Journal of Genetics and Gene Therapy

Extended Abstract

COVID-19 Immunogenetics and Immuno-Epidemiological Parameters

Attapon Cheepsattayakorn

10th Zonal Tuberculosis and Chest Disease Center, Chiang Mai, Thailand
Email: Attapon1958@gmail.com

Abstract

SARS-CoV-2 (COVID-19) infection has evolved to become a pandemic, in contrast to infections with SARS and MERS, whereas SARS-CoV-2 (COVID-19) has demonstrated having the similarities of genome sequence, receptor affinity, pathogenesis, and disease manifestation. Among the most severely affected patients, viral ribonucleic acid (RNA) has been detected in the plasma approximately 15% and viral detection in stool reveals possibility of fecal transmission. COVID-19 has been isolated in human saliva, nasopharynx and lower respiratory tract. Lacking lung biopsies or post-mortem sample investigations leads to an incomplete understanding of the pathogenesis of COVID-19 infection. The innate immune cells are born capable of producing T regulatory cell cytokine (interleukin (IL)-10 and T helper 17 (TH17) cell cytokines (IL-6 and IL-23), but inability of induction of T helper 1 (TH1) cell cytokines (type I interferons (IFNs), IFN-γ, and IL-12). A new lineage of oligoclonal T cells that express natural killer (NK)-related receptors (NKR) is formed, whereas the diversity of the T-cell receptor (TCR) repertoire decreased with age. In addition to the experimental evidence regarding the immunological features in neonates and children that are more prominent than in adults. In consideration of the pro-inflammatory response to the SARS-CoV infection, an overwhelming inflammatory reaction in aging population is a logical possibility. The mean age of COVID-19 patients is 52.4 years, whereas children and adolescents are the least likely group to be infected with the COVID-19, occurring in only 2 % of cases 19 years of age or younger. When the younger-age group get sick, they will get a mild form of COVID-19 without serious complications, with an average death rate of 0.2 %.

COVID-19 (SARS-CoV-2) is the seventh member of the Coronaviridae known to infect humans. Three of these viruses are: 1) SARS CoV-1, 2) MERS and 3) COVID-19 (SARS-CoV-2) can cause severe disease, and 4) HKU1, NL63, OC43 and 229E, are related to mild respiratory symptoms. The two notable characteristics of the COVID-19 genome are: 1) based on structural modelling and early biochemical experiments, COVID-19 is optimized for binding the human ACE2 receptor; and 2) the highly variable spike (S) protein of COVID-19 has a polybasic (furin) cleavage site at the S1 and S2 boundary via the insertion of twelve nucleotides. This event contributes to the acquisition of the three predicted O-linked glycans around the polybasic cleavage site.

Six residuals in the receptor binding domain (RBD) of the spike protein of SARS-CoV and SARS-related coronaviruses, the most variable part of the virus genome appear to be critical for binding to the human ACE2 receptor and the determining host range. Five of these six residuals are mutated in COVID-19 compared to its most closely related virus, RaTG13 sampled from a *Rhinolophus affinis* bat to which it is approximately 96% identical. COVID-19 seems to have an RBD that may bind with high affinity to ACE2 from human, non-human primate, cat, pig, and ferret. COVID-19 may bind less efficiently to ACE2 in other species related to SARS-like viruses, such as rodents and civets. Recent binding studies demonstrated that COVID-19 binds with high affinity to human ACE2. The COVID-19 spike appears to be the result of selection on human or human-like ACE2 permitting another suitable binding solution to occur and this strongly indicate that COVID-19 is not the genetic engineering

product. All COVID-19 sequenced genomes have the well adapted RBD and the polybasic cleavage site, and thus are derived from a common ancestor. Initial analyses demonstrated that Malayan pangolins (*Manis javanica*) illegally imported into Guangdong province, China contain a coronavirus (CoV) that is similar to COVID-19. Nevertheless, no pangolin CoV has been identified to be sufficiently similar to COVID-19 across its entire genome for supporting direct human infection.

Volz., et al. demonstrated their study by analyzing 53 SARS-CoV-2 (COVID-19) whole genome sequences collected up to February 3, 2020. They found a strong association between the time of sample collection and accumulation of genetic diversity of COVID-19. By using Bayesian and maximum likelihood phylogenetic methods revealed that the COVID-19 was introduced into the human population in Wuhan, China in early December 2019 and has an epidemic doubling time of about 7 days. Precise estimated of epidemic size are not possible with current genetic data, the analyses demonstrated substantial heterogeneity in the number of secondary infections caused by each COVID-19-infected case that indicated by a high level of over-dispersion in the reproduction number. Phylogenetic analysis demonstrates a common ancestor to SARS-CoV-2 (COVID-19), human SARS-CoV, and the bat SARS-CoV converge. The four structural viral proteins, envelope (E), membrane (M), nucleocapsid (N), and spike (S), imply a high degree of shared identity in range of 97.7-100 % between the SARS-CoV-2 (COVID-19) and bat coronaviruses that supports the descend of SARS-CoV-2 (VID-19) from an animal.

HLA genotype plays a significant role in differential regulation and activation of T cells as well as disease duration and transmission. A recent study on human leukocyte antigen (HLA) binding affinity of 48,395 unique peptides (possible 8- to 12-mers) from the SARs-CoV-2 (COVID-19) proteome for assessing the potential for cross-protective immunity conferred by previous exposures to common human coronaviruses (i.e. 229E, NL63, OC43, and HKU1) demonstrated that alleles HLA-A*02 : 02, HLA-B*15 : 03, and HLA-C*12 : 03 were the top presenters of conserved peptides. Fifty-six different HLA alleles, especially HLA-B*46 : 01 revealed no appreciable binding affinity (<500 nm) to any of the conserved SARS-CoV-2 (COVID-19) peptides, indicating a concomitant lack of potential for cross-protective immunity from other human coronaviruses. Considering the entire proteome of SARS-CoV-2 (COVID-19), HLA-A and HLA-C alleles expressed the relative largest and smallest capacity to present SARS-CoV-2 (COVID-19) antigens, respectively. No appreciable global correlation between conservation of the SARS-CoV-2 (COVID-19) proteome and its predicted MHC binding affinity, indicating a lack of selective pressure for the capacity to present coronavirus epitopes (p = 0.27, Fisher's exact test). peptide presentation appears to be independent of estimated time of peptide production during SARS-CoV-2 (COVID-19) life cycle, with indistinguishable early and late SARS-CoV-2 (COVID-19) peptide presentation.

A recent study conducted by Zhao *et al* on the correlation compared between the ABO blood group among 1,775 COVID-19-infected patients and 3,694 normal individuals from Wuhan city, China and 23,386 normal persons from Shenzhen city, China revealed that blood group A had a specific significant higher risk for COVID-19 compared to non-blood group A groups (albeit modest effect size, odd ratio (OR) = 1.20, p = 0.02), while blood group O had a significant lower risk for COVID-19 compared to non-O blood groups (OR = 0.67, p < 0.001).

In conclusion, although genomic evidence does not support the belief that COVID-19 is a laboratory construct, currently it is impossible to disprove or prove the theories of its origin. To identify the COVID-19 origin, obtaining virus sequences from immediate non-human animal sources would be the most definite method. In the absence of proper cure of COVID-19, it is necessary to identify the factors that may assist in assessment of the COVID-19 disease severity before rapid progression of the disease.