Int J Clin Exp Med 2015;8(6):9384-9393
www.ijcem.com /ISSN:1940-5901/IJCEM0007342

Original Article

Systematic tracking of dysregulated modules identifies disrupted pathways in narcolepsy

Zhenhua Liu[1], Jiali Zhao[2], Yinyin Tan[2], Minglu Tang[3], Guanzhen Li[3]

[1]Department of Neurology and Shandong Sleep Medical Center, Provincial Hospital Affiliated to Shandong University, Jinan 250021, China; [2]Department of Neurology, Provincial Hospital Affiliated to Shandong University, Jinan 250021, China; [3]Shandong Sleep Medical Center, Provincial Hospital Affiliated to Shandong University, Jinan 250021, China

Received February 25, 2015; Accepted May 25, 2015; Epub June 15, 2015; Published June 30, 2015

Abstract: Objective: The objective of this work is to identify disrupted pathways in narcolepsy according to systematically tracking the dysregulated modules of reweighted Protein-Protein Interaction (PPI) networks. Here, we performed systematic identification and comparison of modules across normal and narcolepsy conditions by integrating PPI and gene-expression data. Methods: Firstly, normal and narcolepsy PPI network were inferred and reweighted based on Pearson correlation coefficient (PCC). Then, modules in PPI network were explored by clique-merging algorithm and we identified altered modules using a maximum weight bipartite matching and in non-increasing order. Finally, pathways enrichment analyses of genes in altered modules were carried out based on Expression Analysis Systematic Explored (EASE) test to illuminate the biological pathways in narcolepsy. Results: Our analyses revealed that 235 altered modules were identified by comparing modules in normal and narcolepsy PPI network. Pathway functional enrichment analysis of disrupted module genes showed 59 disrupted pathways within threshold $P < 0.001$. The most significant five disrupted pathways were: oxidative phosphorylation, T cell receptor signaling pathway, cell cycle, Alzheimer's disease and focal adhesion. Conclusions: We successfully identified disrupted pathways and these pathways might be potential biological processes for treatment and etiology mechanism in narcolepsy.

Keywords: Narcolepsy, dysregulated modules, disrupted pathways, protein-protein interaction network

Introduction

Narcolepsy is a chronic neurological disorder caused by autoimmune destruction of hypocretin-producing neurons inhibiting the brain's ability to regulate sleep-wake cycles normally [1, 2]. People with narcolepsy often experience frequent excessive daytime sleepiness, sleep paralysis, hypnagogic hallucinations and cataplexy, these symptoms may not present in all patients [3]. It severely interferes with every aspect of patients' life, in work and social settings. Narcolepsy occurs in approximately 1 in 2000 individuals, while most cases are sporadic and can be substantially helped but not cured [4].

Studies have revealed that several genes are associated with narcolepsy. Manzotteet al. showed that more than 90% of narcolepsy patients have an association with human leukocyte antigens (HLA) [5, 6]. Meanwhile it has been reported that there is a strong association between narcolepsy and polymorphisms in the T-cell receptor α (*TCRα*) [7], and a weaker association with tumor necrosis factor super family number 4 (*TNFSF4*) [8] and the purinergic receptor subtype *P2Y11* [9]. However, the complex nature of central nervous system and the immune system presents unique challenges in finding a mechanism for narcolepsy [10].

With the advances of high-throughput experimental technologies, large amounts of Protein-Protein interaction (PPI) data are uncovered, which make it possible to study proteins in systematically level [11]. In addition, a PPI network can be modeled as an undirected graph, where vertices represent proteins and edges represent interactions between proteins, to prioritize disease associated genes or pathways and to understand the modus operandi of disease

mechanisms [12]. But it has been noticed that PPI data are often associated with high false positive and false negative rates due to the limitations of the associated experimental techniques and the dynamic nature of protein interaction maps, which may have a negative impact on the performance of complex discovery algorithms [13]. Many computational approaches have been proposed to assess the reliability of protein interactions data. An iterative scoring method proposed by Liu et al was selected to evaluate the reliability and predict new interactions, and it has been shown to perform better than other methods [14]. However, studying multiple diseases simultaneously makes it challenging to discern clearly the intricate underlying mechanisms.

In addition, it is important to effectively integrate omics data into such an analysis, for example, Chu and Chen combined PPI and gene expression data to construct a cancer perturbed PPI network in cervical carcinoma to study gain-and loss-of-function genes as potential drug targets [15]. Magger et al. combined PPI and gene expression data to construct tissue-specific PPI networks for 60 tissues and used them to prioritize disease genes [16]. Beyond straightforward scoring genes in the gene regulatory network, it is crucial to study the behavior of modules across specific conditions in a controlled manner to understand the modus operandi of disease mechanisms and to implicate novel genes [17], since some of important genes may not be identifiable through their own behavior, but their changes are quantifiable when considered in conjunction with other genes (e.g. as modules). What is required, therefore, is a systematic tracking gene and module behavior across specific conditions in a controlled manner. Besides, numbers of human genes has not yet been assigned to definitive pathways, scoring pathways based on module analysis becomes a more reliable analyzing approach.

Therefore, in this paper, we systematically tracked the altered modules of reweighted PPI networks to identify disturbed pathways between normal controls and narcolepsy patients, in order to further to reveal mechanism of narcolepsy. To achieve this, we firstly inferred normal and narcolepsy PPI network based on Pearson correlation coefficient (PCC); then clique-merging algorithm was performed to explore modules in PPI network, and we compared these modules to identify altered modules; finally pathways enrichment analysis of genes in altered modules were carried out based on Expression Analysis Systematic Explored (EASE) test.

Materials and methods

Inferring normal and narcolepsy PPI network

Using human protein interactions, we first constructed generic PPI network, and every interaction had a weight score reflecting the reliability of the interactions.

We utilized a dataset of literature-curated human PPI from the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING), comprising 16730 genes and 1048576 interactions [18]. For STRING, self-loops and proteins without expression value were removed. The remaining largest connected component with score more than 0.8 was kept as the selected PPI network, which consisting of 8590 genes and 53975 interactions.

The microarray expression profiles of E-GEOD-21592 from Array Express database was selected for narcolepsy related analysis. E-GEOD-21592, which consisted of 10 narcolepsy patients and 10 normal controls, existed in Affymetrix Gene Chip Human Genome U133A 2.0 Platform. We performed the standard methods for preprocessing the gene expression profile of each dataset, including “rma” [19], “quantiles” [20], “mas” [21] and “medianpolish” [19].

Briefly, in order to eliminate the influence of nonspecific hybridization, background correction was applied by rma method [19]. The observed Perfect match (PM) probes were modeled as the sum of a normal noise component N (Normal with mean μ and variance σ^2) and an exponential signal component S (exponential with mean α). To avoid any possibility of negatives, the normal was truncated at zero. Given O the observed intensity, Ø and Φ were the standard normal distribution density and distribution functions respectively, and $a = s - \mu - \sigma^2\alpha$, $b = \sigma$, the adjustment as following:

$$E(s \backslash O = o) = a + b \frac{\phi\left(\frac{a}{b}\right) - \phi\left(\frac{o-a}{b}\right)}{\Phi\left(\frac{a}{b}\right) + \Phi\left(\frac{o-a}{b}\right) - 1}$$

Quantiles based algorithm was applied to normalization of the data [20]. The quantile normalization method was a specific case of the transformation $x'_i = F^{-1}(G(x_i))$, where we estimated G by the empirical distribution of each array and F using the empirical distribution of the averaged sample quantiles. PM/Mismatch(MM) correction was conducted by mas method [21]. An ideal mismatch was subtracted from PM. The Ideal MM would always be less than the corresponding PM and thus we could safely subtract it without risk of negative values.

Medianpolish was performed to summarize the probe data [19]. A multichip linear model was fit to data from each probe set. In particular for a probe set k with $i = 1, ..., I_k$ probes and data from $j = 1, ..., J$ arrays we fitted the following model

$$\log_2(PM_{ij}^k) = (\alpha_i^k) + (\beta_j^k) + (\varepsilon_{ij}^k)$$

Where α_i was a probe effect and β_j was the $\log_2$ expression value.

In this paper, gene interactions in narcolepsy patients and normal controls network were reweighted by PCC, which evaluating the probability of two co-expressed gene pairs. PCC is a measure of the correlation between two variables, giving a value between -1 and +1 inclusive [22]. The PCC of a pair of genes (x and y), which encoded the corresponding paired proteins (u and v) interacting in the PPI network, was defined as:

$$PCC(x, y) = \frac{1}{s-1}\sum_{i=1}^{s}\left(\frac{g(x, i) - \bar{g}(x)}{\sigma(x)}\right)\cdot\left(\frac{g(y, i) - \bar{g}(y)}{\sigma(y)}\right)$$

Where s was the number of samples of the gene expression data; $g(x, i)$ or $g(y, i)$ was the expression level of gene x or y in the sample i under a specific condition; $\bar{g}(x)$ or $\bar{g}(y)$ represented the mean expression level of gene x or y and $s(x)$ or $s(y)$ represented the standard deviation of expression level of gene x (or y).

The PCC of a pair of proteins (u and v) was defined as the same as the PCC of their corresponding paired genes (x and y), that was PCC (u, v) = PCC (x, y). If PCC (u, v) has a positive value, there is a positive linear correlation between u and v. in addition, we defined PCC of each gene-gene interaction as weight value of the interaction.

Identifying modules from the PPI network

In this study, module-identification algorithm was based on clique-merging [23, 24]. Our algorithm worked in two steps: in the first step, it found all the maximal cliques from the weighted PPI network and ranked the cliques according to their weighted density, and in the second step, it merged highly overlapped cliques.

Cliques algorithm proposed by Tomia et al was applied to find maximal cliques [25]. It utilized a depth-first search strategy to enumerate all maximal cliques and effectively pruned non-maximal cliques during the enumeration process. The score of a clique C was defined as its weighted interaction density $d_W(C)$:

$$d_W(C) = \frac{\sum_{u\in C, v\in C} w(u,v)}{|C|\cdot(|C|-1)}$$

Where $w(u, v)$ was the weight of the interaction between u and v. Proteins in a larger clique were more likely to have more common neighbors than proteins in a smaller clique, so the edges within a larger clique were likely to have higher weights than those in a smaller clique. Therefore, if the density of two cliques was the same, the weighted density of the larger clique was likely to be higher than that of the smaller clique.

Many of maximal cliques might overlap with one another as thousands of them were generated from a PPI network. The high overlapped cliques should be removed to reduce result size. Besides, merging highly overlapped cliques to form bigger yet still dense sub graphs was also desirable since complexes were not necessarily fully connected and PPI data might be incomplete. The inter-connectivity between two cliques was selected to determine whether two overlapped cliques should be merged together or not. The weighted inter-connectivity I_w between the non-overlapping proteins of C_i and C_j was calculated as follows:

$$I_w(C_i, C_j) = \sqrt{\frac{\sum_{u\in(C_i-C_j)}\sum_{v\in C_j} w(u,v)}{|C_i - C_j|\cdot|C_j|}\cdot\frac{\sum_{u\in(C_j-C_i)}\sum_{v\in C_i} w(u,v)}{|C_j - C_i|\cdot|C_i|}}$$

Given a set of cliques ranked in descending order of their score, denoted as $\{C_1, C_2, ..., C_k\}$, the algorithm removed and merged highly overlapped cliques in the following steps. For every clique C_i, it checked whether there existed clique C_j such that C_j had a lower score than C_i and $|C_i \cap C_j|/|C_j| \geq t$, where $t = 0.5$ was a pre-

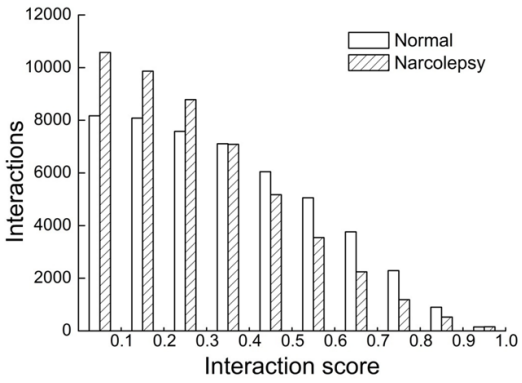

Figure 1. Scorewise distribution of interactions: Normal vs Narcolepsy.

defined threshold for overlapping [17]. If such C_j existed, then using the inter-connectivity scores between C_i and C_j to decide whether to remove C_j or merge C_j with C_i. If inter-score (C_1, C_2) ≥ m, then C_j was merged with C_i, otherwise, C_j was removed. Here, m = 0.25 was a predefined threshold for merging [17].

We captured the effect of differences in interaction weights between normal and narcolepsy through the weighted density-based ranking of cliques. Weighted density assigned higher rank to larger and stronger cliques. Therefore, we expected cliques with lost proteins or weakened interactions to go down the rankings resulting in altered module generation, thereby capturing changes in modules between normal and narcolepsy.

Comparing modules between normal and narcolepsy conditions

H_N and H_D represented the PPI network of normal controls and narcolepsy patients, which identifying the sets of modules S = {S_1, S_2, ..., S_m} and T = { T_1, T_2, ..., T_n} respectively. For each S_i ∈ S, module correlation density d_c (S_i) was defined as:

$$d_c(S_i) = \frac{\sum_{x,y \in S_i} PCC((x, y), M)}{|S_i| \cdot (|S_i| - 1)}$$

Correlation densities of narcolepsy modules (d_c (T_i)) were calculated similarly.

Disrupted or altered module pairs were evaluated by modeling the set γ (S, T) as a maximum weight bipartite matching [26]. Firstly we build a similarity graph M = (V_M, E_M), where V_M = {S U T}, and E_M = U {(S_i, T_j): J (S_i, T_j) ≥ t_J, Δ_c (S_i, T_j) ≥ δ}, where by J (S_i, T_j) = $|S_i \cap T_j|/|S_i \cup T_j|$ was the Jaccard similarity and Δ_c (S_i, T_j) = $|d_c(S_i) - d_c(T_i)|$ was the differential correlation density between S_i and T_j, t_J and δ were thresholds with 2/3 and 0.05 [17]. J (S_i, T_j) weighted every edge (S_i, T_j). We next identified the disrupted module pair's γ (S, T) by detecting the maximum weight matching in M, and we ranked them in non-increasing order of their differential density Δ_c. At last, we inferred genes involved in narcolepsy as Γ = {g: g ∈ S_i ∪ T_j, (S_i, T_j) ∈ γ (S, T)} ranked in non-increasing order of Δ_c (S_i, T_j). To identify altered modules, we matched normal and narcolepsy modules by setting a high t_J, which ensured that the module pairs either had the same gene composition or had lost or gained only a few genes.

Pathway enrichment analysis of genes in altered modules

The Database for Annotation, Visualization, and Integrated Discovery (DAVID) for KEGG pathway enrichment analysis was carried out to further investigate the biological functions of genes in altered modules from normal controls and narcolepsy patients [27]. In addition, pathway analysis of genes only existed in normal or narcolepsy module was also conducted. KEGG pathways with P value thresholds were chosen based on Expression Analysis Systematic Explored (EASE) test applied in DAVID. EASE analysis of the regulated genes indicated molecular functions and biological processes unique to each category [28]. The EASE score was used to detected the significant categories. In both of the functional and pathway enrichment analysis, the threshold of minimum number of genes for the corresponding term > 2 were considered significant for a category.

$$P = \frac{\binom{a+b}{a}\binom{c+d}{c}}{\binom{n}{a+c}}$$

Where n was the number of background genes; a' was the gene number of one gene set in the gene lists; a' + b was the number of genes in the gene list including at least one gene set; a' + c was the gene number of one gene list in the background genes; a' was replaced with a = a'-1.

Results

Analyzing of disruptions in narcolepsy PPI network

We obtained 12493 genes of normal and narcolepsy after preprocessing, then investigated

Table 1. KEGG pathways of genes with interaction score changes > 0.6 between normal and narcolepsy

Pathways	*P* Value
Pathways in cancer	3.53E-22
Focal adhesion	1.37E-15
Cell cycle	2.11E-13
Melanoma	4.58E-13
Glioma	9.30E-13
T cell receptor signaling pathway	4.70E-12
Progesterone-mediated oocyte maturation	2.31E-11
Colorectal cancer	3.01E-11
Fc epsilon RI signaling pathway	6.54E-11
GnRH signaling pathway	7.62E-11
Chemokine signaling pathway	5.33E-10
Chronic myeloid leukemia	9.10E-10
Non-small cell lung cancer	1.13E-09
Renal cell carcinoma	3.68E-09
Neuroactive ligand-receptor interaction	3.79E-09
VEGF signaling pathway	3.81E-09
ErbB signaling pathway	1.00E-08
Pancreatic cancer	1.16E-08
Prostate cancer	2.61E-08
Regulation of actin cytoskeleton	2.88E-08
MAPK signaling pathway	5.76E-08
Gap junction	8.70E-08
Wnt signaling pathway	1.02E-07
Bladder cancer	1.21E-07
Endometrial cancer	1.75E-07
Small cell lung cancer	3.28E-07
ECM-receptor interaction	3.28E-07
Oocyte meiosis	5.24E-07
Melanogenesis	5.64E-07
TGF-beta signaling pathway	1.14E-06
Nucleotide excision repair	2.21E-06
Vascular smooth muscle contraction	2.75E-06
Acute myeloid leukemia	4.55E-06
Purine metabolism	5.58E-06
Long-term potentiation	1.10E-05
Insulin signaling pathway	1.30E-05
Long-term depression	1.64E-05
Neurotrophin signaling pathway	1.72E-05
B cell receptor signaling pathway	1.95E-05
Calcium signaling pathway	3.34E-05
Adherens junction	1.05E-04
Toll-like receptor signaling pathway	1.18E-04
mTOR signaling pathway	3.50E-04
Linoleic acid metabolism	3.57E-04
Jak-STAT signaling pathway	5.88E-04
Dilated cardiomyopathy	5.99E-04
Epithelial cell signaling in Helicobacter pylori infection	6.05E-04

intersections between these gene interaction and STRING PPI network, and identified PPI networks of normal and narcolepsy. The normal H_N and narcolepsy H_D PPI networks displayed equal numbers of nodes (6820) and interactions (41729). Their average scores (weights) were 0.387 and 0.318, respectively. **Figure 1** showed that there were significant differences in the score distribution (0~0.4) of the two networks. Examining these interactions more carefully, we found that scores of 17249 interactions in narcolepsy network were lower compared with normal network, and on the contrary 24480 interactions were higher than these of normal. We extracted those with score changes > 0.6, which included 2210 interactions.

KEGG pathway enrichment analysis of genes involved in these 2210 interactions was performed, and with thresholds of *P* value < 0.001, the results were shown in **Table 1**. These genes were enriched in 47 biological process terms, of which pathways in cancer (P = 3.53E-22), focal adhesion (P = 1.37E-15) and cell cycle (P = 2.11E-13) were the most significant three pathways.

Analyzing of disruptions in narcolepsy modules

Clique-merging algorithm was selected to identify disrupted or altered modules from the normal and narcolepsy PPI network in this paper. In detailed, we performed a comparative analysis between normal *S* and narcolepsy *T* modules to understand disruptions at the module level. Maximal cliques of normal and narcolepsy PPI network were 23208 based on fast depth-first algorithm. With the threshold of nodes > 4, a total of 7070 maximal cliques were identified for module analysis. As shown in **Table 2**, a number of 1800 and 2195 modules were explored in normal and narcolepsy PPI network. For narcolepsy module, average module size was smaller than that of normal module, while maximum

Table 2. Properties of normal and narcolepsy modules

Module set	Numbers of modules	Average module size	Correlation		
			Maximal	Average	Min
Normal *S*	1800	20.406	0.287	-0.075	0.033
Narcolepsy *T*	2195	18.353	0.319	-0.074	0.036

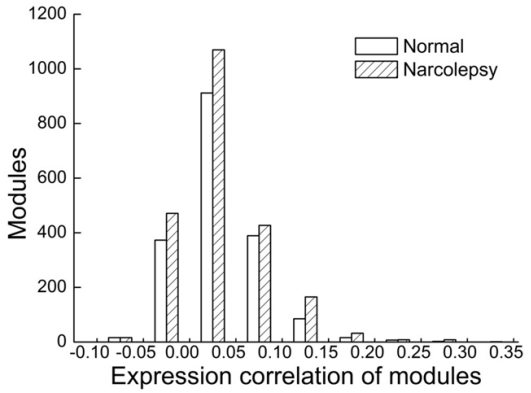

Figure 2. Correlationwise distribution of modules in normal and narcolepsy. Expression correlation of one module was equaled to weighted interaction density in the module.

Table 3. Correlations of matched normal and narcolepsy module pairs

Module pair subset	No. of pairs	Δ_c	
		Maximal	Average
$\gamma'(S, T)$	109	0.226	0.083
$\gamma''(S, T)$	126	0.152	0.074

weighted interaction density was larger than that of normal module. **Figure 2** showed the relationship between numbers of modules and weighted interaction density of modules. There were significant differences of module numbers when the interaction density arranged from -0.05 to 0.15, and modules in normal was higher.

But **Table 2** also showed an overall decrease in correlation in narcolepsy modules, further, this decrease had affected mainly the highly correlated modules (**Figure 2**). Next, we obtained 235 disrupted module pairs and then computed the set of matching modules $\gamma(S, T)$. Giving $|\gamma(S, T)| = 235$, for $t_J = 2/3$ and $\delta = 0.05$. Substantially, we divided $\gamma(S, T)$ into $\gamma'(S, T)$ of module pairs showing higher correlation in normal than narcolepsy, and $\gamma''(S, T)$ of module pairs showing lower correlation in normal than narcolepsy, giving $|\gamma'(S, T)| = 109$, $|\gamma''(S, T)| = 126$. We calculated the absolute differential correlation Δ_c of these subsets, as shown in **Table 3**. Interestingly, this demonstrated a marginal increase in correlation for 126 modules in narcolepsy versus normal, with a maximum increase of 0.152. However, KEGG pathway analysis showed enrichment for similar terms in both $\gamma'(S, T)$ and $\gamma''(S, T)$, which was not specific enough to differentiate the roles of the two subsets and, therefore, whether compensatory or narcolepsy-driving mechanisms coming into play. This prompted further in-depth analysis of the modules.

In-depth analyses of disrupted modules

Of the 235 disrupted module pairs in $\gamma(S, T)$, pathway analysis based on disrupted module gene composition were conducted, **Table 4** showed pathways within threshold P value < 0.001. These genes were enriched in 59 terms, of which oxidative phosphorylation (P = 1.93E-23), T cell receptor signaling pathway (P = 8.44E-21), cell cycle (P = 3.39E-20), Alzheimer's disease (P = 4.13E-17) and focal adhesion (P = 6.89E-17) were the most significant five disrupted pathways. This was consistent with current knowledge on narcolepsy which included Alzheimer's disease, Huntington's disease, Parkinson's disease, glioma and neurotrophin signaling pathway.

In detailed, there were a total of 651 genes presented in normal modules and 643 genes consisted in narcolepsy modules. Among these genes, 198 genes only existed in altered normal modules, thus these genes were gained genes compared with normal, for example, *TOP2A* and *CCNA2* were gained 14 and 10 modules respectively. Whereas 182 genes only consisted in altered narcolepsy modules, which meant 182 genes were missed in the narcolepsy process. There were three genes (*FGD1*, *ARHGEF18* and *MELK*) had missed more than 10 modules. Being consistent with previous enrichment analysis, many genes in gained or missed gene list had been identified in independent wet-lab studies and related to neuro-associated disease, particularly *FGD1* and *ARHGEF18* [29, 30]. For gained genes, cell

Table 4. Pathways based on genes in disrupted normal and narcolepsy modules

Term	*P* Value
Oxidative phosphorylation	1.93E-23
T cell receptor signaling pathway	8.44E-21
Cell cycle	3.39E-20
Alzheimer's disease	4.13E-17
Focal adhesion	6.98E-17
Pathways in cancer	1.44E-15
Huntington's disease	1.84E-15
Prostate cancer	2.04E-15
Long-term depression	2.04E-14
B cell receptor signaling pathway	5.24E-14
Colorectal cancer	6.57E-14
Fc epsilon RI signaling pathway	1.12E-13
Parkinson's disease	1.39E-13
Chemokine signaling pathway	1.95E-13
Chronic myeloid leukemia	3.66E-13
Small cell lung cancer	4.19E-13
Glioma	4.49E-13
VEGF signaling pathway	2.42E-12
Ribosome	8.04E-12
Neurotrophin signaling pathway	1.90E-11
Pancreatic cancer	2.72E-11
Endometrial cancer	4.61E-11
GnRH signaling pathway	6.93E-11
Progesterone-mediated oocyte maturation	1.62E-10
ECM-receptor interaction	4.20E-10
Non-small cell lung cancer	8.62E-10
Linoleic acid metabolism	1.53E-09
Oocyte meiosis	2.33E-09
Acute myeloid leukemia	4.66E-09
Regulation of actin cytoskeleton	7.89E-09
Toll-like receptor signaling pathway	1.46E-08
ErbB signaling pathway	2.34E-08
MAPK signaling pathway	5.84E-08
NOD-like receptor signaling pathway	1.09E-07
Renal cell carcinoma	2.82E-07
Epithelial cell signaling in Helicobacter pylori infection	6.94E-07
Fc gamma R-mediated phagocytosis	7.22E-07
DNA replication	1.12E-06
TGF-beta signaling pathway	1.50E-06
Phosphatidylinositol signaling system	3.45E-06
Wnt signaling pathway	3.46E-06
Inositol phosphate metabolism	4.48E-06
Melanoma	6.30E-06
Leukocyte transendothelial migration	7.19E-06
Arachidonic acid metabolism	8.02E-06
Nucleotide excision repair	2.08E-05

cycle (P = 2.01E-07), long-term depression (P = 2.51E-07) and focal adhesion (P = 3.81E-07) were the most significant 3 of 16 pathways. As for missed genes, top 3 significant terms were prostate cancer (P = 8.46E-12), Fc epsilon RI signaling pathway (P = 9.34E-12) and T cell receptor signaling pathway (P = 2.54E-10).

Discussion

The objective of this work is to identify disrupted pathways in narcolepsy according to systematically tracking the dysregulated modules of reweighted PPI networks. We obtained reweighted normal and narcolepsy PPI network based on PCC and then identified modules in the PPI network. By comparing these modules, 235 dysregulated modules pairs were identified, further, pathway enrichment analyses were conducted utilizing genes in dysregulated modules. The results displayed 59 disrupted pathways, such as T cell receptor signaling pathway, cell cycle and Alzheimer's disease.

A recent genome-wide association study found a strong association between narcolepsy and a T cell receptor gene variant, corroborating the autoimmune hypothesis [7]. In this study, we found that T cell receptor signaling pathway was a disrupted pathway in narcolepsy. Literatures documented that T cell receptor was the major receptor for HLA-peptide presentation, and narcolepsy might be triggered by an acute autoimmune process together with HLA and T cell receptor associations [31]. Chen et al discovered that narcolepsy, a immune-mediated disease, was associated with polymorphisms of the genes encoding T cell receptor alpha chain [32]. Immune-mediated destruction of hypocretin producing neurons may be mediated by microglia/macrophages that become activated either by autoantigen specific CD^{4+} T cells or superantigen stimulated CD^{8+} T cells,

alpha-Linolenic acid metabolism	2.87E-05
Insulin signaling pathway	4.14E-05
Melanogenesis	5.64E-05
Ether lipid metabolism	1.32E-04
Vascular smooth muscle contraction	1.67E-04
Apoptosis	1.74E-04
Alanine, aspartate and glutamate metabolism	1.78E-04
Gap junction	2.45E-04
Retinol metabolism	2.90E-04
Vibrio cholerae infection	4.47E-04
Proteasome	7.64E-04
RIG-I-like receptor signaling pathway	7.88E-04
Bladder cancer	8.83E-04

or independent of T cells, which might lead to the release of neurotoxic molecules and cause selective destruction of hypocretin neurons in the hypothalamus, finally induced narcolepsy [33]. According to our analysis of gene expression profiles, we had also discovered that T cell receptor signaling pathway played an important role in narcolepsy.

Cell cycle is the series of events that take place in a cell leading to its division and duplication, and a dysregulation of the cell cycle components. The cell-division cycle is a vital process by which hair, skin, blood cells, and some internal organs are renewed [34]. Narcolepsy is a disabling sleep disorder and has close relationship with neurons cell. Neurons containing the neuropeptide orexin (hypocretin) were located exclusively in the lateral hypothalamus and send axons to numerous regions throughout the central nervous system, including the major nuclei implicated in sleep regulation [35]. In addition, purinergic signaling played a fundamental role in immune regulation, modulating proliferation, and apoptosis. *P2RY11* expression was widespread, stimulation of *P2RY11* which was one gene marker of narcolepsy delayed cell apoptosis [9, 36, 37]. Therefore, we could conclude that narcolepsy was related to cell cycle tightly.

Alzheimer's disease is a chronic neurodegenerative disease that usually starts slowly and gets worse over time and is characterized by loss of neurons and synapses in the cerebral cortex and certain subcortical regions [38]. The cause of Alzheimer's disease was similar to narcolepsy. Besides, more recent genome wide association studies had found 19 areas in genes that appeared to affect the risk of Alzheimer's disease, for example, *HLA-DRB5*, *CASS4* and *CELF1* [39]. Early studies using low-resolution serological techniques found that narcolepsy was associated with two HLA class II antigens, *DR2* and *DQ1* [40]. The two diseases both had close relationship with *HLA-DR*, and there might be potential linkage between narcolepsy and Alzheimer's disease. Besides, *HLA* family had been proved to be correlated with T cell, and we had discussed that T cell receptor signaling pathway played an important role in narcolepsy.

In conclusion, we successfully identified disrupted pathways, such as T cell receptor signaling pathway, cell cycle and Alzheimer's disease, and these pathways might be potential biological processes for treatment and etiology mechanism in narcolepsy.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Guanzhen Li, Shandong Sleep Medical Center, Provincial Hospital Affiliated to Shandong University, Jingwuweiqi Road, Jinan 250021, China. Tel: 860531-68772077; Fax: 860531-68772071; E-mail: guanzhenlimed@yeah.net

References

[1] Black J, Reaven NL, Funk SE, McGaughey K, Ohayon M, Guilleminault C and Ruoff C. High Rates of Medical Comorbidity in Narcolepsy: Findings from the Burden of Narcolepsy Disease (BOND) Study of 9,312 Patients in the United States. 2013.

[2] Goldbart A, Peppard P, Finn L, Ruoff C, Barnet J, Young T and Mignot E. Narcolepsy and Predictors of Positive MSLTs in the Wisconsin Sleep Cohort. Sleep 2014; 37: 1043.

[3] Kandel ER, Schwartz JH and Jessell TM. Principles of neural science. McGraw-Hill New York 2000.

[4] Silber MH, Krahn LE, Olson EJ and Pankratz VS. The epidemiology of narcolepsy in Olmsted County, Minnesota: a population-based study. Sleep-New York 2002; 25: 197-204.

[5] Li J, Dong X, Yan H, Liu Y, An P, Zhao L, Li Q, Zhang X, Gao Z and Han F. [Positivity analysis of human leukocyte histocompatibility antigen-DQB1* 0602 allele in Chinese patients with narcolepsy]. Zhonghua Yi Xue Za Zhi 2014; 94: 763-765.

[6] Manzotte T, Guindalini C, Mazzotti DR, Palombini L, Souza AL, Poyares D, Bittencourt LR and Tufik S. The human leucocyte antigen DQB1* 0602 allele is associated with electroencephelograph differences in individuals with obstructive sleep apnoea syndrome. J Sleep Res 2013; 22: 217-222.

[7] Hallmayer J, Faraco J, Lin L, Hesselson S, Winkelmann J, Kawashima M, Mayer G, Plazzi G, Nevsimalova S and Bourgin P. Narcolepsy is strongly associated with the T-cell receptor alpha locus. Nat Genet 2009; 41: 708-711.

[8] Faraco J, Lin L, Kornum BR, Kenny EE, Trynka G, Einen M, Rico TJ, Lichtner P, Dauvilliers Y and Arnulf I. ImmunoChip study implicates antigen presentation to T cells in narcolepsy. PLoS Genet 2013; 9: e1003270.

[9] Kornum BR, Kawashima M, Faraco J, Lin L, Rico TJ, Hesselson S, Axtell RC, Kuipers H, Weiner K and Hamacher A. Common variants in P2RY11 are associated with narcolepsy. Nat Genet 2010; 43: 66-71.

[10] Singh AK, Mahlios J and Mignot E. Genetic association, seasonal infections and autoimmune basis of narcolepsy. J Autoimmun 2013; 43: 26-31.

[11] Jordán F, Nguyen TP and Liu WC. Studying protein-protein interaction networks: a systems view on diseases. Brief Funct Genomics 2012; 11: 497-504.

[12] Zhao J, Zhang S, Wu LY and Zhang XS. Efficient methods for identifying mutated driver pathways in cancer. Bioinformatics 2012; 28: 2940-2947.

[13] Wu C, Zhu J and Zhang X. Integrating gene expression and protein-protein interaction network to prioritize cancer-associated genes. BMC Bioinformatics 2012; 13: 182.

[14] Liu G, Li J and Wong L. Assessing and predicting protein interactions using both local and global network topological metrics. Genome Informatics 2008; 22: 138-149.

[15] Chu LH and Chen BS. Construction of a cancer-perturbed protein-protein interaction network for discovery of apoptosis drug targets. BMC Syst Biol 2008; 2: 56.

[16] Magger O, Waldman YY, Ruppin E and Sharan R. Enhancing the prioritization of disease-causing genes through tissue specific protein interaction networks. PLoS Comput Biol 2012; 8: e1002690.

[17] Srihari S and Ragan MA. Systematic tracking of dysregulated modules identifies novel genes in cancer. Bioinformatics 2013; 1-9.

[18] Jensen LJ, Kuhn M, Stark M, Chaffron S, Creevey C, Muller J, Doerks T, Julien P, Roth A and Simonovic M. STRING 8-a global view on proteins and their functional interactions in 630 organisms. Nucleic Acids Res 2009; 37: D412-D416.

[19] Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B and Speed TP. Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 2003; 31: e15-e15.

[20] Bolstad BM, Irizarry RA, Astrand M and Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 2003; 19: 185-193.

[21] Bolstad B. Affy: Built-in Processing Methods. 2013.

[22] Nahler G. Pearson correlation coefficient. Dictionary of Pharmaceutical Medicine 2009; 132-132.

[23] Liu G, Wong L and Chua HN. Complex discovery from weighted PPI networks. Bioinformatics 2009; 25: 1891-1897.

[24] Srihari S and Leong HW. A survey of computational methods for protein complex prediction from protein interaction networks. J Bioinf Comput Biol 2013; 11: 1-27.

[25] Tomita E, Tanaka A and Takahashi H. The worst-case time complexity for generating all maximal cliques and computational experiments. Theor Comput Sci 2006; 363: 28-42.

[26] Gabow HN. An efficient implementation of Edmonds' algorithm for maximum matching on graphs. J ACM 1976; 23: 221-234.

[27] Huang DW, Sherman BT and Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2008; 4: 44-57.

[28] Ford G, Xu Z, Gates A, Jiang J and Ford BD. Expression Analysis Systematic Explorer (EASE) analysis reveals differential gene expression in permanent and transient focal stroke rat models. Brain Res 2006; 1071: 226-236.

[29] Satoh JI, Kawana N and Yamamoto Y. pathway Analysis of chIp-seq-Based nRF1 Target Genes suggests a Logical Hypothesis of their Involvement in the pathogenesis of neurodegenerative Diseases. Gene Regulat Syst Biol 2013; 7: 139.

[30] Arivazhagan A, Kumar DM, Sagar V, Patric IRP, Sridevi S, Thota B, Srividya MR, Prasanna K, Thennarasu K and Mondal N. Higher topoisomerase 2 alpha gene transcript levels predict better prognosis in GBM patients receiving temozolomide chemotherapy: identification of temozolomide as a TOP2A inhibitor. J Neuro-Oncol 2012; 107: 289-297.

[31] Cvetkovic LV, Bayer L, Dorsaz S, Maret S, Pradervand S, Dauvilliers Y, Lecendreux M, Lammers GJ, Donjacour CE and Du Pasquier

RA. Elevated Tribbles homolog 2-specific antibody levels in narcolepsy patients. J Clin Invest 2010; 120: 713-719.
[32] Chen Y, Shi Y, Cheng H, An YQ and Gao GF. Structural immunology and crystallography help immunologists see the immune system in action: how T and NK cells touch their ligands. IUBMB life 2009; 61: 579-590.
[33] Fontana A, Gast H, Reith W, Recher M, Birchler T and Bassetti CL. Narcolepsy: autoimmunity, effector T cell activation due to infection, or T cell independent, major histocompatibility complex class II induced neuronal loss? Brain 2010.
[34] George M. Cells: Building Blocks of Life. The Creative Company, 2002.
[35] Chemelli RM, Willie JT, Sinton CM, Elmquist JK, Scammell T, Lee C, Richardson JA, Williams SC, Xiong Y and Kisanuki Y. Narcolepsy in orexin Knockout Mice: Molecular Genetics of Sleep Regulation. Cell 1999; 98: 437-451.
[36] Vaughan KR, Stokes L, Prince LR, Marriott HM, Meis S, Kassack MU, Bingle CD, Sabroe I, Surprenant A and Whyte MK. Inhibition of neutrophil apoptosis by ATP is mediated by the P2Y11 receptor. J Immunol 2007; 179: 8544-8553.
[37] Di Virgilio F, Boeynaems JM and Robson SC. Extracellular nucleotides as negative modulators of immunity. Curr Opin Pharmacol 2009; 9: 507-513.
[38] Alzheimer's Association. 2012 Alzheimer's disease facts and figures. Alzheimers Dement 2012; 8: 131-68.
[39] Lambert JC, Ibrahim CA, Harold D, Naj AC, Sims R, Bellenguez C, Jun G, DeStefano AL, Bis JC and Beecham GW. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet 2013; 45: 1452-8.
[40] Scammell TE. The neurobiology, diagnosis, and treatment of narcolepsy. Ann Neurol 2003; 53: 154-166.