

ORIGINAL RESEARCH ARTICLE

Association of *TYK2* rs2304256 and rs12720270 single-nucleotide polymorphisms with COVID-19 severity

Arindita Das^{1†*}, Md. Tarek Monoyar Hossain^{1†}, Nazmul Ahsan¹, and Gazi Nurun Nahar Sultana^{2*}

¹Department of Genetic Engineering and Biotechnology, Faculty of Biological Sciences, University of Dhaka, Dhaka, Bangladesh

²Genetic Engineering and Biotechnology Research Laboratory, Centre for Advanced Research in Sciences (CARS), University of Dhaka, Dhaka, Bangladesh

Abstract

Variability in COVID-19 disease severity suggests that host genetic factors influence immune responses and disease outcomes. Tyrosine kinase 2 (*TYK2*), encoded by the *TYK2* gene, helps regulate immune responses and inflammation. However, the impacts of its variants on disease progression remain unclear. This study investigated the association between single-nucleotide polymorphisms (SNPs) in the *TYK2* gene and COVID-19 severity. The study enrolled COVID-19 patients, with one group presenting severe symptoms and another group with mild symptoms. After DNA extraction, a specific region of the *TYK2* gene was amplified and sequenced. Three SNPs were identified within the targeted region: rs2304255 (C>T), rs2304256 (C>A), and rs12720270 (G>A). Among these, rs2304256 (a missense variant) and rs12720270 (an intronic variant) showed significantly different allele frequencies between mild and severe COVID-19 groups. Computational predictions suggested that the identified missense mutations do not directly affect *TYK2* protein function. However, expression quantitative trait locus analysis from the Genotype-Tissue Expression database indicated that rs2304256 and rs12720270 are associated with increased *TYK2* expression. Furthermore, RNA sequencing data analysis suggested that these variants promote exon 8 inclusion in *TYK2* mRNA. As higher levels of full-length *TYK2* potentially cause cytokine over-signaling, SNPs rs2304256 and rs12720270 may contribute to exaggerated inflammatory responses during COVID-19, increasing disease severity.

Keywords: *TYK2* gene; COVID-19 severity; Immune response; Genetic susceptibility

[†]These authors contributed equally to this work.

*Corresponding authors:

Gazi Nurun Nahar Sultana
(nngazi@du.ac.bd);
Arindita Das
(arindita.das@du.ac.bd)

Citation: Das A, Hossain MTM, Ahsan N, Sultana GNN. Association of *TYK2* rs2304256 and rs12720270 single-nucleotide polymorphisms with COVID-19 severity. *Gene Protein Dis.* 2026;5(3):026060005. doi: 10.36922/GPD026060005

Received: February 4, 2026

Revised: August 26, 2026

Accepted: September 1, 2026

Published online: September 18, 2026

Copyright: © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

The 2019 coronavirus disease (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is regarded as one of the deadliest pandemics of modern times. COVID-19 generally results in mild symptoms for many infected individuals, but can lead to severe complications, such as acute respiratory distress syndrome and multi-organ failure for others.¹ While established clinical risk factors such as concomitant chronic diseases (e.g., obesity, diabetes, asthma) are known to predispose individuals to

severe symptoms, it has become increasingly clear that host genetic factors also play a substantial role in determining an individual's susceptibility to SARS-CoV-2 infection and the severity of COVID-19.¹ The significant inter-individual variability observed in host responses to SARS-CoV-2 infection, ranging from asymptomatic carriage to lethal outcomes, strongly suggests that an individual's genetic makeup profoundly shapes their disease trajectory. This foundational understanding is crucial for moving toward personalized risk assessments and targeted interventions in infectious disease management.

Among the genes implicated in immune responses to COVID-19, tyrosine kinase 2 (*TYK2*) has emerged as a gene of significant interest. The expression of *TYK2* is ubiquitous, though the expression level varies among different cell types. *TYK2* is a non-receptor tyrosine protein kinase encoded by the *TYK2* gene on chromosome 19 in humans; it is a member of the Janus kinase (JAK) family and an important component of the JAK/signal transducers and activators of transcription (STAT) signal transduction pathway.²

Tyrosine kinase 2 lacks its own extracellular ligand-binding domain, unlike receptor tyrosine kinases. Instead, it associates with the cytoplasmic domains of various cytokine receptors, including those for type I interferons (IFN) and interleukins (ILs) such as IL-6, IL-10, IL-12, IL-22, and IL-23.³ In addition, type I and type II cytokine receptors lack intrinsic kinase activity and depend on specific binding to selective JAKs for kinase activity. Upon cytokine binding to its cognate receptor, *TYK2*, along with other JAKs, becomes activated. This activation leads to the phosphorylation of specific tyrosine residues on the receptor subunits, which then serve as docking sites for STAT proteins.² Various other cytokines induce *TYK2* phosphorylation, although *TYK2* is not essential for their signal transduction. Notable examples include cytokines that use the gp130 receptor subunit, such as IL-6, leukemia inhibitory factor, oncostatin M, ciliary neurotrophic factor, cardiotrophin-1, IL-11, and IL-27.⁴

The type I IFN family contains more than 20 members, with IFN-alpha subtypes and IFN-beta being the most prominent. These type I IFNs signal through the IFNAR1 and IFNAR2 receptors, interacting with *TYK2* and JAK1, respectively. *TYK2* interacts functionally with the IFNAR1 receptor subunit, enhancing ligand binding, and has a role in stable *IFNAR1* gene expression.⁵ Activity of *TYK2* is crucial for initiating type I IFN responses and combating viral infections during their early stages. Mutant human cell lines that lack *TYK2* show no response to IFN-alpha and have diminished responses to IFN-beta, which is mainly due to a significant decrease in IFNAR1 surface

expression, reducing high-affinity IFN-alpha binding.^{4,6} JAK family proteins, including *TYK2*, are multi-domain structures that share seven homologous regions (JH1–JH7). *In vitro* studies conducted using *TYK2* deletion mutants revealed that the main *TYK2* binding region for IFN alpha/beta receptor chain 1 lies within parts of JH6 and JH7, while JH3–5 regions help in the stability of *TYK2*–IFNAR1 complexes.^{4,7}

Single-nucleotide polymorphisms (SNPs) in *TYK2* can render an individual susceptible to autoimmune and inflammatory diseases, as the gene plays a crucial role in regulating the immune response.⁸ Meta-analyses suggest that various SNPs in the *TYK2* gene, such as rs2304256, rs12720270, rs12720356, rs34536443, and rs35018800, are associated with diverse autoimmune conditions.⁹ A recent study reveals a functional link between *TYK2* and the suppressors of cytokine signaling (SOCS1 and SOCS3) in severe COVID-19 cases. Studies have reported lower *TYK2* expression in COVID-19 patients than in healthy controls.⁸ In contrast, certain genetic variants in the *TYK2* gene that increase *TYK2* expression are associated with an increased risk of severe, life-threatening COVID-19.¹⁰

Thus, *TYK2* plays a pivotal role in mediating responses to IFNs and cytokines that are essential for antiviral immunity. While several studies have highlighted associations between *TYK2* SNPs and immune modulation, their precise impact on COVID-19 severity remains insufficiently explored. Investigating the role of these variants may help bridge this knowledge gap and enhance our understanding of host–genome interactions with SARS-CoV-2. Therefore, this study aimed to determine the frequency of specific *TYK2* gene SNPs among Bangladeshi individuals with severe COVID-19 compared to those with mild symptoms. It also sought to evaluate whether these genetic variations are risk factors for severe COVID-19 and to explore their potential effects on the *TYK2* protein that may contribute to heightened disease severity.

2. Materials and methods

2.1. Samples and data collection

Ethical clearance was obtained from the Ethical Review Committee of the Faculty of Biological Sciences, University of Dhaka, Bangladesh (Ref No. 127/Biol.Sc.). Venous blood samples and data were collected from 72 COVID-19 patients, comprising 34 individuals with mild COVID-19 and 38 with severe COVID-19 outcomes. Individuals with mild COVID-19 had flu-like symptoms, such as a runny nose, cough, and manageable fever. Severe cases were classified as those requiring hospitalization for ventilation support or intensive care in an intensive care unit due to serious complications caused by COVID-19. None of the

patients in the study had received a COVID-19 vaccine. Additionally, patients with comorbidities were excluded to focus solely on genetic factors influencing the severity of COVID-19. All were citizens of Bangladesh and belonged to the same ethnic group. Written informed consent was obtained from each patient. In addition, RNA sequencing (RNA-seq) data of 200 genotyped individuals were collected from the European Genome-Phenome Archive through an official application to the respective authorities.

2.2. DNA extraction, quantification, and amplification by polymerase chain reaction

Total genomic DNA was isolated from blood samples using the Wizard® Genomic DNA Extraction Kit (A1620, Promega, United States). After agarose gel electrophoresis, Quantity One software (version 4.6.6) was used to visualize DNA bands and capture images. Utilizing a NanoDrop® spectrophotometer (Thermo Fisher Scientific, United States), DNA was quantified, and its purity was assessed.

Although *TYK2* is a large gene (30,040 bp), this study focused on a 519-bp segment spanning exon 8 and intron 7. This region was selected because it includes functionally relevant polymorphisms previously implicated in immune regulation and disease susceptibility.^{11,12} Therefore, sequencing this specific exon-intron region was a biologically justified approach rather than an exploratory analysis of the entire gene. To amplify the target region, primer pairs were designed using Primer3 (version 2.6.1) and NCBI Primer-BLAST, and the primers were validated by performing *in silico* polymerase chain reaction (PCR) with two bioinformatics tools, namely UCSC *In-Silico* PCR (version 39x1) and PrimerDigital *In-Silico* PCR.^{13–16} Table 1 provides the necessary primer information.

Polymerase chain reaction was conducted using the ProFlex PCR system (Thermo Fisher Scientific, United States). The amplification process began with an initial denaturation at 95 °C for 5 min, followed by 35 cycles consisting of denaturation at 94 °C for 30 s, annealing at 60 °C for 30 s, and extension at 72 °C for 2 min. The process concluded with a final extension step at 72 °C lasting 7 min. The PCR products were purified using the FavorPrep™ Gel/PCR Purification Mini Kit (FAGCK001-1, Favorgen Biotech Corp., Taiwan) following the manufacturer's protocol. Agarose gel electrophoresis confirmed a single

band of purified PCR product corresponding to the expected 519 bp amplicon size in all samples (Figure S1).

2.3. Genotyping

Cycle sequencing of PCR amplicons was performed using the Applied Biosystems BigDye® Terminator v3.1 sequencing kit (4337455, Thermo Fisher Scientific, United States). The cycle sequencing (sequencing PCR) products were purified using the ethanol/ethylenediaminetetraacetic acid/sodium acetate precipitation method. The SeqStudio™ Genetic Analyzer (Applied Biosystems, Thermo Fisher Scientific, United States) was used to analyze the purified sequencing products through capillary electrophoresis. Software programs such as Chromas v. 2.6.6, BioEdit v. 7.0.9.0, and Geneious Prime were used to view, align, and edit the sequence chromatograms. The reference nucleotide sequence for the targeted region of human non-receptor *TYK2* was obtained from NCBI Nucleotide and compared with the aligned sequences to identify genetic variation.¹⁷

2.4. Bioinformatic analyses of mutation pathogenicity and impact on gene expression

Various computational tools were used in this study, including SIFT (version 6.2.1), PROVEAN (version 1.1.5), PolyPhen-2 (version 2.2.3), and PANTHER-PSEP (version 18.0), to predict the damaging impact of *TYK2* gene mutations.^{18–21} These tools evaluated how amino acid substitutions might affect *TYK2* structure and function. The mutations were considered potentially pathogenic if most of the computational tools predicted them to be deleterious. Additionally, to determine whether these variants contribute to altering global *TYK2* expression, the Genotype-Tissue Expression (GTEx) database (<https://gtexportal.org/>) was explored across multiple tissues, including whole blood. The GTEx Project serves as a comprehensive public resource for human gene expression, regulation, and their connection to genetic variation across a wide range of tissues and individuals.²²

2.5. Analysis of the impacts of the mutations in pre-mRNA splicing

To investigate the impact of the variants on pre-mRNA splicing of *TYK2*, cis-expression quantitative trait loci (cis-eQTL) were analyzed at multiple resolution levels. While

Table 1. Primers used for the amplification of the *TYK2* gene

Primer name	Primer sequence 5'–3'	Amplicon size	Annealing temperature
<i>TYK2</i> -F	GGTGGATGCTGACACAGTGCT	519	60 °C
<i>TYK2</i> -R	TGTCCGCTAGGGCCAGTTTCTT		60 °C

global gene expression analysis may mask variation in specific exon expression, eQTL analysis using RNA-seq data enables quantification of individual splicing events at the transcript, exon, and junction levels. For this purpose, RNA-seq data of primary CD14⁺ monocytes from 200 genotyped individuals were used.²³ In the cis-eQTL study, primary CD14⁺ monocytes were stimulated with influenza A virus to observe their response to a viral attack.²⁴

2.6. Statistical analysis

The sample size was determined using the standard epidemiological formula for unmatched case-control studies comparing two proportions with a desired statistical power of 80% and a two-sided significance level of $\alpha = 0.05$. For rs2304256, the minor allele frequency was 0.18 in the mild group and 0.61 in the severe group, yielding a required minimum sample size of 17 individuals per group. For rs12720270, the minor allele frequency was 0.21 in the mild group and 0.53 in the severe group, yielding a minimum sample size of 32 individuals per group. The present study included 34 mild and 38 severe COVID-19 patients, which exceeds the calculated minimum sample size for both SNPs. From the collected demographic and clinical information on severe and mild COVID-19 patients, the mean, maximum, minimum, and standard deviation (SD) were calculated. Data were expressed as means $\pm$ SD. To assess whether there were statistically significant differences in age and body mass index (BMI) between the mild and severe COVID-19 groups, independent samples *t*-tests were performed; a Chi-square test of independence was used to assess differences in sex distribution. Here, *p*-values were calculated for each variable using GraphPad Prism software (version 10.4.1). Fisher's exact test was performed to identify the association of the *TYK2* variants with severe conditions of COVID-19. Results from R v.4.1.1 included allele and genotype frequencies of the three variants, as well as descriptive statistics, frequencies, and proportions for categorical variables. Additionally, logistic regression analysis was conducted, and the outcomes were presented with genotype frequencies, proportions, odds ratios (ORs), mean differences, and 95% confidence intervals. Five models were evaluated to identify the inheritance model that best fit the study data, where *p*-values were chosen as the selection criteria. The OR was calculated for models showing positive associations. A *p*-value < 0.05 was considered statistically significant. Because five genetic inheritance models were tested for each of the three variants (15 comparisons total), a Bonferroni-corrected significance threshold of $\alpha = 0.05/15 = 0.0033$ was applied in addition to the nominal $\alpha = 0.05$.^{25,26}

Pairwise linkage disequilibrium (*D'* and *r*²) between the three genotyped variants was calculated from

individual-level joint genotype data using the R package (genetics). Haplotype frequencies across the three loci were estimated using an expectation-maximization algorithm implemented in the R package (haplo.stats).²⁷ Haplotype-specific association with COVID-19 severity was assessed using a haplotype score test.²⁸ Hardy-Weinberg equilibrium (HWE) for each of the three variants (rs2304255, rs2304256, rs12720270) was assessed separately in the mild and severe COVID-19 patient groups using chi-square goodness-of-fit tests.²⁹ The mild group was treated as the primary reference for evaluating genotyping quality, as disease-associated loci are expected to deviate from HWE in the affected group by design.

3. Results

3.1. Demographic characteristics of the study population

The patients' demographic information is shown in Table 2. Independent samples *t*-tests revealed no significant differences in mean age (*p*-value = 0.414) or BMI (*p*-value = 0.126) between the two study groups. Similarly, the chi-square test of independence indicated that observed differences in male/female percentages between the mild and severe patient groups were not statistically significant (*p*-value = 0.276). All statistical tests were two-tailed. Based on the statistical analyses of the provided data, age, sex distribution, and BMI were not confounding factors in this study.

Table 2. Demographic information of the study groups

Variable	Mild COVID-19 (<i>n</i> = 34)	Severe COVID-19 (<i>n</i> = 38)	<i>p</i> -value
Age (years), mean $\pm$ SD	34.88 $\pm$ 9.84	37.53 $\pm$ 16.92	0.414
Sex			
Male	82%	68%	0.276
Female	18%	32%	
BMI (kg/m ²), mean $\pm$ SD	23.43 $\pm$ 2.63	24.46 $\pm$ 3.02	0.126

Note: *p*-values for age and BMI were calculated using independent samples *t*-tests; the *p*-value for sex distribution was calculated using a chi-square test of independence.

Abbreviations: BMI: Body mass index; SD: Standard deviation.

3.2. Genotyping results for *TYK2* polymorphisms

By mapping the amplicons' nucleotide sequences generated in this study to the reference *TYK2* nucleotide

sequence from NCBI,¹⁷ variations at 3 positions were identified (Figure 1). Every variation has a unique reference identification number (rsID) in NCBI's dbSNP and ClinVar databases.³⁰ All variants were present in both severe and mild COVID-19 groups (Table 3). Two changes were transitions, whereas the other was a transversion. Figure S2 shows that the identified variants rs2304255 and rs2304256 are in exon 8, whereas rs12720270 is in intron 7 of the *TYK2* gene.

In this study population, the altered allele “T” of rs2304255 showed minimal difference in frequency distribution between severe and mild COVID-19 groups (5% and 3%, respectively) (Table 4). However, there is a significant difference in the frequencies of the altered alleles of rs2304256 and rs12720270 between the two patient groups. The altered allele “A” of rs2304256 had a higher frequency level in the severe group of patients compared to the mild group (61% and 18%, respectively) (Table 4), which is also higher than the frequency noted in the 1000 Genomes and ALFA Allele Frequency databases.³¹ This is similar for the altered allele “A” of rs12720270, where the

frequency in the severe group is 53% compared with 21% in the mild group (Table 4), which is again higher than in the 1000 Genomes and ALFA Allele Frequency databases.³²

3.3. Associations of rs2304256 and rs12720270 with severe COVID-19

The significance of the associations of the three variants with the disease outcome was statistically determined through Fisher's exact test. Two of them, rs2304256 and rs12720270, were found to be statistically significant (p -values = 0.0000002 and 0.000075, respectively), while the remaining variant was not. In addition, the OR of the mutant allele “A” in rs2304256 associated with severe COVID-19 is 7.05 (Table 4), i.e., individuals carrying the mutant allele have almost 7 times higher odds of developing severe COVID-19 compared to those without the mutation. Similarly, the mutant allele “A” in rs12720270 is associated with severe COVID-19, having an OR of 4.24 (Table 4).

According to the inheritance model testing results in R v.4.1.1 (Table 4), the dominant model best fits the data

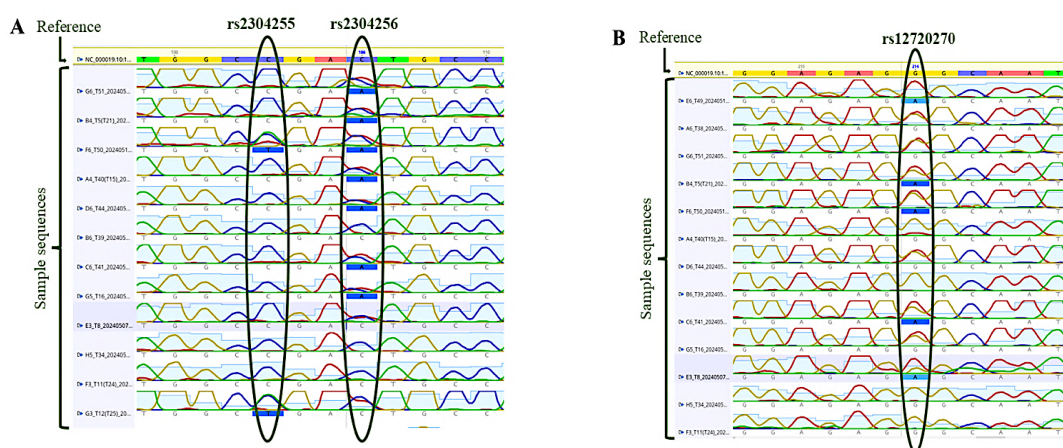


Figure 1. Partial chromatograms of selected *TYK2* sequences of samples (from all sequenced samples in this study) aligned with the reference *TYK2* sequence (from NCBI) using Geneious Prime software. (A) rs2304255 and rs2304256. (B) rs12720270.

Table 3. Variants detected in the *TYK2* gene of the study participants

Nucleotide position (GRCh38.p14)	Nucleotide change	Annotated rsID	Amino acid change	Consequence
chr19:10364973	C>T	rs2304255	G (Gly) > S (Ser)	Missense variant
chr19:10364976	C>A	rs2304256	V (Val) > F (Phe)	Missense variant
chr19:10365084	G>A	rs12720270	–	Intron variant

Table 4. Odds ratio, confidence interval, and *p*-value of the variants from Fisher's exact test and genetic inheritance models

Model	Genotype/Alleles	Mild COVID-19 patients (<i>n</i> = 34)	Severe COVID-19 patients (<i>n</i> = 38)	OR (95% CI)	<i>p</i> -value
rs2304255					
Allelic model ^F	C	66 (97%)	72 (95%)	1.83 (0.25–20.8)	0.684
	T	2 (3%)	4 (5%)		
Additive model	CC	32 (94%)	34 (90%)	1	–
	CT	2 (6%)	4 (10%)	1.88 (0.34–14.25)	0.482
	TT	0	0	NA	NA
Dominant model	CC	32 (94%)	34 (90%)	1	–
	CT & TT	2 (6%)	4 (10%)	1.88 (0.34–14.25)	0.482
Recessive model	CC & CT	34 (100%)	38 (100%)	1	–
	TT	0	0	NA	NA
Over-dominant model	CC & TT	32 (94%)	34 (90%)	1	–
	CT	2 (6%)	4 (10%)	1.88 (0.34–14.25)	0.482
rs2304256					
Allelic model ^F	C	56 (82%)	30 (39%)	7.05 (3.11–16.96)	<0.0001**
	A	12 (18%)	46 (61%)		
Additive model	CC	22 (65%)	4 (11%)	1	–
	CA	12 (35%)	22 (58%)	12.93 (4.58–48.43)	<0.0001**
	AA	0	12 (31%)	636,068,360 (0–∞)	0.992
Dominant model	CC	22 (65%)	4 (10%)	1	–
	CA & AA	12 (35%)	34 (90%)	15.6 (4.85–62.12)	<0.0001**
Recessive model	CC & CA	34 (100%)	26 (68%)	1	–
	AA	0	12 (32%)	151,233,037 (0–∞)	0.992
Over-dominant model	CC & AA	22 (65%)	16 (42%)	1	–
	CA	12 (35%)	22 (58%)	2.52 (0.98–6.7)	0.0574
rs12720270					
Allelic model ^F	G	54 (79%)	36 (47%)	4.24 (1.93–9.73)	<0.0001**
	A	14 (21%)	40 (53%)		
Additive model	GG	20 (59%)	4 (11%)	1	–
	GA	14 (41%)	28 (73%)	11.34 (3.84–43.2)	<0.0001**
	AA	0	6 (16%)	212,724,062 (0–∞)	0.992
Dominant model	GG	20 (59%)	4 (10%)	1	–
	GA & AA	14 (41%)	34 (90%)	12.1 (3.81–47.84)	<0.0001**
Recessive model	GG & GA	34 (100%)	32 (84%)	1	–
	AA	0	6 (16%)	45,203,863 (0–∞)	0.991
Over-dominant model	GG & AA	20 (59%)	10 (35%)	1	–
	GA	14 (41%)	28 (65%)	4 (1.51–11.18)	0.0063

Note: **Statistically significant.

Abbreviations: CI: Confidence interval; F: Fisher's exact test; NA: Not available; OR: Odds ratio.

and describes the association between *TYK2* variants and severe COVID-19 outcomes. In the dominant model, the *p*-values for rs2304256 and rs12720270 were 0.000017 and 0.00008, respectively; both were statistically significant. They remained statistically significant after Bonferroni correction ($\alpha = 0.0033$). No model for rs2304255 reached significance before or after correction. The high ORs of the variant alleles (heterozygous or homozygous) in rs2304256 and rs12720270 in the dominant model (15.6 and 12.1, respectively) (Table 4) suggest that the probability of playing a dominant role by the variant alleles (heterozygous or homozygous) in those two loci in developing severe COVID-19 is much higher.

Pairwise linkage disequilibrium analysis showed moderate-to-strong allelic association between rs2304256 and rs12720270 ($D' = 0.76$, $r^2 = 0.52$, p -value < 0.001), and near-complete D' but low r^2 between rs2304255 and each of the other two variants (Table 5). This is consistent with the low minor allele frequency of rs2304255 ("T" allele frequency = 4%), limiting its informativeness in this cohort.

Expectation-maximization (EM) based haplotype estimation identified six haplotypes with frequencies $> 1\%$ across the three loci. The most common haplotype, C-C-G, had a frequency of 54.3% (Figure S3). The haplotype carrying both risk alleles, C-A-A, had a frequency of 29.5% (Figure S3). Haplotype-based association testing under a dominant model showed a significant global effect (global score statistic = 23.09, $df = 4$, p -value = 0.00012). At the individual haplotype level, C-A-A was significantly associated with severe COVID-19 (score = 4.256, p -value = 0.00002), while the reference haplotype C-C-G was significantly associated with the mild phenotype (score = -3.946, p -value = 0.00008) (Table S1). The C-C-A and C-A-G haplotypes were not significantly associated with severity (p -value = 0.585 and p -value = 0.109, respectively). These findings indicate that rs2304256 and rs12720270 largely co-segregate on a shared risk haplotype, consistent with their moderate-to-strong pairwise linkage disequilibrium ($r^2 = 0.52$).

Table 5. Pairwise linkage disequilibrium among genotyped *TYK2* variants

Variant pair	D'	r^2	p -value
rs2304255 – rs2304256	0.997	0.03	0.040
rs2304255 – rs12720270	0.997	0.07	0.001
rs2304256 – rs12720270	0.76	0.52	< 0.001

All three variants were consistent with HWE in the mild group (rs2304255: $\chi^2 = 0.03$, p -value = 0.86; rs2304256: $\chi^2 = 1.56$, p -value = 0.21; rs12720270: $\chi^2 = 2.29$, p -value = 0.13). These support the absence of genotyping error. In the severe group, rs12720270 showed a statistically significant deviation from HWE ($\chi^2 = 8.67$, p -value = 0.003), consistent with the expected pattern for a locus under disease association, whereas rs2304256 ($\chi^2 = 1.70$, p -value = 0.19) and rs2304255 ($\chi^2 = 0.12$, p -value = 0.73) did not deviate significantly.

3.4. No damaging effect of the two single-nucleotide polymorphisms on tyrosine kinase 2 protein detected in silico

The potential impact of *TYK2* SNPs on the structure and function of the encoded protein was evaluated to determine the pathogenicity of the variants. For this purpose, web-based tools SIFT, PROVEAN, PANTHER-PSEP, and PolyPhen-2 were used. These tools have different terminology and corresponding score ranges, such as SIFT (neutral: score > 0.05 , deleterious: score ≤ 0.05), PROVEAN (neutral: score > 2.5 , deleterious: score ≤ 2.5), PANTHER (neutral: score < 0.5 , disease: score ≥ 0.5), and PolyPhen-2 (benign: score ≤ 0.15 , possibly damaging: $0.15 < \text{score} \leq 0.92$, probably damaging: score > 0.92). All these tools predicted that the query variants have no deleterious impact on the *TYK2* protein (Table 6).

3.5. Association of rs2304256 and rs12720270 with increased *TYK2* transcript expression

Using publicly available data from unrelated individuals in the GTEx database, rs2304256 and rs12720270 were associated with a modest increase in *TYK2* expression in several tissues, including the adrenal gland and whole blood. Violin plots in Figure 2 show that variant alleles (heterozygous or homozygous) at rs2304256 and rs12720270 upregulate *TYK2* expression in both tissues. These low p -values indicate statistically significant and strong associations between these SNPs and *TYK2* expression.

Table 6. Pathogenicity assessment of *TYK2* gene variants using *in silico* tools

Computational tools	Outcome of nucleotide variation	
	rs2304255(C>T)	rs2304256 (C>A)
SIFT	Tolerated (0.26)	Tolerated (0.11)
PROVEAN	Neutral (0.478)	Neutral (0.513)
PANTHER-PSEP	Probably benign (0.19)	Probably benign (0.13)
PolyPhen-2	Benign (0.002)	Benign (0.015)

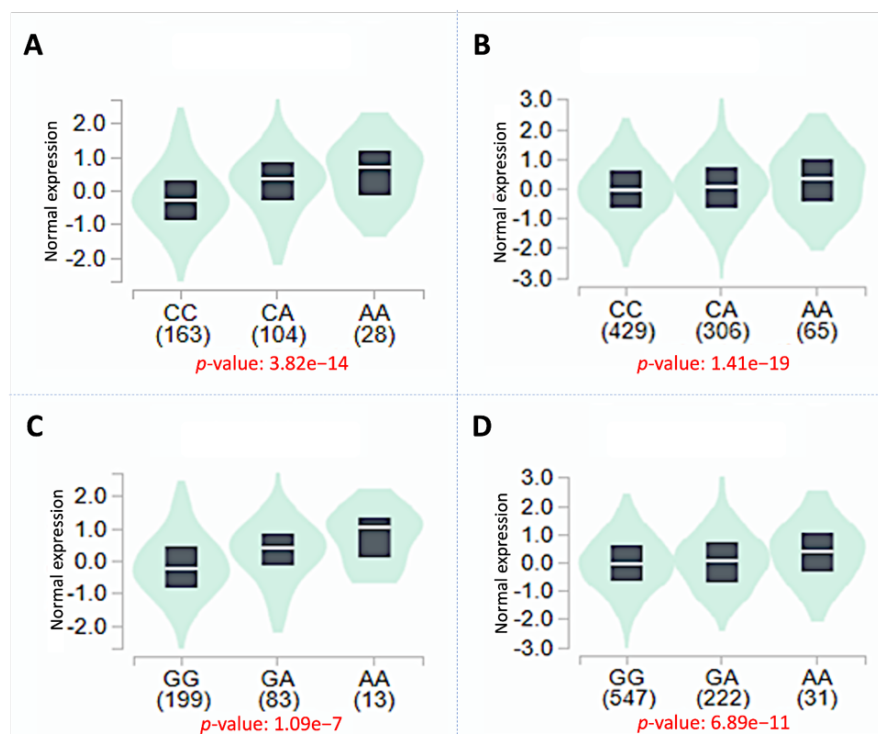


Figure 2. *TYK2* variations are associated with increased *TYK2* expression in two tissues: the adrenal gland (left) and whole blood (right).²² (A) rs23042256 is associated with increased *TYK2* expression in the adrenal gland. (B) rs23042256 is associated with increased *TYK2* expression in whole blood. (C) rs12720270 is associated with increased *TYK2* expression in the adrenal gland. (D) rs12720270 is associated with increased *TYK2* expression in whole blood. The width of the violin plots represents the density of data points. The black bars inside the violin indicate interquartile ranges, and the white line indicates the median.

3.6. Promotion of exon 8 inclusion in *TYK2* transcripts by rs2304256 and rs12720270

After RNA-seq data from primary CD14⁺ monocytes from 200 individuals were stimulated to observe changes during viral attack, cis-eQTL analysis indicated that rs2304256 and rs12720270 have a significant role in pre-mRNA processing. Comparisons of expression at multiple levels (transcript, exon, and junction level) among different genotypes in rs2304256 and rs12720270 suggest that mutant alleles in those two loci are associated with the inclusion of exon 8 in the *TYK2* mRNA during the pre-mRNA splicing event. In addition, these two variants were also found to upregulate global *TYK2* expression.

The notched box plots in Figure 3, generated using R v.4.1.1, describes the expression levels of exon 8, exon 7–9 junction, exon 7–8 junction, and exon 8–9 junction in the final *TYK2* mRNA among different genotypes (CC: $n = 143$, CA: $n = 47$, AA: $n = 10$) of the locus rs2304256, where the colored boxes represent the interquartile range, with the notches indicating confidence intervals around the median. In addition, the black dots represent individual data points, while the red diamonds represent the mean

expression for each genotype. In the mutant genotypes (both heterozygous and homozygous) of rs2304256, exon 8 expression is higher than in the wild-type genotype. So, the altered allele “A” (both heterozygous and homozygous) of rs2304256 stimulates the expression of exon 8 in the *TYK2* mRNA, which is also statistically significant (p -value = $0.0194 < 0.05$) (Figure 3A). A reduction in the expression of the exon 7 and exon 9 junction has been found in the presence of the altered allele “A” (both heterozygous and homozygous) of rs2304256, which is also statistically significant (p -value = $0.00895 < 0.05$) (Figure 3B). This finding indicates reduced exon 8 exclusion and increased inclusion in *TYK2* mRNA in the presence of the altered allele.

Moreover, the altered allele “A” (both heterozygous and homozygous) of rs2304256 is found to be associated with increased expression of the exon 7–exon 8 junction, which indicates the inclusion of exon 8 due to the altered allele (Figure 3C). Again, an increase in the expression of the exon 8 and exon 9 junction in the final *TYK2* mRNA in the presence of the altered allele “A” (both heterozygous and homozygous) of rs2304256 was found, and their association was also statistically significant (p -value =

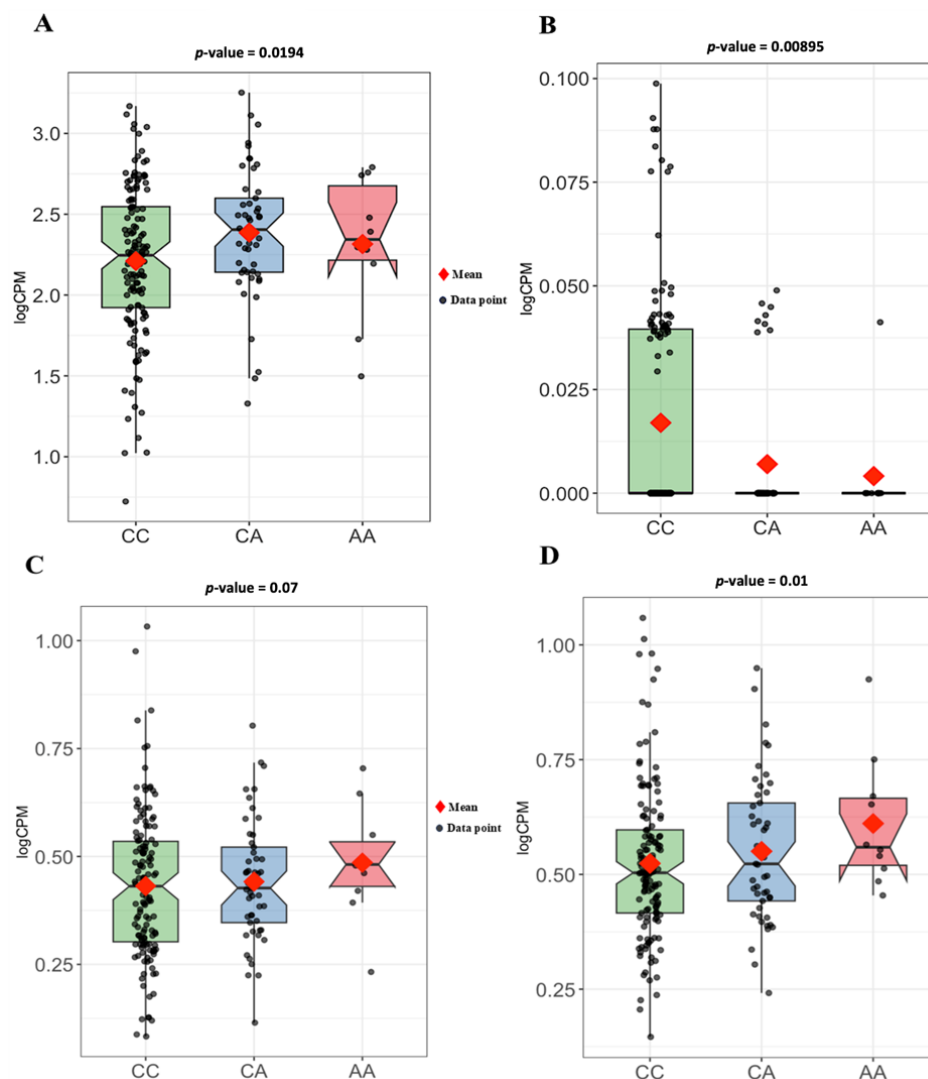


Figure 3. Rs2304256 cis-eQTL analysis on *TYK2* mRNA expression in primary CD14⁺ monocytes from genotyped donors (CC: $n = 143$, CA: $n = 47$, AA: $n = 10$). Effects of the rs2304256 locus on the expression of (A) exon 8, (B) exon 7-8 junction, (C) exon 7-8 junction, and (D) exon 8-9 junction of *TYK2*. In the notched box plot, the x-axis represents the three genotypes (CC, CA, AA), and the y-axis represents logCPM (log counts per million) expression levels of exon 8 in the *TYK2* transcript.

0.01<0.05) (Figure 3D). This also provided evidence of exon 8 inclusion in the *TYK2* mRNA due to the altered allele.

Furthermore, Figure 4 shows that the minor allele “A” of rs2304256 is associated with reduced expression of the *TYK2*-204 (ENST00000525220) transcript lacking exon 8 and increased expression of transcript ENST00000529370 having exon 8, which is a fully functional isoform. These associations were also statistically significant ($p\text{-value} = 0.0091$). Similarly, the altered allele “A” of rs12720270 is also associated with higher expression of exon 8, the exon 7-8 junction, and the exon 8-9 junction, while it is

associated with lower expression of the exon 7-9 junction and transcripts lacking exon 8. Taken together, these findings indicate that rs2304256 and rs12720270 both affect pre-mRNA splicing and promote exon 8 inclusion in the final *TYK2* mRNA.

4. Discussion

The COVID-19 pandemic, caused by SARS-CoV-2, has triggered an unparalleled global health crisis. While many infected individuals exhibit mild to moderate symptoms such as breathing difficulties, cough, fever, chills, headache, muscle pain, and loss of smell or taste, symptom severity

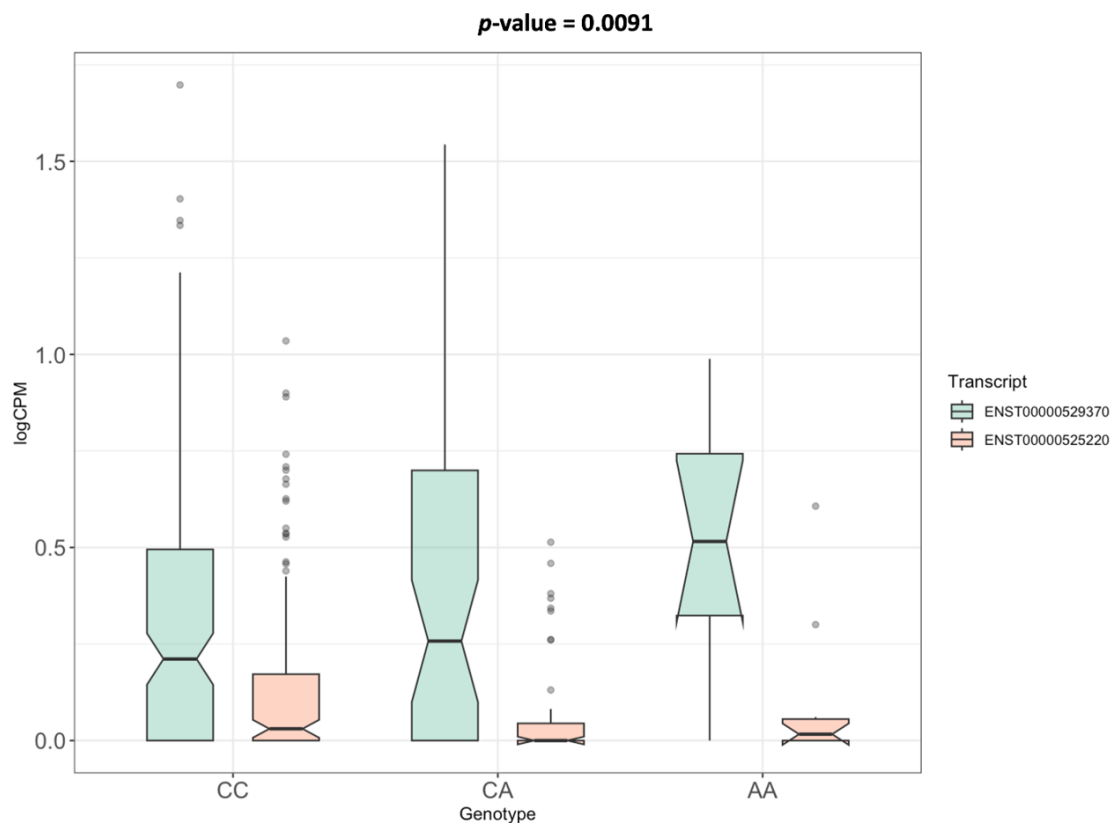


Figure 4. Effects of the genotypes of rs2304256 on the expression of *TYK2* transcripts ENST00000525220 (lacks exon 8) and ENST00000529370 (has exon 8). CPM: Counts per million.

varies significantly from person to person.³³ That is why understanding the relationship between genetic factors and COVID-19 severity has become crucial.

Tyrosine kinase 2 is a member of the JAK protein tyrosine kinase family and plays a key role in signaling by various antiviral and immunoregulatory cytokines in both immune and non-immune cells.¹⁹ Extensive genetic association studies have revealed that *TYK2* is a susceptibility gene for various autoimmune and inflammatory diseases.¹⁹ This study aimed to investigate whether *TYK2* polymorphisms are significantly associated with severe COVID-19 outcomes and whether they affect *TYK2* protein function, thereby contributing to disease severity. To achieve this goal, we compared a targeted region of the *TYK2* gene between patients with severe disease and those with mild symptoms. To specifically examine genetic influences on COVID-19 severity, the study excluded patients who were vaccinated or had pre-existing comorbidities. No significant differences in age (p -value =

0.414), BMI (p -value = 0.126), or sex distribution (p -value = 0.276) were found between the mild and severe COVID-19 patient groups (Table 2), indicating these variables were not confounding factors.

Analyzing the targeted region of the *TYK2* gene, three variants, rs2304255 (C>T), rs2304256 (C>A), and rs12720270 (G>A), were identified (Figures 1 and S2). Among the three variants, the variant allele “A” of rs2304256 (61%) and the variant allele “A” of rs12720270 (53%) were found at higher frequencies in severe patients than in mild patients. These frequencies are also higher than those annotated in the 1000 Genomes and ALFA Allele Frequency database. rs2304256 (C>A) is a non-synonymous SNP causing the amino acid change V362F, whereas rs12720270 (G>A) is an intronic variant (Table 3). Statistical analysis (Fisher’s exact test) revealed a significant association (p -value < 0.05) between two specific gene variants, rs2304256 and rs12720270, and an increased risk of severe COVID-19 (Table 4).

Further analysis using different genetic inheritance models indicated that the dominant model showed the strongest statistical support for this association, based on comparison of nominal *p*-values. In this model, the presence of even one copy of the variant allele at these loci significantly increases the likelihood of severe COVID-19. More specifically, the *p*-values for rs2304256 and rs12720270 were 0.000017 and 0.00008, respectively, and showed statistically significant associations in this model (Table 4). Even after applying a Bonferroni correction for 15 comparisons ($\alpha = 0.05/15 = 0.0033$), the Dominant-model associations for rs2304256 and rs12720270 remained significant by several orders of magnitude. rs2304255 showed no association under any model, before or after correction, consistent with its similar allele frequency between groups. High ORs (15.6 for rs2304256 and 12.1 for rs12720270) in the dominant model indicate that individuals carrying these variant alleles have a significantly higher chance of developing severe COVID-19 than those without the mutations. However, this elevated OR does not mean that every person with the mutation will develop severe COVID-19. It reflects an increased relative likelihood rather than certainty of outcome.

Given the moderate-to-strong linkage disequilibrium between rs2304256 and rs12720270 ($r^2 = 0.52$) (Table 5), we performed haplotype-based analysis to determine whether these variants contribute independent or overlapping signals. The results indicate that the two variants largely co-segregate on a shared risk haplotype (C-A-A, frequency 29.5%), which was significantly associated with severe COVID-19, while the common haplotype carrying neither risk allele (C-C-G) was associated with the mild phenotype (Table S1). This suggests that the two non-synonymous SNP associations reported above likely reflect a single underlying haplotype effect.

In silico analysis using different bioinformatic tools (SIFT, PROVEAN, PolyPhen-2, and PANTHER-PSEP) predicted that the variants rs2304256 and rs12720270 do not have any deleterious effect on the structure and function of the *TYK2* protein (Table 6). Several studies also suggest that these two variants do not alter *TYK2* structure or function in cytokine signaling.¹¹ Subsequently, we sought to determine why these two variants are frequently observed in patients with severe COVID-19 and how they contribute to the risk of developing severe COVID-19 disease. Because rs2304256 and rs12720270 did not show structural or functional differences in the *TYK2* protein, we predicted that these variants might affect pre-mRNA processing. Intronic and exonic sequences can alter gene expression by modulating pre-mRNA processing.³³ Because both variants are close to the intron 7/exon 8

boundary, they may affect exon 8 splicing efficiency, as first reported by Li *et al.*¹¹

In the present study, we sought to corroborate this mechanism using an independent RNA-seq-based cis-eQTL dataset derived from primary monocytes of 200 individuals from the EGA database²³, and to examine, for the first time, whether these established splicing-regulatory variants are associated with COVID-19 severity outcomes. Consistent with the findings of Li *et al.*,¹¹ our analysis confirmed that the minor allele “A” of rs2304256 is associated with reduced expression of the *TYK2-204* transcript (ENST00000525220), which naturally lacks exon 8 (Figure 4). This minor allele also correlated with lower levels of the exon 7–exon 9 junction but higher levels of exon 8, exon 7–exon 8 junction, and exon 8–exon 9 junction (Figure 3). Moreover, it was associated with increased activity at the donor and acceptor splice sites flanking exon 8. These findings suggest that rs2304256 likely influences pre-mRNA splicing by promoting exon 8 inclusion. The minor allele “A” of rs12720270 exhibited similar effects. Thus, both *TYK2* gene variants affect pre-mRNA splicing to promote exon 8 inclusion.

In human cells, different isoforms of *TYK2* are expressed, including the full-length functional *TYK2* and *TYK2-ΔE8* (lacking exon 8). Functional studies show that *TYK2-ΔE8* is catalytically active but cannot bind essential cytokine receptors such as IFNAR1, IL-12Rβ1, and IL-10R2, so it cannot relay JAK/STAT signaling.³⁴ Overall, our analysis and the literature indicate that *TYK2* variants rs2304256 and rs12720270 increase global *TYK2* expression (Figure 2), increase full-length functional *TYK2* by promoting exon 8 inclusion during splicing, and produce fewer *TYK2-ΔE8* variants (Figure 4). This suggests a “dosage” effect, where these variants lead to higher functional *TYK2* levels.

Studies have demonstrated that overexpressed *TYK2* can localize to the nucleus.³⁵ This suggests potential nuclear functions and may promote chronic inflammation, as proposed for JAK1 and JAK2, which can be pathogenic.³⁶ During immune regulation, *TYK2* transmits signals from various cytokine receptors. For example, *TYK2* mediates IL-12 and IL-23 signaling, and over-signaling can play pathogenic roles in autoimmune and autoinflammatory conditions.¹¹ In COVID-19 patients, rs2304256 and rs12720270 variants increased functional *TYK2* protein levels. In a sense, this can help combat SARS-CoV-2 viral attacks through type I IFN antiviral signaling mediated by receptors associated with *TYK2*. However, increased *TYK2* levels also carry risks because *TYK2* mediates signaling of many immunoregulatory cytokines, including IL-10, IL-12, IL-22, and IL-23.¹¹ Increased *TYK2* may lead

to over-signaling of these cytokines. Excessive signaling through these cytokines may contribute to inflammation and cytokine storm, which is associated with worse clinical outcomes in COVID-19 patients. Direct confirmation would require cytokine profiling, *TYK2* protein quantification, and phospho-STAT signaling assays in the patient cohort itself, which were beyond the scope of the present study.

The finding that rs2304256 and rs12720270 are associated with increased *TYK2* expression in both whole blood and the adrenal gland suggests that this upregulation is not restricted to circulating immune cells. It may also occur in tissues involved in the systemic stress and inflammatory response to severe infection. Given the adrenal gland's role in cortisol-mediated regulation of inflammation, this cross-tissue effect may be relevant to the dysregulated inflammatory states observed in severe COVID-19. However, this remains inferential without direct tissue sampling from our patient cohort.

Although the sample size was small (72 mild and severe COVID-19 patients in total) due to resource limitations, it exceeded the calculated minimum sample size for both SNPs, and the study was adequately powered (more than 80%) to detect the observed effect sizes for rs2304256 and rs12720270. Importantly, COVID-19 severity is influenced by multiple genetic and non-genetic factors beyond the two variants examined here, including other *TYK2* and non-*TYK2* polymorphisms, comorbidities, and environmental exposures. Thus, the associations reported should be interpreted as contributing risk factors rather than sole determinants of disease severity. Because this study genotyped a targeted 519 bp region rather than performing genome-wide genotyping, formal population stratification analysis (e.g., principal component analysis) could not be performed. Stratification risk was instead mitigated at the design level: all participants were recruited from a single self-reported ethnic group in Bangladesh, and the two clinical groups did not differ significantly in age, sex distribution, or BMI (Table 2).

This study has several limitations. First, the sample size was relatively limited. Further validation in independent external cohorts is needed to confirm the reproducibility of these findings. Second, all functional inferences in the present study were based on bioinformatics analyses of public databases, including eQTL and RNA-seq data. No cell or molecular functional experiments were performed to verify the exact regulatory mechanisms of the identified SNPs. Third, although the association analysis was adjusted for major confounding factors such as age, sex, and underlying comorbidities, potential residual confounding cannot be completely excluded. Notably, this

study only demonstrates a statistical association between *TYK2* genetic variants and COVID-19 severity, rather than definite causal relationships. Therefore, all mechanistic interpretations should be considered speculative. Further experimental and longitudinal evidence is necessary to clarify how these SNPs affect gene splicing, expression levels, and inflammatory responses during SARS-CoV-2 infection.

5. Conclusion

Understanding disease mechanisms and outcomes is crucial for proper diagnosis, treatment, and risk assessment. This study investigated the relationship between *TYK2* polymorphisms and COVID-19 severity. This study found that the presence of two *TYK2* alleles ("A" allele of rs2304256 and "A" allele of rs12720270) is associated with more severe COVID-19 outcomes, and statistical evidence supports this association. Haplotype analysis further showed that these two alleles largely co-occur on a single risk haplotype (frequency 29.5%). Both rs2304256 and rs12720270 were associated with modest upregulation of *TYK2* expression across multiple tissues, according to the GTEx database. RNA-seq-based cis-eQTL analysis also revealed interesting findings. These findings suggest that the association of these alleles with COVID-19 severity likely reflects higher levels of functional *TYK2*. Altered alleles rs2304256 and rs12720270 in *TYK2* promote inclusion of exon 8, leading to increased functional *TYK2* protein. Increased levels of functional *TYK2* result in over-signaling of cytokines (cytokine storm), which is associated with a greater likelihood of severe outcomes in COVID-19 patients.

Acknowledgments

We thank Professor Maxime Rotival and Etienne Patin of the Pasteur Institute, Paris, France, for their assistance with the valuable RNA-seq data.

Funding

This study was partially supported by the Biotechnology Research Centre (BRC) at the University of Dhaka (BRC-28-2023-24).

Conflict of interest

The authors declare they have no competing interests.

Author contributions

Conceptualization: Gazi Nurun Nahar Sultana

Formal analysis: Arindita Das, Md. Tarek Monoyar Hossain

Funding acquisition: Nazmul Ahsan

Investigation: Arindita Das, Md. Tarek Monoyar Hossain

Methodology: Arindita Das, Md. Tarek Monoyar Hossain

Resources: Gazi Nurun Nahar Sultana, Nazmul Ahsan
Supervision: Gazi Nurun Nahar Sultana, Nazmul Ahsan
Validation: Gazi Nurun Nahar Sultana
Visualization: Md. Tarek Monoyar Hossain
Writing–original draft: Arindita Das, Md. Tarek Monoyar Hossain
Writing–review & editing: Arindita Das, Gazi Nurun Nahar Sultana, Nazmul Ahsan

Ethics approval and consent to participate

The study was conducted in accordance with the Declaration of Helsinki, and ethical clearance was obtained from the Ethical Review Committee of the Faculty of Biological Sciences, University of Dhaka, Bangladesh (Ref No. 127/Biol.Sc.). Written informed consent was obtained from each subject who participated in the study.

Consent for publication

Written informed consent was obtained from each participant for the publication of their data, provided that their names and other potentially identifying information would not be disclosed.

Availability of data

The raw data that support the findings of this study are available upon reasonable request from the corresponding author.

References

- Hasöksüz M, Kiliç S, Saraç F. Coronaviruses and sars-cov-2. *Turk J Med Sci*. 2020;50(SI-1):549-556.
doi: 10.3906/sag-2004-127
- Hu X, Li J, Fu M, Zhao X, Wang W. The JAK/STAT signaling pathway: from bench to clinic. *Signal Transduct Target Ther*. 2021;6(1):402.
doi: 10.1038/s41392-021-00791-1
- Sohn SJ, Barrett K, Van Abbema A, *et al*. A restricted role for TYK2 catalytic activity in human cytokine responses revealed by novel TYK2-selective inhibitors. *J Immunol*. 2013;191(5):2205-2216.
doi: 10.4049/jimmunol.1202859
- Strobl B, Stoiber D, Sexl V, Mueller M. Tyrosine kinase 2 (TYK2) in cytokine signalling and host immunity. *Front Biosci*. 2011;16(1):3224.
doi: 10.2741/3908
- Shemesh M, Lochte S, Piehler J, Schreiber G. IFNAR1 and IFNAR2 play distinct roles in initiating type I interferon-induced JAK-STAT signaling and activating STATs. *Sci Signal*. 2021;14(710):eabe4627.
doi: 10.1126/scisignal.abe4627
- Velasquez L, Fellous M, Stark GR, Pellegrini S. A protein tyrosine kinase in the interferon α/β signaling pathway. *Cell*. 1992;70(2):313-322.
doi: 10.1016/0092-8674(92)90105-L
- Richter MF, Duménil G, Uzé G, Fellous M, Pellegrini S. Specific contribution of Tyk2 JH regions to the binding and the expression of the interferon α/β receptor component IFNAR1. *J Biol Chem*. 1998;273(38):24723-24729.
doi: 10.1074/jbc.273.38.24723
- Zabihi Rizi F, Ghorbani A, Zahtab P, *et al*. TYK2 single-nucleotide variants associated with the severity of COVID-19 disease. *Arch Virol*. 2023;168(4).
doi: 10.1007/s00705-023-05729-2
- Akbari M, Akhavan-Bahabadi M, Shafigh N, *et al*. Expression analysis of IFNAR1 and TYK2 transcripts in COVID-19 patients. *Cytokine*. 2022;153:155849.
doi: 10.1016/j.cyto.2022.155849
- Delgado-Wicke P, Fernández de Córdoba-Oñate S, Roy-Vallejo E, *et al*. Genetic variants regulating the immune response improve the prediction of COVID-19 severity provided by clinical variables. *Sci Rep*. 2024;14(1):20728.
doi: 10.1038/s41598-024-71476-2
- Li Z, Rotival M, Patin E, Michel F, Pellegrini S. Two common disease-associated TYK2 variants impact exon splicing and TYK2 dosage. *PLoS ONE*. 2020;15(1):e0225289.
doi: 10.1371/journal.pone.0225289
- Dendrou CA, Cortes A, Shipman L, *et al*. Resolving TYK2 locus genotype-to-phenotype differences in autoimmunity. *Sci Transl Med*. 2016;8(363):363ra149-363ra149.
doi: 10.1126/scitranslmed.aag1974
- Koressaar T, Remm M. Enhancements and modifications of primer design program Primer3. *Bioinformatics*. 2007;23(10):1289-1291.
doi: 10.1093/bioinformatics/btm091
- Untergasser A, Cutcutache I, Koressaar T, *et al*. Primer3—new capabilities and interfaces. *Nucleic Acids Res*. 2012;40(15):e115.
doi: 10.1093/nar/gks596
- Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinform*. 2012;13(1):134.
doi: 10.1186/1471-2105-13-134
- Yu B, Zhang C. In silico PCR analysis. In: Yu B., Hinchcliffe M., eds. *In Silico Tools for Gene Discovery*. Vol 760. New York, USA: Humana Press 2011:91-107.

- doi: 10.1007/978-1-61779-176-5_6
17. National Center for Biotechnology Information. Homo sapiens chromosome 19, GRCh38.p14 Primary Assembly - Nucleotide - NCBI. NIH.gov. 2025. Accessed January 29, 2025. https://www.ncbi.nlm.nih.gov/nuccore/NC_000019.10?report=fasta&from=10350533&to=10380572&strand=true
 18. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* 2003;31(13):3812-3814.
doi: 10.1093/nar/gkg509
 19. Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics.* 2015;31(16):2745-2747.
doi: 10.1093/bioinformatics/btv195
 20. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet.* 2013;76(1):7-20.
doi: 10.1002/0471142905.hg0720s76
 21. Tang H, Thomas PD. PANTHER-PSEP: predicting disease-causing genetic variants using position-specific evolutionary preservation. *Bioinformatics.* 2016;32(14):2230-2232.
doi: 10.1093/bioinformatics/btw222
 22. GTEx Portal. Broad Institute. 2025. Accessed February 11, 2025. <https://gtexportal.org/home/>
 23. EMBL-EBI, Centre for Genomic Regulation. EGA European Genome-Phenome Archive. Ega-archive.org. 2025. Accessed February 3, 2025. <https://ega-archive.org/>
 24. Rotival M, Quach H, Quintana-Murci L. Defining the genetic and evolutionary architecture of alternative splicing in response to infection. *Nat Commun.* 2019;10(1).
doi: 10.1038/s41467-019-09689-7
 25. Salkind N. Teoria Statistica Delle Classi e Calcolo Delle Probabilità. *Encycl Res Des.* 2010.
doi: 10.4135/9781412961288.n455
 26. Dunn OJ. Multiple comparisons among means. *J Am Stat Assoc.* 1961;56(293):52-64.
doi: 10.1080/01621459.1961.10482090
 27. Liu N, Zhang K, Zhao H. Haplotype-association analysis. *Adv Genet.* 2008;60:335-405.
doi: 10.1016/S0065-2660(07)00414-2
 28. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am J Hum Genet.* 2002;70(2):425-434.
doi: 10.1086/338688
 29. Emigh TH. A comparison of tests for Hardy-Weinberg equilibrium. *Biometrics.* 1980;36(4):627-642.
doi: 10.2307/2556115
 30. National Center for Biotechnology Information. tyk2 homo sapiens - SNP - NCBI. NIH.gov. 2025. Accessed February 1, 2025. <https://www.ncbi.nlm.nih.gov/snp/?term=tyk2+homo+sapiens>
 31. National Center for Biotechnology Information. rs2304256. NIH.gov. 2025. Accessed February 2, 2025. <https://www.ncbi.nlm.nih.gov/snp/rs2304256>
 32. National Center for Biotechnology Information. rs12720270. NIH.gov. 2025. Accessed February 2, 2025. <https://www.ncbi.nlm.nih.gov/snp/rs12720270>
 33. Dirican E, Bal T. COVID-19 disease severity to predict persistent symptoms: a systematic review and meta-analysis. *Prim Health Care Res Dev.* 2022;23.
doi: 10.1017/S1463423622000585
 34. Wang GS, Cooper TA. Splicing in disease: disruption of the splicing code and the decoding machinery. *Nat Rev Genet.* 2007;8(10):749-761.
doi: 10.1038/nrg2164
 35. Ragimbeau J, Dondi E, Vasserot A, Romero P, Uzé G, Pellegrini S. The Receptor Interaction Region of Tyk2 Contains a Motif Required for Its Nuclear Localization. *J Biol Chem.* 2001;276(33):30812-30818.
doi: 10.1074/jbc.M103559200
 36. Ferrao RD, Wallweber HJA, Lupardus PJ. Receptor-mediated dimerization of JAK2 FERM domains is required for JAK2 activation. *eLife.* 2018;7.
doi: 10.7554/ELIFE.38089