

A comparative study of cells in inflammation, EAE and MS using biomedical literature data mining

Mathew Palakal^{1,*}, John Bright², Thomas Sebastian¹ & Shielly Hartanto¹

¹*Department of Computer and Information Science, Indiana University-Purdue University Indianapolis, 723 W. Michigan St. SL 280, Indianapolis, IN, 46202, USA;* ²*Methodist Research Institute, Clarian Health Partners, Indianapolis, IN, 46202, USA*

Received 21 April 2006; accepted 20 September 2006
© 2006 National Science Council, Taipei

Key words: biomedical literature mining, association discovery, association graph, inflammation, multiple sclerosis (MS), experimental allergic encephalomyelitis (EAE)

Summary

Biomedical literature and database annotations, available in electronic forms, contain a vast amount of knowledge resulting from global research. Users, attempting to utilize the current state-of-the-art research results are frequently overwhelmed by the volume of such information, making it difficult and time-consuming to locate the relevant knowledge. Literature mining, data mining, and domain specific knowledge integration techniques can be effectively used to provide a user-centric view of the information in a real-world biological problem setting. Bioinformatics tools that are based on real-world problems can provide varying levels of information content, bridging the gap between biomedical and bioinformatics research. We have developed a user-centric bioinformatics research tool, called BioMap, that can provide a customized, adaptive view of the information and knowledge space. BioMap was validated by using inflammatory diseases as a problem domain to identify and elucidate the associations among cells and cellular components involved in multiple sclerosis (MS) and its animal model, experimental allergic encephalomyelitis (EAE). The BioMap system was able to demonstrate the associations between cells directly excavated from biomedical literature for inflammation, EAE and MS. These association graphs followed the scale-free network behavior (average $\gamma = 2.1$) that are commonly found in biological networks.

Introduction

Biomedical literature contains a wealth of knowledge whose automated extraction is still a tedious task. A simple query on inflammation related diseases returned over 45,000 documents from the MedLine [1] database alone. This provides an indication of the large volume of scientific information available in the form of text. Associations among biological entities such as genes, proteins, molecules, processes, diseases, drugs and chemicals, are one such form of underlying knowledge

available in literature. Deep mining that involves extracting directional, transitive, and hyper associations may unravel novel plausible biological hypotheses.

Bioinformatics tools based on real-world problems can provide varying levels of information content, bridging the gap between biomedical and bioinformatics research. Using inflammatory diseases as a problem domain, in this paper we show that the tools proposed can provide researchers with pertinent information to address specific biological research questions.

There exist numerous public domain biological data sources that concentrate on the management of information pertaining to a particular type of

*To whom correspondence should be addressed. Fax: +1-317-274-9742; E-mail: mpalakal@cs.iupui.edu

biological concept, such as RNA (e.g. RNABase [2]), DNA and protein sequences (e.g. GenBank [3], SwissProt [4]), protein families (e.g. InterPro [5]), gene functions (Gene Ontology [6]), metabolic pathways (e.g. KEGG [7], EcoCyc [8]), or diseases (e.g. OMIM [9]). However, biological analysis tasks generally require these concepts to be considered in their biological context, i.e. their relationships and interactions to/with other objects. Most of these data sources only provide web links for navigational access along the network of biological relationships, making it impossible to perform automated analysis for large data sets across multiple dimensions [10, 11].

There are several commercially available software products that support knowledge discovery from biomedical databases. Ingenuity System [12] has Ingenuity Pathways Knowledge Base, which is the world's largest curated database of biological networks that enables biologists to discover, visualize and explore therapeutically relevant networks significant to their experimental results. The PharmGKB [13] is a publicly available Internet research tool developed by Stanford University that assists researchers in understanding how genetic variation among individuals contributes to differences in reactions to drugs. The TAMBIS tool [14] attempts to provide transparent information retrieval and filtering from biological information services by building a homogenizing layer on top of the different sources. The Incyte's Proteome BioKnowledge Library [15] consists of six database volumes of biological information about proteins. Each volume focuses on a different organism important in pharmaceutical research. PubGene [16] is a proprietary software package designed to identify relationships between genes and proteins based on their co-occurrence in the abstracts of scientific papers, on their sequence homology, and the statistical probability of their co-occurrence with each other and Medical Subject Headings (MeSH) and Gene Ontology (GO) terms, such as diseases and cell processes. ROSETTA [17] is a toolkit designed to support the data mining and knowledge discovery process. The Genostar program also [18] provides researchers with a bioinformatics platform for exploratory genomics.

Some commercial systems find directional relations between objects and build biological pathways. Two major vendors of such systems

are Ariadnegenomics [19] and Ingenuity Systems. The Ariadnegenomics tool, called Pathway Assist, is a pathway analysis software that helps to interpret experiment results in the context of pathways, gene regulation networks and protein interaction maps. It identifies relationships among different biological entities using curated and automatically created databases. The major part of the knowledgebases used by the tool is manually curated by experts in the domain. The part of the system that identifies relations is called Medscan [20, 21]. Ingenuity System provides a number of tools that help biologists in their research. Ingenuity Pathways Knowledgebase has high quality networks created from millions of expert-curated relationships between proteins, genes, complexes, cells, tissues, drugs and diseases.

There is still a need to provide a user-centered and problem specific view of the underlying biological information space. For example, with Pathway Assist, the user is limited to specific knowledgebases provided by the vendor and does not have the ability to create his/her own knowledgebase. Moreover, an important question in biological knowledge management is whether it is possible to generate a novel hypothesis concerning associations between biological objects, based on existing associations as presented in the literature. We showed that such hypothesis generation can be achieved by computing transitive associations using the transitivity property on an underlying association graph [22].

Our efforts in literature mining primarily dealt with *data* $\rightarrow$ *information* and *information* $\rightarrow$ *knowledge* transformation problems. We developed tools such as the rFinder to extract biological entity identification and discover relationships [22–26]. We also investigated the problem of formulating new hypotheses concerning associations using the transitivity property [21, 27]. These tools were integrated into a knowledgebase called BioMap that can perform basic literature mining tasks and create a domain specific knowledge core based on the user's needs. BioMap is currently populated with literature knowledge extracted from MedLine abstracts for the following domains: Inflammation, Tissue regeneration, Arabidopsis, Lung Cancer, and Alzheimer disease.

In this paper we illustrate the utilization of such an integrated tool in the research related to multiple sclerosis (MS), an inflammatory demyelinating

disease of the central nervous system (CNS) that afflicts millions of people worldwide [28]. These findings suggest that early in the disease, ongoing inflammatory activity is present in most patients and is responsible for the relapsing-remitting course of MS, whereas a distinct process might be operative in the progressive phase of the disease, when inflammatory activity diminishes despite the greater presentation of disability [29] by the subject. Activation of immune cells, secretion of inflammatory cytokines and differentiation of encephalitogenic Th1 cells are key processes associated with the pathogenesis of MS [30–32]. Although immunosuppressive agents have been used to treat MS, there is no medical treatment available so far that can cure MS.

The experimental allergic encephalomyelitis (EAE) on the other hand is a prototypic animal model for multiple sclerosis [33]. EAE can be induced in susceptible rodents and primates by immunization with whole brain homogenate or purified neural antigens such as myelin basic protein (MBP), myelin oligodendrocyte glycoprotein (MOG) and proteolipid protein (PLP). The antigenic epitopes of neural antigens have been mapped and adoptive transfer of T cells reactive to these epitopes is sufficient to induce the disease. Using a combination of neural antigens and animal species, it is possible to induce EAE with clinical symptoms similar to those seen in relapsing-remitting or chronic progressive MS. The clinical and pathological features of EAE show close similarity to human MS and therefore EAE has been commonly used as a model system to study the mechanism of MS pathogenesis and to test the efficacy of potential therapeutic agents for the treatment of MS [34–39].

Since the immune and inflammatory responses present in various disease conditions will be mechanistically similar, the therapeutic approaches used in one clinical setting may be useful for others as well. In this study, we have analyzed and compared different types of cells associated with the pathogenesis of inflammation, multiple sclerosis and its animal model, EAE. Our results validate the use of BioMap to excavate molecular components through datamining technology and to identify novel therapeutic targets by analyzing their functional association in complex disease settings such as MS.

Methodology

The overall process of discovering associations among cells in this study involves composing a set of biological terms corresponding to the problem domain of inflammation, resolving synonyms, retrieving abstracts from MedLine, tagging the documents based on parts of speech tags, identifying biological objects, extracting binary associations and, finally, discovering the directional relationships that will result in the creation of the BioMap knowledgebase for the inflammation domain. These steps are explained in detail below.

BioMap system

The BioMap system consists of three levels of data analysis and mining: Level 1, Level 2, and Level 3 (see Figure 1). Level 1 is primarily concerned with extracting knowledge from the biomedical literature database, e.g., MedLine. The user will provide a set of biological terms associated with the problem of interest, for example, inflammation. This is the process through which the system supports user-centered knowledge creation. At this level, the LitMining Engine resolves synonyms, retrieving all possible abstracts from the MedLine database; identifying relevant biological objects and binary relations; and creating a knowledgebase called LitKB [22, 23, 25, 26, 40, 41]. In Level 2, the rFinder discovers directional relationships among objects using the LitKB and stores these in a Relationship Knowledgebase, RelKB. Cross-linking knowledge in LitKB with other biomedical databases is carried out by the Gcell engine using eight different databases, and the context knowledge is stored in ContextKB. Level 3 carries out deep mining of ContextKB and RelKB using the Gcell and TransMiner Engines. The comparative study of cells associated with inflammation, EAE, and MS reported in this paper utilizes the rFinder and TransMiner tools within the BioMap system.

Domains

The domain is a specific biological research area that can be a disease, a compound, an organism, etc. In BioMap, the user can create a unique knowledgebase of a specific domain of interest by providing one or more terms associated with the

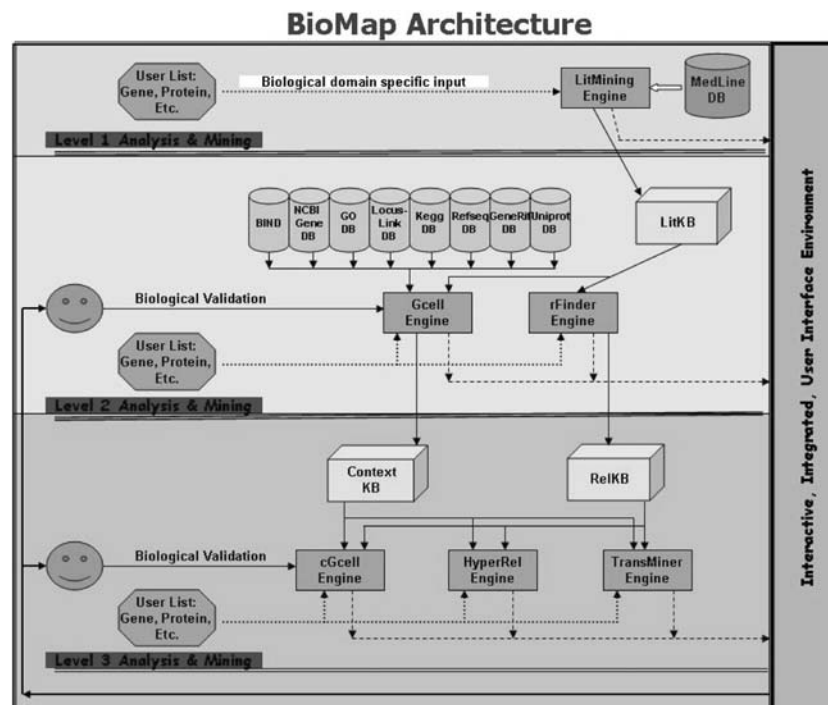


Figure 1. The BioMap System Architecture showing various interacting system components.

domain. The domains used in this study include Inflammation, EAE, and MS.

Data source and data acquisition

The text data source for BioMap is the MedLine database. BioMap uses a HTML wrapper (a parser) to retrieve title and abstract of each article from Medline for a given query. Documents related to the inflammation, EAE and MS domains were obtained from MedLine using terms such as, “acute phase response”, “anti-inflammatory agents”, “chemokines”, “autacoids”, “interleukins”, “monocytes”, “macrophages”, and several genes and proteins involved with inflammation. The total number of documents retrieved was 32,693 for MS, 5,334 for EAE, and 210,235 for inflammation. Each abstract is identified by a unique identification number. A query for each keyword is sent to MedLine in order to acquire all of the identification numbers of the abstracts. Repetition of an identification number can occur if, for instance, two or more keywords appear in the same abstract, but any repetition is ignored. This provides a unique set of abstract identification numbers. Those identification numbers are

used to download the MedLine abstracts one by one in XML format. These abstracts are stored in a local database for further mining and analysis.

BioMap user interface and data analysis

A snapshot of the main BioMap interface is shown in Figure 2. Through this interface the user can create new knowledge domains or run different mining tools on an existing knowledge base. Figure 3 shows the second interface page the user can submit broad or specific queries that would lead to specific pertinent information. In the proposed implementation, the user can select different combinations of input features. For example, one can present a query to the effect of “find all the cells that are present in inflammation but not in MS and EAE”. Different combinations of such queries can be made on the BioMap system. Once the results for such queries are returned from BioMap, the rFinder and the TransMiner tools within the BioMap system can carry out different types of association discovery (direct, transitive, and directional associations) among the biological entities.

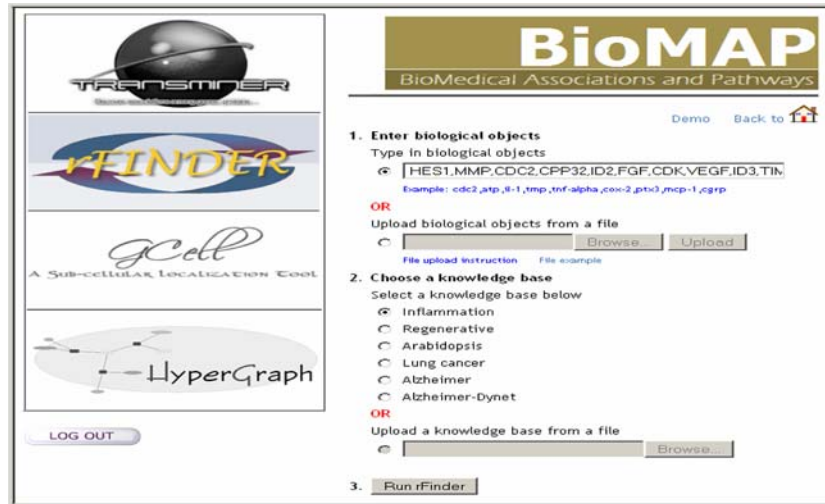


Figure 2. The initial user interface of BioMap showing some of the functionalities.

Discovering binary associations using BioMap

BioMap uses the co-occurrence principle (tf-idf method) [42] to discover direct binary associations among biological objects using the database created during the literature mining process. The algorithm for finding the direct associations and their strength using the knowledgebase LitKB, is as follows:

1. Using SELECT query, select all document identification numbers for documents where the biological objects that are being associated, are present.

```
SELECT document_id
FROM table_name
WHERE objects IN (biological_objects_list);
```

2. Calculate the number of times (frequency) each biological object appears in each document given by the result of step 1.

$tf = T_i[k]$, where: k = biological object k ;

i = document i

3. Calculate the inverse document frequency of each biological object in the document set.

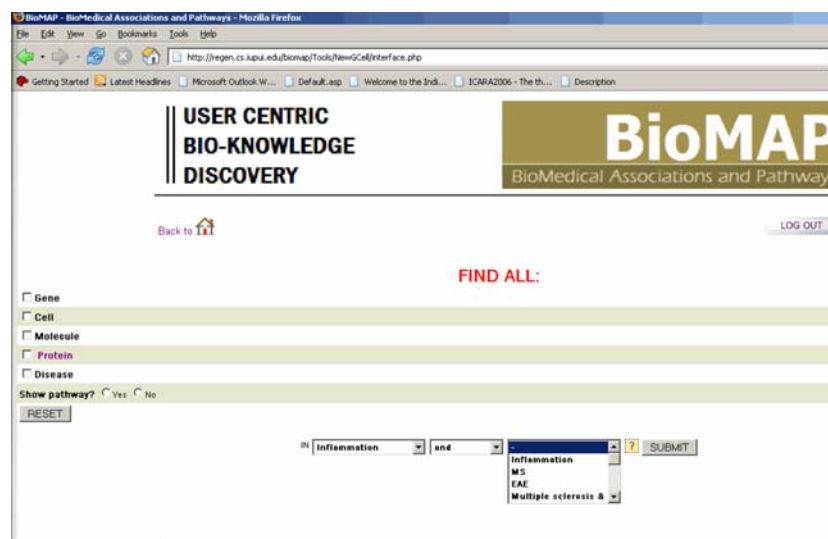


Figure 3. The query selection interface for BioMap.

$$\text{idf} = N/n[k],$$

where: N = total number of documents;
 k = biological object k ; $n[k]$ = total number of documents where biological object k appears.

4. Weight of each biological object is calculated by multiplying tf with idf.

$$\text{Weight} = \text{tf} * \text{idf}$$

$$W_i[k] = T_i[k] * \log(N/n[k])$$

where: k = biological object k ; i = document i ; $w_i[k]$ = weight of biological object k in document i .

5. Direct binary association strength between two biological objects k and l is the total of multiplication of the weight of biological object k and the weight of biological object l in every document:

$$\text{Assoc}[k][l] = \sum_{i=1}^N W_i[k] * W_i[l]$$

The association strength will return a positive number if there is an association between object k and l , no matter how small or weak it is. If there is no association between k and l , the strength will be 0. This will form a symmetric matrix of the number of biological objects. The undirected association graph for direct associations is based on positive, non-diagonal values from the matrix.

Discovering transitive associations using BioMap

Transitive associations that exist in biomedical literature provide plausible new hypotheses that need to be verified and validated experimentally. Studies by others, our own earlier work, and the feedback from our biomedical collaborators indicate that even if the recall of a potential valid association is small, discovering such associations can be extremely valuable [43]. We have developed an efficient way to discover transitive associations among biological objects using the BioMap knowledgebase.

Given a directed graph of associations, TransMiner will form the transitive closure identifying the potential path of associated objects. The transitive associations can be directly discovered by further mining the literature knowledgebase

(LitKB) computed earlier. This process first begins by computing all the direct relations as graph G , using the algorithm described in the previous section.

Relation $R(x, z)$ is transitive if there exists direct relation $R(x, y)$ and direct relation $R(y, z)$ in graph G . For example: if gene A is involved in process B and process B has a role in developing disease C, then gene A and disease C are transitively associated. Warshall's Transitive Closure [44] algorithm is used to explore every possible transitive associations using graph G computed using the previous step.

Algorithm: Transitive Association Discovery

1. Potential transitive associations are the difference between the transitive closure (G^*) and the initial association graph (G).
2. Find the object pair for each potential transitive association in G and calculate the association strength.
3. If the association strength is not zero, the association is direct.
4. If the association strength is zero, the association is transitive.
5. Repeat steps 2, 3, 4, and 5 for all the potential transitive associations discovered to get G' which contain the initial direct associations G and the newly discovered direct associations. ($G^* - G'$) are the indirect transitive associations.

The algorithm utilizes the BioMap knowledgebase to compute the transitive association resulting in an efficient implementation as compared to an earlier implementation where the search for transitive relations were carried out by iteratively fetching the documents from MedLine. This approach would also allow the extraction of all possible associations for the given problem domain as opposed to limiting the document space used in the previous implementation reported in [22].

Discovering directional associations

Finding the directionality of a relationship requires the analysis of different ways in which a single fact can be expressed in natural language. Our study indicated that there are at least 14 different ways to express a single relationship. There exist many attempts to address the prob-

lem of finding relations and their directionality [20, 45–48]. Our solution attempts to develop a fully-automated learning system which can deal with a wide variety of biological objects and is customizable to different biological problems.

As a first step in solving the directionality detection problem, we analyzed several sentence structures. These sentence structures fall into two different classes of direction. The direction of a relation can be either in the forward direction, where the first object has a relation with the second object, or the backward direction, where the second object has a relation with the first object. Analysis of the different sentence structures can reveal the patterns for different rules for the two classes involved in directionality. These rules were captured and represented as a finite state automaton. For example, consider the sentence: “Isoproterenol/NN *inhibits*/VBZ 17beta-estradiol-induced/NNP gene/NN transcription/NNP in/IN A7r5/NNP cells/NNS ./.” This sentence is in simple active voice and the direction is forward in this case. The subject noun “Isoproterenol” (actor) has a forward inhibit relation with “17beta-estradiol-induced gene transcription,” the object noun phrase (patient). After analyzing various sentence structures, rules were developed and learned using a neural network.

We used a connectionist model to learn directionality for several reasons: It is one of the faster learning systems; it can capture the knowledge of direction easily; and most importantly, the neural network is more noise tolerant and has the ability to generalize to a new type of sentence which the neural network has never seen before. A feed forward back propagation neural network was used with 11 inputs, where each input corresponds to a parts-of-speech of a preceding or a succeeding phrase and two outputs representing the two classes. Our initial study on six different types of sentence structures with 600 input sentences gave 88% accuracy on directionality detection [49].

Results

BioMap excavates different immune effector and target cells and their direct relations associated with inflammation, EAE and MS

The BioMap interface shown in Figure 3 was used to generate various results given in this section.

The objective of this work is to carry out a comparative study on cell–cell associations that are common or unique for inflammation, EAE and MS that is reported in the literature data. Specifically, during the induction of MS, Microglial cells play a major role, however the mechanism of the etiology is not known. One of the key objectives of this paper is to identify cellular interaction of microglial cells with other cell components and to analyze the unique associations found in MS with microglial cells and other major significant cellular components like macrophage and monocytes. The association discovery process involves selecting “Cell” followed by “Inflammation and EAE”, or “Inflammation and MS”, or “Inflammation only” for the three different studies reported in the paper using the BioMap system. For example, if the query is, “Cell $\cap$ Inflammation $\cap$ EAE” ($\cap$ is the “AND” operator), the system would first generate a list of all the cells that are associated with inflammation and EAE from the knowledgebase. This list is now fed to the rFinder and TransMiner tools for finding direct, directional and transitive associations. The results shown in Tables 1–7 were summarized based on the outcomes from the rFinder and TransMiner tools. In addition to the textual results of the three different types of associations, these tools also generate association graphs as shown in Figures 4–7. For the sake of brevity, since these association graphs are very large, Figures 4–7 show only partial graphs to highlight the important nodes. In Figures 4–7, the arrow indicates the direction of the reaction and the number associated with the reaction is the frequency of occurrence. The gray shaded arrows indicate “does not” reactions or negations.

Based on the knowledge discovery process described above, we identified 33 cell types as shown in Table 1 after the first query session. Among them are 33 cells that are associated with inflammation, 26 cells associated with EAE and 30 cells involved in MS. The unique cells found in EAE and inflammation were helper T cells and phagocytes, whereas, the unique ones in inflammation and MS were basophils, CD4+ T cell, CD8+ T cell, macroglia, nerve cells and red blood cells. These cells can be broadly classified as immune effector cells and organ-specific target cells. Thus the major immune cells that we have identified include antigen presenting cells such as dendritic cells, macrophage and microglia, lym-

Table 1. Cells in inflammation, EAE and MS discovered by BioMap

Cells	MS	EAE	Inflammation
Astrocyte	+	+	+
Basophils	+	-	+
B cell	+	+	+
CD4+ T-cell	+	-	+
CD+ T-cell	+	-	+
Dendritic cells	+	+	+
Endothelial cells	+	+	+
Eosinophil	+	+	+
Glial cells	+	+	+
Granulocyte	+	+	+
Helper T-cell	-	+	+
Lymphocytes	+	+	+
Macroglia	+	-	+
Macrophage	+	+	+
Mast cell	+	+	+
Microglia	+	+	+
Monocyte	+	+	+
Motor neurons	+	+	+
Nerve cells	+	-	+
Neurons	+	+	+
Neutrophils	+	+	+
NK cells	+	+	+
Oligodendrocyte	+	+	+
Phagocytes	-	+	+
Platelet	+	+	+
Red blood cells	+	-	+
Schwann cells	+	+	+
Stem cell	+	+	+
Suppressor cells	+	+	+
T-cell	+	+	+
Th1 cells	+	+	+
Th2 cells	+	+	+
T lymphocytes	+	+	+

Table 2. Cell-cell transitive associations showing frequency strength for inflammation

Cells	Cells	Strength
Oligodendrocytes	Th2 cells	178
Neurons	Neutrophils	91
Nneurons	NK cells	86
Neurons	Th2 cells	70
Neurons	Th1 cells	57
Neutrophils	NK cells	50
Neutrophils	Th2 cells	50

Table 3. Cell-cell transitive associations with frequency strength for EAE

Cells	Cells	Strength
Neurons	Th1 cell	110
Oligodendrocytes	Th1 cell	109
Neurons	Natural killer cells	52

Table 4. Cell-cell transitive associations with frequency strength for MS

Cells	Cells	Strength
Monocyte	Neurons	279
Oligodendrocytes	Th2 cells	179
Neurons	Neutrophils	91
Neurons	Natural killer cells	86
Neurons	Th2 cells	70
Neurons	Th1 cell	57

phocytes such as CD4+ Th1 and Th2 cells, and CD8+ cytotoxic T cells, B cells, NK cells and granulocytes. We have also detected various target cells such as oligodendrocytes, Schwann cells, neurons, and astrocytes in these diseases. Although there is no significant difference in immune cells associated with different inflammatory diseases, the target cells can vary depending on the tissues and organs involved in the various inflammatory disease. However, these differences are not functionally significant, suggesting that the inflammatory cells involved in EAE and MS are comparable to any other inflammatory diseases. These findings also suggest that the information excavated from biomedical literature through TransMiner is biologically relevant and is reliable for such discoveries in other fields of medicine.

Microglia is the main resident antigen presenting cell in the brain and CD4+ Th1 cells are the important effector cells in EAE/MS [50–56]. We have examined the association of microglial cells with other immune effector and target cells in EAE, MS and inflammation. In MS, we found a direct association of monocytes, NK cells, neurons, neutrophils, oligodendrocytes, phagocytes, schwann cells, stem cells, T cells, T lymphocytes, Th1 cells and Th2 cells with microglial cells. These direct associations that we have identified are significant and can be used in future laboratory experiments to diagnose different stages of dis-

Table 5. Major hubs in direct associations for inflammation, EAE and MS

Inflammation	No. of associations	EAE	No. of associations	MS	No. of associations
Macrophage	1,117	Macrophage	54	Astrocyte	92
Monocyte	1,046	Astrocyte	53	Macrophage	89
Eosinophils	970	Lymphocytes	46	Lymphocyte	87
Neutrophil	913	Microglia	46	Microglia	81
Endothelial cell	634	T-cell	38	Oligodendrocyte	74
Dendritic cell	609	Oligodendrocyte	31	Monocyte	67
Lymphocyte	599	Glial cells	27	T-cell	66
Granulocyte	531	Neuron	23	Neuron	46
Platelet	335	Monocyte	20	Glial cell	42
Basophil	328	Endothelial cells	18	Endothelial cell	27
Astrocyte	299	Dendritic cell	16	Dendritic cell	26
Mast cell	259	Th1 cell	15	Neutrophil	23
Microglia	196	B-cell	13	Th2 cell	22
NK cell	156	Th2 cells	13	Th1 cell	21
Erythrocyte	144	T-lymphocyte	10	NK cell	20
B-cell	124			Mast cell	19
				T-lymphocyte	18
				Granulocyte	15
				B-cell	14
				Schwann cells	14

Table 6 Major hubs in transitive associations for inflammation, EAE and MS

Inflammation	No. of associations	EAE	No. of associations	MS	No. of associations
Dendritic cell	16,014	Macrophage	162	Astrocyte	404
Basophil	15,733	Monocyte	129	Macrophage	319
Endothelial cell	14,442	Astrocytes	95	Monocyte	287
B-cell	13,799	B-cell	81	Endothelial cell	259
Macrophage	10,034	Astrocyte	78	Dendritic cell	249
Monocyte	10,031	Dendritic cells	75	Oligodendrocyte	239
Granulocyte	9,960	Endothelial cells	68	Glial cell	195
Lymphocyte	9,565	Oligodendrocyte	62	B-cell	166
Neutrophil	9,530	Glial cells	58	Lymphocyte	156
Eosinophil	9,094	Neuron	57	Neuron	136
Platelet	8,356	Microglia	51	Microglial cell	114
Mast cell	8,202				
Astrocytes	7,945				

ease pathogenesis and to test the effects of potential therapeutic agents in the treatment of inflammatory diseases.

In EAE, the microglial cells showed direct association with monocytes, NK cells, neurons, oligodendrocytes, T cell, Th1 cells, Th2 cells and T lymphocytes. Whereas, with inflammation, the eosinophils, glial cells, granulocytes, astrocytes,

basophils, dendritic cell, endothelial cells, macroglia, and macrophage showed direct association with microglial cells. A comparative analysis of associations in EAE and inflammation with MS indicated that microglial cells exhibit direct association with dendritic, endothelial, eosinophils, glial cells, granulocyte, macrophage and mast cells in EAE and inflammation.

Table 7. Major hubs in directional (outgoing) associations for inflammation, EAE and MS

Inflammation	No. of associations	EAE	No. of associations	MS	No. of associations
Monocyte	48	Monocyte	31	Neutrophil	36
Macrophages	45	Macrophages	29	Monocyte	34
Neutrophil	41	Neutrophils	26	Macrophages	32
Eosinophils	32	Eosinophils	23	Eosinophil	22
Endothelial cells	24	Endothelial cells	18	Endothelial cells	19
Astrocytes	18	Astrocytes	15	Astrocytes	15
Lymphocyte	17	Microglia	14	Lymphocyte	15

Direct associations among cells that are in Inflammation but not in MS and EAE are neutrophils, basophils, NK cells, platelets, and mast cells. Neutrophils are also known as polymorphonuclear leukocytes (PMN), representing 50–60% of the total circulating leukocytes and constitute the “first line of defence” against infectious agents or “nonself” substances that penetrate the body’s physical barriers. Once an inflammatory response is initiated, neutrophils are the first cells to be recruited to site of infection or injury. Their development in the bone marrow takes about 2 weeks; during this period, they undergo proliferation and differentiation. Basophils are the least common of the granulocytes, representing about 0.5–1% of circulating leukocytes. Like all circulating granulocytes, basophils

can be recruited out of the blood into a tissue when needed. When activated, basophils secrete histamine, several proteoglycans, lipid mediators like leukotrienes, and several cytokines. Histamine and proteoglycans are pre-stored in the cell’s granules while the other secreted substances are newly generated. Each of these substances contributes to inflammation. Recent evidence suggests that basophils are an important source of the cytokine, interleukin-4, perhaps more important than T cells. Interleukin-4 is considered one of the critical cytokines in the development of allergies and the production of IgE antibody by the immune system. There are other substances that can activate basophils to secrete which suggests that these cells have other roles in inflammation. The NK cells constitute a major component of the innate

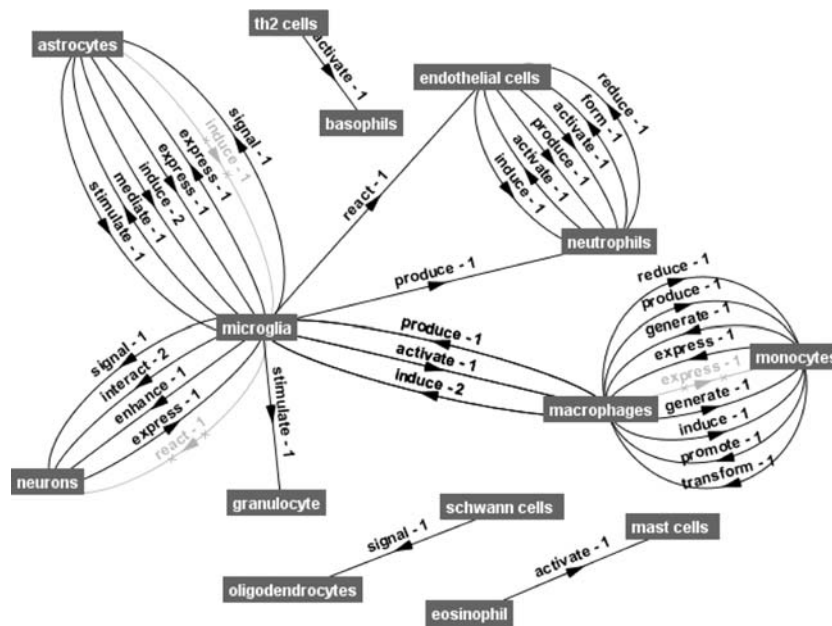


Figure 4. Partial snap shot of the directional relationship graph associated with multiple sclerosis.

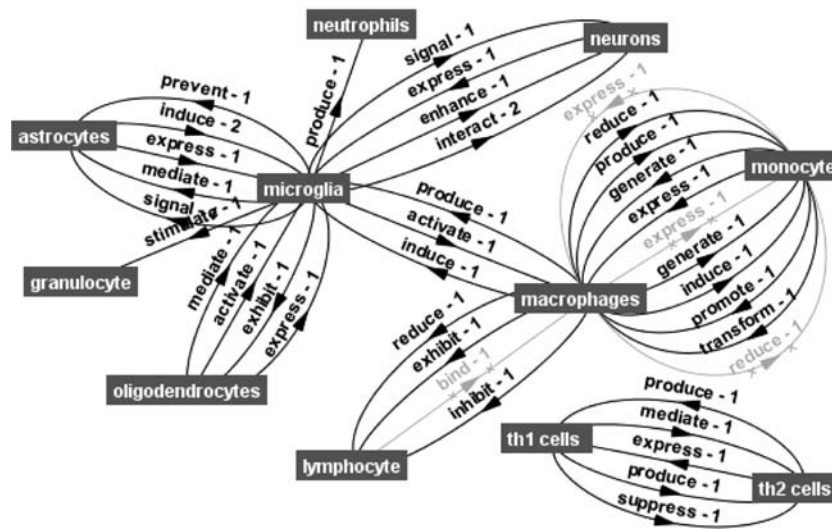


Figure 5. Partial snap shot of the directional relationship graph associated with experimental allergic encephalomyelitis.

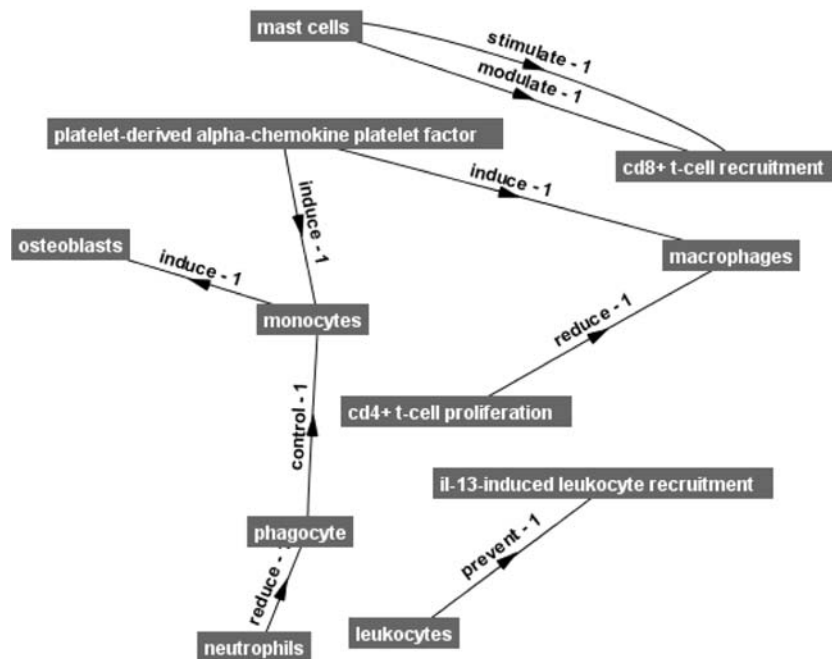


Figure 6. Partial snap shot of the directional relationship graph associated with inflammation.

immune system, which usually express the surface markers CD16 (FcγRIII) and CD56 in humans. NK cells are cytotoxic; small granules in their cytoplasm contain special proteins such as perforin and proteases, known as granzymes. Upon release

in close proximity of a cell to be killed, perforin forms pores in the cell membrane of the target cell through which the granzymes and associated molecules can enter, where they induce apoptosis. NK cells are activated in response to interferons or

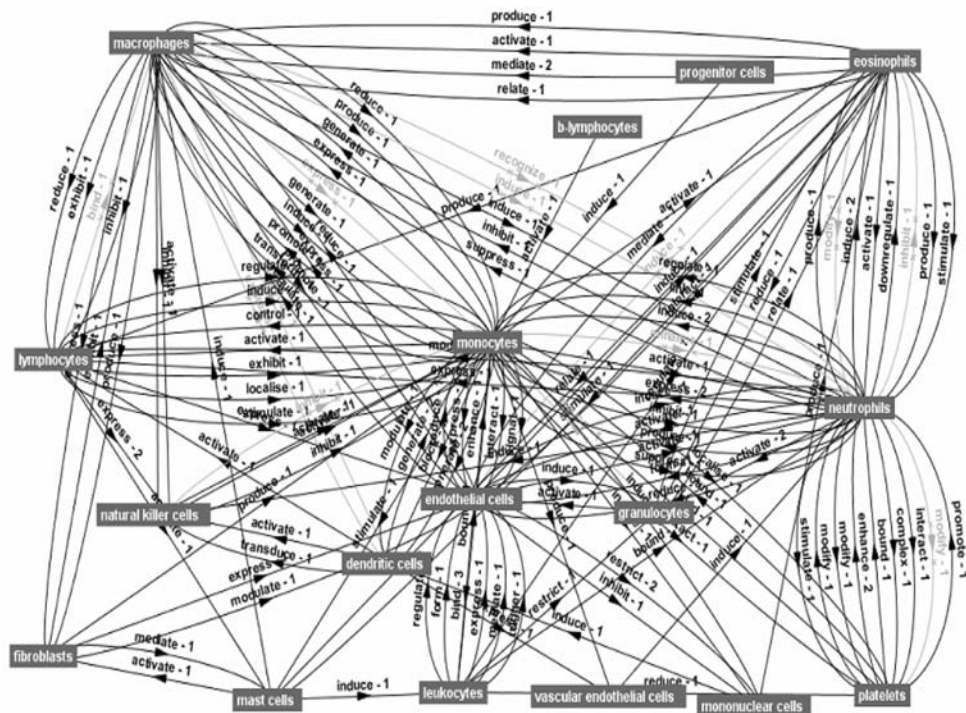


Figure 7. Partial snap shot of the directional relationship graph which shows major hub nodes associated with EAE.

macrophage-derived cytokines. Activated platelets secrete chemotactic substances, they facilitate the binding of leukocytes to the endothelium and their subsequent extravasation, and they may influence the inflammatory responses of leukocytes in both stimulating and inhibiting ways. However, platelets themselves also contain an array of potent proinflammatory substances, and therefore they are regarded as mediator and effector cells in inflammation. The mast cells have long been recognized as potent producers of a large panel of biologically highly active mediators such as biogenic amines, arachidonic acid metabolites, cytokines and chemokines, but most of their biological functions have been elusive and speculative.

Discovering directional association between different immune effector and target cells in inflammation, EAE and MS

The rFinder tool in BioMap was used to discover directional relationships that exist for cellular interactions associated with inflammation, EAE and MS. The unique associations among cells in

inflammation with MS are: eosinophils, glial cells, granulocytes, macrophage, mast cells, nerve cells, phagocytes, platelets and T cells. The unique ones for EAE and MS are eosinophils, granulocytes, macrophage, mast cells, nerve cells, phagocytes, platelets, and T cells. Using rFinder, we have generated the directionality among 28 cell types associated with these inflammatory diseases, of which 24 cells appeared with MS and 22 for EAE.

Directional associations found by rFinder for both MS and EAE showed that macrophages induce/produce monocytes and microglial cells and oligodendrocytes activate microglial cells. It is also interesting to observe that microglia cells interact with neurons through signaling and enhancing. A number of associations were found which are common for both MS and EAE, such as microglia activate macrophage; microglia produces neutrophils and stimulates granulocytes. Those associations also found in MS but not in EAE are: endothelial cells produce neutrophils; eosinophils activate mast cells; macrophage express astrocytes; microglia express astrocytes; schwann cells signal oligodendrocytes; and Th2

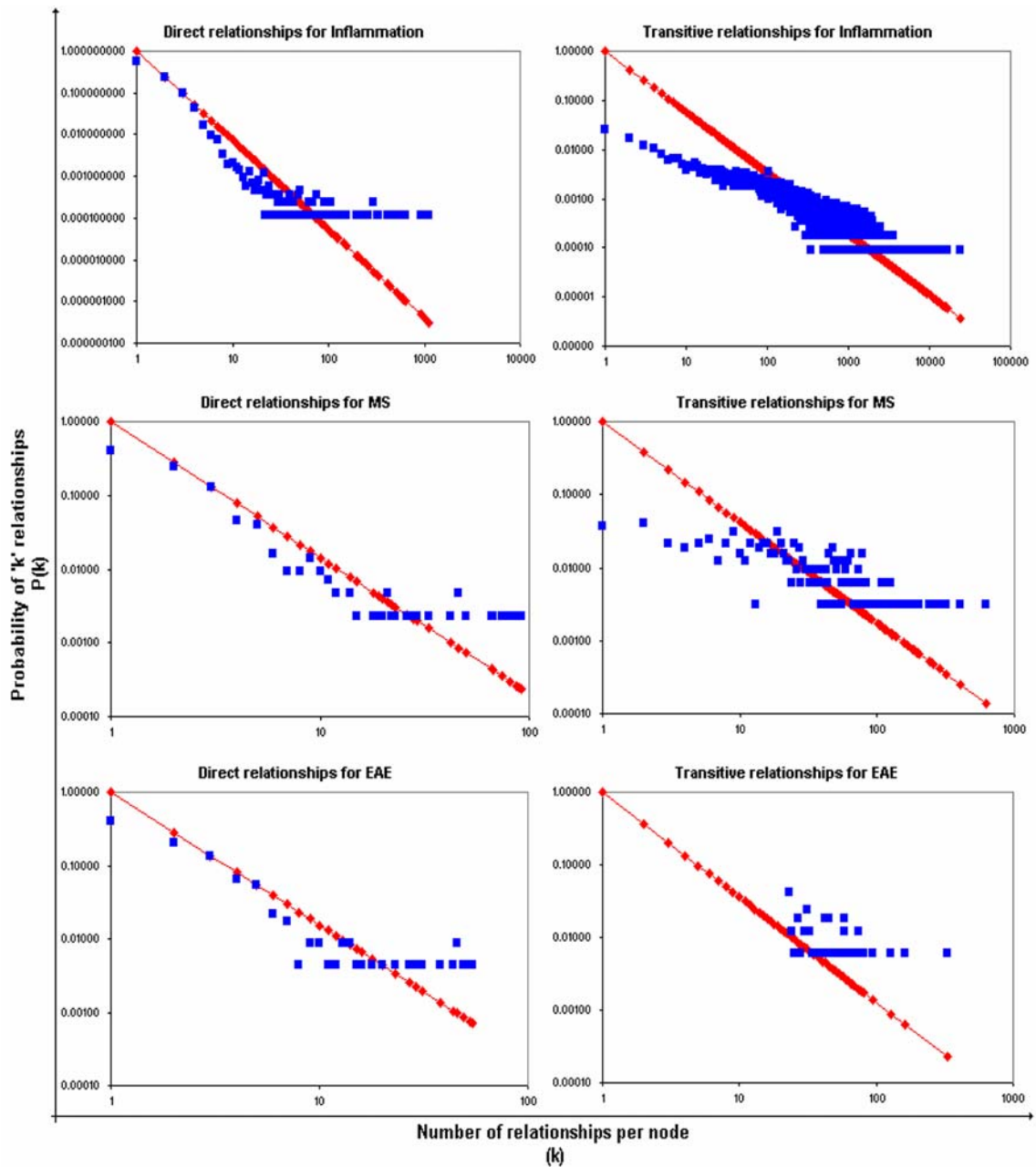


Figure 8. Scale-free topology of Direct Relationship network (left column) and Transitive Relationship network (right column) for Inflammation, EAE and MS.

cells activate basophils. A partial graph depicting the directional associations found for MS is shown in Figure 4.

In disease induction, Th1 cells play a major role; they induce microglia thereby affecting oligodendrocytes and cause demyelination in CNS which results in MS. One of the associations

observed in EAE is that microglial cells prevent astrocytes. However, such a relation could not be observed in MS. During the induction of disease like MS, Th2 cells have pro-inhibitory effects. It has already been established that Th2 cells inhibit microglia as well as Th1 cells which ultimately induce demyelination. Some of the directional

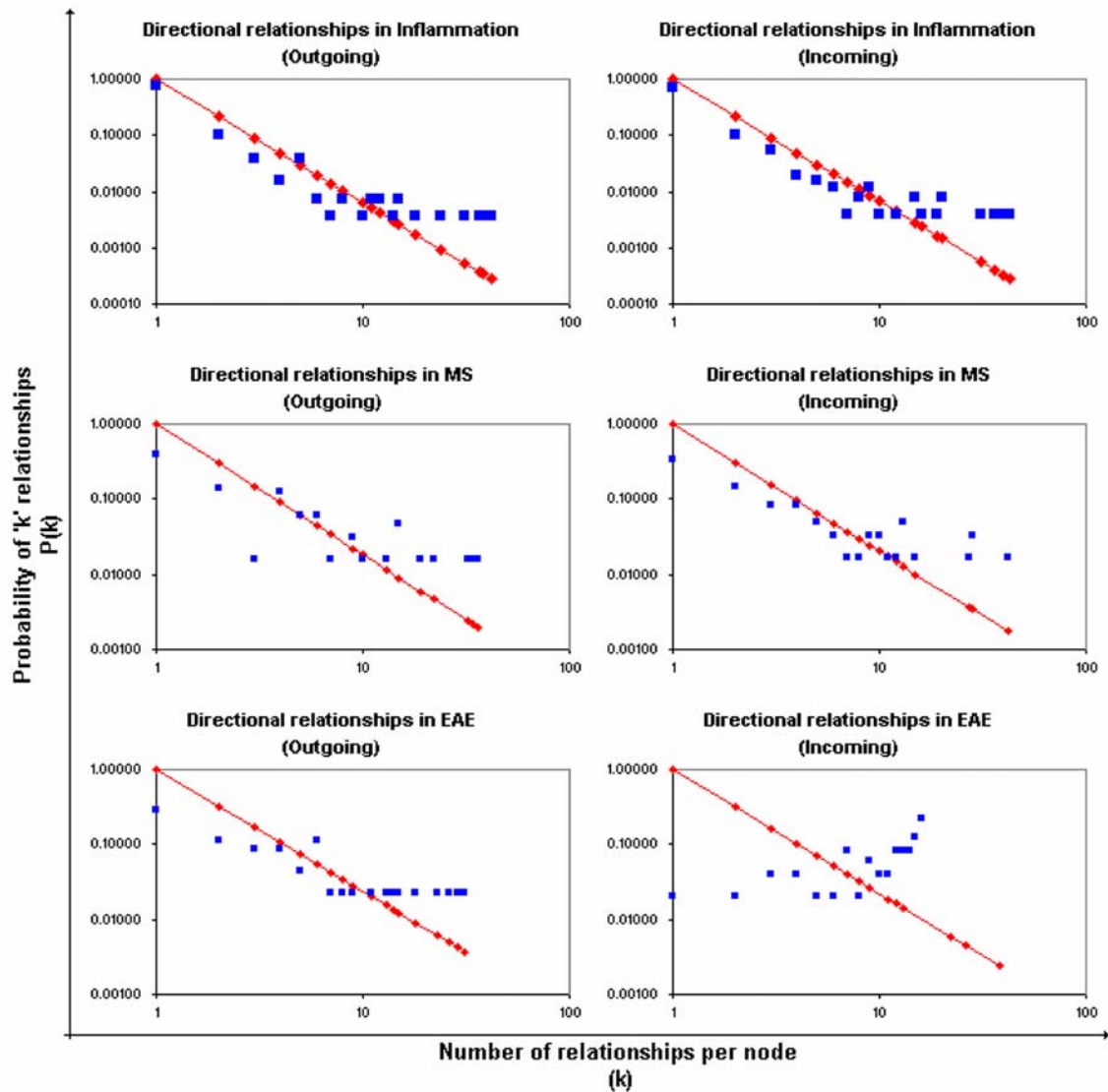


Figure 9. Scale-free topology of Directional Relationship for outgoing and incoming networks for Inflammation, EAE and MS.

associations that were not found in MS but in EAE (see Figure 5) are: macrophage reduce lymphocytes and Th1 cells mediate Th2 cells.

A few unique relations were also found for MS that are not present in inflammation. Some of them include: macrophages and microglia express astrocytes; microglia react with endothelial cells; and, schwann cells signal oligodendrocytes. However the unique associations found in inflammation are extensive, as depicted in the partial graph shown in Figure 6. Some of the significant ones are: CD4⁺ T cell proliferation

reduces macrophages; leukocytes prevent IL-13 induced leukocyte recruitment; mast cell stimulate CD8⁺ cell recruitment; monocytes activate phagocyte differentiation; monocytes induce macrophage differentiation; neutrophils reduce phagocytes; phagocytes control monocytes; and, platelet derived alpha chemokine factor induces macrophages and monocytes.

Microglial cells have a major role in pathogenesis of multiple sclerosis. The comparative analysis of the results obtained through the rFinder tool revealed that, for MS, unique and relevant asso-

ciations with microglial cells are very few. However, the unique associations found in EAE and inflammation but not in MS, such as microglial cells prevent astrocytes; microphages reduce lymphocytes; Th1 cells mediate Th2cells etc., could be significant in MS research.

The major association nodes found through rFinder for Inflammation, EAE, and MS are macrophages, monocytes, eosinophils, astrocytes, lymphocytes, and microglia. In inflammation, macrophages have three major function; antigen presentation, phagocytosis, and immunomodulation through production of various cytokines and growth factors. Macrophages play a critical role in the initiation, maintenance, and resolution of inflammation. They are activated and deactivated in the inflammatory process. The majority of the resident macrophages are derived from circulating monocytes. During an acute inflammation the number of circulating monocytes increases, and a large proportion of these cells migrate to the site of inflammation and differentiate into exudate macrophages. Eosinophils produce cytokines such as interleukin IL-3, IL-5 and granulocyte-macrophage colony stimulating factor (GM-CSF) that contribute to their proinflammatory functions. Astrocytes are sub-type of the glial cells in the brain. Astrocytes are star shaped glial cells that perform a variety of functions in the CNS. Astrocytes provide physical support to neurons and clean up debris within the brain. They also provide neurons with some of the chemicals needed for proper functioning and help control the chemical composition of the fluid surrounding neurons. Finally, astrocytes play a role in providing nourishment to neurons. Lymphocytes is a type of white blood cell involved in the human body's immune system. There are two broad categories of lymphocytes, namely T cells and B cells. Lymphocytes play an important and integral part of the body's defenses. T cells are chiefly responsible for cell-mediated immunity whereas B cells are primarily responsible for humoral immunity (relating to antibodies). Microglia are the smallest of the glial cells. They act as phagocytes cleaning up CNS debris. Microglia are thought to be highly mobile cells that play numerous important roles in protecting the nervous system. They are also thought to have an impact in neurodegenerative disorders like MS. Microglia is responsible for producing an inflammatory reaction to insults.

BioMap detects novel association between immune and target cells in inflammation, EAE and MS

One of the key features of BioMap system is its capability to discover novel associations among biological entities based on the transitivity property using the TransMiner tool. Such transitive associations may be indicators of new hypotheses that be derived based on literature knowledge. In this study, we selected transitive associations that has a strength (frequency of occurrence) greater than 50, for Inflammation, EAE and MS. Table 2 shows cell to cell interaction in inflammation with high strength values excavated through TransMiner. Neurons exhibited high association with neutrophils, NK cells, Th1 and Th2 cells, whereas oligodendrocytes showed highest association with Th2 cells. Table 3 shows the transitive associations with high strength values in EAE. For EAE, neurons with Th1 cells expressed the highest associations, followed by oligodendrocytes with Th1 cells and neurons with NK cells. Cell to cell transitive association with their strength values in MS are given in Table 4. For MS, all the cell components of higher strength values are different from Inflammation and EAE. Monocytes indicated the highest association with neurons followed by oligodendrocytes and Th2 cells.

Considering all cellular participants and giving emphasis to their frequency values in Inflammation, EAE, and MS, TransMiner generated similar cell types with high associations in Inflammation and EAE. However, these were not excavated in MS. Also, transitive associations generated by TransMiner are based on the contextual occurrence of these cells with a high frequency of co-occurrence. The associations that are generated by TransMiner do not give the directionality or the nature of the biological reaction and this could be an inhibiting effect on other cellular participants which ultimately induce or the etiology of the diseases like MS. Thus, those associations with high strength values which are in EAE and inflammation but not in MS can be taken for experimental studies in MS since EAE is an animal model of MS.

Biological entity association network analysis

In addition to the qualitative analysis of the associations which were discovered using BioMap,

we also carried out a statistical analysis of the hub nodes that are present in the association graphs. Numerous studies in the literature indicate that most biological networks are scale-free [57–60]. In a scale-free network, the degree distribution, $P(k)$, shows the probability that a selected node has exactly k links. In a graph with directional links, k_{in} is the in-degree and k_{out} is the out-degree of the node. The degree distribution $P(k)$ approximates the power law, $P(k) \propto k^{-\gamma}$, where γ is the degree exponent and $\propto$ means ‘proportional to’. Various researches have shown that for $2 > \gamma > 3$ strongly indicates that there exists a hierarchy of hubs in biological networks. The hierarchy has the inherent property that a small number hubs are connected with a large number of nodes whereas, a large number of hubs has very few number of nodes [61]. Figure 7 shows a larger network for the directional association for EAE to demonstrate the hub phenomena. Some of the nodes such as “monocytes” and “neutrophils” show many incoming and/or outgoing links, whereas, the “mast cells” has fewer links. There are several more such smaller nodes that were found for EAE that are not shown in this graph for the sake of brevity. This property that has been observed in biological association networks also provides biological meaning to the relationships [62]. We have analyzed the power-law degree distribution of the association graphs obtained for direct, directional and transitive relationships and compared that for Inflammation, EAE, and MS.

Table 5 shows the major hub nodes that are present in Inflammation, MS and EAE, indicating the number of links associated with each node for the direct association network. The power-law degree distribution plots in Figure 8 (left column) show the behavior of the hub nodes that are present in the direct relationship graph. The γ values were 1.29, 1.85, and 1.80 for inflammation, MS and EAE respectively. Table 6 gives the major hub nodes for the transitive association network and Figure 8 (right column) shows the scale-free network for all three diseases. The γ values were 1.24, 1.32, and 1.35 for inflammation, MS and EAE respectively. Figure 9 shows the power-law degree distribution of the in-degree and out-degree for the directional relationship graphs and the respective major hub nodes are given in Table 7. The γ values for both in-degree and out-degree

ranged from 2.16, 1.74, and 1.63 for inflammation, MS and EAE respectively.

Even though some of the γ values are lower than the typical range of $2 > \gamma > 3$ found in scale-free networks, these numbers are close enough to demonstrate the biological significance of the hub nodes. The lower values are due to smaller number of links associated with the hub nodes since we are studying a small biological problem domain. The scale-free property also exist for the transitive association network as these values follow similar ranges found in biological networks [62].

Discussion

The BioMap system consists of several toolsets for biological data mining and integration. In this paper we demonstrated two such tools, the TransMiner and rFinder, can be used to find associations among biological entities. We used BioMap to discover and find associations involved in inflammation, EAE and MS. From the 32 different cells that were found, some were unique for EAE and MS. rFinder provided the directionality of association among cells in inflammation, EAE and MS. Unique similarities and differences observed in EAE or inflammation but not in MS can be taken to the laboratory for further experimental studies to test their relevance in MS. The cell types in these disease systems are rather smaller than those of TransMiner and indicate the different aspects of biological reaction such as directionality and mode of biological associations that researchers look for, such as action between two genes or two proteins or even a gene protein, etc. Transitive relationships established through TransMiner are based on contextual occurrence. The highest strength value indicates the highest frequency of that relationship’s occurrence in the database. Associations not found explicitly in the database are transitive associations and are candidates for hypothesis generation. In effect, TransMiner is able to generate all the associations among the objects and the investigator can then choose the most interesting objects for experimentation.

Another interesting capability of the BioMap tool is its capacity to take established relationships in well-studied disease systems, such as an animal model, and apply these results to less-studied human diseases, such as MS. Taking EAE as a

model of MS, we analyzed the components of all these three diseases using BioMap to observe the cell components and the established relationships among them. For example, cells in inflammation and EAE provide insight to the workings of MS. Through rFinder, the results can be described as:

Monocyte → *induce* → Microglia → *stimulate*
 → Granulocyte → *stimulate*
 → Dendritic cells → *activate* → Inflammation.

Information on all these kinds of associations can ultimately facilitate the organized design of future experiments.

Critical analysis in all these disease domains showed that some associations are unique in both EAE and inflammation and rFinder did not provide any of these associations in MS. Examples of such associations are: microglial cells prevent astrocytes; macrophages reduce lymphocytes; neutrophil reduce phagocytes; phagocytes control monocytes. This type of association excavated through BioMap would be beneficial in providing information to researchers. This is a preliminary study of cell components in MS and its animal model EAE and this could be further extended in other major components such as cytokines, chemokines, transcription factors, kinases, etc. Once all such components involved in the inflammatory disease can be elucidated through the data mining process, one can obtain the different pathways expressed in animal models and take them into the laboratory to identify/validate the real molecular mechanism behind MS.

TransMiner provided most cell types involved in MS and its animal model, EAE. The associations excavated through TransMiner with high strength value can be taken as criteria on which to set up laboratory experiments in MS and EAE. It is also important to note that the researcher needs to validate the associations with low strength values to verify if there would be any significance in the studies before concentrating on associations with high strength values, as some of the relevant ones might be appear only a few times in data sources. The BioMap tool could also be beneficial for young researchers, scientists from various disciplines who deal with inflammatory diseases, students who are working in this area and all patients and individuals interested in understanding the various facets and dimensions of inflammatory diseases like EAE

and MS. We could excavate cell to cell associations established in inflammation and EAE but not in MS that may open up new doors to researchers who are working on Multiple Sclerosis. Thus, this data mining tool could be useful for discovering relationships among biological objects like cells, genes, proteins, molecules etc. from the enormously rich and fabulous quantity of data that human beings cannot examine manually.

The ability of BioMap to directly excavate knowledge from biomedical literature is an important feature, allowing investigators and health professionals to investigate and analyze biomedical literature in-depth. The unique similarities and differences observed through this analysis can be considered in the development of new hypotheses and used to design future laboratory experiments. In this study we have performed a complex analysis of cells associated with inflammation, EAE and MS to validate the use of this innovative tool in biomedical research. Our plan is to further develop the BioMap system to incorporate complex analysis of systems that involve signaling and transcription, which will identify unique molecular and therapeutic targets in various human diseases.

BioMap system is accessible online at: <http://regen.cs.iupui.edu/biomap/>

Acknowledgements

The research reported in this paper was supported in part by National Science Foundation Information Technology Research grant number NSF-IIS/ITR 0081944, NIH BISTI grant, NIH-NIGMS P20 GM66402, and Indiana 21st Century Research Fund.

References

1. <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed>.
2. <http://www.rnabase.org/>.
3. <http://www.ncbi.nlm.nih.gov/Genbank/index.html>.
4. <http://www.pir.uniprot.org/>.
5. <http://www.ebi.ac.uk/interpro/>.
6. <http://www.geneontology.org/>.
7. <http://www.genome.jp/kegg/>.
8. <http://ecocyc.org/>.
9. <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM>.
10. Do H.H., Rahm E., Flexible integration of molecular-biological annotation data: The GenMapper approach, Proc. EDBT 2004, Heraklion, Greece, Springer LNCS, March 2004.

11. Chen L. and Jamil H.M., On using remote user defined functions as wrappers for biological database interoperability. *Int. J. Coop. Inform. Syst. (IJCIS)*, Special Issue on Data Management and Modeling Support in Bioinformatics 12(2) 161–195, 2003.
12. <http://www.ingenuity.com>.
13. <http://pharmgkb.org>.
14. <http://imgproj.cs.man.ac.uk/tambis/details.html>.
15. <http://www.incyte.com/sequence/proteome/index.shtml>.
16. <http://www.pubgene.com>.
17. <http://rosetta.lcb.uu.se>.
18. <http://www.genostar.org/english/index.html>.
19. <http://www.ariadnegenomics.com>.
20. Novichkova S., Egorov S. and Daraselia N., Medscan, a natural language engine for MEDLINE abstracts. *Bioinformatics* 19: 1699–1706, 2003.
21. Daraselia N., Yuryev A., Egorov S., Novichkova S., Nikitin A. and Mazo I., Extracting human protein interactions from MEDLINE using a full-sentence parser. *Bioinformatics* 19: 1–8, 2003.
22. Narayanasamy V., Mukhopadhyay S., Palakal M. and Potter D., TransMiner: Mining transitive associations among biological objects from text. *J. Biomed. Sci.* 11(6) 864–873, 2004.
23. Stephens M., Palakal M., Mukhopadhyay S., Raje R. and Mostafa J., Detecting gene relations from medline abstracts. *PSB (Pacific Symposium on Biocomputing)* 83–495, 2001.
24. Palakal M., Stephens M., Mukhopadhyay S. and Raje R., A Multi-Level Text Mining Method to Extract Biological Relationships. *IEEE CSB*, Stanford, 2002.
25. Palakal M., Stephens M., Mukhopadhyay S., Raje R. and Rhodes S., Identification of biological relationships from text documents using efficient computational methods. *J. Bioinformatics Comput. Biol.* 1(2) 1–34, 2003.
26. Palakal M., Mukhopadhyay S. and Stephens M., Identification of biological relationships from text documents. Book Chapter in *Medical Informatics: Advances in Knowledge Management and Data Mining in Biomedicine*, Ed. H. Chen. Kluwer Publishers, 449–489, 2005.
27. Jayadevaprakash N., Mukhopadhyay S. and Palakal M., An approach to generating association graphs of non-cooccurring text objects using transitive methods. 20th ACM Symposium on Applied Computing-Bioinformatics Track, (ACM SAC), 2005.
28. Dean G., How many people in the world have MS? *Neuroepidemiology* 13: 1–7, 1994.
29. Hemmer B., Archelos J.J. and Hartung H., New concepts in the immunopathogenesis of multiple sclerosis. *Nat. Rev. Neurosci.* 3: 204–304, 2002.
30. Raine C.S., Multiple sclerosis: Immunopathological mechanisms in the progression and resolution of inflammatory demyelination. *Res. Publ. Assoc. Res. Nerv. Ment. Dis.* 68: 37–54, 1990.
31. Prineas J.W. and Raine C.S., Electron microscopy and immunoperoxidase study of early MS lesions. *Neurology* 26: 29–32, 1976.
32. Owens T. and Sriram S., The immunology of MS and of its animal model, experimental allergic encephalomyelitis. *Neurol. Clin. N.A.* 13: 51–73, 1995.
33. Bright J.J., Musuro B.F., Du C. and Sriram S., Expression of IL-12 in CNS and lymphoid organs of mice with experimental allergic encephalomyelitis. *J. Neuroimmunol.* 82: 22–26, 1998.
34. Bright J.J., Du C., Coon M., Sriram S. and Klaus S.J., Prevention of experimental allergic encephalomyelitis via inhibition of IL-12 signaling and IL-12-mediated Th1 differentiation: An effect of the novel anti-inflammatory drug lisofylline. *J. Immunol.* 161: 7015–7022, 1998.
35. Bright J.J., Du C. and Sriram S., Tyrphostin B42 inhibits IL-12-induced tyrosine phosphorylation and activation of janus kinase-2 and prevents experimental allergic encephalomyelitis. *J. Immunol.* 162: 6255–6262, 1999.
36. Natarajan C. and Bright J.J., Curcumin inhibits experimental allergic encephalomyelitis by blocking IL-12 signaling through JAK-STAT pathway in T Cells and differentiation of neural antigen specific Th1 cells. *J. Immunol.* 69: 6506–6513, 2002.
37. Bright J.J., Rodriguez M. and Sriram S., Differential influence of interleukin-12 in the pathogenesis of autoimmune and virus-induced CNS demyelination. *J. Virol.* 3: 1637, 1999.
38. Natarajan C. and Bright J.J., Peroxisome proliferator-activated receptor-gamma agonists inhibit experimental allergic encephalomyelitis by blocking IL-12 production, IL-12 signaling and Th1 differentiation. *Gene. Immunity.* 3: 59–70, 2002.
39. Bright J.J., Targeting autoimmune diseases through nutraceuticals. *Nutrition* 20: 39–43, 2004.
40. Palakal M., Mukhopadhyay S., Mostafa J., Raje R., N'Cho M. and Mishra S.K., An Intelligent Biological Information Management System, 2002 ACM Symposium on Applied Computing, Madrid, Spain, 2002.
41. Kumar K., Palakal M. and Mukhopadhyay S., BioMap: Toward the development of a knowledge base of biomedical literature, 2004 ACM Symposium on Applied Computing, Nicosia, Cyprus, 2004.
42. Salton G., *Automated Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer*. Addison-Wesley, Reading, MA, 1989.
43. Smalheiser N.R., Predicting emerging technologies with the aid of text-based data mining: A micro approach. *Technovation* 21: 689–693, 2001.
44. Warshall S.A., Theorem on Boolean matrices. *JACM* 9(1) 11–12, 1962.
45. Karopka T., Scheel T., Bansemer S. and Glass A., Automatic construction of gene relation networks using text mining and gene expression data, medical informatics and the internet in medicine. 2(29): 169–183, 2004.
46. Rzhetsky A., Iossifov I., Koike T., Krauthammer M., Kra P., Morris M., Yu H., Duboue P.A., Weng W., Wilbur W.J., Hatzivassiloglou V. and Friedman C., GeneWays: A system for extracting, analyzing, visualizing, and integrating molecular pathway data. *J. Biomed. Info.* 37(1) 43–53, 2004.
47. Fiedman C., Kra P., Yu H., Krauthamrner M. and Rzhetsky A., Genies: A natural-language processing system for the extraction of molecular pathways. *Bioinformatics* 17(Suppl. 1) S74–S82, 2001.
48. Pustejovsky J., Castano J., Zhang J., Kotecki M. and Cochran B., Robust relational parsing over biomedical literature: Extracting inhibit relations. *Pacific Symp. Biocomput.* 7: 362–373, 2002.
49. Thomas M., Identification of directional relationships between biological objects from biomedical literature, MS Thesis, Indiana University-Purdue University Indianapolis, Indiana, 2005.
50. Bright J.J., Natarajan C., Sriram S. and Muthian G., Signaling through JAK2-STAT5 pathway is essential for IL-3 induced activation of microglia. *GLIA* 45: 188–196, 2004.

51. Woodroffe M.N., Bellamy A.S., Feldmann M., Davison A.N. and Cuzner M.L., Immunocytochemical characterisation of the immune reaction in the central nervous system in multiple sclerosis. Possible role for microglia in lesion growth. *J. Neurol. Sci.* 74(2–3) 135–152, 1986.
52. Banati R.B., Gehrmann J., Schubert P. and Kreutzberg G.W., Cytotoxicity of microglia. *Glia* 7: 111–118, 1993.
53. Ulvestad E., Williams K., Mork S., Antel J. and Nyland H., Phenotypic differences between human monocytes/macrophages and microglial cells studied in situ and in vitro. *J. Neuropathol. Exp. Neurol.* 53: 492–501, 1994.
54. Bradl M. and Linington C., Animal model of demyelination. *Brain Pathol.* 6: 303–311, 1996.
55. Brosnan C.F. and Raine C.S., Mechanisms of immune injury in multiple sclerosis. *Brain Pathol.* 6: 243–257, 1996.
56. Sriram S. and Rodriguez M., Indictment of microglia as the villain in MS. *Neurology* 48: 464–470, 1997.
57. Gonzalez-Scarano F. and Baltuch G., Microglia as mediators of inflammatory and degenerative diseases. *Annu. Rev. Neurosci.* 22: 219–240, 1999.
58. Barabasi A.L. and Oltvai Z.N., Network Biology: Understanding the cell's functional organization. *Nat. Rev.: Genet.* 5: 101–113, 2004.
59. Guelzim N., Bottani S., Bourguin P. and Kepes F., Topological and causal structure of the yeast transcriptional regulatory network. *Nat. Genet.* 31: 60–63, 2002.
60. Arita M., Scale-freeness and biological networks. *J. Biochem.* 138(1) 1–4, 2005.
61. Calvano S.E., Xiao W., Richards D.R., Felciano R.M., Baker H.V., Cho R.J., Chen R.O., Brownstein B.H., Cobb J.P., Tschoeke S.K., Miller-Graziano C., Moldawer L.L., Mindrinos M.N., Davis R.W., Tompkins R.G. and Lowry S.F., A network-based analysis of systemic inflammation in humans. Large scale collaborative research program. *Nature* 437(13) 1032–1037, 2005.
62. Chen H. and Sharp B.M., Content-rich biological network constructed by mining PubMed abstracts. *BMC Bioinformatics* 5: 147, 2004.