European Review for Medical and Pharmacological Sciences 2014; 18: 2109-2115

Gene expression profiling of gastric cancer

H.-B. JIANG[1], T.-J. YANG[2], P. LU[2], Y.-J. MA[3]

[1]Endoscopy Diagnosis Department, People's Hospital of Zhengzhou, China
[2]Gastrointestinal Surgery Department, People's Hospital of Zhengzhou, China
[3]Digestive System Department, People's Hospital of Zhengzhou, China

Abstract. – OBJECTIVES: Gastric cancer is the second leading cause of cancer-related death worldwide. Gene expression profile facilitates the identification of molecular mechanism of gastric cancer. Previous studies mainly focused on differentially expressed genes (DEGs) without considering MicroRNAs (miRNAs) and transcription factors (TFs). Here we aim to elaborate the mechanism of gastric cancer on transcription level with microarray data from the gene expression omnibus (GEO) database.

MATERIALS AND METHODS: We firstly identified DEGs between gastric cancer and normal tissues. Then the DEGs were mapped in KEGG pathway and gene ontology database to conduct functional categories enrichment analysis. MiRNAs and TFs enriched with target DEGs were also identified.

RESULTS: A total of 977 DEGs were selected, including 492 down regulated and 485 overexpressed genes in gastric cancer tissue. Functional analysis revealed cell cycle, metabolism and ECM related biological processes as the significant items. Eight miRNAs and 20 TFs enriched with target DEGs were detected, including one novel miRNA (miR-557) and four novel TFs (SPI1, NFIC, SPIB and THAP1), which have not been reported to be related to gastric cancer before. All of them might contribute to the pathogenesis since they are all related to other cancers and their target genes have been reported to play important roles in gastric tumorigenesis.

CONCLUSIONS: Our results may facilitate further therapeutic studies of gastric cancer.

Key words:
Gastric cancer, Transcription factors, miRNAs, Gene expression, Pathway, Gene ontology.

Introduction

Gastric cancer, which is the second leading cause of cancer-related death worldwide, affects about one million people per year[1,2]. Although its incidence is decreasing (especially in the West), it is still a major health problem by frequency, aggressiveness and low rate of cure in symptomatic stage. Therefore, improving gastric cancer therapeutic strategies has become a research hotspot.

Gene expression profiles combined with bioinformatics analysis have shown great application

Prospects in explore diagnosis and prognosis markers for complex diseases. Previous gene expression profile studies of gastric cancer have offered great help for understanding the pathogenesis of this disease[3-5]. Most of them mainly focused on the differentially expressed genes (DEGs), without considering MicroRNAs (miRNAs) and transcription factors (TFs) enriched target DEGs. MiRNAs and TFs are the most two important types of regulatory factors that determine gene expression. MiRNAs are 18-22 nucleotide small non-coding RNAs that control various biological processes through binding to the 3' untranslated region of mRNAs and affecting the stability and translation of target mRNAs. MiRNAs have recently been identified as crucial factors in not only tumorigenesis but also tumor aggressiveness[6,7]. It is believed that miRNAs are widely dysregulated in cancer and may be served as potential markers for cancer diagnosis, prognosis and treatment[8]. Previous studies have proposed several miRNAs that play important roles in gastric cancer tumorigenesis and might be potential diagnostic or prognosis biomarkers, such as miR-148a[9], miR-23a[10] and miR-205[11]. TFs are protein molecules which regulate gene expression through binding the cis-elements in target genes' promoter regions. Studies have shown that TFs, such ETS1, is a valuable marker of malignant potential in terms of gastric cancer invasiveness and metastasis[12]. Since both miRNAs and TFs are gene expression regulators, identification of miRNAs and TFs that enriched with target DEGs may provide new targets from further diagnosis and treatment.

In the current study, using microarray data from the gene expression omnibus (GEO) database, we aim to acquire functional categories (pathway and Gene Ontology items), TFs and miRNAs that enriched with DEGs so that to elab-

Corresponding Author: Yingjie Ma, MD; e-mail: mayingjie19@sina.com

orate the mechanism of gastric cancer. Our findings may reveal the possible mechanism of gastric cancer and provide potential therapeutic targets for further studies.

Materials and Methods

Microarray data

In this study, the gene expression profile GSE29272 from the GEO database was used for subsequent analysis. This series represents transcription profile of 268 samples: 134 gastric tumor tissues and 134 adjacent normal glands. The dataset was obtained by using the [HG-U133A] Affymetrix Human Genome U133A Array.

Identification of differentially expressed genes (DEGs)

Normalization of the raw data was performed in R (version 3.0.0) with the Robust Multi-array Analysis (RMA)[13]. The limma package in R was used to identify DEGs. Altered expression of probes was determined using *t*-tests, and the Benjamini-Hochberg method[14] was used for multiple test corrections. Probes with expression changes of $p < 0.05$ and corresponding False Discovery Rate (FDR) < 0.01 were considered to be statistically significant.

Functional enrichment analysis

To explore the functions and pathways of DEGs, we carried out enrichment analysis. DEGs were firstly mapped into the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) databases for annotation. Then hyper geometric distribution test was then used to identify biological processes significantly enriched with DEGs.

MiRNA analysis

For miRNA enrichment analysis, we first use Targetscan[15], miRanda[16] and Pita for miRNA target gene prediction for the DEGs. To avoid false positive results, target gene prediction were selected with the following criterion: 1) for targetscan analysis, float (context score) ≤ 0.3 and int (context score percentile) ≥ 85; 2) for miRanda analysis, prediction score over 500; 3) for Pita analysis, float (score) ≤ –10. In short, miRNA-DEG prediction results supported by all three prediction methods were considered to be confidential. MiRNAs with more than 10 target DEGs were used in further enrichment analysis with the hypergeometric distribution test.

TF analysis

For TF analysis, we first obtained TFs from the JASPAR database (http://jaspar.genereg.net/, version 5.0)[17], which is a collection of transcription factor DNA-binding preferences, modeled as matrices. Then we search TF binding motifs in 2kb upstream of DEGs. For each TF, a pseudocount was used to calculate the position-specific scoring matrix (PSSM). The pseudocount here is defined as sqrt(N)* background[nucleotide], where N represents the total number of sequences used to construct the matrix and a uniform background model over the four bases [0.25, 0.25, 0.25, 0.25] was used. Those with a relative score > 0.9 and FDR < 0.01 were used in further analysis to explore TFs that DEGs significantly enriched.

Results

DEGs between gastric cancer and normal glands

All 268 samples were put into the calculation of DEGs. Compared with normal tissues, a total of 977 DEGs were selected, including 492 down regulated genes and 485 overexpressed genes in gastric cancer tissue.

Results of functional enrichment

In order to explore the disturbed biological functions of DEGs, we conducted KEGG and GO enrichment analysis. All the pathways with corrected $p < 0.05$ and at least 10 DEGs were regarded as significant pathways (Table I). Among the 11 pathways, seven of them are involved in metabolism. In addition, the cell cycle pathway and a pathway in the translation process (ribosome biogenesis in eukaryotes) were included. The ECM-receptor interaction pathway was also indentified to be overrepresented with DEGs.

GO item enrichment analysis results are listed in Table II. All items enriched DEGs mainly involved in the process of metabolism (with the digestive item as the most significant one), cell cycle, ECM and translational process. DEGs enriched biological process items are shown in Figure 1. DEGs enriched cellular component and molecular function items are shown in Figure 2.

MiRNA analysis

MiRNAs can regulate gene expression and control various biological processes[18] through affecting the stability and translation of target mR-

Table I. Significant pathways enriched with DEGs.

KEGG_id	KEGG_description	KEGG_class	*p* value
hsa0071	Fatty acid degradation	Lipid metabolism	6.56E-05
hsa4974	Protein digestion and absorption	Digestive system	6.82E-04
hsa4512	ECM-receptor interaction	Signaling molecules and interaction	2.35E-03
hsa0280	Valine, leucine and isoleucine degradation	Amino acid metabolism	2.36E-03
hsa0980	Metabolism of xenobiotics by cytochrome P450	Xenobiotics biodegradation and metabolism	4.57E-03
hsa4110	Cell cycle	Cell growth and death	8.49E-03
hsa3008	Ribosome biogenesis in eukaryotes	Translation	1.01E-02
hsa5204	Chemical carcinogenesis	Cancers	1.02E-02
hsa4971	Gastric acid secretion	Digestive system	1.47E-02
hsa0982	Drug metabolism - cytochrome P450	Xenobiotics biodegradation and metabolism	1.77E-02
hsa1200	Carbon metabolism	Overview	2.13E-02

NAs. Therefore, it is significant to identify the miRNAs that regulate DEGs. The result of enrichment analysis for miRNAs is listed in Table 3. A total of eight miRNAs were identified with hsa-miR-486 as the most significant miRNA that enriched with target DEGs.

TF analysis

Transcription factors are important regulatory elements for downstream genes. With the identified DEGs, exploring the targets of transcription factor may help disclosure the pathogenesis of gastric cancer. Here we analyzed the upstream sequences of DEGs and presented the TFs enriched with target DEGs. The results of TF analysis is listed in Table IV. A total of 20 TFs were identified, with ETS1 as the most significant one.

Discussion

In this study, based on the GSE29272 from the GEO database, a total of 977 DEGs were identified between gastric cancer tissues and adjacent normal tissues.

DEGs were then used in enrichment analysis and 11 pathways were screened out (Table I). Seven pathways were involved in the metabolism process. Changes in metabolism pathways indicated alteration occurred in metabolites, which make it possible to diagnose gastric cancer. Other pathways included cell cycle, ribosome biogenesis in eukaryotes and the ECM-receptor interaction pathway. Changes in cell cycle pathway confirmed the strong proliferation of gastric cancer cells. Dysregulation of the translation process indicated that routine genetic information process-

Table II. Significant GO items enriched with DEGs.

GO_id	GO_description	GO_class	p value
GO:0007586	Digestion	Process	2.48E-04
GO:0030199	Collagen fibril organization	Process	9.79E-04
GO:0071294	Cellular response to zinc ion	Process	9.83E-03
GO:0044281	Small molecule metabolic process	Process	9.83E-03
GO:0022617	Extracellular matrix disassembly	Process	9.83E-03
GO:0030198	Extracellular matrix organization	Process	1.06E-02
GO:0030574	Collagen catabolic process	Process	1.15E-02
GO:0048407	Platelet-derived growth factor binding	Function	1.18E-02
GO:0005615	Extracellular space	Component	1.21E-02
GO:0005201	Extracellular matrix structural constituent	Function	1.21E-02
GO:0045926	Negative regulation of growth	Process	1.30E-02
GO:0051084	'De novo' posttranslational protein folding	Process	1.32E-02
GO:0031012	Extracellular matrix	Component	1.44E-02
GO:0005604	Basement membrane	Component	2.13E-02
GO:0031145	Anaphase-promoting complex-dependent proteasomal ubiquitin-dependent protein catabolic process	Process	2.25E-02
GO:0000278	Mitotic cell cycle	Process	3.90E-02
GO:0007179	Transforming growth factor beta receptor signaling pathway	Process	4.27E-02

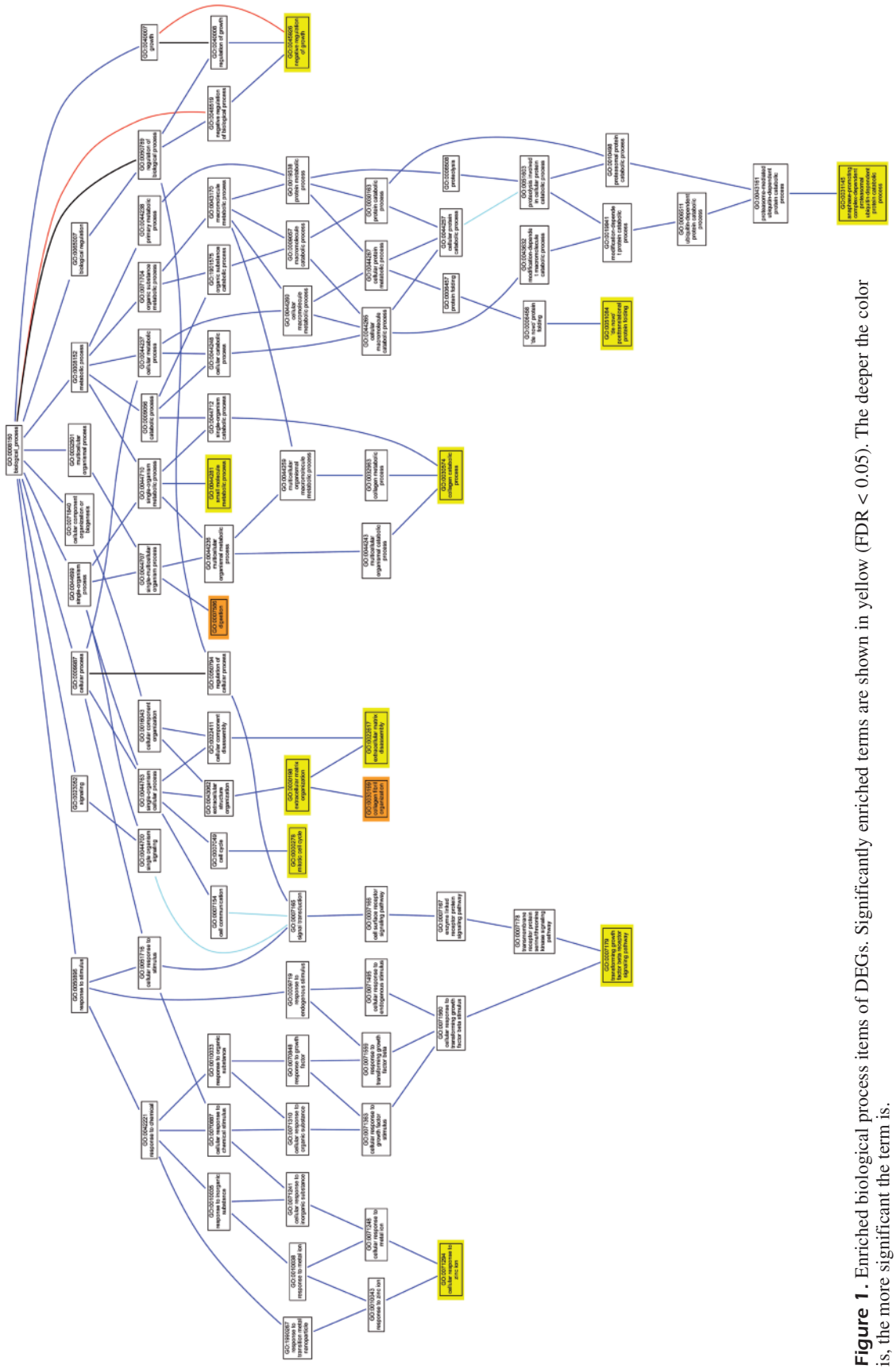

Figure 1. Enriched biological process items of DEGs. Significantly enriched terms are shown in yellow (FDR < 0.05). The deeper the color is, the more significant the term is.

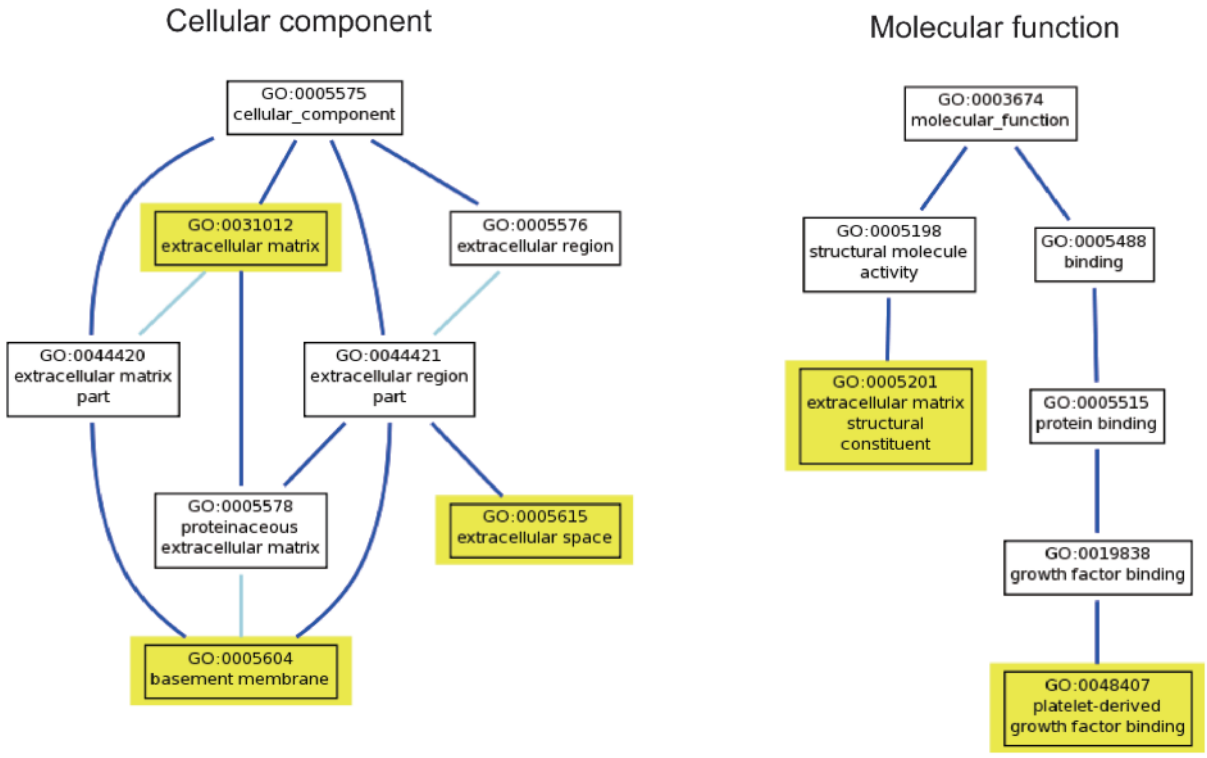

Figure 2. Enriched cellular components and molecular function items. Significantly enriched terms are shown in yellow (FDR < 0.05). The deeper the color is, the more significant the term is.

ing process may be interrupted in gastric cancer. The ECM-receptor interaction pathway has been identified in multiple cancers, suggesting its essential role in cancer biology[19].

GO enrichment analysis results are listed in Table II, including 11 biological process items (Figure 1), three cellular component items and 2 molecular function items (Figure 2). All items involve in metabolism, cell cycle, translation process and ECM, which further confirmed the speculation of pathway enrichment analysis.

MiRNA analysis showed that DEGs significantly share target sites of miRNAs. A total of eight miRNA enriched target DEGs were identified (Table III). Among them, miR-557 has not been reported to be related to gastric cancer before. It has been reported to be deregulated in other cancers, such as hepatocellular carcinoma[20] and breast cancer[21]. In addition, its target DEGs has been reported to be involved in the pathogenesis of gastric cancer. For example, *RUNX* is a well known tumor suppressors in gastric cancer[22] and up regulated expression of *ADAM17* is a prognostic marker in gastric cancer patients[23]. Further studies should be carried out to confirm the potential roles of miR-557 in the disease.

The DEGs were also share targets sites of TFs. Among the 20 TFs that enriched with target DEGs, four TFs including SPI1, NFIC, SPIB and THAP1 have not been reported to be related to gastric cancer before. Both *SPI1* and *SPIB* encode ETS-domain transcription factors that activate gene expression during myelopoiesis. Decrease of SPI1 expression by 80% would lead to acute myeloid leukemia (AML) in mice[24] and this TF has also been reported as potent tumor suppressor in classical Hodgkin lymphoma cells[25]. Although the relationship between *SPI1*

Table III. MiRNAs enriched with target DEGs.

Seed sequence	miRNAs	pvalue
CTGCCCC	hsa-miR-486	8.07E-03
AAACCAG	hsa-miR-29b-1	3.38E-02
TGCACTG	hsa-miR-152	4.05E-02
GTGCAAA	hsa-miR-557	4.22E-02
AATGTGA	hsa-miR-23a, hsa-miR-23b	4.28E-02
TGCACTG	hsa-miR-148a, hsa-miR-148b	4.99E-02

Table IV. Transcription factors enriched with target DEGs.

Name	Family	Class	*p* value
ETS1	Ets	Winged Helix-Turn-Helix	8.56E-65
GATA2	GATA	Zinc-coordinating	3.92E-57
TFAP2A	Helix-Loop-Helix	Zipper-Type	3.56E-45
SPI1	Ets	Winged Helix-Turn-Helix	4.35E-38
SP1	ββα-zinc finger	Zinc-coordinating	5.91E-28
KLF5	ββα-zinc finger	Zinc-coordinating	1.72E-22
NFIC	NFI CCAAT-binding	Other	1.90E-20
ZNF354C	ββα-zinc finger	Zinc-coordinating	1.58E-19
FOXC1	Forkhead	Winged Helix-Turn-Helix	1.60E-17
MZF1	ββα-zinc finger	Zinc-coordinating	5.11E-15
FOXL1	Forkhead	Winged Helix-Turn-Helix	7.67E-15
GATA3	GATA	Zinc-coordinating	6.64E-12
YY1	ββα-zinc finger	Zinc-coordinating	3.12E-11
SPIB	Ets	Winged Helix-Turn-Helix	5.75E-07
FOXP2	Forkhead	Winged Helix-Turn-Helix	7.11E-05
THAP1	THAP	Zinc-coordinating	1.73E-04
EGR1	ββα-zinc finger	Zinc-coordinating	4.08E-04
BRCA1	Other	Other	9.35E-04
FOXA1	Forkhead	Winged Helix-Turn-Helix	9.35E-04
SP2	ββα-zinc finger	Zinc-coordinating	1.44E-02

or *SPIB* and gastric cancer is still unknown, myelograms did show a reduced percentage of cells of the myeloid in gastric patients[26]. Moreover, target DEGs of them, such as *CDC25B*, has been linked to progression of gastric cancers and associated with a poor prognosis[27], indicating the implication of *SPI1* and *SPIB* in gastric tumorigenesis. NFIC encodes a putative tumor suppressor capable of directly repressing the transcription of cyclin D1 (*CCND1*) oncogene[28], which has been well recognized as a gastric cancer risk gene[29], implicating the potential involvement of NFIC in gastric cancer through its regulation of CCND1. *THAP1* encodes a sequence-specific DNA-binding factor which can modulate G1/S cell-cycle progression and cellular proliferation[30]. THAP1 could induce apoptosis in T-cell acute lymphoblastic leukemia though its regulation of the cell cycle and apoptosis regulator 1 (CCAR1) gene expression[31]. In addition, overexpression of one of its target DEGs, *SRI*, would result in multidrug resistance in gastric cancer cells[32]. Further investigations on the involvement of these TFs in gastric cancer pathogenesis are warranted.

Conclusions

With a microarray data set from the GEO database, we identified DEGs in gastric cancer tissues and normal tissues. Functional analysis revealed cell cycle, metabolism and ECM related biological processes as the significant items for gastric cancer. MiRNA and TF analysis identified a novel miRNA and four TFs enriched with DEGs that may play vital roles in the tumorigenesis. Our results may facilitate further therapeutic studies of gastric cancer.

Competing of interest

The authors have no financial conflicts of interest.

Reference

1) Pelucchi C, Lunet N, Boccia S, Zhang ZF, Praud D, Boffetta P, Levi F, Matsuo K, Ito H, Hu J, Johnson KC, Ferraroni M, Yu GP, Peleteiro B, Malekzadeh R, Derakhshan MH, Ye W, Zaridze D, Maximovitch D, Aragones N, Martin V, Pakseresht M, Pourfarzi F, Bellavia A, Orsini N, Wolk A, Mu L, Arzani D, Kurtz RC, Lagiou P, Trichopoulos D, Muscat J, La Vecchia C, Negri E. The stomach cancer pooling (StoP) project: study design and presentation. Eur J Cancer Prev 2014 Feb 20. [Epub ahead of print].
2) Herszenyi L, Tulassay Z. Epidemiology of gastrointestinal and liver tumors. Eur Rev Med Pharmacol Sci 2010; 14: 249-258.
3) Eftang LL, Esbensen Y, Tannaes TM, Blom GP, Bukholm IR, Bukholm G. Up-regulation of CLDN1 in gastric cancer is correlated with reduced survival. BMC Cancer 2013; 13: 586.
4) Pasini FS, Zilberstein B, Snitcovsky I, Roela RA, Mangone FR, Ribeiro U, Jr., Nonogaki S, Brito GC, Cal-

legari GD, Cecconello I, Alves VA, Eluf-Neto J, Chammas R, Federico MH. A gene expression profile related to immune dampening in the tumor microenvironment is associated with poor prognosis in gastric adenocarcinoma. J Gastroenterol 2013; Nov 12. [Epub ahead of print].

5) Wang G, Hu N, Yang HH, Wang L, Su H, Wang C, Clifford R, Dawsey EM, Li JM, Ding T, Han XY, Giffen C, Goldstein AM, Taylor PR, Lee MP. Comparison of global gene expression of gastric cardia and noncardia cancers from a high-risk population in china. PLoS One 2013; 8: e63826.
6) Calin GA, Croce CM. MicroRNA signatures in human cancers. Nat Rev Cancer 2006; 6: 857-866.
7) Waldman SA, Terzic A. Translating MicroRNA discovery into clinical biomarkers in cancer. JAMA 2007; 297: 1923-1925.
8) Garzon R, Marcucci G, Croce CM. Targeting microRNAs in cancer: rationale, strategies and challenges. Nat Rev Drug Discov 2010; 9: 775-789.
9) Xia J, Guo X, Yan J, Deng K. The role of miR-148a in gastric cancer. J Cancer Res Clin Oncol 2014, Mar 23.
10) Zhu LH, Liu T, Tang H, Tian RQ, Su C, Liu M, Li X. MicroRNA-23a promotes the growth of gastric adenocarcinoma cell line MGC803 and downregulates interleukin-6 receptor. FEBS J 2010; 277: 3726-3734.
11) Yin WZ, Li F, Zhang L, Ren XP, Zhang N, Wen JF. Down-regulation of microRNA-205 promotes gastric cancer cell proliferation. Eur Rev Med Pharmacol Sci 2014; 18: 1027-1032.
12) Yu Y, Zhang YC, Zhang WZ, Shen LS, Hertzog P, Wilson TJ, Xu DK. Ets1 as a marker of malignant potential in gastric carcinoma. World J Gastroenterol 2003; 9: 2154-2159.
13) Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 2003; 4: 249-264.
14) Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society. Series B (Methodological) 1995; 57: 289-300.
15) Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol Cell 2007; 27: 91-105.
16) John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS. Human MicroRNA targets. PLoS Biol 2004; 2: e363.
17) Stormo GD. DNA binding sites: representation and discovery. Bioinformatics 2000; 16: 16-23.
18) Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 2004; 116: 281-297.
19) Krupp M, Maass T, Marquardt JU, Staib F, Bauer T, Konig R, Biesterfeld S, Galle PR, Tresch A, Teufel A. The functional cancer map: a systems-level synopsis of genetic deregulation in cancer. BMC Med Genomics 2011; 4: 53.
20) Katayama Y, Maeda M, Miyaguchi K, Nemoto S, Yasen M, Tanaka S, Mizushima H, Fukuoka Y, Arii S, Tanaka H. Identification of pathogenesis-related microRNAs in hepatocellular carcinoma by expression profiling. Oncol Lett 2012; 4: 817-823.
21) Chen L, Li Y, Fu Y, Peng J, Mo MH, Stamatakos M, Teal CB, Brem RF, Stojadinovic A, Grinkemeyer M, McCaffrey TA, Man YG, Fu SW. Role of deregulated microRNAs in breast cancer progression using FFPE tissue. PLoS One 2013; 8: e54213.
22) Zhuang M, Gao W, Xu J, Wang P, Shu Y. The long non-coding RNA H19-derived miR-675 modulates human gastric cancer cell proliferation by targeting tumor suppressor RUNX1. Biochem Biophys Res Commun 2014; 448: 315-322.
23) Shou ZX, Jin X, Zhao ZS. Upregulated expression of ADAM17 is a prognostic marker for patients with gastric cancer. Ann Surg 2012; 256: 1014-1022.
24) Bonadies N, Pabst T, Mueller BU. Heterozygous deletion of the PU.1 locus in human AML. Blood 2010; 115: 331-334.
25) Yuki H, Ueno S, Tatetsu H, Niiro H, Iino T, Endo S, Kawano Y, Komohara Y, Takeya M, Hata H, Okada S, Watanabe T, Akashi K, Mitsuya H, Okuno Y. PU.1 is a potent tumor suppressor in classical Hodgkin lymphoma cells. Blood 2013; 121: 962-970.
26) Dobrodeev GV. [Bone marrow cellular makeup in stomach cancer]. Vopr Onkol 1979; 25: 20-23.
27) Takahashi H, Murai Y, Tsuneyama K, Nomoto K, Okada E, Fujita H, Takano Y. High labeling indices of cdc25B is linked to progression of gastric cancers and associated with a poor prognosis. Appl Immunohistochem Mol Morphol 2007; 15: 267-272.
28) Eeckhoute J, Carroll JS, Geistlinger TR, Torres-Arzayus MI, Brown M. A cell-type-specific transcriptional network required for estrogen regulation of cyclin D1 and cell cycle progression in breast cancer. Genes Dev 2006; 20: 2513-2526.
29) Loh M, Koh KX, Yeo BH, Song CM, Chia KS, Zhu F, Yeoh KG, Hill J, Iacopetta B, Soong R. Meta-analysis of genetic polymorphisms and gastric cancer risk: variability in associations according to race. Eur J Cancer 2009; 45: 2562-2568.
30) Cayrol C, Lacroix C, Mathe C, Ecochard V, Ceribelli M, Loreau E, Lazar V, Dessen P, Mantovani R, Aguilar L, Girard JP. The THAP-zinc finger protein THAP1 regulates endothelial cell proliferation through modulation of pRB/E2F cell-cycle target genes. Blood 2007; 109: 584-594.
31) Lu C, Li JY, Ge Z, Zhang L, Zhou GP. Par-4/THAP1 complex and Notch3 competitively regulated pre-mRNA splicing of CCAR1 and affected inversely the survival of T-cell acute lymphoblastic leukemia cells. Oncogene 2013; 32: 5602-5613.
32) He Q, Zhang G, Hou D, Leng A, Xu M, Peng J, Liu T. Overexpression of sorcin results in multidrug resistance in gastric cancer cells with up-regulation of P-gp. Oncol Rep 2011; 25: 237-243.