UNIVERSITI PUTRA MALAYSIA

A SYSTEMATIC REVIEW ON ASSOCIATION OF ACE2 AND/OR TMPRSS2 POLYMORPHISM WITH SUSCEPTIBILITY TO SARS-COV-2 INFECTIONS

NURUL HUSNA BINTI HASNUL HADI

A SYSTEMATIC REVIEW ON ASSOCIATION OF ACE2 AND/OR TMPRSS2 POLYMORPHISM WITH SUSCEPTIBILITY TO SARS-COV-2 INFECTIONS

NURUL HUSNA BINTI HASNUL HADI

A PROJECT PAPER SUBMITTED AS PARTIAL REQUIREMENT FOR THE DEGREE OF BACHELOR OF SCIENCE (BIOMEDICAL SCIENCES)

DEPARTMENT OF BIOMEDICAL SCIENCES
FACULTY OF MEDICINE AND HEALTH SCIENCES
UNIVERSITI PUTRA MALAYSIA

2021

ABSTRACT

A Systematic Review on Association of ACE2 and/or TMPRSS2 Polymorphism with Susceptibility to SARS-CoV-2 Infections

Nurul Husna Hasnul Hadi[a], Aina Nadheera Abd Rahman[a], Syahril Abdullah[a,b]

[a] Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, Universiti PutraMalaysia

[b] UPM-MAKNA Cancer Research Laboratory, Institute of Bioscience, Universiti Putra Malaysia

Introduction: Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the virus that causes coronavirus disease 2019 (COVID-19). It infects the human cells through binding of subunit 1 of the coronavirus spike (S) protein with the angiotensin-converting enzyme 2 (ACE2) receptor and is activated by transmembrane serine protease 2 (TMPRSS2) through cleavage of the S protein at subunit 2 for fusion with the human cell membrane. Polymorphism is a condition where there are variations in the DNA sequences of an individual which determine the diversity of individuals, groups and populations. These polymorphisms could have an effect on the susceptibility and severity of a disease in individuals. Although there are numerous studies that state the importance of ACE2 and TMPRSS2 in SARS-CoV-2 infection, studies that relate polymorphism of both transmembrane proteins to susceptibility to SARS-CoV-2 infections are still lacking. **Objective:** This systematic study aims to investigate the susceptibility to SARS-CoV-2 infections in individuals with ACE2 and/or TMPRSS2 polymorphisms. **Methodology:** Studies related to ACE2, TMPRSS2 and COVID-19 identified from four databases - Scopus, ScienceDirect, PubMed and EBSCOhost that fit the pre-defined criteria underwent initial and full-text screening. A total of seven studies were then subjected to systematic review. **Results:** Systematic review determining the association of ACE2 and/or TMPRSS2 polymorphism with susceptibility to SARS-CoV-2 infections revealed ambiguous results in the studied population. There was no association observed between the ACE2 polymorphism and susceptibility to SARS-CoV-2 infections. Meanwhile, some of TMPRSS2 variants such as rs7282236, rs75603675 and rs61735789 were associated with COVID-19 test positivity and some other TMPRSS2 variants such as rs12329760, rs200291871, rs61735792 and rs61735794 might be linked with SARS-CoV-2 infections. **Discussion:** There are several factors such as the type and the location of the polymorphism which were found to affect the susceptibility to SARS-CoV-2 infections. Some of the polymorphisms found in the studies were single nucleotide polymorphism (SNP), silent or synonymous polymorphism, missense, intronic and mixture of insertion and deletion (indels). ACE2 polymorphisms were commonly found in the control group compared to the patients. It is determined that ACE2 polymorphisms of all types and locations were not found to affect the

susceptibility to infections but it may affect the severity of diseases in the patients. As for TMPRSS2 polymorphisms, it was frequently found in the COVID-19 patients and most of the variants found were linked to SARS-CoV-2 infections regardless of the types of polymorphism. **Conclusion:** This systematic review reveals mixed findings on the association of ACE2 and/or TMPRSS2 polymorphism with the susceptibility to SARS-CoV-2 infection. Generally, the findings are mostly from small populations and therefore not sufficient to conclude the association of ACE2 and/or TMPRSS2 polymorphism with susceptibility to SARS-CoV-2 infection. There is a need for these studies to be done in a larger population to better understand the possible association of both proteins to SARS-CoV-2 infections.

Keywords: ACE2, TMPRSS2, polymorphism, COVID-19 susceptibility, SARS-CoV-2 infection

ABSTRAK

Kajian Sistematik Berhubung Polimorphisme ACE2 dan/atau TMPRSS2 dengan Kerentanan terhadap Jangkitan SARS-CoV-2

Nurul Husna Hasnul Hadi[a], Aina Nadheera Abd Rahman[a], Syahril Abdullah[a,b]

[a] Jabatan Sains Bioperubatan, Fakulti Perubatan dan Sains Kesihatan, Universiti Putra Malaysia

[b] Makmal Penyelidikan Kanser UPM-MAKNA, Institut Biosains, Universiti Putra Malaysia

Pendahuluan: Sindrom pernafasan akut yang teruk coronavirus 2 (SARS-CoV-2) adalah virus yang menyebabkan penyakit coronavirus 2019 (COVID-19). Ia menjangkiti sel-sel manusia melalui pengikatan subunit 1 pepaku protein koronavirus (S) dengan reseptor enzim penukar angiotensin 2 (ACE2) dan diaktifkan oleh transmembran serine protease 2 (TMPRSS2) melalui pembelahan protein S pada subunit 2 untuk peleburan dengan membran sel manusia. Polimorfisme adalah keadaan di mana terdapat variasi dalam urutan DNA seseorang yang menentukan kepelbagaian individu, kumpulan dan populasi. Polimorfisme ini boleh memberi kesan terhadap kerentanan dan keparahan penyakit pada individu. Walaupun terdapat banyak kajian yang menyatakan pentingnya ACE2 dan TMPRSS2 dalam jangkitan SARS-CoV-2, kajian yang mengaitkan polimorfisme kedua-dua protein transmembran terhadap kerentanan terhadap jangkitan SARS-CoV-2 masih kurang. **Objektif:** Kajian sistematik ini bertujuan untuk mengkaji kerentanan terhadap jangkitan SARS-CoV-2 pada individu dengan polimorfisme ACE2 dan / atau TMPRSS2. **Metodologi:** Kajian yang berkaitan dengan ACE2, TMPRSS2 dan COVID-19 dikenal pasti dari empat pangkalan data - Scopus, ScienceDirect, PubMed dan EBSCOhost yang sesuai dengan kriteria yang telah ditentukan menjalani pemeriksaan awal dan teks penuh. Sebanyak tujuh kajian kemudian dikaji secara sistematik. **Keputusan:** Kajian sistematik yang menentukan hubungan polimorfisme ACE2 dan/atau TMPRSS2 dengan kerentanan terhadap jangkitan SARS-CoV-2 menunjukkan hasil yang tidak jelas pada populasi yang dikaji. Tidak ada hubungan yang diamati antara polimorfisme ACE2 dan kerentanan terhadap jangkitan SARS-CoV-2. Sementara itu, beberapa varian TMPRSS2 seperti rs7282236, rs75603675 dan rs61735789 dikaitkan dengan positif ujian COVID-19 dan beberapa varian TMPRSS2 lain seperti rs12329760, rs200291871, rs61735792 dan rs61735794 mungkin dikaitkan dengan jangkitan SARS-CoV-2. **Perbincangan:** Terdapat beberapa faktor seperti jenis dan lokasi polimorfisme yang didapati mempengaruhi kerentanan terhadap jangkitan SARS-CoV-2. Beberapa polimorfisme yang terdapat dalam kajian adalah polimorfisme nukleotida tunggal (SNP), polimorfisme senyap atau sinonim, kehilangan, intronik dan campuran penyisipan dan penghapusan (indel). Polimorfisme ACE2 biasanya terdapat pada kumpulan kawalan berbanding dengan pesakit. Telah ditentukan bahawa polimorfisme ACE2 dari semua jenis dan lokasi tidak didapati mempengaruhi kerentanan terhadap jangkitan tetapi dapat mempengaruhi keparahan penyakit pada pesakit. Bagi polimorfisme TMPRSS2, ia

sering dijumpai pada pesakit COVID-19 dan kebanyakan varian yang dijumpai berkaitan dengan jangkitan SARS-CoV-2 tanpa mengira jenis polimorfisme. **Kesimpulan:** Kajian sistematik ini menunjukkan penemuan bercampur mengenai hubungan polimorfisme ACE2 dan/atau TMPRSS2 dengan kerentanan terhadap jangkitan SARS-CoV-2. Secara amnya, penemuan ini kebanyakannya berasal dari populasi kecil dan oleh itu tidak mencukupi untuk menyimpulkan hubungan polimorfisme ACE2 dan / atau TMPRSS2 dengan kerentanan terhadap jangkitan SARS-CoV-2. Terdapat keperluan untuk kajian-kajian ini dilakukan pada populasi yang lebih besar untuk lebih memahami kemungkinan perkaitan kedua-dua protein dengan jangkitan SARS-CoV-2.

Kata Kunci: ACE2, TMPRSS2, polimorphisme, kerentanan COVID-19, jangkitan SARS-CoV-2

ACKNOWLEDGEMENT

First and foremost, I would like to express my utmost gratitude to Associate Professor Dr Syahril Abdullah for his supervision throughout this Final Year Project (FYP) journey. His way of thinking and problem solving has motivated me to keep doing this project and do well through ups and downs. With his guidance, I was able to complete my work despite all challenges.

I am also grateful for my friend which is also my project team member, Aina Nadheera Abd Rahman for her support and guidance. Thanks to her detail explanation, assistance and valuable information for every question and problem that I have, I was able to understand better and put together the idea to complete this work. Furthermore, I am thankful for her diligent help and precious strength for my project that enable me to complete my studies.

Lastly, I would like to express my gratitude for my parents, siblings and friends for being understanding and supportive throughout doing this project. Their support, confidence and unconditional loves for me encourages me to do better and stay on track with the others. This dissertation stands as a testament to everyone's support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

%	Percentage
℃	Degree celcius (temperature)
2019-nCoV	2019 novel coronavirus
3D structure	Three-dimensional structure
ABO	Blood grouping system
ACE	Angiotensin-converting enzyme
ACE2	Angiotensin-converting enzyme 2
ACEH	Alternate name of angiotensin converting enzyme 2
ALI	Acute lung injury
ARDS	Acute respiratory distress syndrome
COVID-19	Coronavirus disease 2019
DNA	Deoxyribonucleic acid
E protein	Envelope protein
e.g	Exempli gratia or for example
EBM	Evidence-based medicine
ERGIC	Endoplasmic reticulum - golgi intermediate compartment
Etc	Et cetera
hACE2	Human angiotensin converting enzyme 2
IL-1	Interleukin 1
IL-6	Interleukin 6
IFNs	Interferons
Kb	Kilobases
M protein	Membrane protein

MERS	Middle-east respiratory syndrome
MERS-CoV	Middle east respiratory syndrome coronavirus
MESH	Medical Subject Headings
mRNA	Messenger ribonucleic acid
N protein	Nucleocapsid protein
n.d.	No date
N-glycosylated	Nitrogen-glycosylated
N-terminal	Nitrogen terminal
Nm	Nanometer
NOS	Newcastle-Ottawa Scale
Nsp	Non-structural protein
ORF	Open reading frame
ORF1a/b	Largest gene containing overlapping open reading frames
p22	Short (petit) arm, region 22
PRISMA	Preferred Reporting Items of Systematic Review and Meta-analysis
q22	Long arm, region 22
RAAS	Renin-angiotensin-aldosterone system
RBD	Receptor binding domain
RBM	Receptor binding motif
RdRP	RNA-dependent RNA polymerase
RNA	Ribonucleic acid
ROS	Reactive oxygen species
S protein	Spike protein
SARS	Severe acute respiratory syndrome

SARS-CoV	Severe acute respiratory syndrome coronavirus
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SIRS	Systemic inflammatory respiratory syndrome
SNP	Single nucleotide polymorphism
ssRNA	Single-strand ribonucleic acid
SRCR	Scavenger Receptor Cysteine-Rich domain
TMPRSS2	Transmembrane serine protease 2
TNF-α	Tumor necrosis factor alpha
WHO	World Health Organization

CHAPTER 1

INTRODUCTION

1.1 Background

Coronavirus is a virus that is well-known for having the crown-like structure which are the spikes at the outer surface when viewed under the electron microscope and the virus was named corona which means “crown” according to the Latin words. Coronavirus is under the genus *Betacoronavirus* that is of the family Coronaviridae under the order Nidovirales (Muhammad Adnan et al., 2020). Coronavirus is a single stranded RNA virus that is encapsulated with membrane protein (M), nucleocapsid protein (N), envelope (E) and spike protein (S) and this virus could infect both humans and animals. People that are infected with these coronaviruses will be having respiratory problems as these viruses commonly invade the respiratory system. Among all strains of coronavirus available in this world, there are three strains that are highly infectious and lethal to humans which are SARS-CoV, MERS-CoV and the latest one is SARS-CoV-2.

COVID-19 is the current pandemic that is caused by the severe acute respiratory coronavirus 2 (SARS-CoV-2). This disease was first reported in Wuhan city in China at the end of 2019 when there was an increase in reported

cases of pneumonia-like diseases due to unknown sources. Eventually, the disease was associated with a novel coronavirus and was named as 2019 novel coronavirus (2019-nCoV). Not long after that, the novel coronavirus disease has spread across the continent and affected 216 countries. In March 2020, the World Health Organization (WHO) declared the disease called coronavirus disease 2019 (COVID-19) as a global pandemic with the frequent reported symptoms being fever, cough and fatigue. As reported in December 2020, there are already around 75 million reported cases of COVID-19 across the world with more than 1.5 million mortality cases (WHO, 2020). SARS-CoV-2 is well known for being highly transmissible from human-to-human and the identified transmission route is through the respiratory droplets such as coughing and sneezing, close contact with the infected person and fomite (Yesudhas et al., 2020). The infection of the virus into the cells happened through interaction of the viral spike (S) protein with the host receptor and protein, ACE2 and TMPRSS2 respectively.

For foreign molecules to enter a cell, it will need to be aided by the host's cell receptors or proteins to facilitate and enable the entry to occur. For SARS-CoV-2, it is aided by ACE2 and TMPRSS2 protein. Angiotensin-converting enzyme 2 (ACE2) is a transmembrane protein that binds to SARS-CoV-2 S protein and it is commonly expressed in the cells of the alveoli and the intestinal lining. Apart from that, ACE2 protein could also be found expressed in other extrapulmonary organs such as the epithelial lining of the cardiovascular system, the kidney and the brain and it is not expressed in the organs of the immune system (Xiao et al., 2020). Meanwhile, the transmembrane protease serine 2 (TMPRSS2)

functions to cleave the virus's S protein causing activation of the S protein to fuse with the host cell membrane. TMPRSS2 is found expressed in epithelial cells of the prostate and the lungs (Stopsack et al., 2020).

Genetic polymorphism is a condition where the DNA sequences in the genes of an individual are varied. The variation of the DNA sequences could be two or more and these polymorphisms are involved in determining the diversity of individuals, groups and populations. Single nucleotide polymorphism (SNP) is the most common type of polymorphisms found in the populations alongside the sequence repeat, insertion, deletion and recombination. The SNP involves a variation of a single base pair at a particular site of the gene; this polymorphism might actually be involved with susceptibility and severity of diseases such as diabetes or hypertension. Cargill et al. (1999) as cited in Somaia and Mona (2012) has reported that there are about 300 million of common SNP in human population which one-third of it were used in the research of association of SNP with some diseases (National Human Genome Research Institute, n.d.; Somaia & Mona, 2012). According to Devaux et al. (2020) and Hou et al. (2020), the susceptibility to SARS-CoV-2 infections or COVID-19 disease outcome were plausible to be influenced by the ACE2 gene and protein polymorphism, ACE2 mRNA expression and TMPRSS2 gene polymorphisms. Therefore, the relationship of the susceptibility to SARS-CoV-2 infections with the ACE2 and/or TMPRSS2 polymorphism in an individual is the key interest of this study.

Currently, there are many available studies that have been conducted on factors related with the susceptibility to COVID-19 such as the host genetic factor, ABO blood group, gender, etc (Anastassopoulou et al., 2020). The studies on ACE2 and/or TMPRSS2 with relations to susceptibility to COVID-19 is also available. However, there is a need for further studies on the polymorphism of these host proteins, ACE2 and TMPRSS2, on their association with the susceptibility of population to SARS-CoV-2 infection (Stopsack et al., 2020; Hou et al., 2020). Study on the association between the ACE2 and/or TMPRSS2 polymorphism with susceptibility to SARS-CoV-2 infection is important as this could help for better understanding the disease itself and also for a better development of targeted-therapy (Stopsack et al., 2020). For that, a study of systematic review is used to summarise previous available data on susceptibility to COVID-19 with relations to ACE2 and/or TMPRSS2 polymorphism. Systematic review is a systematic approach to 1) gather, analyse and summarise the findings of the ACE2 and/or TMPRSS2 polymorphisms and susceptibility to SARS-CoV-2 infection, and 2) determine the susceptibility to SARS-CoV-2 infection in individuals or populations with ACE2 and/or TMPRSS2 polymorphisms. By doing systematic review, the relationships and evidence of available studies could clearly be seen alongside with the limitations and disagreement of previous studies. Thus, this paper aims to find and relate the ACE2 and/or TMPRSS2 polymorphism with susceptibility to SARS-CoV-2 infection.

1.2 Objectives

1.2.1 General Objectives

This study aims to investigate the susceptibility to SARS-CoV-2 infection in individuals with ACE2 and/or TMPRSS2 polymorphism.

1.2.2 Specific Objectives

The specific objectives of this study are to use systematic review to 1) identify ACE2 and/or TMPRSS2 polymorphism in COVID-19 patients and 2) determine the susceptibility to SARS-CoV-2 infection in individuals or populations with ACE2 and/or TMPRSS2 polymorphism.

1.3 Hypotheses

It is hypothesized that there is an association between ACE2 and/or TMPRSS2 polymorphism with susceptibility to SARS-CoV-2 infections.

CHAPTER 2

LITERATURE REVIEW

2.1 COVID-19

Coronavirus disease 2019 (COVID-19) is a new respiratory disease that is caused by the coronavirus called SARS-CoV-2. COVID-19 is the fifth reported global pandemic following four previous pandemics caused by the influenza viruses and affecting individuals throughout the globe. This disease was first called Wuhan pneumonia due to the symptoms of the disease that is similar to pneumonia before officially named as coronavirus disease 2019 (COVID-19) in early 2020 (Liu et al., 2020). The most common symptoms of COVID-19 reported are fever, coughing and fatigue, however; there are also some less common symptoms of COVID-19 such as sore throat, headache, muscle or joint pain, nausea, vomiting, diarrhoea, chill, dizziness, skin rashes, anosmia or loss of smell, loss of taste, nasal congestion and conjunctivitis. In some cases, COVID-19 patients reported to have developed severe symptoms including shortness of breath, loss of appetite, confusion, high fever with temperature above 38 °C and chest pain. To note, there are also some rare symptoms of COVID-19 such as irritability, confusion, anxiety, depression, sleep disorders, reduced consciousness which is sometimes associated with seizures and some other neurological complications including brain inflammation, strokes, delirium and nerve damage (Figure 2.1) (WHO, 2020). It is

recommended for all individuals who experienced any of the symptoms to go for medical check-up for early diagnosis and treating the symptoms.

Most of COVID-19 patients, about 80% of them usually recover naturally from COVID-19 without the need for hospital treatment, but there are also some patients whose conditions developed severely that need oxygen support and also intensive care. There are several complications that might develop in individuals with COVID-19 such as respiratory failure, acute respiratory distress syndrome (ARDS), thromboembolism, septic shock, multiorgan failure which includes the heart, liver and kidney injury and the worst case is death that could be due to these complications (Figure 2.1). COVID-19 could develop in any infected individuals but there are some groups of people that are at risk of developing severe illnesses and they are people aged 65 years and older and those of any age with underlying medical conditions like cardiovascular diseases, diabetes, chronic respiratory diseases and cancer (WHO, 2020). Currently, there are many studies done for further understanding of some other hosts' risk factors towards susceptibility to COVID-19 such as the ABO blood group and ACE2 and TMPRSS2 polymorphisms.

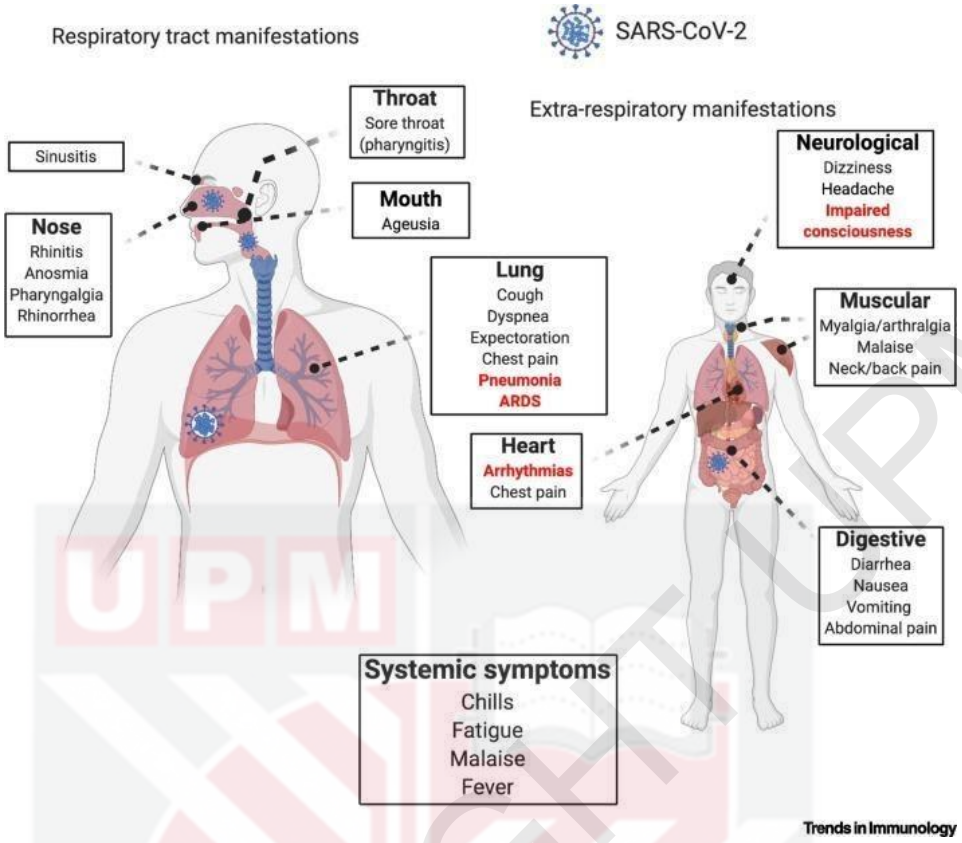

Figure 2.1. Collective symptoms of COVID-19 reported by the infected patients. COVID-19 not only manifests respiratory illnesses but it also affects the other body systems including muscular system, cardiovascular system, nervous system and digestive system. The highlighted red-coloured symptoms are the symptoms that are over-expressed in the severe COVID-19 patients (Harrison et al., 2020).

When a person was exposed to COVID-19, they are most likely to get infected by the virus but it would take some time for the symptoms to appear. The time between the exposure to COVID-19 till the moment the symptoms begin are called incubation period. This incubation period usually ranges from 1 to 14 days but on average, an infected person will experience the symptoms of COVID-19 between 5 to 6 days. Once a person has experienced the symptoms or has been diagnosed with COVID-19, they should isolate or quarantine themselves to avoid transmitting the virus to other susceptible people. This step also works on those who have been in close contact with COVID-19 positive patients. Some

precautions step that needs to be taken by all individuals are wearing a mask while in public especially in a crowded area, physically distancing, keeping room well ventilated, coughing into bent elbow or tissue, avoiding crowds and close contact and cleaning the hands regularly using water and soap or hand sanitizer (WHO, 2020). An important step is to take COVID-19 vaccines as it could help in reducing the risk of getting serious illness due to COVID-19 and most importantly is to prevent from spreading SARS-COV-2 to other susceptible people.

2.2 SARS-CoV-2

Coronavirus is a virus that is well-known for having the crown-like structure which are the spikes at the outer surface when viewed under the electron microscope and the virus was named corona which means “crown” according to the Latin words. Coronavirus is under the genus *Betacoronavirus* of the family Coronaviridae under the order Nidovirales (Shereen et al., 2020). Coronavirus is a positive single stranded RNA virus that is encapsulated with membrane protein (M), nucleocapsid protein (N), envelope (E) and spike protein (S) and this virus could infect both humans and animals (Figure 2.2). The size of the coronaviruses are very small with diameters between 65 to 125 nm and it contains the genetic materials which is the single-strand RNA with the sizes between 26 to 32 kbs in length (Shereen et al., 2020). There are 14 open reading frames (ORFs) of the genetic materials of the coronavirus which one-third of it encodes for four structural proteins (S, E, M and N) and nine accessory proteins (ORF) (Figure 2.3).

The S proteins are known for mediating SARS-CoV entry into the host cells. Two-third of the genome contains the genetic materials for 16 non-structural proteins (nsp 1-16) which function to make up the replicase complex of the coronavirus (Harrison et al., 2020).

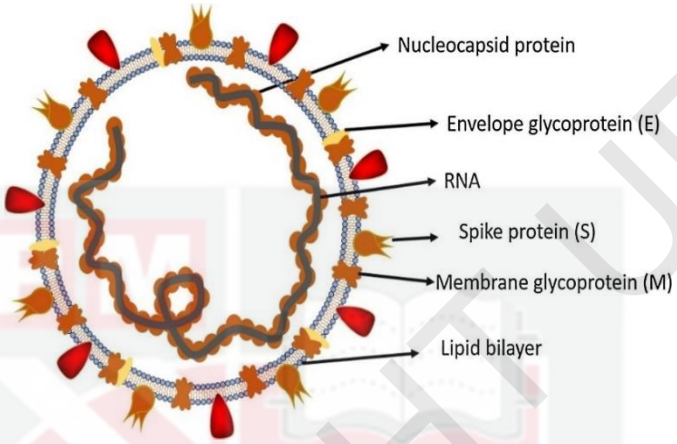

Figure 2.2. Structure of human coronavirus causing respiratory syndrome (Shereen et al., 2020).

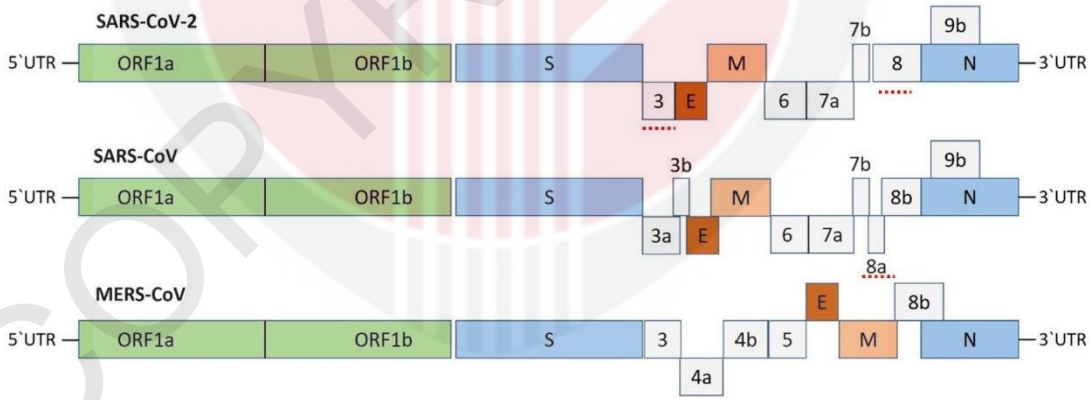

Figure 2.3. Organization of *Betacoronaviruses* (SARS-CoV, MERS-CoV and SARS-CoV-2) genome (Muhammad Adnan et al., 2020).

People that are infected with these coronaviruses will be having respiratory problems such as acute lung injury (ALI) and acute respiratory distress syndrome

(ARDS) as these viruses commonly invade the respiratory system. These respiratory diseases will deteriorate into respiratory failures and could even lead to death if it is not treated well in the early stages (Shereen et al., 2020). Among all strains of coronavirus available in this world, there are three strains of these coronaviruses that are highly infectious and lethal to humans which are SARS-CoV, MERS-CoV and the latest one that is causing pandemic is the SARS-CoV-2. SARS-CoV-2 is well known to be highly transmissible and pathogenic compared to SARS-CoV. The enhanced transmission ability might be due to the genetic recombination on the RBD region on the S protein of SARS-CoV-2 (Shereen et al., 2020).

SARS-CoV-2 infection cases first emerged in Wuhan city, China at the end of 2019 and it is initially linked with Huanan seafood market which the infections were thought to be due to contact with the animals. However, there was human-to-human transmission reported later. The community transmission of the disease occurred within China rapidly which later spread to other 200 countries across the continents (Parasher, 2020). Transmission of the disease globally was most likely to be due to the international travel that allows the further spread of the disease. SARS-CoV-2 infections were thought to have originated from the bats which later spread to the other animals such as the cats, dogs and pangolins (Yesudhas et al., 2020). According to WHO (2020), the main route of transmission of SARS-CoV-2 is through the respiratory droplets and direct contact with the infected person. Some other possible SARS-CoV-2 transmission routes are airborne, fomites, fecal-oral route, bloodborne, mother-to-child and also animal-to-human transmission (Figure 2.4). Now, there are

many asymptomatic cases reported where the individuals have not developed any symptoms following infection to SARS-CoV-2. The disease caused by SARS-CoV-2 is called coronavirus disease 2019 (COVID-19) (Chan et al. as cited in Hoffmann et al., 2020).

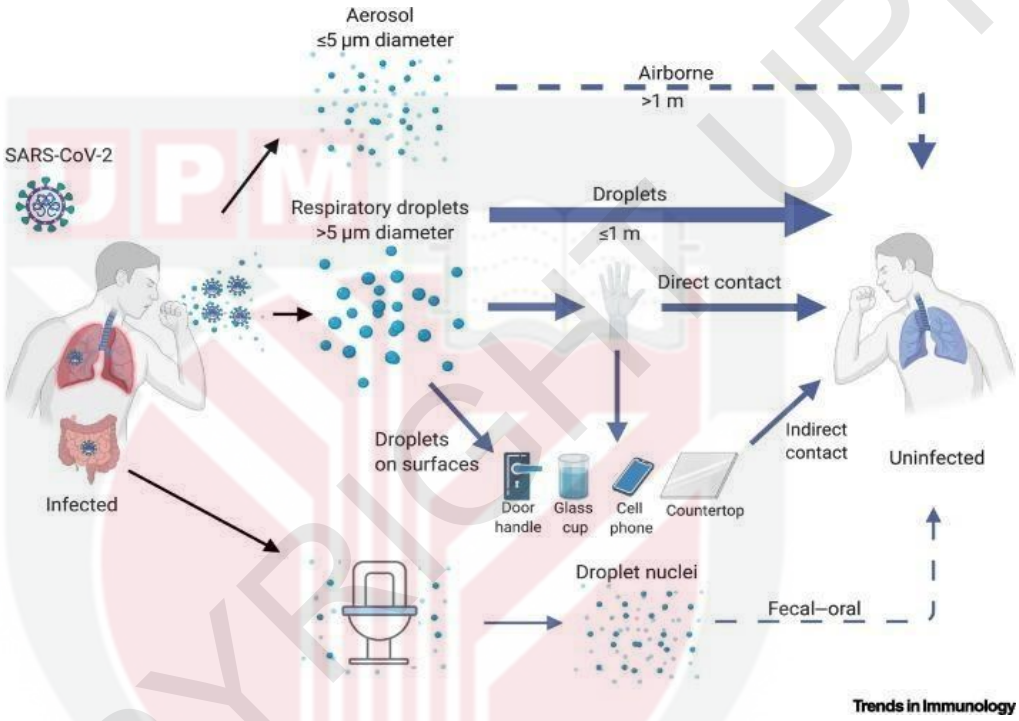

Figure 2.4. Transmission routes of SARS-CoV-2 (Harrison et al., 2020).

The novel coronavirus, SARS-CoV-2, was found to be closely related with SARS-CoV as it shows similar features under microscopic investigations including the clinical manifestations of the disease caused by the virus. Both SARS-CoV and SARS-CoV-2 are of beta lineage of the coronaviruses and the identification of the whole genome of SARS-CoV-2 are about 80% similar with SARS-CoV. To add, the genomic sequence of the spike protein of SARS-CoV-2 are also similar with

the SARS-CoV, with about 76 to 78% similarities. Other similarities found in both coronaviruses are in the receptor-binding motifs (RBM) and receptor-binding domain (RBD) with 50-53% and 73-76% similarities, respectively (Figure 2.2). These similarities in the sequence of the S protein of both coronaviruses explains the possibilities of infection in the host cells is facilitated by the same receptor within the host cells which is the angiotensin-converting enzyme 2 (ACE2) (Yesudhas et al., 2020; Hoffmann et al., 2020). The transmembrane protein, TMPRSS2, was also involved with the viral entry into the host cells.

2.3 ACE2

Angiotensin-converting enzyme 2 (ACE2) is a type I transmembrane protein that is encoded by the ACE2 gene which is located on the p22 region of X chromosome. An ACE2 protein is formed of two terminal which are N-glycosylated N-terminal domain that contains the carboxypeptidase site which is the binding site of SARS-CoV and the other one is the C-terminal that acts as a cytoplasmic tail (Figure 2.5). This ACE2 protein is expressed the most in the epithelial cells of the lung especially in the alveolar, the small intestine, heart, kidneys and other organs that are not of the immune system (Xiao et al., 2020). ACE2 is an enzymatic protein that acts opposite to the ACE enzyme. Generally, ACE2 has a protective role in the cardiovascular system which it aids in maintaining the blood pressure in the renin-angiotensin-aldosterone system (RAAS) through the action of converting the angiotensin I to angiotensin 1–9 and angiotensin II to angiotensin 1–7 that act as

vasodilator thus reducing the blood pressure (Xiao et al., 2020; Ragia & Manolopoulos, 2020).

Not only does ACE2 act as an enzyme, but it is also a protein that enables the entry of SARS-CoVs into the host cells through the binding of the viral S protein with the ACE2 receptor in the host cells. The viral S protein has a great binding affinity with the ACE2 receptor for the cell entry which the S1 subunit of the spike binds with the residue on the ACE2 receptor and the S2 subunit is activated and undergoes cleavage and shedding (Xiao et al., 2020; Yesudhas et al., 2020). This interaction between ACE2 and SARS-CoV-2 is part of the reasons for the higher susceptibility to COVID-19 in patients with underlying disease such as diabetes and hypertension as the level of ACE2 in these patients are increased thus making them more susceptible to COVID-19 (Xiao et al., 2020). To add, the expression of ACE2 gene could be up-regulated in the epithelial cells of the respiratory systems following the release of some cytokines such as the interferons (IFNs) due to the viral infections (Harrison et al., 2020) thus increasing one's susceptibility and severity to a disease such as COVID-19. By learning the interaction of the ACE2 receptor with SARS-CoV-2, researchers are determined that a strategy to treat COVID-19 patients could be developed by interrupting the binding of viral S protein with ACE2 receptor (Ragia & Anastassopoulou, 2020).

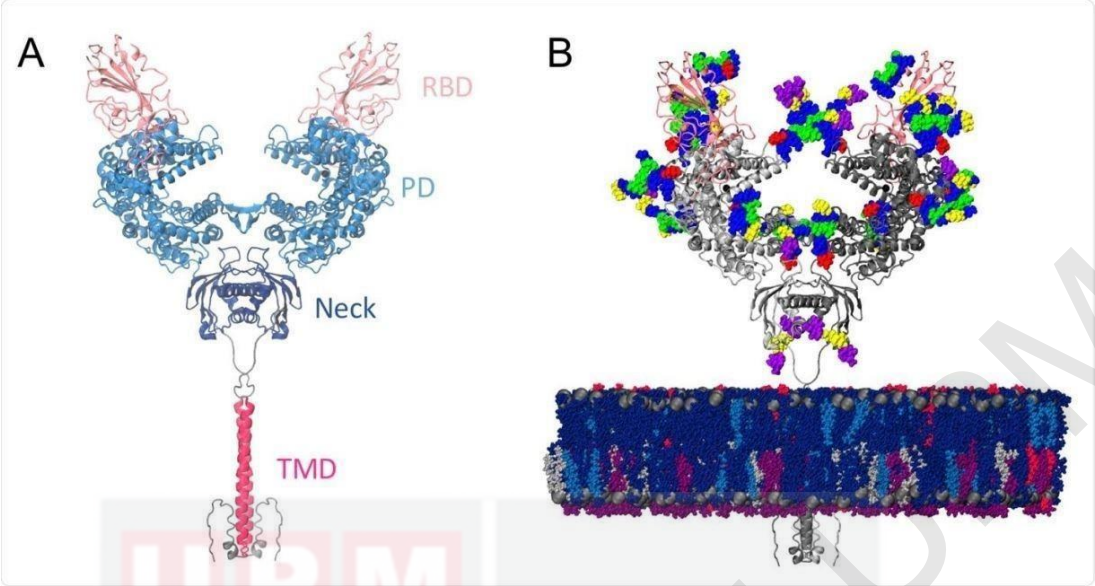

Figure 2.5. Model structure of ACE2 receptor. (**A**) Full length ACE2 protein structure. (**B**) ACE2 protein receptor embedded in the membrane (Robertson, 2020).

2.4 TMPRSS2

Transmembrane protease serine 2 (TMPRSS2) is the type II transmembrane protein encoded by the TMPRSS2 gene on the chromosome 21q22 and this protein is normally manifested as cellular protein in the epithelial cells of the prostate with the function of regulating the normal function of the prostate (Mollica et al., 2020; Shabir, 2020). TMPRSS2 is highly expressed in individuals with prostate cancer as the expression of TMPRSS2 is regulated by the androgens (Shabir, 2020). Not only TMPRSS2 is expressed in the prostate's epithelium but also in the epithelium of the respiratory tract and upper gastrointestinal tract (Stopsack et al., 2020; Shabir, 2020). As TMPRSS2 is also expressed in the respiratory tract epithelium, SARS-CoV-2 infections of the human cells could occur as TMPRSS2 in the cells

will prime and cleave the virus spike protein for fusion into the host cells following the binding of the virus with ACE2 receptor (Yesudhas et al., 2020; Shabir, 2020). The enzyme cleaves the peptide bonds with the nucleophilic amino acids, serine, that is located within the active site of the target protein (Shabir, 2020). The differential susceptibility to influenza virus or coronavirus infections in individuals within the populations could be understood through the understanding of the diverse TMPRSS2 protein expressions in the respiratory tissues (Mollica et al., 2020; Stopsack et al., 2020). Studying the expression of TMPRSS2 protein could also be done in the lungs of the healthy individuals which the expression of TMPRSS2 might be altered following the viral infections. This study follows the previous study that has observed the altered expression of the transmembrane protein on the ACE2 receptor that is due to the SARS-CoV infections (Stopsack et al., 2020). By learning the expressions and role of TMPRSS2 proteins in a disease pathogenesis, it could aid the researchers in further understanding the risk factors of respiratory viral infections such as SARS-CoV-2 and allow them to develop a better treatment for a disease.

2.5 Mechanism SARS-CoV-2 Infection

The mechanisms of SARS-CoV-2 infection is similar with the SARS-CoV (Ragia & Manolopoulos, 2020) and it is suggested to have involved the use of ACE2 receptor along with TMPRSS2 protein for the viral entry (Kupferschmidt & Cohen as cited in Mollica et al., 2020; Stopsack et al., 2020). SARS-CoV-2 infection occurs as soon as the virus enters the body of a susceptible person via any route of transmission and then binds with the host receptor, ACE2. Viral entry into the host cells occurs in two ways, either via endocytosis or membrane fusion. It is facilitated by the attachment and binding of the S protein subunit of the coronaviruses, S1 with the host's cellular receptor (Hoffmann et al., 2020). There is a region in the S protein of SARS-CoV-2 called receptor binding domain (RBD) which is known to bind with the ACE2 receptor on the host cells. The RBD region of the S protein was made up of 3D structure which functions to maintain the van der Waals forces between S protein and host receptor. To add, the critical lysine 31 residue residue on the human ACE2 receptor could recognize the 394 residue (glutamine) in the RBD region (Shereen et al., 2020).

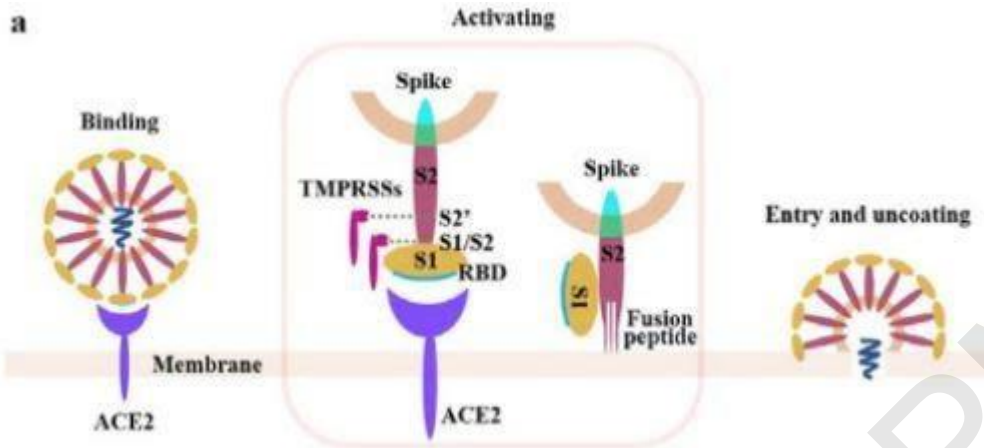

Figure 2.6. Mechanism of SARS-CoV-2 virus entry into the host cells via membrane fusion. The S protein binds with the ACE2 receptor and undergo priming and activation by TMPRSS2 which TMPRSS2 cleaves the S protein at the S1/S2 site and allows membrane fusion to occur (Meng et al., 2020).

After binding of S1 with the host receptor, priming of S protein occurs followed by the cleavage of S1/S2 site by the cellular enzyme, transmembrane serine protease 2 (TMPRSS2). The S2 subunit then undergoes the conformational changes that enable the viral envelope to fuse with the host cell membrane forming the post-fusion state and releasing the viral genomic material, the RNA, into the cytosol of the host cells (Figure 2.6). The viral RNA that consists of ORF1a/b then translated, producing two viral replicase polyproteins, pp1a and pp1ab. Next, the polyproteins are cleaved into smaller products consisting of complexes of replicase non-structural proteins and viral polymerase that is known as RNA-dependent RNA polymerase (RdRP). The cleavage is done by the viral-encoded protease. The positive strand RNA genome is replicated into the negative strand RNA genome. From the negative strand RNA, the subgenomic RNA is produced through transcription and translation producing the viral proteins (S, E, M and N). After that, the viral proteins and the positive strand RNA genome assemble into the ER-

Golgi intermediate compartment (ERGIC) to form the virion. The virion is carried by the vesicles, then the virion is released out of the cells via exocytosis (Yesudhas et al., 2020; Harrison et al., 2020; Muhammad Adnan et al., 2020; Hoffmann et al., 2020). As soon as the virion left the host cells, it could invade the other host cells and the process of the infections is repeated (Figure 2.7).

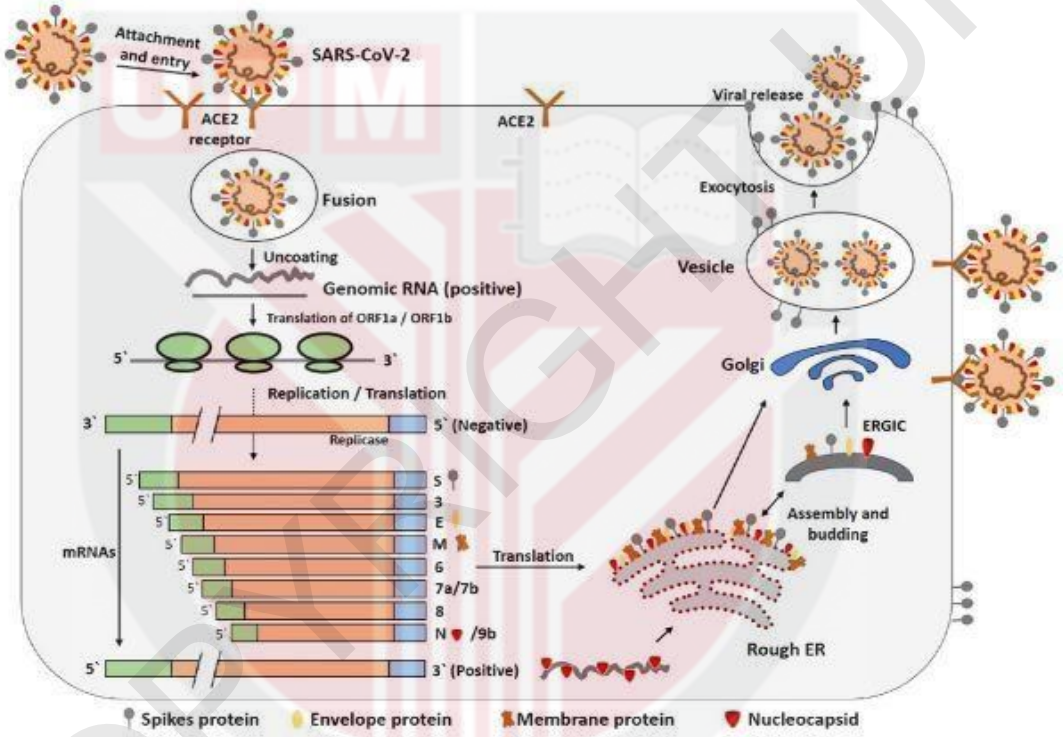

Figure 2.7. The figure shows the life cycle of the SARS-CoV-2 in the host cells (Shereen et al., 2020).

2.6 COVID-19 pathophysiology

Infections of SARS-CoV-2 through the airways enable the access and invasion of the virus to the host cells' receptor, ACE2 in the respiratory tracts

especially in the alveoli. In the alveoli, the virus invades the type 2 pneumocytes that function to produce the surfactant that reduces the surface tension of the alveoli thus preventing the alveoli from collapse. As shown in Figure 2.8, as the virus infects the cells, it causes damages to the cells that leads to the production of inflammatory mediators that activates the alveolar macrophages to release cytokines such as interleukin-1 (IL-1), interleukin-6 (IL-6) and tumour necrosis factor alpha (TNF-α). These cytokines cause the capillaries that surround the alveoli to dilate (vasodilation) and the endothelial cells contract, thus increasing the permeability of the capillaries. These will lead to the leaking of the capillaries which the plasma will move out of the capillaries into the interstitial spaces and accumulate within the area and compressing the alveoli. In some cases, fluid enters the alveoli causing alveolar oedema, lowering the production of surfactant and increasing the surface tension thus leading to alveolar collapses. As this happens, there will be reduced gas exchanges causing hypoxemia and this condition is called acute respiratory distress syndrome (ARDS) (Ninja Nerd Lectures, 2020; Medinaz, 2020).

To add, the cytokines released will also attract the neutrophils into the alveoli and release the reactive oxygen species (ROS) and proteases to destroy the viruses within the infected cells. These ROS and proteases not only destroy the viruses but it also damages the surrounding uninfected cells such as the type 1 and type 2 pneumocytes. The damaged cells will be accumulated within the centre of the alveolar and the accumulation consists of cellular debris, fluid and protein deposition. This accumulation causes consolidation that is bad for the host as it

leads to alteration in gas exchanges, hypoxemia and mechanical irritation that cause coughing (Ninja Nerd Lectures, 2020; Medinaz, 2020).

The released IL-1 and IL-6 will not only circulate within the respiratory systems but also go to the hypothalamus in central nervous systems to send signals for releasing the prostaglandin that will increase the body temperature and cause fever. Due to the hypoxemia, the chemoreceptor within the body will detect and send a signal to the heart to pump more blood to increase the gas exchanges and this action leads to increasing heart rate and respiratory rate. In a severe condition where the inflammatory responses are able to affect the whole body, the condition known as systemic inflammatory response syndrome (SIRS) or commonly called systemic inflammation occurs. Eventually, sepsis or septic shock could happen when the hypotensive condition occurs due to the vasodilation that leads to hypovolemia and reduced peripheral resistance. When hypotension occurs, it leads to a reduced blood flow and multiorgan failure (Ninja Nerd Lectures, 2020; Medinaz, 2020).

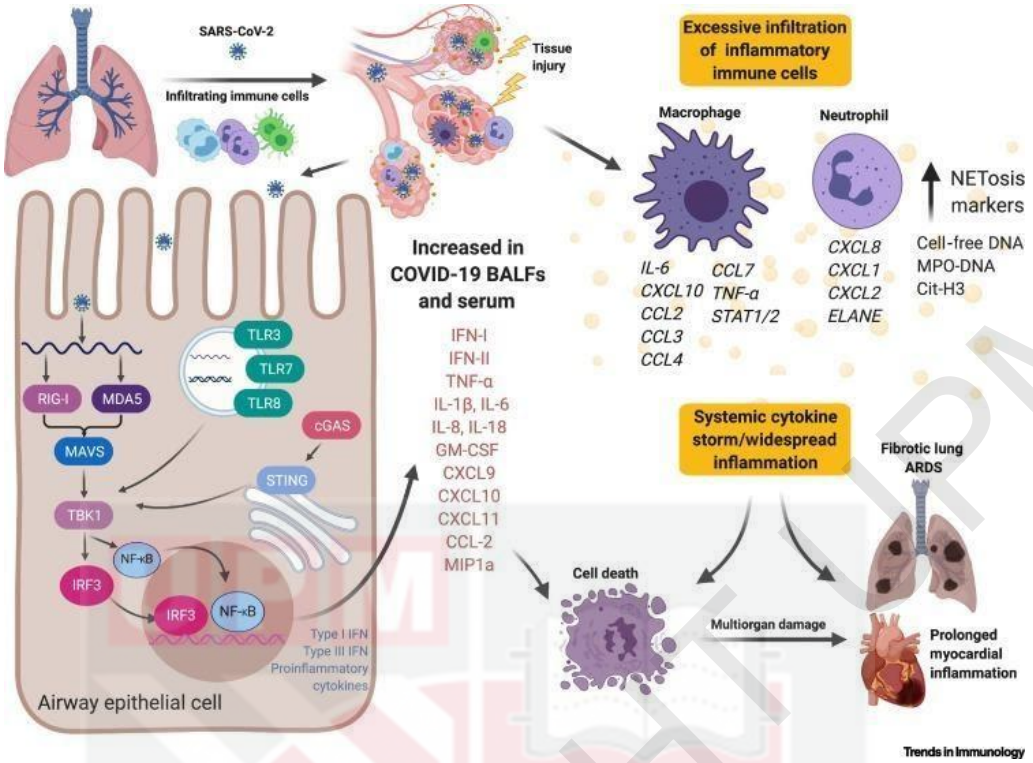

Figure 2.8. Brief overview on pathogenesis of COVID-19 in the lung and other organs (Harrison et al., 2020).

2.7 Polymorphism

Polymorphism is a word that is made up from the combination of Greek words *poly* and *morph* which mean multiple and form, respectively. This word is commonly used in genetic context for describing the diverse appearance of a single gene within an individual or the populations. Generally, genetic polymorphism is used to refer to different forms of a genotype or alleles in a population in which the variations are distinct and could occur in two or more forms. Through this polymorphism, there is variation and diversity in the populations and makes an individual distinct from each other (Wang & He, 2018; Phillips, 2020). There are

many types of polymorphisms such as single nucleotide polymorphism (SNP), tandem repeat polymorphism, insertion, deletion and genetic recombination (Somaia & Mona, 2012; National Human Genome Research Institute (NIH), n.d.) (Figure 2.9). SNP is the most common polymorphism found in a population. There are also other types of polymorphism such as synonymous polymorphisms which do not have an effect on the organism as there are no amino acid changes in the protein produced and it is also known as a silent mutation. Within the SNP, there might be non-synonymous substitution that results in the changes of the codon (Somaia & Mona, 2012).

Figure 2.9. Examples of common types of polymorphisms, the single nucleotide polymorphisms (SNPs) and tandem repeat polymorphism (National Human Genome Research Institute (NIH), n.d.).

A trait will be considered as polymorphisms when the frequency of the allele is more than 1% in the population (Phillips, 2020). The polymorphisms are usually confused with mutation, however; it is notable that the polymorphism first arises from a mutation and it could be inherited throughout generations (Somaia & Mona, 2012). Through the learning of genetic polymorphisms, it enables researchers to study the association of the polymorphisms with susceptibility to diseases, severity of a disease and reaction towards any drug or treatments. To add, the pathogenesis and progression of a disease and the nature of phenotypic variation could also be explored through learning the genetic polymorphisms (Wang & He, 2018). Based on previous study, the susceptibility to viral infection, progression of diseases and death might be contributed by the polymorphism of TMPRSS2 gene or other factors. This needs to be further studied through observational analysis as there is a high susceptibility to influenza virus infection observed in different cohorts of patients with SNP and has higher expression of TMPRSS2 (Stopsack et al., 2020). There is also a study in an Italian cohort to determine the possibility of COVID-19 severity to ACE2 and TMPRSS2 genes' level of expression and polymorphisms (Mollica et al., 2020).

2.8 Systematic Review

Systematic review is defined as a review that uses a reproducible and systematic methodology to search, select and evaluate relevant studies and extract and synthesize data obtained from the eligible studies to answer specific research

questions. According to the hierarchy of evidence, the systematic review is located on top of the pyramid which indicates that the results of the studies have minimal bias effects due to the strict and well-designed methodology (Figure 2.10). Through systematic review, the answers to the questions could be obtained through methods that have been established in advance and reported in the protocol which summarizes all previous research included (Mackenzie et al., 2012). Some medical, nursing and healthcare areas which systematic reviews are commonly done are the clinical testing, adverse effects, study on intervention, public health involvement and cost evaluation (Gopalakrishnan & Ganeshkumar, 2013). Systematic review is essential in some areas as it could guide the public and caregiver in clarifying certain issues that arise and also fill in the research gap as a guide for future studies (Hemingway & Brereton, 2009; Gopalakrishnan & Ganeshkumar, 2013). There are two types of possible results that could be synthesized from previous studies which are quantitative or qualitative.

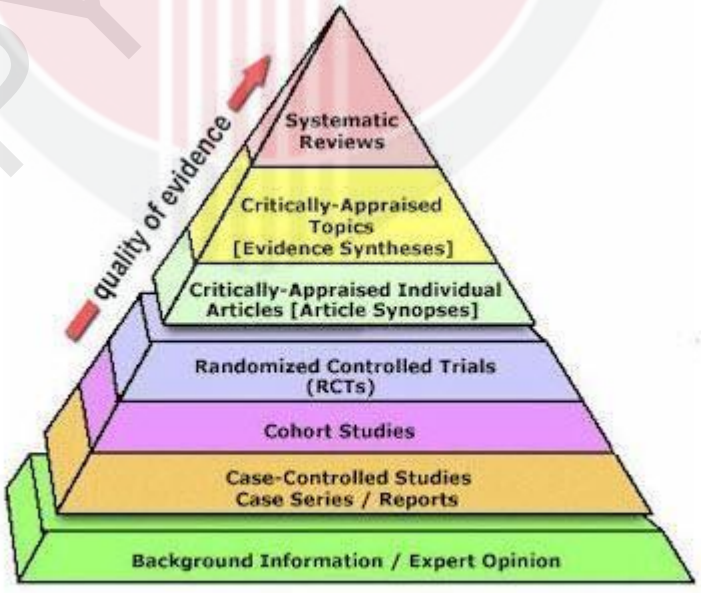

Figure 2.10. The pyramid of research evidence (University Libraries, 2021).

In systematic review, the steps involved are detailed and follow a strategy in order to reduce bias and random errors for achieving good and reliable results on an intervention where the results could be concluded (Liberati et al., 2009; Gopalakrishnan & Ganeshkumar, 2013). A well conducted systematic review is considered to have a high level of evidence as shown by the evidence-based pyramid. Findings from a systematic review could aid the healthcare practitioner in modern evidence-based medicine (EBM) (Tawfik et al., 2019). This allows and eases the researchers in better understanding of an issue and enables them to better understand and make a good decision on an important condition or situation. The most important features of systematic review are the use of clear criteria and methods to seek for relevant studies and extracting and synthesizing the findings of the included studies and this includes the implementation of standard methods to assess the quality of included studies. The Preferred Reporting Items for Systematic Review and Meta-analysis statement (PRISMA) is the basic guidelines that commonly used by researchers to create a priori protocol before they move into further steps in conducting systematic review ((University Libraries, 2021; Liberati et al., 2009).

An initial comprehensive systematic review was done following the outbreaks of the previous coronaviruses, SARS and MERS, in which the study discusses the association of host genetic factors with the susceptibility to infections by coronaviruses (Maria et al., 2020). Thus, to further understand the causes of the current COVD-19 pandemic, the systematic review can be used to discover the

association of host genetic factors, ACE2 and TMPRSS2 polymorphisms, with the susceptibility to SARS-CoV-2 infection. Through systematic review, the association between the variables could be determined and guide the researchers and clinicians for prevention and treatment measures to the patients. To add, the process of identifying the susceptibility alleles and discovering the possible treatment that is not yet determined in human study could be strengthened through this systematic review (LoPresti et al., 2020).

CHAPTER 3

METHODOLOGY

Systematic Review

3.1 Protocol and Registration

The method for conducting this systematic review was performed according to the PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols) 2015 statement (Moher et al., 2015). The protocol for this study was registered in the Prospero before proceeding with further steps of the study. The registration number of this study is CRD42021229963.

3.2 Eligibility Criteria

3.2.1 Types of Studies

Studies with any research design were included in this study with the exception of secondary study (review, systematic review, meta-analysis), letter to the editor, correspondence, medical hypothesis and commentary. The types of studies of interest for this study are observational studies that determined the presence or absence of variances of ACE2 or TMPRSS2

genes in COVID-19 patients and mentioned the types of variance of either gene of interest.

3.2.2 Types of Participants

Participants for this study were the patients that were diagnosed with SARS-CoV-2 infections as the subject of interest with the healthy participants as the control group. There are no restrictions on the age, gender, race and the location of the participants.

3.2.3 Types of Intervention

The intervention includes the observation and analysis on the sample from COVID-19 patient and control groups to observe for the presence of either ACE2 or TMPRSS2 gene polymorphism in them. The method to determine the presence of the gene variants were also reported and the relationship between ACE2 or TMPRSS2 polymorphism on susceptibility to SARS-CoV-2 infections was determined in the study.

3.2.4 Types of Outcome Measures

The outcome measures of these studies were the relationship between ACE2 and/or TMPRSS2 polymorphism and the susceptibility to SARS-CoV-2 infection in individuals were determined and deduced through the study on the presence of the variants in COVID-19 patients.

3.3 Search Strategy and Data Source

A systematic search of articles from four databases, PubMed, ScienceDirect, Scopus and EBSCOhost were done from March 1st 2021 until March 10th 2021 following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. The search term used were (((((((ACE2)) OR (Angiotensin-converting enzyme 2)) OR (hACE2)) OR (ACEH)) OR (Angiotensin I converting enzyme 2) OR ((ACE-related carboxypeptidase)) OR (Metalloprotease MPROT15)) OR (Processed angiotensin-converting enzyme 2)) AND ((((Coronavirus disease 2019) OR (COVID-19)) OR (SARS-CoV-2 infection)) OR (2019 novel coronavirus)), (((((Genetic polymorphism) OR (allelic variation)) OR (genetic variation)) OR (genetic predisposition)) OR (nucleotide variation)) AND ((((Coronavirus disease 2019) OR (COVID-19)) OR (SARS-CoV-2 infection)) OR (2019 novel coronavirus)) and (((((TMPRSS2) OR (transmembrane serine protease 2)) OR (Serine protease 10)) OR (Transmembrane protease serine 2 non-catalytic chain)) OR (Transmembrane protease serine 2

catalytic chain)) AND ((((Coronavirus disease 2019) OR (COVID-19)) OR (SARS-CoV-2 infection)) OR (2019 novel coronavirus)). However, the use of all these search terms were subjected to changes according to the requirements of each database.

In addition, the article searching in PubMed was done by using the 'MESH' term ((("ACE2 protein, human"]) AND "TMPRSS2 protein, human") AND "Polymorphism, Genetic"[Mesh]) AND ("COVID-19"[Mesh] OR "SARS-CoV-2"[Mesh]). Boolean connectors 'OR' was used within the same category while 'AND' was used between different categories of the search term. Parentheses were also used for specifying the searching of the keywords entered. The alternative names of ACE2 and TMPRSS2 were included to increase the range of the related article searching. Language used for the articles was restricted to English and the publication dates were restricted from 2019 until 2021 only in order to reduce the number of irrelevant papers. No restriction was put on the age, gender and the population of the studies. Filters were applied during the search to seek for free full text articles. The studies obtained were screened independently by two reviewers [Nurul Husna Hasnul Hadi (192182), Aina Nadheera Abd Rahman (192388)] following the inclusion and exclusion criteria.

3.4 Study Selection

Available studies obtained from all four databases during the initial search were gathered in the Microsoft Excel Spreadsheet and the duplicates were removed if they have the same author and title. The remaining studies then underwent the title and abstract screening based on the inclusion and exclusion criteria. After screening based on title and abstract were done, all the eligible papers were subjected to full-text screening based on the eligibility criteria to further determine and obtain the relevant studies. Studies that did not meet the eligibility criteria were removed. The screening process was done independently by the two reviewers and the results of the screening were cross-checked within the two reviewers for any differences in the eligible studies and disagreements. Discussion between the reviewers was done for any unclear studies or disagreement to reach the same consensus for the study. The exact details on the flow of the study selection follows the PRISMA flow diagram and the reasons for the excluding studies were recorded.

3.5 Data Extraction

Data extraction was done independently by the first reviewer [Nurul Husna Hasnul Hadi (192182)] and the extracted data were then cross-checked by the second reviewer [Aina Nadheera Abd Rahman (192388)]. Data from the included studies were extracted into the Microsoft Excel Spreadsheet which includes the

relevant data of characteristics of the included publication (e.g first author, year of publication) and the information from the studies (e.g country of the population studied, gene of interest, types of variation, number of variations, position of variation/nucleotide changes/proteins involved, allele frequency, effects caused by variation).

3.6 Quality Assessment

The assessment for the quality of the included studies was performed independently by the two reviewers by using the Newcastle-Ottawa Scale (NOS) that was for observational studies such as the cohort study and case-control study. The Newcastle-Ottawa Scale (NOS) consists of three domains which are the selection, comparability and exposure. Within the dimensions of the scale, there were eight categories which comprised the total of nine stars for full grades if the studies met the criteria of the scale. A maximum score of 4 stars were allocated for the selection domain, 2 stars for comparability domain and 3 stars for the exposure domain (Xiong et al., 2020). The score for the quality assessment was determined through the mean score of the two reviewers. The level of quality of the included studies were considered low if the score was less than 3 (< 3), moderate for score between 4 to 6 and high, if the score was more than 7 (> 7) (Zhang et al., 2021).

CHAPTER 4

RESULT

Systematic Review

4.1 Literature Search

From the literature search done on all four databases, a total of 2521 potentially relevant articles were identified according to the search strategy. The total number of articles were reduced to 2028 after the duplicates of the same title and author within and between databases were removed. The number of articles according to each database (PubMed, ScienceDirect, Scopus, EBSCOhost) were 327, 595, 391 and 715, respectively. These remaining articles underwent the initial screening and were assessed through the title and abstract according to the inclusion and exclusion criteria. The full text of forty-two articles were then screened for eligibility. Thirty-five articles were excluded from the eligible full-text for being a duplicate and the other articles that did not meet the inclusion criteria which the subject of the excluded articles were not from the human populations that consist of the COVID-19 patients and control group. The subjects used in the excluded articles were mostly the data from the human genome databases which were not part of the inclusion criteria. Following this full-text

screening, it resulted with seven eligible full-text for this qualitative study (Figure 4.1).

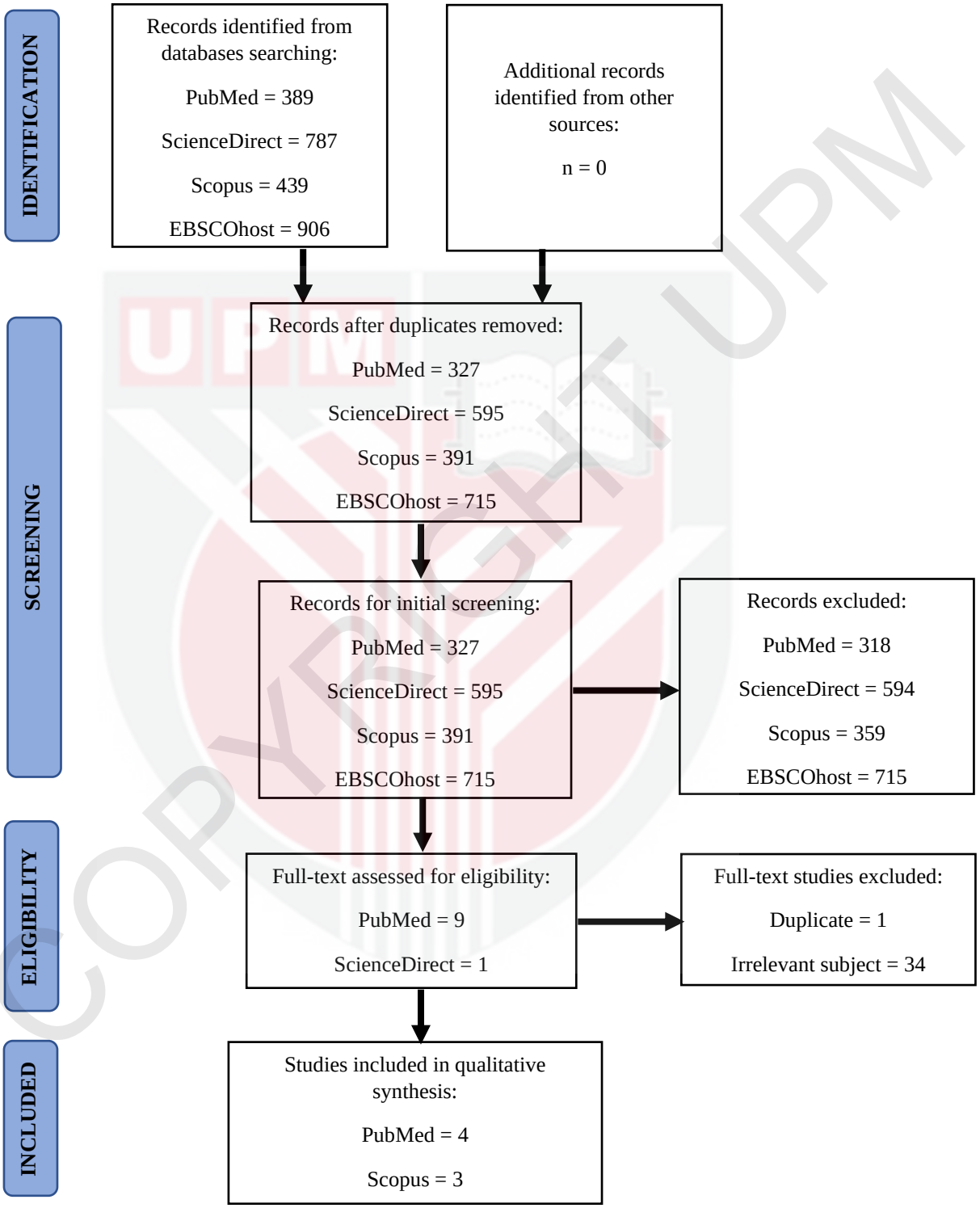

Figure 4.1. Flow of literature search based on PRISMA guidelines. From 2521 articles identified in database searching, only 7 articles fulfilled the inclusion criteria following the initial and full text screening and were included in the qualitative analysis and synthesis.

4.2 Data Extraction

Table 1 summarizes the results of the study selection and the characteristics of seven included studies from PubMed and Scopus for qualitative analysis. All papers were published in 2020. All included studies were observational studies following the case-control study design. All seven studies were conducted in different countries, including the United Kingdom (*n=1*), Italy (*n=3*), Spain (*n=2*) and Russia (*n=1*) (Figure 2). The studies include the genes of interest which are either ACE2 or TMPRSS2 or both. The primary outcome of all included studies depicts the association between the polymorphism of these genes with susceptibility to SARS-CoV-2 infections in the population of the studies. Five studies were focusing on the susceptibility to SARS-CoV-2 infection associated with ACE2 gene variants (Benetti et al. (2021), Novelli et al. (2020), Gómez et al. (2020), Torre-Fuentes et al. (2021), & Shikov et al. (2020), while three studies focus on TMPRSS2 gene variants (Zhou et al. (2021), Torre-Fuentes et al. (2021), & Latini et al. (2020). The effects caused by the variants from each study are included in Table 1, which the results are varied between each study. The consensus for this study was reached through analysis and synthesis of all studies cited in Table 1.

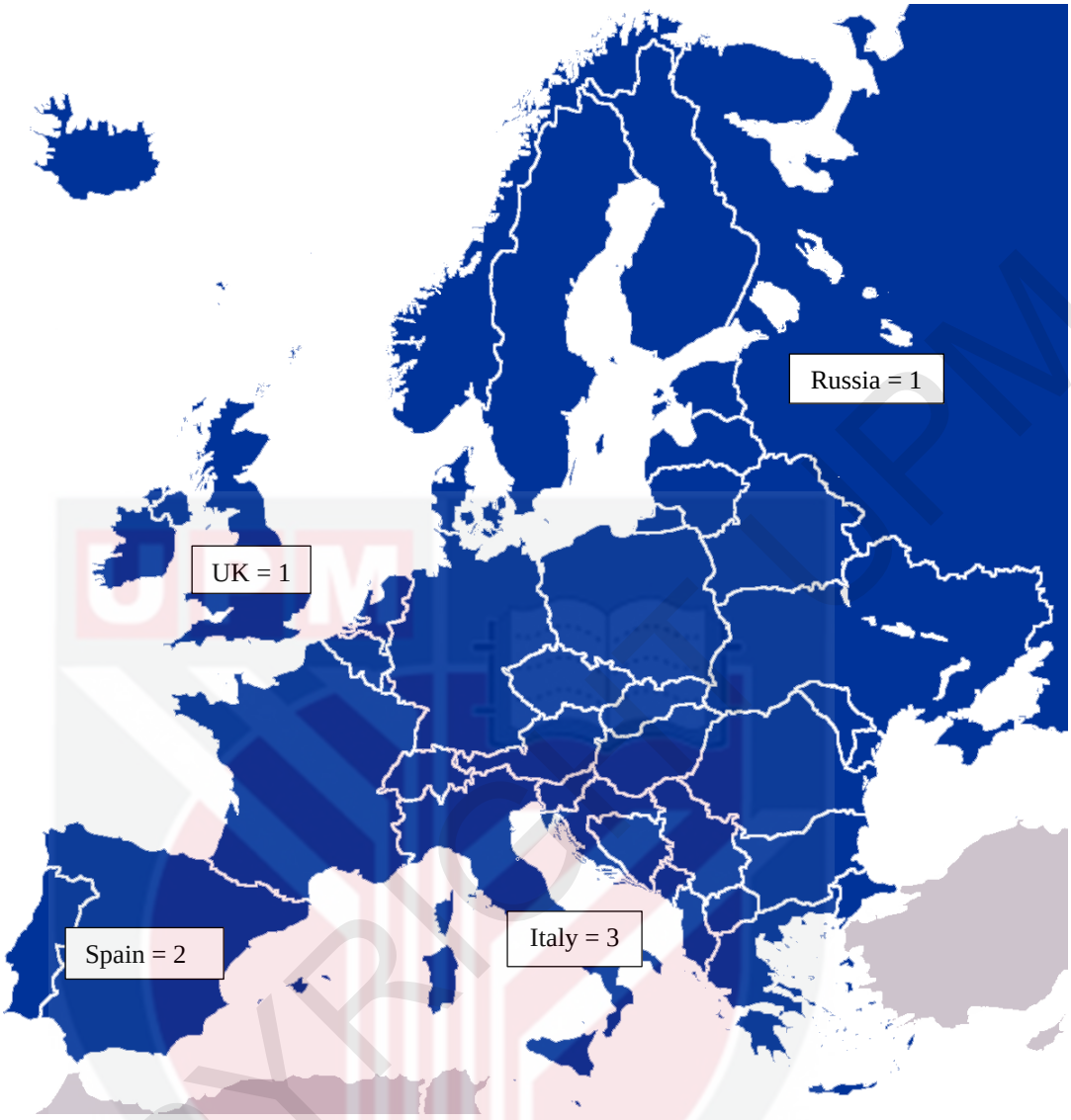

Figure 4.2. Map of Europe with the number of studies in each included country. UK = 1, Spain = 2, Italy = 3 and Russia = 1.

1	2	3	4	5	6	7	8
Author & Year	**Country**	**Gene**	**Types of variation**	**Number of variations**	**dbSNP / coding / nucleotide changes**	**Allele frequency**	**Effects caused by variation**
Zhou et al. (2021)	UK	TMPRSS2	SNP	1	rs7282236 (A/G)	Effect allele frequency (non/neg/pos) 0.751/0.742/0.777	The observed SNP is most likely to have an impact on COVID-19 test positivity which was observed in the increased risk of COVID-19 in the patient compared to the control groups. The clinical manifestation of the genes which immunodeficiency was associated with ACE2 gene while atypical inflammatory spondylopathies, non-infectious gastroenteritis, prostate cancer,

							symptoms involving head and neck and neoplasm of uncertain behaviour were associated with TMPRSS2 gene.
Benetti et al. (2020)	Italy	ACE2	Undetermined	1	c.2158A>G (p.Asn720Asp)	0.012	The effects of allelic variability on ACE2 conformation would slightly affecting the clinical differences between individuals and influencing the severity. This depicts that the susceptibility to viral entry and the disease development would be influenced either by the identified polymorphism
			Silent variant	1	c.2247G>A (p.Val749Val)	0.030	

							or by the cumulative effects of the variants.
Novelli et al. (2020)	Italy	ACE2	Intronic	1	rs2285666 (c.439+4G>A)	Cumulative frequency 0.2353	There is no significant difference found on the association of the variant with COVID-19 severity when compared between both asymptomatic vs mild-moderate-severe and severe vs mild-moderate COVID-19 patients even though the allele is more commonly found in the patient population.
			Missense	2	rs140312271 (c.1888G>C) (p.Asp630His)		The results are undetermined in human studies but the results of in silico study shows

							conflicting results which it predicts the variant of the protein to be benign (3) and pathogenic (2) in different type of tools used.
					rs41303171 (c.2158A>G) (p.Asn720Asp)		The variant is believed to not involved with SARS-CoV-2 S protein binding interaction as it is not located in the binding sites of the virus. In other SNPs, the allelic variant observed only in asymptomatic subject in heterozygous status. The results of the computational study gave a prediction of the variant

							to have a benign effect (4) and pathogenic effects (1) on the infection.
Gómez et al. (2020)	Spain	ACE2	Undetermined / intronic	1	rs2285666	Undetermined	Variant observed with a higher risk in hypertension and hypercholesterolemia patients without significant difference observed between mild and severe COVID-19 cases and no differences observed between COVID-19 patient and the control group. The carrier of ACE2 A allele is at higher risk of developing hypertension.

Torre-Fuentes et al. (2021)	Spain	ACE2	Synonymous	1	rs35803318 (c.2247G>A)	Minor allele frequency 0.038	Variant considered as non-pathogenic. Variant is unavailable in individual with infection and insufficient data available to support its possible protective role.
			Missense	1	rs41303171 (c.2158A>G) (Asn720Asp)	Minor allele frequency 0.016	The variant is less likely could influence viral infection as the codon 720 is not located at the spike protein binding site, however it is highly potential to do so as it is located near to the site of cleavage by TMPRSS2.
		TMPRSS2	Missense	3	rs75603675	Undetermined	The presence of variant is frequent among infected individual although the

							differences is not statistically significant and relates with its possibility in influencing the viral infection even though the pathogenicity is not yet determined.
					rs12329760	Undetermined	The variant might be related with the infection but in the study, it shows no association with infection and the variant is potentially pathogenic.
					rs200291871	Undetermined	The variant might be related with the infection but the in the study, it shows no association with infection.
			Synonymous	8	rs17854725	Undetermined	Undetermined effects.

					rs61735789	Undetermined	The variant is also associated with viral infection although the differences is not statistically significant.
					rs2298659 (c.888C>T)	Minor allele frequency 0.23	Undetermined effects.
					rs3787950	Undetermined	Undetermined effects.
					rs61735794 (c.300C>T)	Minor allele frequency 0.001	Undetermined effects but the variant was detected at significant different frequencies in group of individuals with and without infection.
					rs61735792 (c.300C>T)	Minor allele frequency 0.01	Undetermined effects but the variant was detected at significant different frequencies in group of

							individuals with and without infection.
					rs142750000	Undetermined	Undetermined effects.
					rs141788162	Undetermined	Undetermined effects.
Latini et al. (2020)	Italy	TMPRSS2	Missense	5	rs200291871 (c.22G>C) (p.Gly8Arg)	0.004	Undetermined effects.
					rs75603675 (c.23G>T) (p.Gly8Val)	0.36	Undetermined effects and the frequency of the variant observed in study population is lower than in the GnomAD databases.
					rs61735791 (c.193G>A) (p.Ala65Thr)	0.008	Undetermined effects.
					rs114363287 (c.331G>A) (p.Gly111Arg)	0.004	Undetermined effects and the frequency of the variant observed in study

							population is higher than in the GnomAD databases.
					rs12329760 (c.589G>A) (p.Val197Met)	0.17	Undetermined effects. The variant is located in the catalytic domain of the gene but far from the active site.
Shikov et al. (2020)	Russia	ACE2	SNP & large indels	54 & undetermined	rs879922 (C-G)	Undetermined	The total number of ACE2 eQTLs frequency shows no differences in patient with mild or severe COVID-19. There are certain differences of ACE2 eQTLs frequency in the population but it is less likely affects
					rs1514280 (A-G)		
					rs233575 (G-A)		
					rs4646174 (C-G)		
					rs4240157(C-T)		
					rs4646124 (T-C)		
					rs4646127 (A-G)		
					rs971249 (T-C)		
					rs233574 (T-C)		

					rs757066 (C-T) rs2158083 (C-T) rs4646152 (A-G) rs4646143 (T-C) rs2316903 (G-T) rs4646147 (T-A) rs2316904 (C-T) rs4646153 (C-T) rs4646156 (A-T) r2023802 (G-A) rs2048683 (T-G) rs2048684 (A-C) rs113691336: rs4646158 (C-CATAAG) rs1514279 (G-A) (CAA-C) rs2074192 (C-T)		COVID-19 susceptibility and severity. There are rare singleton variant detected in COVID-19 patient (9 severe, 1 mild).

					rs138373349: rs4646148 (T-TTTAA)		
					rs146122606: rs57823828: rs754565978 (A-AAAGGAAGG)		
					(TTA-T)		
					(TTG-T)		
					rs780722994 (AAC-A)		
					rs4646188 (A-G)		
					rs2285666 (C-T)		
					rs4646142 (G-C)		
					rs4646190 (A-G)		
					rs41303171 (T-C)		
					rs188991352 (T-C)		
					rs41297301 (G-T)		
					rs183546232 (A-G)		

					rs148408803 (C-T) rs137910448 (T-C) rs765716829 (G-A) rs191869625 (C-A) rs146598386 (G-T) rs73195521 (T-A) rs73195520 (C-T) rs72614598 (A-C) rs755766792 (TAAATC-T) rs72614596 (T-A) (A-G) (A-G) rs771204033 (G-A) (T-C) rs768948617 (A-T) rs35803318 (C-T)		

Table 1. Data extracted and analysed from all seven included studies. All studies were published on 2020. Five studies were on ACE2 and three studies were on TMPRSS2 polymorphisms.

4.4 Quality Assessment

The results of the quality assessment using the Newcastle-Ottawa Scale (NOS) for observational studies of case control design were tabulated in Table 2. The quality of all included studies is moderate, with the total score allocated varies from three to eight. Two studies were allocated with three stars, three studies with five stars and the remaining two studies were allocated with eight stars.

AUTHOR	CRITERIA								TOTAL
	SELECTION				COMPARABILITY	EXPOSURE			
	1	2	3	4	1	1	2	3	
Zhou et al. (2021)	*	*	*	*	**	*	*		8/9
Benetti et al. (2020)		*	*		*				3/9
Novelli et al. (2020)	*	*	*		*	*			5/9
Gómez et al. (2020)	*	*	*		*	*			5/9
Torre-Fuentes et al. (2021)	*	*	*	*	**	*	*		8/9
Latini et al. (2020)	*	*				*			3/9
Shikov et al. (2020)	*	*	*		*	*			5/9

Table 2. Quality assessment of included studies using Newcastle-Ottawa Scale (NOS) with a full mark of nine stars. Quality of studies with less than 3 score (< 3) are considered low, between 4 to 6 (4 – 6) are moderate and more than seven (> 7) are high.

CHAPTER 5

DISCUSSION

The transmembrane protein ACE2 and TMPRSS2 are the most well-known proteins that are involved with the mechanism of SARS-CoV-2 viral entry into the host cells. Many previous studies have mentioned the possibilities of the polymorphism of these transmembrane proteins to have an impact on the susceptibility to SARS-CoV-2 infection and there are many available studies have been done on the human genome databases to observe the association of both transmembrane protein polymorphisms with susceptibility to COVID-19.

Based on the previous studies, there is a possible increase in susceptibility to COVID-19 with a certain ACE2 and/or TMPRSS2 polymorphism. The findings were inconclusive as it is observed and tested *in silico* on the collection of genes that are available in the human genome databases and not on the human population; thus, leading to this systematic review which focuses on observing the association of ACE2 and/or TMPRSS2 polymorphism with susceptibility to COVID-19 in the human population that consists of the COVID-19 patient and healthy population. This is the strength of this study, which this is the first systematic review that identifies, analyses

and synthesizes the findings from previous studies that were done *in silico* on the human genome databases to determine the association of ACE2 and/or TMPRSS2 polymorphism with susceptibility to SARS-CoV-2 infections.

The systematic review began with searching and identifying eligible articles from the databases based on the predetermined inclusion and exclusion criteria which resulted in seven eligible articles. From the analysis of seven observational studies, there are mixed findings on the association of ACE2 and/or TMPRSS2 polymorphism with susceptibility to SARS-CoV-2 infection in the human population. It is determined that some polymorphism of the ACE2 gene has minimal association with COVID-19 susceptibility (Torre-Fuentes et al., 2021; Shikov et al., 2020) and in some studies, there is no association mentioned (Benetti et al., 2020; Novelli et al., 2020; Gómez et al.,2020). As for TMPRSS2 gene, there is positive association found in some studies (Zhou et al., 2021) while some did not show the association even though the presence of the variant is more common in COVID-19 patients compared to the control (Torre-Fuentes et al., 2021). There is also no association between TMPRSS2 polymorphism with susceptibility to SARS-CoV-2 mentioned in a study (Latini et al., 2020).

5.1 ACE2

Based on the findings, there are factors that were considered to have an effect on the association of the ACE2 and/or TMPRSS2 polymorphism with SARS-CoV-2 infections which is the type and location of the polymorphism. The presence of the silent or synonymous polymorphism, c.2247G>A (rs35803318), was noted to contribute to the control of gene expression at post-transcriptional level even though it was found having no effect on the protein produced. In another study, it was proposed that this variant has an association with infection, but study by Torre-Fuentes et al. (2020) deemed that it is actually non-pathogenic as supported by the findings of Benetti et al. (2020) that rs35803318 was frequently found in the controls compared to the SARS-CoV-2 infected patients.

The location of the variant in the ACE2 gene is noteworthy to know because it could have a potential impact on binding with SARS-CoV-2 and allows the infections to occur. The intronic variant rs2285666 is a variant that is located at the splice site region of ACE2 gene at the 3' end of the intron. This variant is the only variant that was thought to be able to affect the splicing process to generate the mRNA. Yet, there is no splicing alteration observed when tested *in silico* using Human Splicing Finder (HSF) (Novelli et al., 2020; Gómez et al., 2020). By no means, this indicates that this variant has no effect on susceptibility to SARS-CoV-2 infections. Even so, there might be polymorphisms in the non-coding region on the ACE2 gene that could have an

effect on regulating the gene expression and activity thus affecting the susceptibility to COVID-19 (Novelli et al., 2020).

The missense variant, rs140312271 (c.1888G>C), needs to be further analyzed in a larger study population to confirm its presence and further understand its functional role and impact on ACE2 protein alongside its possible connection to the susceptibility to SARS-CoV-2 infection. It is found that the wild type residue of this gene has undergone numerous mutations throughout evolution. Consequently, the *in silico* studies on the residues shows an irrelevant impact of this residue on the ACE2 protein function and structures following the studies using the ACE2 and its orthologous proteins' sequence alignment (Novelli et al., 2020).

Location of ACE2 gene on the X chromosome is significant to determine the effects of this gene with certain diseases. It could affect the frequency of identified variants in different individuals which in homozygous women, the allele frequency of the identified variant is very low (Benetti et al., 2020). X chromosome inactivation (XCI) is a condition in which the X chromosome is turned off and this condition only occurs in women as women have two X chromosomes. This condition could have affected the alternate expression of ACE2 allele in heterozygous individuals by partially or completely protecting them from the SARS-CoV-2 infection due to the absence of the binding-prone allele. Despite the fact that XCI could have affected the susceptibility to SARS-CoV-2, there is still no exact cause-effect relationship clarified (Benetti et al., 2020). Thus, further studies are needed to validate these findings in

order to understand the effect of XCI with susceptibility to infections alongside the disease development.

Overall, there are several types of polymorphism found in the ACE2 gene of the studied population such as silent or synonymous polymorphism, missense polymorphism, single nucleotide polymorphism, intronic and indels. There is not much association between the ACE2 genetic variants and susceptibility to SARS-CoV-2 infections found from the infection which was most likely due to the small group of studies in a certain population. Still, it is noteworthy that the presence of the variants might have an effect on the severity of the disease as perceived by Shikov et al. (2020) in their study in which they found that there was an increased inflammation in patients with COVID-19. Furthermore, receptor function of ACE2 might be affected by missense variants through the alteration of ACE2 receptor function (Shikov et al., 2020). These findings suggest that the COVID-19 susceptibility and disease severity can be affected by the variations in the ACE2 gene either directly or indirectly. Yet, ACE2 receptor function might not be associated with the effects of the variant on the ACE2 phenotype (Shikov et al., 2020).

Based on the findings, different allelic variations could affect ACE2 arrangement and structures and consequently contribute to the clinical differences between individuals and influence the severity of a disease. Also, the susceptibility to viral entry and the disease development would be influenced either by the identified

polymorphism or by the cumulative effects of the variants. Although most of the variants were frequently found in control groups compared to the patients, extensive research still needs to be done on determining the impact of ACE2 variants on the susceptibility to get SARS-CoV-2 infection. Besides, the risk of infections are affected by the common variants available and genetic variability of ACE2 receptors (Gómez et al., 2020). Therefore, it is determined that ACE2 polymorphism does not have remarkable effects on SARS-CoV-2 infections, instead it influences the severity of diseases in the patients.

5.2 TMPRSS2

The entry of SARS-CoV-2 is facilitated by the enzyme protease and TMPRSS2 protein is the main protein involved in this process. There are three studies done on determining the association of TMPRSS2 polymorphism with susceptibility to COVID-19 (Zhou et al., 2021; Torre-Fuentes et al., 2021; Latini et al., 2020). Based on the studies, it is found that the presence of certain polymorphisms in the gene was associated with COVID-19 test positivity regardless of the variant types. The variants that show association with viral infections are rs7282236, rs75603675 and rs61735789 and these variants were frequently found in the COVID-19 infected patients. Some other variants that might show correlation with SARS-CoV 2 infections are rs12329760, rs200291871, rs61735792 and rs61735794. These variants were found frequently in the COVID-19 patients compared to the control, even though not all variants were detected at significant differences (Zhou et al., 2021; Torre-Fuentes et

al., 2020). In the study by Torre-Fuentes et al. (2020), the variant rs75603675 was considered to have potential to cause overexpression of the protease enzyme in the individuals which could affect the proteolytic cleavage of the S2 protein and allow SARS-CoV-2 viral entry into the host cells. Interestingly, in the study by Latini et al. (2020), this variant was less frequently found in COVID-19 patients when compared to the European reference population obtained from the databases.

Among all variants found in these studies, rs12329760 (c.589G>A) which is located in the catalytic site of the gene could have an effect in the catalysing properties of the gene. Even so, this variant is mapped far from the active site that allows the binding and catalysation of the viral S protein by the TMPRSS2 enzyme (Latini et al., 2020). The viral S protein and ACE2 complex might be able to form a bond and orient themselves within the active site of the TMPRSS2 protein (Madhu, 2018); however, the catalytic action of the enzyme could not be determined to be associated with the distant variant. As the distant variant is still within the catalytic domain of the gene, it might slightly have influence on the catalytic properties of the TMPRSS2 gene, but not up to a noticeable level. This rs12329760 (c.589G>A) variant, which is located in the Scavenger Receptor Cysteine-Rich (SRCR) domain of TMPRSS2, could protect an individual from getting infected by SARS-CoV-2 as it might potentially alter the proteolytic activity of TMPRSS2 (Latini et al., 2020). In addition, even though there are plenty of polymorphisms found in the TMPRSS2 gene, the risk of infection due to these polymorphisms is still unknown.

5.3 Limitations

Firstly, the article searching and screening processes were done only in March 2021 and the articles searching was also done in only four databases, which might not cover as much available studies and latest related research if present. Another limitation is that the included studies were only from a few countries mainly from Europe such as the UK, Russia, Italy and Spain which the data obtained from a certain population thus limiting the findings to only the population involved. Hence, there is a need for future research done in a larger population, in other geographical areas and regions.

To add, there are some studies that are considered as substandard based on the quality assessment, which means that some studies might have less information and are less reliable. There are also studies that did not mention types of polymorphisms observed in their findings. Moreover, not all included studies linked the association of the transmembrane gene, ACE2 and/or TMPRSS2 polymorphisms with susceptibility to SAR-CoV-2 infection. Instead, some of the studies only determined the association of the polymorphism on the clinical manifestation in COVID-19 patients.

Therefore, more research is needed to confirm the association of ACE2 and/or TMPRSS2 polymorphism with susceptibility to SARS-CoV-2 infection, as the

findings from this study is still preliminary and considered not sufficient to conclude the exact link between these two variables. Inclusion studies for future research should also contain valuable information that could explain and determine the association of ACE2 and/or TMPRSS2 polymorphism with SARS-CoV-2 infections. In addition to the polymorphisms analysed in this review, the susceptibility to SARS-CoV-2 infections could be further explored and determined with relations to other proteins or genes that are involved with viral infections into the host cells.

CHAPTER 6

CONCLUSION

Overall, this systematic review has identified three TMPRSS2 polymorphisms, rs7282236, rs75603675 and rs61735789 that are associated with SARS-CoV-2 infections and some other polymorphisms such as rs12329760, rs200291871, rs61735792 and rs61735794 that could be linked to the viral infections. All of these polymorphisms are frequently found in the COVID-19 patients group compared to the control and it affects the rate of infection by altering the expression and catalytic properties of the enzyme TMPRSS2. Meanwhile, there is no ACE2 polymorphism found to be associated with SARS-CoV-2 infections but it is found to affect the disease severity in the infected individuals. However, these studies are limited to certain populations and small groups of subjects. Therefore, there is a need for further studies to be conducted in a large population worldwide to specifically determine the exact role and impact of the ACE2 and/or TMPRSS2 polymorphism with the susceptibility to SARS-CoV-2 infections.

REFERENCES

Anastassopoulou, C., Gkizarioti, Z., Patrinos, G. P., Tsakris, A. (2020). Human genetic factors associated with susceptibility to SARS-CoV-2 infection and COVID-19 disease severity. *Human Genomics, 14*(40). https://doi.org/10.1186/s40246-020-00290-4

Gopalakrishnan, S., & Ganeshkumar, P. (2013). Systematics review and meta-analysis: Understanding the best evidence in primary healthcare. *Journal of Family Medicine and Primary Healthcare, 2*(1), 9-14. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3894019/

Harrison, A. G., Lin, T., & Wang, P. (2020). Mechanisms of SARS-CoV 2 transmission and pathogenesis. *Trends in Immunology, 41*(12), 1100-1115. doi: 10.1016/j.it.2020.10.004.

Hemingway, P., & Brereton, N. (2009). What is a systematic review? *Evidence-based medicine* (2nd ed.). Retrieved from http://www.bandolier.org.uk/painres/download/whatis/Syst-review.pdf

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Kruger, N., Herrler, T., Erichsen, S., Schiergens, T. S., Herrler, G., Wu, N. H., Nitsche, A., Muller, M. A., Drosten, C., & Pohlmann, S. (2020). SARS-CoV 2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell, 181*, 271-280. https://doi.org/10.1016/j.cell.2020.02.052

Hou, Y., Zhao, J., Martin, W., Kallianpur, A., Chung, M. K., Jehi, L., Sharifi, N., Erzurum, A., Eng, C., & Cheng, F. (2020). New insights into genetic susceptibility of COVID-19: An ACE2 and TMPRSS2 polymorphism analysis. *BMC Medicine, 18* (216). https://doi.org/10.1186/s12916-020-01673-z

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration.

Journal of Clinical Epidemiology, 62, e1-e34. doi: 10.1016/j.jclinepi.2009.06.006

Liu, Y. C., Kuo, R. L., & Shih, S. R. (2020). COVID-19: The first documented coronavirus pandemic in history. *Biomed Journal, 43* (4), 328-333. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7199674/

LoPresti, M., Beck, D. B., Duggal, P., Cummings, D. A. T., & Solomon, B. D. (2020). The role of host genetic factors in coronavirus susceptibility: Review of animal and systematic review of human literature. *The American Journal of Human Genetics, 107,* 381-402. https://doi.org/10.1016/j.ajhg.2020.08.007.

Maria, E. D., Latini, A., Borgiani, P., & Novelli, G. (2020). Genetic variants of the human host influencing the coronavirus-associated phenotypes (SARS, MERS and COVID-19): rapid systematic review and field synopsis. *Human Genomics, 14,* 30. https://doi.org/10.1186/s40246-020-00280-6

Medinaz. (2020, March 27). *COVID-19: Corona virus (pathophysiology)* [Video]. YouTube. https://www.youtube.com/watch?v=tmjwd2bHaOs

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). *Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015 statement. Syst Rev*; 4(1):1. doi: 10.1186/2046-4053-4-1

Mollica, V., Rizzo, A., & Massari, F. (2020). The pivotal role of TMPRSS2 in coronavirus disease 2019 and prostate cancer. *Future Oncology*. https://doi.org/10.2217/fon-2020-0571

Muhammad Adnan, S., Suliman, K., Abeer, K., Nadia, B., & Rabeea, S. (2020). COVID-19
infection: Origin, transmission, and characteristics of human coronaviruses. *Journal of Advanced Research, 24,* 91-98. https://doi.org/10.1016/j.jare.2020.03.005

National Human Genome Research Institute (NIH). (n.d.). *Polymorphism*. https://www.genome.gov/genetics-glossary/Polymorphism

Ninja Nerd Lectures. (2020, March 17). *COVID-19 | Coronavirus: Epidemiology, pathophysiology, diagnostics* [Video]. YouTube. https://www.youtube.com/watch?v=PWzbArPgo-o

Parasher, A. (2020). COVID-19: Current understanding of its pathophysiology, clinical presentation and treatment. *Postgrad Medical Journal*. doi:10.1136/postgradmedj-2020-138577

Phillips, T. (2020). *Genetic polymorphism -Different does not mean mutated.* https://journals.lww.com/mejmedgen/Fulltext/2012/07000/Genetic_polymorphism_studies_in_humans.1.aspx

Ragia, G., & Manolopoulos, V. G. (2020). Inhibition of SARS-CoV 2 entry through the ACE2/TMPRSS2 pathway: a promising approach for uncovering early COVID-19 drug therapies. *European Journal of Clinical Pharmacology, 76*, 1623-1630. https://doi.org/10.1007/s00228-020-02963-4

Robertson, S. (2020). *Flexibility of glycosylated membrane-bound ACE2 suggests mechanical energy source in opening of SARS-CoV-2 spike.* https://www.news-medical.net/news/20200917/Flexibility-of-glycosylated-membrane-bound-ACE2-suggests-mechanical-energy-source-in-opening-of-SARS-CoV-2-spike.aspx

Senapati, S., Banerjee, P., Bhagavatula, S., Kushwaha, P. P., & Kumar, S. (2021). Contributions of human ACE2 and TMPRSS2 in determining host-pathogen interaction of COVID-19. *Journal of Genetics*, *100*, 12. https://doi.org/10.1007/s12041-021-01262-w

Shabir, O. (2020). *What is TMPRSS2?* https://www.news-medical.net/health/What-is-TMPRSS2.aspx

Shereen, M. A., Khan, S., Kazmi, A., Bashir, N., & Siddique, R. (2020). COVID-19 infection: Origin, transmission and characteristics of human coronaviruses. *Journal of Advanced Research, 24*, 91-98. https://doi.org/10.1016/j.jare.2020.03.005

Stopsack, K. H., Mucci, L. A., Antonarakis, E. S., Nelson, P. S., & Kantoff, P. W. (2020). TMPRSS2 and COVID-19: Serendipity or opportunity for intervention? *American Association for Cancer Research*. doi: 10.1158/2159-8290.CD-20-0451

Tawfik, G. M., Dila, K. A. S., Mohamed, M. Y. F., Tam, D. N. H., Kien, N. D., Ahmed, A. M., & Huy, N. T. (2019). A step by step guide for conducting a systematic review and meta-analysis with simulation data. *Trop Med Health, 47* (46). https://doi.org/10.1186/s41182-019-0165-6

Thomas, L. (2020). *A closer look at TMPRSS2: This could help treat COVID-19*. https://www.news-medical.net/news/20200528/A-closer-look-at-TMPRSS2-This-could-help-treat-COVID-19.aspx

Universitiy Libraries. (2021). *Systematic reviews & other review types*. https://guides.temple.edu/c.php?g=78618&p=4178713

Wang, Y., & He, W. (2018). Endogenous Mitochondrial Aldehyde Dehydrogenase-2 as an Antioxidant in Liver, *The Liver* (pp. 247–259). Retrieved from doi:10.1016/B978-0-12-803951-9.00021-5

Wells, G. A., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., & Tugwell, P. (n.a.). *The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised study in meta-analyses.* http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

World Health Organization. (2021). Coronavirus disease (COVID-19) pandemic. Retrieved from https://www.who.int/emergencies/diseases/novel-coronavirus-2019

World Health Organization (WHO). (n.d.). *WHO coronavirus (COVID-19) dashboard*. https://covid19.who.int/

Xiao, L., Sakagami, H., & Miwa, N. (2020). ACE2: The key molecule for understanding the pathophysiology of severe and critical conditions of COVID-19: Demon or angel? *Viruses, 12*(491). doi:10.3390/v12050491

Xiong, J., Lipsitz, O., Nasri, F., Lui, L. M. W., Gill, H., Phan, L., Chen-Li, D., Iacobucci, M., Ho, R., Majeed, A., & McIntyre, R. S. (2020). *Impact of COVID-19 pandemic on mental health in general population: A systematic review*. *277*, 55-64. https://www.sciencedirect.com/science/article/pii/S0165032720325891?casa_token=BUpJnf3yyYMAAAAA:21_eiv1oEp4CcXusvJXcXkILPgg1PHkP-qCH4ABj-NuUFNaCrCf59dNMGQP3y8VWlT7kOq_6swY

Yesudhas, D., Srivastava, Ambuj, & Gromiha, M. M. (2020). COVID-19 Outbreak: history, mechanism, transmission, structural studies and therapeutics. *Infection.* https://doi.org/10.1007/s15010-020-01516-2

Zhang, X. M., Jiao, J., Cao, J., Huo, X. P., Zhu, C., Wu, X. J., & Xie, X. H. (2021). Frailty as a predictor of mortality among patients with COVID-19: a systematic review and meta-analysis. *BMC Geriatr, 21, 186.* https://doi.org/10.1186/s12877-021-02138-5

APPENDICES

Appendix A: Registration Form of International Prospective Register of Systematic Review (PROSPERO)

PROSPERO is a database for systematic reviews, rapid reviews and umbrella reviews. PROSPERO accept registrations of reviews mentioned before and it enable researchers to check whether any similar review was already existed so that wastage of effort could be avoided.

ID	Title	Status	Last edited
CRD42021229963	A systematic review on association of ACE2 and/or TMPRSS2 polymorphism with susceptibility to SARS-CoV-2 Infection	Registered	22/04/2021

Figure 1. Status of this systematic review and the ID recorded in the PROSPERO.

Systematic review

To edit the record click *Start an update* below. This will create a new version of the record - the existing version will remain unchanged.

1. * Review title.

Give the title of the review in English

A systematic review on association of ACE2 and/or TMPRSS2 polymorphism with susceptibility to SARS-CoV-2 Infection

2. Original language title.

For reviews in languages other than English, give the title in the original language. This will be displayed with the English language title.

3. * Anticipated or actual start date.

Give the date the systematic review started or is expected to start.

01/02/2021

4. * Anticipated completion date.

Give the date by which the review is expected to be completed.

30/06/2021

5. * Stage of review at time of this submission.

Tick the boxes to show which review tasks have been started and which have been completed.
Update this field each time any amendments are made to a published record.

Reviews that have started data extraction (at the time of initial submission) are not eligible for inclusion in PROSPERO.
If there is later evidence that incorrect status and/or completion date has been supplied, the published PROSPERO record will be marked as retracted.

This field uses answers to initial screening questions. It cannot be edited until after registration.

The review has not yet started: No

Review stage	Started	Completed
Preliminary searches	Yes	No
Piloting of the study selection process	No	No
Formal screening of search results against eligibility criteria	No	No
Data extraction	No	No
Risk of bias (quality) assessment	No	No
Data analysis	No	No

Provide any other relevant information about the stage of the review here.

Review has not yet started. Searching for articles in databases.

Figure 2. Page 1 of PROSPERO Registration Form.

Review has not yet started. Searching for articles in databases.

6. * Named contact.

The named contact is the guarantor for the accuracy of the information in the register record. This may be any member of the review team.

Nurul Husna Hasnul Hadi

Email salutation (e.g. "Dr Smith" or "Joanne") for correspondence:
Ms Nurul Husna

7. * Named contact email.

Give the electronic email address of the named contact.

192182@student.upm.edu.my

8. Named contact address

PLEASE NOTE this information will be published in the PROSPERO record so please do not enter private information, i.e. personal home address

Give the full institutional/organisational postal address for the named contact.

9. Named contact phone number.

Give the telephone number for the named contact, including international dialling code.

+60176857638

10. * Organisational affiliation of the review.

Full title of the organisational affiliations for this review and website address if available. This field may be completed as 'None' if the review is not affiliated to any organisation.

Universiti Putra Malaysia

Organisation web address:

11. * Review team members and their organisational affiliations.

Give the personal details and the organisational affiliations of each member of the review team. Affiliation refers to groups or organisations to which review team members belong.
NOTE: email and country now MUST be entered for each person, unless you are amending a published record.

Miss Nurul Husna Hasnul Hadi. Universiti Putra Malaysia
Miss Aina Nadheera Abd Rahman. Universiti Putra Malaysia
Miss Umaiya Muzaffar. Universiti Putra Malaysia
Assistant/Associate Professor Syahril Abdullah. Universiti Putra Malaysia

12. * Funding sources/sponsors.

Details of the individuals, organizations, groups, companies or other legal entities who have funded or sponsored the review.

None

Grant number(s)
State the funder, grant or award number and the date of award

Figure 3. Page 2 of PROSPERO Registration Form.

13. * Conflicts of interest.

List actual or perceived conflicts of interest (financial or academic).

None

14. Collaborators.

Give the name and affiliation of any individuals or organisations who are working on the review but who are not listed as review team members. **NOTE: email and country must be completed for each person, unless you are amending a published record.**

15. * Review question.

State the review question(s) clearly and precisely. It may be appropriate to break very broad questions down into a series of related more specific questions. Questions may be framed or refined using PI(E)COS or similar where relevant.

Is the susceptibility to SARS-CoV-2 infection in individuals related to the polymorphisms on ACE2 and/or TMPRSS2 gene?

16. * Searches.

State the sources that will be searched (e.g. Medline). Give the search dates, and any restrictions (e.g. language or publication date). Do NOT enter the full search strategy (it may be provided as a link or attachment below.)

Article searching from four databases such as PubMed, ScienceDirect, Scopus & EBSCOhost by using the combination of search terms such as ((((ACE2)) OR (Angiotensin-converting enzyme 2)) OR (hACE2)) OR (ACEH)) AND ((((Coronavirus disease 2019) OR (COVID-19)) OR (SARS-CoV-2 infection)) OR (2019 novel coronavirus)),((((Genetic polymorphism) OR (allelic variation)) OR (genetic variation)) OR (genetic predisposition)) OR (nucleotide variation)) AND ((((Coronavirus disease 2019) OR (COVID-19)) OR (SARS-CoV-2 infection)) OR (2019 novel coronavirus)) and (((((TMPRSS2) OR (transmembrane serine protease 2)) OR (Serine protease 10)) OR (Transmembrane protease serine 2 non-catalytic chain)) OR (Transmembrane protease serine 2 catalytic chain)) AND ((((Coronavirus disease 2019) OR (COVID-19)) OR (SARS-CoV-2 infection)) OR (2019 novel coronavirus)). Boolean connectors 'OR' are used within categories and 'AND' between each categories. This combination will be slightly modified according to the requirement of each databases. The search terms also include the alternative name for ACE2 and TMPRSS2. Publication date are restricted from 2019 until the searching being done and the language is restricted to English papers only.

17. URL to search strategy.

Upload a file with your search strategy, or an example of a search strategy for a specific database, (including the keywords) in pdf or word format. In doing so you are consenting to the file being made publicly accessible.

Or provide a URL or link to the strategy. Do NOT provide links to your search **results**.

Do not make this file publicly available until the review is complete

18. * Condition or domain being studied.

Give a short description of the disease, condition or healthcare domain being studied in your systematic review.

Susceptibility to SARS-CoV-2 infections in individuals with ACE2 and/or TMPRSS2 polymorphism

19. * Participants/population.

Specify the participants or populations being studied in the review. The preferred format includes details of both inclusion and exclusion criteria.

Patients diagnosed with COVID-19.

Inclusion criteria:
Age, gender & race: All included
Study mention the method used to determine the ACE2 and/or TMPRSS2 polymorphism

Figure 4. Page 3 of PROSPERO Registration Form.

Exclusion criteria:
Unclear or insufficient data
Irrelevant data

20. * Intervention(s), exposure(s).

Give full and clear descriptions or definitions of the interventions or the exposures to be reviewed. The preferred format includes details of both inclusion and exclusion criteria.

Determine the ACE2 and/or TMPRSS2 polymorphism in patients with SARS-CoV-2 infection.

Inclusion criteria:
Studies mention the ACE2 and/or TMPRSS2 polymorphism
Studies reported the relationship of SARS-CoV-2 infection with ACE2 and/or TMPRSS2 polymorphism

Exclusion criteria:
Irrelevant data

21. * Comparator(s)/control.

Where relevant, give details of the alternatives against which the intervention/exposure will be compared (e.g. another intervention or a non-exposed control group). The preferred format includes details of both inclusion and exclusion criteria.

Healthy subjects without SARS-CoV-2 infection.

22. * Types of study to be included.

Give details of the study designs (e.g. RCT) that are eligible for inclusion in the review. The preferred format includes both inclusion and exclusion criteria. If there are no restrictions on the types of study, this should be stated.

All type of studies with any research design will be included.

Inclusion criteria:
Countries: All included studies
Year of publication: 2019 - current

Exclusion criteria:
Secondary research paper (systematic review, meta analysis, review paper)
Letter to editor
Medical hypothesis
Duplicate publication
Studies using other than English language

23. Context.

Give summary details of the setting or other relevant characteristics, which help define the inclusion or exclusion criteria.

24. * Main outcome(s).

Give the pre-specified main (most important) outcomes of the review, including details of how the outcome is defined and measured and when these measurement are made, if these are part of the review inclusion criteria.

The relationship between ACE2 and/or TMPRSS2 polymorphism with SARS-CoV-2 infection.

Measures of effect

25. * Additional outcome(s).

List the pre-specified additional outcomes of the review, with a similar level of detail to that required for main outcomes. Where there are no additional outcomes please state 'None' or 'Not applicable' as appropriate to the review

None

Figure 5. Page 4 of PROSPERO Registration Form.

Measures of effect

26. * Data extraction (selection and coding).

Describe how studies will be selected for inclusion. State what data will be extracted or obtained. State how this will be done and recorded.

Article Searching

vailable article obtained from all four databases will be collected and organized into a table in the Microsoft Excel Spreadsheet. The duplicates of the same title and author within the same databases and between databases will be removed.

The remaining search results will be screened by its title and abstract by two reviewer independently, according to the eligibility criteria. The articles that does not fit with the criteria will be removed and the remaining articles will undergo full text screening by both reviewer, independently.

The results from each screening will be cross-checked and will be discussed if there is disagreement between reviewer. There will be involvement of third party if no consensus reached during the discussion.

Data Extraction

The first reviewer will extract the data from the eligible paper into the Microsoft Excel Spreadsheet and the extracted data will be checked by the second reviewer for any addition or correction.

The data that will be extracted consist of:

- Title
- Author name & year of publication
- Link of article
- Location of the studies
- Gene involved
- Number of variation
- Types of variation
- Position/ nucleotide changes/ protein
- Allele frequency
- Effects caused by the variation

27. * Risk of bias (quality) assessment.

State which characteristics of the studies will be assessed and/or any formal risk of bias/quality assessment tools that will be used.

The assessment will be done using the Newscastle Ottawa Scale (NOS).

28. * Strategy for data synthesis.

Describe the methods you plan to use to synthesise data. This **must not be generic text** but should be **specific to your review** and describe how the proposed approach will be applied to your data.
If meta-analysis is planned, describe the models to be used, methods to explore statistical heterogeneity, and software package to be used.

Tabulate the results for included study and summarise the content of each study. Synthesize the results from all included studies to reach consensus on the relationship between the variable and the expected outcome.

29. * Analysis of subgroups or subsets.

State any planned investigation of 'subgroups'. Be clear and specific about which type of study or participant will be included in each group or covariate investigated. State the planned analytic approach.

None

30. * Type and method of review.

Figure 6. Page 5 of PROSPERO Registration Form.

Select the type of review, review method and health area from the lists below.

Type of review

Cost effectiveness	No
Diagnostic	Yes
Epidemiologic	Yes
Individual patient data (IPD) meta-analysis	No
Intervention	No
Living systematic review	No
Meta-analysis	No
Methodology	No
Narrative synthesis	Yes
Network meta-analysis	No
Pre-clinical	No
Prevention	No
Prognostic	No
Prospective meta-analysis (PMA)	No
Review of reviews	No
Service delivery	No
Synthesis of qualitative studies	No
Systematic review	Yes
Other	No

Health area of the review

Alcohol/substance misuse/abuse	No
Blood and immune system	No
Cancer	No
Cardiovascular	No
Care of the elderly	No
Child health	No
Complementary therapies	No
COVID-19	Yes

Figure 7. Page 6 of PROSPERO Registration Form.

For COVID-19 registrations please tick all categories that apply. Doing so will enable your record to appear in area-specific searches

Chinese medicine

Diagnosis

Epidemiological

Genetics

Health impacts

Immunity

Long COVID

Mental health

PPE

Prognosis

Public health intervention

Rehabilitation

Service delivery

Transmission

Treatments

Vaccines

Other

Crime and justice	No
Dental	No
Digestive system	No
Ear, nose and throat	No
Education	No
Endocrine and metabolic disorders	No
Eye disorders	No
General interest	No
Genetics	Yes

Figure 8. Page 7 of PROSPERO Registration Form.

Health inequalities/health equity	No
Infections and infestations	Yes
International development	No
Mental health and behavioural conditions	No
Musculoskeletal	No
Neurological	No
Nursing	No
Obstetrics and gynaecology	No
Oral health	No
Palliative care	No
Perioperative care	No
Physiotherapy	No
Pregnancy and childbirth	No
Public health (including social determinants of health)	No
Rehabilitation	No
Respiratory disorders	Yes
Service delivery	No
Skin disorders	No
Social care	No
Surgery	No
Tropical Medicine	No
Urological	No
Wounds, injuries and accidents	No
Violence and abuse	No

31. Language.

Select each language individually to add it to the list below, use the bin icon to remove any added in error.

English

There is an English language summary.

32. * Country.

Select the country in which the review is being carried out. For multi-national collaborations select all the countries involved.

Figure 9. Page 8 of PROSPERO Registration Form.

Malaysia

33. Other registration details.

Name any other organisation where the systematic review title or protocol is registered (e.g. Campbell, or The Joanna Briggs Institute) together with any unique identification number assigned by them.
If extracted data will be stored and made available through a repository such as the Systematic Review Data Repository (SRDR), details and a link should be included here. If none, leave blank.

34. Reference and/or URL for published protocol.

If the protocol for this review is published provide details (authors, title and journal details, preferably in Vancouver format)

No I do not make this file publicly available until the review is complete

35. Dissemination plans.

Do you intend to publish the review on completion?

Yes

36. Keywords.

Give words or phrases that best describe the review. Separate keywords with a semicolon or new line. Keywords help PROSPERO users find your review (keywords do not appear in the public record but are included in searches). Be as specific and precise as possible. Avoid acronyms and abbreviations unless these are in wide use.

ACE2
TMPRSS2
polymorphism
SARS-CoV-2 infection
COVID-19

37. Details of any existing review of the same topic by the same authors.

If you are registering an update of an existing review give details of the earlier versions and include a full bibliographic reference, if available.

38. * Current review status.

Update review status when the review is completed and when it is published.
New registrations must be ongoing so this field is not editable for initial submission.

Review_Ongoing

39. Any additional information.

Provide any other information relevant to the registration of this review.

40. Details of final report/publication(s) or preprints if available.

Leave empty until publication details are available OR you have a link to a preprint (NOTE: this field is not editable for initial submission).
List authors, title and journal details preferably in Vancouver format.

Figure 10. Page 9 of PROSPERO Registration Form.