

RESEARCH ARTICLE

Open Access

# Network analysis identifies protein clusters of functional importance in juvenile idiopathic arthritis

Adam Stevens<sup>1</sup>, Stefan Meyer<sup>1,2</sup>, Daniel Hanson<sup>1</sup>, Peter Clayton<sup>1\*</sup> and Rachele Donn<sup>3\*</sup>

## Abstract

**Introduction:** Our objective was to utilise network analysis to identify protein clusters of greatest potential functional relevance in the pathogenesis of oligoarticular and rheumatoid factor negative (RF-ve) polyarticular juvenile idiopathic arthritis (JIA).

**Methods:** JIA genetic association data were used to build an interactome network model in BioGRID 3.2.99. The top 10% of this protein:protein JIA Interactome was used to generate a minimal essential network (MEN). Reactome FI Cytoscape 2.83 Plugin and the Disease Association Protein-Protein Link Evaluator (Dapple) algorithm were used to assess the functionality of the biological pathways within the MEN and to statistically rank the proteins. JIA gene expression data were integrated with the MEN and clusters of functionally important proteins derived using MCODE.

**Results:** A JIA interactome of 2,479 proteins was built from 348 JIA associated genes. The MEN, representing the most functionally related components of the network, comprised of seven clusters, with distinct functional characteristics. Four gene expression datasets from peripheral blood mononuclear cells (PBMC), neutrophils and synovial fluid monocytes, were mapped onto the MEN and a list of genes enriched for functional significance identified. This analysis revealed the genes of greatest potential functional importance to be *PTPN2* and *STAT1* for oligoarticular JIA and *KSR1* for RF-ve polyarticular JIA. Clusters of 23 and 14 related proteins were derived for oligoarticular and RF-ve polyarticular JIA respectively.

**Conclusions:** This first report of the application of network biology to JIA, integrating genetic association findings and gene expression data, has prioritised protein clusters for functional validation and identified new pathways for targeted pharmacological intervention.

## Introduction

Juvenile idiopathic arthritis (JIA) is a common chronic disease of childhood. It is not a single disease, but a term that encompasses all forms of arthritis, of unknown origin, that begin before the age of 16 years and persist for longer than 6 weeks. It is the commonest form of chronic arthritis in children and the cause of substantial morbidity. JIA is a complex disease with both environmental factors, currently unknown, and genetic contributions to its

aetiopathogenesis. Of the clinical subtypes defined by the International League Against Rheumatism (ILAR) classification of JIA [1] oligoarticular (both persistent and extended forms) and rheumatoid factor (RF)-negative (RF-ve) polyarticular arthritis are the predominant forms, accounting for approximately half of all presenting cases in the UK [2]. Factors that influence the susceptibility to, and progression of, these major phenotypic forms are currently unknown. Similarly, treatment options remain limited with steroid joint injections and methotrexate being most frequently employed. However, the relapse rate is often high and over 40% of individuals continue to have active disease as adults [3,4].

Replicated genetic loci from JIA candidate gene, genome-wide association studies (GWAS) and fine

\* Correspondence: Peter.clayton@manchester.ac.uk; Rachele.donn@manchester.ac.uk

<sup>1</sup>Manchester Academic Health Sciences Centre, Royal Manchester Children's Hospital, Manchester M13 9WL, UK

<sup>3</sup>Centre for Musculoskeletal Research, University of Manchester, Manchester M13 9PT, UK

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

mapping genotyping are emerging [5-7]. In addition, a limited number of gene expression studies have attempted to establish changes in gene expression patterns with oligoarticular or with RF-ve polyarticular JIA subtypes [8-10]. Single nucleotide polymorphism (SNP) association studies, be it by a candidate gene or a GWAS approach, can identify variations that are associated with a particular complex trait. However, multiple constraints exist in extrapolating these observations forwards to any biological significance. SNP analysis, in the main, is done only at the single locus level. The associated variant may not be the functional polymorphism. Furthermore, SNPs associated with complex traits such as JIA, typically have small effect sizes and account for only a very small fraction of the genetic risk [11]. Thus, for most complex diseases, including JIA, an understanding of the underlying biology, leading to better diagnosis and treatment, is unlikely to arise by detailing the functional consequence of individual SNPs. Therefore, a major challenge for JIA, as for other complex diseases, is the integration of high-throughput omic datasets and subsequent identification and prioritisation of disease-associated loci for additional investigation.

JIA, most likely, arises as a result of abnormalities in genes, but more specifically, via the manifestation of perturbations in multiple protein networks that integrate cellular processes, and those that also link cells within tissues, and tissues within organ systems. An innovative approach to identifying key contributing genetic loci, and potential mechanisms of disease in patients with JIA, is the use of network biology. Network analysis (see Table 1) can allow the summation of various interactions and interdependencies between SNP associations and gene expression data [12]. Defining such a network structure is relevant for biological function. Topologically derived networks have interacting proteins that tend to be co-evolving [13], co-functional [14], and co-expressed [15,16].

It is recognised that for the majority of SNPs with a potential functional role in disease the genome-wide significance threshold for disease association ( $P < 1 \times 10^{-8}$ ) may not be reached [17]. Therefore, despite not individually reaching genome-wide statistical significance, SNPs that cluster in networks can inform the underlying biology of a complex genetic disease such as JIA. Studies now describe how network biology place signature changes within the human interactome, uncovering some of the complexities of human diseases [12,18,19]. Furthermore, network analysis can inform drug discovery and drug targeting for complex genetic diseases such as JIA [20,21].

In this present study we have performed network analysis focussed on oligoarticular and RF-ve polyarticular arthritis, integrating genetic association data with gene expression data. This has been used to derive a JIA interactome, and clusters of proteins of functional relevance.

**Table 1 Glossary**

Term	Definition <sup>a</sup>
Bottleneck	A network position that limits the performance of the system. Described by the mathematical idea of <i>betweenness centrality</i> - a description of the number of paths travelling through a node in a network. An example would be a rate-limiting enzymatic reaction.
Cluster	A measure of the tendency of network nodes to form groups. It manifests from a high density of edges between nodes. In biological networks a cluster could represent a protein complex.
Connectivity/degree	The number of connections (see edge) made by any node within a network.
Edge	An interaction between two nodes. Commonly represents a protein:protein interaction but can also be representative of other biological phenomena such as co-expression.
Interactome	A network representing a whole set of direct and/or indirect interactions related to a specific biological phenomenon.
Minimal Essential Network (MEN)	Top 10% of network protein nodes, as scored by connectivity and bottleneck network properties. Corresponds to highly functionally related positions within the network.
Network analysis	Analysis that relates the structural and mathematical properties of a network to its function.
Node	A vertex within a network. In biological networks a node will usually be a gene, protein or metabolite.
Seed genes	A starting set of genes or metabolites of interest that are used to generate an interactome model by inferring joining interactions algorithmically.

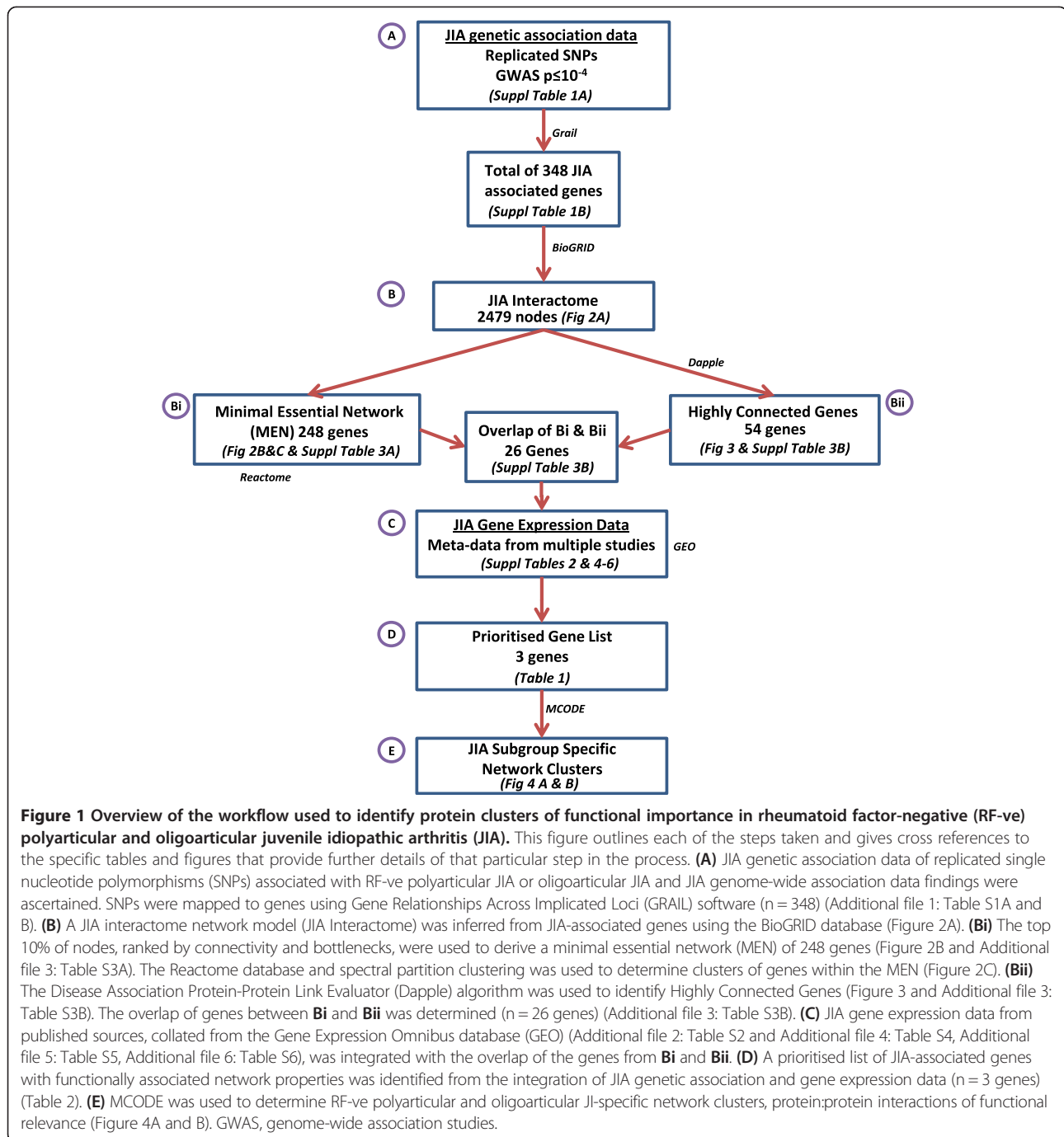
<sup>a</sup>Further technical references are given in Stevens *et al.*, [33].

## Methods

An overview of the approach taken is shown in Figure 1, and the sections below are labelled in correspondence with this figure. These sections detail the key processes and stages in the work flow for the formulation of the JIA subgroup specific network clusters of functional importance.

### Genetic association data

Initially, PubMed [22] and Web of Science [23] searches were conducted and only validated SNP associations from studies of oligoarticular and RF-ve polyarticular JIA were selected. These included multiple well-replicated loci initially associated with adult rheumatoid arthritis found to be significantly associated with JIA [24-30]. In addition, the top findings ( $P \leq 1 \times 10^{-4}$ ) from the largest GWAS study in JIA have been used [6], together with loci replicated by dense genotyping of immune-related disease regions in JIA [7]. SNPs were assigned to genes using an algorithm within Gene Relationships Across Implicated Loci (GRAIL) [31,32] (Additional file 1: Table S1A). This identified a total of 348 JIA-associated genes (Figure 1A).



**Figure 1 Overview of the workflow used to identify protein clusters of functional importance in rheumatoid factor-negative (RF-ve) polyarticular and oligoarticular juvenile idiopathic arthritis (JIA).** This figure outlines each of the steps taken and gives cross references to the specific tables and figures that provide further details of that particular step in the process. **(A)** JIA genetic association data of replicated single nucleotide polymorphisms (SNPs) associated with RF-ve polyarticular JIA or oligoarticular JIA and JIA genome-wide association data findings were ascertained. SNPs were mapped to genes using Gene Relationships Across Implicated Loci (GRAIL) software (n = 348) (Additional file 1: Table S1A and B). **(B)** A JIA interactome network model (JIA Interactome) was inferred from JIA-associated genes using the BioGRID database (Figure 2A). **(Bi)** The top 10% of nodes, ranked by connectivity and bottlenecks, were used to derive a minimal essential network (MEN) of 248 genes (Figure 2B and Additional file 3: Table S3A). The Reactome database and spectral partition clustering was used to determine clusters of genes within the MEN (Figure 2C). **(Bii)** The Disease Association Protein-Protein Link Evaluator (Dapple) algorithm was used to identify Highly Connected Genes (Figure 3 and Additional file 3: Table S3B). The overlap of genes between **Bi** and **Bii** was determined (n = 26 genes) (Additional file 3: Table S3B). **(C)** JIA gene expression data from published sources, collated from the Gene Expression Omnibus database (GEO) (Additional file 2: Table S2 and Additional file 4: Table S4, Additional file 5: Table S5, Additional file 6: Table S6), was integrated with the overlap of the genes from **Bi** and **Bii**. **(D)** A prioritised list of JIA-associated genes with functionally associated network properties was identified from the integration of JIA genetic association and gene expression data (n = 3 genes) (Table 2). **(E)** MCODE was used to determine RF-ve polyarticular and oligoarticular JI-specific network clusters, protein:protein interactions of functional relevance (Figure 4A and B). GWAS, genome-wide association studies.

These 348 JIA-associated genes were the seed genes, that is, known genes that are used to generate an interactome model from all known protein-protein interactions (Additional file 1: Table S1B).

#### Building the JIA interactome

A JIA interactome was derived from the genetic association data using two independent methods (Figure 1Bi and Bii). Interactome modelling uses seed genes, in this

case JIA-associated genes, to infer a network of protein:protein interactions, which are derived from known interactions between the seed genes and their near neighbours [33].

Bi: the first approach used the BioGRID Cytoscape Plugin [34] to create a filter for the BioGRID human interactome model (3.2.99), based on the JIA-associated genes. The resultant JIA interactome was visualised in Cytoscape (version 2.8.3). The Cytohubba Cytoscape Plugin was used

to provide topological analysis of this network and connectivity and bottlenecks were calculated. The JIA interactome was then ranked for both connectivity and bottleneck properties and these were used to evaluate the relative importance of each node [35,36], and to generate a minimal essential network (MEN) [33,37,38]. The top 10% of network nodes was chosen to define the MEN (Figure 1Bi), as this represents the most functional elements of the network [33,37]. The Reactome FI Cytoscape 2.8.3 Plugin [39,40] was then used to determine clusters - distinct groups of protein:protein interactions [41] and to analyse the functional enrichment of biological pathways within the MEN.

Bii: the second approach utilised the Disease Association Protein-Protein Link Evaluator (Dapple) algorithm [42,43] and InWeb database [44]. This approach revealed highly connected genes within the JIA interactome by testing the significance of biological networks using a permutation method (10,000 permutations). The overlap of genes from Bi and Bii was determined using a Venn diagram (Partek Genomics Suite).

#### **Gene expression data was integrated with the JIA interactome**

Gene expression analysis was conducted on a library of gene expression datasets from children with JIA collated from the NCBI Gene Expression Omnibus (GEO) database (Additional file 2: Table S2). Selection was made based on the phenotypic criteria of oligoarticular or RF-ve polyarticular JIA and on control data being accessible.

Gene expression studies were downloaded from GEO [45] and annotation was assessed using QlucoreOmics Explorer 2.2 (Lund, Sweden). Sample comparisons were grouped for disease versus controls, or disease versus disease, using the data in the original peer-reviewed study. Available covariates (confounding factors) as provided in the published data were used. Four separate gene expression datasets were utilised [GEO:GDS711 [9], GSE:11083 [10], GSE:20307 [8], GSE:17755 [46]. The full details of these datasets are given in Additional file 2: Table S2.

Probe-to-gene assignment was made using the appropriate Affymetrix annotation file (Netaffx.com). Dimensional scaling using principal components analysis (PCA) and Iso-map multidimensional scaling [47,48] were used to demonstrate data homogeneity (Qlucore Omics Explorer 2.2) and to identify outliers using cross-validation. Analysis of variance (ANOVA) was used to determine differential gene expression between groups ( $P \leq 0.05$ ) and functional associations were cross-referenced using Ingenuity Pathway Analysis software (IPA). Overlap of gene expression data sets was performed separately for both up- and down-regulated genes using Venn diagrams (Partek).

#### **Deriving a list of prioritised genes of functional relevance**

Genes identified within the MEN by network topology (Figure 1Bi) were overlapped with highly connected genes defined using Dapple ( $P \leq 0.05$ ) (Figure 1Bii). This step identified network positions with putative functional importance using two semi-independent analytical approaches. This overlap was then combined with gene expression data (Figure 1C) to derive a prioritised list of genes of functional relevance.

#### **Identification of JIA subgroup-specific network clusters**

RF-ve polyarticular JIA and oligoarticular JIA clusters were identified from the JIA interactome using the MCODE algorithm plugin for Cytoscape [49]. These clusters identify the most highly related RF-ve polyarticular and oligoarticular JIA subgroup-specific genetic and transcriptomic data within the JIA interactome.

#### **Statistics**

Assessment of pathway associations was performed by the hypergeometric test using the Benjamini-Hochberg false discovery rate (FDR) correction (FDR correction-modified  $P$ -value  $\leq 0.05$ ) [50].

## **Results**

#### **Deriving a minimal essential network of biologically relevant JIA genes**

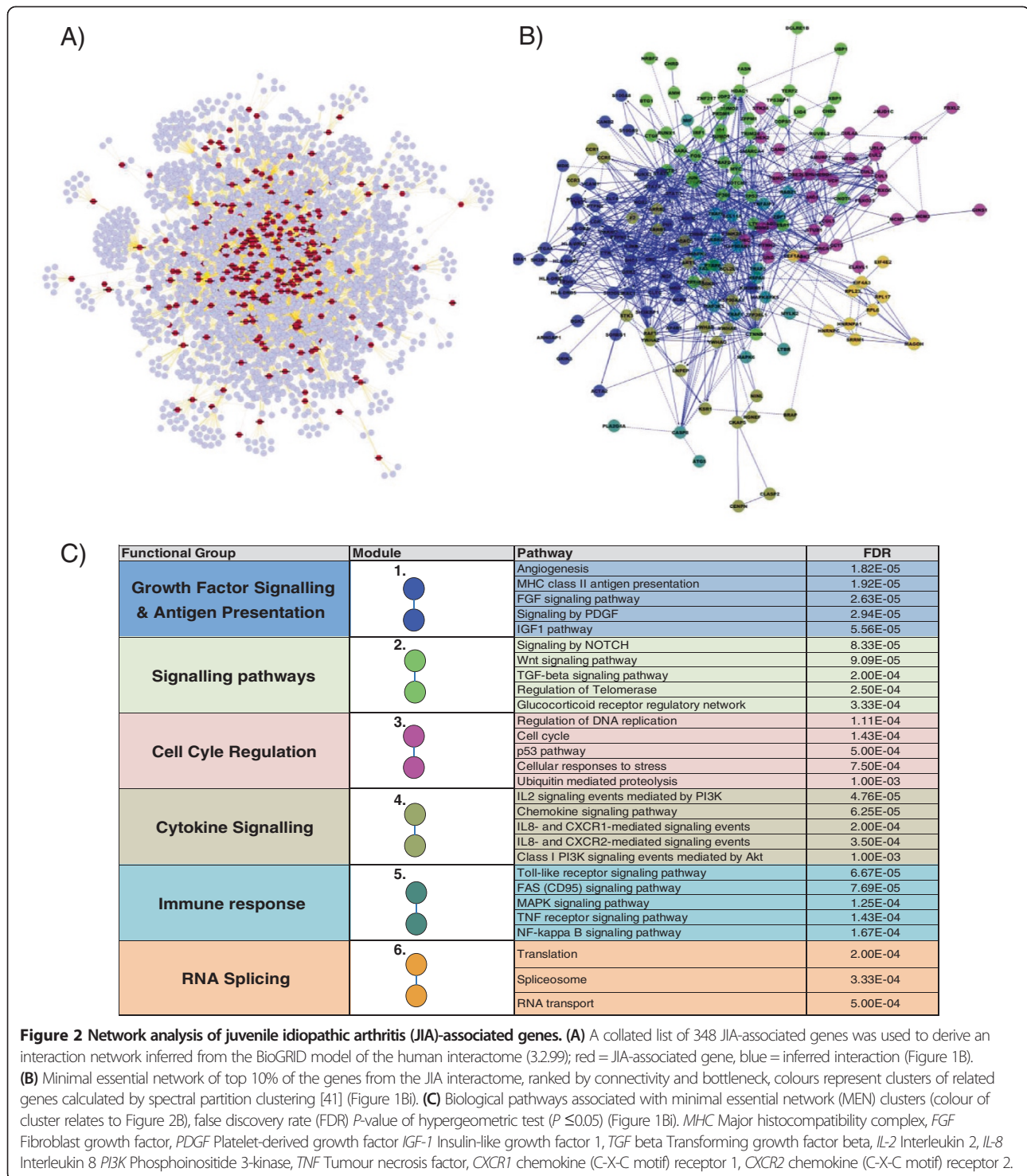
A collated list of 348 JIA-associated seed genes was used to derive a JIA interactome inferred from the BioGRID human interactome database (3.2.99). This consisted of the 348 seed genes (marked red in Figure 2A) along with their immediate inferred interaction partners (protein:protein interactions - purple in Figure 2A). This gave a network of 2,479 proteins and 4,147 edges (Figure 2A). The top 10% of nodes, ranked by connectivity and bottleneck properties were selected to generate a minimal essential network (MEN) (Figure 2B, Additional file 3: Table S3A). The MEN represents the most functionally related components of a network [33,37].

The JIA-associated clusters (same colour code in Figure 2B and C; listed in Additional file 3: Table S3A) within the MEN were mapped onto biological pathways using the Reactome Cytoscape plugin to ascertain function. Six clusters were identified within the MEN, which had clear associated biological pathways (FDR,  $p \leq 0.05$ ). These include pathways related to growth factor signalling and antigen presentation, regulation of the cell cycle and cytokine signalling (Figure 2C).

#### **Permutation analysis of network topology**

Network analysis of the JIA-associated genes was also performed using the Dapple algorithm [42]. This was done to calculate changes in connectivity in the inferred network greater than those expected by chance. First an inferred

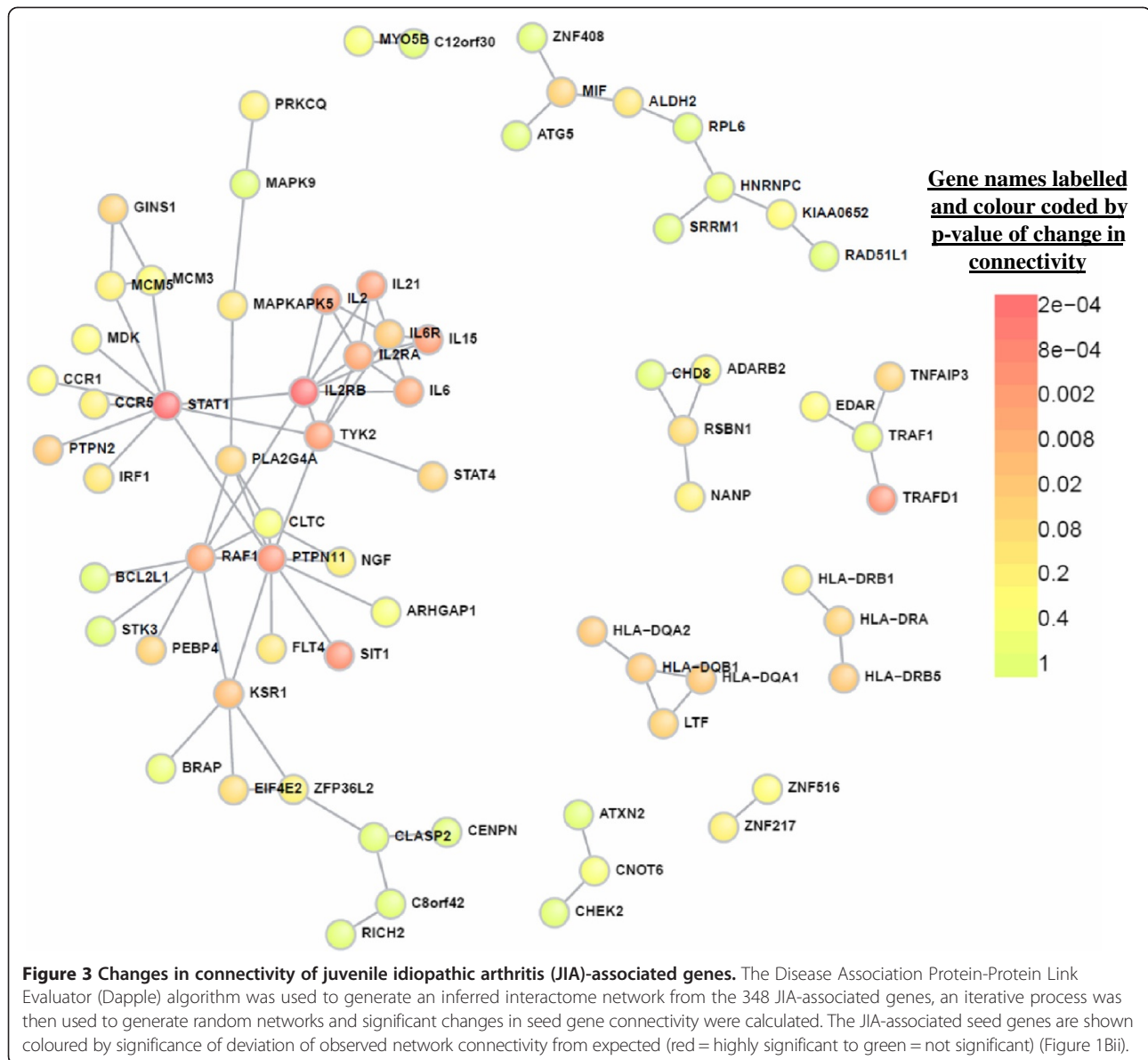




interaction network was derived using the JIA genetic associations as seed genes (Figure 1B). Permutation analysis of this inferred network demonstrated an increase of seed gene connectivity ( $P < 0.002$ ). The seed genes were ranked, by  $P$ -value of increased connectivity (colour scale of ranking, Figure 3; all seed genes listed in Additional file 3: Table S3B).

#### Prioritisation of JIA-associated genes by integration of the JIA interactome with gene expression data

The overlap between the genes present within the MEN and the highly connected loci, identified by Dapple (Figure 1Bi and Bii) was established. This identified 26 genes of the 248 genes that occur in the MEN, that are ranked by Dapple with a  $P$ -value  $\leq 0.05$  (coloured orange



in Additional file 3: Table S3B). These 26 genes represent the most highly connected, and therefore, potentially most functionally important loci, derived from the genetic association findings. Then, in order to prioritise these JIA-associated genes for future functional investigation, gene expression data were aligned (Figure 1C, Figure 2A - red nodes). Four different gene expression series were identified that provided data from RF-ve polyarticular, oligoarticular JIA and controls in previously published studies (GEO). These datasets allowed comparison between gene expression in peripheral blood mononuclear cells (PBMCs), neutrophils or synovial fluid monocytes (SFM) (Additional file 2: Table S2).

#### ***RF-ve polyarticular JIA versus controls***

RF-ve polyarticular JIA gene expression data were compared with controls. Four comparisons were made using PBMC data and one using neutrophils (Additional file 2: Table S2). Comparisons were made separately between up-regulated and down-regulated genes. To increase confidence in the statistical interpretation of the data, and to reduce confounding effects, only genes with significant differential expression in at least three out of the five of these comparisons (Additional file 4: Table S4) were mapped onto JIA-associated genes ranked by network properties. In total this comprised 325 genes; 138 up-regulated and 187 down-regulated (Additional file 4: Table S4).

### Oligoarticular JIA versus controls

Oligoarticular JIA gene expression was compared with controls using PBMC expression data (Additional file 2: Table S2). Comparisons were made separately between up-regulated and down-regulated genes. Genes with significant differential expression in both comparisons (Additional file 5: Table S5) were then mapped onto the list of JIA-associated genes ranked by network properties. There were a total of 150 genes with altered gene expression, 46 being up-regulated and 104 down-regulated (Additional file 5: Table S5).

### Oligoarticular JIA versus RF-ve polyarticular JIA

Oligoarticular was compared to RF-ve polyarticular JIA gene expression data. Two PBMC and one synovial fluid mononuclear cell expression data set were utilised (Additional file 2: Table S2). Comparisons were made separately between up-regulated and down-regulated genes. Genes with significant differential expression in all three comparisons (Additional file 6: Table S6) were then mapped onto the list of JIA associated genes ranked by network properties. A total of 108 genes were changed, 88 were up-regulated and 20 were down-regulated (Additional file 6: Table S6).

### JIA subgroup-specific loci of greatest functional relevance

The JIA subgroup specific loci of greatest functional relevance were identified as those loci that were present within the overlap of the MEN and highly connected genes (Figure 1Bi and Bii) with concomitant variation in gene expression (Figure 1C). These loci are listed in Table 2.

The network analysis identified *KSR1* as the locus of greatest potential functional relevance in RF-ve polyarticular JIA, and *PTPN2* for oligoarticular JIA (Table 2). *Signal transducer and activator of transcription 1 (STAT1)* is differentially involved in oligoarticular versus the polyarticular subgroup. It has increased connectivity, indicating a primary functional role, in oligoarticular JIA compared with RF-ve polyarticular disease (Table 2).

### Prioritised genes form JIA subgroup-specific functional clusters

In order to determine functional clusters, that is, distinct groups of protein:protein interactions in relation to the

key prioritised genes, *KSR1* for RF-ve polyarticular JIA and Protein tyrosine phosphatase, non-receptor type 2 (*PTPN2*) for oligoarticular JIA the MCODE algorithm was applied to the JIA interactome (Figure 1E). The MCODE algorithm identifies highly connected regions of biological networks representing clusters of related function that include protein:protein interaction complexes [49]. This approach identified two high scoring clusters each containing <25 proteins. High-scoring clusters have a high density of local connections and thus they represent a network region with related function [33,51,52]. The cluster score generated by an algorithm such as the MCODE algorithm used in this case, ranks network clusters as a way of prioritising function.

A cluster of 23 proteins contained *PTPN2* and *STAT1* (Figure 4A). Both *PTPN2* and *STAT1* are JIA-associated genes (Additional file 1: Table S1) with concomitant significant changes in gene expression in oligoarticular JIA patients. *PTPN2* was significantly down-regulated in oligoarticular JIA compared with controls; *STAT1* significantly increased in expression in oligoarticular compared to RF-ve polyarticular JIA cases (Additional file 6: Table S6). The main biological pathways associated with this cluster are cytokine signalling ( $P < 1.8 \times 10^{-4}$ ) and DNA repair ( $P < 1.4 \times 10^{-3}$ ).

A second functional cluster of 14 proteins was identified that contained *KSR1* (Figure 4B). *KSR1* is a JIA-associated gene (Additional file 1: Table S1) with significant increased expression in RF-ve polyarticular JIA patients (Additional file 4: Table S4). The main biological pathways found to be associated with the *KSR1*-related cluster are the RAF/MAP kinase cascade ( $P < 8.9 \times 10^{-5}$ ), cell cycle checkpoints ( $P < 5.6 \times 10^{-4}$ ), DNA replication ( $P < 1.3 \times 10^{-4}$ ) and cytokine signalling ( $P < 1.2 \times 10^{-2}$ ).

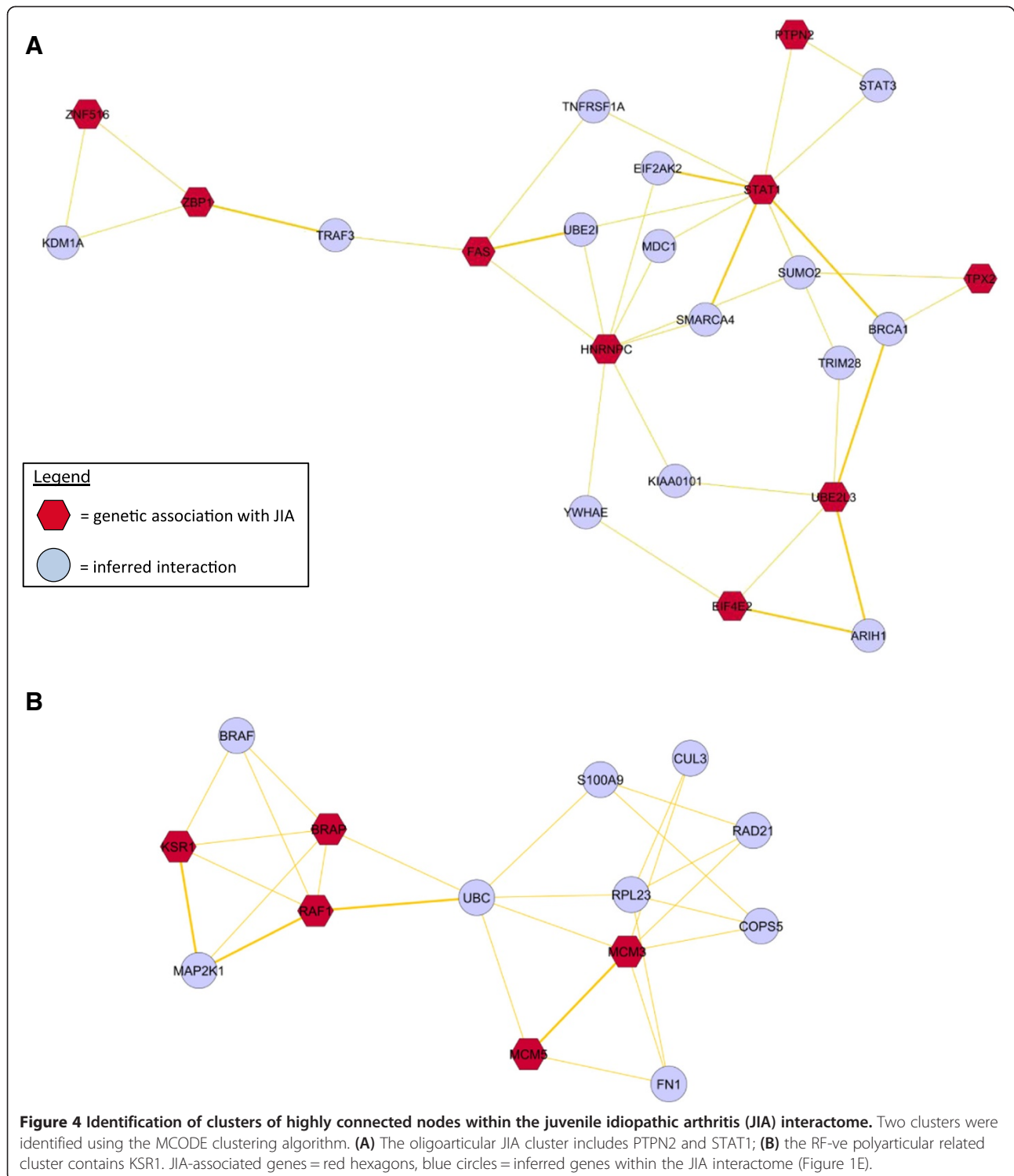
### Discussion

Recognising the necessity to move away from an approach focussed around the determination of SNP function *per se* we have utilised network biology approaches to reveal highly connected regions within a derived JIA interactome in order to identify the interplay of molecular elements that leads to the phenotypic expression of RF-ve polyarticular JIA or to oligoarticular JIA.

**Table 2 Genes of greatest potential functional relevance determined by network analysis**

	JIA-associated genes with correlated expression data that lie within the minimal essential network (MEN)	Dapple P-value
RF-ve polyarticular JIA compared with controls	<i>KSR1</i>	0.0088
Oligoarticular JIA compared with controls	<i>PTPN2</i>	0.0112
Oligoarticular compared with RF-ve polyarticular JIA	<i>STAT1</i>	0.0002

Genes present within the MEN which also show altered gene expression in either rheumatoid-negative (RF-ve) polyarticular juvenile idiopathic arthritis (JIA) compared with controls, oligoarticular JIA compared with controls, or between oligoarticular compared with RF-ve polyarticular JIA are listed. Presence in the MEN indicates functional priority and the Disease association protein-protein link evaluator (Dapple) P-value assesses the statistical significance of network connectivity compared to what would be expected by chance.



We have used a comprehensive set of genetic loci as the starting point to build the network. These loci are replicated SNP associations from candidate genes and the top statistical associations from recent GWAS and fine-mapping studies. There are limited collections of JIA patients worldwide to allow for subsequent GWAS

or future GWAS meta-analysis. Given the premise that gene-gene interactions contribute to complex diseases, combining modest association signals from GWAS analysis with biological data to build networks can help to detect the joint effects of multiple genes [53]. Also, Hua *et al.* recently described the relevance of incorporating



GWAS SNPs with a *P*-value of lower than 0.01 in network analysis [54].

The genetic network was initially built using the BioGRID database. Subsequently, a MEN, representing the core genetic clusters related to the JIA phenotype, was derived using both connectivity and bottleneck network properties [12,33,36-38,55]. Six biological groupings were associated with the MEN, with a predominance of immune-related and cellular signalling pathways being represented (Figure 2). In addition, we utilised the approach described by Rossin *et al.* [42] to provide a statistical measure to the MEN.

The value of network biology is the inter-relational display of multiple omic datasets [56]. An integrated, multi-omic approach reduces noise in the statistical interpretation [57]. Furthermore, network biology focuses on the importance of clusters [58-61] as a measure of similarity rather than simply an overlap of gene sets. It is also an approach that is robust to random variation in the data [62] and to variation over different sizes of networks [60].

Our strategy has been to build the network from genetic association findings and to refine and validate the SNP data by overlaying stringently derived gene expression data. In order to have the most stringent refinement of the network we integrated the gene expression data with the MEN [33,37,38], and then the Dapple algorithm [42] was again employed to rank the gene expression loci that aligned to the MEN, to provide a further level of statistical robustness. This analysis identified JIA subgroup-specific loci.

The key seed gene for RF-ve polyarticular JIA is the scaffold kinase suppressor of Ras (*KSR1*). In three of the five gene expression datasets *KSR1* was over-expressed in RF-ve polyarticular JIA (Additional file 4: Table S4). *KSR1* acts as a location-regulated scaffolding protein connecting MEK to RAF. It promotes MEK and RAF phosphorylation and activity through assembly of an activated signalling complex. By itself, however, it has not been demonstrated to have kinase activity. Importantly, Fusello *et al.* established that *KSR1* knockout mice have reduced susceptibility to rheumatoid arthritis [63]. As *KSR1*-deficient T cells are functionally impaired [64] Fusello tested *Ksr*-deficient mice using a passive transfer model of arthritis. Their findings showed that the induction of arthritis is impaired in the absence of *KSR1* and that this gene plays a role in ERK activation during inflammatory and stress responses both *in vitro* and *in vivo*. Using the MCODE algorithm, *KSR1* interacts with a sub-cluster of 14 other proteins including a functional relationship with RAF1 (Figure 4B). We propose that this sub-cluster is of key functional relevance to the pathogenesis of RF-ve polyarticular JIA.

For oligoarticular JIA *PTPN2* was the seed gene of importance. *PTPN2* was significantly altered (down-regulated)

in expression and mapped to the MEN (Additional file 5: Table S5). The cluster in relation to *PTPN2* is shown, (Figure 4A), and consists of 23 proteins, the majority of which have a role in cytokine signalling and DNA repair.

*PTPN2* encodes for the ubiquitously expressed T cell protein tyrosine phosphatase (TCPTP), a JAK/STAT and growth factor receptor phosphatase that has been linked with the pathogenesis of type 1 diabetes mellitus, rheumatoid arthritis and Crohn's disease by GWAS findings of non-coding SNP associations. Mouse and human studies have shown that reduced expression of TCPTP may drive autoimmune pathologies by enhancing signalling downstream of the T cell receptor (TCR), cytokines, or growth factors to produce a pro-inflammatory cytokine milieu [65].

Hinks *et al.* reported the *C5orf56-IRF1* region to show differential association between oligoarticular and RF-ve polyarticular JIA [7]. *IRF1* is present in the MEN (Figure 2B & Additional file 3: Table S3A) and *IRF1* gene expression is significantly increased in oligoarticular compared to RF-ve polyarticular JIA (Additional file 6: Table S6). It has borderline significance in post iterative modelling (Dapple) ( $P = 0.07$ ). *STAT1*, a signal transducer and transcription activator that mediates cytokine activity and growth factor, mapped within the MEN, was differentially expressed, with higher expression occurring in oligoarticular compared to RF-ve polyarticular disease, and met the FDR cut off for Dapple assignment of  $P \leq 0.05$ . *STAT1* is present in the oligoarticular JIA but not the RF-ve polyarticular JIA sub-cluster. It appears, therefore, to represent a gene that differentiates between the oligoarticular and RF-ve polyarticular clinical subtypes. Further evaluation of *IRF1* and *STAT1* in persistent versus extended oligoarticular JIA is not possible as neither the genetic association nor the gene expression studies have classified oligoarticular JIA cases in this way.

The main observations from the GWAS analysis by Thompson *et al.* [6] relate to the *C3orf1* and *JMJD1C* genes. We find the proteins for these loci to be peripheral within our interactome network and they are not prioritised by Dapple. This implies that although these loci may have a role in JIA pathogenesis their lack of connectivity limits their druggable potential (reviewed in [66]).

Functional assessment of the RF-ve polyarticular and oligoarticular clusters needs to be determined. Mounting evidence indicates that biological systems are organised as modular networks, in which genes, proteins, metabolites and other factors operate in groups rather than as single entities. There is increasing recognition that transcription factors, micro RNAs, DNA methylation and chromatin remodelling regulate the expression of large numbers of genes in concert [67].

It is intriguing that preliminary analysis using the TargetScan human database [68] shows that the three

candidate genes prioritised by this study (*KSR1*, *PTPN2* and *STAT1*) are co-regulated by miR-30a-5p ( $P = 5.3 \times 10^{-4}$ ). Whilst hundreds of miRNAs have been identified to be dysregulated in various disease tissues, only a fraction have been functionally characterised. miR 30a-5p been shown to be differentially expressed and to have biological function regulating expression of genes in a number of diseases including systemic lupus erythematosus (SLE) [69], rodent models of diabetes mellitus [70] glioma growth [71] and carcinoma of the colon [72]. Individual miRNAs can regulate several hundred transcripts with effector molecules that function at various sites within cellular pathways and networks, making them master regulators of the genome. miRNA-based therapy is progressing [73]. miRNAs of potential therapeutic value are those that yield satisfactory efficacy in disease model systems and mechanistic data to allow accurate placement of the miRNA into disease-related pathways. The network analysis we have conducted so far allows us to begin to investigate the capacity of miR 30a-5p in a targeted way in JIA.

Repositioning of drugs is another important therapeutic advantage that can be maximised from network analysis [20,21]. The *KSR1* sub-cluster (Figure 4B) could be therapeutically targeted in RF-ve polyarticular JIA patients, using small-molecule inhibitors of RAF1. Repurposing a drug such as sorafenib (nexavar), currently licensed for the treatment of renal and hepatocellular carcinoma, for the management of RF-ve polyarticular JIA could therefore be considered.

There are limitations to the present analysis. First, we have utilised data from a single GWAS, that is, the only one that currently exists for JIA, and included SNP associations of  $P \leq 1 \times 10^{-4}$ . The widening of input data for the generation of initial interactome networks is, however, acceptable and valid [53,54]. In addition, we have used the data from dense genotyping of immune-related loci [7]. Although only a proportion of the genes potentially involved in JIA pathogenesis may be immune-related, such pathways were highly significant within our network (Figure 2C). Interactome network models can be used to map associated biological function in a sensitive and robust manner via network clusters using stringent downstream analysis [33]. This is the strategy we have applied. Furthermore, if subsequent JIA GWAS or replication of new candidate gene associations do emerge, our network model can readily be adapted to incorporate the additional findings.

Second, there were only four gene expression studies suitable for inclusion. However, these four studies allowed comparisons across different cell populations (PBMCs, neutrophils, synovial fluid monocytes). Including gene expression data, significant across these sites, with the genetic association data and network analysis is a robust approach to identifying major contributing loci

to the disease phenotype. Such integrative approaches have improved statistical power [56,57].

Third, the functional significance of the protein clusters remains to be established. The collection of biological material from oligoarticular and RF-ve polyarticular JIA patients is now underway to facilitate this.

## Conclusions

To our knowledge this is the first use of network analysis to integrate JIA genetic and expression data for the prioritisation of loci for further functional assessment. The value of the use of network biology is to increase confidence in the observations of differentially expressed genes, and genetic findings, by correlation with functionally related network structure. This enhances the identification of functional units or clusters and, critically, can inform the targeting of new therapies. We have used different software-based methods (BioGRID, Ingenuity, Reactome) to infer relationships of differentially expressed genes and genetic association data with known interactions in the literature or protein databases. Dapple has been used at two stages in the analysis to add statistical robustness to our findings.

Deciphering the basis of complex pathologies such as JIA is challenging and although certain progress has been made, much remains to be understood. Deriving a JIA interactome model and sub-clusters, as described herein, offers the potential for a new era in addressing the mechanisms and future interventions for this chronic, disabling condition.

## Additional files

**Additional file 1: Table S1.** Single nucleotide polymorphism (SNP) datasets: replicated loci showing association with oligoarticular and rheumatoid factor-negative (RF-ve) polyarticular juvenile idiopathic arthritis (JIA) were identified from published literature and collated with top genome-wide association studies (GWAS) findings (A). The collated list of genes (B) was used as the seed genes for the network analysis and the generation of the JIA interactome (Figure 1).

**Additional file 2: Table S2.** Gene expression data: juvenile idiopathic arthritis (JIA) gene expression datasets were identified from the Gene Expression Omnibus (GEO). Details of the datasets include sample and control numbers, comparisons made, number of probe sets identified by analysis of variance (ANOVA) as significant ( $P < 0.05$ ), number of individual genes with a gene expression change, confounding factors included as covariants in the ANOVA and type of Affymetrix Chip. PBMC, peripheral blood mononuclear cells; SFM, synovial fluid mononuclear cells; Neut, neutrophils. Overlap groups used for comparisons are colour coded.

**Additional file 3: Table S3.** Network topology of the juvenile idiopathic arthritis (JIA) interactome: (A) Genes from the minimal essential network (MEN), that is, the top 10% of genes from the JIA interactome (ranked by degree (Deg) and bottleneck score (BN)). (B) JIA seed genes from the JIA interactome ranked by significance of deviation of observed network connectivity from expected (Disease Association Protein-Protein Link Evaluator (Dapple) algorithm,  $P \leq 0.05$ ). Orange = also present in minimal essential network (Figure 2B).

**Additional file 4: Table S4.** Rheumatoid factor-negative (RF-ve) polyarticular juvenile idiopathic arthritis (JIA) versus control gene

expression: overlap of genes in RF-ve polyarticular JIA compared to controls present in three of the five gene expression datasets analysed. Datasets described in Additional file 2: Table S2. Red = up-regulated in JIA, green = down-regulated in JIA.

**Additional file 5: Table S5.** Oligoarticular versus control gene expression: overlap of genes in oligoarticular juvenile idiopathic arthritis (JIA) compared to controls present in both gene expression datasets analysed. Datasets described in Additional file 2: Table S2. Red = up-regulated in JIA, green = down-regulated in JIA.

**Additional file 6: Table S6.** Oligoarticular versus rheumatoid factor-negative (RF-ve) polyarticular juvenile idiopathic arthritis (JIA) gene expression: overlap of genes in oligoarticular JIA compared to RF-ve polyarticular JIA present in both gene expression datasets analysed. Datasets described in Additional file 2: Table S2. Red = up-regulated in JIA, green = down-regulated in JIA.

#### Abbreviations

ANOVA: analysis of variance; C3orf1: Chromosome 3 open reading frame 1; C5orf56: Chromosome 5 open reading frame 56; Dapple: Disease Association Protein-Protein Link Evaluator; FDR: false discovery rate; GEO: Gene Expression Omnibus; Grail: Gene Relationships Across Implicated Loci; GWAS: genome-wide association studies; ILAR: International League Against Rheumatism; IPA: Ingenuity pathway analysis; IRF1: Interferon regulatory factor 1; JIA: juvenile idiopathic arthritis; JMJD1C: Jumonji domain containing 1C; KSR1: Kinase suppressor of Ras; MEN: minimal essential network; PBMC: peripheral blood mononuclear cell; PCA: principal components analysis; PTPN2: Protein tyrosine phosphatase, non-receptor type 2; RF: rheumatoid factor; SFM: synovial fluid monocytes; SNP: single nucleotide polymorphism; STAT1: Signal transducer and activator of transcription 1.

#### Competing interests

The authors declare that they have no competing interests.

#### Authors' contributions

AS and RD conceived the idea for integrating genetic and transcriptomic datasets to investigate JIA. AS performed the analyses and developed the minimal essential network as a study tool. DH contributed to and supported network analysis of data. AS, RD, SM and PC designed the analysis plan and the presentation of the data. AS and RD led the writing of the manuscript with input from all authors. All authors read and approved the final manuscript.

#### Acknowledgements

AS is funded in part by a grant from Merck Serono SA, Geneva, Switzerland to assess the Pharmacogenomics of human growth disorders. DH is funded by the University of Manchester Wellcome Trust Institutional Strategic Support Fund. SM is supported by Leukaemia Lymphoma Research, UK, and the Kay Kendall Leukaemia Fund.

#### Author details

<sup>1</sup>Manchester Academic Health Sciences Centre, Royal Manchester Children's Hospital, Manchester M13 9WL, UK. <sup>2</sup>Stem Cell and Leukaemia Proteomics Laboratory, School of Cancer and Imaging Sciences, University of Manchester, Manchester, UK. <sup>3</sup>Centre for Musculoskeletal Research, University of Manchester, Manchester M13 9PT, UK.

Received: 6 January 2014 Accepted: 29 April 2014

Published: 8 May 2014

#### References

1. Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J, He X, Maldonado-Cocco J, Orozco-Alcala J, Prieur AM, Suarez-Almazor ME, Woo P: **International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001.** *J Rheumatol* 2004, **31**:390–392.
2. Hyrich KL, Lal SD, Foster HE, Thornton J, Adib N, Baildam E, Gardner-Medwin J, Wedderburn LR, Chieng A, Davidson J, Thomson W: **Disease activity and disability in children with juvenile idiopathic arthritis one year following presentation to paediatric rheumatology. Results from the Childhood Arthritis Prospective Study.** *Rheumatology (Oxford)* 2010, **49**:116–122.

3. Ravelli A, Martini A: **Juvenile idiopathic arthritis.** *Lancet* 2007, **369**:767–778.
4. Bertilsson L, Andersson-Gare B, Fasth A, Petersson IF, Forsblad-d'Elia H: **Disease course, outcome, and predictors of outcome in a population-based juvenile chronic arthritis cohort followed for 17 years.** *J Rheumatol* 2013, **40**:715–724.
5. Prahalad S, Glass DN: **A comprehensive review of the genetics of juvenile idiopathic arthritis.** *Pediatr Rheumatol Online J* 2008, **6**:11.
6. Thompson SD, Marion MC, Sudman M, Ryan M, Tsoras M, Howard TD, Barnes MG, Ramos PS, Thomson W, Hinks A, Haas JP, Prahalad S, Bohnsack JF, Wise CA, Punaro M, Rosé CD, Pajewski NM, Spigarelli M, Keddache M, Wagner M, Langefeld CD, Glass DN: **Genome-wide association analysis of juvenile idiopathic arthritis identifies a new susceptibility locus at chromosomal region 3q13.** *Arthritis Rheum* 2012, **64**:2781–2791.
7. Hinks A, Cobb J, Marion MC, Prahalad S, Sudman M, Bowes J, Martin P, Comeau ME, Sajuthi S, Andrews R, Brown M, Chen WM, Concannon P, Deloukas P, Edkins S, Eyre S, Gaffney PM, Guthery SL, Guthridge JM, Hunt SE, James JA, Keddache M, Moser KL, Nigrovic PA, Onengut-Gumuscu S, Onslow ML, Rosé CD, Rich SS, Steel KJ, Wakeland EK, et al: **Dense genotyping of immune-related disease regions identifies 14 new susceptibility loci for juvenile idiopathic arthritis.** *Nat Genet* 2013, **45**:664–669.
8. Barnes MG, Grom AA, Thompson SD, Griffin TA, Luyrink LK, Colbert RA, Glass DN: **Biologic similarities based on age at onset in oligoarticular and polyarticular subtypes of juvenile idiopathic arthritis.** *Arthritis Rheum* 2010, **62**:3249–3258.
9. Barnes MG, Aronow BJ, Luyrink LK, Moroldo MB, Pavlidis P, Passo MH, Grom AA, Hirsch R, Giannini EH, Colbert RA, Glass DN, Thompson SD: **Gene expression in juvenile arthritis and spondyloarthropathy: pro-angiogenic ELR + chemokine genes relate to course of arthritis.** *Rheumatology (Oxford)* 2004, **43**:973–979.
10. Frank MB, Wang S, Aggarwal A, Knowlton N, Jiang K, Chen Y, McKee R, Chaser B, McGhee T, Osban J, Jarvis JN: **Disease-associated pathophysiologic structures in pediatric rheumatic diseases show characteristics of scale-free networks seen in physiologic systems: implications for pathogenesis and treatment.** *BMC Med Genomics* 2009, **2**:9.
11. Chang D, Keinan A: **Predicting signatures of "synthetic associations" and "natural associations" from empirical patterns of human genetic variation.** *PLoS Comput Biol* 2012, **8**:e1002600.
12. Lee Y, Li H, Li J, Rebman E, Achour I, Regan KE, Gamazon ER, Chen JL, Yang XH, Cox NJ, Lussier YA: **Network models of genome-wide association studies uncover the topological centrality of protein interactions in complex diseases.** *J Am Med Assoc* 2013, **206**:619–629.
13. Yeang CH, Haussler D: **Detecting coevolution in and among protein domains.** *PLoS Comput Biol* 2007, **3**:e211.
14. Agarwal S, Deane CM, Porter MA, Jones NS: **Revisiting date and party hubs: novel approaches to role assignment in protein interaction networks.** *PLoS Comput Biol* 2010, **6**:e1000817.
15. Han JD, Bertin N, Hao T, Goldberg DS, Berriz GF, Zhang LV, Dupuy D, Walhout AJ, Cusick ME, Roth FP, Vidal M: **Evidence for dynamically organized modularity in the yeast protein-protein interaction network.** *Nature* 2004, **430**:88–93.
16. Lemos B, Meiklejohn CD, Hartl DL: **Regulatory evolution across the protein interaction network.** *Nat Genet* 2004, **36**:1059–1060.
17. Duan S, Luo X, Dong C: **Identification of susceptibility modules for coronary artery disease using a genome wide integrated network analysis.** *Gene* 2013, **531**:347–354.
18. Vidal M, Cusick ME, Barabasi AL: **Interactome networks and human disease.** *Cell* 2011, **144**:986–998.
19. Goh KI, Cusick ME, Valle D, Childs B, Vidal M, Barabasi AL: **The human disease network.** *Proc Natl Acad Sci USA* 2007, **104**:8685–8690.
20. Randhawa V, Bagler G: **Identification of SRC as a potent drug target for asthma, using an integrative approach of protein interactome analysis and in silico drug discovery.** *OMICS* 2012, **16**:513–526.
21. Zhao S, Iyengar R: **Systems pharmacology: network analysis to identify multiscale mechanisms of drug action.** *Annu Rev Pharmacol Toxicol* 2012, **52**:505–521.
22. PubMed. [http://www.ncbi.nlm.nih.gov/pubmed]
23. Web of Science. [www.webofknowledge.com]
24. Donn R, Alourfi Z, Zeggini E, Lamb R, Jury F, Lunt M, Meazza C, De BF, Thomson W, Ray D: **A functional promoter haplotype of macrophage migration inhibitory factor is linked and associated with juvenile idiopathic arthritis.** *Arthritis Rheum* 2004, **50**:1604–1610.



25. Lamb R, Thomson W, Ogilvie E, Donn R: **Wnt-1-inducible signaling pathway protein 3 and susceptibility to juvenile idiopathic arthritis.** *Arthritis Rheum* 2005, **52**:3548–3553.
26. Hinks A, Eyre S, Ke X, Barton A, Martin P, Flynn E, Packham J, Worthington J, Thomson W: **Overlap of disease susceptibility loci for rheumatoid arthritis and juvenile idiopathic arthritis.** *Ann Rheum Dis* 2010, **69**:1049–1053.
27. Thompson SD, Sudman M, Ramos PS, Marion MC, Ryan M, Tsoras M, Weiler T, Wagner M, Keddache M, Haas JP, Mueller C, Prahalad S, Bohnsack J, Wise CA, Punaro M, Zhang D, Rosé CD, Comeau ME, Divers J, Glass DN, Langefeld CD: **The susceptibility loci Juvenile Idiopathic Arthritis shares with other autoimmune diseases extend to PTPN22, COG6 and ANGP11.** *Arthritis Rheum* 2010, **62**:3265–3276.
28. Hinks A, Barton A, John S, Bruce I, Hawkins C, Griffiths CE, Donn R, Thomson W, Silman A, Worthington J: **Association between the PTPN22 gene and rheumatoid arthritis and juvenile idiopathic arthritis in a UK population: further support that PTPN22 is an autoimmunity gene.** *Arthritis Rheum* 2005, **52**:1694–1699.
29. Hinks A, Cobb J, Sudman M, Eyre S, Martin P, Flynn E, Packham J, Childhood Arthritis Prospective Study (CAPS); UK RA Genetics (UKRAG) Consortium; British Society of Paediatric and Adolescent Rheumatology (BSPAR) Study Group, Barton A, Worthington J, Langefeld CD, Glass DN, Thompson SD, Thomson W: **Investigation of rheumatoid arthritis susceptibility loci in juvenile idiopathic arthritis confirms high degree of overlap.** *Ann Rheum Dis* 2012, **71**:1117–1121.
30. Hinks A, Barton A, Shephard N, Eyre S, Bowes J, Cargill M, Wang E, Ke X, Kennedy GC, John S, Worthington J, Thomson W, British Society of Paediatric and Adolescent Rheumatology Study Group: **Identification of a novel susceptibility locus for juvenile idiopathic arthritis by genome-wide association analysis.** *Arthritis Rheum* 2009, **60**:258–263.
31. Raychaudhuri S, Plenge RM, Rossin EJ, Ng AC, Purcell SM, Sklar P, Scolnick EM, Xavier RJ, Altshuler D, Daly MJ: **Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions.** *PLoS Genet* 2009, **5**:e1000534.
32. **GRAIL: Gene Relationships Across Implicated Loci.** [www.broadinstitute.org/mpg/grail/]
33. Stevens A, De Leonibus C, Hanson D, Dowsey AW, Whatmore A, Meyer S, Donn RP, Chatelain P, Banerjee I, Cosgrove KE, Clayton PE, Dunne MJ: **Network analysis: a new approach to study endocrine disorders.** *J Mol Endocrinol* 2013, **52**:R79–R93.
34. Chatr-Aryamontri A, Breitkreutz BJ, Heinicke S, Boucher L, Winter A, Stark C, Nixon J, Ramage L, Kolas N, O'Donnell L, Reguly T, Breitkreutz A, Sellam A, Chen D, Chang C, Rust J, Livstone M, Oughtred R, Dolinski K, Tyers M: **The BioGRID interaction database: 2013 update.** *Nucleic Acids Res* 2013, **41**:D816–D823.
35. Lin CY, Chin CH, Wu HH, Chen SH, Ho CW, Ko MT: **Hubba: hub objects analyzer—a framework of interactive hubs identification for network biology.** *Nucleic Acids Res* 2008, **36**:W438–W443.
36. Yu H, Kim PM, Sprecher E, Trifonov V, Gerstein M: **The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics.** *PLoS Comput Biol* 2007, **3**:e59.
37. Stevens A, Hanson D, Whatmore A, Destenaves B, Chatelain P, Clayton P: **Human growth is associated with distinct patterns of gene expression in evolutionarily conserved networks.** *BMC Genomics* 2013, **14**:547.
38. Stevens A, Clayton P, Tatò L, Yoo HW, Rodriguez-Arnao MD, Skorodok J, Ambler GR, Zignani M, Zieschang J, Della Corte G, Destenaves B, Champigneulle A, Raelson J, Chatelain P: **Pharmacogenomics of insulin-like growth factor-I generation during GH treatment in children with GH deficiency or Turner syndrome.** *Pharmacogenomics J* 2013, **14**:54–62.
39. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T: **Cytoscape 2.8: new features for data integration and network visualization.** *Bioinformatics* 2011, **27**:431–432.
40. Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, Christmas R, Avila-Campilo I, Creech M, Gross B, Hanspers K, Isserlin R, Kelley R, Killcoyne S, Lotia S, Maere S, Morris J, Ono K, Pavlovic V, Pico AR, Vailaya A, Wang PL, Adler A, Conklin BR, Hood L, Kuiper M, Sander C, Schmulevich I, Schwikowski B, Warner GJ, et al: **Integration of biological networks and gene expression data using Cytoscape.** *Nat Protoc* 2007, **2**:2366–2382.
41. Newman ME: **Modularity and community structure in networks.** *Proc Natl Acad Sci USA* 2006, **103**:8577–8582.
42. Rossin EJ, Lage K, Raychaudhuri S, Xavier RJ, Tatar D, Benita Y, Cotsapas C, Daly MJ: **Proteins encoded in genomic regions associated with immune-mediated disease physically interact and suggest underlying biology.** *PLoS Genet* 2011, **7**:e1001273.
43. **The DAPPLE algorithm.** [www.broadinstitute.org/mpg/dapple/dapple.php]
44. **InWeb Database.** [www.cbs.dtu.dk/suppl/dgf/]
45. **Gene Expression Omnibus.** [www.ncbi.nlm.nih.gov/geo/]
46. Lee HM, Sugino H, Aoki C, Nishimoto N: **Underexpression of mitochondrial-DNA encoded ATP synthesis-related genes and DNA repair genes in systemic lupus erythematosus.** *Arthritis Res Ther* 2011, **13**:R63.
47. Tenenbaum JB, De SV, Langford JC: **A global geometric framework for nonlinear dimensionality reduction.** *Science* 2000, **290**:2319–2323.
48. Nilsson J, Fioretos T, Hoglund M, Fontes M: **Approximate geodesic distances reveal biologically relevant structures in microarray data.** *Bioinformatics* 2004, **20**:874–880.
49. Bader GD, Hogue CW: **An automated method for finding molecular complexes in large protein interaction networks.** *BMC Bioinforma* 2003, **4**:2.
50. Benjamini YHY: **Controlling the false discovery rate: a practical and powerful approach to multiple testing.** *J R Stat Soc Ser B Methodological* 1995, **57**:289–300.
51. Ravasz E, Barabasi AL: **Hierarchical organization in complex networks.** *Phys Rev E Stat Nonlin Soft Matter Phys* 2003, **67**:026112.
52. van Dongen S, Breu-Goodger C: **Using MCL to extract clusters from networks.** *Methods Mol Biol* 2012, **804**:281–295.
53. Pedroso I: **Gaining a pathway insight into genetic association data.** *Methods Mol Biol* 2010, **628**:373–382.
54. Hua L, Zhou P, Liu H, Li L, Yang Z, Liu ZC: **Mining susceptibility gene modules and disease risk genes from SNP data by combining network topological properties with support vector regression.** *J Theor Biol* 2011, **289**:225–236.
55. Sun YV: **Integration of biological networks and pathways with genetic association studies.** *Hum Genet* 2012, **131**:1677–1686.
56. Choi H, Pavelka N: **When one and one gives more than two: challenges and opportunities of integrative omics.** *Front Genet* 2011, **2**:105.
57. Ideker T, Dutkowski J, Hood L: **Boosting signal-to-noise in complex biology: prior knowledge is power.** *Cell* 2011, **144**:860–863.
58. Aloy P, Russell RB: **Taking the mystery out of biological networks.** *EMBO Rep* 2004, **5**:349–350.
59. Arrell DK, Terzic A: **Interpreting networks in systems biology.** *Clin Pharmacol Ther* 2013, **93**:389–392.
60. Barabasi AL, Albert R: **Emergence of scaling in random networks.** *Science* 1999, **286**:509–512.
61. Barabasi AL: **Scale-free networks: a decade and beyond.** *Science* 2009, **325**:412–413.
62. Albert R, Jeong H, Barabasi AL: **Error and attack tolerance of complex networks.** *Nature* 2000, **406**:378–382.
63. Fusello AM, Mandik-Nayak L, Shih F, Lewis RE, Allen PM, Shaw AS: **The MAPK scaffold kinase suppressor of Ras is involved in ERK activation by stress and proinflammatory cytokines and induction of arthritis.** *J Immunol* 2006, **177**:6152–6158.
64. Nguyen A, Burack WR, Stock JL, Kortum R, Chaika OV, Afkarian M, Muller WJ, Murphy KM, Morrison DK, Lewis RE, McNeish J, Shaw AS: **Kinase suppressor of Ras (KSR) is a scaffold which facilitates mitogen-activated protein kinase activation in vivo.** *Mol Cell Biol* 2002, **22**:3035–3045.
65. Zikherman J, Weiss A: **Unraveling the functional implications of GWAS: how T cell protein tyrosine phosphatase drives autoimmune disease.** *J Clin Invest* 2011, **121**:4618–4621.
66. Cserehely P, Korcsmaros T, Kiss HJ, London G, Nussinov R: **Structure and dynamics of molecular networks: A novel paradigm of drug discovery: A comprehensive review.** *Pharmacol Ther* 2013, **138**:333–408.
67. Weiss JN, Karma A, MacLellan WR, Deng M, Rau CD, Rees CM, Wang J, Wisniewski N, Eskin E, Horvath S, Qu Z, Wang Y, Lusis AJ: **“Good enough solutions” and the genetics of complex diseases.** *Circ Res* 2012, **111**:493–504.
68. **TargetScan human database.** [http://www.targetscan.org/]
69. Dai Y, Sui W, Lan H, Yan Q, Huang H, Huang Y: **Comprehensive analysis of microRNA expression patterns in renal biopsies of lupus nephritis patients.** *Rheumatol Int* 2009, **29**:749–754.
70. Kim JW, You YH, Jung S, Suh-Kim H, Lee IK, Cho JH, Yoon KH: **miRNA-30a-5p-mediated silencing of Beta2/NeuroD expression is an important initial event of glucotoxicity-induced beta cell dysfunction in rodent models.** *Diabetologia* 2013, **56**:847–855.



71. Wang X, Wang K, Han L, Zhang A, Shi Z, Zhang K, Zhang H, Yang S, Pu P, Shen C, Yu C, Kang C: **PRDM1 is directly targeted by miR-30a-5p and modulates the Wnt/beta-catenin pathway in a Dkk1-dependent manner during glioma growth.** *Cancer Lett* 2013, **331**:211–219.
72. Baraniskin A, Birkenkamp-Demtroder K, Maghnouj A, Zollner H, Munding J, Klein-Scory S, Reinacher-Schick A, Schwarte-Waldhoff I, Schmiegel W, Hahn SA: **MiR-30a-5p suppresses tumor growth in colon carcinoma by targeting DTL.** *Carcinogenesis* 2012, **33**:732–739.
73. Wu Y, Crawford M, Mao Y, Lee RJ, Davis IC, Elton TS, Lee LJ, Nana-Sinkam SP: **Therapeutic delivery of microRNA-29b by cationic lipoplexes for lung cancer.** *Mol Ther Nucleic Acids* 2013, **2**:e84.

doi:10.1186/ar4559

**Cite this article as:** Stevens *et al.*: Network analysis identifies protein clusters of functional importance in juvenile idiopathic arthritis. *Arthritis Research & Therapy* 2014 **16**:R109.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

