International Journal of Genetics
ISSN: 0975-2862 & E-ISSN: 0975-9158, Volume 14, Issue 4, 2022, pp.-850-855.
Available online at https://bioinfopublication.org/pages/jouarchive.php?id=BPJ0000226

Review Article

THE ASSOCIATION BETWEEN *TMPRSS2* GENETIC POLYMORPHISMS AND THE SUSCEPTIBILITY AND SEVERITY OF COVID-19: A SYSTEMATIC REVIEW

SILVA C.S.[1,3*], SILVA M.J.A.[2,3], LIMA K.V.B.[1,2,3], FROTA C.C.[4], SARDINHA D.M.[1,3], LIMA L.N.G.C.[1,2,3]

[1]Postgraduate Program in Parasitic Biology in the Amazon (PPGBPA), University of State of Pará (UEPA), Brazil
[2]Program in Epidemiology and Health Surveillance (PPGEVS) of the Evandro Chagas Institute (IEC), Ananindeua, Pará, Brazil
[3]Bacteriology and Mycology Section of the Evandro Chagas Institute (IEC), Ananindeua, Pará, Brazil
[4]Department of Pathology and Legal Medicine, Faculty of Medicine, Federal University of Ceará, Fortaleza, CE, Brazil
*Corresponding Author: Email - karolinysoares2303@gmail.com

Received: October 01, 2022; Revised: October 26, 2022; Accepted: October 27, 2022; Published: October 30, 2022

Abstract: COVID-19, caused by the SARS-CoV-2 virus, is a highly transmissible disease that has a variety of symptoms. The presence of genetic polymorphisms in genes as ACE-2 and *TMPRSS2* is directly associated with the susceptibility and severity of COVID-19. The objective of this work is to analyze which polymorphisms in the *TMPRSS2* gene are associated with the progression of COVID-19. We identified 35 SNPs associated with disease progression. The high expression of *TMPRSS2* in the lungs was associated with the presence of polymorphisms such as rs383510, rs469390 and rs464397. rs12329760 was the polymorphism most studied, where the presence of the T allele was related to protection against COVID-19. The SNPs found play an important role in determining the prognosis of the disease.

Keywords: *COVID-19, SARS-CoV-2¸ Polymorphism, TMPRSS2, Expression*

Citation: Silva C.S., *et al*., (2022) The Association Between *TMPRSS2* Genetic Polymorphisms and The Susceptibility and Severity of Covid-19: A Systematic Review. International Journal of Genetics, ISSN: 0975-2862 & E-ISSN: 0975-9158, Volume 14, Issue 4, pp.- 850-855.

Academic Editor / Reviewer: I S Chakrapani, Prof Dr Suleyman Cylek

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the pathogen responsible for causing coronavirus disease 2019 (COVID-19). This virus was discovered in 2019 in Wuhan, China, and showed high infectivity, spreading rapidly throughout the world, causing a global pandemic [1].

SARS-CoV-2 spreads primarily through respiratory droplets when people talk, cough, sneeze, or through direct contact with the virus on contaminated surfaces [2]. The pathogen mainly infects the respiratory system, where infected individuals may present several clinical symptoms, such as coughing, fever, dyspnea, skin manifestations, and muscle pain [3]. Approximately 15% of people with COVID-19 develop the severe form of the disease, which can progress to acute respiratory distress syndrome (ARDS), severe pneumonia, kidney damage, and death [4]. The severity of COVID-19 appears to be affected by several risk factors, such as advanced age, cardiovascular disease, obesity, and diabetes [5].

Coronaviruses (CoVs) belong to the order Nidovirales, family Coronoviridae, and subfamily Orthocoronavinae, being subdivided into four genera: Alphacoronavirus, Betacoronavirus, Gammacoronavirus and Deltacoronavirus [6]. SARS-CoV-2 belongs to the group of β -coronaviruses with enveloped, single-stranded, positive-sense RNA [7]. The complete genome of SARS-CoV-2 is approximately 29.9 kb, with a GC content of 38% and 12 open reading frames (ORFs) [8]. CoVs have a spherical outline, with a diameter of the virion ranging from 60 to 140 nm and have characteristic structural proteins: spike (S), membrane (M), envelope (E) and nucleocapsid (N) [9]. In addition, they have 9 to 12 nm spicules that project onto the viral surface, having the appearance of a crown [10].

The entry of SARS-CoV-2 into cells occurs through the angiotensin converting enzyme 2 (ACE2), which is expressed in several organs, such as the lungs, intestine, heart, and kidneys [11]. Binding occurs by the S-protein binding domain of the virus that attaches tightly to human ACE2. Target cell proteases activate the S protein, causing cleavage in the S1 and S2 subunits, allowing fusion between viral and cellular membranes [12].

SARS-CoV-2 uses transmembrane serine protease 2 (*TMPRSS2*) as the enzyme responsible for cleaving and activating the S protein in the fusion process of the human membrane [13].

TMPRSS2 is a well-studied protein because it is involved in the development of prostate cancer due to the overexpression of specific transcription factors transforming erythroblasts (ETS), such as ERG [14]. The *TMPRSS2* gene, responsible for encoding *TMPRSS2*, is related to pathological and physiological processes, such as inflammatory responses, invasion of tumor cell, and fertility. and pain [15]. The *TMPRSS2* gene is found on chromosome 21q22.3, has 15 exons, and an open reading frame of 492 amino acids. This gene harbors androgen-responsive elements in its 5' UTR [16]. Testosterone and dihydrotestosterone are responsible for gene regulation through stimulation of androgen receptors [17].

TMPRSS2 has been shown to be crucial for the entry of SARS-CoV-2 into cells. From this, several studies sought to investigate the hypothesis that genetic variability and expression of the *TMPRSS2* gene are associated with susceptibility to COVID-19 [18].

Host genetic polymorphisms are known to play an important role in determining susceptibility or resistance to viral infections [19]. The host genes play an important role in the entry and replication of SARS-CoV-2 in cells and in the immune response, where the combination of genes may be associated with the pathogenesis of COVID-19 [20]. Studies have shown that gene polymorphisms involved in viral entry into human cells, such as ACE2 and *TMPRSS2*, are associated with susceptibility and severity to disease [21, 22].

Knowing the importance of polymorphisms in genes involved in the entry of SARS-CoV-2 into cells and in determining the susceptibility and severity of COVID-19, the following question arises from the research described here: 'Which SNPs in the *TMPRSS2* gene are associated with susceptibility, aggravation, or protection against COVID-19?'

Materials and Methods

Study design

This is a systematic review of the literature that aims to collect evidence on the reported correlations between the *TMPRSS2* gene polymorphisms and the progression of COVID-19. The review was conducted in accordance with the Preferred Reporting items for Systematic Reviews and Meta-Analyses Protocols (PRISMA Protocol) guidelines. The study followed the following training steps: 1) Elaboration of the guiding question; 2) Stipulation of inclusion and exclusion criteria; 3) Choice of articles; 4) Analysis of articles; 5) Interpretation, discussion, and presentation of the review. The PICO strategy (Population, Intervention, Comparison, Outcome) was used, following the following characteristics: Population: patients with COVID-19; Intervention: Assess the SNPs of the *TMPRSS2* gene in COVID-19 patients; Comparison: COVID-19 and the described SNPs of the *TMPRSS2* gene; Outcome: protection, susceptibility, or severity of COVID-19; Outcome: progression of COVID-19. From this, the following question was generated: What existing SNPs in the *TMPRSS2* gene are associated with susceptibility/worsening or protection against COVID-19?

Data sources

The terms used in the search based on Medical Subject Headings (MeSH) were: "COVID-19"; "SARS-CoV-2"; "*TMPRSS2*" and "Polymorphism". The articles were searched using the Boolean operator 'AND', in the following databases: National Library of Medicine National Institutes of Health of the USA (PubMed), Latin American and Caribbean Literature in Health Sciences (LILACS), Web of Science, and Scientific Electronic Library Online (SCIELO).

Study eligibility criteria

The selected studies were published from the beginning of the publications until June 2022, available complete, clinical studies, comparative studies, cross-sectional studies, case-control studies, cohort studies (prospective and retrospective), in vivo, in vitro trials, and review article narratives. Articles that identified SNPs of the *TMPRSS2* gene associated with susceptibility or progression of cases of COVID-19 were selected. As exclusion criteria, systematic review articles, abstracts, letters to the editor, and those that had subjects not relevant to the research were not included.

Data collection and extraction

Data collection was conducted in July 2022. Two researchers read the articles and independently extracted data from the publications following a predefined protocol. Information such as title, used method (case-control or cohort study), database, and relevant results were collected. The collected data were transferred to a Microsoft Office Excel spreadsheet.

Results

The search strategy resulted in 137 articles, and 81 were excluded since they were duplicates or were not relevant to the investigation. The abstracts of each of the remaining articles were evaluated and 17 were also excluded. At the end, full articles were evaluated and 21 met the inclusion criteria and therefore were added to the search [Fig-1]. The results showing the included studies are summarized in [Table-1]. Thus, 35 SNPs associated with susceptibility, symptom development, and severity of patients infected with SARS-CoV-2 were identified.

Discussion

COVID-19 was initially a disease associated with respiratory symptoms. Subsequently, it was found that the symptoms and severity of the cases can vary, where some patients may have mild symptoms (cough, headache, and fever) or develop acute respiratory illness, which can cause the death of these individuals [44]. The mechanisms involved in the susceptibility and evolution of COVID-19 cases are still unclear; however, the influence of host genetic polymorphisms on the immune response and its association with the severity of the disease are already known [45]. Studies have shown that regulatory genetic variants influence the expression of the *TMPRSS2* gene and change susceptibility to COVID-19, as well as worsening cases [24,45,47].

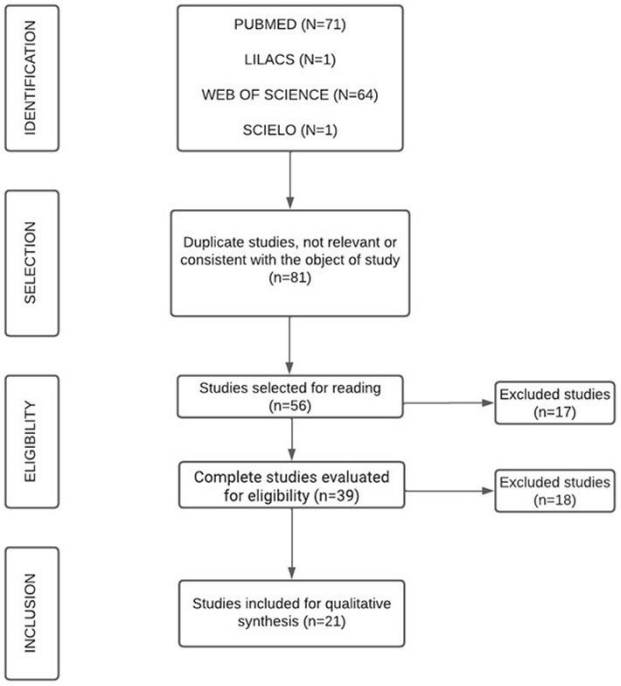

Fig-1 Flowchart of procedures for identification, selection, eligibility, and inclusion of studies for analysis. Belém, PA, Brazil (2022).

This review identified several *TMPRSS2* polymorphisms that are associated with the progression of COVID-19 cases. From the analysis of the selected articles, 35 SNPs were identified that influence the susceptibility, protection, and aggravation of COVID-19 cases. The rs12329760 SNP is described as a substitution of cytosine (C) for thymine (T). This polymorphism was analyzed in several studies, and it was found in most studies that the presence of the mutant T allele is considered protective against the infection and progression of COVID-19, while the wild C allele was associated with a greater number of hospitalizations of patients with COVID-19 [24, 26, 30, 31, 33, 34, 35, 37, 39, 42]. Computational analysis of changes caused by rs12329760 (C>T) in *TMPRSS2* structure shows a change from valine to methionine at position 197 (V197M), helping to form a pocket protein in the structure of *TMPRSS2*, which disfavors its binding to the SARS-CoV-2 S protein [35]. The importance of this SNP is evidenced in the analysis of its frequency in different populations and the advance of COVID-19 in these places. In a study conducted on Italian patients by Monticelli *et al.* (2021), a low frequency of the rs12329760 allele T was detected in this population, where it was suggested that the low frequency of this allele in the SNP had an influence on the high impact of the first wave of the pandemic in Italy [31].

Asselta *et al.* (2020) also analyzed the association of genetic polymorphisms and the severity of COVID-19 among Italians. The SNP rs12329760 and two distinct haplotypes were identified, showing differences in mutation frequency between Italians and East Asians. The presence of rare alleles in both haplotypes was associated with the induction of higher levels of *TMPRSS2*. The first haplotype (SNPs rs463727, rs34624090, rs55964536, rs734056, rs4290734, rs34783969, rs11702475, rs35899679, and rs395041537) was seen to be androgen regulated, which could help explain the severity of COVID-1 virus cases in the Italian population. SNPs belonging to the second haplotype (rs2070788, rs9974589, rs7364083) were correlated with a higher susceptibility to viral infections such as influenza A (H1N1) [24].

Different studies have investigated the presence of other SNPs in populations and analyzed whether the presence of these variants was involved in the susceptibility and worsening of COVID-19 cases [26, 27, 31, 32, 38, 41]. The SNP rs2070788 is characterized by the substitution of a guanine (G) for an adenine (A) and has been investigated in African, American, European, and Asian populations. The frequency of the G allele was highly detected in Americans and Europeans, compared to the frequency of Asians and Africans, and may be associated with high mortality among American and European populations.

Table-1 *Characteristics of the studies included in the systematic review*

Kind of study	SNP	Results	Reference
Case control	rs75603675; rs61735794; rs61735792	The SNP rs75603675 (G>T) was the most frequent variant among those infected with SARS-CoV-2. The SNP rs61735794 (A allele) and rs61735792 (T allele) were also significantly detected in patients with COVID-19.	[23]
Descriptive study	rs12329760 (p.Val160Met) ; Haplotype 1: rs463727, rs34624090, rs55964536, rs734056, rs4290734, rs34783969, rs11702475, rs35899679 and rs35041537; Haplotype 2: rs2070788, rs9974589, rs7364083	SNP rs12329760 (p.Val160Met) was detected at a higher frequency in East Asians than in Europeans. The rare alleles of these haplotypes found, associated with the induction of higher levels of TMPRSS2, are more frequent in the Italian population than in the East Asian population.	[24]
Descriptive study	SNP rs2298659	SNP rs 2298659 (C allele) showed a strong correlation with the fatality rate of COVID-19.	[25]
Cohort	rs7560367; rs200291871, rs12329270	A lower frequency of the variant allele of SNPs rs7560367 (T allele), rs12329270 (A allele) and rs200291871 (C allele) was observed in Italian COVID patients compared to the allele frequency of these SNPs in the European population.	[26]
Descriptive study	SNP rs469390, rs2070788, rs383510 and rs464397	The four SNPs were associated with differential expression of TMPRSS2.The A allele and the AA genotype of SNP rs469390 showed higher gene expression in the lung. The T alleles of rs464397 and rs383510, associated with increased expression of the *TMPRSS2* gene in the lungs, were exhibited with a lower frequency in East Asian populations. The frequency of the G allele for rs2070788 was also lower in the East Asian population compared to the European, African and American population.	[27]
Revision	rs2070788, rs383510	The GG genotype of rs2070788 increases the expression of *TMPRSS2* in the lungs. Regarding to rs383510, the T allele is associated with increased expression of *TMPRSS2* in the lungs.	[28]
Descriptive study	rs12329760	The evaluation of the structural stability of the protein predicted that rs12329760 (T allele) is responsible for destabilizing TMPRSS2, thus compromising the binding affinity of TMPRSS2 to the spike protein of SARS-COV-2 and ACE-2.	[29]
Cross-sectional study	rs12329760	Patients with the SNP rs12329760 CC genotype of the had an association with symptomatic or severe COVID-19; the C allele was associated with a higher viral load and a higher number of deaths.	[30]
Descriptive study	rs12329760	SNP rs12329760 was detected at lower values for East Asians, Finns and Africans. The association of the low-frequency T allele of rs12329760 in the population with the high impact of the first wave of the epidemic in Italy has been suggested.	[31]
Revision	rs713400, rs112657409, rs119110678, rs77675406	The frequency of the T allele at rs713400 influences the expression of TMPRSS2 and is considered higher among Asians than among Europeans, Africans and South Asians.	[32]
Descriptive study	rs781089181, rs570454392, rs867186402, rs12329760, rs118518290, rs762108701	The six variants analyzed were identified decreasing the stability of *TMPRSS2*. The SNPs rs781089181, rs570454392, and rs867186402 were in a highly conserved region and were responsible for crucial changes in the protein. Analyzing the native structure of the protein, 5 disordered regions were formed due to the rs12329760 and rs118518290 variants.	[33]
Descriptive study	rs12329760	The variant rs12329760 (T allele) was predicted to be deleterious and harmful to the expression of *TMPRSS2*, decreasing the stability of the protein.	[34]
***In silico* / descriptive study**	rs12329760 and rs75603675	The SNP rs12329760 (T allele) was considered deleterious by 3 tools used. A new major pocket protein caused by rs12329760 (T allele) was predicted, affecting the structure of TMPRSS2, and affecting its role in the entry of SARS-COV-2. SNP rs75603675 (A allele) has demonstrated the ability to increase protein disorder and influence TMPRSS2 function in facilitating SARS-COV-2 entry.	[35]
Case-control	rs383510	Carriers of the CC genotype in the rs383510 SNP had a 1.73-fold increased risk of infection for SARS-CoV-2.	[36]
Revision	rs12329760	It was observed that the number of cases and high mortality is associated with a low prevalence of the rs12329760 SNP (T allele), this SNP being significantly associated with genetic inactivation of TMPRSS2. From this, the presence of the T allele in rs12329760 is considered a protective factor against infection and disease progression.	[37]
Descriptive study	rs2070788	Significant positive association between the G allele of SNP rs2070788 and the mortality rate of COVID-19 among populations with a high frequency of this variant, for example, the Indian population. PNS eQTL analysis shows that the GG genotype tends to influence significantly higher expression of the *TMPRSS2 gene* in the lung, increasing vulnerability to COVID-19. This variant was found at a high frequency in Americans and at a lower frequency in Africans and East Asians. This may be associated with high mortality among the American population and low severity among Asians.	[38]
Case-control	rs12329760, rs17854725,rs75603675	The T allele of rs12329760 conferred an increased risk of the individual being infected with SARS-CoV-2. The TT genotype was associated with a higher incidence of the severe form of the disease and the TC genotype was associated with an increase in the lesions seen on computed tomography. The presence of the G allele in SNP rs17854725 (A>G) was associated with greater susceptibility, while deaths were more frequent in carriers of the AG genotype. The AA genotype of rs75603675 appears to reduce the risk of a severe form of COVID-19 compared to the CC genotype. The GG genotype of rs4303795 increases the severe form of the disease and the occurrence of lesions.	[39]
Case-control	rs75603675	The presence of snp rs756036675 (T allele) was significantly associated with increased susceptibility to COVID-19 or increased risk of severe disease.	[40]
Descriptive study/ and systematic review	rs2070788, rs383510	The frequency of the G risk allele of the *TMPRSS2* rs2070788 gene was lowest among Africans and highest among Americans. The SNP rs383510 (T allele), responsible for also increasing the expression of *TMPRSS2*, was more frequent in the European population, while the African population was also the least frequent.	[41]
Descriptive study	r s3787946, rs9983330, rs12329760, rs2298661 and rs9985159	The SNPs rs3787946 (C allele), rs9983330 (G allele), rs12329760 (T allele), rs2298661 (A allele) and rs9985159 (T allele) were less recurrent in hospitalized patients, suggesting their protective role. Furthermore, 4 SNPs associated with the reduction of TMPRSS2 in lung tissues were identified: rs9983330 (G allele), rs12329760 (T allele), rs2298661 (A allele) and rs9985159 (T allele).	[42]
Cross-sectional study	rs75603675, rs12329760	The variant rs12329760 had no effect on the severity of the case. Analysis of SNP rs75603675 showed that the CC or CA genotypes compared to AA were associated with more severe COVID-19.	[43]

This is due to the influence of the G allele and the GG genotype on increased expression of *TMPRSS2* in the lungs [27, 38, 41]. The intervention of this genotype was confirmed through the analysis of quantitative expression traits loci (eQTLs) that reported the increase in the expression of *TMPRSS2* in the lungs and its influence on the increase in the expression of MX1, a gene involved in the viral response [28, 48]. Cheng *et al.* (2015) reported that this genotype was associated with a 2x greater risk of developing Influenza A in severe form, due to its ability to increase the expression of TPMRSS2 in the lungs, facilitating viral entry and thus increasing the susceptibility of the individual [49].

SNPs rs713400, rs383510, and rs753675 were also analyzed in the same ethnic groups mentioned above. The SNP rs713400 (C>T) was investigated in a single study, which showed that the presence of the T allele directly interfered with *TMPRSS2* expression. This mutation was identified with high frequency in the East Asian population, a region where the overall case fatality rate (CFR) is significantly lower (3.6%) compared to the global CFR (6.9%), which may suggest a protective character of that allele over the population [32]. In this same study, it was identified that the SNPs rs112657409, rs119110678 and rs77675406 also had a significant influence on the expression of *TMPRSS2*. The authors did not show which variants were involved in the infection, so further studies are needed for full clarification.

The SNP rs383510 (T>C) is located within a functional region with potentiating activity in the expression of *TMPRSS2*, where the presence of the T allele exhibited a higher transcriptional level than the C allele [47]. Database analysis of eQLTs showed that the TT genotype increases lung tissue gene expression [28]. The T allele was detected less frequently in the Asian population and was frequently identified in residents of European countries that had a high mortality rate of approximately 15% during the first wave of the pandemic [27, 41]. The study conducted by Ihram *et al*. (2020) analyzed allele frequencies in European, African, American, East Asian, and Southeast Asian populations. Again, the T allele of rs383510 was associated with increased expression of the *TMPRSS2* gene. However, contrary to the others, Schönfelder *et al*. (2021) found a significant 1.73-fold increase in the risk of SARS-CoV-2 infection in carriers of the CC genotype [36]. SNP rs469390 (G>A) was identified as a missense mutation and is located within an exonic region. The study by Irham *et al*. (2020) detected the influence of the A allele and the AA genotype of the rs469390 SNP with the highest gene expression in the lung. The presence of the heterozygous AG genotype was associated with intermediate expression and the GG genotype with a lower expression, making it evident that the presence of the mutant A allele makes the individual more susceptible to COVID-19. In this study, all SNP alleles related to the highest expression of *TMPRSS2* in the lungs, including the wild-type T allele of rs464397 (T>C), were detected with low frequency in the East Asian population, indicating a possible protection for these individuals [27].

In a study conducted in Spain, rs75603675 (G>T/ p.Gly8Val) was the most frequent SNP of the *TMPRSS2* gene among patients infected with COVID-19 [23]. In Italian population, a lower frequency of the allele variant allele T was observed, compared to data obtained from the Europeans [26]. The presence of the T allele is associated with a decrease in severe cases of COVID-19 [39, 43]. This substitution of G>T may lead to a decrease in the ability of *TMPRSS2* to bind directly to the S protein of SARS-CoV-2. This is due to the substitution of glycine for the long-side-chain hydrophobic valine, which causes a decrease in the functional activity of the proteinase and a reduction in the flexibility of the peptide due to the increase in hydrophobicity. However, the results obtained by Minashkin *et al*., 2020, were discordant, indicating that individuals with the T allele are more susceptible to SARS-CoV-2 infection and have the highest risk of developing the severe form [40]. SNPS rs2298659, rs200291871, rs61735794, and rs61735792 were polymorphisms that were also associated with susceptibility, the development of the severe form of COVID-19, and the appearance of deaths caused by the disease. The study by Kim & Jeong (2021) was the only work that reported a strong correlation of the SNP rs2298659 (C, wild-type allele) with the increased case-fatality rate of COVID-19 [25]. Torre-Fuentes *et al*. (2021) identified that SNPs rs61735794 (A allele) and rs61735792 (T allele) were significantly associated with cases of SARS-CoV infection [23]. Analysis of the rs200291871 SNP in Italian patients with COVID showed the presence of risk allele C with a higher frequency than in other European countries [26].

Andolfo *et al*. (2021) analyzed the polymorphisms of chromosome 21 and found five SNPs within the *TMPRSS2* gene associated with protection against severe COVID-19, they are: rs3787946 (C allele, wild type), rs9983330 (G allele, wild type), rs12329760 (T allele), rs2298661 (A allele, wild type) and rs9985159 (T allele, wild type), and the last four also were significantly correlated with lower expression of *TMPRSS2* in lung tissues. Wild-type alleles of these SNPs were less detected in inpatients compared to healthy controls, indicating a protective role for these variants against disease progression [42].

Rokni *et al*. (2022) analyzed SNPs rs17859725 and rs4303795 in 288 hospitalized patients with COVID-19. The SNP rs17859725 (A>G, Ile256Ile) is a variant synonymous with the exchange of two isoleucine codons at position 256 [34]. In this SNP, the wild-type G allele was identified as associated with a marked increase in the risk of COVID-19 infection and the presence of the AG genotype was more frequent in patients who died [39]. rs4303795 is a functional SNP and is located 2 kb from the promoter region of the *TMPRSS2* gene. The results of allele frequencies detected in this SNP, the G allele, were predominantly found in hospitalized patients with COVID-19. Furthermore, the GG genotype was directly associated with increased development of severe form, as well as the occurrence of lung lesions [39].

Saih *et al*. (2021) used bioinformatics tools to understand the effects of mutations on the *TMPRSS2* protein. A total of six SNPs (rs781089181, rs570454392, rs867186402, rs12329760, rs118518290, rs762108701) were identified by the tools used as responsible for decreasing the stability of *TMPRSS2*, thus being considered harmful variants of the gene. The SNPs rs781089181, rs570454392, and rs867186402 proved to be important polymorphisms due to their location in crucial regions for gene function. The analysis of the native protein and its comparison with the mutant protein showed the formation of 5 disordered regions due the variants rs12329760 and rs118518290, which are responsible for altering the function of *TMPRSS2* [33].

From the results obtained, it is possible to have an overview of which and how the SNPs studied so far are related to the protection, susceptibility, and aggravation of COVID-19 cases. Analysis of these SNPs can help to understand the increase in cases and deaths in several countries where these studies were conducted and provides insights to future research in other countries where studies with SNPs have not yet been conducted. Additionally, the identification of these SNPs helps direct studies aimed at the development of therapeutic interventions that can help people with increased susceptibility or worse progression of the disease.

Conclusion

The *TMPRSS2* gene demonstrated several SNPs associated with the severity and clinical fate of patients infected with SARS-CoV-2. The SNP rs12329760 was the polymorphism most researched by several studies that associated the presence of the T allele with protection against COVID-19. In some SNPs, such as rs383510 and rs75603675, there was disagreement between the studies on which allele would be associated with susceptibility and/or severity of the disease. This divergence makes evident the need for further research to specify which allele is related to the worsening of cases. Furthermore, new studies that analyze *TMPRSS2* polymorphisms can improve the characterization of individuals with greater susceptibility, in the development of pharmaceutical therapies, and in the conception of new vaccines against COVID-19.

Application of research: Characterization of SNPs associated with COVID-19 severity. Determination of SNPs present in different populations and the evolution of cases.

Research Category: Immunogenetics; Infectious Diseases.

Abbreviations: SARS-CoV-2 - severe acute respiratory syndrome coronavirus 2; COVID-19 – Coronavirus disease 2019; CoVs-Coronaviruses; ARDS-acute respiratory distress syndrome; *TMPRSS2* -transmembrane serine protease 2; ETS-transcription factors transforming erythroblasts; CFR-case fatality rate CFR; eQTLs - quantitative expression traits loci; PubMed-National Library of Medicine National Institutes of Health of the USA; LILACS-Latin American and Caribbean Literature in Health Sciences; SCIELO-Scientific Electronic Library Online.

Acknowledgement / Funding: Authors are thankful to Postgraduate Program in Parasitic Biology in the Amazon (PPGBPA), University of State of Pará (UEPA), Brazil

****Principal Investigator or Chairperson of research: Caroliny Soares Silva**
University: University of State of Pará (UEPA), Brazil
Research project name or number: Review study

Author Contributions: CSS was responsible for the conceptualization, formal analysis, data curation, methodology, investigation, validation, visualization, roles/writing – original draft, editing. MJAS and DMS were responsible for the supervision, visualization, writing review and editing. KVBL and CCF were responsible for the supervision, visualization, writing review and editing. LNGCL performed the conceptualization, investigation, project administration, supervision, visualization, writing review and editing. All authors read and approved the final manuscript

Author statement: All authors read, reviewed, agreed and approved the final manuscript. Note-All authors agreed that- Written informed consent was obtained from all participants prior to publish / enrolment

Study area / Sample Collection: University of State of Pará (UEPA), Brazil

Cultivar / Variety / Breed name: Nil

Conflict of Interest: None declared

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.
Ethical Committee Approval Number: Nil

References

[1] De Vito A., Fiore V., Princic E., Geremia N., Panu Napodano C.M., Muredda A.A., Maida I., Madeddu G., Babudieri S. (2021) *PLoS One*, 16(3), 1-14.

[2] Guo Y.R., Cao Q.D., Hong Z.S., Tan Y.Y., Chen S.D., Jin H.J., Tan K.S., Wang D.Y., & Yan Y. (2020) *Military Medical Research*, 7(1), 1-10.

[3] Li J., Wang Y., Liu, Y., *et al*. (2022) *European Journal of Medical Research*, 27(1), 1-10.

[4] Dieter C., Brondani L.A., Leitão C.B., Gerchman F., Lemos N.E., Crispim D. (2022) *Plos One*, 17(7), 1-23.

[5] Lim S., Bae J.H., Kwon H.S. *et al*. (2021) *Nature Reviews Endocrinology*, 17, 11-30.

[6] Dhama K., Khan S., Tiwari R., Sircar S., Bhat S., Malik Y.S., Singh K.P., Chaicumpa W., Bonilla-Aldana D.K., & Rodriguez-Morales A.J. (2020) *Clin Microbiol Rev*, 33(4), 20-28.

[7] Astuti I. & Ysrafil. (2020) *Clinical Research and Reviews*, 14(4), 407-412.

[8] Rahman S., Shishir M. A., Hosen M. I., Khan M. J., Arefin A., & Khandaker A. M. (2022) *Gene Reports*, 27, 1-10.

[9] Glotov O. S., Chernov A. N., Scherbak S. G., & Baranov V. S. (2021) *Russian Journal of Genetics*, 57 (8), 878-892.

[10] Zhu N., Zhang D., Wang W., Li X., Yang B., Song J., Zhao X., Huang B., Shi W., Lu R., Niu P., Zhan F., Ma X., Wang D., Xu W., Wu G., Gao G.F., Tan W. (2020) *New England Journal of Medicine*, 382(8), 727-733.

[11] Fang L., Karakiulakis G., Roth M. (2020) *Lancet Respir Med*., 8(4)

[12] [12] Hoffmann M., Kleine-Weber H., Schroeder S., Krüger N., Herrler T., Erichsen S., Schiergens T.S., Herrler G., Wu N.H., Nitsche A., Müller M.A., Drosten C., & Pöhlmann S. (2020) *Cell*, 181(2), 271-280.

[13] Lima L.N.G.C., Sousa M.S., Lima K.V.B. (2020) *J. Health Biol. Sci*, 8, 1.

[14] Zarubin A., Stepanov V., Markov A., Kolesnikov N., Marusin A., Khitrinskaya I., Swarovskaya M., Litvinov S., Ekomasova N., Dzhaubermezov M., Maksimova N., Sukhomyasova A., Shtygasheva O., Khusnutdinova E., Radzhabov M., Kharkov V. (2021) *Genes*, 12 (1), 1-16.

[15] Lam D.K., Dang D., Flynn A.N., Hardt M., Schmidt B.L. (2015) *Pain*, 156(5), 923-930.

[16] Thunders M. & Delahunt B. (2020) *Journal of Clinical Pathology*, 73(12), 773-776.

[17] Mollica V., Rizzo A., Massari F. (2020) *Future Oncology*, 16(27), 2029-2033

[18] Di Maria E., Latini A., Borgiani P., & Novelli G. (2020) *Human Genomics*, 14(1), 1-14.

[19] Ramos-Lopez O., Daimiel L., Ramírez de Molina A., Martínez-Urbistondo D., Vargas J.A., Martínez J.A. (2020) *International Journal of Genomics*, 1-8.

[20] Debnath M., Banerjee M., & Berk M. (2020) *FASEB Journal*, 34(7), 8787-8795.

[21] Teng S. & Tang Q. (2020) *Computational and Structural Biotechnology Journal*, 18, 2100-2106.

[22] Darbani B. (2020) *International Journal of Environmental Research and Public Health*, 17(10), 3433.

[23] Torre-Fuentes L., Matías-Guiu J., Hernández-Lorenzo L., Montero-Escribano, P., Pytel, V., Porta-Etessam, J., Gómez-Pinedo, U., & Matías-Guiu, J. A. (2021) Journal of Medical Virology, 93(2), 863-869.

[24] Asselta R., Paraboschi E.M., Mantovani A. & Duga S. (2020) *Aging*, 12(11), 10087-10098.

[25] Kim Y.C., & Jeong B.H. (2021) *Genes*, 12(1), 1-9.

[26] Latini A., Agolini E., Novelli A., Borgiani P., Giannini R., Gravina P., Smarrazzo A., Dauri M., Andreoni M., Rogliani P., Bernardini S., Helmer-Citterich M., Biancolella M., Novelli G. (2020) *Genes*, 11(9), 1-8.

[27] Irham L.M., Chou W.H., Calkins M.J., Adikusuma W., Hsieh S.L., Chang W.C. (2020) *Biochemical and Biophysical Research Communications*, 529(2), 263-269.

[28] Piva F., Sabanovic B., Cecati M., Giulietti M. (2021) *Eur J Clin Microbiol Infect Dis*., 40(2), 451-455.

[29] Jeon S., Blazyte A., Yoon C., Ryu H., Jeon Y., Bhak Y., Bolser D., Manica A., Shin E.S., Cho Y.S., Kim B.C., Ryoo N., Choi H. & Bhak J. (2021) *Molecules and Cells*, 44(9), 680-687.

[30] Wulandari L., Hamidah B., Pakpahan C., Damayanti N.S., Kurniati N.D., Adiatmaja C.O., Wigianita M.R., Soedarsono, Husada D., Tinduh D., Prakoeswa C., Endaryanto A., Puspaningsih N., Mori Y., Lusida M.I., Shimizu K., & Oceandy D. (2020) *Human Genomics*, 15 (1), 1-9.

[31] Monticelli M., Hay Mele B., Benetti E., Fallerini C., Baldassarri M., Furini S., Frullanti E., Mari F., Andreotti G., Cubellis M.V., Renieri A. (2021) *Genes*, 12(4), 1-15.

[32] Senapati S., Kumar S., Singh A.K., Banerjee P., & Bhagavatula S. (2020) *Journal of Genetics*, 99(1), 1-5.

[33] Saih A., Bouqdayr M., Baba H., Hamdi S., Moussamih S., Bennani H., Saile R., Wakrim L., & Kettani A. (2021) *Biomed Res Int,* 9982729, 1-17.

[34] Vargas-Alarcón G., Posadas-Sánchez R., & Ramírez-Bello J. (2020) *Life Sci.,* 260 (118313), 1-13.

[35] Paniri A., Hosseini M.M., & Akhavan-Niaki H. (2020) *Journal of Biomolecular Structure and Dynamics*, 1-18.

[36] Schönfelder K., Breuckmann, K., Elsner C., Dittmer U., Fistera D., Herbstreit F., Risse J., Schmidt K., Sutharsan S., Taube C., Jöckel K.H., Siffert W., Kribben A., & Möhlendick B. (2021) *Frontiers in Genetics*, 12, 667231.

[37] Monticelli M., Mele B.H., Andreotti G., Cubellis M.V., & Riccio G. (2021) *European Journal of Medical Genetics*, 64 (6), 104227.

[38] Pandey R.K., Srivastava A., Singh P.P., & Chaubey G. (2022) *Infection, Genetics and Evolution*, 98, 105206.

[39] Rokni M., Heidari Nia M., Sarhadi M., Mirinejad S., Sargazi S., Moudi M., Saravani R., Rahdar S., & Kargar M. (2022) *Applied Biochemistry and Biotechnology*, 194(8), 3507-3526

[40] Minashkin M.M., Grigortsevich N.Y., Kamaeva A.S., Barzanova V.V., Traspov A.A., Godkov M.A., Ageev F.A., Petrikov S.S., & Pozdnyakova N.V. (2022) *Biomedicines*, 10 (3), 549.

[41] Smatti M.K., Al-Sarraj Y.A., Albagha O., & Yassine H.M. (2020) *Frontiers in Genetics*, 11, 578523.

[42] Andolfo I., Russo R., Lasorsa V.A., Cantalupo S., Rosato B.E., Bonfiglio F., Frisso G., Abete P., Cassese G.M., Servillo G., Esposito G., Gentile I., Piscopo C., Villani R., Fiorentino G., Cerino P., Buonerba C., Pierri B., Zollo M., Iolascon A. (2021) *Science*, 2021; 24 (4), 102322.

[43] Villapalos-García G., Zubiaur P., Rivas-Durán R., Campos-Norte P., Arévalo-Román C., Fernández-Rico M., García-Fraile Fraile L., Fernández-Campos P., Soria-Chacartegui P., Fernández de Córdoba-Oñate S., Delgado-Wicke P., Fernández-Ruiz E., González-Álvaro I., Sanz J., Abad-Santos F. & de Los Santos I. (2022) *Life Science Alliance*, 5 (10)

[44] Huang C., Wang Y., Li X., Ren L., Zhao J., Hu Y., Zhang L., Fan G., Xu J., Gu X., Cheng Z., Yu T., Xia J., Wei Y., Wu W., Xie X., Yin W., Li H., Liu M., Xiao, Y. (2020) *The Lancet*, 395 (10223), 497-506.

[45] Ovsyannikova I.G., Haralambieva I.H., Crooke S.N., Poland G.A., & Kennedy R.B. (2020) *Immunological Reviews*, 296(1), 205-219.

[46] Senapati S., Banerjee P., Bhagavatula S., Kushwaha P.P., & Kumar S. (2021) *Journal of Genetics*, 100(1), 12.

[47] Beyerstedt S., Casaro E. B., & Rangel É. B. (2021) *European journal of clinical microbiology & infectious diseases, official publication of the European Society of Clinical Microbiology*, 40(5), 905-919.

[48] Cheng Z., Zhou J., To K.K., Chu H., Li C., Wang D., Yang D., Zheng S., Hao K., Bossé Y., Obeidat M., Brandsma C.A., Song Y.Q., Chen Y., Zheng B.J., Li L., & Yuen K.Y. (2015) *Journal of Infectious Diseases*, 212 (8), 1214-1221.

[49] Bizzotto J., Sanchis P., Abbate M., Lage-Vickers S., Lavignolle R., Toro A., Olszevicki S., Sabater A., Cascardo F., Vazquez E., Cotignola J., & Gueron G. (2020) *Science*, 23(10), 101585.