Original Article

Identification of hub genes in gastric cancer by integrated bioinformatics analysis

Feng Sun, Chen Zhang, Shichao Ai, Zhijian Liu, Xiaofeng Lu

Department of Gastrointestinal Surgery, Nanjing Drum Tower Hospital, the Affiliated Hospital of Nanjing University Medical School, Nanjing, China

Contributions: (I) Conception and design: Z Liu, X Lu. (II) Administrative support: F Sun; (III) Provision of study materials or patients: F Sun, C Zhang; (IV) Collection and assembly of data: F Sun, C Zhang, S Ai; (V) Data analysis and interpretation: F Sun, Z Liu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Zhijian Liu, Xiaofeng Lu. Department of Gastrointestinal Surgery, Nanjing Drum Tower Hospital, the Affiliated Hospital of Nanjing University Medical School, Nanjing, China. Email: njulzj@126.com; lxf_njglyy@sina.com.

Background: Gastric cancer (GC) is one of the most common cancer worldwide. With the high rates of metastasis and recurrence, its overall survival remains poor at the present time. Hence, seeking new potential therapeutic targets of GC is important and urgent.

Methods: We retrieved the gene expression profiles and clinical data from The Cancer Genome Atlas (TCGA) datasets. After screening differentially expressed genes (DEGs), we carried out the survival analysis for overall survival to pick out robust DEGs. To explore the role of these robust DEGs, we conducted Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) functional enrichment analyses. Subsequently, protein interactions network was constructed utilizing the Search Tool for the Retrieval of Interacting Genes (STRING) database. We then presented the module analysis and filtered out hub genes by the Cytoscape software. Finally, Kaplan-Meier analysis was utilized to demonstrate the prognostic role of these hub genes.

Results: According to the gene expression profiles of TCGA and the survival analysis, 238 robust DEGs were filtered out, consisting of 140 up-regulated and 98 down-regulated genes. The up-regulated DEGs were mainly enriched in systemic lupus erythematosus, cytokine activity, and alcoholism, while down-regulated DEGs were mainly enriched in steroid hormone receptor activity, immune response, and metabolism. Through the construction of the protein-protein interaction (PPI) network, eight hub genes were finally screened out, including *CCR8*, *HIST1H3B*, *HIST1H2AH*, *HIST1H2AJ*, *NPY*, *HIST2H2BF*, *GNG7*, and *CCL25*.

Conclusions: Our study picked out eight hub genes, which might be potential prognostic biomarkers for GC and even be treatment targets for clinical implication in the future.

Keywords: Gastric cancer (GC); bioinformatics analysis; The Cancer Genome Atlas (TCGA); hub genes; survival analysis

Submitted Dec 31, 2020. Accepted for publication Apr 23, 2021.
doi: 10.21037/tcr-20-3540
View this article at: https://dx.doi.org/10.21037/tcr-20-3540

Introduction

Gastric cancer (GC) is the fifth most common cancer worldwide, with almost one million new cases every year. Over half of them occur in Eastern Asia, China particularly (1). Although the incidence of GC has declined and the treatment of GC has seen dramatic progress over the years, it remains the third leading cause of cancer-related death worldwide (2). So far, surgical resection is the only way to cure GC. However, 20% of patients lost their chance of surgery at the first clinic visit because of distant

metastasis (3). In those cases, chemotherapy becomes the main treatment they can rely on. At present, chemotherapy for GC is still dominated by conventional chemotherapeutic drugs such as platinum and fluorouracil. The molecularly targeted drugs remain scarce (4). Therefore, it is warranted to learn the underlying biological variation for developing more efficient therapeutic strategies.

Currently, carcinoembryonic antigen (CEA), cancer antigen 19-9 (CA19-9), and cancer antigen 72-4 (CA72-4) are the most common diagnostic markers for GC (5-7). However, the appliance of these biomarkers cannot meet the clinical requirements because of their low sensitivity (8,9). As for the treatment of GC, several tumor-specific proteins have been identified as therapeutic targets, including EGFR, HER-2, VEGFR, mTOR, PD-1, and PD-L1 (10-12). Still to this day, only three molecularly targeted drugs (trastuzumab, ramucirumab, and pembrolizumab) have been approved and marketed for GC treatment worldwide. Besides, a large number of molecules have been reported to be related to clinical outcomes of GC, including cancer-associated genes and non-coding RNAs (13-15). However, there are still no reliable prognosis biomarkers due to the heterogeneity of GC (16). Hence, it is meaningful to seek novel and reliable biomarkers for GC.

Recently, with the advancement of sequencing platforms and the establishment of public databases such as The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO), many bioinformatics analysis studies on GC have been published in these years. For example, using GEO data sets, Zheng *et al.* identified 7 hub genes that affect the prognosis of GC, including *COL4A1*, *COL1A2*, *COL6A3*, *VCAN*, *THBS2*, *TIMP1*, and *SERPINH1* (17). Similarly, another study identified 9 key genes (*COL1A1*, *CDKN3*, *COL1A2*, *COL3A1*, *NDC80*, *TPX2*, *TOP2A*, *TIMP1*, and *CEP55*) correlated with the pathogenesis of GC (18). Beyond these differentially expressed genes, non-coding RNAs, including lncRNAs (long non-coding RNAs) and miRNAs (microRNAs), were also reported to be related to the pathogenesis and prognosis of GC (19-22).

Although an increasing number of integrated bioinformatics analyses on GC have appeared recently, the results differed between these studies because of the different analysis methods. In our study, we focused on DEGs (differentially expressed genes) which affected the prognosis of GC and obtained eight hub genes. This study might provide potential prognostic biomarkers and treatment targets for GC.

We present the following article in accordance with the MDAR checklist (available at https://dx.doi.org/10.21037/tcr-20-3540).

Methods

Data procession

We downloaded the gene expression profiling datasets from TCGA database. The selection criteria were that samples contained complete RNA sequencing data and clinical information. According to the selection criteria, 407 samples were involved in this study, which was composed of 375 GC primary tumor specimens and 32 solid normal tissue specimens. The data were analyzed by the DeSeq2 package and edgeR package in the R language (23,24). Differentially expressed genes (DEGs) were defined by $|\log2FC| \geq 1$, and adjust P value <0.05. Overlapping DEGs between these two kinds of algorithms were retained for further analyses.

Survival analysis

A total of 367 GC patients from TCGA database were enrolled in survival analysis, including 234 (63.8%) men and 133 (36.2%) women. The median age was 67 (range, 35–90 years). Survival outcomes were calculated from the date of surgery to the date of last follow-up or the date of death. Follow-up data were downloaded from the TCGA database. Before statistical analysis, DeSeq2 package was used to transform the raw dataset to the normalized gene expression level. According to the median expression of a specific gene, the patients were divided into two groups. Log-rank test and Kaplan-Meier curve were conducted by the survival package in the R language to evaluate the prognostic value of all overlapping DEGs (25). DEGs which showed a significant correlation with overall survival (P<0.05) were referred to robust DEGs.

Functional enrichment analysis

First, we performed Gene Ontology (GO) analysis on the Database for Annotation, Visualization and Integrated Discovery (DAVID, version 6.8) to elucidate the biological function of these robust DEGs (26,27). BP (Biological process), MF (molecular function), and CC (cellular component) were performed, respectively, P<0.05 and count >2 were considered as statistically significant. Then, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were carried out utilizing the KEGG Orthology-Based

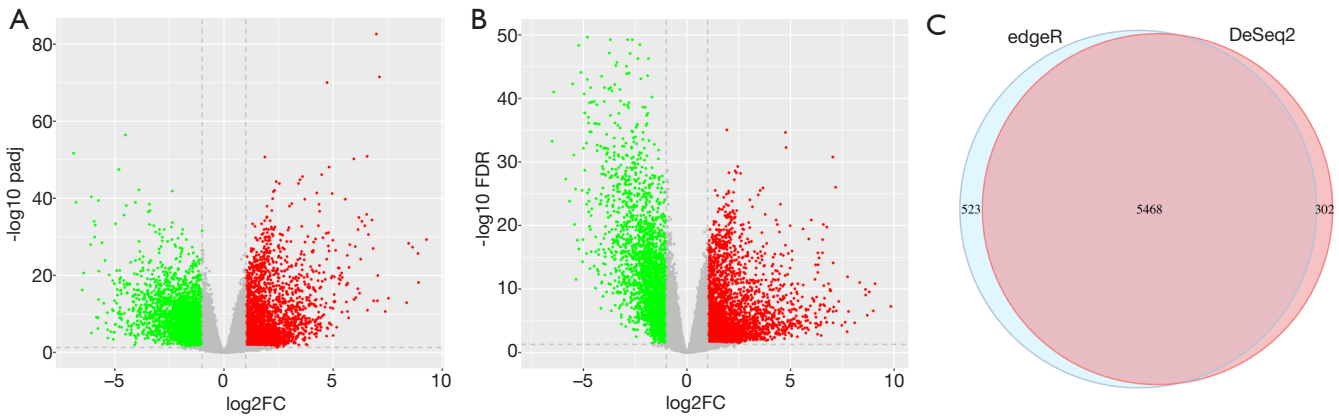

Figure 1 Identification of DEGs. (A) The volcano plots of the DEGs by DeSeq2. (B) The volcano plots of the DEGs by edgeR. (C) Venn diagrams of the DEGs between the DeSeq2 and edgeR. DEG, differentially expressed gene.

Annotation System (KOBAS, version 3.0) (28). Corrected P<0.05 was defined as statistically significant.

PPI network and module screening

For robust DEGs, the Search Tool for the Retrieval of Interacting Genes (STRING, version 10.5) online database was utilized to build protein-protein interaction (PPI) network. Then, we presented it by Cytoscape software (version 3.6.1). Also, the Molecular Complex Detection (MCODE) algorithm was applied to screen neighborhoods of densely connected proteins.

Identification of hub genes

CytoHubba in the Cytoscape software was utilized to screen hub genes among these robust DEGs (29). Both of MCC (maximal clique centrality) and DMNC (density of maximum neighborhood component) were computed, and the overlapping genes were filtered out as hub genes of GC.

Statistical analysis

Survival analysis was performed by the Survival package in the R software. Survival plots were showed by the Kaplan-Meier method, and the significance was calculated by the log-rank test. P<0.05 was defined as statistically significant.

Ethical statement

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). All information from TCGA is available and free for public, so the agreement of the medical ethics committee board is not necessary.

Results

Identification of DEGs

407 samples from TCGA database were enrolled in this study, including 375 GC primary tumor samples and 32 solid normal tissue samples. We screened DEGs using the DeSeq2 package and edgeR package, respectively. The cut-off criteria were corrected P<0.05 and |log2FC| >1. Overall, 5,770 DEGs were screened by DeSeq2 package, including 3,118 up-regulated and 2,652 down-regulated genes. A total of 5,991 DEGs were screened by edgeR package, including 3,479 up-regulated and 2,512 down-regulated genes. *Figure 1A,B* showed the volcano plots of DEGs for each method. We further intersected the results and obtained 5,468 overlapping DEGs, including 2,994 up-regulated and 2,474 down-regulated genes (*Figure 1C*).

Identification of robust DEGs associated with overall survival of GC

Log-rank test for all the overlapping DEGs was performed to explore the robust DEGs that were associated with the survival performance of GC patients. The cut-off criteria were P<0.05. Finally, we obtained 238 significantly robust DEGs, including 140 up-regulated and 98 down-regulated genes (Table S1).

GO and KEGG pathway analysis

GO analysis and KEGG analysis were performed to explore the potential biological function of the above robust DEGs. The GO analysis involved three categories: MF, BP, and CC. In MF category, the up-regulated DEGs were enriched in cytokine activity and protein heterodimerization activity. As for BP, the up-regulated DEGs were significantly enriched in nucleosome assembly, chemokine-mediated signaling pathway, skeletal system development, and immune response. For CC, the up-regulated DEGs were enriched in the nucleosome, centromeric region, extracellular space, and extracellular region (*Figure 2A*). About down-regulated DEGs, they were enriched in steroid hormone receptor activity, lipid binding, and kinase binding in the MF category. In BP category, the down-regulated DEGs were significantly enriched in steroid hormone-mediated signaling pathway, response to calcium ion, response to the drug, and monocyte chemotaxis (*Figure 2B*).

As for KEGG pathway analysis, the most significant pathways of the up-regulated DEGs were systemic lupus erythematosus, cytokine-cytokine receptor interaction, alcoholism, and jak-STAT signaling pathway (*Figure 2C*). For down-regulated DEGs, the most significantly enriched pathways were the intestinal immune network for IgA production, retinol metabolism, metabolism of xenobiotics by cytochrome P450, bile secretion, and chemical carcinogenesis (*Figure 2D*).

PPI network and modules analysis

To characterize the interaction between the robust DEGs, we constructed the PPI network using STRING database and presented it by Cytoscape. Overall, there were 82 edges and 140 nodes in this network, including 48 up-regulated and 34 down-regulated genes (*Figure 3A*). Subsequently, three key modules were extracted utilizing MCODE (*Figure 3B*,*C*,*D*). Module 1 was mainly concerned with the chemokine signaling pathway. Module 2 was mainly concerned with nucleosome assembly. Besides, module 3 involved ECM-receptor interaction.

Screening for hub genes and survival analysis of hub genes

We further screened the hub genes among robust DEGs with cytoHubba. MCC and DMNC were computed to identify hub genes. The top 10 genes were selected based on the two algorithms. Then, we intersected the results by Venn diagram and obtained 8 overlapping hub genes, including *CCR8*, *HIST1H3B*, *HIST1H2AH*, *HIST1H2AJ*, *NPY*, *HIST2H2BF*, *GNG7*, and *CCL25* (*Figure 4A*).

Besides, the survival package in R language was utilized to perform the Kaplan-Meier analysis. As shown in *Figure 4B-I*, patients with higher expression levels of *HIST1H3B*, *HIST1H2AH*, *HIST1H2AJ*, *HIST2H2BF*, and *CCL25* show worse OS, while those with lower expression levels of *CCR8*, *NPY*, and *GNG7* show worse OS.

Discussion

Owing to the high heterogeneity of GC, it still lacks effective therapeutic targets. Although a large number of studies had been performed to identify the driving genes of GC, there were still no reliable biomarkers and drug targets up to now. In recent years, public databases, such as TCGA and GEO, provided a platform to screen the molecular targets of the tumor. Bioinformatics analysis studies of GC have been increasingly reported (17-20). However, because of the limited sample size and impertinent methods, the plentiful DEGs might show no biological roles and clinical significance. Therefore, we introduced the survival analysis at the beginning of our study to obtain clinically significant DEGs. According to the gene expression profiles of TCGA and the survival analysis, 238 robust DEGs were filtered out, consisting of 140 up-regulated and 98 down-regulated genes. The up-regulated DEGs were mainly enriched in systemic lupus erythematosus, cytokine activity, and alcoholism, while down-regulated DEGs were mainly enriched in steroid hormone receptor activity, immune response, and metabolism. Through the construction of the PPI network, eight hub genes were finally screened out, including *CCR8*, *HIST1H3B*, *HIST1H2AH*, *HIST1H2AJ*, *NPY*, *HIST2H2BF*, *GNG7*, and *CCL25*.

CCR8 and CCL25 belong to the CXC subfamily of chemokines, which are important for cell migration. CCR8 (Chemokine (C-C motif) receptor 8) is preferentially expressed in the thymus, participating in regulating monocyte chemotaxis and thymic cell apoptosis. Villarreal *et al.* have shown that CCR8 is elevated in tumor-resident Tregs, and mAb therapy targeting CCR8 significantly inhibited tumor growth in the colorectal cancer model (30). Additionally, CCR8 was shown to enhance cell migration, invasion, and EMT in bladder cancer (31). In GC, a recent study elucidated that CCR8 was highly expressed and associated with the OS of GC patients (32). CCL25 [Chemokine (C-C motif) ligand 25], along with its specific

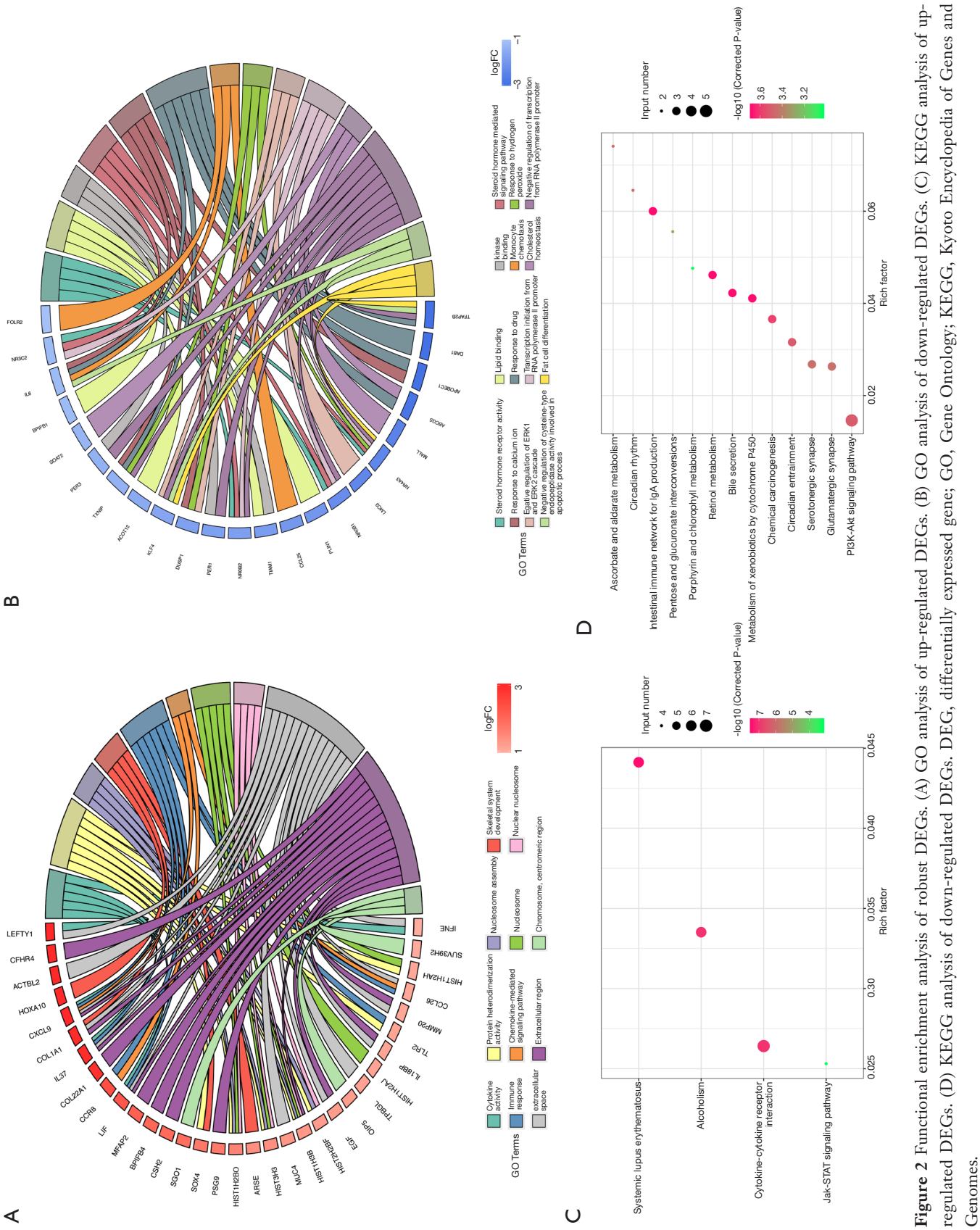

Figure 2 Functional enrichment analysis of robust DEGs. (A) GO analysis of up-regulated DEGs. (B) GO analysis of down-regulated DEGs. (C) KEGG analysis of up-regulated DEGs. (D) KEGG analysis of down-regulated DEGs. DEG, differentially expressed gene; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

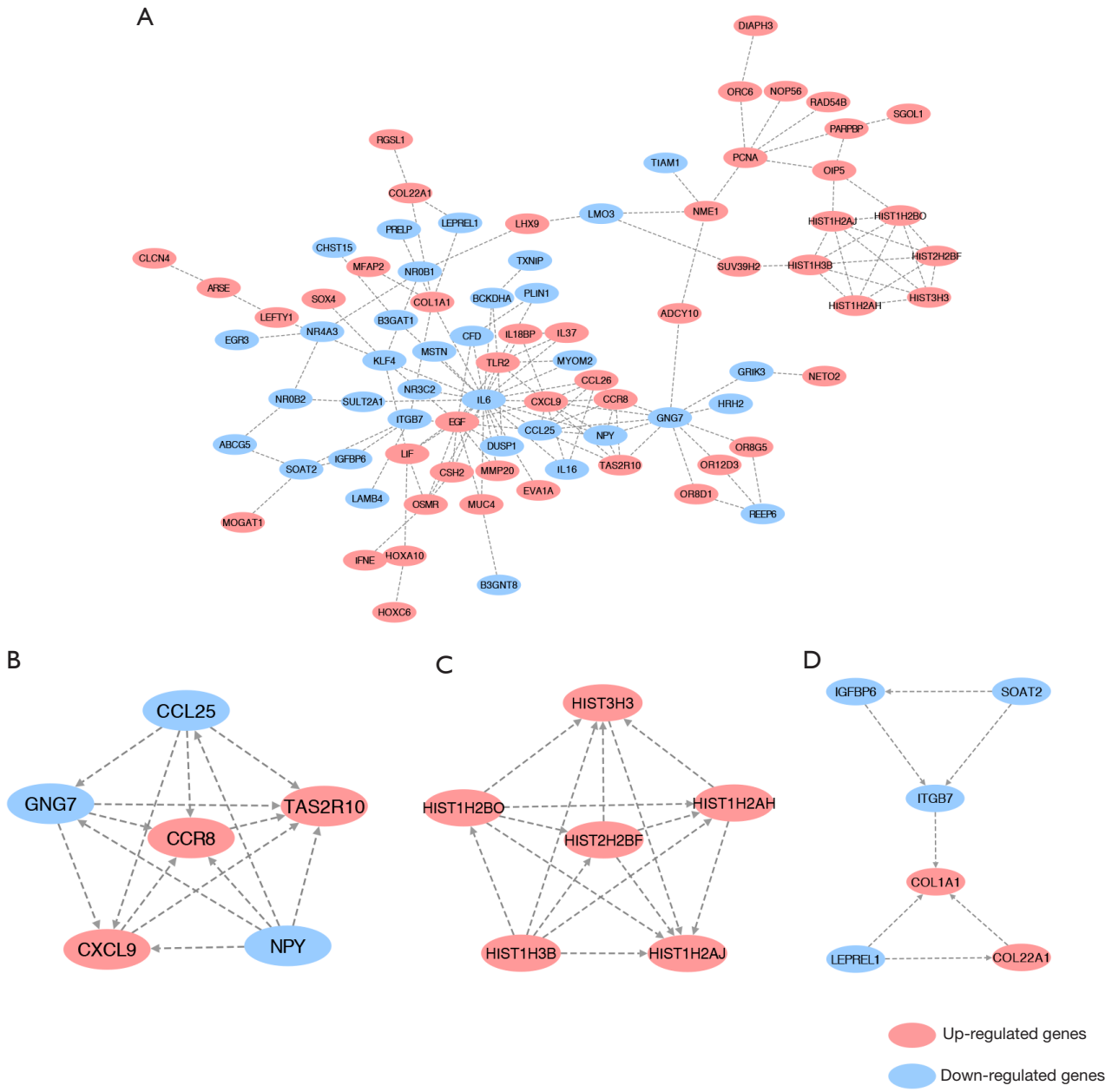

Figure 3 PPI network and hub clustering modules. (A) The PPI network of robust DEGs; (B) module 1; (C) module 2; (D) module 3. Red circles represent up-regulated genes and blue circles represent down-regulated genes. PPI, protein-protein interaction.

receptor CCR9 (chemokine receptor 9), has been well reported to regulate gut mucosal immunity (33). Moreover, several studies demonstrated that CCL25/CCR9 was implicated in proliferation, migration, and anti-apoptosis of cancer cells (34,35). Interestingly, Chen *et al.* revealed that CCL25/CCR9 promoted tumor growth in early-stage of CRC, while suppressed cell invasion and metastasis in the late-stage (36). Further studies are warranted for elucidating the role of CCL25 in GC.

HIST1H3B, HIST1H2AH, HIST1H2AJ, and HIST2H2BF belong to core histones. Two copies of each of the core histones compose histone octamer, which further form nucleosomes along with approximately 146 bp of DNA. It is well established that histone modifications

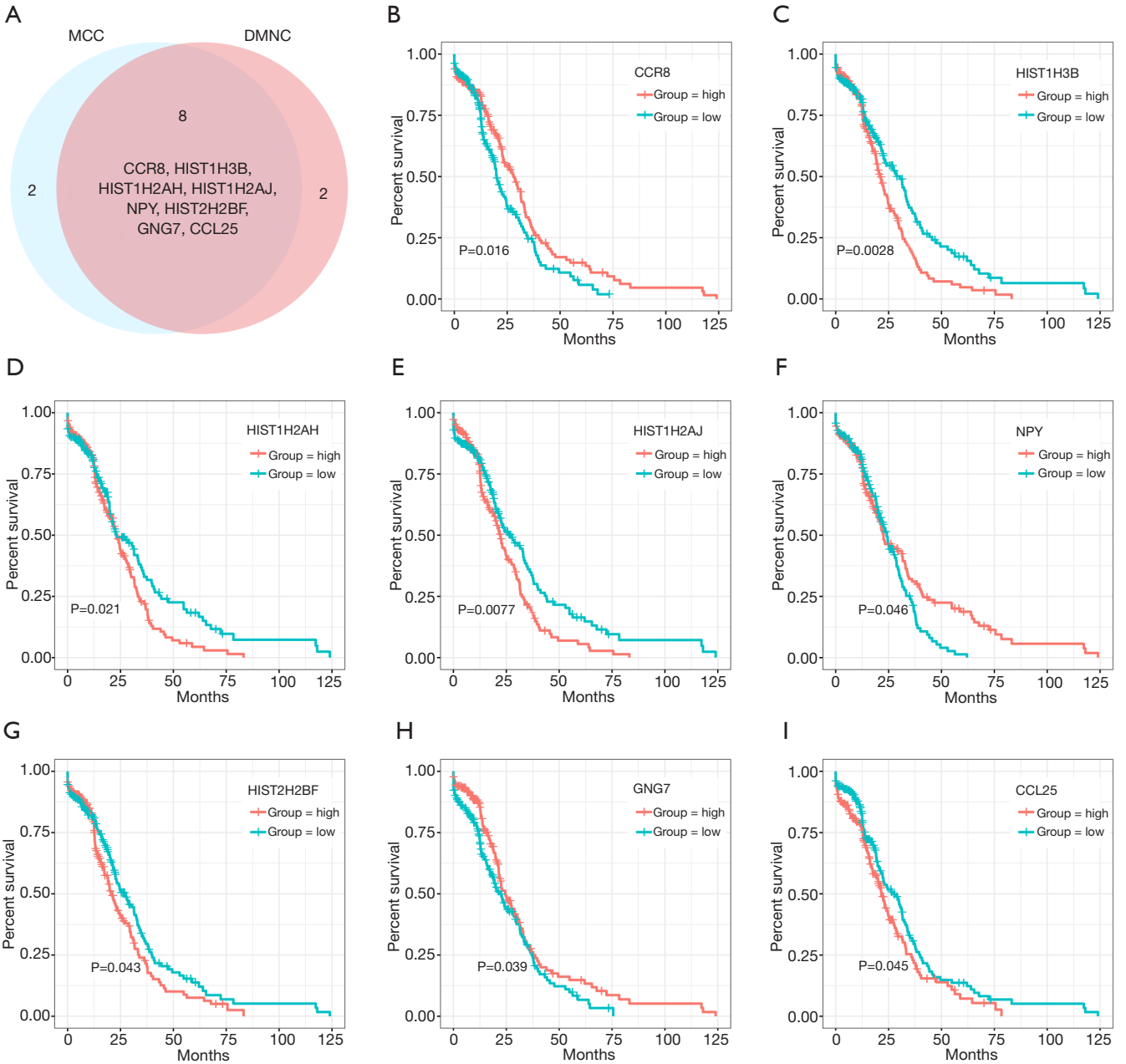

Figure 4 Prognostic value of eight hub genes. (A) Venn diagram of overlapped hub genes based on two methods. (B-I) Kaplan-Meier survival analysis of CCR8, HIST1H3B, HIST1H2AH, HIST1H2AJ, NPY, HIST2H2BF, GNG7, and CCL25.

played a significant role in tumor initiation and progression (37,38). However, the effects of histone expression levels alteration on tumors remain largely unknown (39,40). In the present study, we investigated that HIST1H3B, HIST1H2AH, HIST1H2AJ, and HIST2H2BF were up-regulated in GC and were related to the poor prognosis of GC patients.

Neuropeptide Y (NPY) is a 36 amino acids peptide, which is widely produced in nervous systems. Beyond regulating several physiological processes, such as cognitive function and cardiovascular regulation, NPY was also revealed to promote proliferation, migration, and

vascularization and in several tumors (41-45). Concerning GC, a recent bioinformatics analysis based on GEO database showed that NPY was elevated in GC and was correlated with poor survival (17). However, in our analysis, NYP was down-regulated in GC and high expression of NYP was related to better survival.

G protein γ subunit 7 (GNG7) is a member of the large G protein γ family. There have been compelling suggestions that GNG7 functions as a tumor suppressor gene in many tumors , including pancreatic cancer, esophageal cancer, and renal cell cancer (46-48). Consistent with the previous studies, we elucidated that GNG7 was down-regulated in GC and negatively correlated with overall survival.

In the present study, KEGG functional enrichment analysis indicated that SLE (systemic lupus erythematosus) might be associated with GC. A cohort study observed a significantly higher risk of GC in patients with SLE (49). The correlation was further confirmed by a recent meta-analysis (50). The underlying mechanisms why SLE patients were more likely to develop GC remain unclear. Various types of drugs used in SLE treatment, such as TNF-a inhibitors, corticosteroids, and nonsteroid anti-inflammatory drugs were speculated to be implicated in this process (51,52).

We acknowledged some potential limitations in this study. First, it was a bioinformatics study using TCGA datasets. The lack of clinical validity remains a drawback of our study. Second, the sample size was small and the majority of patients enrolled in this study were white, which might be ethnically homogeneous. Therefore, more studies are warranted for verifying these findings in real-world clinical practice.

Conclusions

Our present study disclosed eight hub genes and several molecular pathways of GC. These findings provided prognostic biomarkers and potential treatment targets for GC. Nevertheless, further molecular biological experiments are warranted to verify the function of the hub genes in GC.

Acknowledgments

We acknowledge and appreciate our colleagues for their valuable suggestions on this paper. We thank Dr. Song Liu and Dr. Xuefeng Xia for the linguistic editing and proofreading.

Funding: None.

Footnote

Reporting Checklist: The authors have completed the MDAR checklist. Available at https://dx.doi.org/10.21037/tcr-20-3540

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://dx.doi.org/10.21037/tcr-20-3540). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

References

1. Stewart B, Wild CP. World cancer report 2014. World 2014. Lyon: International Agency for Research on Cancer, 2014.
2. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394-424.
3. Nashimoto A, Akazawa K, Isobe Y, et al. Gastric cancer treated in 2002 in Japan: 2009 annual report of the JGCA nationwide registry. Gastric Cancer 2013;16:1-27.
4. Wagner AD, Syn NL, Moehler M, et al. Chemotherapy for advanced gastric cancer. Cochrane Database Syst Rev 2017;8:CD004064.
5. Peng W, Zhou Z, Zhong Y, et al. Plasma activity of Thioredoxin Reductase as a Novel Biomarker in Gastric Cancer. Sci Rep 2019;9:19084.

6. Chen C, Chen Q, Zhao Q, et al. Value of Combined Detection of Serum CEA, CA72-4, CA19-9, CA15-3 and CA12-5 in the Diagnosis of Gastric Cancer. Ann Clin Lab Sci 2017;47:260-3.
7. Yu J, Zhang S, Zhao B. Differences and correlation of serum CEA, CA19-9 and CA72-4 in gastric cancer. Mol Clin Oncol 2016;4:441-9.
8. Feng F, Tian Y, Xu G, et al. Diagnostic and prognostic value of CEA, CA19-9, AFP and CA125 for early gastric cancer. BMC Cancer 2017;17:737.
9. Zhou YC, Zhao HJ, Shen LZ. Preoperative serum CEA and CA19-9 in gastric cancer--a single tertiary hospital study of 1,075 cases. Asian Pac J Cancer Prev 2015;16:2685-91.
10. Patel TH, Cecchini M. Targeted Therapies in Advanced Gastric Cancer. Curr Treat Options Oncol 2020;21:70.
11. Sexton RE, Al Hallak MN, Diab M, et al. Gastric cancer: a comprehensive review of current and future treatment strategies. Cancer Metastasis Rev 2020;39:1179-203.
12. Yamashita K, Hosoda K, Niihara M, et al. History and emerging trends in chemotherapy for gastric cancer. Ann Gastroenterol Surg 2021. doi: 10.1002/ags3.12439. [Epub ahead of print].
13. Giuppi M, La Salvia A, Evangelista J, et al. The Role and Expression of Angiogenesis-Related miRNAs in Gastric Cancer. Biology (Basel) 2021;10:146.
14. Roviello G, Aprile G, D'Angelo A, et al. Human epidermal growth factor receptor 2 (HER2) in advanced gastric cancer: where do we stand? Gastric Cancer 2021;24:765-79.
15. Chen W, Yan H, Li X, et al. Circulating tumor DNA detection and its application status in gastric cancer: a narrative review. Transl Cancer Res 2021;10:529-36.
16. Halske C. Intratumoral heterogeneity of gastric cancer-impact on biomarker evaluation. Pathologe 2020;41:76-82.
17. Zheng S, Yang L, Dai Y, et al. Screening and Survival Analysis of Hub Genes in Gastric Cancer Based on Bioinformatics. J Comput Biol 2019;26:1316-25.
18. Liu X, Wu J, Zhang D, et al. Identification of Potential Key Genes Associated With the Pathogenesis and Prognosis of Gastric Cancer Based on Integrated Bioinformatics Analysis. Front Genet 2018;9:265.
19. Li CY, Liang GY, Yao WZ, et al. Identification and functional characterization of microRNAs reveal a potential role in gastric cancer progression. Clin Transl Oncol 2017;19:162-72.
20. Li CY, Liang GY, Yao WZ, et al. Integrated analysis of long non-coding RNA competing interactions reveals the potential role in progression of human gastric cancer. Int J Oncol 2016;48:1965-76.
21. Zhang X, Zhang W, Jiang Y, et al. Identification of functional lncRNAs in gastric cancer by integrative analysis of GEO and TCGA data. J Cell Biochem 2019;120:17898-911.
22. Zhang Z, Dong Y, Hua J, et al. A five-miRNA signature predicts survival in gastric cancer using bioinformatics analysis. Gene 2019;699:125-34.
23. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.
24. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010;26:139-40.
25. Therneau TM, Lumley T. Package 'survival'. R Top Doc 2015:128. Available online: https://github.com/therneau/survival
26. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;4:44.
27. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 2009;37:1-13.
28. Ai C, Kong L. CGPS: A machine learning-based approach integrating multiple gene set analysis tools for better prioritization of biologically relevant pathways. J Genet Genomics 2018;45:489-504.
29. Chin CH, Chen SH, Wu HH, et al. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol 2014;8:S11.
30. Villarreal DO, L'Huillier A, Armington S, et al. Targeting CCR8 Induces Protective Antitumor Immunity and Enhances Vaccine-Induced Responses in Colon Cancer. Cancer Res 2018;78:5340-8.
31. Liu X, Xu X, Deng W, et al. CCL18 enhances migration, invasion and EMT by binding CCR8 in bladder cancer cells. Mol Med Rep 2019;19:1678-86.
32. Yi J, Jiang SJ. Dysregulation of CCL18/CCR8 axis predicts poor prognosis in patients with gastric cancer. Eur J Inflam 2018;16. doi: 10.1177/2058739218796887
33. Fu H, Jangani M, Parmar A, et al. A Subset of CCL25-Induced Gut-Homing T Cells Affects Intestinal Immunity to Infection and Cancer. Front Immunol 2019;10:271.
34. Zhang Z, Sun T, Chen Y, et al. CCL25/CCR9 Signal Promotes Migration and Invasion in Hepatocellular and

Breast Cancer Cell Lines. DNA Cell Biol 2016;35:348-57.
35. Li B, Wang Z, Zhong Y, et al. CCR9-CCL25 interaction suppresses apoptosis of lung cancer cells by activating the PI3K/Akt pathway. Med Oncol 2015;32:66.
36. Chen HJ, Edwards R, Tucci S, et al. Chemokine 25-induced signaling suppresses colon cancer invasion and metastasis. J Clin Invest 2012;122:3184-96.
37. Yang WY, Gu JL, Zhen TM. Recent advances of histone modification in gastric cancer. J Cancer Res Ther 2014;10:240-5.
38. Chervona Y, Costa M. Histone modifications and cancer: biomarkers of prognosis? Am J Cancer Res 2012;2:589-97.
39. Martínez-Ricarte F, Mayor R, Martínez-Sáez E, et al. Molecular Diagnosis of Diffuse Gliomas through Sequencing of Cell-Free Circulating Tumor DNA from Cerebrospinal Fluid. Clin Cancer Res 2018;24:2812-9.
40. Ohshima K, Hatakeyama K, Nagashima T, et al. Integrated analysis of gene expression and copy number identified potential cancer driver genes with amplificationdependent overexpression in 1,454 solid tumors. Sci Rep 2017;7:641.
41. Diaz-delCastillo M, Christiansen SH, Appel CK, et al. Neuropeptide Y is Up-regulated and Induces Antinociception in Cancer-induced Bone Pain. Neuroscience 2018;384:111-9.
42. Li J, Tian Y, Wu A. Neuropeptide Y receptors: a promising target for cancer imaging and therapy. Regenerative Biomaterials 2015;2:215-9.
43. Medeiros PJ, Al-Khazraji BK, Novielli NM, et al. Neuropeptide Y stimulates proliferation and migration in the 4T1 breast cancer cell line. Int J Cancer 2012;131:276-86.
44. Medeiros PJ, Jackson DN. Neuropeptide Y Y5-receptor activation on breast cancer cells acts as a paracrine system that stimulates VEGF expression and secretion to promote angiogenesis. Peptides 2013;48:106-13.
45. Ruscica M, Dozio E, Motta M, et al. Relevance of the neuropeptide Y system in the biology of cancer progression. Curr Top Med Chem 2007;7:1682-91.
46. Long J, Liu Z, Wu X, et al. Gene expression profile analysis of pancreatic cancer based on microarray data. Mol Med Rep 2016;13:3913-9.
47. Ohta M, Mimori K, Fukuyoshi Y, et al. Clinical significance of the reduced expression of G protein gamma 7 (GNG7) in oesophageal cancer. Br J Cancer 2008;98:410.
48. Xu S, Zhang H, Liu T, et al. G Protein γ subunit 7 loss contributes to progression of clear cell renal cell carcinoma. J Cell Physiol 2019;234:20002-12.
49. Yu KH, Kuo CF, Huang LH, et al. Cancer risk in patients with inflammatory systemic autoimmune rheumatic diseases: a nationwide population-based dynamic cohort study in Taiwan. Medicine 2016;95:e3540.
50. Song M, Latorre G, Ivanovic-Zuvic D, et al. Autoimmune Diseases and Gastric Cancer Risk: A Systematic Review and Meta-Analysis. Cancer Res Treat 2019;51:841-50.
51. Hemminki K, Liu X, Ji J, et al. Autoimmune disease and subsequent digestive tract cancer by histology. Ann Oncol 2012;23:927-33.
52. Bernatsky S, Ramsey-Goldman R, Labrecque J, et al. Cancer risk in systemic lupus: an updated international multi-centre cohort study. J Autoimmun 2013;42:130-5.

Cite this article as: Sun F, Zhang C, Ai S, Liu Z, Lu X. Identification of hub genes in gastric cancer by integrated bioinformatics analysis. Transl Cancer Res 2021;10(6):2831-2840. doi: 10.21037/tcr-20-3540

Table S1 Identification of robust DEGs by log-rank test

Genes	P value
Up-regulated genes	
LEFTY1	0.0005
TAS2R10	0.0007
SAMSN1-AS1	0.0008
LINC02657	0.0015
COL22A1	0.0019
SNORD111	0.0023
ARSE	0.0026
XPO5	0.0027
HIST1H3B	0.0028
LINC01123	0.0029
MMP20	0.0034
LOC102723385	0.0035
ACTBL2	0.0039
DDN	0.0041
STEAP2-AS1	0.0042
LOC102725072	0.0043
RBAKDN	0.0044
PSG9	0.0054
LOC286059	0.0069
PTTG3P	0.0071
RAD54B	0.0073
ORC6	0.0073
HIST1H2AJ	0.0077
PLXNA1	0.0078
NOP56	0.0079
LOC101929470	0.0080
SLC22A14	0.0083
TLR2	0.0088
TMEM212	0.0098
RGSL1	0.0103
LACTB2	0.0105
LINC02022	0.0107
MIR663AHG	0.0109
MIR4753	0.0125
KRT40	0.0128
OVOL3	0.0128
PRR5L	0.0131
EVA1A	0.0137
LOC105373764	0.0146
CCNYL2	0.0148
CSH2	0.0152
TPBGL	0.0156
OR8G5	0.0157
CTAG2	0.0158
MINAR2	0.0159
ACRV1	0.0159
CCR8	0.0162
CLDN9	0.0165
PCNA	0.0173
TMEM145	0.0173
MGC32805	0.0174
PNMA2	0.0175
UNC13A	0.0177
LINC01675	0.0177
PTPRU	0.0180
LINC02052	0.0181
GDPD5	0.0194
TTPA	0.0194
CCL26	0.0196
IL37	0.0199
FOXD2-AS1	0.0203
LINC02506	0.0204
SOX4	0.0210
VAC14-AS1	0.0211
HIST1H2AH	0.0211
BPIFB4	0.0211
MIR181A2HG	0.0211
LINC01205	0.0212
HSPE1-MOB4	0.0219
KIAA2012	0.0226
MIR647	0.0228
PRR5-ARHGAP8	0.0236
HIST1H2BO	0.0241
DIAPH3	0.0245
SYCP2L	0.0245
FLJ33534	0.0246
MUC4	0.0247
SGO1	0.0252
TBC1D3P5	0.0256
C6orf99	0.0264
CNGB3	0.0266
CPLX4	0.0268
PARPBP	0.0268
SSX2	0.0270
HOXA10	0.0279
ZAN	0.0280
ZNF724	0.0303
OR12D3	0.0303
LINC00456	0.0305
ADCY10	0.0306
LHX9	0.0309
STAM-AS1	0.0316
TXNDC12-AS1	0.0320
LRRC37A6P	0.0322
CLCN4	0.0323
SCN2A	0.0327
IL18BP	0.0331
ATP11AUN	0.0331
MIR5194	0.0336
LINC00939	0.0337
HOXC6	0.0340
SAGE1	0.0343
FIRRE	0.0347
CXCL9	0.0349
EGF	0.0353
SUV39H2	0.0366
LINC01419	0.0368
LINC01299	0.0385
SCARNA13	0.0388
CCDC34	0.0394
LOC101928708	0.0396
HAR1B	0.0405
LOC100131496	0.0408
HIST3H3	0.0409
SMIM23	0.0410
ALKAL1	0.0411
MFAP2	0.0413
NME1	0.0415
OR8D1	0.0418
OSMR	0.0418
LINC02119	0.0421
HIST2H2BF	0.0428
LINC01285	0.0432
SLC38A4	0.0433
LIF-AS1	0.0439
COL1A1	0.0441
EMBP1	0.0450
MOGAT1	0.0454
CYMP	0.0455
CFHR4	0.0467
IFNE	0.0467
SNORD92	0.0472
OIP5	0.0474
HOXC-AS1	0.0477
LIF	0.0478
VCAN-AS1	0.0488
LACTB2-AS1	0.0493
NETO2	0.0495
LOC105375679	0.0497
STPG3	0.0499
Down-regulated genes	
SNORD9	0.0001
SULT2A1	0.0008
LOC105369748	0.0011
DAB1	0.0028
MSTN	0.0031
KLF4	0.0034
CC2D2A	0.0049
APOBEC1	0.0050
FOLR2	0.0058
NR3C2	0.0058
BCKDHA	0.0062
P3H2	0.0063
TEX26	0.0073
C17orf78	0.0081
MAB21L3	0.0092
MALL	0.0095
ATP5PF	0.0097
CTSF	0.0110
LOC101928881	0.0110
SHISAL2B	0.0111
LINC01354	0.0115
AQP11	0.0118
SMDT1	0.0123
PADI1	0.0127
IL6	0.0133
LOH12CR2	0.0140
PLIN1	0.0151
PLA2G4D	0.0151
B3GAT1	0.0154
PER1	0.0155
PXMP2	0.0167
SKIDA1	0.0168
DUSP1	0.0170
KCND3	0.0170
SMLR1	0.0175
UGT1A4	0.0176
EPB41L3	0.0181
FAM189A2	0.0186
CSTB	0.0191
B3GNT8	0.0206
EGR3	0.0208
RDH5	0.0213
IGFBP6	0.0213
NR0B1	0.0215
LMO3	0.0249
CHST15	0.0260
PHF21B	0.0260
ANKRD63	0.0270
SOAT2	0.0270
PPP2R2C	0.0273
C5orf67	0.0275
LAMB4	0.0276
ARHGEF37	0.0278
KCNJ13	0.0280
MIR1-1HG	0.0284
PER3	0.0288
FNDC4	0.0291
NR0B2	0.0292
BPIFB1	0.0295
GABRA5	0.0296
HRH2	0.0300
CFD	0.0301
SCARA5	0.0306
ACSM6	0.0314
TXNIP	0.0315
CAPN13	0.0325
KLF15	0.0326
TOM1L2	0.0377
ATRNL1	0.0383
GNG7	0.0386
PRAMENP	0.0388
ITGB7	0.0394
TFAP2B	0.0400
ABCG5	0.0400
TIAM1	0.0410
ATP13A5	0.0410
ACOT12	0.0413
MIR215	0.0425
EMILIN3	0.0433
LRRC3B	0.0434
FADS6	0.0439
TP53I3	0.0444
IRX4	0.0447
CCL25	0.0447
DHRS1	0.0448
STOX2	0.0450
ANKRD65	0.0450
PRELP	0.0454
NPY	0.0455
UGT1A10	0.0458
MYOM2	0.0465
LINC01797	0.0469
TMEM236	0.0471
REEP6	0.0474
IL16	0.0475
GRIK3	0.0494
LINC00882	0.0494
NR4A3	0.0496

http://dx.doi.org/10.21037/tcr-20-3540