

Nucleic acid biomarkers of immune response and cell and tissue damage in children with COVID-19 and MIS-C

Authors: Conor J. Loy^{1#}, Alicia Sotomayor-Gonzalez^{2#}, Venice Servellita^{2#}, Jenny Nguyen², Joan Lenz¹, Sanchita Bhattacharya⁴, Meagan E. Williams⁵, Alexandre P. Cheng¹, Andrew Bliss¹, Prachi Saldhi², Noah Brazer², Jessica Streithorst², William Suslovic⁵, Charlotte Hsieh³, Burak Bahar⁵, Nathan Wood³, Abiodun Foresythe², Amelia Gliwa², Kushmita Bhakta⁶, Maria A. Perez⁶, Laila Hussaini⁶, Evan J. Anderson^{6,7}, Ann Chahroudi⁶, Meghan Delaney⁵, Atul J. Butte⁴, Roberta DeBiasi⁵, Christina A. Rostad⁶, Iwijn De Vlaminck^{1#*}, Charles Y. Chiu^{2#*}

Affiliations:

¹Meinig School of Biomedical Engineering, Cornell University, Ithaca, NY, USA

²Department of Laboratory Medicine, University of California, San Francisco, San Francisco, CA, USA

³UCSF Benioff Children's Hospital, Oakland, CA, USA

⁴Bakar Computational Health Sciences Institute, University of California, San Francisco, San Francisco, CA, USA

⁵Children's National Hospital, Washington D.C., USA

⁶Department of Pediatrics and Center for Childhood Infections at Vaccines, Emory University School of Medicine and Children's Healthcare of Atlanta, Atlanta, GA, USA

⁷Department of Medicine, Emory University School of Medicine, Atlanta, GA, US

⁸Department of Medicine, Division of Infectious Diseases, University of California, San Francisco, San Francisco, CA, USA

#These authors contributed equally

*Correspondence: charles.chiu@ucsf.edu, id93@cornell.edu

Key words: SARS-CoV-2, coronavirus disease 2019, cell-free DNA, cell-free RNA, whole blood RNA, RNA sequencing, RNA-Seq, clinical severity, multisystem inflammatory syndrome in children, Nucleic acid sequencing, DNA damage, cell damage, tissue damage, host response, signaling pathways, disease biomarkers, immune response, systems biology, bisulfite sequencing, pediatric

Summary

Differential host responses in coronavirus disease 2019 (COVID-19) and multisystem inflammatory syndrome in children (MIS-C) remain poorly characterized. Here we use next-generation sequencing to longitudinally analyze blood samples from pediatric patients with acute COVID-19 (n=70) or MIS-C (n=141) across three hospitals. Profiling of plasma cell-free nucleic acids uncovers distinct signatures of cell injury and death between these two disease states, with increased heterogeneity and multi-organ involvement in MIS-C encompassing diverse cell types such as endothelial and neuronal Schwann cells. Whole blood RNA profiling reveals upregulation of similar pro-inflammatory signaling pathways in COVID-19 and MIS-C, but also MIS-C specific downregulation of T cell-associated pathways. Profiling of plasma cell-free RNA and whole blood RNA in paired samples yields different yet complementary signatures for each disease state. Our work provides a systems-level, multi-analyte view of immune responses and tissue damage in COVID-19 and MIS-C and informs the future development of new disease biomarkers.

INTRODUCTION

At the onset of the coronavirus disease 2019 (COVID-19) pandemic, SARS-CoV-2 was thought to only cause mild or asymptomatic infection in children. Large-scale surveillance studies have since demonstrated

cases of severe COVID-19-associated pneumonia occurring throughout the pandemic, especially in children with underlying comorbidities^{1,2}. To date, mortality from COVID-19 in children has exceeded the pediatric mortality from influenza in any given season^{3,4}. Furthermore, children are at risk for the post-infectious multisystem inflammatory syndrome in children (MIS-C) associated with COVID-19, which manifests as severe disease with systemic hyperinflammation and multiorgan involvement⁵. The diagnosis of MIS-C currently relies on clinical symptoms, and the pathogenesis of MIS-C is incompletely understood. Of note, cases of MIS-C were initially misdiagnosed as Kawasaki disease (KD)⁶, another systemic inflammatory syndrome, because of overlapping clinical features including fever, hyperinflammation, mucocutaneous involvement, and vascular endothelial dysfunction. However, subsequent studies have demonstrated distinct clinical⁷ and immunological⁸ phenotypes associated with MIS-C as compared to KD. Thus, a better understanding of the pathogenesis of MIS-C is critical to improve clinical diagnosis and inform targeted interventions, particularly as new variants of SARS-CoV-2 continue to emerge.

Initial characterization by proteomics and RNA sequencing revealed that MIS-C has an inflammatory profile similar to KD and severe COVID-19, with key differences, including specific increases in IFN γ , IL-6, IL-17, IL-10, alarmin-related proteins, and other proinflammatory factors⁹⁻¹¹. Autoantibody profiling revealed a unique autoantigen profile targeting organs often injured during MIS-C, but the damage to the organs and tissues has not yet been quantified on a system-wide level¹². Immune cell profiling showed that MIS-C is associated with an expansion of specific subsets of NK cells, T cells, and B cells^{10,13}, and gene expression profiling showed that a high fraction (~24%) of T cells in patients with MIS-C are T cell receptor beta variable 11-2 (TRBV11-2) positive¹³⁻¹⁵. T cell receptor repertoire analyses have led to the hypothesis that the SARS-CoV-2 spike protein has a superantigen effect, causing T cell dysregulation that contributes to the development of MIS-C¹⁵. Finally, flow cytometry analyses have demonstrated that children with MIS-C have reduced virus-specific CD4+ and CD8+ T cells compared to children with COVID-19 and controls¹⁶.

Despite these prior studies, much remains unclear about the pathogenesis of MIS-C, and there is a lack of biomarkers that could be leveraged to develop diagnostic and prognostic assays. Here we performed unbiased profiling of whole blood cellular RNA (wbRNA), plasma cell-free RNA (cfRNA), and plasma cell-free DNA (cfDNA) from 211 pediatric patients diagnosed with COVID-19 or MIS-C and 26 controls across 3 separate pediatric hospital systems in the United States. Longitudinal sampling of these 3 blood analytes enable a complementary and systems-level view of immune responses and cell/tissue damage associated with MIS-C and COVID-19.

RESULTS

Clinical COVID-19 and/or MIS-C cohort

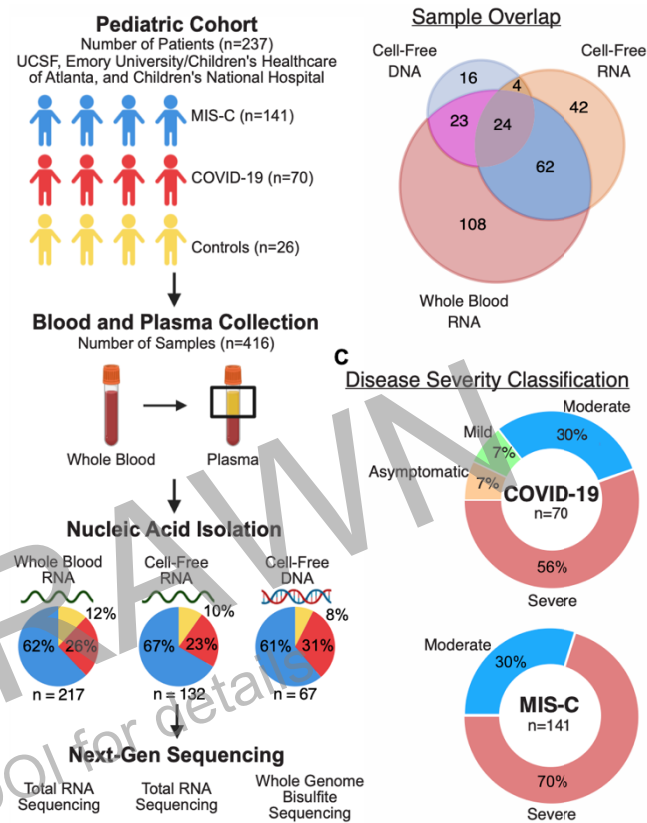


Figure 1. Study design and patient characteristics. (A) Sample collection and processing overview. (B) Distribution of samples across analytes. (C) Distribution of disease severity for each sample group.

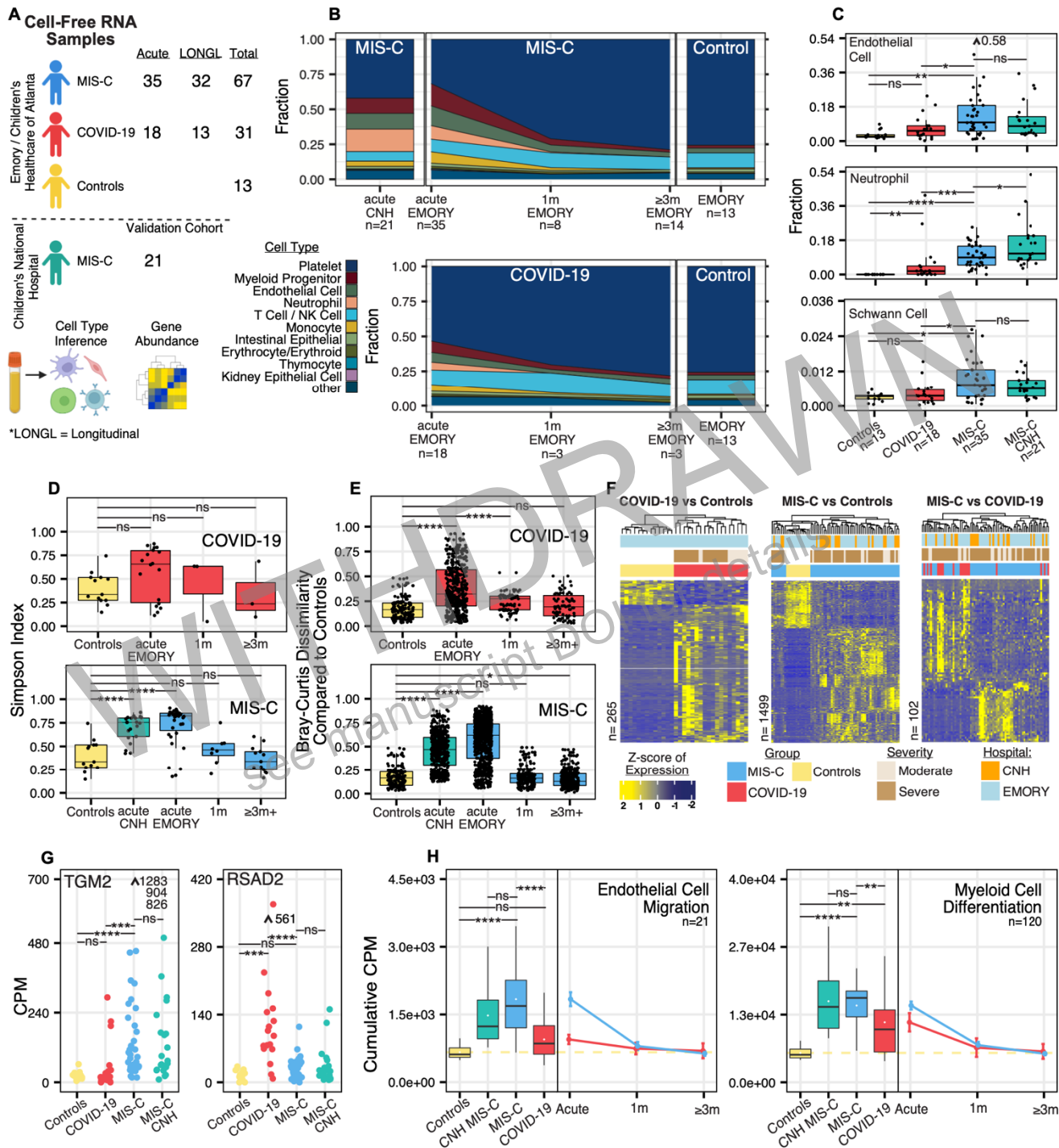
Variables	Group	Overall	COVID-19	MIS-C	Controls
n		237	70	141	26
origin, n (%)	CNH	35 (14.8)		35 (24.8)	
	EMORY	176 (74.3)	49 (70.0)	101 (71.6)	26 (100.0)
	UCSF	26 (11.0)	21 (30.0)	5 (3.5)	
Disease Severity, n (%)	Asymptomatic	5 (2.1)	5 (7.1)		
	Mild	5 (2.1)	5 (7.1)		
	Moderate	63 (26.6)	21 (30.0)	42 (29.8)	
	Severe	138 (58.2)	39 (55.7)	99 (70.2)	
Age, mean (SD)		11 (5)	12 (6)	9 (5)	14 (2)
Gender, n (%)	Female	103 (43.5)	37 (52.9)	49 (34.8)	17 (65.4)
	Male	134 (56.5)	33 (47.1)	92 (65.2)	9 (34.6)
Race, n (%)	American Indian	2 (0.8)			2 (7.7)
	Asian	12 (5.1)	4 (5.7)	2 (1.4)	6 (23.1)
	Black/AA	120 (50.6)	28 (40.0)	86 (61.0)	6 (23.1)
	White	58 (24.5)	17 (24.3)	31 (22.0)	10 (38.5)
	Other/Declined	45 (19.0)	21 (30.0)	22 (15.6)	2 (7.7)
Ethnicity, n (%)	Hispanic	57 (24.1)	23 (32.9)	29 (20.6)	5 (19.2)
	Non-Hispanic	180 (75.9)	47 (67.1)	112 (79.4)	21 (80.8)

Table. Demographic and clinical characteristics of the study cohort

We collected 416 blood and/or plasma samples from 237 patients from the University of California at San Francisco (UCSF), Emory University/Children's Healthcare of Atlanta (EMORY), and Children's National Hospital (CNH) (Fig. 1A and B, Table; Methods). Patients from Emory were prospectively enrolled into a specimen collection and biobanking protocol, permitting the acquisition of longitudinal samples during hospitalization, one month post-hospitalization, and ≥ 3 months post-hospitalization. Remnant clinical samples from patients at the other sites were biobanked and analyzed under approved institutional review board (IRB) protocols with waiver of consent. All samples were stratified by diagnosis, time of collection, and severity of disease (Fig. 1C). Patients were either diagnosed with COVID-19 (without MIS-C) or MIS-C, or were uninfected control subjects. The control subjects were healthy outpatient children prospectively enrolled at Emory. Hospitalization time points were stratified as acute (0-4 days after hospital admission) or post-acute (>4 days after hospital admission), and patients were classified according to clinical severity at time of presentation as having asymptomatic, mild, moderate, or severe disease (Supplementary Table). cfRNA and cfDNA profiling by next-generation sequencing (NGS) were performed from plasma, and transcriptome RNA profiling (RNA-Seq) was performed from whole blood.

Circulating cell-free RNA profiling

We performed transcriptome sequencing of plasma cfRNA on 132 samples from 124 pediatric patients. Of the 132 samples, 88 (67%) were classified as MIS-C, 31 (23%) as moderate-to-severe COVID-19 and 13 (10%) as negative controls (Fig. 2A). Recent work by Vorperian et al. demonstrated the possibility to quantify cell-types-of-origin (CTO) of cfRNA using reference-based deconvolution¹⁷. Here, we implemented BayesPrism and the "Tabula Sapiens" human single-cell transcriptome atlas as a reference to quantify the cfRNA CTO^{18,19}. We consistently observed significant increases in cfRNA from neutrophils, kidney epithelial cells, thymocytes, and solid organ-derived cell types in moderate-to-severe COVID-19 as compared to control individuals, and further increased contributions from these cell types in MIS-C (p-value<0.05 by Mann-Whitney U test) (Fig. 2B and C; Supplemental Fig. 1A). We also observed significant increases in cfRNA from endothelial cells and



neuronal Schwann specifically in children with MIS-C (p -value <0.05 by Mann-Whitney U test) (Fig. 2B and C). Platelet cfRNA was decreased in MIS-C and moderate-to-severe COVID-19 compared to controls. The CTO of cfRNA were highly similar for samples collected at different hospitals, highlighting the potential for cfRNA as a diagnostic biomarker for COVID-19 or MIS-C.

Figure 2. Plasma cell-free RNA profiling. (A) Study design and analysis overview. (B) Average cfRNA deconvolution results for COVID-19, MIS-C, and controls over various time points. (C) Cell-free RNA deconvolution results of endothelial cell, neutrophil, and Schwann cell derived cfRNA (D) Diversity of cell type contributions to the cell-free transcriptome as measured by Simpson's Index. (E) Dissimilarity of samples as compared to controls. Each point represents a comparison to a control, as measured with the Bray-Curtis dissimilarity measurement. (F) Scaled counts per million (CPM) values of significantly differentially abundant genes (DAGs) (DESeq2, Benjamini-Hochberg adjusted p -value < 0.01 , $|\text{Log}_2\text{FoldChange}| > 1.5$). Number of DAGs indicated to the left of the heatmap. Samples and genes are clustered based on correlation. (G) Normalized CPM values of TGM2 and antiviral gene RSAD2 across sample groups at the acute

timepoint. **(H)** Cumulative CPM of genes in significant gene ontology groups (topGO, adjusted p-value < 0.05). Panels A and B show the number of samples in each group. Box plots show cumulative CPM distribution of controls and acute timepoint MIS-C and moderate-to-severe COVID-19. Points represent average cumulative CPM and bars represent standard error. Outliers are indicated with arrows and values. Asterisks indicate statistical significance by Mann-Whitney U test using Benjamini-Hochberg adjusted p-values as follows: ns, non-significant; *, p < 0.05; **, p < 0.01; ***, p < 0.001; ****, p < 0.001.

Next, we sought to investigate longitudinal changes in cfRNA CTO in MIS-C. We observed significantly greater CTO diversity using Simpson's index²⁰ in acute MIS-C versus convalescent MIS-C (≥ 1 month post-hospitalization) or controls (p-value < 0.05 by Mann-Whitney U test) (**Fig. 2D**). In contrast, cfRNA CTO diversity was not significantly different for acute moderate-to-severe COVID-19 compared to controls or convalescent COVID-19 (p-value > 0.05 by Mann-Whitney U test) (**Fig. 2D**). Samples from patients with acute moderate-to-severe COVID-19 separated into 2 groups, one group with high and one with low cfRNA CTO diversity, consistent with the heterogeneity of cell types involved in COVID-19 as previously described²¹. We further compared the cfRNA CTO diversity of post-acute MIS-C and COVID-19. We found that samples from post-acute MIS-C patients had high cfRNA CTO diversity while samples from post-acute moderate-to-severe COVID-19 patients had low cfRNA CTO diversity (**Supplemental Fig. 1B and C**). These trends were also observed using the Shannon diversity index²². Next, we analyzed pairwise dissimilarity in cfRNA CTO between COVID-19 and MIS-C samples and controls. The Bray-Curtis dissimilarity of cfRNA CTO within controls was significantly lower than between controls and acute MIS-C (mean dissimilarity = 0.18 versus 0.56, p-value < 0.05 by Mann-Whitney U test) and between controls and acute moderate-to-severe COVID-19 (mean dissimilarity = 0.18 versus 0.38, p-value < 0.05 by Mann-Whitney U test) (**Fig. 2E**). Among convalescent COVID-19 or MIS-C samples, only the COVID-19 one-month follow up cohort exhibited significant divergence of cfRNA CTO compared to controls (**Fig. 2E**). Patients from 2 different hospitals (Emory and CNH) exhibited similar diversity and pairwise dissimilarity metrics (**Fig. 2D and E**). These findings revealed that acute MIS-C exhibited a higher diversity of cell types than either COVID-19 or controls, and that the observed cell types based on cfRNA gene expression were significantly different between MIS-C and COVID-19.

We identified differentially abundant genes (DAGs) in cfRNA associated with dead or dying cells to gain insight into disease pathogenesis and to characterize potential diagnostic biomarkers²³. Patients diagnosed with acute MIS-C or moderate-to-severe COVID-19 were compared pairwise to controls. Using an absolute log₂ fold change cutoff of 1.5 and a Benjamini-Hochberg adjusted p-value of < 0.01, 1,409 DAGs were identified between MIS-C and controls, 265 DAGs between COVID-19 and controls, and 102 DAGs between MIS-C and COVID-19. Unsupervised clustering revealed distinct gene expression profiles separating MIS-C and COVID-19 from controls (**Fig. 2F**). Samples from acute MIS-C patients were assigned into three groups based on unsupervised clustering, each group with a distinct CTO profile (**Supplemental Fig. 1D**). The three groups consisted of cfRNA predominantly derived from (1) endothelial cells, NK cells, and respiratory ciliated cells, (2) monocytes, neutrophils, and myeloid progenitors, and (3) platelets. The third group clustered with samples from controls and COVID-19 patients and may hence be a technical artifact associated with elevated platelet lysis during sample preparation.

We analyzed significant DAGs between acute MIS-C and moderate-to-severe COVID-19 to obtain insights into disease pathogenesis and differential immune responses associated with these 2 diseases. Acute MIS-C was associated with elevated levels of endothelial cell markers (AKAP12, CNN3, FZD4), neuronal markers (GAS7, FEZ1, VAT1), actin-related genes (FSCN1, AFAP1L1, ITGA9), and an autoantigen also found in patients with celiac disease (TGM2) (**Fig. 2G, Supplemental Fig. 1E**). In contrast, acute COVID-19 was associated with elevated levels of interferon genes (IFI6, IFIT1, IFI44L, IFI27, IFITM1), antiviral genes (RSAD2, MX1, CMP2, LY6E), chemokine genes (CXCL5, CXCL3), and ciliated olfactory cell markers (OR2B6, ENKUR) (**Fig. 2G, Supplemental Fig. 1E**). Next, we performed gene ontology analysis using the R package topGO²⁴. Gene ontology terms enriched in samples from COVID-19 patients included those associated with programmed cell death, response to viral infection, and regulation of the viral life cycle, while those enriched in MIS-C patients included actin cytoskeleton organization, endothelial cell migration, cytokine responses, and

cell migration. To identify disease-specific pathways, we calculated counts per million (CPM) for each gene ontology group by summing the counts of DAGs in that group²⁵. Compared to controls, the cumulative CPM score for endothelial cell migration was significantly increased in acute MIS-C (p-value<0.05 by Mann-Whitney U test), while cumulative CPM scores for myeloid cell differentiation were increased in acute MIS-C and moderate-to-severe COVID-19 (p-value<0.05 by Mann-Whitney U test) (**Fig. 2H**). Control samples were associated with increased gene ontology groups related to cell division and cell communication, consistent with the observation that the baseline cfRNA signal is predominantly derived from extracellular vesicles and cells undergoing apoptosis, as typically occurs during mitosis²⁶. Next, we performed pathway analysis using Ingenuity Pathway Analysis (IPA). Pathways enriched in MIS-C included pyroptosis, synaptogenesis, NF- κ B signaling, IL-6 signaling, IL-8 signaling, antiviral responses including interferon induction, and cholesterol biosynthesis, while a different set of pathways enriched in COVID-19 included macrophage production of nitric oxide, coronavirus pathogenesis, LXR/RXR activation, GP6 signaling, PTEN signaling, summylation, gustation (taste), and HMGB1 signaling (**Supplemental Fig. 1F and G**).

Whole blood RNA profiling by transcriptome sequencing (RNA-Seq)

Whole blood transcriptome profiling was performed on 217 samples from 187 pediatric patients. Of the 217 whole blood samples, 135 (62%) were classified as MIS-C, 56 (26%) as moderate-to-severe COVID-19 and 26 (12%) as negative controls (**Fig 3A**). Samples from Emory and UCSF were used for the differential expression analysis (DEA) and no batch effect was detected via unsupervised clustering (**Supplemental Fig. 2A**). Samples from CNH were used as a validation group (**Fig. 3B-C**). Pairwise comparisons of MIS-C and severe COVID-19 relative to controls showed a large degree of overlap in shared differentially expressed genes (DEGs) between MIS-C and COVID-19 (DESeq2, Benjamini-Hochberg corrected p-value < 0.01, $|\text{Log}_2\text{FoldChange}| > 1.5$) (**Fig. 3B-C, Supplemental Fig. 2A**). The top 2 shared DEGs in both diseases were ADAMTS2, a metalloprotease that processes procollagen (**Fig. 3D and Supplemental Fig. 2B**), and CD177, a neutrophil activator (**Supplemental Fig. 2B and C**). Notably, ADAMTS2 has been previously implicated in severe COVID-19 in a study using peripheral blood mononuclear cell (PBMC) single-cell transcriptome sequencing²⁷, whereas CD177 has been reported to be upregulated in the blood of MIS-C and COVID-19 patients²⁸. We also observed elevated levels of ADAMTS2 during the post-acute stage of MIS-C and COVID-19; however, one month after hospitalization levels returned to baseline in children with MIS-C but were still elevated in children with COVID-19 (**Supplemental Fig. 2D**). Certain inflammatory genes such as interferon-stimulated gene 15 (ISG15) and macrophage-associated sialic acid binding Ig-like lectin 1 (SIGLEC1) were significantly upregulated in COVID-19 but not in MIS-C versus controls (**Fig. 3D and Supplemental Fig. 2B**). In contrast, T-cell receptor beta variable 11-2 (TRBV11-2) was more highly expressed in MIS-C than in COVID-19 versus controls (**Fig. 3D**), a finding that was also seen in a direct head-by-head comparison between MIS-C and COVID-19 (**Supplemental Fig. 2D**). This observation is consistent with two studies showing that TRBV11-2 is overexpressed by T-cells in most MIS-C patients, but not in patients with COVID-19, Kawasaki disease (KD), or toxic shock syndrome (TSS)^{14,29}. These differences in TRBV11-2 expression in MIS-C versus COVID-19 or controls were observed at the post-acute timepoint, but not at the 1 month or ≥ 3 months timepoints (**Supplemental Fig 2D**). Interestingly, expression of gene paralogs KLRF1 and KLRB1, natural killer (NK) cell surface receptors, were found to be significantly decreased in COVID-19 and MIS-C compared to controls (**Fig. 3D, Supplemental Fig. 2C**), consistent with lower expression of these genes reported in severe COVID-19 versus controls³⁰. Decreased KLRF1 and KLRB1 expression in COVID-19 and MIS-C was also observed at the post-acute timepoint, but at 1 month post-hospitalization returned to baseline in MIS-C yet remained decreased in COVID-19 (**Supplemental Fig. 2D**). In the head-to-head comparison between MIS-C and COVID-19, the top upregulated genes in COVID-19 were most likely to be those related to the antiviral type 1 interferon response pathway (e.g., IFIT2, SIGLEC1, IFI27, IFI44L, ISG15, IFIT3) (**Supplemental Fig. 2E; Supplemental Dataset 1**). The top upregulated genes in MIS-C were related to

multiple pathways, likely due to the heterogeneous nature of MIS-C, including those associated with cell-to-cell communication (e.g., ITGA7, CDHR1, CD177, PGF, ERFE, MMP8) (**Supplemental Dataset 1**).

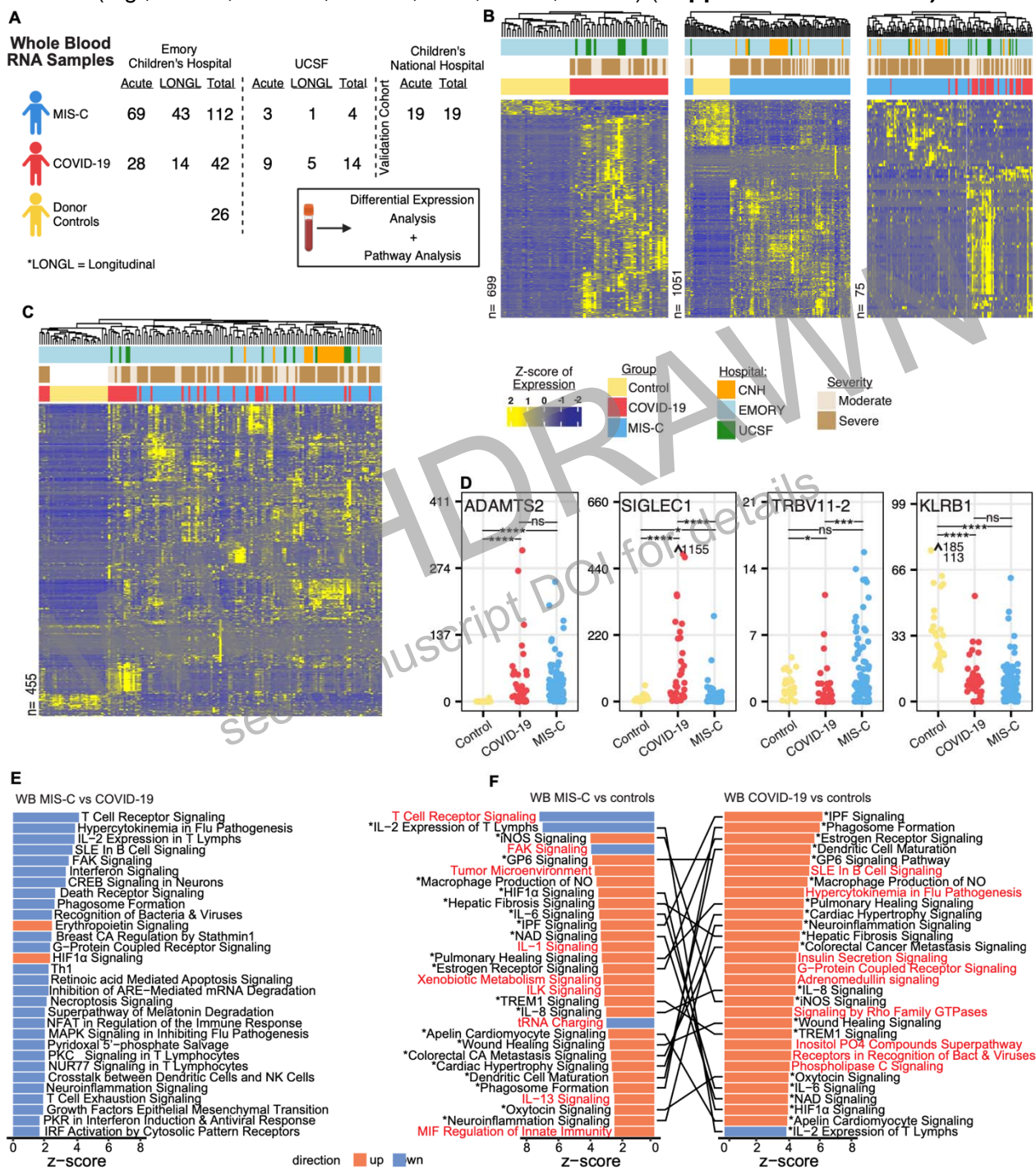


Figure 3. Whole blood RNA profiling. (A) Study design and analysis overview. (B) Scaled CPM values of significantly differentially expressed genes (DESeq2, Benjamini-Hochberg corrected p-value < 0.01, |Log2FoldChange| > 1.5). Differentially expressed genes were discovered using Emory and UCSF samples and the total number is indicated to the left of each heatmap. Samples and genes are clustered based on correlation. (C) Scaled CPM values of the top 300 significantly expressed genes from each comparison (union of genes, DESeq2, Benjamini-Hochberg adjusted p-value < 0.01, ranked by absolute Log2FoldChange). Samples and genes are clustered based on correlation. (D) CPM of ADAMTS2, TRBV11-2, SIGLEC1, and KLRB1 in controls, acute MIS-C, and acute moderate-to-severe COVID-19. Points represent average CPM and bars represent standard error. Asterisks indicate statistical significance by Mann-Whitney U test using Benjamini-Hochberg adjusted p-values as follows: ns, non-significant; *, p < 0.05; **, p < 0.01; ***, p < 0.001; ****, p < 0.001. Panels 3A show the number of samples in each group. (E) Top 20 differential pathways between controls

11 and MIS-C or COVID-19 ranked by activation z-score. Lines connect matching pathways. Pathways highlighted in red are
12 distinct to either MIS-C or COVID-19. **(F)** Top 20 differential pathways between MIS-C and COVID-19 ranked by activation
13 z-score.

14 Next, we compared differentially expressed pathways in MIS-C and COVID-19 relative to controls **(Fig.**
15 **3E)** and to each other **(Fig. 3F)**. In both MIS-C and COVID-19 patients, we found increased activation of
16 immune-related pathways which include phagosome formation, macrophage production of nitric oxide,
17 dendritic cell

18 maturation, TREM1, and IL-6 and IL-8 signaling. Pathways associated with hypoxia (hypoxia-inducible factor 1-
19 alpha, or HIF1-alpha signaling), neuroinflammation, and cardiac hypertrophy were also upregulated in both
20 MIS-C and COVID-19. MIS-C patients showed a marked inhibition of T cell receptor, interleukin-2 (IL-2), and
21 focal adhesion kinase (FAK) signaling pathways, concurrent with an activation of tumor environment, IL-1, and
22 IL-13 signaling pathways. In contrast, COVID-19 patients showed a more pronounced activation of cytokine, B-
23 cell, and adrenomedullin pathways **(Fig. 3F)**. We analyzed differentially expressed pathways and their
24 associations with diseases or biological functions and observed striking differences between MIS-C and
25 COVID-19 compared to controls **(Supplemental Fig. 3)**. MIS-C was characterized by downregulation of
26 multiple pathways related to inflammatory response, cell death and survival, cell-to-cell signaling, and immune
27 cell trafficking, whereas activation of these same pathways was predicted in COVID-19. These results are
28 consistent with the downregulation of exhausted T cells that has been previously reported in children with MIS-
29 C³¹.

31 **Cell-free DNA tissues-of-origin by methylation profiling**

32 We performed whole genome bisulfite sequencing of cfDNA extracted from plasma samples from 67
33 children with MIS-C (n=41), COVID-19 (n=21), or from controls (n=5) **(Fig. 4A)** and compared the cfDNA data
34 to previously published data from an adult COVID-19 cohort³². The highest mean levels of total cfDNA were
35 found in children with MIS-C as compared to pediatric COVID-19, adult COVID-19, or controls (4.12 ng/uL, p-
36 value<0.03, by Mann-Whitney U test) **(Fig. 4B)**. A subset of MIS-C patients (n=3) showed a markedly elevated
37 burden of cfDNA (>10 ng/ul), likely secondary to widespread tissue injury. There were also significantly higher
38 mean levels of cfDNA in more severe COVID-19 pediatric and adult cases as compared to mild or
39 asymptomatic COVID-19 cases and controls (p-value<0.05 by Mann-Whitney U test) **(Fig. 4B)**.

40 We next examined the utility of cfDNA tissues-of-origin (TOO) profiling to identify tissue profiles by
41 comparison to a reference set of methylation profiles of purified cells and tissue samples, as previously
42 described³². We observed significantly elevated levels of solid-organ-derived cfDNA in pediatric acute MIS-C
43 and pediatric moderate-to-severe COVID-19 as compared to pediatric controls, pediatric acute mild or
44 asymptomatic COVID-19, adult controls, and adult mild-to-severe COVID-19 cases (p-value<0.03, by Mann-
45 Whitney U test) **(Fig. 4C)**. Levels of solid-organ-derived cfDNA in acute MIS-C tended to be increased relative
46 to acute moderate-to-severe COVID-19, although this difference was not significant (p-value=0.12, by Mann-
47 Whitney U test) **(Fig. 4C)**. We further observed elevated levels of cfDNA derived from innate immune cell types
48 in acute MIS-C and moderate-to-severe pediatric COVID-19 compared to all other groups; however, this
49 difference was not significant **(Fig. 4C)**. In addition, we identified extensive heterogeneity in the TOO profiles
50 from patients in the moderate-to-severe MIS-C cohort, including elevated levels of eosinophil, neutrophil,
51 erythroblast, liver, heart, kidney, lung, and spleen-derived cfDNA **(Supplemental Fig. 4)**. Although recent
52 studies have reported increases in mitochondrial cfDNA in plasma from patients diagnosed with COVID-19³³,
53 here we observed a significant reduction in the concentration of mitochondrial cfDNA in pediatric patients with
54 MIS-C and COVID-19 (p-value<0.04, by Mann-Whitney U test) **(Fig. 4D)**. Finally, we found that cfDNA metrics
55 associated with affected tissues and organs mirrored organ-specific clinical laboratory parameters. Significant
56 correlations between kidney cfDNA concentration and creatinine levels (Pearson correlation; R=0.40, p-
57 value=0.004), total cfDNA concentration and C-reactive protein (CRP) levels in the blood (Pearson correlation;

R=0.34, p-value=0.022), and liver cfDNA concentration and ALT levels (Pearson correlation; R=0.303, p-value=0.036) were observed (Fig. 4E).

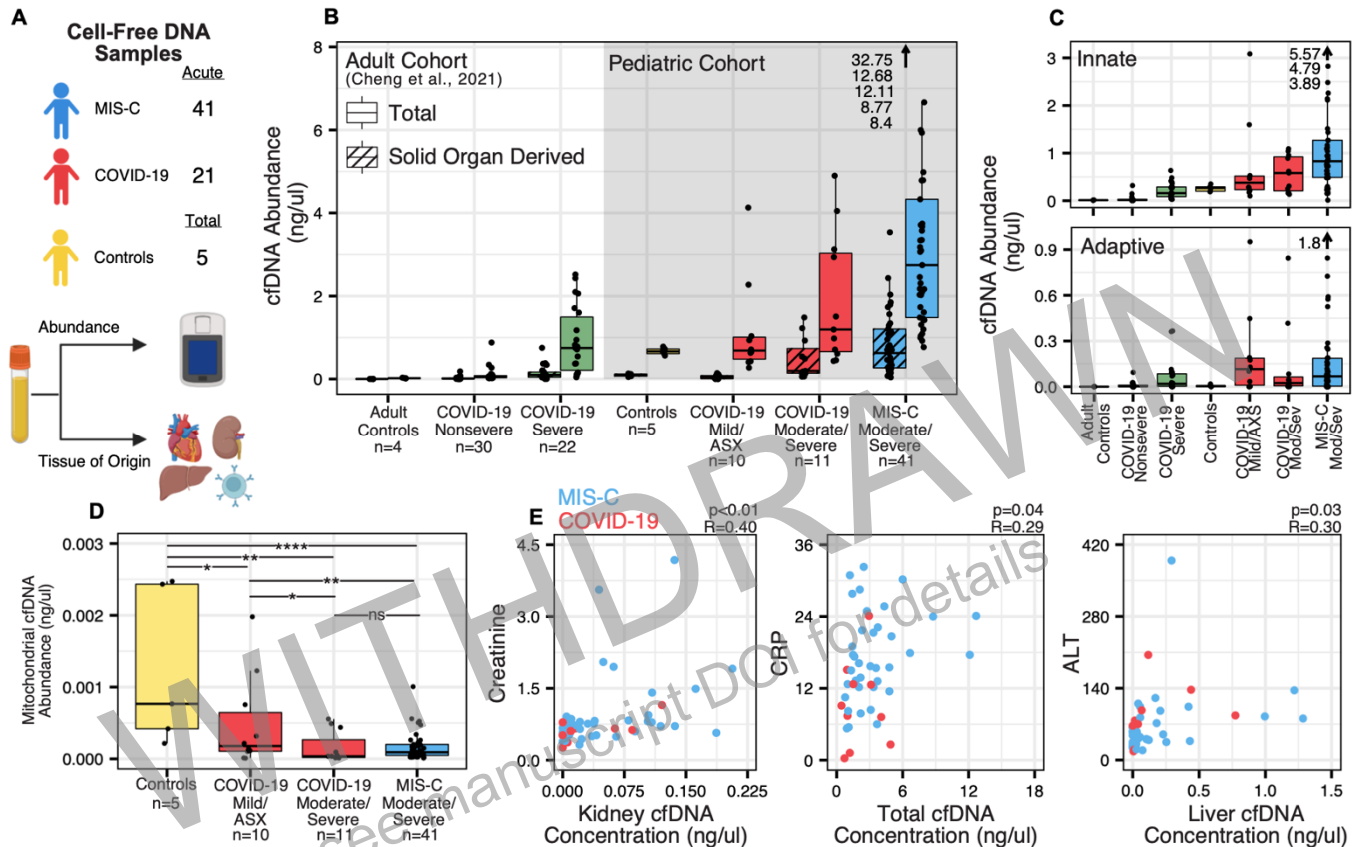


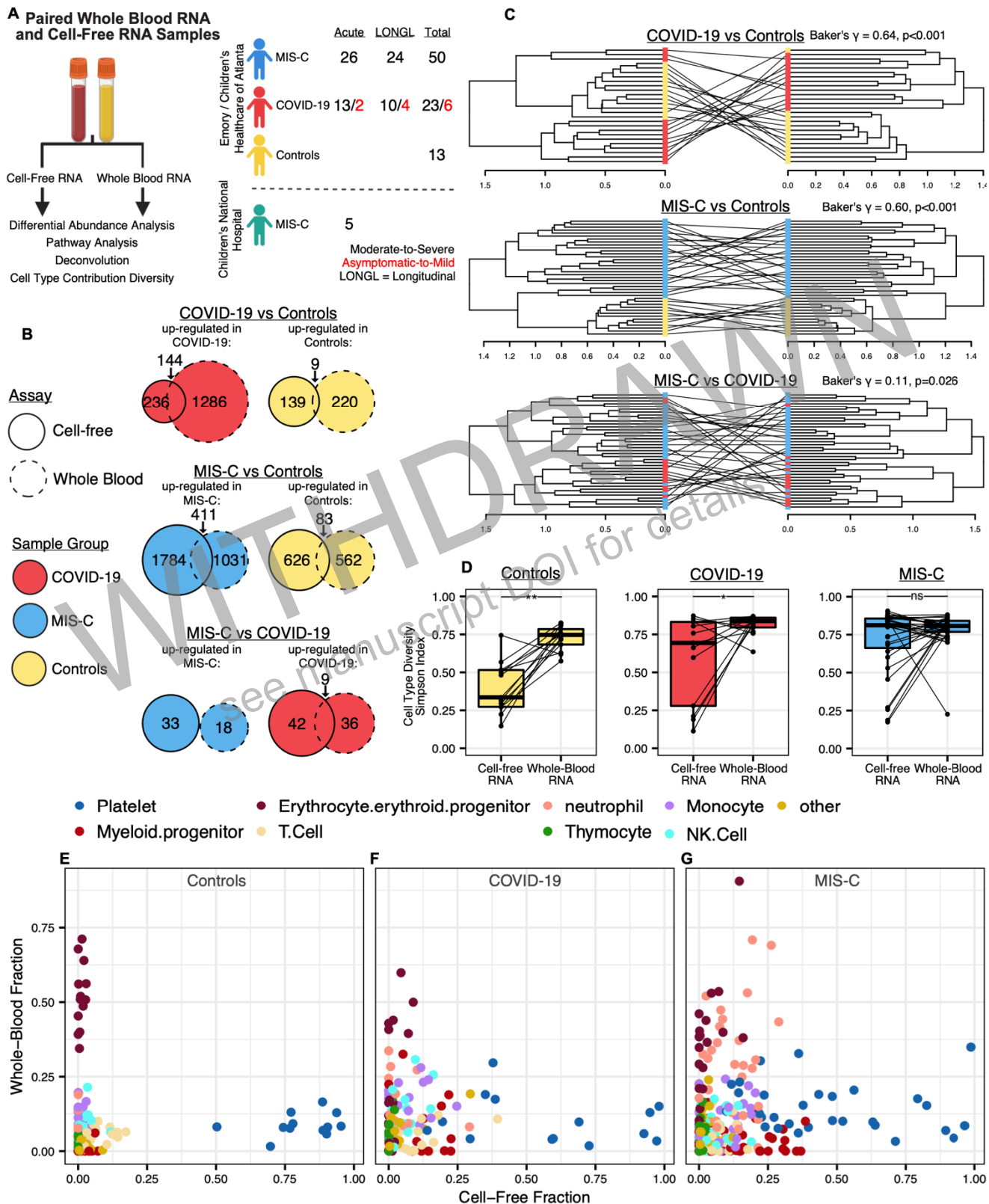
Figure 4. Plasma cfDNA tissues-of-origin by methylation profiling. (A) Study design and analysis overview. (B) Total cfDNA concentration and solid organ derived cfDNA concentration. (C) cfDNA concentration derived from innate and adaptive immune cell types. (D) Abundance of mitochondrial derived cfDNA. Asterisks indicate statistical significance by Mann-Whitney U test using Benjamini-Hochberg adjusted p-values as follows: ns, non-significant; *, p < 0.05; **, p < 0.01; ***, p < 0.001; ****, p < 0.001. (E) Correlation of cfDNA metrics with standard clinical measurements, along with the Pearson correlation Benjamini-Hochberg adjusted p-value for each comparison. Panels A and B show the number of samples corresponding to each group. Abbreviations: ASX, asymptomatic.

Comparative analysis of plasma cell-free RNA and whole blood RNA

A subset of cellular wbrRNA and cfrRNA samples were derived from the same blood draw (n=97), providing the opportunity to directly compare these two different analytes (Fig. 5A). First, we assessed the correlation of cfrRNA and wbrRNA abundance for genes with high average abundance (mean log-transformed CPM>10) across all sample categories. We found 1002 genes with significantly correlated wbrRNA and cfrRNA abundance (Pearson coefficient Benjamini-Hochberg adjusted p-value < 0.05). Positive correlation was observed in 992 (99%) of these 1002 genes, predominated by genes associated with myeloid cell transcription such as BNIP3L, HEMGN and NFKBIA (Supplemental Fig. 5A and B). Randomized permutation testing (n=1000 permutations) using either randomly paired genes or samples yielded on average far fewer significantly correlated genes (mean=250 and <1 with gene and sample randomization, respectively at a Benjamini-Hochberg adjusted p-value of <0.05) and decreased positive correlation (56% and 74% with gene and sample randomization, respectively), confirming the robustness of these observations. Next, we examined the degree of overlap in DEGs/DAGs among patients with either MIS-C (n=26) or moderate-to-severe COVID-19 (n=13), or controls (n=13) (Fig. 5B). We observed a substantial overlap in DEGs/DAGs in wbrRNA and cfrRNA when comparing MIS-C to controls (n=494) and COVID-19 to controls (n=153), but very little overlap

16 when comparing MIS-C to moderate-to-severe COVID-19 (n=9). Of the 9 DEGs/DAGs that overlap when
17 comparing MIS-C to moderate-to-severe COVID-19, seven had been previously reported in association with
18 COVID-19 (IFI6, IFI44L, RSAD2, LY6E, EPST11, XAF1, MX1)³⁴⁻³⁸. Similar to the gene abundance profiles,
19 differential pathway analysis revealed minimal overlap in top enriched pathways between wbRNA and cRNA
20 for each subgroup comparison (**Supplemental Fig. 5C and D**).

WITHDRAWN
see manuscript DOI for details



21 **Figure 5: Comparison of paired whole blood RNA-seq and cfRNA sequencing data. (A)** Study design and analysis
 22 overview. **(B)** Overlap of differentially abundant genes (DAGs)/differentially expressed genes (DEGs) between whole
 23 blood RNA-seq and cfRNA-seq using paired samples (DESeq2, Benjamini-Hochberg adjusted p-value < 0.05,
 24 |Log2FoldChange| > 1). Venn Diagrams represent the overlap of upregulated genes between analytes in each sample
 25 group, as indicated by fill color. **(C)** Comparison of clustering topology between paired whole blood RNA-seq and cfRNA
 26 sequencing samples. Samples were clustered based on correlation of DAGs/DEGs from their respective analyses. Lines
 27 between trees connect paired wbRNA and cfRNA samples. Correlation was calculated using Baker's Gamma and p-

values were calculated using a Monte Carlo permutation test. **(D)** Comparison of cell-types-of-origin diversity between cfRNA and wbRNA as calculated with a Simpson Index. Paired samples are connected with a line. Asterisks indicate statistical significance by Mann-Whitney U test using Benjamini-Hochberg adjusted p-values as follows: ns, non-significant; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.001$. **(E)** Cell-type-of-origin fractions, normalized to blood cell types and split by diagnosis group.

Next, we compared the agreement in unsupervised hierarchical sample clustering based on wbRNA and cfRNA profiling of DEGs (**Fig. 5C**). Samples clustered similarly in the MIS-C versus control and COVID-19 versus control comparisons (Baker's Gamma Index = 0.60 and 0.64, respectively; two-sided p-values < 0.01 for both, $n=39$ and $n=26$, respectively). In contrast, the MIS-C versus COVID-19 comparison yielded clustering that was less similar but still significantly correlated (Baker's Gamma Index = 0.11; two-sided p-value = 0.013, $n=39$). We used subsampling and bootstrapping to determine whether the high Baker's Gamma index in the COVID-19 versus control comparison was an artifact due to a smaller sample size ($n=26$). We did not observe a significant difference between the originally calculated values and bootstrapped distribution of Baker's Gamma Index values after subsampling (p-value = 0.88 for the MIS-C versus control comparison; p-value=0.40 for the MIS-C versus COVID-19 comparison). Thus, despite the lack of overlap in DEGs, cfRNA and wbRNA sample grouping by unsupervised clustering was similar for the MIS-C and COVID-19 samples relative to controls, indicating that these two analytes provide complementary information.

Finally, we compared the cell-types-of-origin (CTO) in wbRNA and cfRNA between paired samples. We observed significantly different wbRNA and cfRNA CTO diversity in paired samples from acute moderate-to-severe COVID-19 patients and paired samples from controls (p-value = 0.032 and p-value = 0.0007, respectively, by paired Wilcoxon test) (**Fig. 5D**). However, the CTO diversity was not significantly different between wbRNA and cfRNA in paired samples from patients with acute MIS-C (p-value = 0.41 by paired Wilcoxon test) (**Fig. 5D**). Furthermore, in unpaired analyses, we observed a significant difference in wbRNA CTO diversity between acute moderate-to-severe COVID-19 and controls, but not between acute MIS-C and controls (Mann-Whitney U Test, Benjamini-Hochberg adjusted p-value = 0.02 and 0.10) (**Supplemental Fig. 5E**). Conversely, we observe a significant difference in cfRNA CTO diversity in acute MIS-C and controls, but not in acute moderate-to-severe COVID-19 and controls (Mann-Whitney U Test, Benjamini-Hochberg adjusted p-value = 0.03 and 0.32) (**Supplemental Fig. 5E**). Finally, we compared patterns of blood-derived CTO in paired wbRNA and cfRNA samples from patients with MIS-C, patients with acute moderate-to-severe COVID-19, and controls (**Fig. 5E-G**). The wbRNA and cfRNA CTO profiles corresponding to MIS-C and COVID-19 (**Fig. 5F and G**), characterized by multiple blood cell types, were distinct from the control group profiles, which consisted of a single predominant baseline cell type (erythroid precursors in wbRNA and platelets in cfRNA) (**Fig. 5E**). Within each disease group, wbRNA and cfRNA CTO profiles were also distinct. For COVID-19, wbRNA CTO profiles had higher contributions from neutrophils, NK cells, T cells, and monocytes, whereas cfRNA CTO profiles had lower proportions of these inflammatory cells (**Fig. 5F**). In contrast, for MIS-C, wbRNA CTO profiles were predominated by neutrophils, whereas cfRNA CTO profiles had relatively higher contributions from myeloid progenitor cells, NK cells, and monocytes (**Fig. 5G**).

DISCUSSION

Here, we report a systems-level, longitudinal analysis of COVID-19 and MIS-C by next-generation sequencing of nucleic acids (cfRNA, wbRNA, and cfDNA) in a large multi-hospital study of 416 blood samples from 237 patients. Using plasma cfRNA profiling, we identify signatures associated with cellular injury and death that distinguish MIS-C and COVID-19, as well as the involvement of previously unreported cell types in MIS-C. wbRNA analysis reveals substantial overlap in pro-inflammatory pathways between MIS-C and COVID-19, yet also reveals pathways that are specific to each disease state. Plasma cfDNA profiling suggests increased cfDNA and solid organ involvement in MIS-C compared to COVID-19 and controls. Comparative

76 analyses of paired cfRNA and wbRNA samples demonstrate that these analytes yield separate, but
77 complementary, signatures associated with MIS-C and COVID-19 that reflect distinct cell types of origin. These
78 results provide novel insights into the differential pathogenesis of MIS-C and COVID-19 and lay the
79 groundwork for the development of minimally invasive diagnostic tests for these two disease states.

80 Whole blood RNA sequencing (RNA-Seq) of cellular RNA has traditionally been considered the gold
81 standard method for assaying gene expression in blood. However, the signal from wbRNA is primarily derived
82 from circulating leukocytes due to primarily sampling cell types found in the blood, thus measuring the
83 activation of a patient's inflammatory and immune response to an infection. In contrast, plasma cfRNA and
84 cfDNA both measure levels and types of cell death from circulating cell types and peripheral tissues^{17,39}. With
85 respect to cells that are turning over or dying, cfDNA enables precise quantification of cell numbers, whereas
86 cfRNA enables characterization of gene expression and pathways⁴⁰⁻⁴². cfRNA profiling can leverage the
87 extensive information regarding cell type specific gene expression that is available from recent large-scale
88 human cell atlas projects, whereas reference data for cfDNA methylation profiling is currently more limited¹⁷.
89 Overall, these cell-associated (wbRNA) and cell-free (cfRNA and cfDNA) approaches mutually complement
90 each other. When combined, they provide a more complete picture of the dynamic, "yin-yang" interplay
91 between host and pathogen or between cell activation / proliferation and cell death than achievable using a
92 single modality alone.

93 cfRNA is an analyte that probes cellular death and immune dynamics on a systems level. Previous
94 analyses of MIS-C and COVID-19 from the blood have relied on single cell or bulk RNA-Seq of whole blood
95 cells, which generally only characterizes the host immune response, or on proteomic and cytokine-based
96 assays, which use a limited number of markers or for which there is a lack of standardized reference data. In
97 contrast, the signals from cfRNA are derived from any cell or vascularized tissue type, and there is a plethora
98 of RNA-Seq reference data that can be used to interpret results. Consistent with prior studies, here we observe
99 increased levels of cfRNA from endothelial cells in MIS-C⁴³ and from neutrophils and thymocytes in MIS-C and
100 COVID-19^{9,44,45} as well as increased signaling from disease-specific pathways in MIS-C (IL-6, IL-8, and NF-
101 kB)^{12,14} and COVID-19 (olfactory, gustation, sumoylation, coronavirus replication, and HMGB1)^{11,13,46-48}.
102 However, the cfRNA data also uncovered several novel features of MIS-C, such as enrichment of neuronal
103 genes associated with synaptogenesis and increased cfRNA burden from Schwann cells. These findings
104 suggest that peripheral nervous system damage may be a common feature of MIS-C. Interestingly, both
105 central and peripheral nervous system involvement in MIS-C have been previously described^{49,50}, although
106 these clinical manifestations are infrequent. Notably, peripheral nervous system damage has also been
107 documented in pediatric and adult COVID-19 and in post-acute sequelae such as long COVID⁵¹⁻⁵³. Future
108 studies are needed to elucidate the mechanisms and clinical spectrum of neurologic involvement in acute MIS-
109 C and their association with long-term neurodevelopment. Furthermore, we observe an enrichment of genes
110 associated with the pyroptosis pathway in MIS-C, likely related to inflammasome activation⁵⁴. Pyroptosis is a
111 form of rapid cellular death that occurs during highly inflammatory states⁵⁵. Previous reports have shown that
112 pyroptosis occurs in vascular endothelial cells in Kawasaki disease, a similar systemic inflammatory
113 syndrome⁵⁶. Based on the observed increase of cfRNA from endothelial cells and cfRNA signatures of
114 pyroptosis, our data thus support the likely critical role of pyroptosis and endothelial cells in MIS-C
115 pathogenesis, and may help explain the overlapping clinical presentations between MIS-C and Kawasaki
116 disease in acutely ill pediatric patients.

117 Whole blood RNA-Seq reveals a high degree of overlap in shared, largely pro-inflammatory genes and
118 pathways between COVID-19 and MIS-C. This is expected as both diseases are caused by SARS-CoV-2 and
119 are highly inflammatory states. However, different levels of expression are observed for certain genes, such as
120 upregulation of ISG15 and SIGLEC1 in COVID-19 and upregulation of TRBV11-2 in MIS-C, as well as for
121 certain pathways, such as inhibition of T-cell receptor, IL-2, and FAK signaling pathways in MIS-C. The latter
122 finding is consistent with the observation of "T-cell exhaustion" associated with downregulation of NK and
123 CD8+ T cells driving a sustained inflammatory response in MIS-C³¹. Genes showing differences in levels of

24 gene expression (e.g. ISG15, SIGLEC1, TRBV11-2, CREB3L) or in persistence of gene expression as
25 revealed by longitudinal data (e.g. ADAMTS2, KLRFB1), may be useful target biomarkers for diagnostic and
26 prognostic assays that can discriminate between MIS-C, COVID-19, and other hyperinflammatory conditions.

27 Like cfRNA, cfDNA allows monitoring and quantification of tissue injury and cell death in a minimally
28 invasive manner from blood. Previous studies have shown that the relative concentrations of cfDNA in specific
29 tissues vary in different disease states, including COVID-19 in adults, solid-organ transplant rejection, graft-
30 versus-host disease following stem cell transplantation, urinary tract infection, and cancer^{32,57-60}. The profiling
31 of cell-type specific methylation in cfDNA can also be used to estimate cfDNA tissues-of-origin (TOO). Here we
32 quantify the concentration of cfDNA and perform cfDNA TOO profiling to assess the extent of cell, tissue, and
33 organ injury related to MIS-C and COVID-19. We observe an increase in cell death and high levels of
34 heterogeneity in TOO of cfDNA in MIS-C compared to COVID-19 and controls, consistent with the systemic
35 inflammatory nature of MIS-C that manifests clinically with the involvement of multiple organs and organ
36 systems. In addition, we observe a decrease in the concentration of mitochondrial cfDNA in MIS-C and
37 COVID-19, which is discrepant from previous studies reporting an increase of mitochondrial cfDNA by PCR⁶¹.

38 Most prior host-biomarker studies of MIS-C and pediatric COVID-19 analyze samples from the acute
39 symptomatic timepoint. Here, we report longitudinal sampling of cfRNA and wbRNA in patients with MIS-C and
40 COVID-19 at acute, post-acute, one-month post-hospitalization, and ≥ 3 months post-hospitalization timepoints.
41 We observe that most, but not all, gene measurements return to “baseline” levels similar to those in control
42 samples by one-month post-hospitalization. In wbRNA we observe opposing dynamics of ADAMTS2 levels in
43 MIS-C and COVID-19, in which elevated ADAMTS2 levels return to baseline in MIS-C but not in COVID-19 at
44 one-month post-hospitalization. Similarly, we observe recovery of wbRNA KLRB1 levels in MIS-C at one-
45 month post-hospitalization, but not in COVID-19. These findings may be related to the post-acute sequelae of
46 SARS-CoV-2 infection (“long COVID”) that has been postulated to be caused by persistent immune
47 dysregulation and that can occur after COVID-19⁶². In contrast, despite the severity of the initial presentation,
48 most clinical and laboratory abnormalities from MIS-C tend to quickly resolve within a few weeks, along with
49 normalization of inflammatory and injury biomarkers⁶³. In cfRNA, we found that most biomarker measurements,
50 such as CTO values and gene modules scores, persist at 1 month but return to baseline ≥ 3 months post-
51 hospitalization. These results are consistent with the generally accepted time frames of recovery after MIS-C⁶⁴.

52 We report a large-scale comparison (n=96) of wbRNA and cfRNA profiles from paired samples in MIS-
53 C and COVID-19. A previous study compared paired wbRNA and plasma cfRNA from healthy individuals but
54 was limited by small sample size (n=3) and lack of a disease group for comparison⁶⁵. Our results reveal
55 distinct, largely nonoverlapping sets of DAGs/DEGs associated with MIS-C, COVID-19, and non-inflammatory
56 controls in wbRNA and cfRNA. Thus, both analytes provide complementary information with regards to the
57 ability to discriminate MIS-C and COVID-19 from controls and from each other. These findings are consistent
58 with the origin of wbRNA and cfRNA; wbRNA being primarily derived from active immune cells in blood and
59 cfRNA from dying cells from the blood and peripheral tissues. They are also consistent with our CTO data
60 showing predominantly erythrocytes in wbRNA and platelets in cfRNA and a greater diversity of cell types from
61 peripheral tissues represented in cfRNA as compared to wbRNA. Overall, our results underscore the potential
62 utility of cfRNA and cfDNA as complementary biomarkers to more traditional diagnostic methods (wbRNA,
63 cytokines, proteomics) in better diagnosing and enhancing our understanding of complex disease states such
64 as MIS-C.

65 **Limitations**

66 This study has multiple limitations. First, we have a limited sample size of asymptomatic-to-mild
67 COVID-19 (wbRNA, n=6; cfRNA, n=6; cfDNA, n=10), as well as a limited number of longitudinally collected
68 samples (MIS-C or COVID-19 during post-acute, one-month, or ≥ 3 months timepoints: wbRNA, n=63; cfRNA,
69 n=45) and controls (cfDNA, n=5). Second, the accuracy of our deconvolution analyses may be limited by the
70 reference set used as a comparator. The cfDNA deconvolution reference set consists of a limited number of
71

cellular and tissue derived methylomes. The cfRNA/wbRNA deconvolution reference set is derived from a single cell RNA sequencing (scRNA-seq) atlas that sequenced poly-adenylated transcripts, while here we performed total RNA-Seq with host ribosomal RNA depletion. Third, we lacked samples from “look-alike” inflammatory conditions (e.g., Kawasaki disease, toxic shock syndrome, macrophage activation syndrome), and thus cannot determine whether the signatures in MIS-C and COVID-19 we observe are distinct from these other conditions. Future studies are needed to address these limitations.

METHODS

Ethics Statement. The protocols for this study were approved locally at each site by the University of California, San Francisco (UCSF) Institutional Review Board (IRB) (#21-33403), San Francisco, CA; Emory University IRB (STUDY00000723), Atlanta, GA; Children’s National Medical Center IRB (Pro00010632), Washington, DC; and Cornell University IRB for Human Participants (2012010003), New York, NY. The protocols at UCSF and Children’s National Medical Center were “no subject contact” sample biobanking protocols under which data was extracted from the medical chart and consent was not obtained. The protocol at Emory University IRB was a prospective enrollment study under which parents provided consent and children assent as appropriate for age. De-identified samples and patient information were shared with collaborating institutions for sample processing (UCSF and Cornell University) and analysis.

Sample Acquisition UCSF. Pediatric hospitalized patients who tested positive and negative for COVID-19 were identified from SARS-CoV-2 real-time PCR (RT-PCR) results from the UCSF Clinical Laboratories daily. Residual whole blood samples were collected in EDTA lavender top tubes from residual blood available. 250 μ l of sample was aliquoted with 250 μ l of 2x DNA/RNA shield (Zymo Research) for a 1:1 ratio. The remaining blood was centrifuged at 2500 rpm for 15 min and the available plasma was obtained. Samples were properly identified and added to the biobanking registry. All samples were stored at -80°C freezer until used.

Sample Acquisition Emory and Children’s Healthcare of Atlanta. At Emory and Children’s Healthcare of Atlanta, pediatric patients with COVID-19, MIS-C, or controls were enrolled into a specimen collection protocol following informed consent and assent, as appropriate for age. For this study, patients were classified as having MIS-C if they met the CDC case definition, and as having COVID-19 if they had any PCR-confirmed SARS-CoV-2 infection. Controls were healthy outpatients with no known history of COVID-19 who volunteered for specimen collection. The specimen collection protocol was approved by the Emory University IRB. Residual whole blood and plasma samples were collected from the clinical laboratory, and prospective blood samples were additionally collected in EDTA lavender top tubes. Longitudinal samples were also collected at 1-month and \geq 3-months timepoints for participants who returned for follow-up. From the EDTA tubes, whole blood was aliquoted, and the remaining blood was centrifuged at 2500 rpm for 15 min to obtain the available plasma. All samples were de-identified and assigned study IDs. Samples were stored at -80°C and shipped on dry ice to either UCSF or Cornell for analysis.

Sample Acquisition Children’s National. Patients with MIS-C were identified by a multidisciplinary task force according to the CDC case definition. Remnant whole blood samples from this population were identified, collected, and processed 12-72 hours after collection. Samples were centrifuged at 1300 xG for 5 minutes at room temperature. Plasma was aliquoted into a cryovial and frozen at -80°C. A DMSO-based cryopreservative (Cryosstor® CS10) was added in a 1:1 ratio to the cell pellet and then frozen at -80°C in a controlled rate freezing container (i.e., Mr. Frosty™). After freezing the pellets with Cryosstor they were transferred to liquid nitrogen cryostorage within 1 week.

19 **Clinical Data.** For the purposes of this study, MIS-C was defined as any patient who met the CDC case
20 definition⁵. Multidisciplinary teams which adjudicated whether a patient met the case definition of MIS-C.
21 COVID-19 was defined as any patient with PCR-confirmed SARS-CoV-2 infection within the preceding 14 days
22 who did not also meet the MIS-C case definition. Clinical data was abstracted from the medical record and
23 entered into a shared REDCap^{66,67} database housed at UCSF.

24
25 **Cell-free RNA Sample Processing and RNA Extraction.** Plasma samples were received on dry ice and
26 stored at -80°C until processed. Prior to extraction, plasma was thawed at room temperature and spun at
27 1300xg for 10min at 4°C. The supernatant was taken and cfRNA was isolated from plasma (115-1000 µl) using
28 the Norgen Plasma/Serum Circulating and Exosomal RNA Purification Mini Kit (51000, Norgen). Extracted
29 RNA was DNase treated with 14 µl of 10 µl DNase Turbo Buffer (AM2238, Invitrogen), 3 µl DNase Turbo
30 (AM2238, Invitrogen), 1µl Baseline Zero DNase (DB0715K, Lucigen-Epicenter) for 30 min at 37°C and then
31 concentrated into 12 µl using the Zymo RNA Clean and Concentrate Kit (R1015, Zymo).

32
33
34 **Cell-free RNA Library Preparation and Sequencing.** Sequencing libraries were constructed from 8 µl of
35 concentrated RNA using the Takara SMARTer® Stranded Total RNA-Seq Kit v2 – Pico Input Mammalian
36 (634418, Takara). Briefly, extracted RNA was reverse transcribed using random priming, barcoded using the
37 SMARTer RNA Unique Dual Index Kit (634451, Takara), rRNA depleted, and further amplified. Library
38 concentration was quantified using a Qubit™ 3.0 Fluorometer (Q33216, Invitrogen) with the dsDNA HS Assay
39 Kit (Q32854, Invitrogen). Libraries were quality-controlled using an Agilent Fragment Analyzer 5200
40 (M5310AA, Agilent) with the HS NGS Fragment kit (DNF-474-0500, Agilent). Libraries were pooled to equal
41 concentrations and sent to the Cornell Genomics core for 150-base pair, paired-end sequencing on an Illumina
42 NextSeq550 machine for an average of 10 million reads per sample.

43
44 **Cell-free DNA Sample Processing and Extraction.** Plasma samples were received on dry ice and stored at -
45 80°C until processed. Prior to extraction, plasma samples (75-650 µl) were thawed at room temperature and
46 spun at 1300xg for 10min at 4°C. The supernatant was taken and cfDNA was isolated from plasma using the
47 Qiagen Circulating Nucleic Acid Kit (55114, Qiagen) and eluted to 45 µl.

48
49 **Cell-free DNA Library Preparation and Sequencing.** Sequencing libraries were constructed from 20 µl of
50 extracted DNA as previously described³². Library concentration was quantified using a Qubit™ 3.0 Fluorometer
51 (Q33216, Invitrogen) with the dsDNA HS Assay Kit (Q32854, Invitrogen). Libraries were quality-controlled
52 using an Agilent Fragment Analyzer 5200 (M5310AA, Agilent) with the HS NGS Fragment kit (DNF-474-0500,
53 Agilent). Libraries were pooled to equal concentrations and sent to the Cornell Genomics core for 150-base
54 pair, paired-end sequencing on an Illumina NextSeq550 machine for an average of 33 million reads per
55 sample.

56
57 **Whole Blood RNA Sample Processing and RNA Extraction.** Whole blood samples were received on dry ice
58 and stored at -80°C until processed. Before extraction, all samples were thawed and pretreated with a 1:1 ratio
59 of 2X RNA/DNA Shield (R1200, Zymo Research) if this was not added prior to freezing. RNA was extracted
60 from whole blood samples (400 µl) using the Quick-RNA Whole Blood kit (R1201, Zymo Research) following
61 manufacturer's instructions. Ribosomal depletion was not performed. RNA was eluted in 15 µl of RNase-free
62 water and stored at -80 °C until use. The concentration of eluted RNA was measured using a Qubit™ Flex
63 Fluorometer (Q33326, Invitrogen) with the RNA HS Assay Kit (Q32852, Invitrogen). RNA Integrity was
64 assessed on a subset of samples using the Agilent Bioanalyzer RNA 6000 Nano/Pico Chip (5067-4626) to
65 determine the RNA Integrity Number (RIN). All samples analyzed were partially degraded with RIN values
66 between 2 and 5.

37

38 **Whole Blood RNA Library Preparation and Sequencing.** Extracted RNA (7 μ l or 10-200 ng) was processed
39 using the NEBNext rRNA Depletion Kit (Human/Mouse/Rat) (E6310, NEB) and the NEBNext Ultra II Directional
40 RNA Library Prep Kit (E7760 & E7765, NEB) following manufacturer's specifications. Briefly, samples were
41 first treated for ribosomal RNA depletion and DNase digestion, fragmented for 8 min, and reverse transcribed.
42 Adapters were ligated to the purified cDNA (1:25 diluted adaptor), followed by library amplification and
43 barcoding using NEBNext Multiplex Oligos (E6609L, NEB) sets 1 to 4. Libraries were purified using NEBNext
44 Sample Purification Beads (E7103, NEB).

45

46 Libraries were quantified using the Qubit™ Flex (Q33327, Invitrogen) with the dsDNA HS Assay Kit (Q32854,
47 Invitrogen). Libraries were pooled and sent to the UCSF Center of Advanced Technology (CAT) for sequencing
48 on an Illumina NovaSeq 6000 Sequencing System using 150-base pair paired-end sequencing. Negative
49 controls (nuclease-free water) were included in every run to monitor for contamination.

50

51 **RNA Bioinformatic Processing.** Sequencing data was processed using a custom bioinformatics pipeline
52 utilizing the Snakemake workflow management system (v7.7.0) for the cfRNA samples and a bash script for
53 the wbRNA samples. Samples were quality filtered and trimmed using BBDUK (v38.90), aligned to the
54 Gencode GRCh38 human reference genome (v38, primary assembly) using STAR (v2.7.0f) default
55 parameters, and features quantified using featureCount (v2.0.0). cfRNA samples were also deduplicated using
56 Picard MarkDuplicates prior to feature quantification (v2.19.2). Mitochondrial, ribosomal, X, and Y chromosome
57 genes were removed prior to analysis.

58

59 **RNA Sample Quality Filtering.** wbRNA and cfRNA samples were filtered using different QC metrics due to
60 differences in RNA concentration and quality. wbRNA samples with less than 10% of reads aligning to the
61 transcriptome were removed from all analyses. cfRNA samples were filtered on the basis of DNA
62 contamination, rRNA contamination, number of counts, and RNA degradation. DNA contamination was
63 estimated by calculating the ratio of reads mapping to introns and exons. Samples with an intron to exon ratio
64 above three were removed. rRNA contamination was measured using Samtools (v1.14). Total counts were
65 calculated using featureCounts. Degradation was estimated by calculating the 5-3' bias as calculated by
66 Qualimap (v2.2.1). Samples with rRNA contamination, total counts, or 5-3' bias greater than three standard
67 deviations from the mean were removed. Also, samples with less than 75,000 total counts were removed.

68

69 **RNA Cell Deconvolution and Diversity.** Cell type deconvolution was performed using BayesPrism (v1.1) with
70 the Tabula Sapiens single cell RNA-seq atlas (Release 1) as a reference. Cells from the Tabula Sapiens atlas
71 were grouped as previously described in Vorperian et al. Cell types with more than 100,000 unique molecular
72 identifiers (UMIs) were included in the reference and subsampled to 300 cells using ScanPy (v1.8.1). Providing
73 an equal number of cells to the deconvolution ensured an unbiased prior for the Bayesian algorithm used in
74 BayesPrism. Cell-type contribution diversity metrics were calculated using the vegan R package (v2.5.7).

75

76 **RNA Differential Abundance/Expression Analysis.** Comparative analysis of DEGs was performed using a
77 negative binomial model as implemented in the DESeq2 package (wbRNA: v1.28.1, cfRNA: v1.34.0) using a
78 Benjamini-Hochberg corrected p-value cutoff <0.01 , unless otherwise stated. Heatmaps were constructed
79 using the pheatmap package in R (v1.0.12), samples and genes were clustered using correlation based
80 hierarchical clustering. Gene ontology analysis was performed using the topGO R package (v2.46.0)²⁴ using a
81 Benjamini-Hochberg-corrected p-value cutoff of <0.05 . Cumulative CPM values for gene ontology terms were
82 calculated by taking the sum of normalized counts from all the significant genes in each gene ontology module.
83 Canonical pathways, diseases and functions were analyzed using QIAGEN Ingenuity Pathway Analysis (IPA)

14 software (v73620684). Pathway analysis was run twice: (1) with all available samples and (2) with samples that
15 have both whole blood and cell-free data available.

16
17 **RNA Paired Sample Correlation Analysis.** Log CPM normalized RNA counts were correlated between
18 paired cfRNA and wbRNA samples. Genes with a low average expression (<10 CPM) in either cfRNA or
19 wbRNA were removed. Samples with extreme cfRNA or wbRNA counts (CPM z-score >3) were removed to
20 eliminate outlier bias. A Pearson correlation was calculated for each gene and a Benjamini-Hochberg-
21 corrected p-value cutoff <0.05 was used to determine significance.

22
23 We performed two permutation tests (n=1000 permutations) to determine the reliability of the number of genes
24 observed to be correlated. First, we randomly shuffled the sample labels and recalculated the total number of
25 significant genes. Second, we randomly shuffled gene labels and recalculated the total number of significant
26 correlations.

27
28 **Entanglement Analysis.** Differential abundance/expression analysis was performed using only paired cfRNA
29 and wbRNA samples as previously described, using a Benjamini-Hochberg-corrected p-value cutoff <0.05.
30 Samples were clustered using correlation based hierarchical clustering. Using the dendextend package in R
31 (v1.15.2), dendrograms were plotted, paired samples were connected by lines, and Baker's gamma correlation
32 coefficient was calculated. Non-exact, two-sided p-values were calculated using a monte carlo permutation
33 test.

34
35 **DNA Concentration.** Eluted DNA was quantified using a Qubit™ 3.0 Fluorometer (Q33216, Invitrogen) with
36 the dsDNA HS Assay Kit (Q32854, Invitrogen). Total cfDNA concentration was estimated using the following
37 formula:

$$cfDNA\ concentration = \frac{(Eluted\ cfDNA\ concentration) * (Elution\ volume)}{(Plasma\ volume)}$$

38
39
40 **DNA Bioinformatic Processing.** Sequencing data was processed using a custom bioinformatics pipeline
41 utilizing the Snakemake workflow management system (v7.7.0). Samples were quality filtered and trimmed
42 using BBDUK (v38.46), aligned to the Gencode GRCh38 human reference genome (v38, primary assembly)
43 and deduplicated using Bismark (v0.22.1) with default parameters, and quality filtered using samtools (v1.14).
44 Prior to any analysis, X and Y chromosome mapped reads were removed.

45
46 **DNA Deconvolution.** DNA tissues of origin deconvolution were performed as previously described³². Briefly, a
47 custom bioinformatic pipeline utilizing the Snakemake workflow management system (v7.7.0) was used to
48 process publicly available methylation references, convert to a standard data format, and normalize across
49 samples. Metline (v0.2-7) was used to discover differentially methylated regions and a quadratic programming
50 algorithm estimated relative contributions of the tissues represented in the methylation references.

51
52 **Quantification and Statistical Analysis.** All statistical analyses were performed using R (cfDNA: v4.1.0,
53 cfRNA: v4.1.0, wbRNA: v4.0.3). Data wrangling and visualization was performed using Python (3.9.1), Pandas
54 (1.3.0) matplotlib-venn (0.11.6), R (v4.1.0), Tidyverse (v1.3.1), and ggplot2 (v3.3.5). Statistical significance was
55 tested using Wilcoxon signed-rank tests and Mann-Whitney U tests in a two-sided manner, unless otherwise
56 stated. All sequencing data was aligned to the GRCh38 Gencode v38 Primary Assembly and features counted
57 using the GRCh38 Gencode v38 Primary Assembly Annotation.

59 **Data Visualization.** Figures were created using Qiagen IPA, Adobe Illustrator, Affinity Designer, and
60 BioRender.com software.

61
62 **Acknowledgements.** We would like to acknowledge staff members at the UCSF Clinical Laboratories and the
63 UCSF Clinical Microbiology Laboratories for their help in identifying and retrieving patient whole blood
64 samples. We thank the Cornell Genomics Center and the UCSF Center for Advanced Technology for help
65 with sequencing libraries. At Emory, we thank Christopher Choi, Caroline Ciric, Khalel De Castro, Theda
66 Gibson, Hui-Mien Hsiao, Wensheng Li, Austin Lu, Lisa Macoy, Kathy Stephens, Madeline Taylor, Ashley
67 Tippet and the Children's Healthcare of Atlanta Research Laboratory for their contributions to specimen and
68 data collection. We thank the patients and their families for contributing their blood to further our understanding
69 of pediatric COVID-19 and MIS-C. This work was supported by the National Institutes of Health (NIH) /
70 National Institute of Child Health and Human Development (NICHD) grant R61HD105618 (R.D., C.A.R., I.D.V.,
71 and C.Y.C.). The funder had no role in study design, data collection and analysis, decision to publish, or
72 preparation of the manuscript.

73 74 **Author Contributions**

75 C.J.L., R.D., C.A.R., I.D.V., and C.Y.C. conceived and designed the study. C.J.L., A.S.-G., V.S., J.L., A.P.C.,
76 and A.B. performed sequencing experiments. J.N., J.L., M.E.W., P.S., N.B., J.S., W.S., C.H., N.W., A.F., A.G.
77 K.B., M.A.P., L.H., E.J.A., A.C., M.D., R.D., C.A.R., and C.Y.C. identified and collected patient samples and
78 clinical metadata. C.J.L., A.S.-G., V.S., B.B., M.D., R.D., C.A.R., I.D.V., and C.Y.C. analyzed sequencing data.
79 M.D., A.B., R.D., C.A.R., I.D.V., and C.Y.C. supervised the study. C.J.L., V.S., S.B., I.D.V., and C.Y.C. wrote
80 the manuscript and prepared the figures. All authors read and edited the manuscript and agree to its contents.

81 82 **Conflicts of Interest**

83 I.D.V. is a member of the Scientific Advisory Board of Karius Inc., Kanvas Biosciences and GenDX. C.Y.C. is a
84 founder and a member of the Scientific Advisory Board of Delve Bio. A.P.C. is listed as an inventor on
85 submitted patents pertaining to cell-free DNA (US patent applications 63/237,367, 63/056,249, 63/015,095,
86 16/500,929) and receives consulting fees from Eurofins Viracor. C.A.R.'s institution has received funding to
87 conduct clinical research unrelated to this manuscript from BioFire Inc., GSK, MedImmune, Micron, Merck,
88 Novavax, PaxVax, Regeneron, Pfizer, and Sanofi-Pasteur. She is co-inventor of patented RSV vaccine
89 technology (International PCT Application No. PCT/US2016/058976, filed 12/28/2016 by Emory University),
90 which has been licensed to Meissa Vaccines, Inc. with royalties received. Her institution has received funding
91 from NIH to conduct clinical trials of Moderna and Janssen COVID-19 vaccines. E.J.A has consulted for Pfizer,
92 Sanofi Pasteur, GSK, Janssen, and Medscape, and his institution receives funds to conduct clinical research
93 unrelated to this manuscript from MedImmune, Regeneron, PaxVax, Pfizer, GSK, Merck, Sanofi-Pasteur,
94 Janssen, and Micron. He also serves on a safety monitoring board for Kentucky BioProcessing, Inc. and Sanofi
95 Pasteur. He serves on a data adjudication board for WCG and ACI Clinical. His institution has also received
96 funding from NIH to conduct clinical trials of Moderna and Janssen COVID-19 vaccines. Atul Butte is a co-
97 founder and consultant to Personalis and NuMedii; consultant to Mango Tree Corporation, and in the recent
98 past, Samsung, 10x Genomics, Helix, Pathway Genomics, and Verinata (Illumina); has served on paid
99 advisory panels or boards for Geisinger Health, Regenstrief Institute, Gerson Lehman Group, AlphaSights,
100 Covance, Novartis, Genentech, and Merck, and Roche; is a shareholder in Personalis and NuMedii; is a minor
101 shareholder in Apple, Meta (Facebook), Alphabet (Google), Microsoft, Amazon, Snap, 10x Genomics, Illumina,
102 Regeneron, Sanofi, Pfizer, Royalty Pharma, Moderna, Sutro, Doximity, BioNtech, Invitae, Pacific Biosciences,
103 Editas Medicine, Nuna Health, Assay Depot, and Vet24seven, and several other non-health related companies
104 and mutual funds; and has received honoraria and travel reimbursement for invited talks from Johnson and
105 Johnson, Roche, Genentech, Pfizer, Merck, Lilly, Takeda, Varian, Mars, Siemens, Optum, Abbott, Celgene,
106 AstraZeneca, AbbVie, Westat, and many academic institutions, medical or disease specific foundations and

associations, and health systems. Atul Butte receives royalty payments through Stanford University, for several patents and other disclosures licensed to NuMedii and Personalis. Atul Butte's research has been funded by NIH, Peraton (as the prime on an NIH contract), Genentech, Johnson and Johnson, FDA, Robert Wood Johnson Foundation, Leon Lowenstein Foundation, Intervallen Foundation, Priscilla Chan and Mark Zuckerberg, the Barbara and Gerson Bakar Foundation, and in the recent past, the March of Dimes, Juvenile Diabetes Research Foundation, California Governor's Office of Planning and Research, California Institute for Regenerative Medicine, L'Oreal, and Progenity. The authors have declared that none of these companies or competing interests had any role in this work or manuscript.

Data and Code Availability

All code will be made available on Github. Processed sequencing data will be deposited in the National Institutes of Health (NIH) / National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) and Gene Expression Omnibus (GEO) repositories under restricted access via Database for Genotypes and Phenotypes (dbGAP).

REFERENCES

1. Woodruff, R. C. *et al.* Risk Factors for Severe COVID-19 in Children. *Pediatrics* **149**, (2021).
2. Shi, D. S. *et al.* Hospitalizations of Children Aged 5–11 Years with Laboratory-Confirmed COVID-19 — COVID-NET, 14 States, March 2020–February 2022. *MMWR. Morbidity and Mortality Weekly Report* **71**, 574–581 (2022).
3. COVID-19 Provisional Counts - Weekly Updates by Select Demographic and Geographic Characteristics. https://www.cdc.gov/nchs/nvss/vsrr/covid_weekly/index.htm.
4. Influenza-associated Pediatric Mortality. <https://gis.cdc.gov/GRASP/Fluview/PedFluDeath.html>.
5. Information for Healthcare Providers about Multisystem Inflammatory Syndrome in Children (MIS-C) | CDC. <https://www.cdc.gov/mis/mis-c/hcp/index.html>.
6. Verdoni, L. *et al.* An outbreak of severe Kawasaki-like disease at the Italian epicentre of the SARS-CoV-2 epidemic: an observational cohort study. *The Lancet* **395**, 1771–1778 (2020).
7. Godfred-Cato, S. *et al.* Distinguishing Multisystem Inflammatory Syndrome in Children From COVID-19, Kawasaki Disease and Toxic Shock Syndrome. *Pediatric Infectious Disease Journal* **41**, 315–323 (2022).
8. Consiglio, C. R. *et al.* The Immunology of Multisystem Inflammatory Syndrome in Children with COVID-19. *Cell* **183**, 968-981.e7 (2020).
9. Gruber, C. N. *et al.* Mapping Systemic Inflammation and Antibody Responses in Multisystem Inflammatory Syndrome in Children (MIS-C). *Cell* **183**, 982-995.e14 (2020).
10. Porritt, R. A. *et al.* The autoimmune signature of hyperinflammatory multisystem inflammatory syndrome in children. *Journal of Clinical Investigation* **131**, (2021).
11. Carter, M. J. *et al.* Peripheral immunophenotypes in children with multisystem inflammatory syndrome associated with SARS-CoV-2 infection. *Nature Medicine* **26**, 1701–1707 (2020).
12. Porritt, R. A. *et al.* The autoimmune signature of hyperinflammatory multisystem inflammatory syndrome in children. *Journal of Clinical Investigation* **131**, (2021).
13. Ramaswamy, A. *et al.* Immune dysregulation and autoreactivity correlate with disease severity in SARS-CoV-2-associated multisystem inflammatory syndrome in children. *Immunity* **54**, 1083-1095.e7 (2021).
14. Sacco, K. *et al.* Immunopathological signatures in multisystem inflammatory syndrome in children and pediatric COVID-19. *Nature Medicine* (2022) doi:10.1038/s41591-022-01724-3.
15. Porritt, R. A. *et al.* HLA class I-associated expansion of TRBV11-2 T cells in multisystem inflammatory syndrome in children. *Journal of Clinical Investigation* **131**, (2021).

16. Singh, V. *et al.* Limited induction of SARS-CoV-2-specific T cell responses in children with multisystem inflammatory syndrome compared with COVID-19. *JCI Insight* **7**, (2022).
17. Vorperian, S. K. *et al.* Cell types of origin of the cell-free transcriptome. *Nat Biotechnol* (2022) doi:10.1038/S41587-021-01188-9.
18. Chu, T., Wang, Z., Pe'er, D. & Danko, C. G. Cell type and gene expression deconvolution with BayesPrism enables Bayesian integrative analysis across bulk and single-cell RNA sequencing in oncology. *Nat Cancer* **3**, 505–517 (2022).
19. Tabula Sapiens Consortium* *et al.* The Tabula Sapiens: A multiple-organ, single-cell transcriptomic atlas of humans. *Science* **376**, eabl4896 (2022).
20. Simpson, E. H. Measurement of Diversity. *Nature* 1949 163:4148 **163**, 688–688 (1949).
21. Suratannon, N., Dik, W. A., Chatchatee, P. & Martin van Hagen, P. COVID-19 in children: Heterogeneity within the disease and hypothetical pathogenesis. *Asian Pacific Journal of Allergy and Immunology* **38**, 170–177 (2020).
22. Shannon, C. E. The mathematical theory of communication. 1963. *MD Comput* **14**, 306–17.
23. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, (2014).
24. Alexa, A. & Rahnenfuhrer, J. topGO: Enrichment Analysis for Gene Ontology. (2020).
25. Rasmussen, M. *et al.* RNA profiles reveal signatures of future health and disease in pregnancy. *Nature* **601**, 422–427 (2022).
26. Pös, O., Biró, O., Szemes, T. & Nagy, B. Circulating cell-free nucleic acids: Characteristics and applications. *European Journal of Human Genetics* **26**, 937–945 (2018).
27. Zhang, Y. *et al.* Identification of Monocytes Associated with Severe COVID-19 in the PBMCs of Severely Infected patients Through Single-Cell Transcriptome Sequencing. *Engineering (Beijing)* (2021) doi:10.1016/J.ENG.2021.05.009.
28. Yang, C. A., Huang, Y. L. & Chiang, B. L. Innate immune response analysis in COVID-19 and kawasaki disease reveals MIS-C predictors. *J Formos Med Assoc* **121**, 623–632 (2022).
29. Moreews, M. *et al.* Polyclonal expansion of TCR Vbeta 21.3 + CD4 + and CD8 + T cells is a hallmark of Multisystem Inflammatory Syndrome in Children. *Sci Immunol* **6**, (2021).
30. Kalfaoglu, B., Almeida-Santos, J., Tye, C. A., Satou, Y. & Ono, M. T-Cell Hyperactivation and Paralysis in Severe COVID-19 Infection Revealed by Single-Cell Analysis. *Frontiers in Immunology* **11**, (2020).
31. Beckmann, N. D. *et al.* Downregulation of exhausted cytotoxic T cells in gene expression networks of multisystem inflammatory syndrome in children. *Nature Communications* **12**, (2021).
32. Cheng, A. P. *et al.* Cell-free DNA tissues of origin by methylation profiling reveals significant cell, tissue, and organ-specific injury related to COVID-19 severity. *Med* **2**, 411-422.e5 (2021).
33. Scozzi, D. *et al.* Circulating mitochondrial DNA is an early indicator of severe illness and mortality from COVID-19. *JCI Insight* **6**, (2021).
34. Zhu, L. *et al.* Single-Cell Sequencing of Peripheral Mononuclear Cells Reveals Distinct Immune Response Landscapes of COVID-19 and Influenza Patients. *Immunity* **53**, 685-696.e3 (2020).
35. Pfaender, S. *et al.* LY6E impairs coronavirus fusion and confers immune control of viral disease. *Nat Microbiol* **5**, 1330–1339 (2020).
36. Gao, X. *et al.* Genome-wide screening of SARS-CoV-2 infection-related genes based on the blood leukocytes sequencing data set of patients with COVID-19. *J Med Virol* **93**, 5544–5554 (2021).
37. Shaath, H., Vishnubalaji, R., Elkord, E. & Alajez, N. M. Single-Cell Transcriptome Analysis Highlights a Role for Neutrophils and Inflammatory Macrophages in the Pathogenesis of Severe COVID-19. *Cells* **9**, (2020).
38. Karami, H. *et al.* Weighted gene co-expression network analysis combined with machine learning validation to identify key modules and hub genes associated with sars-cov-2 infection. *Journal of Clinical Medicine* **10**, (2021).

39. Lehmann-Werman, R. *et al.* Identification of tissue-specific cell death using methylation patterns of circulating DNA. *Proc Natl Acad Sci U S A* **113**, E1826–E1834 (2016).
40. Fox-Fisher, I. *et al.* Remote immune processes revealed by immune-derived circulating cell-free DNA. *Elife* **10**, (2021).
41. Wang, Y. *et al.* Plasma cell-free RNA characteristics in COVID-19 patients. *Genome Res* **32**, 228–241 (2022).
42. Pan, W. *et al.* Simultaneously Monitoring Immune Response and Microbial Infections during Pregnancy through Plasma cfRNA Sequencing. *Clin Chem* **63**, 1695–1704 (2017).
43. Diorio, C. *et al.* Proteomic profiling of MIS-C patients indicates heterogeneity relating to interferon gamma dysregulation and vascular endothelial dysfunction. *Nat Commun* **12**, (2021).
44. Vella, L. A. *et al.* Deep immune profiling of MIS-C demonstrates marked but transient immune activation compared to adult and pediatric COVID-19. *Sci Immunol* **6**, (2021).
45. Reusch, N. *et al.* Neutrophils in COVID-19. *Front Immunol* **12**, (2021).
46. Harikrishnan, P. Gustatory Dysfunction as an Early Symptom in COVID-19 Screening. *The Journal of Craniofacial Surgery* **31**, e656–e658 (2020).
47. Ibrahim, I. H. & Ellakwa, D. E. S. SUMO pathway, blood coagulation and oxidative stress in SARS-CoV-2 infection. *Biochem Biophys Res* **26**, (2021).
48. Chen, R. *et al.* HMGB1 as a potential biomarker and therapeutic target for severe COVID-19. *Heliyon* **6**, e05672 (2020).
49. Hacoheh, Y. *et al.* Neurologic and Radiographic Findings Associated with COVID-19 Infection in Children. *JAMA Neurology* **77**, 1440–1445 (2020).
50. Larovere, K. L. *et al.* Neurologic Involvement in Children and Adolescents Hospitalized in the United States for COVID-19 or Multisystem Inflammatory Syndrome. *JAMA Neurol* **78**, 536–547 (2021).
51. Patti, R., *et al.* Post-Infectious Demyelinating Polyneuropathy Secondary to COVID-19 Infection. *Chest* **160**, A658 (2021).
52. Feldstein, L. R. *et al.* Characteristics and Outcomes of US Children and Adolescents With Multisystem Inflammatory Syndrome in Children (MIS-C) Compared With Severe Acute COVID-19. *JAMA* **325**, 1074–1087 (2021).
53. Balbi, P. *et al.* Peripheral Neuropathy in Patients Recovering from Severe COVID-19: A Case Series. *Medicina (Kaunas)* **58**, 523 (2022).
54. Vora, S. M., Lieberman, J. & Wu, H. Inflammasome activation at the crux of severe COVID-19. *Nat Rev Immunol* **21**, 694–703 (2021).
55. Bergsbaken, T., Fink, S. L. & Cookson, B. T. Pyroptosis: host cell death and inflammation. *Nat Rev Microbiol* **7**, 99–109 (2009).
56. Jia, C. *et al.* Endothelial cell pyroptosis plays an important role in Kawasaki disease via HMGB1/RAGE/cathepsin B signaling pathway and NLRP3 inflammasome activation. *Cell Death Dis* **10**, (2019).
57. Chabon, J. J. *et al.* Integrating genomic features for non-invasive early lung cancer detection. *Nature* **580**, 245–251 (2020).
58. Burnham, P. *et al.* Urinary cell-free DNA is a versatile analyte for monitoring infections of the urinary tract. *Nat Commun* **9**, (2018).
59. Cheng, A. P. *et al.* Cell-free DNA profiling informs all major complications of hematopoietic cell transplantation. *Proc Natl Acad Sci U S A* **119**, (2022).
60. de Vlaminck, I. *et al.* Circulating cell-free DNA enables noninvasive diagnosis of heart transplant rejection. *Sci Transl Med* **6**, (2014).
61. Andargie, T. E. *et al.* Cell-free DNA maps COVID-19 tissue injury and risk of death and can cause tissue injury. *JCI Insight* **6**, (2021).

- 19 62. Phetsouphanh, C. *et al.* Immunological dysfunction persists for 8 months following initial mild-to-
20 moderate SARS-CoV-2 infection. *Nat Immunol* **23**, 210–216 (2022).
- 21 63. Fremed, M. A. & Farooqi, K. M. Longitudinal Outcomes and Monitoring of Patients With Multisystem
22 Inflammatory Syndrome in Children. *Front Pediatr* **10**, (2022).
- 23 64. Multisystem Inflammatory Syndrome in Children (MIS-C): Information for Healthcare Providers About
24 Talking with Families and Caregivers. <https://www.cdc.gov/mis/mis-c/hcp/provider-families.html>.
- 25 65. Ibarra, A. *et al.* Non-invasive characterization of human bone marrow stimulation and reconstitution by
26 cell-free messenger RNA sequencing. *Nat Commun* **11**, (2020).
- 27 66. Harris, P. A. *et al.* The REDCap consortium: Building an international community of software platform
28 partners. *J Biomed Inform* **95**, (2019).
- 29 67. Harris, P. A. *et al.* Research electronic data capture (REDCap)--a metadata-driven methodology and
30 workflow process for providing translational research informatics support. *J Biomed Inform* **42**, 377–381
31 (2009).
- 32

WITHDRAWN
see manuscript DOI for details

Pediatric Cohort

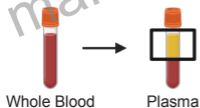
Number of Patients (n=237)

UCSF, Emory University/Children's Healthcare of Atlanta, and Children's National Hospital

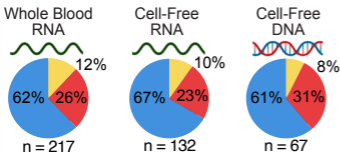


Blood and Plasma Collection

Number of Samples (n=416)



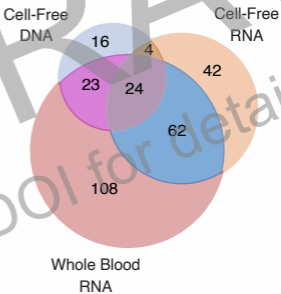
Nucleic Acid Isolation



Next-Gen Sequencing

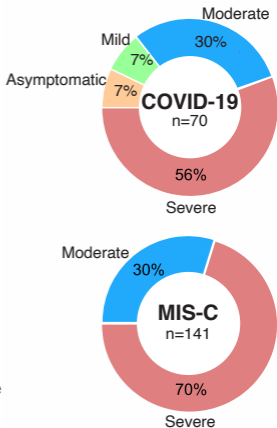
Total RNA Sequencing Total RNA Sequencing Whole Genome Bisulfite Sequencing

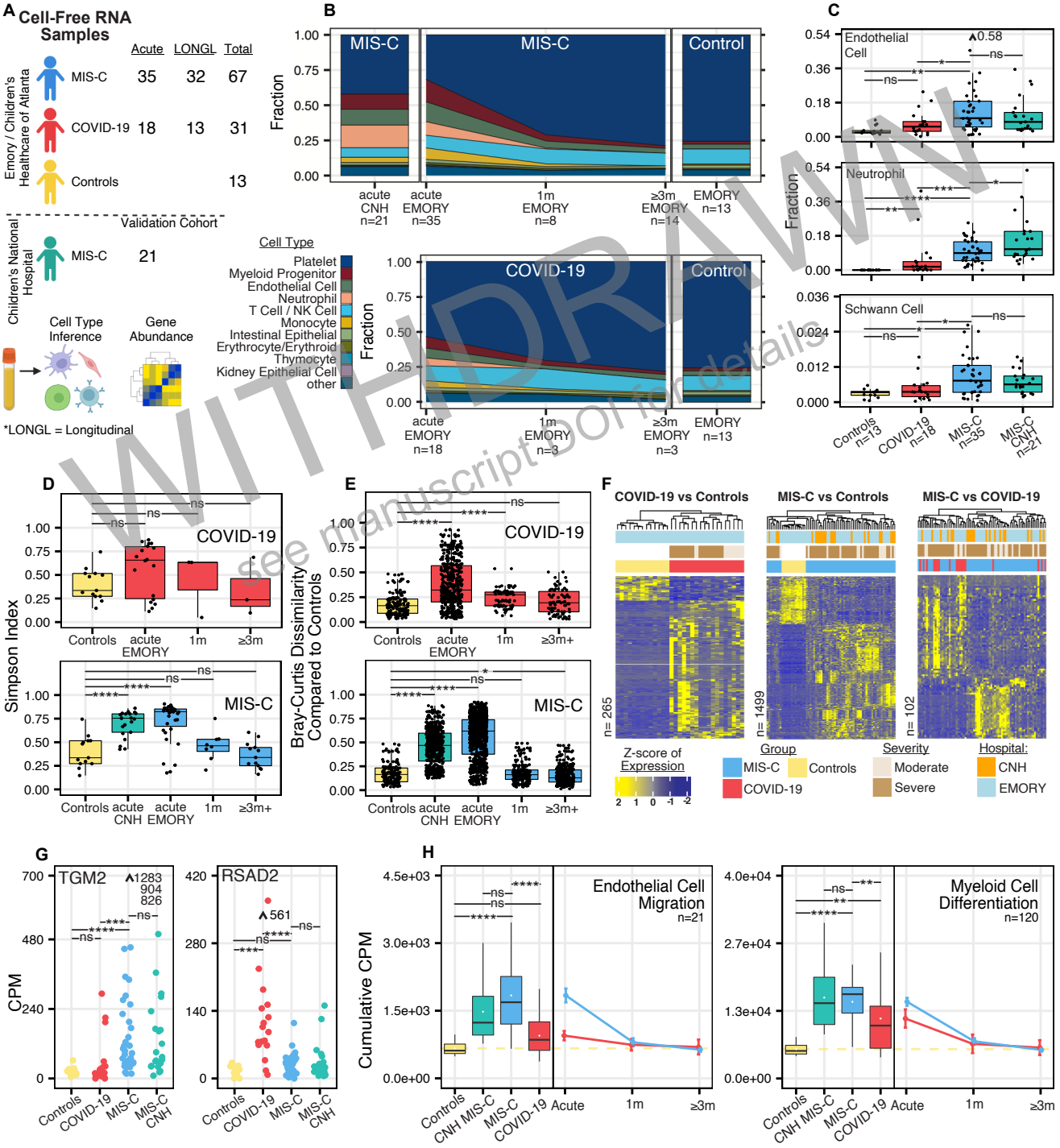
Sample Overlap

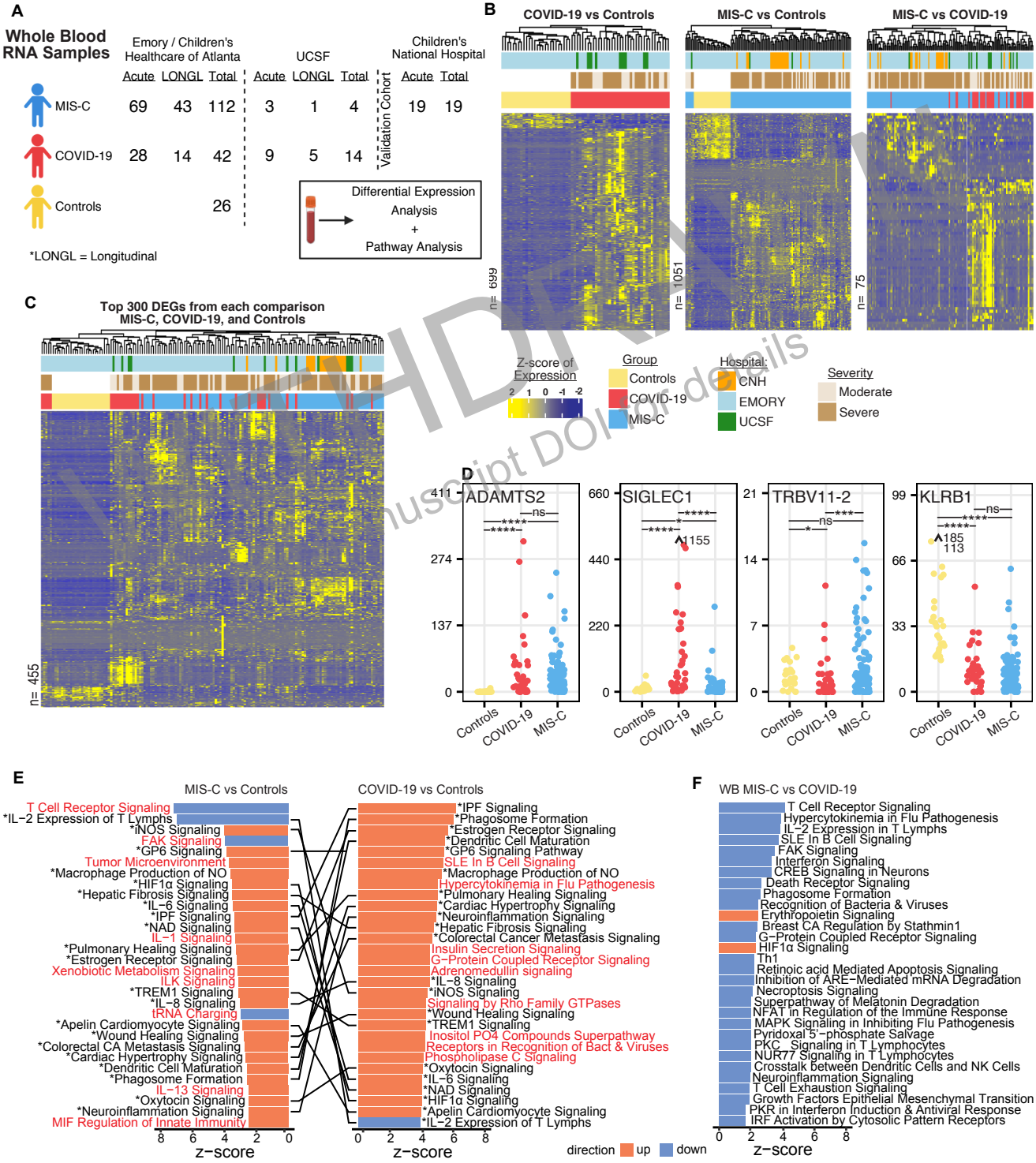


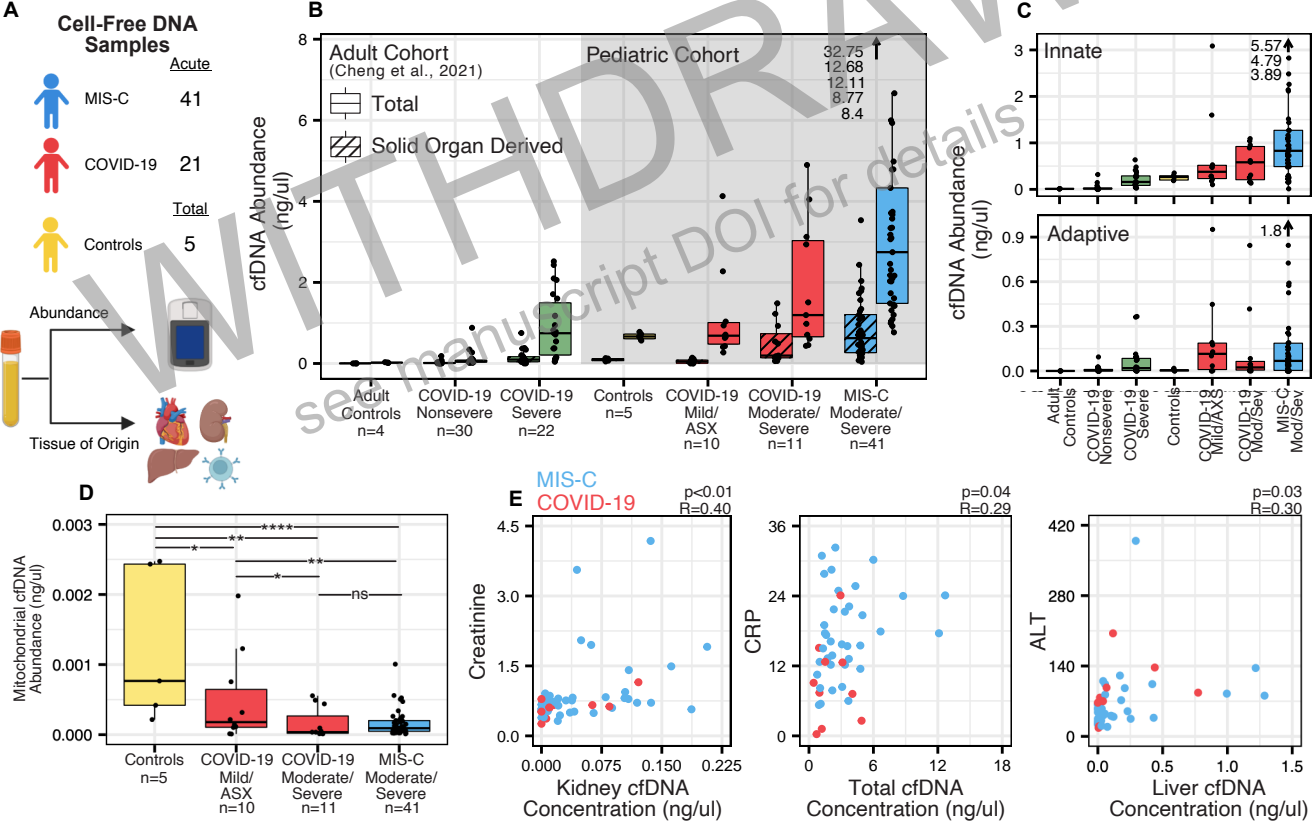
C

Disease Severity Classification

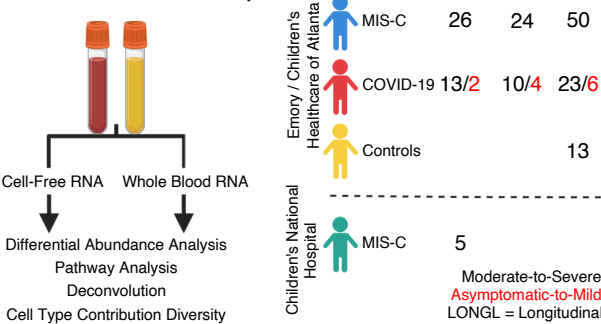








Paired Whole Blood RNA and Cell-Free RNA Samples



	Acute	LONGL	Total
MIS-C	26	24	50
COVID-19	13/2	10/4	23/6
Controls			13
MIS-C	5		

Moderate-to-Severe
Asymptomatic-to-Mild
LONGL = Longitudinal

