

# Association between the HLA genotype and the severity of COVID-19 infection among South Asians

Fatmah M. A. Naemi<sup>1</sup>  | Shurooq Al-adwani<sup>1</sup> | Heba Al-khatibi<sup>2</sup> |  
Ashwaq Al-nazawi<sup>3</sup> 

<sup>1</sup>Laboratory of Histocompatibility and Immunogenetics, King Fahd General Hospital, Ministry of Health, Jeddah, Kingdom of Saudi Arabia

<sup>2</sup>Center of Excellence in Genomic Medicine Research, King Fahd Medical Research Center, King Abdul-Aziz University, Jeddah, Kingdom of Saudi Arabia

<sup>3</sup>Public Health VBD, Ministry of Health, Jeddah, Kingdom of Saudi Arabia

## Correspondence

Fatmah M. A. Naemi, Laboratory of Histocompatibility and Immunogenetics, King Fahd General Hospital, Ministry of Health, Arafat Street, Al Hamra District, Jeddah 8488, Kingdom of Saudi Arabia.  
Email: [fnaemi@moh.gov.sa](mailto:fnaemi@moh.gov.sa)

## Funding information

King Abdulaziz City for Science and Technology, Grant/Award Number: 5-20-01-502-0009

## Abstract

Regional variations are found in the incidence and severity of the COVID-19 infection. Human leukocyte antigen (HLA) polymorphism is one of the genetic factors that might have an impact on the outcome of the disease. This study explored the association between the HLA genotype and the severity of COVID-19 among patients from South Asia. Blood samples from 95 Asians (Bangladeshis, Indians, and Pakistanis) with COVID-19 were collected. The patients were divided according to the severity of their infection: mild ( $N = 64$ ), severe ( $N = 31$ ), and fatal ( $N = 20$ ). DNA was extracted from all samples, and HLA genotyping was performed for both class I (A, B, and C) and class II (DRB1, DQA1, and DQB1) using the PCR-rSSO (polymerase chain reaction–reverse sequence-specific oligonucleotide) molecular method. The frequency of HLA-B\*51 was significantly higher among patients in the fatal group than among those in the mild infection group (15% vs. 4.7%,  $p = 0.027$ ). Additionally, the frequency of HLA-B\*35 was significantly higher in the mild infection group than in the fatal group (21.1% vs. 7.5%,  $p = 0.050$  [a borderline  $p$ -value]). In terms of HLA class II, DRB1\*13 was significantly observed in the fatal group than in the mild infection group (17.5% vs. 11.3%,  $p = 0.049$ ). However, the  $p$ -value for all alleles became insignificant after a statistical correction for the  $p$ -values ( $p_c = 0.216$ ,  $p_c = 0.4$ , and  $p_c = 0.49$ , respectively). Compared with all published data, this study highlights that the association between the HLA system and the COVID-19 outcome might be ethnic-dependent. Genetic variation between populations must be examined on a wider scale to assess infection prognosis and vaccine effectiveness.

## KEYWORDS

COVID-19, HLA, severity, South Asia

## 1 | INTRODUCTION

By January 2021, the positive cases of COVID-19 reported in more than 200 countries have exceeded 90 million in total, with more than 2 million deaths (<https://covid19.who.int/table>). The lack of understanding about the exact mechanism of the virus transmission increases the burden to control the spread of the infection between individuals. However, differences in the number of confirmed cases

between countries and even between areas in the same country have been observed.<sup>1</sup> While the United States reported more than 23 million cases by January 2021, the confirmed cases in other areas have not exceeded half a million (<https://covid19.who.int/table>). Additionally, the severity and fatality of the infection vary between populations, with a 2.6% fatality rate in some areas.<sup>2</sup> The severity of the infection ranges from being asymptomatic to developing into a fatal acute respiratory syndrome.<sup>3,4</sup> South Asian countries, such as

India, are the second most affected area worldwide, with more than 10 million cases and a mortality rate of around 1.3%, as reported in January 2021. Other South Asian countries, such as Pakistan and Bangladesh, reported cases exceeding 500 thousand each, with a mortality rate of around 0.4% (<https://covid19.who.int/table>). The rapid spread of the virus in this region can be attributed to the poor hygiene precautions and a large number of individuals in contact daily.<sup>5</sup> However, other factors, such as genetic variations, can also have an impact on virus transmission and infection severity.<sup>6,7</sup>

Previous studies reported an ethnic variation in the susceptibility to viruses and their severity.<sup>8–10</sup> This can be linked to the variations in the human leukocyte antigen (HLA), the highly polymorphic gene cluster in the human genome. Each HLA allele codes a specific HLA protein with a unique peptide-binding groove. Upon the virus' entry through mucous membranes, it will be randomly degraded into short peptides to be up-loaded into a peptide-binding groove of the HLA molecules and will then be presented to the cell surface.<sup>11</sup> T lymphocytes will recognize the viral peptide presented by self-HLA molecules to elicit an antiviral immune response for virus eradication. The set of HLA alleles inherited by an individual will determine the immune responses to viruses according to the selected peptides that can bind to the peptide-binding groove.<sup>12</sup> HLA class I (A, B, and C) and class II (DRB1, DQA1, and DQB1) alleles have been shown to be associated with the susceptibility to and severity of various infectious diseases.<sup>13–16</sup> In terms of COVID-19, studies from different populations showed an association between various HLA alleles and the susceptibility to or severity of the disease.<sup>17–19</sup> A recent study from Italy showed an increase in the frequency of B\*27:07, DRB1\*15:01, and DQB1\*06:02 among severe COVID-19 patients in a cohort of 99 Italians.<sup>18</sup> However, another study from Sardinia in Italy showed a higher frequency of different allele, HLA-DRB1\*08:01, among COVID-19 hospitalized patients compared to control.<sup>20</sup> A study from Spain, on the other hand, showed the relevance of HLA-A\*11, C\*01, and DQB1\*04 to infection-related mortality in a cohort of 72 severe patients.<sup>21</sup> Additionally, epidemiological and bioinformatics studies from different groups reported different HLA alleles as either protective against or susceptible to infection.<sup>19,22,23</sup> The discrepancies between the studies can be attributed to the study design and the sample size, as well as ethnicity variations. Therefore, it is essential to study the association between HLA genotypes and COVID-19 in different populations worldwide.

In this study, the effect of HLA polymorphism on the severity of COVID-19 among South Asians was assessed and compared to the published data from other countries. This will help in finding the HLA marker related to the severity and fatality of COVID-19, which can be used to identify high-risk individuals. Studying such an association will also have an impact on vaccine innovation and assessing vaccine effectiveness. In Saudi Arabia, several positive cases of COVID-19 reported among non-Saudis reached more than 35% of all cases, adding more pressure on healthcare services.<sup>24</sup> Therefore, it was interesting to study the association between the HLA system and the severity of COVID-19 in the non-Saudi population. Patients from South Asia were selected due to the increased number of cases in the region compared to others, rendering them a good model to assess this association. To our knowledge, this is the first study on the

association between HLA and COVID-19's severity in those populations.

## 2 | METHODS

### 2.1 | Sample collection

This study was conducted after obtaining an approval from the Research Ethics Committee of the Ministry of Health in the Kingdom of Saudi Arabia, with KACST IRB registration number H-02-J-002. Asian patients (Pakistanis, Indians, and Bangladeshis) who were admitted into Ministry of Health hospitals in the Makkah region, Saudi Arabia, because of COVID-19 symptoms were recruited. The enrolled patients were admitted into hospitals between July and September 2020. Consent forms were prepared in the appropriate language and signed by all participants. Positive cases were diagnosed using the molecular detection method for the SARS-CoV-2 virus from nasopharyngeal swabs. Patients were followed up during their hospital stay and divided into three groups according to the severity of their infection: mild infection (patients recovered), severe infection (patients required admission to the intensive care unit [ICU]), and fatal infection (patients died due to COVID-19). The infection severity was classified according to the NIH treatment guideline based on an early study from china.<sup>25</sup> Patients with mild infection were characterized by flu-like symptoms and mild shortness of breath, requiring oxygen supply only. Severely infected patients were characterized mainly by ICU admission due to respiratory support requirements.

### 2.2 | Genomic DNA extraction

Blood samples were collected from all patients during their admission. The samples were collected in EDTA tubes for genomic DNA extraction. DNA was extracted using an automated closed system, the EZ1 Advanced XL instrument (Qiagen) located at King Fahd Research Center, King Abdulaziz University. DNA purities and concentrations were assessed using a NanoDrop spectrophotometer (Thermo Fisher Scientific). DNA purity was assessed, and samples with a DNA concentration higher than 1.8 for a 260/280 ratio were used for the HLA genotyping.

### 2.3 | HLA genotyping

HLA genotyping was performed using molecular methods coupled with a multiplex solid-phase analysis, the PCR-rSSO (polymerase chain reaction–reverse sequence-specific oligonucleotide) method. DNA concentrations were adjusted to 20 ng/μl with molecular grade water for the HLA genotyping molecular assay. The test was performed using LABType SSO kits for each locus HLA-A, -B, -C, -DRB1, and -DQA1/B1 following company instructions (One Lambda). Briefly, genomic DNA was amplified

**TABLE 1** Characteristics of cohort of this study

Character	Specification	Remarks
Nationality	Bangladesh (N = 37) India (N = 27) Pakistan (N = 31)	-
Groups	Mild infection/recovered: N = 64 Admission to ICU: N = 31 Fatal N = 20	Mild/moderate infection: Fever, cough, muscular pain, mild dyspnea Severe: Fever, cough, muscular pain, low oxygen saturation (VAT), Acute Respiratory syndrome
Age in years Mean ± SD	Mild infection: 42.6 ± 12.0 <sup>**</sup> Admission to ICU: 49.8 ± 8.8 <sup>***</sup> Fatal: 54.7 ± 8.4 <sup>****</sup>	p-value <sup>****</sup> <0.001
Gender (Percentage)	Male: 89 (94%) Female: 6 (6%)	-
Outcome	Recovered: 75 (79%) Fatal: 20 (21%)	-

\*Post hoc LSD test ( $p = 0.047$ ) between mild and ICU patients.

\*\*Post hoc LSD test ( $p < 0.001$ ) between mild and fatal patients.

\*\*\*Post hoc LSD test ( $p = 0.245$ ) between ICU and fatal patients.

\*\*\*\*ANOVA (one-way analysis of variance) test.

using a biotinylated locus-specific primer, followed by a detection method using multiplex beads attached to sequence-specific oligonucleotide probes. The beads were then run on Luminex FLEXMAP 3D (Luminex Corp.) at King Fahd Research Center. HLA allele assignment was performed using HLA Fusion software version 4.3 (One Lambda).

## 2.4 | Statistical analysis

The allele frequency in the cohort was assessed for the Hardy-Weinberg equilibrium using Guo and Thompson's method as a part of Python for Population Genomics, PyPop.<sup>26</sup> Mean and standard deviation were used to describe continuous variables, such as age, while frequency and percentage were used to describe categorical variables. For small frequencies, the  $\chi^2$  test or Fisher's exact test was applied to investigate the association between two categorical variables, whereas one-way analysis of variance (ANOVA) was used to compare the three groups. Statistical significance was determined at a  $p$ -value of 0.05. The allele frequency (F%) was determined as a percentage calculated by dividing the absolute count of each allele by the total number of alleles in each group. Statistical analyses were performed using SPSS (Statistical Package for the Social Sciences) software version 26. The Bonferroni correction was applied to the  $p$ -value in that it was multiplied with the allele frequencies of >5% for each HLA locus.

## 3 | RESULTS

### 3.1 | Demographic

In this study, the association between the HLA genotype and the COVID-19 severity was assessed. The study included 95 Asian

patients who were admitted into hospitals because of their symptoms from SARS-CoV-2 viral infection. There were 37 patients from Bangladesh, 27 from India, and 31 from Pakistan. Their hospital stays ranged from 8 to 15 days. According to the severity of their

**TABLE 2** Comparison of HLA-A genotype frequencies in mild, ICU, and fatal groups

	Mild N = 128 F (%)	ICU N = 62 F (%)	Fatal N = 40 F (%)	Mild versus ICU p-value	Mild versus fatal p-value	ICU versus fatal p-value
A*01	14 (10.9)	10 (16.1)	5 (12.5)	0.313 <sup>a</sup>	0.785 <sup>a</sup>	0.613 <sup>a</sup>
A*02	24 (18.8)	10 (16.1)	8 (20.0)	0.659 <sup>a</sup>	0.861 <sup>a</sup>	0.617 <sup>a</sup>
A*03	8 (6.3)	4 (6.5)	3 (7.5)	0.957 <sup>b</sup>	0.780	0.838 <sup>b</sup>
A*11	18 (14.1)	13 (21.0)	9 (22.5)	0.227 <sup>a</sup>	0.205 <sup>a</sup>	0.854 <sup>a</sup>
A*23	1 (0.8)	1 (1.6)	1 (2.5)	0.598 <sup>b</sup>	0.382 <sup>b</sup>	0.752 <sup>b</sup>
A*24	21 (16.4)	4 (6.5)	3 (7.5)	0.057 <sup>b</sup>	0.160 <sup>b</sup>	0.838 <sup>b</sup>
A*26	6 (4.7)	6 (13.6)	3 (7.5)	0.185 <sup>a</sup>	0.490 <sup>b</sup>	0.705 <sup>b</sup>
A*29	3 (2.3)	0 (0.0)	0 (0.0)	NA	NA	NA
A*30	0 (0.0)	1 (1.6)	1 (2.5)	NA	NA	0.752
A*31	5 (3.9)	1 (1.6)	1 (2.5)	0.397 <sup>b</sup>	0.676 <sup>b</sup>	0.752 <sup>b</sup>
A*32	6 (4.7)	1 (1.6)	1 (2.5)	0.291 <sup>b</sup>	0.546 <sup>b</sup>	0.752 <sup>b</sup>
A*33	15 (11.7)	5 (8.1)	2 (5.0)	0.462 <sup>a</sup>	0.219 <sup>b</sup>	0.536 <sup>b</sup>
A*68	7 (5.5)	3 (4.8)	3 (7.0)	0.855 <sup>b</sup>	0.636 <sup>b</sup>	0.577 <sup>b</sup>

Note: The comparison between the two groups was performed with Fisher's exact test or the  $\chi^2$  test, as appropriate, at  $p = 0.05$ ; N, number of alleles; NA, not applicable; F(%), allele frequency.

<sup>a</sup> $\chi^2$  test.

<sup>b</sup>Fisher's exact test.

	Mild N = 128 F (%)	ICU N = 62 F (%)	Fatal N = 40 F (%)	Mild versus ICU p-value	Mild versus fatal p-value	ICU versus fatal p-value
B*07	11 (8.6)	1 (1.6)	1 (2.5)	0.064 <sup>a</sup>	0.191	0.752 <sup>a</sup>
B*08	4 (3.1)	1 (1.6)	0 (0.0)	0.542 <sup>a</sup>	NA	NA
B*13	1 (0.8)	2 (3.2)	2 (5.0)	0.205 <sup>a</sup>	0.079	0.652 <sup>a</sup>
B*14	1 (0.8)	0 (0.0)	0 (0.0)	NA	NA	NA
B*15	16 (12.5)	7 (11.3)	3 (7.5)	0.811 <sup>b</sup>	0.383	0.530 <sup>a</sup>
B*27	2 (1.6)	3 (4.8)	1 (2.5)	0.186 <sup>a</sup>	0.153	0.552 <sup>a</sup>
B*35	27 (21.1)	7 (11.3)	3 (7.5)	0.098 <sup>b</sup>	0.050	0.530 <sup>a</sup>
B*37	7 (5.5)	5 (8.1)	3 (7.5)	0.790 <sup>b</sup>	0.636	0.917 <sup>a</sup>
B*38	3 (2.3)	4 (6.5)	2 (5.0)	0.159 <sup>a</sup>	0.388	0.761 <sup>a</sup>
B*40	13 (10.2)	6 (9.7)	5 (12.5)	0.918 <sup>b</sup>	0.676	0.654 <sup>b</sup>
B*44	9 (7.0)	6 (9.7)	4 (10.0)	0.526 <sup>b</sup>	0.540	0.957 <sup>a</sup>
B*48	1 (0.8)	0 (0.0)	0 (0.0)	NA	NA	NA
B*49	1 (0.8)	0 (0.0)	0 (0.0)	NA	NA	NA
B*51	6 (4.7)	7 (11.3)	6 (15.0)	0.091 <sup>b</sup>	0.027	0.583 <sup>b</sup>
B*52	10 (7.8)	3 (4.8)	3 (7.5)	0.447 <sup>a</sup>	0.948	0.577 <sup>a</sup>
B*55	3 (2.3)	2 (3.2)	2 (5.0)	0.722 <sup>a</sup>	0.388	0.652 <sup>a</sup>
B*56	2 (1.6)	0 (0.0)	0 (0.0)	NA	NA	NA
B*57	4 (3.1)	4 (6.5)	2 (5.0)	0.284 <sup>a</sup>	0.577	0.761 <sup>a</sup>
B*58	7 (5.5)	4 (6.5)	3 (7.5)	0.786 <sup>a</sup>	0.636	0.838 <sup>a</sup>

Note: The comparison between the two groups was performed with Fisher's exact test or the  $\chi^2$  test, as appropriate, at  $p = 0.05$ ; N = number of alleles; NA, not applicable; F%, allele frequency.

<sup>a</sup> $\chi^2$  test.

<sup>b</sup>Fisher's exact test.

infection, the patients were admitted into either an isolation ward or the ICU. All those admitted into the ward recovered ( $N = 64$ ), but only 33% of the ICU patients recovered. Twenty ICU patients died ( $N = 20$ ) due to the complications of the infection, mostly acute respiratory syndrome with multiorgan failure. Patients with mild infection were significantly younger than those admitted into the ICU ( $p = 0.047$ ) and those who died ( $p < 0.001$ , Table 1). There were 89 male (93.7%) and 6 female patients (6.3%) in this random cohort. No association was found between the participants' gender and the severity of their COVID-19 infection as 33% of the women died compared to 20% of the men.

### 3.2 | Frequency of HLA class I genotype among COVID-19 patients

All data fit the Hardy-Weinberg equilibrium after considering the corrected  $p$ -value, as shown in Table S1. The HLA class I genotype was compared between the three groups with varying degrees of infection: mild, severe, and fatal. Table 2 shows the frequency of HLA-A alleles between the groups. The most commonly reported

**TABLE 3** Comparison of HLA-B genotype frequencies in mild, ICU, and fatal groups

HLA-A allele was A\*02 (18.8%) in patients with mild infection; A\*11 (21%) in subjects admitted into the ICU; and A\*11 (22.5%) in the fatal cases. No statistically significant difference was found between the compared groups in terms of the frequency of the HLA-A allele genotype.

Table 3 shows the HLA-B allele frequency between the three COVID-19 groups. The most common HLA-B allele was B\*35 (21.1%) in the mildly infected patients; B\*15 and B\*35 (11.3%) in the ICU patients; and B\*51 (15%) in the fatal cases. HLA-B\*35 was more significantly reported in patients with mild infection than those in the fatal group (21.1% vs. 7.5%). However, the  $p$ -value was borderline ( $p = 0.050$ ) and became insignificant after the correction of the  $p$ -value ( $p_c = 0.4$ ). HLA-B\*51 was more significantly observed in the fatal group than in the mild infection group (15% vs. 4.7%,  $p = 0.027$ ). However, the  $p$ -value was insignificant after the statistical correction ( $p_c = 0.216$ ). Table 4 shows the frequency of HLA-C alleles among the infected Asian patients. The most common HLA-C allele in the patients with mild infection, ICU patients, and fatal group was C\*07 (23.4%, 24.2%, 22.5%, respectively). No statistically significant difference was found between the compared groups in terms of the frequency of the HLA-C allele genotype.

**TABLE 4** Comparison of HLA-C genotype frequencies in mild, ICU, and fatal subjects

	Mild N = 128 F (%)	ICU N = 62 F (%)	Fatal N = 40 F (%)	Mild versus ICU p-value	Mild versus fatal p-value	ICU versus fatal p-value
C*01	5 (3.9)	5 (8.1)	3 (7.5)	0.229 <sup>a</sup>	0.352 <sup>b</sup>	0.276 <sup>b</sup>
C*02	1 (0.8)	1 (1.6)	1 (2.5)	0.598 <sup>b</sup>	0.382 <sup>b</sup>	0.752 <sup>b</sup>
C*03	11 (8.6)	3 (4.8)	2 (5.0)	0.353 <sup>b</sup>	0.458 <sup>b</sup>	0.971 <sup>b</sup>
C*04	26 (20.3)	7 (11.3)	3 (7.5)	0.124 <sup>a</sup>	0.061 <sup>b</sup>	0.530 <sup>b</sup>
C*06	12 (9.4)	11 (17.7)	7 (17.5)	0.097 <sup>a</sup>	0.157 <sup>a</sup>	0.975 <sup>a</sup>
C*07	30 (23.4)	15 (24.2)	9 (22.5)	0.909 <sup>a</sup>	0.902 <sup>a</sup>	0.844 <sup>a</sup>
C*08	11 (8.6)	4 (6.5)	2 (5.0)	0.608 <sup>b</sup>	0.458 <sup>b</sup>	0.761 <sup>b</sup>
C*12	15 (11.7)	4 (6.5)	3 (7.5)	0.257 <sup>b</sup>	0.451 <sup>b</sup>	0.838 <sup>b</sup>
C*14	3 (2.3)	3 (4.8)	3 (7.5)	0.357 <sup>a</sup>	0.125 <sup>b</sup>	0.577 <sup>b</sup>
C*15	13 (10.2)	6 (9.7)	5 (12.5)	0.918 <sup>a</sup>	0.676 <sup>a</sup>	0.654 <sup>a</sup>
C*16	1 (0.8)	3 (4.8)	2 (5.0)	0.068 <sup>b</sup>	0.786 <sup>b</sup>	0.971 <sup>b</sup>

Note: The comparison between the two groups was performed with Fisher's exact test or the  $\chi^2$  test, as appropriate, at  $p = 0.05$ . N, number of alleles; NA, not applicable; F%, allele frequency.

<sup>a</sup> $\chi^2$  test.

<sup>b</sup>Fisher's exact test.

**TABLE 5** Comparison of HLA-DRB1 genotype frequencies in mild, ICU, and fatal subjects

	Mild N = 128 F (%)	ICU N = 62 F (%)	Fatal N = 40 F (%)	Mild versus ICU p-value	Mild versus fatal p-value	ICU versus fatal p-value
DRB1*01	9 (7.0)	4 (6.5)	3 (7.5)	0.882 <sup>a</sup>	0.927 <sup>a</sup>	0.838 <sup>a</sup>
DRB1*03	15 (11.7)	6 (13.6)	4 (10.0)	0.674 <sup>b</sup>	0.764 <sup>a</sup>	0.957 <sup>a</sup>
DRB1*04	7 (5.5)	5 (8.1)	2 (5.0)	0.490 <sup>b</sup>	0.909 <sup>a</sup>	0.550 <sup>a</sup>
DRB1*07	20 (15.6)	11 (17.7)	6 (15.0)	0.711 <sup>b</sup>	0.924 <sup>b</sup>	0.717 <sup>b</sup>
DRB1*08	3 (2.3)	0 (0.0)	0 (0.0)	NA	NA	NA
DRB1*10	11 (8.6)	6 (13.6)	4 (10.0)	0.806 <sup>b</sup>	0.785 <sup>a</sup>	0.957 <sup>a</sup>
DRB1*11	9 (7.0)	1 (1.6)	1 (2.5)	0.117 <sup>a</sup>	0.290 <sup>a</sup>	0.752 <sup>a</sup>
DRB1*12	9 (7.0)	2 (3.2)	1 (2.5)	0.292 <sup>a</sup>	0.290 <sup>a</sup>	0.832 <sup>a</sup>
DRB1*13	9 (7.0)	7 (11.3)	7 (17.5)	0.322 <sup>b</sup>	0.049 <sup>a</sup>	0.374 <sup>b</sup>
DRB1*14	9 (7.0)	4 (6.5)	4 (10.0)	0.882 <sup>a</sup>	0.540 <sup>a</sup>	0.515 <sup>a</sup>
DRB1*15	27 (21.1)	15 (24.2)	7 (17.5)	0.629 <sup>b</sup>	0.621 <sup>b</sup>	0.422 <sup>b</sup>
DRB1*16	0 (0.0)	1 (1.6)	1 (2.5)	NA	NA	NA

Note: The comparison between the two groups was performed with Fisher's exact test or the  $\chi^2$  test, as appropriate, at  $p = 0.05$ ; N, number of alleles; NA, not applicable; F%, allele frequency.

<sup>a</sup> $\chi^2$  test.

<sup>b</sup>Fisher's exact test.

### 3.3 | Frequency of HLA class II genotype among COVID-19 patients

The frequency of HLA class II alleles (DRB1 and DQB1) between mildly and severely infected groups was examined. The most common HLA-DRB1 allele in the mildly infected patients, in the ICU

patients, and in the fatal group was DRB1\*15 (21.1%, 24.2%, and 17.5%, respectively) with no statistical difference between the three groups. Table 5 shows that DRB1\*13 was observed more in the dead subjects than in the mildly infected patients (17.5% vs. 7.0%,  $p = 0.049$ ). However, the  $p$ -value was insignificant after the  $p$ -value correction. There was no significant difference between the ICU and

**TABLE 6** Comparison of HLA-DQA1 genotype frequencies in mild, ICU, and fatal subjects

	Mild N = 128 F (%)	ICU N = 62 F (%)	Fatal N = 40 F (%)	Mild versus ICU p-value	Mild versus fatal p-value	ICU versus fatal p-value
DQA1*01	64 (50.0)	37 (60.0)	26 (65.0)	0.210 <sup>a</sup>	0.097 <sup>a</sup>	0.589 <sup>a</sup>
DQA1*02	21 (16.4)	11 (17.7)	6 (15.0)	0.818 <sup>a</sup>	0.833 <sup>a</sup>	0.717 <sup>a</sup>
DQA1*03	8 (6.3)	5 (8.1)	2 (5.0)	0.642 <sup>a</sup>	0.771 <sup>b</sup>	0.550 <sup>b</sup>
DQA1*04	2 (1.6)	1 (1.6)	1 (2.5)	0.979 <sup>b</sup>	0.696 <sup>b</sup>	0.752 <sup>b</sup>
DQA1*05	24 (18.8)	6 (13.6)	4 (10.0)	0.108 <sup>a</sup>	0.195 <sup>b</sup>	0.957 <sup>b</sup>
DQA1*06	9 (7.0)	2 (3.2)	1 (2.5)	0.292 <sup>b</sup>	0.290 <sup>b</sup>	0.832 <sup>b</sup>

Note: The comparison between the two groups was performed with Fisher's exact test or the  $\chi^2$  test, as appropriate, at  $p = 0.05$ ; N, number of alleles; NA, not applicable; F%, allele frequency.

<sup>a</sup> $\chi^2$  test.

<sup>b</sup>Fisher's exact test.

dead-infected patients in the frequency of this allele (11.3% vs. 17.5%,  $p = 0.347$ , respectively).

Regarding the HLA-DQA1 genotype, the most common allele was DQA1\*01 with 50% frequency in the subjects with mild infection and 65% frequency in both the ICU-admitted subjects and in the fatal cases. No statistically significant difference was found between the three groups in terms of the frequency of the HLA-DQA1 genotype alleles, as illustrated in Table 6. In terms of the HLA-DQB1 genotype, the most common allele was DQB1\*06 (29.7%) in the mildly infected patients while the most common allele in the ICU and in the fatal group patients was DQB1\*05 with allele frequency 37.1% versus 37.5, respectively. No statistically significant difference was found between the compared groups in terms of the frequency of the HLA-DQB1 genotype alleles, as shown in Table 7.

## 4 | DISCUSSION

With the increased number of cases of COVID-19 worldwide, a remarkable increase in confirmed cases was obvious in some populations. Although the variation in infection susceptibility and severity between populations cannot be explained, genetic and environmental factors play a fundamental role in virus pathogenesis. Polymorphism in the HLA system is one of the genetic factors that might be responsible for the variations in virus susceptibility and severity. There is a discrepancy between published studies in identifying an HLA allele as a marker of infection susceptibility or infection severity. This might be attributed to the extensive HLA polymorphism between different populations, in which a common allele in a certain population might be less represented in other populations. On the other hand, different HLA alleles from different populations might share similar peptide-binding sites and can bind the same viral peptides.<sup>27</sup> Due to the high rate of infection among South Asians, examining the HLA polymorphism among those patients was of interest. The association between the HLA system and COVID-19 infection among Saudi patients is under investigation by our group for future comparison.

In this study, patients from South Asia (India, Pakistan, and Bangladesh) infected with SARS-CoV-2 were HLA genotyped. This study showed significant differences in the frequency of HLA class I alleles between the mild and fatal groups. HLA class I antigens are the first line of defense; after the virus entry, viral peptides are complexed with intracellular HLA class I molecules and presented on the cell surface for recognition by cytotoxic T lymphocytes.<sup>28</sup> In the examined cohort, there was a significant increase in the frequency of HLA-B\*51 in the fatal group compared to the mild infection group. On the other hand, there was a significant increase in the frequency of HLA-B\*35 among the mildly infected group compared to the fatal group. Although the p values became insignificant after the statistical correction, the preliminary data can be investigated further with a larger sample size to address these associations. This initial data might suggest that some alleles can have a protective role against infection mortality or can induce infection severity. Interestingly, a study from China examined the capacity of HLA class I antigens to present SARS-CoV-2 peptides; it showed that HLA-B\*35 specific antigens are characterized by a high peptide loading capacity compared to other HLA-B proteins, which might be associated with an efficient immune response.<sup>27</sup> To assess the role of HLA alleles in inducing a significant number of positive cases, it is necessary to assess the allele frequency in a normal population. Unfortunately, data regarding the HLA allele frequency in a normal population from South Asia is limited. Patel et al.<sup>29</sup> identified B\*35 as a common HLA antigen among a group of donors from South India. Another study from North India showed that B\*35:03 and B\*51:01 are common alleles in this area but identified B\*35:01 as a less common HLA genotype.<sup>30</sup> The antigen is also common in some populations in Pakistan but is less common in Bangladesh.<sup>31,32</sup> This might reflect the increase in the number of cases in India compared to other areas. However, this result should be taken with caution due to the small sample size, which can be confirmed by examining a larger sample size and probably at a higher resolution of the HLA typing method. Studies published from other populations, mostly European, reported different HLA alleles associated with COVID-19 severity. A study published in Spain showed a significant increase in the frequency of

**TABLE 7** Comparison of HLA- HLA-DQB1 phenotype frequencies in mild, ICU, and fatal subjects

	Mild N = 128 F (%)	ICU N = 62 F (%)	Fatal N = 40 F (%)	Mild versus ICU p-value	Mild versus fatal p-value	ICU versus fatal p-value
DQB1*02	34 (26.6)	13 (21.0)	5 (12.5)	0.402 <sup>a</sup>	0.066 <sup>a</sup>	0.273 <sup>a</sup>
DQB1*03	23 (18.0)	10 (16.1)	8 (20.0)	0.754 <sup>a</sup>	0.773 <sup>a</sup>	0.617 <sup>a</sup>
DQB1*04	1 (0.8)	0 (0.0)	0 (0.0)	NA	NA	NA
DQB1*05	32 (25.0)	23 (37.1)	15 (37.5)	0.085 <sup>a</sup>	0.124 <sup>a</sup>	0.967 <sup>a</sup>
DQB1*06	38 (29.7)	16 (25.8)	12 (30.0)	0.578 <sup>a</sup>	0.970 <sup>a</sup>	0.643 <sup>a</sup>

Note: The comparison between the two groups was performed with Fisher's exact test or the  $\chi^2$  test, as appropriate, at  $p = 0.05$ ; N, number of alleles; NA, not applicable; F%, allele frequency.

<sup>a</sup> $\chi^2$  test.

HLA-A\*03, -B\*39, and -C\*16 among 72 severely infected patients (recovered and dead) compared to the healthy controls, but it reported a significant increase in the frequency of HLA-A\*11, -C\*01, and -DQB1\*04 in the severe nonsurvivor group compared to the severe survivor patients. However, their data became insignificant after the  $p$ -value correction.<sup>21</sup>

This study showed a significant increase in the frequency of HLA-DRB1\*13 in the fatal group compared to the mildly infected group. Although many studies focused on HLA class I antigens in examining susceptibility to viruses, HLA class II antigens can have a role in the cross-presentation of viral peptides. In this mechanism, degraded viral peptides can be presented by HLA class II antigens to activate helper T lymphocytes, inducing mainly antiviral antibody formations.<sup>33,34</sup> Therefore, HLA class II polymorphism might have a role in infection susceptibility, as reported previously in the case of the influenza virus in the Indian population and other viruses.<sup>35,36</sup> In the context of COVID-19, few studies have reported a contribution of HLA class II antigens to infection severity. A study from Spain showed a significant increase in the frequency of DQB1\*04 in the fatal group compared to the recovered patients.<sup>21</sup> A discrepancy between the studies can be attributed to many factors, including sample size and ethnic variations.

In conclusion, this study highlighted the possibility of a genetic or ethnic variation in the course of COVID-19 severity. An analysis of the combined data from different areas should be performed to identify a set of alleles responsible for infection severity. This will have an impact on the implementation of a screening program to identify individuals at risk for COVID-19.

#### ACKNOWLEDGEMENTS

The authors would like to thank the Ministry of Health's Science and Technology Unit in the Kingdom of Saudi Arabia for their assistance. This study was kindly supported by King Abdulaziz City for Science and Technology (KACST) through Grant number 5-20-01-502-0009 via the Fast Track Funding Path for COVID-19.

#### CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

#### ORCID

Fatmah M. A. Naemi  <https://orcid.org/0000-0002-5701-1865>

Ashwaq Al-nazawi  <http://orcid.org/0000-0001-9528-1893>

#### REFERENCES

- Pareek M, Bangash MN, Pareek N, et al. Ethnicity and COVID-19: an urgent public health research priority. *Lancet*. 2020;395(10234):1421-1422.
- De Sousa E, Ligeiro D, Lérias JR, et al. Mortality in COVID-19 disease patients: correlating the association of major histocompatibility complex (MHC) with severe acute respiratory syndrome 2 (SARS-CoV-2) variants. *Int J Infect Dis*. 2020;98:454-459.
- Chen N, Zhou M, Dong X, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet*. 2020;395(10223):507-513.
- Jordan RE, Adab P, Cheng KK. Covid-19: risk factors for severe disease and death. *BMJ*. 2020;368:m1198.
- Kumar SU, Kumar DT, Christopher BP, Doss CGP. The rise and impact of COVID-19 in India. *Front Med*. 2020;7:250.
- Kenney AD, Dowdle JA, Bozzacco L, et al. Human genetic determinants of viral diseases. *Annu Rev Genet*. 2017;51:241-263.
- Al-Dayyan N, Venugopal D, Dhanasekaran S. Footprint of the COVID-19 pandemic in India: a study of immune landscape and other factors shielding mortality. *Anal Cell Pathol*. 2020;2020:6692739-12.
- Wang JH, Zheng X, Ke X, et al. Ethnic and geographical differences in HLA associations with the outcome of hepatitis C virus infection. *Virology*. 2009;6(1):46.
- Tan AT, Loggi E, Boni C, et al. Host ethnicity and virus genotype shape the hepatitis B virus-specific T-cell repertoire. *J Virol*. 2008;82(22):10986-10997.
- Trowsdale J, Knight JC. Major histocompatibility complex genomics and human disease. *Annu Rev Genom Hum Genet*. 2013;14:301-323.
- Wieczorek M, Abualrous ET, Sticht J, et al. Major Histocompatibility Complex (MHC) class I and MHC class II proteins: conformational plasticity in antigen presentation. *Front Immunol*. 2017;8:292.
- Dendrou CA, Petersen J, Rossjohn J, Fugger L. HLA variation and disease. *Nat Rev Immunol*. 2018;18(5):325-339.
- Blackwell JM, Jamieson SE, Burgner D. HLA and infectious diseases. *Clin Microbiol Rev*. 2009;22(2):370-385.
- Sanchez-Mazas A. A review of HLA allele and SNP associations with highly prevalent infectious diseases in human populations. *Swiss Med Wkly*. 2020;150:w20214.
- Lin M, Tseng HK, Trejaut JA, et al. Association of HLA class I with severe acute respiratory syndrome coronavirus infection. *BMC Med Genet*. 2003;4(1):9.

16. Yuan FF, Velickovic Z, Ashton LJ, et al. Influence of HLA gene polymorphisms on susceptibility and outcome post infection with the SARS-CoV virus. *Viol Sin.* 2014;29(2):128-130.
17. Wang W, Zhang W, Zhang J, He J, Zhu F. Distribution of HLA allele frequencies in 82 Chinese individuals with coronavirus disease-2019 (COVID-19). *HLA.* 2020;96(2):194-196.
18. Novelli A, Andreani M, Biancolella M, et al. HLA allele frequencies and susceptibility to COVID-19 in a group of 99 Italian patients. *HLA.* 2020;96(5):610-614.
19. Iturrieta-Zuazo I, Rita CG, García-Soidán A, et al. Possible role of HLA class-I genotype in SARS-CoV-2 infection and progression: a pilot study in a cohort of Covid-19 Spanish patients. *Clin Immunol.* 2020;219:108572.
20. Littera R, Campagna M, Deidda S, et al. Human leukocyte antigen complex and other immunogenetic and clinical factors influence susceptibility or protection to SARS-CoV-2 infection and severity of the disease course. The Sardinian experience. *Front Immunol.* 2020; 11:605688.
21. Lorente L, Martín MM, Franco A, et al. HLA genetic polymorphisms and prognosis of patients with COVID-19. *Med Intensiv.* 2020;45: 96-103.
22. Pisanti S, Deelen J, Gallina AM, et al. Correlation of the two most frequent HLA haplotypes in the Italian population to the differential regional incidence of Covid-19. *J Transl Med.* 2020; 18(1):352.
23. Nguyen A, David JK, Maden SK, et al. Human leukocyte antigen susceptibility map for severe acute respiratory syndrome coronavirus 2. *J Virol.* 2020;94(13):e00510-e00520.
24. Algaissi AA, Alharbi NK, Hassanain M, Hashem AM. Preparedness and response to COVID-19 in Saudi Arabia: building on MERS experience. *J Infect Public Health.* 2020;13(6):834-838.
25. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. *J Am Med Assoc.* 2020;323(13): 1239-1242.
26. Solberg OD, Mack SJ, Lancaster AK, et al. Balancing selection and heterogeneity across the classical human leukocyte antigen loci: a meta-analytic review of 497 population studies. *Hum Immunol.* 2008;69(7):443-464.
27. Huang S, Tan M. HLA class I genotypes customize vaccination strategies in immune simulation to combat COVID-19. *bioRxiv.* 2020.
28. Sigal LJ. Activation of CD8 T lymphocytes during viral infections. *Encycl Immunobiol.* 2016:286-290.
29. Patel JS, Patel MM, Koringa PG, et al. Human leukocyte antigen alleles, genotypes and haplotypes frequencies in renal transplant donors and recipients from West Central India. *Indian J Hum Genet.* 2013;19(2):219-232.
30. Rani R, Marcos C, Lazaro AM, Zhang Y, Stastny P. Molecular diversity of HLA-A, -B and -C alleles in a North Indian population as determined by PCR-SSOP. *Int J Immunogenet.* 2007;34(3):201-208.
31. Ali ME, Ahmed MU, Alam S, Rahman MH. HLA-A, -B and -DRB1 allele frequencies in the Bangladeshi population. *Tissue Antigens.* 2008;72(2):115-119.
32. Moatter T, Aban M, Tabassum S, Shaikh U, Pervez S. Molecular analysis of human leukocyte antigen class I and class II allele frequencies and haplotype distribution in Pakistani population. *Indian J Hum Genet.* 2010;16(3):149-153.
33. Oxenius A, Bachmann MF, Ashton-Rickardt PG, Tonegawa S, Zinkernagel RM, Hengartner H. Presentation of endogenous viral proteins in association with major histocompatibility complex class II: on the role of intracellular compartmentalization, invariant chain and the TAP transporter system. *Eur J Immunol.* 1995;25(12): 3402-3411.
34. Caillat-Zucman S, Gimenez JJ, Wambergue F, et al. Distinct HLA class II alleles determine antibody response to vaccination with hepatitis B surface antigen. *Kidney Int.* 1998;53(6):1626-1630.
35. Dutta M, Dutta P, Medhi S, Borkakoty B, Biswas D. Polymorphism of HLA class I and class II alleles in influenza A(H1N1)pdm09 virus infected population of Assam, Northeast India. *J Med Virol.* 2018; 90(5):854-860.
36. Sarri CA, Markantoni M, Stamatis C, et al. Genetic contribution of MHC class II genes in susceptibility to West Nile virus infection. *PLOS One.* 2016;11(11):e0165952.

#### SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

**How to cite this article:** Naemi FMA, Al-adwani S, Al-khatibi H, Al-nazawi A. Association between the HLA genotype and the severity of COVID-19 infection among South Asians. *J Med Virol.* 2021;1-8. <https://doi.org/10.1002/jmv.27003>