Structural Analysis of Leukemia Related Gene Network*

Liu Wei[1], Zhang Yuanyuan[2], Xu Xiuzhu[2], Yang Meixi[2], Xu Dashun[3] and Wang Shudong[2]

[1]*Department of Mathematics and Information, Ludong University, Yantai, Shandong 264025, China*
[2.] *College of Information Science and Engineering, Shandong University of Science and Technology, Qingdao, Shandong 266510, China*
[3.] *Department of Mathematics, Southern Illinois University Carbondale, Makanida 62958, USA*
Email: Wangshd2008@yahoo.com.cn ， Liuyiwei1030@yahoo.com.cn

Abstract *Genome forms gene networks in the complicated interactive ways and further study on the cancer-related gene networks can help to understand and predict many unknown biological functions of cancer-related genome, and then obtain the useful informations about molecular mechanism of the formation of cancer. In this research, using the method of mutual information, we construct gene networks corresponding to normal control group and diseased experimental group of acute lymphoblastic leukemia (abbreviated as ALL). Through contrasting the structure difference of two kinds of networks, we find 23 structural key genes of ALL with significant degree-difference, where 21 genes are confirmed to be closely related to the formation of ALL. The match ratio between the prediction by model and literature is up to 91.3%. According to the effectiveness of the method in this work, we can predict that the remaining two genes CXCL1 and TACSTD2 are closely related to ALL. The significant differences of the network structures between control and experimental groups will enlighten the biomedical scientists that the reason of normal organism suffering from leukemia maybe is the great changes of some genes in vivo. The finding of structural key genes of ALL will help the biomedical scientist to further research the pathogenesis of ALL.*

Keywords *Systems biology Gene network Mutual information Leukemia*

1. Introduction

Leukemia is one of the top 10 human cancers. Now many biomedical scientists are engaging in discovering the oncogenes and tumor-suppressor genes of leukemia, hope to know well the pathogenesis of leukemia, and find out the effective methods to treat leukemia. At present, besides doing clinical experiments directly, the main research ways to discover the pathogenesis of leukemia are based on the gene expression profiles data of DNA chip. For instance, the method of cluster analysis was used to analyze the pattern of genes expressed in leukemic blasts from 360 pediatric ALL patients. Distinct expression profiles identified each of the prognostically important leukemia subtypes, including T-ALL, E2A-PBX1, BCR-ABL,

*Supported by National Natural Science Foundation of China (Grant Nos. 60874036, 60503002) and SDUST Research Fund.

TEL-AML1, MLL rearrangement (Yeoh et al., 2002)[1]. The method of pattern recognition was utilized to analyze the gene expression profiles data of patients with leukemia(Yakovlev et al., 2002) [2]. The method of multivariate statistics was used to identify genes modulated by adhesion of human precursor B leukemia cells that regulate proliferation and apoptosis, highlighting new pathways that might provide insights into future therapy aiming at targeting apoptosis of leukemia cells (Astier et al., 2003) [3]. The analysis of neural network internal structure allowed the identification of specific phenotype markers and the extraction of peculiar associations among genes and physiological states. At the same time, the neural network outputs provided assignment to multiple classes, such as different pathological conditions or tissue samples, for previously unseen instances (Silvio et al., 2003)[4]. The method of graphical Gaussian models (GGMs) was used to describe gene association networks and to detect conditionally dependent genes (Korbinian Strimmer et al., 2005)[5]. A robust gene selection approach based on a hybrid between genetic algorithm and support vector machine was formalized and the major goal of this hybridization was to exploit fully their respective merits for identification of key feature genes (or molecular signatures) for a complex biological phenotype (Shaoqi Rao et al., 2005)[6]. The data-mining methods were applied to biomedical research. The results and how these results would affect the diagnosis and treatment of ALL in the future were discussed (Morton, Geoffrey, 2010)[7]. These above-mentioned methods can compensate for the deficiency of clinical trials, and they are of important significance to the prediction, diagnosis and treatment of cancer.

Systems biology is a newly interdiscipline after genomics, proteomics. Now, forward modeling and reverse modeling are two main methods to study systems biology. Modeling and analysis method of complex network is one of the important ways of reverse network modeling and has been widely applied to the study of complex biological system. The complex network models of biological system mainly include genetic networks, protein interaction networks, metabolic networks, signal networks and cellular networks, etc. Harald Lahm et al. researched the complex network of family of endogenous agglutinin in 2004[8], which would help to understand certain types of malignant phenotypes. Adriano V. Werhli et al. constructed Raf signal transduction network using Gaussian and Bayesian model respectively based on the data of system expression profiles in 2006[9]. Theodore J. Perkins et al. analyzed the gene regulatory network of drosophila using the method of reverse network modeling, which can explain the activation of genes[10]. Shudong Wang et al. constructed the logical network of arabidopsis genes under different external stimuli using reverse network modeling of information entropy, simulated and analyzed the dynamical behaviours of the obtained logic network [11].

In this work, based on the gene expression profiles of control group and experimental groups of ALL, we construct the mutual information network of healthy marrow, B_ALL and T_ALL respectively. Through contrasting the changes of network structures under different groups, we discover the significant difference in the networks for control and experimental groups. Furthermore, we find 23 structural key genes of ALL. The paper is organised as follows. We introduce the related background in the first part, discuss the analysis method of complex network in the second one, construct the mutual information gene network and give the main results in the third part, and analyze the biological significances of the experimental results combining with the gene functions in the fourth part.

2. Analysis Method of Complex Network

Complex network can show the complex relationships among large numbers of elements more clearly, and better explain the mutual influence relationships between the structures and functions. Therefore, further study on complex network of the genes of diseases can help to understand pathogenesis of diseases. Biological systems are composed of interactions and mutual regulations of many elements, such as genes, proteins, protein complexes and

transcription factors, etc. If these elements and the interactions and mutual regulations of elements can be simplified as nodes and edges respectively, then the complicated biological systems can be abstracted to complex networks, for example, undirected network, directed network, and weighted network etc. Now, the analysis method of complex network is mainly studying the overall (average path length, clustering coefficient, assortativity coefficient, degree distribution, etc.) and local (community structure, motif etc.) properties of network. In this work, we will analyze the structure of complex life systems to obtain further understanding its function, which is of important reference values for life scientists to predict and cure disease.

Let $G=(V,E)$ be a complex network with node-set $V=\{1,2,\cdots,N\}$ and edge-set E.

The statistics of complex network used in the work are as follows:

(1) Average degree (K)

The degree of a node is defined as the number of nodes adjacent to it. The average degree is the average of the degrees of all nodes in the network, denoted by K.

(2) Average path length (L)

The shortest path length is defined as the distance between any two nodes in the network. The average path length is the average of all of the shortest path lengths in the network, denoted by L.

(3) Average clustering coefficient (C)

The clustering coefficient of node *i* in the network is defined as the number of triangles dividing by the number of triples connected with node *i*. It reflects the tightness of the network connection. The average clustering coefficient is the average of the clustering coefficients of all nodes in the network, denoted by C.

(4) Modularity (Q)

Modularity is defined as

$$Q=\sum_i Q_i=\sum_i (e_{ij}-a_i^2)$$

where e_{ij} denotes the proportion of the edges connecting two different communities to all the edges of network, and a_i is the sum of each element of i^{th} row which denotes the proportion of the edges connecting with i^{th} community to all the edges of network. Modularity is the probability measure of community structure of network, and is mainly used to judge the quality of community structure.

(5) Average coreness (k)

k-core of a graph is defined as the remaining subgraph after repeatedly removing those nodes whose degrees are less than or equal to k. The coreness of a node denotes the depth of the node in the core. The average coreness is the average of the corenesses of the graph, denoted by k.

(6) Average betweenness centrality (b)

The betweenness centrality of a node is defined as the number of the shortest paths through the node in a network. It is an important index which can reflect the topological properties of the network. The average betweenness centrality is the average of the betweenness centralities of all nodes in the network, denoted by b.

(7) Non-isolated node proportion (I)

The isolated node of a network is defined as the node that is not adjacent to any other nodes. It usually doesn't work in the network. The non-isolated node proportion is the number of non-isolated node dividing the number of all nodes. It can reflect the intensive extent of the nodes associated in the network.

3. Construction and Analysis of Gene Network of Mutual Information

3.1 Establishment of Working Database

All the data in this research are from E13159 sample library in GPL570 of NCBI. E13159 sample library is composed of more than 2000 samples. We extract 40 samples from normal samples, B_ALL samples and T_ALL samples respectively. For description convenience, we refer to normal samples as the control group, B_ALL samples and T_ALL samples as two experimental groups to establish three original databases which contain more than 20,000 genes sample data of human genome. It is impossible to establish a gene network of mutual information with more than 20,000 genes using the present computer from the view of computational complexity. Therefore, we should process the original database firstly. According to 301 cancer-related genes we have mastered [12], we select 801 probes corresponding to these genes. If there are several probes corresponding to a gene, then we choose the probe that has the highest expression profiles corresponding to the gene. Through the above processing, we get 286 genes as the research object. Because we focus on gene network structure, the genes whose expression levels are almost 0 or 1 in all of the samples have no contribution to the differences of network structure. Hence, we delete the genes whose the proportions of data expression profiles of samples being 1 are greater than 90% or less than 15% in each library. At last we get 58, 73 and 85 genes from the normal group, the B_ALL group and the T_ALL group respectively to form the corresponding working databases. Next we will construct the gene networks of mutual information based on the 3 working databases.

3.2 Selection of Threshold

3.2.1 Comparison of Network Statistics

Firstly, we should discretize p_value of the working database appropriately. Specific methods are as follows: divide the range of p_value [0, 1] into 20 parts, and mark with 1, 2, …, 20, respectively. This kind of discretization has finer granularity than 0-1 discretization, and it lose less information.

Secondly, we can get a complete connected weighted network of all genes in each library by use of the mutual information formula, and the mutual information value is denoted by the weight. In order to highlight the specificity of network structure and obtain useful biological informations, the mutual information values of the complete connected network are need to be coarse-grained to seek threshold. As the distribution interval of mutual information value in each database is different, all the mutual information values of the 3 working databases are normalized, so that different databases can be comparable.

Finally, we analyze 7 statistics (average degree, average path length, average clustering coefficient, modularity, average coreness, average betweenness centrality, non-isolated node proportion) with the variation of thresholds between 0.1 and 0.9 in steps of 0.01, as shown in Fig. 1.

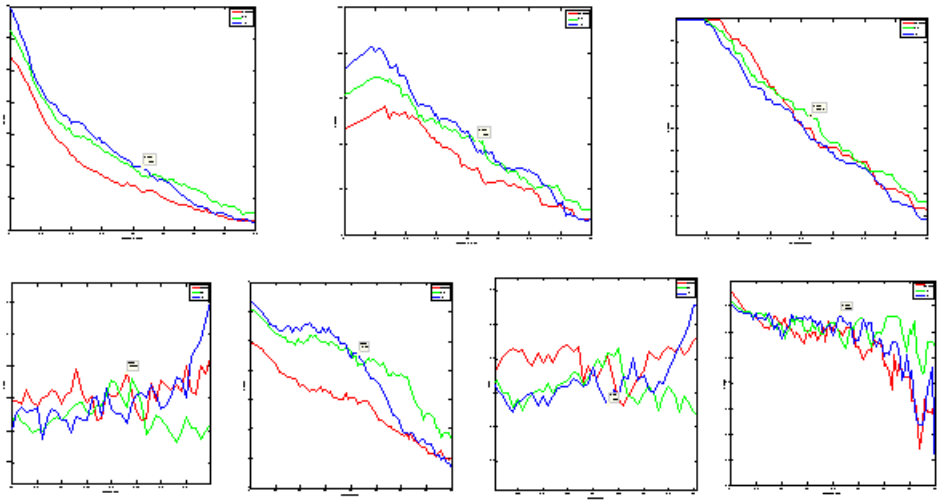

Fig. 1. The change situation of 7 network statistics with the increasing of threshold, where the abscissa and vertical axis denotes the threshold and the corresponding statistics respectively.Red, green, blue curve denotes the change curve of normal group, B_ALL group and T_ALL group respectively.

3.2.2 Construction of Mutual Information Network

We discover that when the threshold is taken as 0.53, 7 statistics of the networks for control and two experimental groups will be separated obviously by studying the statistics curves changing with the increasing of threshold in Fig. 1. So we take 0.53 as the threshold to construct the mutual information gene networks of 3 working databases, as shown in Fig. 2:

(a) Normal (b) B-ALL (c) T-ALL

Fig. 2. When the threshold is taken as 0.53, the mutual information gene networks of 68 coordinate positioning genes of 3 working databases, where (a)，(b)，(c) denotes normal, B_ALL and T_ALL network, respectively.

Obviously, the relationships of many nodes in the three gene networks change in quantity, such as most of isolated nodes in the normal mutual information gene network are no longer isolated in B-ALL or T-ALL network, especially the changes of T-ALL network are significantly. At the same time, many non-isolated nodes in the normal network become isolated in B-ALL or T-ALL network, the connectivity of nodes in the network also changes significantly.

3.3 Selection of Structural Key Genes

A: Some genes are isolated in the normal and T-ALL network, but high connectivity in B-ALL network. Because there are 41 non-isolated nodes in the B_ALL network, take the genes whose degree changes in quantity are more than 20 as the significant genes called the structural key genes. At last, we discover that 8 genes AR, ERBB4, ESR2, Evi-1, PDGFRB, FAT1, CXCL1 and CXCL2 satisfy the condition. Their degrees in 3 gene networks are showed as Table 1.

Table 1 The gene list of type A

Gene	Normal	B_ALL	T_ALL
AR*	0	26	0
ERBB4*	0	27	0
ESR2*	0	26	0
EVI1*	0	33	0
FAT1*	0	28	0
PDGFRB*	0	35	0
CXCL1	0	21	0
CXCL2*	0	25	0

B: Some genes are isolated nodes in normal network and B-ALL network, but high connectivity in the T-ALL network. Because there are 38 non-isolated nodes in the B_ALL network, take the genes whose degree changes in quantity are more than 19 as the structural key genes. At last, we find 8 genes FER, FES, LTA, WNT3, TCL1A, BTK, MSH4 and PTK2 genes satisfy the condition, their degrees in 3 gene networks are showed as Table 2.

Table 2 The gene list of type B

Gene	Normal	B_ALL	T_ALL
FER*	0	0	34
FES*	0	0	27
LTA*	0	0	32
WNT3*	0	0	26
TCL1A*	0	0	31
MSH4*	0	0	25
PTK2*	0	0	24
BTK*	0	0	23

C: The degrees of some genes have changed more significantly in the networks of experimental groups than in the one of normal control group. Take the genes whose degree changes in quantity are more than 18 as the structural key genes. At last, we find 7 genes CDH1, KDR, IGF-1, FOSL1, CTTN, MOS and TACSTD2 satisfy the condition, their degrees in 3 gene networks are showed as Table 3.

Table 3 The gene list of type C

Gene	Normal	B_ALL	T_ALL
CDH1*	0	31	27
TACSTD2	0	30	27
FOSL1*	0	21	21
CTTN*	0	20	20
KDR*	20	0	0
IGF1*	20	0	0
MOS*	19	0	0

4. Conclusion and Analysis

We discover some genes whose degrees change significantly between the networks for control group and experimental groups. For example, some genes are very active (with greater degree) in the normal control group, and abnormal silence (isolated nodes) in the experimental groups, while some others are very active in the experimental groups, and abnormal silence in the

normal control group. That is to say, these genes have great contribute to structure changes of mutual information gene networks for control and experimental groups, so we define those genes whose degrees change significantly between control group and experimental groups as "the structural key gene". Now, we discuss these genes in three conditions as follows:

A: Considering genes AR, ERBB4, ESR2, Evi-1, PDGFRB, FAT1, CXCL1 and CXCL2, they are isolated nodes in normal and T-ALL gene network, but with greater degree in the B-ALL network. Because their degrees in the B-ALL network significantly increase, we can speculate that their functions or the life processes involved in should be to promote cell proliferation and division, and they should promote the development and deterioration of cancer forward. The specific functions of the 8 genes as follows:

Yun Cai et al. investigated the methylation status of the androgen receptor gene (AR) in leukemia cell lines. Results showed the presence of both methylated and unmethylated CpG islands of the AR promotor in leukemia cell lines. In the normal blood samples, only unmethylated bands were observed[13]. The over-expression of ERBB4 gene (Leukemia virus oncogene, homologous chromosome 4) is related with the occurrence and development of tumor. In leukemias, the downstream pathways of ERBB family signaling are frequently activated, and introduction of activated EGFR into hematopoietic cell lines induced proliferation, survival and abrogation of cytokine dependency of the cells. However, the activation mechanisms of the ERBB family signaling are largely unknown[14]. Otabek Imamov et al. demonstrated the novel role for ESR2 in regulating the differentiation of pluripotent hematopoietic progenitor cells, and suggested that the ESR2(*Er* beta)-mouse should be a potential model for myeloid and lymphoid leukemia, and that ESR2 agonists might have clinical value in the treatment of leukemia if the ESR2 is not itself mutated in this disease. The human ESR2 gene has been mapped to chromosome 14q22[15]. Evi-1 gene is frequently overexpressed in leukemias having 3q26 abnormalities such as t(3;3)(q21;q26) and inv(3)(q21 q26), and subjects to structural alteration in t(3;21)(q26;q22). Ogawa S et al. presented another case of structural alteration of Evi-1 gene in a case of inv(3)(q21 q26), in which Evi-1 is truncated and a shorter form of Evi-1 protein is expressed upon rearrangement of the gene. Their result also supports an idea that Evi-1 is a relevant oncogene whose overexpression or structural changes might play a crucial role in development of human leukemias[16]. PDGFRB is constitutively activated by gene fusion with different partners in myeloproliferative disorders with peculiar clinical characteristics. Translocation t(5;12)(q33;p13), resulting in an ETV6/PDGFRB gene fusion, is a recurrent chromosomal abnormality associated with chronic myelomonocytic leukemia (CMML). An analogous translocation was also found in four cell lines with features of pre-B acute lymphoblastic leukemia (ALL) [17]. Thomas Dunwell et al. demonstrated FAT1 gene is significantly more methylated in B-ALL compared to T-ALL[18]. Gill D. et al. demonstrated the prolonged survival of B-CLL in vitro without additional stromal cell support. Furthermore, novel cytokines CCL2 and CXCL2 appear to prolong the survival of B-CLL cells *in vitro*. This culture system and these chemokines may allow us to gain insight into factors modulating B-CLL survival and potentially could lead to targeted therapy as well as serve as an appropriate model to test new therapies[19]. CXCL1 is expressed by macrophages, neutrophils and epithelial cells, and has neutrophil chemoattractant activity. CXCL1 plays a role in spinal cord development by inhibiting the migration of oligodendrocyte precursors and is involved in the processes of angiogenesis, inflammation, wound healing, and tumorigenesis. Based on the specific functions of the above genes, we discover that they are closely related to B_ALL, maybe they are the oncogenes of B_ALL.

B: Considering genes FER, FES, LTA, WNT3, TCL1A, BTK, MSH4 and PTK2, they are isolated nodes in normal and B-ALL network, but with greater degree in the T-ALL network. Because their degrees in the T-ALL network significantly increase, we can speculate that their functions or the life process involved in should be to promote cell proliferation and division, and

they should promote the development and deterioration of cancer forward. The specific functions of the 8 genes as follows:

Expression of FER gene in a wide range of cell types indicates a general role in intracellular signalling or differentiation processes. J. Groffen et al. found that the negative regulation to be the main cause for dysfunctioning of the Fer promoter in a T-cell leukemia cell line[20]. The human FES proto-oncogene is expressed as a transcript of about 3.0 kb in both normal and leukemic myeloid cells. Tesch H. et al. detected truncated FES transcripts of about 0.9 kb in a panel of human lymphoma and lymphoid leukemia cell lines, but not in normal untransformed hematopoietic cells[21]. Zhou MX, et al. discovered that the partially purified LTA (TNF-alpha) obtained from the EU-1 cell line also suppressed the proliferation of TNF-sensitive primary leukemic cells, and this inhibitory activity was abolished by an anti-TNF-alpha specific antibody. The results demonstrated that TNF-alpha is an inhibitor of in vitro proliferation of BCP-ALL cells from most patients[22]. Transfection with a Wnt3 plasmid resulted in a small increase in reporter gene activity, which was augmented by the transfection with the LRP6 coreceptor. Recent microarray analyses have demonstrated that the Wnt3 gene is overexpressed in CLL, compared with normal B and T cells[23]. In T-cell prolymphocytic leukemia (T-PLL), chromosomal imbalances affecting the long arm of chromosome 22 are regarded as typical chromosomal aberrations secondary to a TCRAD-TCL1A fusion due to inv(14) or t(14;14) [24]. Bruton's tyrosine kinase (BTK) deficiency results in a differentiation block at the pre-B cell stage. Likewise, acute lymphoblastic leukemia cells are typically arrested at early stages of B cell development. Janet D. Rowley et al. identified kinase-deficient splice variants of BTK throughout all leukemia subtypes[25]. MSH4 gene belongs to the human DNA mismatch repair system (MMR).The principal function of MMR is to edit replication and to reverse DNA polymerase errors. Inactivation of MMR greatly increases spontaneous mutation rates. H Hirai et al. suggested that disruption of MMR may play an important role in the development of human lymphoid leukemias[26]. Focal adhesion kinase (FAK) is constitutively activated and tyrosine phosphorylated in BCR/ABL-transformed hematopoietic cells. Yi Le et al. suggested that FAK is critical for leukemogenesis and might be a potential target for leukemia therapy[27]. Based on the specific functions of the above genes, we discover that they are closely related to T_ALL, maybe they are the oncogenes of T_ALL.

C: Considering genes CDH1, KDR, IGF-1, FOSL1, CTTN, MOS and TACSTD2, their degrees change more significantly in the networks for experimental groups than those for control group. The specific functions of the 7 genes as follows:

E-cadherin, the gene product of CDH1, plays a key role in cell–cell adhesion. Decrease or loss of E-cadherin expression accompanied by CDH1 promoter methylation has been reported in many human cancers. John R. et al. found that all normal donor samples expressed E-cadherin mRNA, whereas both samples of acute myelogenous leukemia and chronic lymphocytic leukemia had a significant reduction or absence of expression. However, normal blast counterparts expressed only a low level of E-cadherin surface protein[28]. KDR (FLT) gene plays an important role in new angiogenesis，and it is related with tumor grade, growth and prognosis. Kai Neben et al. showed that unique gene expression patterns can be correlated with FLT3-ITD and FLT3-TKD. This might lead to the identification of further pathogenetic relevant candidate genes particularly in AML with normal karyotype[29]. IGF-1 is important in blood formation and regulation and has been shown to stimulate the growth of both myeloid and lymphoid cells in culture. Since infants who develop leukemia are likely to have had at least one transforming event occur in utero, Julie A. Ross et al. hypothesized that high levels of IGF-1 may both produce a larger baby and contribute to leukemogenesis[30]. FOSL1(fra-1) gene, which is a Tax1-inducible fos-related gene, was isolated and Tax1 or serum-responsive cis elements were analyzed to obtain further insight into the mechanism of Tax1 action. Tax1 of human T-cell leukemia virus type 1 stimulates the expression of several cellular immediate-early genes[31]. Patients with lymph node metastasis apt to emerge the amplification of CTTN (EMS1) gene,

accompanied the phenomenon of EMS1 over-expression, tumor cell invasion and metastasis increases simultaneously. The mutation of EMS1 maybe one of the reasons of ALL[32]. Cytogenetic analysis of an infant with Down syndrome with concomitant acute myelogenous leukemia revealed a unique t(8;16)(q22;q24). In situ chromosomal hybridization was used to demonstrate that the protooncogene MOS was translocated from chromosome 8 to chromosome 16. Mark J. et al. reported that the transposition of MOS in association with acute leukemia[33]. TACSTD2 is considered as a cancer-associated antigen, and the antibody of its extracellular domain can reduce the invasiveness of tumor cell. Based on the specific functions of the above genes, we discover that they are closely related to ALL, maybe they are the oncogenes of ALL.

Therefore, comparing the 23 structural key genes of ALL predicted by model with the functions of these genes showed in literatures, we discover that 21 genes (marked with "*" in Table 1, 2 and 3) are confirmed to be closely related to the formation of ALL. The match ratio between the prediction by model and literature is up to 91.3%. Hence, we can predict that the remaining two genes CXCL1 and TACSTD2 are also closely related to ALL based on the effectiveness of the method in this method. The authenticity and reliability of the above speculations, as well as their pathogenic mechanism of ALL are needed to be further validated by biological experiments.

In this research, we use 7 statistics parameters to characterize the network structure from different angles. From the results of the numerical experiments, we can see that the 7 parameters can not measure the network structure fully. Whether or not there is a proper structural parameter, which can give the network structure a more comprehensive characterization? In addition, the number of samples in the databases has a certain degree impact on the result, so how to choose a reasonable number of samples need us to further explore as well.

References

[1] Eng-Juh Yeoh et al.. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. *Cancer Cell*, 2002, 1(2): 133-143.

[2] Yakovlev et al.. Variable selection and pattern recognition with gene expression data generated by the microarray technology. *Mathematical Biosciences*, 2002, 176(1): 71-98.

[3] Anne Laurence Astier et al.. Temporal gene expression profile of human precursor B leukemia cells induced by adhesion receptor: identification of pathways regulating B-cell survival. *Blood*, 2003, 101(3): 1118-1127.

[4] Silvio Bicciato et al.. Pattern identification and classification in gene expression data using an autoassociative neural network model. *Biotechnology and Bioengineering*, 2003, 81(5): 594-606.

[5] Juliane Schäfer and Korbinian Strimmer. An empirical Bayes approach to inferring large-scale gene association networks. *Bioinformatics*, 2005, 21(6): 754-764.

[6] Shaoqi Rao et al.. A robust hybrid between genetic algorithm and support vector machine for extracting an optimal feature gene subset. *Genomics*, 2005, 85(1): 16-23.

[7] Morton, Geoffrey. *Data Mining the Genetics of Leukemia*. QSpace at Queen's Research & Learning Repository, 2010.

[8] Harald Lahm et al.. Tumor galectinology: Insights into the complex network of a family of endogenous lectins. *Glycoconjugate Journal*, 2004, 20(4): 227-238.

[9] Adriano V. Werhli, Marco Grzegorczyk, Dirk Husmeier. Comparative evaluation of reverse engineering gene regulatory networks with relevance networks, graphical Gaussian modelsand bayesian networks. *Bioinformatics*, 2006, 22(20): 2523-2531.

[10] Theodore J. Perkins, Johannes Jaeger, John Reinitz, Leon Glass. Reverse engineering the gap gene network of Drosophila melanogaster. *Plos Comput Biol.*, 2006, 2(5): 0417-0428.

[11] Shudong Wang,Yan Chen, Qingyun Wang, Eryan Li, Yansen Su, Dazhi Meng. Analysis for gene networks based on logical relationships. *Journal of Systems Science and Complexity*, 2010, 23(5): 999-1011.
[12] Lǔsheng Si, Xu Li. *Oncogene, Tumor suppressor gene, Tumor-related genes.* Shanxi Science and Technology Press, Xi'an, 2002.12.
[13] Yun Cai et al.. Androgen Receptor CpG Island Methylation Status in Human Leukemia Cancer Cells. *Leukemia*, 2009, 27(2): 156-162.
[14] Jong Woo Lee et al.. Kinase domain mutation of ERBB family genes is uncommon in acute leukemias. *Leukemia Research*, 2006, 30(2): 241-242.
[15] Otabek Imamov et al.. Estrogen Receptor beta in Health and Disease. *Biology of Reproduction*, 2005, 73(5): 866-871.
[16] Ogawa S et al.. Abnormal expression of Evi-1 gene in human leukemias. *Hum Cell*, 1996, 9(4): 323-332.
[17] J Cools et al.. Chronic myeloproliferative disorders: a tyrosine kinase tale. *Leukemia*, 2006, 20: 200-205.
[18] Thomas Dunwell et al.. A Genome-wide screen identifies frequently methylated genes in haematological and epithelial cancers. *Molecular Cancer*, 2010, 9(44): 1-12.
[19] Gill D., Burgess M., Knop L. et al.. Identification of two novel chemokines (CCL2 and CXCL2) in B-chronic lymphocytic leukemia (B-CLL) and prolonged survival of primary B-CLL in vitro. *Blood*, 2008, 112(11): abstr. 3157.
[20] Carmel M, Shpungin S, Nir U.. Role of positive and negative regulation in modulation of the Fer promoter activity. *Gene*, 2000, 241(1): 87-99.
[21] Tesch H. et al.. Expression of truncated transcripts of the proto-oncogene c-fps/fes in human lymphoma and lymphoid leukemia cell lines. *Oncogene*, 1992, 7(5): 943-952.
[22] Zhou MX, et al.. Effect of tumor necrosis factor-alpha on the proliferation of leukemic cells from children with B-cell precursor-acute lymphoblastic leukemia (BCP-ALL): studies of primary leukemic cells and BCP-ALL cell lines. *Blood*, 1991, 77(9): 2002-2007.
[23] Dennis A. Carson et al.. Activation of the Wnt signaling pathway in chronic lymphocytic leukemia. *PNAS*, 2004, 101(9): 3118-3123.
[24] Stefanie Bug et al.. Recurrent loss, but lack of mutations, of the SMARCB1 tumor suppressor gene in T-cell prolymphocytic leukemia with TCL1A–TCRAD juxtaposition. *Cancer Genetics and Cytogenetics*, 2009, 192(1): 44-47.
[25] Janet D. Rowley et al.. Deficiency of Bruton's tyrosine kinase in B cell precursor leukemia cells. *PNAS*, 2005, 102(37): 13266-13271.
[26] H Hirai et al.. Mutations and loss of expression of a mismatch repair gene, hMLH1, in leukemia and lymphoma cell lines. *Blood*, 1997, 89(5): 1740-1747.
[27] Yi Le et al.. FAK silencing inhibits leukemogenesis in BCR/ABL-transformed hematopoietic cells. *American Journal of Hematology*, 2009, 84(5): 273–278.
[28] John R. et al.. Hypermethylation of E-cadherin in leukemia. *Blood*, 2000, 95(10): 3208-3213.
[29] Kai Neben et al.. Distinct gene expression patterns associated with FLT3- and NRAS-activating mutations in acute myeloid leukemia with normal karyotype. *Oncogene*, 2005, 24: 1580-1588.
[30] Julie A. Ross et al.. Big babies and infant leukemia: a role for insulin-like growth factor-1. *Cancer Causes and Control*, 1996, 7(5): 553-559.
[31] H Tsuchiya et al.. Human T-cell leukemia virus type 1 Tax activates transcription of the human fra-1 gene through multiple cis elements responsive to transmembrane signals. *J Virol*., 1993, 67(12): 7001-7007.
[32] Donald Pasqualea et al.. Clonal evolution with +11q 13, t(1;7) and t(1;4) at relapse in a patient with Ph positive acute lymphocytic leukemia (ALL) treated with single agent front line imatinib followed by dasatinib. *Hematology*, 2007, 12(6): 505 - 509.

[33] Mark J. et al.. Translocation of the MOS gene in a rare t(8;16) associated with acute myeloblastic leukemia and down syndrome. *Cancer Genetics and Cytogenetics*, 1989, 37(2): 221-227.