




ARTICLE



<https://doi.org/10.1038/s41467-022-30357-w>

OPEN

# An Artificial Intelligence-guided signature reveals the shared host immune response in MIS-C and Kawasaki disease

Pradipta Ghosh<sup>1,2,11</sup> , Gajanan D. Katkar<sup>1,11</sup>, Chisato Shimizu<sup>3,4,11</sup> , Jihoon Kim<sup>5</sup>, Soni Khandelwal<sup>6</sup>, Adriana H. Tremoulet<sup>3,4</sup>, John T. Kanegaye<sup>3,4</sup>  & Pediatric Emergency Medicine Kawasaki Disease Research Group\*

Multisystem inflammatory syndrome in children (MIS-C) is an illness that emerged amidst the COVID-19 pandemic but shares many clinical features with the pre-pandemic syndrome of Kawasaki disease (KD). Here we compare the two syndromes using a computational toolbox of two gene signatures that were developed in the context of SARS-CoV-2 infection, i.e., the viral pandemic (ViP) and severe-ViP signatures and a 13-transcript signature previously demonstrated to be diagnostic for KD, and validated our findings in whole blood RNA sequences, serum cytokines, and formalin fixed heart tissues. Results show that KD and MIS-C are on the same continuum of the host immune response as COVID-19. Both the pediatric syndromes converge upon an *IL15/IL15RA*-centric cytokine storm, suggestive of shared proximal pathways of immunopathogenesis; however, they diverge in other laboratory parameters and cardiac phenotypes. The ViP signatures reveal unique targetable cytokine pathways in MIS-C, place MIS-C farther along in the spectrum in severity compared to KD and pinpoint key clinical (reduced cardiac function) and laboratory (thrombocytopenia and eosinopenia) parameters that can be useful to monitor severity.

<sup>1</sup>Department of Cellular and Molecular Medicine, University of California San Diego, San Diego, USA. <sup>2</sup>Department of Medicine, University of California San Diego, San Diego, USA. <sup>3</sup>Department of Pediatrics, University of California San Diego, San Diego, USA. <sup>4</sup>Rady Children's Hospital-San Diego, San Diego, CA, USA. <sup>5</sup>Department of Biomedical informatics, University of California San Diego, San Diego, USA. <sup>6</sup>Department of Computer Science and Engineering, Jacob's School of Engineering, University of California San Diego, San Diego, USA. <sup>11</sup>These authors contributed equally: Pradipta Ghosh, Gajanan D. Katkar, Chisato Shimizu. \*A list of authors and their affiliations appears at the end of the paper. <sup>✉</sup>email: [prghosh@ucsd.edu](mailto:prghosh@ucsd.edu)

**M**ultisystem inflammatory syndrome in children<sup>1</sup> (MIS-C; initially named Pediatric Inflammatory Multisystem Syndrome Temporally associated with SARS-CoV-2, PIMS-TS)<sup>2</sup> is a rare but severe condition that occurs in children and adolescents ~4–6 weeks after exposure to SARS-CoV-2. First reported in April 2020 in a cluster of children in the United Kingdom<sup>3</sup>, followed by other regions of the world<sup>4</sup>, the syndrome is characterized by fever, and variably accompanied by rash, conjunctival injection, gastrointestinal symptoms, shock, and elevated markers of inflammation and antibodies to SARS-CoV-2 in the majority of patients. Myocardial dysfunction and coronary arterial dilation may resemble those seen in another uncommon childhood condition, Kawasaki Disease (KD). KD is an acute inflammatory disorder predominantly seen in young children. Since it was first described in Japan<sup>5</sup> in 1967, KD has emerged as the most common cause of pediatric acquired heart disease in the developed world<sup>6</sup>. Little is known about the definitive triggers of KD; what is most widely accepted is that KD is largely an immune response to a plethora of infectious or environmental stimuli including viruses, fungi (e.g., *Candida* sp.), and bacteria<sup>7–9</sup>. The host genetic background appears to shape this idiosyncratic inflammatory response to an environmental antigen exposure<sup>9</sup>.

On May 14, 2020, the CDC published an online Health Advisory that christened this condition as Multisystem Inflammatory Syndrome in Children (MIS-C) and outlined a case definition<sup>10</sup>. Since then, as the COVID-19 pandemic spread across many countries, cases of MIS-C soared, with features of shock and cardiac involvement requiring inotropic support [in the critical care setting]. But distinguishing MIS-C from KD, KD shock syndrome<sup>11</sup>, and other severe infectious or inflammatory conditions remains a challenge. The need for early diagnostic and prognostic markers of disease severity remains unmet; such markers could objectively guide decisions regarding the appropriateness of the level of care and the timing of initiation of life-saving supportive and therapeutic measures.

As for the immunopathogenesis of MIS-C, limited but key insights have emerged rapidly, most of which focus on the differences between MIS-C and KD. For example, Gruber et al.<sup>12</sup>, and Consiglio et al.<sup>13</sup>, showed that the inflammatory response in MIS-C differs from KD with respect to T cell subsets<sup>13</sup>. These conclusions were generally supported by two other studies, by Vella et al.<sup>14</sup>, and Ramaswamy et al.<sup>15</sup> who also showed that severe MIS-C patients displayed skewed memory T cell TCR repertoires and autoimmunity characterized by endothelium-reactive IgG. Finally, Carter et al.<sup>16</sup>, reported activation of CD4<sup>+</sup>CCR7<sup>+</sup> T cells and  $\gamma\delta$  T cell subsets in MIS-C, which had not been reported in KD, which made them conclude that MIS-C may be a distinct immunopathogenic illness. While these studies further our understanding of MIS-C and the major conclusions of these studies are comprehensively reviewed elsewhere<sup>17</sup>, it is noteworthy that each of these studies had some notable limitations—(i) in Gruber et al.<sup>12</sup>, most of the MIS-C subjects were on immunomodulatory medications when samples were drawn; (ii) in Vella et al.<sup>14</sup>, absence of contemporaneously analyzed healthy pediatric samples which were not available during the early phase of the pandemic; (iii) in Carter et al.<sup>16</sup>, KD subjects were not concurrently studied and the authors themselves acknowledged that such side-by-side immunophenotyping of MIS-C and KD would be necessary to draw conclusions convincingly regarding similarities and differences between these two syndromes; (iv) absence of validation studies in independent cohorts in them all.

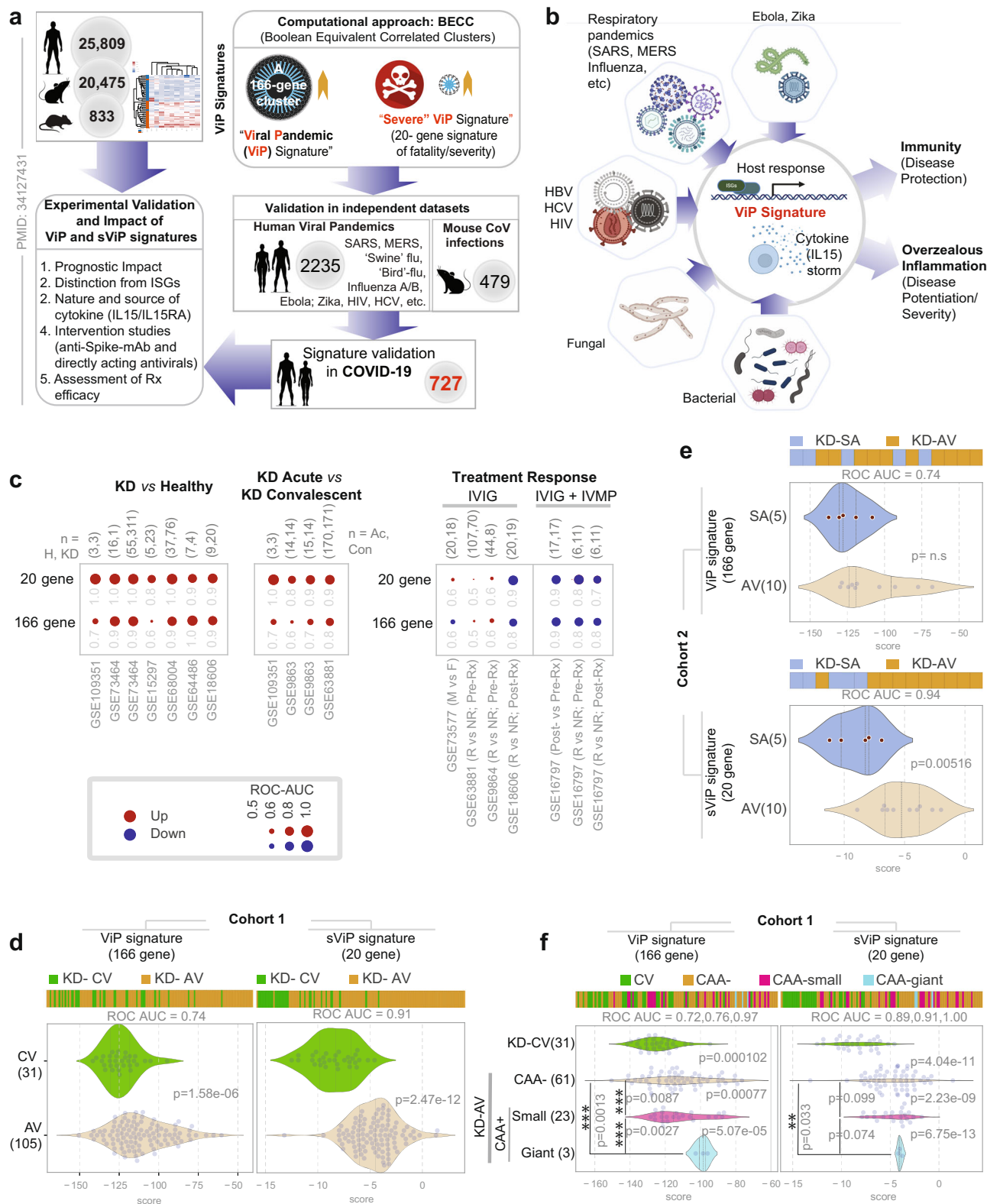
We recently showed that a 166-gene signature is conserved in all *viral* pandemics (ViP), including COVID-19, and a subset of 20-genes within that signature that classifies disease severity<sup>18</sup>. In the absence of a sufficiently large number of COVID-19 datasets at the onset of the COVID-19 pandemic, these ViP signatures

were trained on only two datasets from the pandemics of the past (influenza and avian flu; GSE47963,  $n = 438$ ; GSE113211,  $n = 118$ ) and used without further training to prospectively analyze the samples from the current pandemic (i.e., COVID-19). The ViP signatures appeared to capture the ‘invariant’ host response, i.e., the shared fundamental nature of the host immune response induced by all viral pandemics, including COVID-19. Here we used the ViP signatures as a starting computational framework to navigate the syndrome of MIS-C that is still a relatively poorly understood entity, but none-the-less recognized as an immunologic response to viral exposure. More specifically, we sought to interrogate concurrently the quality and quantity of the shared and unique features in MIS-C and KD. Our results show that the nature of the host immune response in MIS-C is similar to that in the pre-pandemic syndrome of KD, i.e., both are characterized by a *IL15/IL15RA*-centric cytokine storm; however, MIS-C is farther along in the spectrum of disease severity.

## Results

### A gene signature seen in COVID-19 is also induced in KD, and tracks disease severity.

We sought to define the host immune response in KD and compare that to COVID-19 using an artificial intelligence (AI)-based approach. To this end, we took advantage of a recently identified analysis of the host immune response in COVID-19 in which over 45,000 transcriptomic datasets of viral pandemics were analyzed to extract a 166-gene signature<sup>18</sup> (summarized in Fig. 1a). Because publicly available transcriptomic datasets from SARS-CoV-2-infected patients were still relatively few at the onset of the pandemic, the rigor of analysis was increased through the use of an informatics approach, i.e., Boolean equivalent correlated clusters (BECC<sup>19</sup> Fig. 1a) that can identify fundamental invariant (i.e., universally conserved) gene expression relationships underlying any biological domain; in this case, the biological domain of ‘respiratory viral pandemics’ was selected. Unlike some of the mainstream computational approaches (e.g., differential expression, Bayesian, and correlation network analyses, etc.) that are geared to identify the entire spectra of host immune response, BECC exclusively focuses on Boolean equivalent relationships to identify potentially functionally related gene sets that are part of the invariant spectrum of the host response. The resultant 166-gene ViP signature, developed from just two training datasets of pandemics of the past, was found to be conserved in all viral pandemics and outbreaks, including prospective studies on all COVID-19 datasets. The signature reflected the shared fundamental nature of the host immune response to multiple infectious triggers (Fig. 1b summarizes the types of pathogens that were found to induce the ViP signatures<sup>18</sup>). More specifically, the nature of the host immune response was found to be predominantly *IL15/IL15RA*-centric and enabled the formulation of precise therapeutic goals and measurement of therapeutic efficacy. At a molecular level, the ViP signatures were distinct from interferon-stimulated genes (ISGs<sup>20,21</sup>), in that, they revealed the broader and fundamental nature of the host immune response, shared between diverse pathogens and tissue/cell types. This included some tell-tale expected (Type I Interferon and cytokine signaling) and some unique (cellular senescence, exhaustion, chromatin silencing, regulation of apoptosis) pathway enrichments<sup>18</sup>. The latter, i.e., the unique pathways, were specifically enriched in a 20-gene subset of the ViP signature, which we called severe (s)ViP signature; this signature was trained on a large dataset of Influenza A/B-infected adult patients annotated with clinical severity<sup>18</sup>. The sViP signature predicted disease severity in COVID-19 (respiratory failure, need for mechanical ventilation, prolonged hospitalization and/or death)<sup>18</sup>. Consequently, the ViP signatures, but



not ISGs, were found to be prognostic of disease severity in cohorts of COVID-19 datasets<sup>18</sup>.

Because no KD and/or MIS-C samples were used previously in training the model that led to the discovery of the ViP signatures<sup>18</sup>, we used these signatures as is, without further training, as quantitative and qualitative frameworks for measuring the immune response in publicly available historic cohorts of KD predating COVID-19. Both ViP and sViP signatures were

upregulated in blood and tissue samples derived from patients with KD compared to healthy controls (ROC AUC for classification of KD vs. healthy ranged from 0.8 to 1.00 across 7 independent cohorts; Fig. 1c, left), and that such induction was limited to the acute phase of KD and downregulated during convalescence (ROC AUC for classification of KD vs. healthy ranged from 0.6–0.8 for ViP and 0.8–1.00 for sViP across 4 independent cohorts; Fig. 1c, middle).

**Fig. 1 A Viral Pandemic (ViP) signature that is induced in COVID-19<sup>18</sup>, is induced also in epidemic outbreaks of KD.** **a** Schematic displays the computational approach (BECC) and rigor (diversity and number of datasets) used to identify the 166-gene ViP and a subset of 20-gene severe (s)ViP signatures, and the subsequent experimentally validated inferences and impact of the same in a recent study<sup>18</sup>. The numbers in gray circles denote the total number of datasets analyzed in each category. **b** Schematic displays the various pathogenic triggers that induce ViP signatures (many of which are triggers also for KD) and the prominent induction of *IL15/IL15RA* as an invariant nature of the cytokine storm. **c** Bubble plots of ROC-AUC values (radii of circles are based on the ROC-AUC) demonstrating the strength of classification and the direction of gene regulation (Up, red; Down, blue) for the classification based on the 20-gene severe ViP signature (top) and 166-gene ViP signature (bottom) in numerous publicly available historic datasets. ViP signatures classified KD vs. healthy children (left), acute vs. convalescent KD (middle) and treatment response in the setting of combination therapy with IV steroids (MP methylprednisone) and IV IgG alone (IVIg), but not IVIg alone. Numbers on top of bubble plots indicate number of subjects in each comparison group. **d, e** Bar (top) and violin (bottom) plots display the classification of blood samples that were collected during acute (AV), sub-acute (SA; ~10–14 days post-discharge) and convalescent (CV; 1 year post-onset) visits from two independent KD cohorts (**d**; Historic Cohort 1; **e**; Prospective Cohort 2) using ViP (left) or sViP (right) signatures. **f** Bar (top) and violin (bottom) plots display the sub-classification of blood samples in Cohort 1 based on coronary artery aneurysm (CAA) status using ViP (left) or sViP (right) signatures. Welch's two sample unpaired two-sided *t*-test is performed on the composite gene signature score to compute the *p* values. In multi-group setting each group is compared to the first control group and only significant *p* values are displayed on the right. Additional *p* values are displayed on the left.

The strength of association between ViP/sViP signatures and acute KD was also preserved in datasets comprised of KD samples prospectively collected before and after IVIg treatment, and treatment response was annotated as responder (R) or non-responder (NR) (Fig. 1c, right). First, sex had no impact on the induction of signatures (ROC AUC 0.6 in Males vs. Females). Second, although the ViP/sViP signatures did not predict treatment response to IVIg (ROC AUC 0.5–0.6 in pre-treatment samples R vs. NR; GSE63881 and GSE9864), they were reduced in all responders compared to non-responders (ROC AUC 0.8–0.9 in post-treatment samples R vs. NR; GSE18606). Finally, in a study<sup>22</sup> in which the intervention was a combination of IVIg with the intravenous methylprednisolone (IVMP), both ViP signatures were reduced post-Rx (ROC AUC 0.9; GSE16797), and the signatures performed equally well in both pre-treatment and post-treatment samples in differentiating responders from non-responders (ROC AUC 0.7–0.8). These findings suggest that while the IVIg-IVMP combination regimen reduced the signatures effectively among all patients ( $n = 17$ ), responders induced the ViP signatures to a lesser extent than non-responders. The 20-gene sViP signature consistently outperformed the 166-gene ViP signature in its ability to classify samples across all cohorts tested (Fig. 1c).

We next confirmed that both the ViP signatures are induced in acute KD (at presentation,  $\leq 10$  days of illness) compared to convalescent KD (day 289–3240 of illness) in a large new cohort of consecutive patients ( $n = 105$ ) who were diagnosed with the disease prior to the onset of the COVID-19 pandemic (Cohort 1; Supplementary Data 1) (Fig. 1d). Again, the sViP signature outperformed the ViP signature in sample classification (ROC AUC 0.91 vs. 0.74). In an independent cohort (Cohort 2,  $n = 20$ , Supplementary Data 1; Fig. 1e) prospectively enrolled in the current study after the onset of the COVID-19 pandemic, the ViP signatures could differentiate the acute from subacute (~10–14 d after discharge; ~day 17–25 of illness) KD samples. As before, the 20-gene sViP signature outperformed the 166-gene ViP signature.

Finally, we tested the association between sViP signatures and markers of disease severity. Because CAA diameter is a predictor of coronary sequelae (thrombosis, stenosis, and obstruction)<sup>23,24</sup> and subsequent major adverse cardiac events (unstable angina, myocardial infarction, and death<sup>25</sup>), we used the development of coronary artery aneurysms (CAA) as a marker of disease severity. We found that both ViP signatures differentiated acute KD with giant CAAs (defined as a z-score of  $\geq 10$  or a diameter of  $\geq 8$  mm<sup>26,27</sup>) from convalescent KD samples (ROC AUC 0.95 and 0.97 for ViP/sViP signatures, respectively; Fig. 1f). The ViP signature effectively subclassified acute KD patients with giant aneurysms (CAA-giant) from those with either no aneurysms

(CAA—;  $p$  value = 0.0027) or small aneurysms (CAA-small;  $p$  value = 0.0013). Similarly, the sViP signature effectively classified acute KD patients with giant aneurysms (CAA-giant) from those with no aneurysms (CAA—;  $p$  value 0.033). Such an analysis was not possible in Cohort 2 (Supplementary Data 1) because of the smaller cohort size and absence of subjects with giant CAAs.

We conclude that ViP signatures are induced in acute KD, and track disease severity, i.e., risk of developing giant CAAs, much like we observed previously in the setting of adult COVID-19<sup>18</sup>. Because ViP signatures represent the host immune response to diverse pathogens (Fig. 1b), upregulation of ViP signatures in KD is consistent with the hypothesis that KD is triggered by multiple infectious triggers<sup>7,28,29</sup>, some of which may be viral in nature<sup>30–32</sup>.

**Comparison of patients with MIS-C and Kawasaki disease.** Ten children were included who met the CDC definitions for MIS-C, with detectable anti-SARS-CoV-2 nucleocapsid IgG antibodies [Abbott Architect™] and undetectable virus by polymerase chain reaction (PCR; see Table 1). The MIS-C and KD cohorts had notable differences. Although sex and ethnicity were not different, the median age was higher (8.8 years) in the MIS-C cohort than in KD (Table 1), which is in keeping with our original report describing this syndrome in June 2020<sup>2</sup>. Left ventricular ejection fraction (LVEF) was reduced in the MIS-C cohort ( $p = 0.006$ ), consistent with multiple prior reports<sup>33–35</sup>. While all patients had evidence of a marked inflammatory state, the MIS-C cohort had significant cytopenias, including low total WBC, absolute lymphocyte, absolute eosinophil, and platelet counts, with elevation of C-reactive protein level significantly above those observed in the KD cohort (Table 1). Most patients (90%) received intravenous immunoglobulin (IVIg) and 70% were treated with intravenous corticosteroids. One patient received anakinra, and three received infliximab. All patients made a full recovery. In all cases, blood was collected for serum before the initiation of any treatments.

**ViP/sViP signatures place MIS-C and KD on the same host immune continuum, but MIS-C as farther along the spectrum than KD.** We next analyzed whole blood-derived transcriptome and serum cytokine arrays in the current cohort of subjects with KD (Cohorts 2 and 4) and MIS-C (Fig. 2a). When MIS-C and acute KD groups were each compared to the control (subacute KD) samples, both ViP (Fig. 2b) and sViP (Fig. 2c) signatures were found to be induced at significantly higher levels in MIS-C samples compared to acute KD. However, when MIS-C and acute KD were compared to each other, we found that the ViP



**Table 1 Characteristics of patients with Kawasaki disease (KD) and MIS-C analyzed in this study.**

Demographic and clinical parameters	KD		MIS-C	p
	CAA- (n = 10)	CAA+ (n = 10)	MIS-C (n = 10)	
Age, yrs <sup>a</sup>	2.2 (1.8–3.7)	1.8 (1.2–3.5)	8.8 (5.7–11.1)	0.0002
Illness day <sup>b</sup>	6 (5–7)	6 (5–7)	4 (3–4)	NS
Male, n (%)	6 (60)	8 (80)	6 (60)	NS
Ethnicity, n (%)				
Asian	1 (10)	1 (10)	0	NS
African American	1 (10)	1 (10)	2 (20)	
White	3 (30)	2 (20)	1 (10)	
Hispanic	4 (40)	3 (30)	6 (60)	
>2 races	2 (20)	5 (50)	1 (10)	
Zmax	1.0 (0.7–1.2)	3.1 (2.7–3.3)	1.9 (1.5–2.2)	0.0002
LVEF, median (IQR, range)	67 (65–68, 56–76)	70 (62–75, 56–79)	58 (55–62, 31–65)	0.006
Lab data				
WBC, 10 <sup>3</sup> /μL	12.6 (11.6–18.3)	15.8 (13.0–18.5)	11.1 (5.5–11.7)	0.01
ANC, /μL	9159 (6904–106830)	11319 (8896–12539)	8172 (4437–10071)	NS
ALC, /μL	2576 (1739–3658)	2924 (2379–5075)	939 (803–1019)	0.0002
AEC, /μL	184 (104–529)	378 (322–588)	122 (0–231)	0.04
ZHgb	−1.4 (−2.3–0.1)	−1.5 (−1.8 to 0.6)	−2.2 (−2.7 to 0.7)	NS
PLT, 10 <sup>3</sup> /mm <sup>3</sup>	330 (278–396)	363 (338–396)	177 (106–228)	0.002
ESR, mm/h	55 (30–64)	68 (58–76)	44 (36–59)	NS
CRP, mg/dL	6.1 (4.8–12.2)	5.9 (4.0–8.0)	21 (19.6–23.8)	0.01
BNP, pg/mL	ND	ND	33 (19.7–23.6)	NA
Ferritin, ng/mL	ND	ND	323 (223–960)	NA
Troponin, ng/mL	ND	ND	0.02 (0.01–0.13)	NA
D-dimer, μg/mL	ND	ND	1.94 (1.24–2.40)	NA
SARS-CoV-2 testing				
PCR positive	ND	ND	1 (10)	NA
IgG positive <sup>c</sup>	ND	ND	10 (100)	NA
Treatment, n (%)				
IVIg×1	10 (100)	10 (100)	8 (80)	NS
IVIg×2	0	1 (10)	1 (10)	NS
Infliximab	4 (40)	10 (100)	3 (30)	0.003
Anakinra	0	1 (10)	1 (10)	NS
Cyclosporine	0	1 (10)	0	NS
Steroids	0	0	7 (70)	0.0001

Table displays the demographic, clinical and laboratory parameters collected on the KD and patients with MIS-C enrolled into the study. Zmax Maximum Z score (internal diameter normalized for body surface area) for the right and left anterior descending coronary arteries. Laboratory data are pre-treatment. Troponin was measured in nine patients with MIS-C. D-dimer was measured in eight patients with MIS-C. p-values were calculated by Kruskal-Wallis test for continuous variables among three groups and Chi-test for categorical variables.

LVEF left ventricular ejection fraction, WBC white blood cell count, PLT platelets, AEC absolute eosinophil count, ANC absolute neutrophil count, ALC absolute lymphocyte count, CRP C-reactive protein, ESR erythrocyte sedimentation rate, BNP brain natriuretic peptide, ZHgb hemoglobin concentration normalized for age, NS not significant, NA not applicable, ND not done.

<sup>a</sup>Median (Interquartile range (IQR) unless specified.

<sup>b</sup>Illness Day 1 = first day of fever.

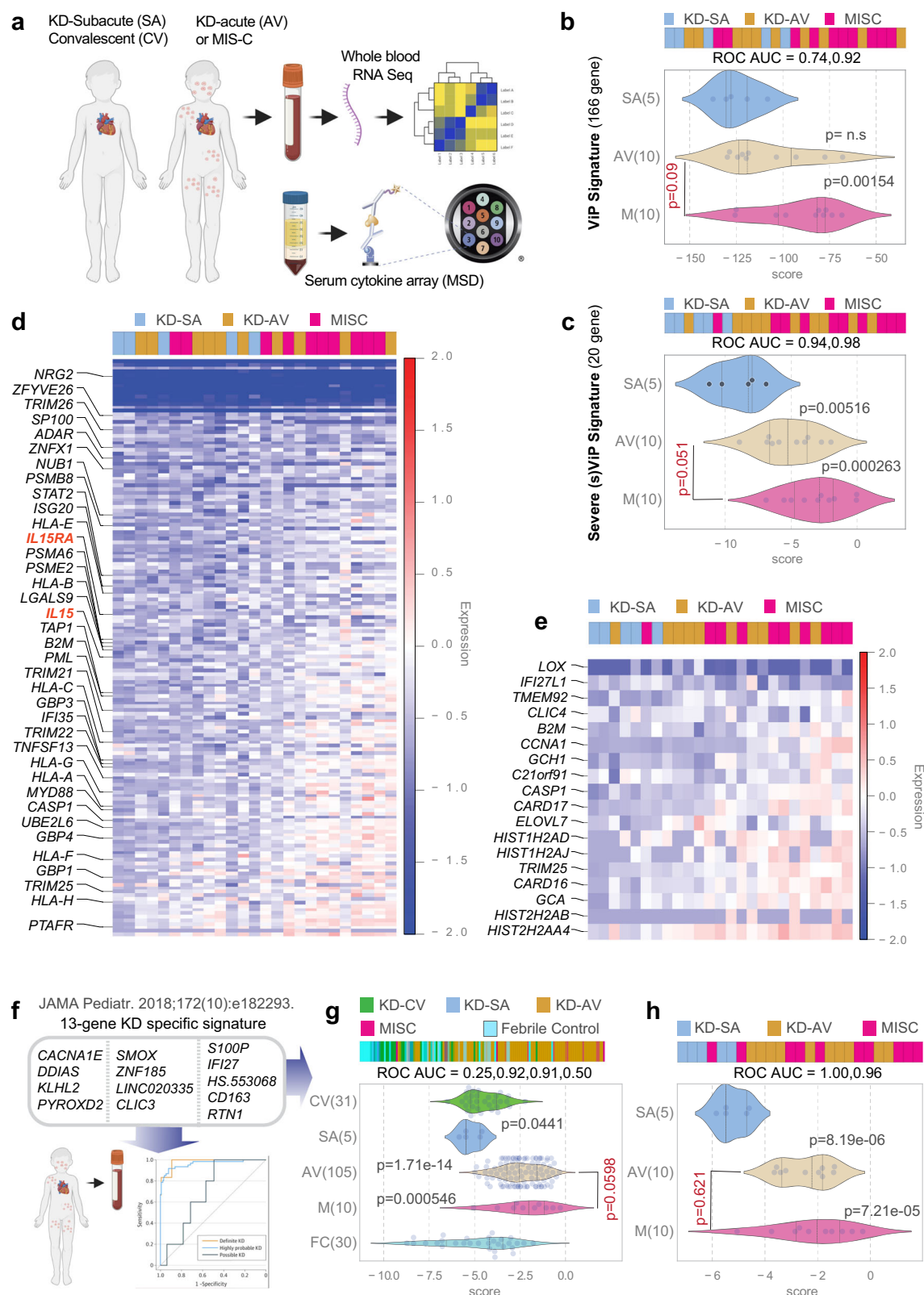
<sup>c</sup>SARS-CoV-2 nucleocapsid IgG positive n = 9 and SARS-CoV-2 peptide array n = 1.

signatures could not distinguish between these samples, indicating that both conditions share a similar host immune response. Heatmaps of patterns of expression (Fig. 2d, e) demonstrate that most of the individual genes contributed to the elevated ViP and sViP signatures observed in MIS-C samples. These genes included *IL15* and *IL15RA* (highlighted in red; Fig. 2d), both components within the cytokine pathway that was previously demonstrated to be consistently elevated in the lungs of patients with fatal COVID-19 and in SARS-CoV-2 challenged hamsters<sup>18</sup>.

Taken together, these analyses led to two key conclusions: (i) that the host immune response, as detected in a qualitative manner using the ViP signatures, is similar in KD and MIS-C and has a *IL15/IL15RA* shared component; (ii) that the degree of such host immune response, as measured quantitatively using the ViP signature scores, is more intense in MIS-C than KD. These findings are consistent with the fact that MIS-C is a host immune response to SARS-CoV-2 exposure, and we previously showed that the interaction of viral spike protein with the host entry receptor, ACE2 is critical for the induction of ViP signatures<sup>18</sup>. Findings are also in keeping with prior work<sup>36</sup> showing that serum levels of *IL15* is significantly elevated in acute KD, ~10-fold

compared with subacute-KD and normal controls, and that such increase correlated with the concomitant increase in serum TNFα.

**A KD-specific signature independently confirms that KD and MIS-C are syndromes on the same host immune response continuum.** To circumvent an over-reliance on one set of signatures (i.e., ViP/sViP), we next analyzed a KD-specific 13 transcript diagnostic signature<sup>37</sup> that was previously shown to be effective in distinguishing children with KD from all other febrile conditions. During validation, the 13-transcript signature mirrored the certainty of clinical diagnosis, i.e., it differentiated definite, highly probable, and possible KD from non-KD with ROC AUCs of 98.1% (95% CI, 94.5–100%), 96.3% (95% CI, 93.3–99.4%), and 70.0% (95% CI, 53.4–86.6%), respectively (Fig. 2f). Unlike the ViP signatures, which has a typical enrichment of interferon and cytokine pathways with a prominent presence of *IL15/IL15RA*, the KD-signature is comprised of a set of non-overlapping genes, some of which relate to major central hubs within the tumor necrosis factor (TNFα) and interleukin 6



(IL6) pathways<sup>37</sup>. When we applied this signature to the historic cohort 1 (Fig. 2g) and to our current cohort (Cohort 2; Fig. 2h), we found that the KD-specific 13 transcript signature could not distinguish between MIS-C and KD in either cohort. Furthermore, a correlation test demonstrated that the two non-overlapping signatures, sViP and KD-13, both of which are significantly induced in KD and MIS-C (Fig. 2c) are independent of

each other (Supplementary fig. S1). This suggests that these two signatures reflect two fundamentally distinct and unrelated biological domains within the host immune response; whether their diagnostic/prognostic abilities may have an additive benefit remains to be explored.

The similar extent to which KD and MIS-C induced the KD-13 signature in two independent cohorts further supports our

**Fig. 2 A KD-specific 13 transcript signature<sup>37</sup> shows that KD and MIS-C are indistinguishable, but ViP/sViP signatures place MIS-C as farther along the spectrum than KD.** **a** Schematic displays the workflow for patient blood collection and analysis by RNA Seq (this figure) and cytokine array by mesoscale (Figs. 4 and 5). **b, c** Bar (top) and violin (bottom) plots display the classification of blood samples that were collected during acute (AV) and sub-acute (SA; ~10–14 days post-discharge) visits of KD subjects and from patients diagnosed with MIS-C. The *p* value for comparison between acute KD (AV) and MIS-C (M) is displayed in red font. **d, e** Heatmaps display the patterns of expression of the 166 genes in ViP (**d**) and 20 gene sViP (**e**) signatures in the KD and MIS-C samples. The only cytokine–receptor pair within the signature, i.e., *IL15/IL15RA*, are highlighted on the left in red font in (**d**). **f** Schematic displays the 13-transcript whole blood signature (no overlaps with ViP signature genes) previously demonstrated to distinguish KD from other childhood febrile illnesses<sup>37</sup>. **g** and **h** Bar (top) and violin (bottom) plots display the classification of blood samples that were collected during acute (AV) and convalescent (CV) visits from two independent KD cohorts (**g**: Historic Cohort 1; **e**: Prospective Cohort 2) using 13-transcript KD signature. FC, febrile control. See also Supplementary Fig. S1 for co-dependence analysis of ViP and KD-13 signatures. Welch's two sample unpaired two-sided *t*-test is performed on the composite gene signature score to compute the *p* values. In multi-group setting each group is compared to the first control group and only significant *p* values are displayed. The *p* value for comparison between acute KD (AV) and MIS-C (M) is displayed in red font.

observation with ViP/sViP signatures that KD and MIS-C share fundamental aspects of host immune response with each other. That KD and MIS-C samples share ViP/sViP signatures with COVID-19 implies that the three diseases represent distinct clinical states on the same host immune response continuum.

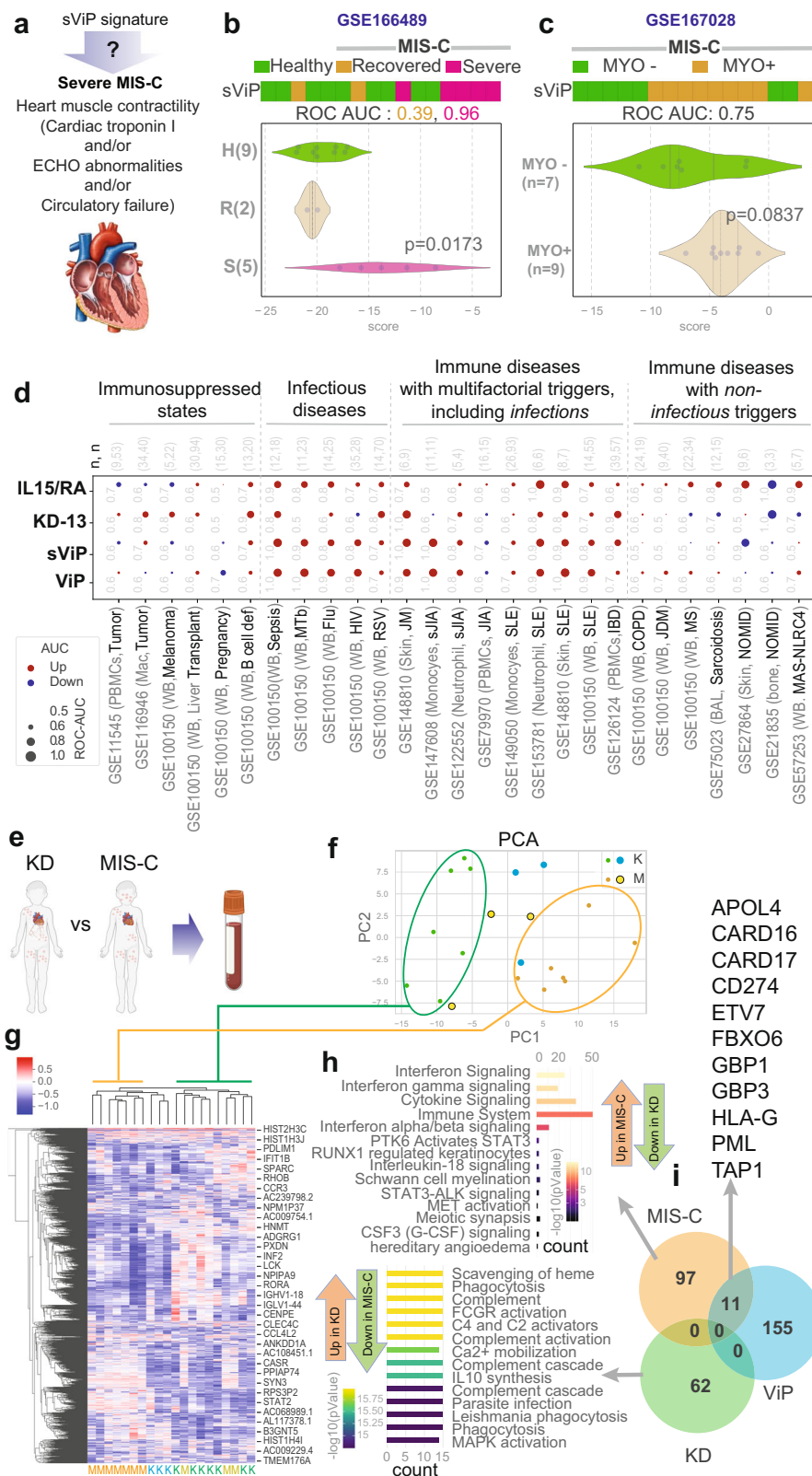
**The sViP signature can recognize severe form of MIS-C that presents with myocardial dysfunction.** Next, we asked if the sViP signature can track disease severity in MIS-C. Because of the limited number of 'severe' cases in our MIS-C cohort, we prospectively analyzed two recently accessible MIS-C cohorts (GSE166489<sup>15</sup> and GSE167028<sup>38</sup>). While both datasets analyzed PBMCs from MIS-C subjects, and both studies used the presence of myocardial dysfunction as basis for severe disease, each study used a slightly different criterion for classification of disease severity (Fig. 3a). de Cevins et al.<sup>38</sup>, classified MIS-C as severe when the patients presented with elevated cardiac troponin I and/or altered ventricular contractility by echocardiography, and clinical signs of heart failure requiring ICU support. Ramaswamy et al.<sup>15</sup>, classified MIS-C as severe if they were critically ill, with cardiac and/or pulmonary failure. In both cohorts, sViP was able to classify severe MIS-C (with myocardial dysfunction; MYO+) from mild-moderate disease (who recovered or presented without myocardial dysfunction; MYO-) (Fig. 3b, c); while the *p* value was significant in GSE166489 (Fig. 3b), a similar trend was conserved in GSE167028<sup>38</sup> (Fig. 3c). These findings show that the sViP signature can identify severe MIS-C who are at risk to develop myocardial dysfunction, just as it did in the case of KD subjects who are at risk of developing giant CAAs (Fig. 1f) and similar to its prior performance in identifying adults with COVID-19 who are at risk of respiratory failure, mechanical ventilation, prolonged hospitalization and/or death<sup>18</sup>.

Taken together with the prior findings, we conclude that the 20-gene sViP signature captures a core set of genes that are expressed in the setting of an overzealous (prolonged or intense, or both) host immune response in all three diseases—KD, MIS-C (this work) and COVID-19<sup>18</sup>—despite the fact that each present with distinct clinical features of severity.

Because all three conditions represent diseases of the immune system that share an 'infectious trigger', we asked if the ViP/sViP signatures are also induced in the setting of other diseases of the immune system. To this end, we analyzed numerous publicly available datasets, ranging from immunosuppressed states (as negative control), infectious diseases (both viral and bacterial; as positive control), and autoimmune diseases (Fig. 3d), and assessed the ability of ViP/sViP signatures to classify control and diseased samples in each dataset. Because the ViP/sViP signatures are able to detect the shared core fundamental host immune response in cell/tissue agnostic manner<sup>18</sup>, we tested diverse samples ranging from whole blood to bronchoalveolar lavage fluid (Fig. 3d). The ViP/sViP signatures performed as

anticipated in the negative and positive control datasets, i.e., neither signature was induced in immunosuppressed conditions, e.g., malignancies, pregnancy, post-transplant immunosuppression, but both were induced in infectious diseases, e.g., sepsis, HIV, RSV, and tuberculosis (left; Fig. 3d). In the case of the autoimmune diseases, the ViP/sViP signatures were induced in some, but not others. The signatures were induced in those conditions that have multifactorial triggers, including potential contributions from infections; for example, mechanistic studies have identified viral link in many of them (EBV-linked autoimmune diseases<sup>39</sup> such as JIA, SLE, IBD). The signatures were not induced in other conditions where the disease triggers remain mysterious (e.g., sarcoidosis) or where the disease is driven by specific mutations, e.g., Neonatal onset multisystem inflammatory disease (NOMID) that is due to mutant *NLR3* and macrophage activation syndrome (MAS) that is due to mutant *NLR4*. These findings lend further support to our finding that ViP/sViP signatures are induced and perform well to identify severe MIS-C, which shares infection as a trigger, much like KD and COVID-19. Intriguingly, numerous infectious and autoimmune diseases shared the *IL15/IL15RA*-centric cytokine response, which is in keeping with prior observations<sup>40</sup>.

**Whole blood transcriptomes and cytokine panels reveal subtle differences between MIS-C and KD.** We next compared the whole blood transcriptomes from KD and MIS-C subjects (Fig. 3e) using a more conventional approach, which involved principal component analysis (PCA; Fig. 3f) in conjunction with hierarchical agglomerative clustering. The PCA analysis showed that 7 of 10 samples in both KD and MIS-C groups formed distinct clusters (circles in Fig. 3f), whereas 3 in each group were outside their respective clusters (yellow and cyan, Fig. 3f). Because agglomerative clustering is the most common type of hierarchical clustering used to group objects in clusters based on their similarity, we next sought this approach to assess grouping of KD and MIS-C samples (Fig. 3g) which also revealed two distinct clusters like PCA. Differential expression analysis was performed with the 7 MIS-C and 7 KD that formed distinct clusters in PCA and hierarchical clustering (Supplementary Data 3). Reactome pathway analyses of the genes upregulated in MIS-C revealed interferon and cytokine signaling (Fig. 3h, top), whereas the genes downregulated in MIS-C mostly enriched pathways concerning the complement cascade and phagocytosis, among others (Fig. 3h, bottom). When we analyzed the overlap between the 166-gene ViP signature and the up- or downregulated list of DEGs in MIS-C, we found that 11 genes overlapped between the ViP signature and upregulated genes in MIS-C (Fig. 3i). However, there was no overlap between the ViP signature and the downregulated genes. These analyses further emphasize the similarities between KD and MIS-C with more extreme features in MIS-C.



We next analyzed a set of 10 serum cytokines using meso scale discovery electrochemiluminescence (MSD-ECL) ultra-sensitive biomarker assay. A panel of 10 target cytokines was prioritized based on a review of the literature for the reported presence and/or relevance of each in either KD and/or MIS-C. An unsupervised clustering of just these 10 cytokines was sufficient to differentiate acute KD and MIS-C from one-year convalescent KD samples

(Fig. 4a; Cohorts #2 and #3, Supplementary Data 1, Supplementary Data 2); the convalescent samples served as baseline ‘healthy’ controls in this case. Regardless of their degree of elevation in the acute setting, all cytokines were virtually undetectable in convalescent samples (Supplementary Fig. S2, Supplementary Data 2). While most cytokines were induced indistinguishably in acute KD and MIS-C (Fig. 4a, b, top), notable exceptions were



**Fig. 3 Performance of ViP/sViP signatures on independent MIS-C datasets and on diverse tissues and in diverse diseases of the immune system.** **a–c** Severe (s)ViP signature can classify severe MIS-C based on in two independent studies ([GSE166489](#)<sup>15</sup> and [GSE167028](#)<sup>38</sup>). Schematic in **a** summarizes the definition of severe MIS-C. **b, c** Classification of blood samples in two cohorts of MIS-C subjects, based on the need for ICU management due to the presence (MYO+) or recovery in the absence (R or MYO–) of myocardial dysfunction using sViP signature. Welch's two sample unpaired two-sided t-test is performed to compute the *p* values. **d** Bubble plots of ROC-AUC values (radii of circles are based on the ROC-AUC) and the direction of gene regulation (Up, red; Down, blue) in publicly available datasets using 4 gene signatures: the 166-gene ViP signature, the 20-gene sViP signature, the KD-13 signature, and finally the *IL15/IL15RA* composite score. Numbers on top of bubble plots indicate number (*n*) of control vs. disease samples in each dataset. Abbreviations: PBMCs peripheral blood mononuclear cells, Mac macrophages, WB whole blood, MTb *M. tuberculosis*, Flu Influenza, HIV human immunodeficiency virus, RSV respiratory syncytial virus, JM juvenile myositis, sjia systemic juvenile idiopathic arthritis, SLE systemic lupus erythematosus, IBD Inflammatory bowel disease, COPD chronic obstructive pulmonary disease, JDM juvenile dermatomyositis, MS multiple sclerosis, BAL bronchoalveolar lavage, NOMID neonatal onset multisystem inflammatory disease, MAS macrophage activation syndrome, NLR4 NLR Family CARD Domain Containing 4. **e** Schematic showing the experimental design for studying differentially expressed genes (DEGs) in between KD and MIS-C subjects. **f, g** PCA (**f**) and a clustered heatmap analysis (**g**) of KD (green, **f**) and MIS-C (orange, **g**) samples are shown based on top 2242 genes according to mean absolute deviation identified using StepMiner algorithm<sup>88</sup>. Source data are provided. **h** Reactome pathway analysis of the DEGs between seven KD and seven MIS-C subjects in **f** (marked on the PCA). **i** Venn diagram between 166-gene ViP signature against the DEGs. Number of genes are indicated for each group in the Venn diagram. 11 overlapping genes between ViP signature and up-regulated in MIS-C are listed at the top.

TNF $\alpha$ , IFN $\gamma$ , IL10, IL8 and IL1 $\beta$ , all of which were elevated to a greater extent in MIS-C compared to KD (Fig. 4b, bottom), either significantly (TNF $\alpha$ , IFN $\gamma$ ; Fig. 4b) or trended similarly, but failed to reach statistical significance (IL10, IL1 $\beta$ , IL8). Gene set enrichment analyses (GSEA) on the transcriptomic dataset for each of the differentially expressed cytokines (Fig. 4c) showed that the gene sets for those pathways were also induced in MIS-C at levels significantly higher than KD (Fig. 4d–h).

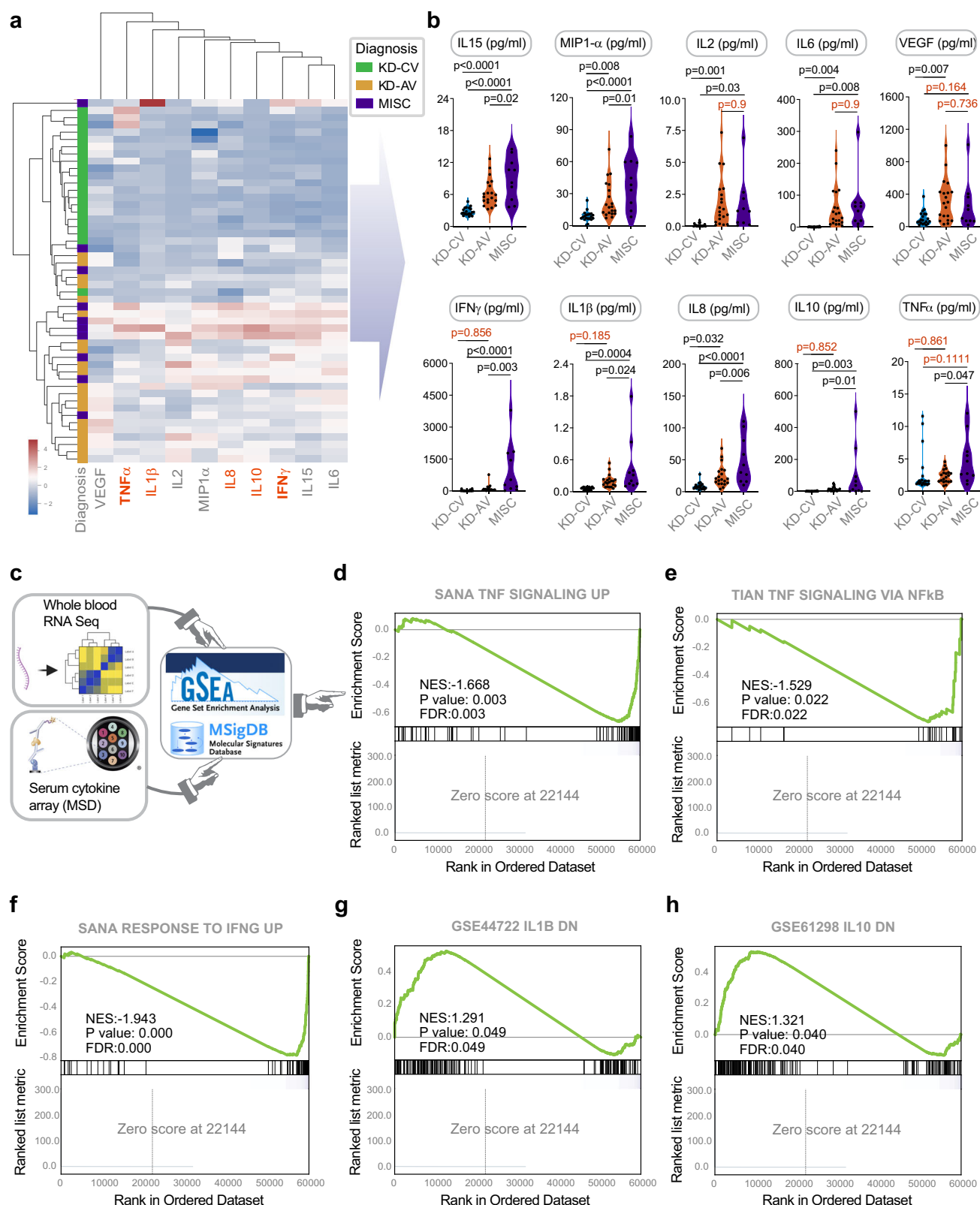
Taken together, these findings suggest that FDA-approved therapeutics targeting TNF $\alpha$  and IL1 $\beta$  pathways may be beneficial to treat MIS-C. The IL-1 receptor is expressed in nearly all tissues and its antagonism by anakinra, a recombinant form of IL-1Ra<sup>41</sup>, prevents receptor binding of either IL-1 $\alpha$  or IL-1 $\beta$ . Similarly, infliximab, a chimeric antibody to TNF $\alpha$ , has been repurposed for COVID-19<sup>42–44</sup>, and our analyses suggest that this agent holds promise as a treatment for MIS-C.

**Integrated analyses of ViP/sViP signatures, cytokine profile, and clinical laboratory parameters reveal unique features of MIS-C and indicators of disease severity.** We next sought to understand how similar host cytokine responses can trigger two distinct clinical syndromes, and how such responses may drive features of clinical severity. To this end, we first carried out an agglomerative hierarchical clustering of the MIS-C and acute KD samples using both cytokine profiling (MSD) and clinical/laboratory parameters. This analysis, coupled with correlation tests (Supplementary Fig. S3) revealed several intriguing observations: (i) visualization by heatmap showed that compared to KD, MIS-C patients had higher cytokine levels and more severe pancytopenia (Fig. 5a); (ii) although platelet counts (PLT), but not absolute eosinophil counts (AEC) were significantly reduced in MIS-C compared to KD (Fig. 5b), there was a strong positive correlation between PLT and AEC in MIS-C, but not KD (Fig. 5c; left) and strong negative correlations of PLT with IL15 in both KD and MIS-C (Fig. 5c; right) and with MIP1 $\alpha$  in MIS-C, but not KD (Fig. 5d); (iii) this is consistent with the fact that IL15 and MIP1 $\alpha$  were found to have a strong positive correlation in MIS-C, but not KD (Fig. 5e). These findings suggest that MIS-C has key distinguishing features of thrombocytopenia and low eosinophil counts, and that both features are negatively correlated with the serum levels of IL15, a key feature of the ViP signature. These findings also held true when we analyzed the two clinical parameters, PLT and AEC, against ViP/sViP signatures, as well as a specific *IL15/IL15RA* composite transcript score from whole blood RNA Seq dataset. We found that PLT and AEC negatively correlated with ViP (Fig. 5f; top), sViP (Fig. 5f; middle) and a *IL15/IL15RA* composite score (Fig. 5f; bottom) in MIS-C, but

such correlations were restricted only to PLT in acute KD. These findings indicate that MIS-C, but not KD, has at least two distinct and interrelated clinical features, thrombocytopenia and eosinopenia, that appear to be related to the degree of induction of ViP signatures and a IL15-predominant cytokine induction. Findings also suggest that MIP1 $\alpha$  is a key contributor to the immune response in MIS-C and that its levels are closely and positively related to the levels of IL15.

These findings reveal key similarities and differences among MIS-C, KD and COVID-19. Thrombocytopenia, which was more pronounced in MIS-C and correlated significantly with IL15 and *IL15/IL15RA* composite transcript score in both KD and MIS-C, has also been reported in COVID-19 and postulated because of various mechanisms<sup>45–49</sup>. In the case of KD, thrombocytopenia has been found to be associated with disease severity<sup>50</sup>. Similarly, in the case of COVID-19, a large meta-analysis confirmed that ~12% of hospitalized patients have thrombocytopenia, which represents a sign of disease severity and poor outcomes<sup>45</sup>. Thrombocytopenia carried a 3-fold enhanced risk of a composite outcome of intensive care unit admission, progression to acute respiratory distress syndrome, and mortality (odds ratio [OR], 3.49; 95% CI, 1.57–7.78), and a subgroup analysis confirmed a significant association with mortality (OR, 7.37; 95% CI, 2.08–26.14). Eosinopenia appears to be a notable shared feature between MIS-C and COVID-19<sup>51</sup>, but not KD. These findings are consistent with the fact that KD is known to present with higher (not lower) eosinophil counts, Th2 cytokines IL-4, IL-5, and eosinophil cationic protein (ECP)<sup>52–56</sup>. As in the case of thrombocytopenia, persistent eosinopenia after admission correlated with COVID-19 severity and low rates of recovery<sup>57</sup>.

**ViP/sViP signatures track the severity of two distinct cardiac phenotypes in MIS-C and KD.** We next analyzed the relationship between ViP signatures and the prominent and unique cardiac phenotype in MIS-C reported by others<sup>33–35</sup> and observed also in our cohort (Fig. 6a), i.e., a significantly reduced LVEF that can present with cardiogenic shock necessitating inotropic support. We found that sViP signature scores, but not ViP or *IL15/IL15RA* composite scores correlate significantly with LVEF (Fig. 6b–d), indicating that LVEF may belong to the domain of clinical indicators of disease severity in MIS-C (alongside platelets and AEC), but it may not be directly related to the *IL15*-centric cytokine signaling. In KD, the ViP and sViP signatures were tested earlier (Fig. 1f) and found to distinguish patients with giant CAA from convalescent samples with ROC AUC > 0.95. A *IL15/IL15RA* composite score performed similarly in distinguishing those samples (Fig. 6e). We hypothesized that



the ViP signatures may be related to two distinct cardiac phenotypes in severe disease: the signatures in KD may signify the nature of the vasculitis that drives the formation of CAAs, whereas the same signature in MIS-C may signify the degree of cardiomyopathy that impairs contractility (Supplementary Fig. S4a). Because we were unable to acquire cardiac tissues from MIS-C-related autopsies, we carried out immunohistochemical

analyses on cardiac tissues from a case of fatal KD. We found that both IL15 and IL15RA were expressed in the cardiomyocytes and coronary arterioles amidst extensive fibrosis, as detected using Masson's trichrome stain (Supplementary Fig. S4b).

Together, these findings suggest that the *IL15/IL15RA* induction we see in COVID-19, KD and MIS-C may have distinct sources and/or target end organs: We previously showed

**Fig. 4 Serum cytokine arrays and whole blood transcriptomes reveal the severity and nature of the cytokine storm in MIS-C that distinguishes it from KD.** **a** Heatmap displays the results of unsupervised clustering of sub-acute and acute KD (KD-SA, KD-AV;  $n = 10$  each) and MIS-C ( $n = 10$ ) subjects using the cytokine profiles determined by mesoscale (MSD). Red = cytokines differentially expressed between MIS-C and KD. See also Supplementary Fig. S2 for violin plots for individual cytokines. **b** Source data are provided as a Supplementary Data 2. Violin plots display the shared (top panels; IL15, MIP1 $\alpha$ , IL2, IL6 and VEGF) and distinct (bottom panels; IFN $\gamma$ , IL1 $\beta$ , IL8, IL10, and TNF $\alpha$ ) features of the cytokine storm in MIS-C vs. KD subjects. Statistical significance was determined by one-way ANOVA followed by Tukey's test for multiple comparisons. **c** Schematic shows the process used to integrate serum cytokine array results with whole blood RNA Seq data; cytokines that were differentially expressed in MIS-C were used to inform GSEA of the corresponding pathways. **d–f** Gene set enrichment analysis (GSEA pre-ranked analysis) of three pathways derived from MSigDB: SANA TNF SIGNALING UP (**d**), TIAN TNF SIGNALING VIA NF $\kappa$ B (**e**), and SANA RESPONSE TO IFNG UP (**f**) demonstrate the significance of TNF (**d, e**) and IFNG (**f**) pathway activation in MIS-C. **g, h** Down-regulated genes after IL1B (**g**) and IL10 (**h**) stimulation were derived from differential expression analysis of GSE44722 ( $n = 269$  genes), and GSE61298 ( $n = 208$  genes) respectively. GSEA pre-ranked analysis to test the significance of IL1B and IL10 pathway is performed like panels **d–f** using the down-regulated genes. GSEA pre-ranked analysis computes nominal pvalue and FDR using an empirical phenotype-based permutation test procedure. No adjustments were made for multiple comparisons because of single hypothesis testing. Source data are provided as a Source Data file.

prominent induction of *IL15/IL15RA* in the lung alveoli of fatal COVID-19 patients<sup>18</sup>, and here we show it in the coronary arteries and cardiomyocytes in KD. However, the latter claim has some notable caveats: (i) it is based on one fatal KD and no MIS-C tissue, hence, may not be representative of what happens in every KD heart; (ii) the induction of *IL-15/IL-15R* may be consistent with systemic hyperinflammation, but immune cell infiltration in the heart was not evident; (iii) because IL-15 has been implicated in a wide range of cardiovascular diseases<sup>58</sup> and was previously shown to be implicated in cardiomyocyte survival<sup>59</sup> during oxidative stress, alternative explanations other than hyperinflammation could be at play. Thus, further studies are required to determine if the findings in KD and MIS-C can be generalized and to determine what may be the impact of *IL-15/IL-15R* expression in the heart.

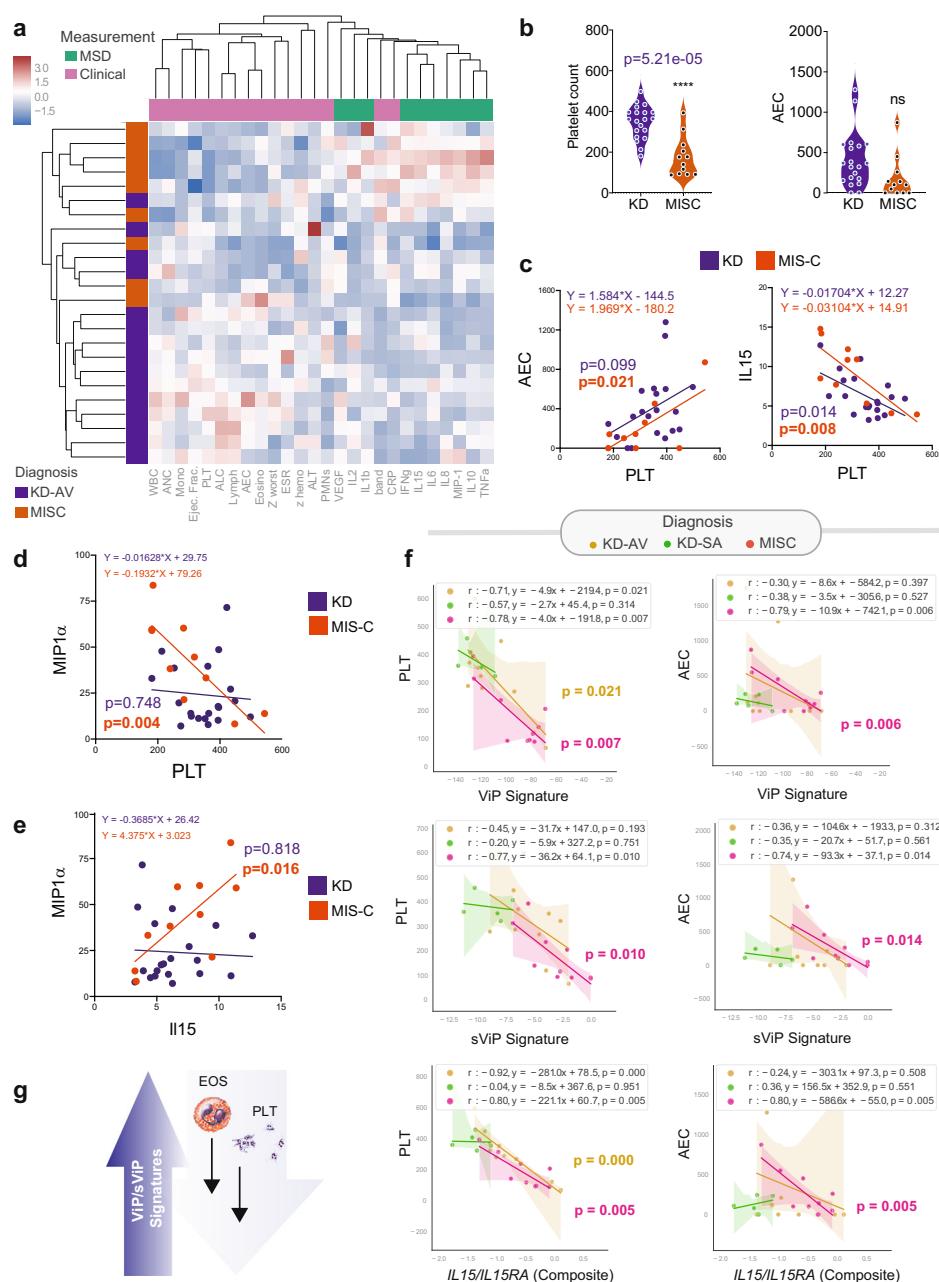
## Discussion

Using a combination of publicly available KD datasets and newly recruited cohorts of KD and MIS-C subjects (summarized in Fig. 7a) and a set of gene signatures we report an unexpected discovery regarding the host immune response in these diagnoses. Our findings show that two distinct clinical syndromes, KD, which predates the current pandemic by 6 decades, and the novel COVID-19, share a similar profile of host immune response. The same host immune response is seen also in MIS-C, a new disease that co-emerged with COVID-19, which has some overlapping features with KD (i.e., clinical presentation, pediatric, etc.), and yet is an immune response to the virus that causes COVID-19 (Fig. 7b). We assessed the quality and intensity of the host immune response in these syndromes with a powerful and unbiased computational tool, the ViP signatures<sup>18</sup>. Challenging to our previous understanding of MIS-C as post-infectious syndrome, recent studies have revealed a prolonged presence of viral replication<sup>60</sup> and dendritic cell exhaustion caused by the persistence of the antigen<sup>61</sup>. The use of this computational tool was rationalized because MIS-C is triggered by exposure to a virus and therefore, the induction of the ViP signatures in the acute phase, followed by their reduction during convalescence was anticipated. Previously we had demonstrated the usefulness of the ViP signatures to define and measure the host immune response in COVID-19, identify the site/source of the cytokine storm, track disease severity, objectively formulate therapeutic goals and track the effectiveness of emerging drugs/biologics<sup>18</sup>. We now show that the same ViP signatures can objectively demonstrate the shared immunophenotypes between all three syndromes (COVID-19, KD and MIS-C), which features an upregulation of the *IL15/IL15RA* pathway. That a 13 transcript KD-specific signature that was previously shown to distinguish KD from other non-KD febrile illnesses<sup>37</sup> failed to distinguish KD from MIS-C, further confirmed that KD and MIS-C share similar molecular

markers of disease and hence, are fundamentally similar at the molecular level (summarized in Fig. 7b). Findings were confirmed also using conventional approaches (PCA and hierarchical agglomerative clustering followed by DEG analyses). Taken together, these results are in keeping with what has been observed by Consiglio et al.<sup>13</sup>, who found KD and MIS-C to be clustered together in a PCA analysis of plasma proteins. These findings suggest that the two clinical syndromes not just share common clinical features (e.g., rash, fever, etc.), but may also share proximal pathways of immunopathogenesis.

Despite the high-level broad similarities, the ViP signatures also helped identify key differences in clinical/laboratory parameters that may help distinguish MIS-C from KD. First, although the ViP signatures placed KD and MIS-C on the same host immune continuum, the degree of host immune response in MIS-C is significantly higher than KD by all measures tested (i.e., ViP, sViP, *IL15/IL15RA* and KD-13 signatures and direct measurement of serum cytokines). Higher ViP signatures in MIS-C tracked three major clinical and/or laboratory parameters (see Fig. 7b): (i) degree of thrombocytopenia in severe cases (all three diseases); (ii) eosinopenia (in COVID-19 and MIS-C, but not KD) and (iii) impaired cardiac contractility (unique to MIS-C; but not KD); (iii) an integrated analysis of serum cytokines and transcriptomics revealed that the proinflammatory MIP1 $\alpha$ , TNF $\alpha$ , and IL1 pathways are significantly induced in MIS-C compared to KD. In light of these findings, a rational approach to MIS-C treatment would be to combine the FDA-approved drugs anakinra<sup>41</sup> and infliximab<sup>42–44</sup>. In fact, during the preparation of this manuscript a new study has already shown favorable outcome in MIS-C with the use of Infliximab<sup>62</sup>. Furthermore, our findings are consistent with the recently released guidelines by the American College of Rheumatology for initial immunomodulatory treatment of MIS-C<sup>63</sup>; it is noteworthy that these guidelines were released while this work was under review.

Finally, our findings reveal a pattern of MIS-C-defining molecular features (*IL15/IL15RA*, MIP1 $\alpha$ , TNF $\alpha$ , and IL1 pathways) and clinical and laboratory parameters (thrombocytopenia, eosinopenia, and reduced myocardial function). For example, MIP1 $\alpha$  elevation shows strong correlations with clinical features of disease (low PLT, high IL15 and low AEC) in MIS-C, but not KD. This suggests two things—(i) that despite shared proximal proximal pathways of immunopathogenesis (i.e., *IL15/IL15RA*-centric cytokine storm), the immunopathogenesis of KD and MIS-C may diverge distally; and (ii) that *IL15/IL15RA*, eosinopenia and thrombocytopenia may be inter-related phenomena in the setting of infection and inflammation. Platelets, besides their role in hemostasis, they are known to participate in the interaction between pathogens and host defense<sup>64–66</sup>. Persistent thrombocytopenia carried higher mortality in sepsis<sup>67,68</sup>, and in COVID-19<sup>69,70</sup>. Our analysis revealed a direct and unusually

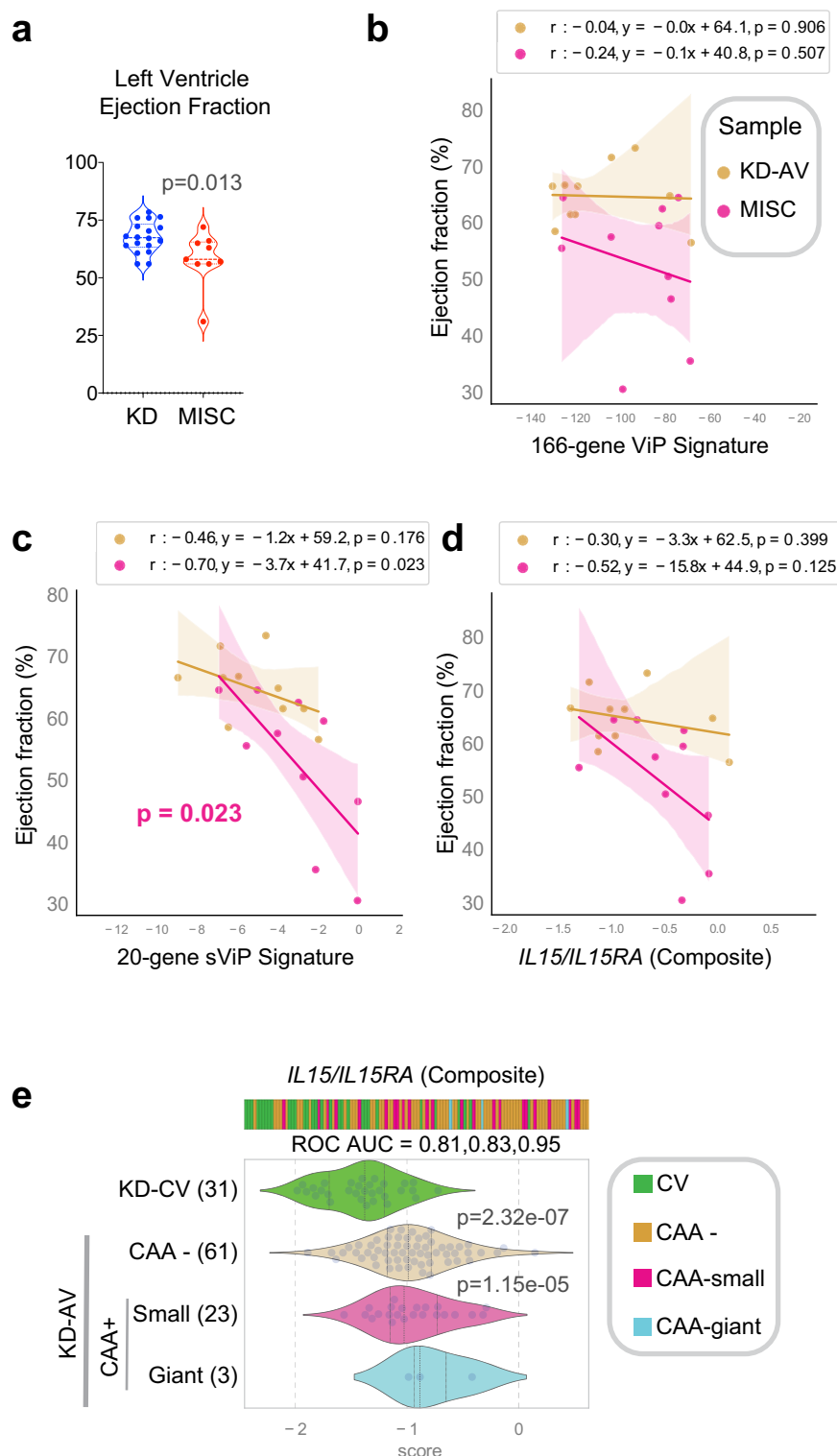


**Fig. 5 An integrated analysis of mesoscale (cytokine) data, ViP/sViP transcriptomic signatures and laboratory and clinical parameters reveals features that are unique to MIS-C.** **a** Heatmap displays the results of hierarchical agglomerative clustering of acute KD (KD-AV;  $n=10$ ) and MIS-C ( $n=10$ ) subjects using the cytokine profiles determined by mesoscale (MSD) and the laboratory features. Source data are provided as a Source Data file. **b** Violin plots display PLT (platelet) and AEC (absolute eosinophil counts) in KD and MIS-C (unpaired two-sided Student's  $t$ -test used to test significance). **c–e** Correlation test (two-sided test of the slope of the regression line compared to zero) between AEC and PLT (**c**; left) and IL15 and PLT (**c**; right), and MIP1α and PLT (**d**) and MIP1α and IL15 (**e**) are shown, and significance was calculated and displayed using GraphPad Prism 9. Significance: ns: non-significant, \*\*\*\* $p < 0.0001$ . See Supplementary Fig. S3 for all possible correlation tests between clinical and cytokine data in KD, MIS-C and COVID-19. **f** Correlation tests between PLT (left) or AEC (right) on the Y-axis and gene signature scores on the X-axis [either ViP (top), sViP (middle) or a IL15/IL15RA composite (bottom)] were calculated and displayed as scatter plots using python seaborn Implots with the  $p$ -values. The confidence interval around the regression line is indicated with shades. **g** Schematic summarizing the findings in MIS-C based on laboratory and RNA seq analysis.

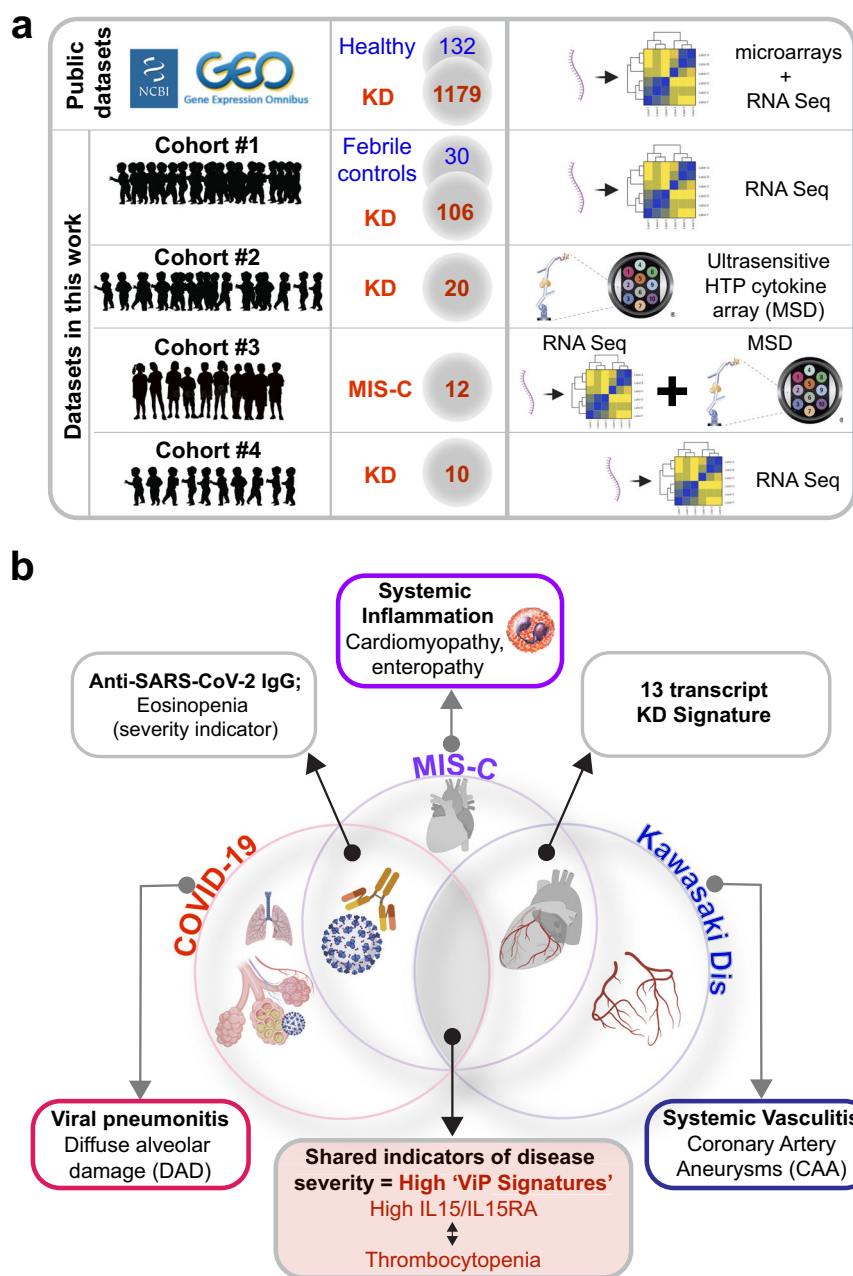
strong correlation between thrombocytopenia and eosinopenia in MIS-C. Eosinophils, on the other hand, as reviewed elsewhere<sup>71</sup>, have important antiviral properties, attributed to their granular proteins (e.g., eosinophil-derived neurotoxin, cationic protein) that display antiviral activities against single-stranded RNA viruses. Eosinophils can also support viral clearance<sup>72</sup>. Eosinopenia, in the setting of acute infection, has been found to be a

direct response to infectious stimuli<sup>73</sup>, TLR4 ligands and chemotactic factors<sup>74</sup>, and has been considered a reliable diagnostic marker of infection<sup>75</sup> in critically ill patients and a predictor of mortality<sup>75–77</sup>. Of relevance to the pediatric syndrome MIS-C, eosinopenia is encountered in about a 1/3rd of the pediatric COVID-19 subjects<sup>78</sup>. It is noteworthy that eosinopenia (defined as an eosinophil count  $<15$  cells/ $\mu$ L and an eosinophil





**Fig. 6 ViP/sViP signatures correlate with two distinct cardiac phenotypes in MIS-C and KD.** **a** Violin plots display the left ventricular ejection functions (LVEF) in KD and MIS-C patients. Statistical significance was determined by unpaired two-sided Student's *t*-test. **b–d** Correlation tests (two-sided test of the slope of the regression line compared to zero) between LVEF (Y-axis) and gene signature scores on the X-axis [either ViP (**b**), sViP (**c**), or a *IL15/IL15RA* composite (**d**)] are displayed as a scatter plot and significance was calculated and displayed as in Fig. 5f. The confidence interval around the regression line is indicated with shades. **e** Bar and violin plots show how a *IL15/IL15RA* composite score varies between KD samples. The score classifies KD-AV with giant CAAs from control (KD-CV) samples with a ROC AUC 0.95. Welch's two sample unpaired two-sided *t*-test is performed on the composite gene signature score to compute the *p* values. In multi-group setting each group is compared to the first control group and only significant *p* values are displayed.



**Fig. 7 Summary of findings and conclusions.** **a** Summary of datasets used (publicly available prior ones and new original cohorts) to support the conclusions in this work. Numbers in circles denote the number of subjects in each cohort. **b** Venn diagram displays the major findings from the current work. ViP/sViP signatures, and more specifically, the *IL15/IL15RA* specific gene induction are shared between patients in all three diagnostic groups. While these signatures are known to be associated with diffuse alveolar damage in the lungs of patients with COVID-19<sup>18</sup>, it is associated with CAA in KD and with reduction in cardiac muscle contractility in MIS-C. Overlapping features between each entity are displayed.

percentage < 0.25%) is a known poor prognostic factor for admissions into the pediatric ICU (hazard ratio [HR]: 2.96;  $P = 0.008^{79}$ ). It is possible that the two related clinical/laboratory parameters (low PLT and AEC) may be useful indicators of disease severity and prognosis in MIS-C and may guide decision-making in therapy and level of care in the hospital setting.

The strength of our study lies in the concurrent analysis of KD and MIS-C samples, our access to relatively large and independent cohorts of patients (in the case of KD), our ability to include age-matched pediatric healthy controls and febrile controls (non-KD and non-MIS-C, both pre-pandemic), and that all samples were drawn prior to the initiation of treatment. In doing so, this

study overcomes some of the limitations of prior studies<sup>12,14,16</sup>. Another strength is the use of a set of ViP signatures (that were validated in COVID-19)<sup>18</sup> and a KD-diagnostic signature<sup>37</sup> as the computational framework to compare the two syndromes. Last, but not the least, the multi-omics approach used here on samples obtained from the same patients allowed us to predict and validate the prominent upregulation of one shared cytokine pathway (i.e., *IL15*) at both transcript and protein level. Notable limitations of our study include a relatively small sample size of MIS-C subjects ( $n = 12$ ), limited number of publicly available MIS-C datasets for independent validation, and our inability to access cardiac tissue from KD and MIS-C subjects. Future studies on

emerging datasets will enable rigorous validation of the analysis presented here.

## Methods

**Kawasaki disease (KD), multisystem inflammatory syndrome in children (MIS-C), febrile control (FC) subjects.** All KD subjects met the American Heart Association (AHA) criteria<sup>80</sup> for complete or incomplete KD and subjects in this study were enrolled before the SARS-CoV-2 pandemic. Demographic and clinical data including echocardiography data and laboratory values were prospectively collected and entered into an electronic database. Coronary artery Z-scores were classified according to the AHA 2017 guidelines as follows: normal <2.0; dilated,  $2 \leq Z < 2.5$ ; aneurysm:  $2.5 \leq Z < 10.0$ ; and giant aneurysm,  $\geq 10.0$ .

All MIS-C patients met the case definition from the Centers for Disease Control and Prevention. Subjects were enrolled prospectively with collection of acute, pre-treatment samples. Demographic and clinical data including echocardiography data and laboratory values were prospectively collected and entered into an electronic database.

Febrile control patients had fever of at least three days duration and at least one mucocutaneous feature of KD including rash, conjunctival injection, or mucosal erythema. All were enrolled prior to the onset of the pandemic.

The final diagnosis for the control patients was adjudicated by a pediatric infectious disease specialist (J.C.B.) and by a pediatric emergency room physician (J.K.) at least 2 months after initial presentation when testing results and clinical outcome were known. The final diagnoses of the 30 FC were defined by PCR or viral culture and included the following infections: 12 adenovirus, 5 EBV, 2 metapneumovirus, 3 rhinovirus, 3 influenza, 2 parainfluenza, 2 RSV, and 1 measles.

The characteristics of patient cohorts that were part of this study are included in Supplementary Data 1. The study protocol was reviewed and approved by the institutional review board at UCSD (UCSD # 14020). Written informed consent from the parents or legal guardians and assent from patients were obtained as appropriate. For all the deidentified human subjects, information including age, gender, and previous history of the disease, was collected from the chart following HIPAA guidelines. The study design and the use of human study participants was conducted in accordance with the criteria set by the Declaration of Helsinki. Patients were not compensated for their participation in this study.

**Collection of blood samples and RNA isolation.** Whole blood was collected into PAXgene<sup>®</sup> tubes (PreAnalytiX) for RNA and into red top tubes for serum before the initiation of any treatment (illness day  $\leq 10$ ) for the KD, MIS-C and FC subjects (illness day <15 for some FC subjects) and at the clinic visit (day 17–25 of illness for subacute and day 289–3240 of illness for late convalescent) for the KD subjects. RNA was extracted following manufacturer's instruction (PAXgene Blood miRNA Kit). Serum was separated immediately by centrifugation and stored at  $-80^{\circ}\text{C}$  until use.

**Tissue samples.** We obtained formalin-fixed, paraffin-embedded tissues from a 4-year-old female who died 9 months after KD onset due to thrombosis of giant aneurysms. Written consent was obtained from the parents. The tissue sampling protocol was reviewed and approved by the Institutional Review Board at UCSD (UCSD# 180587).

**ViP and severe (s)ViP signatures.** ViP (Viral Pandemic) signature is derived from a list of 166 genes using Boolean Analysis of large viral infection datasets (training datasets: GSE47963,  $n = 438$ ; GSE113211,  $n = 118$ ). This 166-gene signature was conserved in all viral pandemics, including COVID-19, inspiring the nomenclatures ViP signature<sup>18</sup>. A subset of 20-genes classified disease severity called severe-ViP signature using an additional cohort (GSE101702,  $n = 159$ )<sup>18</sup>. To compute the ViP signature, first the genes present in this list were normalized according to a modified Z-score approach centered around StepMiner threshold (formula =  $(\text{expr} - \text{SThr})/3 * \text{stddev}$ ). The normalized expression values for every probeset for 166 genes were added together to create the final ViP signature. The severe ViP signature is computed similarly using 20 genes. The samples were ordered finally based on both the ViP and severe-ViP signature. A color-coded bar plot is combined with a violin plot to visualize the gene signature-based classification.

**Data analysis.** Several publicly available microarrays and RNASeq datasets were downloaded from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) website<sup>81–83</sup>. Gene expression summarization was performed by normalizing Affymetrix platforms by RMA (Robust Multichip Average)<sup>84,85</sup> and RNASeq platforms by computing transcripts per millions (TPM)<sup>86,87</sup> values whenever normalized data were not available in GEO. We used  $\log_2(\text{TPM} + 1)$  as the final gene expression value for analyses. GEO accession numbers are reported in figures, and text. KD/MIS-C RNASeq datasets were processed using salmon. Batch correction was performed using ComBat\_seq R package.

**StepMiner analysis.** StepMiner is a computational tool that identifies step-wise transitions in a time-series data<sup>88</sup>. StepMiner analysis is used to identify the threshold to convert continuous gene expression values into Boolean values (High/Low). StepMiner performs an adaptive regression scheme to identify the best possible step up or down based on sum-of-square errors. The steps are placed between time points at the sharpest change between low expression and high expression levels, which gives insight into the timing of the gene expression-switching event. To fit a step function, the algorithm evaluates all possible step positions, and for each position, it computes the average of the values on both sides of the step for the constant segments. An adaptive regression scheme is used that chooses the step positions that minimize the square error with the fitted data. Finally, a regression test statistic is computed using Eq. (1)

$$F \text{ stat} = \frac{\sum_{i=1}^n (\hat{X}_i - \bar{X})^2 / (m - 1)}{\sum_{i=1}^n (X_i - \hat{X}_i)^2 / (n - m)} \quad (1)$$

where  $X_i$  for  $i = 1$  to  $n$  are the values,  $\hat{X}_i$  for  $i = 1$  to  $n$  are fitted values.  $m$  is the degrees of freedom used for the adaptive regression analysis.  $\bar{X}$  is the average of all the values:  $\bar{X} = \frac{1}{n} * \sum_{j=1}^n X_j$ . For a step position at  $k$ , the fitted values  $\hat{X}_i$  are computed by using Eq. (2)

$$\frac{1}{k} * \sum_{j=1}^k X_j \text{ for } i = 1 \text{ to } k \text{ and } \frac{1}{(n - k)} * \sum_{j=k+1}^n X_j \text{ for } i = k + 1 \text{ to } n \quad (2)$$

**Boolean analysis.** Boolean logic is a simple mathematic relationship of two values, i.e., high/low, 1/0, or positive/negative. The Boolean analysis of gene expression data requires the conversion of expression levels into two possible values. The StepMiner algorithm is reused to perform Boolean analysis of gene expression data<sup>89</sup>. The Boolean analysis is a statistical approach which creates binary logical inferences that explain the relationships between phenomena. Boolean analysis is performed to determine the relationship between the expression levels of pairs of genes. The StepMiner algorithm is applied to gene expression levels to convert them into Boolean values (high and low). In this algorithm, first the expression values are sorted from low to high and a rising step function is fitted to the series to identify the threshold. Middle of the step is used as the StepMiner threshold. This threshold is used to convert gene expression values into Boolean values. A noise margin of 2-fold change is applied around the threshold to determine intermediate values, and these values are ignored during Boolean analysis.

**Boolean equivalent correlated clusters (BECC) analysis.** BECC analysis<sup>19</sup> is based on Boolean Equivalent<sup>89</sup> relationships, pair-wise correlation and linear regression analysis. BECC analysis identified ViP and severe-ViP signature using the BooleanNet statistic<sup>18</sup>.

**Heatmaps, hierarchical agglomerative clustering, PCA, differential expression analysis.** Gene expression values were normalized according to a modified Z-score approach centered around StepMiner threshold (formula =  $(\text{expr} - \text{SThr})/3 * \text{stddev}$ ). The samples were ordered according to average of the normalized gene expression values in the largest cluster along the Boolean path. The heatmap use red colors for the high values, white colors for the intermediate values and blue colors for low values. Gene names for few selected genes are highlighted on the left to show their expression patterns. Rows and columns are ordered based on hierarchical agglomerative clustering using python seaborn (version 0.10.1) clustermap function. Dendrograms are displayed for both rows and columns. Principal component analysis (PCA) was performed using sklearn PCA algorithm. PCA and hierarchical clustering algorithm is performed on top genes based on mean absolute deviation. StepMiner threshold was used first on the mean absolute deviation numbers to find high values and a second StepMiner threshold was performed to split the high values into top genes based on mean absolute deviation. Differential expression analysis was performed using DESeq2 in R with adjusted pvalue threshold of 0.1 and  $\log_2$  fold change threshold of 0.5.

**Statistical analyses.** Gene signature is used to classify sample categories and the performance of the multi-class classification is measured by ROC-AUC (receiver operating characteristics area under the curve) values. A color-coded bar plot is combined with a density plot to visualize the gene signature-based classification. All statistical tests were performed using R version 3.2.3 (2015-12-10). Standard  $t$ -tests were performed using python scipy.stats.ttest\_ind package (version 0.19.0) with Welch's two sample  $t$ -test (unpaired, unequal variance (equal\_var = False), and unequal sample size) parameters. Multiple hypothesis correction were performed by adjusting  $p$  values with statsmodels.stats.multitest.multipletests (fdr\_bh: Benjamini/Hochberg principles). The results were independently validated with R statistical software (R version 3.6.1; 2019-07-05). Differential expression analysis was performed in DESeq2 in R. Pathway analysis of gene lists were carried out via the Reactome database and algorithm<sup>90</sup>. Reactome identifies signaling and metabolic molecules and organizes their relations into biological pathways and processes. Kaplan–Meier analysis is performed using lifelines python package version

0.14.6. Violin, Swarm and Bubble plots are created using python seaborn package version 0.10.1. Principal component analysis (PCA) was performed using sklearn. The source code for Boolean analysis framework is available at <https://github.com/sahoo00/BoNE> and <https://github.com/sahoo00/Hegemon>.

**RNA sequencing.** For polyA capture: Total RNA was assessed for quality using an Agilent TapeStation 4200, and samples with an RNA Integrity Number (RIN) greater than 8.0 were used to generate RNA sequencing libraries using the TruSeq Stranded mRNA Sample Prep Kit with TruSeq Unique Dual Indexes (Illumina, San Diego, CA). Samples were processed following manufacturer's instructions, modifying RNA shear time to five minutes. Resulting libraries were multiplexed and sequenced with 100 basepair (bp) paired end reads (PE100) to a depth of approximately 50 million reads per sample on an Illumina NovaSeq 6000. Samples were demultiplexed using bcl2fastq v2.20 Conversion Software (Illumina, San Diego, CA). For ribosomal/globin depletion: Library preparation and sequencing of 30 million 75 or 100 bp paired end reads was conducted using the Illumina's TruSeq RNA Sample Preparation Kit, ribosomal and globin RNA depletion was performed using the Illumina® Ribo-Zero Gold kit and HiSeq 4000 at The Wellcome Center for Human Genetics.

**Human serum cytokines measurement.** Human serum cytokines measurement was performed using the V-PLEX Custom Human Biomarkers from MSD platform. Human serum samples fractionated from peripheral blood of KD and MIS-C patients (all samples collected prior to the initiation of treatments) were analyzed using customized standard multiplex plates as per the manufacturer's instructions.

**Immunohistochemistry.** Formalin-fixed, paraffin-embedded heart tissue sections from COVID19 and KD patients were stained anti-human IL15 receptor A polyclonal antibody (11:200 dilution; proteintech®, Rosemont, IL, USA; catalog# 16744-1-AP) and anti-human IL15 monoclonal antibody (1:10 dilution; Santa Cruz Biotechnology, Inc., Dallas, TX, USA; catalog# sc-8437) after heat-induced antigen retrieval with Tris buffer containing EDTA (pH 9.0). Sections were then incubated with respective HRP-conjugated secondary antibodies followed by DAB and hematoxylin counterstain (Sigma-Aldrich Inc., MO, USA; catalog# MHS1), and visualizing by Leica DM1000 LED (Leica Microsystems, Germany).

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Source data are provided with this paper. All data is available in the main text or the Supplementary materials. The GEO datasets generated in this work can be accessed at GSE178491. Publicly available datasets used: GSE109351; GSE73464; GSE15297; GSE68004; GSE18606; GSE9863; GSE63881; GSE73577; GSE16797; GSE166489; GSE126124; GSE166489; GSE167028; GSE11545; GSE116946; GSE100150; GSE147608; GSE122552; GSE79970; GSE149050; GSE153781; GSE148810; GSE75023; GSE27864; GSE21835; GSE57253.

## Code availability

The software codes are publicly available at the following links: <https://github.com/sahoo00/BoNE><sup>91</sup> and <https://github.com/sahoo00/Hegemon><sup>92</sup>.

Received: 11 April 2021; Accepted: 14 April 2022;

Published online: 16 May 2022

## References

- Levin, M. Childhood multisystem inflammatory syndrome - a new challenge in the pandemic. *N. Engl. J. Med.* **383**, 393–395 (2020).
- Whittaker, E. et al. Clinical characteristics of 58 children with a pediatric inflammatory multisystem syndrome temporally associated with SARS-CoV-2. *JAMA* **324**, 259–269 (2020).
- Riphagen, S., Gomez, X., Gonzalez-Martinez, C., Wilkinson, N. & Theocharis, P. Hyperinflammatory shock in children during COVID-19 pandemic. *Lancet* **395**, 1607–1608 (2020).
- Toubiana, J. et al. Kawasaki-like multisystem inflammatory syndrome in children during the covid-19 pandemic in Paris, France: prospective observational study. *BMJ* **369**, m2094 (2020).
- Kawasaki, T., Kosaki, F., Okawa, S., Shigematsu, I. & Yanagawa, H. A new infantile acute febrile mucocutaneous lymph node syndrome (MLNS) prevailing in Japan. *Pediatrics* **54**, 271–276 (1974).
- Makino, N. et al. Descriptive epidemiology of Kawasaki disease in Japan, 2011–2012: from the results of the 22nd nationwide survey. *J. Epidemiol.* **25**, 239–245 (2015).
- Nakamura, A., Ikeda, K. & Hamaoka, K. Aetiological significance of infectious stimuli in Kawasaki disease. *Front. Pediatr.* **7**, 1–9 (2019).
- Manlhiot, C. et al. Environmental epidemiology of Kawasaki disease: linking disease etiology, pathogenesis and global distribution. *PLoS ONE* **13**, e0191087 (2018).
- Rodó, X. et al. Tropospheric winds from northeastern China carry the etiologic agent of Kawasaki disease from its source to Japan. *Proc. Natl Acad. Sci. USA* **111**, 201400380 (2014).
- Dufort, E. M. et al. Multisystem inflammatory syndrome in children in New York state. *N. Engl. J. Med.* **383**, 347–358 (2020).
- Kanegaye, J. T. et al. Recognition of a Kawasaki disease shock syndrome. *Pediatrics* **123**, e783–e789 (2009).
- Gruber, C. N. et al. Mapping systemic inflammation and antibody responses in multisystem inflammatory syndrome in children (MIS-C). *Cell* **183**, 982–995.e914 (2020).
- Consiglio, C. R. et al. The immunology of multisystem inflammatory syndrome in children with COVID-19. *Cell* **183**, 968–981.e967 (2020).
- 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**, 1–18 (2021).
- 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.e1087 (2021).
- Carter, M. J. et al. Peripheral immunophenotypes in children with multisystem inflammatory syndrome associated with SARS-CoV-2 infection. *Nat. Med.* **26**, 1701–1707 (2020).
- Henderson, L. A. & Yeung, R. S. M. MIS-C: early lessons from immune profiling. *Nat. Rev. Rheumatol.* **17**, 75–76 (2021).
- Sahoo, D. et al. AI-guided discovery of the invariant host response to viral pandemics. *EBioMedicine* **68**, 103390 (2021).
- Dabydeen, S. A., Desai, A. & Sahoo, D. Unbiased Boolean analysis of public gene expression data for cell cycle gene identification. *Mol. Biol. Cell* **30**, 1770–1779 (2019).
- Schoggins, J. W. et al. A diverse range of gene products are effectors of the type I interferon antiviral response. *Nature* **472**, 481–485 (2011).
- Zhang, W. et al. Inhibition of respiratory syncytial virus infection with intranasal siRNA nanoparticles targeting the viral NS1 gene. *Nat. Med.* **11**, 56–62 (2005).
- Ogata, S. et al. Clinical score and transcript abundance patterns identify Kawasaki disease patients who may benefit from addition of methylprednisolone. *Pediatr. Res.* **66**, 577–584 (2009).
- Ben Tsutomu Saji, J. W. N., Jane, C. & Burns, M. T. Guidelines for diagnosis and management of cardiovascular sequelae in Kawasaki disease (JCS 2013). Digest version. *Circ. J.* **78**, 2521–2562 (2014).
- Eleftheriou, D. et al. Management of Kawasaki disease. *Arch. Dis. Child.* **99**, 74–83 (2014).
- Yanagawa, H. et al. Update of the epidemiology of Kawasaki disease in Japan—from the results of 1993–94 nationwide survey. *J. Epidemiol.* **6**, 148–157 (1996).
- McCrindle, B. W. et al. Coronary artery involvement in children with Kawasaki disease: risk factors from analysis of serial normalized measurements. *Circulation* **116**, 174–179 (2007).
- Manlhiot, C., Millar, K., Golding, F. & McCrindle, B. W. Improved classification of coronary artery abnormalities based only on coronary artery z-scores after Kawasaki disease. *Pediatr. Cardiol.* **31**, 242–249 (2010).
- Burns, J. C. et al. Temporal clusters of Kawasaki disease cases share distinct phenotypes that suggest response to diverse triggers. *J. Pediatr.* **229**, 48–53.e41 (2021).
- Rypdal, M. et al. Clustering and climate associations of Kawasaki Disease in San Diego County suggest environmental triggers. *Sci. Rep.* **8**, 16140 (2018).
- Rigante, D. Kawasaki disease as the immune-mediated echo of a viral infection. *Mediterr. J. Hematol. Infect. Dis.* **12**, e2020039 (2020).
- Jordan-Villegas, A., Chang, M. L., Ramilo, O. & Mejías, A. Concomitant respiratory viral infections in children with Kawasaki disease. *Pediatr. Infect. Dis. J.* **29**, 770–772 (2010).
- Chang, L. Y. et al. Viral infections associated with Kawasaki disease. *J. Formos. Med. Assoc.* **113**, 148–154 (2014).
- Belhadj, Z. et al. Acute heart failure in multisystem inflammatory syndrome in children in the context of global SARS-CoV-2 pandemic. *Circulation* **142**, 429–436 (2020).
- Friedman, K. G., Harrild, D. M. & Newburger, J. W. Cardiac dysfunction in multisystem inflammatory syndrome in children: a call to action. *J. Am. Coll. Cardiol.* **76**, 1962–1964 (2020).
- Matsubara, D. et al. Echocardiographic findings in pediatric multisystem inflammatory syndrome associated with COVID-19 in the United States. *J. Am. Coll. Cardiol.* **76**, 1947–1961 (2020).



36. Jang, G. C., Kim, H. Y., Ahn, S. Y. & Kim, D. S. Raised serum interleukin 15 levels in Kawasaki disease. *Ann. Rheum. Dis.* **62**, 264–266 (2003).
37. Wright, V. J. et al. Diagnosis of Kawasaki disease using a minimal whole-blood gene expression signature. *JAMA Pediatr.* **172**, e182293 (2018).
38. de Cevins, C. et al. A monocyte/dendritic cell molecular signature of SARS-CoV-2-related multisystem inflammatory syndrome in children with severe myocarditis. *Medicine (N. Y.)* **2**, 1072–1092 e1077 (2021).
39. Harley, J. B. et al. Transcription factors operate across disease loci, with EBNA2 implicated in autoimmunity. *Nat. Genet.* **50**, 699–707 (2018).
40. McInnes, I. B. & Gracie, J. A. Interleukin-15: a new cytokine target for the treatment of inflammatory diseases. *Curr. Opin. Pharm.* **4**, 392–397 (2004).
41. Dinarello, C. A., Simon, A. & van der Meer, J. W. Treating inflammation by blocking interleukin-1 in a broad spectrum of diseases. *Nat. Rev. Drug Discov.* **11**, 633–652 (2012).
42. Del Valle, D. M. et al. An inflammatory cytokine signature predicts COVID-19 severity and survival. *Nat. Med.* **26**, 1636–1643 (2020).
43. Robinson, P. C. et al. The potential for repurposing anti-TNF as a therapy for the treatment of COVID-19. *Medicine (N. Y.)* **1**, 90–102 (2020).
44. Feldmann, M. et al. Trials of anti-tumour necrosis factor therapy for COVID-19 are urgently needed. *Lancet* **395**, 1407–1409 (2020).
45. Zong, X., Gu, Y., Yu, H., Li, Z. & Wang, Y. Thrombocytopenia is associated with COVID-19 severity and outcome: an updated meta-analysis of 5637 patients with multiple outcomes. *Lab. Med.* **52**, 10–15 (2021).
46. Martincic, Z. et al. Severe immune thrombocytopenia in a critically ill COVID-19 patient. *Int. J. Infect. Dis.* **99**, 269–271 (2020).
47. Iba, T., Levy, J. H., Levi, M. & Thachil, J. Coagulopathy in COVID-19. *J. Thromb. Haemost.* **18**, 2103–2109 (2020).
48. Xu, P., Zhou, Q. & Xu, J. Mechanism of thrombocytopenia in COVID-19 patients. *Ann. Hematol.* **99**, 1205–1208 (2020).
49. Bhattacharjee, S. & Banerjee, M. Immune thrombocytopenia secondary to COVID-19: a systematic review. *SN Compr. Clin. Med.* **2**, 1–11 (2020).
50. Arora, K., Guleria, S., Jindal, A. K., Rawat, A. & Singh, S. Platelets in Kawasaki disease: is this only a numbers game or something beyond. *Genes Dis.* **7**, 62–66 (2020).
51. Tanni, F. et al. Eosinopenia and COVID-19. *J. Am. Osteopath. Assoc.* (2020).
52. Huang, P. Y., Huang, Y. H., Guo, M. M., Chang, L. S. & Kuo, H. C. Kawasaki disease and allergic diseases. *Front. Pediatr.* **8**, 614386 (2020).
53. Tsai, C. M. et al. A novel score system of blood tests for differentiating Kawasaki disease from febrile children. *PLoS ONE* **16**, e0244721 (2021).
54. Liu, X. P. et al. A nomogram model identifies eosinophilic frequencies to powerfully discriminate Kawasaki disease from febrile infections. *Front. Pediatr.* **8**, 559389 (2020).
55. Lin, L. Y. et al. Comparison of the laboratory data between Kawasaki disease and enterovirus after intravenous immunoglobulin treatment. *Pediatr. Cardiol.* **33**, 1269–1274 (2012).
56. Kuo, H. C. et al. Association of lower eosinophil-related T helper 2 (Th2) cytokines with coronary artery lesions in Kawasaki disease. *Pediatr. Allergy Immunol.* **20**, 266–272 (2009).
57. Lindsley, A. W., Schwartz, J. T. & Rothenberg, M. E. Eosinophil responses during COVID-19 infections and coronavirus vaccination. *J. Allergy Clin. Immunol.* **146**, 1–7 (2020).
58. Guo, L. et al. Role of interleukin-15 in cardiovascular diseases. *J. Cell. Mol. Med.* **24**, 7094–7101 (2020).
59. Yeghiazarians, Y. et al. IL-15: a novel prosurvival signaling pathway in cardiomyocytes. *J. Cardiovasc. Pharm.* **63**, 406–411 (2014).
60. Yonker, L. M. et al. Multisystem inflammatory syndrome in children is driven by zonulin-dependent loss of gut mucosal barrier. *J. Clin. Investig.* **131**, 1–12 (2021).
61. Hsieh, L. E. et al. Characterization of SARS-CoV-2 and common cold coronavirus-specific T-cell responses in MIS-C and Kawasaki disease children. *Eur. J. Immunol.* **131**, 1–15 (2021).
62. Abdel-Haq, N. et al. SARS-CoV-2-associated multisystem inflammatory syndrome in children: clinical manifestations and the role of infliximab treatment. *Eur. J. Pediatr.* **180**, 1–11 (2021).
63. Henderson, L. A. et al. American College of Rheumatology Clinical Guidance for multisystem inflammatory syndrome in children associated with SARS-CoV-2 and hyperinflammation in pediatric COVID-19: version 2. *Arthritis Rheumatol.* **73**, e13–e29 (2021).
64. Speth, C., Löffler, J., Krappmann, S., Lass-Flörl, C. & Rambach, G. Platelets as immune cells in infectious diseases. *Future Microbiol.* **8**, 1431–1451 (2013).
65. Assinger, A. Platelets and infection—an emerging role of platelets in viral infection. *Front. Immunol.* **5**, 649 (2014).
66. Seyoum, M., Enawgaw, B. & Melku, M. Human blood platelets and viruses: defense mechanism and role in the removal of viral pathogens. *Thromb. J.* **16**, 16 (2018).
67. Venkata, C., Kashyap, R., Farmer, J. C. & Afessa, B. Thrombocytopenia in adult patients with sepsis: incidence, risk factors, and its association with clinical outcome. *J. Intensive Care* **1**, 9 (2013).
68. Tsirogitis, P. et al. Thrombocytopenia in critically ill patients with severe sepsis/septic shock: prognostic value and association with a distinct serum cytokine profile. *J. Crit. Care* **32**, 9–15 (2016).
69. Bashash, D. et al. The prognostic value of thrombocytopenia in COVID-19 patients; a Systematic Review and Meta-Analysis. *Arch. Acad. Emerg. Med.* **8**, e75 (2020).
70. Liao, D. et al. Haematological characteristics and risk factors in the classification and prognosis evaluation of COVID-19: a retrospective cohort study. *Lancet Haematol.* **7**, e671–e678 (2020).
71. Jesenak, M. et al. Immune parameters and COVID-19 infection—associations with clinical severity and disease prognosis. *Front. Cell. Infect. Microbiol.* **10**, 1–10 (2020).
72. Jesenak, M. & Schwarze, J. Lung eosinophils—a novel “virus sink” that is defective in asthma? *Allergy* **74**, 1832–1834 (2019).
73. Bass, D. A. Reproduction of the eosinopenia of acute infection by passive transfer of a material obtained from inflammatory exudate. *Infect. Immun.* **15**, 410–416 (1977).
74. Bass, D. A. et al. Eosinopenia of acute infection: production of eosinopenia by chemotactic factors of acute inflammation. *J. Clin. Investig.* **65**, 1265–1271 (1980).
75. Abidi, K. et al. Eosinopenia is a reliable marker of sepsis on admission to medical intensive care units. *Crit. Care* **12**, R59 (2008).
76. Shaaban, H., Daniel, S., Sison, R., Slim, J. & Perez, G. Eosinopenia: Is it a good marker of sepsis in comparison to procalcitonin and C-reactive protein levels for patients admitted to a critical care unit in an urban hospital? *J. Crit. Care* **25**, 570–575 (2010).
77. Abidi, K. et al. Eosinopenia, an early marker of increased mortality in critically ill medical patients. *Intensive Care Med.* **37**, 1136–1142 (2011).
78. Du, H. et al. Clinical characteristics of 182 pediatric COVID-19 patients with different severities and allergic status. *Allergy* **76**, 510–532 (2021).
79. Kim, Y. H. et al. Prognostic usefulness of eosinopenia in the pediatric intensive care unit. *J. Korean Med. Sci.* **28**, 114–119 (2013).
80. McCrindle, B. W. et al. Diagnosis, treatment, and long-term management of Kawasaki disease: a scientific statement for health professionals from the American Heart Association. *Circulation* **135**, e927–e999 (2017).
81. Barrett, T. et al. NCBI GEO: mining millions of expression profiles—database and tools. *Nucleic Acids Res.* **33**, D562–D566 (2005).
82. Barrett, T. et al. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* **41**, D991–D995 (2013).
83. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30**, 207–210 (2002).
84. Irizarry, R. A. et al. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res.* **31**, e15 (2003).
85. Irizarry, R. A. et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249–264 (2003).
86. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinforma.* **12**, 323 (2011).
87. Pachter, L. Models for transcript quantification from RNA-Seq. *Genomics (q-bio.GN); Methodology (stat.ME)* **1**, 1–28 (2011).
88. Sahoo, D., Dill, D. L., Tibshirani, R. & Plevritis, S. K. Extracting binary signals from microarray time-course data. *Nucleic Acids Res.* **35**, 3705–3712 (2007).
89. Sahoo, D., Dill, D. L., Gentles, A. J., Tibshirani, R. & Plevritis, S. K. Boolean implication networks derived from large scale, whole genome microarray datasets. *Genome Biol.* **9**, R157 (2008).
90. Fabregat, A. et al. The reactome pathway knowledgebase. *Nucleic Acids Res.* **46**, D649–D655 (2018).
91. Sahoo, D. & Vo, T. D. Artificial intelligence guided discovery of a barrier-protective therapy in inflammatory bowel disease. *GitHub/sahoo00/BoNE* (2021).
92. Sahoo, D. Artificial intelligence guided discovery of a barrier-protective therapy in inflammatory bowel disease. *GitHub/sahoo00/Hegemon* (2021).

## Acknowledgements

This work was supported by the National Institutes for Health (NIH) grants R01-GM138385 (to D.S.), R01-AI141630 (to P.G.), R01DK107585 (to S.D.) and R01-AI155696 (to P.G., D.S., and S.D.), UCOP-RGPO (R01RG3780, R00RG2628 and R00RG2642 to P.G., D.S., and S.D.), iDASH U54HL108460, R01HL140898 (to J.C.B. and A.H.T.), and a UC San Diego Stem Cell Center Pilot award (to P.G., D.S., and S.D.). G.D.K. was supported through The American Association of Immunologists Intersect Fellowship Program for Computational Scientists and Immunologists. Clinical sample collection was supported by the Patient Outcomes Research Institute (PCORI) CER-1602–3447 (to J.C.B.), a Gordon and Marilyn Macklin Foundation grant (J.C.B.), and a UC San Diego Stem Cell Center Pilot award (to P.G., D.S., and S.D.). This publication includes data generated at the UC San Diego IGM Genomics Center Utilizing an Illumina NOVASeq 6000 that was purchased with funding from a National Institutes of

Health SIG grant (#S10 OD026929). Figures were created with BioRender.com and assembled using adobe illustrator.

### Author contributions

J.C.B., A.T.H. recruited the KD patients in cohort 1 and conducted the RNA seq studies; D.S., J.C.B. and P.G. conceptualized the project; C.S., A.T.H., Jo.Ka, J.B., J.C.B., Pediatric Emergency Medicine Kawasaki Disease Research Group\* recruited the KD and MIS-C patients in cohorts 2 and 3 and collected biological samples used in this study; G.D.K. and C.S. carried out the serum cytokine analysis under the supervision of S.D.; S.K. and G.D.K. carried out the immunohistochemical studies under the supervision of D.S. and P.G.; D.S. carried out all computational modeling and analyses and contributed all software used in this work; Ji.Ki. was responsible for the management of transcriptomic datasets and uploading to the NCBI. D.S., G.D.K., and P.G. prepared figures for data visualization; P.G., G.D.K., C.S. wrote the original draft of the manuscript; D.S., G.D.K., A.H., Jo.Ka., C.S., J.C.B. edited and revised the manuscript. All co-authors approved the final version of the manuscript. D.S., S.D., J.C.B., and P.G. supervised various parts of the project and secured funding; D.S. and P.G. administered the project.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-30357-w>.

**Correspondence** and requests for materials should be addressed to Pradipta Ghosh.

**Peer review information** *Nature Communications* thanks Hamid Bassiri, Ying-Hsien Huang and the other, anonymous, reviewer for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022

## Pediatric Emergency Medicine Kawasaki Disease Research Group

Naomi Abe<sup>7</sup>, Lukas Austin-Page<sup>7</sup>, Amy Bryl<sup>8</sup>, J. Joelle Donofrio-Ödmann<sup>7,8</sup>, Atim Ekpenyong<sup>7,8</sup>, Michael Gardiner<sup>7,8</sup>, David J. Gutglass<sup>7,8</sup>, Margaret B. Nguyen<sup>7,8</sup>, Kristy Schwartz<sup>7,8</sup>, Stacey Ulrich<sup>7,8</sup>, Tatyana Vayngortin<sup>7,8</sup>, Elise Zimmerman<sup>7</sup>, Joseph Bocchini<sup>9</sup>, Soumita Das<sup>10</sup>, Jane C. Burns<sup>3,4</sup> & Debashis Sahoo<sup>3,6</sup>

<sup>7</sup>Rady Children's Hospital San Diego, San Diego, CA, USA. <sup>8</sup>Department of Pediatrics, University of California San Diego, San Diego, USA. <sup>9</sup>Willis-Knighton Health System, Shreveport, LA, USA. <sup>10</sup>Department of Pathology, University of California San Diego, San Diego, USA.