

Identifying key stage-specific genes and transcription factors for gastric cancer based on RNA-sequencing data

Yan Wang, MD*

Abstract

Background: To identify gastric cancer (GC)-associated genes and transcription factors (TFs) using RNA-sequencing (RNA-seq) data of Asians.

Materials and methods: The RNA-seq data (GSE36968) were downloaded from Gene Expression Omnibus database, including 6 noncancerous gastric tissue samples, 5 stage I GC samples, 5 stage II GC samples, 8 stage III GC samples, and 6 stage IV GC samples. The gene expression values in each sample were calculated using Cuffdiff. Following, stage-specific genes were identified by 1-way analysis of variance and hierarchical clustering analysis. Upstream TFs were identified using Seqpos. Besides, functional enrichment analysis of stage-specific genes was performed by DAVID. In addition, the underlying protein–protein interactions (PPIs) information among stage IV-specific genes were extracted from STRING database and PPI network was constructed using Cytoscape software.

Results: A total of 3576 stage-specific genes were identified, including 813 specifically up-regulated genes in the normal gastric tissues, 2224 stage I and II-specific genes, and 539 stage IV-specific genes. Also, a total of 9 and 11 up-regulated TFs were identified for the stage I and II-specific genes and stage IV-specific genes, respectively. Functional enrichment showed *SPARC*, *MMP17*, and *COL6A3* were related to extracellular matrix. Notably, 2 regulatory pathways *HOXA4-GLI3-RUNX2-FGF2* and *HMG2-PRKCA* were obtained from the PPI network for stage IV-specific genes. In the PPI network, TFs *HOXA4* and *HMG2* might function via mediating other genes.

Conclusion: These stage-specific genes and TFs might act in the pathogenesis of GC in Asians.

Abbreviations: ANOVA = analysis of variance, BP = biological processes, CC = cellular component, DEG = differentially expressed gene, ECM = extracellular matrix, GC = gastric cancer, PPI = protein–protein interaction, TF = transcription factor.

Keywords: functional enrichment analysis, gastric cancer, protein–protein interaction network, stage-specific genes, transcription factors

1. Introduction

Gastric cancer (GC) is the second-leading cause of cancer-related death worldwide.^[1] In recent decade, even with multiple therapies, the prognosis of GC remains poor with 5-year survival rate ranging between 25% and 35%.^[2] The low survival rate is mainly due to the undetectable characteristics of GC at early stage.^[3] Therefore, it is of great significance to identify the molecular biomarkers pivotal for development and progression of GC.

Recently, several studies have devoted to investigate the mechanisms of GC. Downregulated X-linked inhibitor of apoptosis (*XIAP*) in human GC cells causes apoptosis and chemotherapeutic sensitivities, and it may contribute to develop a new therapeutic strategy for GC.^[4] Previous study indicates that the transcriptional silencing of deleted in liver cancer 1 (*DLC-1*), which is induced by epigenetic mechanism, may play a role in gastric carcinogenesis.^[5] Especially, expression profiles of GC have been extensively investigated, yielding useful insights into the molecular mechanism of carcinogenesis of GC. For example, by performing a genome-wide expression analysis, Junnila et al identified 58 differentially expressed genes (DEGs) between cancerous and non-neoplastic gastric tissues, among which overexpressed *CXCL1* was positively associated with improved survival.^[6] Lei et al screened out 260 DEGs in gastric tissues and further proposed serpin epidase inhibitor, clade H1 (*SERPINH1*) and G protein-coupled receptor family C type 5A (*GPRC5A*) as potential biomarkers for GC.^[7] However, a microarray-based method has those following limitations: background levels of hybridization limit the accuracy of expression measurements, particularly for transcripts with low abundance^[8]; prior knowledge of the genome is used to design probes for hybridization with targets (DNA or RNA), and therefore the unknown genomic regions cannot be interrogated.^[9]

RNA-sequencing (RNA-seq) technology emerges as a new and powerful tool superior to hybridization-based approach in

Editor: Chunxia Cao.

The authors have no funding and conflicts of interest to disclose.

Department of Gastroenterology, Shengjing Hospital of China Medical University, Shenyang, China.

* Correspondence: Yan Wang, Department of Gastroenterology, Shengjing Hospital of China Medical University, No. 36 Sanhao Road, Shenyang 110004, China (e-mail: wycharm@163.com).

Copyright © 2017 the Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Medicine (2017) 96:4(e5691)

Received: 27 July 2016 / Received in final form: 11 November 2016 / Accepted: 29 November 2016

<http://dx.doi.org/10.1097/MD.0000000000005691>

measuring gene expression without those limitations mentioned earlier.^[10] This technology can provide a more detailed and precise view of the entirely expressed transcripts, including low-expressed genes, alternative splice variants, and novel transcripts.^[11] In 2012, by performing RNA-seq, Kim et al obtain a global view of transcriptome in Asian GC and subsequently identify the regulatory effects of AMP-activated protein kinase (AMPK) $\alpha 2$ on energy homeostasis.^[12] While, the obtained results are restricted and RNA-seq data need deeper exploration.

Bioinformatics methods can help excavate more useful information from the large number of microarray data or sequencing data.^[13] Based on bioinformatics methods, this study reanalyzed the published RNA-seq data.^[12] First, the stage-specific genes in different stage GC samples were identified, and the stage-specific genes which have transcriptional regulation function of transcription factors (TFs) were annotated. Then, their potential functions were predicted by functional enrichment analysis. Moreover, the underlying interactions among stage IV-specific genes were investigated by protein-protein interaction (PPI) network.

2. Methods

2.1. RNA-seq data

RNA-seq data of GSE36968 were downloaded from Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo/>) established at the National Center for Biotechnology Information (NCBI) database,^[12] which were previously sequenced using paired-end sequencing based on the GPL9442 AB SOLiD System 3.0 platform. GSE36968 included 6 normal gastric tissue samples, 5 stage I GC samples, 5 stage II GC samples, 8 stage III GC samples, and 6 stage IV GC samples of Asians. The tissue samples were taken from the National Research Resource Bank Program of the Korea Science and Engineering Foundation in the Ministry of Science and Technology. And tumor stages were identified and classified by a pathologist from the Gene Bank at Yonsei University Severance Hospital. The RNA-seq data used in this study were downloaded from GEO database, thus ethical approval and informed consent were not necessary.

2.2. Data preprocessing

The author used Tophat2 (version 2.0.10, <http://ccb.jhu.edu/software/tophat/index.shtml>)^[14] to align the RNA-seq reads against the University of California Santa (UCSC) *hg19* genome sequences (<http://www.genome.ucsc.edu/index.html>) with default parameters, allowing 1 or 2 base mismatches. Also, the sequence alignment should be unique for each read. Following, the author assembled the transcripts and calculated the gene expression levels using Cuffdiff and fragments per kilobase of exon model per million fragments mapped method in Cufflinks, respectively.^[14]

2.3. Stage-specific gene screening

One-way analysis of variance (ANOVA) was applied to screen DEGs among different groups with the cutoff criterion of $P < 0.05$. Subsequently, the screened DEGs were conducted with hierarchical clustering analysis to determine stage-specific genes for specific GC stages.

2.4. Transcription factor screening

Furthermore, the author annotated the stage-specific genes which have transcriptional regulation function of TFs based on the

TRANSFAC database (<http://transfac.gbf.de/TRANSFAC/>).^[15] Besides, the author defined upstream 1.5 kb region of the transcription start site as promoter region and then perform Motif Finding for all stage-specific genes using Seqpos.^[16] $P < 0.00001$ was used as the threshold.

2.5. Functional enrichment analysis

DAVID software was employed to identify significantly enriched Gene Ontology functions in biological processes (BP), molecular function, and cellular component (CC) categories for stage-specific genes.^[17]

2.6. Construction of PPI network

STRING is developed as an online database resource, which provides uniquely comprehensive information for assembling, evaluating, and disseminating PPI in a user-friendly way (<http://string-db.org/>).^[18] To reveal the underlying interactions among stage IV-specific genes, the author further extracted the PPI information from STRING database^[18] and constructed PPI network using Cytoscape software (<http://cytoscape.org/>).^[19]

3. Results

3.1. Stage-specific genes screening

Based on ANOVA and subsequent hierarchical clustering analysis, the author systematically compared the expression profiles at different GC stages, and finally obtained the heat map of stage-specific genes (Fig. 1). A total of 3576 genes with stage-specific expression patterns were identified. Among them, 813 DEGs (23%, 813/3576) were specifically highly expressed in the normal gastric tissues, and 2224 DEGs (62%, 2224/3576; e.g., Kinesin family member C1, *KIFC1*; and septin 2, *SEPT2*) which were specifically highly expressed in stage I had extremely similar expression patterns in stage II GC samples, as well as 539 specifically highly expressed genes (15%, 539/3576; e.g., Neuropilin-2, *NRP2*; collagen triple helix repeat containing-1, *CTHRC1*; secreted protein, acidic, cysteine-rich, osteonectin, *SPARC*; matrix metalloproteinase 17, *MMP17*; and collagen, type VI, alpha 3, *COL6A3*) in stage IV GC samples. However, no stage-specific genes were found in stage III GC samples.

3.2. Transcription factors screening

Based on the TRANSFAC database, a total of 82 TFs were screened from the stage I and II-specific genes and 22 TFs from the stage IV-specific genes. Following, the author further explored the upstream regulators, which might have regulatory effects on stage-specific genes in stage I and II GC and stage IV GC samples. A total of 9 TFs (e.g., high mobility group box-1, *HMGB1*; and homeobox A13, *HOXA13*) were identified to be significantly enriched in the promoter region of highly expressed stage I and II-specific genes, which also experienced over-expression in stage I and II GC and stage IV GC samples (Table 1). Besides, 11 TFs were screened for the stage IV-specific genes, which were up-regulated in stage IV GC samples (Table 1).

3.3. Functional enrichment analysis

Functional enrichment analysis was performed separately for the GC stage-specific genes, and the results were shown in

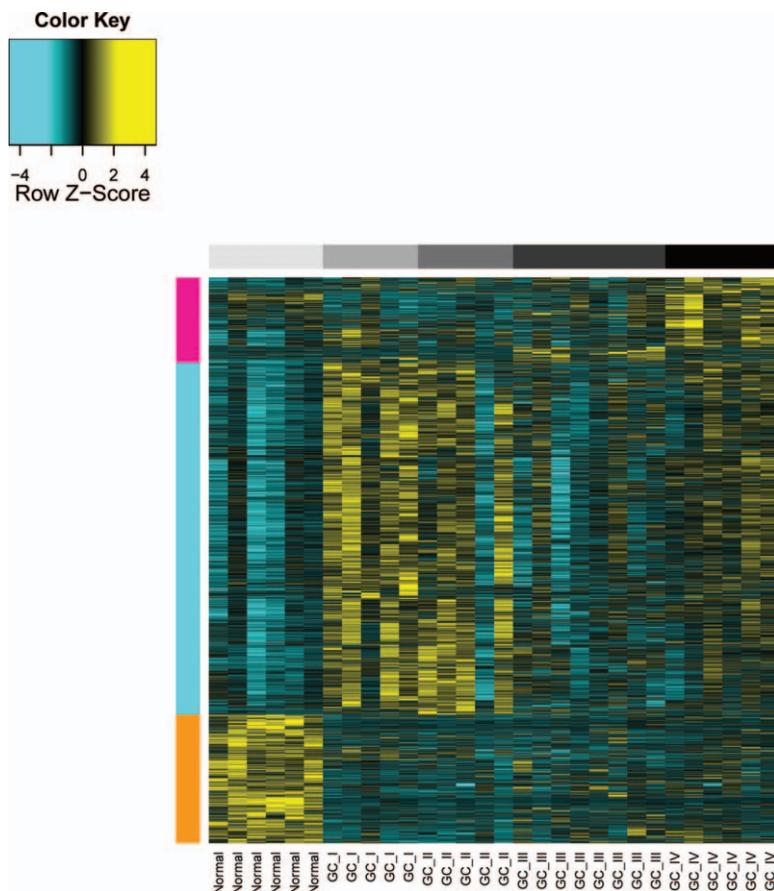


Figure 1. The heat map of hierarchical clustering analysis. The orange, blue, and pink bars represent stage-specific genes in normal stomach tissues, stage I and II GC samples and stage IV GC samples, respectively. GC=gastric cancer.

Fig. 2. For stage-specific genes in normal tissues, functions including digestion, ion transport, and G-protein signaling were significantly enriched. Besides, the stage-specific genes in stage I and II GC samples were mainly associated with mitotic cell cycle and RNA processing. Additionally, the stage-specific genes in stage IV GC samples were dramatically enriched in functions, including vasculature development, skeletal system development, and urogenital system development. Accordingly, GC stage IV-specific genes were significantly enriched in BP such as cell migration and negative regulation of cell differentiation.

Importantly, the author further focused on the involved genes which could encode proteins for CC, including extracellular matrix (ECM) and cell surface. As shown in Table 2, the author identified 24, 7, and 22 genes involved in ECM separately in normal gastric tissues, stage I and II GC samples, and stage IV GC samples, respectively. Besides, 22, 29, and 16 genes were highly

expressed at cell surface in normal, GC stage I and II, and stage IV tissues, respectively.

3.4. PPI network analysis

Based on the PPI information from STRING database, the author further constructed the PPI network for stage IV-specific genes using Cytoscape. A total of 13 nodes were identified with degree >10 in the PPI network (Fig. 3). In addition, homeo box A4 (*HOXA4*), GLI family zinc finger 3 (*GLI3*), runt-related TF 2 (*RUNX2*), fibroblast growth factor 2 (*FGF2*), high mobility group AT-hook 2 (*HMG A2*), and protein kinase C, alpha (*PRKCA*) were involved in 2 regulatory pathways which were considered to play key roles in stage IV GC, including *HOXA4-GLI3-RUNX2-FGF2* and *HMG A2-PRKCA* (Fig. 3). In the first pathway, *HOXA4*, *GLI3*, and *RUNX2* were identified as TFs. And in the second pathway, *HMG A2* was confirmed as TF.

Table 1
Predicted upstream TFs for stage-specific genes.

Stage	TF motif counts	TF genes
GC I and II	9	<i>BDP1, CDC5L, E2F6, HMGB1, HOXA13, SIRT6, SLC20A1, SRF, TGIF1</i>
GC IV	11	<i>BARHL2, BARX1, EN1, HMGA2, HOXA4, NFATC4, PBX3, RUNX1, SHOX2, TEAD1, ZEB1</i>

GC=gastric cancer, TF=transcription factor.

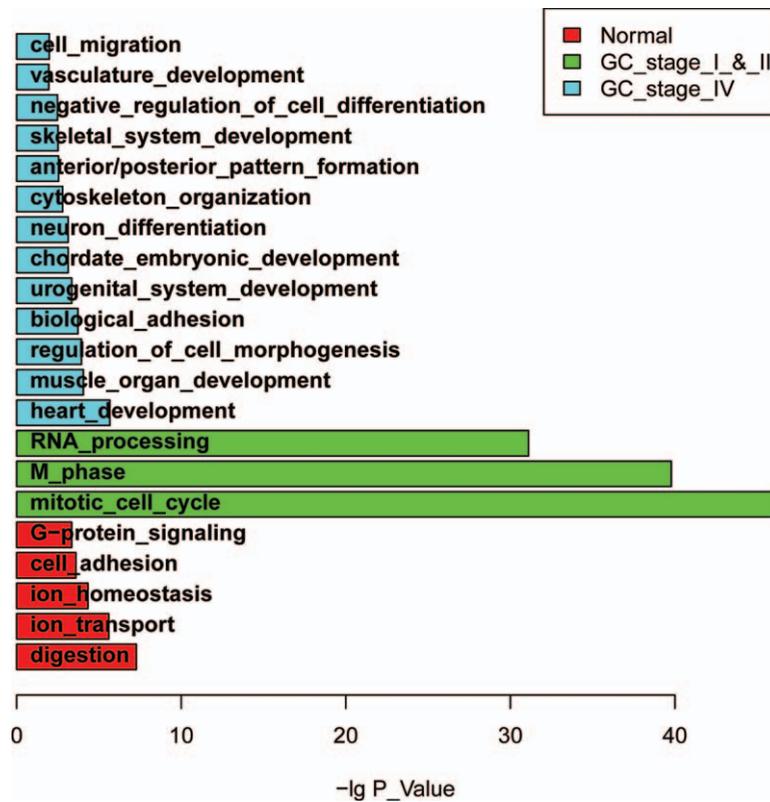


Figure 2. Functional enrichment results of normal gastric-specific genes and stage I and II-specific genes and stage IV-specific genes. GC = gastric cancer.

4. Discussion

A total of 2224 stage I and II-specific genes were identified. Moreover, those genes were mainly associated with mitotic cell cycle and RNA processing, implying their overexpression might promote cancer cells proliferation. For example, *KIF14*, a C-terminal kinesin motor, belongs to the minus-end-directed kinesin-14 family in human cells.^[20] Several studies have shown that *KIF14* plays an essential role in mitosis and meiosis by organizing and stabilizing spindles using its sliding activity along microtubules.^[21,22] The deletion of *KIF14* in cells may lead to eventual apoptosis.^[23] In addition, *SEPT2* has also been proved to play an important role in carcinoma cell division and proliferation through affecting cytoskeletons, which could be suggested as a promising target for GC therapy.^[24] Therefore, the author speculated that the overexpressed of *KIF14* and *SEPT2* in stage I and stage II GC might contribute to the aggravated cell proliferation.

On the other hand, 539 stage IV-specific genes were screened, which were significantly enriched in functions, including vasculature development, skeletal system development, neuron differentiation, and urogenital system development. For instance, *Nrp2* is a transmembrane glycoprotein receptor for vascular endothelial growth factor-C, which is an important lymphangiogenic factor and exerts crucial role in lymph node metastasis of various human cancers including GC.^[25] Several studies have reported the up-regulation of *CTHRC1* in various human solid cancers and associated it with cancer cell adhesion to ECM in the progression of cancer invasion and metastasis.^[26,27] Moreover, the author identified several ECM-related genes with up-regulation in stage IV of GC, including *SPARC*, *MMP17*, and *COL6A3*. The relative expression level of the *SPARC* was shown to be higher in GC tissues than in adjacent normal mucosae, leading to lower 5-year overall survival.^[28] The down-regulation of *SPARC* inhibits invasion and growth of human GC cells.^[29] *MMP-17* belongs to the membrane type-MMPs subfamily anchored to the plasma

Table 2
Extracellular matrix and cell surface-associated stage-specific genes.

Subcellular location	Cancer stage	Gene counts	Gene lists
Extracellular matrix	Normal	24	<i>ODAM, ADAMTSL1, AMTN, SPOCK3, MAMDC2, ADAMTS13, ADAMTSL3, PTPRZ1, FGF9, ADAMTS15</i> , etc.
	GC I and II	7	<i>CD44, SERAC1, SOD1, SMC3, GFOD2, TIMP1, ANXA2</i>
	GC IV	22	<i>ASP, CTHRC1, COL4A1, SPARCL1, TNC, FBN1, MMP17, SPOCK1, SPARC, NID2</i> , etc.
Cell surface	Normal	22	<i>COL23A1, SLC6A2, ADAMTS13, MPP2, SLC9A3, TNFRSF13B, PTPRT, ADA, NCAM1, CD36</i> , etc.
	GC I and II	29	<i>PPFIA3, ENOX2, SEPT2, TNFRSF12A, IDE, ENPEP, IL17RB, HMMR, GSR, ACE</i> , etc.
	GC IV	16	<i>DCBLD2, LY75, PARD3, BMPR2, MMP17, TIMP2, GREM1, ITGB1, THY1, LAYN</i> , etc.

GC = gastric cancer.

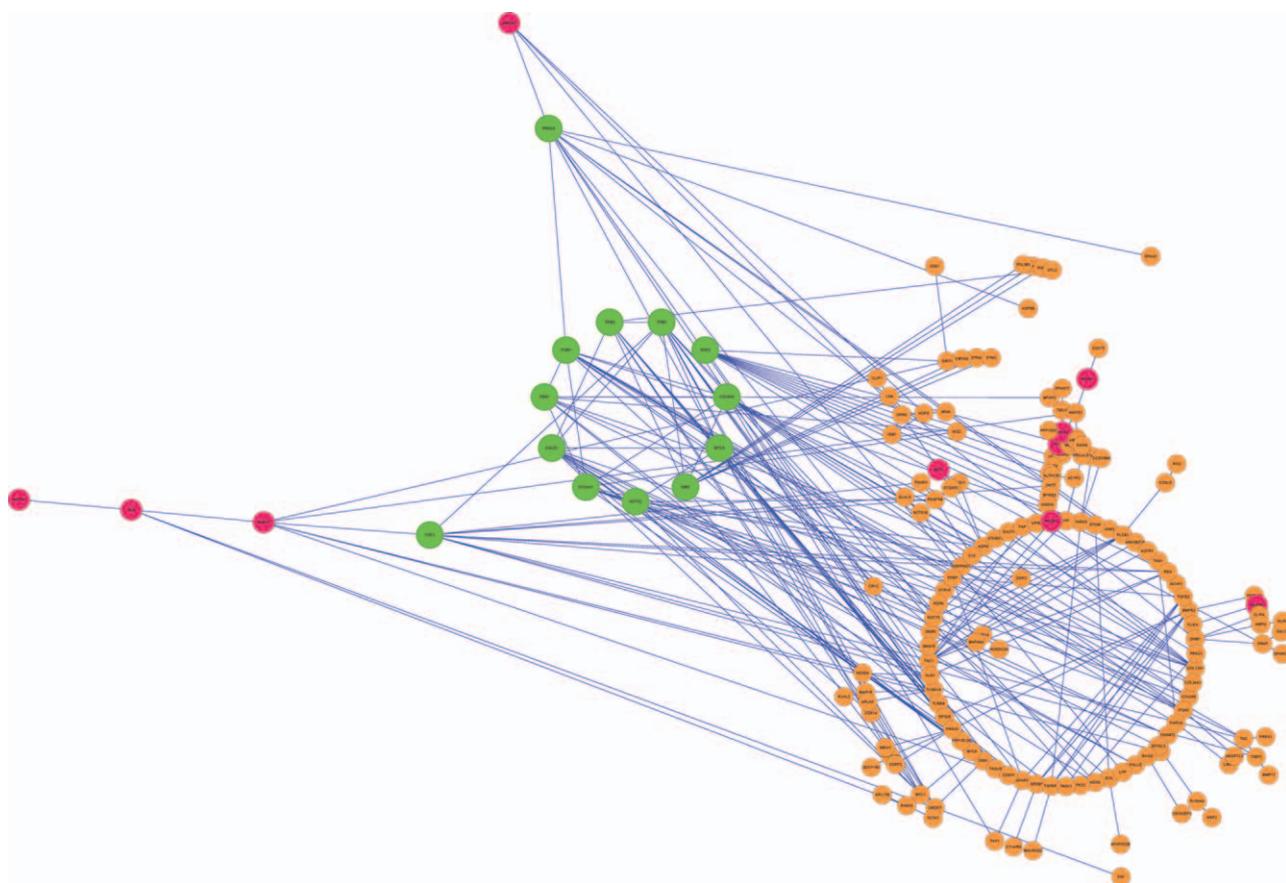


Figure 3. The PPI network of stage IV-specific genes. The red nodes represent transcription factors, the green nodes represent genes with degree >10, and the orange nodes represent other stage IV-specific genes. PPI = protein-protein interaction.

membrane via a glycosyl-phosphatidyl inositol anchor,^[30] which is highly expressed in human cancers and associated with cancer progression.^[30,31] Besides, *COL6A3* has also been found to participate in GC progression^[32] by regulating ECM-receptor interaction pathway.^[33] Thus, the author could speculate that these up-regulated and stage IV-specific genes may contribute to a gradual disappearance of tissue specificity and diverse cell differentiation potentials via affecting various BP.

Besides, by performing Motif Scanning, the author also screened out some key TFs for different stage-specific genes, which can bind the promoter sequence of most stage-specific genes. Herein, 9 TFs were obtained in stage I and II GC, such as *HMGB1* and *HOXA13*, which also experienced over-expression in stage I and II GC and stage IV GC samples. In the previous study, serum *HMGB1* level is significantly associated with invasion, lymph node metastasis, tumor size, and poor prognosis of GC.^[34] *HMGB1* silencing significantly decreases cell proliferation by regulating cell cycle level and cell cycle-related gene cyclin D1 expression, and inhibits cellular metastatic ability by down-regulating *MMP-9* expression in GC MGC-803 cells.^[35] Besides, Homeobox (*HOX*) gene family is known to play crucial roles in tumorigenesis.^[36] Recent study indicates that up-regulated *HOXA13* may be a novel prognostic marker in GC and correlation with TNM stage, histological differentiation, relapse, and overall survival and disease-free survival rate.^[37] The up-regulation of TFs *HMGB1* and *HOXA13* might contribute to the overexpression of genes in stage I and stage II GC.

Furthermore, by constructing PPI network for the stage IV-specific genes, 2 interaction pathways were obtained, including *HOXA4-GLI3-RUNX2-FGF2* and *HMGA2-PRKCA*. Among these interaction networks, *HOXA4*, *GLI3*, *RUNX2*, and *HMGA2* were TFs. Previously, *HOXA4* expression is increased in invasive tumors to inhibit cell invasion.^[38] *HMGA2* silencing induces apoptosis and suppresses proliferation of GC MKN-45 cells.^[39] The first pathway indicated that activation of *RUNX2* was based on *HOXA4-GLI3* interaction and *GLI3-RUNX2* interaction, which may subsequently regulated *FGF2*-mediated expression activation of downstream genes in stage IV GC samples. The second regulatory pathway *HMGA2-PRKCA* suggested that the *HMGA2* might regulate the downstream genes via interacting with *PRKCA* in stage IV GC samples.

5. Conclusion

In conclusion, a total of 3576 stage-specific genes were identified. Also, a total of 9 and 11 up-regulated TFs were identified for the stage I and II-specific genes and stage IV-specific genes, respectively. Besides, *KIF14*, *SEPT2*, *NRP2*, *CTHRC1*, *SPARC*, *MMP17*, *COL6A3*, *HMGB1*, *HOXA13*, *HOXA4*, and *HMGA2* might play important roles in pathogenesis of GC in Asians. The present study provided a novel insight into the molecular mechanisms underlying different GC stages. However, the results were obtained by bioinformatics methods and the

sample sizes were small, thus further experimental validations were still needed.

References

- [1] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin* 2016;66:10–29.
- [2] Yang W, Raufi A, Klempner SJ. Targeted therapy for gastric cancer: molecular pathways and ongoing investigations. *Biochim Biophys Acta* 2014;1846:232–7.
- [3] Orditura M, Galizia G, Sforza V, et al. Treatment of gastric cancer. *World J Gastroenterol* 2014;20:1635–49.
- [4] Wang DG, Sun YB, Ye F, et al. Anti-tumor activity of the X-linked inhibitor of apoptosis (XIAP) inhibitor embelin in gastric cancer cells. *Mol Cell Biochem* 2014;386:143–52.
- [5] He D, Zhang YW, Zhang NN, et al. Aberrant gene promoter methylation of p16, FHIT, CRBP1, WWOX, and DLC-1 in Epstein-Barr virus-associated gastric carcinomas. *Med Oncol* 2015;32:1–6.
- [6] Junnila S, Kokkola A, Mizuguchi T, et al. Gene expression analysis identifies over-expression of CXCL1, SPARC, SPP1, and SULF1 in gastric cancer. *Genes Chromosomes Cancer* 2010;49:28–39.
- [7] Cheng L, Yang S, Yang Y, et al. Global gene expression and functional network analysis of gastric cancer identify extended pathway maps and GPRC5A as a potential biomarker. *Cancer Lett* 2012;326:105–13.
- [8] Gautier L, Cope L, Bolstad BM, et al. affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 2004;20:307–15.
- [9] Ng HK, Roukos DH, Iacopetta B, et al. Next-Generation Sequencing in Cancer Epigenomics and Potential Clinical Applications Next Generation Sequencing in Cancer Research. 2013;Springer, New York, NY:31–53.
- [10] De KE, den Dunnen JT, 't Hoen PA. RNA sequencing: from tag-based profiling to resolving complete transcript structure. *Cell Mol Life Sci* 2014;71:3537–51.
- [11] Han H, Jiang X. Disease biomarker query from RNA-seq data. *Cancer Inform* 2014;13(suppl 1):81–94.
- [12] Kim YH, Liang H, Liu X, et al. AMPK α modulation in cancer progression: multilayer integrative analysis of the whole transcriptome in Asian gastric cancer. *Cancer Res* 2012;72:2512–21.
- [13] Oliver GR, Hart SN, Klee EW. Bioinformatics for clinical next generation sequencing. *Clin Chem* 2014;61:124–35.
- [14] Ghosh S, Chan CK. Analysis of RNA-seq data using TopHat and Cufflinks. *Methods Mol Biol* 2016;1374:339–61.
- [15] Wingender E, Chen X, Hehl R, et al. TRANSFAC: an integrated system for gene expression regulation. *Nucleic Acids Res* 2000;28:316–9.
- [16] He HH, Meyer CA, Shin H, et al. Nucleosome dynamics define transcriptional enhancers. *Nat Genet* 2010;42:343–7.
- [17] Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009;4:44–57.
- [18] Szklarczyk D, Franceschini A, Wyder S, et al. STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* 2015;43:447–52.
- [19] Demchak B, Hull T, Reich M, et al. Cytoscape: the network visualization tool for GenomeSpace workflows. *F1000 Res* 2014;3:151–251.
- [20] Oue N, Mukai S, Trang PTB, et al. Abstract 1413: significance of KIFC1 in spheroid formation of gastric cancer cells. *Cancer Res* 2015;75(15 suppl):1413–513.
- [21] Cai S, Weaver LN, Ems-McClung SC, et al. Kinesin-14 family proteins HSET/XCTK2 control spindle length by cross-linking and sliding microtubules. *Mol Biol Cell* 2009;20:1348–59.
- [22] Cai S, Weaver LN, Ems-McClung SC, et al. Proper organization of microtubule minus ends is needed for midzone stability and cytokinesis. *Curr Biol* 2010;20:880–5.
- [23] Thériault BL, Corson TW. Kif14: A Clinically Relevant Kinesin and Potential Target for Cancer Therapy. Springer, The Netherlands: 2015.
- [24] Yu W, Ding X, Chen F, et al. The phosphorylation of SEPT2 on Ser218 by casein kinase 2 is important to hepatoma carcinoma cell proliferation. *Mol Cell Biochem* 2009;325:61–7.
- [25] Gou H-F, Chen X-C, Zhu J, et al. Expressions of COX-2 and VEGF-C in gastric cancer: correlations with lymphangiogenesis and prognostic implications. *J Exp Clin Cancer Res* 2011;30:14.
- [26] Chen Y-L, Wang T-H, Hsu H-C, et al. Overexpression of CTHRC1 in hepatocellular carcinoma promotes tumor invasion and predicts poor prognosis. *PLoS ONE* 2013;8:e70324.
- [27] Wang P, Wang YC, Chen XY, et al. CTHRC1 is upregulated by promoter demethylation and transforming growth factor- β 1 and may be associated with metastasis in human gastric cancer. *Cancer Sci* 2012;103:1327–33.
- [28] Sato T, Oshima T, Yamamoto N, et al. Clinical significance of SPARC gene expression in patients with gastric cancer. *J Surg Oncol* 2013;108:364–8.
- [29] Yin J, Chen G, Liu Y, et al. Downregulation of SPARC expression decreases gastric cancer cellular invasion and survival. *J Exp Clin Cancer Res* 2010;29:59.
- [30] Sohail A, Sun Q, Zhao H, et al. MT4-(MMP17) and MT6-MMP (MMP25), a unique set of membrane-anchored matrix metalloproteinases: properties and expression in cancer. *Cancer Metastasis Rev* 2008;27:289–302.
- [31] Gencer S, Cebeci A, Irmak-Yazicioglu MB. Matrix metalloproteinase gene expressions might be oxidative stress targets in gastric cancer cell lines. *Chin J Cancer Res* 2013;25:322–33.
- [32] Cui J, Li F, Wang G, et al. Gene-expression signatures can distinguish gastric cancer grades and stages. *PLoS ONE* 2011;6:e17819.
- [33] Hu K, Chen F. Identification of significant pathways in gastric cancer based on protein–protein interaction networks and cluster analysis. *Genet Mol Biol* 2012;35:701–8.
- [34] Zhang J, Zhang R, Lu WW, et al. Clinical significance of hmgb1 expression in human gastric cancer. *Int J Immunopathol Pharmacol* 2014;27:543–51.
- [35] Song B, Song WG, Li ZJ, et al. Effect of HMGB1 silencing on cell proliferation, invasion and apoptosis of MGC-803 gastric cancer cells. *Cell Biochem Funct* 2012;30:11–7.
- [36] Bhatlekar S, Fields JZ, Boman BM. HOX genes and their role in the development of human cancers. *J Mol Med* 2014;92:811–23.
- [37] Chang S, Liu J, Guo S, et al. HOTTIP and HOXA13 are oncogenes associated with gastric cancer progression. *Oncol Rep* 2016;35:3577–85.
- [38] Klausen C, Leung PC, Auersperg N. Cell motility and spreading are suppressed by HOXA4 in ovarian cancer cells: possible involvement of β 1 integrin. *Mol Cancer Res* 2009;7:1425–37.
- [39] Wei CH, Wei LX, Lai MY, et al. Effect of silencing of high mobility group A2 gene on gastric cancer MKN-45 cells. *World J Gastroenterol* 2013;19:1239–46.