

Research Article

Novel Approach for Coexpression Analysis of E2F1–3 and MYC Target Genes in Chronic Myelogenous Leukemia

Fengfeng Wang,¹ Lawrence W. C. Chan,¹ William C. S. Cho,² Petrus Tang,³ Jun Yu,⁴ Chi-Ren Shyu,⁵ Nancy B. Y. Tsui,¹ S. C. Cesar Wong,¹ Parco M. Siu,¹ S. P. Yip,¹ and Benjamin Y. M. Yung¹

¹ Department of Health Technology and Informatics, Hong Kong Polytechnic University, Lee Shau Kee Building, Hung Hom, Kowloon, Hong Kong

² Department of Clinical Oncology, Queen Elizabeth Hospital, 30 Gascoigne Road, Kowloon, Hong Kong

³ Bioinformatics Center, Chang Gung University, Taoyuan 333, Taiwan

⁴ Beijing Institute of Genomics, Chinese Academy of Sciences, Chaoyang District, Beijing 100029, China

⁵ Department of Computer Science, Informatics Institute, University of Missouri, Columbia, MO 65211, USA

Correspondence should be addressed to Lawrence W. C. Chan; wing.chi.chan@polyu.edu.hk

Received 26 June 2014; Accepted 23 July 2014; Published 10 August 2014

Academic Editor: Ryuji Hamamoto

Copyright © 2014 Fengfeng Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. Chronic myelogenous leukemia (CML) is characterized by tremendous amount of immature myeloid cells in the blood circulation. E2F1–3 and MYC are important transcription factors that form positive feedback loops by reciprocal regulation in their own transcription processes. Since genes regulated by E2F1–3 or MYC are related to cell proliferation and apoptosis, we wonder if there exists difference in the coexpression patterns of genes regulated concurrently by E2F1–3 and MYC between the normal and the CML states. **Results.** We proposed a method to explore the difference in the coexpression patterns of those candidate target genes between the normal and the CML groups. A disease-specific cutoff point for coexpression levels that classified the coexpressed gene pairs into strong and weak coexpression classes was identified. Our developed method effectively identified the coexpression pattern differences from the overall structure. Moreover, we found that genes related to the cell adhesion and angiogenesis properties were more likely to be coexpressed in the normal group when compared to the CML group. **Conclusion.** Our findings may be helpful in exploring the underlying mechanisms of CML and provide useful information in cancer treatment.

1. Introduction

Chronic myelogenous leukemia (CML) is a clonal myeloproliferative disorder that is characterized by the premature circulation of many immature myeloid cells in the blood stream [1]. The incidence rate of CML is about 1–2 per 100,000 per year. CML accounts for 20% of all leukemias affecting adults with a median age of 45 to 55 years [2]. The characteristics of CML at the cellular level include increased proliferation, increased resistance to apoptosis, and alterations in adhesion properties of leukemic progenitors [1]. Recently, there are many more studies on the analysis of microarray gene expression profiles in CML. Most of them investigate the function of differentially expressed genes such

as the study to explore the relationship between pathways and differentially expressed genes from untreated CML patients in the chronic phase [3]. However, few studies are available on the coexpression analysis.

Transcription factor (TF), a kind of transacting factor, plays the most vital role in the regulation of gene expression and process of signal transduction [4]. E2F family of transcription factors is important to control cellular proliferation by regulating transcription of various genes involved in DNA replication, DNA repair, mitosis, and cell cycle progression [5]. According to structure-function studies and amino acid sequence analysis, members of the E2F family can be classified into two main subclasses: activators E2F1–3 and repressors E2F4–8 [5]. The transcription activators E2F1,

2, and 3 are vital for cell cycle progression, especially in the G1/S transition process [6]. The protooncogene c-myc encodes a transcription factor (MYC) that can induce both cell proliferation and apoptosis [7]. As a transcription factor, MYC can both activate and repress transcription of target genes. High-throughput techniques have shown that MYC-activated genes are involved in growth, protein synthesis, and mitochondrial function. Most of MYC-repressed genes participate in the interaction and communication between cells and their external environment, and several genes are found to have antiproliferative or antimetastatic properties [8]. In addition, E2F1–3 and MYC are reciprocally regulated in the transcription process to form positive feedback loops among them [9].

Target genes regulated by the same TF tend to be coexpressed, and the coexpression degree is increased if genes share more TFs [10]. Moreover, coexpression analysis has been used to study functionally related genes since the coexpressed genes are more likely to participate in the similar cellular processes and pathways [11]. Furthermore, coexpressed genes are different in different states and cell types [12]. As a result, coexpression pattern analysis is a powerful strategy for grouping genes and further analyzing the underlying mechanisms of diseases. The different coexpression pattern can be regarded as the signature of a disease.

Since target genes regulated by E2F1–3 or MYC are related to cell proliferation and apoptosis, we wonder if there exists difference in the coexpression patterns of genes regulated concurrently by E2F1–3 and MYC between the normal and the CML states. In order to answer this research question, we proposed a method to explore the difference in the coexpression patterns by identifying a disease-specific cutoff point for coexpression levels that classified the coexpressed gene pairs into strong and weak coexpression classes so that the class was best coherent with the disease phenotype. Traditional methods on the coexpression analysis identify significantly coexpressed gene pairs by calculating a *P* value of correlation coefficient for each gene pair individually, which cannot reflect the overall difference between two different groups. Our method calculated all the correlation coefficients in each group to form two different cumulative distributions including all the gene pairs, which can identify the difference between two different groups from the overall structure. Also, the different coexpression pattern reflected the biological alterations in CML compared to the normal state. Annotation of the candidate target genes and mapping the coexpressed gene pairs to the annotated gene pairs from enriched process networks provided important information to understand the underlying mechanisms of the CML and the normal states.

2. Methods

2.1. Microarray Expression Data. Microarray technology is used to monitor the expression levels of thousands of genes in cells simultaneously [13]. Gene expression analysis across different conditions, the normal and the disease states, may contribute much to the exploration of disease mechanisms.

In this study, we analyzed the microarray dataset GSE5550, normalized by variance stabilizing transformations (VSN) method, which is publicly available on the *Gene Expression Omnibus (GEO)* repository [3]. The data were obtained from gene expression measurements of 8,537 unique mRNAs. CD34+ hematopoietic stem and progenitor cells were collected from the bone marrow of patients with untreated CML in the chronic phase and health controls [3]. The subjects recruited for this dataset are Caucasians in Germany. The CML group consisted of nine patient samples, and the control group included eight normal samples. In this dataset, a gene may be interrogated by more than one probe. In this case, we took the average of all the probes for the same mRNA [14, 15].

2.2. Identification of Candidate Target Genes Regulated Concurrently by E2F1–3 and MYC. The interactions between TFs (E2F1, E2F2, E2F3, and MYC) and target genes (TGs) were obtained from *prediction of transcriptional regulatory modules (PReMod)* database [16]. TF binding sites are often clustered together, called cis-regulatory modules (CRMs). *PReMod* database predicts relationships between TFs and their TGs based on the binding affinity and conservation of CRM. It consists of more than 100,000 computationally predicted modules within the human genome [16]. These modules give a description of 229 potential transcription factor families and are the first genome-wide collection of predicted regulatory modules for the human genome [17]. In this study, we called the set of TF binding predictions (TF-TG pairs) from *PReMod* a molecular interaction set. This set was regarded as the reference data. After obtaining the TGs of each TF (E2F1, E2F2, E2F3, and MYC) individually, we identified the common TGs of these four TFs, which were regarded as the candidate target genes for further analysis. The flowchart is shown in Figure S1 in Supplementary Material available online at <http://dx.doi.org/10.1155/2014/439840>.

2.3. Coexpression Measure. We chose Pearson correlation coefficient as the similarity measure. It is represented by the direction cosine between two vectors normalized by the subtraction of their own means, and its value accounts for the angle between two feature vectors instead of the vector lengths. Moreover, Pearson correlation coefficient numerically indicates the biological relationship of two genes but does not vary with the magnitudes of their expression profiles [11, 18]. In general, similarity measure is a kernel function between two feature vectors. In this study, each feature vector consisted of the expression intensity of a gene across all the samples in the normal group or the CML group, respectively. The correlation coefficient of any two genes among the candidate target genes was calculated. We took the absolute value of correlation coefficient ($|r|$) since the coexpression measure output a scalar in the range from 0 to 1 where a high output indicated a strong biological relationship in either positive or negative direction, and a low output indicated a weak biological relationship. The coexpression level was denoted by $C_d(i, j)$ if two expression profiles were extracted from samples of the disease (CML) group and

$C_n(i, j)$ for the normal group, shown in Formulas (1) as follows:

$$\begin{aligned} C_d(i, j) &= |\text{cor}(x_{di}, x_{dj})|, \\ C_n(i, j) &= |\text{cor}(x_{ni}, x_{nj})|, \end{aligned} \quad (1)$$

where $C_d(i, j)$ and $C_n(i, j)$ are defined as the absolute values of correlation coefficients between the expression profiles of genes i and j in the CML group and the normal group, respectively [18]; x_{di} and x_{dj} are the expression profiles of the i th and j th genes in the CML group; x_{ni} and x_{nj} are the expression profiles of the i th and j th genes in the normal group; $\text{cor}(x_{di}, x_{dj})$ and $\text{cor}(x_{ni}, x_{nj})$ are the Pearson correlation coefficients between them in the CML group and the normal group, respectively.

2.4. Classification of Coexpressed Gene Pairs. There was a set of correlation coefficients in either the normal group or the CML group. The two sets of correlation coefficients formed two cumulative distributions. We applied two-sample Kolmogorov-Smirnov (KS) test to identify the difference in the overall distributions of these two conditions (C_d and C_n), including all the gene pairs. The maximum deviation between two cumulative distributions of C_d and C_n was identified (Formulas (2)), at which a threshold was found to classify the coexpressed gene pairs into strong and weak coexpression classes, called the disease-specific cutoff point (C). The cutoff point represented a coexpression level, at which F_d and F_n were extremely deviated. Gene pairs were further classified into four coexpression classes: (i) strongly coexpressed gene pairs in the normal group: pairs with coexpression levels ($|r|$ values) bigger than or equal to C in the normal group; (ii) strongly coexpressed gene pairs in the CML group: pairs with coexpression levels ($|r|$ values) bigger than or equal to C in the CML group; (iii) weakly coexpressed gene pairs in the normal group: pairs with coexpression levels ($|r|$ values) smaller than C in the normal group; and (iv) weakly coexpressed gene pairs in the CML group: pairs with coexpression levels ($|r|$ values) smaller than C in the CML group. Chi-square test was used to determine if the proportions of strongly and weakly coexpressed gene pairs significantly differed between the normal and the CML groups

$$\begin{aligned} D &= \max_C |F_d(C) - F_n(C)|, \\ F_d(C) &= \text{Prob}(C_d \geq C), \\ F_n(C) &= \text{Prob}(C_n \geq C), \end{aligned} \quad (2)$$

where F_d and F_n are the cumulative distribution functions (CDFs) of C_d and C_n , respectively; D is the maximum deviation; C is the cutoff point.

We further identified the specifically coexpressed gene pairs in different groups. Each type of gene pair represented a particular biological meaning. The normal-specific strongly coexpressed gene pairs were the gene pairs strongly coexpressed only in the normal group, which were regarded as the potential molecular interactions maintaining physiological balance in healthy individuals, and the impairment of

these connections may lead to diseases. Obviously, these pairs were the CML-specific weakly coexpressed gene pairs, which were weakly coexpressed only in the CML group. The CML-specific strongly coexpressed gene pairs were the gene pairs strongly coexpressed only in the CML group, which represented the characteristics of the disease and may be the pathogenic alternatives when the corresponding normal-specific gene pairs cannot be coexpressed for responding to stress. Similarly, these pairs were regarded as the normal-specific weakly coexpressed gene pairs.

2.5. Functional Annotation for Candidate Target Genes. We applied *MetaCore* from GeneGo Inc. to annotate the candidate target genes. Specifically, when we uploaded the candidate target genes from Section 2.2 into this database, it mapped these genes to a set of cellular and molecular process networks, which are defined and annotated by Thomson Reuters scientists. In *MetaCore*, each process is defined as a preset network describing the protein interactions among them. In each process network, the annotated target genes were those genes included in both Section 2.2 and this process network. Enrichment analysis for a process network in *MetaCore* is performed based on the P value of hypergeometric intersection between the uploaded candidate target genes and the process-related genes in this database. The lower the P value is obtained, the higher the relevance of this process network to the candidate target genes and the rating of this process network are indicated. Only the top 10 statistically enriched process networks are shown according to the sorted P values in *MetaCore*.

2.6. Mapping Coexpressed Gene Pairs to Annotated Gene Pairs. The annotated target genes in each process network were paired with all the possible combinations to form the annotated gene pairs. The annotated gene pairs from each process network were mapped to the coexpressed gene pairs identified in Section 2.4: the normal-specific strongly coexpressed, the normal-specific weakly coexpressed, the CML-specific strongly coexpressed, and the CML-specific weakly coexpressed gene pairs. We applied Fisher's exact test to identify if there were more mapped normal-specific strongly coexpressed gene pairs than mapped CML-specific strongly coexpressed gene pairs in each process network. In other words, we planned to identify if these genes were more likely to be coexpressed in the normal group compared to the CML group. As a result, one-sided P value was chosen. False discovery rates (FDRs) are usually used to control the expected proportion of false positives for the multiple hypotheses. In this study, the FDRs were calculated based on the P values obtained from Fisher's exact test [19]. A process network was significantly mapped, if its FDR value was smaller than 0.05 [20]. The FDR values were estimated via the *Matlab* function, *mafdr* [21].

3. Results

3.1. Identification of Structural Coexpression Difference. In total, we identified 217 common TGs of E2F1-3 and MYC

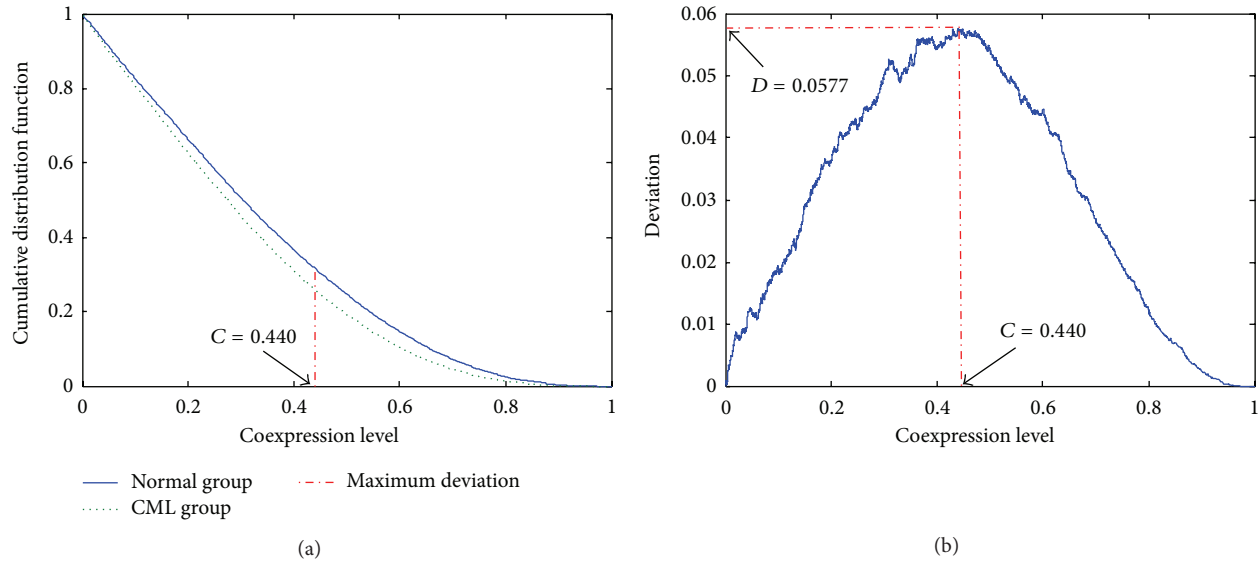


FIGURE 1: Plots of distributions for coexpression analysis. (a) Cumulative distribution functions of coexpression levels in the normal and the CML groups. (b) Deviation distribution against different coexpression cutoff points.

TABLE 1: The coexpressed gene pairs identified by the disease-specific cutoff point.

Group	Number of strongly coexpressed pairs	Number of weakly coexpressed pairs
Normal	7436	16000
CML	6083	17353

that can be found in the microarray dataset GSE5550 (Table S1). We further extracted the available expression profiles of these TGs and calculated the correlation coefficients in both the normal and the CML groups. In each group, there was a set of correlation coefficients of 23,436 gene pairs. We plotted the cumulative distributions of these two sets of data. The distributions between the normal and the CML groups were significantly different (P value = 2.00×10^{-34} for $D = 0.0577$). The disease-specific cutoff point that classified the coexpressed gene pairs into strong and weak coexpression classes was $C = 0.440$ (Figure 1(a)). Figure 1(b) illustrates that the deviation was small at the two extremes, and the peak ($D = 0.0577$) was found at the disease-specific cutoff point. Two coexpression patterns were so distinct that the normal group had more strongly coexpressed (level above ~ 0.440) and less weakly coexpressed (level below ~ 0.440) gene pairs than those in the CML group (Figure 1(a)). The cutoff point classified the gene pairs into four coexpression classes, shown in Table 1. The number of strongly coexpressed gene pairs in the normal group (7436) was larger than that in the CML group (6083). Chi-square test indicated that the proportions of strongly and weakly coexpressed gene pairs significantly differed between the normal and the CML groups (P value = 2.74×10^{-43} for $\chi^2 = 190$).

3.2. MetaCore Analysis for Enriched Process Networks. The top 10 statistically enriched process networks for functional

annotation of the 217 candidate target genes are shown in Table S2. All the P values for hypergeometric intersection test were smaller than 0.05. We got the annotated target genes involved in each process network and mapped the annotated gene pairs to the coexpressed gene pairs. Fisher's exact test was used to identify if there were more mapped normal-specific strongly coexpressed gene pairs than mapped CML-specific strongly coexpressed gene pairs in each process network. The results showed that 8 out of 10 process networks had more mapped normal-specific strongly coexpressed gene pairs (Table 2). Fisher's exact test demonstrated that "*Cell adhesion_Attractive and repulsive receptors*" and "*Development_Regulation of angiogenesis*" process networks were significantly mapped (P values = 0.001 and 0.012, < 0.05 , and FDR values were 0.004 and 0.026, < 0.05).

We further plotted the coexpression networks for the mapped normal-specific strongly coexpressed gene pairs ($a = 6$ and 8) (Figure 2). Both "*Cell adhesion_Attractive and repulsive receptors*" and "*Development_Regulation of angiogenesis*" process networks had ephrin-B2 (EFNB2), ephrin-A5 (EFNA5), and EPH receptor A4 (EPHA4) (Figure S2). From *National Center for Biotechnology Information (NCBI)* database, we obtained the basic information for these genes/proteins. EFNB2 and EFNA5 are the members of the ephrin gene family. EPHA4 protein product is an ephrin receptor. The ephrins (EPH) and EPH-related receptors belong to the largest subfamily of receptor protein-tyrosine kinases, which play a vital role in mediating developmental events. Figure 2 shows that the connection from EFNA5 to EPHA4 was identified as a strongly coexpressed gene pair for these two process networks in the normal group. In addition, protein products from neuropilin 2 (NRP2), transforming growth factor, beta receptor II (TGFB2), and somatostatin receptor 2 (SSTR2) also belong to receptors, which are very important in signal transduction process. The encoded protein from integrin, alpha 2 (ITGA2), plays a vital

TABLE 2: Mapping coexpressed gene pairs to annotated gene pairs from each process network.

Process networks	Fisher's exact test				P value	FDR
	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>		
Development_Neurogenesis in general	14	11	11	14	0.286	0.251
Development_Hedgehog signaling	22	15	15	22	0.081	0.118
Signal transduction_WNT signaling	10	7	7	10	0.247	0.270
Signal transduction_TGF-beta, GDF, and activin signaling	6	5	5	6	0.500	0.365
Cell adhesion_Attractive and repulsive receptors	6	0	0	6	0.001	0.004
Development_Regulation of angiogenesis	8	2	2	8	0.012	0.026
Cardiac development_BMP_TGF_beta_signaling	2	1	1	2	0.500	0.313
Neurophysiological process_Melatonin signaling	3	2	2	3	0.500	0.274

a: mapped normal-specific strongly coexpressed gene pairs; *b*: mapped normal-specific weakly coexpressed gene pairs; *c*: mapped CML-specific strongly coexpressed gene pairs; *d*: mapped CML-specific weakly coexpressed gene pairs.

role in leukocyte intercellular adhesion process. There were three enzymes identified in the coexpression networks: (i) the protein encoded by protein kinase, cAMP-dependent, catalytic, beta (PRKACB) is a protein kinase; (ii) the protein product from prolyl endopeptidase (PREP) is a protease; and (iii) the protein encoded by HIV-1 Tat interactive protein 2 (HTATIP2) is an oxidoreductase required for tumor suppression. From the results, we can infer that these genes/proteins were well connected with each other to transduce signals and maintain physiological balance in healthy individuals. However, in the CML group, these connections were impaired.

4. Discussion and Conclusion

In this study, our developed method successfully identified the difference in the coexpression patterns of those candidate target genes regulated concurrently by E2F1-3 and MYC between the normal and the CML groups from the overall structure (Figure 1). We further found that genes involved in the cell adhesion and angiogenesis properties were more likely to be coexpressed in the normal group compared to the CML group (Table 2 and Figure 2). The alteration in adhesion properties of leukemic progenitors is one CML characteristic at the cellular level [1]. In addition, Bhatia et al. hypothesized that decreased integrin-mediated adhesion of CML progenitors to stroma can lead to continuous cell proliferation [22]. They treated the cells with interferon- α (IFN- α). The results showed that the treatment restored the CML progenitor adhesion to stroma and also the regulation of CML progenitor proliferation [22]. Angiogenesis is the process forming new blood from the preexisting vasculature, including degradation of extracellular matrix proteins, as well as activation, proliferation, and migration of endothelial cells [23]. In leukemia, hematopoietic cells are supported from the normal vascular bed in bone marrow [23]. Increased vascularity was found in acute myeloid leukemia (AML) patients [24]. Importantly, in CML, the number of blood vessels and vascular areas were found to be increased when compared to control bone marrows [23]. Our results showed that the connection from EFNA5 to EPHA4 was identified as a strongly coexpressed gene pair in the normal group ($|r|$ values were 0.720 and 0.013 in the normal group

and the CML group, resp.) (Figure 2). Ephrin-A receptors belong to the largest subfamily of receptor tyrosine kinases that regulate cell shape, mobility, and attachment [25]. Interactions between Ephrin-A receptors and ligands are important in cell-cell communication, initiating unique bidirectional signaling cascades to transduce the information [26]. There may be some relationships between adhesion property and angiogenesis. These two process networks were found to be well controlled in the normal group compared to the CML group. Dysregulation of adhesion and angiogenesis properties is a possible reason leading to CML.

The advantage of our study is the application of coexpression analysis to target genes regulated concurrently by more than one transcription factor under different conditions. We identified different coexpression patterns between the normal and the CML groups. A limitation for differential expression analysis is that it only reflects the upregulation or downregulation of existing components in the well-known pathways under the normal or the disease condition, which cannot identify the functionally associated linkages among genes during signal transduction. In addition, differential expression analysis does not take account of the level of correlations that may exist between gene expression patterns [12]. Coexpression analysis is useful for analyzing the underlying mechanisms of diseases. Moreover, the different coexpression pattern can be regarded as the signature of a disease.

Several methods have been proposed to analyze coexpressed genes. The two-stage screening procedure was applied to select statistically and biologically significant gene pairs in Zhu et al.'s study [27]. Gupta et al. proposed a method for determining the correlation threshold using the clustering coefficient. R^2 metric was used as a measure of similarity between two genes [28]. Previous studies cannot reflect the overall difference between two different groups. Our method calculated all the correlation coefficients in each group (the normal group and the CML group) to form two distributions, which can find the difference between two different groups from the overall structure.

In summary, we have presented a detailed method to identify a disease-specific cutoff point for coexpression levels

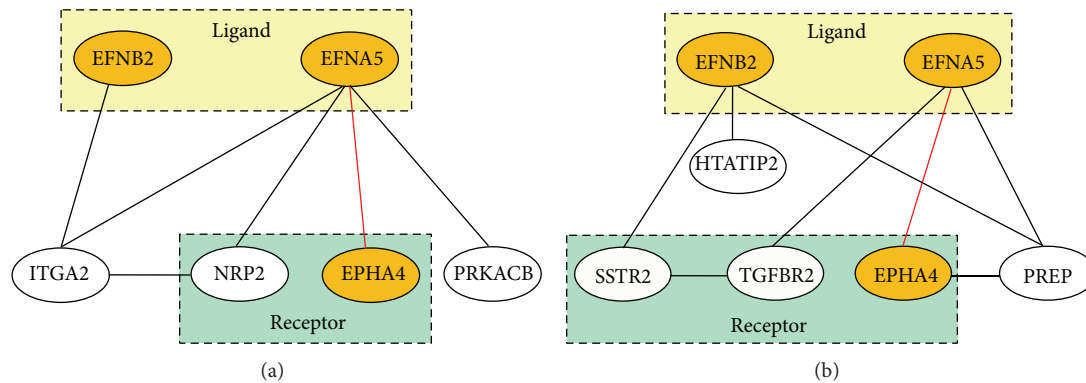


FIGURE 2: Coexpression networks for the mapped normal-specific strongly coexpressed gene pairs. The yellow ellipses are those genes found in both process networks. (a) Mapped normal-specific strongly coexpressed gene pairs in the “Cell adhesion_Attractive and repulsive receptors” process network. (b) Mapped normal-specific strongly coexpressed gene pairs in the “Development_Regulation of angiogenesis” process network.

that classified the coexpressed gene pairs into strong and weak coexpression classes so that the class was best coherent with the disease phenotype. We applied this method to explore the difference in the coexpression patterns of target genes regulated concurrently by E2F1–3 and MYC between the normal and the CML groups. Our method effectively identified the statistical differences between the normal and the CML groups from the overall structure. We further found the potentially altered cell adhesion and angiogenesis properties in the CML state when compared to the normal group. The different coexpression pattern can reflect the biological alterations in CML. Our significant findings will be helpful in exploring the underlying mechanisms of CML and provide useful information in cancer treatment.

Conflict of Interests

The authors declare that there is no conflict of interests.

Acknowledgments

This work was supported by Internal Grants of Hong Kong Polytechnic University (G.55.09.YL61) and (1-ZE17). The authors thank Dr. Thomas Lui for his advice on the analyses using *Matlab*. The *MetaCore* license was provided by the Bioinformatics Center from Chang Gung University.

References

- [1] S. Salesse and C. M. Verfaillie, “Mechanisms underlying abnormal trafficking and expansion of malignant progenitors in CML: BCR/ABL-induced defects in integrin function in CML,” *Oncogene*, vol. 21, no. 56, pp. 8605–8611, 2002.
- [2] R. Frazer, A. E. Irvine, and M. F. McMullin, “Chronic myeloid leukaemia in the 21st century,” *Ulster Medical Journal*, vol. 76, no. 1, pp. 8–17, 2007.
- [3] E. Diaz-Blanco, I. Bruns, F. Neumann et al., “Molecular signature of CD34⁺ hematopoietic stem and progenitor cells of patients with CML in chronic phase,” *Leukemia*, vol. 21, no. 3, pp. 494–504, 2007.
- [4] Q. Cui, Z. Yu, Y. Pan, E. O. Purisima, and E. Wang, “MicroRNAs preferentially target the genes with high transcriptional regulation complexity,” *Biochemical and Biophysical Research Communications*, vol. 352, no. 3, pp. 733–738, 2007.
- [5] C. Timmers, N. Sharma, R. Opavsky et al., “E2f1, E2f2, and E2f3 control E2F target expression and cellular proliferation via a p53-dependent negative feedback loop,” *Molecular and Cellular Biology*, vol. 27, no. 1, pp. 65–78, 2007.
- [6] L. Wu, C. Timmers, B. Malti et al., “The E2F1-3 transcription factors are essential for cellular proliferation,” *Nature*, vol. 414, no. 6862, pp. 457–462, 2001.
- [7] S. Pelengaris, B. Rudolph, and T. Littlewood, “Action of Myc in vivo—Proliferation and apoptosis,” *Current Opinion in Genetics & Development*, vol. 10, no. 1, pp. 100–105, 2000.
- [8] S. Pelengaris and M. Khan, “The many faces of c-MYC,” *Archives of Biochemistry and Biophysics*, vol. 416, no. 2, pp. 129–136, 2003.
- [9] H. A. Collier, J. J. Forman, and A. Legesse-Miller, “MyCed messages: myc induces transcription of E2F1 while inhibiting its translation via a microRNA polycistron,” *PLoS Genetics*, vol. 3, no. 8, p. e146, 2007.
- [10] H. Yu, N. M. Luscombe, J. Qian, and M. Gerstein, “Genomic analysis of gene expression relationships in transcriptional regulatory networks,” *Trends in Genetics*, vol. 19, no. 8, pp. 422–427, 2003.
- [11] M. B. Eisen, P. T. Spellman, P. O. Brown, and D. Botstein, “Cluster analysis and display of genome-wide expression patterns,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 25, pp. 14863–14868, 1998.
- [12] A. Torkamani, B. Dean, N. J. Schork, and E. A. Thomas, “Coexpression network analysis of neural tissue reveals perturbations in developmental processes in schizophrenia,” *Genome Research*, vol. 20, no. 4, pp. 403–412, 2010.
- [13] S. Dudoit, Y. H. Yang, M. J. Callow, and T. P. Speed, “Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments,” *Statistica Sinica*, vol. 12, no. 1, pp. 111–139, 2002.
- [14] T. Breslin, M. Krogh, C. Peterson, and C. Troein, “Signal transduction pathway profiling of individual tumor samples,” *BMC Bioinformatics*, vol. 6, article 163, 2005.
- [15] A. V. Kapp, S. S. Jeffrey, A. Langerød et al., “Discovery and validation of breast cancer subtypes,” *BMC Genomics*, vol. 7, article 231, 2006.

- [16] V. Ferretti, C. Poitras, D. Bergeron, B. Coulombe, F. Robert, and M. Blanchette, "PReMod: a database of genome-wide mammalian cis-regulatory module predictions," *Nucleic Acids Research*, vol. 35, no. 1, pp. D122–D126, 2007.
- [17] M. Blanchette, A. R. Bataille, X. Chen et al., "Genome-wide computational prediction of transcriptional regulatory modules reveals new insights into human gene expression," *Genome Research*, vol. 16, no. 5, pp. 656–668, 2006.
- [18] S. Horvath and J. Dong, "Geometric interpretation of gene coexpression network analysis," *PLoS Computational Biology*, vol. 4, no. 8, Article ID e1000117, 2008.
- [19] J. D. Storey, "A direct approach to false discovery rates," *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, vol. 64, no. 3, pp. 479–498, 2002.
- [20] J. Fu, W. Tang, P. Du et al., "Identifying MicroRNA-mRNA regulatory network in colorectal cancer by a combination of expression profile and bioinformatics analysis," *BMC Systems Biology*, vol. 6, article 68, 2012.
- [21] W. E. Haskins, K. Petritis, and J. Zhang, "MRCQuant-an accurate LC-MS relative isotopic quantification algorithm on TOF instruments," *BMC Bioinformatics*, vol. 12, no. 74, 2011.
- [22] R. Bhatia, J. B. McCarthy, and C. M. Verfaillie, "Interferon- α restores normal $\beta 1$ integrin-mediated inhibition of hematopoietic progenitor proliferation by the marrow microenvironment in chronic myelogenous leukemia," *Blood*, vol. 87, no. 9, pp. 3883–3891, 1996.
- [23] A. Aguayo, H. Kantarjian, T. Manshoury et al., "Angiogenesis in acute and chronic leukemias and myelodysplastic syndromes," *Blood*, vol. 96, no. 6, pp. 2240–2245, 2000.
- [24] J. W. Hussong, G. M. Rodgers, and P. J. Shami, "Evidence of increased angiogenesis in patients with acute myeloid leukemia," *Blood*, vol. 95, no. 1, pp. 309–313, 2000.
- [25] J. P. Himanen, N. Saha, and D. B. Nikolov, "Cell-cell signaling via Eph receptors and ephrins," *Current Opinion in Cell Biology*, vol. 19, no. 5, pp. 534–542, 2007.
- [26] J. Himanen and D. B. Nikolov, "Eph receptors and ephrins," *International Journal of Biochemistry and Cell Biology*, vol. 35, no. 2, pp. 130–134, 2003.
- [27] D. Zhu, A. O. Hero, H. Cheng, R. Khanna, and A. Swaroop, "Network constrained clustering for gene microarray data," *Bioinformatics*, vol. 21, no. 21, pp. 4014–4020, 2005.
- [28] A. Gupta, C. D. Maranas, and R. Albert, "Elucidation of directionality for co-expressed genes: Predicting intra-operon termination sites," *Bioinformatics*, vol. 22, no. 2, pp. 209–214, 2006.

