An integrated analysis of DNA promoter methylation, microRNA regulation, and gene expression in gastric adenocarcinoma

Hongyun Huang[1,2], Lang Xie[1,2], Xiaoxuan Feng[3], Zheng Zheng[1], Juntao Ouyang[3], Yan Li[4], Jinlong Yu[1,2]^

[1]Department of General Surgery of Zhujiang Hospital, Southern Medical University, Guangzhou, China; [2]The Second School of Clinical Medicine, Southern Medical University, Guangzhou, China; [3]The First School of Clinical Medicine, Southern Medical University, Guangzhou, China; [4]Department of Immunology, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou, China

Contributions: (I) Conception and design: J Yu, Y Li; (II) Administrative support: J Yu; (III) Provision of study materials or patients: J Yu, L Xie; (IV) Collection and assembly of data: H Huang, L Xie; (V) Data analysis and interpretation: X Feng, Z Zheng, J Ouyang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Jinlong Yu. Department of General Surgery of Zhujiang Hospital, Southern Medical University, Guangzhou 510280, China; The Second School of Clinical Medicine, Southern Medical University, Guangzhou 510515, China. Email: yujinlong640506@163.com; Yan Li. Department of Immunology, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou 510080, China. Email: liyan336@mail2.sysu.edu.cn.

Background: Gastric adenocarcinoma (GAC), a common type of gastric cancer, poses a significant public health threat worldwide. This study aimed to determine the transcriptional regulatory mechanisms of GAC.

Methods: HTSeq-FPKM raw data were obtained from The Cancer Genome Atlas Stomach Adenocarcinoma data collection. Subsequently, the limma package in R was used to identify differentially expressed genes (DEGs). Differentially methylated genes (DMGs), DEGs, and differentially expressed microRNAs (miRNAs) in normal, and tumor tissues of the same patients were screened and compared using R software tools. A functional enrichment analysis was performed using Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) for various DEGs, DMGs, promoter methylation, and miRNAs. DEG-specific methylation and transcription factors were analyzed using ENCODE ChIP-seq.

Results: DEGs were centrally modified by the histone trimethylation of lysine 27 on histone H3 (H3K27me3). Upstream transcription factors of DEGs were enriched in different ChIP-seq clusters, such as Forkhead Box M1, E2F Transcription Factor 4, and suppressor of zest 12. Integrated regulatory networks of DEGs, promoter methylation, and miRNAs were constructed. Two miRNAs (hsa-mir-1 and hsa-mir-133a) and four DEGs (A disintegrin and metalloproteinase domain 12, transcription factor AP-2 alpha, solute carrier family 5 member 7, and cadherin 19) separately played important roles in the integrated regulatory network. Therefore, these DEGs, DMGs, promoter methylation, and miRNAs may play an important role in GAC pathogenesis.

Conclusions: In summary, the present study results provide insights into the oncogenesis and progression of GAC, thus accelerating the development of novel targeted GAC therapies.

Keywords: Methylation; microRNA (miRNA); gene expression; gastric adenocarcinoma (GAC)

^ ORCID: 0000-0002-4581-1726.

Page 2 of 18

Huang et al. Integrated bioinformatic analysis of GAC

Introduction

Gastric cancer (GC) is the fifth most common malignancy worldwide, with the third-highest incidence and mortality rates of all types of cancer (1). In 2018, more than one million new GC-related cases were diagnosed with an estimated 783,000 global deaths (1). Moreover, incidence rates are markedly elevated in East Asian countries, particularly in Mongolia, Japan, and the Republic of Korea, whereas those in North America, Northern Europe, and African regions are generally low (1). Surgery is considered a potentially curative therapy for early GC; however, the prognosis for patients with advanced adenocarcinoma remains poor, despite improvements in chemotherapy (2).

Epigenetic alteration is an important event during carcinogenesis; it plays a critical role in the transcriptional silencing of tumor suppressor genes (3). Detecting the regulators of gene expression associated with cancer progression is one of the fundamental challenges in cancer research. DNA methylation occurs almost exclusively in the context of CpG islands, defined as regions of more than 200 bases with a G+C content of at least 50% and a ratio of statistically expected CpG frequencies of at least 0.6 (4). Moreover, the methylation of these CpG island shores referred to as regions of lower CpG density that lie close (~2 kb) to CpG islands, is likely to occur in promoter regions near the transcription start site. Additionally, microRNAs (miRNAs) are a class of endogenous small RNAs containing 20–25 nucleotides that regulate mRNA translation and degradation by perfect, or near-perfect, complementarity to the 3' untranslated regions (UTRs) of the target mRNA (5). miRNAs act as regulators of post-transcriptional levels and transcriptome changes and are also major components of the epigenome. Active regulatory functions of nuclear miRNAs have been described in gene promoter regions and exert further effects via epigenetic pathways (6). Hypermethylation and hypomethylation of miRNAs frequently occur in human cancer, representing a new level of complexity in gene regulation. Furthermore, in this research, we found that miRNAs and DNA methylation were in mutual regulation in GAC. Specifically, miRNAs regulate DNA methylation by targeting DNA methyltransferases or methylation-related proteins (7). Nevertheless, the mechanism of the synergistic interactions between DNA methylation and miRNAs as epigenetic regulators of transcriptomic changes, as well as their association with clinical outcomes, have remained largely unexplored in cancer research (8).

Previous studies have identified various genes related to gastric adenocarcinoma (GAC). For instance, Pilehchian Langroudi *et al.* (9) analyzed *FAT4* expression and promoter methylation in GC and reported that *FAT4* is a tumor suppressor gene in cell adhesion. This indicated that a reduced expression of *FAT4* and increased methylation of its promoter might be a key mechanism associated with tumor growth (9). Furthermore, targeted therapies, such as trastuzumab, an antibody against HER2 (also known as ERBB2), and the VEGFR-2 antibody ramucirumab, have recently been initiated (10).

GAC is a major subtype of GC, but the pathogenesis of GAC has yet to be characterized. To reduce the significant morbidity and mortality associated with GAC, it is critical to identify GAC-associated genes and mechanisms. Tumorigenesis is regulated by epigenetic phenomena, including nucleosome remodeling by histone modifications, DNA methylation, and miRNA-mediated targeting of genes associated with various biochemical pathways. Epigenetic abnormalities are heritable as part of gene transcription and collaborate with genetic changes to cause cancer evolution. General hypomethylation and focal hypermethylation of the noncoding region, especially the promoter-related CpG island associated with gene silencing, are the characteristics of cancer cells. This process may be related to the acquisition of histone inhibitory markers (11). Furthermore, mutated genes controlling epigenetic factors have further strengthened the importance of epigenetics in cancer (12). It is worth mentioning that mutations in genes caused by epigenetic changes can be frequently observed (13). Most cancers harbor frequent mutations in genes that encode components of the epigenetic machinery, resulting in abnormalities in the epigenome, which can further affect gene expression patterns and genomic stability (14,15). Hence, conducting an integrated analysis of differentially expressed genes (DEGs) and the expression of regulatory factors, such as methylation, mRNA splicing, transcription factors (TFs), and miRNAs, is an effective strategy for investigating GAC pathogenesis. The present study used microarray technology and bioinformatics analyses to investigate the regulatory network of GAC-associated miRNAs and their target genes in an effort to further explore the function of promoter methylation and miRNAs in the pathogenesis of GAC. We hope these findings will serve as a foundation for future studies. We present the following article in accordance with the MDAR reporting checklist (available at https://dx.doi.org/10.21037/atm-21-3211).

Methods

Datasets

We downloaded the GAC dataset from The Cancer Genome Atlas (TCGA) data portal (https://gdc.cancer.gov/), which included the DNA methylation dataset, the mRNA expression profile dataset, and the miRNA expression profile dataset. In total, we obtained 397 DNA tissue samples (395 GAC, 2 normal), 407 mRNA tissue samples (375 GAC, 32 normal), and 491 miRNA tissue samples (446 GAC, 45 normal). The three datasets were then combined using R software (3.5.2 version, http://cran.r-project.org/src/base/R-3/R-3.5.2.tar.gz). The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Differential expression analysis of mRNA and miRNA

We obtained the differential expressions of mRNAs and miRNAs by using the *limma* package (16) in R [logFC >2, false discovery rate (FDR) <0.01]. The target genes of miRNAs were acquired from mirWalk (http://zmf.umm.uniheidelberg.de/apps/zmf/mirwalk/micrornapredictedtarget.html).

Screening of DEGs and differentially methylated genes (DMGs)

We used the *limma* package in R (16,17) to compare DEGs in GAC and normal tissue samples after adjusting the FDR <0.01 and |log2fold-change (FC)| ≥1 threshold. Differentially methylated regions (DMRs) were discovered using the *minfi* R software (17), with a q value <0.05. By applying *gplots* in the *limma* package (https://cran.r-project.org/web/packages/gplots/), we produced volcano plots to analyze the DEGs with adjusted P values <0.01 and |log2FC| ≥2.

Functional analysis of DEGs

Gene Ontology (GO), consisting of three categories (biological process, BP; cellular component, CC; and molecular function, MF), was used to predict the potential functions of gene products. The GO functional enrichment and annotation of DEGs were computed using Enrichr (18) (https://amp.pharm.mssm.edu/Enrichr/). We used the Kyoto Encyclopedia of Genes and Genomes (KEGG) database, which connects genomic information with functional coding information to systematically analyze gene functions (19). We also used Cytoscape (http://apps.cytoscape.org/apps/keggscape) to reproduce equivalent detailed hand-drawn pathway diagrams (20). The GO and KEGG enrichment analyses were performed for the DEGs, DMGs, and the target genes of miRNAs, as well as promoter methylation. The *P* values of the enriched terms were corrected by the Holm Bonferroni method (21). An adjusted P value <0.05 was applied as the cut-off criterion. ChIP-X enrichment analysis (ChEA) (22) and ENCODE ChIP-seq (23) were used to screen the enriched TFs located upstream of the DEGs. Lists of the mammalian gene symbols, for which the program computes over-represented TFs from the ChIP-X database, were then imported using ChEA.

Correlation analysis of DEGs, differential miRNAs, and promoter methylation

We used a network graph to determine the correlations between GAC DEGs, differential miRNAs, and promoter methylation. Funrich (24) (http://www.mybiosoftware.com/funrich-functional-enrichment-analysis-tool.html), a stand-alone software tool, was used to produce Venn diagrams to show the functional enrichment and interaction network analysis of the identified genes (25). The data visualization tool MEXPRESS (https://mexpress.be/) was used to analyze further the precise methylation locus and expression (26).

Construction of the risk score-based prognostic model

The GAC tumor samples from the TCGA-STAD cohort (n=350) were enrolled as the training cohort to construct the prognostic model, and the trimmed mean of M-values (TMM) normalized expression matrix of 17 interested genes were obtained from *Figure 1*. Ten-fold cross-validation were applied to the multivariate cox regression model to train the minimum lambda value, with 17 genes as the parameters. The minimum lambda value ($\lambda_{min} \approx 0.046$) was applied to the original multivariate cox model, returning a subset of variables containing 5 variables (*ADAM12*, *CDH19*, *RIMS1*, *SHOX2*, and *IGF2BP1*). Furthermore, the risk score model was built on the Lasso-Cox regression result, with 5 genes selected based on the subset with λ_{min} applied. The coefficients of the risk score model were weighted by the coefficients of the lasso multivariate Cox regression model. The formula for calculating the risk score is: risk score = $0.058316123 \times EXP_{ADAM12} + 0.014146433 \times EXP_{CDH19} + 0.015680240 \times EXP_{RIMS1} + 0.032075667 \times EXP_{SHOX2} +$
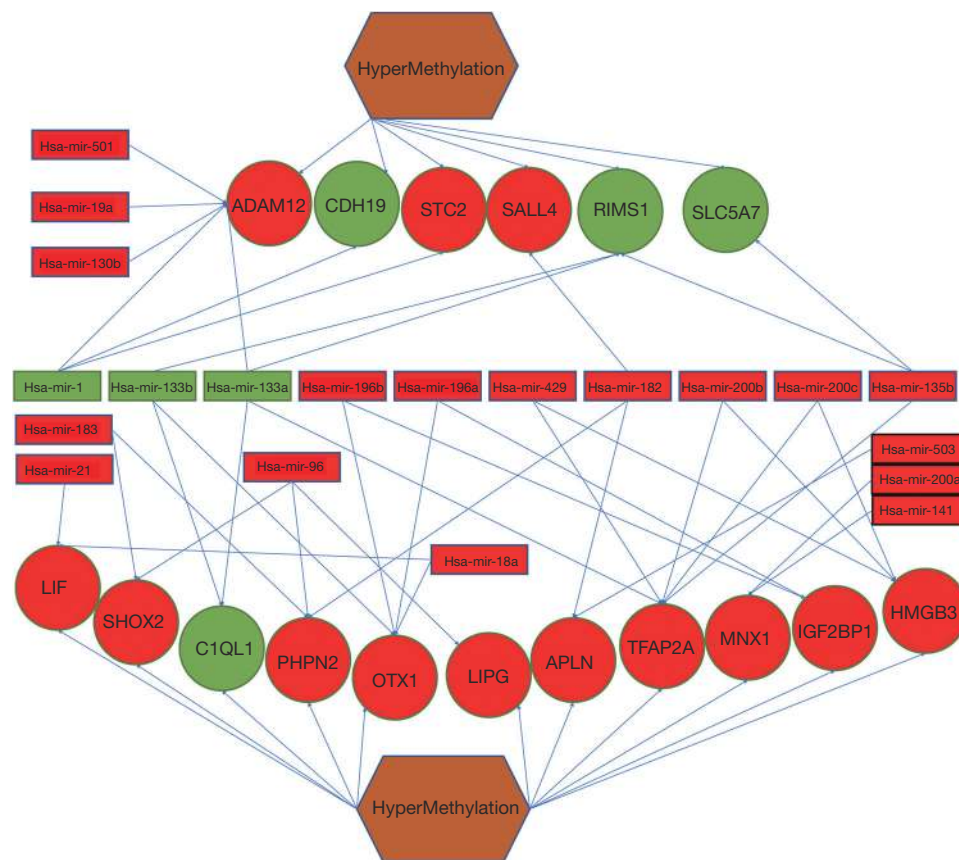
Page 4 of 18

Huang et al. Integrated bioinformatic analysis of GAC

Figure 1 Interrelated networks in the context of DEGs, differentially expressed miRNAs, and gene expression. Methylation and microRNA regulatory network of 31 DEGs. The downregulated and upregulated genes are shown in green and red circles, respectively. The upregulated and downregulated miRNAs are displayed in red and green rectangles, respectively. DEGs, differentially expressed genes

$0.007903263 \times \text{EXP}_{IGF2BP1}$.

Validation of the prognostic model in the training and validation cohort

In the training cohort, time-dependent receiver operating characteristic (ROC) analysis was adopted to evaluate the performance of the risk score-based prognostic signature to predict 1-, 3- and 5-year survival. Kaplan-Meier (KM) survival curves combined with a log-rank test were used to test the differences in prognosis between the high- and low-risk groups. Gastric tumor samples from the GSE15459 dataset (n=192) were extracted as the validation cohort. Similarly, time-dependent ROC analysis and KM survival analysis were also performed on the validation cohort.

Construction and assessment of a predictive nomogram

Univariate and multivariate analyses were successively performed on the 5 genes-based risk score prognostic model alone and combined with other clinical characteristics (including TNM stage, age, and sex) of patients with GAC to test their predictive efficacy. In the training cohort, univariate Cox regression analysis was used to identify clinical characteristics that were significantly associated with overall survival (OS). Characteristics with P<0.05 in the univariate analysis were included in the multivariate Cox regression model. Subsequently, nomogram for 1-, 3-, and 5-year OS were constructed using the *rms* R package (27) as a visualizing aid to obtain predicted values manually from the multivariate analysis. The performance of the

nomogram was evaluated using Harrell's concordance index (c-index) and calibration curves.

Overall survival analysis

To identify which interactions between miRNAs and methylation are significantly associated with GAC prognosis, we used the Kaplan-Meier (KM) overall survival analysis from UALCAN (http://ualcan.path.uab.edu/index.html) (28). This program aids in investigating gene expression patterns related to clinical parameters. A set of KM plots (29) from both the univariate analysis (considering only gene expression) and multivariate analysis (considering clinical parameters along with gene expression) were provided from the output page of the survival analysis after using the "scan by gene" option in UALCAN. A KM survival plot was generated for every gene in each TCGA cancer type, using the "*survival*" (https://cran.r-project.org/web/packages/survival/index.html) and "*survminer*" (http://mirrors.dotsrc.org/pub/cran/web/packages/survminer/index.html) packages. The survival curves of samples with high gene expressions and low/medium gene expressions were compared using the log-rank test (30). Based on the KM plots, we then ran survival analyses for the two extreme subgroups; the first subgroup comprised the lowest quantiles of miRNAs and methylation profiles, and the second subgroup comprised the highest quantiles of both profiles. Observation of the KM plots allowed us to assess the relationship between gene expression and overall survival in GAC patients.

Statistical analysis

Statistical analyses were performed using R software (https://mirrors.tuna.tsinghua.edu.cn/CRAN/). Kaplan-Meier survival analysis was conducted together with the log-rank test to compare the prognostic risk of GAC patients with different average expression levels. Unless otherwise noted, a P value <0.05 was considered statistical significance.

Results

Identification of DEGs, target genes of miRNAs, and promotor methylation

The overall study design is outlined in a flow chart in *Figure 2*. A total of 397 DNA samples, 407 mRNA samples,

and 491 miRNA samples from TCGA were analyzed using the *minifi* package in R with a q value <0.05 to identify DMRs; in turn, 8,615 DMGs were identified using *annover*. Finally, 2,059 differential promoter methylation genes were identified. Using the *limma* package in R, 1,160 downregulated target genes and 3,480 upregulated target genes of miRNAs were identified. Moreover, 507 downregulated genes and 919 upregulated genes were identified. The DEGs were selected by volcano plot filtering (P value <0.05 and |log2FC| ≥2) (*Figure 3*).

GO enrichment analysis of methylation-related genes

To further investigate the function of methylation-related genes in GAC, we used the GO enrichment analysis in *Enrichr*, and the KEGG pathway enrichment analysis in Cytoscape. Four specific variables were used in the analyses: DEGs, DMGs, the target genes of miRNAs, and promoter methylation. The top seven GO pathways are shown in *Figure 4A-4D*. Specifically, DEGs were enriched in the extracellular matrix organization (ECM) (BP), the spindle of cell components (CC), and endopeptidase activity (MF) (*Figure 4A*). For DMGs, the predominantly enriched genes included the development of the nervous system (BP), exon (CC), and transcriptional activator and factor activity (MF) (*Figure 4B*). The target genes of miRNAs were enriched in the regulation of transcription from RNA polymerase II promoter (BP), nuclear body (CC), and protein kinase activity (MF) (*Figure 4C*). Promoter methylation genes showed enrichment in the regulation of transcription and DNA-template (BP), an integral component of the plasma membrane (CC), and transcription regulatory region DNA binding (MF) (*Figure 4D*).

KEGG pathway enrichment analysis of methylation-related genes

Based on the KEGG pathway enrichment analysis, DEGs were significantly enriched in the IL-17 signaling pathway, cytokine-cytokine receptor interaction, and cell cycle (Figure S1), while other pathways exhibited little correlation with GAC. DMGs were associated with the neuroactive ligand-receptor interaction, cAMP signaling pathway, RAP1 signaling pathway, exon guidance, cell adhesion molecules, ribosome biogenesis in eukaryotes, and ribosomes (Figure S2). The target genes of miRNAs were correlated with endocytosis, signaling pathways regulating pluripotency of stem cells, proteoglycans in cancer, the
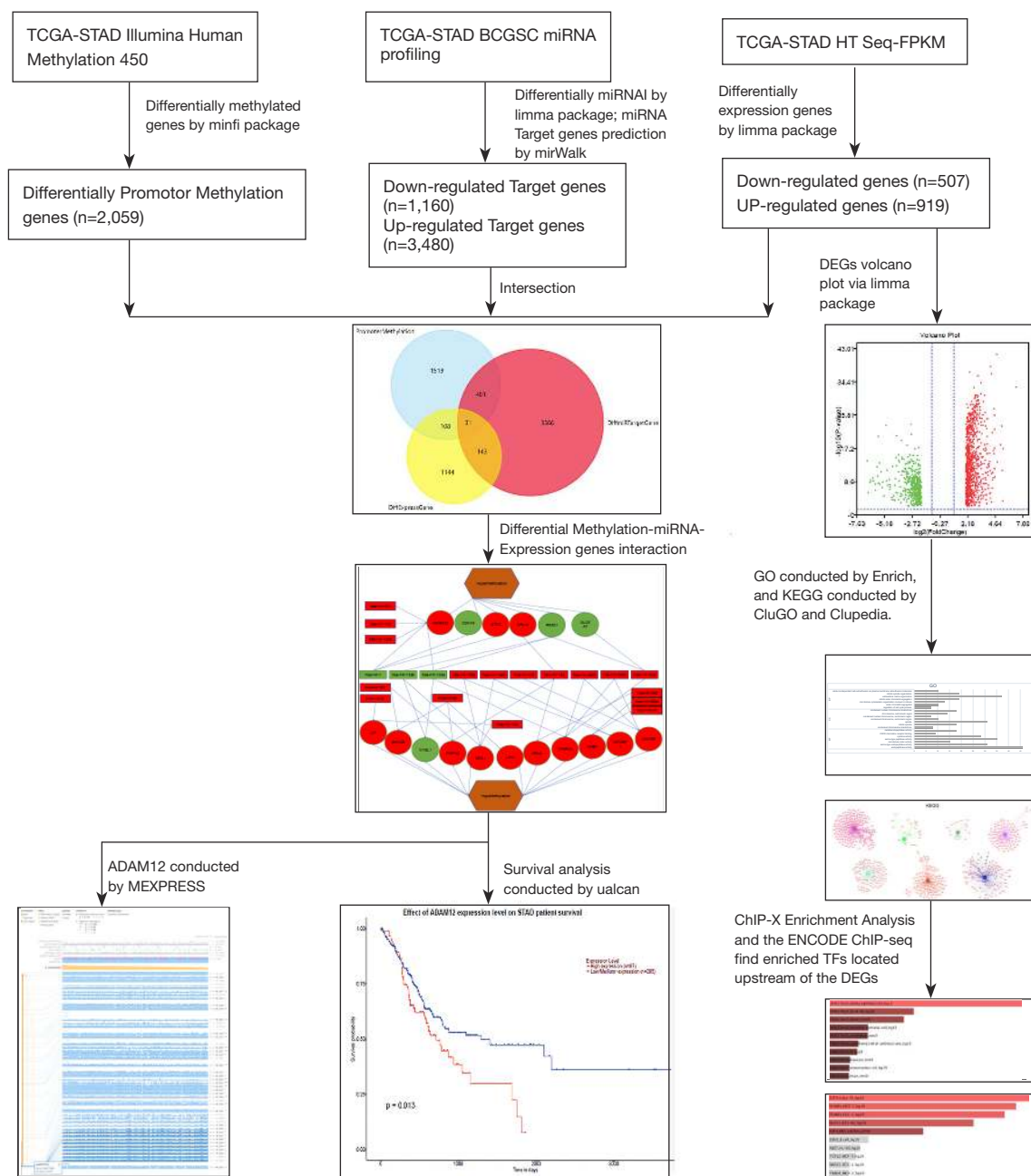
Page 6 of 18

Huang et al. Integrated bioinformatic analysis of GAC

Figure 2 Flow chart representing the detailed analysis process applied in this work. TCGA-STAD, The Cancer Genome Atlas Stomach Adenocarcinoma; DEGs, differentially expressed genes; KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, Gene Ontology.

RAP1 signaling pathway, olfactory transduction, axon guidance, the mitogen-activated protein kinase (MAPK) signaling pathway, pathways in cancer, and RAS signaling pathways (Figure S3). Promotor methylation genes were strongly related to neuroactive ligand-receptor interactions and cell adhesion molecules (Figure S4).

Transcriptional analysis of DEGs

ChEA2 analyses obtained from the ENCODE database (*Figure 5A*) indicated that the screened DEGs were modified and regulated by multi-cancer cell line histones, including trimethylation of lysine 27 on histone H3 (H3K27me3). The binding patterns of the upstream TFs were not as
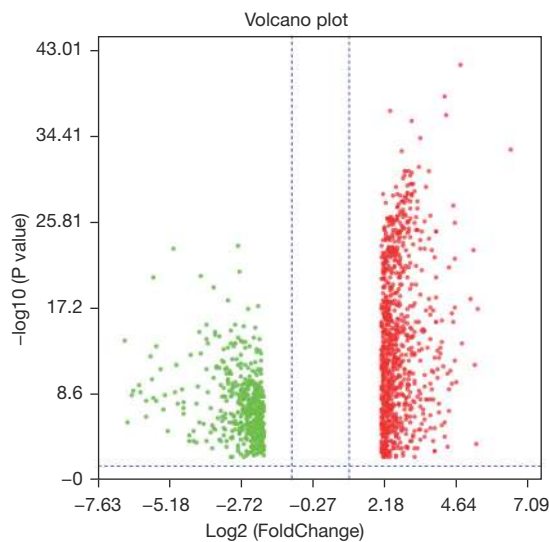
Figure 3 Volcano plot depicting the differentially expressed genes between normal gastric and GAC samples. The red dots indicate upregulated genes, the green dots indicate downregulated genes, and the white block represents the remaining genes—|log2FC| <2. The Y-axis represents FDR, and the X-axis represents the value of log2FC. DEGs, differentially expressed genes; GAC, gastric adenocarcinoma.

clustered as those of the histone modification; however, they were enriched in different ChIP-seq TF clusters in different cell lines (*Figure 5B*), such as forkhead box M1 (FOXM1) in MCF7 cells and endometrial cells (ECC1), E2F transcription factor 4 (E2F4) in HeLaS3 cells and murine erythroleukemia (MEL) cells, and the suppressor of zest 12 (SUZ12) in NTERA2 clone D1 (NT2D1).

### Integrated analysis of DEGs, differential miRNAs, and promoter methylation

In the Venn diagram shown in *Figure 6*, a total of 31 genes appeared in the overlapped area between differentially expressed miRNA target genes, DEGs, and promoter methylated genes. In addition, based on the hypermethylation and hypomethylation levels, potential regulatory networks associated with methylation, miRNA expression, and gene expression were generated (*Figure 1*). Within this network were 17 upregulated and 3 downregulated miRNAs and 13 upregulated and 4 downregulated genes. In addition, 6 hypermethylated genes and 11 hypomethylated genes were included. In total, 20 miRNAs were found in the network from which miR-133a (degree =4) had the highest degree.

In addition, A disintegrin and metalloproteinase domain 12 (ADAM12) (31), orthodenticle homeobox 1 (OTX1), and transcription factor AP-2 alpha (TFAP2A) (32) had relatively high degrees. Moreover, ADAM12 was associated with upregulation of hsa-mir-501, hsa-mir-19a, and hsa-mir-130b, and downregulation of hsa-mir-1 and hsa-mir-133a (*Figure 1*). TFAP2P correlated with upregulation of hsa-mir-429, hsa-mir-200b, hsa-mir-200c, and hsa-mir-135b, as well as the downregulation of hsa-mir-133a.

Establishment of the risk score and prognostic model of the nomogram

The method section has presented the formula to calculate the risk score based on 5 genes (*ADAM12*, *CDH19*, *RIMS1*, *SHOX2*, and *IGF2BP1*), the method section has presented the formula to calculate the risk score. Based on the median f the risk score, the 350 GAC patients in the training cohort were divided into a high-risk group (n=175) and a low-risk group (n=175). Time-dependent ROC curves (*Figure 7A*) were drawn to evaluate the predictive efficacy of the constructed risk score-based prognostic model, with AUCs for the prediction of 1 year, 3 years, and 5 years of OS were 0.607, 0.632, and 0.735, respectively. Kaplan-Meier survival analysis (*Figure 7B*) indicated that patients in the high-risk group had worse prognoses than the low-risk group (P=0.001). The GSE15459 dataset (n=192) was used as the validation cohort to evaluate the performance of the risk score-based model; it comprised 96 patients in the high-risk group and 96 patients in the low-risk group, respectively. The time-dependent ROC curve and Kaplan-Meier survival curve were also plotted. As shown in *Figure 7C,7D*, the AUC for the prediction of OS at 1 year, 3 years, 5 years in the validation cohort were 0.568, 0.587 and 0.615, respectively; worse prognoses were revealed in patients in the high-risk group than in the low-risk group (P=0.0089).

To construct a prognostic model of nomogram, the univariate and multivariate analyses of OS regarding canonical clinicopathologic characteristics were performed in the training cohort (Table S1). In the univariate analysis, OS was significantly correlated with risk score (P=0.001), TNM-T stage (P=0.018), TNM-N stage (P=0.005), TNM-M stage (P=0.014), and age at diagnosis (P=0.023). However, in the multivariate analysis, only the risk score (P=0.007), TNM-N stage (P=0.017), TNM-M stage (P=0.017), and age at diagnosis (P=0.012) remained significantly associated with OS. As shown in *Figure 7E*, a

Huang et al. Integrated bioinformatic analysis of GAC
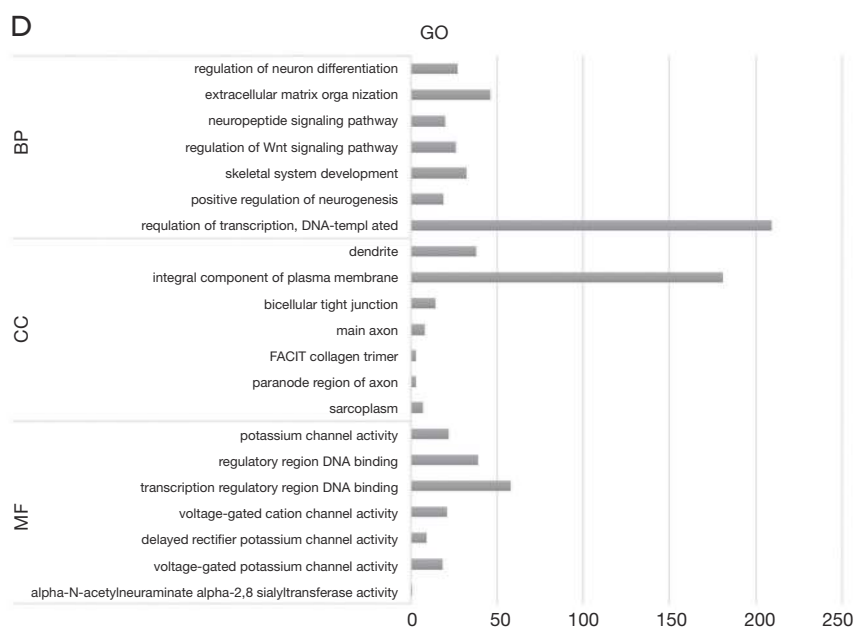
A

B

C
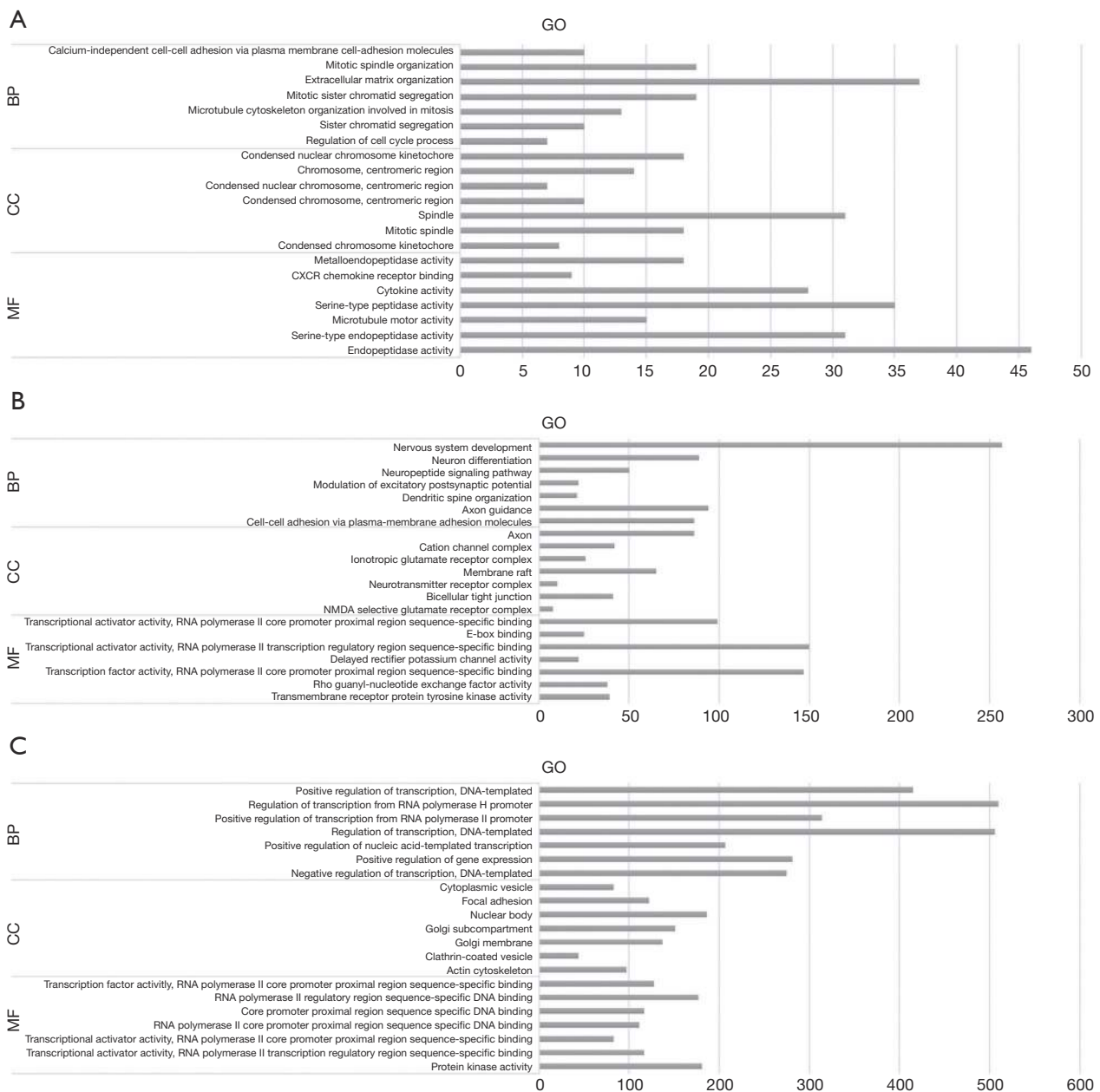
Figure 4 GO enrichment analysis of methylation-related genes. GO enrichment results for the (A) DEGs, (B) DMGs, (C) miRNA target genes, (D) and promotor methylation. The data are represented as the mean ± SEM. The Student's *t*-test was used to analyze significant differences, P<0.05 *vs.* shCtrl. shCtrl, negative control cells; GO, Gene Ontology; DEGs, differentially expressed genes; DMGs, differentially methylated genes; BP, biological process; CC, cellular component; MF, molecular function.
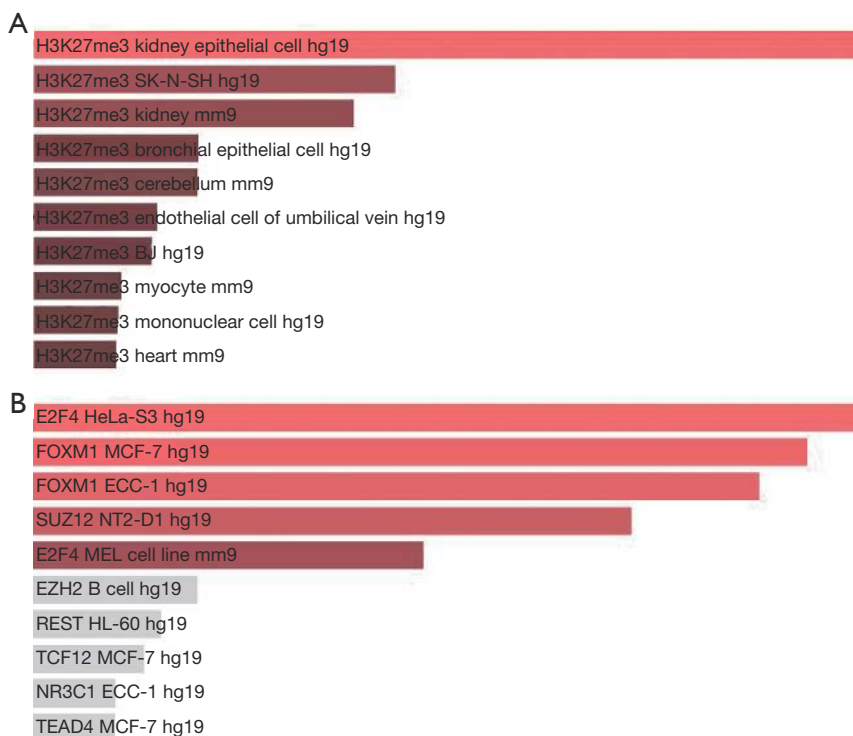

Figure 5 Transcriptional analysis of DEGs. Enriched TFs are located upstream of the DEGs, as identified by ChIP-X enrichment analysis (A) and ENCODE ChIP-seq of histone modifications and TFs (B). TFs, transcription factors; DEGs, differentially expressed genes.
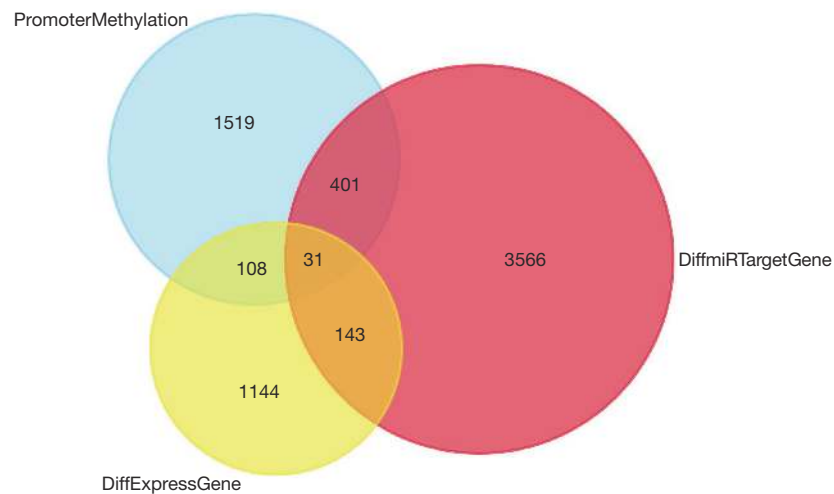
Page 10 of 18

Huang et al. Integrated bioinformatic analysis of GAC

Figure 6 Correlation analysis of the DEGs, differentially expressed miRNAs, and promotor methylation. The Venn diagram representing the functional enrichment and interaction network formed in the context of the DEGs, miRNA target genes, and promoter methylation are also shown. DEGs, differentially expressed genes.

nomogram was further constructed to predict 1-, 3-, and 5-year OS in the training cohort (TCGA-STAD dataset) to provide an accurate prediction of the prognosis. The estimated total points were calculated from clinical factors significantly associated with OS in the univariate analysis. For example, a high-risk score patient (67.5 points) aged older than 60 years old (65 points) with pathologic T3/T4 (27.5 points), N1/2/3 (75 points), M1 stage (100 points) would score 335 points in total for OS, and his/her predicted 1-, 3-, 5-year OS would be approximately 50%, 12.5%, and less than 10% respectively. Additionally, the prognostic nomogram model was internally validated by assessing both discriminations using the c index and plotting calibration curves (*Figure 7F*). The c-index of the combined nomogram (0.640, 95% CI: 0.615–0.666) was higher than that of other clinicopathological characteristics alone.

Analysis of the relationships between key gene methylation and expression

To further analyze the precise methylation loci and expression, MEXPRESS was used to investigate the relationships between ADAM12 and TFAP2A methylation and expression (Figures S5,S6). The dark blue regions in Figure S5 demonstrate a negative correlation with ADAM12 gene expression (Pearson's correlation coefficients for each probe are shown on the right), indicating that ADAM12 methylation silences gene expression. The dark

blue regions in Figure S6 demonstrate a positive correlation with TFAP2A gene expression, indicating that TFAP2A methylation activates gene expression.

Survival outcomes

Survival outcomes were analyzed using UALCAN to investigate whether the gene expression levels of the hypermethylation key drivers ADAM12, SLC5A7, and CDH19, and the hypomethylation driver TFAP2A, play important roles in GAC patient prognosis. The log-rank P values of ADAM12, SLC5A7, CDH19, and TFAP2A were 0.0044, 0.013, 0.036, and 0.0066, respectively. Therefore, regulation of oncogenes caused by aberrant methylation could result in different overall survival. Specifically, patients with a high TFAP2A expression (*Figure 8A*) had better prognoses, while those with a low/medium expression of ADAM12 (*Figure 8B*), CDH19 (*Figure 8C*), and SLC5A7 (*Figure 8D*) had poorer prognoses.

Discussion

In this study, we performed an integrated analysis of DEGs, DMGs, the target genes of miRNAs, and promoter methylation. Several DEGs, miRNAs, and TFs were screened with respect to metabolism and cell apoptosis found in various diseases, and which are known to play a significant role in the progression of GAC. Zhao *et al.* (33)
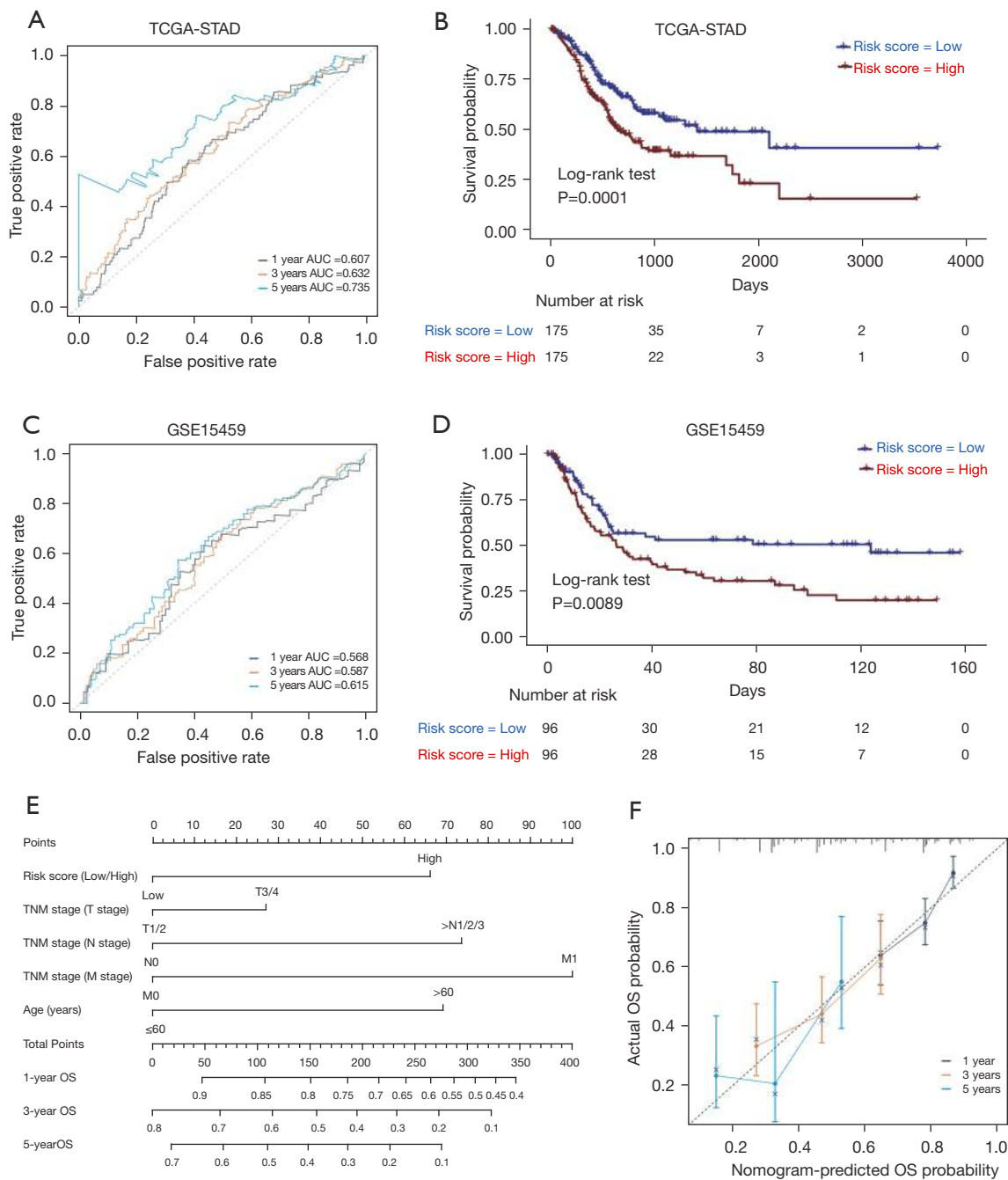
Figure 7 Exploration of the risk score and prognostic model of the nomogram in the training and validation cohort. (A) ROC curve analysis of risk score-based prognostic model in the training cohort. (B) Kaplan-Meier survival curve of the high- and low-risk subgroups in the training cohort. (C) ROC curve analysis of risk score-based prognostic model in the validation cohort. (D) Kaplan-Meier survival curve of the high- and low-risk subgroups in the validation cohort. (E) Nomogram for predicting the OS rates at 1-, 3- and 5-year in the training cohort. (F) Calibration curve of the nomogram for predicting 1-, 3-, and 5-year OS in the training cohort. AUC, area under the curve; OS, overall survival; ROC, receiver operating characteristic.
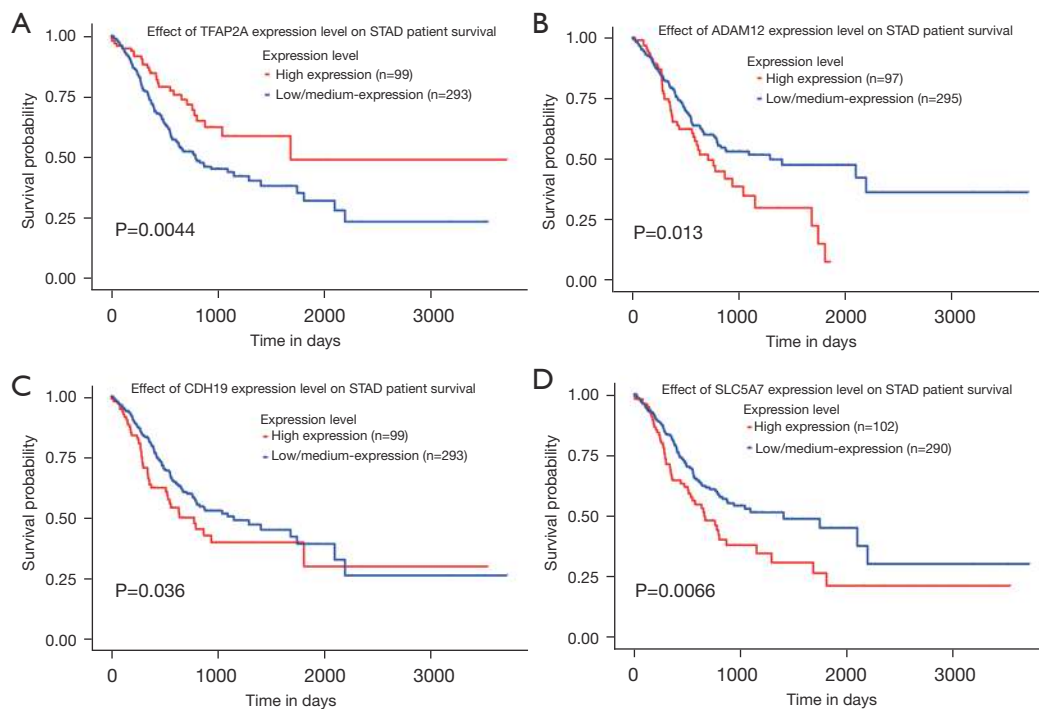
Page 12 of 18

Huang et al. Integrated bioinformatic analysis of GAC

Figure 8 Survival analyzes to investigate the impact of the key methylation driver genes for the prognoses of GAC patients. Kaplan-Meier survival curves demonstrating the prognostic value of (A) *TFAP2A*, (B) *ADAM12*, (C) *CDH19*, (D), and *SLC5A7*. A two-sided P value of <0.05 was regarded as statistically significant; of note, the P value was FDR-corrected. P<0.05 *vs.* shCtrl. shCtrl, negative control cells. ADAM12, A disintegrin and metalloproteinase domain 12; TFAP2A, transcription factor AP-2 alpha; SLC5A7, solute carrier family 5 member 7; CDH19, cadherin 19; FDR, false discovery rate.

identified several novel genetic aberrations, gene network modules, and miRNA target interactions within TCGA. However, their results differed slightly from ours, likely due to the bioinformatics tools used. Additionally, as the pathogenesis of GAC remains unclear, the results presented here may help to clarify the molecular mechanisms of this disease.

Our GO analysis demonstrated enrichment of DEGs in the ECM. Extracellular composition and organization are regulated to control cell behavior and differentiation; however, dysregulation of ECM dynamics leads to the development of various diseases, such as cancer (34). Moreover, the DMGs in our study were enriched in transcriptional activator and TF activity. For example, the promoter region contains putative binding sites for multiple TFs, including the signal transducer and activator of transcription 6 (STAT6), a downstream effector of IL-4 (35). The GO enrichment analysis also revealed that changes in the target genes of miRNAs were primarily enriched in protein kinase activity, which is known to be related to

cancer proliferation, cell invasion, and migration (36-40). In terms of promoter methylation, GO was enriched in the transcription regulatory region for DNA binding. Changes in DNA methylation associated with GC have been previously studied and provide additional clues into the pathogenesis of the disease (41).

The KEGG analysis showed that DEGs were primarily enriched in the IL-17 signaling pathway, cytokine-cytokine receptor interaction, and cell cycling, which is consistent with previous reports (42). It has been reported that IL-17 plays an important role in the upregulation of VEGF to promote tumor angiogenesis. We also identified that DMGs were enriched in the cAMP signaling pathway, RAP1 signaling pathway, and cell adhesion molecules. cAMP interacts with the HERG protein by binding to the cAMP-binding domain of HERG protein and subsequently impacts HERG in GAC (43). Meanwhile, reduced RAP1 signaling, which has been reported in prostate cancer, is related to cell adhesion, migration, and survival associated with metastasis. Cell adhesion molecules were in accordance with the studies

above (44). Furthermore, the target genes of miRNAs were enriched in proteoglycans in cancer, the RAP1 signaling pathway, MAPK signaling pathway, pathways in cancer, and RAS signaling pathways. Proteoglycans perform multiple functions, including the regulation of tumor cell growth, survival, adhesion, metastasis, and angiogenesis through interactions between their charged glycosaminoglycan chains and effector proteins (45).

The core protein of proteoglycans can also work with other proteins, such as integrins, to accommodate their signaling (46). The MAPK signaling pathway is thought to be involved in a myriad of mechanisms associated with eukaryotic cell regulation by coordinating with diverse receptor families, such as growth factors closely related to the immune system (47). In addition, MAPK regulates the activation of gene transcription, protein synthesis, cell cycle machinery, cell death, and differentiation (47). RAS also plays a key role in malignant cells, including the deregulation of tumor cell growth, programmed cell death, and the ability to invade and induce new blood vessel formation, making it a new therapeutic target for GAC (48,49). Promoter methylation was also significantly enriched in cell adhesion molecules. These results provide the foundation for further GAC research; however, additional studies are required to verify our findings.

TF expression and histone modifications are under strict control. Hence, during dynamic packaging, maintaining stable and ordered chromatin is vital to ensure normal cellular homeostasis. Factors influencing the number of chromatin-associated cellular events, including transcription, warped histones, and DNA, are subject to covalent post-translational modifications (50). In the present study, a comparative analysis of histone modifications in tumor and normal tissues was conducted, which revealed that the DEGs were primarily regulated by H3K27me3 in several cancer cell lines. H3K27me3, a post-translational modification, is highly correlated with genomic silencing (51). Many studies have reported that cytosine methylation of DNA regulatory sequences is associated with the transcriptional inactivation of genes, while hypomethylation supports transcriptional activation (52,53). Zhang *et al.* explored the relevance of H3K27me3 and DNA methylation in terms of molecular mechanisms to better understand the relationship between DNA methylation and histone modifications in cancer-associated gene silencing (53). Our results suggest that abnormal modification of H3K27me3 may play an important role in GAC with regulation predominantly executed by H3K27me3. Other studies have reported that hypomethylated genes in H3K27me3 activate transcription, whereas hypermethylated genes inactivate transcription (52,54).

The enrichment analysis performed with ChIP-seq revealed that DEGs were enriched in different TF clusters, including E2F4, FOXM1, and SUZ12. E2F4, a member of E2F, regulates various cellular functions related to the cell cycle and apoptosis (55). In fact, GAC frequently exhibits a mutation of the adenosine-guanine-cytosine (AGC) repeat of the E2F4 gene. Furthermore, E2F4 appears to promote tumor progression in GAC (56,57). FOXM1 participates in cell cycling by regulating both the transition from G1 to S and progression to mitosis (58). Moreover, FOXM1 mediates the promotion of human GAC angiogenesis, growth, and metastasis, and enhanced regulation of FOXM1 leads to the acceleration of GAC (58,59). SUZ12 is primarily involved in histone modification (60), whereas HOTAIR expression positively correlates with SUZ12 expression levels and, therefore, may affect the epigenetic state of cancer tissues. A high expression of HOTAIR is associated with higher stages and lymph node metastasis in GAC (61). Consequently, a high expression of SUZ12 may contribute to GAC progression.

The constructed differential miRNAs and promoter methylation networks showed that the hub nodes included SLC5A7 and CDH19. Among these hub genes, ADAM12 and TFAP2A showed the highest node degrees. Significant changes in the expression of key upstream genes may affect a large number of downstream target genes. Indeed, altered DNA methylation in the promoter region of genes has been shown to cause the inactivation of tumor suppressors and other cancer-related genes and is regarded as the most well-defined epigenetic characteristic in GC (62). ADAM12 is significantly upregulated in GAC, inhibiting cancer proliferation, invasion, and metastasis (63). In addition, the upregulation of hypomethylated ADAM12 has also been reported in breast, pancreatic, and ovarian cancers, and may serve as a valuable biomarker for diagnosis, treatment, and prognosis (64,65). In the cellular context, TFAP2A is individually related to cell differentiation and development and cancer progression/regression. Wang *et al.* suggested that a reduced expression of TFAP2P is associated with GAC prognosis and may be a potential marker for improved prognosis in GAC patients (66). Moreover, ZNF471 recruits KAP1 to the promoter of target genes, thereby inducing H3K9me3 enrichment and repressing oncogenic TFAP2A (67). However, few studies have examined SLC5A7 in the context of GAC. The downregulated expression

Page 14 of 18

Huang et al. Integrated bioinformatic analysis of GAC

of *SLC5A7* in various cancers is reportedly associated with poor clinical outcomes (68). For instance, SLC5A7 was found to be downregulated and hypermethylated in lung adenocarcinoma (LUAD). However, it was shown to be upregulated and hypomethylated in squamous cell carcinoma (LUSC) (69). The overall survival (OS) of patients with high promotor region methylation of LUSC was better than for those with low methylation; however, no significant correlation was observed between methylation and OS in LUAD (69). CDH19 encodes a tyrosine phosphatase associated with adherent junctions. As GC develops from the same glands as colorectal tumors, CDH19 may also be related to GAC (70). Furthermore, new human antibodies targeting CDH19 have been applied to prevent, treat, and ameliorate melanoma and metastatic melanoma disease.

The generated network of DEGs, differential miRNAs, and promoter methylation constructed in this study identified two significant differentially expressed miRNAs, miR-1 and miR133a, with miR-133a showing the highest node degree (five). miRNAs can distinguish between cancer and normal tissues or various cancers (71,72), and may regulate diverse protein-coding genes and participate in different molecular pathways associated with tumor evolution and progression (73-75). Recent studies have claimed that miR-1 and miR-133a function as tumor suppressors, and their oncogenes have been reported in different cell environments (76) . Indeed, overexpression of miR-1 can inhibit cell growth, differentiation, and migration, while its downregulation promotes cell proliferation and migration in endothelial cells of GC and activates proangiogenesis-associated signaling (77). Moreover, the potential anti-tumorigenic function of miR-1 in primary HCCs is related to the methylation-mediated silencing of miR-1 and upregulation of the potential oncogenic target genes of miR-1 (78). Recently, miR-1 has been suggested as a therapeutic target for prostate cancer and NSCLC (79,80).

Furthermore, reduced miR-133a expression, caused by hypermethylation of its promoter or negative regulation of TFs, was associated with larger and more invasive GAC tumors (81,82), resulting in metastasis of peripheral organs (83). In addition, altered miR-133a expression has been reported in various types of cancer, including esophageal squamous cell carcinoma and bladder cancer, mainly regulating the expression of FSCN1 as a tumor suppressor (84,85). DNA methylation could, directly and indirectly, affect the drug response. Directly, DNA methylation involves gene transcription and thus impacts the methylation level of epigenome targets. Target genes and the expression of proteins affected by epigenome targets is the indirect way that is imposed on the response to agents of cancer cells (86). Therefore, as the key DEGs, miRNAs, and TFs mentioned above and based on the relationship with methylation, the outcome of treatment with GAC could be predicted by detecting hypermethylated or hypomethylated levels of the corresponding targets.

Clinical or/and pathological staging at diagnosis and treatment are practical components that determine the prognosis of GAC patients (87). Methylation-driven gene expression could be predicted by hypomethylation and hypermethylation of corresponding genes exactly based on bioinformatic analysis (88). Nowadays, Gastrointestinal endoscopy (GIS) has served as one of the most widely used diagnostic tools for pathological analysis, is neither suitable for early diagnosis nor free from the risk of morbidity, and detecting miRNAs in body fluids can be a potential method in the diagnosis of GAC. It has been demonstrated that endogenous miRNA in serum or plasma remains stable. Whether in extreme conditions such as boiling, very low/high pH levels, extended storage time, and multiple freeze-thaw cycles (89,90), or stored for more than a decade in archival tissues and human serum specimens, miRNA can be easily detected (91). The stability and resistance to degradation of miRNAs make them more useful biomarkers for cancer diagnosis. Q-RT-PCR is widely utilized to identify previously unknown new miRNAs, which is easier to perform and is a cost-effective technique. Blood-based (92) and gastric juice-based (93) miRNAs are also considered potential biomarkers for early detection of GAC. miRNA-125b was identified as a crucial miRNA in GC development from miRNA-gene network analyses related to gastric oncogenesis (93). Another study found that hsa-miR-421/hsa-miR-29b-1-5p targeted CREBZF and could play an important role in MKN-74 cell migration, suggesting that increased CREBZF by inhibition of hsa-miR-421/hsa-miR-29b-1-5p could be pivotal in preventing gastric cancer in its early stage (94). Moreover, proliferation, as well as migration and invasion abilities of GAC cells, were restricted by the overexpression of HAND2-AS1 and HIF3A and enhanced by miR-184. Furthermore, HAND2-AS1 rescued enhanced GAC cell proliferation, migration, and invasion abilities as well as the glycolytic process caused by hypoxia via miR-184/HIF3A (95). Therefore, once preclinical screening and validation are resolved, it will promote the development of miRNA biomarkers in body

fluids for clinical applications.

Conclusions

Our study reports on an integrated analysis of high-throughput sequencing data for DEGs, differential miRNAs, and promoter methylation associated with GAC. Four possible GAC-related DEGs, including ADAM12, TFAP2A, SCL5A7, and CDH19, and two possible GAC-related miRNAs, miR-133a and miR-1, were identified. Moreover, DMRs were screened and integrated. GAC-related promoter methylation may be under comparable transcriptional regulation. Furthermore, several TFs and miRNAs, as well as promoter methylation, may play critical roles in GAC tumorigenesis. These results provide the foundation for further GAC research and will need to be confirmed by additional experiments.

Footnote

Reporting Checklist: The authors have completed the MDAR reporting checklist. Available at https://dx.doi.org/10.21037/atm-21-3211

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://dx.doi.org/10.21037/atm-21-3211). 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).

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. 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.
2. Zhao S, Fan N, Li H, et al. Apatinib combined with paclitaxel-based chemotherapy in patients with taxane-resistant advanced gastric cancer: a single-arm exploratory study. Ann Transl Med 2020;8:1233.
3. Galm O, Esteller M. Beyond genetics--the emerging role of epigenetic changes in hematopoietic malignancies. Int J Hematol 2004;80:120-7.
4. Sepich-Poore GD, Zitvogel L, Straussman R, et al. The microbiome and human cancer. Science 2021;371:eabc4552.
5. Sharma PC, Gupta A. MicroRNAs: potential biomarkers for diagnosis and prognosis of different cancers. Transl Cancer Res 2020;9:5798-818.
6. Dejea CM, Fathi P, Craig JM, et al. Patients with familial adenomatous polyposis harbor colonic biofilms containing tumorigenic bacteria. Science 2018;359:592-7.
7. Wang S, Wu W, Claret FX. Mutual regulation of microRNAs and DNA methylation in human cancers. Epigenetics 2017;12:187-97.
8. Liu H, Lei C, He Q, et al. Nuclear functions of mammalian MicroRNAs in gene regulation, immunity and cancer. Mol Cancer 2018;17:64.
9. Pilehchian Langroudi M, Nikbakhsh N, Samadani AA, et al. FAT4 hypermethylation and grade dependent downregulation in gastric adenocarcinoma. J Cell Commun Signal 2017;11:69-75.
10. Van Cutsem E, Sagaert X, Topal B, et al. Gastric cancer. Lancet 2016;388:2654-64.
11. Ando M, Saito Y, Xu G, et al. Publisher Correction: Chromatin dysregulation and DNA methylation at transcription start sites associated with transcriptional repression in cancers. Nat Commun 2019;10:2415.
12. Biswas S, Rao CM. Epigenetics in cancer: Fundamentals and Beyond. Pharmacol Ther 2017;173:118-34.
13. Baylin SB, Jones PA. Epigenetic Determinants of Cancer. Cold Spring Harb Perspect Biol 2016;8:a019505.
14. Garraway LA, Lander ES. Lessons from the cancer genome. Cell 2013;153:17-37.

Page 16 of 18

Huang et al. Integrated bioinformatic analysis of GAC

15. Shen H, Laird PW. Interplay between the cancer genome and epigenome. Cell 2013;153:38-55.

16. Smyth GK, Michaud J, Scott HS. Use of within-array replicate spots for assessing differential expression in microarray experiments. Bioinformatics 2005;21:2067-75.

17. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics 2014;30:1363-9.

18. Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res 2016;44:W90-7.

19. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000;28:27-30.

20. Nishida K, Ono K, Kanaya S, et al. KEGGscape: a Cytoscape app for pathway data integration. F1000Res 2014;3:144.

21. McLaughlin MJ, Sainani KL. Bonferroni, Holm, and Hochberg corrections: fun names, serious changes to p values. PM R 2014;6:544-6.

22. Lachmann A, Xu H, Krishnan J, et al. ChEA: transcription factor regulation inferred from integrating genome-wide ChIP-X experiments. Bioinformatics 2010;26:2438-44.

23. ENCODE Project Consortium. The ENCODE (ENCyclopedia Of DNA Elements) Project. Science 2004;306:636-40.

24. Pathan M, Keerthikumar S, Ang CS, et al. FunRich: An open access standalone functional enrichment and interaction network analysis tool. Proteomics 2015;15:2597-601.

25. Kestler HA, Müller A, Gress TM, et al. Generalized Venn diagrams: a new method of visualizing complex genetic set relations. Bioinformatics 2005;21:1592-5.

26. Koch A, De Meyer T, Jeschke J, et al. MEXPRESS: visualizing expression, DNA methylation and clinical TCGA data. BMC Genomics 2015;16:636.

27. Iasonos A, Schrag D, Raj GV, et al. How to build and interpret a nomogram for cancer prognosis. J Clin Oncol 2008;26:1364-70.

28. Chandrashekar DS, Bashel B, Balasubramanya SAH, et al. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. Neoplasia 2017;19:649-58.

29. Bradburn MJ, Clark TG, Love SB, et al. Survival analysis part II: multivariate data analysis--an introduction to concepts and methods. Br J Cancer 2003;89:431-6.

30. Wu W, Sargent DJ. Chapter 12 - Statistics and Clinical Trials. In: Gunderson LL, Tepper JE, editors. Clinical Radiation Oncology (Third Edition). Philadelphia: W.B. Saunders, 2012:223-37.

31. Asakura M, Kitakaze M, Takashima S, et al. Cardiac hypertrophy is inhibited by antagonism of ADAM12 processing of HB-EGF: metalloproteinase inhibitors as a new therapy. Nat Med 2002;8:35-40.

32. Hilger-Eversheim K, Moser M, Schorle H, et al. Regulatory roles of AP-2 transcription factors in vertebrate development, apoptosis and cell-cycle control. Gene 2000;260:1-12.

33. Zhao R, Zhao L, Xu X, et al. Analysis of microRNA expression profiles reveals a 5-microRNA prognostic signature for predicting overall survival time in patients with gastric adenocarcinoma. Oncol Rep 2019;41:2775-89.

34. Walker C, Mojares E, Del Río Hernández A. Role of Extracellular Matrix in Development and Cancer Progression. Int J Mol Sci 2018;19:3028.

35. Mazzone A, Gibbons SJ, Bernard CE, et al. Identification and characterization of a novel promoter for the human ANO1 gene regulated by the transcription factor signal transducer and activator of transcription 6 (STAT6). FASEB J 2015;29:152-63.

36. Akter H, Park M, Kwon OS, et al. Activation of matrix metalloproteinase-9 (MMP-9) by neurotensin promotes cell invasion and migration through ERK pathway in gastric cancer. Tumour Biol 2015;36:6053-62.

37. Bae JS, Park SH, Kim KM, et al. CK2α phosphorylates DBC1 and is involved in the progression of gastric carcinoma and predicts poor survival of gastric carcinoma patients. Int J Cancer 2015;136:797-809.

38. Fu H, Wang C, Yang D, et al. Curcumin regulates proliferation, autophagy, and apoptosis in gastric cancer cells by affecting PI3K and P53 signaling. J Cell Physiol 2018;233:4634-42.

39. Liu J, Xu B, Xu G, et al. Reduced EphB6 protein in gastric carcinoma and associated lymph nodes suggests EphB6 as a gastric tumor and metastasis inhibitor. Cancer Biomark 2017;19:241-8.

40. Wang BY, Liu QY, Cao J, et al. Selective CDK7 inhibition with BS-181 suppresses cell proliferation and induces cell cycle arrest and apoptosis in gastric cancer. Drug Des Devel Ther 2016;10:1181-9.

41. Zheng ZH, Sun XJ, Ma MC, et al. Studies of promoter methylation status and protein expression of E-cadherin gene in associated progression stages of gastric cancer. Yi Chuan Xue Bao 2003;30:103-8.

42. Wu X, Yang T, Liu X, et al. IL-17 promotes tumor angiogenesis through Stat3 pathway mediated

upregulation of VEGF in gastric cancer. Tumour Biol 2016;37:5493-501.

43. Shao XD, Guo XZ, Ren LN, et al. The mechanism of COX-2 regulating HERG channel in gastric cancer cells. Bratisl Lek Listy 2014;115:487-91.

44. Bailey CL, Kelly P, Casey PJ. Activation of Rap1 promotes prostate cancer metastasis. Cancer Res 2009;69:4962-8.

45. Sanderson RD, Yang Y, Purushothaman A, et al. Proteoglycans and cancer. Cell-Extracellular Matrix Interactions in Cancer 2010:191-215.

46. Kyriakis JM, Avruch J. Mammalian mitogen-activated protein kinase signal transduction pathways activated by stress and inflammation. Physiol Rev 2001;81:807-69.

47. Downward J. Targeting RAS signalling pathways in cancer therapy. Nat Rev Cancer 2003;3:11-22.

48. Shields JM, Pruitt K, McFall A, et al. Understanding Ras: 'it ain't over 'til it's over'. Trends Cell Biol 2000;10:147-54.

49. Barski A, Cuddapah S, Cui K, et al. High-resolution profiling of histone methylations in the human genome. Cell 2007;129:823-37.

50. Meng X, Zhao Y, Liu J, et al. Comprehensive analysis of histone modification-associated genes on differential gene expression and prognosis in gastric cancer. Exp Ther Med 2019;18:2219-30.

51. Richardson B, Yung R. Role of DNA methylation in the regulation of cell function. J Lab Clin Med 1999;134:333-40.

52. Nakao M. Epigenetics: interaction of DNA methylation and chromatin. Gene 2001;278:25-31.

53. Zhang L, Zhong K, Dai Y, et al. Genome-wide analysis of histone H3 lysine 27 trimethylation by ChIP-chip in gastric cancer patients. J Gastroenterol 2009;44:305-12.

54. Richardson BC. Role of DNA methylation in the regulation of cell function: autoimmunity, aging and cancer. J Nutr 2002;132:2401S-5S.

55. Evangelou K, Havaki S, Kotsinas A. E2F transcription factors and digestive system malignancies: how much do we know? World J Gastroenterol 2014;20:10212-6.

56. Schwemmle S, Pfeifer GP. Genomic structure and mutation screening of the E2F4 gene in human tumors. Int J Cancer 2000;86:672-7.

57. Souza RF, Yin J, Smolinski KN, et al. Frequent mutation of the E2F-4 cell cycle gene in primary human gastrointestinal tumors. Cancer Res 1997;57:2350-3.

58. Okada K, Fujiwara Y, Takahashi T, et al. Overexpression of forkhead box M1 transcription factor (FOXM1) is a potential prognostic marker and enhances chemoresistance for docetaxel in gastric cancer. Ann Surg Oncol 2013;20:1035-43.

59. Li Q, Zhang N, Jia Z, et al. Critical role and regulation of transcription factor FoxM1 in human gastric cancer angiogenesis and progression. Cancer Res 2009;69:3501-9.

60. Lee NK, Lee JH, Kim WK, et al. Promoter methylation of PCDH10 by HOTAIR regulates the progression of gastrointestinal stromal tumors. Oncotarget 2016;7:75307-18.

61. Hajjari M, Behmanesh M, Sadeghizadeh M, et al. Up-regulation of HOTAIR long non-coding RNA in human gastric adenocarcinoma tissues. Med Oncol 2013;30:670.

62. Qu Y, Dang S, Hou P. Gene methylation in gastric cancer. Clin Chim Acta 2013;424:53-65.

63. Carl-McGrath S, Lendeckel U, Ebert M, et al. The disintegrin-metalloproteinases ADAM9, ADAM12, and ADAM15 are upregulated in gastric cancer. Int J Oncol 2005;26:17-24.

64. Ahn JH, Ju W, Sung HY, inventors. Composition for diagnosing ovarian cancer metastasis by using cpg methylation in gene, and use thereof. United States Patent Application 20160265061. Available online: http://www.freepatentsonline.com/y2016/0265061.html

65. Yu J, Walter K, Omura N, et al. Unlike pancreatic cancer cells pancreatic cancer associated fibroblasts display minimal gene induction after 5-aza-2'-deoxycytidine. PLoS One 2012;7:e43456.

66. Wang W, Lv L, Pan K, et al. Reduced expression of transcription factor AP-2α is associated with gastric adenocarcinoma prognosis. PLoS One 2011;6:e24897.

67. Cao L, Wang S, Zhang Y, et al. Zinc-finger protein 471 suppresses gastric cancer through transcriptionally repressing downstream oncogenic PLS3 and TFAP2A. Oncogene 2018;37:3601-16.

68. Li M, Sun Q, Wang X. Transcriptional landscape of human cancers. Oncotarget 2017;8:34534-51.

69. Yang Y, Wang M, Liu B. Exploring and comparing of the gene expression and methylation differences between lung adenocarcinoma and squamous cell carcinoma. J Cell Physiol 2019;234:4454-9.

70. Bujko M, Kober P, Mikula M, et al. Expression changes of cell-cell adhesion-related genes in colorectal tumors. Oncol Lett 2015;9:2463-70.

71. Ling H, Fabbri M, Calin GA. MicroRNAs and other non-coding RNAs as targets for anticancer drug development. Nat Rev Drug Discov 2013;12:847-65.

72. Gray SG. Squaring the circle: sponging microRNAs in gastric cancer. Transl Cancer Res 2019;8:S183-7.

73. Li X, Zhang Y, Zhang Y, et al. Survival prediction of

Page 18 of 18

Huang et al. Integrated bioinformatic analysis of GAC

gastric cancer by a seven-microRNA signature. Gut 2010;59:579-85.

74. Volinia S, Calin GA, Liu CG, et al. A microRNA expression signature of human solid tumors defines cancer gene targets. Proc Natl Acad Sci U S A 2006;103:2257-61.

75. Moreira FC, Assumpção M, Hamoy IG, et al. MiRNA expression profile for the human gastric antrum region using ultra-deep sequencing. PLoS One 2014;9:e92300.

76. Sempere LF, Freemantle S, Pitha-Rowe I, et al. Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in murine and human neuronal differentiation. Genome Biol 2004;5:R13.

77. Xie M, Dart DA, Guo T, et al. MicroRNA-1 acts as a tumor suppressor microRNA by inhibiting angiogenesis-related growth factors in human gastric cancer. Gastric Cancer 2018;21:41-54.

78. Datta J, Kutay H, Nasser MW, et al. Methylation mediated silencing of MicroRNA-1 gene and its role in hepatocellular carcinogenesis. Cancer Res 2008;68:5049-58.

79. Ambs S, Prueitt RL, Yi M, et al. Genomic profiling of microRNA and messenger RNA reveals deregulated microRNA expression in prostate cancer. Cancer Res 2008;68:6162-70.

80. Hu Z, Chen X, Zhao Y, et al. Serum microRNA signatures identified in a genome-wide serum microRNA expression profiling predict survival of non-small-cell lung cancer. J Clin Oncol 2010;28:1721-6.

81. Nohata N, Hanazawa T, Enokida H, et al. microRNA-1/133a and microRNA-206/133b clusters: dysregulation and functional roles in human cancers. Oncotarget 2012;3:9-21.

82. Liu Y, Zhang X, Zhang Y, et al. Identification of miRNomes in human stomach and gastric carcinoma reveals miR-133b/a-3p as therapeutic target for gastric cancer. Cancer Lett 2015;369:58-66.

83. Xie M, Dart DA, Owen S, et al. Insights into roles of the miR-1, -133 and -206 family in gastric cancer (Review). Oncol Rep 2016;36:1191-8.

84. Chiyomaru T, Enokida H, Tatarano S, et al. miR-145 and miR-133a function as tumour suppressors and directly regulate FSCN1 expression in bladder cancer. Br J Cancer 2010;102:883-91.

85. Kano M, Seki N, Kikkawa N, et al. miR-145, miR-133a and miR-133b: Tumor-suppressive miRNAs target FSCN1 in esophageal squamous cell carcinoma. Int J Cancer 2010;127:2804-14.

86. Vural S, Palmisano A, Reinhold WC, et al. Association of expression of epigenetic molecular factors with DNA methylation and sensitivity to chemotherapeutic agents in cancer cell lines. Clin Epigenetics 2021;13:49.

87. Hohenberger P, Gretschel S. Gastric cancer. Lancet 2003;362:305-15.

88. Huang WY, Hsu SD, Huang HY, et al. MethHC: a database of DNA methylation and gene expression in human cancer. Nucleic Acids Res 2015;43:D856-61.

89. Mitchell PS, Parkin RK, Kroh EM, et al. Circulating microRNAs as stable blood-based markers for cancer detection. Proc Natl Acad Sci U S A 2008;105:10513-8.

90. Chen X, Ba Y, Ma L, et al. Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. Cell Res 2008;18:997-1006.

91. Iorio MV, Croce CM. MicroRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review. EMBO Mol Med 2012;4:143-59.

92. Cai H, Yuan Y, Hao YF, et al. Plasma microRNAs serve as novel potential biomarkers for early detection of gastric cancer. Med Oncol 2013;30:452.

93. Cui L, Zhang X, Ye G, et al. Gastric juice MicroRNAs as potential biomarkers for the screening of gastric cancer. Cancer 2013;119:1618-26.

94. Kim YJ, Jeong S, Jung WY, et al. miRNAs as potential biomarkers for the progression of gastric cancer inhibit CREBZF and regulate migration of gastric adenocarcinoma cells. Int J Med Sci 2020;17:693-701.

95. Xu Z, Lv H, Wang Y, et al. HAND2-AS1 Inhibits Gastric Adenocarcinoma Cells Proliferation and Aerobic Glycolysis via miRNAs Sponge. Cancer Manag Res 2020;12:3053-68.

Table S1 Univariate and multivariate analyses of variables associated with overall survival

Characteristics	Number of patients	Univariate analysis		Multivariable analysis	
		HR (95% CI)	P value	HR (95% CI)	P value
Risk (High *vs.* low)	350	1.729 (1.242, 2.407)	0.001	1.614 (1.138, 2.290)	0.007
Pathologic T stage (T3/4 *vs.* T1/2)	346	1.643 (1.088, 2.481)	0.018	1.217 (0.774, 1.913)	0.395
Pathologic N stage (N1/2/3 *vs.* N0)	339	1.805 (1.199, 2.716)	0.005	1.703 (1.100, 2.638)	0.017
Pathologic M stage (M1 *vs.* M0)	335	2.051 (1.157, 3.636)	0.014	2.060 (1.140, 3.722)	0.017
Age (>60 *vs.* ≤60)	347	1.531 (1.060, 2.212)	0.023	1.649 (1.119, 2.431)	0.012
Sex (male *vs.* female)	350	1.324 (0.930, 1.890)	0.119		
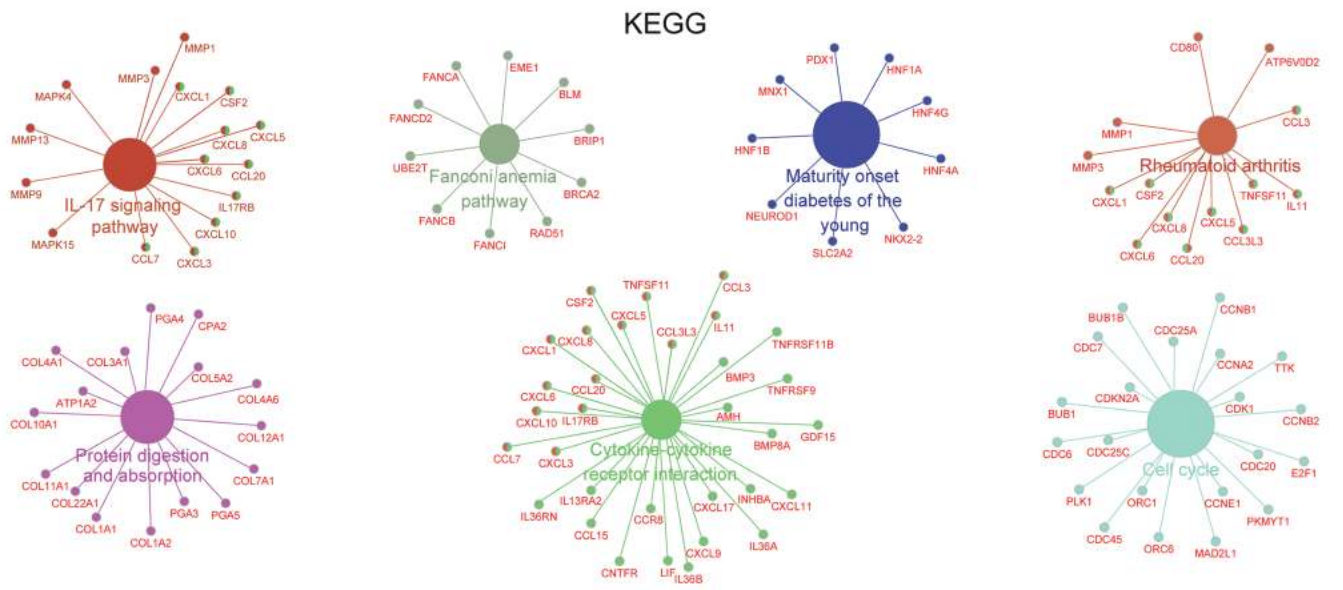
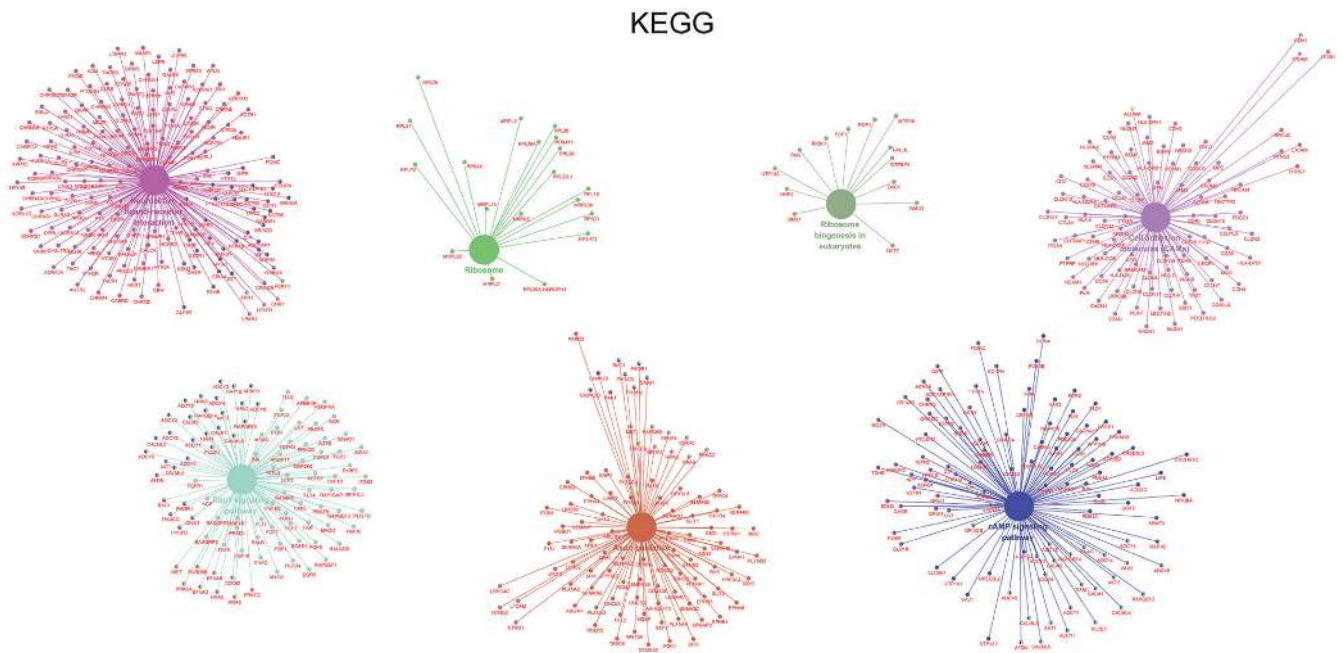**Figure S1** KEGG enrichment results for DEGs.

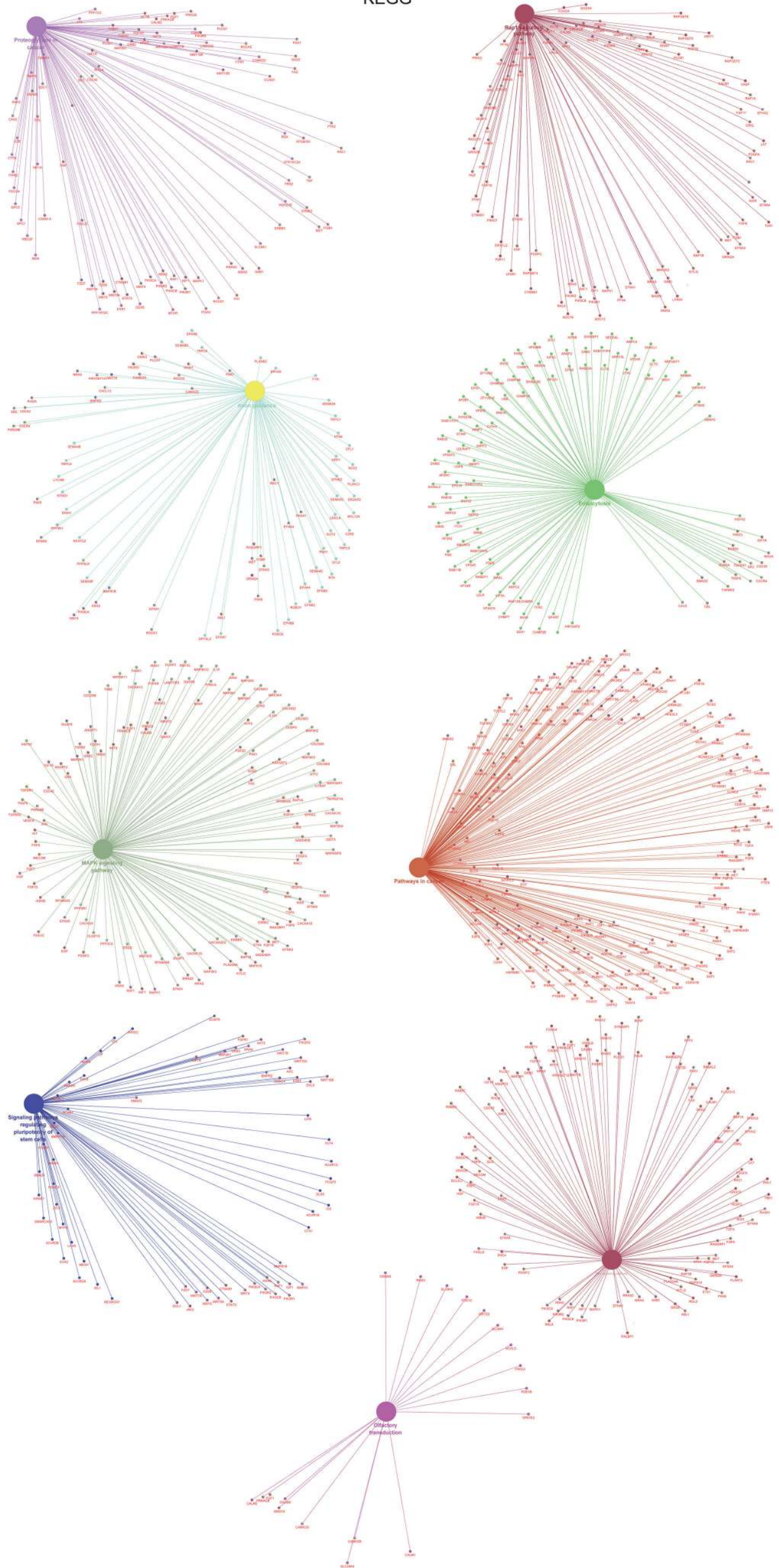**Figure S2** KEGG enrichment results for DMGs.

KEGG

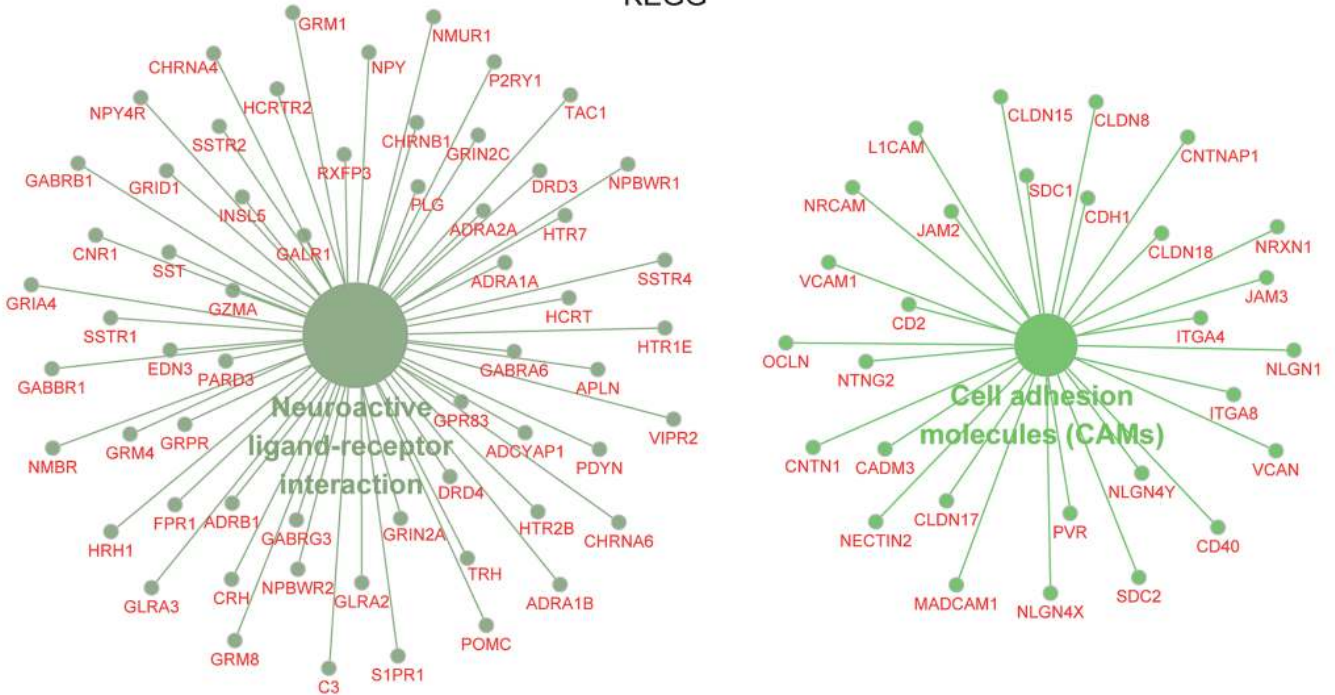**Figure S3** KEGG enrichment results for miRNA target genes.

Figure S4 KEGG enrichment results for differential promoter methylation.
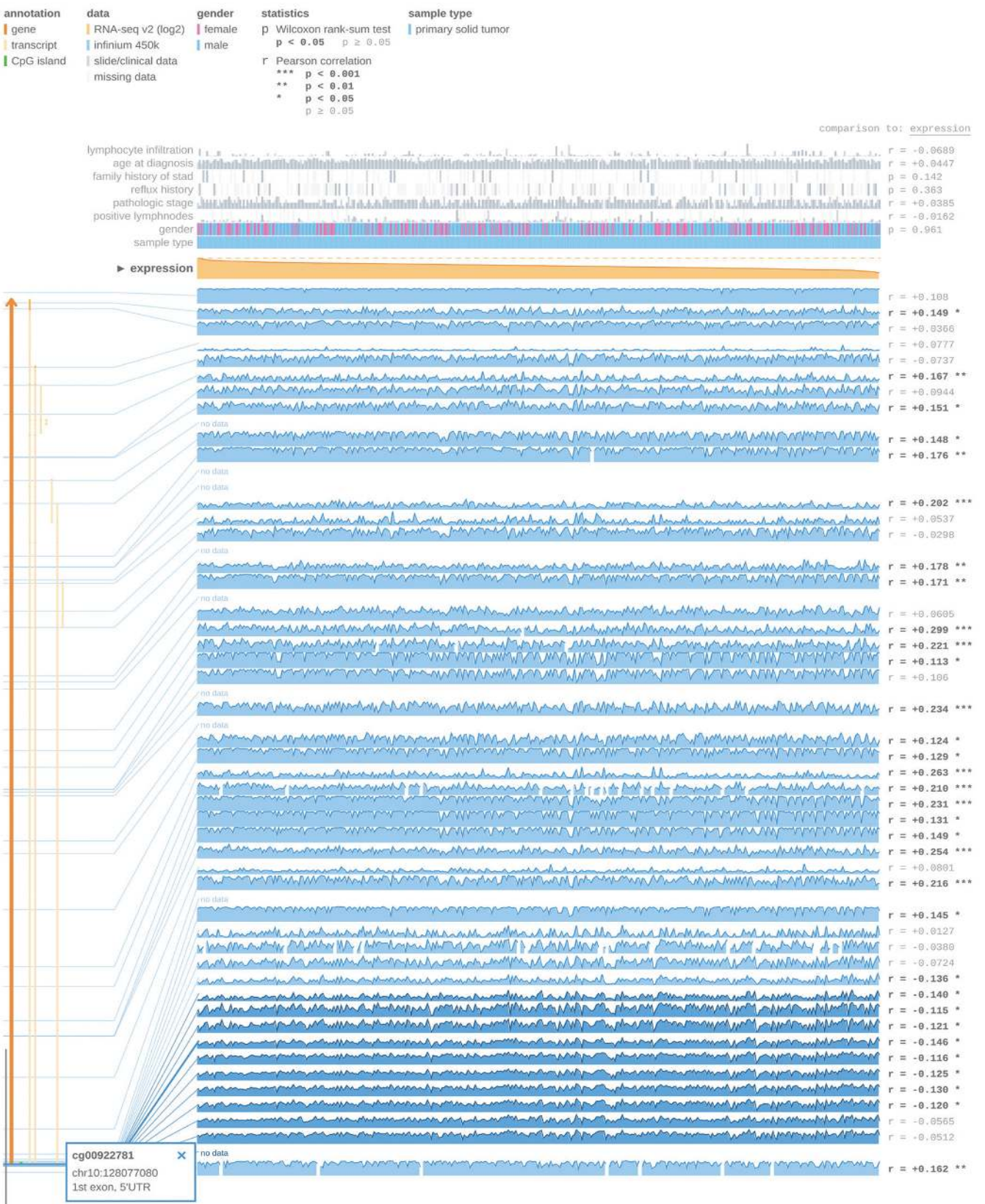

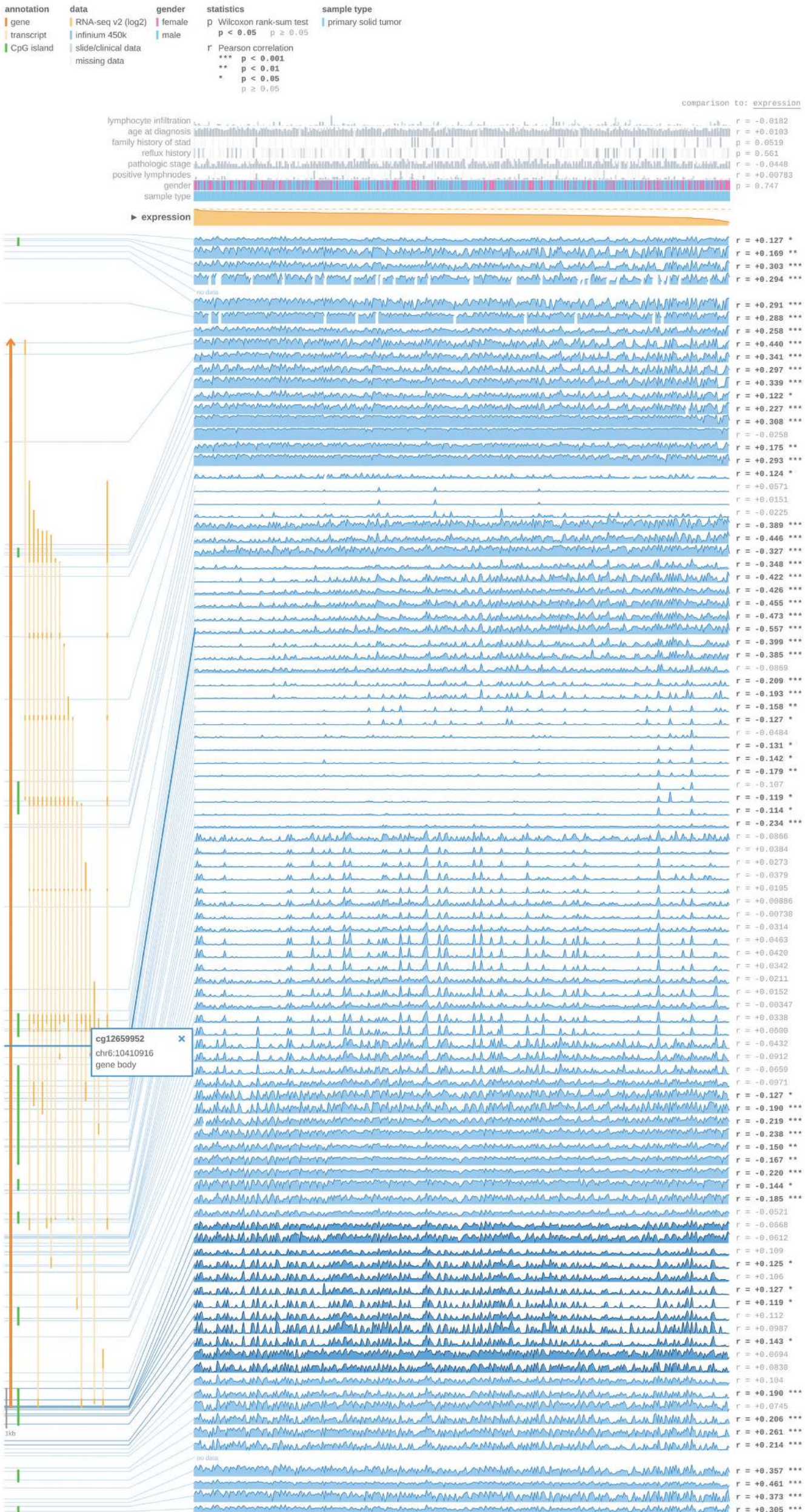**Figure S5** Correlation analysis for ADAM12 methylation locus and expression.

Figure S6 Correlation analysis for TFAP2P methylation locus and expression.