

RESEARCH ARTICLE

Open Access



Whole blood expression profiling from the TREAT trial: insights for the pathogenesis of polyarticular juvenile idiopathic arthritis

Kaiyu Jiang¹, Laiping Wong¹, Ashley D. Sawle², M. Barton Frank³, Yanmin Chen¹, Carol A. Wallace^{4,5} and James N. Jarvis^{1,5*}

Abstract

Background: The Trial of Early Aggressive Therapy in Juvenile Idiopathic Arthritis (TREAT trial) was accompanied by a once-in-a-generation sample collection for translational research. In this paper, we report the results of whole blood gene expression analyses and genomic data-mining designed to cast light on the immunopathogenesis of polyarticular juvenile idiopathic arthritis (JIA).

Methods: TREAT samples and samples from an independent cohort were analyzed on Affymetrix microarrays and compared to healthy controls. Data from the independent cohort were used to validate the TREAT data. Pathways analysis was used to characterize gene expression profiles. Furthermore, we correlated differential gene expression with new information about functional regulatory elements within the genome to develop models of aberrant gene expression in JIA.

Results: There was a strong concordance in gene expression between TREAT samples and the independent cohort. In addition, rheumatoid factor (RF)-positive and RF-negative patients showed only small differences on whole blood expression profiles. Analysis of the combined samples showed 158 genes represented by 176 probes that showed differential expression between TREAT subjects at baseline and healthy controls. None of the differentially expressed genes were encoded within linkage disequilibrium blocks containing single nucleotide polymorphisms known to be associated with risk for JIA. Functional analysis of these genes showed functional associations with multiple processes associated with innate and adaptive immunity, and appeared to reflect overall suppression of STAT1–3/interferon response factor-mediated pathways.

Conclusions: Despite their limitations, whole blood expression profiles clearly distinguish children with polyarticular JIA from healthy controls. Whole blood expression profiles identify several immunologic pathways of biologic relevance that will need to be pursued in homogeneous cell populations in order to clarify mechanisms of pathogenesis.

Trial registration: ClinicalTrials.gov registry #NCT00443430, originally registered 2 March 2007 and last updated 30 May 2013.

Keywords: Juvenile idiopathic arthritis, Microarray, Whole blood, Gene expression, Pathogenesis

* Correspondence: jamesjar@buffalo.edu

¹Department of Pediatrics, University at Buffalo Jacobs School of Medicine and Biomedical Sciences, Clinical & Translational Research Center, 875 Ellicott St., Buffalo, NY, USA

⁵Genetics, Genomics, and Bioinformatics Program, University at Buffalo Jacobs School of Medicine and Biomedical Sciences, Buffalo, NY, USA

Full list of author information is available at the end of the article



Background

The Trial of Early Aggressive Therapy in Juvenile Idiopathic Arthritis (TREAT; ClinicalTrials.gov registry #NCT00443430) was an National Institute of Health (NIH) funded clinical trial [1] that compared two therapeutic regimens for treatment of newly-diagnosed polyarticular juvenile idiopathic arthritis (JIA). One treatment arm used methotrexate (MTX) as an initial therapy, while the other used a combined regimen of MTX, the tumor necrosis factor (TNF) inhibitor etanercept (ET), and oral prednisone. As part of the TREAT trial, whole blood was collected for RNA expression studies at specific time points during the course of the first year of therapy. Using whole blood expression data from the TREAT subjects, we have previously reported on the feasibility of developing expression-based prognostic biomarkers for children with the polyarticular form of JIA [2]. However, the whole blood gene expression data also provide a window through which we might also gain valuable insights into both the pathogenesis of JIA and the underlying biology of treatment response, both of which are currently poorly understood.

While whole blood (and buffy coat) expression data are inherently “noisy” (among other things, they reflect gene expressions in multiple cells and cell subsets), there are both technical [3, 4] and computational approaches [5] that can be used to improve the signal-to-noise ratio in whole blood expression data and derive meaningful mechanistic insights. Furthermore, projects like the NIH Encyclopedia of Functional DNA Elements (ENCODE) and Roadmap Epigenomics have provided investigators with a wealth of information from which to derive mechanistic insights from gene expression data. In this study, we used whole blood gene expression data derived from baseline samples from children enrolled in the TREAT study, coupled with data-mining from public resources, to identify novel pathways that contribute to JIA disease pathogenesis.

Methods

Patient samples

We have previously described the TREAT baseline samples [2]. Patients entered the TREAT trial between October 2007 and November 2009 [1]. All children fit international criteria for polyarticular-onset JIA [6]. Parents of these children gave informed written consent to provide samples for translational uses, and the protocol to use these specimens was approved by the TREAT study oversight committee and the University of Oklahoma Institutional Review Board. The patients included 19 boys and 45 girls, aged 2–14 years. ANA were detected in 30 of the girls and 10 of the boys. Approximately 2.5 ml of blood was collected at the time of enrollment (month 0) prior to treatment in a PAXgene tube (PreAnalytiX GmbH, Hilden, Germany). Samples were stored at -80 °C. A

summary of patient ages and characteristics is shown in Table 1.

Healthy control samples

Controls consisted of 8 healthy girls and 11 healthy boys between the ages of 7 and 13 years recruited from the OU Children’s Physicians General Pediatrics clinic. The protocol for obtaining these specimens was approved by the University of Oklahoma IRB (#13205). Anesthesia for the phlebotomy was provided using topical lidocaine/prilocaine solution. These samples are hereafter referred to as healthy children (HC).

Independent cohort

In addition to the above, whole blood PAXgene specimens were obtained from an independent cohort of children with newly diagnosed, polyarticular-onset JIA recruited from the University of Oklahoma Health Sciences Center pediatric rheumatology clinic. These children ranged in age from 3 years to 15 years and consisted of 4 boys and 6 girls. All patients in this cohort were rheumatoid factor (RF)-negative. These samples are hereafter referred to as the Oklahoma cohort (OK). Table 2 summarizes the characteristics of these patients.

All research procedures were carried out strictly following the IRB-approved protocols.

RNA processing

RNA was purified from whole blood PAXgene specimens using a PAXgene Blood RNA kit (Qiagen, Valencia, CA, USA) as recommended by the manufacturer, including a DNase (Qiagen) step to remove genomic DNA. Globin transcripts, which reduce labeling efficiency of whole blood cell RNA and decrease signal-to-noise ratios on microarrays [7], were reduced using GLOBINclear-Human (Ambion, Austin, TX, USA). Final RNA preparations were suspended in RNase-free water, quantified spectrophotometrically, and analyzed for RNA integrity by

Table 1 Phenotypic characteristics of patients at month 0 (baseline)

	Arm 1	Arm 2
Female	19 (63.3 %)	26 (76.5 %)
Male	11 (36.7 %)	8 (23.5 %)
Aged 2–6 years	7	5
Aged 7–11 years	11	15
Aged 12–16 years	12	14
Age (years), mean ± SD	10.5 ± 4.5	11.4 ± 3.6
RF-positive	9 (30 %)	12 (35.3 %)
ANA-positive	22 (73.3 %)	18 (52.9 %) ^a

Values are given as n (%) unless otherwise indicated

^aOne missing value

ANA Antinuclear antibody, RF rheumatoid factor, SD standard deviation

Table 2 Characteristics of JIA patients from the Oklahoma cohort

Patient number	Age	Sex	RF status	ANA status
1	9 years	F	Negative	Negative
2	12 years	F	Negative	Negative
3	9 years	M	Negative	Negative
4	11 years	F	Negative	Positive
5	7 years	F	Negative	Negative
6	3 years	F	Negative	Positive
7	8 years	M	Negative	Negative
8	14 years	M	Negative	Negative
9	15 years	F	Negative	Positive
10	4 years	M	Negative	Negative

ANA Antinuclear antibody, F female, JIA juvenile idiopathic arthritis, M male, RF rheumatoid factor

capillary gel electrophoresis (Agilent 2100 Bioanalyzer; Agilent Technologies, Palo Alto, CA, USA).

Microarray analysis

cRNA was produced from reverse transcribed cDNA using the Illumina® TotalPrep RNA Amplification Kit (Ambion, Inc., Austin, TX, USA), hybridized to Illumina WG-6 v3 or Illumina HT-12 v4 human whole genome microarrays, and stained according to the manufacturer's directions. Array hybridizations were undertaken in three separate batches. The first batch consisted of the 19 healthy controls and 26 baseline samples hybridized on Illumina WG-6 v3 arrays. The second batch consisted of the remaining patient samples hybridized to Illumina HT-12 v4 arrays. The independent cohort of OK samples were hybridized on Illumina WG-6 v3 slides. Gene microarray data have been made available to the scientific public (GEO Accession Number GSE55319).

Statistical analysis

All statistical analyses were carried out in R (www.r-project.org). To facilitate statistical analyses relative to healthy controls, it was necessary to combine data from different array batches. Due to the difference in the arrays it was necessary to create combined datasets using only those probes that were present on both array formats. Illumina probe IDs were used to identify 39,426 common probes. Datasets were variance stabilized and normalized using robust spline normalization via the *lumi* package [8, 9]. Prior to statistical analysis non-responding probes were filtered out of the datasets using the detection *p* value provided by the Illumina quality control metrics to eliminate probes not responding at higher than background levels.

Differential gene expression analysis was performed using the *limma* package [10, 11]. The false discovery rate (FDR) was estimated using the method described by

Benjamini and Hochberg [11]. Statistical significance of gene expression was determined at $FDR \leq 0.05$. Gene lists of interest were exported from R and uploaded to Ingenuity IPA (Ingenuity Systems, Inc., Redwood City, USA) for further functional analysis.

Network analyses of differentially expressed genes from whole blood expression profiles

We used the Ingenuity Pathway Analysis (IPA) software (Ingenuity Systems®, Redwood City, CA, USA; IPA Summer Release, June, 2015) for network analysis, with differentially expressed (DE) genes as input following the default setting of a maximum 35 genes per network. IPA network generation transforms the query genes into a set of relevant networks based on its extensive, curated Ingenuity Pathway Knowledge Base database [12, 13]. IPA utilizes a multi-stage heuristic algorithm in constructing networks through an iterative process that optimizes both interconnectivity and number of query genes under the constraint of network size. Briefly, IPA constructs networks using gene connectivity with other genes under the assumption that the gene with the highest number of connections is the most important and, thus, has the most influence. These connections represent regulatory interactions that may be either direct (e.g., the protein product of gene "A" directly regulates the expression of gene "B") or indirect (genes "A" and "B" are regulated by the same transcription factor). The last step of IPA network construction is score calculation using the Fisher Exact test on a hypergeometric distribution. The Fisher Exact test in IPA defines the null hypothesis as being a similar proportion of query genes map to a network in the same proportion as the entire reference gene set map to the network. A network score is derived from a *p* value ($\text{score} = -\log_{10} p \text{ value}$) that indicates the probability of the query genes (defined as target molecules in IPA) in a network being randomly associated within a connecting network. It is important to note that the network score does not infer network quality; it simply indicates the fitness between a reference network and the network of query genes.

Linking genetic and expression data

The recent completion of a genome-wide fine mapping study for JIA [14] extends the list of previously known disease-associated genetic variants [15] and provided us with the opportunity to determine whether there is a genetic linkage to gene expression in JIA. Using the bedtools program [16] we intersected DE genes identified in the comparison between JIA and HC with linkage disequilibrium (LD) blocks of JIA-associated single nucleotide polymorphisms (SNPs) extracted from the JIA fine mapping study [14] as well as other known regions of risk as reviewed by Hersh and Prahalad [15]. We

obtained LD block information from the SNAP database (<http://www.broadinstitute.org/mpg/snap>) using the 1000 genome project pilot1 and HapMap3 with a cutoff of $r^2 < 0.9$ and a distance limit of 500 bases.

Results

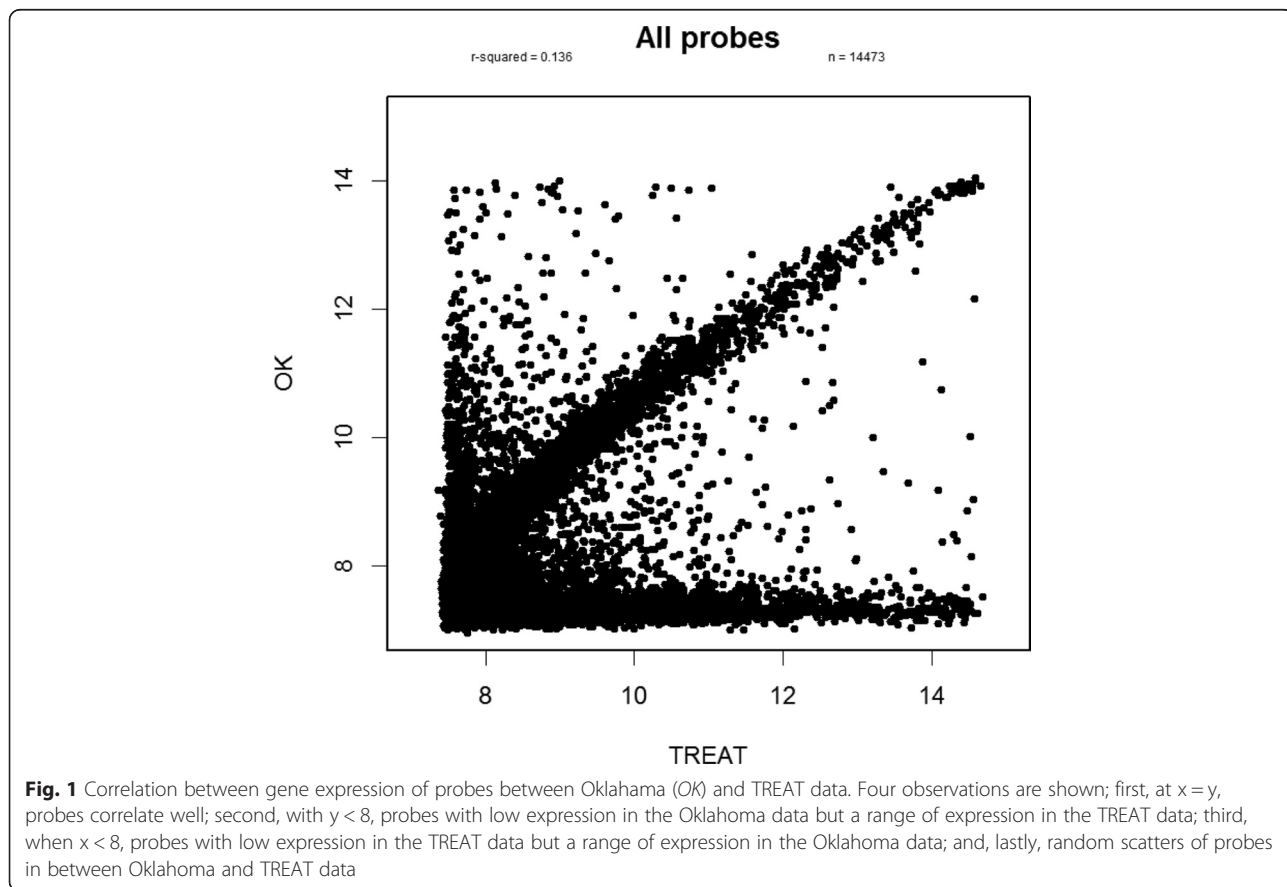
Correlation of TREAT samples with the Oklahoma cohort

From the TREAT study we analyzed 44 baseline samples from 28 RF-negative patients and 16 RF-positive patients, plus 19 control samples from healthy children. In addition, we analyzed independent samples from 10 RF-negative patients (the OK cohort). Oklahoma and TREAT data were first normalized without using the COMBAT algorithm that we previously applied to the TREAT data to filter out batch effects [2], as there is only one condition (baseline) in the OK subjects for comparison and, thus, the algorithm cannot differentiate between biological variation and technical variation (batch effects). Our assumption was that batch effects would weaken the statistical correlation relative to the true biological situation, and any correlation we might find using this approach would thus be significant.

We first created a scatter plot correlating the OK and TREAT expression values for all probes (Fig. 1). As

shown in the figure, there are four distinct groups of probes: 1) probes that appeared to match well (roughly, $x = y$); 2) probes with low expression in the Oklahoma cohort but demonstrating a range of expression in TREAT subjects (seen along the bottom of the graph); 3) probes that showed low expression in the TREAT subjects but a range of expression values in the Oklahoma patients (a small number running up the left-hand side of the plot); and 4) a small scattering of probes in between.

We then asked whether those probes that are low in one set but high in the other may reflect poor quality in the arrays, especially for low-expression probes. To investigate this possibility, we used Illumina array GenomeStudio software, which calculates and reports a detection p value representing the confidence that a given transcript is expressed above the background defined by negative control probes. The cut-off we used for trimming probes from the arrays prior to differential expression analysis was to only select probes with $p \leq 0.01$. The histograms in Fig. 2 show the distribution of the probe qualities in the two datasets. For each probe we counted the number of arrays on which it had a detection p value ≤ 0.01 . Using this approach, we identified multiple poor quality probes on the both array sets.



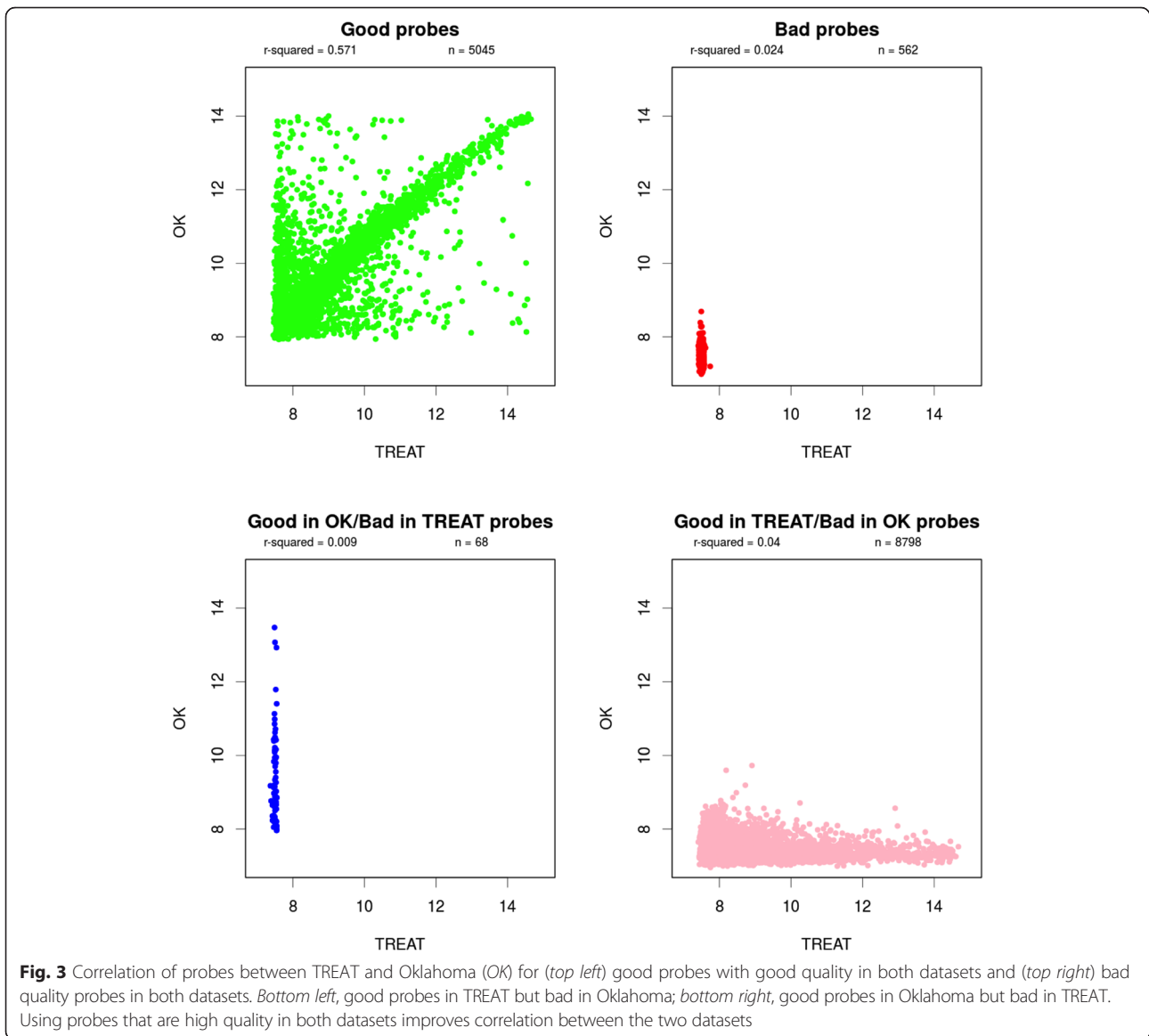
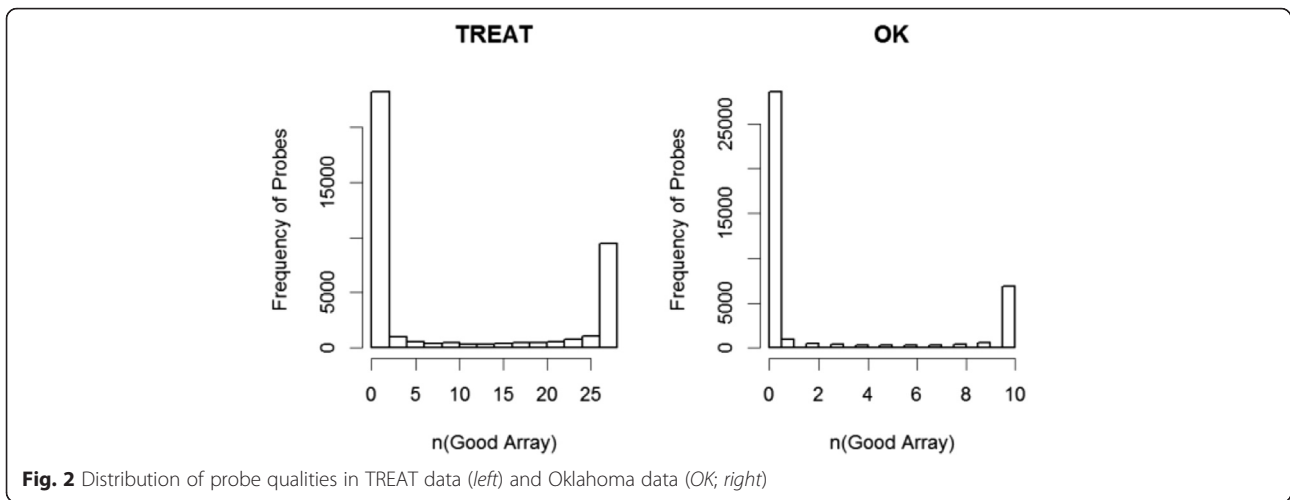


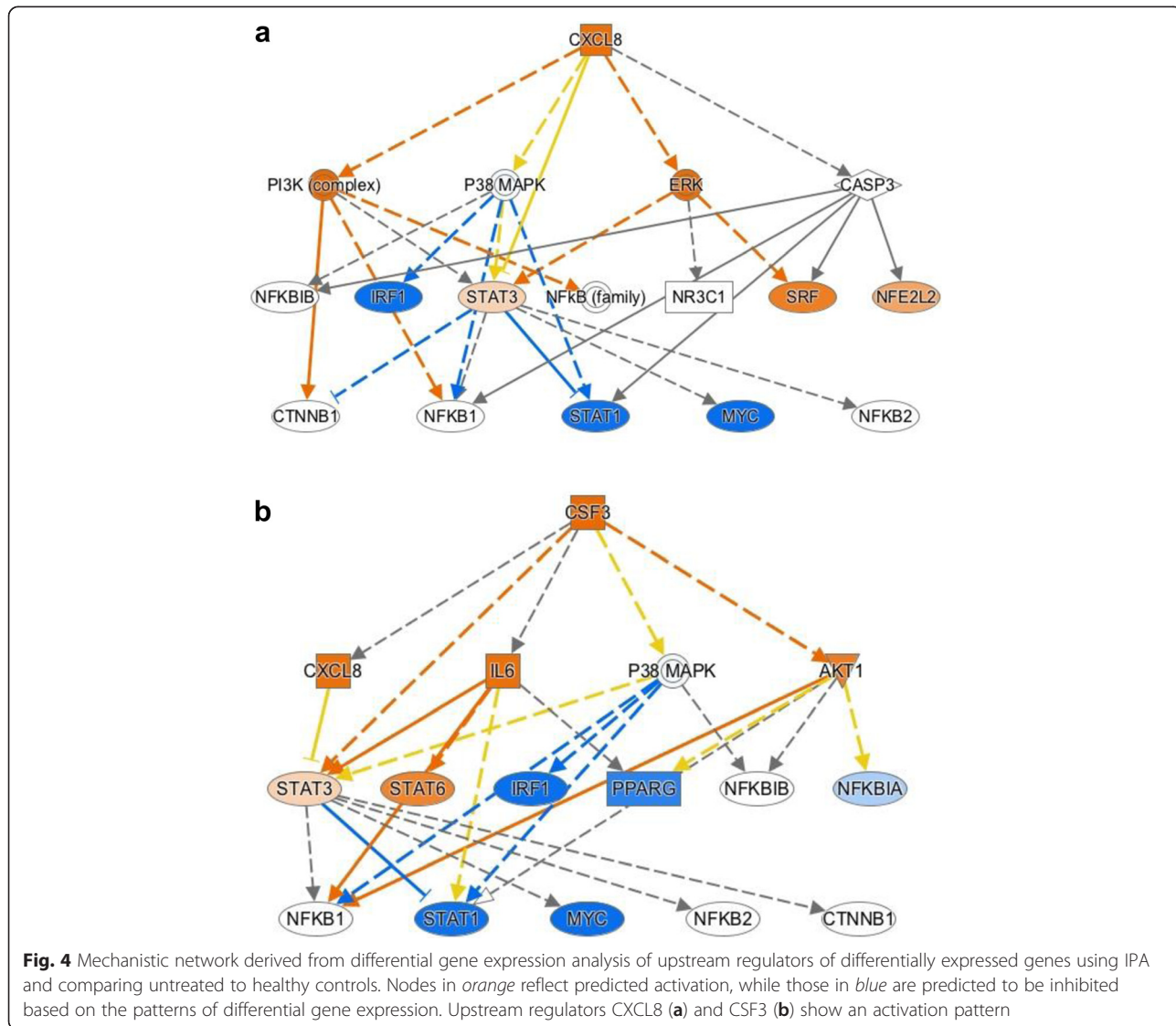
Table 3 Differentially expressed genes in juvenile idiopathic arthritis between RF-positive and RF-negative patients

Probe_ID	Symbol	Mean expression RF ⁻ (Log2)	Mean expression RF ⁺ (Log2)	Fold change (Log2)	FDR
ILMN_1690443	C14orf82	8.775	9.284	-0.509	0.046
ILMN_1717594	DKFZp761E198	8.567	9.061	-0.493	0.028
ILMN_1721113	HLA-C	9.43	10.126	-0.696	0.04
ILMN_1778202	FLJ40722	11.76	12.259	-0.499	0.032
ILMN_1806165	HSPA6	10.943	11.429	-0.486	0.036
ILMN_1912662		11.38	11.896	-0.516	0.038
ILMN_2243516	C11orf63	10.335	10.939	-0.604	0.036
ILMN_2285713	TDP1	10.931	11.421	-0.49	0.015

FDR false discovery rate, RF⁻ rheumatoid factor negative, RF⁺ rheumatoid factor positive

Upon splitting the probes into groups based on their quality, we discerned that there were technical issues creating the two unmatched tails. In the comparison, we required that all 10 of the Oklahoma arrays be good quality (Fig. 3). Using this approach, we found that

good probes (top left) displayed good quality in both datasets and bad quality probes were bad in both. Selecting only probes that are high quality in both datasets significantly improves the r^2 (0.627 for our two datasets).



We also note that in our biomarker paper [2] we were able to corroborate gene expression in the TREAT and Oklahoma cohorts using quantitative polymerase chain reaction (qPCR) [11]. Having assured ourselves of the reproducibility of the array results, we proceeded to data analysis.

Differential gene expression

We first compared RF-positive ($n = 16$) and RF-negative ($n = 28$) samples. We found only 7 genes represented on 8 probes that showed differential expression between these groups; these genes are shown in Table 3. These results are consistent with our previous observation that RF-positive and RF-negative samples group together when analyzed through hierarchical cluster analyses [2]. For this reason, we determined that it would be reasonable to analyze all 44 samples (RF-positive and RF-negative) as a group.

There were 158 genes represented by 176 probes that showed differential expression with at least a 1.4-fold difference, with an FDR of 0.05, when TREAT study subjects were compared with healthy control children. The analysis of these genes using the Ingenuity database showed particular enrichment for genes regulating leukocyte adhesion and extravasation. This pattern is reflected in the network shown in Fig. 4a, demonstrating a pattern of interleukin (IL)-8 CXCL8 activation. IL-8 is produced by multiple leukocyte subsets, including macrophages [17, 18], as well as other cell types such as endothelial cells [19]. IL-8 acts through multiple G protein-coupled receptors and is a potent chemoattractant for neutrophils. It is interesting to note that we have recently demonstrated that G protein-coupled receptor signaling networks showed extensive re-organization when TREAT study subjects responded to therapy [5]. Patterns of gene

expression reflect regulation by CSF3, which, like IL-8, is a potent activator of neutrophils (Fig. 4b). We have previously shown that neutrophils in JIA show aberrant patterns of activation that are linked to the metabolic pathways through which neutrophils produce myeloperoxidase and superoxide ion [20].

The baseline TREAT expression data also reflect activation of adaptive immune responses through CD3–T-cell receptor-mediated signaling, as shown in Fig. 5. T cells have long been accepted as mediators of JIA pathogenesis [21], and thus this finding was expected.

It is interesting to note that, at the same time, the expression data simultaneously reflect a pattern that we have previously identified from RNA sequencing data in human neutrophils [22]: an apparent suppression of interferon response factor (IRF)-mediated pathways leading to a suppression of type1 and type 2 interferons, as shown in Fig. 6. This pattern of suppression reflects a broader suppression of Toll-like receptor (TLR)9 activation, as shown in Fig. 7. The TREAT expression data suggest a previously unrecognized role for transforming growth factor (TGF)B1 in JIA, although its role in adult rheumatoid disease has been long recognized [23], and we have shown that TGFBI is overexpressed in children with active JIA on therapy compared with children who have achieved clinical remission on medication [24]. These findings suggest aberrant patterns of activation and/or gene regulation within the adaptive immune system.

Linking genetic and expression data

The search for overlaps between the DE genes identified in the comparison between healthy controls and baseline TREAT subjects and LD blocks containing JIA-associated SNPs yielded no DE genes within any of the observed LD blocks. We next determined the closest LD blocks to

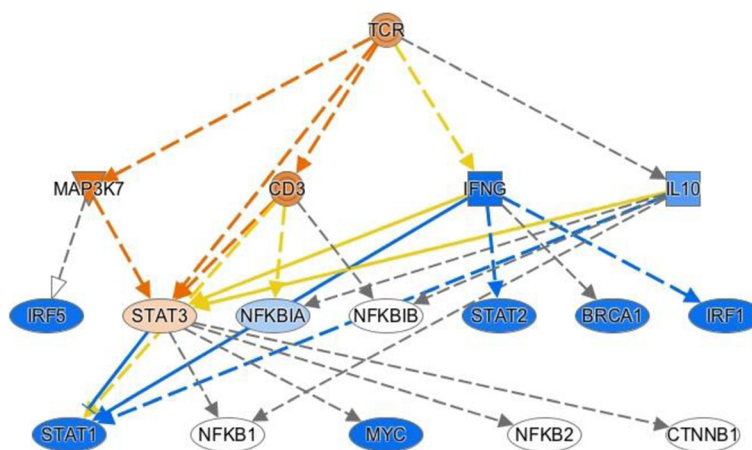


Fig. 5 Mechanistic network derived from differential gene expression analysis of upstream regulators of differentially expressed genes using IPA and comparing untreated JIA to healthy controls. Nodes in orange reflect predicted activation, while those in blue are predicted to be inhibited based on the patterns of differential gene expression. CD3–T-cell receptor (TCR) activation is predicted from the pattern of gene expression

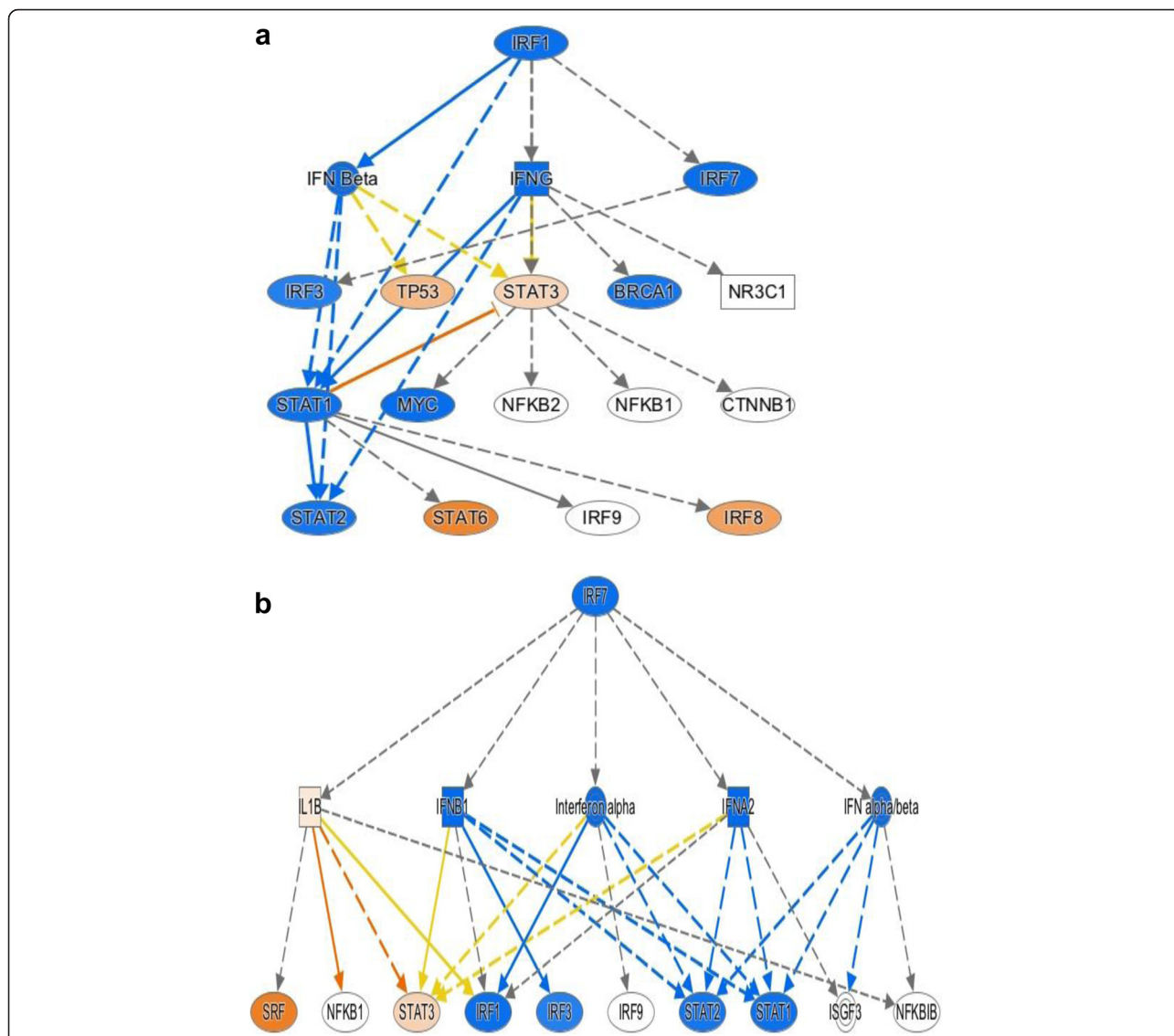


Fig. 6 Mechanistic network derived from differential gene expression analysis using IPA and comparing untreated JIA to healthy controls. Nodes in orange reflect predicted activation, while those in blue are predicted to be inhibited based on the patterns of differential gene expression. Suppression of IRF1 (a) and IRF7 (b) regulated networks is predicted from this analysis

the DE genes, and identified 75 LD blocks containing JIA-associated SNPs situated near 38 DE genes, with distances ranging between 345 and 390 kilobases. Next, we interrogated the upstream regulators of the DE genes as identified on Ingenuity analysis, i.e., IRF1, IRF3, IRF5, STAT1, STAT2, and STAT3. We queried the LD blocks containing the JIA-associated SNPs using the bedtools program. We downloaded ENCODE transcription factor binding site (TFBS) data from UCSC Genome Browser ENCODE data at <http://hgdownload.cse.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeRegTfbsClustered/> and extracted TFBSs of the regulators of interest. Intersection analysis demonstrated that 53 LD blocks containing JIA-associated SNPs overlapped 232 TF binding sites for the

regulators of interest. We applied Fisher’s Exact test for enrichment analysis of TF binding sites overlapping with the LD blocks containing JIA-associated SNPs using all ENCODE TF binding sites as background. This analysis showed no statistically significant evidence for enrichment of IRF1, IRF3, IRF5, STAT1, STAT2, or STAT3 binding within the LD blocks containing JIA-associated SNPs (Fisher’s Exact test *p* value 0.091).

Discussion

In this paper, we analyzed the TREAT whole blood gene microarray data in an attempt to gather insights into disease mechanisms in polyarticular JIA. We found that the whole blood expression profiles reflect complex

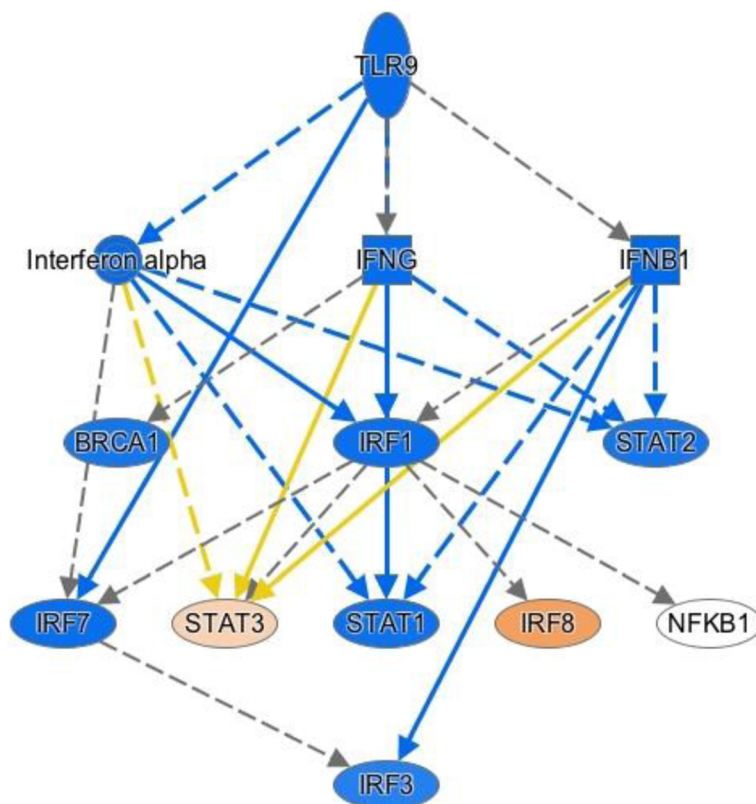


Fig. 7 Mechanistic network derived from differential gene expression analysis using IPA and comparing untreated JIA to healthy controls. Nodes in orange reflect predicted activation, while those in blue are predicted to be inhibited based on the patterns of differential gene expression. Suppression of TLR9-regulated networks is predicted from this analysis

interactions between innate and adaptive immunity, a finding that is consistent with our previous reports that both peripheral blood mononuclear cell (PBMC) [24] and neutrophil [25] gene expression profiles are abnormal in untreated children with JIA. We were not able to answer the question of whether the innate or adaptive immune aberrations are primary. While an increasing body of data shows that neutrophils are important mediators of adaptive immune responses [26–31], it is equally plausible that the transcriptional aberrations we see in neutrophils are the result of an altered cytokine milieu generated from altered T-cell function. In support of the idea that the neutrophil defect is a primary aspect of the disease is our finding that neutrophil gene expression profiles in JIA patients remain distinctly abnormal even after PBMC profiles begin to resemble those of healthy children in response to therapy [32].

These findings are also significant for what they do not tell us. Among the 158 genes that showed differential expression between baseline TREAT samples and those of healthy controls, none was located within LD blocks where there is known genetic risk for JIA [14, 15], nor are the identified LD blocks enriched for binding sites for the TFs that the whole blood expression data suggest may be

important regulators of the differentially expressed genes. This finding corroborates published work demonstrating that most of the genetic risk for JIA lies within the non-coding genome [14]. We have recently demonstrated that most of the regions identified by Hinks et al. [14] are enriched (above genome background) for H3K4me1/H3K27ac-marked enhancers that can be identified in both neutrophils and CD4+ T cells [33]. Thus, if polyarticular JIA, like many complex diseases, can be characterized by the presence of so-called expression quantitative trait loci (eQTL) [34], it seems likely that the loci that most strongly influence expression will be located in non-promoter regulatory regions (e.g., enhancers, insulators, and so forth) and reflect complex layers of transcriptional control rather than perturbed function of the protein products of specific genes.

Our findings here suggest potentially useful targets for therapy in JIA. For example, the broad suppression of type 1 and type 2 interferon responses in JIA may reflect overall suppression of TLR9-mediated processes. TLR9 is an intracellular pattern recognition receptor that detects highly methylated DNA, which is common in bacterial and viral genomes (and relatively rare in mammalian genomes) [35]. In recent years, TLR9 has become

an attractive therapeutic target for immune modulation in immune diseases [36, 37], as well as cancer [38–42]. Given these findings, it is hardly surprising that hydroxychloroquine, which suppresses TLR9 pathways [43, 44], has been shown to be ineffective in JIA. The TREAT whole blood expression data suggest that strategies to augment TLR9 responses might be more promising.

There are obviously limitations to these data and their interpretation. The first is the inherent “noisiness” of whole blood expression profiles. Whole blood expression profiles represent an amalgam of peripheral blood cells, including abundant leukocyte subtypes such as neutrophils and platelets, and less abundant cells such as monocytes, natural killer (NK) cells, and even circulating CD34+ cells [45]. Furthermore, while the complete blood counts of the TREAT baseline subjects did not deviate from the range typically seen in children of the same age, it is possible that the differences in expression profiles reflect expansion of small leukocyte subsets not typically identified on complete blood counts performed in a standard clinical laboratory. The noisiness and relative insensitivity of whole blood expression profiling can be reduced by removing globin genes before the RNA labeling step [3, 46], as we did here, but this step reduces only a small portion of the complexity that limits the utility of whole blood expression data. Leukocyte subset transcriptomes show a considerable degree of specificity, reflecting the specific immunologic functions of each cell type. Thus, while there are considerable commonalities in the transcriptomes and regulatory regions of peripheral blood leukocytes [47–49], it is likely that there are critical elements of leukocyte function/dysfunction in polyarticular JIA (e.g., B cells, monocytes, Th17 cells) that simply cannot be identified on whole blood expression profiling. Thus, while whole blood or blood leukocyte expression profiling has been invaluable in allowing us to develop a mechanistic understanding of significant pathologic disturbances, such as sepsis or blunt trauma [50], it seems likely that complex functional genomics approaches of specific leukocyte subsets will be required to fully elucidate the pathogenesis of more subtle phenotypes where inflammation is chronic and more indolent, as is the case in polyarticular JIA [22].

Conclusions

Whole genome expression profiling of untreated children with polyarticular JIA reveals complex transcriptional differences when compared with healthy controls. Activation of leukocyte chemotaxis/extravasation pathways and neutrophil activation by CSF3 are reflected in whole blood transcription analyses. At the same time, suppression of STAT1–3/IRF pathways, as we have previously reported in JIA neutrophils [22], is revealed in the whole blood expression profile. None of the genes that showed differential expression between children with JIA and healthy

control children is encoded within LD blocks containing known JIA-associated SNPs. These findings suggest that genetic risk loci for JIA either exert their effects before the disease phenotype emerges or involves more subtle and complex layers of transcriptional regulation (e.g., by *trans*-acting enhancers [33]) than can be discerned from whole blood expression profiles.

Abbreviations

DE, differentially expressed; ET, etanercept; FDR, false discovery rate; HC, healthy children; IL, interleukin; IPA, Ingenuity Pathway Analysis; IRF, interferon response factor; JIA, juvenile idiopathic arthritis; LD, linkage disequilibrium; MTX, methotrexate; NIH, National Institute of Health; OK, independent Oklahoma cohort; PBMC, peripheral blood mononuclear cell; RF, rheumatoid factor; SNP, single nucleotide polymorphism; TFBS, transcription factor binding site; TGF, transforming growth factor; TLR, Toll-like receptor; TREAT, Trial of Early Aggressive Therapy in Juvenile Idiopathic Arthritis

Acknowledgments

This study was supported by NIH grants R01 AI-084200, R01 AR-060604 (JNJ) and R01-AR 049762 (CAW).

Authors' contributions

KJ and YC performed the wet lab experiments. This included sample processing, RNA purification, and ascertainment of RNA purity. LW, ADS, and KJ performed data analysis. This included differential gene expression analysis and corroboration of the OK and TREAT cohorts (ADS), querying of LD blocks for DE genes and TF binding (LW), and Ingenuity Pathway analysis (KJ). MBF was responsible for performing microarray procedure and assisted in data analysis and interpretation. This included RNA labeling and hybridization procedures as well as preliminary statistical procedures to monitor quality of hybridization reactions. CAW was the principal investigator on the TREAT study and assisted in data interpretation. JNJ designed and organized this study and assisted in data analysis and interpretation. All authors assisted in drafting and revising the manuscript, and read and approved the manuscript.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Pediatrics, University at Buffalo Jacobs School of Medicine and Biomedical Sciences, Clinical & Translational Research Center, 875 Ellicott St., Buffalo, NY, USA. ²Irving Cancer Institute, Columbia University College of Physicians and Surgeons, 1130 Saint Nicholas Ave., New York, NY 10032, USA. ³Oklahoma Medical Research Foundation, Arthritis & Clinical Immunology Program, 800 NE 13th St., Oklahoma City, OK 73104, USA. ⁴Division of Rheumatology, Seattle Children's Hospital and Research Institute, 4800 Sand Point Way NE, MA.7.110, Seattle, WA 98105, USA. ⁵Genetics, Genomics, and Bioinformatics Program, University at Buffalo Jacobs School of Medicine and Biomedical Sciences, Buffalo, NY, USA.

Received: 21 March 2016 Accepted: 22 June 2016

Published online: 07 July 2016

References

- Wallace CA, Giannini EH, Spalding SJ, et al. Trial of early aggressive therapy in polyarticular juvenile idiopathic arthritis. *Arthritis Rheum.* 2012;64:2012–21.
- Jiang K, Sawle AD, Frank MB, et al. Whole blood gene expression profiling predicts therapeutic response at six months in patients with polyarticular juvenile idiopathic arthritis. *Arthritis Rheumatol.* 2014;66:1363–71.
- Liu J, Walter E, Stenger D, Thach D. Effects of globin mRNA reduction methods on gene expression profiles from whole blood. *J Mol Diagn.* 2006;8:551–8.
- Chai V, Vassilakos A, Lee Y, Wright JA, Young AH. Optimization of the PAXgene blood RNA extraction system for gene expression analysis of clinical samples. *J Clin Lab Anal.* 2005;19:182–8.
- Du N, Jiang K, Sawle AD, et al. Dynamic tracking of functional gene modules in treated juvenile idiopathic arthritis. *Genome Med.* 2015;7:109.

6. Petty RE, Southwood TR, Manners P, et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. *J Rheumatol*. 2004;31:390–2.
7. Cobb JP, Mindrinos MN, Miller-Graziano C, et al. Application of genome-wide expression analysis to human health and disease. *Proc Natl Acad Sci U S A*. 2005;102:4801–6.
8. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. *Bioinformatics*. 2008;24:1547–8.
9. Lin SM, Du P, Huber W, Kibbe WA. Model-based variance-stabilizing transformation for Illumina microarray data. *Nucleic Acids Res*. 2008;36:e11.
10. Smyth GK. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol*. 2004;3:Article3.
11. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B*. 1995;57:289–300.
12. Calvano SE, Xiao W, Richards DR, et al. A network-based analysis of systemic inflammation in humans. *Nature*. 2005;437:1032–7.
13. Ficenc D, Osborne M, Pradines J, et al. Computational knowledge integration in biopharmaceutical research. *Brief Bioinform*. 2003;4:260–78.
14. Hinks A, Cobb J, Marion MC, et al. Dense genotyping of immune-related disease regions identifies 14 new susceptibility loci for juvenile idiopathic arthritis. *Nat Genet*. 2013;45:664–9.
15. Hersh AO, Prahalad S. Immunogenetics of juvenile idiopathic arthritis: a comprehensive review. *J Autoimmun*. 2015;64:113–24.
16. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26:841–2.
17. Bhardwaj RS, Eblenkamp M, Berndt T, Tietze L, Klosterhalfen B. Role of HSP70i in regulation of biomaterial-induced activation of human monocytes-derived macrophages in culture. *J Mater Sci Mater Med*. 2001;12:97–106.
18. Xu H, Lai W, Zhang Y, et al. Tumor-associated macrophage-derived IL-6 and IL-8 enhance invasive activity of LoVo cells induced by PRL-3 in a KCNN4 channel-dependent manner. *BMC Cancer*. 2014;14:330.
19. Teofoli P, Lotti T. Cytokines, fibrinolysis and vasculitis. *Int Angiol*. 1995;14:125–9.
20. Jarvis JN, Petty HR, Tang Y, et al. Evidence for chronic, peripheral activation of neutrophils in polyarticular juvenile rheumatoid arthritis. *Arthritis Res Ther*. 2006;8:R154.
21. Grom AA, Hirsch R. T-cell and T-cell receptor abnormalities in the immunopathogenesis of juvenile rheumatoid arthritis. *Curr Opin Rheumatol*. 2000;12:420–4.
22. Jiang K, Sun X, Chen Y, Shen Y, Jarvis JN. RNA sequencing from human neutrophils reveals distinct transcriptional differences associated with chronic inflammatory states. *BMC Med Genomics*. 2015;8:55.
23. Pohlers D, Beyer A, Koczan D, et al. Constitutive upregulation of the transforming growth factor-beta pathway in rheumatoid arthritis synovial fibroblasts. *Arthritis Res Ther*. 2007;9:R59.
24. Knowlton N, Jiang K, Frank MB, et al. The meaning of clinical remission in polyarticular juvenile idiopathic arthritis: gene expression profiling in peripheral blood mononuclear cells identifies distinct disease states. *Arthritis Rheum*. 2009;60:892–900.
25. Jarvis JN, Jiang K, Frank MB, et al. Gene expression profiling in neutrophils from children with polyarticular juvenile idiopathic arthritis. *Arthritis Rheum*. 2009;60:1488–95.
26. Beauvillain C, Delneste Y, Scotet M, et al. Neutrophils efficiently cross-prime naive T cells in vivo. *Blood*. 2007;110:2965–73.
27. Puga I, Cols M, Barra CM, et al. B cell-helper neutrophils stimulate the diversification and production of immunoglobulin in the marginal zone of the spleen. *Nat Immunol*. 2012;13:170–80.
28. Sun J, Dodd H, Moser EK, Sharma R, Braciale TJ. CD4+ T cell help and innate-derived IL-27 induce Blimp-1-dependent IL-10 production by antiviral CTLs. *Nat Immunol*. 2011;12:327–34.
29. Hock BD, Taylor KG, Cross NB, et al. Effect of activated human polymorphonuclear leucocytes on T lymphocyte proliferation and viability. *Immunology*. 2012;137:249–58.
30. Harwood NE, Barral P, Batista FD. Neutrophils—the unexpected helpers of B-cell activation. *EMBO Rep*. 2012;13:93–4.
31. Odobasic D, Kitching AR, Yang Y, et al. Neutrophil myeloperoxidase regulates T-cell-driven tissue inflammation in mice by inhibiting dendritic cell function. *Blood*. 2013;121:4195–204.
32. Jiang K, Frank M, Chen Y, Osban J, Jarvis JN. Genomic characterization of remission in juvenile idiopathic arthritis. *Arthritis Res Ther*. 2013;15:R100.
33. Jiang K, Zhu L, Buck MJ, et al. Disease-associated single-nucleotide polymorphisms from noncoding regions in juvenile idiopathic arthritis are located within or adjacent to functional genomic elements of human neutrophils and CD4+ T cells. *Arthritis Rheumatol*. 2015;67:1966–77.
34. Gibson G, Powell JE, Marigorta UM. Expression quantitative trait locus analysis for translational medicine. *Genome Med*. 2015;7:60.
35. Yasuda K, Richez C, Uccellini MB, et al. Requirement for DNA CpG content in TLR9-dependent dendritic cell activation induced by DNA-containing immune complexes. *J Immunol*. 2009;183:3109–17.
36. Kline JN. Eat dirt: CpG DNA and immunomodulation of asthma. *Proc Am Thorac Soc*. 2007;4:283–8.
37. Klinman DM. Adjuvant activity of CpG oligodeoxynucleotides. *Int Rev Immunol*. 2006;25:135–54.
38. Rosa R, Melisi D, Damiano V, et al. Toll-like receptor 9 agonist IMO cooperates with cetuximab in K-ras mutant colorectal and pancreatic cancers. *Clin Cancer Res*. 2011;17:6531–41.
39. Berger R, Fiegl H, Goebel G, et al. Toll-like receptor 9 expression in breast and ovarian cancer is associated with poorly differentiated tumors. *Cancer Sci*. 2010;101:1059–66.
40. Vaisanen MR, Vaisanen T, Jukkola-Vuorinen A, et al. Expression of toll-like receptor-9 is increased in poorly differentiated prostate tumors. *Prostate*. 2010;70:817–24.
41. Ronkainen H, Hirvikoski P, Kauppila S, et al. Absent Toll-like receptor-9 expression predicts poor prognosis in renal cell carcinoma. *J Exp Clin Cancer Res*. 2011;30:84.
42. Takala H, Kauppila JH, Soini Y, et al. Toll-like receptor 9 is a novel biomarker for esophageal squamous cell dysplasia and squamous cell carcinoma progression. *J Innate Immun*. 2011;3:631–8.
43. Katz SJ, Russell AS. Re-evaluation of antimalarials in treating rheumatic diseases: re-appreciation and insights into new mechanisms of action. *Curr Opin Rheumatol*. 2011;23:278–81.
44. Brewer EJ, Giannini EH, Kuzmina N, Alekseev L. Penicillamine and hydroxychloroquine in the treatment of severe juvenile rheumatoid arthritis. Results of the USA-USSR double-blind placebo-controlled trial. *N Engl J Med*. 1986;314:1269–76.
45. Stroncek DF, Clay ME, Smith J, et al. Composition of peripheral blood progenitor cell components collected from healthy donors. *Transfusion*. 1997;37:411–7.
46. Field LA, Jordan RM, Hadix JA, et al. Functional identity of genes detectable in expression profiling assays following globin mRNA reduction of peripheral blood samples. *Clin Biochem*. 2007;40:499–502.
47. Weinstein JS, Lezon-Geyda K, Maksimova Y, et al. Global transcriptome analysis and enhancer landscape of human primary T follicular helper and T effector lymphocytes. *Blood*. 2014;124:3719–29.
48. Lin H, Joehanes R, Pilling LC, et al. Whole blood gene expression and interleukin-6 levels. *Genomics*. 2014;104:490–5.
49. Powell JE, Henders AK, McRae AF, et al. Genetic control of gene expression in whole blood and lymphoblastoid cell lines is largely independent. *Genome Res*. 2012;22:456–66.
50. Xiao W, Mindrinos MN, Seok J, et al. A genomic storm in critically injured humans. *J Exp Med*. 2011;208:2581–90.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

