**PRIMARY RESEARCH**                                                                                          **Open Access**

# Initial study on TMPRSS2 p.Val160Met genetic variant in COVID-19 patients

Laksmi Wulandari[1†], Berliana Hamidah[2†], Cennikon Pakpahan[2,3†], Nevy Shinta Damayanti[4], Neneng Dewi Kurniati[5], Christophorus Oetama Adiatmaja[6,7], Monica Rizky Wigianita[6], Soedarsono[1], Dominicus Husada[8], Damayanti Tinduh[9], Cita Rosita Sigit Prakoeswa[10], Anang Endaryanto[8], Ni Nyoman Tri Puspaningsih[11,12], Yasuko Mori[13], Maria Inge Lusida[14,15], Kazufumi Shimizu[13,16] and Delvac Oceandy[17*]

## Abstract

**Background:** Coronavirus disease 2019 (COVID-19) is a global health problem that causes millions of deaths worldwide. The clinical manifestation of COVID-19 widely varies from asymptomatic infection to severe pneumonia and systemic inflammatory disease. It is thought that host genetic variability may affect the host's response to the virus infection and thus cause severity of the disease. The SARS-CoV-2 virus requires interaction with its receptor complex in the host cells before infection. The transmembrane protease serine 2 (TMPRSS2) has been identified as one of the key molecules involved in SARS-CoV-2 virus receptor binding and cell invasion. Therefore, in this study, we investigated the correlation between a genetic variant within the human *TMPRSS2* gene and COVID-19 severity and viral load.

**Results:** We genotyped 95 patients with COVID-19 hospitalised in Dr Soetomo General Hospital and Indrapura Field Hospital (Surabaya, Indonesia) for the TMPRSS2 p.Val160Met polymorphism. Polymorphism was detected using a TaqMan assay. We then analysed the association between the presence of the genetic variant and disease severity and viral load. We did not observe any correlation between the presence of TMPRSS2 genetic variant and the severity of the disease. However, we identified a significant association between the p.Val160Met polymorphism and the SARS-CoV-2 viral load, as estimated by the Ct value of the diagnostic nucleic acid amplification test. Furthermore, we observed a trend of association between the presence of the C allele and the mortality rate in patients with severe COVID-19.

**Conclusion:** Our data indicate a possible association between TMPRSS2 p.Val160Met polymorphism and SARS-CoV-2 infectivity and the outcome of COVID-19.

**Keywords:** COVID-19, TMPRSS2, Polymorphism

---

* Correspondence: delvac.oceandy@manchester.ac.uk
†Laksmi Wulandari, Berliana Hamidah and Cennikon Pakpahan contributed equally to this work.
[17]Division of Cardiovascular Sciences, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK
Full list of author information is available at the end of the article

## Background

Coronavirus disease 2019 (COVID-19) is the biggest pandemic in the twenty-first century so far. Since the declaration of the pandemic by the World Health Organization (WHO), more than 110 million cases with more than 2.4 million deaths worldwide have been recorded as per mid-February 2021 [1]. COVID-19 is caused by an infection with the SARS-CoV-2 virus, which typically infects cells in the respiratory tract. The clinical presentations of COVID-19 range widely from asymptomatic infection to lethal pneumonia. It is known that three major factors, i.e. age, gender and the presence of underlying diseases, play a major role in determining COVID-19 severity [2–4]. However, it is not clear whether genetic variability contributes significantly to the clinical outcomes of COVID-19 patients.

One important factor that may play a crucial role in determining COVID-19 severity is the interaction between the virus and the host cells. SARS-CoV-2 infects the host cells by binding with its receptor on the surface of the host cell membrane. The main receptor for the SARS-coronavirus family is the angiotensin-converting enzyme 2 (ACE2) [5]. It is known that the spike (S) protein of SARS-CoV-2 mediates the binding of the virus to the ACE2 protein [6]. Considering the importance of virus receptor binding during the infection, it is logical to hypothesise that genetic variations within the gene encoding ACE2 may be associated with the degree of infection and hence the severity of the disease. Surprisingly, studies have reported no correlation between the genetic variations in the human *ACE2* gene and the severity of COVID-19 [7], as well as the previous severe acute respiratory syndrome (SARS) [8].

In addition to ACE2, several other molecules are also involved in the process of SARS-CoV-2 virus entry. For example, transmembrane protease serine 2 (TMPRSS2) [9] and neuropilin-1 (NRP1) [10] have been identified as co-receptors for SARS-CoV-2 that play a crucial role during virus entry. These molecules are important in mediating virus entry; for example, TMPRSS2 is known to facilitate the cleavage of the S protein, enabling membrane fusion and endocytic entry of the virus particles [11]. This has prompted us to hypothesise that genetic variability within the *TMPRSS2* gene may play a role in determining SARS-CoV-2 infection.

A recent analysis based on computational modelling suggested that out of more than eleven thousand single nucleotide polymorphisms (SNPs) within the human *TMPRSS2* gene (dbSNP, NCBI) only 21 SNPs with minor allele frequency (MAF) between 0.01 and 0.95 were predicted to affect the function of the protein [12]. Of these 21 SNPs, only two SNPs are missense variants (rs12329760 and rs75603675). The rs12329760 polymorphism (also known as p.Val160Met variant) has been

shown in several studies to play a role in mediating the risk for prostate cancer, confirming clinical consequences of this genetic variant [13–16]. Therefore, in this study, we focused on studying the association between the p.Val160Met variant of the TMPRSS2 gene and the severity, viral load and clinical outcomes of COVID-19 patients. Although we did not find any correlation between the p.Val160Met polymorphism and disease severity, we observed a possible association between the TMPRSS2 pVal160Met variant and the viral load in COVID-19 patients.

## Results

### Characteristics of patients

Characteristics of COVID-19 patients included in this study are described in Table 1. Age distributions were significantly different between patients with asymptomatic and mild versus moderate and severe COVID-19. There was a significant difference in the sex distribution with a higher proportion of male patients in the symptomatic groups. Significant differences were observed in the proportions of patients with underlying diseases between the asymptomatic and mild patients versus moderate and severe groups. As expected, the patients in the asymptomatic and mild COVID-19 groups displayed a significantly lower frequency of underlying diseases compared with the moderate and severe groups, including diabetes (*P* value < 0.001), cardiovascular disease (*P* value = 0.009) and liver disease (*P* value = 0.007).

### TMPRSS2 p.Val160Met polymorphism and COVID-19 severity

The TMPRSS2 p.Val160Met polymorphism (rs12329760) was successfully detected in all patients. The genotype and allele frequencies of this SNP are shown in Additional file 1. We observed a deviation of the allele frequency from Hardy–Weinberg equilibrium ($\chi^2$ = 6.72, *P* value = 0.035). However, the frequency of C allele (61.6%) and T allele (38.4%) in this study population seemed to be comparable with the reported frequency in the Asian population (dbSNP, NCBI) (Additional file 1).

The cross-tab analysis for the genotype and severity groups indicated no significant difference in the distribution of TMPRSS2 p.Val160Met polymorphism among the four groups of patients (Table 2). There were higher odds ratios in subjects with CC and CT genotypes to get symptomatic or more severe COVID-19 than those with TT genotype although they did not reach statistical significance (Table 3).

### TMPRSS2 p.Val160Met polymorphism and viral load

Next, we analysed the association between polymorphism and the viral load. All of the patients had positive results of the nucleic acid amplification testing (NAAT)

**Table 1** Demographic and baseline characteristics

| Variables | Asymptomatic (N=21) | Mild (N=12) | Moderate (N=32) | Severe (N=30) | All patients (N=95) | P value chi-square test (unless otherwise stated) |
|---|---|---|---|---|---|---|
| Age (years) | 33.9 ± 2.4[*] | 35.6 ± 2.7[*] | 52.3 ± 2.1 | 48.8 ± 1.5 | 44.7 ± 1.3 | [*]$P<0.001$ versus moderate and severe groups (ANOVA) |
| Gender (%) | | | | | | |
| Male | 8 (38.1%) | 10 (83.3%) | 19 (59.4%) | 23 (76.7%) | 60 (63.2%) | 0.016 |
| Female | 13 (61.9%) | 2 (16.7%) | 13 (40.6%) | 7 (23.3%) | 35 (36.8%) | |
| Underlying diseases (%) | | | | | | |
| Diabetes | 0 | 0 | 9 (28.1%) | 12 (40%) | 21 (22.1%) | <0.001 |
| CVD | 2 (9.5%) | 0 | 13 (40.6%) | 10 (33.3%) | 25 (26.3%) | 0.009 |
| Liver diseases | 0 | 0 | 4 (12.5%) | 9 (30%) | 13 (13.7%) | 0.007 |
| Kidney diseases | 0 | 0 | 4 (12.5%) | 1 (3.3%) | 5 (5.3%) | 0.144 |
| Lung diseases | 0 | 0 | 3 (9.4%) | 0 | 3 (3.2%) | 0.107 |
| Others | 0 | 0 | 3 (9.4%) | 2 (6.7%) | 5 (5.3%) | 0.386 |

for the SARS-CoV-2 virus. The Ct value was used as the semi-quantitative predictor of the viral load. Since Ct values vary depending on the qPCR system and the methodology of the NAAT, we only focused our analysis on patients with moderate and severe COVID-19. All of the patients in these groups were hospitalised in Dr Soetomo General Academic Hospital; hence, the NAAT was conducted in the same place, i.e. the Clinical Pathology and Microbiology Laboratory, Dr Soetomo General Academic Hospital. We analysed the Ct values of the first NAAT, which were conducted at the time when the patients were admitted to the hospital. A low Ct value is likely associated with a high viral load, whereas a high Ct value is likely to be associated with a low viral load. As illustrated in Fig. 1a, a significant difference was observed in the Ct value between patients with a TT genotype and patients with a CC genotype ($P = 0.04$), indicating a possible association of this genotype with a higher viral load. The Pearson correlation analysis also indicated a trend of decreasing Ct value with the presence of the C allele ($P = 0.08$). In contrast, we did not

observe any difference and correlation between Ct value and patients' gender (Fig. 1b) as well as between Ct value and age (Fig. 1c). Additionally, there was no difference in Ct value between patients with moderate and severe COVID-19 (Fig. 1d).

### TMPRSS2 polymorphism and patients' outcome
During the course of the study, all of the patients with mild COVID-19 recovered, whereas 9.4% of the patients with moderate COVID-19 and 60% of the patients with severe COVID-19 died. When we analysed the association between TMPRSS2 p.Val160Met polymorphism and the patients' outcomes, we did not find any association between the polymorphism and mortality in the moderate COVID-19 group (Table 4). However, we observed a trend of association in the severe group, in which a higher proportion of patients who died of COVID-19 had a CC genotype ($P = 0.042$ using the linear-by-linear association chi-squared test) (Table 4). We also observed an increasing trend of odds

**Table 2** Genotype and allele frequencies of the TMPRSS2 p.Val160Met polymorphism according to COVID-19 severity

| | Asymptomatic (N=21) | | Mild (N=12) | | Moderate (N=32) | | Severe (N=30) | | Chi-square test |
|---|---|---|---|---|---|---|---|---|---|
| | N | % | N | % | N | % | N | % | |
| Genotype (amino acids) | | | | | | | | | |
| CC (Val/Val) | 8 | 38.1 | 4 | 33.3 | 17 | 53.1 | 13 | 43.4 | $\chi^2 = 3.11$ |
| CT (Val/Met) | 7 | 33.3 | 5 | 41.7 | 11 | 34.4 | 10 | 33.3 | $P=0.79$ |
| TT (Met/Met) | 6 | 28.6 | 3 | 25 | 4 | 12.5 | 7 | 23.3 | |
| Allele | | | | | | | | | |
| C Allele | 23 | 54.8 | 13 | 54.2 | 45 | 70.3 | 36 | 60 | $\chi^2 = 3.51$ |
| T Allele | 19 | 45.2 | 11 | 45.8 | 19 | 29.7 | 24 | 40 | $P=0.32$ |

Wulandari *et al. Human Genomics*        (2021) 15:29

Page 4 of 9

**Table 3** Analysis of odds ratio for the risk of symptomatic or more severe COVID-19 in each genotype

| Genotype | Asymptomatic (*N*=21) vs all symptomatic cases (*N*=74) | | | Asymptomatic and mild (*N*=33) vs moderate and severe cases (*N*=62) | | |
|---|---|---|---|---|---|---|
| | Odds ratio | 95% CI | *P*-value | Odds ratio | 95% CI | *P*-value |
| CC | 1.821 | 0.533–6.219 | 0.335 | 2.045 | 0.676–6.185 | 0.255 |
| CT | 1.592 | 0.447–5.664 | 0.471 | 1.432 | 0.462–4.437 | 0.573 |
| TT | 1 | Reference | | 1 | Reference | |

ratios of mortality in subjects with CC and CT genotypes (Table 5).

## Discussion

This is the first study to demonstrate a possible association between TMPRSS2 p.Val160Met polymorphism and the degree of SARS-CoV-2 viral load as indicated by the Ct value of NAAT in patients with COVID-19.

Patients with a CC genotype, which corresponds to the presence of valine amino acid, tend to display a lower Ct value (high viral load). We also found a trend of association between a CC genotype and mortality in a group of patients with severe COVID-19.

It is widely known that the SARS-CoV-2 virus enters the host cells via binding with ACE2, which acts as the main receptor for the viral particles [6, 9, 17]. The spike
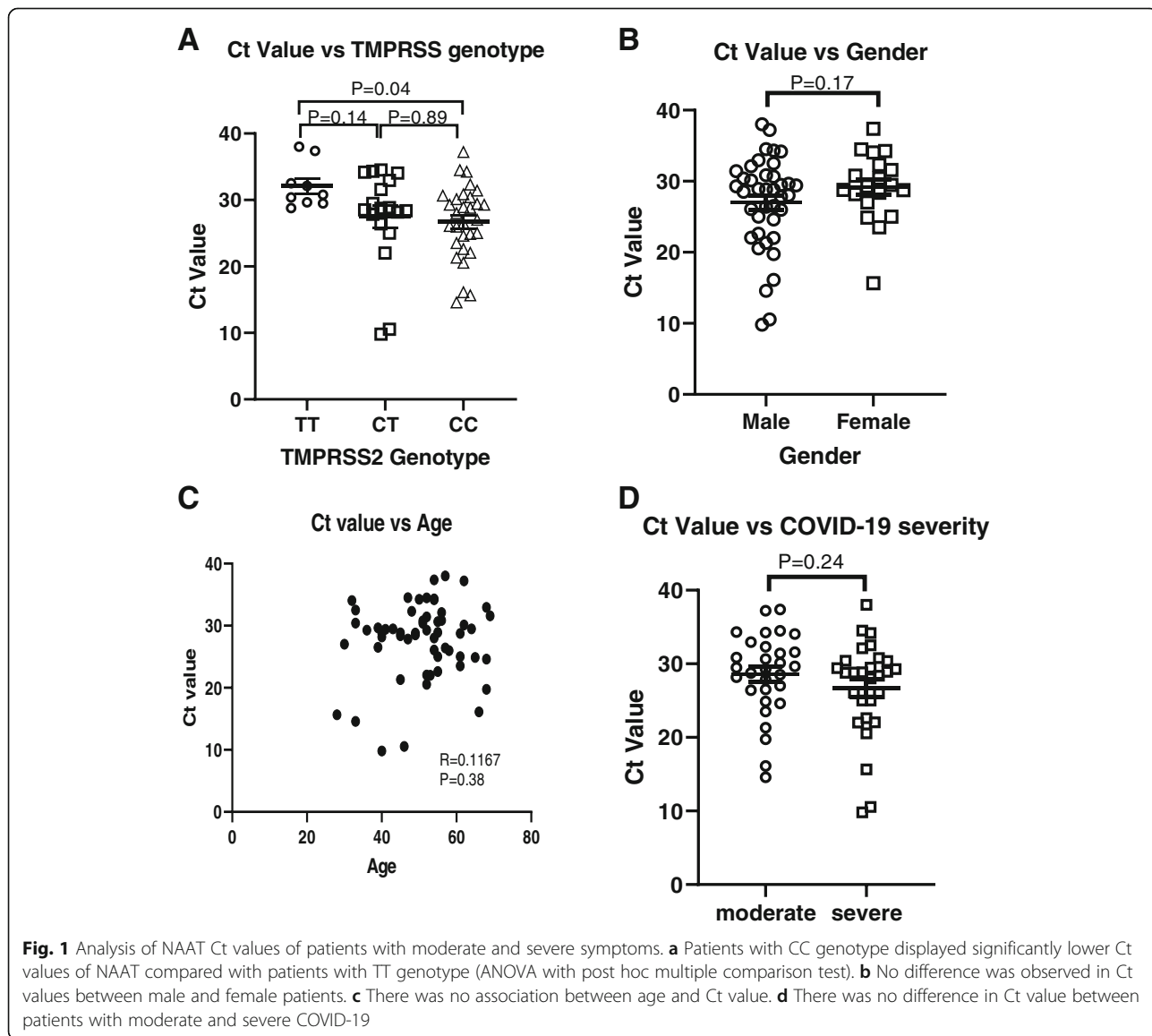


**Fig. 1** Analysis of NAAT Ct values of patients with moderate and severe symptoms. **a** Patients with CC genotype displayed significantly lower Ct values of NAAT compared with patients with TT genotype (ANOVA with post hoc multiple comparison test). **b** No difference was observed in Ct values between male and female patients. **c** There was no association between age and Ct value. **d** There was no difference in Ct value between patients with moderate and severe COVID-19

Wulandari *et al. Human Genomics*        (2021) 15:29

Page 5 of 9

**Table 4** Association between TMPRSS2 polymorphism with mortality/survival in patients with moderate-severe COVID-19

|  | Moderate COVID-19 (*N*=32) | | Severe COVID-19 (*N*=30) | |
|---|---|---|---|---|
|  | Recovered (*N*=29, 90.6%) | Died (*N*=3, 9.4%) | Recovered (*N*=12, 30%) | Died (*N*=18, 60%) |
| Genotype (amino acids) | | | | |
| CC (Val/Val) | 16 (50%) | 1 (3.1%) | 3 (10 %) | 10 (33.3%) |
| CT (Val/Met) | 10 (31.3%) | 1 (3.1%) | 4 (13.3%) | 6 (20 %) |
| TT (Met/Met) | 3 (9.4%) | 1 (3.1%) | 5 (16.7%) | 2 (6.7%) |
|  | Chi-square test *P*=0.498 | | Chi-square test *P*=0.109 | |
|  | Linear-by-linear chi-square association test *P*=0.299 | | Linear-by-linear chi-square association test *P*=0.042 | |

(S) protein of the SARS-CoV-2 virus consists of two sub-units: the S1 sub-unit, which is important for virus attachment, and the S2 sub-unit, which is essential for membrane fusion. ACE2 molecule can bind to the S1 protein to promote virus invasion into the host cells [6, 18]. In human, ACE2 is expressed in many organs, such as the upper respiratory tract, alveolar epithelial cells, vascular endothelial cells and macrophages [5].

In addition to ACE2, several other molecules are involved in SARS-CoV-2 virus binding and cell penetration. The S protein needs to be cleaved to activate the endocytic route of virus entry and to enable membrane fusion. It has been reported that several host proteases are involved in the process of S protein breakdown. These include TMPRSS2, cathepsin L, furin [9] and NRP1 [10].

TMPRSS2 is a serine protease that can prime the S protein of SARS-CoV-2 to enable cell penetration [9, 19]. The expression of TMPRSS2 in VeroE6 cells facilitates SARS-CoV-2 virus entry and promotes virus invasion [9]. Notably, treatment with the TMPRSS2 inhibitor (camostat mesylate) significantly reduced SARS-CoV-2 virus infection [9]. Moreover, TMPRSS2 is also involved in SARS-CoV-1 virus infection [20], supporting the idea of the critical role of this molecule in mediating virus entry.

The human *TMPRSS2* gene is located in chromosome 21.q22.3. It encodes protein that contains a transmembrane domain, low-density lipoprotein receptor class A (LDLRA) domain, scavenger receptor cysteine-rich (SRCR) domain and serine protease catalytic domain [21]. At least six nucleotide variants within the human *TMPRSS2* coding region that cause amino acid substitutions have been identified. These include p.Val160Met,

p.Gly181Arg, p.Arg240Cys, p.Gly259Ser, p.Pro335Leu and p.Gly432Ala [22]. Of these variants, the p.Val160-Met variant is often associated with diseases, notably prostate cancer. A study conducted on a Japanese population indicated that the TMPRSS2 p.Val160Met variant (also known as Met160Val polymorphism) was associated with the risk of sporadic prostate cancer [16]. Also, a study conducted on 214 patients with prostate cancer demonstrated that the T allele of this variant, which is associated with the presence of Met amino acid, was associated with TMPRSS2-ERG fusion and, thus, might be important in prostate cancer pathogenesis [23].

Our data indicate that in our study population the proportion of the genotypes deviates from the Hardy–Weinberg equilibrium. This deviation could be due to natural selection, non-random mating, genetic drift, or gene flow [24]. In our study population, the deviation was likely due to the higher number of subjects with CC and TT genotypes than the expected values (Additional file 1). However, the frequency of subjects with heterozygous genotype (CT) was lower than the expected value. Therefore, it is unlikely that the deviation was due to natural selection or advantages of a specific allele because the number of subjects with homozygous genotypes of both the T and C alleles was higher than the expected values, implying that there was no specific advantage of either the T or C allele. Thus, it is more likely that the deviation was due to non-random mating or genetic drift, which is more likely to occur in a small study population.

Recent bioinformatic analysis studying the functional effects of nucleotide variants within the human *TMPRSS2* gene revealed that the p.Val160Met variant was the

**Table 5** Analysis of odds ratio for the mortality in moderate and severe COVID-19 patients

|  | Odds ratio of mortality in moderate COVID-19 | | | Odds ratio of mortality in severe COVID-19 | | |
|---|---|---|---|---|---|---|
|  | Odds ratio | 95% CI | *P*-value | Odds ratio | 95% CI | *P*-value |
| Genotype | | | | | | |
| CC | 0.188 | 0.009-3.895 | 0.241 | 8.33 | 1.034-67.14 | 0.062 |
| CT | 0.3 | 0.014-6.382 | 0.423 | 3.75 | 0.473-29.75 | 0.335 |
| TT | 1 | Reference | | 1 | Reference | |

Wulandari *et al. Human Genomics*        (2021) 15:29

Page 6 of 9

most likely variant that might affect TMPRSS2 protein function and stability [12]. Furthermore, a computational analysis to predict the effects of polymorphism on protein structure suggested that the Val160Met substitution might create a pocket protein by influencing several amino acid residues, which might affect TMPRSS2 structure and its role in SARS-CoV-2 cell entry [12]. Possible changes in TMPRSS2 function and/or structure due to the Val160Met substitution might explain our findings on the association of this SNP with the viral load in COVID-19 patients. Alteration of TMPRSS2 function/structure will likely affect the binding of the S protein to ACE2 or the membrane fusion process. Reduction in TMPRSS2 enzymatic activity may decrease the furin cleavage of the S1 protein, which may subsequently decrease S2 fusion to the host's cell membrane. However, further studies at the molecular level are required to prove this hypothesis, for example, by generating recombinant TMPRSS2 proteins bearing the variants and testing them in an in vitro model of SARS-CoV-2 cell infection.

Despite the association between pVal160Met polymorphism and the Ct value, we did not find any correlation between the variant and COVID-19 severity. This might be due to other confounding factors that strongly contribute to the severity of COVID-19. It is believed that factors, such as age [3], gender [25] and pre-existing diseases (hypertension, diabetes, CVD and lung disease) [4], strongly correlate with the risk of severe COVID-19. Further analysis with a larger study population is required to control these confounding variables. Interestingly, in patients with severe COVID-19, we observed a trend of association between this polymorphism and the mortality of COVID-19 patients. However, this requires further confirmation in studies with a larger sample size.

Several studies have found associations between genetic variations in the patient's genome and COVID-19 severity. Many of the reported polymorphisms were related to genes involved in the development of inflammatory response, for example, polymorphisms in genes related to type 1 interferon immunity [26], polymorphisms in X-chromosomal TLR7 [27] and polymorphisms within genes involved in the interleukin 1 signalling pathway [28]. Our finding indicates a correlation between polymorphism in the gene encoding the virus receptor complex, i.e. TMPRSS2, and COVID-19 severity. Our data are consistent with previous reports on TMPRSS2 polymorphisms and COVID-19 severity. Initial whole genome analysis of 322 COVID-19 patients in a Chinese population observed a decreasing allele frequency of the TMPRSS2 rs12329760 variant among patients with severe disease compared with patients with mild COVID-19 [28], indicating the importance of TMPRSS2 in COVID-19. Consistently, studies on Italian

COVID-19 patients also demonstrated the role of genetic polymorphisms within the TMPRSS2 gene in determining COVID-19 severity. An observation on 133 COVID-19 patients found a difference in frequency of this variant in the COVID-19 patient cohort compared with the frequency in the reference databases [29]. In particular, in-depth analysis of available data from the COVID-19 Host Genetic Initiative (HGI) [30] suggested a significant association between TMPRSS2 gene polymorphism rs12329760 with severe/hospitalised COVID-19 [31]. However, further analysis on the available HGI data (https://app.covid19hg.org/) indicated that there was no significant association/difference of polymorphic allele between all SARS-CoV-2-infected subjects and the general population ($P=0.569$). This data is in line with our finding that the polymorphism may have more significant effects in severe cases of COVID-19.

Together, all of the data indicate a crucial involvement of the TMPRSS2 genetic variation, the p.Val160Met (rs12329760) in particular, in mediating the severity of COVID-19. This will contribute to the growing body of evidence on the crucial involvement of the host's genetic factor in determining susceptibility to and/or severity of COVID-19.

## Conclusions

In summary, this is the first study to demonstrate a possible association between TMPRSS2 p.Val160Met polymorphism and higher viral load in COVID-19 patients. The main limitation of our study is its small sample size. Further large-scale studies are required to validate our findings. Also, by using the Ct value, we can only have an estimate of the viral load. Precise determination of the viral RNA copy number using standard curve qPCR is required to accurately determine the viral load. Mechanistic analysis using a cell culture system is also important to confirm the effects of p.Val160Met on TMPRSS2 protein function. Nevertheless, our finding may provide new insights into the possibility of using this polymorphism as a biomarker or predictor for COVID-19 severity/ clinical outcome. Furthermore, our data may also support the idea of targeting TMPRSS2 in COVID-19 therapy, as has been done in some clinical trials [32].

## Methods

### Study design, patients and data collection

This study was a cross-sectional study conducted from June to August 2020. During this period, a total of 95 patients with COVID-19 were enrolled. Patients with moderate and severe COVID-19 ($n = 62$, 65.3%) were hospitalised in Dr Soetomo General Academic Hospital, Surabaya, Indonesia, whilst 33 patients (34.7%) with asymptomatic or mild symptoms were treated in Indrapura KOGABWILHAN II Hospital, Surabaya, Indonesia.

Wulandari *et al. Human Genomics*     (2021) 15:29

Page 7 of 9

The diagnosis was confirmed using the nucleic acid amplification test (NAAT) of the oro-nasopharyngeal swab specimens. For patients with moderate and severe symptoms, the NAAT was performed in the Clinical Pathology and Microbiology Laboratory, Dr Soetomo General Academic Hospital, whereas for asymptomatic patients and patients with mild symptoms, the NAAT was conducted in the Centre for Health Laboratory, Surabaya, as part of the standard procedure for COVID-19 management in East Java Province, Indonesia. This study obtained ethical approval from the Local Ethics Committee of Dr Soetomo General Academic Hospital, Surabaya, Indonesia (0006/LOE/301.4.2/V/2020). All patients have signed the informed consent and agreed to participate in this study.

We clustered patients in three categories of disease severity based on criteria according to the WHO Guideline for COVID-19 Management [33] as follows: (i) mild: characterised by the presence of COVID-19 symptoms that meet the case definition of COVID-19 (fever, persistent cough, fatigue, anorexia, shortness of breath, myalgia, sore throat, nose congestion, headache, diarrhoea, nausea and vomiting, anosmia, ageusia) without evidence of viral pneumonia and hypoxia; (ii) moderate: characterized by the presence of the clinical signs of pneumonia but without any signs of hypoxia ($SpO_2 \geq 93\%$); and (iii) severe: characterized by the presence of the clinical signs of pneumonia and one of the clinical signs of respiratory distress (respiratory rate > 30×/min, severe respiratory distress, or $SpO_2 < 93\%$).

### DNA isolation
Heparinized peripheral blood samples were collected and stored in a −80°C freezer before use. DNA extraction was performed using the QIAamp® Blood DNA Midi kit (cat #51185, Qiagen) according to the manufacturer's recommended protocol. DNA concentrations were determined using a microvolume spectrophotometer (NanoDrop Lite, Thermo Fisher Scientific). The procedures were conducted in the Biosafety Level 3 (BSL 3) Laboratory in the Institute of Tropical Disease, Universitas Airlangga, to reduce the risk of COVID-19 transmission.

### Polymorphism detection
The TMPRSS2 polymorphism (rs12329760, TMPRSS2 p.Val160Met also known as TMPRSS2 Met160Val polymorphism) was detected using a TaqMan SNP genotyping assay (Cat #4351379, Applied Biosystems, USA) in accordance with the protocol recommended by the manufacturer. Genotyping was performed using real-time polymerase chain reaction (RT-PCR) with VIC and FAM fluorescent reporters to indicate allelic discrimination. The 7500 Fast Real-Time PCR System (Applied Biosystems) was used in conjunction with the 7500 software v2.3 (Life Technologies™, Applied Biosystems) to create the allelic discrimination plot.

### Data analysis
Statistical analyses were performed using the IBM SPSS Statistics Software ver. 23 (IBM Corp.) or GraphPad Prism ver. 8 (GraphPad Software, LLC). A chi-squared test was used to examine the Hardy–Weinberg equilibriums and to determine the association between categorical variables in the cross-tabulation data. ANOVA with post hoc multiple comparisons was used to analyse numerical data. A $P$ value less than 0.05 was considered to be statistically significant.

## Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s40246-021-00330-7.

---

**Additional file 1.** Genotype and allele frequencies of the TMPRSS2 p.Val160Met polymorphism in all patients.

---

### Authors' contributions
L.W, designed the study, managed the funding, supervised the biological sample collection and collected and analysed the clinical data; B.H, performed the DNA extraction and polymorphism detection and wrote the manuscript; C.P: performed the DNA extraction and polymorphism detection and wrote the manuscript; N.S.D, collected blood samples and clinical data; N.D.K, collected blood samples; M.R.W, collected the clinical data; C.O.A, collected blood samples; S, supervised the blood sample collection, D.H, supervised the blood sample collection; DT, designed the study and supervised ethical clearance; CRSP, designed the study and supervised the project; A.E, supervised the project; N.N.T.P, supervised the project; Y.M, supervised the project; M.I.L, supervised the project; K.S, supervised the DNA isolation and polymorphism detection; DO, conceived the original idea, designed the study, performed the data analysis and wrote and edited manuscript. The authors read and approved the final manuscript.

### Availability of data and materials
All data generated or analysed during this study are included in this published article and its supplementary information files.

## Declarations

### Ethics approval and consent to participate
The protocol of the study including biological sample collection and storage of materials and information was approved by the Ethics Committee of Dr Soetomo General Academic Hospital, Surabaya, Indonesia (0006/LOE/301.4.2/V/2020). All patients signed the informed consent form and agreed to participate in this study.

### Consent for publication
Not applicable

### Competing interests
The authors declare that they have no competing interests.

### Author details
[1]Department of Pulmonology and Respiratory Medicine, Faculty of Medicine, Universitas Airlangga/Dr Soetomo General Academic Hospital, Surabaya, Indonesia. [2]Department of Biomedical Sciences, Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia. [3]Andrology Program, Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia. [4]Indrapura KOGABWILHAN II Hospital, Surabaya, Indonesia. [5]Department of Medical Microbiology, Faculty of Medicine, Universitas Airlangga/Clinical Microbiology Unit, Central Laboratory Installation, Dr Soetomo General Academic Hospital, Surabaya, Indonesia. [6]Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia. [7]Clinical Pathology Program, Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia. [8]Department of Child Health, Faculty of Medicine, Universitas Airlangga/Dr Soetomo General Academic Hospital, Surabaya, Indonesia. [9]Department of Physical Medicine and Rehabilitation, Faculty of Medicine, Universitas Airlangga/Dr Soetomo General Academic Hospital, Surabaya, Indonesia. [10]Department of Dermatology Venerology, Faculty of Medicine, Universitas Airlangga/Dr. Soetomo General Academic Hospital, Surabaya, Indonesia. [11]Department of Chemistry, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia. [12]Laboratory of Proteomic, University CoE-Research Center for Bio-Molecule Engineering, Universitas Airlangga, Surabaya, Indonesia. [13]Center for Infectious Diseases, Kobe University Graduate School of Medicine, Kusunoki-cho, Chuo-ku, Kobe, Japan. [14]Institute of Tropical Disease, Universitas Airlangga, Surabaya, Indonesia. [15]Department of Microbiology, Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia. [16]CRC-ERID, Institute of Tropical Disease, Universitas Airlangga, Surabaya, Indonesia. [17]Division of Cardiovascular Sciences, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK.

## References
1. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis. 2020;20(5):533–4. https://doi.org/10.1016/S1473-3099(20)30120-1.
2. Docherty AB, Harrison EM, Green CA, Hardwick HE, Pius R, Norman L, et al. Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. BMJ. 2020;369:m1985.
3. Wu C, Chen X, Cai Y, Xia J, Zhou X, Xu S, et al. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. JAMA Intern Med. 2020;180(7):934–43. https://doi.org/10.1001/jamainternmed.2020.0994.
4. Yang J, Zheng Y, Gou X, Pu K, Chen Z, Guo Q, et al. Prevalence of comorbidities and its effects in patients infected with SARS-CoV-2: a systematic review and meta-analysis. Int J Infect Dis. 2020;94:91–5. https://doi.org/10.1016/j.ijid.2020.03.017.
5. Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature. 2003;426(6965):450–4. https://doi.org/10.1038/nature02145.
6. Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature. 2020;581(7807):215–20. https://doi.org/10.1038/s41586-020-2180-5.
7. Novelli A, Biancolella M, Borgiani P, Cocciadiferro D, Colona VL, D'Apice MR, et al. Analysis of ACE2 genetic variants in 131 Italian SARS-CoV-2-positive patients. Hum Genomics. 2020;14(1):29. https://doi.org/10.1186/s40246-020-00279-z.
8. Chiu RW, Tang NL, Hui DS, Chung GT, Chim SS, Chan KC, et al. ACE2 gene polymorphisms do not affect outcome of severe acute respiratory syndrome. Clin Chem. 2004;50(9):1683–6. https://doi.org/10.1373/clinchem.2004.035436.
9. Hoffmann M, Kleine-Weber H, Schroeder S, Kruger N, Herrler T, Erichsen S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell. 2020;181(2):271–80 e278. https://doi.org/10.1016/j.cell.2020.02.052.
10. Cantuti-Castelvetri L, Ojha R, Pedro LD, Djannatian M, Franz J, Kuivanen S, et al. Neuropilin-1 facilitates SARS-CoV-2 cell entry and infectivity. Science. 2020;370(6518):856–60. https://doi.org/10.1126/science.abd2985.
11. Bestle D, Heindl MR, Limburg H, Van Lam Van T, Pilgram O, Moulton H, Stein DA, Hardes K, Eickmann M, Dolnik O et al: TMPRSS2 and furin are both essential for proteolytic activation of SARS-CoV-2 in human airway cells. Life Sci Alliance. 2020;3(9):e202000786. https://doi.org/10.26508/lsa.202000786.
12. Paniri A, Hosseini MM, Akhavan-Niaki H. First comprehensive computational analysis of functional consequences of TMPRSS2 SNPs in susceptibility to SARS-CoV-2 among different populations. J Biomol Struct Dyn. 2020. https://doi.org/10.1080/07391102.2020.1767690.
13. Bhanushali A, Rao P, Raman V, Kokate P, Ambekar A, Mandva S, et al. Status of TMPRSS2-ERG fusion in prostate cancer patients from India: correlation with clinico-pathological details and TMPRSS2 Met160Val polymorphism. Prostate Int. 2018;6(4):145–50. https://doi.org/10.1016/j.prnil.2018.03.004.
14. Giri VN, Ruth K, Hughes L, Uzzo RG, Chen DY, Boorjian SA, et al. Racial differences in prediction of time to prostate cancer diagnosis in a prospective screening cohort of high-risk men: effect of TMPRSS2 Met160Val. BJU Int. 2011;107(3):466–70. https://doi.org/10.1111/j.1464-410X.2010.09522.x.
15. Lubieniecka JM, Cheteri MK, Stanford JL, Ostrander EA. Met160Val polymorphism in the TRMPSS2 gene and risk of prostate cancer in a population-based case-control study. Prostate. 2004;59(4):357–9. https://doi.org/10.1002/pros.20005.
16. Maekawa S, Suzuki M, Arai T, Suzuki M, Kato M, Morikawa T, et al. TMPRSS2 Met160Val polymorphism: significant association with sporadic prostate cancer, but not with latent prostate cancer in Japanese men. Int J Urol. 2014;21(12):1234–8. https://doi.org/10.1111/iju.12578.
17. Wang Q, Zhang Y, Wu L, Niu S, Song C, Zhang Z, et al. Structural and functional basis of SARS-CoV-2 entry by using human ACE2. Cell. 2020;181(4):894–904 e899. https://doi.org/10.1016/j.cell.2020.03.045.
18. Cavasotto CN, Lamas MS, Maggini J. Functional and druggability analysis of the SARS-CoV-2 proteome. Eur J Pharmacol. 2021;890:173705. https://doi.org/10.1016/j.ejphar.2020.173705.
19. Matsuyama S, Nao N, Shirato K, Kawase M, Saito S, Takayama I, et al. Enhanced isolation of SARS-CoV-2 by TMPRSS2-expressing cells. Proc Natl Acad Sci U S A. 2020;117(13):7001–3. https://doi.org/10.1073/pnas.2002589117.
20. Shulla A, Heald-Sargent T, Subramanya G, Zhao J, Perlman S, Gallagher T. A transmembrane serine protease is linked to the severe acute respiratory syndrome coronavirus receptor and activates virus entry. J Virol. 2011;85(2):873–82. https://doi.org/10.1128/JVI.02062-10.
21. Thunders M, Delahunt B. Gene of the month: TMPRSS2 (transmembrane serine protease 2). J Clin Pathol. 2020;73(12):773–6. https://doi.org/10.1136/jclinpath-2020-206987.
22. Hou Y, Zhao J, Martin W, Kallianpur A, Chung MK, Jehi L, et al. New insights into genetic susceptibility of COVID-19: an ACE2 and TMPRSS2 polymorphism analysis. BMC Med. 2020;18(1):216. https://doi.org/10.1186/s12916-020-01673-z.
23. FitzGerald LM, Agalliu I, Johnson K, Miller MA, Kwon EM, Hurtado-Coll A, et al. Association of TMPRSS2-ERG gene fusion with clinical characteristics and outcomes: results from a population-based study of prostate cancer. BMC Cancer. 2008;8(1):230. https://doi.org/10.1186/1471-2407-8-230.
24. Wigginton JE, Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. Am J Hum Genet. 2005;76(5):887–93. https://doi.org/10.1086/429864.
25. Gebhard C, Regitz-Zagrosek V, Neuhauser HK, Morgan R, Klein SL. Impact of sex and gender on COVID-19 outcomes in Europe. Biol Sex Differ. 2020;11(1):29. https://doi.org/10.1186/s13293-020-00304-9.
26. Zhang Q, Bastard P, Liu Z, Le Pen J, Moncada-Velez M, Chen J, et al. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. Science. 2020;370(6515):eabd4570. https://doi.org/10.1126/science.abd4570.

27.  van der Made CI, Simons A, Schuurs-Hoeijmakers J, van den Heuvel G, Mantere T, Kersten S, et al. Presence of genetic variants among young men with severe COVID-19. JAMA. 2020;324(7):663–73.

28.  Wang F, Huang S, Gao R, Zhou Y, Lai C, Li Z, et al. Initial whole-genome sequencing and analysis of the host genetic contribution to COVID-19 severity and susceptibility. Cell Discov. 2020;6(1):83. https://doi.org/10.1038/s41421-020-00231-4.

29.  Latini A, Agolini E, Novelli A, Borgiani P, Giannini R, Gravina P, et al. COVID-19 and genetic variants of protein involved in the SARS-CoV-2 entry into the host cells. Genes (Basel). 2020;11(9):1010. https://doi.org/10.3390/genes11091010.

30.  COVID-19 Host Genetics Initiative. The COVID-19 Host Genetics Initiative, a global initiative to elucidate the role of host genetic factors in susceptibility and severity of the SARS-CoV-2 virus pandemic. Eur J Hum Genet. 2020; 28(6):715–8.

31.  Andolfo I, Russo R, Lasorsa VA, Cantalupo S, Rosato BE, Bonfiglio F, et al. Common variants at 21q22.3 locus influence MX1 and TMPRSS2 gene expression and susceptibility to severe COVID-19. iScience. 2021;24(4): 102322.

32.  Hofmann-Winkler H, Moerer O, Alt-Epping S, Brauer A, Buttner B, Muller M, et al. Camostat mesylate may reduce severity of coronavirus disease 2019 sepsis: a first observation. Crit Care Explor. 2020;2(11):e0284. https://doi.org/10.1097/CCE.0000000000000284.

33.  World Health Organization: Clinical management of COVID-19: World Health Organization; 2020.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.