

Published in final edited form as:

J Infect Dis. 2009 August 15; 200(4): 657–666. doi:10.1086/603538.

Gene Transcript Abundance Profiles Distinguish Kawasaki Disease from Adenovirus Infection

Stephen J. Popper^{1,2}, Virginia E. Watson^{3,4}, Chisato Shimizu³, John T. Kanegaye^{3,4}, Jane C. Burns^{3,4}, and David A. Relman^{1,2,5}

¹Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford

²Department of Medicine, Stanford University School of Medicine, Stanford

³Department of Pediatrics, University of California San Diego School of Medicine, La Jolla

⁴Rady Children's Hospital, San Diego

⁵Veterans Affairs Palo Alto Health Care System, Palo Alto, California

Abstract

Background—Acute Kawasaki disease (KD) is difficult to distinguish from other illnesses that involve acute rash or fever, in part because the etiologic agent(s) and pathophysiology remain poorly characterized. As a result, diagnosis and critical therapies may be delayed.

Methods—We used DNA microarrays to identify possible diagnostic features of KD. We compared gene expression patterns in the blood of 23 children with acute KD and 18 age-matched febrile children with 3 illnesses that resemble KD.

Results—Genes associated with platelet and neutrophil activation were expressed at higher levels in patients with KD than in patients with acute adenovirus infections or systemic adverse drug reactions, but levels in patients with KD were not higher than those in patients with scarlet fever. Genes associated with B cell activation were also expressed at higher levels in patients with KD than in control subjects. A striking absence of interferon-stimulated gene expression in patients with KD was confirmed in an independent cohort of patients with KD. Using a set of 38 gene transcripts, we successfully predicted the diagnosis for 21 of 23 patients with KD and 7 of 8 patients with adenovirus infection.

Conclusions—These findings provide insight into the molecular features that distinguish KD from other febrile illnesses and support the feasibility of developing novel diagnostic reagents for KD based on the host response.

Efforts to discern the etiology of acute febrile disease are hampered by the paucity of reliable discriminating clinical features, difficulties in obtaining appropriate specimens, insensitive methods for detecting known causative agents, and the lack of diagnostic tests for some conditions associated with fever, such as autoimmune diseases and adverse drug reactions. As a result, many acute febrile illnesses remain unexplained, especially in the early days after onset of clinical signs and symptoms. Molecular profiling of the host response offers an

2009 by the Infectious Diseases Society of America. All rights reserved.

Reprints or correspondence: Dr. Stephen J. Popper, Dept of Microbiology and Immunology, Stanford University School of Medicine, MC5124, Stanford, CA 94305 (sopper@stanford.edu)..

Potential conflicts of interest: none reported.

Presented in part: Pediatric Academic Societies and Asian Society for Pediatric Research Joint Meeting, Honolulu, Hawaii, 2–6 May 2008 (abstract 6430.1).

approach for classifying acutely ill hosts that complements traditional diagnostic approaches based on microbial detection [1]. Studies of human genome-wide transcript abundance patterns in peripheral blood suggest that these patterns might provide useful information about the disease mechanism, outcome, nature of the infectious agent, and diagnosis [2–6]. This last possibility has not been adequately explored, especially in the setting of a clinical syndrome that presents important diagnostic dilemmas.

Kawasaki disease (KD) is an acute, self-limited inflammatory illness of infants and children [7]; ~25% of untreated patients develop coronary artery aneurysms or ectasia. Intravenous immunoglobulin (IVIG) reduces the rate of coronary artery aneurysms to ~5% when administered within the first 10 days of illness, but KD remains the leading cause of acquired pediatric heart disease in developed nations.

Despite 30 years of research, no etiologic agent has been identified for KD. In the absence of a specific diagnostic test, KD is diagnosed according to clinical criteria, many of which are shared by other illnesses characterized by rash or fever, including adenovirus infections, streptococcal scarlet fever, and systemic drug reactions. Many children with KD consequently receive an erroneous or late diagnosis, which leads to delays in treatment and an increased risk of coronary artery aneurysm formation [8,9].

We recently examined whole-blood genomewide transcript abundance patterns in patients with KD and identified specific transcript levels associated with the risk of subsequent failure to respond to IVIG therapy [4]. In the current study, we compared patterns of whole-blood gene expression in patients during the acute phase of KD with patterns found in patients during the early phase of 3 illnesses that have similar clinical presentations but well-defined alternative etiologies. We identified patterns of gene expression and corresponding biological programs that were different in patients with KD, compared to patients with the other illnesses, and we were able to distinguish between KD and adenovirus infection on the basis of gene expression patterns. The results from this study indicate that comparative analysis of host gene expression profiles is a promising approach for better understanding febrile illnesses and that such analysis may contribute to the development of a test for KD that enables more accurate and timely diagnosis.

MATERIALS AND METHODS

Study population and sample collection

Three groups of pediatric patients—patients with KD ($n = 23$), febrile control subjects ($n = 18$), and healthy, nonfebrile control subjects ($n = 10$)—were enrolled at 2 clinical sites (Rady Children's Hospital–San Diego and Children's Hospital Boston) after obtaining informed consent from parents or legal guardians. A human subjects research protocol was reviewed and approved by the institutional review boards at the University of California, San Diego, Children's Hospital Boston, and Stanford University. All patients with KD had fever and ≥ 4 of the 5 principal clinical criteria for KD (ie, rash, conjunctival injection, cervical lymphadenopathy, changes in the oral mucosa, and changes in the extremities) or 3 criteria in combination with coronary artery abnormalities documented by echocardiography [10]. Coronary artery dimensions were recorded for all patients with KD.

Nasopharyngeal (NP) and stool samples for viral cultures were obtained from all febrile control subjects. Control subjects classified as having acute adenovirus infection had fever for ≥ 3 days, conjunctival or mucocutaneous changes, a negative result from culture of a throat sample for group A β -hemolytic *Streptococcus* (GAS), and an NP culture positive for adenovirus. Control subjects classified as having acute streptococcal scarlet fever had fever, a diffuse scarlatiniform rash, clear conjunctivae, NP and stool samples negative for virus on culture, and a positive

rapid test result for GAS. Control subjects with systemic drug reactions had fever and systemic signs associated with ingestion of a drug known to cause hypersensitivity reactions, a throat culture negative for GAS, and NP and stool samples negative for virus on culture. Healthy pediatric control subjects ($n = 10$) were children <6 years of age undergoing minor elective surgery for polydactyly. Clinical data, including sex, ethnicity, age, day of illness (the first day of fever was defined as day 1 of illness), results of laboratory testing, response to IVIG therapy, and coronary artery status, were recorded for all subjects (Appendix). Blood samples were obtained for determination of complete blood count and differential; erythrocyte sedimentation rate; levels of C-reactive protein, alanine aminotransferase, and γ -glutamyl transferase; and RNA studies (PAXgene Blood RNA System; PreAnalytiX GmbH).

DNA microarray analysis

RNA transcripts in the samples and a standard reference RNA (Universal Human Reference RNA; Stratagene) were amplified using the MessageAmp aRNA amplification kit (Ambion). Sample and reference transcripts were then reverse-transcribed, labeled with fluorescent dyes (cyanine [Cy] 5 and Cy3, respectively), mixed together, and hybridized to complementary DNA (cDNA) microarrays (Appendix). The data are available from the Stanford Microarray Database [11] and NCBI's Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?accpGSE15297>).

GeneTrail [12], the Molecular Signature Database [13], and NextBio (<http://nextbio.com>) were used to identify Gene Ontology (GO) terms, biological pathways from the Kyoto Encyclopedia of Genes and Genomes (<http://www.genome.jp/kegg/>), and other array data sets associated with gene sets in our analysis. Terms in GeneTrail were identified using a false discovery rate of 5%. Significance analysis of microarrays was used to identify genes associated with differences between patients with KD and each control group at a median false discovery rate of 1% [14]. Prediction analysis of microarrays with 10-fold cross-validation was used to identify and assess the predictive values of genes distinguishing patients with KD from those with adenovirus infection [15]. Potential differences in clinical parameters were tested using nonparametric statistical analyses and Stata software (version 7; StataCorp).

Quantitative reverse-transcription polymerase chain reaction (RT-PCR)

Levels of ISG15, LY6E, and MX1 mRNA were measured using TaqMan 5'-nuclease Gene Expression Assays (Applied Biosystems) (Appendix).

RESULTS

Comparison of transcript profiles in patients with KD or other febrile illnesses

RNA from the peripheral blood of 23 children with KD and 18 children with a well-defined illness that had a similar clinical presentation was processed and hybridized to cDNA arrays. The demographic and clinical characteristics of the subjects in these groups are presented in table 1.

To focus on the transcripts with the greatest differences in abundance among the subjects, we identified 808 transcripts that varied ≥ 3 -fold from the median in ≥ 2 of the 41 arrays, and we used unsupervised clustering to organize the samples (figure 1). Sample cluster A formed 1 of the 2 main branches of the dendrogram and consisted primarily of samples from patients with KD; 14 of the 19 samples were from patients with KD, and 4 of the remaining 5 samples were from subjects with scarlet fever. Conversely, sample cluster B contained 7 of 8 samples obtained from subjects with adenovirus infection, along with 3 samples from subjects with systemic drug reactions and 1 from a patient with KD. The grouping of samples suggested that disease-specific gene expression was a prominent feature of the overall transcript profiles. We

then compared the correlation coefficients for pairs of KD and non-KD samples from a given control group with the correlation coefficients for comparisons between KD samples; the transcript profiles of the subjects with adenovirus infection and or systemic drug reactions differed significantly from those of the patients with KD, but the profiles of subjects with scarlet fever did not (figure 2, which appears only in the electronic version of the *Journal*).

Identification of KD-associated transcript patterns

We used significance analysis of microarrays to identify specific transcripts that were significantly more or less abundant in the children with KD than in subjects from each of the 3 control groups [14] (table 2, which is available only as a spreadsheet in the online version of the *Journal*). In patients with KD, 130 transcripts were differentially expressed in comparison to subjects with adenovirus infections, 135 transcripts were differentially expressed in comparison to subjects with systemic drug reactions, and 4 transcripts were differentially expressed in comparison to subjects with scarlet fever; 29 transcripts were common to the first 2 comparisons. To identify additional transcripts whose abundance was similar across control groups but different from that in the KD group, we compared transcript levels in the patients with KD with those in the combined set of febrile control subjects; 31 additional transcripts were identified in this manner.

To identify putative functional processes associated with these transcripts, we clustered the set of 271 differentially expressed transcripts (figure 3) and applied several complementary approaches. We identified GO terms and annotated biological pathways that were overrepresented, we identified other array data sets that were enriched for genes present in each of these 4 gene clusters, and we incorporated information about the temporal pattern of expression for these genes from a previous study of KD [4].

Genes associated with cell adhesion and innate immune responses in patients with KD

There were 4 main clusters among the KD-associated genes, each with ≥ 20 transcripts and an average Pearson correlation coefficient of 0.5. The first of these clusters (C1) (figure 3) consisted of 45 transcripts that were more abundant in patients with KD than in control subjects. Forty-two transcripts were significantly less abundant in subjects with adenovirus infection than in those with KD, and 13 were less abundant in subjects with drug reactions. Genes represented in the C1 cluster were enriched for the GO term of contractile fibers ($P < .01$); these included genes encoding vinculin and the myosin light chain 4 and 9 subunits. Each of these genes is involved in modulating cell adhesion, morphology, and migration, as are a number of other genes in this cluster, including *SPARC*, *ITGA2B*, *ITGB5*, *TUBB1*, and *TBCC*. *SPARC*, *SELP*, *ITGA2B*, and *TUBB1* are also highly expressed in platelets; this gene set was significantly enriched for genes expressed during in vitro differentiation of stem cells into megakaryocytes ($P < .01$) [16] and for genes associated with hematopoietic cell maturation ($P < 10^{-5}$) [17].

Cluster 2 (C2) (figure 3) also contained transcripts that were more abundant in patients with KD than in subjects with adenovirus infections or drug reactions. Approximately half of the transcripts (54 of 100) were among those previously identified as defining the acute phase of KD in our study of temporal patterns of gene expression in KD [4]. These included a number of genes encoding antimicrobial peptides (*BPI*, *SLPI*, *DEFA1*, *PI3*) and other genes associated with early innate immune responses (*S100* genes *S100P* and *S100A12*, *MMP9*, *SORL1*, and *PBEF*). No GO terms met our criteria for significance, but we identified 2 overlapping Kyoto Encyclopedia of Genes and Genomes pathways, “apoptosis” and “insulin signaling,” that were significantly associated with this gene cluster ($P < .01$). Genes associated with these GO terms encoded proteins in signaling pathways linked to immune responses and inflammation, including interleukin (IL)–1 receptor type 1, IL-1 receptor–associated protein, IKBKG,

PIK3R1, and PRKAR1A. Additional cytokine receptors (IL-4 receptor and interferon [IFN]- γ receptor subunit 1) were also present in this gene cluster. C2 also included many genes known to be expressed at higher levels in neutrophils than in other leukocyte subsets [18], as well as genes associated with differentiation of both CD34⁺ myeloid cells and monocytes ($P < 10^{-14}$ and $P < 10^{-4}$, respectively) [19,20].

B cell activation program in patients with KD

The 33 transcripts in cluster 3 (C3) (figure 3) were also significantly more abundant in patients with KD than in control subjects. Twenty-nine achieved statistical significance only when the pooled set of febrile control subjects was compared with the patients with KD, suggesting a consistent difference between patients with KD and the febrile control subjects within this cluster, but a less dramatic difference than within clusters 1 and 2. This impression was reinforced by a power calculation indicating that similar sample sizes would be needed to identify significant differences for each control group (figure 4, which appears only in the electronic version of the *Journal*).

C3 was highly enriched for genes whose expression is associated with B cells [18]. The pattern of transcript abundance was associated with a gene expression program that characterizes a nonplasma cell stage of activation: *PAX5*, which plays a central role in the early activation and differentiation of B cells and whose expression must be repressed for development of plasma cells, was more highly expressed in patients with KD, as were genes known to be direct targets of *PAX5* transcriptional activation (*CD79A*, *MS4A2*, *IGHM*, *SPIB*, *HLA-DNA*, *TRAF5*) [21]. Unlike the expression of genes in C1 and C2, the expression in C3 did not diminish when the acute phase of the disease ended; the average transcript abundance for individual patients with KD was very similar weeks to months later, during the convalescent phase ($P = .92$; Student's paired *t* test).

IFN-responsive genes in patients with KD or adenovirus infection

The fourth cluster (C4) (figure 3) consisted of 60 transcripts that were less abundant in patients with KD than in those with adenovirus infection. The defining feature of this gene cluster was the presence of canonical IFN-induced genes, such as *MX1*, *MX2*, *ISG15* (*G1P2*), *IFIT2*, *OAS1*, and *OAS2*. As a group, C4 was highly enriched for genes known to be expressed after cell stimulation with type I IFNs, both in vitro and in subjects with hepatitis C treated with pegylated IFN (all $P < 10^{-15}$) [22–24]. To verify this dramatic difference in IFN-induced gene expression in a separate group of subjects, we used RT-PCR to measure levels of *ISG15*, *LY6E*, and *MX1* in an independent cohort of 10 patients with KD, 12 subjects with adenovirus infection, and 8 healthy control subjects, matched for age and sex. Levels of these transcripts in the patients with KD were again significantly lower than those in subjects with adenovirus infection and similar to those in healthy children (figure 5).

Distinguishing KD and adenovirus infection by transcript levels

The presence of multiple expression programs that differed between adenovirus infection and KD suggested that the 2 diseases might be distinguished using transcript levels. We used prediction analysis of microarrays to identify a set of genes that might differentiate between these cohorts. Using a set of 38 genes and 10-fold cross-validation, we were able to correctly identify 21 of 23 patients with KD and 7 of 8 with adenovirus infection (figure 6A). Sixteen genes were among the IFN-induced gene set, and the other 22 were distributed among the gene sets associated with cell adhesion or motility and innate immune responses (figure 6B).

DISCUSSION

We analyzed patterns of gene transcript abundance in the blood of subjects with acute KD, adenovirus or streptococcal toxin-mediated infections, or systemic drug reactions. These 4 clinical cohorts were chosen because experience and the published literature indicate that they can be clinically confused in the populations from which the patients with KD were recruited [25,26]. We identified sets of genes whose transcript abundance levels were associated specifically with KD; the corresponding functional annotations point to cellular processes that may contribute to the unique pathogenic process that defines KD. A subset of these genes was used to distinguish patients with KD from those with adenovirus infections, with high sensitivity and specificity.

Two sets of genes were expressed at higher levels in patients with KD than in subjects with adenovirus infections or systemic drug reactions. The first was associated with expression programs that regulate and promote cellular adhesion and motility and are also highly expressed in platelets. The second contained many genes associated with innate immune responses generated by neutrophils. High numbers of platelets and neutrophils are characteristic of acute KD, but transcript levels of these genes are unlikely to result simply from differences in the relative number of these 2 cell subsets. Although platelet counts in the patients with KD in this study were higher than those in subjects with adenovirus infection or systemic drug reactions, they were also higher than in those with scarlet fever, despite the absence of a significant difference in gene expression. Similarly, there were no significant differences in relative neutrophil abundance among the different subject groups in this study that could explain the differences in expression of genes associated with innate immune responses. Despite known temporal variation in the expression of genes associated with innate immune responses within the acute phase of KD [4], the similarities in day of illness for the patients with KD, adenovirus infections, and scarlet fever also make it unlikely that this variable alone explains differences in the abundance of transcripts for these genes (table 1).

The expression of these genes may instead reflect the activation state of the cell subsets involved and may help explain the cellular events that contribute to coronary artery damage in patients with KD. Platelet activation has long been recognized as associated with coronary artery aneurysm formation in KD [27], and platelets can respond to and generate inflammatory signals in ways that modulate their interaction with both neutrophils and blood vessel endothelia (reviewed in Zimmerman et al [28]). Similarly, it has become clear that variation in neutrophil gene transcription is associated with differences in neutrophil function [29].

A third set of transcripts that were more abundant in patients with KD than in each of the 3 control groups was enriched for genes associated with B cell differentiation and activation [21]. The presence of *PAX5* and *IGHM* in this set of genes is consistent with an increased proportion of either mature naive B cells or immunoglobulin M (IgM)-positive memory cells in KD. Patterns of transcript abundance in memory B cells and naive cells share similarities when compared with plasma cells [30]. Kuo et al found a subset of the genes associated with KD among those expressed at higher levels in memory B cells than in plasma cells [31]. Our findings are also consistent with a recent report that *CD180* (*Ly64*) is expressed in a higher fraction of B cells in patients with KD than in febrile control subjects [32]; *CD180* is also expressed at higher levels in memory B cells than in either naive or plasma cells [33]. Studies of purified B cell populations are needed to determine whether the expression pattern we observed in this study is a memory B cell signature, but these findings may help explain previous reports that peripheral blood mononuclear cells from children with KD spontaneously secreted immunoglobulin at higher levels than those from febrile control subjects [34]; memory B cells are primed for rapid activation and differentiation into antibody-secreting plasma cells,

and recent studies suggest that IgM-positive memory B cells respond to both T cell-independent and T cell-dependent antigens [35].

The fourth set of transcripts that distinguished patients with KD from febrile control subjects consisted of IFN-stimulated genes, which were expressed at lower levels in patients with KD than in those with adenovirus infections. Lower expression levels of these genes in subjects with acute KD than in both subjects recovering from KD and healthy control subjects suggest that an IFN response is absent in patients with KD. It is unlikely that the pattern of expression we observed represents a defect in the immune repertoire of patients with KD; unlike individuals with known defects in IFN signaling, children who have had KD do not suffer from a general susceptibility to infections. The absence of an IFN response does not necessarily imply that KD is not caused by a virus. Many viruses evade the IFN response by a wide variety of mechanisms, including inhibition of IFN production and of IFN signaling, both of which would decrease the expression of IFN-stimulated genes (reviewed in Randall et al [36]).

Adenovirus infections are particularly problematic in the differential diagnosis of KD because the currently available rapid test for this common group of viruses is very insensitive, unlike the one for streptococcal disease [25]. A PCR-based test using broadly cross-hybridizing primers is highly sensitive and specific but remains a research tool [37]. Immunofluorescence-based tests suffer from poor sensitivity (~50%), although these tests are highly specific, and a positive test result for cells from the nasopharynx suggests current infection [38]. Many of the laboratory features associated with acute adenovirus infection, such as elevated white blood cell counts with neutrophil predominance and high levels of acute-phase reactants (eg, fibrinogen and C-reactive protein), parallel those associated with acute KD [25]. Using a subset of the differentially expressed genes identified in this study, we were able to correctly identify 21 of 23 subjects with acute KD and 7 of 8 subjects with adenovirus infection.

We recognize several limitations to our study. First, small sample size and a restricted set of control illnesses may limit the ability to generalize from our results; larger studies that include additional diseases are needed. Second, although subjects were classified on the basis of generally accepted clinical and laboratory criteria, some may have been misclassified.

Our findings have implications for future research into the etiology and diagnosis of KD. First, the search for an etiologic agent should focus on agents that either suppress or fail to elicit a robust type I IFN response in vivo. Second, adenovirus infections are among the diseases most often confused with KD (25); efforts to develop improved diagnostic tests for KD might be well served by including both human gene products that are highly expressed in adenovirus infection but not in KD and gene products with the inverse differential abundance profile. Combinations of markers, some derived from host response and some from possible causative agents, may provide the most effective approach for diagnosis and classification in acute febrile subjects.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We wish to thank Jane Newburger for providing samples and clinical information, Patrick Brown and Elizabeth Joyce for comments and discussion, Joan Panchari for sample collection, DeeAnna Scherrer for RNA isolation, and Kristy Coolley for assistance with sample preparation and microarray hybridization.

Financial support: National Institutes of Health (grant NHLBI R01-HL69413 to J.C.B. and J.T.K.); Horn Foundation (D.A.R.).

APPENDIX

SUPPLEMENTARY METHODS

Additional clinical information

IVIG nonresponse was defined as persistent or recrudescent fever ($\geq 38^{\circ}\text{C}$ [100.4°F] rectally or orally) 48 h after completion of the IVIG infusion (2 g/kg). The internal diameters of the right coronary and left anterior descending arteries were classified by echocardiography as normal (z score, <2.5 [standard deviations from the mean, normalized for body surface area [39]], dilated (z score, 2.5 to <4.0), or aneurysmal (saccular or fusiform dilatation of a coronary artery segment; z score, ≥ 4.0).

Additional microarray analysis details

PAXgene samples were stored at 4°C on the day of collection, and RNA was extracted within 1 week, in accordance with the manufacturer's directions. Full details of the protocols used for array hybridization are available at <http://cmgm.stanford.edu/pbrown/protocols>. Images of hybridized arrays were obtained using a GenePix 4000B microarray scanner and analyzed with GenePix software (version 5.0; Axon Instruments). The arrays used for these studies contain 37,632 spots derived from cDNA clones representing ~18,000 unique human genes [40].

Data were filtered to include only clones that met the following criteria for $\geq 80\%$ of the samples tested: signal intensity 2.5-fold above background in either the Cy5 (sample) or Cy3 (reference) channel and a regression correlation for the 2 channels of ≥ 0.6 across each measured element. A normalization factor was applied so that the mean \log_2 ratio for each array (sample) was 0, and data for each clone were then median centered across all observations. Selected data were organized using a hierarchical clustering algorithm based on a Pearson correlation metric, with average linkage clustering [41], and visualized using TreeView software, version 1.1.1 [42].

Additional RT-PCR information

cDNA was synthesized using Oligo(dT)20 (Invitrogen) and Superscript III reverse transcriptase (Invitrogen), in accordance with the manufacturer's instructions. PCRs were prepared using the TaqMan Universal PCR Master Mix (Applied Biosystems) and cDNA derived from 20 ng of total RNA. The relative abundance of the target transcripts was calculated by comparison with a standard curve and normalized to the expression level of TATA box binding protein-associated factor, RNA polymerase I, B (*TAF1B*). The Applied Biosystems assay catalog numbers are as follows: *ISG15*, Hs00192713_m1; *LY6E*, Hs00158942_m1; *MX1*, Hs00182073_m1; *TAF1B*, Hs00374547_m1.

References

1. Liu M, Popper SJ, Rubins KH, Relman DA. Early days: genomics and human responses to infection. *Curr Opin Microbiol* 2006;9:312–9. [PubMed: 16679048]
2. Boldrick JC, Alizadeh AA, Diehn M, et al. Stereotyped and specific gene expression programs in human innate immune responses to bacteria. *Proc Natl Acad Sci USA* 2002;99:972–7. [PubMed: 11805339]
3. Griffiths MJ, Shafi MJ, Popper SJ, et al. Genomewide analysis of the host response to malaria in Kenyan children. *J Infect Dis* 2005;191:1599–611. [PubMed: 15838786]
4. Popper SJ, Shimizu C, Shike H, et al. Gene-expression patterns reveal underlying biological processes in Kawasaki disease. *Genome Biol* 2007;8:R261. [PubMed: 18067656]
5. Simmons CP, Popper SJ, Doloczek C, et al. Patterns of host genome-wide gene transcript abundance in the peripheral blood of patients with acute dengue haemorrhagic fever. *J Infect Dis* 2007;195:1097–107. [PubMed: 17357045]

6. Ramilo O, Allman W, Chung W, et al. Gene expression patterns in blood leukocytes discriminate patients with acute infections. *Blood* 2007;109:2066–77. [PubMed: 17105821]
7. Burns JC, Glode MP. Kawasaki syndrome. *Lancet* 2004;364:533–44. [PubMed: 15302199]
8. Minich LL, Sleeper LA, Atz AM, et al. Delayed diagnosis of Kawasaki disease: what are the risk factors? *Pediatrics* 2007;120:e1434–40. [PubMed: 18025079]
9. Wilder MS, Palinkas LA, Kao AS, Bastian JF, Turner CL, Burns JC. Delayed diagnosis by physicians contributes to the development of coronary artery aneurysms in children with Kawasaki syndrome. *Pediatr Infect Dis J* 2007;26:256–60. [PubMed: 17484225]
10. Newburger JW, Takahashi M, Gerber MA, et al. Diagnosis, treatment, and long-term management of Kawasaki disease: a statement for health professionals from the Committee on Rheumatic Fever, Endocarditis and Kawasaki Disease, Council on Cardiovascular Disease in the Young, American Heart Association. *Circulation* 2004;110:2747–71. [PubMed: 15505111]
11. Demeter J, Beauheim C, Gollub J, et al. The Stanford Microarray Database: implementation of new analysis tools and open source release of software. *Nucleic Acids Res* 2007;35:D766–70. [PubMed: 17182626]
12. Backes C, Keller A, Kuentzer J, et al. GeneTrail—advanced gene set enrichment analysis. *Nucleic Acids Res* 2007;35:W186–92. [PubMed: 17526521]
13. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 2005;102:15545–50. [PubMed: 16199517]
14. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA* 2001;98:5116–21. [PubMed: 11309499]
15. Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci USA* 2002;99:6567–72. [PubMed: 12011421]
16. Shim MH, Hoover A, Blake N, Drachman JG, Reems JA. Gene expression profile of primary human CD34⁺CD38^{lo} cells differentiating along the megakaryocyte lineage. *Exp Hematol* 2004;32:638–48. [PubMed: 15246160]
17. Keller MA, Addya S, Vadigepalli R, et al. Transcriptional regulatory network analysis of developing human erythroid progenitors reveals patterns of coregulation and potential transcriptional regulators. *Physiol Genomics* 2006;28:114–28. [PubMed: 16940433]
18. Jeffrey KL, Brummer T, Rolph MS, et al. Positive regulation of immune cell function and inflammatory responses by phosphatase PAC-1. *Nat Immunol* 2006;7:274–83. [PubMed: 16474395]
19. Rutella S, Bonanno G, Procoli A, et al. Hepatocyte growth factor favors monocyte differentiation into regulatory interleukin (IL)-10⁺⁺IL-12^{low/neg} accessory cells with dendritic-cell features. *Blood* 2006;108:218–27. [PubMed: 16527888]
20. Takeda A, Goolsby C, Yaseen NR. NUP98–HOXA9 induces long-term proliferation and blocks differentiation of primary human CD34⁺ hematopoietic cells. *Cancer Res* 2006;66:6628–37. [PubMed: 16818636]
21. Schebesta A, McManus S, Salvaggio G, Delogu A, Busslinger GA, Busslinger M. Transcription factor Pax5 activates the chromatin of key genes involved in B cell signaling, adhesion, migration, and immune function. *Immunity* 2007;27:49–63. [PubMed: 17658281]
22. Taylor MW, Tsukahara T, Brodsky L, et al. Changes in gene expression during pegylated interferon and ribavirin therapy of chronic hepatitis C virus distinguish responders from nonresponders to antiviral therapy. *J Virol* 2007;81:3391–401. [PubMed: 17267482]
23. Sanda C, Weitzel P, Tsukahara T, et al. Differential gene induction by type I and type II interferons and their combination. *J Interferon Cytokine Res* 2006;26:462–72. [PubMed: 16800785]
24. Indraccolo S, Pfeffer U, Minuzzo S, et al. Identification of genes selectively regulated by IFNs in endothelial cells. *J Immunol* 2007;178:1122–35. [PubMed: 17202376]
25. Barone SR, Pontrelli LR, Krilov LR. The differentiation of classic Kawasaki disease, atypical Kawasaki disease, and acute adenoviral infection: use of clinical features and a rapid direct fluorescent antigen test. *Arch Pediatr Adolesc Med* 2000;154:453–6. [PubMed: 10807294]
26. Burns JC, Mason WH, Glode MP, et al. Clinical and epidemiologic characteristics of patients referred for evaluation of possible Kawasaki disease. United States Multicenter Kawasaki Disease Study Group. *J Pediatr* 1991;118:680–6. [PubMed: 2019921]

27. Burns JC, Glode MP, Clarke SH, Wiggins J Jr, Hathaway WE. Coagulopathy and platelet activation in Kawasaki syndrome: identification of patients at high risk for development of coronary artery aneurysms. *J Pediatr* 1984;105:206–11. [PubMed: 6235335]
28. Zimmerman GA, Weyrich AS. Signal-dependent protein synthesis by activated platelets: new pathways to altered phenotype and function. *Arterioscler Thromb Vasc Biol* 2008;28:s17–24. [PubMed: 18296586]
29. Theilgaard-Monch K, Porse BT, Borregaard N. Systems biology of neutrophil differentiation and immune response. *Curr Opin Immunol* 2006;18:54–60. [PubMed: 16343884]
30. Bhattacharya D, Cheah MT, Franco CB, et al. Transcriptional profiling of antigen-dependent murine B cell differentiation and memory formation. *J Immunol* 2007;179:6808–19. [PubMed: 17982071]
31. Kuo TC, Shaffer AL, Haddad J Jr, Choi YS, Staudt LM, Calame K. Repression of BCL-6 is required for the formation of human memory B cells in vitro. *J Exp Med* 2007;204:819–30. [PubMed: 17403935]
32. Imayoshi M, Yamamoto S, Watanabe M, et al. Expression of CD180, a Toll-like receptor homologue, is up-regulated in children with Kawasaki disease. *J Mol Med* 2006;84:168–74. [PubMed: 16389554]
33. Luckey CJ, Bhattacharya D, Goldrath AW, Weissman IL, Benoist C, Mathis D. Memory T and memory B cells share a transcriptional program of self-renewal with long-term hematopoietic stem cells. *Proc Natl Acad Sci USA* 2006;103:3304–9. [PubMed: 16492737]
34. Leung DY, Siegel RL, Grady S, et al. Immunoregulatory abnormalities in mucocutaneous lymph node syndrome. *Clin Immunol Immunopathol* 1982;23:100–12. [PubMed: 7047030]
35. Moens L, Wuyts M, Meyts I, De Boeck K, Bossuyt X. Human memory B lymphocyte subsets fulfill distinct roles in the anti-polysaccharide and anti-protein immune response. *J Immunol* 2008;181:5306–12. [PubMed: 18832686]
36. Randall RE, Goodbourn S. Interferons and viruses: an interplay between induction, signalling, antiviral responses and virus countermeasures. *J Gen Virol* 2008;89:1–47. [PubMed: 18089727]
37. Shike H, Shimizu C, Kanegaye J, Foley JL, Burns JC. Quantitation of adenovirus genome during acute infection in normal children. *Pediatr Infect Dis J* 2005;24:29–33. [PubMed: 15665707]
38. Arnold JC, Singh KK, Spector SA, Sawyer MH. Undiagnosed respiratory viruses in children. *Pediatrics* 2008;121:e631–7. [PubMed: 18310182]
39. de Zorzi A, Colan SD, Gauvreau K, Baker AL, Sundel RP, Newburger JW. Coronary artery dimensions may be misclassified as normal in Kawasaki disease. *J Pediatr* 1998;133:254–8. [PubMed: 9709715]
40. Alizadeh A, Eisen M, Davis RE, et al. The lymphochip: a specialized cDNA microarray for the genomic-scale analysis of gene expression in normal and malignant lymphocytes. *Cold Spring Harb Symp Quant Biol* 1999;64:71–8. [PubMed: 11232339]
41. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 1998;95:14863–8. [PubMed: 9843981]
42. Saldanha AJ. Java Treeview—extensible visualization of microarray data. *Bioinformatics* 2004;20:3246–8. [PubMed: 15180930]

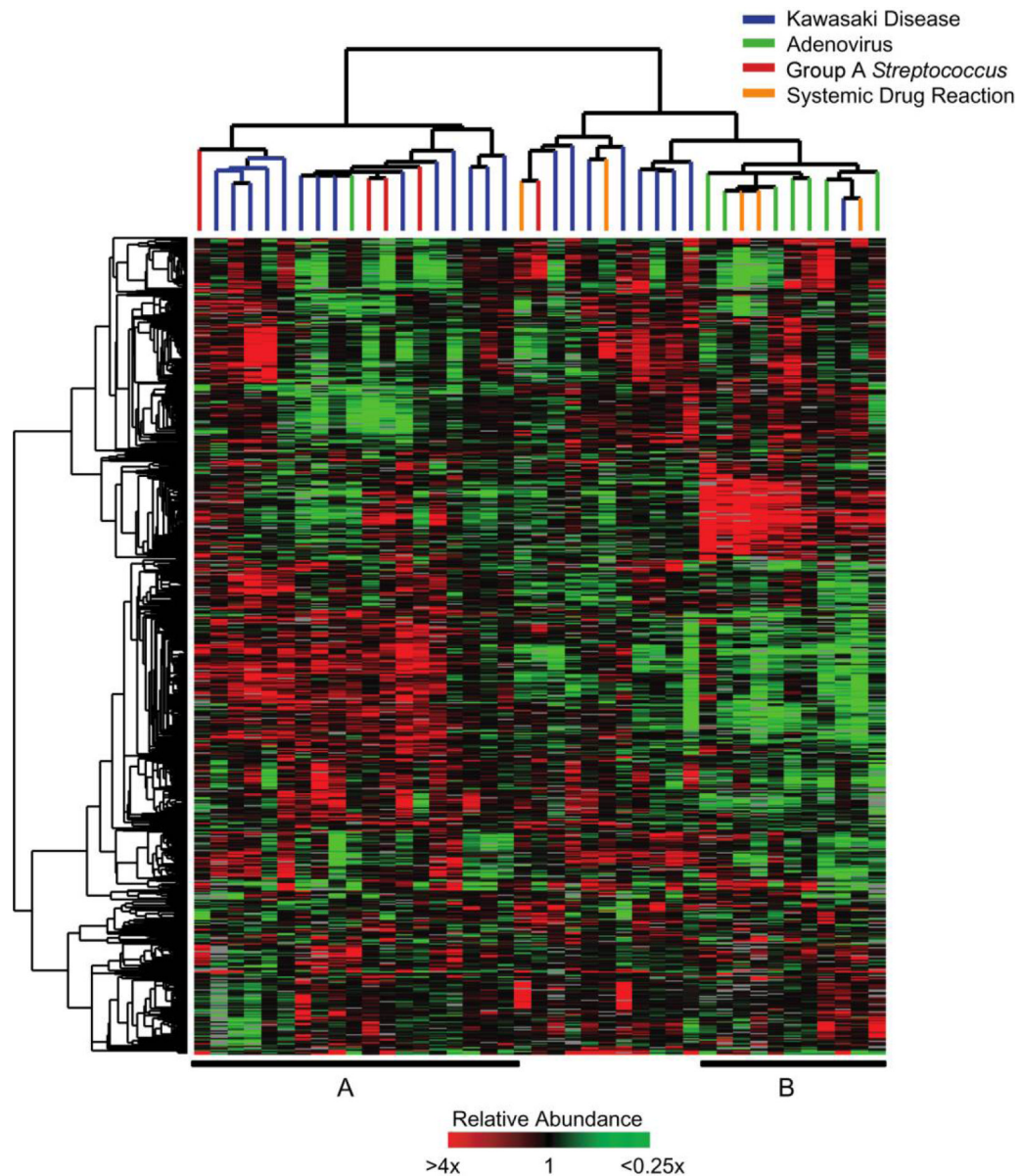


Figure 1.

Interindividual variation in patients with Kawasaki disease or febrile illnesses with similar clinical presentation. Transcripts with a ≥ 3 -fold change from the median expression level in ≥ 2 of 41 samples are shown. Genes and samples are organized using hierarchical clustering; each row represents a single transcript (complementary DNA clone), and each column a sample. Color scale at bottom reflects range of expression relative to median expression for each transcript. Black indicates median level of expression; red, greater expression; green, less expression; gray, missing data. Horizontal bars at bottom indicate sample clusters referred to in the text.

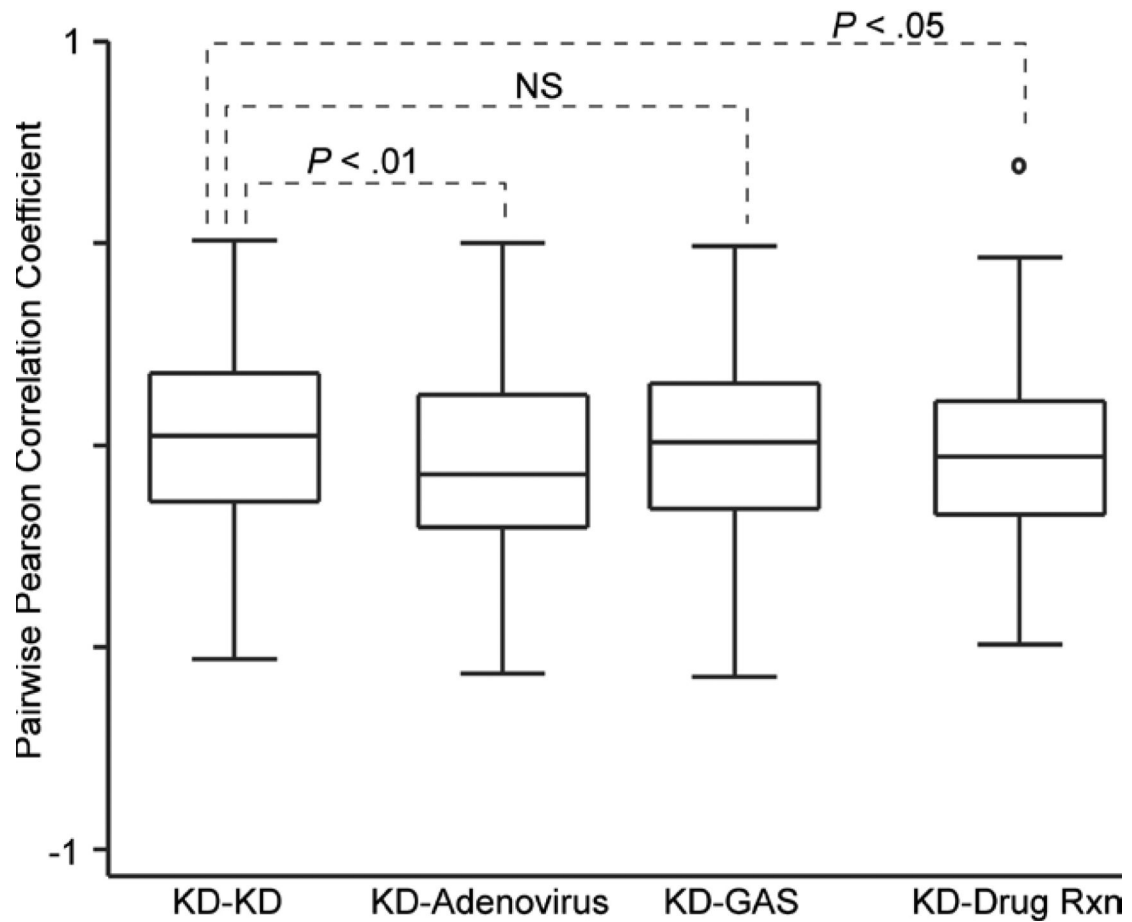


Figure 2.

Comparison of whole-blood transcription profiles from patients with Kawasaki disease (KD) and control groups. Correlation coefficients were calculated for each KD sample and each control sample in the data set, and they were plotted by comparison group. The central horizontal line in each box represents the median correlation coefficient, the upper and lower borders of each box indicate the 25th–75th percentiles, and the whiskers indicate the 10th–90th percentiles. *Open circle*, data point outside the 90th percentile. *P* values were calculated with the nonparametric Spearman rank sum test; NS, not significant. GAS, group A β -hemolytic *Streptococcus* infection; Rxn, reaction.

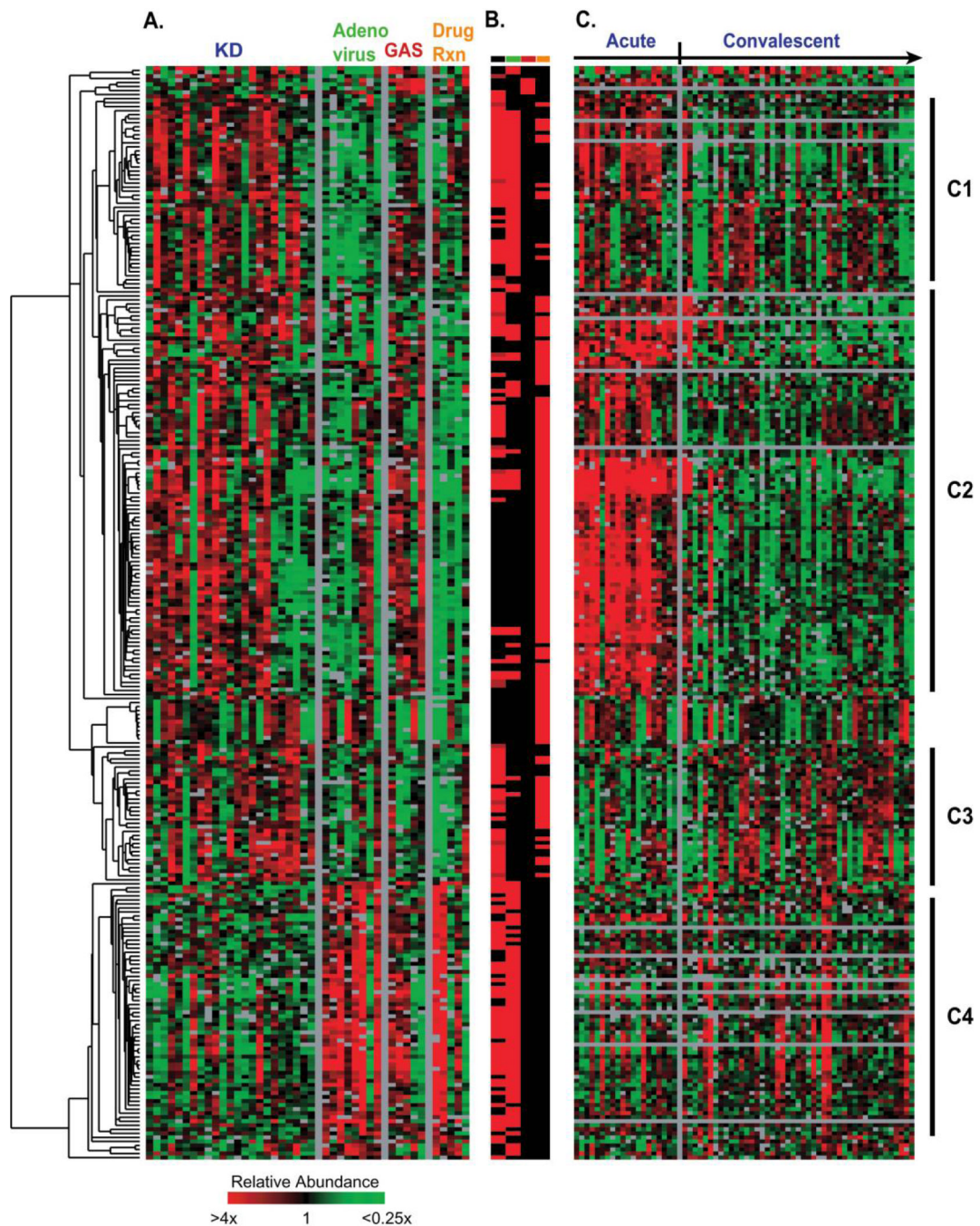


Figure 3.

Kawasaki disease (KD)-associated transcripts. Transcripts with significant differences in abundance were organized using hierarchical clustering. Color scale at bottom reflects range of expression relative to median expression for each transcript; black indicates median level of expression; red, greater expression; green, less expression; gray, missing data. C1–C4 and vertical bars denote gene clusters discussed in the text. *A*, Samples sorted by diagnosis and day of illness. *B*, Diagnosis-specific differences in transcript level. From left to right, columns represent patients with KD versus all control subjects and patients with adenovirus infection, group A *Streptococcus* (GAS) infection, or drug reactions (Rxn). Red indicates statistically significant difference compared with patients with KD; black, indicates no statistically

significant difference. C, Whole-blood transcript levels from 20 patients with KD sampled in both acute and convalescent phases of disease [4].

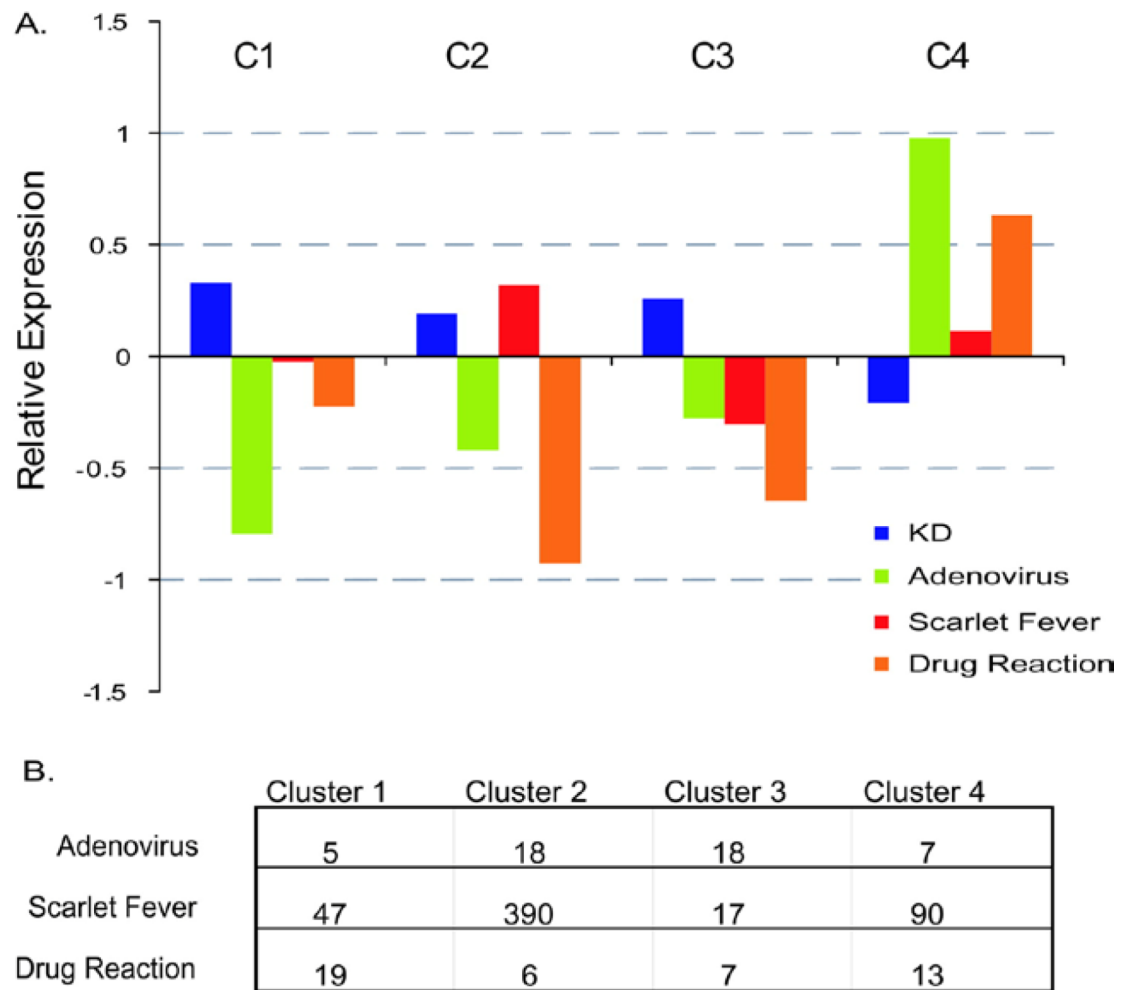


Figure 4.

Average expression of gene clusters associated with Kawasaki disease (KD). *A*, Average expression of gene clusters identified in figure 3. *B*, Estimated number of samples (both KD and control) needed to identify an existing difference in expression levels between patients with KD and a specific control group, given the average difference in expression observed in this data set. Cluster numbers correspond to clusters identified in figure 3. The type I error (α) was set at .05; power (β) at .8.

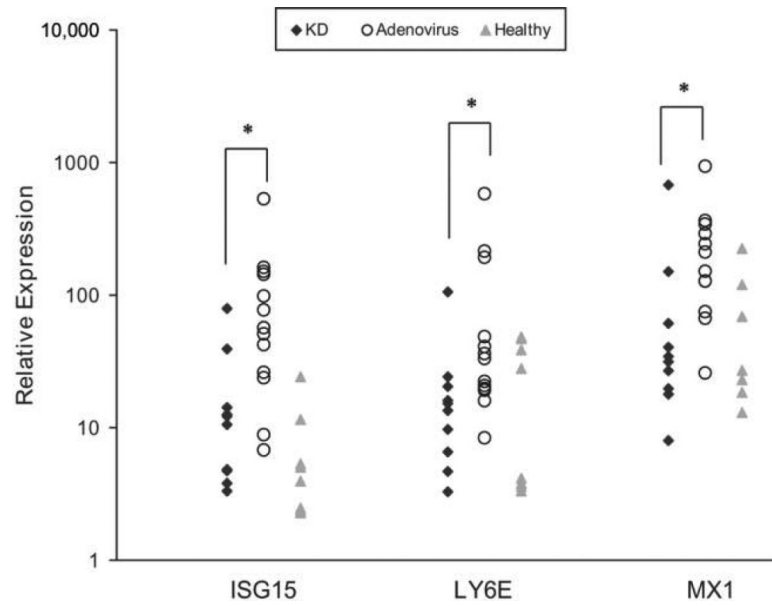


Figure 5. Interferon-responsive gene transcript levels. Transcript levels in whole blood were measured using TaqMan real-time reverse-transcription polymerase chain reaction and normalized to levels of *TAF1B* transcript. KD, Kawasaki disease. * $P < .05$.

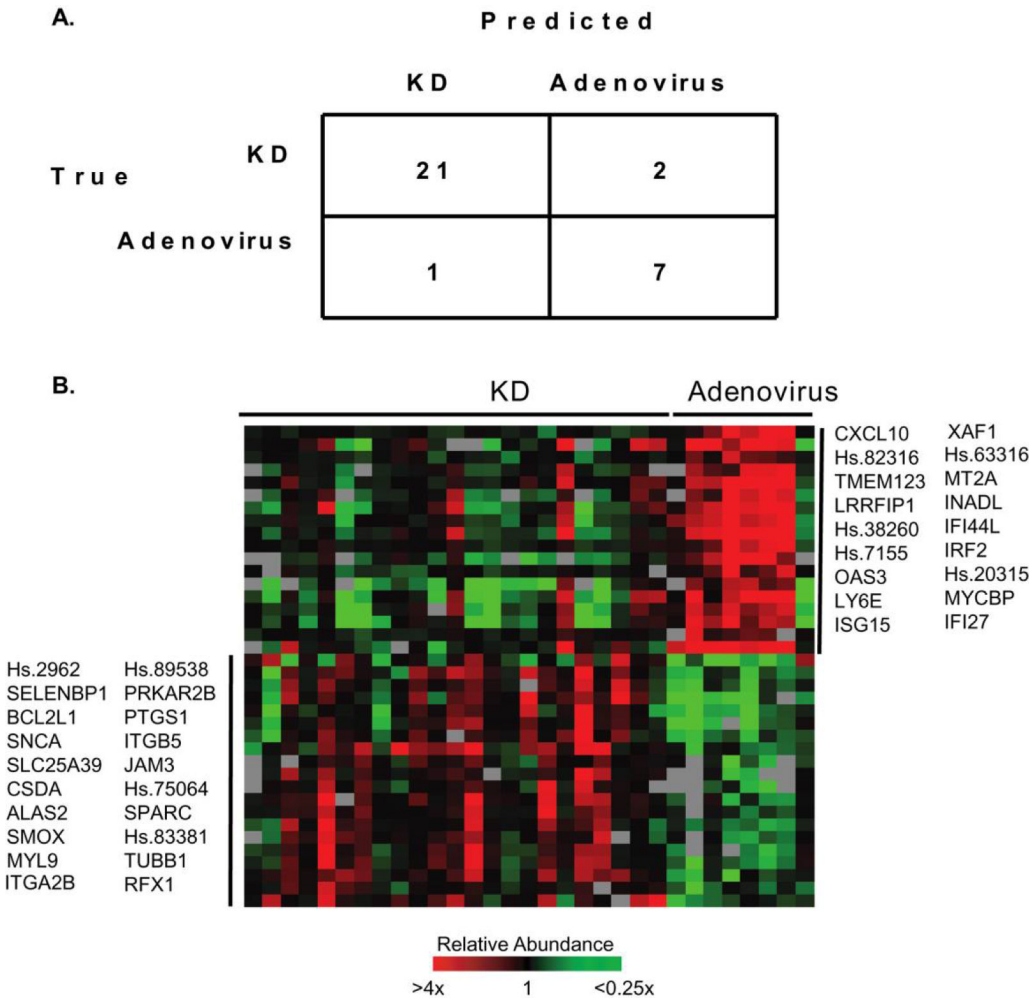


Figure 6. Prediction of Kawasaki disease (KD) and adenovirus infection using transcript abundance levels. The prediction analysis of microarrays algorithm was used to identify 38 transcripts that predicted KD infections in a 10-fold cross-validation scheme. *A*, Predicted and actual classification. *B*, Gene expression patterns for the 38 transcripts used to predict infection class. Gene symbols or Unigene cluster identification numbers are presented for each transcript. Samples from patients with KD and subjects with adenovirus infection are sorted by study identification number. Color scale at bottom reflects range of expression relative to median expression for each transcript; black indicates median level of expression; red, greater expression; green, less expression; gray, missing data.

Table 1

Clinical Characteristics of Study Patients with Kawasaki Disease (KD) and Febrile Control Subjects

Characteristic	Patients with KD (<i>n</i> = 23)	Control subjects, by illness		
		Adenovirus infection (<i>n</i> = 8)	Group A <i>Streptococcus</i> infection (<i>n</i> = 5)	Drug reaction (<i>n</i> = 5)
Age at illness onset, years	2.5 (0.4–12.8)	2.6 (0.5–8.1)	3.8 (1.9–6.8)	8.6 (0.9–15.9)
Sex	15 male, 8 female	5 male, 3 female	3 male, 2 female	5 male
Day of illness ^a	6 (4–12)	5.5 (3–11)	3 (2–21) ^b	6 (4–14)
CA status	13 N, 4 A, 6 D	NA	NA	NA
IVIG response	18 R, 3 NR, 2 LT	NA	NA	NA
White blood cell count, ×10 ³ /mm ³	13.2 (6.5–23.5)	14.5 (6–22.5)	13.6 (6.8–14.8)	5.7 (3.2–15.3) ^b
Neutrophils, %	49 (12–69)	50 (29–74)	60 (20–74)	47 (21–72)
Bands, %	12 (0–53)	8 (0–26)	8 (5–66)	9 (4–15)
Absolute neutrophil count, ×10 ³ /mm ³	8.6 (2.8–19.0)	6.6 (4.0–16.9)	11.0 (4.5–12.1)	2.8 (1.8–4.3) ^b
Monocytes, %	5 (0–11)	11 (2–15) ^b	6 (3–13)	5 (4.6–6.0)
Eosinophils, %	3 (0–11)	0 (0–1) ^b	1 (0–11)	4 (0–22)
Lymphocytes, %	27 (8–53)	19 (13–65)	11 (5–25) ^b	37 (17–43)
zHg	1.0 (–3.5 to 3.8)	–0.4 (–1.6 to 0.0)	–1.0 (–2.5 to 0.7)	1.0 (–1.1 to 2.7) ^b
Hematocrit, %	32.7 (28.9–42.7)	33.6 (32.7–37.2)	35.2 (30.2–38.8)	39.9 (36.5–45.7) ^b
ESR, mm/h	74 (23–140)	57.5 (13–88)	5	10 (4–16) ^b
CRP mg/dL	8.3 (0.8–36.9)	3.3 (2.0–7.8) ^b	2.4	5 (1.4–7.5)
Platelet count, ×10 ³ /mm ³	443 (217–595)	304 (174–455) ^b	298 (209–376) ^b	300 (148–586)
ALT, IU/L	52 (3–550)	13 (0.3–66) ^b	ND	15 (6–31) ^b
GGT, IU/L	75 (17–233)	19 (12–44) ^b	ND	34.5 (27–50)

NOTE. Data are medians (ranges), unless otherwise indicated. A, aneurysm; ALT, alanine aminotransferase; CA, coronary artery; CRP, C-reactive protein; D, dilated; ESR, erythrocyte sedimentation rate; GGT, γ -glutamyl transferase; IVIG, intravenous immunoglobulin; LT, treated after day 10; N, normal; NA, not applicable; ND, not done; NR, IVIG nonresponder; R, IVIG responder; zHg, z scores for hemoglobin level, normalized by age.

^aDay on which samples were obtained.

^b*P* < .05 for comparison with KD group (rank sum test).