with VITT in the selected genes.

Patient	Gene Symbol and Transcript ID	dbSNP ID	Genetic Variant Nomenclature	Minor Allele Frequency	In Silico Prediction					ACGM Variant Classification
					FATHMM	MutationTaster	Provean	Polyphen	Sift	
VITT02	ITGA2B NM_000419	rs199641871	c.457G>A (p.Ala153Thr)	0.0002339	D	D.C.	N	D	T	VUS
VITT02	ITGAD NM_005353	rs147321998	c.736C>T (p.Arg246Ter)	0.002425						VUS
VITT02	THBD NM_000361	rs888161210	c.1465G>C (p.Asp489His)	N/A	T	P	N	P.D.	D	VUS
VITT02	FGA NM_000508	rs1257998751	c.1132T>A (p.Ser378Thr)	N/A	T	P	N	B	T	VUS
VITT02	CTNNA1 NM_001903	rs371054484	c.377G>C (p.Arg126Pro)	0.00001548	T	D.C.	D	P.D.	D	VUS
VITT02	PIEZO1 NM_001142864	rs776672249	c.2545C>T (p.Arg849Cys)	0.00001668	T	D.C.	D	D	D	VUS
VITT02	GLCE NM_015554	rs199559391	c.1357A>G (p.Thr453Ala)	0.0001583	T	D.C.	N	D	T	VUS
VITT02	ENPP6 NM_153343	rs147979605	c.541C>G (p.Arg181Gly)	0.00001765	T	D.C.	N	B	T	VUS
VITT02	SERPINA12 NM_173850	rs61758960	c.608T>C (p.Leu203Pro)	0.004	D	D.C.	D	D	D	VUS

VITT03	EPB41 NM_001166005	rs150835844	c.1228G>T (p.Asp410Tyr)	0.0002479	D	D.C.	D	D	D	VUS
VITT03	PDIA6 NM_005742	rs768426261	c.256C>T (p.His86Tyr)	0.00007778	T	D.C.	D	D	T	VUS
VITT03	DST NM_001144769	NA	c.901C>T (p.Pro301Ser)	N/A	D	D.C.	D	-	D	VUS
VITT03	SMPD1 NM_000543	rs142215226	c.340G>A (p.Val114Met)	0.001060	D	D.C.	N	-	D	VUS
VITT05	STAB2 NM_017564	rs1030976854	c.2065C>G (p.Pro689Ala)	0.000008793	T	D.C.	D	D	D	VUS
VITT05	CD47 NM_001777	rs369671811	c.311A>G (p.Asp104Gly)	0.00005315	T	P	D	P.D.	D	VUS
VITT06	FGA NM_000508	N/A	c.2068C>G (p.Gln690Glu)	N/A	D	D.C.	N	B	T	VUS
VITT06	VASP NM_003370	rs202205375	c.1130G>C (p.Arg377Pro)	0.00001771	T	D.C.	N	D	D	VUS
VITT06	TBXAS1 NM_001166253	rs759354516	c.1673T>C (p.Leu558Pro)	0.00005285	T	D.C.	D	D	T	VUS
VITT06	PIK3CA NM_006218.4	N/A	c.834G>A (p.Met278Ile)	NA	T	D.C.	N	B	T	VUS
VITT06	LIMS2 NM_017980	rs770469332	c.254G>A ((p.Arg85Gln)	0.0000177	D	D.C.	D	P.D.	D	VUS
VITT06	VCL NM_014000.3	rs754046223	c.1298G>A (p.Arg433His)	N/A	T	D.C.	N	P.D.	D	VUS

VITT06	ENTPD1 NM_001164178	rs142591047	c.1226T>C (p.Ile409Thr)	0.000007742	T	P	D	B	T	VUS
VITT18	GP6 NM_001083899	rs779660245	c.830G>A (p.Arg277Gln)	<0.0001	T	P	D	D	D	VUS
VITT18	SERPINA12 NM_173850	rs146053420	c.1062C>G (p.His354Gln)	0.0008485	D	D.C.	D	D	D	VUS
VITT18	MFSD2B NM_001346880	rs138903557	c.1000C>G (p.Pro334Ala)	0.003837	D	D.C.	D	D	D	VUS
VITT18	NID2 NM_007361	N/A	c.2353A>G (p.Thr785Ala)	N/A	D	P	D	B	D	VUS
VITT18	COLGALT1 NM_024656	rs764429704	c.487C>G (p.Leu163Val)	0.00002324	D	D.C.	N	P.D.	D	VUS
VITT18	WDFY4 NM_020945	rs765608663	c.193C>T (p.Arg65Cys)	0.0001555	T	P	D	D	D	VUS
VITT18	SAMD9L NM_152703	rs776780720	c.1076G>A (p.Arg359Gln)	0.00001765	T	P	N	D	D	VUS
VITT021	LRTM2 NM_001163926	rs919752328	c.1046G>A (p.Arg349His)	N/A	T	D.C.	N	D	D	VUS
VITT021	COL5A1 NM_001278074	rs201997623	c.404C>T (p.Ser135Phe)	0.00001763	T	D.C.	D	D	D	VUS

FATHMM: T=Tolerated, D=Damaging **Mutation Taster:** P=Polymorphism; D.C.=Disease Causing **Provean:** N=Neutral; D=Damaging **Polyphen:** B=Benign Variant; P.D.=Possibly/Probably Damaging; D=Damaging **SIFT:** T=Tolerated; D=Damaging **ACMG classification:** VUS= Variant of Uncertain Clinical Significance; **LB=** Likely Benign Variant, **B=** Benign Variant

6. DISCUSSION

6.1. Covid-19 Study

In this thesis 80 Covid-19 patients, among whom 40 subjects admitted to ordinary hospital wards (non-ICU) and 40 patients admitted to intensive care unit or died (ICU/death), suffering from Covid-19, were analysed by high-throughput sequencing. NGS analysis results showed the presence of 212 genetic variants in the 7 genes included in the virus entry-related panel. Forty-two out of the 212 variants were rare ($MAF \leq 1\%$) and have been identified in 34 out of the 80 COVID-19 patients. The most affected genes were *TMPRSS2* and *CTSB*, accounting for 31% and 24% of the total rare variants, respectively. Among rare variants identified in *TMPRSS2* gene, the greater percentage was represented by missense variants (61.5%). Moreover, *TMPRSS2* gene accounted the largest number of VUS/LP variants.

We further analyzed population subgroups based on the severity of clinical outcomes (non-ICU vs ICU/death patients). *TMPRSS2* showed a higher percentage of variants in the non-ICU group (40%) compared to the ICU/death group (23%). Conversely, *CTSB* exhibited an opposite trend, with an increase from 15% in the non-ICU group to 32% in the ICU/death group. Despite fewer *TMPRSS2* variants in the ICU/death group, this gene presented a higher proportion of VUS/LP variants in patients with worse outcomes (1/8 vs 2/5). Additionally, the *HSPA5* gene displayed two VUS/LP variants in the ICU/death group and none in the non-ICU group.

These three genes, in particular, characterize two interesting cases. The first one, COV174, a 62-year-old male with comorbidities of hypertension and dyslipidemia, was admitted to intensive care due to severe pulmonary involvement and presented with two missense variants in the *TMPRSS2* and *HSPA5* genes. Both variants were partially classified as deleterious by *in silico* predictors and classified as variants of uncertain significance (VUS) according to ACMG criteria. Specifically, the heterozygous variant *TMPRSS2* c.415G>T leads to the substitution of Alanine at position 139 with Serine (p.Ala139Ser) in the amino acid sequence, and it is located in the transmembrane region of the protein. The protein has predominantly proteolytic activity but has also been shown to play an important role in facilitating the fusion of viral pathogens (influenza, CoV) to the host cell membrane. In fact, data from literature reported that, like the SARS virus CoV-1, the CoV-2 virus enters cells by attachment to ACE2 receptors and the viral spike protein is processed (or primed) by the cellular protease *TMPRSS2*. The authors also showed that inhibitors of *TMPRSS2* could block viral entry in cell culture, reinforcing its importance in the life cycle of the virus⁹⁴.

As concerns the other variant identified in COV174 patient, it is a heterozygous variant in the

HSPA5 gene (c.1778C>T) which determines the substitution of Threonine with Isoleucine at position 593 of the protein (p.Thr593Ile), not previously reported in the literature. Heat Shock Protein 5 (HSPA5), a member of the Hsp70 family of chaperones, is localized in the endoplasmic reticulum and plays a key role in regulating the Unfolded Protein Response (UPR), a process that reorganizes or degrades misfolded proteins accumulating in the ER¹⁹³. Under specific intracellular stress conditions, such as the accumulation of misfolded proteins, HSPA5 may translocate to the cytoplasm and cell membrane. HSPA5 proteins exposed on the cell surface (CS-HSPA5) are characteristic of several types of cancer, including breast, ovarian, pancreatic, and colon cancers, and are also capable of facilitating the entry of viral pathogens such as Zika, Dengue, Hepatitis C, Papilloma virus, Ebola, and MERS-CoV¹⁹⁴. Based on these premises, studies have hypothesized that the binding of CS-HSPA5 to the SARS-CoV-2 spike protein may facilitate viral attachment to the membrane and subsequent entry into the cell. Significantly higher expression levels of CS-HSPA5 have been observed in lung cancer patients, who are more susceptible to developing severe forms of SARS-CoV-2 infection. It has been suggested that the virus may have a greater propensity to invade tissues with a high concentration of HSPA5 on the cell surface, indicating it as a potential therapeutic target¹⁵⁷. Ibrahim and collaborators have recently identified the binding sites of the spike protein on the HSPA5 receptor through cellular docking studies combined with structural bioinformatics analysis, highlighting that regions III (residues C391-C525) and I (C480-C488) may facilitate spike protein binding¹⁹⁵. The mutation we identified is located near region III, which includes the substrate-binding domain (amino acids 420-500), for which the crystal structure is available. Simulation using Pymol software indicates that following the mutation, the affected amino acid, Thr593, loses one of its bonds with glutamic acid at position 589 (figure 38). Both residues are approximately 100 amino acids away from the portion of the protein involved in SARS-CoV-2 binding, so it is not possible to determine the effect of this mutation on viral attachment to the host cell. Nonetheless, this is the first potentially deleterious variant identified in this gene.

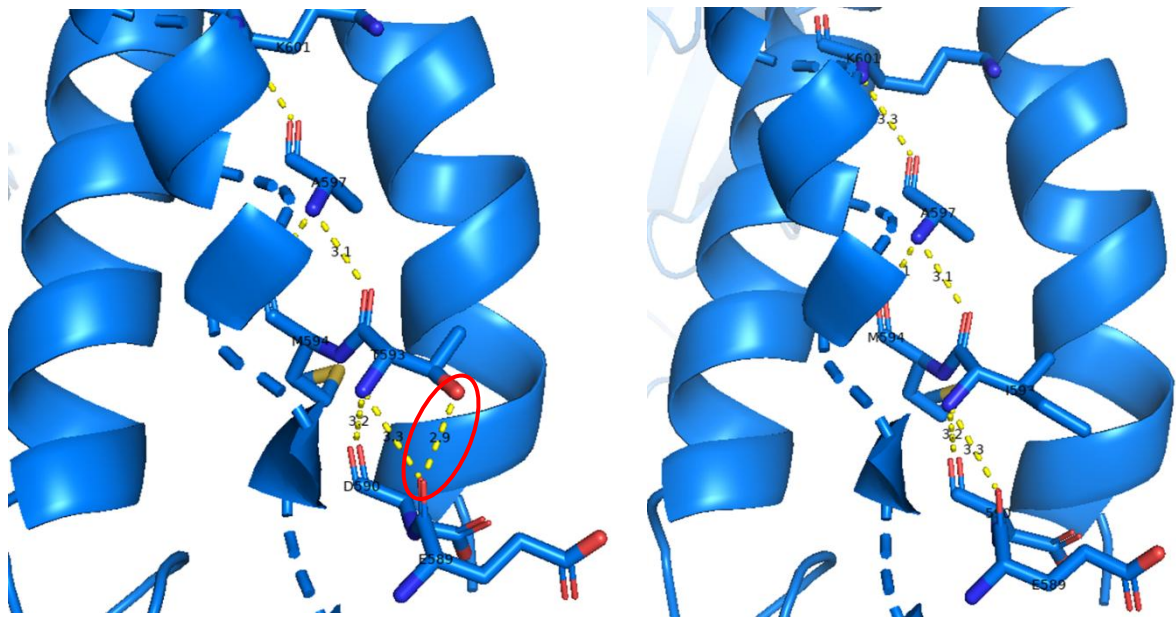


Figure 38. Structural model of the wild-type and mutated (WT) protein (p.Thr593Ile). Pictured on the left is the WT form of the protein, and on the right the mutated form with the replacement of a Threonine by an Isoleucine. In red, in the WT protein, the hydrogen bond between the Threonine 593 residue and the Glutamate 589 residue is highlighted, which is lost in the mutated form with the T593I substitution (Images created with PyMOL software).

The second interesting case is patient COV048, a 65-year-old male who was hospitalized in intensive care for an extended period due to deep vein thrombosis and pulmonary embolism, with no pre-existing comorbidities. He presented with two rare heterozygous variants in the *CTSB* and *TMPRSS2* genes. The *CTSB* (c.886G>A) variant, located in exon 9 of the gene, results in the substitution of Valine at position 296 with Isoleucine (p.Val296Ile). This variant is recorded in dbSNP (rs200851570) and is situated in the portion of the protein corresponding to the cathepsin B heavy chain. It is classified as a variant of uncertain significance (VUS) according to ACMG criteria. Cathepsins B (as cathepsin L) is a lysosomal protease involved in the entry of the SARS-CoV-2 virus into human cells. Studies have shown that the virus can enter cells through proteolytic cleavage of the spike protein, which is facilitated not only by the *TMPRSS2* protease, but also by cathepsins B and L. These proteins are crucial because they allow the virus to enter the cytoplasm of cells after fusion with the membrane.^{160,161}

The patient also carried another rare variant in the *TMPRSS2* gene (c.470_477dupGCGACTCC), which leads to an amino acid substitution from Serine to Alanine at position 160, and the introduction of a premature stop codon after 81 aminoacids (p.Ser160AlafsTer81). Protein breakdown occurs in the Cysteine-rich domains that are likely to mediate protein-protein interactions and ligand binding. This variant is classified as likely pathogenic based on ACMG criteria. Interestingly, patient COV048 also presented with two rare variants in genes involved in

the coagulation cascade. In fact, a parallel study conducted by our research group analyzed the genetic profiles of the same 80 patients using the same NGS technology, employing a panel of 11 genes directly involved in haemostatic mechanisms. The first of the two variants identified in patient COV048 is a missense mutation [c.218C>T; (p.Pro73Leu)] in the *SERPINC1* gene [Serpine peptidase inhibitor, clade C (antithrombin) member 1, OMIM*107300]. This variant is classified as likely pathogenic according to ACMG criteria. The protein encoded by this gene primarily functions as a natural anticoagulant within the coagulation system, inhibiting several activated factors, including IXa, Xa, XIa, XIIa, and thrombin. Heterozygous mutations in this gene can lead to impaired inhibition of factor X and thrombin, subsequently increasing the generation and activity of the latter¹⁹⁶. More than 250 genetic alterations have been identified in patients with antithrombin deficiency, including missense and nonsense point mutations, as well as small or large insertions and deletions¹⁹⁷. A particular subgroup of these alterations is due to mutations in heparin-binding sites¹⁹⁷. The variant identified in our study falls within this category of congenital antithrombin defects, as it involves the highly conserved proline residue at position 73 of the polypeptide chain, which, following substitution, loses its ability to bind heparin, thus interfering with its anticoagulant activity. This same mutation was first described in 1985¹⁹⁸, and was identified in 2012 in a patient with recurrent thrombosis at a young age¹⁹⁹, and in 2014 in 37 out of 42 families with type II antithrombin deficiency in a study conducted on the Finnish population²⁰⁰.

The patient also had a missense mutation [c.596C>T (p.Thr199Ile)] in the *F10* gene [Coagulation factor X, OMIM*613872], which encodes the vitamin K-dependent coagulation factor X, responsible for catalyzing the conversion of prothrombin to thrombin in the presence of activated factor V (Va), calcium ions, and phospholipids during clot formation. Factor X is converted into its activated form (Xa) via cleavage by factor IXa (in the intrinsic coagulation pathway) or factor VIIa (in the extrinsic pathway). The mutation identified in patient COV048 has a minor allele frequency of 0.0000649 in the European population according to gnomAD and is classified as VUS by ACMG guidelines. Threonine at position 199 plays a crucial role in the protein's structure, as it undergoes post-translational modification (O-glycosylation), involving the addition of an O-N-acetylgalactosamine (O-GalNac) chain to this residue, catalyzed by the enzyme GalNac transferase.

Given the importance of this residue in protein structure, the concurrent presence of this mutation with the pathogenic mutation in *SERPINC1* previously described, and the known regulatory mechanisms exerted by the protein encoded by *SERPINC1* on the activated form of factor X, the role of the *F10* variant in modulating the patient's phenotype cannot be ruled out. Therefore, the presence of these two variants together with variants identified in *CTSB* and *TMPRSS2* genes (viral entry NGS gene panel) likely facilitating viral entry and viral cycle, might have contributed

to the worsening of the patient's clinical condition.

As concerns other genes of the viral entry NGS gene panel investigated, *ACE2* gene represents the most studied gene in the literature, as the protein it encodes is the one most involved in the mechanism of virus entry into the host cell. In the COVID-19 cohort investigated (n=80), 3 rare variants were found. The only VUS variant (c.1019A>G, p.Gln340Arg) was identified in a patient with a more favourable clinical course. The other two variants [c.77A>G (p.Lys26Arg) and c.1665-7del] are classified as Likely Benign according to ACMG criteria and were found in patients admitted to the ICU. In particular, the patient COV452 (81-year-old woman, deceased in ICU) exhibited a homozygous rare variant (c.77A>G, p.Lys26Arg) variant, located at a peptidase site and previously associated with a possible effect on the ability of ACE2 to bind Spike glycoprotein on the viral wall¹²⁴. Furthermore, the same variant has been found more frequently in the literature in the Italian population than in the Asian population, and this might contribute to explain the higher mortality rates in Italy during the onset of the pandemic¹²². In COV452 subject, the homozygous polymorphism rs2285666 (c.439+4G>A), previously associated with increased disease severity and IL-6 overproduction, was also detected¹³¹. In fact, the subject had high serum levels of IL-6 at the time of admission (94 pg/ml). Given the information in the literature and the concomitant presence of these two variants in COV452 patient, it is not possible to exclude a role for them in modulating the virus entry and life cycle mechanism with consequent potential modulation of the patient's phenotype.

Regarding the other genes in the panel (*IL6*, *FURIN*), the following VUS variants were identified: *IL6* c.11_13del (rs1034241131) and *FURIN* c.230T>C (p.Leu77Pro, rs143659625). No VUS variants were identified in *CTSL* gene.

It is known from literature data the biological importance of these genes in the mechanism of viral infection. Furin is a member of the subtilisin-like peptidase family that, together with TMPRSS2, appears to play a key role in the proteolytic activation of SARS-CoV-2, indicating their selective inhibition as a potential therapeutic target for the treatment of infection¹¹⁷. As discussed above, intronic (c.372+5G>A) and missense variants (c.128C>T, p.Ala43Val; c.436G>A, p.Gly146Ser; c.893G>A, p.Arg298Gln; c.1906A>G, p.Ile636Val) at the level of this gene have been identified in SARS-CoV-2-positive patients in Italy¹⁵³. Some studies have suggested that polymorphisms in *FURIN* might influence the risk and severity of COVID-19 infection. For instance, variants such as rs6226 and rs4702 have been identified in genetic analyses as potential factors associated with a higher risk of developing severe forms of COVID-19. However, the evidence regarding the precise impact of these variants is still not completely clear and further studies are needed to better understand their role in the disease^{154,155}. Our results are in agreement with reports in the literature that variants in this gene do not seem to be associated with a worse progression of the clinical picture from SARS-CoV-2 infection.

IL-6 is part of a superfamily of cytokines involved in cell signalling and mediating multiple inflammatory processes in the lung with both anti- and pro-inflammatory activities. An association between genetic variants of the interleukin-6 (*IL6*) gene and the severity of COVID-19 has been reported in some works. In particular, the polymorphism on the *IL6* promoter region rs1800795 (c.-174G > C) was studied in relation to the severity of COVID-19. This polymorphism has been associated with an increased production of IL-6, which may contribute to the so-called 'cytokine storm'.¹³² Furthermore, another study in paediatric patients showed an association between specific *IL6* polymorphisms and systemic inflammation, demonstrating that the severity of COVID-19 may also be related to these genetic variants in children¹³⁶. However, research is still ongoing and other genetic and environmental factors contribute to the complexity of the response to the virus, including inflammatory markers such as C-reactive protein (CRP) and D-dimer, which have been associated with disease severity in several studies¹³⁵.

Based on the abovementioned information concerning the role of proteins encoded by these genes in the viral infection, the possible role of variant identified in the COVID-19 cohort studied cannot be excluded, although needing further confirmation concerning the functional role of variants identified.

Overall, the findings highlight how genetic variants in the seven genes (*ACE2*, *TMPRSS2*, *CTSL*, *CTSB*, *HSPA5*, *IL6*, *FURIN*), selected based on their involvement in SARS-CoV-2 entry into host cells and their documented associations with COVID-19, underscore the presence of genetic variability among individuals that may contribute to modulating susceptibility to viral infection.

In certain cases, the presence of loss-of-function mutations, which can be correlated with a loss or reduction of activity in various molecular components critical for viral entry into the cell, may confer a potential protective effect. Conversely, gain-of-function variants or those that positively modulate the expression level of these genes may be associated with increased susceptibility to infection. Further investigation into the roles of the different variants identified, along with a comprehensive assessment of the potential contribution of common genetic variants, may enable the identification of an allelic burden of risk, helping to identify individuals with higher or lower susceptibility to SARS-CoV-2 infection.

Lp(a) and Covid-19

Lipoprotein(a) [Lp(a)] is an atherogenic lipid particle composed of a polymorphic glycoprotein called apo(a), covalently bound to a low-density lipoprotein (LDL) portion containing apoB100. Given its potential prothrombotic and proinflammatory properties, Lp(a) has been suggested as a

possible contributor to a worse clinical outcome in patients with COVID-19.^{63,64}

Given the conflicting opinions in the literature.³⁻⁵, with the aim of detecting any relationship between genetic determinants of Lp(a) and Covid-19 severity, on the same COVID population the KIV2 repeat polymorphism was studied. Genetic characterisation of the LPA KIV2 polymorphism produced an estimate of the total number (sum of repeats on both alleles) of KIV2 repeats. Since over 90% of the variance in Lp(a) levels is genetically determined, with the KIV2 repeat polymorphism being the main determinant⁷⁵, we have decided to evaluate this CNV in our cohort with ddPCR technique. As concluded in a work from our laboratory in fact, a significant inverse correlation between Lp(a) levels and CNV values was confirmed using ddPCR⁷⁵.

The analysis carried out into COVID-19 cohort (n=80) showed no relationship between KIV2 repeats and disease severity (ICU admission/death and non-ICU admission, thrombotic event) in the 80 patients in our cohort. No differences were observed for the number of KIV2 repeats according to disease severity, neither comparing ICU/death patients with non-ICU patients, nor comparing patients with thrombosis with patients without thrombosis.

Data obtained are in keeping with those reported by Bianconi et al. who reported that Lp(a) does not influence biomarkers of plasma thrombotic activity and systemic inflammation nor has any impact on thrombotic events and adverse clinical outcomes in patients hospitalized for COVID-19³. Conversely, conflicting results with data reported by Pawlos and co-workers were observed. They concluded that patients with elevated Lp(a) level had significantly longer hospitalization time and higher oxygen demand on admission. Indeed, elevated Lp(a) was also significantly associated with intubation, ICU and Death rate⁴.

Some limitations of the study should be mentioned, such as the lack of plasma Lp(a) levels and the sample size of the population investigated. An additional limitation is the inability of the technique to discriminate CNVs specific to each allele. In fact, ddPCR can accurately detect the total number of CNVs present across both alleles, but it does not allow to obtain information concerning the protein encoded by each allele (higher or lower molecular weight isoforms).

This study could be a further incentive to continue with the genetic evaluation of the KIV2 repeats polymorphism, given the difficulties in apo(a) isoforms determination. Actually, the high degree of dimensional heterogeneity of apo(a) causes measurement bias, making the development of accurate methods for measuring plasma Lp(a) levels extremely complex.^{69,70}

In addition, confirming the use of ddPCR would lead to a considerable reduction in time and total costs compared to the use of “gold standard” techniques such as Western Blot and Pulsed Field Gel Electrophoresis (PFGE).

6.2. VITT Study

This study addresses for the first time whether rare genetic variants could contribute to the development of VITT, a clinical condition mainly associated with COVID-19 adenoviral vector vaccine administration and its adverse complications through a WES approach.

Due to the rare occurrence of VITT, in this study, we focused the attention on rare variants. At a whole analysis of the rare variants identified in the 6 patients with VITT, no significant enrichment of mutated genes in biological/functional pathways common to different analyzed subjects was found. Therefore, we focused on potentially damaging mutations on the basis of ACMG criteria, occurring in selected genes involved in biological and functional processes associated with blood coagulation and fibrinolysis, platelet activation and aggregation, integrin-mediated signaling pathway, inflammation with particular attention to those involved in vascular damage, and autoimmune thrombocytopenia. Actually, selected pathways have been described to contribute to a major push in VITT and in its complication severity^{9,11,201,202}.

For each patient, we identified different VUS ranging from 4 to 13 variants. VITT05 and VITT18 patients exhibited the worst clinical outcome.

In VITT05, a rare VUS in the *STAB2* gene [c.2065C>G (p.Pro689Ala)] has been classified as harmful by most in silico predictors used. *STAB2* gene encodes stabilin-2, expressed predominantly in the sinusoidal endothelium of the liver and spleen and acting as a systemic scavenger receptor for heparin, chondroitin sulfate, dermatan sulfate, nonglycosaminoglycan, acetylated low-density lipoprotein, procollagen propeptides, and advanced glycation end products^{203,204}. Previous WES data also showed an association of *STAB2* locus with venous thromboembolic disease²⁰⁵; really, previous data demonstrated that stabilin-2, beyond its role as a scavenger receptor for the aforementioned components, acts as a clearance receptor for VWF, supporting the connection between stabilin-2 and VWF plasma levels²⁰⁵. Therefore, it could be hypothesized that *STAB2* genetic variants might affect stabilin-2 cell surface expression by possibly inducing protein misfolding or nonsense-mediated decay, thus influencing molecular pathways stabilin-2 mediated and, in turn, contributing to the development of a prothrombotic phenotype. Indeed, it has been observed a role for stabilin-2 in influencing the thrombus incidence in association with elevated VWF plasma levels, potentially by promoting the interaction of platelets and leukocytes with the vessel wall²⁰⁶.

VITT05 patient was also a carrier of another rare variant in the *CD47* gene [c.311A>G (p.Asp104Gly)], encoding the leukocyte surface antigen CD47, for which a variable prediction of pathogenicity by in silico tools was evidenced. The contribution of the *CD47* gene might be

considered critical, as well; really, the leukocyte surface antigen CD47 plays a role as a receptor for thrombospondin, a secreted matricellular protein observed to be upregulated in vascular cells after injury as well as in chronic disease, and modulator of integrin signaling. Indeed, the CD47-TSP1 binding has been previously shown to influence thrombus formation leading events, including the adhesion of platelets to the endothelial wall, nitric oxide/cGMP (cyclic guanosine monophosphate) signaling induction, platelet activation, and aggregation^{207–209}. VITT05 also exhibited two further VUS in genes involved in inflammation pathways. A rare variant in *COL6A1* gene [c.2191C>T (p.Arg731Cys)], encoding for alpha-1 chain of collagen VI, for which in silico tools used for pathogenicity prediction established a damaging effect, was detected. Data from the literature showed that the *COL6A1* gene was upregulated in patients with nonemphysematous chronic obstructive pulmonary disease, so supporting the profibrotic mast cell phenotype, through collagen VI deposition²¹⁰. Moreover, a further variant was found in the *IL5RA* gene [c.913G>A p.(Asp305Asn)], encoding the interleukin-5 receptor subunit alpha, involved in the survival, differentiation, and chemotaxis of eosinophils, contributing to modulate the innate immune system²¹¹.

VITT18 is a patient with a severe disability as a consequence of cerebral stroke with intraparenchymal hemorrhage. Interestingly, this patient carries a missense variant [c. 830G>A (p.Arg277Gln)] in *GP6* gene encoding for platelet glycoprotein VI (GPVI), a collagen receptor involved in collagen-induced platelet adhesion and activation, contributing to platelet procoagulant activity and thrombin and fibrin formation. The GPVI represents a receptor expressed on platelets and megakaryocytes and acts as a receptor for collagen as well as several plasma and vascular proteins (i.e., laminin, fibronectin, and galectin-3). Data from the literature suggested the contribution of platelet GPVI in influencing venous thrombotic complications; really GPVI might sustain the thrombo-inflammatory status by inducing neutrophil granular release. Moreover, it has been suggested the role of a GPVI-thrombin-fibrin feedforward loop supporting the thrombus formation²¹². Apart from the *GP6* gene variant, VITT18 patient also carries two rare missense variants in *SERPINA12* and *MFSD2B* genes [c.1062C>G (p.His354Gln) and c.1000C>G (p.Pro334Ala)], both classified as damaging by all in silico prediction tools. *SERPINA12* gene product, vaspin, is a member of the large family of serine protease inhibitors and has been observed to inhibit target protease kallikrein 7 (KLK7), the process accelerated by the presence of heparin²¹³. Actually, it has been evidenced that serum vaspin is an independent prognostic marker of major adverse cardiac events (cardiovascular death, recurrent acute myocardial infarction (AMI), or hospitalization for heart failure) in AMI patients. In particular, low vaspin levels have been found to be associated with increased inflammation in AMI patients

²¹⁴.

As concerns *MFSD2B*, it encodes the sphingosine-1-phosphate transporter, able to mediate the export of sphingosine-1-phosphate in red blood cells and platelets²¹⁵. A mouse model highlighted that *Mfsd2b* knock-out could influence platelet intrinsic functions, thus supporting its possible role in limiting thrombus formation²¹⁶. In particular, the mouse model suggested that sphingosine-1-phosphate signaling could be involved in platelet biogenesis. *Mfsd2b* has been observed to be required for sphingosine-1-phosphate release in resting and activated platelets, really *Mfsd2b* knock-out is associated with a reduced platelet intrinsic function and, in turn, a reduced thrombosis in mice. Therefore, gain-of-function variants in the *MFSD2B* gene might contribute to a prothrombotic phenotype. VITT18 phenotype might be also supported by the presence of two further VUS: a variant in the *BCL6* gene, encoding a transcriptional repressor, able to influence T cell proliferative capacity and fate, and a variant in *TRPV1* gene, encoding a ligand-activated nonselective calcium permeant cation channel involved in the detection of noxious chemical and thermal stimuli. Data from mouse models reported that *Bcl6* expression was reduced in the presence of mir-155 (microRNA 155) deficiency in advanced atherosclerosis, supporting its involvement in the modulation of the atherosclerotic process²¹⁷. Moreover, *BCL6* has been observed to be involved in live-attenuated influenza vaccine (LAIV)-induced T follicular helper cell differentiation, influencing the antibody response²¹⁸. As concerns the *TRPV1* gene, it has been shown that brain channels are activated by ischemic stroke, causing neurological and motor deficits as well as infarction after brain ischemia²¹⁹. Moreover, it has been reported that *TRPV1* is present in platelets, thus representing a substrate for platelet activation by inflammatory mediators in atherosclerotic plaques²²⁰.

In VITT02 patient, in whom both venous and arterial thrombosis occurred, a VUS with all damaging predictors of pathogenicity [c.608T>C (p.Leu203Pro)] was found in the *SERPINA12* gene. In addition, in VITT02 patient, there has been evidence of the presence of VUS in genes involved in two strictly related biological processes: (cell-matrix adhesion [*ITGA2B*, *ITGAD*, and *FGA*, gene ontology term GO:0007160] and blood coagulation [*THBD* and *FGA*, GO:0072377]). Interestingly, VITT02 exhibited the presence of two variants in genes encoding for integrin family members, such as *ITGA2B*, encoding for the integrin alpha-IIb and associated with Glanzmann thrombasthenia (OMIM 273800), and *ITGAD*, encoding for the integrin alpha-D. Altogether, these loci might contribute to explain the thrombotic profile. Among other variants identified, two have not been previously described in the general population; the first one is a rare missense variant (c.1465G>C) in *THBD* gene, encoding for thrombomodulin, an endothelial-specific type I membrane receptor involved in the regulation of the coagulation process, determining the substitution of the Aspartic Acid at position 489 with an Histidine (p.Asp489His) located in the extracellular region in the serine/threonine-rich domain, a binding side for alpha-L/beta-2 and

alpha-M/beta-2 integrin²²¹; the second one is a rare missense variant [c.1132T>A (p.Ser378Thr)] in the *FGA* gene, encoding for the fibrinogen α -chain. As concerns the *THBD* variant, the serine/threonine-rich domain represents a critical domain for protein anticoagulant activity, and data concerning other variants previously identified in the same domain possibly support its role. Further VUSs have been detected in the *PIEZO1* gene, encoding the piezo-type mechanosensitive ion channel component 1, which was shown to be a key modulator of abnormal platelet activation and thrombosis under hypertension²²², and *GLCE* gene, encoding the D-glucuronyl C5-epimerase, involved in heparan sulfate/heparin biosynthesis²²³. In VITT02 patient, uncertain significant rare variants have also been detected in the *CARD8* and *EGFR* genes, both involved in inflammation pathways. *CARD8* protein represents an inflammasome sensor, mediating the activation of the inflammasome according to pathogen-associated signals response. In particular, it has been reported that *CARD8* participates in the regulation of cytokines and chemokines expression in endothelial cells and atherosclerotic lesions²²⁴. As concerns the *EGFR* gene, it encodes a receptor tyrosine kinase binding ligands of the EGF family. Data from RNA sequencing analysis, performed in glioblastoma tissue specimens according to the thrombotic phenotype, showed a higher *EGFR* expression in the group with many thrombi with respect to that with few thrombi²²⁵.

Similarly to VITT02, a VUS in the *FGA* gene was also detected in VITT06 patient, also exhibiting both venous thrombosis and stroke. The missense variant identified in this patient [c.2068C>G (p.Gln690Glu)] has not previously been described and exhibited a variable pathogenicity prediction with the in silico tools used, thus supporting the possible contribution of fibrinogen chain alterations to a prothrombotic phenotype. This patient also exhibited VUS in *VASP*, *TBXAS1*, and *ENTPD1* genes. *VASP* gene encodes the vasodilator-stimulated phosphoprotein, actin-associated proteins involved in a range of processes dependent on cytoskeleton remodeling and cell polarity; in particular, it regulates actin dynamics in platelets and plays an important role in regulating platelet activation/aggregation^{226,227}. *TBXAS1* gene encodes the thromboxane-A synthase 1, an enzyme responsible for the conversion of prostaglandin H₂ (PGH₂) to thromboxane A₂ (TXA₂), a potent inducer of blood vessel constriction and platelet aggregation²²⁸. Data from the literature also evidenced the association of thromboxane A synthase 1 gene expression and promotor haplotypes with the risk of large artery-atherosclerosis stroke in the Iranian population²²⁹. As concerns *ENTPD1*, it encodes the ectonucleoside triphosphate diphosphohydrolase 1 observed to be implicated in the prevention of platelet aggregation by hydrolyzing platelet-activating ADP to AMP²³⁰. VITT06 patient also exhibited uncertain significance rare variants in *FFAR2*, *NDST1*, and *NLRP2* genes. *FFAR2* gene encodes a G protein-coupled receptor that is activated by a major product of dietary fiber digestion, the short-chain fatty acids (SCFAs), and that plays a role in the regulation of whole-body energy homeostasis and

in intestinal immunity. Previous data from the literature reported its role as a potential biomarker to predict AMI occurrence²³¹. *NDST1* encodes a member of the heparan sulfate/heparin GlcNAc N-deacetylase/N-sulfotransferase family. It has been previously reported in mouse models that modification of the heparan sulfate fine structure by *Ndst1* deletion contributes to decrease vascular smooth muscle cells (VSMC) proliferation, as well as vascular remodeling²³². *NDST1* deficiency has also been previously shown to correlate with allergic airway inflammation²³³. The *NLRP2* gene encodes a member of the nucleotide-binding and leucine-rich repeat receptor (NLR) family. It interacts with components of the I κ B kinase (IKK) complex and can regulate both caspase-1 and NF- κ B (nuclear factor kappa-light-chain-enhancer of activated B cells) activity. Previous data from the literature reported higher expression levels both in vitro and in vivo under ischemic stroke conditions²³⁴.

In VITT03, in whom a deep vein thrombosis was observed, a missense VUS [c.1228G>T(p.Asp410Tyr)] classified by all in silico predictors as damaging has been identified in *EPB41* gene, encoding for a structural element of the erythrocyte membrane skeleton, for which gene expression studies evidenced its contribution in the thrombus formation network²³⁵. Actually, a decreased erythrocyte membrane flexibility results in an increased red erythrocyte rigidity that might contribute to a thrombogenic potential²³⁶. Apart from *EPB41*, VITT03 also carries a further missense VUS in the *PDIA6* gene, encoding for the protein disulfide-isomerase A6, that, beyond its role as a molecular chaperone, able to inhibit aggregation of misfolded proteins, also plays a role in platelet aggregation²³⁷. VITT03 patient also carries uncertain significance rare variants in *AKT1* and *NOTCH2* genes. *AKT1* gene encodes one of the three members of the human AKT serine-threonine protein kinase family. It has been demonstrated to contribute to dendritic cell survival and maturation²³⁸. Moreover, *AKT1* mutations are associated with Proteus syndrome, characterized by an increased risk of thromboembolism²³⁹. *NOTCH2* gene encodes a member of the Notch family, playing a role in a variety of developmental processes by controlling cell fate decisions. Data from the literature reported that anti-PF4/heparin antibody production is impaired in B-cell-specific Notch2-deficient mice, lacking marginal zone (MZ) B cells²⁴⁰.

In VITT21 VUS in *COL5A1* and *LRTM2* loci have been observed; interestingly, the *LRTM2* gene encoded the leucine-rich repeat and transmembrane domain-containing protein 2 also involved in heparin binding. Moreover, this patient also displayed a further VUS in the *COL5A1* gene, encoding the alpha 1 chain of the type V collagen, able to bind thrombospondin as well as heparin^{241,242}, thus supporting its possible involvement in the clinical phenotype. VITT21 also

carries a VUS in the *ADAM8* gene. *ADAM8* gene encodes a member of the ADAM (a disintegrin and metalloprotease domain) family, a family of transmembrane glycoproteins with heterogeneous expression profiles and proteolytic, cell-adhesion, -fusion, and -signaling properties. Data from the literature reported that *ADAM8* has immunomodulation properties; in particular, it has been observed to be associated with the function of neutrophils during inflammatory responses²⁴³.

In conclusion, WES analysis performed in 6 patients with VITT identifies a number of rare variants which are conditions possibly favouring the prothrombotic state triggered by the exposure to the vaccine.

To the best of our knowledge, no other studies are present in the literature sequencing patients with VITT. Only a recent paper addressed the genetic characteristics of a VITT patient by genotyping more than 870,000 known SNPs in the human genome with an approach not able to identify unknown and rare variants²⁴⁴.

The observation of a burden of VUS for each VITT patient might raise the issue of whether one or more VUS are causative variants, and whether, instead, we should hypothesize a burden of susceptibility genetic factors in the modulation of the clinical phenotype after an acquired condition, such the exposure to adenovirus.

Data of the present study, to be confirmed and implemented, need functional studies and/or an extension to a larger number of patients, thus allowing a more comprehensive definition of these molecular pathways in modulating clinical manifestations of this condition also in consideration of a possible evaluation of the combined contribution of rare and common variants.

7. BIBLIOGRAPHY

1. Cosar, B. *et al.* SARS-CoV-2 Mutations and their Viral Variants. *Cytokine Growth Factor Rev.* **63**, 10–22 (2022).
2. MacLean, O. A., Orton, R. J., Singer, J. B. & Robertson, D. L. No evidence for distinct types in the evolution of SARS-CoV-2. *Virus Evol.* **6**, veaa034 (2020).
3. Bianconi, V. *et al.* Lipoprotein(a) Does Not Predict Thrombotic Events and In-Hospital Outcomes in Patients with COVID-19. *J. Clin. Med.* **12**, 3543 (2023).
4. Pawlos, A., Gorzelak-Pabiś, P., Staciwa, M. & Broncel, M. Elevated Lp(a) and course of COVID-19: Is there a relationship? *PLOS ONE* **17**, e0266814 (2022).
5. Kaltoft, M. *et al.* Lipoprotein(a) during COVID-19 hospitalization: Thrombosis, inflammation, and mortality. *Atherosclerosis* **357**, 33–40 (2022).
6. Sticchi, E. *et al.* Apolipoprotein(a) Kringle-IV Type 2 Copy Number Variation Is Associated with Venous Thromboembolism. *PLOS ONE* **11**, e0149427 (2016).
7. Sticchi, E. *et al.* Role of lipoprotein (a) and LPA KIV2 repeat polymorphism in bicuspid aortic valve stenosis and calcification: a proof of concept study. *Intern. Emerg. Med.* **14**, 45–50 (2019).
8. See, I. *et al.* US Case Reports of Cerebral Venous Sinus Thrombosis With Thrombocytopenia After Ad26.COV2.S Vaccination, March 2 to April 21, 2021. *JAMA* **325**, 2448 (2021).
9. Marcucci, R. & Marietta, M. Vaccine-induced thrombotic thrombocytopenia: the elusive link between thrombosis and adenovirus-based SARS-CoV-2 vaccines. *Intern. Emerg. Med.* **16**, 1113–1119 (2021).
10. EMA - Vaxzevria report.
11. Greinacher, A., Selleng, K. & Warkentin, T. E. Autoimmune heparin-induced thrombocytopenia. *J. Thromb. Haemost.* **15**, 2099–2114 (2017).
12. Jaax, M. E. *et al.* Complex formation with nucleic acids and aptamers alters the antigenic properties of platelet factor 4. *Blood* **122**, 272–281 (2013).
13. Giusti, B. *et al.* Whole Exome Sequencing in Vaccine-Induced Thrombotic Thrombocytopenia (VITT). *BioMed Res. Int.* **2024**, 2860547 (2024).
14. Huang, C. *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet* **395**, 497–506 (2020).
15. Zhu, N. *et al.* A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727–733 (2020).
16. Wu, F. *et al.* A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265–269 (2020).
17. Coronaviridae Study Group of the International Committee on Taxonomy of Viruses *et al.* The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat. Microbiol.* **5**, 536–544 (2020).
18. World Health Organization (WHO). <https://www.who.int>.
19. Cereda, D. *et al.* The early phase of the COVID-19 epidemic in Lombardy, Italy. *Epidemics* **37**, 100528 (2021).
20. Zhang, H., Penninger, J. M., Li, Y., Zhong, N. & Slutsky, A. S. Angiotensin-converting enzyme 2 (ACE2) as a SARS-CoV-2 receptor: molecular mechanisms and potential therapeutic target. *Intensive Care Med.* **46**, 586–590 (2020).
21. The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team, null. The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19) - China, 2020. *China CDC Wkly.* **2**, 113–122 (2020).
22. Wu, Z. & McGoogan, J. M. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA* **323**, 1239 (2020).

23. Clerkin, K. J. *et al.* COVID-19 and Cardiovascular Disease. *Circulation* **141**, 1648–1655 (2020).
24. Bonow, R. O., Fonarow, G. C., O’Gara, P. T. & Yancy, C. W. Association of Coronavirus Disease 2019 (COVID-19) With Myocardial Injury and Mortality. *JAMA Cardiol.* **5**, 751 (2020).
25. Hendren, N. S., Drazner, M. H., Bozkurt, B. & Cooper, L. T. Description and Proposed Management of the Acute COVID-19 Cardiovascular Syndrome. *Circulation* **141**, 1903–1914 (2020).
26. Lindner, D. *et al.* Association of Cardiac Infection With SARS-CoV-2 in Confirmed COVID-19 Autopsy Cases. *JAMA Cardiol.* **5**, 1281 (2020).
27. Varga, Z. *et al.* Endothelial cell infection and endotheliitis in COVID-19. *The Lancet* **395**, 1417–1418 (2020).
28. Menter, T. *et al.* Postmortem examination of COVID-19 patients reveals diffuse alveolar damage with severe capillary congestion and variegated findings in lungs and other organs suggesting vascular dysfunction. *Histopathology* **77**, 198–209 (2020).
29. Lala, A. *et al.* Prevalence and Impact of Myocardial Injury in Patients Hospitalized With COVID-19 Infection. *J. Am. Coll. Cardiol.* **76**, 533–546 (2020).
30. Rettew, A., Garrahy, I., Rahimian, S., Brown, R. & Sangha, N. COVID-19 Coagulopathy. *Life* **14**, 953 (2024).
31. Gómez-Mesa, J. E., Galindo-Coral, S., Montes, M. C. & Muñoz Martin, A. J. Thrombosis and Coagulopathy in COVID-19. *Curr. Probl. Cardiol.* **46**, 100742 (2021).
32. Garigliany, M. *et al.* SARS-CoV-2 Natural Transmission from Human to Cat, Belgium, March 2020. *Emerg. Infect. Dis.* **26**, 3069–3071 (2020).
33. Halfmann, P. J. *et al.* Transmission of SARS-CoV-2 in Domestic Cats. *N. Engl. J. Med.* **383**, 592–594 (2020).
34. Oreshkova, N. *et al.* SARS-CoV-2 infection in farmed minks, the Netherlands, April and May 2020. *Eurosurveillance* **25**, (2020).
35. Wang, W. *et al.* Detection of SARS-CoV-2 in Different Types of Clinical Specimens. *JAMA* (2020) doi:10.1001/jama.2020.3786.
36. Van Doremalen, N. *et al.* Aerosol and Surface Stability of SARS-CoV-2 as Compared with SARS-CoV-1. *N. Engl. J. Med.* **382**, 1564–1567 (2020).
37. International Committee on Taxonomy of Viruses: ICTV.
38. Thachil, J. *et al.* ISTH interim guidance on recognition and management of coagulopathy in COVID-19. *J. Thromb. Haemost.* **18**, 1023–1026 (2020).
39. Lodigiani, C. *et al.* Venous and arterial thromboembolic complications in COVID-19 patients admitted to an academic hospital in Milan, Italy. *Thromb. Res.* **191**, 9–14 (2020).
40. Singh, B. *et al.* COVID-19 Infection and Arterial Thrombosis: Report of Three Cases. *Ann. Vasc. Surg.* **70**, 314–317 (2021).
41. Choudry, F. A. *et al.* High Thrombus Burden in Patients With COVID-19 Presenting With ST-Segment Elevation Myocardial Infarction. *J. Am. Coll. Cardiol.* **76**, 1168–1176 (2020).
42. Gheblawi, M. *et al.* Angiotensin-Converting Enzyme 2: SARS-CoV-2 Receptor and Regulator of the Renin-Angiotensin System: Celebrating the 20th Anniversary of the Discovery of ACE2. *Circ. Res.* **126**, 1456–1474 (2020).
43. Fried, J. A. *et al.* The Variety of Cardiovascular Presentations of COVID-19. *Circulation* **141**, 1930–1936 (2020).
44. Tadros, A., Hughes, D. P., Dunmore, B. J. & Brindle, N. P. J. ABIN-2 protects endothelial cells from death and has a role in the antiapoptotic effect of angiopoietin-1. *Blood* **102**, 4407–4409 (2003).
45. Favaloro, E. J. Laboratory testing for suspected COVID-19 vaccine-induced (immune) thrombotic thrombocytopenia. *Int. J. Lab. Hematol.* **43**, 559–570 (2021).
46. Santos, R. A. S. *et al.* The ACE2/Angiotensin-(1–7)/MAS Axis of the Renin-Angiotensin System: Focus on Angiotensin-(1–7). *Physiol. Rev.* **98**, 505–553 (2018).

- 47.Englisch, C. N., Tschernig, T., Flockerzi, F., Meier, C. & Bohle, R. M. Lesions in the lungs of fatal corona virus disease Covid-19. *Ann. Anat. - Anat. Anz.* **234**, 151657 (2021).
- 48.Mackman, N., Antoniak, S., Wolberg, A. S., Kasthuri, R. & Key, N. S. Coagulation Abnormalities and Thrombosis in Patients Infected With SARS-CoV-2 and Other Pandemic Viruses. *Arterioscler. Thromb. Vasc. Biol.* **40**, 2033–2044 (2020).
- 49.Khismatullin, R. R. *et al.* Pathology of lung-specific thrombosis and inflammation in COVID-19. *J. Thromb. Haemost.* **19**, 3062–3072 (2021).
- 50.SIIA - Società Italiana dell'Ipertensione Arteriosa Lega Italiana contro l'Ipertensione Arteriosa.
- 51.Eaton, D. L. *et al.* Partial amino acid sequence of apolipoprotein(a) shows that it is homologous to plasminogen. *Proc. Natl. Acad. Sci.* **84**, 3224–3228 (1987).
- 52.Kratzin, H., Armstrong, V. W., Niehaus, M., Hilschmann, N. & Seidel, D. Structural Relationship of an Apolipoprotein (a) Phenotype (570 kDa) to Plasminogen: Homologous Kringle Domains are Linked by Carbohydrate-Rich Regions. *Biol. Chem. Hoppe. Seyler* **368**, 1533–1544 (1987).
- 53.Kamstrup, P. R. Lipoprotein(a) and ischemic heart disease—A causal association? A review. *Atherosclerosis* **211**, 15–23 (2010).
- 54.Kamstrup, P. R., Tybjaerg-Hansen, A. & Nordestgaard, B. G. Elevated Lipoprotein(a) and Risk of Aortic Valve Stenosis in the General Population. *J. Am. Coll. Cardiol.* **63**, 470–477 (2014).
- 55.Kamstrup, P. R. & Nordestgaard, B. G. Elevated Lipoprotein(a) Levels, LPA Risk Genotypes, and Increased Risk of Heart Failure in the General Population. *JACC Heart Fail.* **4**, 78–87 (2016).
- 56.Dentali, F. *et al.* Lipoprotein(a) as a Risk Factor for Venous Thromboembolism: A Systematic Review and Meta-analysis of the Literature. *Semin. Thromb. Hemost.* **43**, 614–620 (2017).
- 57.Maloberti, A. *et al.* Lipoprotein(a): Cardiovascular Disease, Aortic Stenosis and New Therapeutic Option. *Int. J. Mol. Sci.* **24**, 170 (2022).
- 58.Argraves, K. M., Kozarsky, K. F., Fallon, J. T., Harpel, P. C. & Strickland, D. K. The atherogenic lipoprotein Lp(a) is internalized and degraded in a process mediated by the VLDL receptor. *J. Clin. Invest.* **100**, 2170–2181 (1997).
- 59.Wu, H. D. *et al.* High lipoprotein(a) levels and small apolipoprotein(a) sizes are associated with endothelial dysfunction in a multiethnic cohort. *J. Am. Coll. Cardiol.* **43**, 1828–1833 (2004).
- 60.Kronenberg, F. & Utermann, G. Lipoprotein(a): resurrected by genetics. *J. Intern. Med.* **273**, 6–30 (2013).
- 61.Kronenberg, F. Human Genetics and the Causal Role of Lipoprotein(a) for Various Diseases. *Cardiovasc. Drugs Ther.* **30**, 87–100 (2016).
- 62.Hopewell, J. C. *et al.* Lipoprotein(a) Genetic Variants Associated With Coronary and Peripheral Vascular Disease but Not With Stroke Risk in the Heart Protection Study. *Circ. Cardiovasc. Genet.* **4**, 68–73 (2011).
- 63.Moriarty, P. M. *et al.* Lipoprotein(a) and Its Potential Association with Thrombosis and Inflammation in COVID-19: a Testable Hypothesis. *Curr. Atheroscler. Rep.* **22**, 48 (2020).
- 64.Vuorio, A., Kaste, M. & Kovanen, P. T. Elevated Lipoprotein(a) and Cerebral Venous Sinus Thrombosis in COVID-19. *J. Stroke Cerebrovasc. Dis.* **30**, 105865 (2021).
- 65.Pirro, M. *et al.* Lipoprotein(a) and inflammation: A dangerous duet leading to endothelial loss of integrity. *Pharmacol. Res.* **119**, 178–187 (2017).
- 66.Ferretti, G. *et al.* Lipoprotein(a): A missing culprit in the management of athero-thrombosis? *J. Cell. Physiol.* **233**, 2966–2981 (2018).
- 67.Kamstrup, P. R. Genetically Elevated Lipoprotein(a) and Increased Risk of Myocardial Infarction. *JAMA* **301**, 2331 (2009).
- 68.Kamstrup, P. R., Tybjaerg-Hansen, A. & Nordestgaard, B. G. Extreme Lipoprotein(a) Levels and Improved Cardiovascular Risk Prediction. *J. Am. Coll. Cardiol.* **61**, 1146–1156 (2013).
- 69.Marcovina, S. M. & Albers, J. J. Lipoprotein (a) measurements for clinical application. *J. Lipid Res.* **57**, 526–537 (2016).

70. Kronenberg, F. Lipoprotein(a) measurement issues: Are we making a mountain out of a molehill? *Atherosclerosis* **349**, 123–135 (2022).
71. Tsimikas, S. *et al.* NHLBI Working Group Recommendations to Reduce Lipoprotein(a)-Mediated Risk of Cardiovascular Disease and Aortic Stenosis. *J. Am. Coll. Cardiol.* **71**, 177–192 (2018).
72. Lanktree, M. B. *et al.* Determination of lipoprotein(a) kringle repeat number from genomic DNA: copy number variation genotyping using qPCR. *J. Lipid Res.* **50**, 768–772 (2009).
73. Hindson, B. J. *et al.* High-Throughput Droplet Digital PCR System for Absolute Quantitation of DNA Copy Number. *Anal. Chem.* **83**, 8604–8610 (2011).
74. Mazaika, E. & Homsy, J. Digital Droplet PCR: CNV Analysis and Other Applications. *Curr. Protoc. Hum. Genet.* **82**, (2014).
75. Barbieri, G. *et al.* Digital droplet PCR versus quantitative PCR for lipoprotein (a) kringle IV type 2 repeat polymorphism genetic characterization. *J. Clin. Lab. Anal.* **38**, e24998 (2024).
76. AIFA.
77. Thrombotic Thrombocytopenia after ChAdOx1 nCoV-19 Vaccination. *N. Engl. J. Med.* (2021).
78. Gresele, P. *et al.* Interactions of adenoviruses with platelets and coagulation and the vaccine-induced immune thrombotic thrombocytopenia syndrome. *Haematologica* **106**, 3034–3045 (2021).
79. Kadkhoda, K. Post-adenoviral-based COVID-19 vaccines thrombosis: A proposed mechanism. *J. Thromb. Haemost.* **19**, 1831–1832 (2021).
80. Chakraborty, S. *et al.* Proinflammatory IgG Fc structures in patients with severe COVID-19. *Nat. Immunol.* **22**, 67–73 (2021).
81. Azzarone, B., Veneziani, I., Moretta, L. & Maggi, E. Pathogenic Mechanisms of Vaccine-Induced Immune Thrombotic Thrombocytopenia in People Receiving Anti-COVID-19 Adenoviral-Based Vaccines: A Proposal. *Front. Immunol.* **12**, 728513 (2021).
82. Wang, L. *et al.* Inference of person-to-person transmission of COVID-19 reveals hidden super-spreading events during the early outbreak phase. *Nat. Commun.* **11**, 5006 (2020).
83. Tan, W. *et al.* A Novel Coronavirus Genome Identified in a Cluster of Pneumonia Cases — Wuhan, China 2019–2020. *China CDC Wkly.* **2**, 61–62 (2020).
84. Wu, Y. *et al.* SARS-CoV-2 is an appropriate name for the new coronavirus. *The Lancet* **395**, 949–950 (2020).
85. Desogus, A.; S. Lazzari and G. Kiesner Kamps, B.S.; and C. Hoffmann. *COVID Reference 2021.6 ITA.* (2021).
86. Coutard, B. *et al.* The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. *Antiviral Res.* **176**, 104742 (2020).
87. Lurie, N., Saville, M., Hatchett, R. & Halton, J. Developing Covid-19 Vaccines at Pandemic Speed. *N. Engl. J. Med.* **382**, 1969–1973 (2020).
88. Callaway, E. The race for coronavirus vaccines: a graphical guide. *Nature* **580**, 576–577 (2020).
89. Suthar, M. S. *et al.* Rapid Generation of Neutralizing Antibody Responses in COVID-19 Patients. *Cell Rep. Med.* **1**, 100040 (2020).
90. Wang, C. *et al.* A human monoclonal antibody blocking SARS-CoV-2 infection. *Nat. Commun.* **11**, 2251 (2020).
91. Robson, B. COVID-19 Coronavirus spike protein analysis for synthetic vaccines, a peptidomimetic antagonist, and therapeutic drugs, and analysis of a proposed achilles' heel conserved region to minimize probability of escape mutations and drug resistance. *Comput. Biol. Med.* **121**, 103749 (2020).
92. Ho, T., Wu, S., Chen, J., Li, C. & Hsiang, C. Emodin blocks the SARS coronavirus spike protein and angiotensin-converting enzyme 2 interaction. *Antiviral Res.* **74**, 92–101 (2007).
93. Wan, Y., Shang, J., Graham, R., Baric, R. S. & Li, F. Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *J. Virol.* **94**, e00127-20 (2020).

94. Hoffmann, M. *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **181**, 271–280.e8 (2020).
95. Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C. & Garry, R. F. The proximal origin of SARS-CoV-2. *Nat. Med.* **26**, 450–452 (2020).
96. Wrapp, D. *et al.* Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260–1263 (2020).
97. Walls, A. C. *et al.* Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**, 281–292.e6 (2020).
98. Ruch, T. R. & Machamer, C. E. The Coronavirus E Protein: Assembly and Beyond. *Viruses* **4**, 363–382 (2012).
99. Naqvi, A. A. T. *et al.* Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach. *Biochim. Biophys. Acta BBA - Mol. Basis Dis.* **1866**, 165878 (2020).
100. Sardar, R., Satish, D., Birla, S. & Gupta, D. Integrative analyses of SARS-CoV-2 genomes from different geographical locations reveal unique features potentially consequential to host-virus interaction, pathogenesis and clues for novel therapies. *Heliyon* **6**, e04658 (2020).
101. Romano, M., Ruggiero, A., Squeglia, F., Maga, G. & Berisio, R. A Structural View of SARS-CoV-2 RNA Replication Machinery: RNA Synthesis, Proofreading and Final Capping. *Cells* **9**, 1267 (2020).
102. Nishiga, M., Wang, D. W., Han, Y., Lewis, D. B. & Wu, J. C. COVID-19 and cardiovascular disease: from basic mechanisms to clinical perspectives. *Nat. Rev. Cardiol.* **17**, 543–558 (2020).
103. Zhao, Z., Bashiri, S., Ziora, Z. M., Toth, I. & Skwarczynski, M. COVID-19 Variants and Vaccine Development. *Viruses* **16**, 757 (2024).
104. 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).
105. Zhou, W. *et al.* N439K Variant in Spike Protein Alter the Infection Efficiency and Antigenicity of SARS-CoV-2 Based on Molecular Dynamics Simulation. *Front. Cell Dev. Biol.* **9**, 697035 (2021).
106. Ashwaq, O., Manickavasagam, P. & Haque, S. M. V483A: an Emerging Mutation Hotspot of SARS-CoV-2. *Future Virol.* **16**, 419–429 (2021).
107. Zhang, L. *et al.* The D614G mutation in the SARS-CoV-2 spike protein reduces S1 shedding and increases infectivity. Preprint at <https://doi.org/10.1101/2020.06.12.148726> (2020).
108. Zhang, L. *et al.* SARS-CoV-2 spike-protein D614G mutation increases virion spike density and infectivity. *Nat. Commun.* **11**, 6013 (2020).
109. Zeng, B., Gao, L., Zhou, Q., Yu, K. & Sun, F. Effectiveness of COVID-19 vaccines against SARS-CoV-2 variants of concern: a systematic review and meta-analysis. *BMC Med.* **20**, 200 (2022).
110. Chakraborty, C., Sharma, A. R., Bhattacharya, M. & Lee, S.-S. A Detailed Overview of Immune Escape, Antibody Escape, Partial Vaccine Escape of SARS-CoV-2 and Their Emerging Variants With Escape Mutations. *Front. Immunol.* **13**, 801522 (2022).
111. Cui, J., Li, F. & Shi, Z.-L. Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol.* **17**, 181–192 (2019).
112. Du, L. *et al.* The spike protein of SARS-CoV — a target for vaccine and therapeutic development. *Nat. Rev. Microbiol.* **7**, 226–236 (2009).
113. Sanders, J. M., Monogue, M. L., Jodlowski, T. Z. & Cutrell, J. B. Pharmacologic Treatments for Coronavirus Disease 2019 (COVID-19): A Review. *JAMA* (2020) doi:10.1001/jama.2020.6019.
114. Li, G. & De Clercq, E. Therapeutic options for the 2019 novel coronavirus (2019-nCoV). *Nat. Rev. Drug Discov.* **19**, 149–150 (2020).
115. Sigrist, C. J., Bridge, A. & Le Mercier, P. A potential role for integrins in host cell entry by SARS-CoV-2. *Antiviral Res.* **177**, 104759 (2020).
116. Ulrich, H. & Pillat, M. M. CD147 as a Target for COVID-19 Treatment: Suggested Effects of Azithromycin and Stem Cell Engagement. *Stem Cell Rev. Rep.* **16**, 434–440 (2020).

117. Yildirim, Z., Sahin, O. S., Yazar, S. & Bozok Cetintas, V. Genetic and epigenetic factors associated with increased severity of Covid-19. *Cell Biol. Int.* **45**, 1158–1174 (2021).
118. Xu, W. *et al.* Causal association of epigenetic aging and COVID-19 severity and susceptibility: A bidirectional Mendelian randomization study. *Front. Med.* **9**, 989950 (2022).
119. Hou, Y. *et al.* New insights into genetic susceptibility of COVID-19: an ACE2 and TMPRSS2 polymorphism analysis. *BMC Med.* **18**, 216 (2020).
120. Cao, Y. *et al.* Comparative genetic analysis of the novel coronavirus (2019-nCoV/SARS-CoV-2) receptor ACE2 in different populations. *Cell Discov.* **6**, 11 (2020).
121. Khayat, A. S. *et al.* ACE2 polymorphisms as potential players in COVID-19 outcome. *PLOS ONE* **15**, e0243887 (2020).
122. Xiao, F. *et al.* Characterization of Angiotensin-Converting Enzyme 2 Ectodomain Shedding from Mouse Proximal Tubular Cells. *PLoS ONE* **9**, e85958 (2014).
123. Fawzy, M. S. *et al.* The role of angiotensin-converting enzyme 2 (ACE2) genetic variations in COVID-19 infection: a literature review. *Egypt. J. Med. Hum. Genet.* **23**, 97 (2022).
124. Benetti, E. *et al.* ACE2 gene variants may underlie interindividual variability and susceptibility to COVID-19 in the Italian population. *Eur. J. Hum. Genet.* **28**, 1602–1614 (2020).
125. Fujikura, K. & Uesaka, K. Genetic variations in the human severe acute respiratory syndrome coronavirus receptor *ACE2* and serine protease *TMPRSS2*. *J. Clin. Pathol.* **74**, 307–313 (2021).
126. Lippi, G., Lavie, C. J., Henry, B. M. & Sanchis-Gomar, F. Do genetic polymorphisms in angiotensin converting enzyme 2 (*ACE2*) gene play a role in coronavirus disease 2019 (COVID-19)? *Clin. Chem. Lab. Med. CCLM* **58**, 1415–1422 (2020).
127. Quiles, J. L. *et al.* Do nutrients and other bioactive molecules from foods have anything to say in the treatment against COVID-19? *Environ. Res.* **191**, 110053 (2020).
128. Chaudhary, M. COVID-19 susceptibility: potential of ACE2 polymorphisms. *Egypt. J. Med. Hum. Genet.* **21**, 54 (2020).
129. Vausort, M., Wagner, D. R. & Devaux, Y. Long Noncoding RNAs in Patients With Acute Myocardial Infarction. *Circ. Res.* **115**, 668–677 (2014).
130. Devaux, C. A. & Fantini, J. ACE2 receptor polymorphism in humans and animals increases the risk of the emergence of SARS-CoV-2 variants during repeated intra- and inter-species host-switching of the virus. *Front. Microbiol.* **14**, 1199561 (2023).
131. Makled, A. F. *et al.* Angiotensin-Converting Enzyme-2 (ACE-2) with Interferon-Induced Transmembrane Protein-3 (IFITM-3) Genetic Variants and Interleukin-6 as Severity and Risk Predictors among COVID-19 Egyptian Population. *Int. J. Microbiol.* **2023**, 1–12 (2023).
132. Verma, S. *et al.* Genetic polymorphisms of IL6 gene –174G > C and –597G > A are associated with the risk of COVID-19 severity. *Int. J. Immunogenet.* **50**, 5–11 (2023).
133. Linnik, J. E. & Egli, A. Impact of host genetic polymorphisms on vaccine induced antibody response. *Hum. Vaccines Immunother.* **12**, 907–915 (2016).
134. Doyle, W. J. *et al.* The Interleukin 6 –174 C/C Genotype Predicts Greater Rhinovirus Illness. *J. Infect. Dis.* **201**, 199–206 (2010).
135. Nouredine, R. *et al.* The Interleukin-6 gene variants may protect against SARS-CoV-2 infection and the severity of COVID-19: a case-control study in a Moroccan population. *BMC Med. Genomics* **17**, 139 (2024).
136. AbdelAziz, R. A., Abd-Allah, S. T., Moness, H. M., Anwar, A. M. & Mohamed, Z. H. Role of interleukin 6 polymorphism and inflammatory markers in outcome of pediatric Covid- 19 patients. *BMC Pediatr.* **24**, 625 (2024).
137. Huang, S.-W. & Wang, S.-F. SARS-CoV-2 Entry Related Viral and Host Genetic Variations: Implications on COVID-19 Severity, Immune Escape, and Infectivity. *Int. J. Mol. Sci.* **22**, 3060 (2021).
138. Stopsack, K. H., Mucci, L. A., Antonarakis, E. S., Nelson, P. S. & Kantoff, P. W. *TMPRSS2* and COVID-19: Serendipity or Opportunity for Intervention? *Cancer Discov.* **10**, 779–782 (2020).

139. Posadas-Sánchez, R. *et al.* Association of the Transmembrane Serine Protease-2 (TMPRSS2) Polymorphisms with COVID-19. *Viruses* **14**, 1976 (2022).
140. Rokni, M. *et al.* Association of TMPRSS2 Gene Polymorphisms with COVID-19 Severity and Mortality: a Case-Control Study with Computational Analyses. *Appl. Biochem. Biotechnol.* **194**, 3507–3526 (2022).
141. Fernández-de-las-Peñas, C. *et al.* Genetic Association between ACE2 (rs2285666 and rs2074192) and TMPRSS2 (rs12329760 and rs2070788) Polymorphisms with Post-COVID Symptoms in Previously Hospitalized COVID-19 Survivors. *Genes* **13**, 1935 (2022).
142. Yaghoobi, A. *et al.* TMPRSS2 polymorphism (rs12329760) and the severity of the COVID-19 in Iranian population. *PLOS ONE* **18**, e0281750 (2023).
143. Kaidashev, I. *et al.* Polymorphism of tmprss2 (rs12329760) but not ace2 (rs4240157), tmprss11a (rs353163) and cd147 (rs8259) is associated with the severity of COVID-19 in the Ukrainian population. *Acta Biomed. Atenei Parm.* **94**, e2023030 (2023).
144. David, A. *et al.* A common TMPRSS2 variant has a protective effect against severe COVID-19. *Curr. Res. Transl. Med.* **70**, 103333 (2022).
145. Pandey, R. K., Srivastava, A., Singh, P. P. & Chaubey, G. Genetic association of TMPRSS2 rs2070788 polymorphism with COVID-19 case fatality rate among Indian populations. *Infect. Genet. Evol.* **98**, 105206 (2022).
146. De Andrade, C. C. *et al.* A Polymorphism in the TMPRSS2 Gene Increases the Risk of Death in Older Patients Hospitalized with COVID-19. *Viruses* **14**, 2557 (2022).
147. Schönfelder, K. *et al.* Transmembrane serine protease 2 Polymorphisms and Susceptibility to Severe Acute Respiratory Syndrome Coronavirus Type 2 Infection: A German Case-Control Study. *Front. Genet.* **12**, 667231 (2021).
148. Abdelsattar, S. *et al.* ACE2 and TMPRSS2 SNPs as Determinants of Susceptibility to, and Severity of, a COVID-19 Infection. *Br. J. Biomed. Sci.* **79**, 10238 (2022).
149. Bestle, D. *et al.* TMPRSS2 and furin are both essential for proteolytic activation of SARS-CoV-2 in human airway cells. *Life Sci. Alliance* **3**, e202000786 (2020).
150. Fernandez, C. *et al.* Plasma levels of the proprotein convertase furin and incidence of diabetes and mortality. *J. Intern. Med.* **284**, 377–387 (2018).
151. Torre-Fuentes, L. *et al.* ACE2, TMPRSS2, and Furin variants and SARS-CoV-2 infection in Madrid, Spain. *J. Med. Virol.* **93**, 863–869 (2021).
152. Klaassen, K. *et al.* Functional prediction and comparative population analysis of variants in genes for proteases and innate immunity related to SARS-CoV-2 infection. *Infect. Genet. Evol.* **84**, 104498 (2020).
153. Latini, A. *et al.* COVID-19 and Genetic Variants of Protein Involved in the SARS-CoV-2 Entry into the Host Cells. *Genes* **11**, 1010 (2020).
154. De Araújo, J. L. F., Menezes, D., Aguiar, R. S. D. & Souza, R. P. D. IFITM3, FURIN, ACE1, and TNF- α Genetic Association With COVID-19 Outcomes: Systematic Review and Meta-Analysis. *Front. Genet.* **13**, 775246 (2022).
155. Deng, H., Yan, X. & Yuan, L. Human genetic basis of coronavirus disease 2019. *Signal Transduct. Target. Ther.* **6**, 344 (2021).
156. Dobrindt, K. *et al.* Common genetic variation in humans impacts *in vitro* susceptibility to SARS-CoV-2 infection. Preprint at <https://doi.org/10.1101/2020.09.20.300574> (2020).
157. Fu, J. *et al.* Evaluation and characterization of HSPA5 (GRP78) expression profiles in normal individuals and cancer patients with COVID-19. *Int. J. Biol. Sci.* **17**, 897–910 (2021).
158. Zhu, J., Wu, C. & Wu, L. Associations Between Genetically Predicted Protein Levels and COVID-19 Severity. *J. Infect. Dis.* **223**, 19–22 (2021).
159. Onoja, A. *et al.* An explainable model of host genetic interactions linked to COVID-19 severity. *Commun. Biol.* **5**, 1133 (2022).
160. Smatti, M. K., Al-Sarraj, Y. A., Albagha, O. & Yassine, H. M. Host Genetic Variants Potentially

Associated With SARS-CoV-2: A Multi-Population Analysis. *Front. Genet.* **11**, 578523 (2020).

161. Huang, Y. *et al.* Comprehensive and Integrative Analysis of Two Novel SARS-CoV-2 Entry Associated Proteases CTSB and CTSL in Healthy Individuals and Cancer Patients. *Front. Bioeng. Biotechnol.* **10**, 780751 (2022).
162. COVID-19 Host Genetics Initiative *et al.* Mapping the human genetic architecture of COVID-19. *Nature* **600**, 472–477 (2021).
163. Velavan, T. P. *et al.* Host genetic factors determining COVID-19 susceptibility and severity. *eBioMedicine* **72**, 103629 (2021).
164. Degenhardt, F. *et al.* Detailed stratified GWAS analysis for severe COVID-19 in four European populations. *Hum. Mol. Genet.* **31**, 3945–3966 (2022).
165. The COVID-19 Host Genetics Initiative & Ganna, A. A second update on mapping the human genetic architecture of COVID-19. Preprint at <https://doi.org/10.1101/2022.12.24.22283874> (2022).
166. Ben Shachar, S. *et al.* MHC Haplotyping of SARS-CoV-2 Patients: HLA Subtypes Are Not Associated with the Presence and Severity of COVID-19 in the Israeli Population. *J. Clin. Immunol.* **41**, 1154–1161 (2021).
167. Novelli, A. *et al.* HLA allele frequencies and susceptibility to COVID -19 in a group of 99 Italian patients. *HLA* **96**, 610–614 (2020).
168. Arshad, N. *et al.* SARS-CoV-2 accessory proteins ORF7a and ORF3a use distinct mechanisms to down-regulate MHC-I surface expression. *Proc. Natl. Acad. Sci.* **120**, e2208525120 (2023).
169. Dieter, C. *et al.* Polymorphisms in ACE1, TMPRSS2, IFIH1, IFNAR2, and TYK2 Genes Are Associated with Worse Clinical Outcomes in COVID-19. *Genes* **14**, 29 (2022).
170. Baldassarri, M. *et al.* Gain- and Loss-of-Function CFTR Alleles Are Associated with COVID-19 Clinical Outcomes. *Cells* **11**, 4096 (2022).
171. Zhang, Q. *et al.* Human genetic and immunological determinants of critical COVID-19 pneumonia. *Nature* **603**, 587–598 (2022).
172. Gusev, E., Sarapultsev, A., Solomatina, L. & Chereshev, V. SARS-CoV-2-Specific Immune Response and the Pathogenesis of COVID-19. *Int. J. Mol. Sci.* **23**, 1716 (2022).
173. Lee, D. *et al.* Inborn errors of OAS–RNase L in SARS-CoV-2–related multisystem inflammatory syndrome in children. *Science* **379**, eabo3627 (2023).
174. Matuozzo, D. *et al.* Rare predicted loss-of-function variants of type I IFN immunity genes are associated with life-threatening COVID-19. *Genome Med.* **15**, 22 (2023).
175. Kousathanas, A. *et al.* Whole-genome sequencing reveals host factors underlying critical COVID-19. *Nature* **607**, 97–103 (2022).
176. Namkoong, H. *et al.* DOCK2 is involved in the host genetics and biology of severe COVID-19. *Nature* **609**, 754–760 (2022).
177. Jia, G. *et al.* Genome- and transcriptome-wide association studies of 386,000 Asian and European-ancestry women provide new insights into breast cancer genetics. *Am. J. Hum. Genet.* **109**, 2185–2195 (2022).
178. Ferreira, M. A. *et al.* Genome-wide association and transcriptome studies identify target genes and risk loci for breast cancer. *Nat. Commun.* **10**, 1741 (2019).
179. Momozawa, Y. & Mizukami, K. Unique roles of rare variants in the genetics of complex diseases in humans. *J. Hum. Genet.* **66**, 11–23 (2021).
180. Gao, H. *et al.* The landscape of tolerated genetic variation in humans and primates. *Science* **380**, eabn8153 (2023).
181. The GenOMICC Investigators *et al.* Genetic mechanisms of critical illness in COVID-19. *Nature* **591**, 92–98 (2021).
182. Badla, B. A. *et al.* Genetic determinants of severe COVID-19 in young Asian and Middle Eastern patients: a case series. *Sci. Rep.* **13**, 20294 (2023).

183. Boukhari, R. *et al.* ABO Blood Group Incompatibility Protects Against SARS-CoV-2 Transmission. *Front. Microbiol.* **12**, 799519 (2022).
184. Ellis, P. J. I. Modelling suggests ABO histo-incompatibility may substantially reduce SARS-CoV-2 transmission. *Epidemics* **35**, 100446 (2021).
185. Zietz, M., Zucker, J. & Tatonetti, N. P. Associations between blood type and COVID-19 infection, intubation, and death. *Nat. Commun.* **11**, 5761 (2020).
186. Zhao, J. *et al.* Relationship Between the ABO Blood Group and the Coronavirus Disease 2019 (COVID-19) Susceptibility. *Clin. Infect. Dis.* **73**, 328–331 (2021).
187. Ellinghaus, D. COVID-19 host genetics and ABO blood group susceptibility. *Camb. Prisms Precis. Med.* **1**, e10 (2023).
188. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
189. Thorvaldsdottir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).
190. McKenna, A. *et al.* The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
191. Richards, S. *et al.* Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **17**, 405–424 (2015).
192. Pavord, S. *et al.* Clinical Features of Vaccine-Induced Immune Thrombocytopenia and Thrombosis. *N. Engl. J. Med.* **385**, 1680–1689 (2021).
193. Gething, M.-J. & Sambrook, J. Protein folding in the cell. *Nature* **355**, 33–45 (1992).
194. Elfiky, A. A., Baghdady, A. M., Ali, S. A. & Ahmed, M. I. GRP78 targeting: Hitting two birds with a stone. *Life Sci.* **260**, 118317 (2020).
195. Ibrahim, I. M., Abdelmalek, D. H., Elshahat, M. E. & Elfiky, A. A. COVID-19 spike-host cell receptor GRP78 binding site prediction. *J. Infect.* **80**, 554–562 (2020).
196. Middeldorp, S. & Van Hylckama Vlieg, A. Does thrombophilia testing help in the clinical management of patients? *Br. J. Haematol.* **143**, 321–335 (2008).
197. Lane, D. A., Kunz, G., Olds, R. J. & Thein, S. L. Molecular genetics of antithrombin deficiency. *Blood Rev.* **10**, 59–74 (1996).
198. J Y Chang & T H Tran. Antithrombin III Basel. Identification of a Pro-Leu substitution in a hereditary abnormal antithrombin with impaired heparin cofactor activity.
199. Szilágyi, S., Péter, A., Tünde Magyar, M., Balogh, S. & Bereczky, Z. Recurrent Arterial Thrombosis Associated With the Antithrombin Basel Variant and Elevated Lipoprotein(a) Plasma Level in an Adolescent Patient. *J. Pediatr. Hematol. Oncol.* **34**, 276–279 (2012).
200. Puurunen, M., Salo, P., Engelbarth, S., Javela, K. & Perola, M. Type II antithrombin deficiency caused by a founder mutation Pro73Leu in the Finnish population: clinical picture. *J. Thromb. Haemost.* **11**, 1844–1849 (2013).
201. Hwang, J. *et al.* Predictors of mortality in thrombotic thrombocytopenia after adenoviral COVID-19 vaccination: the FAPIC score. *Eur. Heart J.* **42**, 4053–4063 (2021).
202. Greinacher, A. Heparin-induced thrombocytopenia. *J. Thromb. Haemost.* **7**, 9–12 (2009).
203. Harris, E. N., Weigel, J. A. & Weigel, P. H. The Human Hyaluronan Receptor for Endocytosis (HARE/Stabilin-2) Is a Systemic Clearance Receptor for Heparin. *J. Biol. Chem.* **283**, 17341–17350 (2008).
204. Harris, E. N., Baggenstoss, B. A. & Weigel, P. H. Rat and human HARE/stabilin-2 are clearance receptors for high- and low-molecular-weight heparins. *Am. J. Physiol.-Gastrointest. Liver Physiol.* **296**, G1191–G1199 (2009).
205. Desch, K. C. *et al.* Whole-exome sequencing identifies rare variants in STAB2 associated with venous thromboembolic disease. *Blood* **136**, 533–541 (2020).

206. Michels, A. *et al.* Stabilin-2 deficiency increases thrombotic burden and alters the composition of venous thrombi in a mouse model. *J. Thromb. Haemost.* **19**, 2440–2453 (2021).
207. Roberts, D. D., Miller, T. W., Rogers, N. M., Yao, M. & Isenberg, J. S. The matricellular protein thrombospondin-1 globally regulates cardiovascular function and responses to stress via CD47. *Matrix Biol.* **31**, 162–169 (2012).
208. Rogers, N. M., Sharifi-Sanjani, M., Csányi, G., Pagano, P. J. & Isenberg, J. S. Thrombospondin-1 and CD47 regulation of cardiac, pulmonary and vascular responses in health and disease. *Matrix Biol.* **37**, 92–101 (2014).
209. Jeanne, A. *et al.* Towards the Therapeutic Use of Thrombospondin 1/CD47 Targeting TAX2 Peptide as an Antithrombotic Agent. *Arterioscler. Thromb. Vasc. Biol.* **41**, (2021).
210. Imamoto, T. *et al.* Different Transcriptome Features of Peripheral Blood Mononuclear Cells in Non-Emphysematous Chronic Obstructive Pulmonary Disease. *Int. J. Mol. Sci.* **25**, 66 (2023).
211. Sussman, T. A., Abazeed, M. E., McCrae, K. R. & Khorana, A. A. RNA expression and risk of venous thromboembolism in lung cancer. *Res. Pract. Thromb. Haemost.* **4**, 117–123 (2020).
212. Perrella, G., Nagy, M., Watson, S. P. & Heemskerk, J. W. M. Platelet GPVI (Glycoprotein VI) and Thrombotic Complications in the Venous System. *Arterioscler. Thromb. Vasc. Biol.* **41**, 2681–2692 (2021).
213. Oertwig, K. *et al.* Glycosylation of human vaspin (SERPINA12) and its impact on serpin activity, heparin binding and thermal stability. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **1865**, 1188–1194 (2017).
214. Zhou, X. *et al.* Serum Vaspin as a Predictor of Adverse Cardiac Events in Acute Myocardial Infarction. *J. Am. Heart Assoc.* **8**, e010934 (2019).
215. Vu, T. M. *et al.* Mfsd2b is essential for the sphingosine-1-phosphate export in erythrocytes and platelets. *Nature* **550**, 524–528 (2017).
216. Chandrakanthan, M. *et al.* Deletion of Mfsd2b impairs thrombotic functions of platelets. *Nat. Commun.* **12**, 2286 (2021).
217. Wei, Y. *et al.* Regulation of *Csf1r* and *Bcl6* in Macrophages Mediates the Stage-Specific Effects of MicroRNA-155 on Atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* **35**, 796–803 (2015).
218. Aljurayyan, A. *et al.* Activation and Induction of Antigen-Specific T Follicular Helper Cells Play a Critical Role in Live-Attenuated Influenza Vaccine-Induced Human Mucosal Anti-influenza Antibody Response. *J. Virol.* **92**, e00114-18 (2018).
219. Miyanohara, J., Shirakawa, H., Sanpei, K., Nakagawa, T. & Kaneko, S. A pathophysiological role of TRPV1 in ischemic injury after transient focal cerebral ischemia in mice. *Biochem. Biophys. Res. Commun.* **467**, 478–483 (2015).
220. Harper, A. G. S., Brownlow, S. L. & Sage, S. O. A role for TRPV1 in agonist-evoked activation of human platelets. *J. Thromb. Haemost.* **7**, 330–338 (2009).
221. Kawamoto, E. *et al.* LFA-1 and Mac-1 integrins bind to the serine/threonine-rich domain of thrombomodulin. *Biochem. Biophys. Res. Commun.* **473**, 1005–1012 (2016).
222. Zhao, W. *et al.* Piezo1 initiates platelet hyperreactivity and accelerates thrombosis in hypertension. *J. Thromb. Haemost.* **19**, 3113–3125 (2021).
223. Li, J. Glucuronyl C5-Epimerase. in *Progress in Molecular Biology and Translational Science* vol. 93 59–78 (Elsevier, 2010).
224. Paramel, G. V. *et al.* Expression of CARD8 in human atherosclerosis and its regulation of inflammatory proteins in human endothelial cells. *Sci. Rep.* **10**, 19108 (2020).
225. Furuta, T. *et al.* Intratumoral thrombosis as a histological biomarker for predicting epidermal growth factor receptor alteration and poor prognosis in patients with glioblastomas. *J. Neurooncol.* **164**, 633–641 (2023).
226. Benz, P. M. *et al.* Vasodilator-Stimulated Phosphoprotein (VASP)-dependent and -independent pathways regulate thrombin-induced activation of Rap1b in platelets. *Cell Commun. Signal.* **14**, 21 (2016).

227. Aszodi, A. The vasodilator-stimulated phosphoprotein (VASP) is involved in cGMP- and cAMP-mediated inhibition of agonist-induced platelet aggregation, but is dispensable for smooth muscle function. *EMBO J.* **18**, 37–48 (1999).
228. Park, S.-A. *et al.* Association of polymorphisms in thromboxane A2 receptor and thromboxane A synthase 1 with cerebral infarction in a Korean population. *BMB Rep.* **42**, 200–205 (2009).
229. Ramazi, S., Heydari-Zarnagh, H., Goudarzian, M., Khalaj-Kondori, M. & Bonyadi, M. Thromboxane A synthase 1 gene expression and promotor haplotypes are associated with risk of large artery-atherosclerosis stroke in Iranian population. *J. Cell. Biochem.* **120**, 15222–15232 (2019).
230. Kaczmarek, E. *et al.* Identification and Characterization of CD39/Vascular ATP Diphosphohydrolase. *J. Biol. Chem.* **271**, 33116–33122 (1996).
231. Ruan, J. *et al.* Low Expression of *FFAR2* in Peripheral White Blood Cells May Be a Genetic Marker for Early Diagnosis of Acute Myocardial Infarction. *Cardiol. Res. Pract.* **2020**, 1–8 (2020).
232. Adhikari, N. *et al.* Heparan sulfate Ndst1 regulates vascular smooth muscle cell proliferation, vessel size and vascular remodeling. *J. Mol. Cell. Cardiol.* **49**, 287–293 (2010).
233. Ge, X. N. *et al.* Regulation of eosinophil recruitment and allergic airway inflammation by heparan sulfate proteoglycan (HSPG) modifying enzymes. *Exp. Lung Res.* **44**, 98–112 (2018).
234. Puleo, M. G. *et al.* Molecular Mechanisms of Inflammasome in Ischemic Stroke Pathogenesis. *Pharmaceuticals* **15**, 1168 (2022).
235. Morgan, S. *et al.* Identifying novel mechanisms of abdominal aortic aneurysm via unbiased proteomics and systems biology. *Front. Cardiovasc. Med.* **9**, 889994 (2022).
236. Weisel, J. W. & Litvinov, R. I. Red blood cells: the forgotten player in hemostasis and thrombosis. *J. Thromb. Haemost.* **17**, 271–282 (2019).
237. Jordan, P. A. *et al.* A role for the thiol isomerase protein ERP5 in platelet function. *Blood* **105**, 1500–1507 (2005).
238. Park, D., Lapteva, N., Seethammagari, M., Slawin, K. M. & Spencer, D. M. An essential role for Akt1 in dendritic cell function and tumor immunotherapy. *Nat. Biotechnol.* **24**, 1581–1590 (2006).
239. Keppler-Noreuil, K. M., Lozier, J. N., Sapp, J. C. & Biesecker, L. G. Characterization of thrombosis in patients with Proteus syndrome. *Am. J. Med. Genet. A.* **173**, 2359–2365 (2017).
240. Zheng, Y. *et al.* Critical role for mouse marginal zone B cells in PF4/heparin antibody production. *Blood* **121**, 3484–3492 (2013).
241. Mumby, S. M., Raugi, G. J. & Bornstein, P. Interactions of thrombospondin with extracellular matrix proteins: selective binding to type V collagen. *J. Cell Biol.* **98**, 646–652 (1984).
242. Yaoi, Y. *et al.* Primary structure of the heparin-binding site of type V collagen. *Biochim. Biophys. Acta BBA - Gen. Subj.* **1035**, 139–145 (1990).
243. Gómez-Gavero, M. *et al.* Expression and Regulation of the Metalloproteinase ADAM-8 during Human Neutrophil Pathophysiological Activation and Its Catalytic Activity on L-Selectin Shedding. *J. Immunol.* **178**, 8053–8063 (2007).
244. Mendes-de-Almeida, D. P. *et al.* A case report of vaccine-induced immune thrombotic thrombocytopenia (VITT) with genetic analysis. *Front. Cardiovasc. Med.* **10**, 1189320 (2023).