Published in final edited form as:
Arthritis Rheum. 2009 July ; 60(7): 2102–2112. doi:10.1002/art.24601.

Subtype-specific peripheral blood gene expression profiles in recent onset juvenile idiopathic arthritis

MG Barnes[1], **AA Grom**[1], **SD Thompson**[1], **TA Griffin**[1], **P Pavlidis**[2], **L Itert**[1], **N Fall**[1], **DP Sowders**[1], **CH Hinze**[1], **BJ Aronow**[1], **LK Luyrink**[1], **S Srivastava**[1], **N Ilowite**[3], **B Gottlieb**[4], **J Olson**[5], **D Sherry**[6], **DN Glass**[1], and **RA Colbert**[1]

[1]Cincinnati Children's Hospital Medical Center, Cincinnati, OH

[2]University of British Columbia Vancouver, BC

[3]Albert Einstein College of Medicine, Bronx, NY; Peoria

[4]Schneider Children's Hospital, New Hyde Park, NY

[5]Medical College of Wisconsin and Children's Research Institute, Milwaukee, WI

[6]Children's Hospital of Philadelphia, Philadelphia, PA

Abstract

Objective—A multi-center study of recent onset juvenile idiopathic arthritis (JIA) subjects prior to treatment with DMARDS or biologics was undertaken to identify peripheral blood gene expression differences between JIA subclasses and controls.

Methods—PBMC from 59 healthy children and 136 JIA subjects (28 enthesitis-related arthritis [ERA], 42 persistent oligoarthritis, 45 RF- polyarthritis, and 21 systemic) were isolated over Ficoll. Poly-A RNA was labeled using NuGEN Ovation and gene expression profiles were obtained using Affymetrix HG-U133 plus 2.0 Arrays.

Results—9,501 differentially expressed probe sets were identified among JIA subtypes and controls (ANOVA, FDR 5%). Specifically, 193, 1036, 873 and 7595 probe sets were different between controls and ERA, persistent oligoarthritis, RF- polyarthritis and systemic JIA samples respectively. In persistent oligoarthritis, RF- polyarthritis and systemic JIA subtypes, up-regulation of genes associated with IL-10 signaling was prominent. A hemoglobin cluster was identified that was under-expressed in ERA patients but over-expressed in systemic JIA. The influence of JAK/STAT, ERK/MAPK, IL-2 and B cell receptor signaling pathways was evident in persistent oligoarthritis. In systemic JIA, up-regulation of innate immune pathways, including IL-6, TLR/IL1R, and PPAR signaling were noted, along with down-regulation of gene networks related to NK and T cells. Complement and coagulation pathways were up-regulated in systemic JIA with a subset of these genes differentially-expressed in other subtypes as well.

Conclusions—Expression analysis identified differentially expressed genes in PBMCs between subtypes of JIA early in disease and controls, thus providing evidence for immunobiologic differences between these forms of childhood arthritis.

Juvenile idiopathic arthritis (JIA) encompasses chronic childhood arthritis of unknown etiology and is manifest by diverse clinical symptoms and outcomes (1). Patients with persistent oligoarthritis have cumulative involvement of fewer than five joints, whereas extended oligoarthritis indicates involvement of five or more joints some time after six months of disease. Polyarthritis involves five or more joints within the first six months of disease and is subdivided by the presence or absence of rheumatoid factor (RF+ or RF- polyarthritis). Enthesitis-related arthritis (ERA) typically affects older (>6 years) males who frequently have HLA-B27 and may have a family history of spondyloarthropathy. Systemic

JIA involves chronic arthritis and associated systemic features that may include quotidian fevers, erythematous rash, generalized lymphadenopathy and hepatosplenomegaly.

Heterogeneity of JIA can be partially accounted for by interactions of complex genetic and environmental factors. While some genetic associations reported in other autoimmune diseases are also found in JIA, there are additional genetic factors unique to JIA and specific for JIA subtypes (reviewed in (2)). These include well-documented HLA Class II associations (reviewed in (3)) and subtype-specific genetic linkages (4). Understanding interacting genetic traits may one day contribute to determining diagnosis and prognosis of JIA.

Outcomes in JIA are variable and range from complete recovery to persistent active arthritis with subsequent joint destruction and/or ankylosis that produce significant disability. For most patients, long-term outcome is difficult to predict early on, and identification of patients that would benefit from early aggressive treatment is uncertain. Whole-genome gene expression analysis has significant discovery potential regarding JIA classification, prognosis, and pathogenesis. The genome-wide coverage of this technology offers an unbiased view of disease processes and can generate novel hypotheses since it does not involve investigating specific genes of interest based on previous understanding of disease. This comprehensive approach has been successfully applied to several rheumatologic conditions including SLE (5) and some forms of JIA (6–10). In the present study, we report the analysis of peripheral blood mononuclear cell (PBMC) gene expression in a large cohort of recent onset JIA subjects prior to treatment with DMARDs (disease modifying anti-rheumatic drugs) or biologics. We find that for each subtype of JIA PBMC, gene expression patterns can largely distinguish patients from normal controls. To our knowledge this is the first comparison of PBMC gene expression profiles of multiple subtypes of recent onset JIA.

SUBJECTS, MATERIALS AND METODS

Subjects and Controls

Following informed consent, patients with recent onset JIA were enrolled at five clinical sites (see Table 1 for additional information) and followed for up to two years. The clinical sites were Cincinnati Children's Hospital Medical Center (CCHMC) (61 patients), Schneider Children's Hospital (28 patients), Children's Hospital of Philadelphia (26 patients), Children's Hospital of Wisconsin (14 patients), and Toledo Children's Hospital (7 patients). Subjects were classified by ILAR criteria (11) using the cumulative clinical and laboratory information available from all study visits. Patients were generally enrolled early in disease (median 5 mo; 69% < 6 mo; 90% < 12 mo), and those with relatively long disease duration were retained because of slow disease evolution (3 ERA) or delayed initiation of DMARD therapy (1 persistent oligoarthritis; 9 RF- polyarthritis; 1 systemic). Subjects had not received DMARDs or biologics (antimalarials, azathioprine, cyclosporine, tacrolimus, gold salts, leflunomide, methotrexate, penicillamine, sulfasalazine, tacrolimus, adalimumab, etanercept, infliximab, or other biologics) prior to sample acquisition. Most subjects were on NSAIDs and a limited number were on other medications: atenolol (n=1), homatropine eyedrops (n=1), corticosteroid eye drops (n=3), omeprazole (n=2) and oral corticosteroids (n=4; range 0.5–1 mg/kg). Five subjects received intra-articular corticosteroids within30 days prior to sampling (4 days [RF- polyarticular], 24 days [ERA], 27 days [oligoarticular], 27 days [RF- polyarticular], and 29 days [RF- polyarticular]) and all had active joints at time of sample. Of note, of the nine patients who had previously taken steroids, eight clustered with their respective subtypes with only one being an outlier. Subjects did not appear to have other inflammatory disease in addition to JIA. The 59 controls were apparently healthy children from the Cincinnati area. With the wide range of demographics among JIA subtypes it was impossible for controls to perfectly match each subtype. To account for this

difference in characteristics a broad age range of controls was included, most controls were Caucasian since most JIA subjects were also Caucasian, sex-related probe sets were removed from the analysis (see Quality Control and Data Management).

Sample Preparation

Peripheral blood was collected using acid citrate dextrose (ACD). PBMC were isolated over Ficoll and RNA was immediately stabilized in TRIzol Reagent (Invitrogen; Carlsbad, CA). Processing was accomplished as quickly as possible as measured by Time to Freezing (TTF; the length of time between phlebotomy and freezing in TRIzol). Samples were stored at −80°C at the collecting site prior to shipment to CCHMC on dry ice. RNA was purified at CCHMC on RNeasy columns then stored in water at −80°C. RNA samples were randomized into groups of eleven and a universal standard (US) was included in each group to provide a technical replicate to measure batch-to-batch variation. The US was comprised of pooled PBMC RNA from 35 healthy adult volunteers.

Labeling

RNA quality was assessed using an Agilent 2100 Bioanalyzer according to standard protocols in the CCHMC Affymetrix GeneChip® core. 100 ng of RNA was labeled using NuGEN Ovation Version 1. Labeled cDNA was hybridized to Affymetrix HG U133plus 2.0 arrays and scanned with an Agilent G2500A GeneArray Scanner. Data from GeneChips® were assessed for quality using a combination of positive and negative spike-in controls, percent present calls, and average background.

Data Analysis

Microarray data were imported into GeneSpring GX 7.3 (Agilent Technologies; Palo Alto, CA) and pre-processed using Robust Multi-Array Average (RMA) followed by normalization of each probe to the median of all samples. Distance-Weighted-Discrimination (DWD) was used to align centroids of pre-defined groups (12–16) to control for batch-to-batch variation. GeneChip® data are available through NCBI's Gene Expression Omnibus (17), series accession GSE13501 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE13501). A supervised analysis was performed using ANOVA (Benjamini-Hochberg false-discovery rate of 5%) followed by Tukey post-hoc testing to identify genes with differential expression between pre-defined groups. Hierarchical clustering of samples using genes selected by supervised analysis was performed using Pearson correlation. Clustering using Spearman correlation gave similar results and the stronger gene clusters were stable when using different correlations (data not shown). Gene lists were analyzed using Ingenuity Pathways Analysis (Ingenuity Systems; Redwood City, CA) to identify biological pathways with differential expression.

Real Time Polymerase Chain Reaction (PCR)—PBMC RNA was reverse transcribed using a blend of oligo (dT) and random hexamers provided in the iScript™ cDNA Synthesis Kit (Bio-Rad, Hercules, CA). Real time PCR reactions were performed in a 20μl volume using an iCycler instrument (Bio-Rad), gene-specific primers and TaqMan™ probes for haptoglobin, IL-10, MS4A4A and SOCS3 (TaqMan Assays-on-Demand™, Applied Biosystems, Forest City, CA). Raw data was normalized and expressed relative to a housekeeping gene, Tubulin.

RESULTS

Quality Control and Data Management

Quality Control (QC)—In preparation for a study of this magnitude and duration, with multiple investigators and centers, extensive quality control measures were instituted to reduce variation and ensure reliability of results (Figure 1). Prior to sample collection and processing, a standardized protocol was established and adopted in each center and in-person training was provided. Emphasis was placed on minimizing delay in sample processing. A pilot study was performed where several identical aliquots of peripheral blood were kept at room temperature for varying periods of time, then processed and analyzed in parallel. In agreement with other studies (18,19), extended TTF was associated with expression changes in a subset of genes (unpublished observations). Thus, samples with TTF > 240 min (approximately 5% of samples) were excluded from analysis.

Prior to labeling and hybridization, RNA quality was assessed according to standard protocols of the CCHMC Affymetrix GeneChip® Core (Figure 1, RNA QC). Poor quality RNA was infrequent, but resulted in removal of approximately 2% of samples. An additional 8% failed to label or hybridize properly based on initial evaluation of microarray results (Figure 1, Core Microarray QC), but virtually all of these samples were re-run successfully.

Gene expression data were subjected to RMA pre-processing then normalized to the median of all samples. Technical variation between batches of samples was monitored using the US (Materials and Methods). Since the US is pooled RNA and identical for multiple runs, it monitors batch-to-batch variation in labeling and hybridization. Distance Weighted Discrimination (DWD), a method that aligns centroids of pre-defined cohorts, was applied to adjust for this variation (14). This variation may not have been identified without the US, suggesting that this experimental design should be considered for other large microarray studies.

Comparing samples from males and females, 216 probe sets identified genes whose expression differed according to sex. Interestingly, 70 of these probe sets were not encoded on either X or Y chromosomes. These 216 probe sets were removed from further analysis to reduce differences in gene expression caused by changes in sex ratios. Therefore, our study used 54,459 probe sets from the Affymetrix HG U133plus2.0 GeneChip®.

Genome-wide Expression Differences

Comparison across samples—PBMC were collected from JIA patients (ERA, persistent oligoarthritis, RF- polyarthritis and systemic) early in disease and prior to treatment with DMARDs or biologics. Information used to classify subjects included all clinical and laboratory data that became available during the two year study. For instance, a patient could have been classified as `probable oligoarthritis' at baseline, but later re-classified as `RF- polyarthritis' as more joints became involved. Comparing samples from subjects and controls (ANOVA; FDR 5%) identified 9501 differentially expressed probe sets. The Tukey post-hoc test identified 193, 1036, 873, and 7595 probe sets different between controls and ERA, persistent oligoarthritis, RF- polyarthritis and systemic JIA, respectively (Supplemental Table 1). These probe set lists represent 5671, 148, 703, 608 and 4643 (total ANOVA, ERA, persistent oligoarthritis, RF- polyarthritis, and systemic JIA respectively) unique Gene Symbol annotations according to NetAffx (www.affymetrix.com). These gene numbers should be interpreted with caution since many probe sets are annotated with more than one gene, are not annotated with any gene, refer to predicted or hypothetical genes or proteins, or may hybridize to additional genes than are

annotated. The overlap between probe set lists was relatively small, with each subtype having many unique probe sets (154, 649, 479 and 6741 for ERA, persistent oligoarthritis, RF- polyarthritis and systemic arthritis, respectively), indicating subtype-specific gene expression differences. Only seven probe sets (six genes) were common to all four lists. Amyloid beta precursor protein binding protein 2 (APPBP2), two Zinc finger proteins (ZNF230, ZND451), and two orfs (C15orf17, C14orf012) were under-expressed in patients compared to controls. Monocyte to macrophage differentiation associated protein (MMD) was over-expressed in all patient groups.

Results were confirmed by an independent method using samples from 14 subjects with systemic arthritis and 16 controls. Real time PCR was performed for four mRNAs identified by GeneChip® as over-represented in subjects with systemic arthritis. Compared to controls, the relative expression of target mRNAs in systemic arthritis was increased by 3.3 (haptoglobin), 5.7 (IL-10), 2.6 (MS4A4A) and 4.4 (SOCS3). This is consistent with increased expression measured by microarray analysis (6.6, 1.2, 3.7 and 2.7).

Supervised hierarchical clustering was applied to each list of differentially expressed probe sets (Figure 2). This procedure assists with visualization and is expected to produce groupings of largely homogeneous samples according to the sample designations that were used for gene selection. For ERA, persistent oligoarthritis and RF- polyarthritis gene expression patterns were similar among many JIA subjects as shown by co-clustering, however a number of JIA subjects co-clustered with controls (Figure 2A, B, C)., The gene expression patterns for systemic JIA subjects were more homogeneous with only two outliers (Figure 2D).

Biological Meaning of Differentially-expressed Genes (Canonical pathways)— Bioinformatic approaches were used to identify cohesive biological themes in the dataset (6,20,21). Over-representation of 163 “canonical pathways” (defined by Ingenuity) was investigated (Figure 3). These pathways are derived from the literature and Kyoto Encyclopedia of Genes and Genomes (KEGG) (22) and offer starting points to investigate themes in gene expression datasets. The gene lists for ERA, persistent oligoarthritis, and RF-polyarthritis were separated into two groups for this analysis (Figure2A, a and b; Figure 2B, c and d; Figure 2C, e and f). The gene list for systemics is much larger and was separated into 13 groups based on expression patterns exhibited in Figure 2D. To be considered over-represented, a pathway had to be ranked in the top 5 and have $p \leq 0.05$ (Fischer's exact test; without multiple testing correction for discovery purposes).

This analysis identified 46 over-represented pathways (Figure 3) primarily related to immunity and inflammation. Gene lists for these pathways are available from the authors or Ingenuity, Inc. As expected from gene expression differences (Figure 2), the relative contributions of each pathway differed between JIA subtypes. In persistent oligoarthritis, pathways for IL-2, B cell receptor and JAK/STAT signaling pathways were over-represented. Persistent oligoarthritis, RF- polyarthritis and systemic subtypes all exhibited over-representation of the IL-10 signaling pathway and the glucocorticoid receptor signaling pathway. The number of over-represented pathways was highest in systemic JIA with 34 pathways relating to this subtype. Most notably, up-regulation of innate immune pathways including IL-6, TLR/IL1R, and PPAR signaling pathways, as well as the complement system and coagulation cascade were over-represented in systemic JIA. Conversely, NK cell, T cell and antigen-presentation pathways were down-regulated in systemic JIA.

Specific Clusters of Interest

In addition to the Ingenuity pathways analysis, we selected gene groups of interest and further investigated specific genes by literature searches. The lists presented in Table 2 are

those that were previously identified by ANOVA. Full gene lists are available upon request. Several groups of differentially expressed genes are quite intriguing and offer hypotheses for future consideration.

Hemoglobin/erythrocyte Cluster—Expression of several hemoglobin (Hb) genes was found to be differentially regulated in JIA subtypes (Table 2). In samples from ERA subjects, adult Hb α (HBA1/A2, HBA2) and β (HBB) were down-regulated. In systemic JIA, in addition to up-regulation of adult Hb, there was increased expression of several fetal Hbs including γ (HBG1/G2), δ (HBD), ε (HBE1) and μ (HBM) in cluster `o' (Figure 2D) which also included several erythrocyte structural proteins, cell surface molecules and enzymes.

Coagulation Cascade—The coagulation cascade was an over-represented canonical pathway (Table 2). Expression of several genes whose products have anti-coagulant properties were modulated, including tissue factor pathway inhibitor (TFPI) up-regulated in RF- polyarthritis and systemics, thrombomodulin (THBD) increased in systemics and protein C [inactivator of coagulation factors Va and VIIIa] (PROC) down-regulated in RF- polyarthritis. Additionally, genes for a number of pro-coagulant proteins exhibited increased expression in JIA with the largest effect seen in systemic JIA.

Complement Cascade—While most complement protein synthesis occurs in the liver, local production in areas of inflammation also occurs by circulating cells (macrophages, dendritic cells and monocytes (23–29)). Many factors of the complement cascade were found to be differentially-expressed in JIA subjects (Table 2). Persistent oligoarthritis, RF- polyarthritis and systemic JIA show differential expression of several complement inhibitory proteins (CR1, CR2, CD55, CD59). Interestingly, ERA had decreased expression of many immunoglobulins, including the complement fixing IgG1, while there was a slight increase in IgG1 expression in systemic JIA (data not shown). Additionally, several factors from the classical pathway of complement activation (C1q, C2, and C4) displayed increased expression in systemic JIA.

DISCUSSION

To our knowledge this is the first study to comprehensively evaluate PBMC gene expression patterns in several subtypes of JIA early in disease and prior to initiation of DMARDs or biologics. Extensive QC measures were observed during sample collection and processing, and data analysis to ensure validity of results. The results demonstrate that JIA subtypes can be distinguished from healthy controls using PBMC gene expression patterns. The most striking differences are found in systemic JIA, and a number of pathways differentially affected in each JIA subtype provide a framework for future investigations of pathogenesis.

An important consideration in the interpretation of gene expression profiles obtained from complex mixtures of cells is the influence of cellular composition. Genes that are referred to as `up- or down-regulated' may actually be `over- or under-represented' due to differences in abundance of cell populations. This was nicely demonstrated by Bennett et al., who found a granulopoiesis signature in SLE (5), and is likely responsible for some of the striking differences seen in systemic JIA (6).

In the current study, there are several examples of disease-specific pathways that are altered. For example, in persistent oligoarthritis IL-2, B cell receptor, JAK/STAT and ERK/MAPK signaling pathways are over-represented. In RF- polyarthritis pathways representing G-protein coupled receptor and cAMP-mediated signaling, and NRF-2-mediated oxidative stress response signaling were prominent. IL-10 signaling was over-represented in several

subtypes including persistent oligoarthritis, RF- polyarthritis, and systemic JIA, perhaps reflecting activation of anti-inflammatory mechanisms. In systemic JIA, IL-6, TLR/IL1R and PPAR pathways were affected, consistent with results recently reported by Ogilvie et al. (8). Additionally, networks related to NK cells and T cells were down-regulated in systemic JIA, a pattern that is remarkably similar to what has been seen in septic shock (30). These observations support the concept that innate immune activation plays a prominent role in systemic JIA pathogenesis, and provide evidence for immunobiological differences between persistent oligoarthritis and RF- polyarthritis.

While ERA has not been extensively examined previously, several studies have reported gene expression differences in systemic JIA (6,8,10), a limited number in RF- polyarthritis (7,9), and a single study of persistent oligoarthritis(7), although these JIA subtypes were not generally examined in combination. In systemic JIA, Pascual et al. emphasized the importance of IL-1β by noting its increased production by patient-derived PBMCs and dramatic responses to IL-1β inhibition in 7/9 patients (10). Conversely, Ogilvie et al. did not see a prominent IL-1β gene expression signature (8), which may be consistent with recent work of Gattorno et al., who identified two subsets of systemic JIA based on differential responsiveness to IL-1 blockade(31). In the present study, we did not see prominent over-expression of IL-1-responsive genes, but we found evidence for over-representation of TLR/IL1R pathway genes, raising the possibility of excessive TLR stimulation. Taken together, these observations are consistent with systemic JIA being a heterogeneous disease and further exploration of disease subgroups is warranted.

More recently, Allantaz et al. defined a set of 12 differentially expressed genes that distinguish systemic JIA from other systemic illnesses such as viral or bacterial infection and SLE (10). In our data set, seven of these 12 genes were differentially expressed in systemic JIA, with direction and magnitude of changes similar to that reported (10). Five genes exhibit increased expression (CLIC-2, TLOC1, WNK1, C18orf10, and UBB3'UTR), while two are reduced (WHDC1 and C18orf17). Differential expression of these genes was specific to systemic JIA, and not seen in other subtypes. There are several possible explanations for the lack of differential expression of the additional 5 genes noted by Allantaz et al., including exposure of their subjects to medications such as corticosteroids and methotrexate, disease duration, and disease heterogeneity. Note that comparisons between our results and those of Ogilvie et al. are limited by study design (8). They examined active and inactive systemic JIA patients with well-established disease who were treated with combinations of corticosteroids, methotrexate and TNF-α inhibitors. They found evidence of IL-6 and IL-10 over-expression, which is consistent with our Ingenuity analysis.

Jarvis et al. have reported peripheral blood leukocyte gene expression changes in subjects classified by ACR criteria as having polyarticular juvenile rheumatoid arthritis (JRA) compared to healthy controls (9). Comparisons with this study are limited because different cell populations were studied and the microarrays were different. Notably, buffy coat-derived leukocytes contain neutrophils, which are generally absent from Ficoll-purified PBMC.

Barnes et al. (7) previously reported gene expression differences using PBMC from pauciarticular and polyarticular course JRA. The main finding was up-regulation of pro-angiogenic CXCL chemokines in polyarticular JRA compared to pauciarticular or healthy controls. The current study does not identify these genes as being over-expressed. This is not unexpected since the most prominent findings of the earlier study were based on comparisons of PBMC to synovial fluid mononuclear cells, which was not part of the current study. In addition, the previous samples were obtained from patients with long-

standing disease (average 9.3 years), were not collected with rigorous attention to QC, and had been stored in many instances for several years. Consequently, extensive comparisons with the current study are not possible.

A traditional way to visualize microarray data is supervised hierarchical clustering. This method returns clusters of the patient subgroups used to derive the gene lists. In the current study, visual inspection of clustering trees suggests that gene expression patterns identify distinct subgroups within each of the JIA subtypes. Thus, current JIA subtypes may include more heterogeneity than previously appreciated. Studies are ongoing to assess this heterogeneity and its impact on classification and prognosis.

The decreased adult hemoglobin gene expression in patients with ERA was unexpected. Patients with ankylosing spondylitis and other types of spondyloarthritis rarely are anemic (32). In fact, in our study, patients with ERA had the highest hemoglobin levels and systemic patients had the lowest (data not shown). We hypothesize that the decreased expression of hemoglobin in ERA might be a response to TGF-β, which has been reported to be over-expressed in ankylosing spondylitis (33) and can act through an AP-1 binding site near the hemoglobin gene.

One of the most highly correlated gene groups distinguishing systemic JIA from controls is designated `o' in Figure 2D. Examination of this cluster revealed a large group of genes encoding red blood cell structural proteins and enzymes, similar to the erythropoietic signature described previously (6) where it was suggested that the signature was due to the presence of increased immature erythrocyte precursors in peripheral circulation. This cluster also contains various hemoglobins including those normally expressed during embryonic and fetal stages of human development (hemoglobins γ, δ, ε, μ, and θ). Under normal conditions, genes encoding fetal hemoglobins undergo silencing soon after birth (34) although these proteins may sometimes be present in adults (35). We hypothesize that the increased expression of adult and fetal hemoglobins in patients with systemic JIA may be explained by the presence of immature erythrocyte precursors.

The “Coagulation System Pathway” was identified as over-represented in systemic JIA although many of the genes are modulated in other subtypes. These findings are consistent with the clinical observations of mildly increased levels of D-dimers in a majority of JIA patients with the largest increase seen in systemic JIA, which suggests a coagulopathy that may correlate with disease activity (36,37). In innate inflammatory responses, activation of coagulation and fibrin deposition is an important mechanism that helps contain inflammatory activity to the site of injury or infection; however, when not localized, coagulation can have a deleterious effect in patients suffering from systemic inflammatory disorders such as septic shock or macrophage activation syndrome, a well-known complication of systemic JIA. We hypothesize that the over-representation of several anti-coagulant proteins such as TFPI in RF-polyarthritis and systemic JIA indicates an attempt to down-regulate systemic inflammation-induced coagulation and fibrin deposition.

Many factors from the complement cascade were over-expressed in JIA patients (Table 2). While a large proportion of complement synthesis occurs in the liver, production also occurs locally in areas of inflammation by circulating cells (macrophages, dendritic cells and monocytes (23–29)). Three subtypes of JIA had increased expression of some complement inhibitory proteins (CR1, CR2, CD55, CD59). CR1 is present on blood cells including erythrocytes, neutrophils, monocytes, eosinophils and T and B cells and is involved in immune clearance and inhibition of the complement cascade by acting as a cofactor for cleavage of C3b. CR2 is found mainly on B cells and is involved in B cell activation and immune clearance. CD55 and CD59 are markedly up-regulated on activated macrophages

and less so on CD4 and CD8 positive lymphocytes. This up-regulation could indicate the release of complement factors during poorly controlled inflammation in the joint.

High-throughput technologies such as Affymetrix GeneChips® produce vast amounts of data that can be analyzed and interpreted in many different ways. Making large datasets publicly available allows investigators to apply alternative methods of analysis and increase data utilization. High-throughput technologies are by their nature hypothesis generating, and the hypotheses presented in this study must be validated in other cohorts and the biological relevance of gene expression differences identified here remain to be determined at protein and organism levels.

In conclusion, this study has identified gene expression differences in PBMC between JIA subtypes and controls related to immunity and inflammation. Several differentially expressed genes replicate findings in other studies providing further confidence in our results which have the potential to greatly expand our understanding of JIA. Expression patterns indicate that defined JIA subtypes have strong internal similarities although clustering also indicates some degree of heterogeneity within subtypes. These findings will likely provide important mechanistic information with respect to JIA subtypes and lead to an improved molecular definition of JIA.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We wish to acknowledge specific individuals for their efforts: Wendy Bommer (clinical research coordinator; Cincinnati Children's Hospital Medical Center). Jane Boyd (study coordinator; Peoria). Joseph Couri, MD (Methodist Medical Center of Illinois). Mark Getz, MD (Order of Saint Francis Medical Center). Jesse Gillis, Ph.D. (data pre-processing; University of British Columbia). Anne Johnson (clinical research coordinator; Cincinnati Children's Hospital Medical Center). Marsha Malloy, RN, MBA (study coordinator; Medical Center of Wisconsin). Beth Martin (study coordinator; Toledo). Marilyn Orlando (study coordinator; Schneider). David Wilson (study coordinator; Children's Hospital of Pittsburgh). Sara Jane Wilson (study coordinator; Children's Hospital of Philadelphia). Jeremy Zimmermann (data collection, Medical Center of Wisconsin). We also wish to acknowledge our sources of funding: Cincinnati Children's Hospital Research Foundation; The Arthritis Foundation, Ohio Valley Chapter; NIH grants P01AR048929, P30AR47353, P60AR47784.

REFERENCES

1. Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J, et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol. 2004; 31(2):390–2. [PubMed: 14760812]
2. Prahalad S, Glass DN. A comprehensive review of the genetics of juvenile idiopathic arthritis. Pediatr Rheumatol Online J. 2008; 6(1):11. [PubMed: 18644131]
3. Donn RP, Ollier WE. Juvenile chronic arthritis--a time for change? Eur J Immunogenet. 1996; 23(3):245–60. [PubMed: 8803538]
4. Thompson SD, Moroldo MB, Guyer L, Ryan M, Tombragel EM, Shear ES, et al. A genome-wide scan for juvenile rheumatoid arthritis in affected sibpair families provides evidence of linkage. Arthritis Rheum. 2004; 50(9):2920–30. [PubMed: 15457461]
5. Bennett L, Palucka AK, Arce E, Cantrell V, Borvak J, Banchereau J, et al. Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. J Exp Med. 2003; 197(6):711–23. [PubMed: 12642603]
6. Fall N, Barnes M, Thornton S, Luyrink L, Olson J, Ilowite NT, et al. Gene expression profiling of peripheral blood from patients with untreated new-onset systemic juvenile idiopathic arthritis reveals molecular heterogeneity that may predict macrophage activation syndrome. Arthritis Rheum. 2007; 56(11):3793–804. [PubMed: 17968951]

7. Barnes MG, Aronow BJ, Luyrink LK, Moroldo MB, Pavlidis P, Passo MH, et al. Gene expression in juvenile arthritis and spondyloarthropathy: pro-angiogenic ELR+ chemokine genes relate to course of arthritis. Rheumatology (Oxford). 2004; 43(8):973–9. [PubMed: 15150433]
8. Ogilvie EM, Khan A, Hubank M, Kellam P, Woo P. Specific gene expression profiles in systemic juvenile idiopathic arthritis. Arthritis Rheum. 2007; 56(6):1954–65. [PubMed: 17530721]
9. Jarvis JN, Dozmorov I, Jiang K, Frank MB, Szodoray P, Alex P, et al. Novel approaches to gene expression analysis of active polyarticular juvenile rheumatoid arthritis. Arthritis Res Ther. 2004; 6(1):R15–R32. [PubMed: 14979934]
10. Allantaz F, Chaussabel D, Stichweh D, Bennett L, Allman W, Mejias A, et al. Blood leukocyte microarrays to diagnose systemic onset juvenile idiopathic arthritis and follow the response to IL-1 blockade. J Exp Med. 2007; 204(9):2131–44. [PubMed: 17724127]
11. Petty RE, Southwood TR, Baum J, Bhettay E, Glass DN, Manners P, et al. Revision of the proposed classification criteria for juvenile idiopathic arthritis: Durban, 1997. J Rheumatol. 1998; 25(10):1991–4. [PubMed: 9779856]
12. Marron JS, Todd MJ, Ahn J. Distance Weighted Discrimination. Journal of the American Statistical Association. 2007; 102:1267–1271.
13. Lu Y, Lemon W, Liu PY, Yi Y, Morrison C, Yang P, et al. A gene expression signature predicts survival of patients with stage I non-small cell lung cancer. PLoS Med. 2006; 3(12):e467. [PubMed: 17194181]
14. Benito M, Parker J, Du Q, Wu J, Xiang D, Perou CM, et al. Adjustment of systematic microarray data biases. Bioinformatics. 2004; 20(1):105–14. [PubMed: 14693816]
15. Hu Z, Fan C, Oh DS, Marron JS, He X, Qaqish BF, et al. The molecular portraits of breast tumors are conserved across microarray platforms. BMC Genomics. 2006; 7:96. [PubMed: 16643655]
16. Herschkowitz JI, Simin K, Weigman VJ, Mikaelian I, Usary J, Hu Z, et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. Genome Biol. 2007; 8(5):R76. [PubMed: 17493263]
17. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 2002; 30(1):207–10. [PubMed: 11752295]
18. Feezor RJ, Baker HV, Mindrinos M, Hayden D, Tannahill CL, Brownstein BH, et al. Whole blood and leukocyte RNA isolation for gene expression analyses. Physiol Genomics. 2004; 19(3):247–54. [PubMed: 15548831]
19. Baechler EC, Batliwalla FM, Karypis G, Gaffney PM, Moser K, Ortmann WA, et al. Expression levels for many genes in human peripheral blood cells are highly sensitive to ex vivo incubation. Genes Immun. 2004; 5(5):347–53. [PubMed: 15175644]
20. Burczynski ME, Peterson RL, Twine NC, Zuberek KA, Brodeur BJ, Casciotti L, et al. Molecular classification of Crohn's disease and ulcerative colitis patients using transcriptional profiles in peripheral blood mononuclear cells. J Mol Diagn. 2006; 8(1):51–61. [PubMed: 16436634]
21. Greene JG, Greenamyre JT, Dingledine R. Sequential and concerted gene expression changes in a chronic in vitro model of parkinsonism. Neuroscience. 2007
22. Kanehisa M, Goto S, Kawashima S, Nakaya A. The KEGG databases at GenomeNet. Nucleic Acids Res. 2002; 30(1):42–6. [PubMed: 11752249]
23. Alper CA, Johnson AM, Birtch AG, Moore FD. Human C'3: evidence for the liver as the primary site of synthesis. Science. 1969; 163(864):286–8. [PubMed: 4883617]
24. Cao W, Bobryshev YV, Lord RS, Oakley RE, Lee SH, Lu J. Dendritic cells in the arterial wall express C1q: potential significance in atherogenesis. Cardiovasc Res. 2003; 60(1):175–86. [PubMed: 14522421]
25. Castellano G, Woltman AM, Nauta AJ, Roos A, Trouw LA, Seelen MA, et al. Maturation of dendritic cells abrogates C1q production in vivo and in vitro. Blood. 2004; 103(10):3813–20. [PubMed: 14726389]
26. Lu J, Wu X, Teh BK. The regulatory roles of C1q. Immunobiology. 2007; 212(4–5):245–52. [PubMed: 17544810]
27. Muller W, Hanauske-Abel H, Loos M. Biosynthesis of the first component of complement by human and guinea pig peritoneal macrophages: evidence for an independent production of the C1 subunits. J Immunol. 1978; 121(4):1578–84. [PubMed: 701808]

28. Vincent F, de la Salle H, Bohbot A, Bergerat JP, Hauptmann G, Oberling F. Synthesis and regulation of complement components by human monocytes/macrophages and by acute monocytic leukemia. DNA Cell Biol. 1993; 12(5):415–23. [PubMed: 8517928]

29. Vincent F, Eischen A, de la Salle H, Bergerat JP, Faradji A, Hauptman G, et al. Synthesis of complement components C2 and C4 by human monocyte-derived macrophages during in vitro differentiation in serum-free culture conditions. Pathobiology. 1991; 59(3):136–9. [PubMed: 1883508]

30. Shanley TP, Cvijanovich N, Lin R, Allen GL, Thomas NJ, Doctor A, et al. Genome-level longitudinal expression of signaling pathways and gene networks in pediatric septic shock. Mol Med. 2007; 13(9–10):495–508. [PubMed: 17932561]

31. Gattorno M, Piccini A, Lasiglie D, Tassi S, Brisca G, Carta S, et al. The pattern of response to anti-interleukin-1 treatment distinguishes two subsets of patients with systemic-onset juvenile idiopathic arthritis. Arthritis Rheum. 2008; 58(5):1505–15. [PubMed: 18438814]

32. Golding DN. Hematology and biochemistry of ankylosing spondylitis. Br Med J. 1973; 2(5867): 663. [PubMed: 4714857]

33. Braun J, Bollow M, Neure L, Seipelt E, Seyrekbasan F, Herbst H, et al. Use of immunohistologic and in situ hybridization techniques in the examination of sacroiliac joint biopsy specimens from patients with ankylosing spondylitis. Arthritis Rheum. 1995; 38(4):499–505. [PubMed: 7718003]

34. Little JA, Dempsey NJ, Tuchman M, Ginder GD. Metabolic persistence of fetal hemoglobin. Blood. 1995; 85(7):1712–8. [PubMed: 7535584]

35. Umemura T, Al-Khatti A, Papayannopoulou T, Stamatoyannopoulos G. Fetal hemoglobin synthesis in vivo: direct evidence for control at the level of erythroid progenitors. Proc Natl Acad Sci U S A. 1988; 85(23):9278–82. [PubMed: 2461566]

36. Bloom BJ, Tucker LB, Miller LC, Schaller JG. Fibrin D-dimer as a marker of disease activity in systemic onset juvenile rheumatoid arthritis. J Rheumatol. 1998; 25(8):1620–5. [PubMed: 9712110]

37. Gallistl S, Mangge H, Neuwirth G, Muntean W. Activation of the haemostatic system in children with juvenile rheumatoid arthritis correlates with disease activity. Thromb Res. 1998; 92(6):267–72. [PubMed: 9870893]

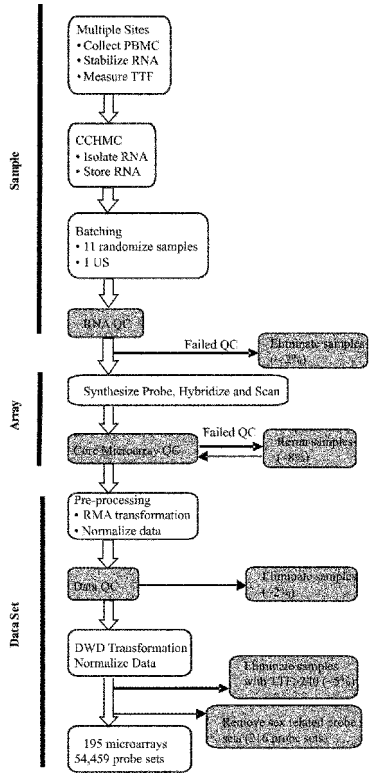

Figure 1. Quality control

In a study of this size and duration, quality control of sample collection and processing is essential. The QC plan is split into three general categories indicated on the left side: sample collection, sample processing, and data processing. The percentage of samples removed at each step is indicated. Samples were collected at multiple sites and all RNA was isolated at CCHMC. RNAs were sent to the Affymetrix GeneChip® core where probes were synthesized and hybridized to Affymetrix U133 plus 2.0 GeneChips®. Data was pre-processed and quality controlled. Data from 54,459 probe sets and 195 arrays which passed all QC measures were used in all further analyses.

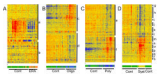

Figure 2. Hierarchical clustering of differentially expressed genes by JIA subtype

Differentially expressed genes were identified using ANOVA (false discovery rate 5%) followed by Tukey post-hoc testing. The normalized expression level for each gene (rows) in each sample (columns) is indicated by color. Red, yellow, and blue rectangles reflect expression levels that are greater than, equal to, or less than the mean expression level in all samples, respectively. The colored lines below the clusterings indicates diagnoses (ERA: orange; persistent oligoarthritis: light blue; RF- polyarthritis: dark blue; systemic: yellow; controls: green). Bars and letters to the right of each heat map indicate how the gene lists were divided for Ingenuity analysis (Figure 3).

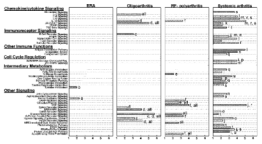

Figure 3. Global analysis of differentially expressed genes

Gene lists based on strong patterns in the heat maps (Figure 2) were evaluated via Ingenuity's “Canonical Pathways” function. To be included in the list of differentially expressed pathways, the pathway had to be over-represented in the list of interest ($p \leq 0.05$) as well as being among the top 5 pathways for that list. The negative log of the p-value is indicated on each horizontal axis and the name of the pathway is on the vertical axis. Pathways are separated into broad categories to simplify interpretation. Letters and numbers indicate which portion of the heat map from Figure 2 contributes to the respective pathway. This figure should be used to identify clusters of under- or over-expressed genes. If a pathway was found in more than one cluster, only the minimal p value (maximal negative decadic logarithm) is represented.

Table 1

Selected Patient Characteristics

JIA Subtype	Control	ERA	Persistent Oligoarthritis	RF-Polyarthritis	Systemic
N*	59	28	42	45	21
% Female†	58%	11%	71%	82%	38%
Age At Onset (years)‡	NA§	12.65 (6.4–16.9)	4.35 (1.1–13.8)	7.6 (1.2–16.3)	4.5 (0.8–15.7)
Age At Sample (years)‡	8.7 (1.8 – 23.8)	13.25 (6.9–17.2)	4.85 (1.3–14.4)	8.8 (1.5–16.8)	4.5 (1–15.9)
Duration Of Disease At Sample (years)‡	NA	0.4 (0–2.7)	0.35 (0–1.2)	0.5 (0.1–3.3)	0.2 (0–2.5)
Joint Count At Sample¶	NA	0.5 (0–17)	1 (0–4)	9 (0–45)	8 (0–64)
Primary Race#	53:4:0:2	25:3:0:0	40:1:1:0	42:2:1:0	18:2:1:0
Joint Count = 0**	NA	14	4	1	1

*Number of subjects in each category

†percent of group that is female

‡median (range) given in years

§Not applicable

¶Median number of joints active at time of sample (range)

#Primary race of patients in the study presented as Caucasian:African American: Asian: Unknown.

**number of patients for whom active joint count was zero at time of sample

Table 2

Selected gene groups of interest

Probe ID†	Gene Symbol‡	Fold Change* p-value§	ERA	P oligo¶	RF-Poly #	Systemic
Hemoglobin						
206697_s_at	HP	7.45E-15	0.87	1.41	1.06	**6.61**
208470_s_at	HP/HPR	2.38E-18	0.88	1.38	1.12	**6.18**
211745_x_at	HBA1	4.73E-09	0.53	0.68	0.74	**4.30**
204018_x_at	HBA1/A2	3.62E-09	**0.52**	0.68	0.72	**4.52**
214414_x_at	HBA2	4.36E-07	**0.56**	0.69	0.74	**3.00**
217232_x_at	HBB	4.12E-07	**0.58**	0.74	0.77	**2.91**
216063_at	HBBP1	1.34E-15	0.94	0.93	0.97	**1.84**
206834_at	HBD	3.10E-19	0.59	0.70	0.77	**14.18**
205919_at	HBE1	2.99E-12	0.98	0.99	1.02	**1.59**
204848_x_at	HBG1/G2	1.72E-09	0.55	0.99	1.02	**8.63**
240336_at	HBM	1.73E-25	0.77	1.01	1.02	**8.38**
220807_at	HBQ1	7.07E-27	0.91	1.05	0.99	**4.65**
Coagulation						
204714_s_at	F5	2.22E-08	1.14	1.00	1.13	**1.92**
203305_at	F13A1	2.02E-02	1.19	1.06	1.26	**1.68**
214916_x_at	F13B	8.08E-03	**0.69**	1.19	0.83	0.94
210664_s_at	TFPI	3.91E-08	1.22	1.13	**1.34**	**2.40**
206259_at	PROC	1.97E-02	0.94	0.93	**0.89**	0.96
207808_s_at	PROS1	5.52E-03	1.34	1.04	1.50	**2.42**
203887_s_at	THBD	1.81E-03	0.99	1.25	1.01	**1.70**
201108_s_at	THBS1	2.06E-04	1.33	1.17	**1.50**	**2.29**
206493_at	ITGA2B	8.69E-07	1.19	1.04	1.41	**2.95**
204626_s_at	ITGB3	3.85E-05	1.07	1.09	1.23	**2.02**
207389_at	GP1BA	4.14E-04	1.18	1.04	1.21	**1.57**
220336_s_at	GP6	7.51E-03	1.06	1.05	1.13	**1.48**

		Fold Change*				
Probe ID†	**Gene Symbol‡**	**p-value§**	**ERA**	**P oligo¶**	**RF-Poly #**	**Systemic**
206883_x_at	GP9	2.94E-03	1.04	1.04	1.01	**1.29**
202628_s_at	SERPINE1	1.74E-02	1.01	1.03	1.02	**1.20**
1554997_a_at	PTGS2	2.05E-02	1.29	**1.74**	1.36	**1.68**
206157_at	PTX3	1.82E-02	1.07	1.15	0.84	**1.43**
Complement						
218232_at	C1QA	3.85E-08	0.91	1.07	0.89	**1.78**
202953_at	C1QB	6.24E-07	0.97	1.04	0.98	**1.75**
225353_s_at	C1QC	6.87E-09	0.97	0.99	0.96	**1.50**
203052_at	C2	2.42E-02	1.06	1.18	1.06	**1.44**
209906_at	C3AR1	4.36E-02	0.98	1.04	0.92	**1.46**
214428_x_at	C4A	9.92E-16	0.94	1.02	0.99	**2.10**
205654_at	C4BPA	1.80E-02	1.01	0.97	1.00	**1.16**
220088_at	C5AR1	3.39E-02	1.11	**1.22**	1.13	**1.38**
201926_s_at	CD55	7.72E-11	1.12	**1.19**	**1.14**	**1.55**
200983_x_at	CD59	3.49E-07	0.84	1.02	0.83	**1.68**
211920_at	CFB	1.55E-03	0.95	1.07	0.99	**1.17**
217484_at	CR1	7.03E-03	1.11	**1.17**	1.10	**1.28**
205544_s_at	CR2	2.60E-02	1.03	**1.25**	**1.31**	1.12
205785_at	ITGAM	8.43E-03	1.00	1.03	0.96	**1.24**
200986_at	SERPING1	3.41E-02	1.00	1.19	1.09	**2.07**

*Fold change as ratio of geometric means (geometric mean patients/geometric mean of controls), of selected genes identified as significantly differential expressed by ANOVA. Bold indicates specific differences identified by Tukey post-hoc analysis. Note that probe sets with _x_at or _s_at do not follow the usual rules for probe set design and are considered lower quality by Affymetrix than those that end _at and may measure multiple transcripts.

†Affymetrix U133 plus 2.0 GeneChip® probe set IDs

‡Full gene names are in supplementary table 1

§false discovery rate adjusted p-value from ANOVA analysis

¶persistent oligoarthritis

#RF- polyarthritis