MEDIRES
Symbol of Quality Research

Archives of Immunology Research and Therapy
info@mediresonline.org

ISSN: 2836-5003 RESEARCH ARTICLE

Prevalence of rs11385942 and rs657152 single nucleotide polymorphisms in the world population susceptible to severe COVID-19

Mohitosh Biswas

Department of Pharmacy, University of Rajshahi, Rajshahi-6205, Bangladesh

***Corresponding Author**: Mohitosh Biswas, Department of Pharmacy, University of Rajshahi, Rajshahi-6205, Bangladesh.

Received: 12 December 2022; **Accepted:** 20 December 2022; **Published:** 28 December 2022.

Citation: Mohitosh Biswas, (2022). Prevalence of rs11385942 and rs657152 single nucleotide polymorphisms in the world population susceptible to severe COVID-19. Archives of Immunology Research and Therapy. 1(1). DOI: 10.58489/2836-5003/006

Abstract

This study aimed to investigate the predictive prevalence of rs11385942 and rs657152 single nucleotide polymorphisms (SNPs) in the world populations since these were associated with severe COVID-19. Genetic data of rs11385942 and rs657152 SNPs of the 26 populations were obtained from the 1000 Genomes project. Phenotypes were assigned as high-risk, medium-risk or no-risk based on the carrying of characteristics risk alleles. It was demonstrated that the prevalence of risk allele associated with rs11385942 SNP for developing severe COVID-19 was significantly different in various ethnic groups (Chi-square test, $p<0.00001$) with highly prevalent in South Asia (29.6%; 95% CI 27%-32%), followed by Europe (8.1%; 95% CI 6%-10%), Africa (5.3%; 95% CI 4%-7%), America (4.6%; 95% CI 3%-6%) and East Asia (0.5%; 95% CI 0%-1%), respectively. However, prevalence of risk allele associated with rs657152 SNP was not significantly different in various ethnic groups (Chi-square test, $p=0.06$) but was highly prevalent in South Asia (46.1%; 95% CI 44%-48%), followed by Africa (43.6%; 95% CI 43%-44%), Europe (36.8%; 95% CI 36%-37%), East Asia (36.3%; 95% CI 35%-38%) and America (30.5%; 95% CI 30%-31%), respectively. High-risk phenotypes associated with carrying two copies of rs11385942 variant alleles were significantly different in various ethnic groups (Chi-square test, $p=0.005$) with highly prevalent in South Asia (9.4%; 95% CI 7%-12%), followed by Africa (0.6%; 95% CI 0%-1%) and America (0.3%; 95% CI 0%-1%), respectively. Genetic associations of rs11385942 and rs657152 with severe COVID-19 should urgently assess in different ethnicities for the pathogenesis of this pandemic disease.

Keywords: *COVID-19, single nucleotide polymorphisms, severe COVID-19, pathogenesis*

Introduction

Coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) was discovered in Wuhan, China, in early 2019, and since then rapidly evolved into a global pandemic [1,2]. As of 4th July 2020, there were approximately 10.1 million confirmed SARS-CoV-2 cases worldwide, with total deaths over 0.5 million [3]. Interestingly, COVID-19 has varied clinical manifestations, having only mild symptoms in the large majority of infected patients while many patients had no symptoms at all [2,4]. Mortality rates were predominantly higher in the subgroup of patients who had severe respiratory failure considered as severe COVID-19 which may require early and prolonged support by mechanical ventilation [2,5].

The pathogenesis of COVID-19 and it's associated respiratory failure is poorly understood to date. Although, higher mortality is consistently associated with older age, male sex, and comorbidities [6-8], however, genetic risk factors associated with the severity of COVID-19 have not yet been fully clarified. A recent genome-wide association study (GWAS) attempted to delineate host genetic factors contributing to severe COVID-19 with respiratory failure [2]. This study after analyzing 8,582,968 single-nucleotide polymorphisms (SNPs) and conducting a meta-analysis of the two case–control panels involving 1980 patients at seven hospitals in

the Italian and Spanish epicenters of the SARS-CoV-2 pandemic in Europe, identified two novel SNPs i.e., rs11385942 and rs657152 with severe COVID-19. The rs11385942 SNP at locus 3p21.31, had an association signal spanning six genes *SLC6A20, LZTFL1, CCR9, FYCO1, CXCR6*, and *XCR1* whereas the rs657152 SNP at locus 9q34.2, the association signal coincided with the ABO blood group locus [2]. This study concluded that these two SNPs (rs11385942 and rs657152) were associated with severe COVID-19 as a genetic susceptibility locus.

These findings emphasized that it should explore at the population level to identify risk phenotypes associated with carrying these SNPs risk alleles for identifying people who may be at greater risk for severe COVID-19 in different ethnicities. The current study was therefore aimed to investigate the predictive prevalence of the rs11385942 and rs657152 SNPs and associated risk phenotypes susceptible to severe COVID-19 in the world populations using 1000 Genomes genetic data.

Materials and Methods

Study participants

This study utilized all genetic data pertaining to rs11385942 and rs657152 obtained from five continental groups, the Americas (AMR), Europe (EUR), Africa (AFR), South Asia (SAS) and East Asia (EAS) participated in the 1000 Genomes project as described elsewhere [9,10]. In the continent of America, four ethnic groups (CLM= Colombian in Medellin, Colombia; MXL= Mexican Ancestry in Los Angeles, California; PEL= Peruvian in Lima, Peru; PUR= Puerto Rican in Puerto Rico) consisting of 347 healthy volunteers participated in the 1000 Genomes project. Similarly, in Europe, five ethnic groups (CEU= Utah residents with Northern and Western European ancestry; FIN=Finnish in Finland; GBR= British in England and Scotland; IBS=Iberian populations in Spain; TSI= Toscani in Italy) consisting of 503 healthy volunteers participated in the 1000 Genomes project. In Africa, seven ethnic groups (ACB=African Caribbean in Barbados; ASW=African Ancestry in Southwest US; ESN= Esan in Nigeria; GWD=Gambian in Western Division, The Gambia; LWK=Luhya in Webuye, Kenya; MSL=Mende in Sierra Leone; YRI= Yoruba in Ibadan, Nigeria) consisting of 661 healthy volunteers participated in the 1000 Genomes project. In South Asia, five ethnic groups (BEB, Bengali in Bangladesh; GIH= Gujarati Indian in Houston, TX; ITU= Indian Telugu in the UK; PJL= Punjabi in Lahore, Pakistan; STU= Sri Lankan Tamil in the UK) consisting of 489 healthy volunteers participated in the 1000 Genomes project. In East Asia, five ethnic groups (CDX= Chinese Dai in Xishuangbanna, China; CHB=Han Chinese in Bejing; CHS=China, Southern Han Chinese, China; JPT= Japanese in Tokyo, Japan; KHV= Kinh in Ho Chi Minh City, Vietnam) consisting of 504 healthy volunteers participated in the 1000 Genomes project.

Allele frequency and determination of risk phenotypes

Allele and genotype frequency of rs11385942 and rs657152 were obtained from the 1000 Genomes project Phase III in line with Fort Lauderdale principles 11. Since no genotype information was available for rs657152 SNP, therefore, phenotypes were assigned based on the carrying of the rs11385942 risk allele only. Participants carrying two copies of rs11385942 SNP were predicted to consider high-risk phenotypes. In a comparative fashion, participants carrying one copy of the rs11385942 SNP were predicted to consider medium-risk phenotypes. However, participants carrying no mutation of the rs11385942 SNP were predicted to consider as no-risk/normal phenotypes.

Statistical analysis

Prevalence of rs657152 risk allele is presented as descriptive statistics as frequency and with 95% CI calculation in five continental groups consisting of 26 populations that participated in the 1000 Genomes project. However, the prevalence of rs11385942 risk allele as well as associated risk phenotypes as determined by the method described above is presented as descriptive statistics also as frequency and with 95% CI calculation in five continental groups consisting of 26 populations participated in 1000 Genomes project.

Human ethics approval

All the genetic data of humans presented in this study were obtained from the 1000 Genomes project Phase III in line with Fort Lauderdale principles. By this principle, no further human ethics approval is required to publish any results using 1000 Genomes project data as it has already been published elsewhere [9,10].

Validation of data analysis

All data analysis was primarily carried out by the principal investigator. However, all data analysis was also carried out by an independent reviewer anonymously. Finally, the author double-checked the

analyzed data provided by this independent researcher and amended it if found any anomalies.

Data availability

The datasets generated during and/or analyzed during this current study are available in the 1000 Genomes data repository (https://www.internationalgenome.org/).

Result

Using rs11385942 SNP data from the 1000 Genomes project, it was found that the variant allele was prevalent in 9.5% (95% CI 9%-10%) of the total 26 populations that participated in this genomic project. It was also found that the variant allele was prevalent highest in South Asia (29.6%; 95% CI 27%-32%), followed by Europe (8.1%; 95% CI 6%-10%), Africa (5.3%; 95% CI 4%-7%), America (4.6%; 95% CI 3%-6%) and East Asia (0.5%; 95% CI 0%-1%), respectively as shown in Figure 1. The risk allele associated with the rs11385942 SNP was prevalent and significantly different in various ethnicities as assessed by the Chi-square test, $p<0>$

However, as described in the Method section, high-risk phenotypes due to carrying two copies of the rs11385942 variant alleles were highly prevalent in South Asia (9.4%; 95% CI 7%-12%), followed by Africa (0.6%; 95% CI 0%-1%) and America (0.3%; 95% CI 0%-1%), respectively as shown in Figure 2. No high-risk phenotypes were found in East Asia and the European population. The high-risk phenotypes associated with the rs11385942 SNP were prevalent and significantly different in various ethnicities as assessed by the Chi-square test, $p=0.005$. However, medium-risk phenotypes were found highest in South Asia (40.3%; 95% CI 36%-45%), followed by Europe (16.1%; 95% CI 13%-19%), Africa (9.4%; 95% CI 7%-12%), America (8.6%; 95% CI 6%-12%) and East Asia (1%; 95% CI 0%-2%), respectively as shown in Figure 2. Medium-risk phenotypes associated with the rs11385942 SNP was also prevalent and significantly different in various ethnicities as assessed by the Chi-square test, $p<0>$

Using rs657152 SNP data from the 1000 Genomes project, it was found that the variant allele was prevalent in 38.6% (95% CI 38%-39%) of the total 26 populations that participated in this genomic project. It was also found that variant allele was prevalent highest in South Asia (46.1%; 95% CI 44%-48%), followed by Africa (43.6%; 95% CI 43%-44%), Europe (36.8%; 95% CI 36%-37%), East Asia (36.3%; 95% CI 35%-38%) and America (30.5%; 95% CI 30%-31%), respectively as shown in Figure 3. Although the risk allele associated with the rs657152 SNP was not prevalent significantly different in various ethnicities (Chi-square test, $p=0.06$) but was considerably high in different ethnic groups.

The results of this study hypothesized that ~10-39% of the world population was being at susceptible to severe COVID-19 due to carrying rs11385942 and rs657152 genetic polymorphisms, warranting urgent future studies to assess these genetic associations with prognosis and severity of COVID-19.

In summary, it was found that ~10-39% of the total 26 populations that participated in the 1000 Genomes project were identified as being as carrying risk alleles due to rs11385942 and rs657152 SNPs and were predicted to develop severe COVID-19.

Discussion

Although demographic characteristics e.g., age and sex or comorbidities of COVID-19 patients have been associated with the clinical severity of COVID-19, however, genetic associations have not been widely studied. The findings of the present study may therefore assist in designing genetic association studies providing evidence that at least two SNPs were predictively associated with severe COVID-19 and were prevalent in a considerable proportion of the world's populations.

The rs11385942 on chromosome 3p21.31 covered a cluster of six genes i.e., *SLC6A20*, *LZTFL1*, *CCR9*, *FYCO1*, *CXCR6*, and *XCR1*. Several of these genes have functions that are potentially relevant to COVID-19. For example, *SLC6A20* encodes the sodium–imino acid transporter 1 (SIT1) which interacts with angiotensin-converting enzyme 2 (ACE2) functionally [2]. It has already been well-established that ACE2 acts as a cell-surface receptor for host entry of SARS-CoV-2 [12-14]. Another example is the C-X-C motif chemokine receptor 6 (CXCR6) encoded by the *CXCR6* gene which regulates the specific location of lung CD8 T cells throughout the sustained immune response to airway pathogens. As such, a recent GWAS concluded that the rs11385942 SNP on chromosome 3p21.31 locus was involved in COVID-19 susceptibility with a possible enrichment in patients with severe COVID-19 [2].

Considerably higher and significantly different frequencies of the risk allele and associated risk phenotypes due to carrying homozygous or heterozygous of the rs11385942 SNP among the

world populations as identified in the current study suggesting that patients with this variant may develop severe COVID-19 than patients with homozygous for the non-risk allele. The findings of the current study also suggest that South Asian people are at greater risk for developing severe COVID-19 than other ethnicities.

The ABO blood groups have also been implicated in susceptibility to SARS-CoV-2 infection as reported in recent preprint studies [15,16]. Recent GWAS confirmed that blood group O was associated with a lower risk of acquiring COVID-19 than that non-O blood groups. In contrast, blood group A was associated with a higher risk of acquiring COVID-19 than non-A blood groups indicating that the rs657152 on chromosome 9q34.2 coincided with the ABO blood group locus holds considerable risk for population stratification [2]. The findings of the present study revealed that the risk allele associated with rs657152 SNP was prevalent and substantially high in various ethnicities in which South Asian people are at greater risk for developing severe COVID-19.

From the findings of the current study, it is also hypothesized that these two SNPs i.e., rs11385942 and rs657152 may also involve not only the severity of COVID-19 but also SARS-CoV-2 test positivity. This is because recent findings indicated that SARS-CoV-2 test positivity was significantly different for various ethnic COVID-19 patients [17]. Importantly, only a small proportion of suspected patients were positive for SARS-CoV-2 testing [17-19] rendering that some genetic factors may contribute to this variability in SARS-CoV-2 test positivity.

The great novelty of the present study is that it is the first study predictively identified risk alleles and associated risk phenotypes considering 1000 Genomes of genetic data of 26 populations and indicating that the risk allele and also risk phenotypes for developing severe COVID-19 were prevalent in significantly different in various ethnicities. The pharmacogenomic considerations of the two SNPs i.e., rs11385942 and rs657152 in the pathogenesis of COVID-19 may ensure precision public health in terms of providing effective prognosis, prevention, and control measures in advance. These findings warranted urgent clinical studies for assessing these genetic associations with the severity of COVID-19 for effective pathogenesis, control, and prevention of the SARS-CoV-2 pandemic virus.

Limitations

There are some limitations of this study. Firstly, it is a predictive study, therefore, the association of risk alleles due to carrying rs11385942 and rs657152 SNPs with the severity of COVID-19 may vary in real-world clinical assessment. Secondly, genotype–phenotype association of current findings may also vary. Thirdly, adjustments for all potential sources of bias e.g., underlying cardiovascular and metabolic factors, demographics i.e., age/sex was not considered for assessing the severity of COVID-19. Future clinical studies should consider and adjust these potential confounders for assessing genetic associations with the severity of COVID-19.

Conclusion

The prevalence of risk alleles and associated risk phenotypes due to carrying rs11385942 and rs657152 SNPs was significantly different in various ethnic groups where South Asian people may have a greater risk for developing severe COVID-19 as compared to other ethnicities. Genetic associations of rs11385942 and rs657152 SNPs with severe COVID-19 should urgently assess at population levels in different ethnicities for effective pathogenesis, control, and prevention of the SARS-CoV-2 pandemic virus.

Conflict Of Interest

The authors declare no conflict of interest.

Data availability

All relevant data used in this analysis is freely available upon request.

Author contribution

This is a single-authored article; therefore, it was conceived, designed, and performed all analyses by MB. This manuscript was drafted by MB also.

Acknowledgment

Department of Pharmacy, University of Rajshahi, Rajshahi-6205, Bangladesh.

Reference

1. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, Zhao X, Huang B, Shi W, Lu R, Niu P, Zhan F, Ma X, Wang D, Xu W, Wu G, Gao GF, Tan W; China Novel Coronavirus Investigating and Research Team (2019). A Novel Coronavirus from Patients with Pneumonia in China, N Engl J Med. 2020 Feb 20;382(8):727-733
2. Severe Covid-19 GWAS Group. (2020). Genomewide association study of severe Covid-

19 with respiratory failure. *New England Journal of Medicine*, *383*(16), 1522-1534.
3. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports
4. Wu, Z., & McGoogan, J. M. (2020). 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. *jama*, *323*(13), 1239-1242
5. Marini JJ, Gattinoni L. (2020), Management of COVID-19 Respiratory Distress. JAMA - J Am Med Assoc; 323(22):2329-2330. doi:10.1001/jama.2020.6825.
6. Zhou F, Yu T, Du R, et al. (2020), Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet.
7. Li, X., Xu, S., Yu, M., Wang, K., Tao, Y., Zhou, Y., ... & Zhao, J. (2020). Risk factors for severity and mortality in adult COVID-19 inpatients in Wuhan. *Journal of Allergy and Clinical Immunology*, *146*(1), 110-118.
8. Biswas, M., Rahaman, S., Biswas, T. K., Haque, Z., & Ibrahim, B. (2020). Effects of sex, age and comorbidities on the risk of infection and death associated with COVID-19: a meta-analysis of 47807 confirmed cases. *Age and Comorbidities on the Risk of Infection and Death Associated with COVID-19: A Meta-Analysis of*, *47807*.
9. Sudmant, P. H., Rausch, T., Gardner, E. J., Handsaker, R. E., Abyzov, A., Huddleston, J., ... & Korbel, J. O. (2015). An integrated map of structural variation in 2,504 human genomes. *Nature*, *526*(7571), 75-81.
10. 1000 Genomes Project Consortium. (2015). A global reference for human genetic variation. *Nature*, *526*(7571), 68.
11. 1000 Genomes Database. Http://Www.1000Genomes.Org/. https://www.internationalgenome.org/. Published 2020. Accessed April 13, 2020.
12. Shang, J., Ye, G., Shi, K., Wan, Y., Luo, C., Aihara, H., ... & Li, F. (2020). Structural basis of receptor recognition by SARS-CoV-2. *Nature*, *581*(7807), 221-224.
13. Wan Y, Shang J, Graham R, Baric RS, Li F. (2020), Receptor recognition by novel coronavirus from Wuhan: An analysis based on decade-long structural studies of SARS. J Virol.
14. Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S., ... & Wang, X. (2020). Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *nature*, *581*(7807), 215-220.
15. Zhao J, Yang Y, Huang H-P, et al. (2020), Relationship between the ABO Blood Group and the COVID-19 Susceptibility. medRxiv: 2020.03.11.20031096.
16. Zeng, X., Fan, H., Lu, D., Huang, F., Meng, X., Li, Z., ... & Hu, X. (2020). Association between ABO blood groups and clinical outcome of coronavirus disease 2019: Evidence from two cohorts. *Medrxiv*.
17. Martinez, D. A., Hinson, J. S., Klein, E. Y., Irvin, N. A., Saheed, M., Page, K. R., & Levin, S. R. (2020). SARS-CoV-2 positivity rate for Latinos in the Baltimore–Washington, DC region. *Jama*, *324*(4), 392-395.
18. Mehta, N., Kalra, A., Nowacki, A. S., Anjewierden, S., Han, Z., Bhat, P., ... & Chung, M. K. (2020). Association of use of angiotensin-converting enzyme inhibitors and angiotensin II receptor blockers with testing positive for coronavirus disease 2019 (COVID-19). *JAMA cardiology*, *5*(9), 1020-1026.
19. Reynolds, H. R., Adhikari, S., Pulgarin, C., Troxel, A. B., Iturrate, E., Johnson, S. B., ... & Hochman, J. S. (2020). Renin–angiotensin–aldosterone system inhibitors and risk of Covid-19. *New England Journal of Medicine*, *382*(25), 2441-2448.