Identification of potential key genes associated with diffuse large B-cell lymphoma based on microarray gene expression profiling

X. LUO[1], F. SHI[2], H. QIU[1], Y. TONG[2], X. GAO[2,*]

[1]Department of Hematology, The First Affiliated Hospital of Soochow University, Suzhou, Jiangsu 215006, P.R. China; [2]Department of Hematology, Jinhua People's Hospital, Jinhua, Zhejiang 321000, P.R. China

*Correspondence: gaoxinfang80@hotmail.com

The study aimed to screen potential key genes, and their targeted miRNAs and transcription factors (TFs) that were related to diffuse large B-cell lymphoma (DLBCL), and explore potential therapeutic targets for the progression of DLBCL. Dataset GSE56315 extracted from human tonsils was downloaded from Gene Expression Omnibus. Limma package was used to identify differential expression genes (DEG) between DLBCL and normal human tonsils samples, and the function and pathway enrichment analyses were performed. Then, functional interaction (FI) networks analyses of DEGs were implemented, and modules were extracted. Additionally, DLBCL-related miRNAs were predicted based on miR2disease database. Thereafter, TF-target DEGs and miRNAs targeted genes were respectively obtained. Finally, the integrated network of TF-target-miRNA was constructed. A total of 4,495 DEGs were identified between DLBCL and NHT samples. Among them, 114 up-regulated DEGs were contained in 8 modules of FI network, while 189 down-regulated DEGs were contained in 12 sub-modules. In addition, most DEGs were enriched in the function of "DNA binding" and pathways of "chemokine signaling pathway", "phosphatidylinositol signaling system" and "RNA degradation". Moreover, 19 miRNAs related with DLBCL were downloaded from Mirwalk2. Furthermore, miRNAs of miR-21-5p, miR-155 and miR-17-5p, the TF of STAT1, and DEGs such as *NUF2*, *CCR1*, *PIK3R1*, *SMC1A*, *FOXK1* and *CNOT6L* had high degrees in the integrated networks of TF-target-miRNA. DEGs like *NUF2*, *CCR1*, *PIK3R1*, *SMC1A*, *FOXK1* and *CNOT6L* might be closely associated with the pathogenesis of DLBCL.

Key words: Diffuse large B-cell lymphoma, differentially expressed gene, transcription factor, miRNA, functional interaction

Diffuse large B-cell lymphoma (DLBCL), a heterogeneous group of malignant proliferation of mature B cells, is one common kind of Non-Hodgkin's lymphoma (NHL) and accounts for 30% of NHL [1, 2]. In addition, DLBCL can be classified into two subgroups, germinal center B cell-like (GCB) and activated B cell-like (ABC) subgroups, based on the cDNA microarrays [3]. The incidence rate of DLBCL is about 10-15 of 100,000 people annually in the United States and most of them are men [4]. Additionally, inflammatory response is an important action in the development of DLBCL, and other genetic and environmental factors may also contribute to the acceleration of lymphomagenesis in DLBCL. Current treatment of DLBCL is typically combined chemotherapy. However, the relapse rate of DLBCL patients is nearly 40% [5]. Therefore, it is necessary to further develop new methods after comprehensive understanding pathogenic mechanisms of DLBCL from molecular level.

At present, the DLBCL related gene markers and molecular basis are partially understood. Several gene markers have been reported to be associated with the development of DLBCL. For instance, V-Myc avian myelocytomatosis viral oncogene homolog (*MYC*) is overexpressed in patients with DLBCL and contributes to tumorigenesis via repression of miR-17-92 cluster expression [6]. In addition, the down-regulated expression of miR-34a which increases its targets of *p53*, forkhead box p1 (*FOXP1*) and b-cell CLL/lymphoma 2 (*BCL2*), serves as a crucial prognostic marker of DLBCL in daily clinical work [7]. Similarly, Banham *et al* have found that the high expression of FOXP1 transcription factor (TF) may be closely linked with pathogenesis of DLBCL [8]. Furthermore, B lymphocyte-induced maturation protein 1 (*BLIMP1*), as a tumor suppressor gene, is inactivated by deletions or mutation in patients with DLBCL, and the inactivation of *BLIMP1* lead to lymphomagenesis by obstruction B cells differentiation [9].

Moreover, some pathways have been recently identified to be associated with the lymphomagenesis of DLBCL. The activation of nuclear factor kappa B (NF-κB) pathway is a hallmark of the ABC subtype of DLBCL, and may result in a change of the tumor cell proliferation and survival of DLBCL [10, 11].

Despite these informative findings, the gene and molecular mechanism involved in DLBCL is still obscure, and it needs to be further elucidated, which contributes to the clinical therapy and diagnosis of DLBCL. A previous study by Dybkaer *et al* hypothesizes to provide new diagnostic and prognostic tests for DLBCL based on subset-specific B-cell-associated gene signatures (BAGS) [12], and they have demonstrated that *CD58*, *LMO2*, several histocompatibility complex class II-signature and stromal-1-signature genes that may have a positive influence on prognosis of BAGS-assigned centrocyte subtype. However, they failed to comprehensively use those BAGS data to screen the key genes of DLBCL. In the present study, we re-analyzed their expression profile via a series of bioinformatics methods and the purpose of this study was to identify the key DLBCL related gene targets for providing important reference of further therapy and diagnosis of DLBCL.

Materials and methods

Data resource. All the data used in this study were downloaded from Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) website with the accession number of GSE56315 [12], which based on the platform of Affymetrix Human Genome U133 Plus 2.0 Array. A total of 122 samples were consisted in this profile, including 89 DLBCL samples and 33 normal human tonsils (NHT) samples.

Data preprocessing. Firstly, the raw probel-level data in CEL format were downloaded. Then, preprocessing of expression profile data were conducted by using Affy package in R language, such as format transformation of raw data, missing data filling, background correction by MAS method and normalization using quartile method [13].

Differentially expressed genes (DEGs) analysis and hierarchical clustering. Following the data preprocessing, the DEGs between DLBCL and NHT samples were analyzed by non-paired t test implemented in Linear Models for Microarray Analysis (limma, http://www.bioconductor.org/packages/release/bioc/html/limma.html) package in R [14], and multiple testing adjustment were performed by Benjamini & Hochberg method to adjust the original P-value into the false discovery rate (FDR) [15]. The FDR < 0.01 and |log2 fold change (FC)| ≥ 2 were chosen as the cut-off criteria to filter DEGs. Then, the pheatmap package in R (http://cran.r-project.org/web/packages/pheatmap/index.html) was used to perform clustering analysis [16].

Function and pathway enrichment analyses of DEGs. The Multifaceted Analysis Tool for Human Transcriptome (MATHT, http://www.biocloudservice.com) is an online tool for Gene Ontology (GO) function [17], and Kyoto Encyclopedia of Genes and Genomes (KEGG, http://www.genome.jp/kegg/) pathway analyses [18]. In this study, the MATHT tool was used to perform GO and KEGG enrichment analyses for the up-regulated and down-regulated DEGs, respectively, both with the cut-off criterion of P < 0.05.

Gene functional interaction (FI) analysis. The ReactomeFIViz is a Cytoscape app that allows further network analysis based on Reactome gene pathway annotations, and the networks were divided into modules with the ReactomeFIViz built-in "cluster FI network" tool [19]. In addition, ReactomeFIViz can calculate the pearson correlation coefficient (PCC) of each module. Then, modules were obtained with the parameter of MCLclustering size = 10. Moreover, the Database for Annotation, Visualization and Integrated Discovery (DAVID, http://david.abcc.ncifcrf.gov) was used to perform pathway enrichment analysis for each module gene set [20].

The integrated network construction of TF-target-miRNA. At first, the miRNAs related to DLBCL were searched from miR2disease (http://www.mir2disease.org/) database, and the validated target genes of 19 DLBCL-related miRNAs were downloaded from Mirwalk2 [21]. Then, these target genes and up-regulated or down-regulated DEGs were overlapped, respectively, using Venn diagram analysis, to select potential miRNA-DEG interactions. The iRegulon plugin (http://iregulon.aertslab.org.) in Cytoscape is aimed to detect enriched TF motifs and their optimal sets of direct targets [22]. In our study, iRegulon plugin was used to search the TFs of the miRNA-targeted genes based on the following parameters: minimum identity between orthologous genes: 0.05; maximum FDR on motif similarity: 0.001. The output result was Normalized Enrichment Score (NES), and the higher the score was, the more reliable of the prediction. The TF-target interactions with NES > 4 were selected, and then these targets were overlapped with DEGs, to identify the TFs of DEGs. Finally, the predicted TF-target-miRNA networks were constructed for up-regulated and down-regulated DEGs, respectively, by using Cytoscape software (http://www.cytoscape.org/) [23].

Results

DEGs screening and hierarchy cluster analysis. Based on the aforementioned criteria, a total of 4,495 DEGs were identified between DLBCL samples and NHT samples, including 2,224 up-regulated genes and 2,271 down-regulated genes in DLBCL samples. Additionally, the hierarchy cluster analysis revealed that the DEGs could correctly differentiate the two kinds of samples with correlated expression profiles, indicating that the DEGs could be applied to further analysis (Figure 1).

Function and pathway enrichment analyses of DEGs. Using the MATHT online tool, enrichment analyses indicated that the up-regulated DEGs were significantly related to functions such as "immune response", "ribosome", "chemotaxis" and "taxis", and pathways such as "chemokine signaling pathway" (e.g., *CCR1*) and "cytokine-cytokine receptor interaction" (Figure 2). In addition, the down-
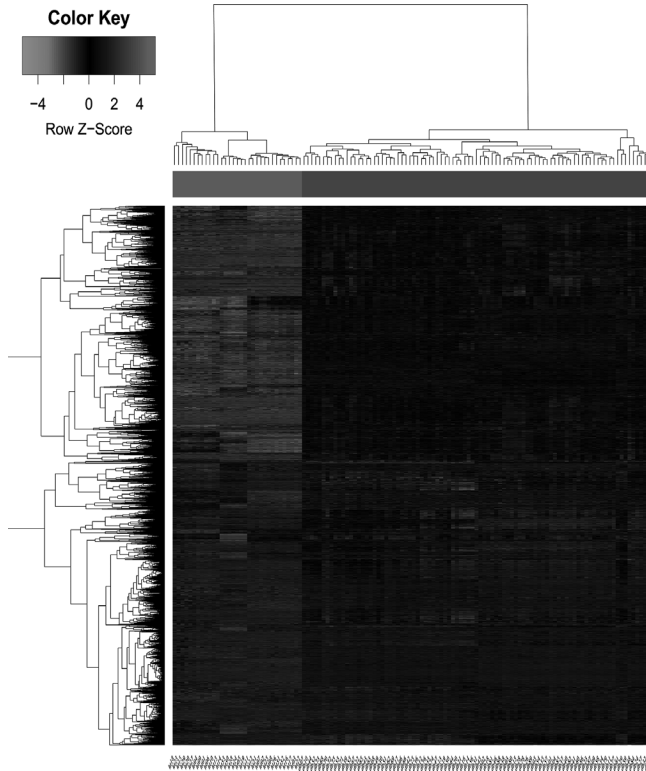
Figure 1. The hierarchical cluster map of DEGs. The horizontal axis shows names of each sample, and the right vertical axis shows clusters of DEGs. The color towards to red represents high expression values, and color towards to green represents lower expression values. Furthermore, the change of color from green to red stands for the change in expression values from low to high. DEGs: differentially expressed genes.

regulated DEGs were significantly related to functions such as "non-membrane-bounded organelle" and "DNA binding" (e.g., *FOXK1*), and pathways such as "ribosome" and "spliceosome" (Figure 3, Table 1).

Gene FI network of DEGs. As a result, a total of 8 modules of the FI network were obtained for the up-regulated DEGs (Figure 4A), including 114 nodes corresponding with 357 interactions. In those modules, the top 3 important modules were purple module with 24 DEGs (average PCC = 0.7801), green module with 18 DEGs (average PCC = 0.7362) and rose red module with 17 DEGs (average PCC = 0.8412). While, a total of 12 modules were obtained for the down-regulated DEGs, including 189 nodes corresponding with 1391 interactions (Figure 4B). Among those down-regulated modules, the top 3 important modules were respectively purple module with 39 DEGs (average PCC = 0.8211), green module with 24 DEGs (average PCC = 0.8532) and rose red module with 17 DEGs (average PCC = 0.8762) (Table 2).

Pathway enrichment analysis of the module gene sets. Based on the DAVID database, only gene sets from 10 modules were enriched in corresponding KEGG pathways, among which 5 modules were from up-regulated DEGs and 5 modules were from down-regulated DEGs. As a result, up-regulated DEGs were remarkably enriched in pathways such as "cell cycle" (e.g., *ESPL1*) and "phosphatidylinositol signaling system" (e.g., *PIK3R1*). Meanwhile, down-regulated DEGs were closely associated with pathway of "RNA degradation" (e.g., *CNOT6L*) (Table 3).

The integrated TF-target-miRNA network. In the integrated TF-target-miRNA network of the up-regulated DEGs,
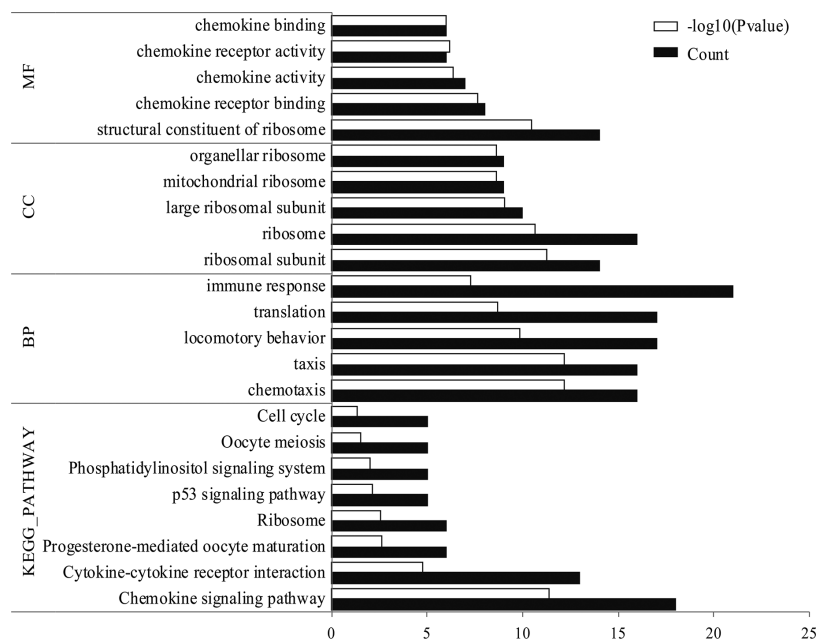
Figure 2. Result of pathway enrichment and top 5 categories of up-regulated DEGs. MF: molecular function; CC: cellular component; BP: biological process; DEGs: differentially expressed genes.
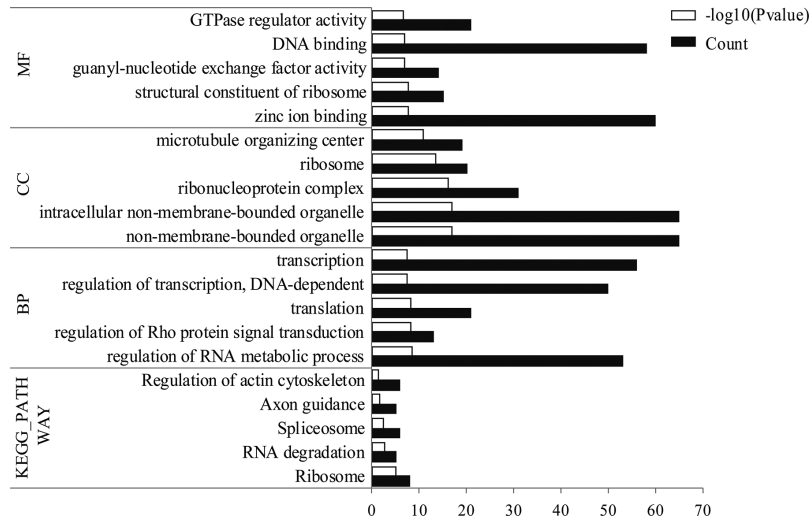
Figure 3. Result of pathway enrichment and top 5 categories of down-regulated DEGs. MF: molecular function; CC: cellular component; BP: biological process; DEGs: differentially expressed genes.

only one TF, STAT1, was found to have relationships with miRNA-targeted DEGs, such as *NUF2*, *CCR1* and *DTX3L* (Figure 5). *PIK3R1* was another crucial node in this network with high degree. By contrast, no TF was searched for the down-regulated DEGs, and DEGs such as *SMC1A*, *FOXK1* and *CNOT6L* were highlighted with high connection degrees were integrated in the down-regulation network (Figure 6). Additionally, two miRNAs, hsa-miR-21-5p (degree = 10) and hsa-miR-17-5p (degree = 25), which related to DLBCL had high connection degrees in above two TF-target-miRNA networks. Moreover, ten DEGs (e.g., *CCR1*, *PIK3R1* and *SMC1A*) were identified to be the targets of hsa-miR-21-5p, whilst twenty five DEGs such as *FOXK1* and *CNOT6L* were targets of hsa-miR-17-5p.

Discussion

In the present study, a total of 114 up-regulated and 189 down-regulated DEGs were identified. Eight modules of the FI network were constructed for the up-regulated DEGs, while 12 were for the down-regulated DEGs. In addition, DEGs such as *NUF2*, *CCR1*, *PIK3R1*, *SMC1A*, *FOXK1* and *CNOT6L*, the miRNAs of hsa-miR-21-5p, miR-155 and hsa-miR-17-5p, and the TF of STAT1 were highlighted in the TF-target-miRNA networks. Moreover, most DEGs were predicted to be related to the function of "DNA binding" (e.g., *FOXK1*) and pathways like "phosphatidylinositol signaling system" (e.g., PIK3R1), "chemokine signaling pathway" (e.g., *CCR1*) and pathway of "RNA degradation" (e.g., *CNOT6L*).

Table 1. The remarkably enriched pathway of differentially expressed genes

	PathwayID	PathwayName	Count	P Value	Genes
Up-regulated genes	hsa04062	Chemokine signaling pathway	18	4.18E-12	*CCR1, STAT1, CCR7, CCR5*, PIK3R1…
	hsa04060	Cytokine-cytokine receptor interaction	13	1.74E-05	*CCR1, CXCL2, CXCL9, CX3CL1, CCR7…*
	hsa04914	Progesterone-mediated oocyte maturation	6	2.32E-03	*CCNB1, CDK1, CCNB2, PIK3R3, PIK3R1…*
	hsa03010	Ribosome	6	2.57E-03	*RPL31, RPL22, RPLP0, RPS24, RPL29…*
	hsa04115	p53 signaling pathway	5	6.78E-03	*CCNB1, CDK1, CCNB2, CHEK1, PTEN*
	hsa04070	Phosphatidylinositol signaling system	5	9.12E-03	*SYNJ1, PIK3R3, PTEN, PIK3R1, PIP4K2B*
	hsa04114	Oocyte meiosis	5	3.12E-02	*CCNB1, CDK1, CCNB2, SLK, ESPL1*
	hsa04110	Cell cycle	5	4.70E-02	*CCNB1, CDK1, CCNB2, CHEK1, ESPL1*
Down-regulated genes	hsa03010	Ribosome	8	7.07E-06	*RPL14, RPL27A, RPLP2, RPS27L, RPL37A…*
	hsa03018	RNA degradation	5	1.39E-03	*CNOT8, CNOT6L, CNOT3, CNOT7, CNOT4*
	hsa03040	Spliceosome	6	4.05E-03	*TRA2A, CDC40, CDC5L, HNRNPC, SNRNP70…*
	hsa04360	Axon guidance	5	2.47E-02	*ABLIM1, CDC42, PLXNB2, RAC1, ARHGEF12*
	hsa04810	Regulation of actin cytoskeleton	6	3.32E-02	*CDC42, ARHGEF7, RAC1, ARHGEF12, FGD3…*

Notes: Count stands for the number of differentially expressed genes which were enriched in the corresponding functional category.

Table 2. The modules of differentially expressed genes

	Module	Genes in Module	Average Correlation	Gene
Modules for up-regulated genes	0	24	0.7801	CDK1, CEP41, NSL1, NUF2, SLK…
	1	18	0.7362	APLNR, C5AR1, CCL21, CXCL13, CXCL2…
	2	17	0.8412	CISH, IFI27, PTPN2, STAT1, WARS…
	3	13	0.7503	PIK3R1, PIK3R3, PTEN, RCAN2, SYNJ1…
	4	11	0.876	GFM1, MRPL12, MRPL23, MRPL3, MRPL32…
	5	11	0.6575	CCR1, CCR2, CCR5, CCR7, CX3CL1…
	6	10	0.7523	BATF3, BCL2A1, CEBPB, EMSY, ENPP2…
	7	10	0.6683	ETF1, RPL22, RPL29, RPL31, RPLP0 …
Modules for down-regulated genes	0	39	0.8211	ZFP36L2, ZNF107, ZNF117, ZNF141, ZNF160…
	1	24	0.8532	AKAP9, BUB1, CENPJ, CEP135, CEP290…
	2	17	0.8762	ABLIM1, ABR, AKAP13, ARHGAP30, ARHGAP33…
	3	15	0.8881	EIF5B, GRSF1, RPL14, RPL27A, RPL37A…
	4	15	0.8436	AURKAIP1, GADD45GIP1, GFM2, MRPL20, MRPL38…
	5	14	0.731	CDC40, CDC5L, FUS, HNRNPC, HNRNPL…
	6	13	0.7959	BIRC5, CENPT, ESCO1, MAU2, NSL1…
	7	12	0.8796	CNOT3, CNOT4, CNOT6L, CNOT7, CNOT8…
	8	10	0.9093	ARFGEF2, IPO8, KPNA5, NEK9, NUP188…
	9	10	0.9001	ARID4B, FOXK1, ING2, KLF13, MNT…
	10	10	0.855	ARHGAP23, CDC42, CDC42BPG, FGD2, FGD3…
	11	10	0.837	AKIRIN2, EMSY, HMBOX1, PHF12, RNF141…

**Table 3. The pathway enrichment analysis of differentially expressed genes in modules**

	Module in Network	Pathway ID	Pathway Name	Count	P Value	Gene
Up-regulated DEGs	Module 0	hsa04114	Oocyte meiosis	5	3.61E-06	CCNB1, CDK1, CCNB2, SLK, ESPL1
		hsa04110	Cell cycle	5	6.12E-06	CCNB1, CDK1, CCNB2, CHEK1, ESPL1
		hsa04115	p53 signaling pathway	4	5.54E-05	CCNB1, CDK1, CCNB2, CHEK1
		hsa04914	Progesterone-mediated oocyte maturation	3	4.48E-03	CCNB1, CDK1, CCNB2
	Module 1	hsa04080	Neuroactive ligand-receptor interaction	6	8.73E-05	S1PR3, P2RY13, C5AR1, P2RY14, FPR3…
		hsa04062	Chemokine signaling pathway	4	5.53E-03	GNAI3, CCL21, CXCL13, CXCL2
	Module 3	hsa04070	Phosphatidylinositol signaling system	5	1.09E-05	SYNJ1, PIK3R3, PTEN, PIK3R1, PIP4K2B
		hsa05213	Endometrial cancer	3	5.06E-03	PIK3R3, PTEN, PIK3R1
		hsa00562	Inositol phosphate metabolism	3	5.45E-03	SYNJ1, PTEN, PIP4K2B
		hsa05214	Glioma	3	7.36E-03	PIK3R3, PTEN, PIK3R1
		hsa05218	Melanoma	3	9.28E-03	PIK3R3, PTEN, PIK3R1
		hsa04370	VEGF signaling pathway	3	1.01E-02	SH2D2A, PIK3R3, PIK3R1
		hsa05222	Small cell lung cancer	3	1.28E-02	PIK3R3, PTEN, PIK3R1
		hsa05215	Prostate cancer	3	1.43E-02	PIK3R3, PTEN, PIK3R1
		hsa04666	Fc gamma R-mediated phagocytosis	3	1.59E-02	PIK3R3, PIK3R1, PIP4K2B
		hsa04660	T cell receptor signaling pathway	3	2.03E-02	ICOS, PIK3R3, PIK3R1
		hsa04650	Natural killer cell mediated cytotoxicity	3	2.85E-02	PIK3R3, PIK3R1, HCST
	Module 5	hsa04062	Chemokine signaling pathway	10	1.55E-13	CCR7, CCR5, CXCL14, CCR1, CCR2…
		hsa04060	Cytokine-cytokine receptor interaction	10	3.41E-12	CCR7, CCR5, CXCL14, CCR1, CCR2…
	Module 7	hsa03010	Ribosome	6	1.06E-08	RPL31, RPL22, RPLP0, RPS4Y1, RPS24…
Down-regulated DEGs	Module 1	hsa00565	Ether lipid metabolism	2	2.85E-02	PAFAH1B1, PAFAH1B2
	Module 2	hsa04360	Axon guidance	4	3.75E-04	ABLIM1, PLXNB2, RAC1, ARHGEF12
		hsa04810	Regulation of actin cytoskeleton	4	1.55E-03	ARHGEF7, RAC1, IQGAP3, ARHGEF12
	Module 3	hsa03010	Ribosome	8	4.10E-12	RPL14, RPL27A, RPLP2, RPS27L, RPL37A…
	Module 7	hsa03018	RNA degradation	5	9.46E-08	CNOT8, CNOT6L, CNOT3, CNOT7, CNOT4
	Module 10	hsa04810	Regulation of actin cytoskeleton	2	4.43E-02	CDC42, FGD3

Notes: DEG represents differentially expressed genes; Count stands for the number of DEGs which were enriched in the corresponding functional category.

Signal transducer and activator of transcription 1 (STAT1) belongs to the STAT family, which is activated by tyrosine phosphorylation and crucial to develop human neoplasias via participating in the regulation of cell proliferation and survival [24]. Previous studies have showed that STAT1 plays an important role in promoting apoptosis [25]. In the present study, the STAT1 was the only TF in the TF-target-miRNA network for up-regulated DEGs, and was predicted to regulate the expres-
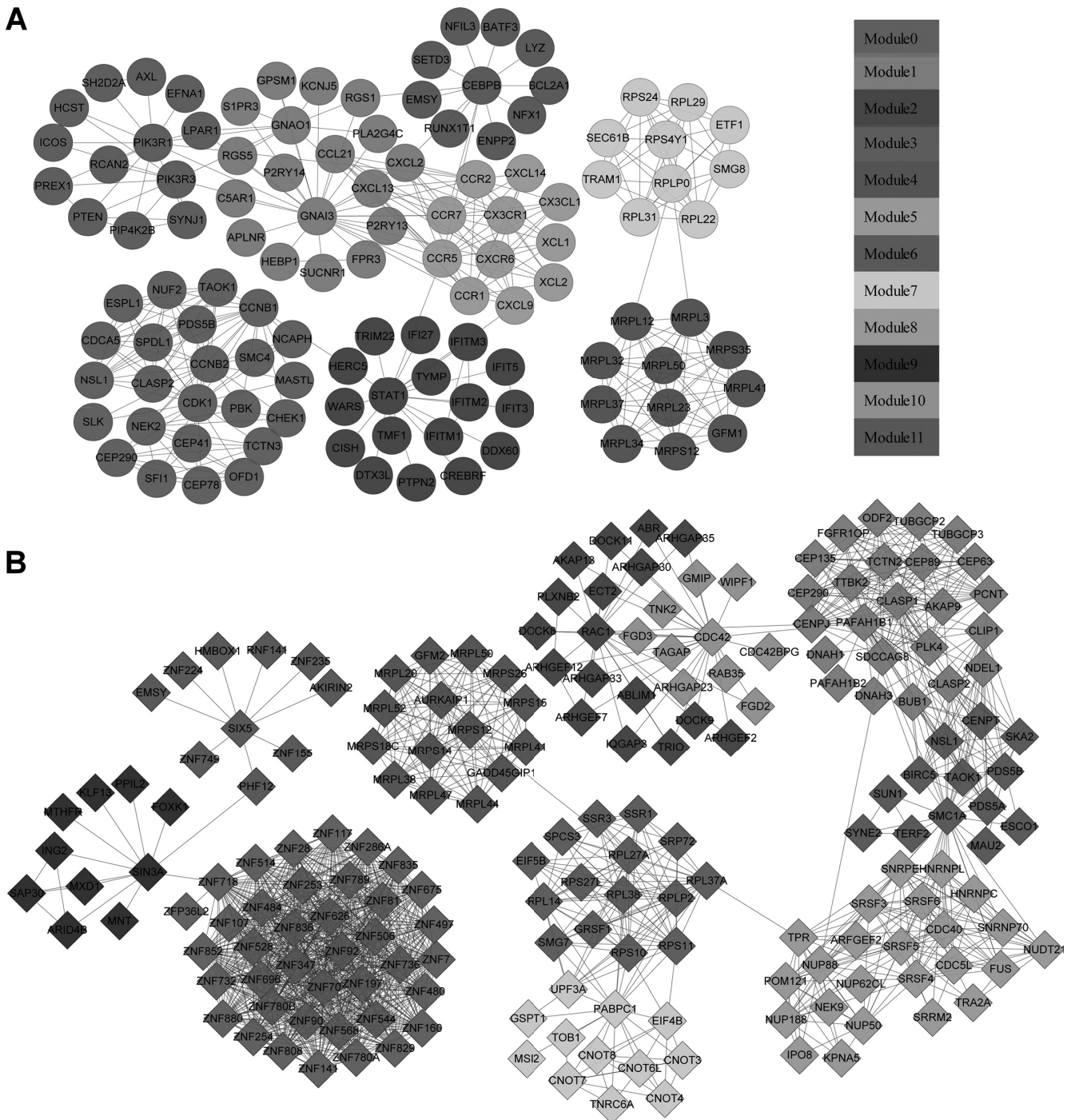
Figure 4. The modules identified from the functional interactions of DEGs. A. The modules identified from functional interactions of up-regulated DEGs. B. The modules identified from functional interactions of down-regulated DEGs. The circular-shaped nodes represent up-regulated DEGs, and the diamond-shaped nodes represent down-regulated DEGs. Different colors are used to distinguish different clustering modules. DEGs: differentially expressed genes.
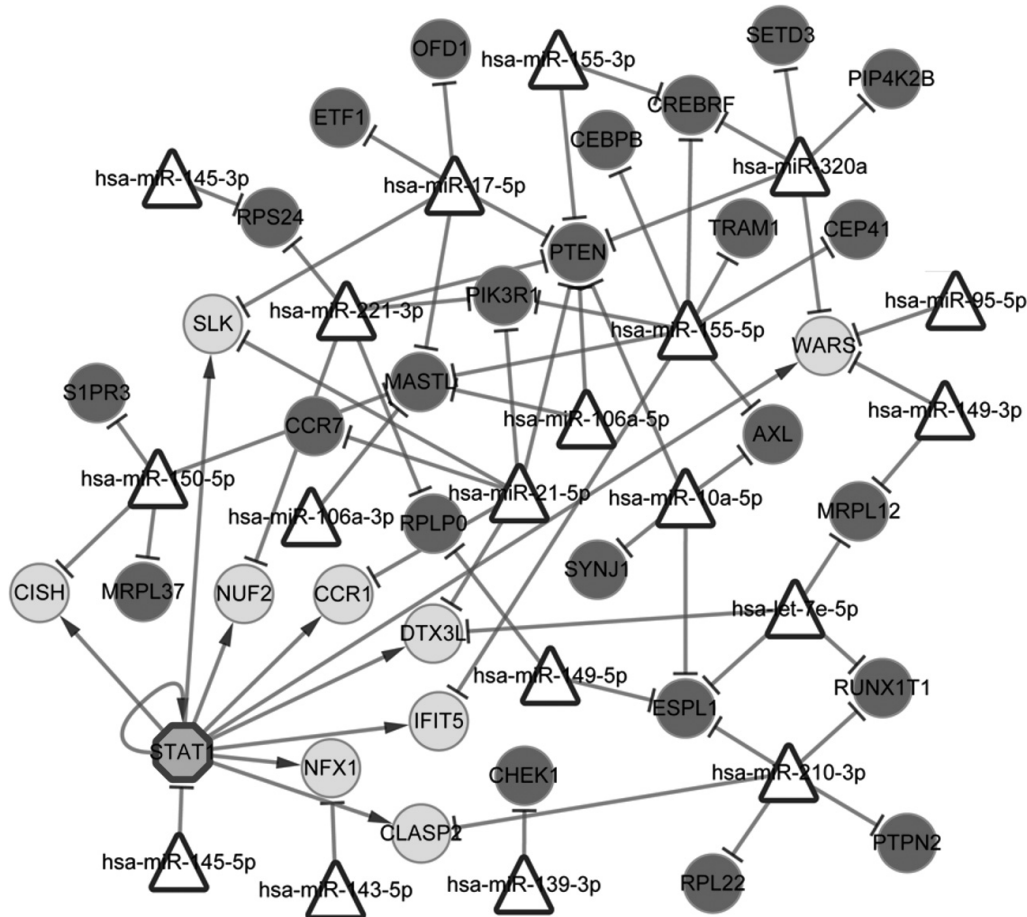
Figure 5. Integrative regulatory network of TF-target-miRNA for the up-regulated DEGs. The triangle-shaped nodes stand for miRNA, and all the remaining nodes in red circles stand for up-regulated DEGs. The hexagon-shaped nodes represent TF. The T-shaped lines represent interactions between genes and miRNAs, and the arrows represent interactions between TF and target genes. DEGs: differentially expressed genes; TF: transcription factor.

sion of *NUF2*, suggesting it might have an inhibitory role in DLBCL development. In addition, it is indicated that NDC80 kinetochore complex component (*NUF2*) was up-regulated in human glioma cells and pancreatic tissues, and *NUF2* plays an important role in regulation cell apoptosis [26, 27]. Therefore, it could be inferred that the overexpression of STAT1 might induce the up-regulation of *NUF2*. These regulations might contribute to inhibit the progression of DLBCL via promoting cell apoptosis. However, there are some different views on the role of STAT1 in DLBCL. Bhatt *et al* have demonstrated that interleukin 21 can induce cell apoptosis in a subset of mantle cell lymphoma cells via activating the STAT3-cMyc pathway and not by activating the STAT1 signaling pathway [29]. In addition, Camicia *et al* have suggested that *BAL1* as a oncogenic survival factor in DLBCL may induce the phosphorylation of STAT1, and STAT1 acts as an oncogene in DLBCL, not as tumor suppressor [30]. Therefore, it is still needed further experimental data to verify the function of STAT1 in DLBCL.

Numerous studies have proved that miR-21 has great function in the development of DLBCL and other cancers via

regulating cell proliferation and apoptosis [31-33]. In our study, *CCR1*, *PIK3R1* and *SMC1A* were predicted as target genes of miR-21-5p, and *CCR1* and *PIK3R1* had high degrees in the up-regulated network of TF-target-miRNA. The chemokine receptor 1 (*CCR1*) encodes a member of the beta chemokine receptor family, and it can bind to C-X-C motif chemokine to have function in the progression of tumors [34]. Interestingly, a previous study indicated that the expression of *CCR1* was up-regulated in the non-germinal center subtype of DLBCL, and *CCR1* may contribute to the progression of DLBCL via chemokine signaling pathway [35], which was consistent with our findings. In addition, phosphoinositide-3-kinase regulatory subunit 1 (*PIK3R1*) is negatively regulated by miR-21-5p, and miR-21-5p targeted *PIK3R1* inhibits cell growth and invasion via suppressing PI3K/AKT signaling activation in breast cancer [36]. Moreover, PI3K/AKT pathway is involved in the development of DLBCL [37]. Similarly, chemokine signaling pathway and PI3K/AKT signaling pathway were predicted to be significant pathways in this study. Therefore, it might be supposed that *CCR1* and *PIK3R1* might function as
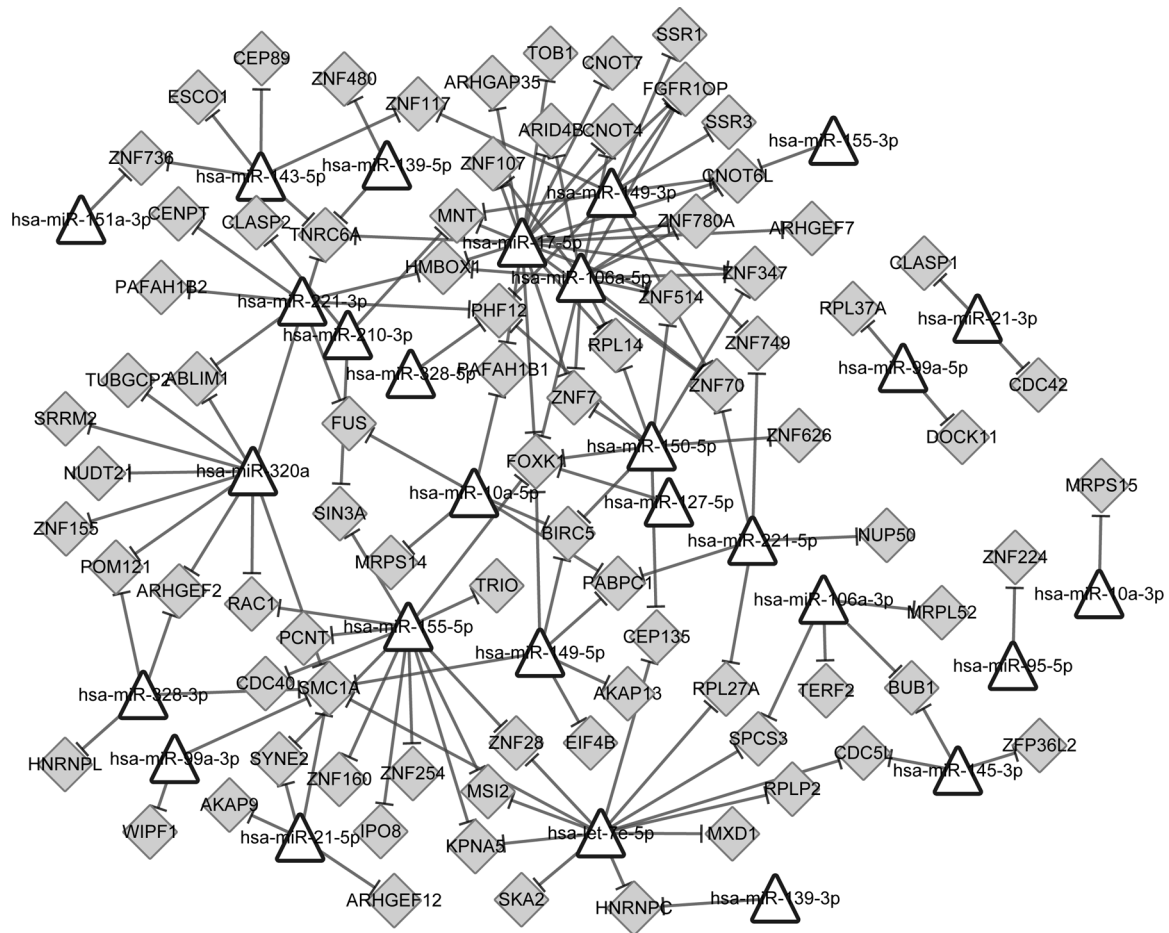
Figure 6. Integrative regulatory network of miRNA-target for the down-regulated DEGs. The triangle-shaped nodes stand for miRNA, and the diamond-shaped nodes represent down-regulated DEGs. The T-shaped lines represent interactions between genes and miRNAs. DEGs: differentially expressed genes.

important genes to participate the development of DLBLC via those two pathways. Structural maintenance of chromosomes 1A (*SMC1A*) that encodes a subunit of the cohesin protein complex is involved in chromosome cohesion during cell cycle and DNA repair [38]. On the other hand, in the present study, *SMC1A* showed high degree in the down-regulated network. The knockdown of *SMC1A* suppresses the growth of glioblastoma cells [38, 39]. Collectively, we supposed that miR-21, as well as its target genes like *CCR1*, *PIK3R1* and *SMC1A* might control the growth of DLBCL cells. Additionally, it has been indicated that miR-155 overexpression plays important roles in the progression of DLBCL [40]. In our study, *SMAD4* as the target gene of miR-155 was predicted to be down-regulated in the DLBCL cells. In consistent with our result, Go *et al* have found that the loss of *SMAD4* expression may be related with the proliferation of DLBCL cells by enhancing transforming growth factor β (TGF-β) pathway signaling pathway [41]. In addition, it has been reported that miR-155 may regulate *SMAD4* in TGF-β1/SMAD signaling pathway in human breast cancer cells [42]. Therefore, we supposed that miR-155 and its

target gene *SMAD4* might play key role in the occurrence of lymphoma of DLBCL via TGF-β1 pathway.

Moreover, several studies have proved that miR-17-5p plays a role as a tumor suppressor in malignant cells such as breast cancer cell, lung cancer cell and gastric cancer cell [43-45]. In our study, *FOXK1* and *CNOT6L*, which were two crucial genes with a high connection degree in the integrated network of down-regulated genes, were predicted as the targeted genes of miR-17-5p. Myocyte nuclear factor (*FOXK1*) contains a fork head DNA binding domain which is correlative with cell growth and metabolism, and *FOXK1* suppression lead to apoptosis and promote cell susceptibility to 5-fluorouracil-induced apoptosis in colorectal cancer cell [46]. In addition, CCR4-NOT transcription complex subunit 6 like (*CNOT6L*) is a deadenylase subunit of CCR4-NOT complex [47]. Moreover, *CNOT6L* is linked to cell death and survival by recognizing the RNA substrate and causing mRNA degradation [48]. Furthermore, miR-17 can bind to *CNOT6L* to down-regulate the level of phosphatase and tensin homolog (*PTEN*), and thus to facilitate tumor growth [49]. In consistent with our results,

FOXK1 and *CNOT6L* were enriched in the pathways of "DNA binding" and "RNA degradation", respectively. Consequently, we inferred that *FOXK1* and *CNOT6L* which participated in "DNA binding" and "RNA degradation" are most likely to play a major role in the progression of DLBCL via regulation of cell growth and apoptosis of DLBCL cells by targeting miR-17.

In summary, the DEGs like *NUF2*, *CCR1*, *PIK3R1*, *SMC1A*, *SMAD4*, *FOXK1* and *CNOT6L*, the TF of STAT1, and the miRNAs such as miR-21-5p, miR-155 and miR-17-5p might be closely involved in the progression of DLBCL via regulation the cell proliferation and apoptosis. Moreover, they might be used as potential therapeutic and diagnostic biomarkers for DLBCL. However, further studies should be designed to confirm these results.

References

[1] ALIZADEH AA, EISEN MB, DAVIS RE, MA C, LOSSOS IS et al. Distinct types of diffuse large b-cell lymphoma identified by gene expression profiling. Nature 2000; 403: 503–511. https://doi.org/10.1038/35000501

[2] LOSSOS IS. Molecular pathogenesis of diffuse large B-cell lymphoma. J Clin Oncol 2005; 23: 6351–6357. https://doi.org/10.1200/JCO.2005.05.012

[3] WRIGHT G, TAN B, ROSENWALD A, HURT EH, WIESTNER A et al. A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. Proc Natl Acad Sci U S A 2003; 100: 9991–9996. https://doi.org/10.1073/pnas.1732008100

[4] MORTON LM, WANG SS, DEVESA SS, HARTGE P, WEISENBURGER DD et al. Lymphoma incidence patterns by WHO subtype in the United States, 1992–2001. Blood 2006; 107: 265–276. https://doi.org/10.1182/blood-2005-06-2508

[5] VAIDYA R, WITZIG TE. Prognostic factors for diffuse large B-cell lymphoma in the R(X)CHOP era. Ann Oncol 2014; 25: 2124–2133. https://doi.org/10.1093/annonc/mdu109

[6] CHANG TC, YU D, LEE YS, WENTZEL EA, ARKING DE et al. Widespread microRNA repression by Myc contributes to tumorigenesis. Nat Genet 2008; 40: 43–50. https://doi.org/10.1038/ng.2007.30

[7] HE M, GAO L, ZHANG S, TAO L, WANG J et al. Prognostic significance of miR-34a and its target proteins of FOXP1, p53, and BCL2 in gastric MALT lymphoma and DLBCL. Gastric Cancer 2013; 17: 431–441. https://doi.org/10.1007/s10120-013-0313-3

[8] BANHAM AH, CONNORS JM, BROWN PJ, CORDELL JL, OTT G et al. Expression of the FOXP1 transcription factor is strongly associated with inferior survival in patients with diffuse large B-cell lymphoma. Clin Cancer Res 2005; 11: 1065–1072.

[9] PASQUALUCCI L, COMPAGNO M, HOULDSWORTH J, MONTI S, GRUNN A et al. Inactivation of the PRDM1/BLIMP1 gene in diffuse large B cell lymphoma. J Exp Med 2006; 203: 311–317. https://doi.org/10.1084/jem.20052204

[10] KOZLOSKI GA, JIANG X, BHATT S, RUIZ J, VEGA F et al. MiR-181a negatively regulates NF-κB signaling and affects activated B-cell like diffuse large B-cell lymphoma pathogenesis. Blood 2016; 127: 2856–2866. https://doi.org/10.1182/blood-2015-11-680462

[11] BOGNAR MK, VINCENDEAU M, ERDMANN T, SEEHOLZER T, GRAU M et al. Oncogenic CARMA1 couples NF-κB and β-catenin signaling in diffuse large B-cell lymphomas. Oncogene 2016; 35: 4269–4281. https://doi.org/10.1038/onc.2015.493

[12] DYBKAER K, BOGSTED M, FALGREEN S, BODKER JS, KJELDSEN MK et al. Diffuse large B-cell lymphoma classification system that associates normal B-cell subset phenotypes with prognosis. J Clin Oncol 2015; 33: 1379–1388. https://doi.org/10.1200/JCO.2014.57.7080

[13] GAUTIER L, COPE L, BOLSTAD BM, IRIZARRY RA. Irizarry RA: affy–analysis of Affymetrix GeneChip data at the probe level. Bioinformatics 2004; 20: 307–315. https://doi.org/10.1093/bioinformatics/btg405

[14] SMYTH GK. limma: Linear Models for Microarray Data, p 397–420. In: R. Gentleman, V. J. Carey, W. Huber, R. A. Irizarry, S. Dudoit (Eds.), Bioinformatics and Computational Biology Solutions Using R and Bioconductor, Springer New York, New York, NY, 2005, pp 465. ISBN 9870387251462.

[15] BENJAMINI Y, HOCHBERG Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Statist Soc B 1995; 57: 289–300.

[16] WANG L, CAO C, MA Q, ZENG Q, WANG H et al. RNA-seq analyses of multiple meristems of soybean: novel and alternative transcripts, evolutionary and functional implications. BMC Plant Biol 2014; 14: 169. https://doi.org/10.1186/1471-2229-14-169

[17] ASHBURNER M, BALL CA, BLAKE JA, BOTSTEIN D, BUTLER H et al. Gene Ontology: tool for the unification of biology. Nat Genet 2000; 25: 25–29. https://doi.org/10.1038/75556

[18] KANEHISA M, GOTO S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000; 28: 27–30. https://doi.org/10.1093/nar/28.1.27

[19] WU G, DAWSON E, DUONG A, HAW R, STEIN L. ReactomeFIViz: a Cytoscape app for pathway and network-based data analysis. Version 2. F1000Res. 2014; 3: 146. https://doi.org/10.12688/f1000research.4431.2

[20] HUANG DW, SHERMAN BT, TAN Q, COLLINS JR, ALVORD WG et al. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol 2007; 8: R183. https://doi.org/10.1186/gb-2007-8-9-r183

[21] DWEEP H, GRETZ N. miRWalk2.0: a comprehensive atlas of microRNA-target interactions. Nat Methods 2015; 12: 697. https://doi.org/10.1038/nmeth.3485

[22] JANKY RS, VERFAILLIE A, IMRICHOVA H, VAN DE SANDE B, STANDAERT L et al. iRegulon: From a Gene List to a Gene Regulatory Network Using Large Motif and Track Collections. PLoS Comput Biol 2014; 10: e1003731. https://doi.org/10.1371/journal.pcbi.1003731

[23] KOHL M, WIESE S, WARSCHEID B. Cytoscape: software for visualization and analysis of biological networks. Methods Mol Biol 2011; 696: 291–303. https://doi.org/10.1007/978-1-60761-987-1_18

[24] CALO V, MIGLIAVACCA M, BAZAN V, MACALUSO M, BUSCEMI M et al. STAT proteins: From normal control of cellular events to tumorigenesis. J Cell Physiol 2003; 197: 157–168. https://doi.org/10.1002/jcp.10364

[25] STEPHANOU A, KNIGHT R, D S. (Eds.) STAT1 and STAT3: Opposing roles in Cell Death and Cell Cycle Regulation. JAK-STAT Pathway in Disease. Landes Bioscience: Texas, 2009. ISBN 978–1-58706–315–2

[26] HUANG SK, QIAN JX, YUAN BQ, LIN YY, YE ZX et al. SiRNA-mediated knockdown against NUF2 suppresses tumor growth and induces cell apoptosis in human glioma cells. Cell Mol Biol (Noisy-le-grand) 2014; 60: 30–36.

[27] HU P, SHANGGUAN J, ZHANG L. Downregulation of NUF2 inhibits tumor growth and induces apoptosis by regulating lncRNA AF339813. Int J Clin Exp Pathol 2015; 8: 2638–2648.

[28] BACHMANN SB, FROMMEL SC, CAMICIA R, WINKLER HC, SANTORO R et al. DTX3L and ARTD9 inhibit IRF1 expression and mediate in cooperation with ARTD8 survival and proliferation of metastatic prostate cancer cells. Mol Cancer 2014; 13: 125. https://doi.org/10.1186/1476-4598-13-125

[29] BHATT S, SAROSIEK KA, LOSSOS IS. Abstract LB-328: Preclinical activity of interleukin 21 in mantle cell lymphoma. Cancer Res 2013; 73: https://doi.org/10.1158/1538-7445.AM2013-LB-328

[30] CAMICIA R, BACHMANN SB, WINKLER HC, BEER M, TINGUELY M et al. BAL1/ARTD9 represses the anti-proliferative and pro-apoptotic IFNgamma-STAT1-IRF1-p53 axis in diffuse large B-cell lymphoma. J Cell Sci 2013; 126: 1969–1980. https://doi.org/10.1242/jcs.118174

[31] SONG GQ, GU L, HE BS, PAN YQ, WANG SK. Expression of MiRNA-21 in diffuse large B cell lymphoma and its significance. Zhongguo Shi Yan Xue Ye Xue Za Zhi 2014; 22: 1603–1609. doi: 10.7534/j.issn.1009–2137.2014.06.019

[32] YAN LX, HUANG XF, SHAO Q, HUANG MY, DENG L et al. MicroRNA miR-21 overexpression in human breast cancer is associated with advanced clinical stage, lymph node metastasis and patient poor prognosis. RNA 2008; 14: 2348–2360. https://doi.org/10.1261/rna.1034808

[33] FRANKEL L, CHRISTOFFERSEN N, JACOBSEN A, LINDOW M, KROGH A et al. Programmed cell death 4 (PDCD4) is an important functional target of the microRNA miR-21 in breast cancer cells. J Biol Chem 2008; 283: 1026–1033. https://doi.org/10.1074/jbc.M707224200

[34] ZHU Q, HAN X, PENG J, QIN H, WANG Y. The role of CXC chemokines and their receptors in the progression and treatment of tumors. J Mol Histol 2012; 43: 699–713. https://doi.org/10.1007/s10735-012-9435-x

[35] ANDERSON MW, ZHAO S, AI WZ, TIBSHIRANI R, LEVY R et al. C-C chemokine receptor 1 expression in human hematolymphoid neoplasia. Am J Clin Pathol 2010; 133: 473–483. https://doi.org/10.1309/AJCP1TA3FLOQTMHF

[36] YAN LX, LIU YH, XIANG JW, WU QN, XU LB et al. PIK3R1 targeting by miR-21 suppresses tumor cell migration and invasion by reducing PI3K/AKT signaling and reversing EMT, and predicts clinical outcome of breast cancer. Int J Oncol 2016; 48: 471–484. https://doi.org/10.3892/ijo.2015.3287

[37] BAOHUA Y, XIAOYAN Z, TIECHENG Z, TAO Q, DAREN S. Mutations of the PIK3CA gene in diffuse large B cell lymphoma. Diagn Mol Pathol 2008; 17: 159–165. https://doi.org/10.1097/PDM.0b013e31815d0588

[38] Zhang Y, Jiang R, LI JD, ZHANG XY, ZHAO P et al. SMC1A knockdown induces growth suppression of human lung adenocarcinoma cells through G1/S cell cycle phase arrest and apoptosis pathways in vitro. Oncol Lett 2013; 5: 749–755.

[39] LI J, FENG W, CHEN L, HE J. Downregulation of SMC1A inhibits growth and increases apoptosis and chemosensitivity of colorectal cancer cells. J Int Med Res 2016; 44: 67–74. https://doi.org/10.1177/0300060515600188

[40] SHI JS, ZHANG J, LI J. [Role of miR-155 in pathogenesis of diffuse large B cell lymphoma and its possible mechanism]. Zhongguo shi yan xue ye xue za zhi 2014; 22: 869–872.

[41] GO JH. Expression Pattern of Smad Proteins in Diffuse Large B-cell Lymphomas. Korean J Pathol 2004; 38: 301–305.

[42] NIKIPHOROVA ZN, KUDRYAVTSEV IA, SHEVCHENKO VE. miR-21 and miR-155 in the regulation of TGF-β1/SMAD signaling pathway of the line breast cancer cells with different metastatic potential. Women Reproductive System Tumors 2015; 11: 15–21. https://doi.org/10.17650/1994-4098-2015-11-3-15-21

[43] HOSSAIN A, KUO MT, SAUNDERS GF. Mir-17–5p Regulates Breast Cancer Cell Proliferation by Inhibiting Translation of AIB1 mRNA. Mol Cell Biol 2006; 26: 8191–8201. https://doi.org/10.1128/MCB.00242-06

[44] MATSUBARA H, TAKEUCHI T, NISHIKAWA E, YANAGISAWA K, HAYASHITA Y et al. Apoptosis induction by antisense oligonucleotides against miR-17–5p and miR-20a in lung cancers overexpressing miR-17–92. Oncogene 2007; 26: 6099–6105. https://doi.org/10.1038/sj.onc.1210425

[45] WANG M, GU H, WANG S, QIAN H, ZHU W ET AL. Circulating miR-17–5p and miR-20a: molecular markers for gastric cancer. Mol Med Rep 2012; 5: 1514–1520.

[46] WU Y, XIE R, LIU X, WANG J, PENG Y et al. Knockdown of FOXK1 alone or in combination with apoptosis-inducing 5-FU inhibits cell growth in colorectal cancer. Oncology Rep 2016; 36: 2151–2159. https://doi.org/10.3892/or.2016.5041

[47] MITTAL S, ASLAM A, DOIDGE R, MEDICA R, WINKLER GS. The Ccr4a (CNOT6) and Ccr4b (CNOT6L) deadenylase subunits of the human Ccr4–Not complex contribute to the prevention of cell death and senescence. Mol Biol Cell 2011; 22: 748–758. https://doi.org/10.1091/mbc.E10-11-0898

[48] WANG H, MORITA M, YANG X, SUZUKI T, YANG W et al. Crystal structure of the human CNOT6L nuclease domain reveals strict poly(A) substrate specificity. EMBO J 2010; 29: 2566–2576. https://doi.org/10.1038/emboj.2010.152

[49] QU J, LI M, ZHONG W, HU C. Competing endogenous RNA in cancer: a new pattern of gene expression regulation. Int J Clin Exp Med 2015; 8: 17110–17116.