COVID 19: An Epidemiological and Host Genetics Appraisal

Arup Ratan Bandyopadhyay¹, Diptendu Chatterjee², Kusum Ghosh³, Pranabesh Sarkar⁴

¹Professor, ²Assistant Professor (Grade III), ³Junior Research Fellow (UGC-NET), ⁴Senior Research Fellow (UGC-NET), Department of Anthropology, University College of Science, University of Calcutta, 35, Ballygunge Circular Road, Kolkata – 700019, India

Submitted: 23-04-2020 Revised: 29-04-2020 Published: 01-05-2020

ABSTRACT

Coronaviruses (CoVs) is a single single-strand RNA genome approximately 26 - 32 kb in size. Out of the seven coronaviruses, three HCoVs (Human CoVs) have been discovered that causes severe pneumonia such as Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS) and recently recognized SARS-CoV-2, which possesses varying degrees of lethality worldwide and happened to be bioterrorism in terms of the recent outbreak through human-to-human transmission from China to all over the world. Epidemiological and Clinical study on SARS-COV-2 have recently been reported world-wide but lack of data on prognosis factors including effective medicine or vaccine are yet to be clinically approved to prevent this infectious disease. Human pathogenic coronaviruses SARS-CoV-2 bind to their target cells through ACE2, which is expressed by epithelial cells of the lung, intestine, kidney, and blood vessels. The difference in distribution, maturation, and functioning of viral receptors could be considered as a possible reason for the genetic heterogeneity of ACE2, and age and sex related difference in the incidence of the disease such as, the positive correlation with ACE2 expression and age including the severity of the infection of SARS-CoV-2. Since the ACE2 location in X chromosome, therefore, the males presumable might have more morbidity and mortality by SARS-CoV-2 than females due to sex-based immunological differences like greater observable circulating level of ACE2 in males or else it may be due to the patterns of life style variables such as prevalence of smoking among the males. Additionally, the Angiotensin-Converting Enzyme 1 (ACE1) is characterized by a genetic insertion/deletion (I/D) polymorphism in intron 16, which is associated with alterations in circulating and tissue concentrations of ACE, where the study reported as D allele is associated with a reduced expression of ACE2. Nevertheless, studies from different states of Indian population on ACE I/D gene polymorphism shows higher frequency of I allele which might explain the lower prevalence of SARS-CoV-2 in Indian population and consequently be subject matter of research of SARS-CoV-2 on epidemiological and public health issues.

Key words: SARS-CoV-2; genetic polymorphism; epidemiology; public health; gender

INTRODUCTION

Coronaviruses (CoVs), belongs to the family Coronaviridae. They are enveloped viruses with a single-strand, having a RNA genome approximately 26 - 32 kb in size, which is the largest known genome for an RNA virus. The term ‘coronavirus’ is coined due to the spike projections of the virus membrane which resemblance with a crown, or corona in Latin word. CoVs are separated into four genera based on phylogeny: alpha-CoV, beta-CoV, gamma-CoV and delta-CoV.¹

Types of coronavirus

From the large family of Coronaviruses seven is known to infect Human population. Among them four (229E, NL63, OC43, and HKU1) are known to cause common cold that rarely cause death first described in 1960s.² But the other three HCoVs (Human CoVs) have been discovered that...
causes severe pneumonia such as Severe Acute Respiratory Syndrome (SARS)\(^7\) and Middle East Respiratory Syndrome (MERS)\(^8\) and recently recognized SARS-CoV-2,\(^5\) which possesses varying degrees of lethality worldwide and happened to be bioterrorism to entire human population.

**Out break history and variation**

The recent outbreak of a novel coronavirus disease (SARS-CoV-2) occurred on early December 2019, in Wuhan city, China\(^6\) with a rapid transmission from animal-to-human and from human-to-human from China and even all over the world has put the world on alert\(^2\) and declared as pandemic by WHO. Although, epidemiological\(^8,9\) and clinical\(^5,10\) features of patients with SARS-CoV-2 have recently been reported Worldwide but there are a very few data on prognostic factors of SARS-CoV-2.\(^6\) Yet, no effective drugs or vaccine are clinically approved to prevent this infectious disease. So, it is become very important to find out the strategies for prevention and it is urgently needed to understand the variations in severity and fatality of the disease in human populations.\(^10\)

The variation in severity, morbidity, mortality of disease (both communicable and non-communicable) are generally be perceived by analyzing the host genetic aspects, which is ethnic specific in general and individual in specific. Nevertheless, age and gender also being imperative factors for all kind of health and disease studies.

**Pathogenesis**

Human pathogenic coronaviruses SARS-CoV and SARS-CoV-2 bind to their target cells through ACE2, which is expressed by epithelial cells of the lung, intestine, kidney, and blood vessels.\(^11\) Patients with type 1 and/or type 2 diabetes, who are treated with ACE inhibitors and Angiotensin II type-I Receptor Blockers (ARBs) expression of ACE2 is substantially increased among them.\(^12\) Consequently, the increased expression of ACE2 would facilitate infection with COVID-19. Therefore, it can be assumed that in case of diabetes and hypertension treatment with ACE2-stimulating drugs may increase the risk of developing severe and substantial increase of mortality in COVID-19.

**Epidemiology**

Although recent studies\(^13,14,11\) claimed that, ACE-I and ARBs potentially increases the risk of COVID-19 but there is no clinically published report till date which can shed light on the question whether, ACE-I or ARB either improve or worsen the susceptibility to COVID-19 infection.\(^15\) In a study among COVID-19 patients, 17% found to have a history of co-morbidities. Among them only 7% had the fatality risk although not reported that they were treated with ACE-I and ARBs or not.\(^16\) Eventually, it is unclear without confirmation about the benefit or risk of the medication people who are using these should be stopped or not. Moreover, ACE-I and ARBs are life-saving treatments for those who are suffering from chronic conditions and discontinuation of these medications will lead to an adverse effect to their health condition leading to mortality.

**Virus and host genetics polymorphism**

The spike protein of SARS-CoV-2 had a strong binding affinity with ACE2 in human cell based on structural modeling.\(^17\) Recent reports suggested that, SARS-CoV-2 protein binds to ACE2 through Leu455, Phe486, Gln493, Asn501, and Tyr505 and postulated that residues near lysine 31, and tyrosine 41, 82, 353, 355, and 357 of the ACE2 receptor interacting with SARS-CoV-2 spike protein and helps in binding.\(^9\) Cleavage of the C-terminal segment of ACE2 by Transmembrane protease serine 2 (TMPRSS2), enhances the spike protein-driven viral entry in the host.\(^18\) Hence, the possibility of host genetic variability of the ACE2 receptor could modulate the virus intake and ultimately severity of the disease.

Not only has ACE2 facilitated the invasion of SARS virus for rapid replication, but at the same time ACE2 is depleted from the cell membrane and therefore the possibility of damaging effects of Ang II are enhanced, which eventually, result in acute deterioration of lung tissues.\(^19\) In addition to that, the Angiotensin-Converting Enzyme 1 (ACE1) is characterized by a genetic Insertion/Deletion (I/D) polymorphism in intron 16, which is associated with alterations in circulating and tissue concentrations of ACE, where the D allele is associated with a reduced expression of ACE2. Interestingly ACE2 (Xp22) and ACE (17q23.3) share only 42% of amino acid identity, and they both act as carboxypeptidases to cleave amino acids from the peptides’ carboxyl terminal.\(^20\) A further aspect that might be subject matter of investigation is the genetic predisposition for an increased risk of SARS-CoV-2 infection, related to ACE2 polymorphisms that associated with diabetes mellitus, cerebral stroke, and hypertension, specifically in Asian populations.

The difference in distribution, maturation, and functioning of viral receptors is considered as a possible reason for the age-related difference in the incidence of the disease. Previous studies demonstrated the positive correlation of ACE2 expression and the infection of SARS-CoV-2\(^2,21,6\) specifically in lung adenocarcinoma (LUAD) which demonstrated that, ACE2 expression was positively correlated with age. Middle-aged and
older adults in LUAD samples discerned an increment by ~1.2 times with every 10-year increase of age. This may partly explain the observation that the elderly are more susceptible to SARS-CoV-2 than the children.\textsuperscript{20,21,28} Although this does not necessarily mean that children are less susceptible to the infection, but infants can be infected by SARS-CoV-2 but the mortality rate is very low.\textsuperscript{23} Along with the age sexual dimorphism in affected condition may be another curtain raising issue and found to be neglected while deliberating protocols for disease control but it is very important to explore the most useful prognostic factor for individualized assessment.\textsuperscript{8}

Previous epidemiological data from the 2002-2003 for SARS and MERS epidemic indicated that there may be sex-dependent differences in disease outcomes.\textsuperscript{24} Similar trend is observed from the analysis of recent report from Chinese health authority, who confirmed the number of SARS-CoV-2 patients on the Chinese mainland has reached 76,936, and 2,442 people have died of the disease as on Feb 23, 2020 and among the deceased patients, most were old and two-thirds were males,\textsuperscript{25} which might elevate the issue of ACE2 receptor, which is located in X chromosome (Xp22) and eventually indicate sex biasness. While, using single-cell sequencing, study reported a positive correlation\textsuperscript{26} with expression of ACE2 and found to have more frequency in Asian men, which in turn can be considered as a reason for the higher prevalence of SARS-CoV-2 in this subgroup of patients than in women. In addition to that, using the large Genotype Tissue Expression (GTEx) data reported a strong positive ACE2 Expression of Quantitative Loci (eQTLs) in East Asians females which again in turn indicating Asian females are more protected against SARS-CoV and as well as related severe symptoms rather than males, who are supposed to being more susceptible and morbid. It is also revealed that women have remarkably elevated basal level of ACE2. SARS-CoV-2 infection decreases the level of ACE2 in females but the repression of ACE2 might be counteracted by higher basal ACE2 level and this higher basal ACE2 is inducible by higher sex hormone levels (Estrogen and Androgen) which protect them from this infectious disease even though they have the similar susceptibility like the male,\textsuperscript{21} eventually it opens up the question of morbidity status of menopausal women regarding the status of sex hormonal levels.

All the above-mentioned data suggests that males have more morbidity and mortality by SARS-CoV-2 than females. One reason may be due to sex-based immunological differences\textsuperscript{27} or else it may be due to the patterns of life style variables such as prevalence of smoking\textsuperscript{28} among the males. Earlier study confirmed that differences in this disease prevalence and severity are related to higher expression of ACE2, because ACE2 is the host cell receptor for SARS-CoV-2.\textsuperscript{29} Several factors such as, smoking, differences in ACE2 localization and its density in alveolar cells, hormonal asset examined to explain this difference but apart from this it is very interesting fact that ACE2 is located on the X-chromosome. This causes circulating level of ACE2 to always be higher in men.\textsuperscript{8} But the X-inactivation mechanism give rise to the alternate expression of the two alleles which in turn guarantee a heterogeneous population of ACE2 molecules in case of female. Some of which might plays a protective role towards the infection if the X-inactivation skewed towards the less SARS-CoV-2 binding prone allele.\textsuperscript{29,30} However, it may be possible that, genetic heterogeneity of ACE2 molecule can modulate infection and disease progression. These observations suggest that may be the differences in immunity, gene expression or even genetic background may contribute to the differential susceptibility and severity of SARS-CoV-2 infection in different ethnic population.\textsuperscript{31}

It is confirmed from previous studies, virus binds to the ACE-2 receptor in order to gain entry. The receptor is present in ciliated epithelial cells in the upper and lower airway, as well as in alveoli in the lower airway,\textsuperscript{2} and in some other tissues (Heart, Ovary, Brain etc.) as well.\textsuperscript{10} Higher ACE2 expression level, plays a major role in the spread of the infection, mainly by binding to ACE2 receptors in the lung.\textsuperscript{26} Therefore, the expression level and expression pattern of human ACE2 in different tissues and their genetic predisposition, may justify the route of susceptibility, symptoms, and outcome of SARS-CoV-2 infection.\textsuperscript{28} On this background concurrent study\textsuperscript{2} on the genetic polymorphism of ACE2 (rs6632677) among three populations; Asia, Europe and Africa demonstrated the virus is likely to infect the people with rs6632677 polymorphism, which is notable polymorphism among the Chinese especially those who are in SE China, but both the Europeans and Africans are devoid of such kind of polymorphism.

However, Expression Quantitative Trait Loci (eQTLs) study\textsuperscript{28} on the basis of Genotype Tissue Expression database\textsuperscript{32} (GTEx) and potential functional coding variants\textsuperscript{33} in ACE2 among populations were systematically investigated for the candidate functional coding variants in ACE2 and the Allele Frequency (AF) differences among populations. The analysis of all the 1700 variants in ACE2 gene region from the China MAP (China Metabolic Analysis Project) and 1000 Genomes Project revealed among them AFs of 62 variants located in the coding regions of ACE2 in China MAP and 1KGP (1000 Genome Project). Although mutations in these regions were not
found among the studied three populations which indicated that there was a lack of natural mutations that will resist coronavirus S-protein binding in populations. But the East Asian populations have much higher AFs in the eQTL variants associated with higher ACE2 expression in tissues compared to others (European, African, South Asian, and Ad-Mixed American). According this study, Europeans have less susceptibility to this disease but it has been found that after China Italy became the epicenter for SARS-CoV-2.An attempt to find out the ACE2 genetic variants that may be responsible for this situation through Network of Italian Genomes (NIG), three common (p.Lys26Arg, p.Gly211Arg, p.Asn720Asp) and 30 rare missense variants were identified. All these three variants represented in the Italian and European populations but extremely rare in the Asian population and these genetic variants may be of consideration as an ethnic variation and as well as might be the basis for the recent outbreak in Italy, Europe.

As mentioned earlier ACE2 (Xp22) and ACE (17q23.3) share only 42% of amino acid identity, they both act as carboxypeptidases to cleave amino acids from the peptides’ carboxyl terminal, therefore, along with ACE2, ACE polymorphism might be imperative to study worldwide in general, and specifically in Indian scenario. Therefore, Epidemiological studies are urgent and it has already been reported that, different age, sex and ethnic groups have different susceptibility of morbidity and mortality to this infectious disease along with the genetic polymorphisms.

The possible contribution of host genetic factors to the susceptibility of SARS-CoV-2 has been investigated recently from a number of possible points of view. Among them a recently reported dimension is the association of ACE I/D polymorphism and prevalence of SARS-CoV-2 infection. Since D allele is associated with a number of comorbidities including, hypertension, diabetes, cancer and heart failure and these are all are regarded as a risk factor for severe SARS-CoV-2 and the association of D allele with SARS-CoV-2 as well could be subject matter of research.

The ACE I/D polymorphism discerned important geographical variation. Although the outbreak of SARS-CoV-2 emerged in Asian countries but after that European particularly Southern European countries experienced much higher incidence of morbidity and mortality rate. Perhaps the ACE I/D genotype distribution (polymorphism) in different geographical location might raise issues regarding important role in the prevalence of the SARS-CoV-2 infection. In this context two recent studies found a positive association of cumulative incidence of SARS-CoV-2 and deletion (D) in the angiotensin-converting 1 (ACE1) gene among the European Population. The frequency of the D-allele was found to be higher in Europe and Southern Europe compared to Asia and hoist intuitive phenomenon behind the higher frequency of morbidity and mortality among the Europeans than in the Asians.

However, the scenario in Indian context apart from East Asian (China and Korea) populations characterized with low D allele frequencies, similar characteristic found in South Asian (India) population as well. Nevertheless, studies from different states of Indian population on ACE I/D gene polymorphism shows higher frequency of I allele which can be a reason for lower prevalence of SARS-CoV-2 in Indian population and might be the subject matter of research on epidemiological and public health perspectives in terms of differential spread of SARS-CoV-2 in different regions.

However, along with the genetic polymorphism study, another study also postulated that the incidence of recurrent coronavirus outbreaks in China may be due to the high diversity of coronaviruses and their animal hosts in China, and as well as concomitant food and dietary culture despite of low frequency of ACE D alleles. So till now the actual reason behind this outbreak in Asia is unclear. Recent reports of the ACE2 expression analysis in lung tissues from Asian, European and African populations are still inconclusive. So, it would also be of importance to investigate the co-receptors of SARS-CoV or SARS-CoV-2 and find out whether their abundance varies among different populations or not. Based on the available data, in conclusion, it is fairly possible to envisage the occurrence and development of SARS-CoV-2 depend on the interaction between the virus and the virus host combinations. Viral factors include virus type, mutation, viral load, viral titer, and viability of the virus in vitro. On the other hand, the individual’s characteristics factor includes genetic factors in terms of genetic polymorphisms, age, gender, nutritional and life style status, neuroendocrine-immune regulation, physical status and as a whole differential in different ethnic groups. These factors all contribute to whether an individual is infected with the virus, the duration and severity of the disease, and the reinfection.

REFERENCES


https://doi.org/10.1056/NEJMoa030747


https://doi.org/10.1056/NEJMoa1211721


https://doi.org/10.1016/S0140-6736(20)30183-5


https://doi.org/10.1101/2020.02.23.20026864


https://doi.org/10.1056/NEJMoa2001316


https://doi.org/10.1016/S0140-6736(20)30154-9


https://doi.org/10.1016/S0140-6736(20)30211-7


https://doi.org/10.1128/JVI.00127-20


https://doi.org/10.1128/JVI.00127-20


https://doi.org/10.1136/bmj.m810


https://doi.org/10.1016/S2213-2600(20)30116-8


https://doi.org/10.3760/cma.j.issn.0254-6450.2020.02.003


https://doi.org/10.1007/s11427-020-1637-5


https://doi.org/10.1128/JVI.02062-10


https://doi.org/10.1161/101.res.87.5.e1


https://doi.org/10.1016/j.jmii.2015.10.008


https://doi.org/10.1001/jama.2020.2131


https://doi.org/10.4049/jimmunol.1601896


https://doi.org/10.1101/2020.01.26.919985


https://doi.org/10.1136/jech-2016-207805


https://doi.org/10.1038/s41421-020-0147-1


https://doi.org/10.1038/sj.emboj.7600640


https://doi.org/10.20944/preprints202002.0258.v1

31. Chen Y, Shan K and Qian W. Asians and other races express the SARS-CoV-2 cell-entry receptor. Preprints 2020;

https://doi.org/10.1007/s11427-020-1637-5


Authors Contribution:
ARB-Concept and design, manuscript preparation, statistically analyzed and interpreted, critical revision of the manuscript; DC-Concept and Design of the study, Review of Literature; KG-Review of Literature, data handling, statistically analyzed; PS-Review of Literature, data handling.

Work attributed to:
Department of Anthropology, University of Calcutta.

Orcid ID:
Prof. Arup Ratan Bandyopadhyay - https://orcid.org/0000-0003-3119-0022

Source of Support: None. Conflicts of Interest: None.