European Review for Medical and Pharmacological Sciences 2015; 19: 4811-4826

Dysregulated module approach identifies disrupted genes and pathways associated with acute myelocytic leukemia

Y. FANG[1], L.-N. XIE[1], X.-M. LIU[1], Z. YU[1], F.-S. KONG[1], N.-X. SONG[1], F. ZHOU[1]

[1]Department of Hematology, Jinan Military General Hospital, Jinan, China

Abstract. – OBJECTIVE: To identify disrupted genes and pathways involved in acute myelocytic leukemia (AML) by systematically tracking the dysregulated modules across normal and AML conditions.

MATERIALS AND METHODS: In this study, we firstly integrated the protein interaction data and expression profiles to infer and reweight the normal and AML networks using Pearson correlation coefficient (PCC). Next, clustering-based on maximal cliques (CMC) approach and a maximum weight bipartite matching method were implemented to infer the condition-specific modules and capture the disturbed modules, respectively, from two conditional networks. Then, the gene compositions and functional enrichment analysis were performed to identify the dysregulated genes and pathways. Finally, reverse transcription polymerase chain reaction (RT-PCR) was implemented to study the expression level of several key genes in AML patients.

RESULTS: In two conditional-specific networks, universal changes of gene correlations were revealed, making the differential correlation density among disrupted module pairs. In this work, a total of 84 altered modules were identified by comparing modules in normal and AML networks. Functional enrichment analysis showed that genes in altered modules mainly involved in cell cycle, nucleic acids and cancer signaling process, and differentially expressed genes (DEGs) and changed gene correlations were mainly participated in natural killer cell-mediated cytotoxicity and acute myeloid leukemia pathway. The key genes, such as MYC, EGFR, MAPK1 and CCNA1, were all significantly differentially expressed in AML patients.

CONCLUSIONS: This module approach effectively identifies dysregulated pathways and genes associated with AML. The considerable differences of gene correlations yield to these dysfunctional modules, and the coordinated disruption of these very modules contributes to leukemogenesis.

Key words:
Acute myelocytic leukemia, Protein-protein interaction network, Modules, Pathway.

Introduction

In biological cell, proteins generally interact with other proteins to accomplish particular cellular tasks rather than act alone[1]. Multiprotein complexes (modules or pathways) form fundamental units responsible for driving key biological processes within cells[2]. The disruptions of these functional modules might contribute to the oncogenesis, thus, studying the module behavior cross different conditions is important for understudying principles of functional dysregulation. Also, the compositions of modules will vary according to cellular requirements. Analyzing the gene components of modules provides insight into how the ensemble of expressed proteins is organized into functional modules.

With the advance of high-throughput techniques in molecular biology, significant amount of protein interactions have been produced, making it possible to systematically mine modules from protein-protein interactions (PPI) network. Several computational algorithms based on graph clustering, dense region finding and clique finding have been developed for the prediction of modules from PPI networks, and the comparison of them are also presented[3,4]. Bader et al[5] proposed MCODE method, a graph clustering algorithm based solely on PPI network, to detect modules based on the proteins' connectivity values in the PPI network. Several methods predict modules by incorporating gene expres-

Corresponding Author: Fang Zhou, MD; e-mail: fangzhoumed@yeah.net

sion data or functional information with PPI data, such as CMC[6] and STM[7]. To reveal the pathogenesis, apart from predicting modules, it is necessary to identify the disrupted modules from condition-specific networks. Srihari and Ragan[8] provided a straightforward yet systematic method to identify and compare modules by integrating clique-merging and maximum weight bipartite matching model, performing well for assessing the reliability of protein interactions. Leukemia is a malignant cancer of the blood-forming organs, characterized by distorted proliferation and development of leukocytes and their precursors[9]. Acute myelocytic leukemia (AML) is one of the most common and deadly forms of hematopoietic malignancies through genetic alteration and pathways deregulation[10]. Also, distinct phenotypes of AML patients present different clinical outcome[11]. Stirewalt et al[12] revealed seven over-expressed AML-related genes by comparing expression profiles between AML and normal controls, and considered these genes might be potential therapeutic targets. Beghini et al[13] indicated that the regeneration-associated ligand-dependent Wnt signaling module was activated in AML cells. Previous studies have revealed several AML-related molecular or functional unites[14,15]. It is necessary to track disrupted gene and module behavior across specific conditions using a systematic method in a controlled manner.

Therefore, in order to identify the disrupted genes and pathways associated with AML in the present study, we performed a systematically tracking of altered modules across normal and AML conditions. To achieve this, we firstly integrated the PPI data and expression profiles to infer the normal- and AML-specific PPI network based on Pearson correlation coefficient (PCC); then, we inferred the condition-specific modules from two conditional networks using clustering-based on maximal cliques (CMC) approach, and captured the disturbed modules by a comparative study. Moreover, differential expression and gene composition analysis in altered modules were implemented, and functional enrichment analysis of active genes in altered modules and differentially expressed genes (DEGs) were carried out to identify the dysregulated pathways. Finally, reverse transcription polymerase chain reaction (RT-PCR) was implemented to study the expression level of several key genes. This study could effectively identify the disruption of AML-related modules and genes, and provide an alternative and compensatory way to counter genome destabilizing agents.

Materials and Methods

Generic PPI Data

In this section, we constructed the human generic PPI network based on the search tool for the retrieval of interacting genes/proteins (STRING) database (http://string-db.org)[14]. STRING is a carefully curated database which provides a comprehensive, quality-controlled collection of gene/protein associations for a large number of organisms with a global perspective[14]. It comprises several type of gene associations, such as gene co-expression, experimentally derived evidence, gene neighborhood and curated pathway databases[15]. STRING could assess and integrate all these data to gain a confidence score for all gene/protein interactions. Each interaction is given a weight value (between 0 and 1) reflecting the strength and reliability of the interaction. The human generic PPI data with combined scores are then obtained, in which there are 1048576 interactions.

Expression data recruitment and preprocessing

In the present study, a search on transcription profiling data between AML patients and healthy donors was conducted. Three microarray datasets of AML deposited in ArrayExpress database (http://www.ebi.ac.uk/arrayexpress/) were recruited under the accession number of E-GEOD-9476[12], E-GEOD-37307, and E-MTAB-220[13]. From these datasets, a total of 89 AML samples and 67 normal controls were collected in the current study. The characteristics of these three studies were shown in Table I.

Prior to analysis, the original information of each dataset from all conditions was subjected to data preprocessing. First, the annotations for probes were obtained from the manufacturer's documents. Next, background correction and normalization were conducted by robust multi-chip average (RMA) method[16] and quantiles based algorithm[17], respectively. The perfect match and mismatch values were revised via micro array suite 5.0 (MAS 5.0) algorithm[18]. Then, the data were screened using the feature filter method. Each probe was mapped to one gene, and the probe was discarded if it could not match

Table I. Characteristics of the studies related to acute myeloid leukemia in our study.

Accession number	Year	Sample size	Gene number	Platform
		Total (cases/controls)		
E-GEOD-9476	2008	64 (26/38)	12493	Affymetrix HG-U133A
E-GEOD-37307	2012	49 (30/19)	12493	Affymetrix HG-U133A
E-MTAB-220	2012	43 (33/10)	20102	Affymetrix HG-U133Plus2

any genes. After data preprocessing, a total of 12493, 12493, and 20102 genes were included in these three dataset, respectively. Finally, the distance-weighted discrimination (DWD) method[19] employed inSilicoMerging package (Brussels, Belgium. Support at https://insilicodb.com/) was implemented to eliminate undesired batch effects in the gene expression values and merge these three studies into one prior to further analysis.

Inferring AML and normal PPI networks

Prior to inferring condition-specific networks, original PPI network was performed data preprocessing. Protein identifiers in STRING were matched up with gene symbols using the getBM algorithm of the biomaRt R package[20] (an open source support at http://www.biomart.org/). Each identifier was mapped to one gene. The protein identifier would be discarded if it could not match up with any gene symbol. To minimize false-positives, only high-scoring interactions (score value ≥ 0.8) were remained to result PPI network which comprised 8590 nodes and 53975 interactions. Mapping the genes from microarray data to this network, the final PPI network involved in AML was constructed.

By integrating expression data with the resulted PPI network, the two tissue condition-specific PPI networks were constructed as follows. The correlations of gene pairs were identified using PCC, a well-established measure of correlation between two variables, giving a value range of +1 (perfect positive correlation) to −1 (perfect negative correlation) with 0 denoting the absence of a relationship[21]. We first calculated the PCC of gene interactions based on gene expression values in different conditions (AML and normal). The PCC of a pair of genes X and Y (encoding the corresponding paired proteins u and v) 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 represents the number of samples in the expression data, g(X,i) or g (Y,i) represents the gene expression level in the sample i under a specific condition, $\bar{g}(X)$ or $\bar{g}(Y)$ is the mean expression level of gene X or Y, and σ(X) or σ(Y) is the standard deviation of expression level of gene X or Y.

Then, the correlation wise frequency distribution was computed for interacting gene pairs in two conditions, and the expression landscapes through the correlation wise frequency distribution of gene pairs were also captured. Finally, the two conditional PPI networks were inferred by reweighting the interactions in the generic PPI network based the gene correlations.

Inferring AML and normal modules

Based on the AML and normal PPI networks, we inferred the condition-specific modules. In the present study, the CMC approach, proposed by Liu et al[6], was implemented to extract modules from the two conditional networks. The CMC approach first extracts all maximal cliques using the cliques mining algorithm[22] from the weighted PPI network. In the present study, the cliques with protein number ≥ 4 were screened to perform further analysis.

Next, CMC assigns a score (weighted density) to each clique C, whose score is calculated as:

$$\text{score}(C) = \frac{\sum_{u\in C, v\in C} p(u,v)}{|C|\cdot(|C|-1)}$$

, where $p(u,v)$ is the weight of the interaction between u and v calculated using PCC method. All cliques were ranked in descending order based on their score.

Then, we repeatedly screen the highly overlapping cliques which were defined as the overlap value of two cliques was greater than a predefined overlap-threshold T_o. The procedure was that, for every clique C_x, we looked for a clique C_y that the score of C_x > C_y and the overlap $|C_x \cap C_y|/|C_y| \geq T_o$. If such a C_y was found, CMC calculate the inter-connectivity between the non-overlapping proteins of C_x and C_y to determine

whether merging these two cliques as a bigger one or simply removing the one with a lower score C_y.

$$interscore(C_x,C_y)=\sqrt{\frac{\sum_{u\in(C_x-C_y)}\sum_{v\in C_y}p(u,v)}{|C_x-C_y|\cdot|C_y|}\cdot\frac{\sum_{u\in(C_y-C_x)}\sum_{v\in C_x}p(u,v)}{|C_y-C_x|\cdot|C_x|}}$$

If the interscore (C_x,C_y) was greater than a predefined merge-threshold T_m, then C_y was merged into C_x to form a module, otherwise C_y was removed.

In this study, the predefined thresholds T_o = 0.5, and T_m = 0.25. Finally, the modules in different conditions (AML and normal) were captured from the weighted PPI networks respectively.

Capturing disturbed AML modules

In order to capture the disturbed modules, a comparative study was performed to study the alterations between AML and normal modules. Let M_N and M_T be the sets of modules from normal and AML conditions respectively. We first consider the normal condition. The score value (correlation density) for each module was computed as,

$$Score(M_x)=\frac{\sum_{u\in M_x,v\in M_x}PCC(u,v)}{|M_x|\cdot(|M_x|-1)},$$

, where $M_x \in M_N$.

Similarly, the score value (correlation density) for each module from AML condition was computed. After all module correlation densities in different conditions were computed, we embarked on module comparison to capture the disrupted modules pairs using maximum weight bipartite matching algorithm[23]. First, we used the Jaccard similarity, a simple and natural similarity index, to measure the degree of association between modules in two conditions. To screen altered modules, a high thresholds of $J(M_x,M_y)$ was set to ensure the module pairs had same gene composition or had lost (or gained) only a few genes. It may be expressed as follows,

$$J(M_x,M_y)=\frac{|M_x\cap M_y|}{|M_x\cup M_y|}$$

, where $M_x \in M_N$, and $M_y \in M_T$.

Then, we calculated the differential correlation density between M_x and M_y, defined as $\Delta(M_x,M_y)$ = |Score(M_x)-Score(M_y)|. Next, we identified the disrupted module pairs by the maximum weight bipartite matching, and ranked them in descending order based on the differential correlation density between M_x and M_y. In the present study, module pairs under the thresholds of $J(M_x,M_y) \geq 2/3$ and $\Delta(M_x,M_y) \geq 0.15$ were considered as disrupted modules.

Gene composition analysis in altered modules

Generally, the difference among altered modules in different conditions is due to the altered gene composition or gene interactions. Thus, we performed the gene composition analysis in altered modules. First, all genes in the disrupted module pairs were extracted. Next, compared with each module from the disrupted module pairs in normal condition, we screened the altered genes in the corresponding disrupted AML module, including missed genes (the genes appeared in normal modules but not in disease ones) and added genes (the genes appeared in disease modules but not in normal ones). Sometimes, some genes were missed (or added) in one condition, and also added (or missed) in another condition. Thus, these both added and missed genes in disrupted module pairs were also analyzed. Moreover, we calculated the missed or added frequency of each gene in disrupted disease modules.

Identifying DEGs between AML patients and normal subjects

It has been well confirmed that the propensity of many diseases can be reflected in the difference of gene expression levels[24]. In the present study, DEGs between AML patients and normal controls were identified based on the transcript profiles above mentioned. The expression differences were detected by the empirical Bayes method (F test) implemented in the linear models for microarray data (LIMMA) package and adjusted by the false discovery rate (FDR) method. An adjusted $p < 0.05$ and |logFoldChange| $\geq$ 2.0 were used to screen DEGs.

Pathway enrichment analysis

Further, pathway enrichment analysis was carried out to further investigate the biological functions of genes involved in AML, including genes in altered modules and DEGs between AML and normal subjects. Also, pathway analyses of active genes in altered module and gene interactions with score changes > 0.7 were also conducted. In this study, the significant enriched pathways were identified under the threshold value of $p < 0.05$ based on expression analysis systematic explored (EASE) test[25].

RT-PCR validation

In order to verify several genes identified from our study, a total of 12 peripheral blood samples from AML patients were obtained in our hospital

between September 2014 and March 2015, including 7 males and 5 females. Normal samples (n = 15) were recruit from health people undergoing physical examination in our hospital. Written informed consents were obtained from those who agreed to participate in this study before collection of tissue. The ethical clearance was approved by the Institutional Ethical Committee.

Total RNA was isolated from peripheral blood using the Trizol reagent (Invitrogen, Carlsbad, CA, USA) following the manufacturer's instructions. cDNA was synthesized using the AMV reverse transcriptase and oligo (dT18) primers (Invitrogen). Amplification of the reverse transcribed RNA was performed in a total volume of 20 μL containing 2 μL of diluted cDNA, 8 μL dNTPs, 1 μL of Taq DNA Polymerase High Fidelity (Invitrogen), 3 μL of each forward and reverse primer. The data were normalized to the β-actin reference gene. The primers for key genes and β-actin were as follows: MYC: F-CGACTACGACTCGGTGCAGCC, R-CCGAGAAGCCGCTCCACATAC; EGFR: F-CGGGACATAGTCAGCAGTGA, R-ACTGGTTGTGGCAGCAGTC; MAPK1: F-GCGCTACACCAACCTCTCGT, R-CACGGTGCAGAACGTTAGCTG; CCNA1: F-ATTCATTAAGTGAAATTGTGC, R-CTTCCATTCAGAAACTTATTG; β-actin: F-AAGTACTCCGTGTGGATCGG, R-TCAAGTTGGGGGACAAAAAG. Each reaction was performed in three replicate samples. RT-PCR products were separated through 1.5% agarose gels electrophoresis and analyzed by Quantity One software of gel imaging analyzer (Bio-Rad,, Hercules, CA, USA). All data were presented as mean ± standard deviation (SD), and statistical analysis was carried out using SPSS 19.0 software (SPSS, Chicago, IL, USA). The Student's *t*-test was performed to determine the statistical significance of differences between groups. A value of $p < 0.05$ was considered statistically significant.

Results

PPI and expression data

First, from the human genetic PPI data, the high-scoring interactions (score value ≥ 0.8) were screened, in which there were 8590 nodes and 53975 interactions. After data preprocessing and merging of three expression datasets, a total of 12493 genes were recruited in our study. Mapping the 12493 genes to the high-scoring PPI network, the final PPI network comprised 6820 nodes and 41729 interactions was constructed.

Disruptions in AML PPI network

By calculating the correlations of gene pairs in AML patients and normal subjects based on the PPC, the AML- and normal-specific PPI networks were inferred by re-weighting the interactions, respectively. The AML- and normal-specific PPI networks displayed equal numbers of nodes (6820) and interactions (41729). However, the correlation wise frequency distribution was different between two conditions (Figure 1A), and the average scores in AML (0.417) was slightly lower than that in normal condition (0.434). We further examined these interactions

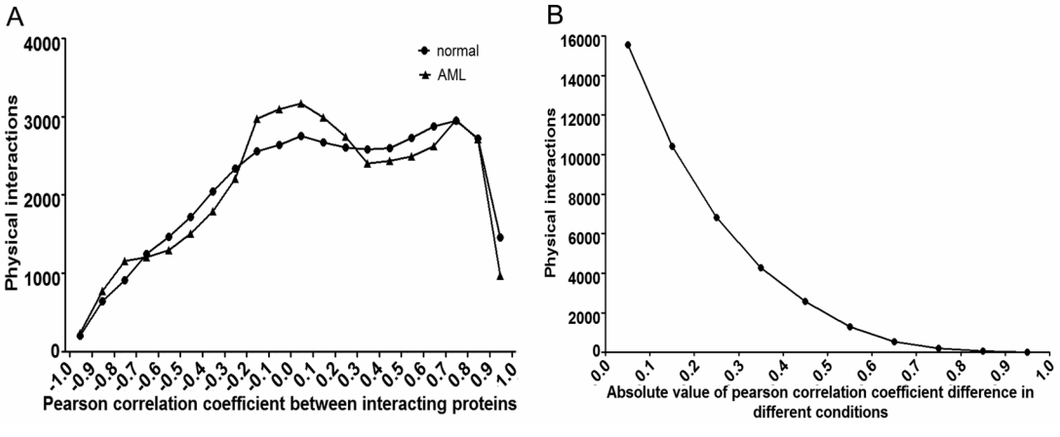

Figure 1. The correlation wise frequency distribution of gene physical interactions in normal- and AML-specific networks. **A**, the expression correlation wise distribution between interacting proteins; **B**, the distribution of changed expression correlation wise between two conditions.

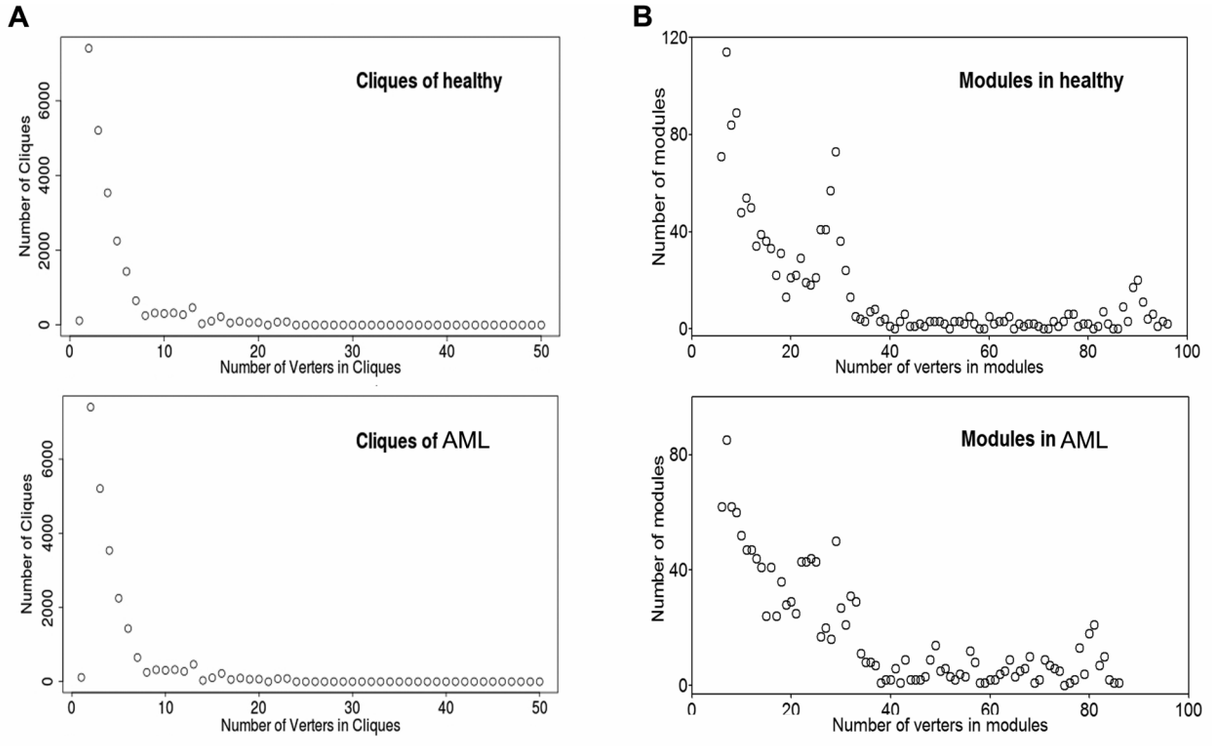

Figure 2. The frequency distribution of cliques and modules in normal and AML conditions. **A**, Distribution tendency of cliques; **B**, Distribution tendency of modules.

deeply and found that 21864 interactions showed lower scores and 19865 interactions presented higher scores in AML compared with normal network. Figure 1B showed the absolute value of PCC difference between AML and normal condition. Those gene interactions with score changes > 0.7 were extracted for further analysis, including 250 interactions and 352 genes.

Disrupted modules in AML

In the present study, in order to infer the disrupted modules, all maximal cliques were extracted from the AML and normal networks. After discarding the cliques with protein number < 4, a total of 7070 maximal cliques were screened in both AML and normal network. Using clique-merging algorithm, a total of 1379 and 1344 modules were identified from AML and normal networks, respectively. The frequency distributions of clique and module characteristics were shown in Figure 2A and 2B. Meanwhile, modules in different conditions showed different frequency distribution of module correlation density (Figure 3). As displayed in Table II, the average module size of AML modules (26.4) was bigger than that of normal modules (24.6), and the average module correlation density was slightly larger in AML (0.058) than that in normal condition (0.053). Under the thresholds of $J(M_x,M_y) \geq 2/3$ and $\Delta(M_x,M_y) \geq 0.15$, a total of 84 disrupted modules were obtained (Table S I). Notably, all these altered modules showed higher correlation in AML than that in normal, presenting a mean $\Delta(M_x,M_y)$ of 0.19.

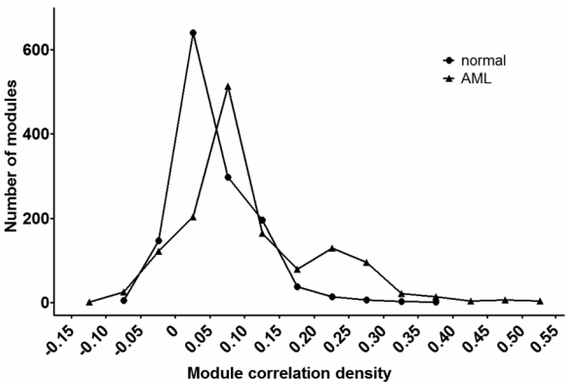

Figure 3. The distribution characteristics of module correlation density.

Table II. The characteristics of modules in acute myeloid leukemia and normal network

	Module size		Module correlation density	
	Normal	AML	Normal	AML
Maximal	96	86	0.366	0.312
Average	24.6	26.4	0.053	0.058
Minimum	6	6	-0.071	-0.057

Gene composition analysis in altered modules

The altered gene composition or gene interactions contribute to the difference among disrupted module pairs. Of the 84 disrupted module pairs, 10 module pairs had the same gene composition and the other ones showed the changed gene compositions. In these 10 module pairs with the same gene composition, the PCC values of gene interactions were changed in AML modules universally relative to the corresponding normal modules, making the differential correlation density among disrupted module pairs. Taking module pair M_{N429} and M_{T442} as an example, the module M_{N429} showed a weak correlation density of 0.219 in normal condition, indicating inactivity of the module. However, in tumor, the correlation of the corresponding disrupted module M_{T442} had strengthened to 0.542, presenting an increase of 0.323 (Figure 4). In detail, the gene correlations with *EGFR* showed most obvious increases, from -0.416, -0.053, -0.360 and 0.138 in normal condition to -0.045, 0.404, 0.153 and 0.143 in AML. This might be the main reason making the differential correlation density between disrupted module pair M_{N429} and M_{T442}.

In the 74 disrupted module pairs with changed gene compositions, the module pair M_{N431} and M_{T443} showed the highest differential correlation density ($\Delta = 0.331$). When analyzing the gene compositions of this disrupted module pair, an interesting swapping phenomenon was revealed. In M_{T443} a new gene *MET* replaced the existing gene *IGF1R*, forming new physical interactions with the remaining ones in tumor (Figure 5). Moreover, the module correlation density was increased from 0.214 in M_{N431} to 0.545 in M_{T443}. Also, the PCC values of gene interactions were changed universally between two conditions. Apparently, both changed gene correlations and gene swapping contributed to the differential correlation density in disrupted module pairs with changed gene compositions.

In the 84 disrupted module pairs, a total of 395 genes were extracted. Compared with each module from the disrupted module pairs in normal condition, we also screened the active genes in the corresponding disrupted AML modules, including 62 added genes and 103 missed genes. Meanwhile, a total of 15 genes were both added and missed genes in disrupted module pairs. Moreover, the missed or added frequency of each

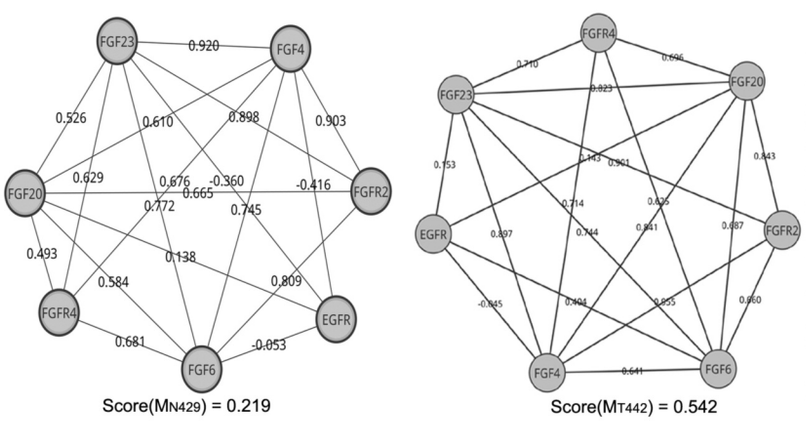

Figure 4. A disrupted module pair MN429 and MT442 with the same gene composition. MT442 presents an increased module correlation (0.323) relative to MT442. The gene correlations with EGFR shows obvious increases in AML relative to that in normal condition.

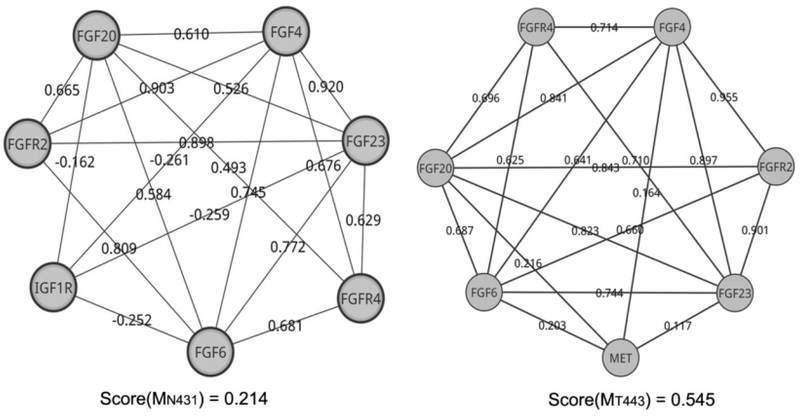

Figure 5. A disrupted module pair MN431 and MT443 with changes in gene composition. The module correlation density of MN431 was increased 0.331 relative to that of MT443. In the matched modules, a new gene MET replaced the existing gene IGF1R.

gene in disrupted disease modules was shown in Table S II. These genes swapping events in modules might be indicative of tumor disruptions and relevant to understand the development and progress involved in AML.

DEGs between AML patients and normal subjects

In the present study, the difference of gene expression levels between AML patients and normal subjects were also identified. Under the threshold value of $p < 0.05$ and $|\text{logFoldChange}| \geq 2.0$, a total of 385 DEGs were screened. By overlapping DEGs and genes in altered modules, we identified 21 DEGs in disrupted modules.

Pathway analysis

In order to further investigate the biological functions of genes involved in AML, pathway enrichment analysis was carried out for genes in altered modules and DEGs. Under the threshold value of $p < 0.05$, genes in altered modules significantly participated in 33 signaling pathways which mainly involved in cell cycle, nucleic acids and cancer signaling process (Table S III), and the most three significant pathways were cell cycle (p = 1.25E-47), nucleotide excision repair (p = 3.12E-25) and DNA replication (p = 1.16E-21). Similar to the result of genes in altered modules, the active genes in disrupted modules were also enriched in cell cycle, nucleotide excision repair and cancer-related pathways (Table S IV and Table S V).

The DEGs were significantly enriched in 4 signaling pathways, including hematopoietic cell lineage, natural killer cell mediated cytotoxicity, ribosome, acute myeloid leukemia. Those gene interactions with score changes > 0.7 were significantly involved in 46 pathway terms (Table SVI), of which neuroactive ligand-receptor interaction (p = 1.36E-10), chemokine signaling pathway (p = 3.76E-07) and acute myeloid leukemia pathway (p = 5.64E-07) were the most significant pathways. Moreover, natural killer cell mediated cytotoxicity and acute myeloid leukemia were the common pathways of DEGs and changed gene correlations.

Experimental validation

In the present study, the active genes (*MYC* and *EGFR*) and the DEGs (*MAPK1* and CCNA1) in the disrupted modules were selected to perform the expression validation by RT-PCR in AML patients. We found that the mRNA expression levels of *MAPK1*, *MYC* and *EGFR* in AML patients were significant higher than that in normal controls ($p < 0.05$); while the expression of *CCNA1* was significant lower in AML than that in normal ($p < 0.05$) (Figure 6). The differential expression of these genes might contribute the uncontrollable cellular process in hematopoietic system.

Discussion

Although recent researches into the genetics of AML have revealed several potential molecules involved in the prognostic factors and therapeutic targets, deeper study of AML pathogenesis is imperative. In the present study, a systematic identification of disrupted modules across AML and normal conditions was performed by integrating PPI and gene expression data. Based on the disrupted modules, the dysregulated pathways and genes were identified. We revealed the universally changed expres-

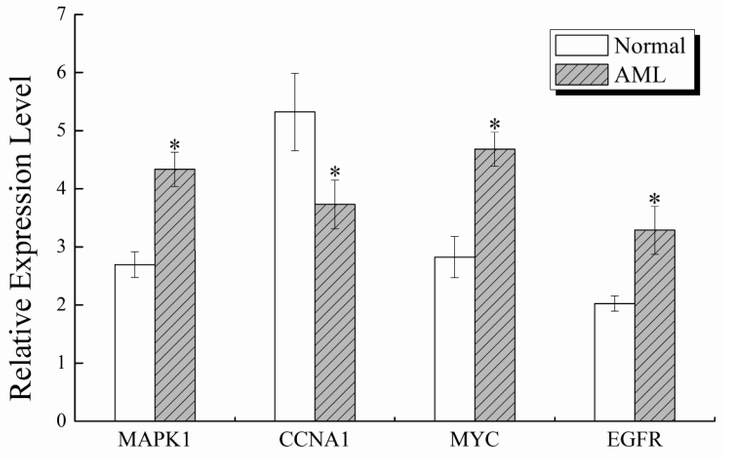

Figure 6. Relative expressions of dyeregulated genes in AML patients. * $p < 0.05$.

sion correlations and gene compositions in disrupted module pairs between two conditions.

Traditionally, the changes in individual gene expression permit identifying genes that may regulate tumor growth and differentiation. However, genes seldom accomplish their biological functions independently. Modules of physically interacting genes, as one of the fundamental functional units, are responsible for driving key biological mechanisms within the cell. Previous studies have provided several computational techniques to identify the highly connected clusters or modules in a network of protein interactions[2,3,26]. It has been indicated that the PI3K/PKB signaling module is related to the regulation of leukemia, and its activation is thought to correlate with poor prognosis and enhanced drug resistance[27]. Moreover, a study of Milella et al[28] revealed that disruption of the MEK/MAPK signal transduction module might represent an effective and relatively specific therapeutic strategy for AML. In brief, studying on pathogenesis and therapeutic target in module level is more efficient relative to that in the level of single gene.

In this work, from the AML- and normal-specific PPI networks, the correlations of gene pairs were changed widely, making the significant differential correlation density among disrupted module pairs. For example, the gene correlations with EGFR in disrupted module M_{T442} were higher than that in the corresponding normal module M_{N429}. Epidermal growth factor receptor (EGFR, for short), also known as ErbB-1, is a cell-surface receptor, and the over-expression or mutation of which has been associated with a number of cancers, including lung cancer[29], glioblastoma multiforme[30], and AML[31]. Also, several anticancer therapeutic drugs for EGFR inhibitors have yielded to cancer treatment[32]. Previous studies have indicated that EGFR mutations presented sensitivity to gefitinib in non-small cell lung carcinomas[33,34]. Several researches also reported the efficacy of EGFR-specific inhibitors in the therapy of AML[35-38]. Moreover, it has been indicated that the downstream signaling proteins of EGFR could initiate MAPK signal transduction pathway which has been reported as an effective target for AML[28,39]. In addition, gene swapping events existed in most disrupted module pairs, which might be indicative of tumor disruptions and relevant to understand the development and progress involved in AML.

Based on the disrupted genes involved in AML, functional enrichment analysis was conducted. The results showed genes in altered modules and active genes mainly participated in cell cycle, nucleic acids and cancer-related signaling process, and genes with changed expressions and correlations were significantly enriched in hematopoietic cell lineage, natural killer cell mediated cytotoxicity and acute myeloid leukemia pathway. Cell cycle is the series of events that take place in a cell leading to cell division and duplication. The dysregulation of the cell cycle components could cause the cell to multiply uncontrollably, further leading to tumor formation[40]. Cell cycle pathway and cell cycle-related molecular have been considered as the targets in cancer treatment[41,42]. Chim et al[43] indicated the important role of cell cycle dysregulation in oncogenesis of AML. Natural killer cell mediated cytotoxicity contributes to the innate immune response against numerous malignancies, including leukemias[44]. Natural killer cell treated with specific activating cytokines has been taken as an effective therapeutic strategy in AML treatment[45-47]. The other pathways identified in our study also showed close relationships with the de-

velopment and progress of AML, such as hematopoietic cell lineage and acute myeloid leukemia pathway, indicating the feasibility of this systematic method in this study.

Conclusions

This module approach effectively identifies dysregulated pathways and genes associated with AML by systematic tracking disrupted module pairs between two conditions, and works in alternative and compensatory ways to counter genome destabilizing agents. Obviously, the considerable differences of gene correlations yield to these dysfunctioning modules, and the coordinated disruption of these very modules contributes to leukemogenesis. The availability of specific pathways and genes involved in AML should be further explored in animal models and in human clinical trials.

Acknowledgements

This work was supported by Natural Science Foundation of Shandong Province, China (2013ZRB02007). Meanwhile, we would like to thank Ji'nan Evidence Based Medicine Science-Technology Center for the technical assistance and critical reading of the manuscript.

Conflict of Interest

The Authors declare that there are no conflicts of interest.

Table S I. The altered module pairs.

MN	MT	J	Δ	Score (M_N)	Score (M_T)
431	443	0.750	0.331	0.214	0.545
431	442	0.750	0.328	0.214	0.542
429	443	0.750	0.326	0.219	0.545
429	442	1.000	0.323	0.219	0.542
430	443	1.000	0.286	0.259	0.545
430	442	0.750	0.283	0.259	0.542
1189	1247	0.684	0.266	0.159	0.425
1035	1247	0.700	0.264	0.161	0.425
1152	1211	0.909	0.253	0.262	0.515
1133	1194	0.846	0.252	0.171	0.423
1191	1248	1.000	0.242	0.286	0.528
882	933	0.692	0.239	0.036	0.275
881	933	0.692	0.238	0.037	0.275
646	719	1.000	0.237	0.222	0.459
439	451	0.778	0.236	0.044	0.280
438	451	0.889	0.225	0.055	0.280
957	1007	1.000	0.225	0.079	0.304
456	468	0.700	0.225	0.095	0.320
1135	1197	0.944	0.213	0.173	0.386
1135	1198	0.944	0.213	0.173	0.386
1112	1087	0.696	0.207	0.129	0.336
1173	1227	0.750	0.206	0.087	0.293
1190	1247	1.000	0.205	0.220	0.425
968	1197	0.800	0.199	0.187	0.386
968	1198	0.800	0.199	0.187	0.386
1135	1196	1.000	0.198	0.173	0.371
1133	1195	0.923	0.197	0.171	0.368
1034	1246	0.818	0.191	0.136	0.327
1033	1246	0.818	0.187	0.140	0.327
719	848	0.742	0.184	0.110	0.294
968	1196	0.850	0.184	0.187	0.371
814	870	0.909	0.184	0.155	0.339
1112	1255	0.762	0.184	0.129	0.313
1136	1197	0.941	0.183	0.203	0.386
1136	1198	0.941	0.183	0.203	0.386
1137	1197	0.941	0.183	0.203	0.386
1137	1198	0.941	0.183	0.203	0.386
363	360	1.000	0.180	0.113	0.293
1038	1087	0.850	0.178	0.158	0.336
1034	1083	0.909	0.178	0.136	0.314
1135	1025	0.900	0.177	0.173	0.350
1208	1269	0.857	0.177	0.084	0.261
1033	1083	0.826	0.174	0.140	0.314
815	873	0.750	0.174	0.158	0.332
1048	1100	0.667	0.173	0.076	0.249
390	390	0.706	0.172	0.068	0.240
819	874	0.667	0.170	0.105	0.275
883	932	0.929	0.168	0.088	0.256
1136	1196	0.889	0.168	0.203	0.371
1137	1196	0.889	0.168	0.203	0.371
1189	1246	1.000	0.168	0.159	0.327
456	467	0.727	0.167	0.095	0.262
1035	1246	0.900	0.166	0.161	0.327
816	873	0.933	0.166	0.166	0.332
968	1025	0.950	0.163	0.187	0.350
756	848	0.852	0.163	0.131	0.294
387	390	0.706	0.162	0.078	0.240
1048	1102	0.667	0.162	0.076	0.238
389	390	0.706	0.162	0.078	0.240
596	671	0.950	0.161	0.179	0.340
397	401	0.941	0.161	0.219	0.380
904	955	0.667	0.160	-0.053	0.107
452	461	0.711	0.159	0.115	0.274
1102	1175	0.720	0.158	0.106	0.264
388	390	0.706	0.157	0.083	0.240
1132	848	0.704	0.156	0.138	0.294
427	440	0.875	0.156	0.184	0.340
834	884	0.857	0.156	0.165	0.321
1103	1175	0.692	0.155	0.109	0.264
838	885	0.857	0.155	0.198	0.353
1099	1175	0.760	0.155	0.109	0.264
1189	1083	0.900	0.155	0.159	0.314
815	872	0.929	0.155	0.158	0.313
1038	1255	0.842	0.155	0.158	0.313
1035	1083	1.000	0.153	0.161	0.314
355	409	0.676	0.153	0.113	0.266
355	410	0.676	0.153	0.113	0.266
608	684	0.727	0.153	0.110	0.263
357	409	0.676	0.152	0.114	0.266
357	410	0.676	0.152	0.114	0.266
354	353	0.895	0.152	0.153	0.305
847	894	0.667	0.152	0.019	0.171
837	885	0.857	0.152	0.201	0.353
1101	1175	0.731	0.151	0.113	0.264

Table S II. The active genes in disrupted module pairs

Added genes		Missed genes		Both added and missed genes		
Symbol	Time	Symbol	Time	Symbol	Missed time	Added time
IMP3	7	CLSPN	6	MYC	2	3
CENPA	4	RFC3	6	EGFR	1	2
KIF2C	4	RBL2	6	MET	1	2
CDKN3	4	RFC2	4	POLA1	4	1
AURKB	4	TIMELESS	4	RAD17	2	1
UBE2C	4	HUS1	4	IGF1R	2	1
KIF4A	4	TOPBP1	4	CCNB1	1	2
NCAPG2	4	RAD9A	4	SKP1	3	2
NEK2	4	RFC5	4	RB1	1	3
ESPL1	4	DSCC1	4	FBL	3	1
KCNA4	4	FANCE	4	DCAF13	3	1
KCNV2	4	RPA1	4	NOL10	4	2
KCNC2	4	RPA2	4	RBM28	4	1
HEATR1	4	TYMS	4	DDX27	2	3
RB1	3	FEN1	4	RPA3	1	1
DDX27	3	ORC6	4			
MYC	3	ORC2	4			
RBL1	3	ORC3	4			
MET	2	POLA2	4			
EGFR	2	POLA1	4			
TBL1XR1	2	ORC1	4			
XPC	2	POLE	4			
ELL	2	DBF4	4			
FBXL5	2	LIG1	4			
E2F3	2	RNASEH2A	4			
CCNB1	2	KCNA3	4			
SKP1	2	NOL10	4			
NOL10	2	KIAA0020	4			
RAD17	1	DDX24	4			
MAD2L1	1	RBM28	4			
IGF1R	1	CCND1	4			
CHUK	1	CCND2	4			
NDEL1	1	CCNA2	3			
CLIP1	1	SKP1	3			
PARP4	1	SSRP1	3			
MCM3	1	FBL	3			
RRM2	1	DCAF13	3			
POLA1	1	YWHAZ	2			
UBTF	1	RAD17	2			
PRKDC	1	IGF1R	2			
MCM2	1	TLR3	2			
FBXO4	1	CLASP2	2			
USP47	1	UBC	2			
EGLN1	1	SKP2	2			
TERF2IP	1	POLR2C	2			
SMC3	1	GTF2F1	2			
ACD	1	GNA15	2			
TFDP1	1	CHRM2	2			
TBP	1	EDN1	2			
RPA3	1	NVL	2			
POLR2E	1	DDX27	2			
POLR1C	1	RSL24D1	2			
GTF2H2	1	NOC3L	2			
GTF2H1	1	PES1	2			
ERCC3	1	PDCD11	2			
FBL	1	RPF1	2			
DCAF13	1	MCM7	2			
SMAD4	1	DES	2			

Continued

Table S II. The active genes in disrupted module pairs

Added genes		Missed genes		Both added and missed genes		
Symbol	Time	Symbol	Time	Symbol	Missed time	Added time
MED15	1	NEB	2			
PPARA	1	CDKN1B	2			
RBM28	1	MAX	2			
AP2A2	1	MYC	2			
		GINS4	2			
		KCNS1	1			
		KCNG2	1			
		KCNH6	1			
		KCND3	1			
		FGF3	1			
		EGFR	1			
		MET	1			
		CENPN	1			
		CLASP1	1			
		CENPE	1			
		NUSAP1	1			
		TTK	1			
		KIF11	1			
		BUB3	1			
		DLGAP5	1			
		CCNB1	1			
		KIF23	1			
		PAFAH1B1	1			
		KLHL21	1			
		FBXO5	1			
		CCNA1	1			
		POLR2I	1			
		POLR2D	1			
		POLR2L	1			
		ERCC6	1			
		POLR1D	1			
		MCM6	1			
		RB1	1			
		LIPC	1			
		CEL	1			
		PNLIPRP1	1			
		EPAS1	1			
		HIF3A	1			
		CXCR1	1			
		CXCR2	1			
		GCGR	1			
		RPA3	1			
		MYBPC2	1			
		TNNI3	1			
		PTRF	1			

Table S3. The significant pathways based on the genes in the disrupted modules.

ID	Term	Count	*p* value
hsa04110	Cell cycle	56	1.25E-47
hsa03420	Nucleotide excision repair	26	3.12E-25
hsa03030	DNA replication	22	1.16E-21
hsa05200	Pathways in cancer	51	5.55E-19
hsa05218	Melanoma	24	2.02E-16
hsa05215	Prostate cancer	21	4.33E-11
hsa05222	Small cell lung cancer	17	5.26E-08
hsa03430	Mismatch repair	10	6.11E-08
hsa04114	Oocyte meiosis	19	8.81E-08
hsa05219	Bladder cancer	12	2.31E-07
hsa05220	Chronic myeloid leukemia	15	4.84E-07
hsa04115	p53 signaling pathway	14	9.46E-07
hsa00240	Pyrimidine metabolism	16	1.77E-06
hsa05214	Glioma	13	2.62E-06
hsa04010	MAPK signaling pathway	27	4.96E-06
hsa05212	Pancreatic cancer	13	1.12E-05
hsa03020	RNA polymerase	8	5.53E-05
hsa05211	Renal cell carcinoma	11	2.35E-04
hsa03022	Basal transcription factors	8	2.51E-04
hsa04120	Ubiquitin mediated proteolysis	15	5.31E-04
hsa05223	Non-small cell lung cancer	9	7.79 E-03
hsa04810	Regulation of actin cytoskeleton	19	1.00 E-03
hsa00561	Glycerolipid metabolism	8	1.24 E-03
hsa04210	Apoptosis	11	1.37 E-03
hsa00230	Purine metabolism	15	1.59 E-03
hsa05210	Colorectal cancer	10	3.83 E-03
hsa04310	Wnt signaling pathway	14	3.96 E-03
hsa04623	Cytosolic DNA-sensing pathway	8	4.04 E-03
hsa04914	Progesterone-mediated oocyte maturation	10	4.49 E-03
hsa03410	Base excision repair	6	9.01 E-03
hsa04620	Toll-like receptor signaling pathway	10	1.26 E-02
hsa04350	TGF-beta signaling pathway	9	1.54 E-02
hsa05120	Epithelial cell signaling in Helicobacter pylori infection	7	4.05 E-02

Table S4. The significant pathways based on the genes in the disrupted modules.

ID	Term	Count	*p* value
hsa04110	Cell cycle	14	3.83E-13
hsa03420	Nucleotide excision repair	5	1.87E-04
hsa04114	Oocyte meiosis	6	7.30E-04
hsa05200	Pathways in cancer	9	1.14 E-03
hsa05218	Melanoma	5	1.18 E-03
hsa05212	Pancreatic cancer	5	1.24 E-03
hsa05220	Chronic myeloid leukemia	5	1.45 E-03
hsa03030	DNA replication	4	1.68 E-03
hsa05210	Colorectal cancer	5	2.20 E-03
hsa04350	TGF-beta signaling pathway	5	2.51 E-03
hsa05219	Bladder cancer	4	2.62 E-03
hsa05215	Prostate cancer	5	2.72 E-03
hsa05214	Glioma	4	8.25 E-03
hsa04520	Adherens junction	4	1.43 E-02
hsa05222	Small cell lung cancer	4	1.80 E-02
hsa03022	Basal transcription factors	3	2.25 E-02
hsa00240	Pyrimidine metabolism	4	2.49 E-02

Table S5. The significant pathways based on the missed genes in the disrupted modules.

ID	Term	Count	*p* value
hsa03030	DNA replication	14	2.47E-17
hsa04110	Cell cycle	17	2.26E-13
hsa03420	Nucleotide excision repair	9	2.24E-08
hsa03430	Mismatch repair	7	1.47E-07
hsa00240	Pyrimidine metabolism	9	9.81E-06
hsa03020	RNA polymerase	5	2.54E-04
hsa05200	Pathways in cancer	12	9.97E-04
hsa05218	Melanoma	6	1.19 E-03
hsa00230	Purine metabolism	8	1.61 E-03
hsa05222	Small cell lung cancer	6	2.52 E-03
hsa05219	Bladder cancer	4	1.18 E-02
hsa05210	Colorectal cancer	5	1.45 E-02
hsa05215	Prostate cancer	5	1.81 E-02
hsa05214	Glioma	4	3.45 E-02
hsa04114	Oocyte meiosis	5	3.60 E-02
hsa03440	Homologous recombination	3	4.00 E-02
hsa05120	Epithelial cell signaling in Helicobacter pylori infection	4	4.19 E-02

Table S6. The significant pathways based on the gene pairs with score change > 0.7

ID	Term	Count	*p* value
hsa04080	Neuroactive ligand-receptor interaction	40	1.36E-10
hsa04062	Chemokine signaling pathway	28	3.76E-07
hsa05221	Acute myeloid leukemia	15	5.64E-07
hsa05210	Colorectal cancer	17	2.66E-06
hsa05200	Pathways in cancer	37	3.94E-06
hsa04660	T cell receptor signaling pathway	19	4.69E-06
hsa04020	Calcium signaling pathway	25	4.90E-06
hsa04270	Vascular smooth muscle contraction	18	3.11E-05
hsa04370	VEGF signaling pathway	14	6.68E-05
hsa04912	GnRH signaling pathway	16	8.30E-05
hsa00591	Linoleic acid metabolism	8	3.34E-04
hsa04510	Focal adhesion	23	3.55E-04
hsa04810	Regulation of actin cytoskeleton	24	3.64E-04
hsa04664	Fc epsilon RI signaling pathway	13	4.08E-04
hsa04730	Long-term depression	12	5.19E-04
hsa05212	Pancreatic cancer	12	7.53E-04
hsa04916	Melanogenesis	14	1.12 E-03
hsa05218	Melanoma	11	2.45 E-03
hsa00980	Metabolism of xenobiotics by cytochrome P450	10	2.58 E-03
hsa04310	Wnt signaling pathway	17	3.16 E-03
hsa00982	Drug metabolism	10	3.24 E-03
hsa05213	Endometrial cancer	9	3.76 E-03
hsa04620	Toll-like receptor signaling pathway	13	4.05 E-03
hsa04630	Jak-STAT signaling pathway	17	4.11 E-03
hsa05215	Prostate cancer	12	4.31 E-03
hsa05211	Renal cell carcinoma	10	7.35 E-03
hsa04722	Neurotrophin signaling pathway	14	8.22 E-03
hsa04512	ECM-receptor interaction	11	8.31 E-03
hsa00010	Glycolysis / Gluconeogenesis	9	9.09 E-03
hsa04012	ErbB signaling pathway	11	1.06 E-02
hsa05220	Chronic myeloid leukemia	10	1.14 E-02
hsa05214	Glioma	9	1.21 E-02
hsa04540	Gap junction	11	1.23 E-02
hsa04010	MAPK signaling pathway	23	1.27 E-02
hsa05216	Thyroid cancer	6	1.30 E-02
hsa05219	Bladder cancer	7	1.66 E-02
hsa05223	Non-small cell lung cancer	8	1.66 E-02
hsa05120	Epithelial cell signaling in Helicobacter pylori infection	9	1.87 E-02
hsa04666	Fc gamma R-mediated phagocytosis	11	1.89 E-02
hsa00350	Tyrosine metabolism	7	2.06 E-02
hsa00360	Phenylalanine metabolism	5	2.16 E-02
hsa05222	Small cell lung cancer	10	2.28 E-02
hsa04914	Progesterone-mediated oocyte maturation	10	2.62 E-02
hsa04662	B cell receptor signaling pathway	9	3.17 E-02
hsa04650	Natural killer cell mediated cytotoxicity	13	3.21 E-02
hsa04150	mTOR signaling pathway	7	4.28 E-02

Reference

1. Bu D, Zhao Y, Cai L, Xue H, Zhu X, Lu H, Zhang J, Sun S, Ling L, Zhang N. Topological structure analysis of the protein-protein interaction network in budding yeast. Nucleic Acids Res 2003; 31: 2443-2450.
2. Nguyen PV, Srihari S, Leong HW. Identifying conserved protein complexes between species by constructing interolog networks. BMC Bioinformatics 2013; 14 Suppl 16: S8.
3. Srihari S, Leong HW. A survey of computational methods for protein complex prediction from protein interaction networks. J Bioinform Comput Biol 2013; 11: 1230002.
4. Li X, Wu M, Kwoh CK, Ng SK. Computational approaches for detecting protein complexes from protein interaction networks: a survey. BMC Genomics 2010; 11 Suppl 1: S3.
5. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics 2003; 4: 2.
6. Liu G, Wong L, Chua HN. Complex discovery from weighted PPI networks. Bioinformatics 2009; 25: 1891-1897.
7. Hwang W, Cho YR, Zhang A, Ramanathan M. A novel functional module detection algorithm for protein-protein interaction networks. Algorithms Mol Biol 2006; 1: 24.
8. Srihari S, Ragan MA. Systematic tracking of dysregulated modules identifies novel genes in cancer. Bioinformatics 2013; 29: 1553-1561.
9. Stass S, Harold R, Rock WR. Handbook of hematologic pathology. Marcel Dekker, 2000.
10. Rubnitz JE, Gibson B, Smith FO. Acute myeloid leukemia. Hematol Oncol Clin North Am 2010; 24: 35-63.
11. Callea V, Morabito F, Martino B, Stelitano C, Oliva B, Nobile F. Diagnostic and prognostic relevance of the immunophenotype in acute myelocytic leukemia. Tumori 1991; 77: 28-31.
12. Stirewalt DL, Meshinchi S, Kopecky KJ, Fan W, Pogosova-Agadjanyan EL, Engel JH, Cronk MR, Dorcy KS, McQuary AR, Hockenbery D, Wood B, Heimfeld S, Radich JP. Identification of genes with abnormal expression changes in acute myeloid leukemia. Genes Chromosomes Cancer 2008; 47: 8-20.
13. Beghini A, Corlazzoli F, Del Giacco L, Re M, Lazzaroni F, Brioschi M, Valentini G, Ferrazzi F, Ghilardi A, Righi M, Turrini M, Mignardi M, Cesana C, Bronte V, Nilsson M, Morra E, Cairoli R. Regeneration-associated WNT signaling is activated in long-term reconstituting AC133bright acute myeloid leukemia cells. Neoplasia 2012; 14: 1236-1248.
14. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ, von Mering C. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 2015; 43: D447-452.
15. Taboada B, Verde C, Merino E. High accuracy operon prediction method based on STRING database scores. Nucleic Acids Res 2010; 38: e130.
16. Ma L, Robinson LN, Towle HC. ChREBP*Mlx is the principal mediator of glucose-induced gene expression in the liver. J Biol Chem 2006; 281: 28721-28730.
17. Rifai N, Ridker PM. Proposed cardiovascular risk assessment algorithm using high-sensitivity C-reactive protein and lipid screening. Clin Chem 2001; 47: 28-30.
18. Pepper SD, Saunders EK, Edwards LE, Wilson CL, Miller CJ. The utility of MAS5 expression summary and detection call algorithms. BMC Bioinformatics 2007; 8: 273.
19. Huang H, Lu X, Liu Y, Haaland P, Marron JS. R/DWD: distance-weighted discrimination for classification, visualization and batch adjustment. Bioinformatics 2012; 28: 1182-1183.
20. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat Protoc 2009; 4: 1184-1191.
21. Adler J, Parmryd I. Quantifying colocalization by correlation: the Pearson correlation coefficient is superior to the Mander's overlap coefficient. Cytometry A 2010; 77: 733-742.
22. Tomita E, Tanaka A, Takahashi H. The worst-case time complexity for generating all maximal cliques and computational experiments. Theoret Comput Sci 2006; 363: 28-42.
23. Gabow HN. An efficient implementation of Edmonds' algorithm for maximum matching on graphs. J ACM 1976; 23: 221-234.
24. Zhao J, Yang TH, Huang Y, Holme P. Ranking candidate disease genes from gene expression and protein interaction: a Katz-centrality based approach. PLoS One 2011; 6: e24306.
25. Ford G, Xu Z, Gates A, Jiang J, 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.
26. Spirin V, Mirny LA. Protein complexes and functional modules in molecular networks. Proc Natl Acad Sci U S A 2003; 100: 12123-12128.
27. Polak R, Buitenhuis M. The PI3K/PKB signaling module as key regulator of hematopoiesis: implications for therapeutic strategies in leukemia. Blood 2012; 119: 911-923.
28. Milella M, Kornblau SM, Estrov Z, Carter BZ, Lapillonne H, Harris D, Konopleva M, Zhao S, Estey E, Andreeff M. Therapeutic targeting of the MEK/MAPK signal transduction module in acute myeloid leukemia. J Clin Invest 2001; 108: 851-859.
29. Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, Harris PL, Haserlat SM, Supko JG, Haluska FG, Louis DN, Christiani DC, Settleman J, Haber DA. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med 2004; 350: 2129-2139.
30. Kuan CT, Wikstrand CJ, Bigner DD. EGF mutant receptor vIII as a molecular target in cancer therapy. Endocr Relat Cancer 2001; 8: 83-96.

31. Sun JZ, Lu Y, Xu Y, Liu F, Li FQ, Wang QL, Wu CT, Hu XW, Duan HF. Epidermal growth factor receptor expression in acute myelogenous leukaemia is associated with clinical prognosis. Hematol Oncol 2012; 30: 89-97.
32. Zhang H, Berezov A, Wang Q, Zhang G, Drebin J, Murali R, Greene MI. ErbB receptors: from oncogenes to targeted cancer therapies. J Clin Invest 2007; 117: 2051-2058.
33. Paez JG, Janne PA, Lee JC, Tracy S, Greulich H, Gabriel S, Herman P, Kaye FJ, Lindeman N, Boggon TJ, Naoki K, Sasaki H, Fujii Y, Eck MJ, Sellers WR, Johnson BE, Meyerson M. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science 2004; 304: 1497-1500.
34. Lara-Guerra H, Chung CT, Schwock J, Pintilie M, Hwang DM, Leighl NB, Waddell TK, Tsao MS. Histopathological and immunohistochemical features associated with clinical response to neoadjuvant gefitinib therapy in early stage non-small cell lung cancer. Lung Cancer 2012; 76: 235-241.
35. Lainey E, Sebert M, Thepot S, Scoazec M, Bouteloup C, Leroy C, De Botton S, Galluzzi L, Fenaux P, Kroemer G. Erlotinib antagonizes ABC transporters in acute myeloid leukemia. Cell Cycle 2012; 11: 4079-4092.
36. Thepot S, Boehrer S, Seegers V, Prebet T, Beyne-Rauzy O, Wattel E, Delaunay J, Raffoux E, Hunault M, Jourdan E, Chermat F, Sebert M, Kroemer G, Fenaux P, Ades L. A phase I/II trial of Erlotinib in higher risk myelodysplastic syndromes and acute myeloid leukemia after azacitidine failure. Leuk Res 2014; 38: 1430-1434.
37. Stegmaier K, Corsello SM, Ross KN, Wong JS, Deangelo DJ, Golub TR. Gefitinib induces myeloid differentiation of acute myeloid leukemia. Blood 2005; 106: 2841-2848.
38. Boehrer S, Ades L, Galluzzi L, Tajeddine N, Tailler M, Gardin C, de Botton S, Fenaux P, Kroemer G. Erlotinib and gefitinib for the treatment of myelodysplastic syndrome and acute myeloid leukemia: a preclinical comparison. Biochem Pharmacol 2008; 76: 1417-1425.
39. Oda K, Matsuoka Y, Funahashi A, Kitano H. A comprehensive pathway map of epidermal growth factor receptor signaling. Mol Syst Biol 2005; 1: 2005 0010.
40. Champeris Tsaniras S, Kanellakis N, Symeonidou IE, Nikolopoulou P, Lygerou Z, Taraviras S. Licensing of DNA replication, cancer, pluripotency and differentiation: an interlinked world? Semin Cell Dev Biol 2014; 30: 174-180.
41. Rao RN. Targets for cancer therapy in the cell cycle pathway. Curr Opin Oncol 1996; 8: 516-524.
42. Newman DJ, Cragg GM, Holbeck S, Sausville EA. Natural products and derivatives as leads to cell cycle pathway targets in cancer chemotherapy. Curr Cancer Drug Targets 2002; 2: 279-308.
43. Chim CS, Wong AS, Kwong YL. Epigenetic inactivation of INK4/CDK/RB cell cycle pathway in acute leukemias. Ann Hematol 2003; 82: 738-742.
44. Romanski A, Bug G, Becker S, Kampfmann M, Seifried E, Hoelzer D, Ottmann OG, Tonn T. Mechanisms of resistance to natural killer cell-mediated cytotoxicity in acute lymphoblastic leukemia. Exp Hematol 2005; 33: 344-352.
45. Brune M, Hansson M, Mellqvist UH, Hermodsson S, Hellstrand K. NK cell-mediated killing of AML blasts: role of histamine, monocytes and reactive oxygen metabolites. Eur J Haematol 1996; 57: 312-319.
46. Rohner A, Langenkamp U, Siegler U, Kalberer CP, Wodnar-Filipowicz A. Differentiation-promoting drugs up-regulate NKG2D ligand expression and enhance the susceptibility of acute myeloid leukemia cells to natural killer cell-mediated lysis. Leuk Res 2007; 31: 1393-1402.
47. Ingram W, Chan L, Guven H, Darling D, Kordasti S, Hardwick N, Barber L, Mufti GJ, Farzaneh F. Human CD80/IL2 lentivirus-transduced acute myeloid leukaemia (AML) cells promote natural killer (NK) cell activation and cytolytic activity: implications for a phase I clinical study. Br J Haematol 2009; 145: 749-760.