Analysis of Genetics Risk Factors Associated with COVID-19 Severity

By

Maria Saleem

Department of Biochemistry

Faculty of Biological Sciences

Quaid-i-Azam University

Islamabad, Pakistan

2024

Analysis of Genetics Risk Factors Associated with COVID-19 Severity

A thesis submitted in partial fulfillment of requirements for the degree of

Master of Philosophy

In

Biochemistry/Molecular Biology

By

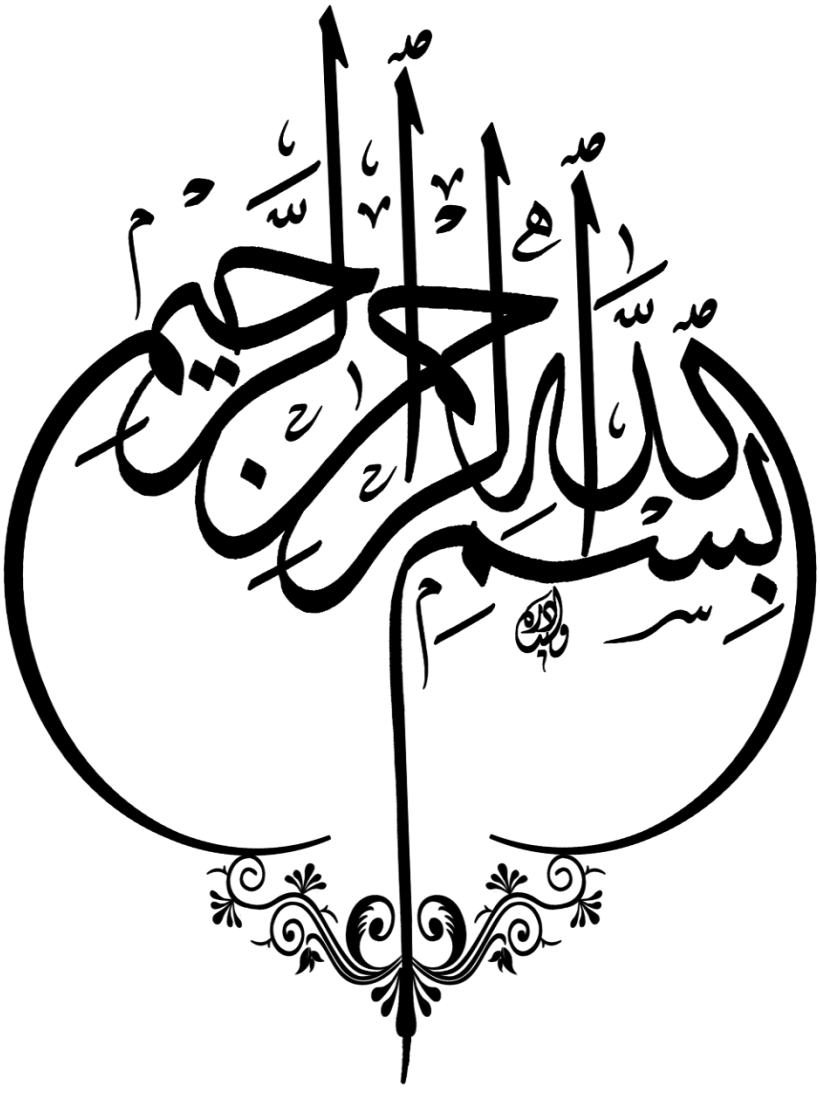
Maria Saleem

Department of Biochemistry

Quaid-i-Azam University

Islamabad, Pakistan

2024


بسم الله الرحمن الرحيم

Quaid-i-Azam University, Islamabad, PAKISTAN
Department of Biochemistry
Faculty of Biological Sciences

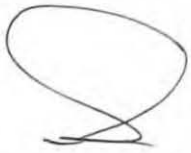
CERTIFICATE

This thesis, submitted by **Ms. Maria Saleem** to the Department of Biochemistry, Faculty of Biological Sciences, Quaid-i-Azam University, Islamabad, Pakistan, is accepted in its present form as satisfying the thesis requirement for the Degree of Master of Philosophy in Biochemistry/Molecular Biology.

Examination Committee:

1. **External Examiner:** Signature: _____

2. **Supervisor:**
   Prof. Dr. Muhammad Ansar Signature: _____



3. **Chairperson:**
   Prof. Dr. Iram Murtaza Signature: _____

Dated:

Acknowledgment

All praise is due to **Allah Almighty**, the Compassionate and the Merciful, who has bestowed upon me the strength, wisdom, and perseverance to complete this thesis. I am deeply grateful for His guidance and blessings. I extend my sincerest gratitude to the **Prophet Hazrat Muhammad** (peace be upon him), the exemplar of compassion and wisdom, whose teachings continue to illuminate my path and inspire my journey.

I would like to express my heartfelt appreciation to my worthy supervisor, **Dr. Muhammad Ansar** (Professor, Department of Biochemistry). His scholastic supervision, constructive feedback, encouragement, valuable suggestions, unwavering support and timely correction have been instrumental in shaping the course of this thesis. I am truly fortunate to have had the opportunity to benefit from his mentorship.

To my beloved **parents**, whose boundless love, encouragement, and sacrifices have been the foundation of my academic pursuit. Your unwavering support has been my source of strength and inspiration.

A special mention goes to my dear siblings, **Areeba Saleem, Aliza Saleem, Daniyal Saleem**, and **Muhammad Abdullah**. Your encouragement, understanding, and support have been a constant source of motivation.

I extend my gratitude to Senior **Samra Akram** for her guidance, wisdom, and continuous support. Her mentorship has been invaluable in navigating the complexities of this academic journey. To my dear friends, **Areesha, Somia, Hadiqa, Khadeeja, Sundas, Faseeha** and **Laiba**, thank you for the camaraderie, shared moments, and unwavering encouragement. Your friendship has made this academic endeavor more enriching.

Special thanks to **Muneeba Asif** and **Ayesha Shabbir** for their guidance and motivation. Their insightful input and encouragement have been a driving force behind my academic achievements. I would also like to express my appreciation to my lab fellows **Mahnoor Sakina, Hamza Hameed**, and **Ali Asghar**, for their, support, and shared commitment to academic excellence. To everyone who has contributed, no matter how big or small, to this academic venture, I am sincerely thankful for your involvement.

Maria Saleem

Dedication

I dedicate these efforts

To

My **parents** and **grandparents**

Your unwavering love and enduring support have been my guiding stars. This thesis is a humble tribute to your influence, wisdom, and the immeasurable impact you've had on my journey. Thank you for being the pillars of strength in my life.

Declaration

I declare that the content within this thesis is my own original work. I am the sole author, and none of the material has been submitted elsewhere for any degree or certificate.

Maria Saleem

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ABO	Blood Type System
ACE	Angiotensin-Converting Enzyme
ACE I/D	Angiotensin-Converting Enzyme Insertion/Deletion
ACE2	Angiotensin-Converting Enzyme 2
ARE	Androgen Response Element
AXL	AXL Receptor Tyrosine Kinase
BLAT	Blast-Like Alignment Tool
COVID-19	Coronavirus Disease 2019
CCR2	C-C Chemokine Receptor Type 2
CCR5	C-C Chemokine Receptor Type 5
CLD	Chronic Liver Disease
DMV	Double-Membrane Vesicle
DPP4	Dipeptidyl Peptidase 4
E PROTEIN	Envelope Protein
EtBr	Ethidium Bromide
EXOSAP	Exonuclease I and Shrimp Alkaline Phosphatase
FURIN	Furin, Paired Basic Amino Acid Cleaving Enzyme
GRP78	Glucose-Regulated Protein 78
HLA	Human Leukocyte Antigen
IFNAR2	Interferon Alpha and Beta Receptor Subunit 2
LDLRA	Low-Density Lipoprotein Receptor-Related Protein Associated
M PROTEIN	Membrane Protein
MHC I	Major Histocompatibility Complex Class I
MHC II	Major Histocompatibility Complex Class II
N PROTEIN	Nucleocapsid Protein
NRP1	Neuropilin 1
Nsp	Non-structural Protein
OAS	2'-5'-Oligoadenylate Synthase
ORFs	Open Reading Frames
PAR2	Protease-Activated Receptor 2
PCR	Polymerase Chain Reaction

PPPIR15A	Protein Phosphatase 1 Regulatory Subunit 15A
RAAS	Renin-Angiotensin-Aldosterone System
RARE	Retinoic Acid Response Element
RdRp	RNA-dependent RNA Polymerase
S PROTEIN	Spike Protein
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SLC6A20	Solute Carrier Family 6 Member 20
SNPs	Single Nucleotide Polymorphisms
SRDR	Scavenger Receptor Cysteine-Rich Domain
TLR3	Toll-like Receptor 3
TLR7	Toll-like Receptor 7
Tm	Melting Temperature
TMPRSS2	Transmembrane Protease, Serine 2
TYK2	Tyrosine Kinase 2

Abstract

The COVID-19 pandemic was caused by the severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2). This complex disease manifests different symptoms across different ethnicities, influenced by both environmental and genetic factors. This study focused on understanding of the genetic risk factors associated withCOVID-19 severity in the Pakistani population. The study was performed on a cohort of 131 individuals categorized into mild (N=53) and severe (N=78) groups. The study examined ACE insertion/deletion polymorphism, a TMPRSS2 variant (rs3787950), and HLA-A*11 genotype in the cohort of 131 individuals. Genotyping was conducted using agarose gel for ACE and HLA-A*11, but Sanger sequencing was performed for TMPRSS2 variant analysis. ACE D/D genotype is more common in individuals within severe group as compared to individuals in the mild group. Statistical analysis, including Fisher's Exact Test (0.04676) and Chi Square test (0.0444) for genotypes, Fisher's Exact Test (0.03153) and Chi Square test (0.03467) for alleles of ACE revealed significant association with severity of COVID-19. To explore the impact of the ACE I allele under recessive and dominant models, further analysis yielded a significant result for recessive effect. This indicates that individuals having I allele are less prone to develop severe disease. TMPRSS2 rs3787950 was the only variant found in our study cohort and the frequency of minor allele is 12%. The HLA-A*11 allele genotyping shows its frequent presence in both severe and mild group. However, no significant associations were found for TMPRSS2 and HLA-A*11 in relation to the disease severity. These findings enhance our understanding of host genetics in the Pakistani population, but additional studies are required to identify additional host genetic factors.

Key words: COVID-19, Host genetic Factors, ACE, TMPRSS2, HLA

Chapter 1

Introduction

1. Introduction

Severe Acute Respiratory Syndrome Coronavirus 2 is responsible for the onset of Coronavirus Disease 2019 and the initial emergence occurred in Wuhan, China, in December 2019. It swiftly evolve into a global pandemic due to its high virulence and contagion (Wang et al., 2020). On January 30, 2020, the World Health Organization (WHO) declared a public health emergency, which remained in effect until May 5, 2023. As of October 25, 2023, there have been 771,549,718 cases confirmed globally, with 6,974,473 recorded deaths (https://covid19.who.int/table). Initially the virus was named novel Corona Virus (nCoV-19), but later SARS-CoV-2 was assigned due to genetic similarities with the earlier SARS coronavirus (International Committee on Taxonomy of Viruses, January 7, 2020).

1.1. Clinical and Pathological Features of COVID-19

Depending upon host genetics, manifestations of COVID-19 are diverse but commonly encompass symptoms such as headache, Elevated body temperature, throat irritation, exhaustion, respiratory issues, anosmia and ageusia. Onset of symptoms may occur within one to fourteen days following exposure to the virus. A significant portion of individuals (at least one-third) who contact the virus do not exhibit noticeable symptoms (Oran and Topol, 2021). Among those who do manifest symptoms to a degree requiring classification as patients, the majority (81%) experience mild to moderate manifestations, including mild pneumonia. About 14% develop more severe symptoms characterized by Shortness of breath and reduced oxygen levels or imaging indicating over 50% lung involvement. A smaller proportion, 5%, progress to critical symptoms, marked by respiratory collapse, abrupt circulatory failure, or multiorgan failure (Wu and McGoogan, 2020).

SARS-CoV-2 primarily targets respiratory epithelial cells, causing a spectrum of COVID-19 manifestations from mild cold-like symptoms to severe conditions like acute respiratory distress syndrome and interstitial pneumonia. Severe cases can lead to fatal outcomes and organ damage, affecting the gastrointestinal tract, kidneys, and heart (Zaim et al., 2020). In COVID-19, Multiorgan failure is linked to systemic effects of the virus, including hypoxemia, Disturbed immune system responses, Vascular damage, coagulation abnormalities, and tissue fibrosis (Mokhtari et al., 2020).

1.2. SARS- CoV 2

1.2.1. Genomic Classification and Phylogenetic Analysis

SARS-CoV-2 belongs to the extensive family of coronaviruses, identified by a helical nucleocapsid and Single-stranded RNA with a positive polarity (Naqvi et al., 2020). Through genome phylogenetic analysis, it is classified in the subgenus Sarbecovirus within the Beta coronavirus genus, sharing genetic relatedness with SARS-CoV and bat-derived SARS-related coronaviruses (SARS-CoV). Despite this relatedness, coronaviruses remain distinct entity (Gorbalenya et al., 2020).

1.2.2. Genomic Identity, Variants and Structural Composition

The genetic sequence of SARS-CoV-2 shares around 82% identity with that of MERS-CoV and SARS-CoV and while structural proteins and crucial enzymes display over 90% sequence similarity (Naqvi et al., 2020). WHO recognized five SARS-CoV-2 variants named alpha, beta, gamma, and delta. The structure of coronavirus particles involves a helical nucleocapsid enveloped by a lipid bilayer. This nucleocapsid results from the interaction between RNA and Phosphorylated nucleocapsid (N) proteins. Within the lipid bilayer, there are three to four types of embedded structural proteins. (Siu et al., 2008).

1.2.3. Genome Organization and Open Reading Frames (ORFs)

The RNA sequence of SARS-CoV-2 is notably lengthy, spanning approximately 30kb that encodes a polypeptide of 9860 amino acids (Gorkhali et al., 2021). The genome organization reveals ORF1a, ORF1b, S, M, E, and N as six primary open reading frames and ORF3a/b, ORF6, ORF7a/b, ORF8, ORF9b/c, and ORF10 as accessory open reading frames (Wu et al., 2020) as shown in figure 1.1. The ORFs collectively encode 29 proteins, encompassing both structural and nonstructural proteins.

1.2.4. Nonstructural proteins

The primary open reading frame (ORF1a/b) in SARS-CoV-2 produces polyproteins pp1a and pp1b, encompassing approximately two-thirds of the viral genome. pp1a and pp1b are cleaved by Papain-Like Protease and 3C-Like Protease which yields 16 nonstructural proteins (nsp1–16), including essential components like RNA-dependent RNA polymerase and helicase (Wu et al., 2020). These nonstructural proteins play a pivotal role in viral translation, assembly and replication, enhancing pathogenicity

through modulation of transcriptional regulation and resistance to antiviral responses. The intricate functions of these proteins contribute to the overall virulence of SARS-CoV-2 (Chan et al., 2020).

1.2.5. Structural Proteins

Four structural proteins (spike, envelope, membrane, and nucleocapsid) are encoded at the 3' end of the SARS-CoV-2 genome. (Gorkhali et al., 2021). The spike protein is a substantial glycoprotein that recognizes ACE2 receptors, facilitating virus entry and evading the immune system (Rajpal et al., 2022). It crucially controls viral host range and the host immune response with precise structural arrangements. The assembly involves the membrane glycoprotein (M), while protein E functions as a viroporin, creating ion channels (Siu et al., 2008). SARS-CoV-2 also encodes eight accessory proteins, contributing to its structural complexity (Zhang et al., 2021).

1.2.6. Infection Mechanism and Pathogenesis

The progression of the SARS-CoV-2 infection involves distinct and well-defined stages. These include the virus entering host cells, the replication and transcription of its genetic material, the synthesis of viral proteins, the assembly of new virions, and the eventual release of these virions to start infection in nearby cells. These stages are depicted in figure 1.2 and specific details are provided in below sections;

1.2.6.1. Receptor Recognition

The initiation of infection involves the recognition of the virus by host cell surface receptors. The virus enters host cells by either fusing with the plasma membrane or through endosomes. In both scenarios, viral spike (S) protein binds specifically to the extracellular peptidase domain of the ACE2 (Hoffmann et al., 2020). This binding is facilitated by the S1 domain of the viral S protein, specifically through its RBD.

1.2.6.2. Membrane Fusion

After the activation of the S protein, the S2 domain plays a role in mediating the fusion of the viral and host cell membranes (Huang et al., 2020). In endosomes, cathepsin L activates the spike protein. Cellular serine protease TMPRSS2 activates S protein at the plasma membrane surface by selective cleavage between S1 and S2 domain (Majchrzak and Poręba, 2022).

1.2.6.3. Translation of viral replication machinary

Upon entering the host cell, viral genomic RNA is released into the cytoplasm. By hijacking host machinery it undergoes translation and produces pp1a and pp1ab polyproteins. These polyproteins undergo cleavage by enzymes such as Papain-like protease and 3C-like protease and yield functional nonstructural proteins like Helicase and the RNA replicase–transcriptase complex (RdRp).

1.2.6.4. Replication and Translation of Viral Structure Proteins

RdRp (Nsp12) directs the synthesis of RNA encoding structural proteins (S, E, M, and N) structural proteins which is subsequently translated by ribosomes attached to the endoplasmic reticulum. Viral RNA is replicated in double-membrane vesicles generated from ER. These DMVs provides shield from the host's innate immune system.

1.2.6.5. Virion assembly and Release

Nucleocapsid proteins (N) are assembled in the cytoplasm by interaction between viral genomic RNA and Phosphorylated nucleocapsid (N) proteins (Lebeau et al., 2020). The assembly of mature virions takes place as viral proteins, including S, M, E, and the nucleocapsid, come together. This results in the formation of progeny viruses. Small vesicles carry the mature virions from the ER to the cell surface through the Golgi apparatus. In the final stage, the infection cycle is completed when virions are discharged from the infected cell through exocytosis and ready to find out a new host cells.

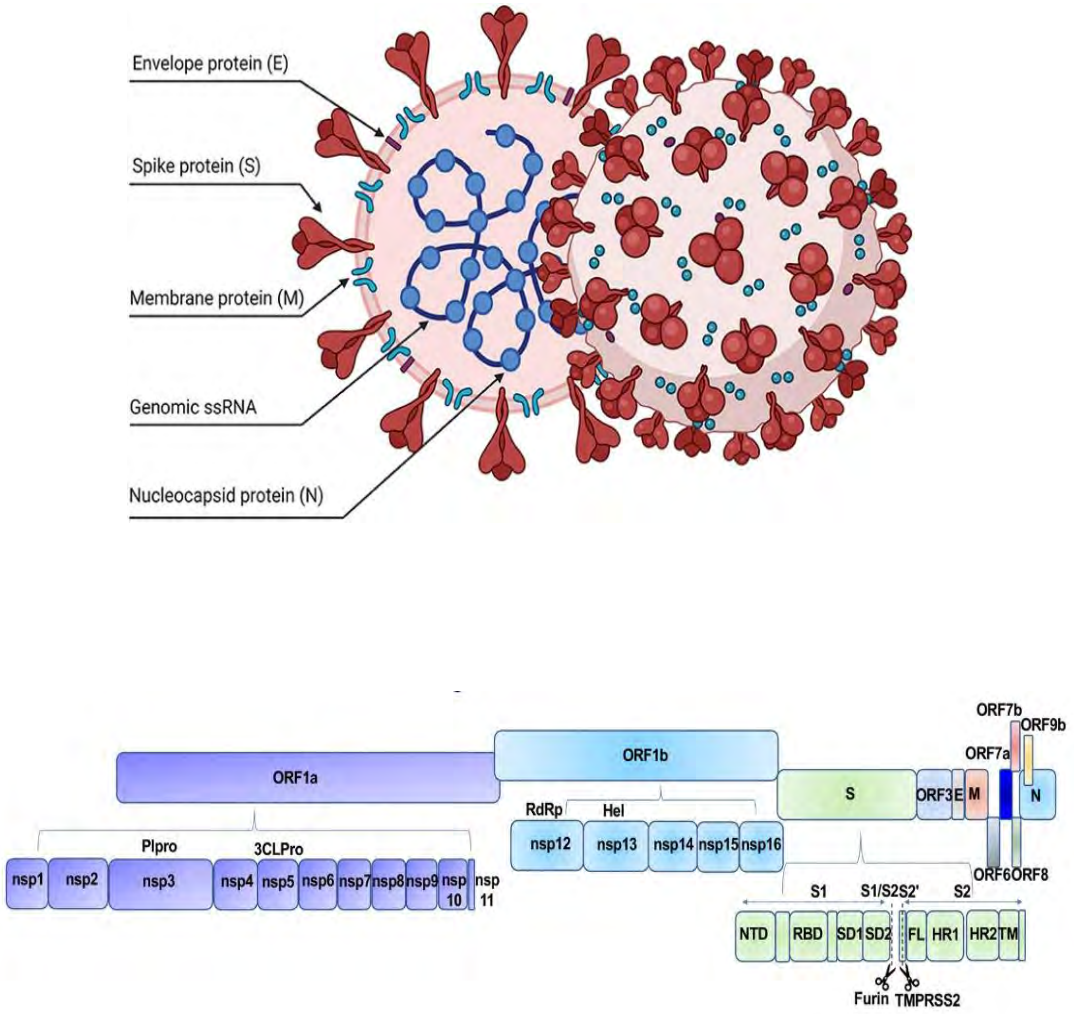


Figure 1.1. SARS-CoV 2 Virion Structure and Genomic Organization (Adopted from Zhang et al., 2021 and Pizzato et al., 2022).

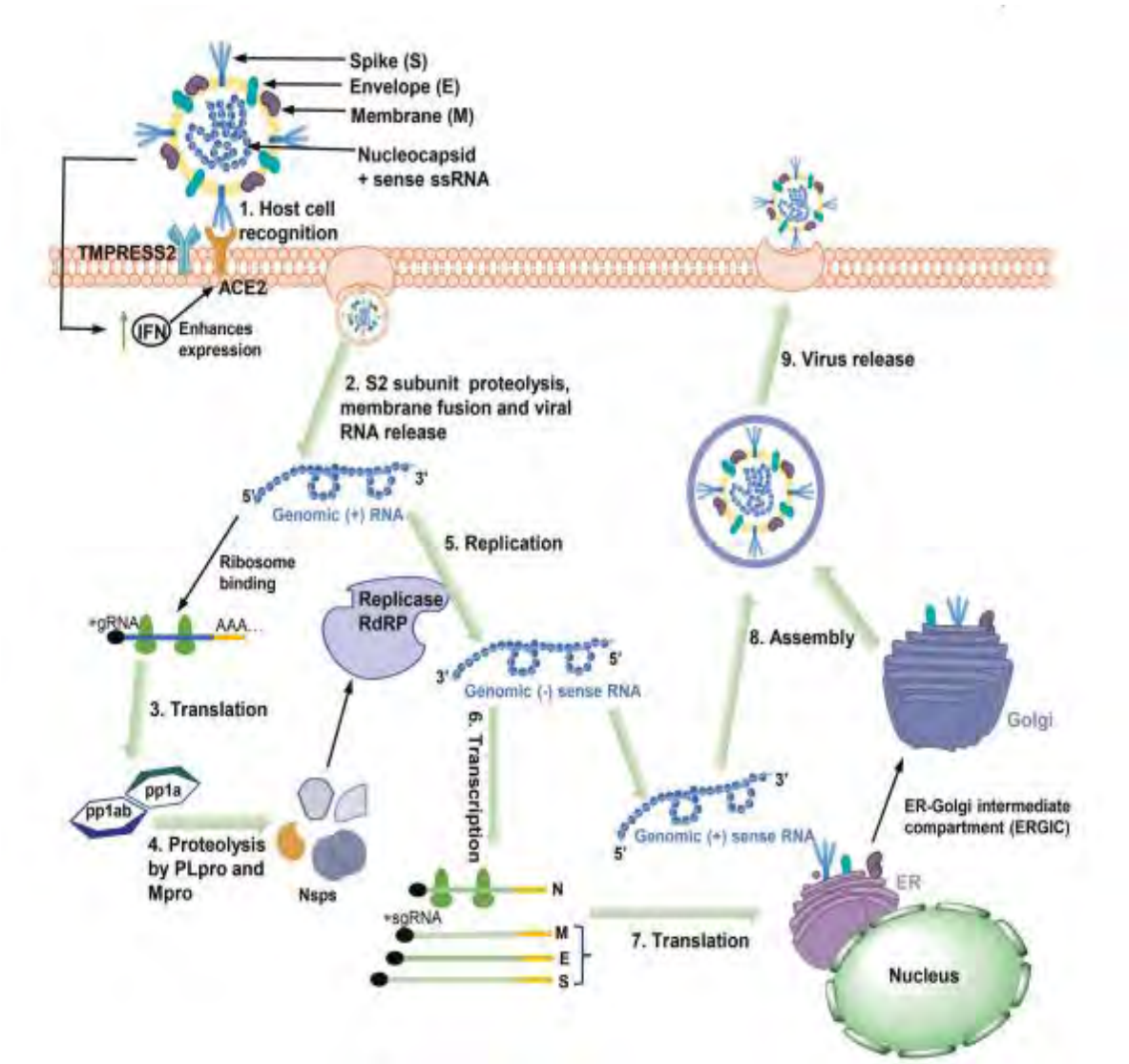


Figure 1.2. SARS-CoV 2 Infection Mechanism and Pathogenesis (Adopted and Modified Yan et al., 2022)

1.3. Genetic Factors Influencing COVID-19 Severity

The clinical manifestations of COVID-19 are determined by interplay of both genetic and environmental factors. Genome-wide association studies have revealed significant associations between host genetic factors (e.g., ACE2, ABO, TMPRSS2, TYK2, PPP1R15A, SLC6A20, IFNAR2, OAS, CCR2, CCR5, TLR7, SLC6A20, HLA) and Susceptibility to SARS-CoV-2 infection (Khouri et al., 2023). These genetic factors impact the severity of infection through their roles in SARS-CoV2 entry receptors (ACE2), accessory proteases (TMPRSS2 and Furin), and immune system genes (HLA, cytokines). ACE2 and TMPRSS2 play crucial roles in determining susceptibility by affecting the binding of the spike protein. Polymorphisms in these genes have been linked to COVID-19 susceptibility, and studies have identified associations with ACE1 polymorphism (Atiku et al., 2023) and variations in HLA alleles (Dobrijević et al., 2023) as additional contributors to susceptibility.

1.3.1. Angiotensin-I-converting enzyme 2 (ACE2)

ACE2 functions as a primary receptor and attachment site for SARS-CoV-2, orchestrating the initial stages of viral engagement. In 2000, ACE2 was identified as the primary homologue of ACE. Its gene is located on chromosome Xp22.2 spanning 39.98 kb of genomic DNA and comprising 18 exons (Donoghue et al., 2000). ACE2 is a zinc metalloprotease, a Type 1 integral membrane glycoprotein with a genomic structure similar to the initial 17 exons of ACE (Tipnis et al., 2000). It is widely expressed in tissues, with highest levels in the testis, small intestine, heart, kidneys, thyroid, and adipose tissue. Moderated expression is seen in the bladder, colon, liver, lungs, and adrenal glands, while ACE2 is relatively lower in the blood, blood vessels, bone marrow, brain, spleen and muscle, indicating diverse distribution with varying concentrations across specific tissues and organs (Lubbe et al., 2020).

1.3.1.1. Structure and Functions

ACE2 comprises 805 amino acids with a 17-amino-acid N-terminal signal peptide and a peptidase domain (aa 19-615) hosting the crucial HEXXH zinc-binding metalloprotease motif. The C-terminal collectrin-like domain (CLD) spans residues 616 to 768, featuring a ferredoxin-like folded neck domain (615-726). The structure includes a 22-amino-acid transmembrane hydrophobic helix region and a 43-amino-acid intracellular segment (Devaux et al., 2020). Resembling collectrin, ACE2's

conserved HEXXH motif in the PD characterizes zinc metalloproteinase that play a pivotal role in catalysis by chelating a zinc ion (Figure 1.3).

ACE2 converts angiotensin I to Ang 1-9 and subsequently to Ang 1-7. The ACE2/Ang 1-7/Mas pathway, mediated by the Mas receptor, promotes vasodilation and anti-proliferative effects, counteracting the vasoconstrictor actions of the ACE/Ang II/AT1R pathway (Carpenter et al., 2022). This pathway offers cardiovascular and pulmonary benefits, protecting against inflammation, apoptosis, and remodeling in respiratory disorders (Liu et al., 2020a). It has protease-independent functions, including cleaving apelin-13 and regulating vasoactive peptides (Camargo et al., 2009).

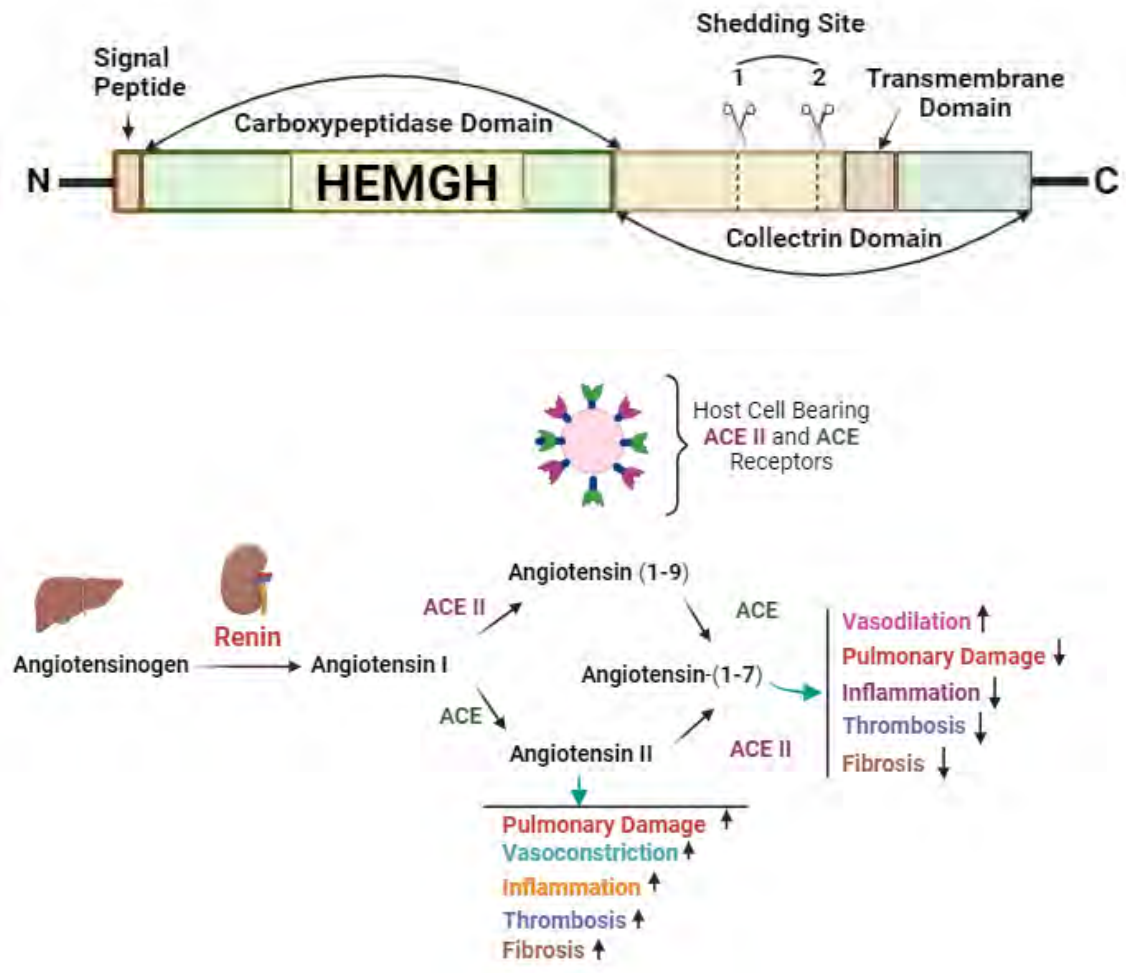


Figure 1.3. ACE2 Protein Structure and Function in RAAS

1.3.1.2. Role in SARS-CoV-2 Pathogenesis

Angiotensin-converting enzyme 2 is significantly involved in the pathogenesis of SARS-CoV-2 (Kuba et al., 2021). As the primary cellular receptor for the virus, ACE2 facilitates its entry into host cells, particularly those in the respiratory system. The interaction of the viral spike protein with ACE2 facilitates internalization of the virus through a series of steps, including membrane fusion and endocytosis (Shirbhate et al., 2021). Furthermore, ACE2 is widely dispersed across a wide range of tissues, including the kidneys, heart, gastrointestinal tract, and lungs (Beyerstedt et al., 2021), this distribution may potentially explain the observed multi-organ involvement in severe cases of COVID-19. Additionally, the Interaction of SARS-CoV-2 with ACE2 receptor disrupts the renin-angiotensin-aldosterone system (RAAS) and this disruption can lead to elevated levels of angiotensin II, potentially contributing to inflammation and vasoconstriction (Ni et al., 2020). The resultant dysregulation of the RAAS could contribute to the severity of COVID-19, especially in triggering inflammatory reactions linked to acute respiratory distress syndrome as demonstrated in figure 1.4.

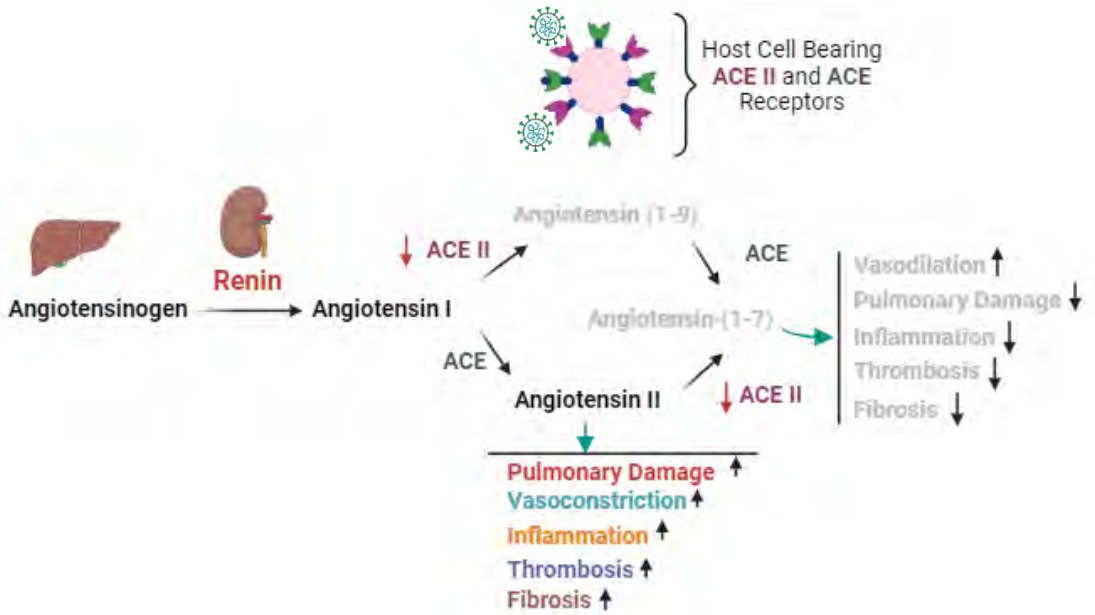


Figure 1.4. ACE2 Significances in COVID 19 Pathogenesis

1.3.1.3. Polymorphism of ACE2 And Covid-19 Severity

The ACE2 gene displays frequent genetic polymorphism, which may potentially influence the association between ACE2 and the S protein of SARS-CoV-2, impacting virus entry and mitigating lung and systemic damage in COVID-19 (Calcagnile et al., 2021). In silico study of 290,000 samples identified ACE2 variants associated with increased susceptibility (S19P, T92I, I21V, K26R, T27A, N64K, I21V, Q102P, E23K, and H378R) and protective mutations (K31R, N33I, H34R, E35K, E37K, D38V, Y50F, N51S, M62V, K68E, F72V, Y83H, G326E, G352V, D355N, Q388L, and D509Y) (Suryamohan et al., 2021). Biochemical analyses revealed altered S-protein affinity in specific variants, with K31R and E37K showing decreased and K26R and T92I showing enhanced affinity. Soluble ACE2 variants K26R and T92I demonstrated increased efficacy in inhibiting S-protein pseudo typed virus entry, highlighting the role of ACE2 variations in SARS-CoV-2 susceptibility (Suryamohan et al., 2021).

In a molecular docking study, six ACE2 missense variants (I21T, A25T, K26R, E37K, T55A, E75G) showed increased affinity for the receptor-binding domain of SARS-CoV-2 S protein compared to the wild type, while 11 variants (I21V, E23K, K26E, T27A, E35K, S43R, Y50F, N51D, N58H, K68E, M82I) displayed decreased affinity (Calcagnile et al., 2021). In another study focusing on ACE2 gene expression within the nasal epithelium, it was observed that children under 10 years of age displayed the lowest levels of ACE2 expression. This observation suggests a plausible correlation with milder presentations of COVID-19 in the pediatric demographic (Bunyavanich et al., 2020).

In a research investigation involving individuals from Spain, the study highlighted the association between the C/C genotype at rs2106809 and the presence of the A allele at rs2285666 within the ACE2 gene, indicating their identification as risk factors for susceptibility to COVID-19 (Sabater Molina et al., 2022). However, a study on the Turkish population found no correlation between ACE2 rs2106809 and rs2285666 polymorphisms and COVID-19 severity (Karakaş Çelik et al., 2021). On the contrary, findings from a German study demonstrated a correlation between the G/G genotype at ACE2 rs2285666 or the presence of the G allele and an almost twofold higher likelihood of SARS-CoV-2 infection. Additionally, individuals with this genotype or

allele exhibited a threefold elevated risk of experiencing severe disease or succumbing to COVID-19-related complications (Möhlendick et al., 2021).

Through SNP genotyping, a research investigation identified five intronic SNPs (rs4646142, rs4240157, rs-6632677, rs2048683, and rs2074192) situated within ACE2, Notably, among these, rs6632677 and rs4646142 emerged as particularly promising genetic markers associated with increased susceptibility to COVID-19-related cardiovascular complications (Ma et al., 2022). In Saudi Arabian population, the correlation between the rs4240157 polymorphism in ACE2 and the severity of COVID-19 was observed, where a high prevalence of alleles C was linked to severe outcomes. Additionally, a study from the Iranian Pasteur Institute, involving 1078 COVID-19 patients, associated the ACE2 rs2285666 C/C genotype with increased mortality rates (Khalilzadeh et al., 2022). These findings highlight the intricate connection between severity of COVID-19 and ACE2 polymorphisms across diverse populations.

1.3.2. Angiotensin-I-converting enzyme (ACE)

ACE, a vital component of the renin-angiotensin aldosterone system (RAAS), significantly contributes to the pathogenesis of COVID-19. The angiotensin-converting enzyme (ACE) is a vital metalloenzyme which is encoded by a gene on chromosome 17q23.3. The gene comprised of 26 exons and 25 introns with a total span of 21,313 bp (Riordan, 2003). The ACE gene gives rise to two distinct isoforms. Exons 1–12 and 14–26 encode somatic ACE, which is widely expressed in a variety of cell types. However, germinal or testicular ACE is transcribed from exons 13–26. (Riordan, 2003).

1.3.2.1. Structural features and functions

Human somatic ACE is a type-I membrane-bound protein, presenting a heavily glycosylated 1227-residue extracellular part, a 22-residue transmembrane section, and 28-residue cytosolic tail (Figure 1.5). The extracellular component comprises two catalytically active domains (N and C), each exhibiting homology and containing a zinc-binding motif. (Soubrier et al., 1988).

ACE functions as a central regulator in diverse physiological processes. Primarily, it hydrolyzes angiotensin I to angiotensin II, contributing to the angiotensin-renin system and influencing blood pressure through vasoconstriction. ACE also participates in the kinin-kallikrein system, degrading vasodilator bradykinin. In neuropeptide regulation, ACE modulates synaptic transmission by cleaving hormones like substance P and

enkephalins (Skidgel et al., 1984). Additionally, it plays a role in opioid signaling, regulates hematopoietic stem cell differentiation, is implicated in cannabinoid signaling, and contributes to amyloid-beta metabolism (Hemming and Selkoe, 2005).

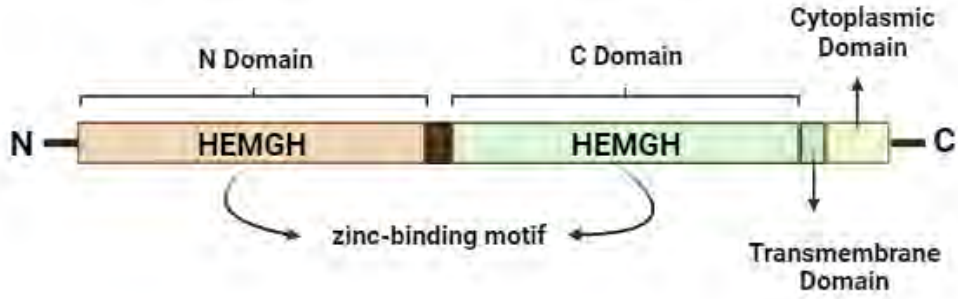


Figure 1.5. ACE Protein Structure

1.3.2.2. ACE Role in SARS COV 2 Infection

ACE is not directly involved COVID-19 pathogenesis. ACE converts Angiotensin-1 to Angiotensin-2, causing adverse effects through AT1 receptor binding. However, in SARS-CoV-2 infection, the virus engages ACE2 receptors, reducing ACE2 activity and disrupting the delicate ACE-ACE2 balance in RAAS. This disruption can lead to elevated ACE activity or reduced ACE2 expression, resulting in uncontrolled Angiotensin-2 levels. This imbalance potentially contributes to acute lung injury, characterized by increased vascular permeability, vasoconstriction, apoptosis, fibrosis, and heightened release of proinflammatory mediators (Zheng and Cao, 2020b).

1.3.2.3. ACE Polymorphism and COVID-19 severity

Genetic variations in the ACE gene, particularly ACE insertion/deletion polymorphism, referred as ACE I/D, are associated with COVID-19 distinct outcomes (Gomes et al., 2022). The ACE D/D genotype is identified as a genetic risk factor for thromboembolic events in COVID-19 pneumonia patients (Calabrese et al., 2021). Multiple studies in diverse populations consistently showed a significant correlation between the ACE insertion/deletion and the clinical severity of COVID-19. The ACE I/I genotype appears to be protective, while the ACE D/D genotype is associated with an increased

risk of severe outcomes (Gunal et al., 2021; Gómez et al., 2020 and Verma et al., 2021). However, some studies did not find a significant association (Karakaş Çelik et al., 2021 and Möhlendick et al., 2021). There are racial differences in ACE gene polymorphisms. Eastern Asian have a greater frequency of I allele, whereas African Americans exhibit the greatest prevalence of the D allele. The COVID-19 outcomes are consistent with these genetic variances; populations with a low occurrence of ACE I allele and high D/D genotype have higher fatality rates, while populations with a high frequency of ACE I allele and a low prevalence of ACE D/D genotype have significantly lower fatality rates (Zheng and Cao, 2020a).

Research investigating the correlation between ACE gene polymorphisms and outcomes in COVID-19 cases, particularly rs1799752 and rs4341 variants, have yielded inconsistent and inconclusive results. Möhlendick et al., (2021) found no impact of ACE rs1799752 polymorphism, while Calabrese et al., (2021) suggested a connection between ACE D/D genotypes and elevated ACE levels in COVID-19. Saad et al., (2021) linked ACE rs1799752 polymorphism, particularly ACE I/D genotype, to an increased risk of COVID-19 and worse outcomes. Martinez-Gomez et al., (2022) in Mexico found no consistent evidence for rs1799752 and rs4344 polymorphisms in the ACE gene influencing COVID-19 outcomes. Alimoradi et al., (2022) identified ACE1 rs4343 GG homozygote as a potential contributor to COVID-19 susceptibility. Iniguez et al., (2021) suggested that rs4343 polymorphisms could be predictive markers for COVID-19 severity, emphasizing the need for further investigation.

1.3.3. Transmembrane serine protease 2 (TMPRSS2)

TMPRSS2 plays a crucial role in COVID-19 as it facilitates the priming of the viral spike protein, promoting viral entry into host cells. TMPRSS2 is a transmembrane protease serine 2. It is encoded by a gene that spans approximately 43.59 kilobases (kb) on chromosome 21q22.3 and structurally organized into 14 exons and 13 introns (Paoloni-Giacobino et al., 1997). It exhibits two alternative splicing variants, generating transcripts of 3.25 kb and 3.21 kb (Park, 2010). Notably, TMPRSS2 features multiple androgen receptor elements (AREs) upstream of the transcription start site and within the first intron. This indicate potential regulatory influence by androgen signaling on TMPRSS2 expression, implicating its role in processes modulated by androgen activity (Shen et al., 2017).

1.3.3.1. Structure and Functions

TMPRSS2 is a 492-amino acid long protein that exhibits a unique structure with a single-pass type II membrane arrangement (Paoloni-Giacobino et al., 1997). It is predominantly located on the cell membrane, has an external C-terminal serine protease domain and an N-terminal intracellular region. Notable structural elements include a catalytic triad in the protease domain, a Scavenger Receptor Cysteine-Rich Domain (SRDR), and a Low-Density Lipoprotein Receptor Class A Domain (LDLRA). TMPRSS2, initially expressed as a single-chain zymogen, undergoes self-cleavage to activate its proteolytic function (Wettstein et al., 2022). The majority of mature proteases remain membrane-bound, with a fraction possibly released extracellularly. This intricate process highlights TMPRSS2's regulatory role in cellular functions (Figure 1.6).

The exact physiological role of TMPRSS2 is uncertain, but it appears to participate in anti-inflammatory processes in the lungs and prostate. In the lungs, it regulates epithelial sodium currents by cleaving an epithelial sodium channel. In the prostate, TMPRSS2 cleaves Proteinase Activated Receptor 2 (PAR2), a G-protein coupled receptor (Wilson et al., 2005). The phenotypic analysis of mice lacking the TMPRSS2-encoded protease reveals that mice without TMPRSS2 appear to exhibit normal function and unaffected phenotype (Kim et al., 2006).

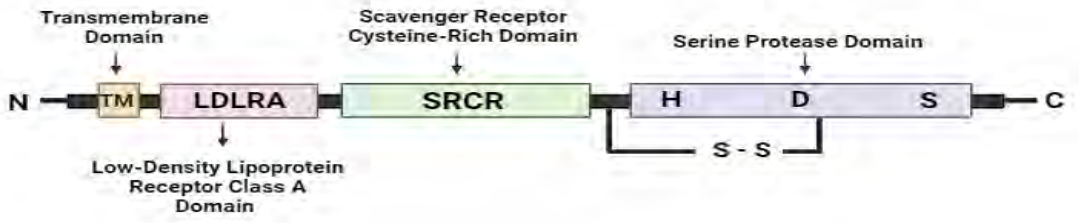


Figure 1.6. TMPRSS2 Protein Structure

1.3.3.2. TMPRSS2 Role in SARS-CoV-2 Pathogenesis

TMPRSS2 expression in the respiratory system plays a vital role in facilitating the virus interaction with host cells in the lungs, there by contributing significant role in the pathogenesis of COVID-19 (Liu et al., 2020b). Functioning as a cell surface protease,

TMPRSS2 plays a pivotal role in initiating the activation of viral S protein after its interaction with ACE2 receptors. This proteolytic activation is crucial for the subsequent fusion of the host cell membrane with virus and enable the virus to efficiently enter host cells and initiating the process of infection (Hoffmann et al., 2020). COVID-19 severity is influenced by age and gender-related TMPRSS2 variations (Ranjan et al., 2021). Pediatric patients experience milder disease, attributed to robust innate immune responses and lower steroid hormones (Mihalopoulos et al., 2020). The male predominance in COVID-19 may be influenced by androgen-regulated TMPRSS2, while preadolescents exhibit a lower incidence of the virus. Hamon and Ween, 2022 findings suggest that the tissue-specific up regulation of TMPRSS2 in smokers has the Possibility to enhance the COVID-19 Severity.

1.3.3.3. TMPRSS2 Polymorphism and COVID-19 Severity

TMPRSS2 genetic variations play a central role in susceptibility of COVID-19, revealing the intricate interplay of host genetic factors. Analysis of TMPRSS2 gene expression reveals four variations (rs464397, rs469390, rs2070788, and rs383510) influencing lung tissue expression (Irham et al., 2020). Increased prevalence of up regulating TMPRSS2 variants in Europeans and Americans, compared to Asians, suggest an elevated COVID-19 infection risk. The T/T genotype of rs464397 is associated with increased TMPRSS2 expression (Sahranavard-Pirbazari et al., 2023). Four regulatory variations (rs112657409, rs713400, rs77675406, and rs11910678) are reported to have impact on TMPRSS2 expression highlighting their potential pathological relevance (Torre-Fuentes et al., 2021). In a German cohort, an intronic variant known as rs383510 within the TMPRSS2 gene showed an association with an increased susceptibility to COVID-19 infection (Schönfelder et al., 2021).

The rs12329760 variant (C to T) is strongly correlated with Increased severity of COVID-19 in Irani population and this study highlighted study the T allele as a risk factor (Yaghoobi et al., 2023). Polymorphisms, including rs2070788, rs7364083, rs9974589, and a sex related rs8134378 in TMPRSS2, related to sex, play pivotal role in both priming of viral proteins and cell adhesion. (Alshahawey et al., 2020). Androgen receptor-regulated TMPRSS2 expression, higher in male prostates, potentially explains sex-related COVID-19 differences, indicating a protective role for females (Asselta et

al., 2020). This contributes to increased virus entry, influencing the higher Fatality rate observed in male COVID-19 patients compared to females.

A prevalent European haplotype linked to the TMPRSS2 gene is absent in Asians, but this haplotype is associated with androgen-responsive enhancement. This haplotype, involving specific SNPs (rs35041537, rs35899679 , rs34783969 , rs734056, rs4290734, rs55964536, rs11702475, rs34624090, rs35074065 and rs463727), is expected to selectively up regulate TMPRSS2 expression in response to androgens (Asselta et al., 2020). Another identified variant, , enhances viral infectivity (Russo et al., 2020).

Specific variants, including rs77675406, rs464397, rs469390, and rs2070788, were associated with varying TMPRSS2 expression levels. The missense mutation in rs469390 showed the Increased pulmonary expression of TMPRSS2 in those with A/A carriers, while homozygous T/T genotypes of rs464397 and rs383510 exhibited the highest lung expression (Irham et al., 2020). For rs2070788, G/G genotypes displayed higher pulmonary TMPRSS2 expression. Variant allele frequencies related to elevated pulmonary TMPRSS2 expression were higher in American, African, and European populations compared to East Asian populations, potentially contributing to increased COVID-19 susceptibility (Irham et al., 2020).

1.3.4. Human Leukocyte Antigens (HLA)

The Human Leukocyte Antigen (HLA) system is a vital component of the immune system, playing a crucial role in providing immunity against infectious agents including SARS-CoV-2 infection. The immune system is a complex network designed to protect the body from various foreign antigens. It comprises two main types of defenses: innate immunity and adaptive immunity. Innate immunity provides immediate, antigen-independent defense through anatomical, physiological, and inflammatory barriers, involving immune cells like phagocytes, natural killer cells, and dendritic cells. It lacks immunologic memory but can initiate an adaptive immune response. Adaptive immunity functions in an antigen-specific manner, featuring memory for faster responses upon re-exposure. T cells, originating in the bone marrow and maturing in the thymus, circulate and collaborate with antigen-presenting cells through unique receptors (TCRs). B cells, originating from bone marrow, directly recognize antigens, producing antibodies and forming memory cells for the humoral immune response. Both innate and adaptive immune responses are crucial for overall immune defense.

1.3.4.1. HLA System

The HLA system, commonly referred to as the major histocompatibility complex, constitutes a crucial aspect of the immune system and regulated by genes situated on chromosome 6p21. These genes encode cell surface molecules designed for presenting immunogenic to the T-cell receptor on T lymphocytes. There are three main classes of HLA molecules:

1.3.4.1.1. HLA Class I

HLA Class I molecules, integral components of the HLA system, are essential for immune surveillance. Comprising classical genes (HLA-A, HLA-B, and HLA-C) and non-classical genes (HLA-E and HLA-F), they present endogenous antigens to CD8+ cytotoxic T cells. Classical genes initiate immune responses against abnormal cells, featuring high polymorphism for diverse antigen recognition. Non-classical genes, like HLA-E, modulate NK cell activity, and HLA-F contributes to immune regulation. The combined functions of classical and non-classical Class I HLA genes underscore the immune system's adaptability and efficacy in monitoring, responding to, and regulating immune reactions.

1.3.4.1.2. HLA Class II

HLA-II molecules are pivotal in presenting exogenous antigens to CD4+ helper T lymphocytes for triggering the immune system. Comprising classical genes (HLA-DP, HLA-DQ, and HLA-DR) and non-classical genes (HLA-DM, HLA-DO), these molecules have alpha and beta polypeptide chains forming a peptide-binding groove. Classical genes exhibit significant polymorphism, diversifying immune recognition, while non-classical genes, like HLA-DM and HLA-DO, contribute specialized functions for efficient and regulated immune responses. The coordinated actions of classical and non-classical HLA Class II genes emphasize their crucial role in orchestrating adaptive immunity.

1.3.4.1.3. HLA class III

The genomic HLA class III region is responsible for encoding numerous molecules crucial in the inflammatory response, encompassing complement components such as factor B, C2, and C4 as well as three heat shock proteins, lymphotoxin, and tumor necrosis factor-alpha.

1.3.4.2. Integral Role in Adaptive and Innate Immunity

HLA system is vital for both adaptive and innate immunity. In adaptive immunity, HLA molecules play a crucial role in presenting antigens to T cells, facilitating activation, proliferation, and diversity. This polymorphic nature enhances the immune system's capacity to recognize a broad range of antigens. In innate immunity, HLA molecules regulate immune responses by interacting with NK cells, influencing their activity based on changes in infected cells. Moreover, the MHC class III region of HLA contributes to inflammatory responses, shaping innate immune reactions.

1.3.4.3. HLA Significance in SARS-CoV-2

Following infection, antigen-presenting cells process the spike antigen into peptides and present them via HLA class I to cytotoxic T cells, leading to the elimination of infected cells. Moreover, HLA class II presentation to CD4+ helper cells trigger B cell differentiation, generating memory B cells and plasma cells that produce neutralizing antibodies, thereby blocking viral interaction with the ACE-2 receptor (Figure 1.7). This coordinated immune response is pivotal in defending against SARS-CoV-2.

1.3.4.4. HLA Polymorphism and COVID-19 Severity

HLA gene variations significantly influence susceptibility, severity, and immune responses to SARS-CoV-2 infection. Analyzing individuals' HLA profiles reveals diverse manifestations of COVID-19, with specific HLA types associated with varying susceptibility.

In the Sardinian population, the extended HLA-A*02:05, HLA-B*58:01, HLLA-C*07:01 and HLA-DRB1*03:01 haplotype has been linked to provide immunity against COVID-19 infection, while the presence of HLA-DRB1*08:01 in the same population is linked with a negative impact on COVID-19 outcomes (Littera et al., 2020). Specifically, in Italy, an increased occurrence of HLA-DRB1*08 was detected in COVID-positive individuals, correlating with increased mortality (Littera et al., 2020). In another study peptide binding prediction analyses indicated that DRB1*08 alleles lacked high-affinity binding to viral peptides (Amoroso et al., 2021).

In the Chinese population, studies identified associations between increased susceptibility to adverse COVID-19 outcomes and the HLA-A*11:01, HLA-B*51:01, and HLA-C*14:02 alleles (Wang et al., 2020), while the HLA-B22 serotype, encompassing alleles B*54:01, B*55:01, B*55:07, B*55:12, and B*56:01, was

suggested as a potential susceptibility marker due to compromised binding to SARS-CoV-2 (Yung et al., 2021). In Japan, presence of the DRB109:01 allele was linked to heightened susceptibility to severe disease outcomes (Anzurez et al., 2021), and a Spanish study found associations Among particular HLA alleles and severity of COVID-19, including HLA-A32 being more common in healthy individuals and HLA-B39 and HLA-C16 more prevalent in individuals with COVID-19 (Lorente et al., 2021).

In Greece, positive associations were identified for HLA-DRB1*01 alleles in total COVID-19 patients, with hospitalized patients showing increased frequencies of HLA-A*11, HLA-A*23, and DRB1*09 alleles (Weiner et al., 2021). In Iran, certain HLA alleles were elevated in moderate and critical cases, with a negative association between DRB104 and severe disease (Farahani et al., 2021). In the United States, HLA-B*51:01 and HLA-A*26:01 alleles were potentially protective, while HLA-A*03:01, HLA-DRB1*15:01, and the B44 serotype genotypes were positively associated, indicating increased severity in COVID-19 cases. An Indian study also associated specific HLA alleles with Severe manifestations of COVID-19 (Alnaqbi et al., 2022).

In another Indian population study, HLA genotyping in 54 SARS-CoV-2-positive patients and 42 asymptomatic individuals revealed specific HLA alleles associated with severe COVID-19 cases, including HLA-C*04:01:01:01, HLA-DRB5*01:01:01:02, DQA1*03:01:01:01, HLA-DPB1*04:01:01:41, and HLA-DPA1*01:03:01:02, (Vishnubhotla et al., 2021). Additionally, in South Asian populations, higher frequencies of HLA-B*51 were associated with fatal cases, while HLA-B*35 was more prevalent in mild infections. The presence of DRB113 was significantly observed in fatal cases within the HLA class II group (Naemi et al., 2021).

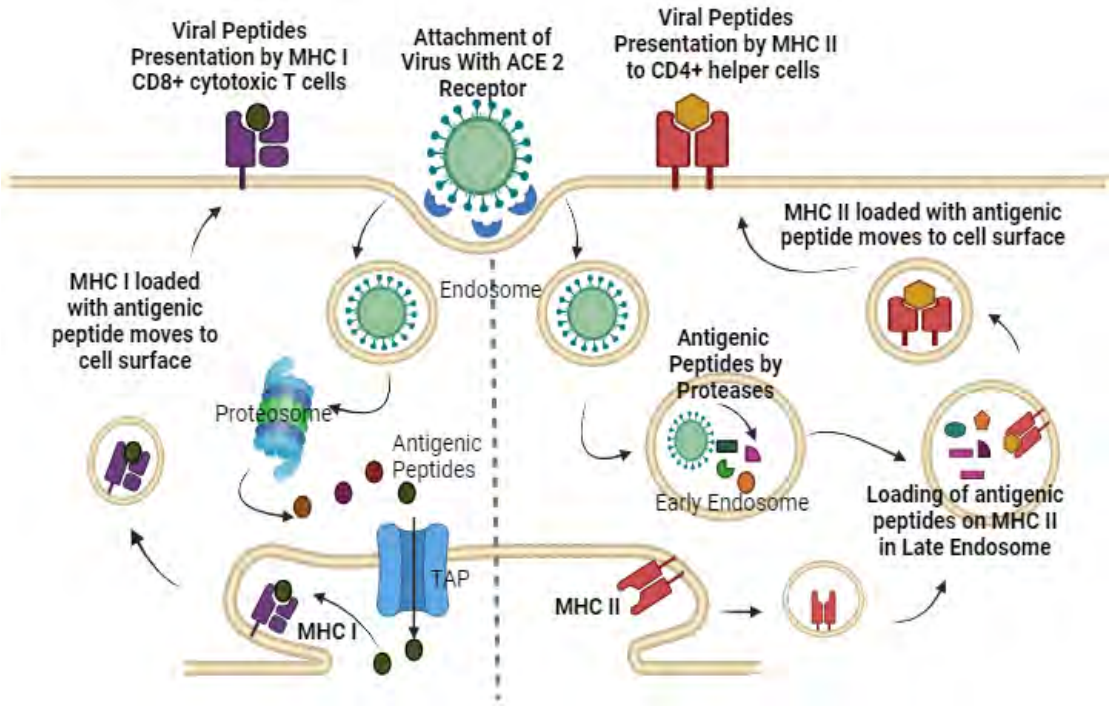


Figure 1.7. Role of MHC I and MHC II in viral peptides presentation to T cells

1.4. A Pakistani Perspective on COVID-19 Research

As of November 22, 2023, the global impact of COVID-19 has been substantial, with reported cases reaching 772,166,517 worldwide. South Asia, including Pakistan, reported a total of 61,210,423 cases, displaying comparatively lower infection and mortality rates than Europe and North America.

From January 3, 2020, to November 22, 2023, Pakistan documented 1,580,631 COVID-19 cases and 30,656 deaths, positioning itself as the third highest in reported cases among Asian nations. The first case was reported in Karachi on February 26, 2020. Several factors contributed to the unique impact of COVID-19 in Pakistan (Raza et al., 2022).

Genetic insights from research in the Pakistani population revealed a nuanced understanding of genetic variants influencing COVID-19 impact. Notable findings included infrequent ACE2 variations, with specific variants such as p.Pro86Leu and 15C > G showing significant associations (Waryah et al., 2023). ACE2 polymorphisms, especially rs781378335 and rs2285666, independently correlated with severe COVID-19, highlighting the genetic complexities of disease susceptibility (Sidhwani et al., 2023). The study in the Asian population links Allele D of ACE insertion/deletion polymorphism to infection and mortality rates, indicating a potential genetic influence on COVID-19 susceptibility and outcomes (Pati et al., 2020).

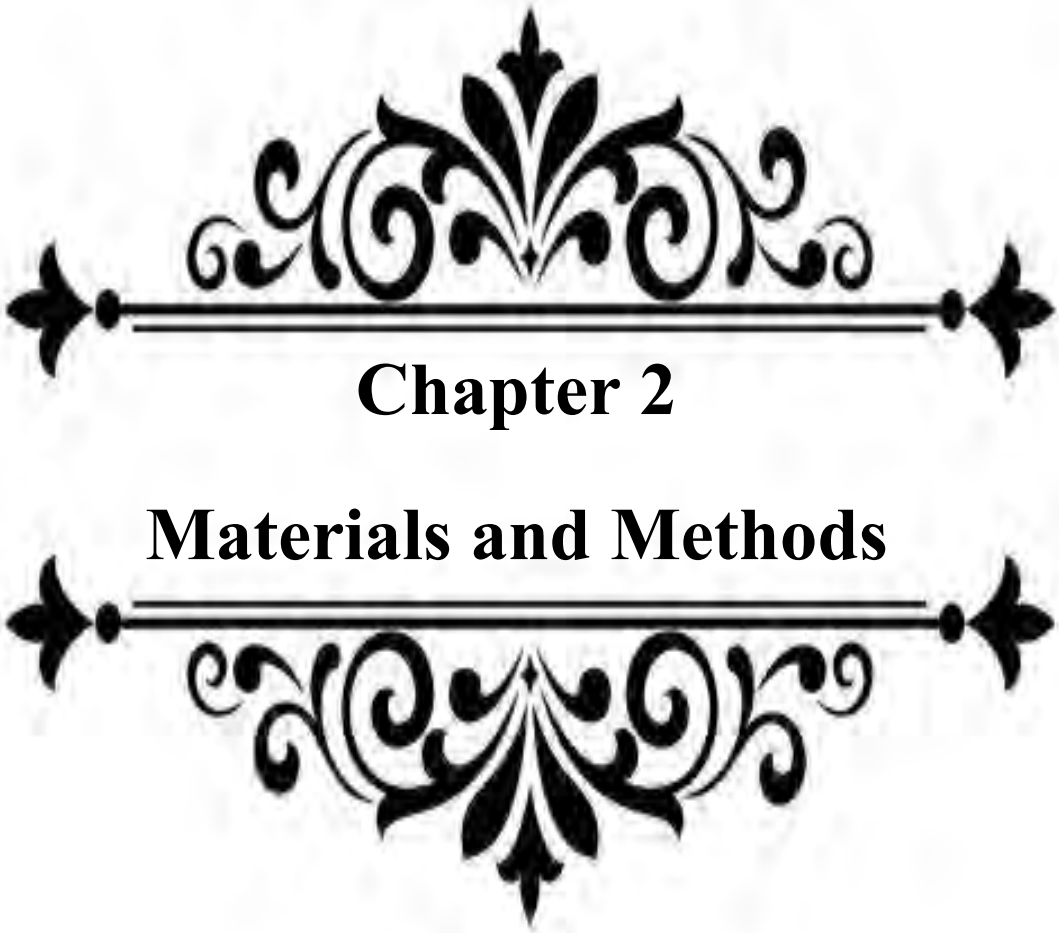
Studies on South Asian COVID-19 patients, including Pakistanis, underscored a significant correlation between HLA genotypes and infection severity. In fatal cases, HLA-B*51 frequency was notably higher, while HLA-B*35 prevailed in mild cases (Naemi et al., 2021). Additionally, DRB1*13 in HLA class II showed a significant increase in fatal cases (Naemi et al., 2021). Another investigation involving 92 individuals diagnosed with COVID-19 emphasized a notable association between HLA-C*04 and HLA-B*35 alleles and COVID-19 severity, emphasizing the potential contributions of adaptive and innate immune responses (Tay et al., 2023).

Understanding the intricacies of genetic variations will be crucial to understand host response against viral infection.

1.5. Aim and Objectives

This study aims to identify genetic risk factors for COVID-19 severity in the Pakistani population through three main objectives:

1. To explore severity-associated Single Nucleotide Polymorphisms (SNPs) in TMPRSS2 gene.
2. To investigate the role of ACE insertion/deletion polymorphism in COVID-19 severity.
3. To explore potential link of HLA Class-I alleles with COVID-19 severity.



Chapter 2

Materials and Methods

2. Materials and Methods

2.1. Ethical Approval

Ethical approval was needed for investigation, which involves acquiring blood samples from COVID-19 patients. In order to meet the ethical standards, the permission from Quaid-e-Azam University's Bioethics Committee (BEC).

2.2. Sample Acquisition and Preservation

To determine the association of genetic risk factors, patients were stratified into mild and severe groups, according to WHO definitions for the COVID-19 severity. The 10ml sterile syringes were used to draw blood samples from 131 COVID-19 patients belonging to different racial groups. To prevent coagulation these collected blood samples were transferred into EDTA tubes and stored at 4°C until DNA extraction.

2.3. Genomic DNA extraction

Genomic DNA extraction from blood is a process of isolating the genetic material from nucleated cells, primarily white blood cells. This involves breaking cell membranes, releasing cellular contents, and removing impurities, resulting in purified genomic DNA. Thermofisher Scientific's Genomic DNA Purification Kit was used to extract COVID-19 patient DNA using the following protocol:

i. **Lysis and Denaturation:** 200 µl of patient blood sample was taken in an Eppendorf tube and 40 µl of lysis solution was added to it. The mixture was incubated for 5-10 minutes at 65°C.
ii. **Chloroform Extraction:** After lysis 600 µl of chloroform was added. The solution was thoroughly mixed by inverting the tube several times and centrifuged at 10,000 rpm for 2 minutes.
iii. **DNA Precipitation:** The aqueous layer was transferred to another tube. 720 µl of double distilled water was added to 80 µl of the concentrated Precipitation Solution. The aqueous phase was mixed with 800 µl of precipitation solution for 1-2 minutes by inverting the tube then centrifuged at 10,000 rpm for two minutes.
iv. **Solubilizing DNA Pellet:** After discarding the supernatant, the DNA pellet was dissolved in 100 µl of NaCl solution by vortexing it.

v. **DNA Ethanol Precipitation:** 300 ul of ice-cold ethanol was added to precipitate the DNA by letting it sit for 10 minutes at -20°C. Then centrifuged at 10,000 rpm for 3-4 minutes.

vi. **Washing and Solubilizing:** Residual ethanol was removed then DNA pellet was washed with 70 % cold ethanol and finally dissolved in 100 ul TE buffer.

2.4. Quantification and Normalization

Nanodrop (Titerteck-Berthold Technologies, Colibri, Germany) spectrophotometer is used for the quantitative assessment of extracted DNA. The Nanodrop works on the principle of Beer-Lambert Law, and it measures the concentration of extracted DNA at 260 nm meanwhile it also takes absorbance at 280 nm and 230 nm for proteins and salts contamination respectively. The ratios it provides (A260/ 280 and A260 / 230) tell about purity of DNA samples.

Following the quantification process, a calculated volume of nuclease-free water was added to dilute the DNA samples uniformly. The following formula was used to normalize the samples.

Initial concentration x Initial volume = Final concentration x Final volume

$$C1V1 = C2V2$$

2.5. Agarose Gel Electrophoresis (1%)

Qualitative assessment of stock and dilutions of the extracted DNA was done by using 1 % agarose gel electrophoresis according to the following protocol:

i. **Agarose Gel Preparation:** 1g of agarose was weighed and mixed into 100ml of 1X TBE buffer in a conical flask. The mixture was microwaved for 2-3 minutes until the agarose was totally dissolved. To stain and highlight the DNA, 10µl of ethidium bromide (EtBr) was added to the gel after it had cooled for one minute.

ii. **Gel Casting:** Gel caster was set, and comb was placed in it. The tray was filled with the melted agarose mixture. The air bubbles in the gel were removed by using a pipette tip. The gel was allowed to cool at room temperature so that it could polymerize. When the gel was polymerized, it was put into the gel tank with 1X TBE running buffer.

iii. **Sample Loading:** 3ul of DNA samples were mixed with 3ul of loading dye. Samples were loaded into the wells using a micropipette.

iv. **Electrophoresis:** After connecting the gel apparatus to the power supply, the gel was run for 40-45 minutes at 100-110 volts until the DNA bands migrated sufficiently.

v. **Gel Visualization:** The gel was visualized in the Syngene Bioimaging gel documentation system (Syngene, UK).

2.6. Primer Designing

2.6.1. Genotyping primers for ACE Polymorphism

ACE insertion/deletion of Alu elements in intron 16 was first identified by (Rigat et al., 1990). Tiret et al., (1993) designed two oligonucleotides for detection of 190bp fragment in absence and 490bp fragment in presence of Alu elements. Gómez et al., (2020) used these classical primers to detect the effect of ACE insertion/deletion in COVID-19. For ACE insertion/deletion investigation in COVID-19 patient included in the study same primer pair was used and in listed in table 2.1.

2.6.2. Sequencing Primers for TMPRSS2

SNPs in TMPRSS2 can help in explaining the difference in susceptibility of COVID-19 in mild and severe group. Exon 3 having SNPs rs61735793, rs61735791 and rs3787950 have been selected due to earlier reports in several studies (Sahranavard-Pirbazari et al., 2023). The Ensembl Genome Browser (https://asia.ensemblorg/index.html) was used to retrieve exon sequences and genomic locations of the chosen SNPs.

The Primer3 software (https://primer3.ut.ee/) and NCBI Primer BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi) were utilized to pick primers for the chosen exon. The primer requirements, which included primer length, melting temperature (Tm), and GC content, were specified beforehand.

2.6.3. Primer Quality Assurance

To assess ACE and TMPRSS2 primer's characteristics, including specificity, secondary structures, and melting temperature following online tools are used:

i. **BLAST-Like Alignment Tool (BLAT)** aligns primers to a target genome or sequence database, which helps determine the specificity of the primers

(https://genome.ucsc.edu/cgibin/hgBlat). By locating probable off-target binding sites, this type of analysis aids in confirming that the primers precisely and specifically amplify the targeted genomic areas.

ii. **OligoCalc** (http://biotools.nubic.northwestern.edu/OligoCalc.html) is a useful tool for primer assurance tests since it provides important information about primer design and optimization. Melting temperature (Tm) and GC content are two important details it provides that help with the accuracy of primer properties.

iii. **In silico PCR** (https://genome.ucsc.edu/cgi-bin/hgPcr) is a computational approach. It simulates DNA amplification and predicts expected products. It determines primer specificity by comparing it to the target genome. This way, potential off-target sites are detected. Knowing original amplicon size helps to determine expected product sizes.

2.6.4. HLA-A*11 Sequence Specific Primers

Several studies have reported that certain HLA alleles are linked to a higher or lower risk of contracting SARS-CoV-2. For this study specific primers were designed to detect HLA-A*11 alleles according to phototyping protocol by (Bunce et al., 1995)

2.6.5. Primers Quality Assurance

Non-specific binding in the HLA sequence-specific primers was avoided by using two tools from the IPD-IMGT/HLA Database of the EMBL-EBI. The IPD-IMGT/HLA (https://www.ebi.ac.uk/ipd/imgt/hla/) Database is a specialized and comprehensive resource that provides information on the human leukocyte antigen (HLA) system.

i. **The Sequence Alignment Tool** played a crucial part in the HLA allele analysis (https://www.ebi.ac.uk/ipd/imgt/hla/alignment/). The polymorphisms unique to every HLA allele were systematically identified using this method.

ii. **Probe and Primer Search Tool** (https://www.ebi.ac.uk/ipd/imgt/hla/probe/) played a crucial role in ensuring the specificity of HLA allele-specific primers. Its functionality extends to ruling out nonspecific binding by highlighting the specific HLA alleles to which a particular probe or primer can potentially bind.

2.7. PCR Amplification

Prior to ACE genotyping, TMPRSS2 sequencing and HLA typing, sample DNA was amplified through Polymerase Chain Reaction (PCR). In 1983 Kary Mullis invented PCR, a molecular biology technique (Mullis, K.B., 1990). It amplifies DNA, allowing the precise amplification of specific genomic regions as specified above.

2.7.1. ACE, TMPRSS2 Amplification

Prior to sequencing exon 3 of TMPRSS2 and ACE insertion/deletion genotyping, DNA sample of each included individual was used to amplify through procedure specified below;

i. For the analysis of the 131 samples, 25 µL reactions were prepared using the reagent combination listed in table 2.2.
ii. In 200 µL PCR tubes, a mixture comprising PCR buffer, MgCl2, dNTPs, forward and reverse primers, Taq polymerase and nuclease-free water, with volumes and concentrations as outlined in table 2.2, was prepared.
iii. Following centrifugation, the tubes were transferred to a thermocycler (Biometra Germany). The thermocycler was programmed to run based on the specifications detailed in table 2.3.

2.7.2. HLA Typing PCR with abm MasterMix

The abm MegaFi™ Pro Fidelity 2X PCR MasterMix was used to perform HLA typing of 130 individuals in compliance with the protocol.

i. 13µL reactions were prepared using the concentrations and volume requirements listed in table 2.3.
ii. Reaction components were mixed and briefly centrifuged, then transferred to a thermocycler (Biometra Germany). The thermocycler was programmed to run based on the specifications detailed in table 2.4.
iii. To validate the amplification, the PCR product was analyzed through electrophoresis on a 2% agarose gel by following protocol mentioned in section 2.5 with slight modifications.

2.8. ExoSAP-IT Purification of PCR Product

ExoSAP-IT is a reagent used for purifying PCR products. It includes Exonuclease I and Shrimp Alkaline Phosphatase to remove leftover primers and nucleotides which

improves the accuracy of sequencing. The amplified PCR product underwent purification using the ExoSAP IT. PCR Product Cleanup Reagent following the protocol outlined below:

i. **Reaction Mixture Preparation:** 7µl purification reaction was prepared by mixing 5µl PCR product with 2µl ExoSAP-IT.
ii. **Enzymatic Cleanup:** The reaction was incubated at 37°C in the thermocycler for 15 minutes to eliminate leftover primers and nucleotides.
iii. **Enzyme Inactivation:** Following the incubation period, ExoSAP-IT was heated in the thermocycler for 15 minutes to 80°C to deactivate it.
iv. **Product Purification:** The cleaned PCR product was directly used for DNA sequencing.

2.9. Sequencing

The BigDye™ Terminator v3.1 Cycle Sequencing Kit (Thermofisher Scientific) was utilized for the sequencing of ACE2 and TMPRSS2 exons. 4µl (30–40 ng) of the ExoSAP purified PCR product, 4.0µl of the Terminator Ready Reaction Mix, and 1µl of the matching primer were combined in a 0.2 ml tube. To make up the remaining volume, 10µl of ddH2O was added. After mixing, centrifuging, and incubating for 25 cycles at 96°C for 10 seconds, 50°C for 5–10 seconds, and 60°C for 4 minutes, the reaction was placed in the thermocycler. We kept the tubes in storage at 4°C.

2.10. Sequencing Product Purification

The BigDye XTerminator™ was used to purify the sequencing reaction in accordance with the subsequent protocol:

i. The SAM solution was combined with the BigDye XTerminator™ beads after they had been vortexed for 8 to 10 seconds.
ii. To prepare a 110µl SAM/BigDye XTerminator™ bead working solution, 45 µL of SAM solution and 90 µL of BigDye XTerminator™ bead solution was mixed together.
iii. Each sample received the bead mix, and MicroAmp™ Clear Adhesive Film was used to seal the plate.
iv. After centrifuging at 1000 g for two minutes, the 96-well plate was vortexed for 20 minutes at 1,800 rpm.

2.11. Data Analysis

2.11.1. ACE Genotyping by gel electrophoresis

Genotyping of ACE insertion/deletion variations was conducted through 2% agarose gel electrophoresis. In order to genotype 10 µL of PCR product was mixed with 10 µL of loading dye and loaded on gel. The gel was subjected to electrophoresis for 45-50 minutes at 100 to 110 volts. The amplified samples were resolved along with 100bp ladder and resultant bands were analyzed to detect insertion or deletion of Alu repeat.

2.12.2. Sequencing Data Analysis

The sequencing data was analyzed with several different software tools. BioEdit (http://www.mbio.ncsu.edu/BioEdit/bioedit.html), Sequencher® version 5.4.6 by Gene Codes and CLC Genomics Workbench (http://www.clcbio.com/products/clc-genomics-workbench) were utilized for comprehensive examination. These tools allow thorough assessment interpretation and analysis of sequencing data.

2.12.3. HLA Typing Gel Electrophoresis

HLA typing of the selected samples were investigated by using 2% agarose gel electrophoresis in accordance with the comprehensive technique described in section 4.8. A modification to the protocol involved filling the wells with 5µL of loading dye and 5µL of the HLA PCR product. The purpose of this modification was to improve the clarity of differentiation between any non-specific HLA bands and specific ones. The gel electrophoresis procedure took 45–50 minutes to complete, with an electrical voltage of between 100 and 110 volts.

2.13. Statistical test

2.13.1. Non-Parametric Test:

For the data analysis, non-parametric tests, Fisher's exact test and the chi-square test are employed. These tests are distribution-free and require few distributional assumptions, particularly used when dealing with ordinal or categorical data, small sample sizes and situations where normality assumptions do not hold.

2.13.2. Fisher's Exact Test:

Fisher's exact test is used to determine significant associations between two categorical variables. The test is commonly used for analyzing data in contingency tables,

providing an exact p-value and odd ratios along with confidence intervals for determining statistical significance.

2.13.3. Chi-Square Test:

The chi-square test is a statistical test that is used to determine whether there is a significant association between two categorical variables. It is often applied to contingency tables, which display the joint distribution of the two variables.

2.13.4. Contingency Tables:

To facilitate the application of Fisher's exact test and the chi-square test, data is organized into either 2 by 3 contingency table for ACE and TMPRSS2 genotypes or 2 by 2 contingency table to explore the association between ACE, TMPRSS2, and HLA-A*11 alleles with the severity of COVID-19.

2.13.5. Statistical Analysis Software:

To perform Fisher's exact test and the chi-square the R project (https://www.r-project.org/) and R Studio (https://posit.co/products/open-source/rstudio/) are used. R is a specialized programming language used for statistical analysis and it provides wide range of libraries and packages for specific analyses or visualizations. R Studio serves as an integrated development environment (IDE), designed to enhance the R coding experience by providing a user-friendly interface with tools for real-time code execution, plotting, and project management. Together, R and R Studio create a powerful environment for statistical computing, data analysis, and visualization.

Table 2.1. ACE, TMPRSS2 and HLA Primers

Primer	Sequence	Length	Tm (°C)	Product Size
hACE Intron-16-F	CTGGAGACCACTCCCATCCTTTCT	24bp	66.1	191bp
hACE Intron-16-R	GATGTGGCCATCACATTCGTCAGAT	25bp	68.4	
hTMPRSS2 Exon-3-F	CACCCAGCCTTGTAGTACCTT	21bp	59.37	609 bp
hTMPRSS2 Exon-3-R	CAGGAAGAGAGTGGAAAGGCA	21bp	59.65	
HLAA*11 Exon-2-3-F	GACACGGAATGTGAAGGCCCAG	22bp	58.6	552bp
HLAA*11 Exon-2-3-R	GAGCCACTCCACGCACCG	18bp	59.7	
HLA Typing Control Primer				
hAPC EX-15-F	TCTTAATTTACCAGTGAGGGACG	23bp	57.92	251bp
hAPC EX-15-R	TAAGAGTGCCAACCAAAAATGC	22bp	58.01	

Table 2.2. PCR Reagents

Reagent	Volume Required (per 25µl Reaction)
10X Taq. Buffer with (NH4)2SO4	2.5µl
Magnesium Chloride (25Mm)	1.5-2µl
dNTPs (10Mm)	0.5µl
Forward Primer (10Mm)	1µl
Reverse Primer (10Mm)	1µl
Template DNA	2µl
Taq. Polymerase	0.2µl
Nuclease Free Water	Remaining Volume (15.8-16.8µl)
HLA Typing (MegaFi™ Pro Fidelity 2X PCR MasterMixper 13µl reaction)	
MegaFi™ Pro Fidelity 2X PCR MasterMix	6µl
Template DNA	2µl (40-60 ng/µl)
Primer Mix (10nM conc. per primer)	1µl
Nuclease Free Water	Remaining Volume (6.8µl)

Table 2.3. PCR Profile for ACE and TMPRSS2

Step	Cycles	Temperature (°C)	Time
Initial Denaturation		95°	10 min
Denaturation	40	95°	1 min
Annealing		60° - 64°	1 min
Extension		72°	20s - 1 min
Final Extension		72°	10 min
Pause		4°	∞

Table 2.4. PCR Profile for HLA Typing

Step	Cycles	Temperature (°C)	Time
Initial Denaturation		95°	5 min
Denaturation	35	95°	15 min
Annealing		60°	15 min
Extension		72°	30s
Final Extension		72°	5min
Pause		4°	∞

Table 2.5. Agarose Gel Reagents

Reagent	Required Volume (per 100ml gel)
Agarose	1-2g
1X Tris-borate-EDTA (TBE) buffer	100ml
Ethidium bromide (EtBr)	10µl
Loading dye	3µl

Chapter 3

Results

3. Results

3.1. Study Cohort

This study comprises a cohort of 131 samples who were tested positive for COVID-19. The collected samples were divided into Mild and Severe groups based on their symptoms and disease severity after COVID-19 infection. The mild group has 53 samples and severe group has 78 samples, respectively. The severe group include patients who exhibit pronounced symptoms, including severe pneumonia, shortness of breath, high fever, persistent cough, lack of oxygen and organ failure. The patients in Mild group are characterized by milder symptoms such as mild respiratory symptoms, general discomfort like fatigue and body aches, low-grade fever, and no or mild pneumonia.

3.2. ACE Genotyping

To explore the association of COVID-19 severity with ACE insertion/deletion (I/D) polymorphism (rs4646994), genotyping was conducted by polymerase chain reaction (PCR) followed by gel electrophoresis. The genotyping was carried out on 131 samples belonging to mild and severe groups. The ACE allele status was judged by the analysis of bands obtained on agarose gel. A 190bp PCR product indicates deletion or absence of Alu element, whereas 490bp PCR product indicates presence of Alu repeat (Figure 3.1).

Among 131 samples, in 40 individuals are homozygous for single 190bp band and are thus labeled as D/D genotype. In 63 individuals, the presence of both 190bp and 490bp bands suggests the heterozygous I/D genotype, whereas in 28 individuals a single 490bp band, indicated homozygous I/I genotype.

In the subgroup analysis of mild group having 53 individuals, 17 were homozygous I/I, 23 showed the heterozygous I/D genotype, and 13 had the homozygous D/D genotype. Among 78 severe cases, 11 individuals had the homozygous I/I genotype, 40 had heterozygous I/D genotype, and 27 had the homozygous D/D genotype.

3.3. ACE Sequencing

To confirm the presence and absence of Alu repeat, Sanger sequencing was performed on two representative samples. The sequencing result of a D/D homozygote individual

confirmed the absence of Alu repeat in 190bp amplified product (Figure 3.2), whereas the sequencing of I/I homozygote individual identified the insertion of a 289 base pairs of a Alu repeat sequence in intron 16 of ACE gene (Figure 3.3). Further sequence comparison of identified insertion of 289bp revealed features consistent to previously reported sequence of I/I allele (Figure 3.4) (Manson et al., 2010). Alu repeat have left arm towards 5' end, in middle linker sequence, right arm towards 3' and Poly-A tail, structural features includes Retinoic Acid Receptor response element (RARE) sequence, ERE-like sequence identified in the 3' and 5' arms (Figure 3.4).

3.4. TMPRSS2 Sequencing

To investigate the association of two missense variants (rs61735793, rs61735791) and a synonymous variant rs3787950 present in TMPRSS2 gene in 131 individuals, comprising 78 severe and 53 mild cases as mentioned in section 3.1.

The synonymous variant rs3787950 (g.41494369T>C) was detected in 27 individuals, constituting 20.7% of the cohort of 131. This specific variant involves an A to G substitution at the 225th position in exon 3 of TMPRSS2. Importantly, this substitution in the codon does not induce a change in the protein sequence (p. Thr75=).

Further analysis revealed that rs3787950 was observed in 17 patients from the severe infection group and 10 patients from the mild infection group. In severe group this SNP is present in the heterozygous (A/G) state whereas in mild group 2 out of 10 individuals were homozygous G/G and remaining 8 have heterozygous A/G genotype.

3.5. HLA Typing

This study also looked at HLA-A*11 alleles in 130 COVID-19 samples. The results revealed distinct HLA-A*11 patterns among the severe group. 552bp PCR product indicates presence of HLA-A*11 and 251bp PCR product indicate control APC allele. Notably, 38 out of 78 individuals (48.7%) in the severe group exhibited the presence of HLA-A11, whereas in mild case 33 out of 53 individuals (62.3%) showed the HLA-A*11 type.

3.6. Statistical Analysis

3.6.1. ACE

3.6.1.2. Genotypic association

To investigate association of ACE insertion/deletion (I/D) polymorphism with COVID-19 severity, Fisher's Exact Test and Pearson's Chi-squared Test were used. Fisher's Exact Test yielded a p-value of 0.04676, indicating sufficient evidence for ACE I/D polymorphism association with COVID-19 severity. Similarly, Pearson's Chi-squared test reported a statistic of 0.978 with 2 degrees of freedom and a p-value of 0.0444, and also supports a significant association between ACE I/D polymorphism association with COVID-19 severity.

3.6.1.2. Allelic Association

The Investigation of ACE allelic association with disease severity, also yielded a p-value of 0.03153 with Fisher's Exact Test and thus indicate a significant association between ACE allelic status and C0VID-19 severity. The odds ratio estimate was 1.7598, with a 95% confidence interval ranging from 1.0391 to 2.9945. Pearson's Chi-squared Test yielded a statistic of 4.4615 with 1 degree of freedom and a p-value of 0.03467 and support significant association.

Overall, the results from both tests consistently point to a statistically significant association between ACE allelic status and disease severity. The observed p-values falling below the conventional significance level of 0.05 support the conclusion of association.

3.6.1.3. Dominant Model II + ID vs. DD

Dominant model is designed to determine if the I allele with dominant effect is associated with severity of COVID-19. Fisher's Exact Test (p-value 0.25) and Chi-squared Test are performed for dominant model for the ACE I allele and they do not provide evidence for a significant association with disease severity (Table 3.1). The odds ratio of 0.616 indicates a negligible effect, and the 95% confidence interval 0.257 to 1.424 further supports the non-significant effect (Table 3.1).

Similarly, the Chi-squared Test produced a chi-squared value of 1.0756 with 1 degree of freedom yielded a p-value of 0.2997, and further supports the absence of a significant

association. Therefore, based on the statistical analysis the dominant effect of ACE I allele does not appear to exert a considerable influence on the likelihood of severe outcomes in COVID-19.

3.6.1.4. Recessive Model II vs. (DD + ID)

Recessive model is designed to determine whether the Alu insertion in the ACE has recessive effect or not. During statistical analysis, Fisher's Exact Test yielded a p-value of 0.01735 and Chi-square test resulting chi-squared value of 5.0434 with 1 degree of freedom yielded a p-value of 0.02472, and thus revealed significant evidence for recessive effect of ACE I allele (Table 3.1). The Alu deletion is more frequent (67/78=85.8%) in individuals in severe group (either homozygous or heterozygous), indicating deletion of Alu has significant effect. Thus, individuals with I allele exhibit a reduced likelihood to develop severe disease, indicating the potential role of ACE I/D polymorphism in influencing disease outcome.

3.6.2. TMPRSS2 (rs3787950)

Similar analysis of TMPRSS2 variant (rs3787950) genotype association with COVID-19 disease severity, with Fisher's Exact Test produced a p-value of 0.1744, surpassing the significance threshold of 0.05, indicating no significant association between TMPRSS2 genotype and disease severity (Table 3.2). The Chi-Square test is not applicable as frequency of G/G genotype is less than 5 (Table 3.2).

Fisher's Exact Test reported a p-value of 1, suggesting no significant association for alleles and disease severity. The odds ratio estimate was 1.0436, with a 95% confidence interval ranging from 0.4332 to 2.4419 further support it. The Chi-Square Test statistic 0.13939 with 1 degree of freedom and a p-value of 1, signifying no significant association. Both tests showed lack of evidence for a statistically significant association between TMPRSS2 allelic status and disease severity (Table 3.2).

3.6.3. HLA -A*11

The Fisher's Exact Test for the HLA-A*11 allele association with disease severity yielded a p-value of 0.154, odds ratio is 0.578194, with a 95% confidence interval from 0.2649773 to 1.2411943. The p-value is above the 0.05 significance threshold, indicating no effect of HLA-A*11 allele on disease severity. Pearson's Chi-squared Test

resulted in a p-value of 0.1774, also exceeding 0.05, leading to the same conclusion (Table 3.2).

In summary, the combined results from Fisher's Exact Test and Pearson's Chi-squared Test, there is no compelling evidence supporting a significant relationship between the presence of the HLA-A*11 allele and disease severity. The p-values exceeding the conventional significance level of 0.05 indicate a lack of statistical significance, and the odds ratio estimate with its confidence interval reinforces the absence of a association between the HLA A11 allele and disease severity in cohort of 131 samples included in this study.

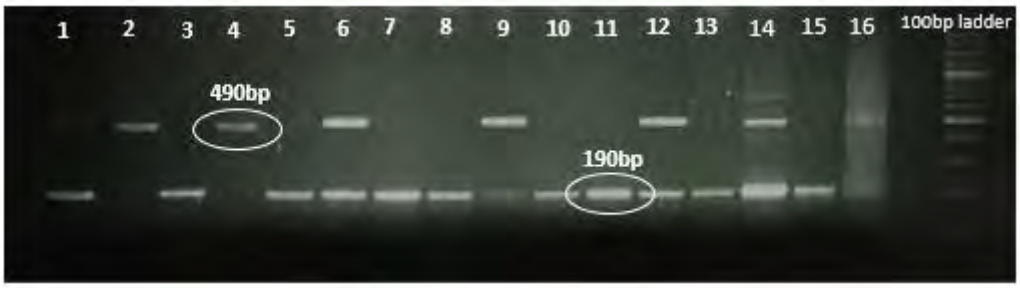


Figure 3.1. 2% Agrose gel representing ACE allele. Lane 11 with 190bp band represents the homozygous deletion of Alu repeat, but lane 4 with 490bp band represent homozygous insertion of Alu repeat and lane 6 with both bands represents heterozygous status of respective individual.

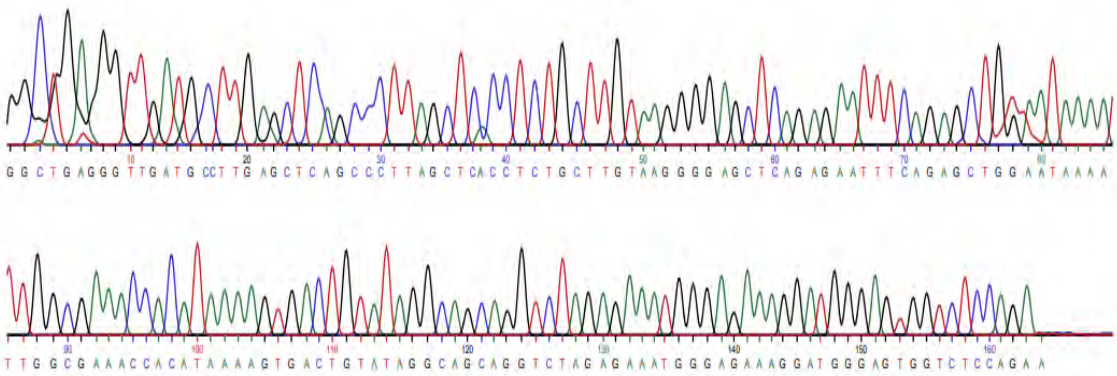


Figure 3.2. Sanger sequencing chromatogram for ACE allele without Alu repeat. This sequence was obtained from an individual homozygous for ACE Alu deletion and has a PCR product size of 190bp.

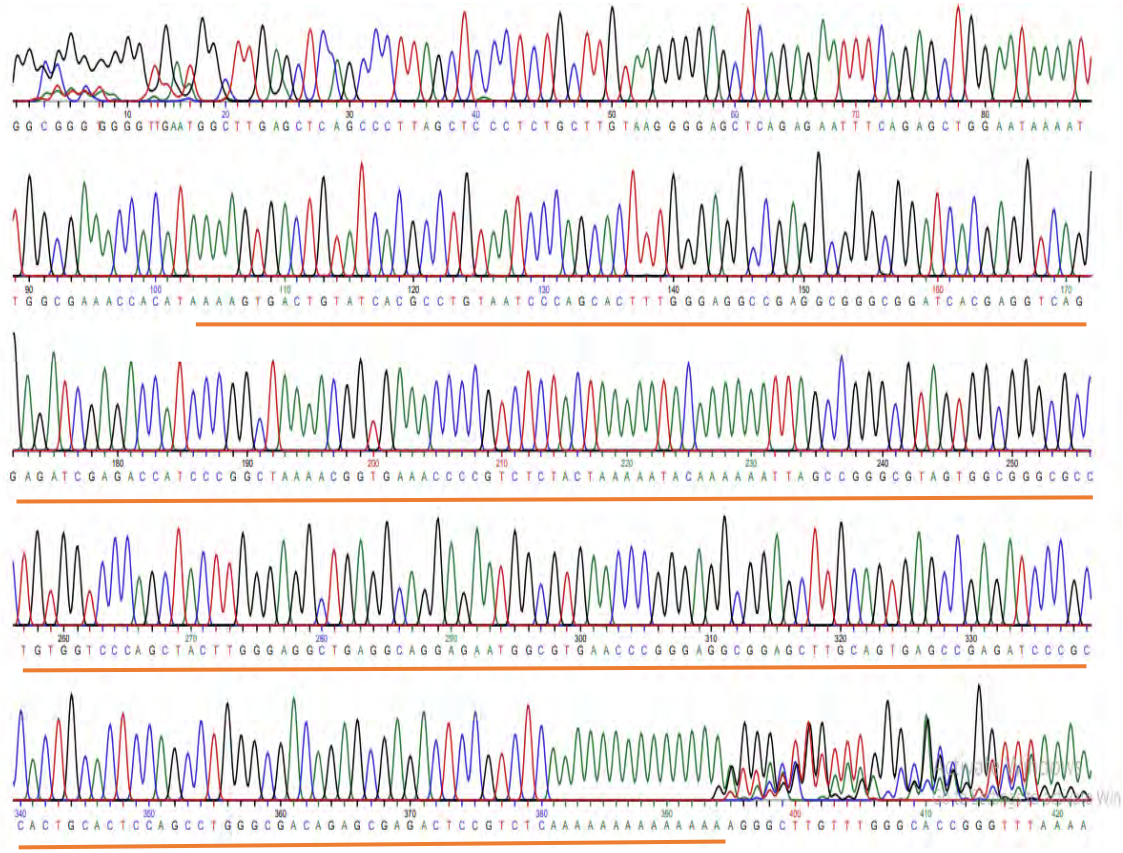


Figure 3.3. Sanger sequencing chromatogram for ACE allele with Alu repeat. This sequence was obtained from an individual homozygous for ACE Alu insertion. Underline sequence represents insertion and has a PCR product size of 490bp.

Sequence of allele ACE with deletion of Alu repeat element

TGGAGAGCCACTCCCATCCTTTCTCCCATTTCTCTAGACCTGCTGCCATACAGTCACT
TTTATGTGGTTTCGCCAATTTTATTCCAGCTCTGAAATTCTCTGAGCTCCCCTTACAA
GCAGAGGTGAGCTAAGGGCTGGAGCTCAAGGCATTCAAACCCCTACCAGATCTGAC
GAATGTGATGGCCACGTC

Sequence of ACE allele with insertion of Alu repeat element

GGCGGGTGGGGTTGAATGGCTTGAGCTCAGCCCTTAGCTCCCTCTGCTTGTAAGGGG
AGCTCAGAGAATTTCAGAGCTGGAATAAAATTGGCGAAACCACAT{AAAAGTGACT
GTATCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGCGGATCACGAGGT
CAGGAGATCGAGACCATCCCGGCTAAAACGGTGAAACCCCGTCTCTACTAAAAAT
ACAAAAAATTAGCCGGGCGTAGTGGCGGGCGCCTGTGGTCCCAGCTACTTGGGAGG
CTGAGGCAGGAGAATGGCGTGAACCCGGGAGGCGGAGCTTGCAGTGAGCCGAGAT
CCCGCCACTGCACTCCAGCCTGGGCGACAGAGCGAGACTCCGTCTCAAAAAAAAAA
A}AAAAGGGCTTGTTTGGGCACCGGGTTTAAAAAAGGGGAAAANGGGGGGGAGGG
GCCCCCCCAAAACCTC

Structural features of Alu repeat elements

TAAAAATACAAA	adenosine rich connector
GGATCACGAGGTCA	Alu-Retinoic Acid Receptor response element (RARE) sequence
AGACCATCCCGGCTA	ERE-like sequence (red) identified in the 5' monomer of the Alu element
AGCTTGCAGTGAGCC	ERE-like sequence (red) identified in the 3' monomer of the element

Figure 3.4. ACE alleles sequence with and without Alu repeat. The characteristic features of Alu repeat are also shown.

Figure.3.5. HLA-A*11 allele (552bp) and Control (APC, 251bp) on 2% agarose gel

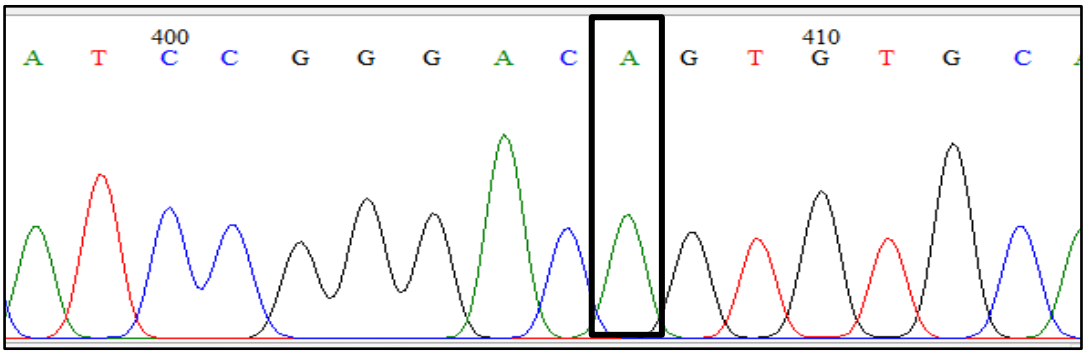


Figure 3.6. TMPRSS2 variant rs3787950. No A to G substitution (c.225A>G) was detected in this individual and have homozygous A/A genotype

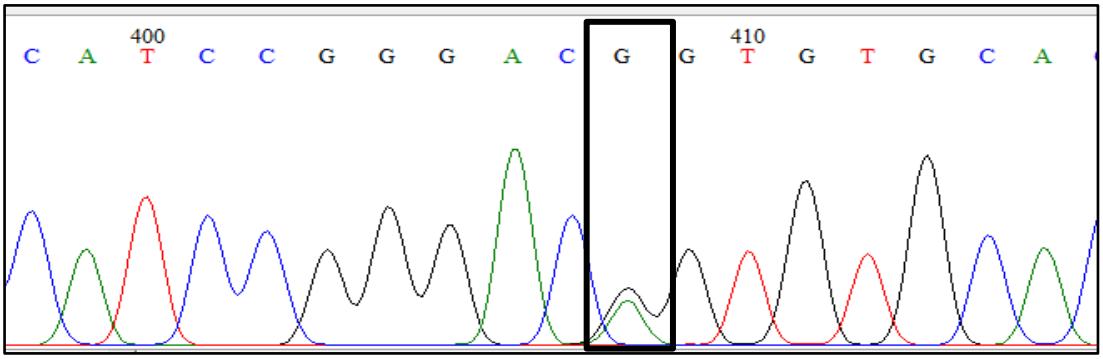


Figure 3.7. TMPRSS2 variant rs3787950. The representative individual has A to G substitution (c.225A>G) and is heterozygous A/G.

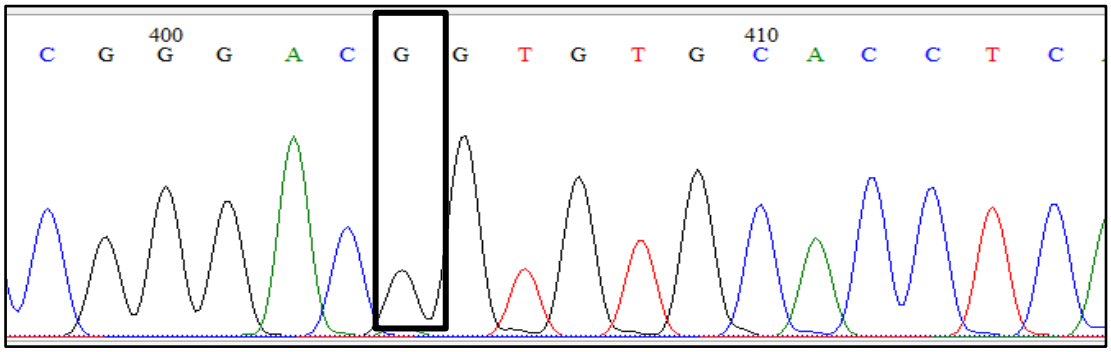


Figure 3.8. TMPRSS2 variant (rs3787950). The representative individual has A to G substitution (c.225A>G), and is homozygous G/G

Table.3.1. Statistical analysis of ACE I/D Polymorphism (rs4646994)

	Severe N=78	Mild N=53	Fisher' Exact Test			Chi Square Test		
			p-value	Odd Ratio	95% CI	X-square	df	p-value
ACE Genotypes								
D/D	27	13						
I/D	40	23	**0.04676**	-	-	0.978	2	**0.0444**
I/I	11	17						
ACE Alleles								
D	94	49	**0.03153**	1.759811	1.039107 - 2.994487	4.4615	1	**0.03467**
I	62	57						
II + ID vs. DD Dominant Model								
I/I +I/D	51	40	0.25	0.616154	0.2570157-1.4236068	1.0756	1	0.2997
D/D	27	13						
II vs. (DD + ID) Recessive Model								
I/I	11	17	**0.01735**	0.350662	0.1327574-0.8911837	5.0434	1	**0.02472**
D/D+I/D	67	36						

Table 3.2 Statistical Analysis of TMPRSS2 and HLA-A*11

	Severe N=78	Mild N=53	Fisher' Exact Test			Chi Square Test		
			p-value	Odd Ratio	95% CI	X-square	df	p-value
TMPRSS2 (rs3787950) Genotypes								
A/A	61	41	0.2889	-	-	Not Applicable		
A/G	17	8						
G/G	0	2						
TMPRSS2 (rs3787950) Alleles								
A	139	94	0.566	1.24318	0.5387176- 2.8282974	0.13939	1	0.7089
G	17	12						
HLA-A*02								
Present	38	33	0.154	0.57819	0.2649773-1.2411943	1.819	1	0.1774
Absent	33	20						

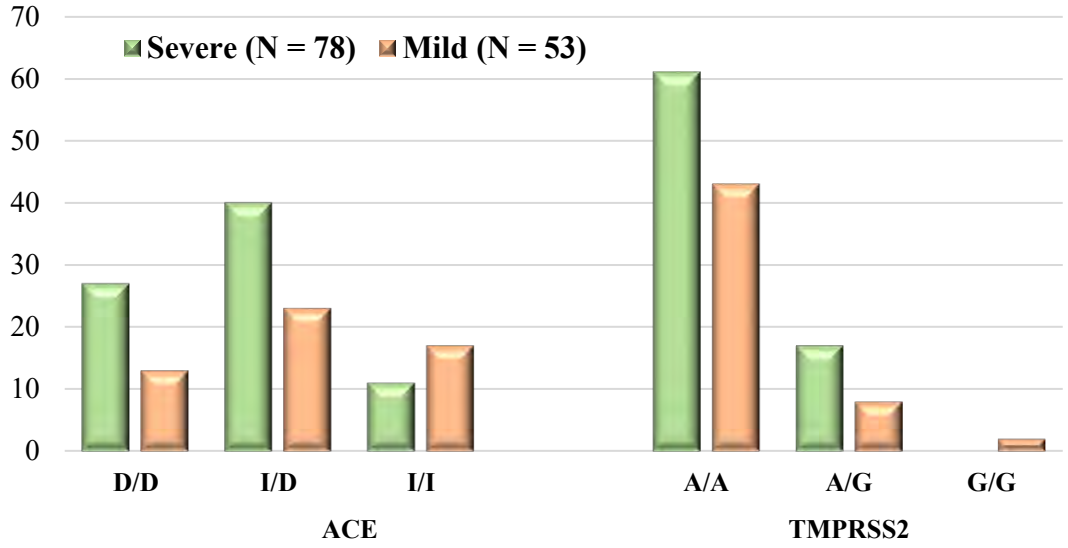


Figure 3.9. Graphic representation of ACE and TMPRSS2 genotype data in Severe and Mild groups

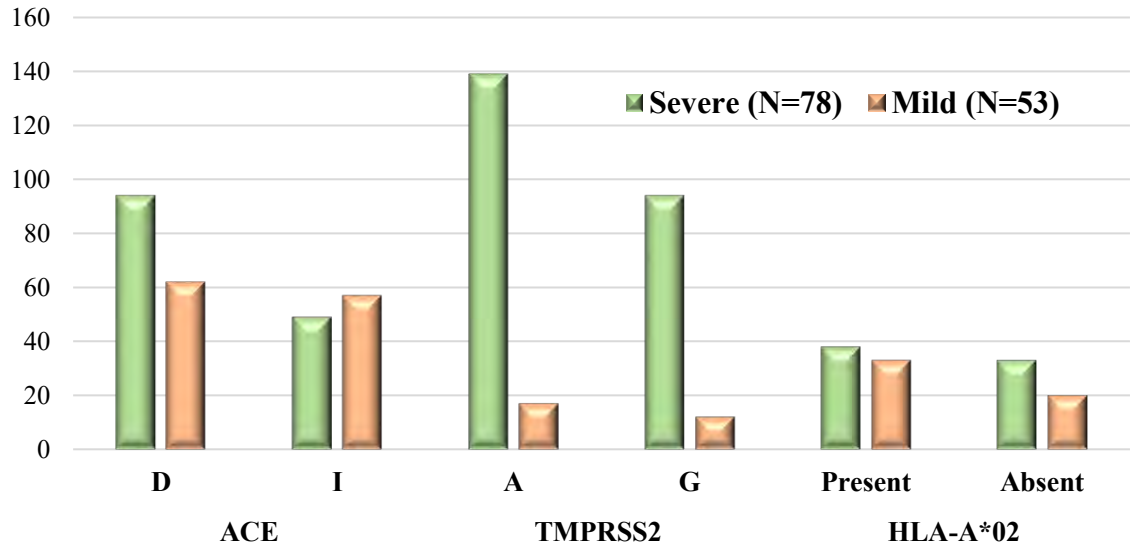


Figure 3.10. Graphic representation of allele frequencies of ACE, TMPRSS2 and HLA-A*1

Chapter 4

Discussion

4. Discussion

Coronavirus Disease 2019 (COVID-19) is a recently emerged infectious disease that is caused by a novel corona virus named as Severe Acute Respiratory Syndrome Coronavirus 2 (Wang et al., 2020). This virus was first discovered in China in December 2019 and soon after its identification, it spread globally and caused pandemic. The virus shows variations in infection patterns among individuals resulting in different disease phenotypes as some individuals are highly susceptibility to the virus, while others may remain resistant or asymptomatic. The dynamic nature of COVID-19 indicates the importance of genetic and environmental risk factors and intricate interplay of both factors results in different outcomes of COVID-19 related clinical presentation. There are several genetic risk factors and their associated genetic variations which play an important role in the severity of disease and host immune response against infection. The virus entry into human cells depends on ACE2, DPP4, GRP78, AXL, TMPRSS2, FURIN, NRP1 and Cathepsin L (Deng et al., 2021). TLR3, IRF7, IRF9, TICAM1/TRIF, UNC93, TRAF3, TBK1, IRF3, NEMO/IKBKG, IFNAR1, IFNAR2, STAT1, STAT2 and HLA alleles play role in adaptive and innate immunity (Velavan et al., 2021). This study intends to explore the association of TMPPRSS2, HLA-A*11 and ACE polymorphisms with disease severity.

The Renin-Angiotensin-Aldosterone System (RAAS) plays a crucial role in regulating blood pressure and maintaining homeostasis (Fountain et al., 2023). During infection virus engages with ACE2 through its spike to gain entry in host cells and in this way normal function of ACE2 get compromised. As a result, angiotensin-II level is increased which causes inflammation, increased blood pressure due to vasoconstriction and other cardiovascular problems as commonly seen in COVID-19 patients. RAAS homeostasis also get disturbed due to Genetic variations in ACE and ACE2 genes, such as the ACE insertion/deletion (ID) polymorphism (rs4646994) (Gómez et al., 2020). The ACE I/D polymorphism involves the insertion or deletion of approximately 290bp of Alu repeat in intron 16. Alu repeats, a mobile DNA elements, covering 11% of the genome (Deininger 2011). These are capable of self-transcription through the reverse transcription process and randomly get inserted in coding as well as in non-coding region and can potentially affect gene expression.

Insertion/deletion ACE polymorphisms exhibit significant geographical and ethnical variability, where the frequency of I allele is increased in Asian and African populations compared to Caucasians (Hubacek et al., 2021). Several studies have explored the ACE I/D polymorphism and they found that this genetic variation has significant effect on susceptibility and severity of COVID-19. Yamamoto et al., (2020) explored the impact ACE insertion/deletion polymorphism on SARS-CoV-2 infection and found ACE I/I genotype negatively correlation with both the number of SARS-CoV-2 cases and deaths, indicating a potential influence on COVID-19 prevalence and clinical outcomes. Verma et al., (2021) also investigated the association of ACE I/D polymorphism with COVID-19 severity in 269 confirmed cases and he found higher frequency of ACE DD genotype and D allele in severe COVID-19 patients. These results suggest that ACE genotype may influence COVID-19 clinical outcome, potentially serving as a predictive marker for risk and severity. Pati et al., (2020) studied data from various Asian countries and explored the association between ACE gene insertion/deletion (I/D) polymorphism and SARS-CoV-2 infection outcomes and reveals a significant positive correlation of the D allele with both SARS-CoV-2 infection rate and mortality. But another study conducted by Hubacek et al., (2021), contradicted their findings. He investigated the ACE insertion/deletion (I/D) polymorphism in 408 SARS-CoV-2-positive individuals (163 asymptomatic, 245 symptomatic) compared with a control group of 2,559 subjects. The frequency of ACE I/I homozygotes is significantly higher in symptomatic COVID-19 patients, indicating a potential link to increased severity. This suggests that ACE I/D polymorphism may serve as a predictor for the severity of COVID-19, with I/I homozygotes having an elevated risk of symptomatic infection.

Our study also explored the effect of ACE I/D polymorphism in a cohort of 131 individuals, categorized into mild and sever groups based on their clinical status. ACE genotyping results showed that deletion of ALU repeat is more common in individuals who exhibited severe symptoms than individuals with mild symptoms. Statistical analysis results for both ACE genotypic association (Fisher'Excat Test p-value = 0.04676; Chi Square Test p-value = 0.0444) and ACE allelic association (Fisher'Excat Test p-value = 0.03153; Chi Square Test p-value = 0.03467) show significant association between COVID 19 severity and ACE I/D polymorphism. According to these results individuals in which Alu repeat is deleted are prone to severe outcomes of

COVID-19 disease. Our results are consistent with earlier studies of Pati et al., (2020), Verma et al., (2021) and Yamamoto et al., (2020).

In order to explore whether the effect of insertion of ALU repeat element is dominant or recessive, dominant (II + ID vs. DD) and recessive (II vs. DD + ID) models were analyzed. The Dominant model indicates that presence of Alu repeat in single allele is enough for causing severe outcome while recessive models requires the presence of both alleles to cause its effects. During statistical analysis, non-significant results obtained for dominant model (Fisher'Excat Test p-value = 0.25; Chi Square Test p-value = 0.2997) but significant results were obtained for recessive model (Fisher'Excat Test p-value =0.01735; Chi Square Test p-value = 0.02472. These results indicate that individuals homozygous for Alu repeat insertion are less likely to experience severe outcome of infection. In COVID-19 patient ACE level may also rise due to deletion of Alu repeat. As these repeats have ability to regulate gene expression due to presence of regulatory element and high abundance of CpGs dinucleotide which are under epigenetic control.

HLA system also plays a pivotal role in the adaptive immune response to SARS-CoV-2 in COVID-19(Augusto et al., 2023). The diverse HLA genetic polymorphisms in class I and class II influence antigen presentation to CD8+ and CD4+ T cells, facilitating a robust and diverse T-cell response against viral proteins. HLA Class I alleles present 9-8 amino acid viral peptides to T- cells. Some alleles of HLA class I are highly efficient in presenting diverse T-cell epitopes and produce strong immunity against viral infection. As SARS-COV-2 show similarity with other coronaviruses both at genome and protein level (Abdelrahman et al., 2020). The individuals that are already exposed to other corona virus are more likely to develop strong immunity against infection as they already develop T-cell mediated immunity upon exposure of SARS-Cov-2.

Several Insilco and observational studies have been conducted to explore the effect of HLA-A*11 and other HLA alleles with severe disease outcome. Nguyen et al., (2020) conducted an extensive Insilco study that covered 145 HLA genotypes and assessed viral peptides binding strength across a diverse range of HLA alleles. HLA-A*11:01 allele binds numerous SARS-CoV-2 peptides (n=750) and experimentally validated with GLMWLSYFV binding. Tomita et al., (2020) found a higher COVID-19 risk associated with HLA-A*02:01, suggesting lower efficacy in presenting SARS-CoV-2

antigens compared to HLA-A*11:01 and HLA-A*24:02. These studies suggest that individuals with HLA-A11:01 genotypes may have a more efficient T-cell-mediated antiviral response and act as protective allele.

On the contrary, studies discussed below proposed that HLA-A*11 has a non-protective effect and is associated with increased severity and mortality of the disease. Chinese population study analyzed 332 COVID-19 patients across different severity levels and links HLA-A*11:01, B*51:01, C*14:02 alleles to increased COVID-19 severity and susceptibility (Warren and Birol, 2020). Warren and Birol, (2021) analyzed 126 patients (100 COVID-19 positive, 26 COVID-negative) and found DPA1*02:02 allele associated with COVID-19 (29%, p=0.0174) but not linked to hospitalization risk. Their study also showed that A*11:01, C04:01 and DQA1*01:02 alleles were associated with COVID-19 severity (p=0.0078, p=0.0087, p=0.0121). Observational study on 3886 healthy controls and 72 COVID-19 patients conducted in Spain (Lorente et al., 2021) also showed that HLA-A*32 being more prevalent in controls and HLA-B*39 and HLA-C*16 in COVID-19 in patients. Japanese study on 137 mild and 53 severe COVID-19 cases identified HLA-A*11:01:01:01 and HLA-C*12:02:02:01-HLA-B*52:01:01:02 as significantly associated with COVID-19 severity using high-resolution sequencing-based typing for eight HLA genes (Khor et al., 2021).

Our study also explored the class I HLA-A*11 allele association with mild and severe group of COVID-19. Genotyping results showed that HLA-A*11 is frequently present in 131 cohort of COVID-19 positive individuals in 48% individuals in severe group and in 62% individuals in mild group. But statistical results (Fisher'Excat Test p-value = 0.154; Chi square Test p-value = 0.1774) showed non-significant effect of HLA-A*11 with disease severity.

TMPRSS2 is another important host genetic risk factor that may play pivotal role in SARS-COV-2 pathogenicity (de Andrade et al., 2022). Virus use ACE2 receptor in synergy with TMPRSS2 to gain entry in host cell. (Bestle et al., 2020). Genetic variation in TMPRSS2 may play an important role in variable outcomes of COVID-19 in individuals belonging to different ethnicities.

Sekiya et al., (2022) found common missense variant rs12329760 as a molecular detriment of COVID-19 susceptibility in East Asian population as this variant replace normal valine with methionine at 160 of TMPRSS2 protein. Contrary to findings in

European populations study, in Iranian COVID-19 patients found a significant association between the TMPRSS2 gene variant rs12329760 (C to T), specifically the minor T allele, and increased severity of the disease (Yaghoobi et al., 2023). Polymorphisms, including rs2070788, rs7364083, rs9974589 and rs8134378 in TMPRSS2, play crucial roles in viral protein priming and cell adhesion (Alshahawey et al., 2020). Androgen receptor-regulated TMPRSS2 expression, higher in male prostates, potentially explains sex-related COVID-19 differences, indicating a protective role for females (Asselta et al., 2020).TMPRSS2 gene polymorphisms, including rs12329760 (T allele), are associated with increased risk of COVID-19 susceptibility and severity. Rokni et al., (2022) observed differences between severe and non-severe cases, and specific TMPRSS2 genotypes (rs17854725 A > G, rs12329760 C > T, rs4303795 A > G) and linked to a more invasive disease pattern. The AG genotype of rs17854725 A > G was associated with higher mortality(Rokni et al., 2022).

Our study explores the role of two missense and one synonymous variant in Pakistani population. Sanger sequencing of 131 samples showed variation only for synonymous variants. Minor alleles are present in only 11% of individuals. A/A genotype is more common than A/G and G/G. Statistical results for genotypic association (Fisher'Excat Test p-value = 0.1744; Chi Square Test p-value = 0.1557) and allelic association (Fisher'Excat Test p-value = 1; Chi Square Test p-value = 1) show no significant association as in both cases p-values are non-significant.

5. Conclusion

This study concludes that ACE functional insertion/deletion polymorphism has a significant impact on COVID-19 severity and individuals having deletion of Alu repeat in ACE gene are at higher risk of developing severe outcomes of COVID-19 disease than individual with Alu insertion. TMPRSS2 rs3787950 variant is detected in 11% individuals, but this variant is also not significantly associated with disease severity. HLA-A*11 distribution is not different among mild and severe group and is also not associated with COVID-19 severity in our cohort. In future, expanding genetic investigations will uncover additional variants and risk factors that may contribute to COVID-19 severity. A broader exploration of the genetic landscape, coupled with larger sample sizes, can provide a more comprehensive understanding of the interplay between host genetics and COVID-19 outcomes.

6. References

Abdelrahman, Z., LI, M. & Wang, X. 2020. Comparative review of SARS-CoV-2, SARS-CoV, MERS-CoV, and influenza a respiratory viruses. Frontiers in immunology, 2309.

Alimoradi, N., Sharqi, M., Firouzabadi, D., Sadeghi, M. M., Moezzi, M. I. & Firouzabadi, N. 2022. SNPs of ACE1 (rs4343) and ACE2 (rs2285666) genes are linked to SARS-CoV-2 infection but not with the severity of disease. Virology Journal, 19, 48.

Alnaqbi, H., Tay, G. K., Jelinek, H. F., Francis, A., Alefishat, E., Chehadeh, S. E. H., Saeed, A. T., Hussein, M., Salameh, L. & Mahboub, B. H. 2022. HLA repertoire of 115 UAE nationals infected with SARS-CoV-2. Human Immunology, 83, 1-9.

Alshahawey, M., Raslan, M. & Sabri, N. 2020. Sex-mediated effects of ACE2 and TMPRSS2 on the incidence and severity of COVID-19; the need for genetic implementation. Current Research in Translational Medicine, 68, 149.

Amoroso, A., Magistroni, P., Vespasiano, F., Bella, A., Bellino, S., Puoti, F., Alizzi, S., Vaisitti, T., Boros, S. & Grossi, P. A. 2021. HLA and AB0 polymorphisms may influence SARS-CoV-2 infection and COVID-19 severity. Transplantation, 105, 193-200.

Anzurez, A., Naka, I., Miki, S., Nakayama-Hosoya, K., Isshiki, M., Watanabe, Y., Nakamura-Hoshi, M., Seki, S., Matsumura, T. & Takano, T. 2021. Association of HLA-DRB1* 09: 01 with severe COVID-19. Hla, 98, 37-42.

Asselta, R., Paraboschi, E. M., Mantovani, A. & Duga, S. 2020. ACE2 and TMPRSS2 variants and expression as candidates to sex and country differences in COVID-19 severity in Italy. Aging (albany NY), 12, 10087.

Atiku, S. M., Kasozi, D. & Campbell, K. 2023. Single Nucleotide Variants (SNVs) of Angiotensin-Converting Enzymes (ACE1 and ACE2): A Plausible Explanation for the Global Variation in COVID-19 Prevalence. Journal of the Renin-Angiotensin-Aldosterone System, 2023.

Augusto, D. G., Murdolo, L. D., Chatzileontiadou, D. S., Sabatino JR, J. J., Yusufali, T., Peyser, N. D., Butcher, X., Kizer, K., Guthrie, K. & Murray, V. W. 2023. A common allele of HLA is associated with asymptomatic SARS-CoV-2 infection. Nature, 620, 128-136.

Bestle, D., Heindl, M. R., Limburg, H., Pilgram, O., Moulton, H., Stein, D. A., Hardes, K., Eickmann, M., Dolnik, O. & Rohde, C. 2020. TMPRSS2 and furin are both essential for proteolytic activation of SARS-CoV-2 in human airway cells. Life science alliance, 3.

Beyerstedt, S., Casaro, E. B. & Rangel, É. B. 2021. COVID-19: angiotensin-converting enzyme 2 (ACE2) expression and tissue susceptibility to SARS-CoV-2 infection. European journal of clinical microbiology & infectious diseases, 40, 905-919.

Bian, J. & Li, Z. 2021. Angiotensin-converting enzyme 2 (ACE2): SARS-CoV-2 receptor and RAS modulator. Acta Pharmaceutica Sinica B, 11, 1-12.

Bunce, M., O'Neill, C., Barnardo, M., Krausa, P., Browning, M., Morris, P. & Welsh, K. 1995. Phototyping: comprehensive DNA typing for HLA-A, B, C, DRB1, DRB3, DRB4, DRB5 & DQB1 by PCR with 144 primer mixes utilizing sequence-specific primers (PCR-SSP). Tissue antigens, 46, 355-367.

Bunyavanich, S., Do, A. & Vicencio, A. 2020. Nasal gene expression of angiotensin-converting enzyme 2 in children and adults. Jama, 323, 2427-2429.

Calabrese, C., Annunziata, A., Coppola, A., Pafundi, P. C., Guarino, S., Di Spirito, V., Maddaloni, V., Pepe, N. & Fiorentino, G. 2021. ACE gene I/D polymorphism and acute pulmonary embolism in COVID19 pneumonia: a potential predisposing role. Frontiers in medicine, 7, 631148.

Calcagnile, M., Forgez, P., Iannelli, A., Bucci, C., Alifano, M. & Alifano, P. 2021. Molecular docking simulation reveals ACE2 polymorphisms that may increase the affinity of ACE2 with the SARS-CoV-2 Spike protein. Biochimie, 180, 143-148.

Camargo, S. M., Singer, D., Makrides, V., Huggel, K., Pos, K. M., Wagner, C. A., Kuba, K., Danilczyk, U., Skovby, F. & Kleta, R. 2009. Tissue-specific amino

acid transporter partners ACE2 and collectrin differentially interact with hartnup mutations. Gastroenterology, 136, 872-882. e3.

Carpenter, R. M., Young, M. K., Petri, W. A., Lyons, G. R., Gilchrist, C., Carey, R. M. & Petri jr, W. A. 2022. Repressed Ang 1–7 in COVID-19 is inversely associated with inflammation and coagulation. Msphere, 7, e00220-22.

Chan, J. F.-W., Kok, K.-H., Zhu, Z., Chu, H., To, K. K.-W., Yuan, S. & Yuen, K.-Y. 2020. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. Emerging microbes & infections, 9, 221-236.

Chen, T., Wu, D., Chen, H., Yan, W., Yang, D., Chen, G., Ma, K., Xu, D., Yu, H. & Wang, H. 2020. Clinical characteristics of 113 deceased patients with coronavirus disease 2019: retrospective study. bmj, 368.

Cheung, K. S., Chan, E. W., Seto, W. K., Wong, I. C. & Leung, W. K. 2020. ACE (angiotensin-converting enzyme) inhibitors/angiotensin receptor blockers are associated with lower colorectal cancer risk: a territory-wide study with propensity score analysis. Hypertension, 76, 968-975.

Creary, L. E., Gangavarapu, S., Caillier, S. J., Cavalcante, P., Frangiamore, R., Lie, B. A., Bengtsson, M., Harbo, H. F., Brauner, S. & Hollenbach, J. A. 2021. Next-generation sequencing identifies extended HLA class I and II haplotypes associated with early-onset and late-onset myasthenia gravis in Italian, Norwegian, and Swedish populations. Frontiers in Immunology, 12, 667336.

De Andrade, C. C., Silva, A. T. P., Vasconcelos, L. R. S., OLiveira, P. R. S., De Souza, C. D. F., Da Costa Armstrong, A. & Do Carmo, R. F. 2022. A Polymorphism in the TMPRSS2 Gene Increases the Risk of Death in Older Patients Hospitalized with COVID-19. Viruses, 14, 2557.

Deininger, P. 2011. Alu elements: know the SINEs. Genome biology, 12, 1-12.

Deng, H., Yan, X. & Yuan, L. 2021. Human genetic basis of coronavirus disease 2019. Signal transduction and targeted therapy, 6, 344.

Devaux, C. A., Rolain, J.-M. & Raoult, D. 2020. ACE2 receptor polymorphism: Susceptibility to SARS-CoV-2, hypertension, multi-organ failure, and COVID-

19 disease outcome. Journal of Microbiology, Immunology and Infection, 53, 425-435.

Dobrijević, Z., Gligorijević, N., Šunderić, M., Penezić, A., Miljuš, G., Tomić, S. & Nedić, O. 2023. The association of human leucocyte antigen (HLA) alleles with COVID-19 severity: A systematic review and meta-analysis. Reviews in Medical Virology, 33, e2378.

Donoghue, M., Hsieh, F., Baronas, E., Godbout, K., Gosselin, M., Stagliano, N., Donovan, M., Woolf, B., Robison, K. & Jeyaseelan, R. 2000. A novel angiotensin-converting enzyme–related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1-9. Circulation research, 87, e1-e9.

Farahani, R. H., Esmaeilzadeh, E., Asl, A. N., Heidari, M. F. & Hazrati, E. 2021. Frequency of HLA Alleles in a Group of Severe COVID-19 Iranian patients. Iranian Journal of Public Health, 50, 1882.

Fountain, J. H., Kaur, J. & Lappin, S. L. 2023. Physiology, renin angiotensin system. StatPearls [Internet]. StatPearls Publishing.

Gómez, J., Albaiceta, G. M., García-Clemente, M., López-Larrea, C., Amado-Rodríguez, L., Lopez-Alonso, I., Hermida, T., Enriquez, A. I., Herrero, P. & Melón, S. 2020. Angiotensin-converting enzymes (ACE, ACE2) gene variants and COVID-19 outcome. Gene, 762, 145102.

Gorbalenya, A. E., Baker, S. C., Baric, R. S., De Groot, R. J., Drosten, C., Gulyaeva, A. A., Haagmans, B. L., Lauber, C., Leontovich, A. M., Neuman, B. W., Penzar, D., Perlman, S., Poon, L. L. M., Samborskiy, D. V., Sidorov, I. A., Sola, I., Ziebuhr, J. & Coronaviridae Study Group Of The International Committee On Taxonomy Of, V. 2020. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nature Microbiology, 5, 536-544.

Gorkhali, R., Koirala, P., Rijal, S., Mainali, A., Baral, A. & Bhattarai, H. K. 2021. Structure and function of major SARS-CoV-2 and SARS-CoV proteins. Bioinformatics and biology insights, 15, 11779322211025876.

Gunal, O., Sezer, O., Ustun, G. U., Ozturk, C. E., Sen, A., Yigit, S. & Demirag, M. D. 2021. Angiotensin-converting enzyme-1 gene insertion/deletion polymorphism

may be associated with COVID-19 clinical severity: a prospective cohort study. Annals of Saudi Medicine, 41, 141-146.

Hamon, R. & Ween, M. P. 2022. E-Cigarette Vapour Increases ACE2 and TMPRSS2 Expression in a Flavour-and Nicotine-Dependent Manner. International Journal of Environmental Research and Public Health, 19, 14955.

Hemming, M. L. & Selkoe, D. J. 2005. Amyloid β-protein is degraded by cellular angiotensin-converting enzyme (ACE) and elevated by an ACE inhibitor. Journal of Biological Chemistry, 280, 37644-37650.

Hoffmann, M., Kleine-Weber, H., Krüger, N., Müller, M., Drosten, C. & Pöhlmann, S. 2020. The novel coronavirus 2019 (2019-nCoV) uses the SARS-coronavirus receptor ACE2 and the cellular protease TMPRSS2 for entry into target cells. BioRxiv, 2020.01. 31.929042.

Hooper, N. M. 1991. Angiotensin converting enzyme: implications from molecular biology for its physiological functions. International journal of biochemistry, 23, 641-647.

Huang, Y., Yang, C., Xu, X.-F., Xu, W. & Liu, S.-W. 2020. Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. Acta Pharmacologica Sinica, 41, 1141-1149.

Hubacek, J. A., Dusek, L., Majek, O., Adamek, V., Cervinkova, T., Dlouha, D. & Adamkova, V. 2021. ACE I/D polymorphism in Czech first-wave SARS-CoV-2-positive survivors. Clinica Chimica Acta, 519, 206-209.

Imai, Y., Kuba, K., Ohto-Nakanishi, T. & Penninger, J. M. 2010. Angiotensin-converting enzyme 2 (ACE2) in disease pathogenesis. Circulation Journal, 74, 405-410.

Iniguez, M., Pérez-Matute, P., Villoslada-Blanco, P., Recio-Fernandez, E., Ezquerro-Pérez, D., Alba, J., Ferreira-Laso, M. L. & Oteo, J. A. 2021. ACE gene variants rise the risk of severe COVID-19 in patients with hypertension, dyslipidemia or diabetes: a spanish pilot study. Frontiers in Endocrinology, 12, 688071.

Irham, L. M., Chou, W.-H., Calkins, M. J., Adikusuma, W., Hsieh, S.-L. & Chang, W.-C. 2020. Genetic variants that influence SARS-CoV-2 receptor TMPRSS2

expression among population cohorts from multiple continents. Biochemical and Biophysical Research Communications, 529, 263-269.

Jami, G., Ataee, M., Esmaeili, V., Chamani, S., Rezaei, A. & Naghizadeh, A. 2023. Characterization of the angiotensin-converting enzyme 2 (ACE2), the main receptor for the SARS-CoV-2 virus. American Journal of Clinical and Experimental Immunology, 12, 24.

Karakaş Çelik, S., Çakmak Genç, G., Pişkin, N., Açikgöz, B., Altinsoy, B., Kurucu İşsiz, B. & Dursun, A. 2021. Polymorphisms of ACE (I/D) and ACE2 receptor gene (Rs2106809, Rs2285666) are not related to the clinical course of COVID-19: A case study. Journal of Medical Virology, 93, 5947-5952.

Khalilzadeh, F., Sakhaee, F., Sotoodehnejadnematalahi, F., Zamani, M. S., Ahmadi, I., Anvari, E. & Fateh, A. 2022. Angiotensin-converting enzyme 2 rs2285666 polymorphism and clinical parameters as the determinants of COVID-19 severity in Iranian population. International Journal of Immunogenetics, 49, 325-332.

Khor, S.-S., Omae, Y., Nishida, N., Sugiyama, M., Kinoshita, N., Suzuki, T., Suzuki, M., Suzuki, S., Izumi, S. & Hojo, M. 2021. HLA-A* 11: 01: 01: 01, HLA-C* 12: 02: 02: 01-HLA-B* 52: 01: 02: 02, age and sex are associated with severity of Japanese COVID-19 with respiratory failure. Frontiers in Immunology, 12, 658570.

Khouri, B. F., De Souza Candido, I. P., Poli-Frederico, R. C. & Bignardi, P. R. 2023. Host Genetics and COVID-19: Genes Underlying the Patterns of Susceptibility and Prognosis. Gene Expression, 22, 222-231.

Kim, T. S., Heinlein, C., Hackman, R. C. & Nelson, P. S. 2006. Phenotypic analysis of mice lacking the Tmprss2-encoded protease. Molecular and cellular biology, 26, 965-975.

Kuba, K., Yamaguchi, T. & Penninger, J. M. 2021. Angiotensin-Converting Enzyme 2 (ACE2) in the Pathogenesis of ARDS in COVID-19. Frontiers in Immunology, 12, 732690.

Lam, D. K., Dang, D., Flynn, A. N., Hardt, M. & Schmidt, B. L. 2015. TMPRSS2, a novel membrane-anchored mediator in cancer pain. Pain, 156, 923.

Lebeau, G., Vagner, D., Frumence, E., Ah-Pine, F., Guillot, X., Nobecourt, E., Raffray, L. & Gasque, P. 2020. Deciphering SARS-CoV-2 virologic and immunologic features. International journal of molecular sciences, 21, 5932.

Littera, R., Campagna, M., Deidda, S., Angioni, G., Cipri, S., Melis, M., Firinu, D., Santus, S., Lai, A. & Porcella, R. 2020. Human leukocyte antigen complex and other immunogenetic and clinical factors influence susceptibility or protection to SARS-CoV-2 infection and severity of the disease course. The Sardinian experience. Frontiers in immunology, 11, 605688.

Liu, M.-Y., Zheng, B., Zhang, Y. & Li, J.-P. 2020a. Role and mechanism of angiotensin-converting enzyme 2 in acute lung injury in coronavirus disease 2019. Chronic Diseases and Translational Medicine, 6, 98-105.

Liu, Y., Qu, H.-Q., Qu, J., Tian, L. & Hakonarson, H. 2020b. Expression pattern of the SARS-CoV-2 entry genes ACE2 and TMPRSS2 in the respiratory tract. Viruses, 12, 1174.

Lopez-Sublet, M., Di Lanzacco, L. C., Danser, A. J., Lambert, M., Elourimi, G. & Persu, A. 2018. Focus on increased serum angiotensin-converting enzyme level: from granulomatous diseases to genetic mutations. Clinical biochemistry, 59, 1-8.

Lorente, L., Martín, M., Franco, A., Barrios, Y., Cáceres, J., Solé-Violán, J., Perez, A., Y Ramos, J. M., Ramos-Gómez, L. & Ojeda, N. 2021. HLA genetic polymorphisms and prognosis of patients with COVID-19. Medicina intensiva, 45, 96-103.

Lubbe, L., Cozier, G. E., Oosthuizen, D., Acharya, K. R. & Sturrock, E. D. 2020. ACE2 and ACE: structure-based insights into mechanism, regulation and receptor recognition by SARS-CoV. Clinical Science, 134, 2851-2871.

Lucas, J. M., Heinlein, C., Kim, T., Hernandez, S. A., Malik, M. S., True, L. D., Morrissey, C., Corey, E., Montgomery, B. & Mostaghel, E. 2014. The androgen-regulated protease TMPRSS2 activates a proteolytic cascade involving components of the tumor microenvironment and promotes prostate cancer metastasis. Cancer discovery, 4, 1310-1325.

M.H., Cheung, E. and Kallen, C.B., 2010. Location analysis for the estrogen

Ma, Y., Li, Q., Chen, J., Liu, S., Liu, S., He, X., Ling, Y., Zheng, J., Corpe, C. & Lu, H. 2022. Angiotensin-Converting Enzyme 2 SNPs as Common Genetic Loci and Optimal Early Identification Genetic Markers for COVID-19. Pathogens, 11, 947.

Maclachlan, R., Kehoe, P. G. & Miners, J. S. 2022. Dysregulation of ACE-1 in normal aging and the early stages of Alzheimer's disease. The Journals of Gerontology: Series A, 77, 1775-1783.

Majchrzak, M. & Poręba, M. 2022. The roles of cellular protease interactions in viral infections and programmed cell death: A lesson learned from the SARS-CoV-2 outbreak and COVID-19 pandemic. Pharmacological Reports, 74, 1149-1165.

Martinez-Gomez, L. E., Herrera-Lopez, B., Martinez-Armenta, C., Ortega-Pena, S., Camacho-Rea, M. D. C., Suarez-Ahedo, C., Vazquez-Cardenas, P., Vargas-Alarcon, G., Rojas-Velasco, G. & Fragoso, J. M. 2022. ACE and ACE2 gene variants are associated with severe outcomes of COVID-19 in men. Frontiers in immunology, 13, 812940.

Mason, C.E., Shu, F.J., Wang, C., Session, R.M., Kallen, R.G., Sidell, N., Yu, T., Liu,

Mihalopoulos, M., Levine, A. C., Marayati, N. F., Chubak, B. M., Archer, M., Badani, K. K., Tewari, A. K., Mohamed, N., Ferrer, F. & Kyprianou, N. 2020. The resilient child: sex-steroid hormones and COVID-19 incidence in pediatric patients. Journal of the Endocrine Society, 4, bvaa106.

Möhlendick, B., Schönfelder, K., Breuckmann, K., Elsner, C., Babel, N., Balfanz, P., Dahl, E., Dreher, M., Fistera, D. & Herbstreit, F. 2021. ACE2 polymorphism and susceptibility for SARS-CoV-2 infection and severity of COVID-19. Pharmacogenetics and genomics.

Mokhtari, T., Hassani, F., Ghaffari, N., Ebrahimi, B., Yarahmadi, A. & Hassanzadeh, G. 2020. COVID-19 and multiorgan failure: A narrative review on potential mechanisms. Journal of molecular histology, 51, 613-628.

Naemi, F. M., Al-Adwani, S., Al-Khatabi, H. & Al-Nazawi, A. 2021. Association between the HLA genotype and the severity of COVID-19 infection among South Asians. Journal of Medical Virology, 93, 4430-4437.

Naqvi, A. A. T., Fatima, K., Mohammad, T., Fatima, U., Singh, I. K., Singh, A., Atif, S. M., HARIPRASAD, G., HASAN, G. M. & HASSAN, M. I. 2020. Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach. Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease, 1866, 165878.

Ni, W., Yang, X., Yang, D., Bao, J., Li, R., Xiao, Y., Hou, C., Wang, H., Liu, J. & Yang, D. 2020. Role of angiotensin-converting enzyme 2 (ACE2) in COVID-19. Critical Care, 24, 1-10.

Oran, D. P. & Topol, E. J. 2021. The proportion of SARS-CoV-2 infections that are asymptomatic: a systematic review. Annals of internal medicine, 174, 655-662.

Paoloni-Giacobino, A., Chen, H., Peitsch, M. C., Rossier, C. & Antonarakis, S. E. 1997. Cloning of the TMPRSS2 gene, which encodes a novel serine protease with transmembrane, LDLRA, and SRCR domains and maps to 21q22. 3. Genomics, 44, 309-320.

Park, Y. 2010. TMPRSS2 (transmembrane protease, serine 2). Atlas of Genetics and Cytogenetics in Oncology and Haematology, 14, 1163-1165.

Pati, A., Mahto, H., Padhi, S. & Panda, A. K. 2020. ACE deletion allele is associated with susceptibility to SARS-CoV-2 infection and mortality rate: An epidemiological study in the Asian population. Clinica Chimica Acta, 510, 455-458.

Pizzato, M., Baraldi, C., Boscato Sopetto, G., Finozzi, D., Gentile, C., Gentile, M. D., Marconi, R., Paladino, D., Raoss, A. & Riedmiller, I. 2022. SARS-CoV-2 and the host cell: A tale of interactions. Frontiers in Virology, 1.

Rajpal, V. R., Sharma, S., Sehgal, D., Singh, A., Kumar, A., Vaishnavi, S., Tiwari, M., Bhalla, H., Goel, S. & Raina, S. N. 2022. A comprehensive account of SARS-CoV-2 genome structure, incurred mutations, lineages and COVID-19 vaccination program. Future Virology, 17, 687-706.

Ranjan, J., Ravindra, A. & Mishra, B. 2021. Gender and genetic factors impacting COVID-19 severity. Journal of Family Medicine and Primary Care, 10, 3956.

Raza, M.A.A., Yan, C., Abbas, H.S.M. and Ullah, A., 2022. COVID-19 pandemic control and administrative issues in Pakistan: How Pakistan mitigated both pandemic and administration issues?. Journal of Public Affairs, 22, p.e2760.

receptor-α reveals binding to diverse ERE sequences and widespread binding

Reinke, L. M., Spiegel, M., Plegge, T., Hartleib, A., Nehlmeier, I., Gierer, S., Hoffmann, M., Hofmann-Winkler, H., Winkler, M. & Pöhlmann, S. 2017. Different residues in the SARS-CoV spike protein determine cleavage and activation by the host cell protease TMPRSS2. PLoS One, 12, e0179177.

Rigat, B., Hubert, C., Alhenc-Gelas, F., Cambien, F., Corvol, P. & Soubrier, F. 1990. An insertion/deletion polymorphism in the angiotensin I-converting enzyme gene accounting for half the variance of serum enzyme levels. The Journal of clinical investigation, 86, 1343-1346.

Riordan, J. F. 2003. Angiotensin-I-converting enzyme and its relatives. Genome biology, 4, 1-5.

Rokni, M., Heidari Nia, M., Sarhadi, M., Mirinejad, S., Sargazi, S., Moudi, M., Saravani, R., Rahdar, S. & Kargar, M. 2022. Association of TMPRSS2 gene polymorphisms with COVID-19 severity and mortality: a case-control study with computational analyses. Applied biochemistry and biotechnology, 194, 3507-3526.

Russo, R., Andolfo, I., Lasorsa, V. A., Iolascon, A. & Capasso, M. 2020. Genetic analysis of the coronavirus SARS-CoV-2 host protease TMPRSS2 in different populations. Frontiers in genetics, 11, 872.

Saad, H., Jabotian, K., Sakr, C., Mahfouz, R., Akl, I. B. & Zgheib, N. K. 2021. The role of angiotensin converting enzyme 1 insertion/deletion genetic polymorphism in the risk and severity of COVID-19 infection. Frontiers in Medicine, 8, 798571.

Sabater Molina, M., Nicolás Rocamora, E., Bendicho, A. I., Vázquez, E. G., ZOrio, E., Rodriguez, F. D., Gil Ortuño, C., Rodríguez, A. I., Sánchez-López, A. J. & Jara Rubio, R. 2022. Polymorphisms in ACE, ACE2, AGTR1 genes and severity of COVID-19 disease. PLoS One, 17, e0263140.

Sahranavard-Pirbazari, P., Khoshghiafeh, A., Kamali, M. J., Esfandiar, H., Bakhtiari, M. & Ahmadifard, M. 2023. A comprehensive review of ACE2, ACE1, TMPRSS2 and IFITM3 gene polymorphisms and their effect on the severity of COVID-19. Advances in Medical Sciences, 68, 450-463.

Sakai, K., Ami, Y., Nakajima, N., Nakajima, K., Kitazawa, M., Anraku, M., takayama, I., Sangsriratanakul, N., Komura, M. & Sato, Y. 2016. TMPRSS2 independency for haemagglutinin cleavage in vivo differentiates influenza B virus from influenza A virus. Scientific reports, 6, 29430.

Schmid, C. W. 1998. Does SINE evolution preclude Alu function? Nucleic acids research, 26, 4541-4550.

Schönfelder, K., Breuckmann, K., Elsner, C., Dittmer, U., Fistera, D., Herbstreit, F., Risse, J., Schmidt, K., Sutharsan, S. & Taube, C. 2021. Transmembrane serine protease 2 polymorphisms and susceptibility to severe acute respiratory syndrome coronavirus type 2 infection: a German case-control study. Frontiers in genetics, 12, 667231.

Sekiya, T., Ogura, Y., Kai, H., Kawaguchi, A., Okawa, S., Hirohama, M., Kuroki, T., Morii, W., Hara, A. & Hiramatsu, Y. 2022. TMPRSS2 gene polymorphism common in East Asians confers decreased COVID-19 susceptibility. Frontiers in Microbiology, 13, 943877.

Shen, L. W., Mao, H. J., Wu, Y. L., Tanaka, Y. & Zhang, W. 2017. TMPRSS2: A potential target for treatment of influenza virus and coronavirus infections. Biochimie, 142, 1-10.

Shirato, K., Kawase, M. & Matsuyama, S. 2013. Middle East respiratory syndrome coronavirus infection mediated by the transmembrane serine protease TMPRSS2. Journal of virology, 87, 12552-12561.

Shirbhate, E., Pandey, J., Patel, V. K., Kamal, M., Jawaid, T., Gorain, B., Kesharwani, P. & Rajak, H. 2021. Understanding the role of ACE-2 receptor in pathogenesis of COVID-19 disease: a potential approach for therapeutic intervention. Pharmacological Reports, 1-12.

Sidhwani, S. K., Mirza, T., Khatoon, A., Shaikh, F., Khan, R., Shaikh, O. A. & Nashwan, A. J. 2023. Angiotensin-converting enzyme 2 (ACE2)

polymorphisms and susceptibility of severe SARS-CoV-2 in a subset of Pakistani population. Virology Journal, 20, 1-12.

Siu, Y., Teoh, K., Lo, J., Chan, C., Kien, F., Escriou, N., Tsao, S., Nicholls, J., altmeyer, R. & Peiris, J. 2008. The M, E, and N structural proteins of the severe acute respiratory syndrome coronavirus are required for efficient assembly, trafficking, and release of virus-like particles. Journal of virology, 82, 11318-11330.

Skidgel, R. A., Engelbrecht, S., Johnson, A. R. & Erdös, E. G. 1984. Hydrolysis of substance P and neurotensin by converting enzyme and neutral endopeptidase. Peptides, 5, 769-776.

Soubrier, F., Alhenc-Gelas, F., Hubert, C., Allegrini, J., John, M., Tregear, G. & Corvol, P. 1988. Two putative active centers in human angiotensin I-converting enzyme revealed by molecular cloning. Proceedings of the National Academy of Sciences, 85, 9386-9390.

Suryamohan, K., Diwanji, D., Stawiski, E. W., Gupta, R., Miersch, S., Liu, J., Chen, C., Jiang, Y.-P., Fellouse, F. A. & Sathirapongsasuti, J. F. 2021. Human ACE2 receptor polymorphisms and altered susceptibility to SARS-CoV-2. Communications biology, 4, 475.

Tay, G. K., Alnaqbi, H., Chehadeh, S., Peramo, B., Mustafa, F., Rizvi, T. A., Mahboub, B. H., Uddin, M., Alkaabi, N. & Alefishat, E. 2023. HLA class I associations with the severity of COVID-19 disease in the United Arab Emirates. Plos one, 18, e0285712.

Tipnis, S. R., Hooper, N. M., Hyde, R., Karran, E., Christie, G. & Turner, A. J. 2000. A human homolog of angiotensin-converting enzyme: cloning and functional expression as a captopril-insensitive carboxypeptidase. Journal of Biological Chemistry, 275, 33238-33243.

Tiret, L., Nicaud, V., Kee, F., Evans, A., Cambou, J., Arveiler, D., Luc, G., Amouyel, P., Poirier, O. & Lecerf, L. 1993. Deletion polymorphism in angiotensin-converting enzyme gene associated with parental history of myocardial infarction. The Lancet, 341, 991-992.

Tomita, Y., Ikeda, T., Sato, R. & Sakagami, T. 2020. Association between HLA gene polymorphisms and mortality of COVID-19: An in silico analysis. Immunity, inflammation and disease, 8, 684-694.

Torre-Fuentes, L., Matías-Guiu, J., Hernández-Lorenzo, L., Montero-Escribano, P., Pytel, V., Porta-Etessam, J., Gómez-Pinedo, U. & Matías-Guiu, J. A. 2021. ACE2, TMPRSS2, and Furin variants and SARS-CoV-2 infection in Madrid, Spain. Journal of medical virology, 93, 863-869.

Uversky, V., Kretsinger, R. & Permyakov, E. 2013. Encyclopedia of metalloproteins. Springer: New York, NY, USA.

Velavan, T. P., Pallerla, S. R., Rüter, J., Augustin, Y., Kremsner, P. G., Krishna, S. & Meyer, C. G. 2021. Host genetic factors determining COVID-19 susceptibility and severity. EBioMedicine, 72.

Verma, S., Abbas, M., Verma, S., Khan, F. H., Raza, S. T., Siddiqi, Z., Ahmad, I. & Mahdi, F. 2021. Impact of I/D polymorphism of angiotensin-converting enzyme 1 (ACE1) gene on the severity of COVID-19 patients. Infection, Genetics and Evolution, 91, 104801.

Vishnubhotla, R., Sasikala, M., Ketavarapu, V. & Reddy, D. N. 2021. High-resolution HLA genotyping identifies alleles associated with severe COVID-19: A preliminary study from India. Immunity, Inflammation and Disease, 9, 1781-1785.

Wang, F., Huang, S., Gao, R., Zhou, Y., Lai, C., Li, Z., Xian, W., Qian, X., Li, Z. & Huang, Y. 2020. Initial whole-genome sequencing and analysis of the host genetic contribution to COVID-19 severity and susceptibility. Cell discovery, 6, 83.

Warren, R. L. & Birol, I. 2020. Retrospective in silico HLA predictions from COVID-19 patients reveal alleles associated with disease prognosis. MedRxiv.

Warren, R. L. & Birol, I. 2021. HLA alleles measured from COVID-19 patient transcriptomes reveal associations with disease prognosis in a New York cohort. PeerJ, 9, e12368.

Waryah, Y. M., Khidri, F. F., Nigar, R., Devrajani, T., Rajput, A. R., Waryah, A. M. & Ujjan, I. D. 2023. Impact of ACE2 gene variations on COVID-19 pathogenicity in Pakistani patients. Saudi Journal of Biological Sciences, 30, 103813.

Weiner, J., Suwalski, P., Holtgrewe, M., Rakitko, A., Thibeault, C., Mueller, M., Patriki, D., Quedenau, C., Krueger, U. & Ilinsky, V. 2021. Increased risk of severe clinical course of COVID-19 in carriers of HLA-C* 04: 01. EClinicalMedicine, 40.

Wettstein, L., Kirchhoff, F. & Münch, J. 2022. The transmembrane protease TMPRSS2 as a therapeutic target for COVID-19 treatment. International Journal of Molecular Sciences, 23, 1351.

Wilson, S., Greer, B., Hooper, J., Zijlstra, A., Walker, B., Quigley, J. & Hawthorne, S. 2005. The membrane-anchored serine protease, TMPRSS2, activates PAR-2 in prostate cancer cells. Biochemical Journal, 388, 967-972.

within repetitive DNA elements. Nucleic acids research, 38(7), pp.2355-2368.

Wu, A., Peng, Y., Huang, B., Ding, X., Wang, X., Niu, P., Meng, J., Zhu, Z., Zhang, Z. & Wang, J. 2020. Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. Cell host & microbe, 27, 325-328.

Wu, Z. & Mcgoogan, J. M. 2020. 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, 1239-1242.

Yaghoobi, A., Lord, J. S., Rezaiezadeh, J. S., Yekaninejad, M. S., Amini, M. & Izadi, P. 2023. TMPRSS2 polymorphism (rs12329760) and the severity of the COVID-19 in Iranian population. PLoS One, 18, e0281750.

Yamamoto, N., Ariumi, Y., Nishida, N., Yamamoto, R., Bauer, G., Gojobori, T., Shimotohno, K. & Mizokami, M. 2020. SARS-CoV-2 infections and COVID-19 mortalities strongly correlate with ACE1 I/D genotype. Gene, 758, 144944.

Yung, Y. L., cheng, C. K., Chan, H. Y., Xia, J. T., Lau, K. M., Wong, R. S., Wu, A. K., Chu, R. W., Wong, A. C. & Chow, E. Y. 2021. Association of HLA-B22

serotype with SARS-CoV-2 susceptibility in Hong Kong Chinese patients. Hla, 97, 127-132.

Zaim, S., Chong, J. H., Sankaranarayanan, V. & Harky, A. 2020. COVID-19 and multiorgan response. Current problems in cardiology, 45, 100618.

Zhang, Q., Xiang, R., Huo, S., Zhou, Y., Jiang, S., Wang, Q. & Yu, F. 2021. Molecular mechanism of interaction between SARS-CoV-2 and host cells and interventional therapy. Signal transduction and targeted therapy, 6, 233.

Zheng, H. & Cao, J. J. 2020a. ACE gene polymorphism and severe lung injury in patients with COVID-19.

Zheng, H. & Cao, J. J. 2020b. Angiotensin-converting enzyme gene polymorphism and severe lung injury in patients with coronavirus disease 2019. The American journal of pathology, 190, 2013-2017.

7. Electronic References

BLAT (https://genome.ucsc.edu/cgibin/hgBlat)

Ensembl Genome Browser (https://asia.ensemblorg/index.html)

In silico PCR (https://genome.ucsc.edu/cgi-bin/hgPcr).

IPD-IMGT/HLA (https://www.ebi.ac.uk/ipd/imgt/hla/)

NCBI Primer Blast (https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi)

OlicoCalc (http://biotools.nubic.northwestern.edu/OligoCalc.html)

Primer 3 Software (https://primer3.ut.ee/)

Probe and Primer Search Tool (https://www.ebi.ac.uk/ipd/imgt/hla/probe/)

The Sequence Alignment Tool (https://www.ebi.ac.uk/ipd/imgt/hla/alignment/).

WHO (data.who.int/dashboards/covid19/cases?n=c)

PROFESSOR
Department of Biochemistry
Quaid-i-Azam University, Islamabad

Analysis of Genetics Risk Factors Associated with COVID-19 Severity

ORIGINALITY REPORT

12%	8%	8%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to University of Sydney Student Paper	1%
2	"Virus Infected Cells", Springer Science and Business Media LLC, 2023 Publication	<1%
3	Parisa Sahranavard-Pirbazari, Azin Khoshghiafeh, Mohammad Javad Kamali, Hanieh Esfandiar et al. "A comprehensive review of ACE2, ACE1, TMPRSS2 and IFITM3 gene polymorphisms and their effect on the severity of COVID-19", Advances in Medical Sciences, 2023 Publication	<1%
4	etheses.bham.ac.uk Internet Source	<1%
5	static.frontiersin.org Internet Source	<1%
6	Abd Rahman, Amirah Assyiqqin. "Determination of Pathogenicity Novel	<1%