Network-based analysis of differentially expressed genes in cerebrospinal fluid (CSF) and blood reveals new candidate genes for multiple sclerosis

Nahid Safari-Alighiarloo[1], Mostafa Rezaei-Tavirani[1], Mohammad Taghizadeh[2], Seyyed Mohammad Tabatabaei[3] and Saeed Namaki[4]

[1] Proteomics Research Center, Department of Basic Science, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[2] Bioinformatics Department, Institute of Biochemistry and Biophysics, Tehran University, Tehran, Iran
[3] Medical Informatics Department, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[4] Immunology Department, Faculty of Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

ABSTRACT

Background. The involvement of multiple genes and missing heritability, which are dominant in complex diseases such as multiple sclerosis (MS), entail using network biology to better elucidate their molecular basis and genetic factors. We therefore aimed to integrate interactome (protein–protein interaction (PPI)) and transcriptomes data to construct and analyze PPI networks for MS disease.

Methods. Gene expression profiles in paired cerebrospinal fluid (CSF) and peripheral blood mononuclear cells (PBMCs) samples from MS patients, sampled in relapse or remission and controls, were analyzed. Differentially expressed genes which determined only in CSF (MS *vs.* control) and PBMCs (relapse *vs.* remission) separately integrated with PPI data to construct the Query-Query PPI (QQPPI) networks. The networks were further analyzed to investigate more central genes, functional modules and complexes involved in MS progression.

Results. The networks were analyzed and high centrality genes were identified. Exploration of functional modules and complexes showed that the majority of high centrality genes incorporated in biological pathways driving MS pathogenesis. Proteasome and spliceosome were also noticeable in enriched pathways in PBMCs (relapse *vs.* remission) which were identified by both modularity and clique analyses. Finally, STK4, RB1, CDKN1A, CDK1, RAC1, EZH2, SDCBP genes in CSF (MS *vs.* control) and CDC37, MAP3K3, MYC genes in PBMCs (relapse *vs.* remission) were identified as potential candidate genes for MS, which were the more central genes involved in biological pathways.

Discussion. This study showed that network-based analysis could explicate the complex interplay between biological processes underlying MS. Furthermore, an experimental validation of candidate genes can lead to identification of potential therapeutic targets.

Subjects Bioinformatics, Computational Biology, Immunology, Neurology
Keywords Protein–protein interaction network (PPIN), Transcriptome, Topology, Modularity, Clique analysis, Multiple sclerosis

INTRODUCTION

Multiple sclerosis (MS) is a complex disease affecting the central nervous system (CNS) in which genetic, environmental and immunological factors are considered as its etiology (*Ebers, 2008*; *Svejgaard, 2008*). Although MS shows both autoimmune and neurodegenerative features, the pathophysiological processes which may occur both within and outside of the CNS remain obscure and don't have an uniform distribution within the MS population (*Brynedal et al., 2010*). To study such complex diseases, which involved noticeably missing heritability (*Goris & Liston, 2012*; *Manolio et al., 2009*), it is more efficient to describe perturbed processes and dysregulated pathways rather to identify individual genes (*Kim, Wuchty & Przytycka, 2011*).

Transcriptome analysis of the target organ, i.e., the central nervous system, should reflect an unbiased survey of expression profiles for genes with altered transcript levels in disease states. Since it is difficult to achieve CNS samples, cerebrospinal fluid (CSF) cells have been used in many studies as a surrogate for the target organ in CNS disorders (*Brynedal et al., 2010*). Furthermore, peripheral blood mononuclear cells (PBMCs) are being considered as an easily accessible and informative source of biological material in MS transcriptome studies (*Achiron et al., 2004*; *Bomprezzi et al., 2003*; *Singh et al., 2007*). In this line, it has been reported that the study linking peripheral and CSF immune responses were essential to understand the immunopathogenesis of MS (*Christensen et al., 2012*).

Since the expression level change of a gene in a transcriptomic profile may be a result of an expression change of another gene and may not be the direct cause of the cellular phenotype, additional information is required to put them in context (*Wachi, Yoneda & Wu, 2005*). Network-based analyses of protein–protein interaction (PPI) or interactome delineate the known associations among proteins in the context of biochemistry, signal transduction and biomolecular networks (*Rezaei-Tavirani et al., 2016*; *Zali & Rezaei-Tavirani, 2014*; *Wu et al., 2009*). The integration analysis of large scale gene expression and PPI data will place the differentially expressed genes in the biological context (*Bapat et al., 2010*; *Li et al., 2012*; *Safari-Alighiarloo et al., 2014*; *Safari-Alighiarloo et al., 2016*). Protein networks reflect the functional grouping of interacting up/down regulated genes. The roles of the subsets of these genes, therefore, may be resolved using the combined data (*Wachi, Yoneda & Wu, 2005*).

Recently, the topological analyses have been applied to PPI networks by tools or algorithms such as modularity and centrality analyses by which the biological significance of proteins has been determined (*Huang et al., 2013*; *Lee et al., 2011*). Graph centrality measures like degree, betweenness and closeness centrality are very useful in the identification of nodes that are functionally crucial in the network by ranking the elements of the network (*Hindumathi et al., 2014*). In the PPI network the nodes with high degree defined as hub proteins and the nodes with high betweenness defined as bottleneck proteins, which both play a pivotal role in networks (*Yu et al., 2007*). The interest gene sets which usually presented as gene modules, protein complexes or pathways have been analyzed in integrative databases such as the Database for Annotation, Visualization and Integrated Discovery (DAVID) (*Huang, Sherman & Lempicki, 2008*), the Kyoto Encyclopedia of

Genes and Genomes (KEGG) and Reactome to identify the sets of biological processes and molecular pathways of genes.

This study integrated transcriptome-interactome data to construct PPI networks for MS using abnormally expressed genes in paired CSF and PBMCs samples. Topological analyses were performed to determine the significant network biomarkers. Underlying biological processes and pathways have been sought by modularity and clique analyses. Finally, potential disease markers were identified, which were high centrality genes significantly involved in functional modules or complexes.

METHODS

Transcriptome data collection and processing

Gene expression profiles in both CSF cells and PBMCs were obtained from the ArrayExpress Database under the accession number of EMTAB- 69 based on the Human Genome 133 plus 2.0 arrays (*Brynedal et al., 2010*). Accordingly, this study consisted of 26 multiple sclerosis patients, of whom 12 and 14 patients were sampled during relapse and remission, respectively. The MS patients were selected from a large cohort of newly diagnosed MS patients, and none of the patients had ever received immunomodulatory drugs. Control population included 18 subjects with other neurological diseases to assess MS specific transcriptome. The microarray raw data were converted to gene expression values using the RMA algorithm by the affy package within R software (*Gautier et al., 2004*). After preprocessing, each expression profile containing 54, 675 probe sets that ones with less discriminative power were removed according to the measurement of overall variance by the varFilter function using the genefilter package from the Bioconductor project within R software (*Gentleman et al., 2011*). After the preprocessing, a total of 27,336 probe sets from each sample were used for further analysis. To identify differential expression of the selected probes, the limma package in R software was used to perform the moderated t-test (*Smyth, 2005*). Where a gene had more than one probe on the microarray, the average expression value of all the related probes was used to estimate expression level of the gene.

Interactome data

The human PPI network was gathered from four major IMEx (*Orchard et al., 2007*) public databases: IntAct (*Kerrien et al., 2012*), MINT (*Ceol et al., 2010*), DIP (*Xenarios et al., 2002*) and InnateDB (*Lynn et al., 2008*). Indeed, public PPI databases which only stored experimentally verified interactions used to eliminate possible spurious interactions and avoid misleading conclusions. Our recent study showed IMEx databases (especially IntAct and DIP databases) had a greater number of significant correlations for their proteins' topological features than the all other paired comparisons between BIND, HPRD, MINT, IntAct and DIP databases (*Safari-Alighiarloo, Taghizadeh & Rezaei-Tavirani, 2015*).

QQPPI networks construction and topological analysis

To construct Query-Query PPI (QQPPI) networks, the differentially expressed genes in CSF (MS *vs.* control) and PBMCs (relapse *vs.* remission) were separately located on human PPI network. QQPPI networks only consistent of the query genes as the nodes and direct

interactions among them. The subnetworks of QQPPI were constructed and visualized by Cytoscape software (*Shannon et al., 2003*). Centrality parameters of QQPPI networks were analyzed using the Cytoscape and CentiBin softwares (*Junker, Koschützki & Schreiber, 2006*). The following parameters were calculated to determine biologically significant nodes (*Zhang, 2009*). Degree: the number of links to a given node. The Betweenness centrality of node v is calculated as:

$$C_B(v) = \sum_{s \neq t \neq v \in V} \frac{\rho st(v)}{\rho st}, \tag{1}$$

where the number of all shortest paths between node s and t regarded as ρst, and the number of shortest paths which passing through a node v out of ρst regarded as $\rho st(v)$. Indeed, this formula represents the ratio of the number of shortest paths passing through node v to the number of all shortest paths between s and t. The current flow betweenness centrality of a node v is the average of the current flow over all source–target pairs. Closeness centrality is defined as the reciprocal of the total distance from a node v, to all other nodes. Therefore, high values of closeness should indicate that all other nodes are in proximity to node v.

$$C_C(v) = \frac{1}{\sum_{u \in v} \text{dis}(u, v)}. \tag{2}$$

The centroid value is the most complex node centrality index and is computed by focusing the calculus on couples of nodes (v, w). The centroid value of an individual node 'v' is calculated by considering the number of nodes that have minimum shortest path which are closer to 'v' than 'w'. A node v with the highest centroid value is the node with the highest number of neighbors separated by the shortest path to v,

$$C_{cen}(v) = \min(f(v, w)), \tag{3}$$

where $f(v, w) = yv(w) - yw(v)$ and $yv(w)$ denotes the number of nodes that are closer to v than w. Eigen vector centrality assign the relative significance of all nodes in the network by weighting connections to highly important nodes more than connections to nodes of low importance.

$$\lambda C_{IV} = AC_{IV}, \tag{4}$$

where C_{IV} donates the Eigen vector and λ donates the Eigen value.

Hub and bottleneck nodes were extracted from the networks in two steps; (1) In the networks, nodes with degree greater than or equal to the sum of mean and twice the standard deviation (S.D.), i.e., mean $+ 2^\star$S.D. of the degree distribution, were considered as hubs (*Ray, Ruan & Zhang, 2008*). (2) We defined bottlenecks as the proteins that were in the top 5% in terms of betweenness centrality.

Identification and annotation of functional modules

Network clustering was implemented by Clustering with overlap neighborhood expansion (ClusterONE) algorithm in order to identify the connected regions within the networks

with possible overlap (*Nepusz, Yu & Paccanaro, 2012*). The modules were identified to have a minimum density of >0.05 and a degree of >5. A cluster with a *p*-value of <0.05 was determined to be a module. The functional meaning for identified modules was further explored, and they considered as candidate functional modules if their genes were significantly enriched in the biological process of Gene Ontology (GO) annotation or KEGG pathway.

Identification of complexes containing clique

CFinder software was applied to extract biologically meaningful protein complexes from the PPI networks (*Adamcsek et al., 2006*). CFinder (http://www.cfinder.org/) was downloaded and implemented locally. Cliques with 3 nodes and 4 nodes (3-cliques, 4-cliques) were identified in the QQPPI networks by this software. The cliques were searched against CORUM database (*Ruepp et al., 2010*) to find significant protein complexes. Then, all the proteins associated with a specific complex were identified using the in house algorithm. Complexes containing 3 or more query proteins, as a cut-off, were listed in this study.

Functional enrichment analysis

An enrichment analysis was performed using Functional Annotation Chart in DAVID bioinformatics. To determine functional modules, only the enriched GO terms and pathways with *p*-values < 0.05 were considered significant. Furthermore, Cytoscape Enrichment Map plugin was used to visualize significant terms enriched in entire networks by following parameters: *p*-value cut-off = 0.001, *q*-value cut-off = 0.05, overlap coefficient cut-off = 0.6 (*Merico et al., 2010*).

RESULTS

Expression analysis

We used the Limma package to analyze gene expression profile, E-MTAB-69, for comparison of four transcriptomes in MS (CSF: MS *vs*. controls and relapse *vs*. remission, PBMCs: MS *vs*. controls and relapse *vs*. remission). There were 3,062 genes with FDR < 0.05 whose expression was different in the CSF of MS patients as compared to the controls, but none in the respective PBMCs comparison. The number of up and down regulated genes was 1,080 and 1,982, respectively. On the contrary, when MS patients in relapse to those in remissions were compared, 1,163 differential expression genes with FDR ≤ 0.1 were seen in PBMCs, but none in the CSF. The number of up and down regulated genes was 301 and 762, respectively. The full lists of annotated differentially expressed probe sets are shown in Table S1 for the MS *vs*. control comparison in CSF cells, and in Table S2 for the relapse versus remission in PBMCs cells.

Networks' topological analysis

We used only direct interactions of differentially expressed genes to construct QQPPI networks. The CSF PPI network consisted of 1,440 nodes and 3,500 edges and PBMCs PPI network involved 483 nodes and 941 edges. Topological features were processed to characterize the biology network from the random network. The power law of node degree
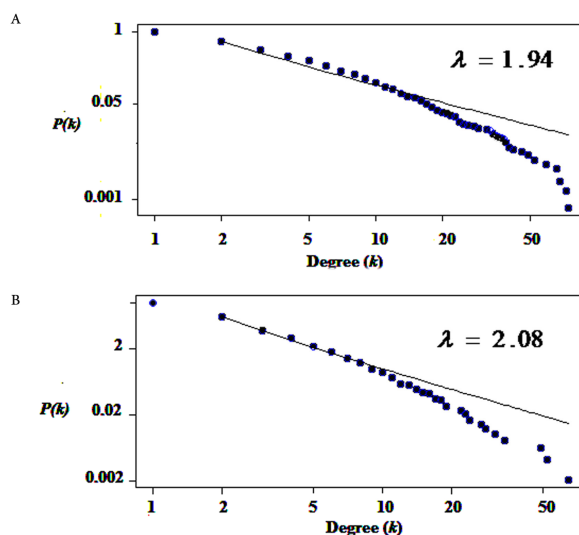
Figure 1 The degree distribution of nodes followed power law distribution. (A) Degree distribution of differentially expressed genes in CSF QQPPI network. (B) PBMCs QQPPI network. The graph represents a decreasing trend of degree distribution with an increase in the number of links showing scale-free topology.

Table 1 Hub-bottleneck identification. Cut-off determination for hubs & number of hubs and bottlenecks.

	Mean (M)	Standard Deviation (S.D)	Cut-off (M + 2*S.D)	Number of hubs	Number of bottlenecks
CSF	4.86	7.3	19.4	56	72
PBMCs	3.89	6.06	16.01	20	25

distribution is one of most important criteria (*Maslov & Sneppen, 2002*; *Zhu, Gerstein & Snyder, 2007*). The distribution of node degree approximately followed power law distributions, where $P(k)$ is a distribution of node degree, k is a degree and λ is a degree exponent, with $\lambda = 1.94$ and $\lambda = 2.08$ for CSF and PBMCs networks, respectively, and Fig. 1 indicates that the QQPPI networks were scale-free. The hubs and bottlenecks were extracted from the QQPPI networks by the criterion described in the method section (Table 1). Besides, we calculated four others centrality measurements involving closeness centrality, centroid value, Eigen vector centrality and current flow betweenness centrality and identified more central genes in the networks. The list of all nodes and their centrality measurements are prepared in Tables S3 and S4 for CSF and PBMCs in which candidate markers have been highlighted. The graphical structure of CSF and PBMCs PPI networks containing 5% top central genes are represented in Fig. 2.

Modularity analysis

The ClusterONE algorithm was selectively implemented on CSF and PBMCs networks to mine the functional modules, which may reveal a lot of hidden biological significant processes. We further performed GO and pathway analysis using DAVID tool to characterize functional modules; 14 and six functional modules were discovered for
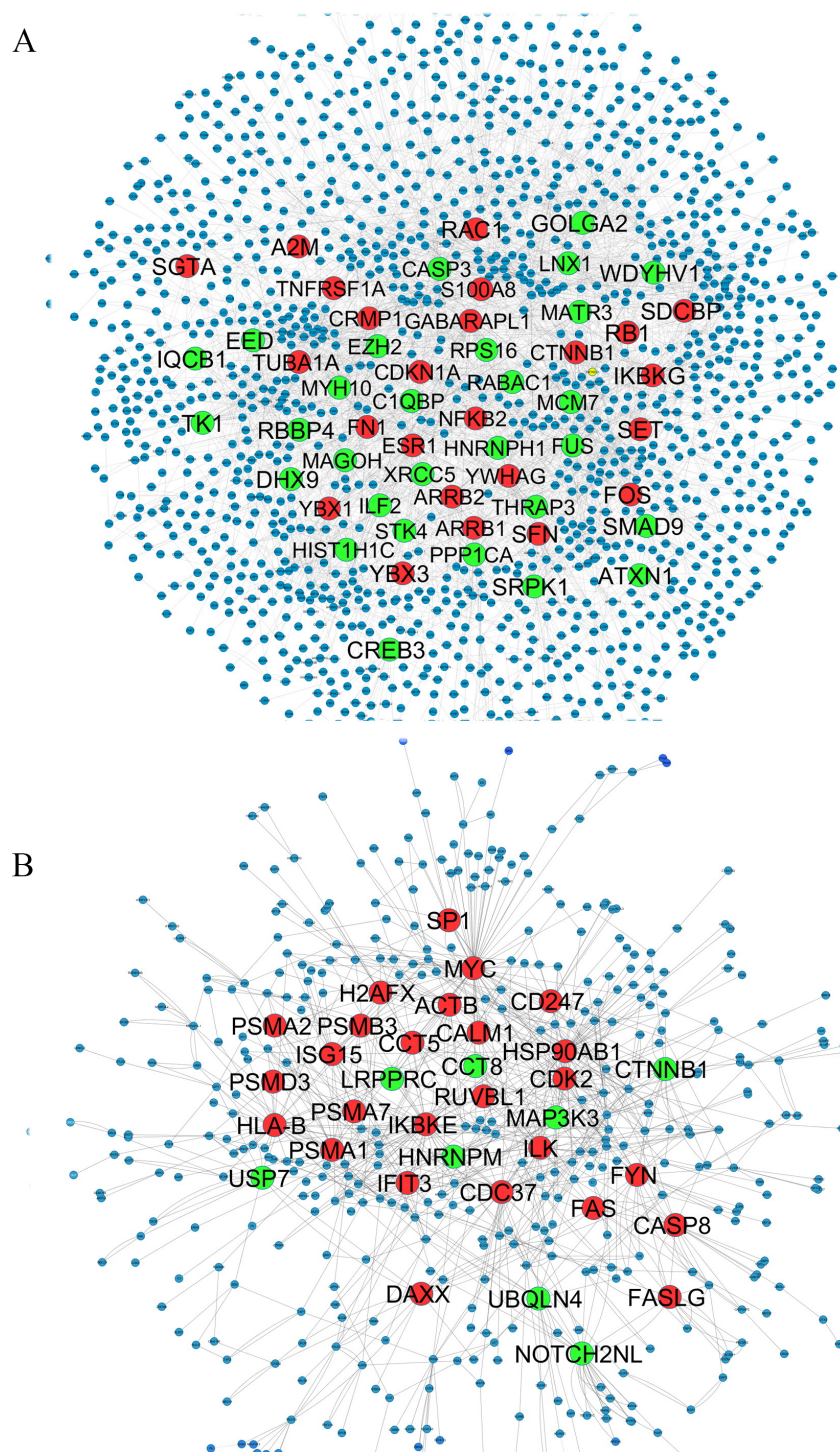
Figure 2 QQPPI networks generation by mapping of differentially expression genes on PPI data. (A) CSF QQPPI network. (B) PBMCs QQPPI network. Nodes with high centrality measures are shown by bigger size than others. Green and red nodes represent proteins encoded by up- and down-regulated genes, respectively. Graphical representation of nodes was implemented by "Spring Embedded" layout in Cystoscape.

Table 2 Modularity analysis. The list of pathways enriched in modules for CSF (MS *vs.* controls).

Module ID	Pathway	*p*-value
M1	hsa04062 : chemokine signaling pathway	7.7E–3
	hsa04060 : cytokine–cytokine receptor interaction	1.5E–2
	hsa04672 : intestinal immune network for IgA production	3.8E–2
M2	hsa05010 : Alzheimer's disease	3.0E–3
	hsa05014 : Amyotrophic Lateral Sclerosis (ALS)	3.1E–2
	hsa04720 : long-term potentiation	4.0E–2
M4	hsa04640 : hematopoietic cell lineage	4.7E–6
	hsa04060 : cytokine–cytokine receptor interaction	1.4E–4
	hsa04210 : apoptosis	8.6E–4
	hsa05020 : prion diseases	2.1E–2
	hsa05332 : graft-versus-host disease	2.3E–2
	hsa04940 : type I diabetes mellitus	2.5E–2
M7	hsa00590 : arachidonic acid metabolism	3.3E–2
M9	hsa04620 : toll-like receptor signaling pathway	1.7E–5
M13	hsa05217 : basal cell carcinoma	2.3E–3
	hsa04340 : hedgehog signaling pathway	2.3E–3
M14	hsa04010 : MAPK signaling pathway	1.5E–2
M15	hsa04110 : cell cycle	1.2E–2
M20	hsa04115 : p53 signaling pathway	3.3E–10
	hsa04110 : cell cycle	2.5E–8
	hsa04914 : progesterone-mediated oocyte maturation	1.8E–3
M22	hsa04810 : regulation of actin cytoskeleton	1.6E–2
	hsa04666 : Fc gamma R-mediated phagocytosis	2.4E–2
	hsa04530 : tight junction	4.4E–2
M23	hsa04144 : endocytosis	1.5E–2
	hsa04540 : gap junction	2.7E–2
M25	hsa04510 : focal adhesion	2.7E–2
	hsa04512 : ECM-receptor interaction	3.1E–2

CSF and PBMCs (*p*-value < 0.05), respectively. In the case of CSF, enriched modules were relevant to the comparison of MS versus controls in which modules correlated remarkably with many immune-related pathways such as, cytokine–cytokine receptor interaction, chemokine signaling pathway, Toll-like receptor signaling pathway, T cell receptor signaling pathway and Hematopoietic cell lineage. Further to them, some modules were enriched for apoptosis, p53 signaling pathway, MAPK signaling pathway, Hedgehog signaling pathway and Fc gamma R-mediated phagocytosis. The other major enriched pathways in modules included focal adhesion, cell cycle, endocytosis, gap junction, tight junction, ECM-receptor interaction, regulation of actin cytoskeleton (Table 2).

For PBMCs, enriched modules corresponded to the comparison of relapse versus remission in which the majority of enriched pathways contributed to immune-related pathways like antigen processing and presentation, primary immunodeficiency, RIG-I-like receptor signaling pathway, Toll-like receptor signaling pathway and cytosolic DNA-sensing

Table 3 Modularity analysis. The list of pathways enriched in modules for PBMCs (relapse *vs.* remission).

Module ID	Pathway	*p*-value
M1	hsa04612 : antigen processing and presentation	6.6E–8
	hsa05340 : primary immunodeficiency	2.7E–2
	hsa05332 : graft-versus-host disease	3.0E–2
	hsa02010 : ABC transporters	3.4E–2
M7	hsa04115 : p53 signaling pathway	3.5E–3
	hsa04110 : cell cycle	1.2E–2
M8	hsa04623 : cytosolic DNA-sensing pathway	2.5E–4
	hsa04622 : RIG-I-like receptor signaling pathway	5.2E–4
	hsa04620 : toll-like receptor signaling pathway	1.5E–3
M9	hsa04120 : ubiquitin mediated proteolysis	2.3E–2
M10	hsa03050 : proteasome	1.3E–9
M11	hsa03040 : spliceosome	1.0E–3
	hsa04350 :TGF-beta signaling pathway	4.1E–2

pathway. TGF-beta signaling pathway and p53 signaling pathway were the two noticeable signaling pathways in modules. The last enriched pathways were spliceosome, proteasome, ubiquitin mediated proteolysis and cell cycle (Table 3).

Identification of cliques and complexes

CFinder software was implemented to identify several 3-cliques and 4-cliques in the QQPPI networks. The corresponding complexes were retrieved from the CORUM database and shown in Table 4. In the case of CSF (MS *vs.* controls), these complexes mediated various biological functions such as protein processing (proteolytic), proteasomal degradation, stress response, protein binding, protease activator (ID: 32, 192 and 193), DNA conformation modification, transcription repression, protein modification by acetylation, deacetylation (ID:58), DNA conformation modification, transcription repression and posttranscriptional control (ID:105, 995,996 and 974), mitotic cell cycle and protein modification (ID:310 and 313), chromosome segregation/division (ID:1464), cell junction (ID:1839), actin cytoskeleton organization and biogenesis (ID:3008), ribosome biogenesis (ID:3055), protein modification and cellular signaling (MAPKKK cascade (ID:5909 and 5886).

For PBMCs (relapse *vs.* remission), the identified complexes involved in many biological processes like protein processing (proteolytic), proteasomal degradation, stress response, protein binding, protease activator (ID: 38, 39, 181, 191, 192, 193, 194), RNA processing and RNA binding (ID: 351 and ID: 1181), protein targeting, sorting and translocation, protein transport and homeostasis (ID:437), protein kinase (ID:5199), NIK-I-kappaB/NF-kappaB cascade and cytokine activity (ID: 5269), apoptosis (ID: 5473, 5860, 5859, 5799 and 5800).

Functional enrichment analysis of the networks

To gain a full view of the networks potential functions, the networks' nodes were annotated using the Functional Annotation Chart in DAVID and visualized using the Enrichment

Table 4   Clique analysis. The list of complexes enriched for CSF and PBMCs.

Gene symbol	Complex
CSF (MS *vs.* controls)	
PSMA3, PSMB1, PSMB3, PSMB9, PSME1, PSMD7	Proteasome (ID:39, 192,193)
GPS2, NCOR2, TBL1X	SMRT complex (ID:58)
EED, EZH2, RBBP4	Polycomb repressive complex 2,3 (PRC 2,3) (ID:105, 996,995), EED-EZH2 complex (ID:974)
CCNB1, CCNB2, CCND1, CDK1, CDKN1A	Cell cycle kinase complex CDC2 (ID:310)
CCND1,CCND3, CDKN1A	Cell cycle kinase complex CDK5 (ID:313)
CBX5, DSN1, ZWINT	Mis12 centromere complex (ID:1464)
CTNNA1, CTNNB1, SDCBP	SDCBP-CTNNB1-CTNNA1-CDH1 complex (ID:1839)
TUBA1A, TUBA1B, TUBA1C	60S APC containing complex (ID:3008)
IGF2BP1, ILF2, NOLC1, RPLP2, RPS11, RPS16, SRPK1, TUBA1A, YBX1	Nop56p-associated pre-rRNA complex (ID:3055)
MAP2K1, SFN, YWHAG	Ksr1 complex (Ksr1, Mek, 14-3-3, Mapk), EGF stimulated (ID:5909, 5886)
PBMCs (relapse *vs.* remission)	
PSMA1, PSMA2, PSMA7, PSMB10, PSMB3, PSMD3, PSMD4	Proteasome (ID: 38, 39, 181, 191, 192, 193, 194)
DHX15, PABPC1, PRPF19, SF3B3, SNRPB	Spliceosome (ID:351)
ACTB, ANXA6, MYH9, SPTAN1	PA700-20S-PA28 complex (ID:437)
HNRNPH1, HNRNPM, PABPC1, PRPF19, SF3B3, SNRPB	C complex spliceosome (ID:1181)
CDC37, HSP90AB1, MAP3K3	Kinase maturation complex 1 (ID:5199)
CDC37, HSP90AB1, IKBKE	TNF-alpha/NF-kappa B signaling complex 8 (ID: 5269)
CASP8 FADD FAS	FAS-FADD-CASP8 complex (ID: 5473, 5860), FAS-FADD-CASP8-CASP10 complex (ID: 5859), Death induced signaling complex DISC (ID: 5799, 5800)
ACTB, MYH9, SPTAN1	Emerin complex 1 (ID: 5604)

Map plugin in Cytoscape. As shown in Fig. 3, each node represented one functional annotation term. Nodes with more enriched genes were larger. Edge width was indicated the extent of overlapping between these categories (overlap coefficient cut-off 0.6). In case of CSF (MS *vs.* control), the ten most enriched entries in Gene Ontology (GO) biological processes were GO:0009611~Response to wounding, GO:0006955~Immune response, GO:0007242~Intracellular signaling, GO:0010604~Positive regulation of macromolecule metabolic processes, GO:0007049~Cell cycle, GO:0002684~Positive regulation of immune system processes, GO:0042981~Regulation of apoptosis, GO:0006954~Inflammatory response, GO:0016044~Membrane organization, GO:0008283~Cell proliferation. In case of PBMCs (relapse *vs.* remission), ten most enriched terms included GO:0043068~Positive regulation of programmed cell death, GO:0006974~Response to DNA damage stimulus, GO:0043065~Positive regulation of apoptosis, GO:0006917~Induction of apoptosis, GO:0044265~Cellular macromolecule metabolic processes, GO:0033554~Cellular response to stress, GO:0022402~Cell cycle processes, GO:0010033~Response to organic substance, GO:0002684~Positive regulation of immune system process, GO:0051249~Regulation of lymphocyte activation.
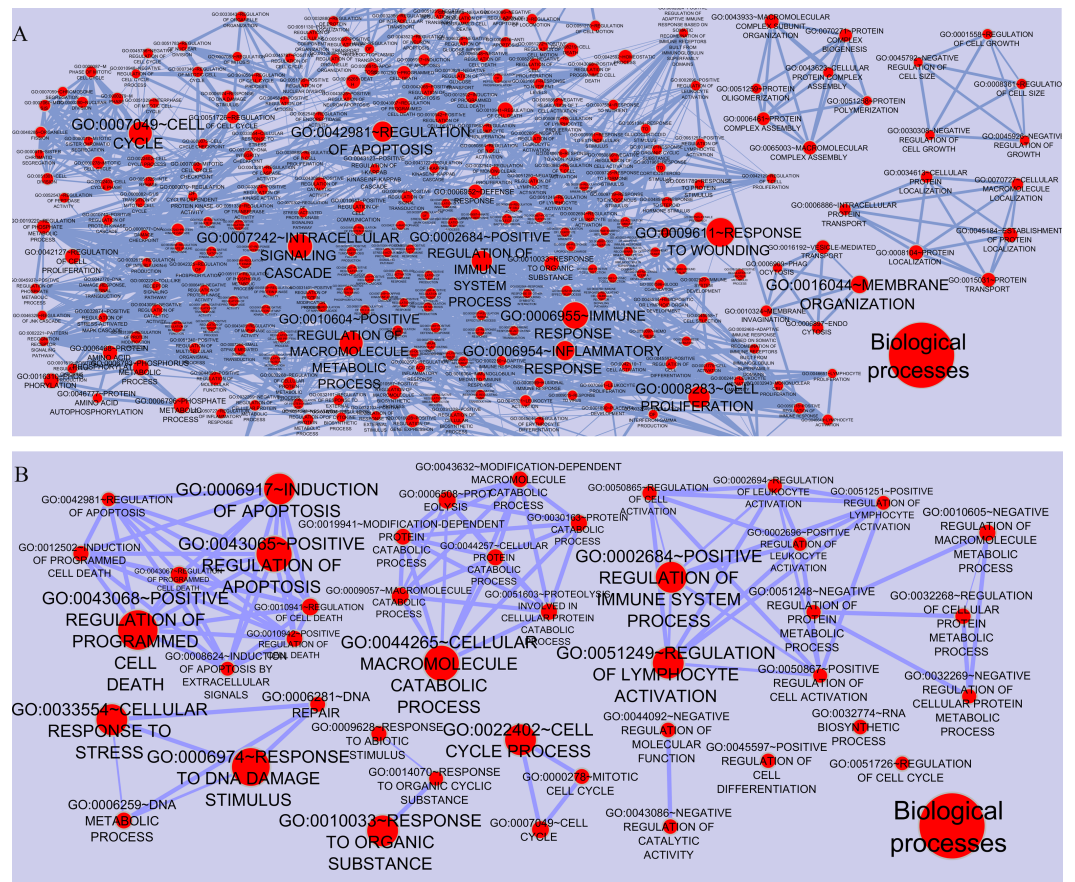
Figure 3 **Functional categories of the networks were visualized using the Enrichment map plugin of the Cytoscape.** Significant biological processes are represented by one node in (A) CSF QQPPI network. (B) PBMCs QQPPI network. Nodes' sizes indicate the significance of the enrichment (*p*-value). Edges show gene overlap between nodes and thickness indicates the number of overlapping enriched genes.

Mining and identification disease markers in modules and complexes

We screened more central nodes in CSF and PBMCs to investigate detailed analysis about their association in functional modules and complexes. Interestingly it was found that eight genes (STK4, RB1, CDKN1A, CDK1, RAC1, ARRB2, ARRB1, FN1) were located in functional modules and 15 genes (RBBP4, ILF2, RPS16, PSMA3, EED, EZH2, CDK1, CDKN1A, CTNNB1, SDCBP, SRPK1, TUBA1A, YBX1, SFN, YWHAG) were associated in complexes in CSF-QQPPI network. In case of PBMCs-QQPPI network, four genes (CDK2, PSMA1, IKBKE and MYC) were located in functional modules and genes (PSMA1, PSMA2, PSMA7, PSMB3, PSMD3, HNRNPM, FAS, ACTB, CDC37, HSP90AB1, MAP3K3, IKBKE, CASP8) were associated in complexes. Besides, it was obvious from the mining of functional modules and complexes that more central genes incorporated in driving pathways of MS pathogenesis (Table 5, also see Tables 2, 3 and 4). The more central genes and their expression values in CSF and PBMCs networks are illustrated in Fig. 4 and candidate markers are represented in Fig. 5. Since these differentially expressed genes in microarray dataset corresponded to topologically significant nodes in PPI networks,

Table 5 Central genes. The list of more central genes enriched in functional modules and complexes for CSF and PBMCs.

Module/complex ID	Gene symbol
CSF(MS *vs.* controls)	
M13	SMAD1
M14	STK4
M15	RB1
M20	CDKN1A, CDK1
M22	RAC1
M23	ARRB2, ARRB1
M25	FN1
ID:39, 192,193	PSMA3
ID:105, 974,996,995	EED, EZH2, RBBP4
ID:310,313	CDK1, CDKN1A
ID:1839	CTNNB1, SDCBP
ID:3008	TUBA1A
ID:3055	SRPK1, YBX1, ILF2, RPS16
ID:5909, 5886	SFN, YWHAG
PBMCs (relapse *vs.* remission)	
M7	CDK2
M8	IKBKE
M10	PSMA1
M11	MYC
ID: 38, 39, 181, 191, 192, 193, 194	PSMA1, PSMA2, PSMA7, PSMB3, PSMD3
ID:1181	HNRNPM
ID:437,5604	ACTB
ID:5199, 5269	CDC37, HSP90AB1, MAP3K3, IKBKE
ID: 5473, 5860, 5859, 5799, 5800	CASP8,FAS

which have functional importance because of their involvement in functional modules and complexes, they called as candidate disease markers in our study.

DISCUSSION

Although myriad genetic studies investigate the MS pathogenesis, our understanding have remained incomplete about the exact mechanism and its genetics (*Baranzini et al., 2009*; *Zhang et al., 2011*). The analysis of network-based biological data provides prominent tool to decipher the genetic basis of complex diseases by unraveling genes and processes not recognizable by genetic association approaches (*Sharma et al., 2013*; *Yu et al., 2013*). Due to implication of both intrathecal and peripheral immune activation in MS pathogenesis (*Brynedal et al., 2010*; *Christensen et al., 2012*; *Veroni et al., 2015*), we constructed PPI networks of abnormally expressed genes in paired tissues (CSF and PBMCs) for MS by integrating interactome and transcriptomes data. Using the analysis of the networks, we not only recognized several underlying biological processes, we also identified some important candidate markers for MS.
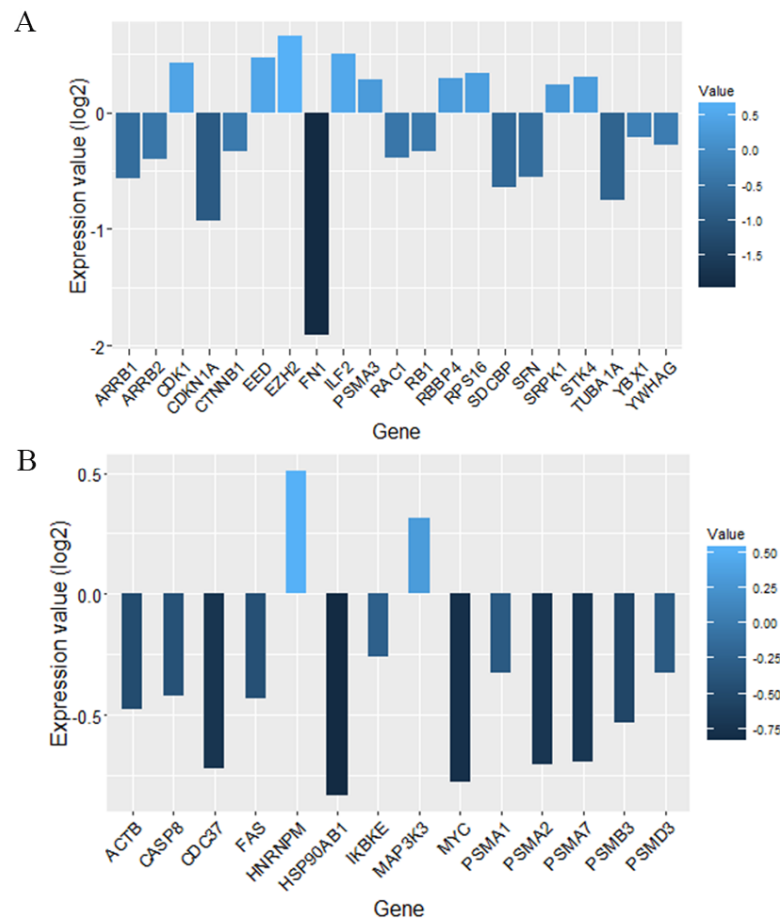
Figure 4 Nodes with high centrality measures which involved in significant biological pathways and their expression values. More central nodes in (A) CSF QQPPI network. (B) PBMCs QQPPI network.

By studying two different tissues, neither being the actually affected tissue in MS, largely different but potentially significant and noticeable results have been achieved. Since the CNS is embedded in CSF, it is presupposed that CSF would represent more of the processes occurring within the CNS (*Brynedal et al., 2010*). Besides, the implication of peripheral immune response in MS progression has been substantiated (*Kebir et al., 2009*). The simultaneously obtained CSF and PBMCs samples showed only differentially expressed genes in CSF, comparing MS patients versus controls, and PBMCs comparing relapse versus remission. Lack of differential expression in CSF (relapse *vs.* remission) implies that outside events of the CNS such as pathogens and other environmental triggers not influencing the cells of the CSF, may primarily affect MS bouts (*Brynedal et al., 2010*). To further explore underling biological processes in MS, functional modules and complexes were characterized in these different tissues. More central genes involved in immune response, apoptosis, cell cycle and cell adhesion pathways, which considered as main biological processes in MS pathogenesis. Furthermore, in PBMCs (relapse *vs.* remission), proteasome and spliceosome complexes were enriched by both modularity and clique analyses and in CSF (MS *vs.* control) only proteasome enriched by clique analysis. The
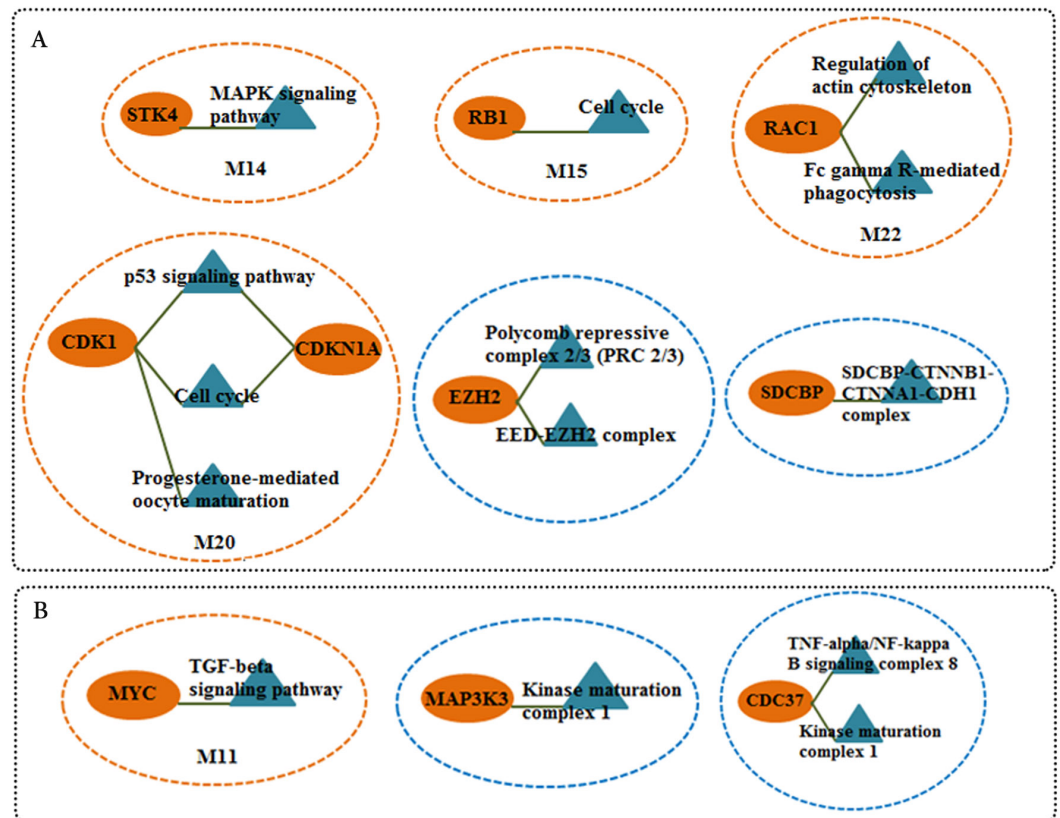
Figure 5 **Candidate markers involved in functional modules and complexes.** The functional enrichment of candidate markers in (A) CSF QQPPI network. (B) PBMCs QQPPI network. Modules and complexes illustrated by brown and blue dotted circles, respectively.

proteasome has crucial role in cell cycle progression and immune response (*Basler et al., 2015*). The immunoproteasome is a cytokine-induced variant of the 20S proteasome, which involved in the pathogenesis of autoimmune diseases and in the modulation of T helper cell differentiation (*Basler, Kirk & Groettrup, 2013*). Inhibition of the immunoproteasome subunit LMP7 (β5i) in animal models for autoimmune diseases including MS protected against these diseases (*Basler et al., 2014*). Spliceosome as basic machinery splicing, only enriched in relapse phase of MS in PBMCs. There is evidence to delineate the important role of alternative splicing in autoimmunity (*Evsyukova et al., 2010*).

In this work, we also identified the candidate genes for MS employing of topological analysis on the constructed networks in CSF and PBMCs tissues. These markers were further analyzed through a literature survey to confirm their potential contributions in MS pathogenesis. The compendium annotations for some of the most relevant candidate markers in terms of their expression are followed. In case of CSF, the first candidate was serine/threonine kinase 4 (STK4), also known MST1, enriched in MAPK signaling pathway (M14), and it was upregulated. The study of Konstantin et al. showed that genetic deletion of Mst1 altered T cell function and protected against autoimmunity as deletion of Mst1 reduced the severity of experimental autoimmune encephalomyelitis (EAE). Their

results indicated that Mst1 regarded as a critical regulator of adaptive immune responses, Th1/Th2-dependent cytokine production and as a potential therapeutic target for immune disorders (*Salojin et al., 2014*). The second marker, RB transcriptional corepressor 1 (RB1) participated in cell cycle pathway (M7) and was downregulated. This gene is a negative regulator of the cell cycle, and acts as a transcriptional regulator (*Indovina et al., 2013*). It has been reported that Rb-mediated gene expression repression of E2F2 (transcription factor 2), by acting to tether Rb to specific E2F promoter sites, was crucial in T cells, and mutation of E2F2 in mice resulted in enhanced T lymphocyte proliferation leading to the development of autoimmunity (*Murga et al., 2001*). The third marker, cyclin-dependent kinase inhibitor 1A (CDKN1A) also named p21, was enriched in p53 signaling pathway, cell cycle (M20), and it was downregulated. The evidence exists for the p21 as a cell-cycle inhibitor that suppressed autoimmunity (*Trakala et al., 2009*). Indeed, recent studies disclosed that p21 as a specific regulator of the homeostasis of memory/activated T lymphocytes (*Arias et al., 2007*). The fourth marker, cyclin-dependent kinase 1 (CDK1) was enriched in p53 signaling pathway, cell cycle and progesterone-mediated oocyte maturation pathway (M20), and it was upregulated. The study of *Yoshida et al. (2013)* implicated that cyclin-dependent kinases were important regulators and potential targets for modulation of T cell immunity and tolerance. In this line, their results showed that CDK (including CDK1) inhibitors prohibited Th17 differentiation and expedited iTreg (induced regulatory T cells) development, which induced improving of experimental autoimmune encephalomyelitis in mice. The fifth marker, rho family, small GTP binding protein Rac1 (RAC1) was enriched in regulation of actin cytoskeleton, Fc gamma R-mediated phagocytosis (M22), and it was downregulated. Rac1 as a major player of the Rho family of small GTPases, have a key regulatory role in both actin and microtubule cytoskeletal dynamics, which it is central to axonal growth and stability, as well as dendrite and spine structural plasticity in the nervous system. Besides, it is also a substantive regulator of NADPH-dependent membrane oxidase (NOX) which is a main source of reactive oxygen species (ROS). Thereby, Rac1 plays a principle role in the inflammatory response and neurotoxicity mediated by microglia cells in the nervous system (*D'Ambrosi et al., 2014*). The study results displayed the relevance of Rac1 dysregulation in the pathogenesis of Amyotrophic Lateral Sclerosis (ALS) (*D'Ambrosi et al., 2014*). The sixth marker, enhancer of zeste homologue 2 (EZH2) involved in polycomb repressive complexes 2, 3 and EED-EZH2 complex, and it was upregulated. EZH2 catalyzes the addition of three methyl groups to lysine 27 of histone H3 (H3K27) in target gene promoters, which caused gene silencing. The *Li & Jiang (2015)* study showed that the activity of EZH2 must be consistently inhibited in neurons to evade re-entrance into a cell cycle process, and thus its overexpression could begin a pathway that ended in CNS neurodegeneration. The seventh marker, the syndecan binding protein (SDCBP) involved in SDCBP-CTNNB1-CTNNA1-CDH1 complex and the enriched biological process, was the cell junction; it was downregulated. SDCBP was shown to interact with syndecans, which aided cell adhesion and enhanced attraction and concentration of growth factors at the cell surface (*Beekman & Coffer, 2008*). Lopez-Ramirez et al. study represented that human brain endothelial permeability has been controlled by miR-155, which targeted molecules involved in cell-to-cell interactions

such as SDCBP. They results indicated that $miR\text{-}155\_^{/}\_$ mice showed lower levels of blood–brain barrier leakage in experimental autoimmune encephalomyelitis and an acute model of systemic inflammation (*Lopez-Ramirez et al., 2014*).

In case of PBMCs, the candidate markers corresponded to the comparison of relapse versus remission. The first candidate marker was cell division cycle 37 (CDC37). It was incorporated in complex TNF-alpha/NF-kappa B signaling complex 8 and downregulated. CDC37 and HSP90 is a member of IKK complex that disruption of the interaction between CDC37/HSP90 and IKK complexes impaired the activation of IKK and NF-κB in a TNF-dependent manner (*Chen, Cao & Goeddel, 2002*; *Chen & Goeddel, 2002*). NF-κB acts as a central mediator of immune and inflammatory responses, and it is involved in regulation of cell proliferation and apoptosis (*Oeckinghaus & Ghosh, 2009*). In this line, some studies disclosed that apoptosis was suppressed during acute relapse and this issue may lead to prolonged survival of autoreactive T cells (*Achiron et al., 2007*; *Achiron et al., 2004*). Downregulation of CDC37 could be remarkable issue during relapse stage with apoptosis suppression for further studies. The second marker, mitogen-activated protein kinase 3 (MAP3K3) involved in Kinase maturation complex 1, and it was upregulate. It is a member of MAPKs which implicated in all aspects of immune responses, from the initiation phase of innate immunity to activation of adaptive immunity (*Dong, Davis & Flavell, 2002*). The last marker, v-myc avian myelocytomatosis viral oncogene homolog (MYC) incorporated in TGF-beta signaling pathway (M11), and it was downregulated. *Achiron et al. (2004)* study showed that all components of the TGF-beta signaling pathway were underexpressed during MS pathogenesis. Since this pathway is known to prohibit cell proliferation and increase susceptibility to apoptosis induced by TGF-beta, their underexpression may be relevant to autoreactive T-cell expansion in MS patients.

CONCLUSIONS

This study showed the necessity of network-based analysis to get more insights in MS pathogenesis at post-genomic era. In summary, QQPPI networks of abnormally expressed genes in paired CSF (MS *vs.* control) and PBMCs (relapse *vs.* remission) samples were constructed for MS, and centrality, modularity and clique analyses have been implemented. Our results indicated that genes with high centrality in the networks incorporated into the main biological processes in MS progression at CSF and PBMCs. Furthermore, we identified several candidate genes via the systems biology viewpoint which might facilitate the identification of potential targets for the treatment of MS.

ADDITIONAL INFORMATION AND DECLARATIONS

Safari-Alighiarloo et al. (2016), *PeerJ*, DOI 10.7717/peerj.2775

16/22

REFERENCES

Achiron A, Feldman A, Mandel M, Gurevich M. 2007. Impaired expression of peripheral blood apoptotic-related gene transcripts in acute multiple sclerosis relapse. *Annals of the New York Academy of Sciences* **1107**:155–167 DOI 10.1196/annals.1381.017.

Achiron A, Gurevich M, Friedman N, Kaminski N, Mandel M. 2004. Blood transcriptional signatures of multiple sclerosis: unique gene expression of disease activity. *Annals of Neurology* **55**:410–417 DOI 10.1002/ana.20008.

Adamcsek B, Palla G, Farkas IJ, Derényi I, Vicsek T. 2006. CFinder: locating cliques and overlapping modules in biological networks. *Bioinformatics* **22**:1021–1023 DOI 10.1093/bioinformatics/btl039.

Arias CF, Ballesteros-Tato A, García MI, Martín-Caballero J, Flores JM, Martínez-A C, Balomenos D. 2007. p21CIP1/WAF1 controls proliferation of activated/memory T cells and affects homeostasis and memory T cell responses. *The Journal of Immunology* **178**:2296–2306 DOI 10.4049/jimmunol.178.4.2296.

Bapat SA, Krishnan A, Ghanate AD, Kusumbe AP, Kalra RS. 2010. Gene expression: protein interaction systems network modeling identifies transformation-associated molecules and pathways in ovarian cancer. *Cancer Research* **70**:4809–4819 DOI 10.1158/0008-5472.CAN-10-0447.

Baranzini SE, Galwey NW, Wang J, Khankhanian P, Lindberg R, Pelletier D, Wu W, Uitdehaag BM, Kappos L, Polman CH. 2009. Pathway and network-based analysis of genome-wide association studies in multiple sclerosis. *Human Molecular Genetics* **18**:2078–2090 DOI 10.1093/hmg/ddp120.

Basler M, Kirk CJ, Groettrup M. 2013. The immunoproteasome in antigen processing and other immunological functions. *Current Opinion in Immunology* **25**:74–80 DOI 10.1016/j.coi.2012.11.004.

Basler M, Mundt S, Bitzer A, Schmidt C, Groettrup M. 2015. The immunoproteasome: a novel drug target for autoimmune diseases. *Clinical and Experimental Rheumatology* **33**:74–79.

Basler M, Mundt S, Muchamuel T, Moll C, Jiang J, Groettrup M, Kirk CJ. 2014.
Inhibition of the immunoproteasome ameliorates experimental autoimmune encephalomyelitis. *EMBO Molecular Medicine* **6**:226–238.

Beekman JM, Coffer PJ. 2008. The ins and outs of syntenin, a multifunctional intracellular adaptor protein. *Journal of Cell Science* **121**:1349–1355 DOI 10.1242/jcs.026401.

Bomprezzi R, Ringnér M, Kim S, Bittner ML, Khan J, Chen Y, Elkahloun A, Yu A, Bielekova B, Meltzer PS. 2003. Gene expression profile in multiple sclerosis patients and healthy controls: identifying pathways relevant to disease. *Human Molecular Genetics* **12**:2191–2199 DOI 10.1093/hmg/ddg221.

Brynedal B, Khademi M, Wallström E, Hillert J, Olsson T, Duvefelt K. 2010. Gene expression profiling in multiple sclerosis: a disease of the central nervous system, but with relapses triggered in the periphery? *Neurobiology of Disease* **37**:613–621 DOI 10.1016/j.nbd.2009.11.014.

Ceol A, Aryamontri AC, Licata L, Peluso D, Briganti L, Perfetto L, Castagnoli L, Cesareni G. 2010. MINT, the molecular interaction database: 2009 update. *Nucleic Acids Research* **38**:D532–D539.

Chen G, Cao P, Goeddel DV. 2002. TNF-induced recruitment and activation of the IKK complex require Cdc37 and Hsp90. *Molecular Cell* **9**:401–410 DOI 10.1016/S1097-2765(02)00450-1.

Chen G, Goeddel DV. 2002. TNF-R1 signaling: a beautiful pathway. *Science* **296**:1634–1635 DOI 10.1126/science.1071924.

Christensen JR, Börnsen L, Hesse D, Krakauer M, Sørensen PS, Søndergaard HB, Sellebjerg F. 2012. Cellular sources of dysregulated cytokines in relapsing-remitting multiple sclerosis. *Journal of Neuroinflammation* **9**:1–12 DOI 10.1186/1742-2094-9-215.

D'Ambrosi N, Rossi S, Gerbino V, Cozzolino M. 2014. Rac1 at the crossroad of actin dynamics and neuroinflammation in Amyotrophic Lateral Sclerosis. *Frontiers in Cellular Neuroscience* **8**:279.

Dong C, Davis RJ, Flavell RA. 2002. MAP kinases in the immune response. *Annual Review of Immunology* **20**:55–72 DOI 10.1146/annurev.immunol.20.091301.131133.

Ebers GC. 2008. Environmental factors and multiple sclerosis. *The Lancet Neurology* **7**:268–277 DOI 10.1016/S1474-4422(08)70042-5.

Evsyukova I, Somarelli JA, Gregory SG, Garcia-Blanco MA. 2010. Alternative splicing in multiple sclerosis and other autoimmune diseases. *RNA Biology* **7**:462–473 DOI 10.4161/rna.7.4.12301.

Gautier L, Cope L, Bolstad BM, Irizarry RA. 2004. affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**:307–315 DOI 10.1093/bioinformatics/btg405.

Gentleman R, Carey V, Huber W, Hahne F. 2011. Genefilter: methods for filtering genes from microarray experiments. R package version 1. *Available at http://www.bioconductor.org/packages/2.6/bioc/html/genefilter.html*.

Goris A, Liston A. 2012. The immunogenetic architecture of autoimmune disease. *Cold Spring Harbor Perspectives in Biology* **4**:a007260 DOI 10.1101/cshperspect.a007260.

Hindumathi V, Kranthi T, Rao S, Manimaran P. 2014. The prediction of candidate genes for cervix related cancer through gene ontology and graph theoretical approach. *Molecular BioSystems* **10**:1450–1460 DOI 10.1039/c4mb00004h.

Huang DW, Sherman BT, Lempicki RA. 2008. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols* **4**:44–57 DOI 10.1038/nprot.2008.211.

Huang K-C, Yang K-C, Lin H, Tsun-Hui TT, Lee W-K, Lee S-A, Kao C-Y. 2013. Analysis of schizophrenia and hepatocellular carcinoma genetic network with corresponding modularity and pathways: novel insights to the immune system. *BMC Genomics* **14**:S10 DOI 10.1186/1471-2164-14-S5-S10.

Indovina P, Marcelli E, Casini N, Rizzo V, Giordano A. 2013. Emerging roles of RB family: new defense mechanisms against tumor progression. *Journal of Cellular Physiology* **228**:525–535 DOI 10.1002/jcp.24170.

Junker BH, Koschützki D, Schreiber F. 2006. Exploration of biological network centralities with CentiBiN. *BMC Bioinformatics* **7**:1 DOI 10.1186/1471-2105-7-1.

Kebir H, Ifergan I, Alvarez JI, Bernard M, Poirier J, Arbour N, Duquette P, Prat A. 2009. Preferential recruitment of interferon-γ–expressing TH17 cells in multiple sclerosis. *Annals of Neurology* **66**:390–402 DOI 10.1002/ana.21748.

Kerrien S, Aranda B, Breuza L, Bridge A, Broackes-Carter F, Chen C, Duesbury M, Dumousseau M, Feuermann M, Hinz U. 2012. The IntAct molecular interaction database in 2012. *Nucleic Acids Research* **40**:D841–D846.

Kim Y-A, Wuchty S, Przytycka TM. 2011. Identifying causal genes and dysregulated pathways in complex diseases. *PLOS Computational Biology* **7**:e1001095 DOI 10.1371/journal.pcbi.1001095.

Lee S-A, Tsao TT, Yang K-C, Lin H, Kuo Y-L, Hsu C-H, Lee W-K, Huang K-C, Kao C-Y. 2011. Construction and analysis of the protein–protein interaction networks for schizophrenia, bipolar disorder, and major depression. *BMC Bioinformatics* **12**:S20 DOI 10.1186/1471-2105-12-S13-S20.

Li J, Jiang D. 2015. The role of epigenomics in the neurodegeneration of ataxia-telangiectasia. *Epigenomics* **7**:137–141 DOI 10.2217/epi.14.81.

Li M, Wu X, Wang J, Pan Y. 2012. Towards the identification of protein complexes and functional modules by integrating PPI network and gene expression data. *BMC Bioinformatics* **13**:109 DOI 10.1186/1471-2105-13-109.

Lopez-Ramirez MA, Wu D, Pryce G, Simpson JE, Reijerkerk A, King-Robson J, Kay O, De Vries HE, Hirst MC, Sharrack B. 2014. MicroRNA-155 negatively affects blood–brain barrier function during neuroinflammation. *The FASEB Journal* **28**:2551–2565 DOI 10.1096/fj.13-248880.

Lynn DJ, Winsor GL, Chan C, Richard N, Laird MR, Barsky A, Gardy JL, Roche FM, Chan TH, Shah N. 2008. InnateDB: facilitating systems-level analyses of the mammalian innate immune response. *Molecular Systems Biology* **4**:218.

Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A. 2009. Finding the missing heritability of complex diseases. *Nature* **461**:747–753 DOI 10.1038/nature08494.

Maslov S, Sneppen K. 2002. Specificity and stability in topology of protein networks. *Science* **296**:910–913 DOI 10.1126/science.1065103.

Merico D, Isserlin R, Stueker O, Emili A, Bader GD. 2010. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLOS ONE* **5**:e13984 DOI 10.1371/journal.pone.0013984.

Murga M, Fernández-Capetillo O, Field SJ, Moreno B, Luis R, Fujiwara Y, Balomenos D, Vicario A, Carrera AC, Orkin SH. 2001. Mutation of E2F2 in mice causes enhanced T lymphocyte proliferation, leading to the development of autoimmunity. *Immunity* **15**:959–970 DOI 10.1016/S1074-7613(01)00254-0.

Nepusz T, Yu H, Paccanaro A. 2012. Detecting overlapping protein complexes in protein–protein interaction networks. *Nature Methods* **9**:471–472 DOI 10.1038/nmeth.1938.

Oeckinghaus A, Ghosh S. 2009. The NF-κB family of transcription factors and its regulation. *Cold Spring Harbor Perspectives in Biology* **1**:a000034 DOI 10.1101/cshperspect.a000034.

Orchard S, Kerrien S, Jones P, Ceol A, Chatr-aryamontri A, Salwinski L, Nerothin J, Hermjakob H. 2007. Submit your interaction data the IMEx way. *Proteomics* **7**:28–34 DOI 10.1002/pmic.200700286.

Ray M, Ruan J, Zhang W. 2008. Variations in the transcriptome of Alzheimer's disease reveal molecular networks involved in cardiovascular diseases. *Genome Biology* **9(10)**:R148 DOI 10.1186/gb-2008-9-10-r148.

Rezaei-Tavirani M, Zamanian-Azodi M, Rajabi S, Masoudi-Nejad A, Rostami-Nejad M, Rahmatirad S. 2016. Protein clustering and interactome analysis in Parkinson and Alzheimer's diseases. *Archives of Iranian Medicine* **19(2)**:101–109.

Ruepp A, Waegele B, Lechner M, Brauner B, Dunger-Kaltenbach I, Fobo G, Frishman G, Montrone C, Mewes H-W. 2010. CORUM: the comprehensive resource of mammalian protein complexes—2009. *Nucleic Acids Research* **38**:D497–D501 DOI 10.1093/nar/gkp914.

Safari-Alighiarloo N, Taghizadeh M, Rezaei-Tavirani M. 2015. Cytoscape retrieved protein–protein interaction (PPI) networks: seeking the correlation of human proteins'topological features between major public PPI databases due to their medical importance. *International Journal of Analytical, Pharmaceutical and Biomedical Sciences* **4**:137–146.

Safari-Alighiarloo N, Taghizadeh M, Rezaei-Tavirani M, Goliaei B, Peyvandi AA. 2014. protein–protein interaction networks (PPI) and complex diseases. *Gastroenterology and Hepatology from Bed to Bench* **7**:9–16.

Safari-Alighiarloo N, Taghizadeh M, Tabatabaei SM, Shahsavari S, Namaki S, Khodakarim S, Rezaei-Tavirani M. 2016. Identification of new key genes for type 1 diabetes through construction and analysis of the protein–protein interaction networks based on blood and pancreatic islet transcriptomes. *Journal of Diabetes* Epub ahead of print Nov 21 2016.

Salojin KV, Hamman BD, Chang WC, Jhaver KG, Al-Shami A, Crisostomo J, Wilkins C, Digeorge-Foushee AM, Allen J, Patel N. 2014. Genetic deletion of Mst1 alters T

cell function and protects against autoimmunity. *PLOS ONE* **9**:e98151
DOI 10.1371/journal.pone.0098151.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research* **13**:2498–2504
DOI 10.1101/gr.1239303.

Sharma A, Gulbahce N, Pevzner SJ, Menche J, Ladenvall C, Folkersen L, Eriksson P, Orho-Melander M, Barabási A-L. 2013. Network-based analysis of genome wide association data provides novel candidate genes for lipid and lipoprotein traits. *Molecular & Cellular Proteomics* **12**:3398–3408 DOI 10.1074/mcp.M112.024851.

Singh M, Scott T, LaFramboise W, Hu F, Post J, Ehrlich G. 2007. Gene expression changes in peripheral blood mononuclear cells from multiple sclerosis patients undergoing β-interferon therapy. *Journal of the Neurological Sciences* **258**:52–59
DOI 10.1016/j.jns.2007.02.034.

Smyth GK. 2005. Limma: linear models for microarray data. In: *Bioinformatics and computational biology solutions using R and Bioconductor*. Berlin Heidelberg: Springer, 397–420.

Svejgaard A. 2008. The immunogenetics of multiple sclerosis. *Immunogenetics* **60**:275–286 DOI 10.1007/s00251-008-0295-1.

Trakala M, Arias CF, García MI, Moreno-Ortiz MC, Tsilingiri K, Fernández PJ, Mellado M, Díaz-Meco MT, Moscat J, Serrano M. 2009. Regulation of macrophage activation and septic shock susceptibility via p21 (WAF1/CIP1). *European Journal of Immunology* **39**:810–819 DOI 10.1002/eji.200838676.

Veroni C, Marnetto F, Granieri L, Bertolotto A, Ballerini C, Repice AM, Schirru L, Coghe G, Cocco E, Anastasiadou E. 2015. Immune and Epstein-Barr virus gene expression in cerebrospinal fluid and peripheral blood mononuclear cells from patients with relapsing-remitting multiple sclerosis. *Journal of Neuroinflammation* **12**:132 DOI 10.1186/s12974-015-0353-1.

Wachi S, Yoneda K, Wu R. 2005. Interactome-transcriptome analysis reveals the high centrality of genes differentially expressed in lung cancer tissues. *Bioinformatics* **21**:4205–4208 DOI 10.1093/bioinformatics/bti688.

Wu J, Vallenius T, Ovaska K, Westermarck J, Mäkelä TP, Hautaniemi S. 2009. Integrated network analysis platform for protein–protein interactions. *Nature Methods* **6**:75–77 DOI 10.1038/nmeth.1282.

Xenarios I, Salwinski L, Duan XJ, Higney P, Kim S-M, Eisenberg D. 2002. DIP, the database of interacting proteins: a research tool for studying cellular networks of protein interactions. *Nucleic Acids Research* **30**:303–305 DOI 10.1093/nar/30.1.303.

Yoshida H, Kotani H, Kondo T, Tani I, Wei X, Tsuruta S, Kimura A, Asakawa M, Ito M, Nagai S. 2013. CDK inhibitors suppress Th17 and promote iTreg differentiation, and ameliorate experimental autoimmune encephalomyelitis in mice. *Biochemical and Biophysical Research Communications* **435**:378–384 DOI 10.1016/j.bbrc.2013.04.096.

Yu H, Kim PM, Sprecher E, Trifonov V, Gerstein M. 2007. The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics. *PLOS Computational Biology* **3**:e59 DOI 10.1371/journal.pcbi.0030059.

Yu F-D, Yang S-Y, Li Y-Y, Hu W. 2013. Co-expression network with protein–protein interaction and transcription regulation in malaria parasite Plasmodium falciparum. *Gene* **518**:7–16 DOI 10.1016/j.gene.2012.11.092.

Zali H, Rezaei-Tavirani M. 2014. Meningioma protein–protein interaction network. *Archives of Iranian Medicine* **17**:262–272.

Zhang A. 2009. *Protein interaction networks: computational analysis.* New York: Cambridge University Press.

Zhang F, Shi Y, Wang L, Sriram S. 2011. Role of HDAC3 on p53 expression and apoptosis in T cells of patients with multiple sclerosis. *PLOS ONE* **6**:e16795 DOI 10.1371/journal.pone.0016795.

Zhu X, Gerstein M, Snyder M. 2007. Getting connected: analysis and principles of biological networks. *Genes & Development* **21**:1010–1024 DOI 10.1101/gad.1528707.