

---

---

## Analysis of Coronaviral Spike Proteins and Virus–host Interactions

Miran Kim<sup>1,2</sup>, Myeongji Cho<sup>1</sup>, Ji-Hae Lee<sup>3,4</sup>, Hayeon Kim<sup>5,\*</sup> and Hyeon S. Son<sup>1,4,\*</sup>

<sup>1</sup> Laboratory of Computational Biology & Bioinformatics, Graduate School of Public Health, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul, 08826, Korea

<sup>2</sup> Special Chemistry, Laboratory Medicine, Asan Medical Center, 88, Olympic-ro 43-gil, Songpa-gu, Seoul 138-736, Korea

<sup>3</sup>Department of Bioinformatics, Industrialization Support Research Team, Nakdonggang National Institute of Biological Resources, Sangju, Gyeongsangbukdo, Korea

<sup>4</sup>SNU Bioinformatics Institute, Interdisciplinary Graduate Program in Bioinformatics, College of Natural Science, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul, 08826, Korea

<sup>5</sup>Department of Biomedical Laboratory Science, Kyungdong University, 815 Gyeonhwon-ro, Munmak, Wonju, Gangwondo, 24695, Korea

---

### Abstract

**Objectives:** Recent outbreaks caused by Middle East respiratory syndrome coronavirus (MERS-CoV), such as the May 2015 outbreak in Korea, highlight the urgency of studying new mutant viruses that may be introduced in the future and preparing effective countermeasures to prevent large-scale infections. Most coronaviruses that infect humans cause only mild respiratory infections, but research on coronaviruses has increased in the wake of large-scale epidemics of MERS and severe acute respiratory syndrome (SARS). Therefore, we conducted a comparative analysis of the major human coronaviruses using genetic information and performed a network analysis to investigate the interactions between viruses and host immune proteins.

**Methods:** Phylogenetic and structural analyses were performed using the spike protein sequences of six coronaviruses (HCoV-OC43, HCoV-HKU1, MERS-CoV, SARS-CoV, HCoV-229E, and HCoV-NL63) causing diseases in humans. Network analysis was performed using Cytoscape 3.3.0 to analyze the interactions of viral proteins with host immune proteins.

**Results:** The phylogenetic analysis showed that although HCoV-OC43 and MERS belong to different lineages (lineages A and C, respectively), the evolutionary distance between spike proteins in these two viruses is relatively close. The structural analysis confirmed that the structure of the spike protein of HCoV-OC43 was similar to those of SARS-CoV and MERS-CoV, which are both highly transmissible. Finally, the network analysis confirmed interactions between the human IC1 protein, which is involved in the activation of the C1 complex, and non-structural proteins in SARS-CoV.

**Conclusion:** The similarity of SARS-CoV and MERS-CoV with other coronaviruses suggests the need for continued study of mutations in coronaviral genomes. In particular, the IC1 protein, which interacts with non-structural proteins in SARS-CoV, may have a major effect on the host immune response to viral infection because it has functions related to complement system activation. Studies of these viruses and host immune proteins will assist the development of vaccines and therapeutic agents against coronaviruses by enhancing our understanding of the detailed immune mechanisms against viral infections.

**keywords:** *Bioinformatics, Coronavirus, Database, Immune, Network analysis*

### Introduction

An outbreak caused by Middle East respiratory syndrome coronavirus (MERS-CoV) occurred in Korea in May 2015, resulting in 186 infections and

36 deaths, as well as the isolation of more than 16,000 people and healthcare workers until the end of the outbreak in January 2016 [1]. Such events highlight the urgent need to analyze new variant viruses that may be introduced in the future and

---

\* Co-corresponding authors:

Hayeon Kim (hykim1984@kduniv.ac.kr)

Department of Biomedical Laboratory Science, Kyungdong University, 815 Gyeonhwon-ro, Munmak, Wonju, Gangwondo, 24695, Korea

Hyeon S. Son (hss2003@snu.ac.kr, 02-880-2746)

Graduate School of Public Health, Seoul National University, 599 Gwanak-ro, Gwanak-gu, Seoul 151-742, Korea.

prepare effective countermeasures to prevent large-scale infections and the spread of these viruses. Interestingly, the MERS infections in Korea differed from those in the Middle East. In the Middle East, the virus was transmitted mainly by camels and other animals and resulted in a high mortality rate of 35–40%; however, the propagation rate was low, and most infections required close contact, generally resulting in reduced pathogenicity as propagation progressed [2,3]. In contrast, the MERS outbreak in Korea was facilitated by human–human transmission with a high propagation rate, as confirmed tertiary infections and super-spreaders. Moreover, the mortality rate was approximately 20%, lower than that in the Middle East [4]. These differences suggest that the same virus may differ in terms of transmission rate and infection symptoms, and that individual health status or immunity as well as environmental factors may influence these differences [5].

According to a report by the International Commission on Taxonomy of Viruses, human coronaviruses belong to genera Alphacoronavirus and Betacoronavirus (order Nidovirales, family Coronaviridae, subfamily Coronavirinae), although the subfamily Coronavirinae includes two additional genera, Gammacoronavirus and Deltacoronavirus. In particular, HCoV-229E and HCoV-NL63 in the genus Alphacoronavirus, and HCoV-HKU1, HCoV-OC43, MERS-CoV, and severe acute respiratory syndrome coronavirus (SARS-CoV) in the genus Betacoronavirus, are representative viruses that cause infections in humans [6]. Coronaviruses are enveloped positive-sense single-stranded RNA viruses, and have the largest genomes among the RNA viruses [7]. The five major structural proteins are spike (S), membrane (M), envelope (E), nucleocapsid (N), and hemagglutinin-esterase (HE). HCoV-229E, HCoV-NL63, and SARS-CoV possess four genes encoding S, M, N, and E, whereas HCoV-OC43 and HCoV-HKU1 possess five genes encoding S, M, N, E, and HE (Table 1).

Most coronaviruses that infect humans cause only mild respiratory infections, such as mild colds; however, active research on coronaviruses has increased since the large-scale SARS epidemic. Because coronaviruses can spread among humans,

**Table 1.** Functions of coronaviral structural proteins

Protein	Functions
S (spike)	Receptor interaction, cell–cell fusion, viral entry [21]
M (membrane)	Virion membrane, viral assembly, infectivity [22]
N (nucleocapsid)	Binding of the viral genome in a beads-on-a-string conformation, induction of cell-cycle delay or arrest [23, 24]
E (envelope)	Viral assembly, virus-like particle formation with M protein [25]
HE (hemagglutinin-esterase)	Acetyl-esterase activities enhance cell entry [26]

they have the potential to cause large-scale epidemics. Therefore, studies of infection mechanisms are essential for predicting the risk of infection and preparing countermeasures. It is also necessary to understand the immune response to viral infections, a complex network produced by the interactions of biochemical molecules. As bioinformatics techniques using advanced computational skills have been introduced into the immunology, a large amount of experimental data of immune responses involved in various infections have been stored and analyzed, and analysis tools have been developed to understand the detailed immunoregulatory mechanism. Therefore, bioinformatics-based approaches can provide new insights and important scientific evidences for the identification of host immune mechanisms and life phenomena that remain unknown. In this study, we investigated host immune mechanisms against viral infection, considering that the response of host immune cells at the early stage of infection may be the main factor in determining disease severity and viral pathogenicity. In this analysis, genomic information on coronaviruses and information related to the host immune mechanism were used. A phylogenetic tree was constructed using the amino acid sequence of the S protein of coronaviruses that infect humans. Furthermore, a structural analysis was performed to determine whether amino acid sequence differences reflect structural differences, and interactions between viruses and immune proteins were analyzed.

## Methods

### Data collection

Coronaviral protein data obtained from GenBank (<https://www.ncbi.nlm.nih.gov/genbank>) and ViPR (<http://www.viprbrc.org>) were categorized as structural proteins (i.e., S, M, E, N, and HE) and non-structural proteins (e.g., non-structural protein [Nsp] 1–16, leader protein, RNA-dependent RNA polymerase). A list of host proteins involved in the immune response to viral infections was prepared via literature reviews, and data were collected from ImmPort (<http://www.immport.org>) and GenBank (<https://www.ncbi.nlm.nih.gov/genbank>). Immune-related proteins were divided into three categories (cytokines and receptors, natural killer cells, and antigen processing and presentation) to collect genetic information. The collected data were stored in a secondary database constructed for this study, HCoV-IMDB. Finally, data were collected from VirHostNet (<http://virhostnet.prabi.fr>) and VirusMentha (<http://virusmentha.uniroma2.it>) to analyze the interactions between coronaviruses and host immune proteins.

### Phylogenetic analysis

We investigated the evolutionary relationship among six coronaviruses (HCoV-OC43, HCoV-HKU1, MERS-CoV, SARS-CoV, HCoV-229E, and HCoV-NL63) that infect humans using the amino acid sequence of the S protein. Excluding partial sequences, sequences 1,000–1,500 amino acids in

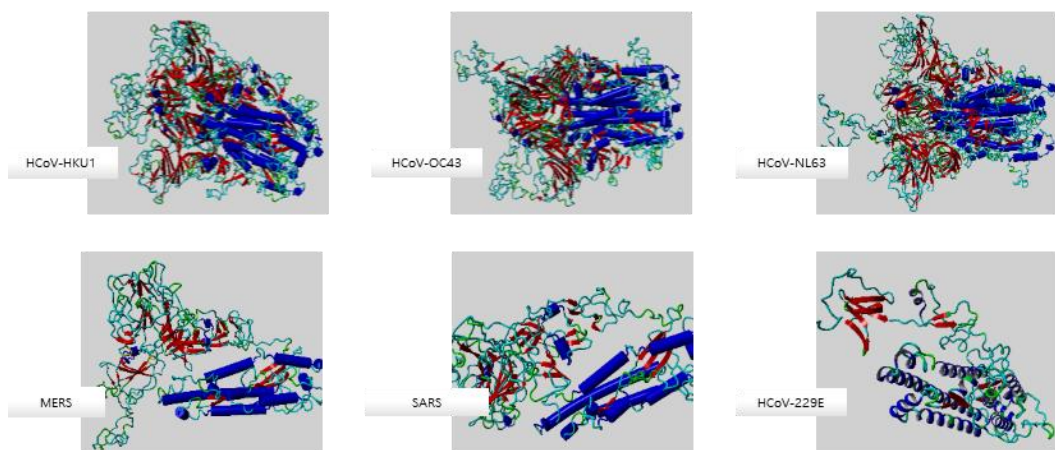
length that did not differ from the length of the reference sequence were analyzed. The phylogenetic analysis was performed using one HCoV-OC43, five HCoV-NL63, four CoV-229E, seven HCoV-HKU1, four MERS-CoV, and eight SARS-CoV S protein sequences, which were selected by UniProt mapping. The phylogenetic tree was constructed using the neighbor-joining method in the program MEGA ver. 7.0, and bootstrapping was performed with 1,000 iterations.

### Structural analysis

Structural analysis was performed to determine whether sequence differences among the six coronaviruses reflect structural differences. First, protein structure models were generated using the amino acid sequences of the S proteins of six coronaviruses, which were downloaded into a PDB file via the homology modeling server SWISS-MODEL (Figure 1). Next, the PDB file of each protein was uploaded to YASARA, structural alignments were performed for each virus, and the similarity of the root-mean-square deviation (RMSD) values and sequence similarity were measured.

### Network analysis

Data on the interactions between SARS-CoV and host proteins were downloaded from the constructed database HCoV-IMDB, and network analysis was performed using Cytoscape ver. 3.3.0 [8]. The coronaviral proteins and human immune



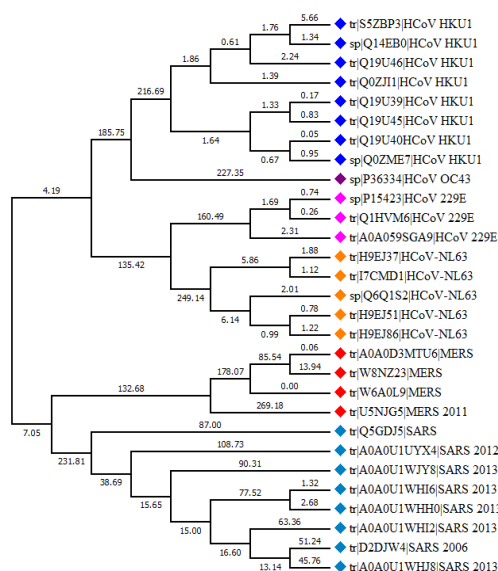
**Figure 1.** Structural analysis of six coronaviral spike (S) proteins. Structural models of S proteins were predicted based on amino acid sequences using SWISS-MODEL.

proteins were designated as the target node and the source node, respectively. The network was converted into an organic layout, and the two types of nodes were color-coded to facilitate classification (coronaviral protein node: pink; human immunoprotein node: blue).

## Results

### Phylogenetic analysis of six coronaviruses

Phylogenetic analysis was performed to investigate the evolutionary relationship of the infection characteristics among six coronaviruses (HCoV-OC43, HCoV-HKU1, MERS-CoV, SARS-CoV, HCoV-229E, and HCoV-NL63) known to cause disease in humans and the sequence differences among their S proteins. The S proteins of the six coronaviruses were classified by genus (Figure 2). The evolutionary distance between HCoV-229E and HCoV-NL63, both belonging to the genus Alphacoronavirus, was 0.407, and that between HCoV-OC43 and HCoV-HKU1, belonging to lineage A in the genus Betacoronavirus, was



**Figure 2.** Phylogenetic analysis of the spike (S) proteins of six coronaviruses. A phylogenetic tree of the S protein sequences of the six investigated coronaviruses (HCoV-OC43, HCoV-HKU1, MERS-CoV, SARS-CoV, HCoV-229E, and HCoV-NL63) causing infectious disease in humans was constructed using the neighbor-joining method.

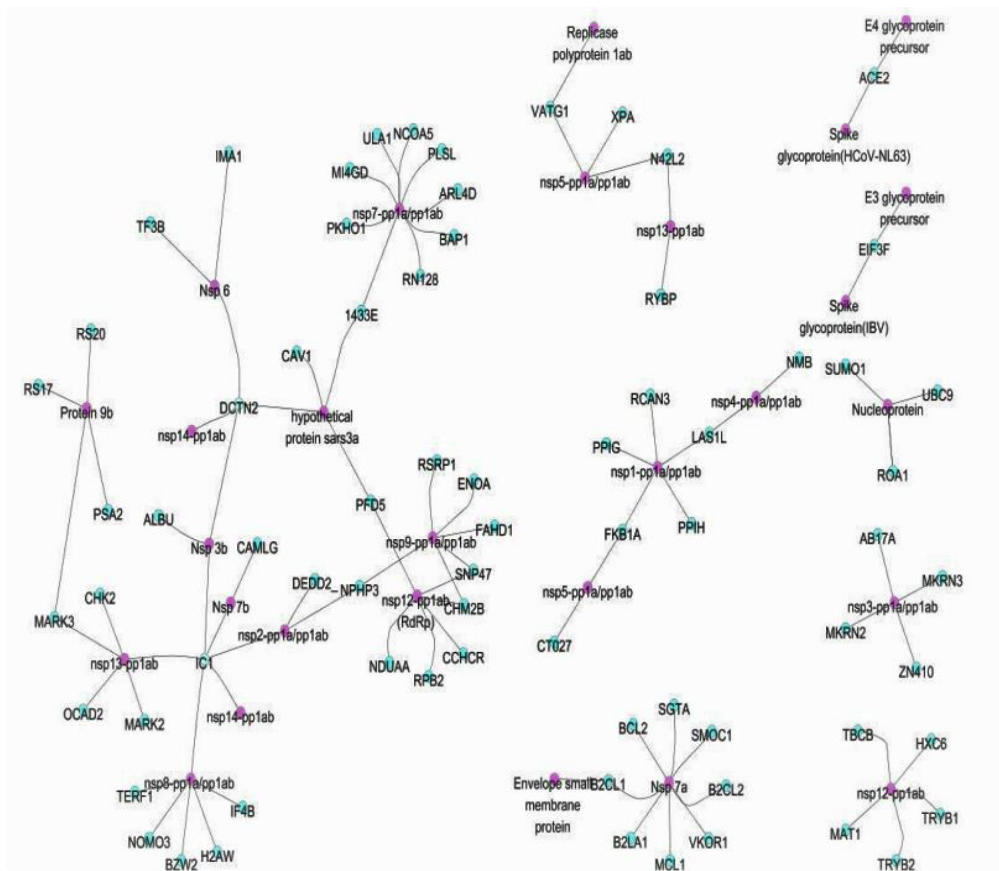
**Table 2.** Root-mean-square deviation (RMSD) and sequence identity values of coronaviral spike proteins

Compared viruses	RMSD (Å)	Sequence identity (%)
HCoV-OC43 and HCoV-HKU1	0.282	68.18
HCoV-NL63 and HCoV-229E	0.574	77.75
HCoV-OC43 and MERS	0.565	35.39
MERS and SARS	0.710	34.64
HCoV-OC43 and SARS	0.661	31.64
HCoV-HKU1 and SARS	0.681	32.27

0.374. Unexpectedly, the evolutionary distance between the S proteins of HCoV-OC43 and MERS-CoV, which belong to lineages A and C, respectively, was 0.946, which was lower than the mean of 1.067. Therefore, these two viruses were found to be relatively similar compared with the other viruses. Distance of phylogenetic tree shows the relationships between the operational taxonomic units (OTUs) based on the calculated genetic distance by sorting the nucleotide sequence of each species and comparing the genes between sequences at all positions. In the neighbor-joining method used in this study, the distance-based method is applied to the multiple alignments. Distance matrices are used to calculate the difference of the sequences between the clustered groups. This method reconstructs phylogenetic trees on the assumption that each sequence has a different evolutionary rate. The initial tree is a star-like unrooted tree with initial node. Each pair of taxon is evaluated for being joined and the sum of all branch lengths is calculated in the result tree. A pair that yields the smallest sum is considered the closest neighbor and is combined. A new branch is inserted between the taxa and the total branch length is recalculated. This process is repeated until all OTUs are assigned to the terminal node.

### RMSDs of S proteins

The RMSD is a measure of the average distance between the atoms of aligned proteins, which enables prediction of the similarity among the three-dimensional protein structures. Among the RMSD values, HCoV-OC43 and HCoV-HKU1 had



**Figure 3.** Network of the interactions between coronaviral proteins and human immune proteins. The pink and blue nodes represent viral proteins and human immune-related proteins, respectively. Three to five immune proteins showed interactions with viral non-structural proteins.

the most similar structures (0.282 Å). The RMSD values for the pairs HCoV-NL63/HCoV-229E, HCoV-OC43/MERS-CoV, MERS-CoV/SARS-CoV, HCoV-OC43/SARS-CoV, and HCoV-HKU1/SARS-CoV were 0.574 Å, 0.565 Å, 0.710 Å, 0.661 Å, and 0.681 Å, respectively (Table 2). These results showed that the structural similarity among S proteins did not correlate with the sequence similarity results. Notably, the structure of the S protein of HCoV-OC43 was similar to those of SARS-CoV and MERS-CoV, which are both highly transmissible.

#### Virus–host protein interactions

The interactions of viral proteins with host immune proteins were visualized by network analysis (Figure 3). The results revealed that three to five human proteins interacted with viral non-structural proteins. The network revealed

interactions of the Nsp7a protein of SARS-CoV with human B-cell lymphoma 2 (BCL2), BCL2-like protein 1 (B2CL1), B2CL2, myeloid cell leukemia 1 (MCL1), BCL2-related protein A1 (B2LA1), small glutamine-rich tetratricopeptide repeat containing alpha (SGTA), SPARC-related modular calcium binding 1 (SMOC1), and vitamin K epoxide reductase complex subunit 1 (VKOR1). These results were consistent with those of the study by Marta [9], which showed that Nsp7a affected the decrease in NF- $\kappa$ B and type I interferon expression in relation to the inflammatory response. BCL2, B2CL1, and MCL1 are involved in apoptosis and inflammatory responses, suggesting that Nsp7a is involved in the apoptosis mechanism [10–12]. B2LA1 affects interleukin 3, thus delaying apoptosis, and has a major role in endothelial-cell maintenance during

infection and exogenous signaling in hematopoietic cells. B2CL2 promotes cell survival by inhibiting apoptosis-promoting proteins [13]. SGTA binds directly to heat shock cognate 71 kDa protein and the heat shock protein 70 family members, which assist with protein folding, inhibit aggregation, and mediate ATPase activity [14]. The interactions of viral Nsp7a protein with these host proteins could prevent normal functioning of the host cell due to alterations in the chaperones that regulate endogenous cellular proteins. The network also indicated interactions of the Nsp1-pp1a/pp1ab protein of SARS-CoV with human ribosomal biogenesis protein LAS1L, regulator of calcineurin family member 3 (RCAN3), peptidylprolyl isomerase H (PPIH), peptidylprolyl isomerase G (PPIG), peptidyl-prolyl cis-trans isomerase (FKB1A), and peptidylprolyl isomerase A (PPIA) proteins. The Nsp1 protein of SARS-CoV binds to the 40S ribosome, causing transformation of capped mRNA in the host cell, thereby inhibiting protein translation and synthesis [15]. Nsp1 interacts with these proteins to allow binding of viral proteins to the ribosome.

Furthermore, interactions between the human plasma protease C1 inhibitor (IC1) protein with the Nsp3b, Nsp7b, Nsp13-pp1ab (ZD, NTPase/HEL), Nsp8-pp1a/pp1ab, Nsp2-pp1a/pp1ab, and Nsp14-pp1ab (nuclease ExoN homolog) proteins of SARS-CoV were confirmed. The IC1 protein is involved in activation of the C1 complex and acts in conjunction with the C1r or C1s protease to chemically inactivate the complex. IC1 is also involved in the physiological mechanisms of complement system activity, blood clotting, fibrin degradation, and kinin synthesis [18]. These interactions are consistent with the upregulation patterns of immune-related genes observed in a study of SARS-CoV-infected hosts [19]. In addition, the interactions of the Nsp7-pp1a/pp1ab protein of SARS-CoV with human ULA1, NCOA5, PLSL, ARL4D, BAP1, RN128, 1433E, PKHO1, and MI4GD proteins were confirmed. Finally, the network showed interactions of human DCTN2 protein with Nsp6, Nsp14-pp1ab, Nsp3b, Nsp6, and hypothetical protein sars3a of SARS-CoV. The findings confirmed that human proteins interact with SARS-CoV proteins Nsp9-pp1a/pp1ab and Nsp12-pp1ab.

## Discussion

The host immune response to viral infection differs among pathogens and is influenced by the activation of host immune proteins. Depending on the extent of the immune response, viruses that enter host cells can easily spread to other hosts or worsen the symptoms of infectious diseases. In this study, virus and human immune proteins were analyzed to investigate virus–host interactions related to host immune mechanisms. Although coronaviruses typically cause mild colds, the need for further research is increasing given that SARS and MERS can cause serious symptoms. Therefore, we analyzed six major coronaviruses known to infect humans. The results of the phylogenetic analysis of coronavirus S proteins indicated that MERS-CoV (in lineage A) and HCoV-OC43 (in lineage C) exhibited the closest evolutionary distance among the coronaviruses evaluated. This result is consistent with a previous study in adults in South Asia suggesting that HCoV-OC43 causes acute respiratory infections, necessitating careful monitoring of genetic mutations in this virus [20]; therefore, continued investigation and monitoring of HCoV-OC43 infections and mutations should be conducted. In addition, the RMSD of the S proteins of the six coronaviruses showed that the sequence and structure of the HCoV-OC43 S protein were similar to those of SARS and MERS, which are both highly contagious in humans. The network analysis confirmed the interactions of the human IC1 protein with Nsp3b, Nsp7b, Nsp13-pp1ab (ZD, NTPase/HEL), Nsp8-pp1a/pp1ab, Nsp2-pp1a/pp1ab, and Nsp14-pp1ab (nuclease ExoN homolog) proteins of SARS-CoV. This is consistent with the upregulation patterns of immune-related genes in a previous SARS-CoV study [19], suggesting that the IC1 protein has an important role in the immune response of coronaviruses.

This findings highlight the importance of understanding the relationships between coronavirus proteins and human immune proteins for clarifying the mechanism of viral immunity and for the development of vaccines or therapeutic drugs. Our results may be a meaningful attempt in respect of utilization of diverse bioinformatic approaches, including phylogenetic analysis, structural analysis, and network analysis. It can be used to understand detailed mechanisms of immune

responses to viral infections and to monitor and predict the outbreaks and spread of viral diseases.

## Conclusion

In this study, we analyzed structural and non-structural proteins of coronaviruses and their interactions with host immune proteins using bioinformatics techniques. To clarify the effect of the host immune system on coronaviral infections, we performed phylogenetic analysis, sequence and structural homology analyses, and network analysis of virus–host interactions. Coronaviral epidemics not only present as periodic persistent infections but are also associated with unpredictable epidemiological patterns due to the emergence of new or mutated viruses. As a result of various factors, coronaviruses are not confined to a single region but rather spread across regions; therefore, it is imperative to identify trends in outbreaks to prepare effective countermeasures. The network analysis performed in this study confirmed the interaction between the host IC1 protein, which affects complement system activation, and non structural proteins of SARS-CoV. This finding implies that the virus can affect host immune proteins, which may interfere with normal immune mechanisms due to excessive or suppressed immune responses by the infected host. Although experimental validations are required to clearly demonstrate that ic1 protein plays an important role in protein-protein interactions, further validations and advanced studies on this protein may help to elucidate the detailed mechanism of immune responses observed in infections caused by coronaviruses.

## Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT and MOE) (No. 2016R1C1B2015511 and No. 2017R1D1A1B03033413). Support from CMB is gratefully acknowledged.

## References

1. Korea Centers for Disease Control and Prevention. Middle East Respiratory Syndrome Information. June 2015. Available at:

<http://www.mers.go.kr/mers/html/jsp/main.jsp>  
(Accessed 1 Oct, 2015)

2. Butler D. South Korean MERS outbreak spotlights lack of research. *Nature*. 2015;522:139-140.
3. Banik GR, Khandaker G, Rashid H. Middle East Respiratory Syndrome Coronavirus "MERS-CoV": Current Knowledge Gaps. *Paediatric respiratory reviews*. 2015;16:197-120.
4. Je SM, Bae WJ, Kim JY, Seok SH, Hwang ES. The Preparedness for Re-emerged and Emerging Infectious Diseases: A Lesson Through Outbreak of Middle East Respiratory Syndrome Coronavirus (MERS-CoV) in South Korea. *Journal of Bacteriology and Virology*. 2015;45:339–353.
5. Anneke E, Lenny H, Jan S. Pathogen–host–environment interplay and disease emergence. *Emerging Microbes & Infections*. 2013; 2:e5.
6. Chan JF, Lau SK, To KK. Middle East respiratory syndrome coronavirus: another zoonotic betacoronavirus causing SARS like disease. *Clinical microbiology reviews*. 2015;28:465.
7. Masters PS, Kuo L, Ye R. Genetic and molecular biological analysis of protein-protein interactions in coronavirus assembly. *Advances in Experimental Medicine and Biology*. 2006;581:163.
8. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research*. 2003;13(11):2498-2504.
9. Marta LD, Jose LN, Regla JM, Jimenez G, Raul FB, Craig F, Carlos CR, Stanley P, Luis E. Inhibition of NF- $\kappa$ B-Mediated Inflammation in Severe Acute Respiratory Syndrome Coronavirus-Infected Mice Increases Survival. *Journal of virology*. 2014;88:913-924.
10. Maurer U, Charvet C, Wagman AS, DeJardin E, Green DR. Glycogen synthase kinase-3 regulates mitochondrial outer membrane permeabilization and apoptosis by destabilization of MCL-1. *Molecular cell*. 2006;21:749-760.
11. Bruey JM, Bruey-Sedano N, Luciano F, Zhai D, Balpai R, Xu C, Kress CL, Bailly-Maitre B, Li X, Osterman A, Matsuzawa S, Tersikh AV, Faustin

- B, Reed JC. Bcl-2 and Bcl-XL regulate proinflammatory caspase-1 activation by interaction with NALP1. *Cell*. 2007;129:45-56.
12. Wei Y, Pattingre S, Sinha S, Bassik M, Levine B. JNK1-mediated phosphorylation of Bcl-2 regulates starvation-induced autophagy. *Molecular cell*. 2008;30:678-688.
  13. Gibson L, Holmgren SP, Huang DC, Bernard O, Copeland NG, Jenkins NA, Sutherland GR, Baker E, Adams JM, Cory S. Bcl-w, a novel member of the Bcl-2 family, promotes cell survival. *Oncogene*. 1996;13:665-675.
  14. Dutta S, Tan YJ. Structural and functional characterization of human SGT and its interaction with Vpu of the human immunodeficiency virus type 1. *Biochemistry*. 2008;47:10123-10131.
  15. Kamitani W, Huang C, Narayanan K, Lokugamage KG, Makino S. A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein. *Nature structural & molecular biology*. 2009;16:1134–1140.
  16. Castle CD, Cassimere EK, Lee J, Denicourt C. Las1L is a nucleolar protein required for cell proliferation and ribosome biogenesis. *Molecular and cellular biology*. 2010;30:4404-4414.
  17. Teigelkamp S, Achsel T, Mundt C, Goethel SF, Cronshagen U, Lane WS, Marahiel M, Luehrmann R. The 20kD protein of human [U4/U6.U5] tri-snRNPs is a novel cyclophilin that forms a complex with the U4/U6-specific 60kD and 90kD proteins. *RNA*. 1998;4:127-141.
  18. Aulak KS, Davis AE. III, Donaldson VH, Harrison RA. Chymotrypsin inhibitory activity of normal C1-inhibitor and a P1 Arg to His mutant: evidence for the presence of overlapping reactive centers. *Protein Science*. 1993;2:727-732.
  19. Gralinski LE, Baric RS. Molecular pathology of emerging coronavirus infections. 2015;235:185-195.
  20. Maryam NA, Kim TN, Xiang YO, Yong KP, Yutaka T, Jack BC, Nik SH, Adeeba K, Kok KT. Molecular epidemiology and evolutionary histories of human coronavirus OC43 and HKU1 among patients with upper respiratory tract infections in Kuala Lumpur, Malaysia. *Virology journal*. 2016;13:33.
  21. Weiss SR, Navas-Martin S. Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus. *Microbiology and Molecular Biology Reviews*. 2005;69(4):635-664.
  22. Haan CA, Wit M, Kuo L, Montalto-Morrison C, Haagmans BL, Weiss SR, Masters PS, Rottier PJ. The glycosylation status of the murine hepatitis coronavirus M protein affects the interferogenic capacity of the virus in vitro and its ability to replicate in the liver but not the brain. *Virology*. 2003;312:395-406.
  23. Chang CK, Sue SC, Yu TH, Hsieh CM, Tsai CK, Chiang YC, Lee SJ, Hsiao HH, Wu WJ, Chang WL, Lin CH, Huang TH. Modular organization of SARS coronavirus nucleocapsid protein. *Journal of biomedical science*. 2006;13(1):59–72.
  24. Wurm T, Chen H, Hodgson T, Britton P, Brooks G, Hiscox JA. Localization to the nucleolus is a common feature of coronavirus nucleoproteins, and the protein may disrupt host cell division. *Journal of virology*. 2001;75:9345-9356.
  25. Vennema H, Godeke GJ, Rossen JW, Voorhout WF, Horzinek MC, Opstelten DJ, Rottier PJ. Nucleocapsid-independent assembly of coronavirus-like particles by coexpression of viral envelope protein genes. *The EMBO journal*. 2006;15:2020-2028.
  26. Kazi L, Lissenberg A, Watson R, de Groot RJ, Weiss SR. Expression of hemagglutinin esterase protein from recombinant mouse hepatitis virus enhances neurovirulence. *Journal of virology*. 2005;79(24):15064-15073.