



Biological analysis of chronic lymphocytic leukemia: integration of mRNA and microRNA expression profiles

L. Dong¹, K.H. Bi¹, N. Huang¹ and C.Y. Chen²

¹Department of Hematology, QianFoShan Hospital Affiliated to Shandong University, Jinan, China

²Department of Hematology, Qilu Hospital of Shandong University, Jinan, China

Corresponding author: C.Y. Chen

E-mail: chunyanchenql@yeah.net

Genet. Mol. Res. 15 (1): gmr.15017170

Received August 8, 2015

Accepted October 15, 2015

Published January 8, 2016

DOI <http://dx.doi.org/10.4238/gmr.15017170>

ABSTRACT. Chronic lymphocytic leukemia (CLL) is a disease that involves progressive accumulation of nonfunctioning lymphocytes and has a low cure rate. There is an urgent requirement to determine the molecular mechanism underlying this disease in order to improve the early diagnosis and treatment of CLL. In this study, genes differentially expressed between CLL samples and age-matched controls were identified using microRNA (miRNA) and mRNA expression profiles. Differentially expressed (DE) miRNA targets were predicted by combining five algorithms. Common genes were obtained on overlapping the DE mRNA and DE miRNA targets. Then, network and module analyses were performed. A total of 239 miRNA targets were predicted and 357 DE mRNAs were obtained. On intersecting miRNA targets and DE mRNAs, 33 common genes were obtained. The protein-protein interaction network and module analysis identified several crucial genes and modules that might be associated with the development of CLL. These DE mRNAs were significantly enriched in the hematopoietic cell lineage ($P = 2.58E-4$), mitogen-activated protein kinase signaling pathway ($P = 0.0025$), and leukocyte transendothelial migration pathway

($P = 0.0026$). Thus, we conducted biological analysis on integration of DE mRNAs and DE miRNAs in CLL, determined gene expression patterns, and screened out several important genes that might be related to CLL.

Key words: Chronic lymphocytic leukemia; MicroRNA; Module analysis; Differentially expressed genes

INTRODUCTION

Chronic lymphocytic leukemia (CLL) is one of the most commonly diagnosed leukemias managed by practicing hematologists (Byrd et al., 2004). It has unique epidemiologic, biologic, and clinical features (Caligaris-Cappio and Ghia, 2008). To date, many studies have reported the importance of genes in the diagnosis and treatment of CLL. Rosenwald et al. (2001) reported that CLL can be characterized by a common gene expression “signature”, irrespective of the immunoglobulin (Ig) mutational status. CLL can be diagnosed by the mutational status of immunoglobulin heavy-chain variable-region (IgV_H) genes in leukemic cells (Crespo et al., 2003). The clinical course and overall survival of patients with CLL can be predicted by genomic features, including unmutated IgV_H genes, del(11q22.3), del(17p13.1), and p53 mutations (Grever et al., 2007). Gene expression profiling has potential pathogenetic and clinical relevance (Klein et al., 2001). The relationship between different pathways and CLL has also been studied previously. It was indicated that the Akt/Mcl-1 pathway plays a key role in CLL, which might inhibit multiple environmental signals that promote leukemic cell survival *in vivo* and may therefore provide a novel strategy for the treatment of CLL (Longo et al., 2008). Inhibitors of pathways, including protein kinase C and PI3-kinase, could be combined with the drugs used in the treatment of B-cell CLL (Barragán et al., 2002).

MicroRNAs (miRNAs) comprise a large family of highly conserved noncoding genes thought to be involved in temporal and tissue-specific gene regulation (Calin et al., 2002). They are involved in various biological processes, including cell proliferation and cell death during development, stress resistance, and fat metabolism, through the regulation of gene expression (Ambros, 2003). miRNAs can regulate the expression of target genes by binding to complementary sites in their transcripts, thereby leading to translational repression or transcript degradation (Griffiths-Jones et al., 2007). Differentially expressed (DE) miRNAs can contribute to the development and progression of cancers (Volinia et al., 2006). It was reported that miRNA expression patterns had relevance to the biological and clinical behaviors of CLL (Calin et al., 2004). Fulci et al. (2007) also identified a set of DE miRNAs that were associated with prognostic factors and disease progression in CLL.

Genome-wide discovery of signatures could provide important insights into the biological mechanisms underlying CLL. High throughput microarray technology, a promising clinical tool, provides methods to study disease-specific transcriptional changes simultaneously. Use of data that have been deposited in publically available data repositories recently can enable secondary use of existing public microarray data to predict outcome and biomarkers for CLL (Wray et al., 2014). The gene expression profiles of mRNA and miRNA will be beneficial for the diagnosis and targeted treatment of CLL in clinical practice (Yoo et al., 2009). However, although human tumors have been profiled extensively by genomics-based studies, little is known about how their global expression alterations form a network that contributes to aggressive disease and poor outcomes. Therefore, system biology analysis is required for such a study. The five key components are the

global measurements of biomolecules, the integration of biological information, the identification of molecular responses to perturbations, the building of testable models, and the refinement of models through testing of these hypotheses (Cho et al., 2011).

To better understand the complex pathology of CLL and identify molecular networks involved in CLL, miRNA and mRNA expression profiles of CLL cases and various normal B cell subpopulations were compared to identify DE miRNA and DE mRNA. In order to improve the accuracy of miRNA target prediction, five algorithms, that is, miRanda (Enright et al., 2003), miRDB (Wang, 2008), miRWalk (Dweep et al., 2011), RNA22 (Miranda et al., 2006), and TargetScan (Lewis et al., 2003), were used to predict miRNA target genes simultaneously. On intersecting the miRNA targets and DE mRNAs, the overlapped genes were selected as common genes, which are considered as more robust genetic markers. Next, a systematic approach was employed to investigate genes differentially expressed between CLL cases and normal controls; this approach involved the protein-protein interaction (PPI) network, centralities analysis, module analysis, and functional enrichment analysis. Thus, this study provides insights into CLL development at the molecular level and provides methods for advanced diagnostics and determining the prognosis of CLL.

MATERIAL AND METHODS

Samples

The microarray miRNA and mRNA expression profiles of CLL (Accession numbers: E-MTAB-1454 and E-GEOD-22529, respectively) were deposited in the ArrayExpress (<http://www.ebi.ac.uk/arrayexpress/>) database. For E-MTAB-1454, the miRNA expression profiles of 217 CLL cases and 37 samples of normal B cell subpopulations were compared. For E-GEOD-22529, 41 CLL tumor samples and 11 age-matched controls were used.

Identification of DE miRNAs and DE mRNAs

The files named CLL-Normalized-data-QLUCORE and GPL96-tbl-1 in the platform annotation files supported by Affymetrix (Santa Clara, CA, USA) were used to map the relationship between the probes and gene symbols. The average value of the gene symbol with multiple probes was obtained for further analysis.

Two comparison experiments were conducted using the linear models for microarray data (LIMMA) package. DE miRNAs and mRNAs were identified by assimilating a set of gene-specific *t*-tests with the threshold of false-discovery rate (FDR) ≤ 0.05 and $\log_{2}FC > 2$.

miRNA target gene predictions

The DE miRNAs identified in the last step were used for target gene prediction. The miRNA targets were predicted by an online software miRWalk (Dweep et al., 2011) (<http://mirwalk.uni-hd.de/>), a comprehensive database on miRNAs. In the present study, five algorithms were used for miRNA target gene prediction, that is, miRanda (Enright et al., 2003), miRDB (Wang, 2008), miRWalk (Dweep et al., 2011), RNA22 (Miranda et al., 2006), and Targetscan (Lewis et al., 2003). Only if a gene was confirmed by all five algorithms (SUM = 5), the gene could be identified as a target gene. Target genes were selected as common genes if they overlapped with DE miRNAs in E-GEOD-22529.

Identification of a complex network

Generally, proteins seldom accomplish their biological functions independently; therefore, it is necessary to comprehend their interactions by studying larger functional groups of proteins. In the present study, the PPI data were downloaded from the STRING database (<http://string.embl.de/>). Then, the DE genes identified were imported into the interaction network, and the interactions were screened with both end nodes having genes.

In the network, the regulatory impact factor (RIF) metric was implemented to identify the key transcription factors (TFs). The RIF metric was originally developed to identify causal regulators from gene expression data for two contrasting conditions (e.g., healthy vs disease) (Fortes et al., 2011). RIF can combine the change between the TFs and the DE genes. The equation for RIF is as follows:

$$RIF(TF_j) = \frac{1}{n_{de}} \sum_{i=1}^{n_{de}} \left[(e1_j \times r_{1i})^2 - (e2_j \times r_{2i})^2 \right] \quad (\text{Equation 1})$$

where n_{de} is the number of DE genes; $e1$ ($e2$) is the expression value of DE gene j in condition 1 (condition 2); $r1_{ij}$ ($r2_{ij}$) is the correlation of TF_j and DE gene i in condition 1 (condition 2). Then, the interaction network of the DE genes with the RIF value was constructed.

Centrality-based analysis of complex networks

To identify relevant nodes of a biological network, protocols for analysis integrating centrality analysis and laboratory experimental data are needed. In this study, stress (Scardoni and Laudanna, 2012), as a node centrality index, was used for network analysis. Stress can be calculated by measuring the number of shortest paths passing through a node. The stress of a node in a biological network can indicate the relevance of a protein as being functionally capable of holding together communicating nodes. The higher this value, the higher the relevance of the protein in connecting regulatory molecules. The stress could simply indicate that a molecule is highly involved in cellular processes. To calculate the stress of a node v , first, all shortest paths are calculated. Then, the number of shortest paths passing through v is counted. A “stressed” node is a node traversed by a high number of shortest paths. The stress of a node was calculated as follows:

$$C_{st}(v) = \sum_{s \neq v} \sum_{t \neq v} \sigma_{st}(v) \quad (\text{Equation 2})$$

where s_{st} is the number of shortest paths between vertexes s and t , and $s_{st}(v)$ is the number of shortest paths between s and t passing through the vertex v .

In our study, nodes with stress threshold < 3467 (average score) were filtered.

Module detection for the interaction network

In order to explore the highly interconnected regions and their functions in the network, we implemented the module analysis. In the present study, the Molecular Complex Detection (MCODE) algorithm (Bader and Hogue, 2003) was employed to perform module detection in the

PPI network. The MCODE algorithm mainly includes three stages: vertex weighting, complex prediction, and optional post-processing. At the stage of vertex weighting, all vertices based on their local network density were weighted using the highest k-core of the vertex neighborhood. At the stage of complex prediction, the vertex-weighted graph was first taken as input and a complex with the highest weighted vertex was seeded and then recursively moved outward from the seed vertex. The complex included vertices whose weight were above a given threshold. The threshold was a given percentage away from the weight of the seed vertex. At the third stage, complexes with core less than 2 (graph of minimum degree 2) were filtered. Next, the “fluff” option and “haircut” option can be run as options. The “fluff” option was used to increase the size of the complex according to a given “fluff” parameter between 0 and 1. The “haircut” option was used to remove the vertices that were singly connected to the core complex, and the complexes obtained were all 2-cored. When both options are performed, “fluff” is run first and then “haircut”. The software for the MCODE algorithm was obtained from <http://baderlab.org/Software/MCODE>.

Functional enrichment and pathway-enrichment analysis

The Gene Ontology (GO) annotation system can provide a set of expert-curated terms that describe biological entities with respect to three aspects, biological process (BP), molecular function (MF), and cellular component (CC), and are organized into a hierarchical structure. Kyoto Encyclopedia of Genes and Genomes (KEGG) is a knowledge base for systematic analysis of gene functions, linking genomic information with higher order functional information. In this study, the GO functional enrichment and KEGG pathway-enrichment analysis for DE genes were performed by using the online tool Database for Annotation, Visualization, and Integrated Discovery (DAVID). GO terms and KEGG pathways with P values less than 0.01 were selected based on the expression analysis systematic explored (EASE) test implemented in the DAVID. The principle of EASE is as follows:

$$p = \frac{\binom{a+b}{a} \binom{c+d}{c}}{\binom{n}{a+c}} \quad (\text{Equation 3})$$

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

RESULTS

Identification of DE miRNAs, DE mRNAs, and miRNA targets

For E-MTAB-1454, 10 DE miRNAs (3 upregulated miRNAs and 7 downregulated miRNAs) were obtained, for which 239 target genes (included 71 upregulated and 168 downregulated targets) were predicted using the miRWalk software with SUM = 5. The number of DE miRNA targets predicted by five target gene prediction algorithms is shown in Table 1. For E-GEO-22529, 357 DE mRNAs were obtained with the criterion of FDR \leq 0.05, including 203 upregulated genes

and 154 downregulated genes. On overlapping the DE miRNA targets and DE mRNAs, 33 genes were identified as common genes. The common genes were selected for further research.

Table 1. Number of target genes of differentially expressed miRNAs predicted by five target gene prediction algorithms.

Expression change	DE miRNA	Target number
Upregulated	hsa-miR-150	20
	hsa-miR-34a	51
	hsa-miR-126	0
Downregulated	hsa-miR-193b	7
	hsa-miR-365	1
	hsa-miR-181b	69
	hsa-miR-33b	0
	hsa-miR-582-5p	0
	hsa-miR-181a	56
	has-miR-363	35

Comprehensive analysis of the complex network

The PPI network with 264 nodes and 1101 edges was identified using Cytoscape 3.0 (Figure 1). In the RIF metric, nine genes were obtained with the RIF value and eight were TFs. These eight genes were *JUNB*, *MYC*, *TCF4*, *CEBPB*, *FOS*, *NFIL3*, *XBP1*, and *CUX1*.

When analyzing the centralities of the PPI network, the stress index was implemented. The higher the stress value, the higher the relevance of the protein in connecting regulatory molecules. The PPI network constructed based on the stress index is shown in Figure 2. Genes with high stress (threshold >3467) in the network are marked in Figure 2. The top 5 ranked genes were *MYC*, *CD14*, *FOS*, *SMAD3*, and *CCND2*.

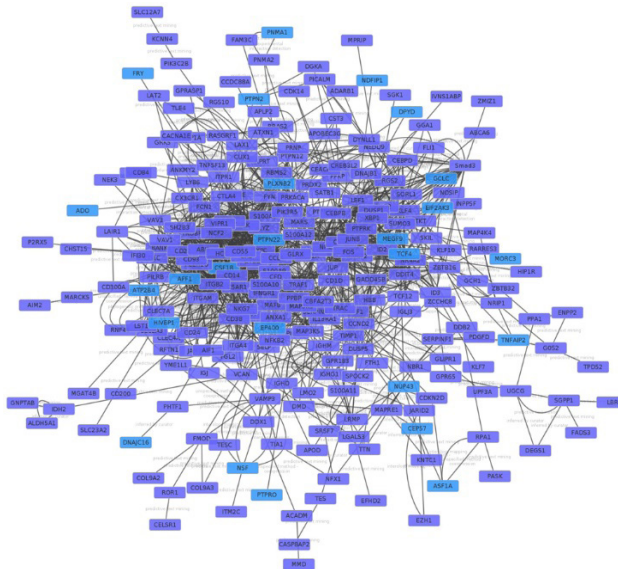


Figure 1. Protein-protein interaction network of the differentially expressed mRNAs constructed in chronic lymphocytic leukemia. The blue node represents differentially expressed mRNA. The green node represents target genes of both miRNA and differentially expressed mRNA, i.e., common target genes.

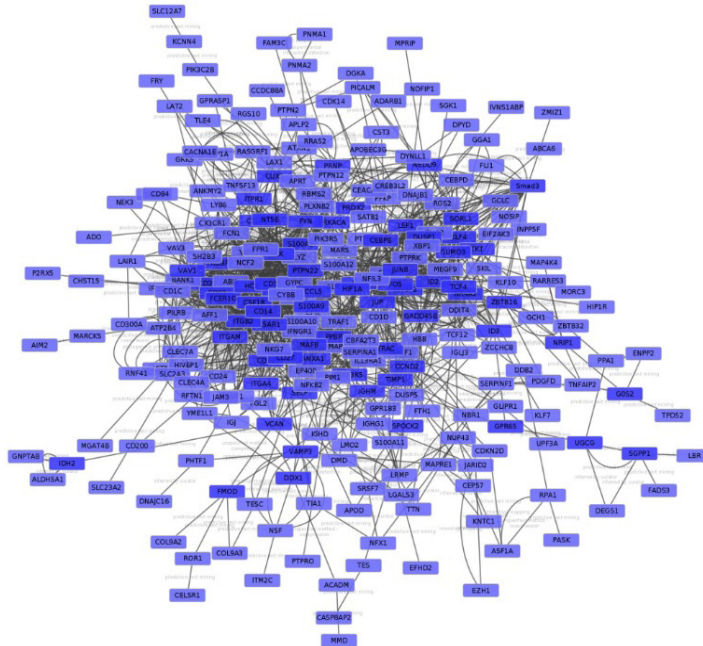


Figure 2. Protein-protein interaction network constructed based on the stress index for centrality analysis. The blue node represents differentially expressed mRNA. The dark blue node represents a gene that has high stress value (threshold >3467) in the network.

Module detection for the interaction network

The MCODE algorithm was used to identify the clusters in the PPI network. On setting Degree Cutoff = 2, K-core = 4, Max. depth = 100, we obtained five modules, which are shown in Figure 3. Module 1 had the highest degree and the most number of genes, and it included one of the common genes *CSF1R* and three TFs (Figure 3A). Module 2 had one target gene, *CSF1R*, and four TFs (Figure 3B). Module 4 included two miRNA targets and two TFs (Figure 3D). There were no target genes and TFs in module 3 and module 5 (Figure 3C and E).

Functional enrichment and pathway-enrichment analysis

Using GO analysis for the DE genes, we identified 126 significant terms, which were classified in three GO categories, that is, BP (102), MF (17), and CC (7). In BP, the DE genes were mostly related to the immune response ($P = 1.34E-12$), defense response ($P = 3.09E-10$), inflammatory response ($P = 1.37E-7$), leukocyte activation ($P = 2.64E-7$), and lymphocyte activation ($P = 2.79E-7$). The most significant terms of MF and CC were intrinsic to plasma membrane ($P = 4.35E-6$) and protein dimerization activity ($P = 1.61E-6$), respectively. The most significantly enriched GO terms are shown in Table 2.

Pathway enrichment analysis based on KEGG database yielded three significant enrichment pathways (Table 3). They were hematopoietic cell lineage ($P = 2.58E-4$), mitogen-activated protein kinase (MAPK) signaling pathway ($P = 0.0025$), and leukocyte transendothelial migration ($P = 0.0026$).

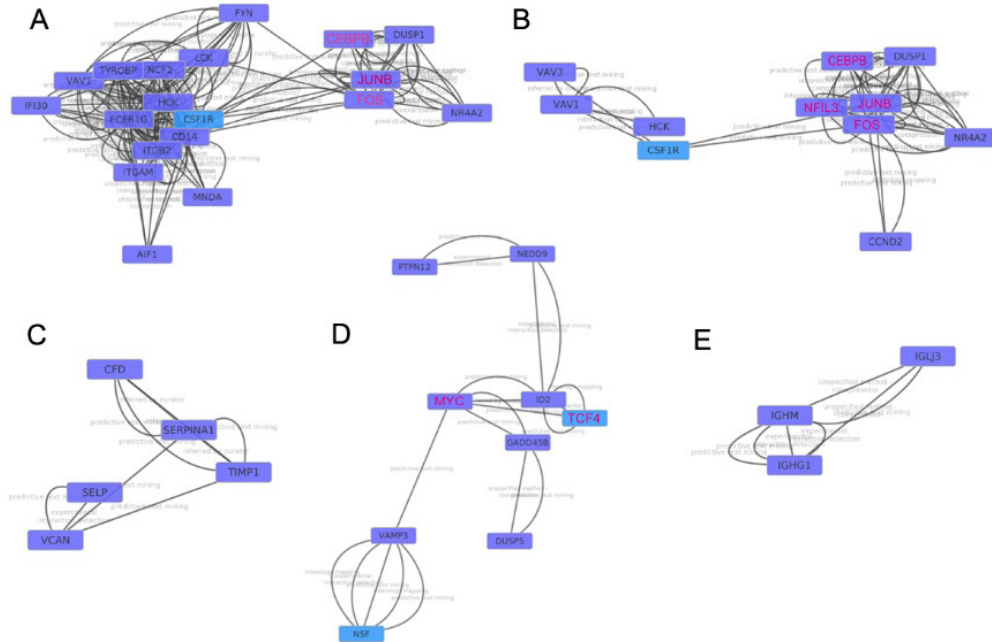


Figure 3. Module detection of the protein-protein interaction network. **A.** Module 1; **B.** Module 2; **C.** Module 3; **D.** Module 4; **E.** Module 5. The blue node represents differentially expressed mRNA. The green node represents a common gene. The red node represents a transcription factor.

Table 2. Significant functional enrichment of differentially expressed genes.

	ID	Term	P value	
BP	GO:0006955	Immune response	1.34E-012	
	GO:0006952	Defense response	3.09E-010	
	GO:0006954	Inflammatory response	1.37E-007	
	GO:0045321	Leukocyte activation	2.64E-007	
	GO:0046649	Lymphocyte activation	2.79E-007	
	GO:0009611	Response to wounding	3.85E-007	
	GO:0001775	Cell activation	9.97E-007	
	GO:0032844	Regulation of homeostatic process	3.30E-005	
	GO:0002684	Positive regulation of immune system process	5.56E-005	
	GO:0010033	Response to organic substance	7.58E-005	
	MF	GO:0046983	Protein dimerization activity	1.61E-006
		GO:0003823	Antigen binding	1.25E-004
GO:0019901		Protein kinase binding	1.93E-004	
GO:0001846		Opsonin binding	2.68E-004	
GO:0019900		Kinase binding	2.79E-004	
GO:0042802		Identical protein binding	5.47E-004	
GO:0016564		Transcription repressor activity	6.90E-004	
CC	GO:0031226	Intrinsic to plasma membrane	4.35E-006	
	GO:0005887	Integral to plasma membrane	5.29E-006	
	GO:0005886	Plasma membrane	6.64E-005	
	GO:0044459	Plasma membrane part	3.51E-004	
	GO:0000267	Cell fraction	9.76E-004	

Table 3. Significantly enriched pathways of differentially expressed genes.

ID	Term	P value	Genes
hsa04640	Hematopoietic cell lineage	2.58E-4	<i>CD38, CD55, CR2, CD1C, CD22, ITGA4, ITGAM, CD14, CSF1R, CD1D</i>
hsa04010	MAPK signaling pathway	2.47E-3	<i>PTPN7, NFKB2, DUSP5, MAP4K4, FOS, MAP3K5, DUSP1, ARRB2, MAP3K3, RASGRF1, RRAS2, PRKACB, GADD45B, MYC, CACNA1A, CD14</i>
hsa04670	Leukocyte transendothelial migration	2.56E-3	<i>CYBB, VAV3, NCF2, SIPA1, PIK3R5, ITGB2, ITGA4, VAV1, JAM3, ITGAM</i>

DISCUSSION

To understand the pathology of CLL, study on the molecular mechanism underlying the disease is needed. A deeper understanding of these mechanisms may help identify new diagnostic methods and targets for therapy. In this study, the miRNA and mRNA expression profiles of CLL cases and normal controls were used for bioinformatic analysis.

In recent years, studies on expression profile-based prognostic markers have expanded. Gene expression profiling studies have identified a large number of genes with differential expression between prognostic subgroups in CLL (Sevov et al., 2012). Kaderi et al. (2011) indicated that high lipoprotein lipase and *CLLU1* mRNA expression were strong markers of poor clinical outcome in CLL. A study of different immunogenetic subsets of CLL revealed several DE miRNAs and indicated that specific DE miRNAs among CLL subgroups might modulate the biological and clinical behaviors of the CLL clones. Most recently, Tavolaro et al. (2015) revealed that CLL progression is sustained by the upregulation of miR-132 and miR-212. Previous studies demonstrated significant differences in the expression of multiple genes of CLL. In order to obtain more robust genetic markers for CLL, we attempted to improve our study in two ways: improvement of predictive accuracy for miRNA targets by combining multiple bioinformatic strategies and improvement in determining the genes common between miRNA targets and DE mRNA.

When analyzing the stress index of the PPI network, several genes, such as *MYC*, *CD14*, *FOS*, *SMAD3*, and *CCND2*, showed high stress values, which implied high relevance in connecting regulatory molecules. *MYC* (stress = 52070), as a TF included in module 4, showed the highest stress value, and was involved in the MAPK signaling pathway that we identified. *MYC* is a multifunctional gene that plays a role in cell cycle progression, apoptosis, and cellular transformation (de Alboran et al., 2001). A study by Nie et al. (2012) demonstrated that *MYC* is a universal amplifier of genes expressed in lymphocytes and embryonic stem cells. Genome-wide expression studies showed that CLL cells responded with increased expression of *MYC* and other genes associated with B-cell receptor activation amplified by *MYC* (Pede et al., 2013).

In module 4, *MYC* was directly connected to several genes, including *TCF4*, *ID2*, *GADD45B*, and *VAMP3*. *TCF4* was both a TF and a common gene. *TCF4* has been reported to play an important role in leukemia progression, and its mutants could downregulate the target gene expression (Chang et al., 2006). A study on the Wnt pathway showed that the *TCF4* gene, a known downstream target of this pathway, was overexpressed in CLL (Gutierrez et al., 2010). Moreover, *TCF4* could promote neoplastic transformation (Kolligs et al., 2002) and was deregulated in MLL-rearrangement leukemia (Li et al., 2009). Our finding, combined with these previous results, indicated that *MYC* and *TCF4* played an important role in CLL.

In the present study, we predicted the target genes of DE miRNAs associated with CLL. On comparison with DE mRNAs, we obtained a total of 33 common genes; 25 of them were mapped on the PPI network. These common genes were considered as a more reliable genetic signature.

CSF1R, included in modules 1 and 2, was a common gene and a target of several TFs. *CSF1R* is a tyrosine kinase receptor that is essential for macrophage differentiation and thus occupies a central role in hematopoiesis (Bonifer and Hume, 2008). *CSF1R* can mediate the effect of *CSF1*, which regulates the survival proliferation, and differentiation of mononuclear phagocytic cells (Cecchini et al., 1994). Previous gene function analysis had indicated that *CSF1R* was aberrantly expressed in acute myeloid leukemia (Casas et al., 2003). The PU.1-mediated upregulation of *CSF1R* might be a useful therapeutic target for leukemia (Aikawa et al., 2010). Combined with the crucial role of *CSF1R* in the network and modules, it was predicted to be associated with the development and progression of CLL. In addition, TF-encoding *JUNB*, *CEBPB*, *FOS*, and the common gene *NSF*, also showed high stress values, indicating their potential roles in CLL.

In this study, module analysis was used to build groups of the PPI network, which contained functionally related genes, such as genes in a specific pathway. In our study, we obtained five modules. Modules 1, 2, and 4 included several common genes and TFs and had more DE genes and higher stress values than the other modules. These genes were mainly enriched in the hematopoietic cell lineage, MAPK signaling pathway, and leukocyte transendothelial migration, indicating their potential important role in CLL.

In conclusion, we obtained 33 common genes by intersecting DE miRNA target genes and DE mRNAs between CLL patients and normal controls. The PPI network of DE genes was constructed, and several common genes and TFs were mapped on it. These DE genes mainly participate in the hematopoietic cell lineage, MAPK signaling pathway, and leukocyte transendothelial migration pathway. Modules 1, 2 and 4 were predicted to be associated with CLL. In conclusion, this study can help us to understand the pathogeny of CLL, which is helpful for the diagnosis and treatment of CLL.

Conflicts of interest

The authors declare no conflict of interest.

REFERENCES

- Aikawa Y, Katsumoto T, Zhang P, Shima H, et al. (2010). PU.1-mediated upregulation of *CSF1R* is crucial for leukemia stem cell potential induced by MOZ-TIF2. *Nat. Med.* 16: 580-585, 1p, 585. <http://dx.doi.org/10.1038/nm.2122>
- Ambros V (2003). MicroRNA pathways in flies and worms: growth, death, fat, stress, and timing. *Cell* 113: 673-676. [http://dx.doi.org/10.1016/S0092-8674\(03\)00428-8](http://dx.doi.org/10.1016/S0092-8674(03)00428-8)
- Bader GD and Hogue CW (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* 4: 2. <http://dx.doi.org/10.1186/1471-2105-4-2>
- Barragán M, Bellosillo B, Campàs C, Colomer D, et al. (2002). Involvement of protein kinase C and phosphatidylinositol 3-kinase pathways in the survival of B-cell chronic lymphocytic leukemia cells. *Blood* 99: 2969-2976. <http://dx.doi.org/10.1182/blood.V99.8.2969>
- Bonifer C and Hume DA (2008). The transcriptional regulation of the Colony-Stimulating Factor 1 Receptor (*csf1r*) gene during hematopoiesis. *Front. Biosci.* 13: 549-560. <http://dx.doi.org/10.2741/2700>
- Byrd JC, Stilgenbauer S and Flinn IW (2004). Chronic lymphocytic leukemia. *Hematology (Am Soc Hematol Educ Program)* 2004: 163-183. <http://dx.doi.org/10.1182/asheducation-2004.1.163>
- Caligaris-Cappio F and Ghia P (2008). Novel insights in chronic lymphocytic leukemia: are we getting closer to understanding the pathogenesis of the disease? *J. Clin. Oncol.* 26: 4497-4503. <http://dx.doi.org/10.1200/JCO.2007.15.4393>
- Calin GA, Dumitru CD, Shimizu M, Bichi R, et al. (2002). Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. USA* 99: 15524-15529. <http://dx.doi.org/10.1073/pnas.242606799>
- Calin GA, Liu C-G, Sevignani C, Ferracin M, et al. (2004). MicroRNA profiling reveals distinct signatures in B cell chronic

- lymphocytic leukemias. *Proc. Natl. Acad. Sci. USA* 101: 11755-11760. <http://dx.doi.org/10.1073/pnas.0404432101>
- Casas S, Nagy B, Elonen E, Aventin A, et al. (2003). Aberrant expression of HOXA9, DEK, CBL and CSF1R in acute myeloid leukemia. *Leuk. Lymphoma* 44: 1935-1941. <http://dx.doi.org/10.1080/1042819031000119299>
- Cecchini MG, Dominguez MG, Mocchi S, Wetterwald A, et al. (1994). Role of colony stimulating factor-1 in the establishment and regulation of tissue macrophages during postnatal development of the mouse. *Development* 120: 1357-1372.
- Chang HR, Cheng TL, Liu TZ, Hu HS, et al. (2006). Genetic and cellular characterizations of human TCF4 with microsatellite instability in colon cancer and leukemia cell lines. *Cancer Lett.* 233: 165-171. <http://dx.doi.org/10.1016/j.canlet.2005.03.011>
- Cho J-H, Gelinax R, Wang K, Etheridge A, et al. (2011). Systems biology of interstitial lung diseases: integration of mRNA and microRNA expression changes. *BMC Med. Genomics* 4: 8. <http://dx.doi.org/10.1186/1755-8794-4-8>
- Crespo M, Bosch F, Villamor N, Bellosillo B, et al. (2003). ZAP-70 expression as a surrogate for immunoglobulin-variable-region mutations in chronic lymphocytic leukemia. *N. Engl. J. Med.* 348: 1764-1775. <http://dx.doi.org/10.1056/NEJMoa023143>
- de Alboran IM, O'Hagan RC, Gärtner F, Malynn B, et al. (2001). Analysis of C-MYC function in normal cells via conditional gene-targeted mutation. *Immunity* 14: 45-55. [http://dx.doi.org/10.1016/S1074-7613\(01\)00088-7](http://dx.doi.org/10.1016/S1074-7613(01)00088-7)
- Dweep H, Sticht C, Pandey P and Gretz N (2011). miRWalk--database: prediction of possible miRNA binding sites by "walking" the genes of three genomes. *J. Biomed. Inform.* 44: 839-847. <http://dx.doi.org/10.1016/j.jbi.2011.05.002>
- Enright AJ, John B, Gaul U, Tuschl T, et al. (2003). MicroRNA targets in Drosophila. *Genome Biol.* 5: R1. <http://dx.doi.org/10.1186/gb-2003-5-1-r1>
- Fortes MR, Reverter A, Nagaraj SH, Zhang Y, et al. (2011). A single nucleotide polymorphism-derived regulatory gene network underlying puberty in 2 tropical breeds of beef cattle. *J. Anim. Sci.* 89: 1669-1683. <http://dx.doi.org/10.2527/jas.2010-3681>
- Fulci V, Chiaretti S, Goldoni M, Azzalin G, et al. (2007). Quantitative technologies establish a novel microRNA profile of chronic lymphocytic leukemia. *Blood* 109: 4944-4951. <http://dx.doi.org/10.1182/blood-2006-12-062398>
- Grever MR, Lucas DM, Dewald GW, Neuberger DS, et al. (2007). Comprehensive assessment of genetic and molecular features predicting outcome in patients with chronic lymphocytic leukemia: results from the US Intergroup Phase III Trial E2997. *J. Clin. Oncol.* 25: 799-804. <http://dx.doi.org/10.1200/JCO.2006.08.3089>
- Griffiths-Jones S, Saini HK, van Dongen S and Enright AJ (2007). miRBase: tools for microRNA genomics. *Nucleic Acids Res.* 36: D154-D158. <http://dx.doi.org/10.1093/nar/gkm952>
- Gutierrez A, Jr., Tschumper RC, Wu X, Shanafelt TD, et al. (2010). LEF-1 is a prosurvival factor in chronic lymphocytic leukemia and is expressed in the preleukemic state of monoclonal B-cell lymphocytosis. *Blood* 116: 2975-2983. <http://dx.doi.org/10.1182/blood-2010-02-269878>
- Kaderi MA, Kanduri M, Buhl AM, Sevov M, et al. (2011). LPL is the strongest prognostic factor in a comparative analysis of RNA-based markers in early chronic lymphocytic leukemia. *Haematologica* 96: 1153-1160. <http://dx.doi.org/10.3324/haematol.2010.039396>
- Klein U, Tu Y, Stolovitzky GA, Mattioli M, et al. (2001). Gene expression profiling of B cell chronic lymphocytic leukemia reveals a homogeneous phenotype related to memory B cells. *J. Exp. Med.* 194: 1625-1638. <http://dx.doi.org/10.1084/jem.194.11.1625>
- Kolligs FT, Nieman MT, Winer I, Hu G, et al. (2002). ITF-2, a downstream target of the Wnt/TCF pathway, is activated in human cancers with beta-catenin defects and promotes neoplastic transformation. *Cancer Cell* 1: 145-155. [http://dx.doi.org/10.1016/S1535-6108\(02\)00035-1](http://dx.doi.org/10.1016/S1535-6108(02)00035-1)
- Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, et al. (2003). Prediction of mammalian microRNA targets. *Cell* 115: 787-798. [http://dx.doi.org/10.1016/S0092-8674\(03\)01018-3](http://dx.doi.org/10.1016/S0092-8674(03)01018-3)
- Li Z, Luo RT, Mi S, Sun M, et al. (2009). Consistent deregulation of gene expression between human and murine MLL rearrangement leukemias. *Cancer Res.* 69: 1109-1116. <http://dx.doi.org/10.1158/0008-5472.CAN-08-3381>
- Longo PG, Laurenti L, Gobessi S, Sica S, et al. (2008). The Akt/Mcl-1 pathway plays a prominent role in mediating antiapoptotic signals downstream of the B-cell receptor in chronic lymphocytic leukemia B cells. *Blood* 111: 846-855. <http://dx.doi.org/10.1182/blood-2007-05-089037>
- Miranda KC, Huynh T, Tay Y, Ang YS, et al. (2006). A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell* 126: 1203-1217. <http://dx.doi.org/10.1016/j.cell.2006.07.031>
- Nie Z, Hu G, Wei G, Cui K, et al. (2012). c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. *Cell* 151: 68-79. <http://dx.doi.org/10.1016/j.cell.2012.08.033>
- Pede V, Rombout A, Vermeire J, Naessens E, et al. (2013). CLL cells respond to B-Cell receptor stimulation with a microRNA/mRNA signature associated with MYC activation and cell cycle progression. *PLoS One* 8: e60275. <http://dx.doi.org/10.1371/journal.pone.0060275>
- Rosenwald A, Alizadeh AA, Widhopf G, Simon R, et al. (2001). Relation of gene expression phenotype to immunoglobulin mutation genotype in B cell chronic lymphocytic leukemia. *J. Exp. Med.* 194: 1639-1648. <http://dx.doi.org/10.1084/jem.194.11.1639>
- Scardoni G and Laudanna C (2012). Centralities based analysis of complex networks. In: *New frontiers in graph theory* (Zhang Y, ed.). InTech Press, Rijeka, 323-348.

- Sevov M, Rosenquist R and Mansouri L (2012). RNA-based markers as prognostic factors in chronic lymphocytic leukemia. *Expert Rev. Hematol.* 5: 69-79. <http://dx.doi.org/10.1586/ehm.11.80>
- Tavolaro S, Colombo T, Chiaretti S, Peragine N, et al. (2015). Increased chronic lymphocytic leukemia proliferation upon IgM stimulation is sustained by the upregulation of miR-132 and miR-212. *Genes Chromosomes Cancer* 54: 222-234. <http://dx.doi.org/10.1002/gcc.22236>
- Volinia S, Calin GA, Liu C-G, Ambs S, et al. (2006). A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc. Natl. Acad. Sci. USA* 103: 2257-2261. <http://dx.doi.org/10.1073/pnas.0510565103>
- Wang X (2008). miRDB: a microRNA target prediction and functional annotation database with a wiki interface. *RNA* 14: 1012-1017. <http://dx.doi.org/10.1261/rna.965408>
- Wray CJ, Ko TC and Tan FK (2014). Secondary use of existing public microarray data to predict outcome for hepatocellular carcinoma. *J. Surg. Res.* 188: 137-142. <http://dx.doi.org/10.1016/j.jss.2013.12.013>
- Yoo BK, Emdad L, Su ZZ, Villanueva A, et al. (2009). Astrocyte elevated gene-1 regulates hepatocellular carcinoma development and progression. *J. Clin. Invest.* 119: 465-477. <http://dx.doi.org/10.1172/JCI36460>