Supporting Information

Hu et al. 10.1073/pnas.1302968110

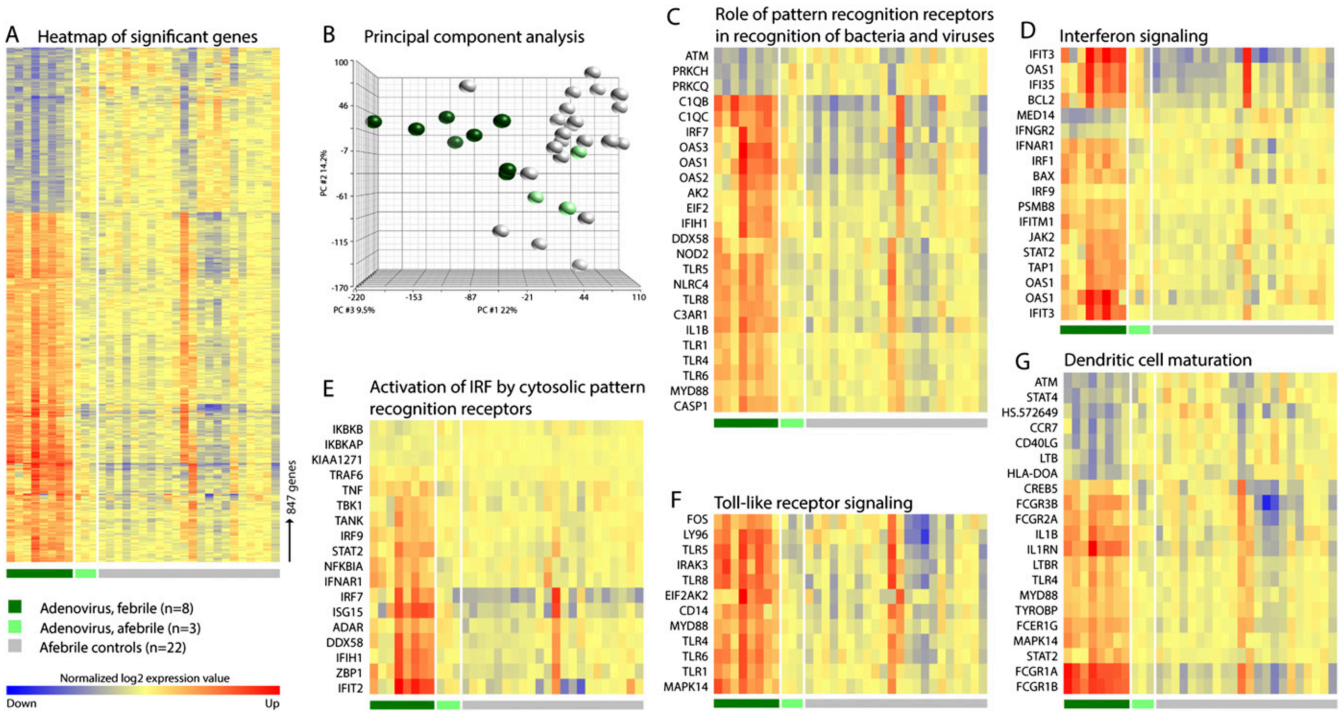


Fig. S1. Blood transcriptional profiles of febrile adenovirus-positive children are distinctively different from the profiles of healthy children and afebrile children with adenovirus infections. Microarray analysis was conducted on RNA extracted from blood samples of 11 children with confirmed adenovirus infection (8 febrile and 3 afebrile children) and 22 afebrile virus-negative children. (*A*) Hierarchical clustering of all probe sets with a statistically significant and greater than twofold difference between adenovirus-positive febrile children and virus-negative afebrile controls [false discovery rate (FDR) at 5%]. (*B*) Principal component analysis of differentially expressed genes, with each oval representing one child. (*C*–*G*) Hierarchical clustering of differentially expressed genes from *A* according to their expression intensity in five Ingenuity canonical pathways of particular interest, which are among the most strongly activated pathways in adenovirus-positive febrile children. Each row represents a gene with expression value that is normalized to the mean of the afebrile virus-negative control group. Gene names are listed to the left. Each column represents one individual. Red represents up-regulation, and blue represents down-regulation.

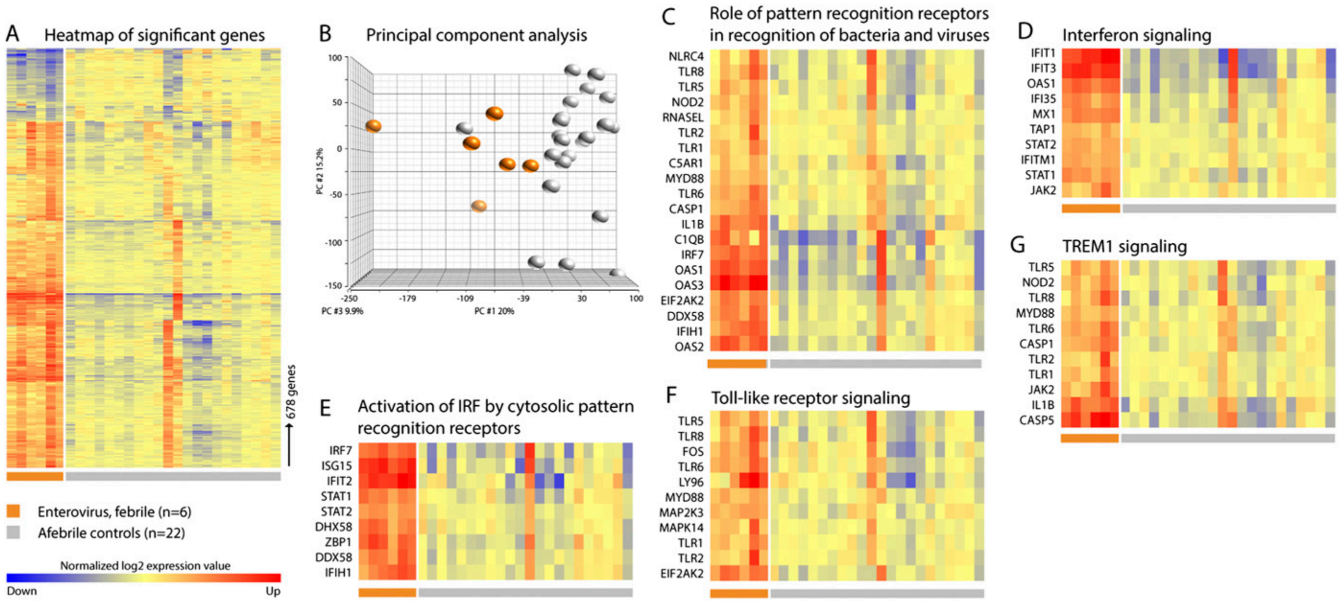
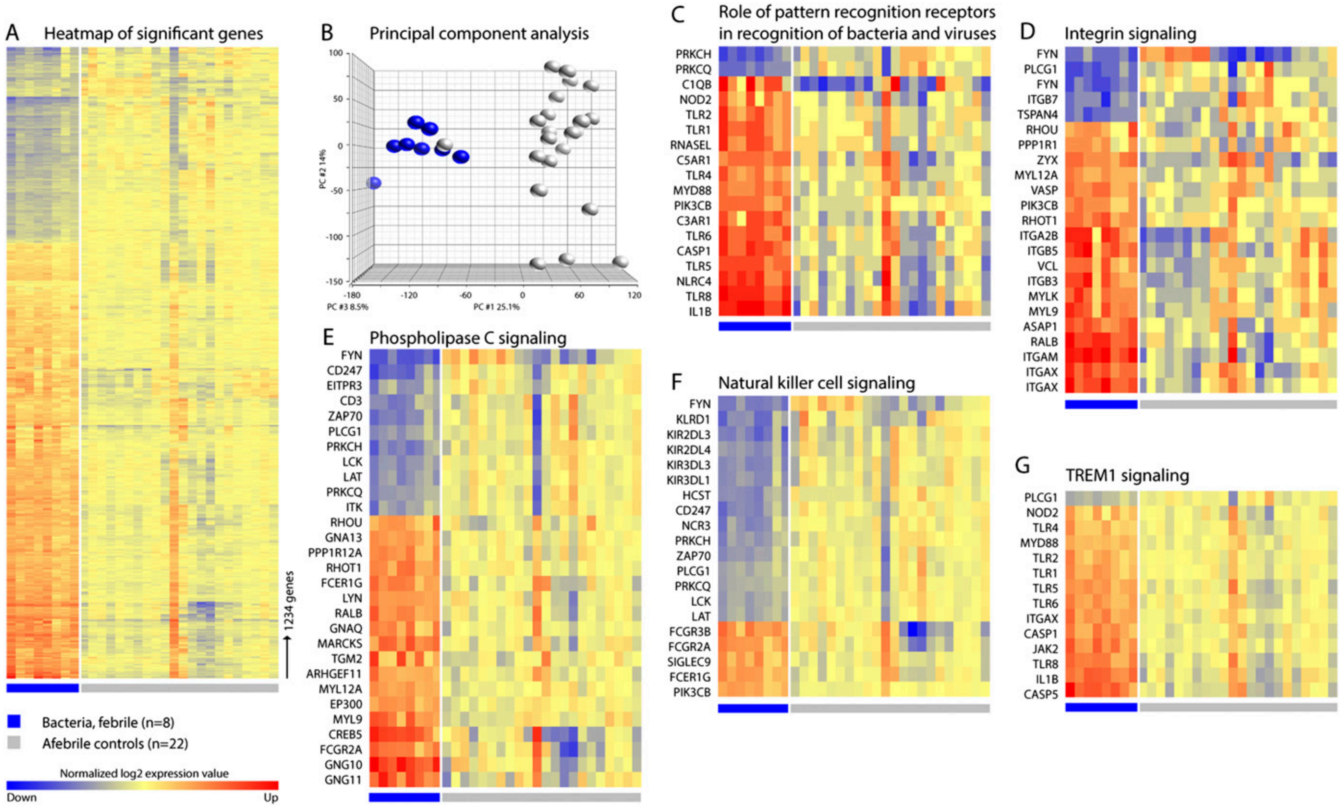


Fig. S2. Blood transcriptional profiles of enterovirus-positive febrile children are distinctly different the profiles of virus-negative afebrile children. Microarray analysis was conducted on RNA extracted from whole-blood samples of 6 enterovirus-positive febrile and 22 virus-negative afebrile children. (*A*) Hierarchical clustering of all probe sets with a statistically significant greater than twofold difference between enterovirus-positive febrile children and virus-negative afebrile controls ($P < 0.05$, FDR at 20%). (*B*) Principal component analysis of differentially expressed genes, with each oval representing one child. (*C*–*G*) Hierarchical clustering of differentially expressed genes in *A* according to their expression intensity in five Ingenuity canonical pathways of particular interest, which are among the most strongly activated pathways in febrile children with enterovirus infection. Each row represents a gene with expression value that is normalized to the mean of the afebrile virus-negative control group. Gene names are listed to the left. Each column represents one individual. Red represents up-regulation, and blue represents down-regulation.


A Heatmap of significant genes

B Principal component analysis

C Role of pattern recognition receptors in recognition of bacteria and viruses

D Integrin signaling

E Phospholipase C signaling

F Natural killer cell signaling

G TREM1 signaling

Fig. S3. Blood transcriptional profiles of febrile children with acute bacterial infections are distinctively different from the profiles of virus-negative afebrile children. Microarray analysis was conducted on RNA extracted from whole-blood samples of 8 febrile children with confirmed bacterial infection and 22 virus-negative afebrile children. (*A*) Hierarchical clustering of all probe sets with a statistically significant greater than twofold difference between febrile children with acute bacterial infection and virus-negative afebrile controls (FDR at 5%). (*B*) Principal component analysis of differentially expressed genes, with each oval representing one child. (*C*–*G*) Hierarchical clustering of differentially expressed genes in *A* according to their expression intensity in five Ingenuity canonical pathways of particular interest, which are among the most strongly activated pathways in febrile children with bacterial infection. Each row represents a gene with expression value that is normalized to the mean of the afebrile virus-negative control group. Gene names are listed to the left. Each column represents one individual. Red represents up-regulation, and blue represents down-regulation.

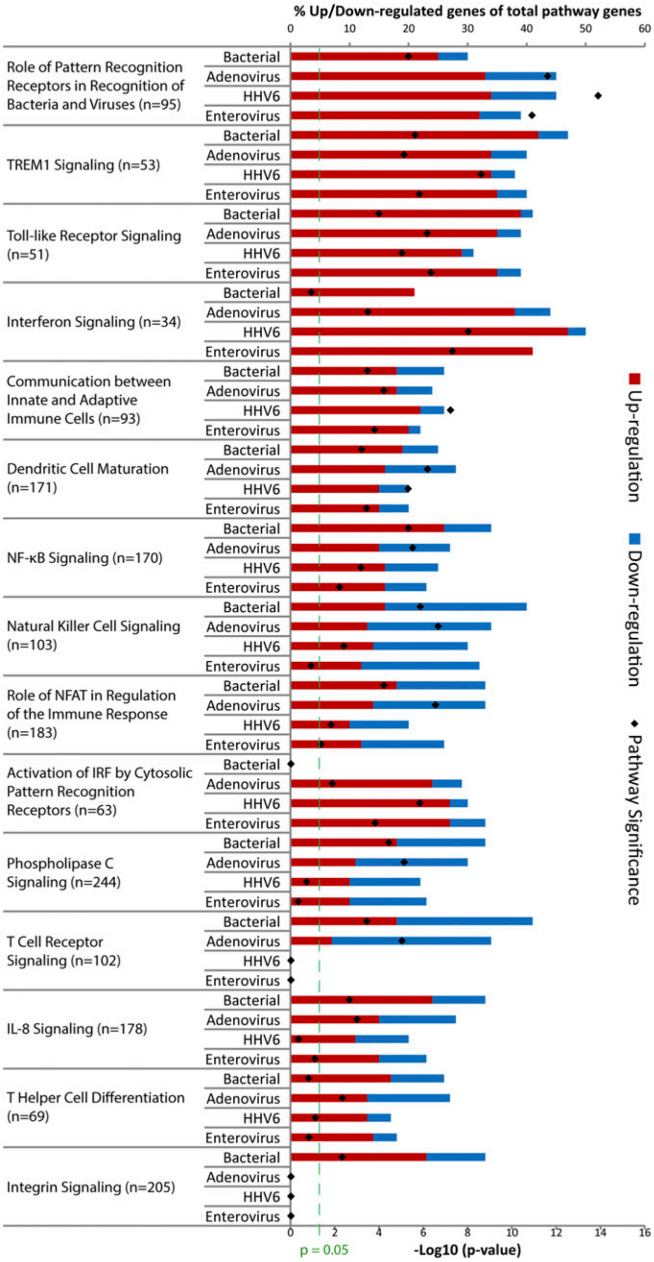
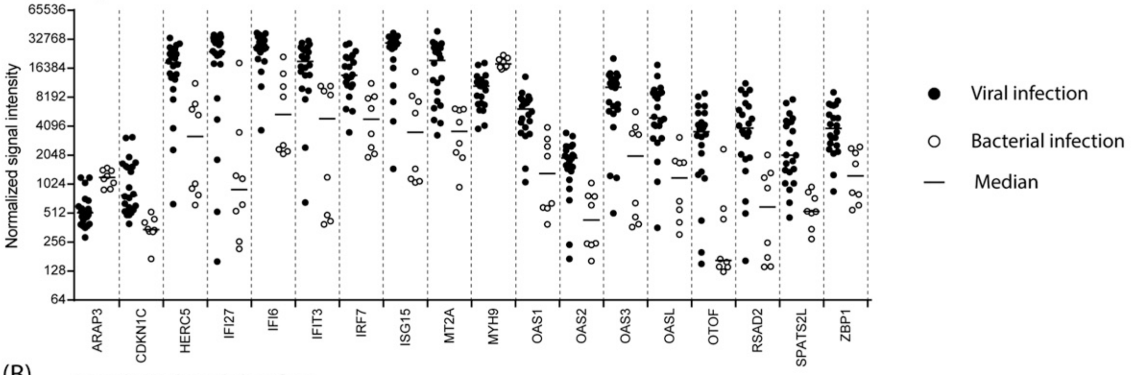
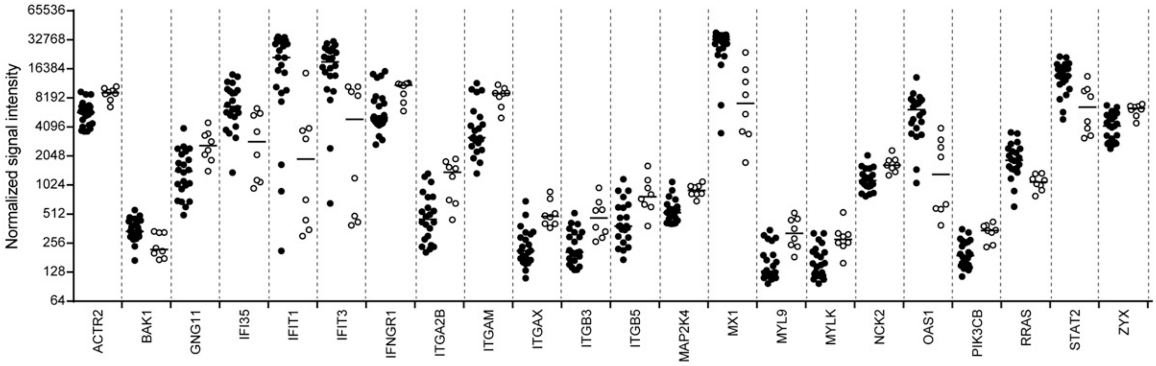


Fig. S4. Selected significantly up- and down-regulated Ingenuity canonical pathways identified for febrile children positive for adenovirus, human herpesvirus 6 (HHV-6), or enterovirus and febrile children with acute bacterial infections. The pathways were arranged in ascending order by average *P* value of four infections for a pathway (i.e., the most significantly up- or down-regulated pathway is at the top).

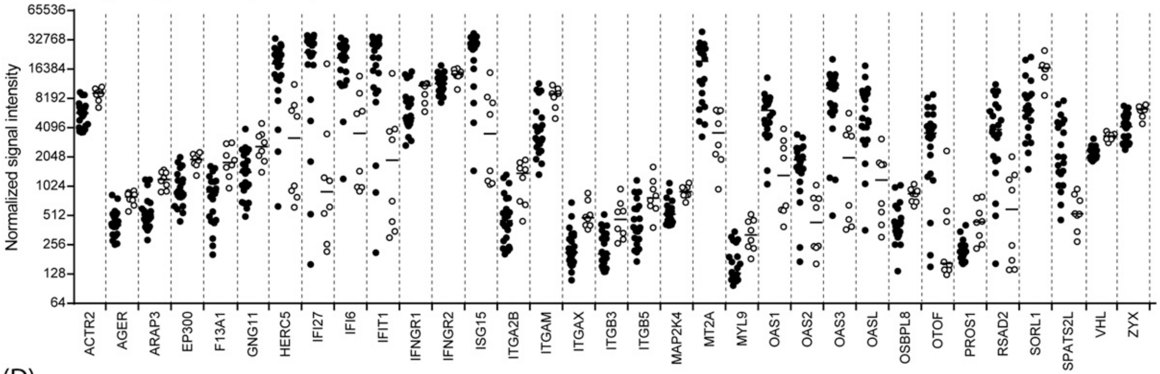
(A) 18 gene-based classifiers



(B) 22 pathway-based classifiers



(C) 33 hybrid gene- and pathway-based classifiers



(D) RT-qPCR validation data for 9 classifiers

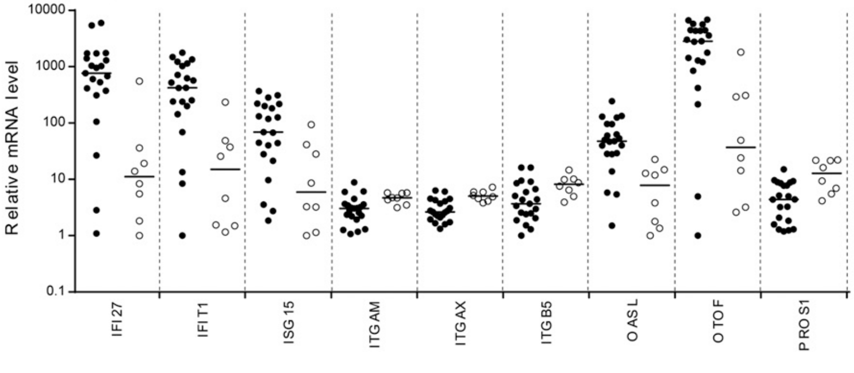


Fig. S5. Quantile-normalized raw signal intensity of the classifier probes in 30 febrile children (22 virus-positive children and 8 children with acute bacterial infection) in our study: (*A*) 18 classifiers identified from 260 viral- and 1,321 bacterial-specific probes, (*B*) 22 classifiers identified from 34 genes in the Ingenuity IFN signaling pathway and 205 genes in the Ingenuity integrin signaling pathway, (*C*) 33 classifiers identified from using both gene-level and pathway-based approaches, and (*D*) relative expression data of nine key classifier genes, which were generated in quantitative RT-PCR (RT-qPCR) validation assays for 29 of 30 febrile children (one RNA sample with HHV-6 infection was not available for the assays). The expression level was calculated using ΔΔCt method and normalized to endogenous reference GAPDH. Each dot represents one sample.

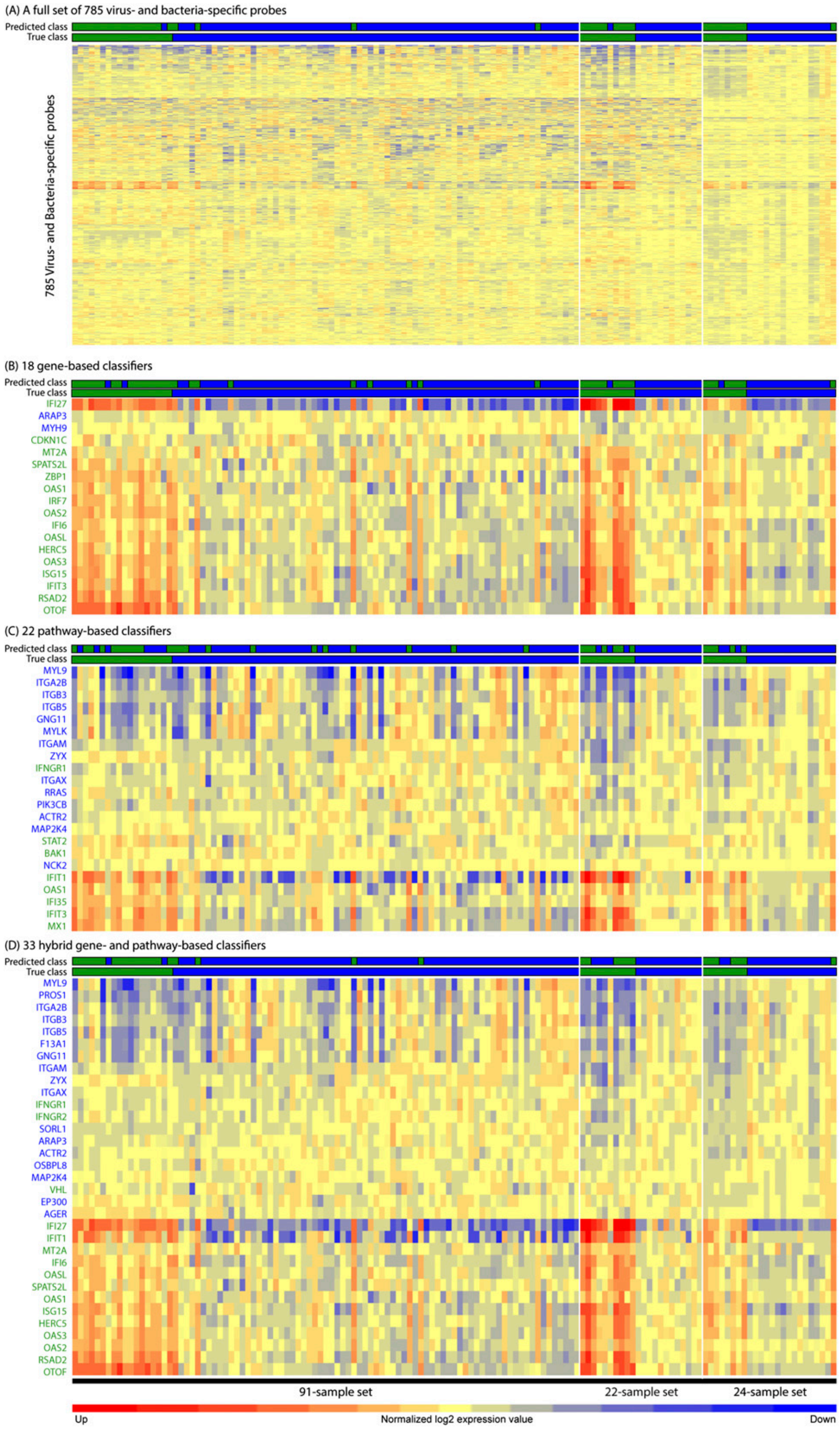


Fig. S6. Validation of three sets of classifier probes discriminating virus-positive febrile children from febrile children with acute bacterial infection using three independent cohorts of patients. The 91-sample validation set included 18 subjects with influenza A, 29 subjects with *Escherichia coli*, 31 subjects with *Staphylococcus aureus*, and 13 subjects with *Streptococcus pneumoniae*. This set was profiled with the Affymetrix Human Genome U133A Array platform. The 22-sample validation set consisted of seven children with influenza A, three children with influenza B, six children with *S. aureus*, and six children with *S. pneumoniae*, and it was profiled with the Affymetrix Human Genome U133 Plus 2.0 Array platform. The 24-sample validation set was composed of 5 subjects with influenza A, 3 subjects with influenza B, 13 subjects with *S. aureus*, and 3 subjects with *S. pneumoniae*, and it was profiled with the Illumina Sentrix Human-6 Expression BeadChip platform. Overall prediction accuracy was 95% (130/137), 88% (120/137), 88% (121/137), and 91% (124/137) with (*A*) a full set of

Legend continued on following page

785 probes overlapped across all three datasets with 1,581 virus- and bacteria-specific probes, (*B*) gene-based classifiers ($n = 18$), (*C*) pathway-based classifiers ($n = 22$), and (*D*) hybrid gene- and pathway-based classifiers ($n = 33$), respectively. Patient groups are indicated by colored stripes at the top of the heatmap. True class indicates status determined by virus-specific PCR assays and bacterial cultures, and it was assigned to these cases in the original study (1). Predicted class was determined by prediction made with the classifier probes, and it is labeled with green for viral or blue for bacterial infection. Gene names in green signify genes selected from the viral-specific gene set (or the IFN signaling pathway genes), and gene names in blue represent genes selected from the bacterial-specific gene set (or the integrin signaling pathway genes). Expression values presented in the heatmaps were normalized to the mean of those cases with bacterial infection within each dataset. Heatmap rows are gene probes, whereas columns are individual subjects.

1. Ramilo O, et al. (2007) Gene expression patterns in blood leukocytes discriminate patients with acute infections. *Blood* 109(5):2066–2077.

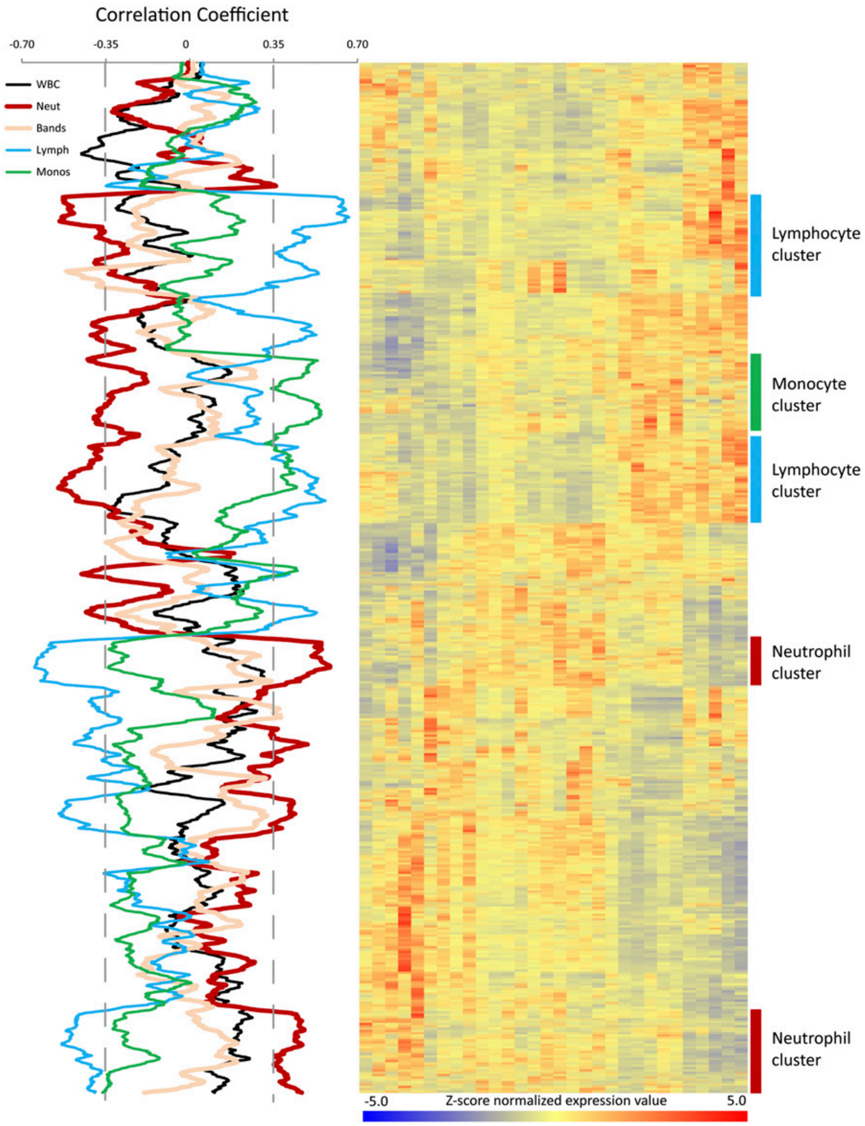


Fig. S7. Correlation of transcriptional changes and leukocyte subpopulations in febrile young children. Whole-blood samples were from 30 febrile young children with confirmed viral/bacterial infection. Probe sets with at least a 1.5-fold change in level of expression over virus-negative afebrile controls are shown. The expression pattern of the corresponding 4,716 probe sets is displayed in hierarchical cluster format, where rows represent genes and columns represent individual samples. Correlation coefficients were calculated between the expression level of each probe set and white blood cell counts (total, neutrophil, lymphocyte, bands, and monocyte counts) across 30 patients. The correlation values are plotted as moving averages of 50 probe sets (along the vertical axis). Dashed lines indicate the lowest values of correlation coefficients significant (adjusted $P < 0.05$) for each parameter.

Table S1. Individual virus- and bacteria-specific profile gene probes with strongest effects

		Adenovirus vs. control		HHV-6 vs. control		Enterovirus vs. control		Bacteria vs. control	
Gene symbol	Accession	Fold change	Adjusted *P* value	Fold change	Adjusted *P* value	Fold change	Adjusted *P* value	Fold change	Adjusted *P* value
RETN	NM_020415.2	**6.626**	***0.001***	1.746	0.273	1.554	0.475	1.997	0.172
OLFM4	NM_006418.3	**5.776**	***0.011***	1.029	0.655	1.570	0.539	3.150	0.117
LYZ	NM_000239.1	**2.130**	***0.011***	−1.050	0.622	−1.225	0.530	1.045	0.493
RPS6KA5	NM_004755.2	−**1.701**	***0.037***	−1.179	0.467	1.027	0.672	1.223	0.331
TSPYL2	NM_022117.1	−**1.692**	***0.002***	−1.113	0.437	−1.197	0.390	−1.250	0.160
ITPR1	NM_002222.4	−**1.625**	***0.008***	−1.143	0.415	1.092	0.580	1.345	0.120
CCL8	NM_005623.2	2.353	0.104	**12.885**	***0.000***	2.003	0.379	1.059	0.505
CCL2	NM_002982.3	2.179	0.113	**9.752**	***0.000***	2.271	0.294	−1.197	0.448
LRRC50	NM_178452.3	1.212	0.334	**3.426**	***0.000***	1.347	0.439	−1.074	0.470
VPS28	NM_016208.2	−1.069	0.434	−**1.625**	***0.015***	−1.048	0.640	1.084	0.423
NME4	NM_005009.2	1.108	0.363	−**1.562**	***0.020***	1.163	0.491	−1.116	0.376
MBNL3	NM_133486.1	1.305	0.346	1.349	0.440	**3.190**	***0.096***	1.649	0.228
HAGH	NM_005326.4	1.342	0.212	−1.103	0.567	**2.162**	***0.096***	1.100	0.449
TMPRSS9	NM_182973.1	1.236	0.238	1.048	0.610	**2.057**	***0.059***	1.301	0.226
KLC3	NM_177417.1	1.076	0.453	−1.142	0.498	**1.999**	***0.083***	1.307	0.245
ST6GALNAC4	NM_175039.3	1.286	0.195	−1.056	0.601	**1.945**	***0.079***	−1.045	0.482
PROS1	NM_000313.1	1.162	0.294	−1.029	0.628	1.265	0.382	**2.473**	***0.001***
SCGB1C1	NM_145651.2	−1.017	0.522	1.152	0.492	1.230	0.495	**2.463**	***0.004***
ARAP3	NM_022481.5	1.005	0.535	−1.104	0.545	1.268	0.446	**2.201**	***0.007***
STXBP5	NM_139244.2	−1.284	0.160	−1.281	0.244	1.094	0.595	**2.050**	***0.004***
GZMH	NM_033423.3	−1.929	0.106	−1.733	0.234	−2.088	0.260	−**5.258**	***0.003***
KIR2DL3	NM_014511.3	−1.016	0.529	−1.092	0.601	−1.642	0.346	−**3.389**	***0.007***
KIR2DL4	NM_002255.3	1.165	0.384	1.062	0.613	−1.408	0.405	−**2.646**	***0.008***
KIR3DL1	NM_013289.1	−1.018	0.524	−1.101	0.572	−1.354	0.433	−**2.572**	***0.008***
MT1X	NM_005952.2	1.107	0.414	1.291	0.309	1.170	0.545	−**1.787**	***0.041***
OSBPL5	NM_145638.1	−1.014	0.519	1.187	0.325	−1.101	0.566	−**1.767**	***0.007***

These probes were selected because they had significant up- or down-regulation in children positive for only one virus or with acute bacterial infection using an adjusted *P* value of 0.05. Gene transcription in children positive for only one virus and children with acute bacterial infection were each compared with gene transcription in afebrile virus-negative control children. Candidate genes were derived using the shrunken centroid algorithm procedure in the Prediction of Microarray Analysis tool from Stanford University (http://www-stat.stanford.edu/~tibs/SAM/). Probes were sorted within each virus/bacteria group by descending fold change (when up-regulated) or ascending fold change (when down-regulated). Bold and italic fonts indicate genes that differed significantly at adjusted *P* value < 0.05.

Table S2. Correlation in expression level between RT-qPCR and microarray results

Classifier	Pearson correlation of coefficient	*P* value
IFI27	0.722	2.17E-07
IFIT1	0.903	3.55E-15
ISG15	0.775	6.79E-09
ITGAM	0.925	4.29E-17
ITGAX	0.813	3.11E-10
ITGB5	0.886	6.86E-14
OASL	0.937	1.66E-18
OTOF	0.927	2.55E-17
PROS1	0.705	5.47E-07

Table S3. Demographics of 65 cases in the study

Patient ID	Pathogen	Fever	Sex	Age (mo)	Ethnicity	Antibiotics used	WBC (×1,000)	Neut (%)	Bands (%)	Lymph (%)	Mono (%)	WBC status by age-specific normal values	WBC status by cutoff of 15,000
9006	Adenovirus	Yes	Female	4	White	Yes	28.7	78	5	14	3	Increased	>15,000
9010	Adenovirus	Yes	Male	24	Black	Yes	25.6	66	9	15	10	Increased	>15,000
9021	Adenovirus	Yes	Female	2	Black	Yes	25.2	33	7	31	27	Increased	>15,000
9170	Adenovirus	Yes	Female	13	Black	Yes	18.2	78	10	3	6	Increased	>15,000
9203	Adenovirus	Yes	Female	9	White	Yes	30.9	54	12	23	10	Increased	>15,000
9282	Adenovirus	Yes	Female	17	White	Yes	17.1	72	10	10	7	Normal	>15,000
9289	Adenovirus	Yes	Female	15	Other	No	24.6	84	0	12	4	Increased	>15,000
9340	Adenovirus	Yes	Male	2	White	Yes	15.2	37	1	44	16	Normal	>15,000
9081	Adenovirus	No	Male	7	White								
9097	Adenovirus	No	Female	14	White								
9134	Adenovirus	No	Female	26	White								
9022	HHV-6	Yes	Female	10	Other	Yes	15.1	85	0	13	2	Normal	>15,000
9023	HHV-6	Yes	Female	12	Black	No	10.1	70	2	23	5	Normal	Not >15,000
9032	HHV-6	Yes	Male	3	Black	Yes	8.9	41	2	42	12	Normal	Not >15,000
9064	HHV-6	Yes	Female	7	Black	No	6.1	66	0	25	8	Normal	Not >15,000
9156	HHV-6	Yes	Male	3	Black	No	7	15	9	48	22	Normal	Not >15,000
9300	HHV-6	Yes	Male	2	Black	No	6.1	52	3	26	11	Normal	Not >15,000
9416	HHV-6	Yes	Male	12	White	No	5.7	47	0	41	11	Decreased	Not >15,000
9575	HHV-6	Yes	Male	25	White	Yes	5.6	42	1	48	6	Normal	Not >15,000
9437	HHV-6	No	Male	17	White								
9515	HHV-6	No	Male	18	Black								
9008	Enterovirus	Yes	Male	8	White	No	12.1	58	0	29	13	Normal	Not >15,000
9016	Enterovirus	Yes	Male	29	White	No	12.2	70	0	24	6	Normal	Not >15,000
9267	Enterovirus	Yes	Male	2	White	Yes	7.9	42	1	52	4	Normal	Not >15,000
9450	Enterovirus	Yes	Female	16	Black	Yes	15	49	1	44	5	Normal	>15,000
9467	Enterovirus	Yes	Male	32	Black	Yes	11.7	80	7	9	4	Normal	Not >15,000
9587	Enterovirus	Yes	Female	10	Black								
9087	Rhinovirus	No	Male	3	White								
9113	Rhinovirus	No	Male	7	White								
9118	Rhinovirus	No	Male	26	White								
9133	Rhinovirus	No	Male	5	White								
9149	Rhinovirus	No	Male	6	White								
9150	Rhinovirus	No	Female	12	White								
9151	Rhinovirus	No	Female	30	White								
9163	Rhinovirus	No	Female	32	White								
9298	*E. coli*	Yes	Female	30	White	Yes	30.6	84	0	11	5	Increased	>15,000
9359	Bacteria	Yes	Male	16	Black	Yes	20.3	73	1	22	2	Increased	>15,000
9397	MRSA	Yes	Male	3	Black	Yes	16	66	1	25	8	Normal	>15,000
9468	*Salmonella*	Yes	Male	25	White	No	17.8	55	0	26	5	Increased	>15,000
9501	*E. coli*	Yes	Male	4	White	Yes	12.2	67	2	20	8	Normal	Not >15,000
9519	MSSA	Yes	Male	32	White	Yes	16	65	0	25	6	Increased	>15,000
9523	MSSA	Yes	Female	10	Black	Yes	19.8	53	0	38	6	Increased	>15,000
9602	MRSA	Yes	Male	20	Black	Yes	17.1	72	0	20	8	Normal	>15,000
9050	Control	No	Male	10	White								
9051	Control	No	Male	17	Black								
9057	Control	No	Male	10	White								
9059	Control	No	Male	9	White								
9061	Control	No	Male	4	Black								
9062	Control	No	Male	17	White								
9066	Control	No	Female	6	White								
9067	Control	No	Female	4	White								
9075	Control	No	Male	3	White								
9091	Control	No	Male	11	White								
9093	Control	No	Male	10	White								
9110	Control	No	Female	12	Black								
9114	Control	No	Male	25	White								
9115	Control	No	Male	2	White								
9116	Control	No	Male	17	White								
9117	Control	No	Male	28	White								
9125	Control	No	Female	33	White								
9137	Control	No	Female	20	White								

Table S3. Cont.

Patient ID	Pathogen	Fever	Sex	Age (mo)	Ethnicity	Antibiotics used	WBC (×1,000)	Neut (%)	Bands (%)	Lymph (%)	Mono (%)	WBC status by age-specific normal values	WBC status by cutoff of 15,000
9146	Control	No	Male	22	White								
9147	Control	No	Female	9	White								
9187	Control	No	Female	5	White								
9294	Control	No	Male	11	White								

Lymph, lymphocyte; mono, monocyte; MRSA, methicillin-resistant *Staphylococcus aureus*; MSSA, methicillin-sensitive *Staphylococcus aureus*; neut, neutrophil; WBC, white blood cell.