Bioinformatics analysis of single-cell mRNA-seq of SARS-CoV-1 and SARS-CoV-2 infection compared to MERS-CoV from Sequence Read Archive (SRA) database reveals novel targets for therapies

Mengyao Wang
Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, The University of Tokyo

Peng Lu (porterlu@g.ecc.u-tokyo.ac.jp)
Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, The University of Tokyo  https://orcid.org/0000-0002-1405-5294

Research Article

Keywords: COVID-19, SARS, MERS, cytokine, ISGylation, innate immunity

DOI: https://doi.org/10.21203/rs.3.rs-62454/v1

License: ☒ This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

The global pandemic of COVID-19 caused by SARS-CoV-2 is still threatening the world. By May 13, 2020, more than 40 million people have been infected by SARS-CoV-2 and almost 300 thousand deaths were reported. The discovery and development of anti-viral drugs and vaccines are being conducted worldwide and the understanding of the molecular responses of a single cell to SARS-CoV-2 is in urgent need. The comparative analysis of gene expression in SARS-CoVs and MERS-CoV infected Calu-3 cells reveals that although the coronaviruses cause similar acute respiratory distress syndromes, the molecular responses of Calu-3 cells to SARS-CoVs infections showed a unique signature. A total of 64 correlated differentially expressed genes (DEGs) were identified in this study. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses indicated that the DEGs were significantly involved in biological process of ‘Response to interferon-γ’, ‘Viral life cycle’, ‘Phagosome’ and ‘Epstein-Barr virus infection’. STRING analysis showed that the DEGs that were up-regulated after SARS-CoVs infections but down-regulated after MERS-CoV infections showed a strong interaction network. Molecular Complex Detection (MCODE) analysis further refined a unique network consisted of eight hub genes out of 64 DEGs, which are involved in cytokine response (CXCL8, CCL20, and CSF2), ISGylation (ISG15), macrophage activation (ITGAM), complement system (C3), and NFκB signaling pathway (TRAF1 and NFkB2). The unique network identified here will be a potential fingerprint for distinguishing SARS-CoVs infections from MERS-CoV infection. The identification of the eight hub genes will lead to the discovery of new possible therapeutic targets for fighting COVID-19.

1 Introduction

Since December, 2019, a novel coronavirus, known as severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2), spreads rapidly and globally. To date, the coronavirus disease 2019 (COVID-19) caused by SARS-CoV-2 has become a world pandemic with a high mortality1.

The global outbreak of severe acute respiratory syndrome (SARS) happened in 2002 from Guangdong Province, China². At that time, SARS-CoV-1 was not closely related to any known human or animal coronavirus or torovirus (Drosten et al., 2003; Ksiazek et al., 2003). Patients’ age ranged from 1 to 100 years². Fever of 38ºC or more, cough, myalgia and shortness of breath were the most frequent symptoms (Peiris et al., 2003a). SARS-CoV-1 was reported to have affected 8096 people and caused 774 deaths with the fatality ratio of 9.6%². SARS-CoV-1 belongs to the coronavirus genus in the Coronaviridae family and has a positive-sense RNA genome of 29.7 kb (He et al., 2004a). Similar to other coronaviruses, the spike (S) proteins of SARS-CoV-1 are responsible for virus binding, fusion and entry, and are major inducers of neutralizing antibodies (Gallagher and Buchmeier, 2001; Ho et al., 2004). It has been demonstrated that membrane-associated angiotensin-converting enzyme 2 (ACE2), an essential regulator of heart function, is a functional receptor in target cells for SARS-CoV-1 (Dimitrov, 2003; Li et al., 2003; He et al., 2004b).
The latest local outbreak of SARS was reported in April 2004, in Beijing and Anhui Province, China. Then SARS-CoV-1 seemed to disappear in human world\(^2\).

Ten years after the emergence of SARS, Middle East respiratory syndrome (MERS) was first identified in Saudi Arabia in 2012\(^3\). MERS-CoV is a novel member of the betacoronavirus and distinct from SARS-CoV-1 (Peiris et al., 2003b; van Boheemen et al., 2012). The typical presentation of MERS-CoV disease is fever, cough and shortness of breath, similar to SARS\(^3\). But infection with MERS-CoV can be asymptomatic\(^3\). At the end of January 2020, a total 2519 cases of MERS, including 866 associated deaths were reported globally. The fatality ratio is quite high at around 34.3\(^{\%}\). MERS-CoV has a large, positive-sense RNA genomes of 30.1 kb. The MERS-CoV also utilizes S proteins for interaction with and entry into target cells (Raj et al., 2013). Raj et al. identified that dipeptidyl peptidase 4 (DPP4) functions as a cellular receptor for MERS-CoV (Raj et al., 2013). However, DPP4, expressed on the surface of several cell types, does not share any sequence or structural similarities to ACE2 (Li, 2005; Wang et al., 2013). To date, there are still about 200 MERS cases reported every year, mainly in Saudi Arabia\(^3\). Efficient vaccines against MERS-CoV are still under development (Haagmans et al., 2016; Jiaming et al., 2017).

Now, human is facing with the new challenge of COVID-19 caused by SARS-CoV-2. Typical clinical symptoms of COVID-19 are fever, dry cough, breathing difficulties, headache and pneumonia (Zhou et al., 2020). Most COVID-19 patients with other diseases, such as diabetes mellitus, hypertension, cardiac diseases, obesity or cancer, are prone to get severe (Liang et al., 2020; Wang et al., 2020). By May 13, 2020, the number of confirmed COVID-19 cases is 4,170,424, including 287,399 deaths with the fatality ratio of 6.9\(^{\%}\). And the number is still increasing. SARS-CoV-2 is closely related to SARS-CoV-1, sharing 79.6% sequence identity (Zhou et al., 2020). SARS-CoV-2 uses the ACE2 for entry and the serine protease TMPRSS2 for S proteins priming, same with SARS-CoV-1 (Matsuyama et al., 2010; Hoffmann et al., 2020; Walls et al., 2020). On the other side, the receptor-binding domain (RBD) in the S proteins of SARS-CoV-2 differs from SARS-CoV-1. Thus SARS-CoV-2 is supposed to have an RBD that binds with high affinity to ACE2 (Andersen et al., 2020; Wan et al., 2020). The threat of SARS-CoV-2 is much serious than previous SARS-CoV-1 and MERS-CoV. However, the origin, transmission path and infection mechanism of SARS-CoV-2 are still elusive. Effective medicines and vaccines against SARS-CoV-2 are still under investigation. Consequently, it is urgent to understand the molecular responses of a single cell to coronaviruses infection. Here, we compared the public single-cell (Calu-3 cell) mRNA-seq data from Sequence Read Archive (SRA) database among MERS-CoV, SARS-CoV-1, and SARS-CoV-2 and identified 64 differentially expressed genes (DEGs). We further refined eight potential hub proteins in responding to SARS-CoVs. The results will provide clues for understanding the molecular mechanism of SARS-CoV-2 infection and provide potential strategies for combating the ongoing COVID-19 outbreak and preventing future emergence of lethal coronaviruses.

2 Results

Differentially expressed genes in the infected Calu-3 cells
In the MERS-CoV infected datasets, a total of 282 up-regulated genes and 1181 down-regulated genes were identified (Figure 1A & Figure 2). In the SARS-CoVs infected datasets, more genes were up-regulated (Figure 1B&C). In detail, a total of 761 up-regulated genes and 55 down-regulated genes were identified in SARS-CoV-1 infected cells (Figure 2) and a total of 1229 up-regulated genes and 396 down-regulated genes were identified in SARS-CoV-2 infected cells (Figure 2). The intersections from the upset-plot demonstrated that 64 DEGs were changed in all three coronaviruses infection groups compared to control groups (Figure 2). The list of 64 DEGs was illustrated in Figure 3. Seven genes were up-regulated both in the MERS-CoV infected cells and SARS-CoVs infected cells, whereas 15 genes were down-regulated both in the MERS-CoV infected cells and SARS-CoVs infected cells (Figure 2). In addition, 41 genes were down-regulated in MERS-CoV infected cells, but up-regulated in SARS-CoVs infected cells (Figure 2). Only one gene was up-regulated in MERS-CoV infected cells but down-regulated in SARS-CoVs infected cells (Figure 2). Interestingly, no gene was identified that is up-regulated in SARS-CoV-1 but down-regulated SARS-CoV-2, and no gene was identified that is down-regulated in SARS-CoV-1 but up-regulated SARS-CoV-2 (Figure 2).

**GO enrichment and KEGG pathway analyses**

The GO enrichment analyses suggested that the DEGs were significantly involved in biological process of ‘Response to interferon-γ’ and ‘Viral life cycle’ (Figure 4A & Table S1). In addition, KEGG analysis results indicated that the DEGs were mainly enriched in the pathways of ‘Phagosome’ and ‘Epstein-Barr virus infection’ (Figure 4B & Table S2).

**PPI network and hub genes detection**

STRING database analysis identified two networks with 38 DEGs involved in the interaction. In the main PPI network, 63 nodes and 88 edges were identified, which included up-regulated and down-regulated DEGs that either share the same or opposite trend in MERS-CoV infected cells and SARS-CoVs infected cells (Figure 5A). The DEGs that were involved in the main PPI network contained 26 genes that were down-regulated in MERS-CoV infected cells but up-regulated in SARS-CoVs infected cells, six genes that were down-regulated both in MERS-CoV infected and SARS-CoVs infected cells, two genes that were up-regulated both in MERS-CoV infected and SARS-CoVs infected cells, and one gene that was up-regulated in MERS-CoV infected cells but down-regulated in SARS-CoVs infected cells. A total of seven genes had more than 5 nodes and were highlighted by black boxes in the PPI network, including ICAM1, HLA-B, EGR1, CXCL8, ITGAM, OASL, and ISG15 (Figure 5A). In STRING analysis, all the co-expressed DEGs were involved in immune response, response to cytokine and stress. Molecular Complex Detection (MCODE) analysis further identified ten top hub genes, including CXCL8, ITGAM, C3, CCL20, CSF2, EGR1, ISG15, TRAF1, NFKB2, and IFITM3 (Figure 5B). Eight out of ten predictive genes were also predicted by GO and KEGG analyses. However, EGR1 only predicted by KEGG analysis and IFITM3 only predicted by GO analysis.

**Diseases related to the DEGs**
The results returned by DAVID database showed that a total of 83 diseases were related to the DEGs with a $p$-value < 0.05 (Table S3). Type II Diabetes, edema, rosiglitazone, pharmacogenomic, infection, and cancer had the ratio (gene counts in each disease to 64 input DEGs) > 30% and were considered as mainly related diseases (Figure 6).

3 Materials And Methods

Data collection

Data objects were obtained from Sequence Read Archive (SRA) database in National Center for Biotechnology Information (NCBI) by sra-tools (SRA Handbook [Internet]. Using the SRA, 2010). The sra files, returned by the ‘prefetch’ function, were converted to fastq format using the ‘fasterq-dump’ function. The following high-throughput sequencing data of the mRNA from human lung cancer cell line, Calu-3 cells were selected as the study target. All samples of treatment groups have been post-infected by the respective coronaviruses for 24 h.

SRR10357366, SRR10357367, SRR10357368, SRR10357369, SRR10357370, and SRR10357371 from the BioProject PRJNA580021 (Zhang et al., 2020) contain the sequencing data of Calu-3 cells infected by MERS-CoV. SRR10357369, SRR10357370, and SRR10357371 correspond to the data of Calu-3 cells infected by MERS-CoV. SRR10357366, SRR10357367 and SRR10357368 correspond to the data of control groups. SRR11550033, SRR11550034, SRR11550039, SRR11550040, SRR11550045, and SRR11550046 from the BioProject PRJNA625518 (Lamers et al., 2020) contain the sequencing data of Calu-3 cells infected by SARS coronaviruses. SRR11550039 and SRR11550040 correspond to the data of Calu-3 cells infected by SARS-CoV-1. SRR11550045 and SRR11550046 correspond to the data of Calu-3 cells infected by SARS-CoV-2. SRR11550033 and SRR11550034 correspond to the data of control groups.

Data process

Quality control checks on the raw sequence data were performed by FastQC. The mapping of the sequence data were performed by HISAT2 (Kim et al., 2015). The expression level of mRNA was counted by featureCount (Liao et al., 2014), using human genome index GRCh38 provided by HISAT2 (Kim et al., 2019). The expression matrices created by featureCount were then estimated by MultiQC (Ewels et al., 2016) and analyzed in R software (version R 4.0.0) using the DESeq2 R package (Love et al., 2014). The fold change (FC) of the original data was log$_2$ transformed and the log$_2$FC shrinkage was performed by apeglm (Zhu et al., 2019) R package. The significant DEGs were filtered out using the threshold with Benjamini-Hochberg (Benjamini and Hochberg, 1995) adjusted $p$-value < 0.05 and with log$_2$FC shrinkage values > 1.0 or < −1.0. The heat-map was created by pheatmap R package based on the normalized expression values in transcripts per million (TPM) of DEGs. DEGs were visualized in volcano plots using limma (Ritchie et al., 2015) and ggplot2 R package. The overlapping of DEGs from different datasets were visualized by an upset-plot, using UpSetR (Lex et al., 2014) R package.
Gene enrichment analysis

The correlated DEGs in MERS-CoV, SARS-CoV-1, and SARS-CoV-2 infected Calu-3 cells were identified and searched in the Gene Ontology (GO) database (Ashburner et al., 2000; Carbon et al., 2019) for enrichment analysis, including molecular function (MF), cellular component (CC) and biological process (BP). The enrichment analysis of pathways was performed by searching the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Kanehisa, 2000, 2019; Kanehisa et al., 2019). The GO and KEGG enrichment analyses were performed by clusterprofiler R package (Yu et al., 2012). The adjusted $p$-value < 0.05 was considered significant.

Protein–protein interaction (PPI) analysis

To better understand the interactions between the proteins that correspond to the DEGs, STRING database was used to visualize the PPI network (Szklarczyk et al., 2015). Cytoscape software (Shannon, 2003) was used to refine the PPI network by selecting the interactions with a combined score of >0.4 and the MCODE (Bader and Hogue, 2003) algorithm was used to extract the hub genes. The parameters of MCODE used in this study were as follows: the degree of cut-off = 2; cluster finding = haircut; node score cut-off = 0.2; k-core = 2; and the maximum depth = 100.

DEGs in other diseases

The relationship between the DEGs and other diseases were analyzed by searching in the DAVID (Huang et al., 2009b, 2009a) using database ‘GAD_DISEASE’, ‘GAD_DISEASE_CLASS’, and ‘OMIM_DISEASE’. The $p$-value < 0.05 was considered significant.

4 Discussion

The present study identified 64 correlated DEGs in SARS-CoVs infected and MERS-CoV infected cells (Figure 3). No gene was identified that has opposite expression trend between SARS-CoV-1 infected and SARS-CoV-2 infected cells (Figure 2). The comparison of the differences and similarities of DEGs among MERS-CoV infected, SARS-CoV-1 infected and SARS-CoV-2 infected cells demonstrated that the molecular responses of Calu-3 cells to SARS-CoV-2 infection are quite similar to SARS-CoV-1 infection but different from MERS-CoV infection. Genes that were down-regulated in the MERS-CoV infected cells but up-regulated in SARS-CoVs infected cells are responsible for telling the difference between MERS-CoV and SARS-CoVs infections. CXCL8, CCL20, CSF2, ITGAM, C3, ISG15, TRAF1, and NFKB2 were identified as the main hub genes since all of them were predicted important in STRING, GO and KEGG analyses (Figure 4 & Figure 5).

CXCL8 (also known as IL-8) and CCL20 (also known as LARC or MIP3A) are two kinds of chemokine that take a great part in the immune cell recruitment during inflammation or pro-inflammation and function as cytokines (Zlotnik et al., 2011; Zlotnik and Yoshie, 2012). CXCL8 and CCL20 belong to the C-X-C motif (CXCL) family and C-C motif (CC) family, respectively. CXCL8 has been considered as a potential
biomarker for acute respiratory distress syndromes (García-Laorden et al., 2017). The up-regulation of the CXCL8 gene expression by SARS-CoV-1 is resulted from a direct effect of the virus at the cellular level (Spiegel and Weber, 2006). On the contrary, a study using human lung tissues has found that CXCL8 transcription was up-regulated only by SARS-CoV-1, but not SARS-CoV-2 infection (Chu et al., 2020), whereas a study using mouse lung and brain tissues have reported that a significant increase of CXCL8 expression after infection with MERS-CoV (Hao et al., 2019). However, in the present study, by analyzing the single Calu-3 cell mRNA sequencing data, the up-regulation of the CXCL8 was only found in the cases of SARS-CoV-1 and SARS-CoV-2 infections but not in the case of MERS-CoV infection. In the case of CCL20, a study using human peripheral blood mononuclear cells on CCL20 has reported that an early enhancement in the expression of CCL20 after the infection of SARS-CoV-1 (Ng et al., 2004). However, the study on the expression of CCL20 after the infection of SARS-CoV-2 and MERS-CoV is still insufficient.

CSF2, which is called colony-stimulating factor 2, is a white blood cell growth factor that also functions as a cytokine (Francisco-Cruz et al., 2014). CSF2 fights infection by rapidly growing its number. Interestingly, CSF2 is reported to exhibit a significantly increased serum level at the early stage during SARS-CoV-2 infection (Huang et al., 2020). But the up-regulation of CSF2 is not observed at the early stage of SARS-CoV-1 and MERS-CoV infections, whereas only a moderate up-regulation is observed at the late stage (Sun et al., 2020). The infection mechanism of SARS-CoV-2 may be different with SARS-CoV-1. Therefore, CSF2 is a potential unique biomarker for SARS-CoV-2.

ITGAM (also known as CR3A) is one protein subunit that forms macrophage-1 antigen (Mac-1), which is involved in the innate immune system and mediates inflammation (Solovjov et al., 2005). Studies have reported an increase of ITGAM expression in SARS-CoV-1 (Reghunathan et al., 2005; Kong et al., 2009) and SARS-CoV-2 infected cases (Wangzhou et al., 2020), but such increase has not been reported in MERS-CoV infected cases.

C3 (complement component 3), which is a pivotal complement component in the complement system and contributes to innate immunity (de Bruijn and Fey, 1985), involving coronaviruses infection (Li et al., 2020) and pro-inflammation (Mastellos et al., 2019). The binding between C3 and ITGAM has been experimentally verified by affinity chromatography (Micklem and Sim, 1985). SARS-CoV-1 has been reported to induce the up-regulation of C3 (Danesh et al., 2011). Furthermore, because C3-deficient mice infected with SARS-CoV-1 were reported to exhibit less respiratory dysfunction (Gralinski et al., 2018), the inhibition of C3 was also considered a potential target to alleviate the inflammatory lung complications of SARS-CoV-2 infection (Mastaglio et al., 2020; Risitano et al., 2020).

ISG15 is an ubiquitin-like protein and is called interferon-stimulated gene 15. It is tightly related to specific signaling pathways that are involved in innate immunity. The main cellular function of ISG15 is ISGylation but also has a wide range of anti-viral activity (Morales and Lenschow, 2013; Dzimianski et al., 2019). The papain-like proteases (PLpro) derived from MERS-CoV and SARS-CoV-1 have deISGylating activities and are reported to bind ISG15 for the antagonism of the innate immune response (Lindner et al., 2007; Bailey-Elkin et al., 2014; Mielech et al., 2014; Ratia et al., 2014). Additionally, a current study has
demonstrated that the infection of SARS-CoV-2 will induce the expression of interferon-induced transmembrane family genes, including ISG15 (Khoury et al., 2020). But in the case of SARS-CoV-2, evidence has been showed that PLpro can bind with ISG15.

TRAF1 (TNF receptor-associated factor 1) and NFKB2 (nuclear factor NFκB p100 subunit) are involved in the NFκB signaling pathway that is linked to inflammation and is also predicted as a target pathway for the COVID-19. The expression level of TRAF1 was reported to increase during the SARS-CoV-1 infection (Mitchell et al., 2013) and SARS-CoV-2 infection (Rian et al., 2020a). The expression response of TRAF1 to MERS-CoV infection has not been reported yet. Meanwhile, an increased expression of NFKB2 was reported in COVID-19 patients (Yan et al., 2020), SARS-CoV-2 (Rian et al., 2020), and SARS-CoV-1 infected cells (Tang et al., 2005), but the expression level of NFKB2 were not significantly changed in MERS-CoV infected Calu-3 cells (Selinger et al., 2014), which is in consistent with the data analyzed in the present study (Figure 3).

5 Conclusion

In summary, the eight hub genes identified in the present study are involved in the cytokine response (CXCL8, CCL20, and CSF2), ISGylation (ISG15), macrophage activation (ITGAM), complement system (C3), and NFκB signaling pathway (TRAF1 and NFKB2). Those genes are also highly related to other diseases, such as Type II Diabetes, edema, rosiglitazone, and cancer, which brought a further understanding of the molecular mechanism that COVID-19 patients with those diseases will be more likely to have severe symptoms and mortality. Although the expression level change of CXCL8 has been reported differently in MERS-CoV and SARS-CoV-2 infection, the other hub genes, according to the data published to date, showed the same expression level change as in the present study. The unique network of the eight hub genes will be a potential fingerprint for telling SARS-CoVs infections from MERS-CoV infection and exploit new possible therapeutic targets for fighting COVID-19.

Declarations

Conflict of Interest

None

Author Contributions

P.L provided study design and collected data. M.W performed statistical analysis and data interpretation. M.W and P. L wrote this paper.

Acknowledgments

We thank Xi Zhang for sharing the transcriptome data of MERS-CoV infected Calu-3 cells and Mart M. Lamers for sharing the transcriptome data of SARS-CoV-1/2 infected Calu-3 cells to Sequence Read
Archive (SRA) database.

Footnotes

1. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/

2. https://www.who.int/csr/sars/en/

3. https://www.who.int/en/news-room/fact-sheets/detail/middle-east-respiratory-syndrome-coronavirus-(mers-cov)

4. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

References

Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C., and Garry, R. F. (2020). The proximal origin of SARS-CoV-2. Nat. Med. 26, 450–452. doi:10.1038/s41591-020-0820-9.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: Tool for the unification of biology. Nat. Genet. 25, 25–29. doi:10.1038/75556.

Bader, G. D., and Hogue, C. W. V (2003). An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics 4, 2. doi:10.1186/1471-2105-4-2.

Bailey-Elkin, B. A., Knaap, R. C. M., Johnson, G. G., Dalebout, T. J., Ninaber, D. K., van Kasteren, P. B., et al. (2014). Crystal Structure of the Middle East Respiratory Syndrome Coronavirus (MERS-CoV) Papain-like Protease Bound to Ubiquitin Facilitates Targeted Disruption of Deubiquitinating Activity to Demonstrate Its Role in Innate Immune Suppression. J. Biol. Chem. 289, 34667–34682. doi:10.1074/jbc.M114.609644.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. Ser. B 57, 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x.

Carbon, S., Douglass, E., Dunn, N., Good, B., Harris, N. L., Lewis, S. E., et al. (2019). The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res. 47, D330–D338. doi:10.1093/nar/gky1055.

Chu, H., Chan, J. F.-W., Wang, Y., Yuen, T. T.-T., Chai, Y., Hou, Y., et al. (2020). Comparative replication and immune activation profiles of SARS-CoV-2 and SARS-CoV in human lungs: an ex vivo study with implications for the pathogenesis of COVID-19. Clin. Infect. Dis. doi:10.1093/cid/ciaa410.

Danesh, A., Cameron, C. M., León, A. J., Ran, L., Xu, L., Fang, Y., et al. (2011). Early gene expression events in ferrets in response to SARS coronavirus infection versus direct interferon-alpha2b stimulation. Virology 409, 102–112. doi:10.1016/j.virol.2010.10.002.
de Bruijn, M. H., and Fey, G. H. (1985). Human complement component C3: cDNA coding sequence and derived primary structure. *Proc. Natl. Acad. Sci.* 82, 708–712. doi:10.1073/pnas.82.3.708.

Dimitrov, D. S. (2003). The Secret Life of ACE2 as a Receptor for the SARS Virus. *Cell* 115, 652–653. doi:10.1016/S0092-8674(03)00976-0.

Drosten, C., Günther, S., Preiser, W., van der Werf, S., Brodt, H.-R., Becker, S., et al. (2003). Identification of a Novel Coronavirus in Patients with Severe Acute Respiratory Syndrome. *N. Engl. J. Med.* 348, 1967–1976. doi:10.1056/NEJMoa030747.

Dzimianski, J. V., Scholte, F. E. M., Bergeron, É., and Pegan, S. D. (2019). ISG15: It's Complicated. *J. Mol. Biol.* 431, 4203–4216. doi:10.1016/j.jmb.2019.03.013.

Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32, 3047–3048. doi:10.1093/bioinformatics/btw354.

Francisco-Cruz, A., Aguilar-Santelises, M., Ramos-Espinosa, O., Mata-Espinosa, D., Marquina-Castillo, B., Barrios-Payan, J., et al. (2014). Granulocyte–macrophage colony-stimulating factor: not just another haematopoietic growth factor. *Med. Oncol.* 31, 774. doi:10.1007/s12032-013-0774-6.

Gallagher, T. M., and Buchmeier, M. J. (2001). Coronavirus Spike Proteins in Viral Entry and Pathogenesis. *Virology* 279, 371–374. doi:10.1006/viro.2000.0757.

García-Laorden, M. I., Lorente, J. A., Flores, C., Slutsky, A. S., and Villar, J. (2017). Biomarkers for the acute respiratory distress syndrome: how to make the diagnosis more precise. *Ann. Transl. Med.* 5, 283–283. doi:10.21037/atm.2017.06.49.

Gralinski, L. E., Sheahan, T. P., Morrison, T. E., Menachery, V. D., Jensen, K., Leist, S. R., et al. (2018). Complement Activation Contributes to Severe Acute Respiratory Syndrome Coronavirus Pathogenesis. *MBio* 9. doi:10.1128/mBio.01753-18.

Haagmans, B. L., van den Brand, J. M. A., Raj, V. S., Volz, A., Wohlsein, P., Smits, S. L., et al. (2016). An orthopoxvirus-based vaccine reduces virus excretion after MERS-CoV infection in dromedary camels. *Science (80-. ).* 351, 77–81. doi:10.1126/science.aad1283.

Hao, X., Lv, Q., Li, F., Xu, Y., and Gao, H. (2019). The characteristics of hDPP4 transgenic mice subjected to aerosol MERS coronavirus infection via an animal nose-only exposure device. *Anim. Model. Exp. Med.* 2, 269–281. doi:10.1002/ame2.12088.

He, R., Dobie, F., Ballantine, M., Leeson, A., Li, Y., Bastien, N., et al. (2004a). Analysis of multimerization of the SARS coronavirus nucleocapsid protein. *Biochem. Biophys. Res. Commun.* 316, 476–483. doi:10.1016/j.bbrc.2004.02.074.
He, Y., Zhou, Y., Liu, S., Kou, Z., Li, W., Farzan, M., et al. (2004b). Receptor-binding domain of SARS-CoV spike protein induces highly potent neutralizing antibodies: implication for developing subunit vaccine. *Biochem. Biophys. Res. Commun.* 324, 773–781. doi:10.1016/j.bbrc.2004.09.106.

Ho, T.-Y., Wu, S.-L., Cheng, S.-E., Wei, Y.-C., Huang, S.-P., and Hsiang, C.-Y. (2004). Antigenicity and receptor-binding ability of recombinant SARS coronavirus spike protein. *Biochem. Biophys. Res. Commun.* 313, 938–947. doi:10.1016/j.bbrc.2003.11.180.

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., et al. (2020). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* 181, 271-280.e8. doi:10.1016/j.cell.2020.02.052.

Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., et al. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 395, 497–506. doi:10.1016/S0140-6736(20)30183-5.

Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 37, 1–13. doi:10.1093/nar/gkn923.

Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57. doi:10.1038/nprot.2008.211.

Jiaming, L., Yanfeng, Y., Yao, D., Yawei, H., Linlin, B., Baoying, H., et al. (2017). The recombinant N-terminal domain of spike proteins is a potential vaccine against Middle East respiratory syndrome coronavirus (MERS-CoV) infection. *Vaccine* 35, 10–18. doi:10.1016/j.vaccine.2016.11.064.

Kanehisa, M. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 28, 27–30. doi:10.1093/nar/28.1.27.

Kanehisa, M. (2019). Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* 28, 1947–1951. doi:10.1002/pro.3715.

Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K., and Tanabe, M. (2019). New approach for understanding genome variations in KEGG. *Nucleic Acids Res.* 47, D590–D595. doi:10.1093/nar/gky962.

Khouri, M., Cuenca, J., Cruz, F. F., Figueroa, F. E., Rocco, P. R. M., and Weiss, D. J. (2020). Current Status of Cell-Based Therapies for Respiratory Virus Infections: Applicability to COVID-19. *Eur. Respir. J.* doi:10.1183/13993003.00858-2020.

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–60. doi:10.1038/nmeth.3317.

Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37, 907–915. doi:10.1038/s41587-019-
Kong, S. L., Chui, P., Lim, B., and Salto-Tellez, M. (2009). Elucidating the molecular physiopathology of acute respiratory distress syndrome in severe acute respiratory syndrome patients. *Virus Res.* 145, 260–269. doi:10.1016/j.virusres.2009.07.014.

Ksiazek, T. G., Erdman, D., Goldsmith, C. S., Zaki, S. R., Peret, T., Emery, S., et al. (2003). A Novel Coronavirus Associated with Severe Acute Respiratory Syndrome. *N. Engl. J. Med.* 348, 1953–1966. doi:10.1056/NEJMoa030781.

Lamers, M. M., Beumer, J., van der Vaart, J., Knoops, K., Puschhof, J., Breugem, T. I., et al. (2020). SARS-CoV-2 productively infects human gut enterocytes. *Science (80-. ).*, eabc1669. doi:10.1126/science.abc1669.

Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R., and Pfister, H. (2014). UpSet: Visualization of Intersecting Sets. *IEEE Trans. Vis. Comput. Graph.* 20, 1983–1992. doi:10.1109/TVCG.2014.2346248.

Li, F. (2005). Structure of SARS Coronavirus Spike Receptor-Binding Domain Complexed with Receptor. *Science (80-. ).* 309, 1864–1868. doi:10.1126/science.1116480.

Li, G., Fan, Y., Lai, Y., Han, T., Li, Z., Zhou, P., et al. (2020). Coronavirus infections and immune responses. *J. Med. Virol.* 92, 424–432. doi:10.1002/jmv.25685.

Li, W., Moore, M. J., Vasilieva, N., Sui, J., Wong, S. K., Berne, M. A., et al. (2003). Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* 426, 450–454. doi:10.1038/nature02145.

Liang, W., Guan, W., Chen, R., Wang, W., Li, J., Xu, K., et al. (2020). Cancer patients in SARS-CoV-2 infection: a nationwide analysis in China. *Lancet Oncol.* 21, 335–337. doi:10.1016/S1470-2045(20)30096-6.

Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. doi:10.1093/bioinformatics/btt656.

Lindner, H. A., Lytvyn, V., Qi, H., Lachance, P., Ziomek, E., and Ménard, R. (2007). Selectivity in ISG15 and ubiquitin recognition by the SARS coronavirus papain-like protease. *Arch. Biochem. Biophys.* 466, 8–14. doi:10.1016/j.abb.2007.07.006.

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi:10.1186/s13059-014-0550-8.

Mastaglio, S., Ruggeri, A., Risitano, A. M., Angelillo, P., Yancopoulou, D., Mastellos, D. C., et al. (2020). The first case of COVID-19 treated with the complement C3 inhibitor AMY-101. *Clin. Immunol.* 215, 108450.
Mastellos, D. C., Ricklin, D., and Lambris, J. D. (2019). Clinical promise of next-generation complement therapeutics. *Nat. Rev. Drug Discov.* 18, 707–729. doi:10.1038/s41573-019-0031-6.

Matsuyama, S., Nagata, N., Shirato, K., Kawase, M., Takeda, M., and Taguchi, F. (2010). Efficient Activation of the Severe Acute Respiratory Syndrome Coronavirus Spike Protein by the Transmembrane Protease TMPRSS2. *J. Virol.* 84, 12658–12664. doi:10.1128/JVI.01542-10.

Micklem, K. J., and Sim, R. B. (1985). Isolation of complement-fragment-iC3b-binding proteins by affinity chromatography. The identification of p150,95 as an iC3b-binding protein. *Biochem. J.* 231, 233–236. doi:10.1042/bj2310233.

Mielech, A. M., Kilianski, A., Baez-Santos, Y. M., Mesecar, A. D., and Baker, S. C. (2014). MERS-CoV papain-like protease has deISGylating and deubiquitinating activities. *Virology* 450–451, 64–70. doi:10.1016/j.virol.2013.11.040.

Mitchell, H. D., Eisfeld, A. J., Sims, A. C., McDermott, J. E., Matzke, M. M., Webb-Robertson, B.-J. M., et al. (2013). A Network Integration Approach to Predict Conserved Regulators Related to Pathogenicity of Influenza and SARS-CoV Respiratory Viruses. *PLoS One* 8, e69374. doi:10.1371/journal.pone.0069374.

Mieres, D. J., and Lenschow, D. J. (2013). The Antiviral Activities of ISG15. *J. Mol. Biol.* 425, 4995–5008. doi:10.1016/j.jmb.2013.09.041.

Ng, L. F., Hibberd, M. L., Ooi, E.-E., Tang, K.-F., Neo, S.-Y., Tan, J., et al. (2004). A human in vitro model system for investigating genome-wide host responses to SARS coronavirus infection. *BMC Infect. Dis.* 4, 34. doi:10.1186/1471-2334-4-34.

Peiris, J., Lai, S., Poon, L., Guan, Y., Yam, L., Lim, W., et al. (2003a). Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet* 361, 1319–1325. doi:10.1016/S0140-6736(03)13077-2.

Peiris, J. S. M., Yuen, K. Y., Osterhaus, A. D. M. E., and Stöhr, K. (2003b). The Severe Acute Respiratory Syndrome. *N. Engl. J. Med.* 349, 2431–2441. doi:10.1056/NEJMra032498.

Qihong Yan, Pingchao Li, Xianmiao Ye, Xiaohan Huang, Xiaoneng Mo, Qian Wang, Yudi Zhang, Kun Luo, Zhaoming Chen, Jia Luo, Xuefeng Niu, Ying Feng, Tianxing Ji, Bo Feng, Jinlin Wang, Feng Li, Fuchun Zhang, Fang Li, Jianhua Wang, Liqiang Feng, Zhilong Chen, L. C. (2020). Longitudinal peripheral blood transcriptional analysis of COVID-19 patients captures disease progression and reveals potential biomarkers. *medRxiv* (preprint).

Raj, V. S., Mou, H., Smits, S. L., Dekkers, D. H. W., Müller, M. A., Dijkman, R., et al. (2013). Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. *Nature* 495, 251–254. doi:10.1038/nature12005.
Ratia, K., Kilianski, A., Baez-Santos, Y. M., Baker, S. C., and Mesecar, A. (2014). Structural Basis for the Ubiquitin-Linkage Specificity and deISGylating Activity of SARS-CoV Papain-Like Protease. *PLoS Pathog.* 10, e1004113. doi:10.1371/journal.ppat.1004113.

Reghunathan, R., Jayapal, M., Hsu, L. Y., Chng, H. H., Tai, D., Leung, B. P., et al. (2005). Expression profile of immune response genes in patients with severe acute respiratory syndrome. *BMC Immunol.* 6. doi:10.1186/1471-2172-6-2.

Rian, K., Esteban-Medina, M., Hidalgo, M. R., Cubuk, C., Falco, M. M., Loucera, C., et al. (2020a). Mechanistic modeling of the SARS-CoV-2 disease map. *bioRxiv* (preprint) 19, 2020.04.12.025577. doi:10.1101/2020.04.12.025577.

Rian, K., Esteban-Medina, M., Hidalgo, M. R., Cubuk, C., Falco, M. M., Loucera, C., et al. (2020b). Mechanistic modeling of the SARS-CoV-2 disease map. *bioRxiv* (preprint) 19, 2020.04.12.025577. doi:10.1101/2020.04.12.025577.

Risitano, A. M., Mastellos, D. C., Huber-Lang, M., Yancopoulou, D., Garlanda, C., Ciceri, F., et al. (2020). Complement as a target in COVID-19? *Nat. Rev. Immunol.* doi:10.1038/s41577-020-0320-7.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43, e47–e47. doi:10.1093/nar/gkv007.

Selinger, C., Tisoncik-Go, J., Menachery, V. D., Agnihotram, S., Law, G., Chang, J., et al. (2014). Cytokine systems approach demonstrates differences in innate and pro-inflammatory host responses between genetically distinct MERS-CoV isolates. *BMC Genomics* 15, 1161. doi:10.1186/1471-2164-15-1161.

Shannon, P. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* 13, 2498–2504. doi:10.1101/gr.1239303.

Solovjov, D. A., Pluskota, E., and Plow, E. F. (2005). Distinct Roles for the α and β Subunits in the Functions of Integrin αMβ2. *J. Biol. Chem.* 280, 1336–1345. doi:10.1074/jbc.M406968200.

Spiegel, M., and Weber, F. (2006). Inhibition of cytokine gene expression and induction of chemokine genes in non-lymphatic cells infected with SARS coronavirus. *Virol. J.* 3, 17. doi:10.1186/1743-422X-3-17.

SRA Handbook [Internet]. Using the SRA (2010). *Bethesda Natl. Cent. Biotechnol. Inf.*

Sun, J., Ye, F., Wu, A., Yang, R., Pan, M., and Sheng, J. (2020). Comparative transcriptome analysis reveals the intensive early-stage responses of host cells to SARS-CoV-2 infection. *medRxiv* (preprint).

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., et al. (2015). STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 43, D447–D452. doi:10.1093/nar/gku1003.
Tang, B. S. F., Chan, K., Cheng, V. C. C., Woo, P. C. Y., Lau, S. K. P., Lam, C. C. K., et al. (2005). Comparative Host Gene Transcription by Microarray Analysis Early after Infection of the Huh7 Cell Line by Severe Acute Respiratory Syndrome Coronavirus and Human Coronavirus 229E. *J. Virol.* 79, 6180–6193. doi:10.1128/JVI.79.10.6180-6193.2005.

van Boheemen, S., de Graaf, M., Lauber, C., Bestebroer, T. M., Raj, V. S., Zaki, A. M., et al. (2012). Genomic Characterization of a Newly Discovered Coronavirus Associated with Acute Respiratory Distress Syndrome in Humans. *MBio* 3. doi:10.1128/mBio.00473-12.

Walls, A. C., Park, Y.-J., Tortorici, M. A., Wall, A., McGuire, A. T., and Veesler, D. (2020). Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* 181, 281-292.e6. doi:10.1016/j.cell.2020.02.058.

Wan, Y., Shang, J., Graham, R., Baric, R. S., and Li, F. (2020). Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *J. Virol.* 94. doi:10.1128/JVI.00127-20.

Wang, N., Shi, X., Jiang, L., Zhang, S., Wang, D., Tong, P., et al. (2013). Structure of MERS-CoV spike receptor-binding domain complexed with human receptor DPP4. *Cell Res.* 23, 986–993. doi:10.1038/cr.2013.92.

Wang, W., Xu, Y., Gao, R., Lu, R., Han, K., Wu, G., et al. (2020). Detection of SARS-CoV-2 in Different Types of Clinical Specimens. *JAMA*. doi:10.1001/jama.2020.3786.

Wangzhou, A., Paige, C., Neerukonda, S. V, Dussor, G., Ray, P. R., and Price, T. J. (2020). A pharmacological interactome platform for discovery of pain mechanisms and targets. *bioRxiv* (preprint), 2020.04.14.041715. doi:10.1101/2020.04.14.041715.

Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *Omi. A J. Integr. Biol.* 16, 284–287. doi:10.1089/omi.2011.0118.

Zhang, X., Chu, H., Wen, L., Shuai, H., Yang, D., Wang, Y., et al. (2020). Competing endogenous RNA network profiling reveals novel host dependency factors required for MERS-CoV propagation. *Emerg. Microbes Infect.* 9, 733–746. doi:10.1080/22221751.2020.1738277.

Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., et al. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579, 270–273. doi:10.1038/s41586-020-2012-7.

Zhu, A., Ibrahim, J. G., and Love, M. I. (2019). Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinformatics* 35, 2084–2092. doi:10.1093/bioinformatics/bty895.
Zlotnik, A., Burkhardt, A. M., and Homey, B. (2011). Homeostatic chemokine receptors and organ-specific metastasis. *Nat. Rev. Immunol.* 11, 597–606. doi:10.1038/nri3049.

Zlotnik, A., and Yoshie, O. (2012). The Chemokine Superfamily Revisited. *Immunity* 36, 705–716. doi:10.1016/j.immuni.2012.05.008.

Figures
Figure 1

Volcano plot of DEGs in (A) MERS-CoV infected Calu-3 cells, (B) SARS-CoV-1 infected Calu-3 cells, and (C) SARS-CoV-2 infected Calu-3 cells. Gray, red and blue color represented relatively equal, up and down regulated genes respectively, based on the threshold of log2FC shrinkage value > 1.0 or < –1.0 and -log10 (p-value) >1.3.
Figure 2

Upset-plot of the intersections of DEGs from two independent datasets (MERS-CoV infection and SARS-CoVs infections). A total of 64 overlapping DEGs were identified. The bar plot on the lower left presented the gene counts in each DEG group. The bar plot on the upper right showed the intersections of the gene counts among different DEG groups.
Figure 3

Heat-map showing the normalized expression values in transcripts per million (TPM) of the overlapping DEGs identified from MERS-CoV infection and SARS-CoVs infections. DEGs were classified into four groups. '↑↑↑' represented genes that were up-regulated in both MERS-CoV infection and SARS-CoVs infections. '↑↓↓' represented the gene that was up-regulated in MERS-CoV infection but were down-regulated in SARS-CoVs infections. '↓↑↑' represented genes that were down-regulated in MERS-CoV infection but were up-regulated in SARS-CoVs infections. '↓↓↓' represented genes that were down-regulated in both MERS-CoV infection and SARS-CoVs infections.
infection but up-regulated in SARS-CoVs infections. ‘↓↓’ represented genes that were down-regulated both in MERS-CoV infection and SARS-CoVs infections. The SARS-CoVs infections included the infection of SARS-CoV-1 and SARS-CoV-2.

Figure 4

GO (A) and KEGG pathway (B) enrichment analyses of the overlapping 64 DEGs. The top ten biofunctions and pathways were presented, ranked by gene ratio (gene counts in each pathway to 64 input DEGs).
Figure 5

Potential co-regulatory network of the DEGs. (A) The PPI network consisting of 38 DEGs predicted by STRING database. (B) The top 10 hub genes and their interaction network identified by MCODE algorithm. The red rectangles represented the genes that were up-regulated in both MERS-CoV infection and SARS-CoVs infections. The orange rectangles represented the genes that were down-regulated in MERS-CoV infection but up-regulated in SARS-CoVs infections. The blue rectangles represented the genes that were down-regulated both in MERS-CoV infection and SARS-CoVs infections. The rose red rectangle represented the gene that was up-regulated in MERS-CoV infection but were down-regulated in SARS-CoVs infections. The genes that had more than 5 nodes were highlighted by black boxes and their numbers of nodes were marked at the lower right corner. The SARS-CoVs infections included the infection of SARS-CoV-1 and SARS-CoV-2.
Figure 6

Diseases enrichment analysis. Top 20 diseases enriched in DAVID database and ranked by -log (p-value). The red dotted line represented $p = 0.05$. The blue columns that over the red dotted line denoted significant diseases that are related to the DEGs. The orange dots represented the ratio of gene counts in each disease to 64 input DEGs.
Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplemaentarytables.docx