

Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

Nanxi Guo

Date

Transcriptome Sequencing of Peripheral Blood Reveals Distinct Expression Landscapes
of COVID-19 and MIS-C Patients

By

Nanxi Guo

Master of Science in Public Health

Biostatistics and Bioinformatics

Zhaohui “Steve” Qin, Ph.D.

(Thesis Advisor)

Hao Wu, Ph.D.

(Reader)

Transcriptome Sequencing of Peripheral Blood Reveals Distinct Expression Landscapes
of COVID-19 and MIS-C Patients

By

Nanxi Guo

B.S.

Hunan University

2019

Thesis Advisor: Zhaohui “Steve” Qin, Ph.D.

Reader: Hao Wu, Ph.D.

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University in
partial fulfillment of the requirements for the degree of
Master of Science in Public Health in
Biostatistics and Bioinformatics

2021

Abstract

Transcriptome Sequencing of Peripheral Blood Reveals Distinct Expression Landscapes of COVID-19 and MIS-C Patients

By Nanxi Guo

Background: The pandemic spread of the coronavirus disease 2019 (COVID-19) has been proclaimed a severe public health emergency of international concern (PHEIC) by the WHO. Of note, children were reported symptoms similar to severe COVID-19, which has then been confirmed as a rare complication of COVID-19 in children, termed as multisystem inflammatory syndrome (MIS-C). However, little is known about its genetic mechanisms compared to COVID-19.

Methods: RNA-sequencing of blood transcriptome were performed from COVID-19 and MIS-C patients in both mild and severe scenarios, as well as healthy control samples. Then bioinformatic approaches were performed to identify distinct expression landscapes of COVID-19 and MIS-C patients.

Results: Peripheral blood transcriptomes of 8 COVID-19 patients, 15 MIS-C patients and 8 healthy controls were depicted here. Immune response of COVID-19 patients compared to healthy controls were first detected. Neutrophil activation-associated terms, lymphocyte differentiation terms, and cardiomyopathy-associated terms were more enriched in comparison between COVID-19 and MIS-C, in severe illness levels. Further analyses on immune molecular signatures by peripheral blood mononuclear cells (PBMC) revealed increasing of T cells (CD8⁺ T memory cell markers), CD14 cells in COVID-19 vs MIS-C comparison.

Conclusions: This study supported the clinical differentiation between MIS-C patients and COVID-19 patients in genetic way. Therefore, several evidence were provided for diagnostic biomarkers in COVID-19 and MIS-C patients, for better screening and diagnostics, and to treat patients with a personalized approach.

Transcriptome Sequencing of Peripheral Blood Reveals Distinct Expression Landscapes
of COVID-19 and MIS-C Patients

By

Nanxi Guo

B.S.

Hunan University

2019

Thesis Advisor: Zhaohui “Steve” Qin, Ph.D.

Reader: Hao Wu, Ph.D.

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University in
partial fulfillment of the requirements for the degree of
Master of Science in Public Health in
Biostatistics and Bioinformatics
2021

Background

Pandemic spread of coronavirus, severe acute respiratory syndrome-coronavirus 2 (SARS-CoV-2), has resulted in over 130 million confirmed cases and over 2 million deaths worldwide (WHO, covid19.who.int, as of April 11, 2021), called 2019-novel coronavirus diseases (COVID-19) [1]. Current clinical research efforts primarily focus on adult cohort studies for functional vaccine development and pharmacological interventions, with patients greater than or equal to 85 years old showing the highest hospitalization (70.3%) and case fatality (27.3%) rates [2]. However, there are documented rare cases where children can be severely affected without discernable underlying conditions [3-7]. In April of 2020, United Kingdom reported the first case of presentation in children similar to incomplete Kawasaki disease (KD) or toxic shock syndrome [8]. Then more and more cases have been reported worldwide. This condition has been termed as multisystem inflammatory syndrome (MIS-C) in children by the CDC, which has also been confirmed that MIS-C is associated with the virus that causes COVID-19. MIS-C appears to be a rare complication of COVID-19 in children, occurring in less than 1 percent of children with confirmed SARS-CoV02 infection [9]. The pathophysiology of MIS-C is not well understood. As for now, documents were limited in the epidemiology, pathophysiology, clinical presentation and diagnosis of MIS-C.

Hence, to investigate the mechanisms of the immune response and immune molecular signatures in COVID-19 patients and MIS-C patients is currently an active topic.

In this work, bioinformatics approaches were conducted on a total of 31 whole blood transcriptomes in order to identify host genomic, transcriptomic and immune molecular signature associated with differentiation in severe COVID-19 and MIS-C. This study not only represent the genetic risk factors of these two scenarios, but also be urgently needed for identification of diagnostic biomarkers, therapeutic of disease and development of vaccines.

Methods

Human cohorts

A total of 8 children confirmed with SARS-CoV-2 infection by RT-PCR, serology, or antigen test, with mild or severe symptoms but no underlying conditions, were recruited in this study. Another 15 children meeting MIS-C criteria, with positive for current or recent SARS-CoV-2 infection, with mild or severe symptoms were enrolled in this study as well. MIS-C was diagnosed by the criteria provided by the CDC. Another 8 healthy children (<15 yrs) were recruited as non-disease controls. They were subjected to testing the nasopharyngeal secretion for SARS-CoV-2 and were all confirmed to be asymptomatic and seronegative. All the whole blood samples were collected within 72h of hospital admission.

RNA-sequencing

RNA-sequencing was conducted on a total of 31 subjects, on a high-throughput next-generation sequencing (NGS) Illumina NovaSeq platform using 2X150bp reads with >98% coverage of the 22,000 genes in the genome at $\geq 30X$. Then the 31 whole blood samples were ready for RNA-sequencing analysis.

RNA-sequencing analysis

Fastq sequencing files were aligned to the hg38 human reference genome using the STAR aligner (v2.7.7a). Counts matrix were imported using `DESeqDataSetFromMatrix` function from DESeq2 (v1.28.1). DESeq2 was used for counts normalization for each transcript using default parameters, resulting in 28,277 normalized transcripts.

Differential expression analysis was conducted for the scenario status (COVID-19 vs controls, COVID-19 vs MIS-C), mild/severe (severe COVID-19 vs severe MIS-C, mild COVID-19 vs mild MIS-C). Then principal component analysis was conducted using all transcripts as input. Differentially expressed genes (DEGs) were selected with the fold change of log₂-transformed expression level > 1 and $p < 0.05$. DEGs were used as input for volcano plots and visualized as bar plots. Gene Ontology enrichment analysis (GOEA) was then performed using R package ClusterProfiler (v3.16.1) to determine significant enrichment of biological processes related to each comparison group (p value < 0.05 ; q value < 0.05). Gene Set Enrichment Analysis (GSEA) was conducted to further investigate the gene sets with concordant differences from the whole expression matrix.

GSEA was applied to KEGG and Reactome gene sets. Differential cell type change analysis using Logistic/linear regression (LRcell) was performed to identify cell type of human body immunity that drives the changes observed in the differential gene expression experiment [24]. Peripheral blood mononuclear cells (PBMC) were chosen as marker genes when applying Logistic Regression on the top DEGs with p-values. Then the cell types for which were responsible in DEGs were identified.

Results

Immune Response of COVID-19 Patients Compared to Healthy Controls

To investigate the monogenic causes, immune response and immune molecular signature of COVID-19 patients and MIS-C patients, blood samples were analyzed from 8 healthy controls, 8 COVID-19 patients, including 3 severe cases and 5 mild cases, and 15 MIS-C patients, including 10 severe cases and 5 mild cases, by RNA sequencing (Table 1).

Uniquely mapped reads of 31 samples were obtained after aligning to the hg38 human reference genome (Table 2), with 65% to 73% percent uniquely mapped reads. Principal component analysis (PCA) was performed to represent two-dimensional data, showing separation of healthy samples and COVID-19, MIS-C samples (Figure 1A). Based on PCA, the two PCs account for 47% and 13%, respectively, of the total variation.

Distribution of the COVID-19, MIS-C, and healthy control samples in the PCA revealed heterogeneity in the transcriptomic profiles (Figure 1A). In order to further investigate it,

differentially expressed genes (DEGs) were then compared between COVID-19 patients and healthy controls, and 1639 upregulated and 920 downregulated genes were identified ($FC > |2|$, $p_{adj} < 0.05$; Figure 1B). Significantly, ID1 was the most prominently upregulated genes [10], with the lowest p value (Figure 1C). Heightened expression was found for several pro-inflammatory genes and anti-inflammatory markers, such as Early Growth Response 1 (EGR1), transcobalamin 2 (TCN2), Tumor Necrosis Factor Alpha-Induced Protein 2 (TNFAIP3), Complement C1q Subcomponent Subunit C (C1QC), and Myeloperoxidase (MPO) [11,12], in the macrophages (Figure 1C). Downregulated genes included several T lymphocyte-related and antibody factors, such as C-X-C Motif Chemokine Receptor 3 (CXCR3), Charged Multivesicular Body Protein 7 (CHMP7). Gene ontology enrichment analysis (GOEA) on the differentially expressed genes (DEGs) revealed neutrophil activation-associated terms, lymphocyte differentiation, inflammatory response, and T cell differentiation terms (Figure 1D). GSEA showed that 15 and 107 pathways were significantly enriched in KEGG and Reactome, respectively, when the adjusted p value < 0.05 (Table 3). Specifically, leukocyte transendothelial migration, extracellular matrix organization, cytokine signaling in immune system, neutrophil degranulation, and infection disease pathways were upregulated and shared in both GOEA and GSEA, which were the gene sets that need to be studied.

Table 1. Clinical Parameters of COVID-19 Patients and MIS-C Patients

	Total COVID-19 patients N=8	COVID Severe N=3	COVID Mild N=5	Total MIS-C patients N=15	MIS-C Severe N=10	MIS-C Mild N=5
Sex (n, m/f)	3/5	1/2	2/3	6/9	4/6	2/3
Age (years, mean \pm SD)	12.9 (2.8)	12 (4.4)	13.4 (1.8)	7.6 (4.4)	8.4 (4.7)	6.0 (3.4)
Laboratory values (mean \pm SD)						
Total blood						
neutrophil count (10 ³ cells/ μ L)	5.7 (2.7)	6.8 (4.3)	4.8 (0.7)	8.1 (6.2)	8.2 (7.1)	7.9 (4.4)
Total blood						
lymphocyte count (10 ³ cells/ μ L)	1.5 (1.3)	2.0 (1.7)	1.1(0.9)	0.8 (0.3)	0.6 (0.2)	1.1 (0.3)

Table 2. Uniquely mapped reads of 31 samples generated by STAR aligner

Sample Name	Uniquely mapped reads (millions, %)	Sample Name	Uniquely mapped reads (millions, %)
Healthy controls		Mild MIS-C patients	
PHC4	28.6 (67.4)	PTCOV20	31.9 (70.7)
PHC5	33.9 (66.7)	PTCOV22	81.5 (72.8)
PHC6	56.4 (69.8)	PTCOV34	29.1 (72.2)
PHC7	39.2 (67.7)	PTCOV42	43.4 (69.7)
PHC8	46.9 (68.2)	PTCOV60	41.4 (67.7)
PHC9	54.4 (69.0)	Severe MIS-C patients	
PHC10	36.3 (69.3)	PTCOV28	36.7 (70.2)
PHC12	34.8 (64.9)	PTCOV32	49.0 (72.8)
Mild COVID-19 patients		PTCOV33	56.2 (71.5)
PTCOV18	25.9 (67.2)	PTCOV45	43.4 (70.6)
PTCOV24	49.7 (67.7)	PTCOV48	55.0 (69.8)
PTCOV25	36.1 (66.8)	PTCOV51	40.8 (66.3)
PTCOV27	42.1 (67.6)	PTCOV54	40.2 (72.6)
PTCOV44	29.3 (67.8)	PTCOV59	40.1 (67.0)
Severe COVID-19 patients		PTCOV68	36.9 (69.0)
PTCOV19	65.6 (73.0)	PTCOV72	24.4 (66.0)
PTCOV29	47.6 (68.4)		
PTCOV31	32.4 (66.7)		

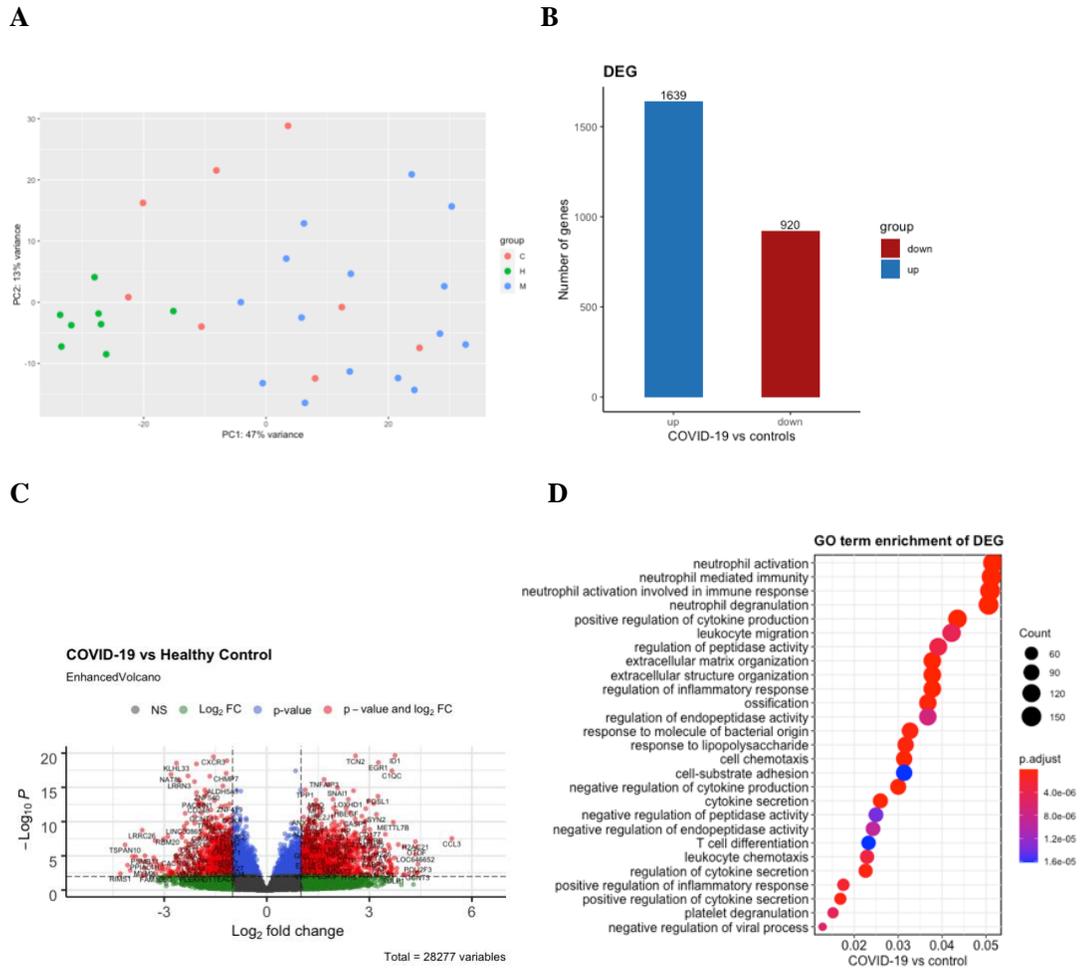


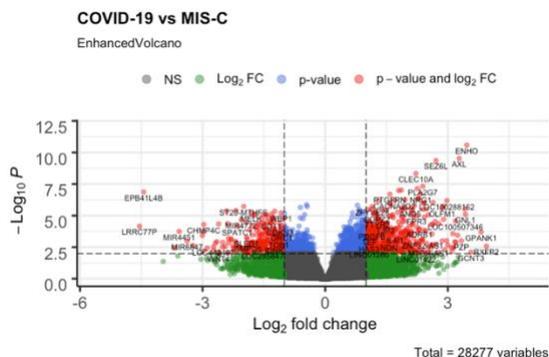
Fig 1. Immune response of COVID-19 patients compared to healthy controls. **a** PCA plot depicting relationship of all samples included mild and severe COVID-19 and MIS-C samples as well as healthy control samples. **b** Bar plot of significantly upregulated and downregulated genes ($FC > |2|$, $adj.p \text{ value} < 0.05$) comparing COVID-19 and healthy control samples. **c** Volcano plot depicting \log_2 fold change (FC) and adjusted p values comparing COVID-19 and healthy control samples. Differentially expressed genes were shown and selected genes are labeled via gene names. **d** Dot plot of the top 27 most enrichment GO terms for differentially expressed genes comparing COVID-19 and healthy control samples.

Differences in Gene Expression in COVID-19 Patients and MIS-C Patients

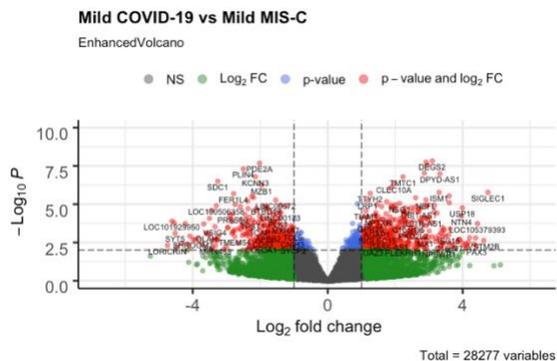
Given the heterogeneous clinical presentation of COVID-19 and MIS-C, samples were classified as “mild” or “severe” by disease severity based on WHO ordinal scale. Overall, there was a smaller number of DEGs in blood samples from COVID-19 patients than in MIS-C patients when compared to healthy controls. Upregulated genes showed greater fold changes than the downregulated genes. Several DEGs were found in the COVID-19 vs MIS-C comparison (Figure 2A), which were different from that in the COVID-19 vs healthy control comparison, including homeostasis, lung cancer-associated genes, and anti-inflammation genes, such as energy homeostasis associated gene (ENHO), SEZ6L, and AXL [13]. Mild COVID-19 versus mild MIS-C shared lymphohematopoietic-associated gene with all the COVID-19 versus MIS-C, such as CLEC10A expression, representing a candidate to better define CD1c⁺ DCs as well as a candidate receptor for future antigen-targeting approaches [14], among the upregulated DEGs (Figure 2B). Comparison between COVID-19 and MIS-C patients in severity condition has the smallest number of DEGs, and 5 upregulated genes and 8 downregulated genes were identified ($|FC| > 2$; $p_{adj} < 0.05$). A candidate cardiomyopathy gene, SORBS2, was one of the most significant upregulated gene with the largest fold change (Figure 2C). DEGs were also found for dystonia-associated gene, SPR, with the lowest p value (Figure 2C). GOEA reflected these findings as well. GOEA in the mild COVID-19 vs mild MIS-C sample comparison (Figure 2E) included negative regulation of locomotion, muscle

organ development, down regulation of viral life process, and inflammatory response terms, which shared the similar terms in COVID-19 vs MIS-C comparison (Figure 2D). Whereas these terms were not differentially regulated in severe COVID-19 vs severe MIS-C samples, regulation of vasculature and heart contraction terms were more pronounced here (Figure 2F). GSEA identified similar findings with that in GOEA. GSEA of COVID-19 versus MIS-C samples identified that 31 gene sets were significantly enriched in the Reactome signature database (Table 4), including cell cycle, neutrophil degranulation pathway. Moreover, pathways included cell cycle, infectious disease, influenza infection, and extracellular matrix organization were significantly enriched in the Reactome signature database comparing mild COVID-19 and mild MIS-C patients (Table 5). Although all samples from COVID-19 versus MIS-C patients showed functional enrichment for inflammatory response, infectious disease, neutrophil degranulation and influenza infection associated terms in general, direct comparison of severe COVID-19 and MIS-C patients revealed additional cardiomyopathy-associated characteristic (hypertrophic cardiomyopathy hcm and dilated cardiomyopathy) of the pathways in severe COVID-19 and MIS-C (Table 6). Overall, differences in gene expression were not merely restricted to neutrophil and T lymphocyte factors, but included cardiomyopathy- and olfactory -associated pathways, indicating complicated transcriptionally altered in COVID-19 and MIS-C patients [15].

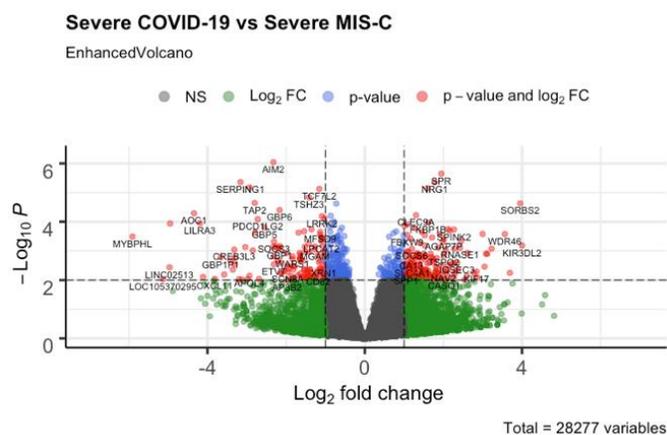
A



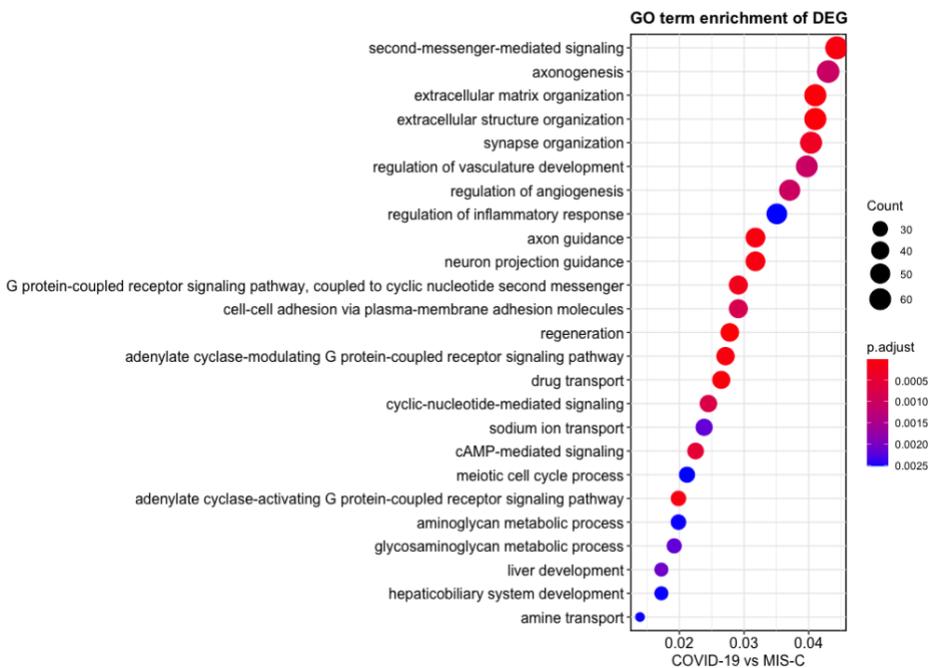
B



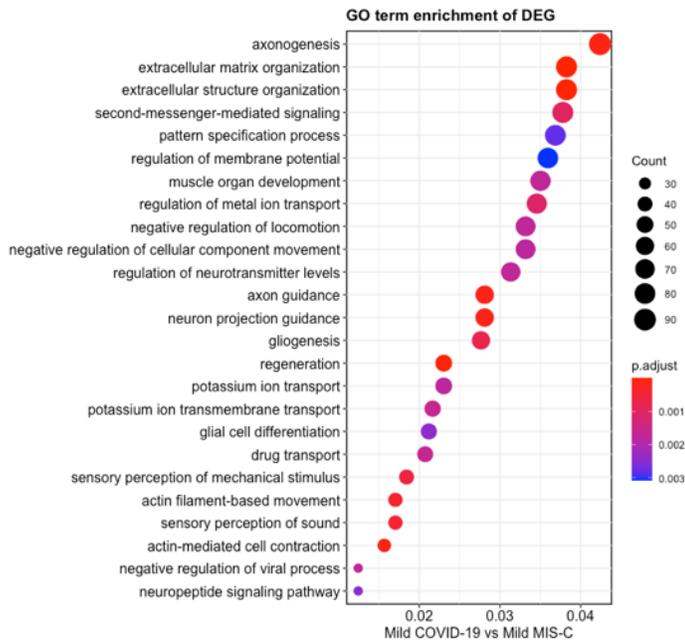
C



D



E



F

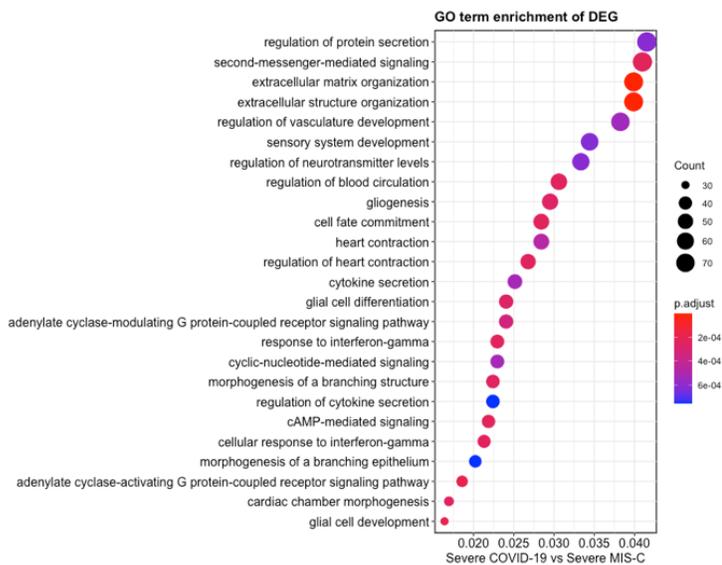
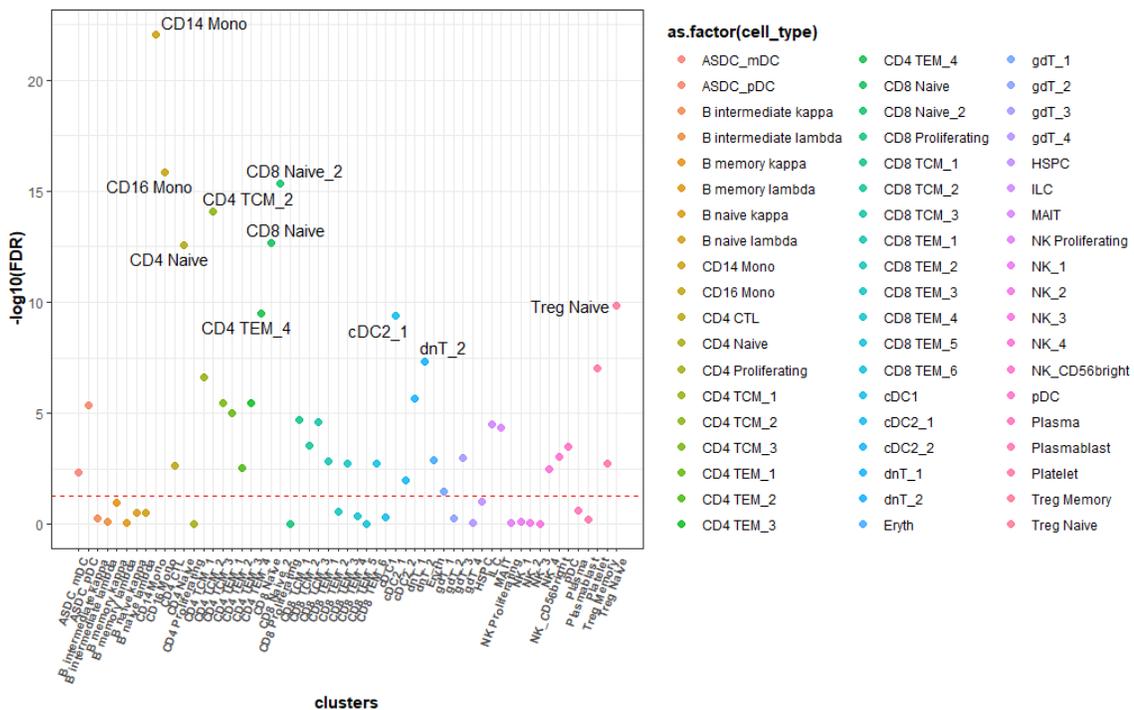


Fig 2. Differences in gene expression in COVID-19 patients and MIS-C patient. **a-c** Volcano plot depicting \log_2 fold change (FC) and adjusted p values comparing COVID-19 and MIS-C samples, mild COVID-19 and mild MIS-C samples, severe COVID-19 and severe MIS-C samples, respectively. Differentially expressed genes were shown and selected genes are labeled via gene names. **d-f** Dot plot of the top 25 most enrichment GO terms for differentially expressed genes comparing COVID-19 and MIS-C samples, mild COVID-19 and mild MIS-C samples, severe COVID-19 and severe MIS-C samples, respectively.

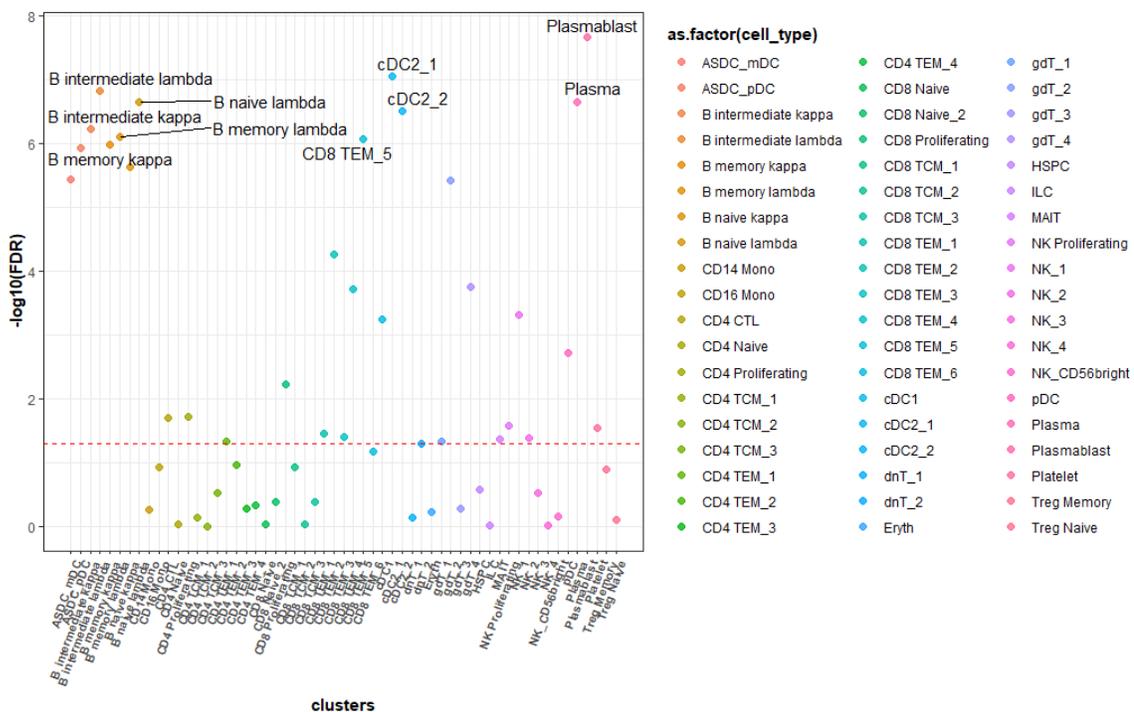
Immune Molecular Signatures of COVID-19 Patients Compared to MIS-C Patients

Differential cell type change analysis using Logistic/linear regression (LRcell) was used to further identify immune molecular that drives the top differential gene expression in RNA-seq analysis. Several T cells were identified in COVID-19 verses MIS-C sample comparison when PBMC was used as marker genes (Figure 3A), such as CD8⁺ T memory cell markers (CD8 TEM_2, CD8 TEm_4, CD8 TEM_5), which were antigen specific and provided an enhanced protective response when the same antigen is encountered again [16], and Natural killer T cells (NK_1), which shared properties of T cells, but their T-cell receptors were far more limited in diversity [17]. For example, the SARS-CoV2-reactive memory CD8⁺ T cells were more pronounced in patients with mild COVID-19 illness compared to those with MIS-C mild illness [18]. These differences were more significant in severe COVID-19 patients versus severe MIS-C patients. Moreover, a cell surface receptor and differentiation marker [19], CD14, was involved that caused the difference of gene enrichment in mild COVID-19 vs mild MIS-C sample comparison (Figure 3B).

A



B



C

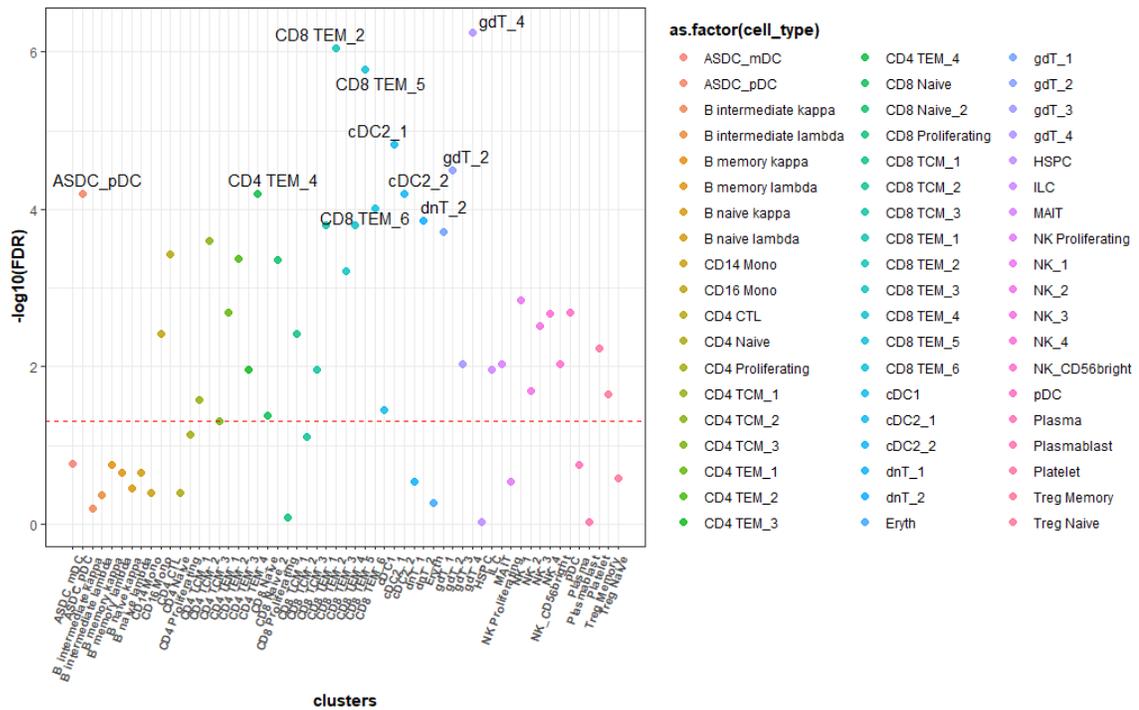


Fig 3. Immune molecular signatures of COVID-19 patients compared to MIS-C patients. **a-c** Plot of cell types that drives the top differential gene expression when using PBMC as marker genes in LRCeL analysis comparing COVID-19 and MIS-C patients, mild COVID-19 and mild MIS-C patients, and severe COVID-19 and severe MIS-C patients.

Discussion

The emerging of new SARS-CoV-2 virus has led to severe pulmonary disease and complications such as MIS-C, with significant morbidities and mortalities [20]. The pathophysiology of MIS-C is not well understood. It has been confirmed that the clinical features of MIS-C and severe COVID-19 overlap [9]. However, by applying bioinformatics approaches on blood transcriptomes of samples, strong evidence was provided for the existence of genetic risk factors behind the current clinical and

immunological performance of COVID-19 and MIS-C, which potentially helping differentiate MIS-C from severe COVID-19. Notedly, a number of myocarditis cases associated with severe SARS-CoV2 illness have been reported in children worldwide, which were recently detected as a typical clinical finding of MIS-C complication. Several results from GSEA and GOEA showed cardiomyopathy-associated characteristic in patients with severe COVID-19 and MIS-C, providing evidence of the association between MIS-C and SARS-CoV2 among children [21], supporting the clinical findings of more severe myocarditis in MIS-C patients [22], and revealing important biological basis for patient stratification and better selection of individual treatment. In addition, dramatic transcriptional changes were detected in neutrophils, inflammatory response, T cell differentiation, and platelet function, when comparing COVID-19 samples with healthy controls. Furthermore, a relevant target in SARS-CoV2, CD14, which was enhanced in COVID-19 versus MIS-C, was identified as an important initial point in host recognition of viral and host-derived products in the lungs [18], and has also been treated as a rational and feasible therapeutic approach to therapy for patients who are seriously ill with SARS-CoV2 infection. It was interested to know that an inhibitory monoclonal antibody against CD14 has been used in more than 165 human subjects without increasing secondary bacterial infections [18]. Moreover, differential activation of anti-viral CD8+ T cells response, which has been detected among COVID-19 and MIS-C patients, was thought to be critical for control of viral infections, but currently our knowledge of CD8+ T cell

response against SARS-CoV2 was limited [19]. A long-term follow up studies will be required to capture the activation of CD8+ T cells response during the SARS-CoV2 infections, and to clarify if it was contributed to the differences in clinical features among COVID-19 patients and MIS-C patients. Collectively, this study provided evidence of distinct expression landscapes of COVID-19 and MIS-C patients using transcriptome sequencing of peripheral blood, for depicting the immune response and identifying diagnostic biomarkers, therapeutic of disease and development of vaccines.

There were several factors that limited the interpretation of this study. First, not adequate numbers of samples were recruited in this study, especially for severe COVID-19 (three patients). Certainty of findings, which were detected between COVID-19 and MIS-C, is limited, need to be further validity by adequate number of samples. Furthermore, this study only focused on cell marker genes of PBMC which was acquired from sing-cell RNA-sequencing. The interpretation and conclusion of the analysis will be more systematic if other sets of tissues such as heart and lung can be involved into the analysis. Collectively, this study provided and visualized evidence for potentially immune response revealed in COVID-19 patients from that in MIS-C patients, which can be utilized in large cohorts for disease risk or outcome prediction in differentiation of severe MIS-C from COVID-19. These results included the immune molecular characters may

contribute to diagnostic biomarkers, for better informed the therapies and vaccines, better screening and diagnostics, and to treat patients with a personalized approach.

Table 3. Gene sets enriched in KEGG comparing COVID-19 and Healthy control

Pathway	padj	Enrichment Score	Size
KEGG_LYSOSOME	0.00835	0.53914	116
KEGG_COMPLEMENT_AND_COAGULATION_CASCADES	0.00835	0.58880	56
KEGG_SYSTEMIC_LUPUS_ERYTHEMATOSUS	0.00835	0.61650	116
KEGG_TOLL_LIKE_RECEPTOR_SIGNALING_PATHWAY	0.00943	0.47452	86
KEGG_FC_GAMMA_R_MEDIATED_PHAGOCYTOSIS	0.00943	0.48192	91
KEGG_PARKINSONS_DISEASE	0.01157	0.46417	104
KEGG_RIBOSOME	0.01253	-0.42121	86
KEGG_OXIDATIVE_PHOSPHORYLATION	0.01384	0.43773	105
KEGG_REGULATION_OF_ACTIN_CYTOSKELETON	0.01384	0.39522	191
KEGG_LEISHMANIA_INFECTION	0.01384	0.48232	71
KEGG_ALZHEIMERS_DISEASE	0.03005	0.40237	147
KEGG_FOCAL_ADHESION	0.03452	0.37830	183
KEGG_NOD_LIKE_RECEPTOR_SIGNALING_PATHWAY	0.03452	0.48111	59
KEGG_BASAL_CELL_CARCINOMA	0.04262	-0.44344	48
KEGG_LEUKOCYTE_TRANSENDOTHELIAL_MIGRATION	0.04474	0.422095	104

Positive and negative value of enrichment score indicates the genes in the signature database (KEGG) will be mostly at the top and bottom of our gene list, respectively. Size represents the total number of genes in the signature database (KEGG).

Table 4. Selected gene sets enriched in Reactome comparing COVID-19 and MIS-C

Pathway	padj	Enrichment Score	Size
REACTOME_CELL_CYCLE_CHECK POINTS	0.00045	-0.44741	284
REACTOME_NEUTROPHIL_DEGRA NULATION	0.00060	-0.40626	461
REACTOME_CELL_CYCLE_MITOTI C	0.00068	-0.40390	542
REACTOME_CELL_CYCLE	0.00079	-0.38915	650

Positive and negative value of enrichment score indicates the genes in the signature database (Reactome) will be mostly at the top and bottom of our gene list, respectively. Size represents the total number of genes in the signature database (Reactome).

Table 5. Selected gene sets enriched in Reactome comparing mild COVID-19 and mild MIS-C

Pathway	padj	Enrichment Score	Size
REACTOME_CELL_CYCLE	0.00024	-0.40440	645
REACTOME_EXTRACELLULAR_M ATRIX_ORGANIZATION	0.00045	-0.38749	252
REACTOME_INFLUENZA_INFECTI ON	0.00044	-0.42724	154
REACTOME_INFECTIOUS_DISEAS E	0.00098	-0.29680	671

Positive and negative value of enrichment score indicates the genes in the signature database (Reactome) will be mostly at the top and bottom of our gene list, respectively. Size represents the total number of genes in the signature database (Reactome).

Table 6. Selected gene sets enriched in KEGG comparing severe COVID-19 and severe MIS-C

Pathway	padj	Enrichment Score	Size
KEGG_HYPERTROPHIC_CARDIOMYOPATHY_HCM	0.00035	0.51186	72
KEGG_DILATED_CARDIOMYOPATHY	0.00036	0.50957	76

Positive and negative value of enrichment score indicates the genes in the signature database (KEGG) will be mostly at the top and bottom of our gene list, respectively. Size represents the total number of genes in the signature database (KEGG).

References

1. World Health Organization. (n.d.). *WHO Director-General's remarks at the media briefing on 2019-nCoV on 11 February 2020*. World Health Organization.
<https://www.who.int/director-general/speeches/detail/who-director-general-s-remarks-at-the-media-briefing-on-2019-ncov-on-11-february-2020>.
2. Bialek, S., Boundy, E., Bowen, V., Chow, N., Cohn, A., Dowling, N., ... Sauber-Schatz, E. (2020). Severe Outcomes Among Patients with Coronavirus Disease 2019 (COVID-19) — United States, February 12–March 16, 2020. *MMWR. Morbidity and Mortality Weekly Report*, 69(12), 343–346. <https://doi.org/10.15585/mmwr.mm6912e2>
3. Dieterich, D., & Rodriguez Rivas, C. (2020). Faculty Opinions recommendation of Virological assessment of hospitalized patients with COVID-2019. *Faculty Opinions – Post-Publication Peer Review of the Biomedical Literature*.
<https://doi.org/10.3410/f.737654646.793573074>
4. Kam, K.-qian, Yung, C. F., Cui, L., Tzer Pin Lin, R., Mak, T. M., Maiwald, M., ... Thoon, K. C. (2020). A Well Infant With Coronavirus Disease 2019 With High Viral Load. *Clinical Infectious Diseases*, 71(15), 847–849. <https://doi.org/10.1093/cid/ciaa201>
5. Zou, L., Ruan, F., Huang, M., Liang, L., Huang, H., Hong, Z., ... Wu, J. (2020). SARS-CoV-2 Viral Load in Upper Respiratory Specimens of Infected Patients. *New*

England Journal of Medicine, 382(12), 1177–1179.

<https://doi.org/10.1056/nejmc2001737>

6. Lai, C.-C., Liu, Y. H., Wang, C.-Y., Wang, Y.-H., Hsueh, S.-C., Yen, M.-Y., ...

Hsueh, P.-R. (2020). Asymptomatic carrier state, acute respiratory disease, and pneumonia due to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2): Facts and myths. *Journal of Microbiology, Immunology and Infection*, 53(3), 404–412.

<https://doi.org/10.1016/j.jmii.2020.02.012>

7. Wei, M., Yuan, J., Liu, Y., Fu, T., Yu, X., & Zhang, Z.-J. (2020). Novel Coronavirus Infection in Hospitalized Infants Under 1 Year of Age in China. *JAMA*, 323(13), 1313.

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

8. Riphagen, S., Gomez, X., Gonzalez-Martinez, C., Wilkinson, N., & Theocharis, P.

(2020). Hyperinflammatory shock in children during COVID-19 pandemic. *The Lancet*, 395(10237), 1607–1608. [https://doi.org/10.1016/s0140-6736\(20\)31094-1](https://doi.org/10.1016/s0140-6736(20)31094-1)

9. Mary Beth F Son, Kevin Friedman. COVID-19: Multisystem inflammatory syndrome in children (MIS-C) clinical features, evaluation, and diagnosis. *UpToDate*,

<https://www.uptodate.com/contents/covid-19-multisystem-inflammatory-syndrome-in-children-mis-c-clinical-features-evaluation-and-diagnosis>

10. Islam, A. B., & Khan, M. A.-A.-K. (2020). Lung transcriptome of a COVID-19 patient and systems biology predictions suggest impaired surfactant production which may be druggable by surfactant therapy. *Scientific Reports*, 10(1). <https://doi.org/10.1038/s41598->

020-76404-8

11. IT2122. G. (n.d.). https://www.gbiosciences.com/Protein-Research/Protein-Detection-Analysis-Systems/ELISA/Human_ELISA_Kit/IT2122.

12. Goud, P. T., Bai, D., & Abu-Soud, H. M. (2021). A Multiple-Hit Hypothesis Involving Reactive Oxygen Species and Myeloperoxidase Explains Clinical Deterioration and Fatality in COVID-19. *International Journal of Biological Sciences*, 17(1), 62–72.

<https://doi.org/10.7150/ijbs.51811>

13. U.S. National Library of Medicine. (n.d.). *AXL AXL receptor tyrosine kinase [Homo sapiens (human)] - Gene - NCBI*. National Center for Biotechnology Information.

<https://www.ncbi.nlm.nih.gov/gene/558>.

14. Heger, L., Balk, S., Lühr, J. J., Heidkamp, G. F., Lehmann, C. H. K., Hatscher, L., ...

Dudziak, D. (2018, March 26). *CLEC10A Is a Specific Marker for Human CD1c+ Dendritic Cells and Enhances Their Toll-Like Receptor 7/8-Induced Cytokine Secretion*.

Frontiers. <https://www.frontiersin.org/articles/10.3389/fimmu.2018.00744/full>.

15. Aschenbrenner, A.C., Mouktaroudi, M., Krämer, B. *et al.* Disease severity-specific neutrophil signatures in blood transcriptomes stratify COVID-19 patients. *Genome*

Med **13**, 7 (2021). <https://doi.org/10.1186/s13073-020-00823-5>

16. Samji, T., & Khanna, K. M. (2017). Understanding memory CD8⁺T

cells. *Immunology letters*, 185, 32–39. <https://doi.org/10.1016/j.imlet.2017.02.012>

17. Simeonidis, S., Castagliuolo, I., Pan, A., Liu, J., Wang, C.-C., Mykoniatis, A., ... Pothoulakis, C. (2003). Regulation of the NK-1 receptor gene expression in human macrophage cells via an NF- κ B site on its promoter. *Proceedings of the National Academy of Sciences*, *100*(5), 2957–2962. <https://doi.org/10.1073/pnas.0530112100>
18. Zanoni, I., Tan, Y., Di Gioia, M., Springstead, J. R., & Kagan, J. C. (2017). By Capturing Inflammatory Lipids Released from Dying Cells, the Receptor CD14 Induces Inflammasome-Dependent Phagocyte Hyperactivation. *Immunity*, *47*(4). <https://doi.org/10.1016/j.immuni.2017.09.010>
19. Petitdemange, C., Kasturi, S. P., Kozlowski, P. A., Nabi, R., Quarnstrom, C. F., Reddy, P. B., ... Pulendran, B. (2019). Vaccine induction of antibodies and tissue-resident CD8+ T cells enhances protection against mucosal SHIV-infection in young macaques. *JCI Insight*, *4*(4). <https://doi.org/10.1172/jci.insight.126047>
20. Zhu L;Yang P;Zhao Y;Zhuang Z;Wang Z;Song R;Zhang J;Liu C;Gao Q;Xu Q;Wei X;Sun HX;Ye B;Wu Y;Zhang N;Lei G;Yu L;Yan J;Diao G;Meng F;Bai C;Mao P;Yu Y;Wang M;Yuan Y;Deng Q;Li Z;Huang Y;Hu G;Liu Y;Wang X;Xu Z;Liu P;Bi Y;Shi Y;Zhang S;Chen Z;Wang J;Xu X;Wu G;Wa. (n.d.). Single-Cell Sequencing of Peripheral Mononuclear Cells Reveals Distinct Immune Response Landscapes of COVID-19 and Influenza Patients. *Immunity*. <https://pubmed.ncbi.nlm.nih.gov/32783921/>.
21. Carter, M. J., Fish, M., Jennings, A., Doores, K. J., Wellman, P., Seow, J., ... Shankar-Hari, M. (2020). Peripheral immunophenotypes in children with multisystem

inflammatory syndrome associated with SARS-CoV-2 infection. *Nature*

Medicine, 26(11), 1701–1707. <https://doi.org/10.1038/s41591-020-1054-6>

22. Lee PY;Day-Lewis M;Henderson LA;Friedman KG;Lo J;Roberts JE;Lo MS;Platt CD;Chou J;Hoyt KJ;Baker AL;Banzon TM;Chang MH;Cohen E;de Ferranti SD;Dionne A;Habiballah S;Halyabar O;Hausmann JS;Hazen MM;Janssen E;Meidan E;Nelson RW;Nguyen AA;Sundel RP;Dedeoglu F; (n.d.). Distinct clinical and immunological features of SARS-CoV-2-induced multisystem inflammatory syndrome in children. *The Journal of clinical investigation*. <https://pubmed.ncbi.nlm.nih.gov/32701511/>.

23. Vella, L. A., Giles, J. R., Baxter, A. E., Oldridge, D. A., Diorio, C., Kuri-Cervantes, L., ... Wherry, E. J. (2020). Deep Immune Profiling of MIS-C demonstrates marked but transient immune activation compared to adult and pediatric COVID-19.

<https://doi.org/10.1101/2020.09.25.20201863>

24. Marvinquiet. (n.d.). *marvinquiet/LRcell*. GitHub.

<https://github.com/marvinquiet/LRcell/>.