

---

## **Computational identification of potential microRNA network biomarkers for the progression stages of gastric cancer**

---

Le Lu, Yanda Li and Shao Li\*

MOE Key Laboratory of Bioinformatics  
and Bioinformatics Division,  
TNLIST/Department of Automation,  
Tsinghua University,  
Beijing 100084, China  
Fax: +86 010 62786911  
E-mail: lule@mail.tsinghua.edu.cn  
E-mail: dalyd@tsinghua.edu.cn  
E-mail: shaoli@mail.tsinghua.edu.cn  
\*Corresponding author

**Abstract:** MicroRNAs (miRNAs) are potential biomarkers in the diagnosis of human disease. In this study, a novel concept, the miRNA network biomarker, was proposed for the selection of biomarkers. Each miRNA network biomarker contains miRNA targets, as well as Transcription Factors (TFs), that affect the miRNA expression. The obtained biomarkers were applied to classifying expression data sets in different progression stages from chronic gastritis to gastric cancer. Furthermore, these biomarkers could accurately (94%) discriminate gastric cancer samples from normal samples in another data set. Angiogenesis-related pathways and genes were found to be enriched in these network biomarkers.

**Keywords:** miRNA; microRNA; biomarker; TFs; transcription factors; TSS; transcription start site; miRNA target; angiogenesis; gastric cancer; chronic gastritis; TFBS; transcription factor binding site.

**Reference** to this paper should be made as follows: Lu, L., Li, Y. and Li, S. (2011) 'Computational identification of potential microRNA network biomarkers for the progression stages of gastric cancer', *Int. J. Data Mining and Bioinformatics*, Vol. 5, No. 5, pp.519–531.

**Biographical notes:** Le Lu got her BS in Dept. Biotechnology, Peking University and PhD in School of Biological Sciences, Nanyang Technological University. She is now a postdoc in Dept. Automation, Tsinghua University.

Yanda Li is a Professor in Department of Automation, Tsinghua University. He is an academician of the Chinese Academy of Sciences. He is specialist in signal processing, intelligent control and bioinformatics.

Shao Li is a Professor in Department of Automation, Tsinghua University. He is also the Deputy Director of Bioinformatics Division, Tsinghua National Laboratory of Information Science and Technology, and heads a research group concentrating on network biology, network-based complex disease research and drug discovery, and traditional Chinese medicine systems bioinformatics.

## 1 Introduction

Until recently, biomarkers were usually individual genes and their expression pattern could distinguish samples between different stages during the development of the disease (Ramaswamy et al., 2003; Ein-Dor et al., 2006). However, how the biomarker genes function in the development of the disease remained unknown. Studies on complex biological networks, which integrated genomics, transcriptomics, proteomics and metabolomics, have provided new insights into the understanding of the molecular basis of human diseases (Anastassiou, 2007). Network biomarkers are defined as a group of genes or gene products with identified or predicted functional relationships that can successfully discriminate samples from different stages of the development of a disease (Li et al., 2007; Li, 2009). The functional relationships between these network biomarkers are experimentally identified protein–protein interactions or predictions based on current systems biology approaches. For example, Chuang et al. (2007) proposed a network-based classification approach to identify biomarkers for breast cancer metastasis. They found that the network biomarkers were more reproducible than individual biomarkers selected without network information. Wu et al. (2008) showed that the combination of a human protein network and a phenotype network could correctly identify disease genes.

MicroRNAs (miRNAs) are 19–25 nt non-coding RNAs that play important roles in gene regulation, either through the inhibition of protein translation or by mRNA degradation. Many miRNAs act as tumour suppressors or oncogenes in human diseases; hence, their expressions are altered in cancer development and progression (Esquela-Kerscher and Slack, 2006). In this study, we proposed a novel concept, the miRNA network biomarker. Each miRNA network contains the TFs that regulate certain miRNAs and potential miRNA targets. The miRNA network biomarkers are coding genes that belong to prior constructed miRNA networks and have the predictive power to discriminate samples from various disease stages.

MiRNAs are transcribed independently or from within their host genes, sharing the same promoter regions. To understand the miRNAs' regulatory mechanism, a fundamental step is to correctly locate their Transcription Start Sites (TSSs). Wang et al. (2009b) proposed a novel approach, termed CoreBoost\_HM, to predict RNA polymerase II core-promoters. On the basis of genome-wide histone modifications and DNA sequence features, they further applied CoreBoost\_HM to detect the promoters of miRNAs and found that CoreBoost\_HM could accurately identify known miRNA promoters in the intergenic regions (Wang et al., 2009b).

MiRNAs function through the regulation of their target genes. Animal miRNA target prediction has attracted much attention recently (Krek et al., 2005; Lewis et al., 2005; Betel et al., 2007; Friedman et al., 2009). Among the available target prediction algorithms, PicTar (Krek et al., 2005), miRanda (Betel et al., 2007) and TargetScan (Grimson et al., 2007) are three widely used programmes. It has been suggested that a combination of these three programmes has the highest sensitivity, based on a systematic comparison (Sethupathy et al., 2006) between the performance of five mammalian miRNA target prediction algorithms, as well as various combinations of them.

In this study, we first constructed an miRNA network for each known human miRNA, based on prior annotation about TF/miRNA and miRNA/target regulation. The miRNA network integrated two types of nodes, namely the TF nodes that regulate

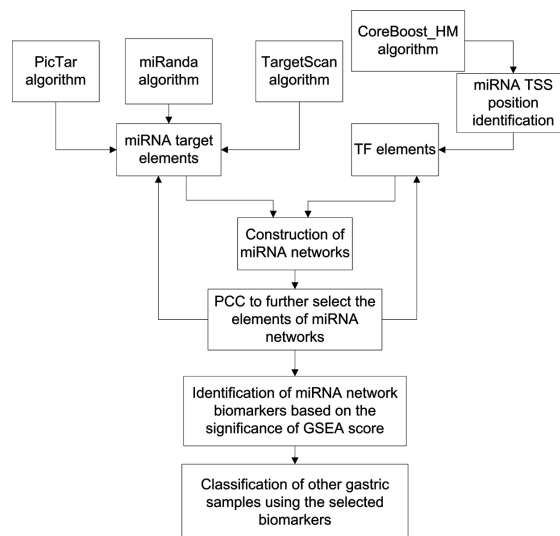
miRNA and the miRNA target nodes. The miRNA network biomarkers that can distinguish between the progression stages of gastric cancer were selected by Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005), which was based on the expression profile of each miRNA network in Gse2669 microarray data set. The identified biomarkers were used to classify another gastric cancer data set, Gse13911, to test their reproducibility.

## 2 Materials and methods

### 2.1 Construction of miRNA networks

A flow chart of the miRNA network biomarkers selection approach is shown in Figure 1. A total of 555 human miRNAs were downloaded from miRBase (<http://microrna.sanger.ac.uk/sequences/>) released in September 2008. Thus, we constructed 555 miRNA networks. An miRNA network contains TFs that bind to the miRNA TSS, as well as the miRNA targets. The TSSs were predicted by CoreBoost\_HM (Wang et al., 2009b). The promoter region (upstream – 3000 bps to downstream 1000 bps) of each predicted miRNA TSS was mapped to the conserved and experimentally identified Transcription Factor Binding Sites (TFBSs). The chromosomal coordinates of the TFBSs were downloaded from UCSC (<http://genome.ucsc.edu>). Potential targets from three widely used miRNA targets prediction algorithms, namely PicTar (Krek et al., 2005), miRanda (Betel et al., 2007) and TargetScan (Grimson et al., 2007), were combined. Only the targets that were predicted by all the three algorithms were taken into further consideration.

**Figure 1** Flow chart of the miRNA network biomarkers identification approach. Each miRNA network integrated both TFs and target nodes. The expression value of each node was retrieved from Gse2669. GSEA was applied to evaluate whether an miRNA network shows significantly different expressions between the progression stages of gastric cancer. The identified miRNA network biomarkers were used to classify samples in Gse13911



## 2.2 Pearson Correlation Coefficient

The expression series (Gse2669 and Gse13911), from gastritis to the progression stages of gastric cancer, were retrieved from Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/>). Gse2669 comprised 124 gastric biopsies with normal, chronic gastritis, intestinal metaplasia and carcinomas. Gse13911 (D'Errico et al., 2009) comprised 69 samples from 38 tumour biopsies and 31 adjacent normal tissue samples. On the basis of the Gse2669 data set, we further determined the nodes of each miRNA network. The Pearson Correlation Coefficient (PCC)  $P_{X,Y}$  indicates the linear relationship between two gene expression patterns. It is defined as:

$$P_{x,y} = \frac{E((X - \mu_X)(Y - \mu_Y))}{\sigma_X \sigma_Y}$$

where  $X$  and  $Y$  are two certain gene expression patterns, with mean expression values  $\mu_X$  and  $\mu_Y$ , respectively, and  $\sigma_X$  and  $\sigma_Y$  are the standard deviations. A node will be removed from the network if none of its  $P_{X,Y}$  values are greater than the cut-off value.

## 2.3 The construction of random networks

To evaluate the significance of the GSEA score and identify the miRNA network biomarkers, we randomly selected the same number of TFs and target nodes as that of each miRNA network. The obtained random networks were also scored by the GSEA. The R package of the GSEA software was downloaded from <http://www.broad.mit.edu/gsea/>

## 2.4 Assessment of the classification performance

Initially proposed by Vapnik, Support Vector Machine (SVM) is a general machine-learning algorithm (Vapnik, 1998). Here, we applied LIBSVM (Fan et al., 2005) (<http://www.csie.ntu.edu.tw/~cjlin/libsvm>) to classifying the gastric cancer and normal samples in the Gse13911 data set. The input for SVM is the expression profiles of the gastric cancer and normal samples, with the format described as in the LIBSVM webpage. The class labels are binary variables, either two different disease stages (Chronic gastritis; Laurens: mixed or diffused; Intestinal metaplasia and Laurens: intestinal) or normal sample vs. a disease stage.

The output for the SVM is the model (parameters) used in the sample classification. Leave-one-out cross-validation was performed to access the classification performance of the identified miRNA network biomarkers. The sensitivity (Sen), specificity (Sp), accuracy (Acc) and Matthew's Correlation Coefficient (MCC) are defined as:

$$\text{Sen} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\%$$

$$\text{Sp} = \frac{\text{TN}}{\text{TN} + \text{FP}} \times 100\%$$

$$\text{Acc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}} \times 100\%$$

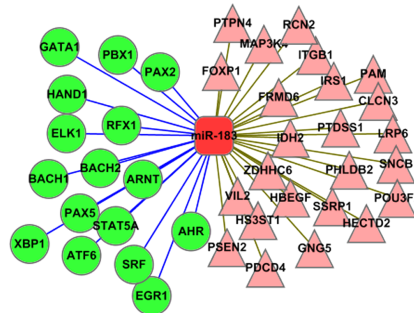
$$MCC = \frac{TP \times TN - FN \times FP}{\sqrt{(TP + FN) \times (TP + FP) \times (TN + FN) \times (TN + FP)}}$$

where TP, TN, FP and FN indicate the number of true positive, true negative, false positive and false negative predictions, respectively.

### 2.5 The construction of an angiogenesis network

The angiogenesis network was constructed based on literature mining approach (Li et al., 2006), with key words ‘Angiogenesis or Neovascularisation’ in PubMed (<http://www.ncbi.nlm.nih.gov/pubmed/>). The functional relationship between two genes was retrieved from HPRD (<http://www.hprd.org/>) (Mishra et al., 2006) and KEGG (<http://www.genome.jp/kegg/>) (Kanehisa et al., 2008). If two genes are annotated to be related, we added an edge between them in the angiogenesis network (Li et al., 2006). Figure 2 was drawn by the software Cytoscape (Shannon et al., 2003).

**Figure 2** An example of an miRNA network. The miR-183 network contained 40 nodes, namely 23 targets (the triangle orange nodes) and 17 TFs (the round green nodes). MiR-183 was regarded as a bridge node to connect the TFs and targets, and its own expression was not taken into consideration. The blue edges represented the transcriptional regulation between TFs and miR-183, and the khaki edges represented the posttranscriptional regulation between miR-183 and its targets (see online version for colours)



## 3 Results

The miRBase (Griffiths-Jones et al., 2006) (Release 12) comprised 555 human miRNA records. We applied CoreBoost\_HM (Wang et al., 2009b) to detect their potential TSSs. Along the human genome, there were 519 positions predicted by CoreBoost\_HM to be the TSSs of miRNAs with high sensitivity. We mapped these positions with the chromosomal coordinates of conserved TFBSs. If a certain TFBS was within the promoter region of an identified miRNA TSS, it would be taken as a TF node of this miRNA network.

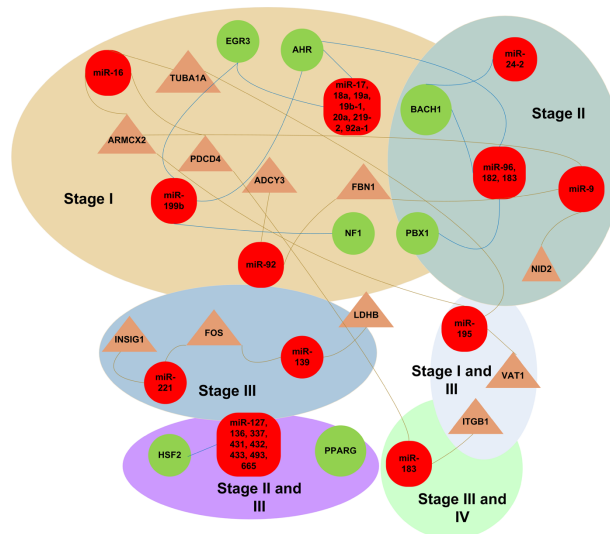
There are various animal miRNA target prediction algorithms. However, the potential targets predicted by different algorithms are not consistent. To reduce the number of false positives, we combined the prediction results of PicTar, miRanda and TargetScan. If a certain target was predicted by all three algorithms, we adopted it as a target node.

An miRNA network was required to contain at least two TFs or target nodes. In total, there were 70 TFs and 1276 target nodes integrated into 323 miRNA networks. Figure 2 illustrates the network for miR-183. Each miRNA network was evaluated for its predictive power in discriminating between the progression stages of gastric cancer.

The Gse2669 series contains 124 samples under five different conditions (Normal stomach; Chronic gastritis; Laurens: mixed or diffused; Intestinal metaplasia and Laurens: intestinal). The PCC cut-off values for TF and target nodes were 0.5 and 0.75, respectively. A node was removed from the miRNA network if none of its PCCs were greater than the cut-off value. After the PCC selection step, 96 miRNA networks were obtained and their expression values in Gse2669 were scored by GSEA.

GSEA was used to evaluate the predictive power of each miRNA network for discriminating between the progression stages of gastric cancer. We used the default False Positive Rate (FDR)  $q$ -value setting of GSEA (FDR  $q$ -value  $<0.25$ ) as the cut-off value. The 69 miRNA networks were taken into further consideration. To test whether the GSEA score was due to the lack of expression coherence between samples, we randomly selected the same number of TF and target nodes as that in these 69 networks and scored the random networks by GSEA for 100 rounds. If the FDR  $q$ -value of an miRNA network was significantly ( $p$ -value  $<0.05$ ) less than that of random networks, it was identified as an miRNA network biomarker. Some miRNAs, such as the miR17~92 cluster (Hayashita et al., 2005), shared both promoters and TFs, hence we grouped the miRNA network biomarkers according to their overlapped nodes. In total, 12 miRNA network biomarkers were obtained that could distinguish between the progression stages of gastric cancer (Table 1 and Figure 3).

**Figure 3** The selected miRNA network biomarkers that could distinguish between the progression stages of gastric cancer. The round rectangle red nodes represented miRNAs (bridge nodes); the triangle orange nodes represented their targets. TFs were represented by the round green nodes. The blue edges represented the transcriptional regulation between TFs and miRNAs, and the khaki edges represented the posttranscriptional regulation between miRNA and their targets. Stage I: Chronic gastritis; Stage II: Laurens: mixed or diffused; Stage III: Intestinal metaplasia; Stage IV: Laurens: intestinal; Default control: Normal samples (see online version for colours)



To evaluate the reproducibility of these network biomarkers, we used them to classify the normal and gastric cancer samples from another data set (Gse13911). The expression values of the selected miRNA network biomarkers for gastric cancer were retrieved, a total of nine genes (Table 1: Stage II and Stage III). The resulting sensitivity and specificity were 0.92 and 0.96, respectively. The total accuracy and MCCs were 0.94 and 0.88, respectively. Furthermore, we randomly selected the same number of genes as the miRNA network biomarkers and used them to classify the samples in Gse13911. The results for 100 rounds suggested that the classification performance of the miRNA network biomarkers was significantly better than that of the genes selected without network information ( $p$ -value <0.02).

**Table 1** Twelve miRNA network biomarkers for the progression stages of gastric cancer

<i>miRNA network biomarkers</i>		<i>The progression stages of gastric cancer*</i>
<i>MiRNA</i>	<i>Target gene symbol</i>	
miR-16	TUBA1A PDCD4 ARM CX2	Stage I
miR-17, 18a, 19a, 19b-1, 20a, 219-2, 92a-1	AHR EGR3	Stage I
miR-92	ADCY3 FBN1	Stage I
miR-9	FBN1 NID2 ARM CX2	Stage II
miR-221	INSIG1 FOS	Stage III
miR-139	LDHB FOS	Stage III
miR-24-2	AHR BACH1	Stage I; Stage II
miR-96, 182, 183	AHR PBX1 BACH1	Stage I; Stage II
miR-199b	AHR EGR3 NF1	Stage I; Stage I and III
miR-195	TUBA1A VAT1 ARM CX2	Stage I and III
miR-127, 136, 337, 431, 432, 433, 493, 665	PPARG HSF2	Stage II and III
miR-183	ITGB1 PDCD4	Stage III and IV

\*Stage I: Chronic gastritis; Stage II: Laurens: mixed or diffused; Stage III: Intestinal metaplasia; Stage IV: Laurens: intestinal; Default control: Normal samples.

Angiogenesis is one of the fundamental steps in the transition of solid tumours, and the study of angiogenesis has attracted much attention during recent decades (Li et al., 2003, 2006; Fan et al., 2006). To study angiogenesis globally, we constructed an angiogenesis network based on literature mining approaches. The gene interaction annotations were retrieved from the HPRD and KEGG databases. The constructed angiogenesis network contained 1893 genes (nodes) and 7598 edges. As with most biological networks, the angiogenesis network is also a scale-free network (Barabási and Oltvai, 2004). In a scale-network, the 'hub' nodes have many more connections than other nodes. The average connection in the angiogenesis network for each node is eight; however, some hub node genes have more than 100 connections, such as TP53 with 111 connections.

Angiogenesis is one of the critical steps in the gastric cancer development, especially in the transition of chronic gastritis to gastric tumour (Concetta et al., 2005; Edhi et al., 2002; Jones et al., 1999; Hull et al., 1996). Recent research (Bonauer et al., 2009) showed that miR-92a controlled the angiogenesis process, and the overexpression of miRNA-92a in endothelial infarction cells blocked angiogenesis in vitro and in vivo (Bonauer et al., 2009). Furthermore, miR-221 is known to inhibit angiogenesis in vitro (Poliseno et al.,

2006). In addition, miRNA-18 a/miRNA-19a have proangiogenic effects in tumour angiogenesis (Dews et al., 2006). All these previous studies are consistent with our results, namely that the miR-92a, miR-18a and miR-19a networks were identified as biomarkers for chronic gastritis (Stage I), and miR-221 was identified as a biomarker for intestinal metaplasia (Stage III).

Among the identified biomarkers, FOS is known to be involved in the progress of *Helicobacter pylori* induced invasion and in angiogenesis of gastric cells (Chang et al., 2005), with 37 connections in the angiogenesis network. Furthermore, ITGB1 (Carlson et al., 2008) and PPARG (Masamune and Shimosegawa, 2009) are both established as regulators of angiogenesis, with 53 and 20 connections, respectively. To understand the possible involvement of miRNAs in gastric angiogenesis, we embedded the miRNA network biomarkers into the angiogenesis network and obtained 126 directly connected genes. The GO annotation from the DAVID (<http://david.abcc.ncifcrf.gov/home.jsp>) bioinformatics database showed that these genes were significantly associated in 10 KEGG pathways (Bonferroni corrected  $p$ -values  $< 0.05$ ; Table 2). Among them, the Toll-like receptor signalling pathway (van Beijnum et al., 2008), MAPK signalling pathway (Wang et al., 2009a) and Focal adhesion pathway (Skuli et al., 2009) are all angiogenesis-related pathways.

**Table 2** Functional enrichment for 126 directly connected genes in the angiogenesis network to the miRNA network biomarker genes

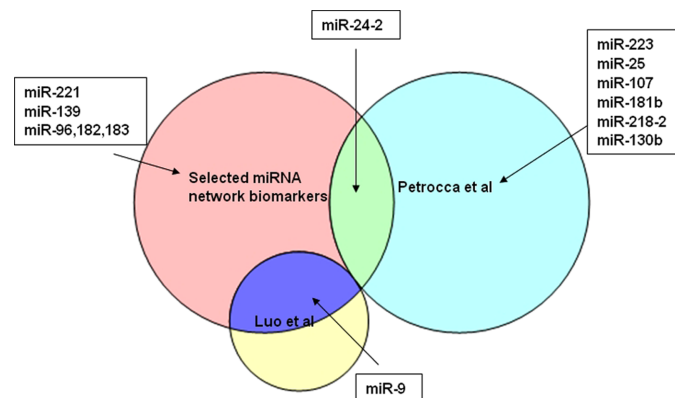
<i>KEGG pathways</i>	<i>Bonferroni corrected p-value</i>
Focal adhesion	3.20E-30
ECM-receptor interaction	9.40E-23
Cell communication	1.10E-09
Toll-like receptor signalling pathway	3.30E-06
ErbB signalling pathway	4.10E-02
MAPK signalling pathway	9.50E-03
IGF-1 Signalling Pathway	3.20E-03
IL 6 signalling pathway	2.80E-02
TSP-1 Induced apoptosis in microvascular endothelial cell	3.20E-02
Angiotensin II mediated activation of JNK Pathway via Pyk2 dependent signalling	4.60E-02

#### 4 Discussion

Gastric cancer is the second most widely diagnosed malignancy in the world and 12% patients die every year (Petrocca et al., 2006; Uemura et al., 2001). Chronic superficial gastritis is induced by *Helicobacter pylori* infection and may lead to gastric carcinoma after several identifiable clinical steps (chronic atrophic gastritis, intestinal metaplasia and dysplasia) (Uemura et al., 2001). Most gastric tumours can be classified into the histological types of intestinal or diffuse (Petrocca et al., 2006; Lauren, 1965). There has been some evidence that miRNAs are differently expressed in gastric cancer tissues (Petrocca et al., 2006, 2008; Luo et al., 2009). Luo et al. (2009) used a hierarchical

clustering method to analyse the expression profiles of 24 malignant and three normal gastric tissues. Their results suggested that miR-433 and miR-9 were significantly down-regulated in gastric cancer tissues. Petrocca et al. (2006) examined the expression of 245 miRNAs in 20 cases of human stomach carcinoma and normal gastric mucosa. They found that seven miRNAs (miR-223, miR-25, miR-107, miR-24-2, miR-181b, miR-218-2 and miR-130b) could discriminate normal and cancer samples with a misclassification error of 17%. However, we found that these two miRNA biomarker sets shared no biomarkers in common, suggesting that the reproducibility of individual miRNA biomarkers might be relatively low. According to our results, miR-9 and miR-24-2 were identified as miRNA network biomarkers, partially consistent with Luo et al. and Petrocca et al., respectively (Figure 4).

**Figure 4** Venn diagram illustrating the comparison between similar studies on the miRNA biomarkers in gastric cancer. Among the five miRNA network biomarkers identified by our study, miR-9 and miR-24-2 were reported to be differentially expressed in gastric cancer tissues in Luo et al. (2009) and Petrocca et al. (2006), respectively. Furthermore, miR-9 was one of the two miRNAs reported by Luo et al. (2009) with significantly different expression. There were no miRNAs reported to be differentially expressed both in Luo et al. (2009) and Petrocca et al. (2006) (see online version for colours)



MiRNAs are potential biomarkers, though the expression data sets of miRNAs might be lacking for certain disease stages. Hence, miRNA targets had been identified to be potential biomarkers (Liang, 2008). Liang (2008) used computationally predicted target genes of three miRNAs, miR-34b/34c/449, to discriminate lung adenocarcinoma and squamous cell carcinoma (two major histological subtypes of lung cancer), and obtained relatively high accuracy. However, miRNA target genes that were regulated at the protein level, but not at the transcription level, might not show significant expression changes accordingly. Hence, the integration of TFs might help to solve this problem because the TFs/miRNA regulation is at the level of transcription.

In our study, each miRNA network contained TFs and target nodes; both were based on prediction algorithms due to the lack of experimental data. To reduce the number of false positives, we put strict requirements in each selection step:

- 1 only the intersection of the three target prediction algorithms was taken
- 2 we considered the predicted TSSs with high sensitivity as well as conserved TFBSs

- 3 expression profiles of the miRNA network nodes were required to be correlated or anti-correlated
- 4 the GSEA score of each miRNA network was compared with that of random networks.

Each of these selection steps could help to reduce false positives. For instance, the GSEA scores without PCC selection were not significant compared with those of random networks. Hence, the direct use of static gene regulatory networks (based on prior knowledge) might cause difficulties in uncovering biological interactions under certain conditions. In this study, the static gene regulatory networks were based on sequence analysis results, which suggested that an miRNA might interact with a target or be under the regulation of a TF. However, it does not mean that all these potential functional relationship actually occurred under certain conditions. Hence, we further identified the elements of each network by the expression profiles of each potential node based on the assumption that genes under regulation of the same miRNA will have similar expression profiles, and thus the PCC was used. The detection of miRNA network biomarkers that can distinguish between disease progression stages might shed light on the identification of biomarkers for certain conditions and improve their reproducibility.

### Acknowledgements

This work was supported by NSFC grants (Nos. 60934004 and 60805009) and China Postdoctoral Science Foundation funded project.

### References

- Anastassiou, D. (2007) 'Computational analysis of the synergy among multiple interacting genes', *Mol. Syst. Biol.*, Vol. 3, No. 83, p.3.
- Barabási, A-L. and Oltvai, Z.N. (2004) 'Network biology: understanding the cell's functional organization', *Nat. Rev. Genet.*, Vol. 5, pp.101–113.
- Betel, D., Wilson, M., Gabow, A., Marks, D.S. and Sander, C. (2007) 'The microRNA.org resource: targets and expression', *Nucleic Acids Res.*, Vol. 36, Database issue, pp.D149–D153.
- Bonauer, A., Carmona, G., Iwasaki, M., Mione, M., Koyanagi, M., Fischer, A., Burchfield, J., Fox, H., Doebele, C., Ohtani, K., Chavakis, E., Potente, M., Tjwa, M., Urbich, C., Zeiher, A.M. and Dimmeler, S. (2009) 'MicroRNA-92a controls angiogenesis and functional recovery of ischemic tissues in mice', *Science*, Vol. 324, pp.1710–1713.
- Carlson, T.R., Hu, H., Braren, R., Kim, Y.H. and Wang, R.A. (2008) 'Cell-autonomous requirement for  $\beta$ 1 integrin in endothelial cell adhesion, migration and survival during angiogenesis in mice', *Development*, Vol. 135, pp.2193–2202.
- Chang, Y-J., Wu, M-S., Lin, J-T. and Chen, C-C. (2005) 'Helicobacter pylori-induced invasion and angiogenesis of gastric cells is mediated by cyclooxygenase-2 induction through TLR2/TLR9 and promoter regulation', *J. Immunol.*, Vol. 175, pp.8242–8252.
- Chuang, H., Lee, E., Liu, Y., Lee, D. and Ideker, T. (2007) 'Network-based classification of breast cancer metastasis', *Mol. Syst. Biol.*, Vol. 3, No. 140, p.3.
- Concetta, T., Antonio, C., Alba, R., Erika, M., Stefania, S., Massimo, M., Antonietta, G.G., Gerardo, N., Vittorio, R., Fortunato, C., Camillo del Vecchio, B. and Marco, R. (2005) 'Vascular endothelial growth factor and neo-angiogenesis in *H. pylori* gastritis in humans', *J. Pathol.*, Vol. 207, pp.277–284.

- D'Errico, M., Rinaldis, E.D., Blasi, M.F., Viti, V., Falchetti, M., Calcagnile, A., Sera, F., Saieva, C., Ottini, L., Palli, D., Palombo, F., Giuliani, A. and Dogliotti, E. (2009) 'Genome-wide expression profile of sporadic gastric cancers with microsatellite instability', *Eur. J. Cancer*, Vol. 45, pp.461–469.
- Dews, M., Homayouni, A., Yu, D., Murphy, D., Sevignani, C., Wentzel, E., Furth, E.E., Lee, W.M., Enders, G.H., Mendell, J.T. and Thomas-Tikhonenko, A. (2006) 'Augmentation of tumor angiogenesis by a Myc-activated microRNA cluster', *Nat. Genet.*, Vol. 38, pp.1060–1065.
- Edhi, G., Shingo, T., Masahiko, T., Arata, K., Wei-Hao, S., Hitoshi, S., Yoshiaki, S., Yutaka, S., Sunao, K. and Masatsugu, H. (2002) 'Influences of *Helicobacter pylori* on gastric angiogenesis and ulcer healing in mice', *J. Gastroenterol. Hepatol.*, Vol. 17, pp.960–965.
- Ein-Dor, I., Zuk, O. and Domany, E. (2006) 'Thousands of samples are needed to generate a robust gene list for predicting outcome in cancer', *Proc. Natl. Acad. Sci. USA*, Vol. 103, pp.5923–5928.
- Esquela-Kerscher, A. and Slack, F.J. (2006) 'Oncomirs – microRNAs with a role in cancer', *Nat. Rev. Cancer*, Vol. 6, pp.259–269.
- Fan, R.E., Chen, P.H. and Lin, C.J. (2005) 'Working set selection using second order information for training SVM', *J. Mach. Learn. Res.*, Vol. 6, pp.1889–1918.
- Fan, T., Yeh, J., Leung, K.W., Yue, P.Y.K. and Wong, R.N.S. (2006) 'Angiogenesis: from plants to blood vessels', *Trends Pharmacol. Sci.*, Vol. 27, pp.297–309.
- Friedman, R.C., Farh, K.K-H., Burge, C.B. and Bartel, D.P. (2009) 'Most mammalian mRNAs are conserved targets of microRNAs', *Genome Res.*, Vol. 19, pp.92–105.
- Griffiths-Jones, S., Grocock, R.J., van Dongen, S., Bateman, A. and Enright, A.J. (2006) 'miRBase: microRNA sequences, targets and gene nomenclature', *Nucleic Acids Res.*, Vol. 34, pp.D140–144.
- Grimson, A., Farh, K.K-H., Johnston, W.K., Garrett-Engele, P., Lim, L.P. and Bartel, D.P. (2007) 'MicroRNA targeting specificity in mammals: determinants beyond seed pairing', *Mol. Cell*, Vol. 27, pp.91–105.
- Hayashita, Y., Osada, H., Tatematsu, Y., Yamada, H., Yanagisawa, K., Tomida, S., Yatabe, Y., Kawahara, K., Sekido, Y. and Takahashi, T. (2005) 'A polycistronic MicroRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation', *Cancer Res.*, Vol. 65, pp.9628–9632.
- Hull, M.A., Hewett, P.W., Brough, J.L. and Hawkey, C.J. (1996) 'Isolation and culture of human gastric endothelial cells', *Gastroenterology*, Vol. 111, pp.1230–1240.
- Jones, M.K., Itani, R.M., Wang, H., Tomikawa, M., Sarfeh, I.J., Szabo, S. and Tarnawski, A.S. (1999) 'Activation of VEGF and Ras genes in gastric mucosa during angiogenic response to ethanol injury', *Am. J. Physiol. Gastrointest Liver Physiol.*, Vol. 276, pp.G1345–1355.
- Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T. and Yamanishi, Y. (2008) 'KEGG for linking genomes to life and the environment', *Nucleic Acids Res.*, Vol. 36, pp.D480–D484.
- Krek, A., Grun, D., Poy, M.N., Wolf, R., Rosenberg, L., Epstein, E.J., Macmenamin, P., Da Piedade, I., Gunsalus, K.C., Stoffel, M. and Rajewsky, N. (2005) 'Combinatorial microRNA target predictions', *Nat. Genet.*, Vol. 37, pp.495–500.
- Lauren, P. (1965) 'The two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma. an attempt at a histo-clinical classification', *Acta Pathol. Microbiol. Scand.*, Vol. 64, pp.31–49.
- Lewis, B.P., Burge, C.B. and Bartel, D.P. (2005) 'Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are MicroRNA targets', *Cell*, Vol. 120, pp.15–20.
- Li, S. (2009) 'Network systems underlying traditional Chinese medicine syndrome and herb Formula', *Current Bioinformatics*, Vol. 4, pp.188–196.

- Li, S., Lu, A-P., Wang, Y-Y. and Li, Y-D. (2003) 'Suppressive effects of a Chinese herbal medicine qing-luo-yin extract on the angiogenesis of collagen-induced arthritis in rats', *Am. J. Chin. Med.*, Vol. 31, pp.713–720.
- Li, S., Wu, L. and Zhang, Z. (2006) 'Constructing biological networks through combined literature mining and microarray analysis: a LMMA approach', *Bioinformatics*, Vol. 22, pp.2143–2150.
- Li, S., Zhang, Z.Q., Wu, L.J., Zhang, X.G., Li, Y.D. and Wang, Y.Y. (2007) 'Understanding Zheng in traditional Chinese medicine in the context of neuro-endocrine-immune network', *IET Syst. Biol.*, Vol. 1, pp.51–60.
- Liang, Y. (2008) 'An expression meta-analysis of predicted microRNA targets identifies a diagnostic signature for lung cancer', *BMC Med. Genomics*, Vol. 1, p.61.
- Luo, H., Zhang, Z., Zhang, X., Ning, B., Guo, J., Nie, N., Liu, B. and Wu, X. (2009) 'MicroRNA expression signature in gastric cancer', *Chin. J. Cancer. Res.*, Vol. 21, pp.74–80.
- Masamune, A. and Shimosegawa, T. (2009) 'Signal transduction in pancreatic stellate cells', *J. Gastroenterol.*, Vol. 44, pp.249–260.
- Mishra, G.R., Suresh, M., Kumaran, K., Kannabiran, N., Suresh, S., Bala, P., Shivakumar, K., Anuradha, N., Reddy, R., Raghavan, T.M., Menon, S., Hanumanthu, G., Gupta, M., Upendran, S., Gupta, S., Mahesh, M., Jacob, B., Mathew, P., Chatterjee, P., Arun, K.S., Sharma, S., Chandrika, K.N., Deshpande, N., Palvankar, K., Raghavnath, R., Krishnakanth, R., Karathia, H., Rekha, B., Nayak, R., Vishnupriya, G., Kumar, H.G.M., Nagini, M., Kumar, G.S.S., Jose, R., Deepthi, P., Mohan, S.S., Gandhi, T.K.B., Harsha, H.C., Deshpande, K.S., Sarker, M., Prasad, T.S.K. and Pandey, A. (2006) 'Human protein reference database – 2006 update', *Nucleic Acids Res.*, Vol. 34, pp.D411–D414.
- Petrocca, F., Pillozzi, E., Rapazzotti, M., Aurello, P., Mercantini, P., Volinia, S., Ruco, L.P., Croce, C.M. and Vecchione, A. (2006) 'MicroRNAs deregulation in gastric cancer', *AACR Meeting*, Vol. 47, p.1338.
- Petrocca, F., Visone, R., Onelli, M.R., Shah, M.H., Nicoloso, M.S., De Martino, I., Iliopoulos, D., Pillozzi, E., Liu, C-G., Negrini, M., Cavazzini, L., Volinia, S., Alder, H., Ruco, L.P., Baldassarre, G., Croce, C.M. and Vecchione, A. (2008) 'E2F1-regulated MicroRNAs impair TGF[beta]-dependent cell-cycle arrest and apoptosis in gastric cancer', *Cancer Cell*, Vol. 13, pp.272–286.
- Poliseno, L., Tuccoli, A., Mariani, L., Evangelista, M., Citti, L., Woods, K., Mercatanti, A., Hammond, S. and Rainaldi, G. (2006) 'MicroRNAs modulate the angiogenic properties of HUVECs', *Blood*, Vol. 108, pp.3068–3071.
- Ramaswamy, S., Ross, K.N., Lander, E.S. and Golub, T.R. (2003) 'A molecular signature of metastasis in primary solid tumors', *Nat. Genet.*, Vol. 33, pp.49–54.
- Sethupathy, P., Megraw, M. and Hatzigeorgiou, A.G. (2006) 'A guide through present computational approaches for the identification of mammalian microRNA targets', *Nat. Meth.*, Vol. 3, pp.881–886.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T. (2003) 'Cytoscape: a software environment for integrated models of biomolecular interaction networks', *Genome Res.*, Vol. 13, pp.2498–2504.
- Skuli, N., Monferran, S., Delmas, C., Favre, G., Bonnet, J., Toulas, C. and Cohen-Jonathan Moyal, E. (2009)  $\alpha_v\beta_3/\alpha_v\beta_5$  integrins-FAK-RhoB: a novel pathway for hypoxia regulation in glioblastoma', *Cancer Res.*, Vol. 69, pp.3308–3316.
- Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S. and Mesirov, J.P. (2005) 'Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles', *Proc. Natl. Acad. Sci., USA*, Vol. 102, pp.15545–15550.
- Uemura, N., Okamoto, S., Yamamoto, S., Matsumura, N., Yamaguchi, S., Yamakido, M., Taniyama, K., Sasaki, N. and Schlemper, R.J. (2001) 'Helicobacter pylori infection and the development of gastric cancer', *Engl. J. Med.*, Vol. 345, pp.784–789.

- van Beijnum, J., Buurman, W. and Griffioen, A. (2008) 'Convergence and amplification of Toll-Like Receptor (TLR) and Receptor For Advanced Glycation End products (RAGE) signaling pathways via High Mobility Group B1 (HMGB1)', *Angiogenesis*, Vol. 11, pp.91–99.
- Vapnik, V.N. (1998) *Statistical Learning Theory*, Wiley, New York.
- Wang, D., Weng, Q., Zhang, L., He, Q. and Yang, B. (2009a) 'VEGF and Bcl-2 interact via MAPKs signaling pathway in the response to hypoxia in neuroblastoma', *Cell Mol. Neurobiol.*, Vol. 29, pp.391–401.
- Wang, X., Xuan, Z., Zhao, X., Li, Y. and Zhang, M.Q. (2009b) 'High-resolution human core-promoter prediction with CoreBoost\_HM', *Genome Res.*, Vol. 19, pp.266–275.
- Wu, X., Jiang, R., Zhang, M.Q. and Li, S. (2008) 'Network-based global inference of human disease genes', *Mol. Syst. Biol.*, Vol. 4, No. 189, p.4.