MEDICAL SCIENCE MONITOR

CLINICAL RESEARCH

e-ISSN 1643-3750
© Med Sci Monit, 2020; 26: e918906
DOI: 10.12659/MSM.918906

Received: 2019.07.21
Accepted: 2019.10.16
Published: 2020.01.01

Integrative Gene Expression Profiling Analysis to Investigate Potential Prognostic Biomarkers for Colorectal Cancer

Authors' Contribution:
Study Design A
Data Collection B
Statistical Analysis C
Data Interpretation D
Manuscript Preparation E
Literature Search F
Funds Collection G

ABCDEF 1 Xinkui Liu
C 2,3,4 Zhitong Bing
AG 1 Jiarui Wu
C 1 Jingyuan Zhang
C 1 Wei Zhou
C 1 Mengwei Ni
C 1 Ziqi Meng
C 1 Shuyu Liu
C 2,3 Jinhui Tian
C 1 Xiaomeng Zhang
C 5 Yingfei Li
C 1 Shanshan Jia
C 1 Siyu Guo

1 Department of Clinical Chinese Pharmacy, School of Chinese Materia Medica, Beijing University of Chinese Medicine, Beijing, P.R. China
2 Evidence Based Medicine Center, School of Basic Medical Science, Lanzhou University, Lanzhou, Gansu, P.R. China
3 Key Laboratory of Evidence Based Medicine and Knowledge Translation of Gansu Province, Lanzhou, Gansu, P.R. China
4 Institute of Modern Physics of Chinese Academy of Sciences, Lanzhou, Gansu, P.R. China
5 Center for Drug Metabolism and Pharmacokinetics (DMPK) Research of Herbal Medicines, Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing, P.R. China

Corresponding Author: Jiarui Wu, e-mail: exogamy@163.com
Source of support: The study was financially supported by National Natural Science Foundation of China (Grant nos. 81473547 and 81673829)

Background: Despite noteworthy advancements in the multidisciplinary treatment of colorectal cancer (CRC) and deeper understanding in the molecular mechanisms of CRC, many of CRC patients with histologically identical tumors present different treatment response and prognosis. Thus, more evidence on novel predictive and prognostic biomarkers for CRC remains urgently needed.
This study aims to identify potential prognostic biomarkers for CRC with integrative gene expression profiling analysis.

Material/Methods: Differential expression analysis of paired CRC and adjacent normal tissue samples in 6 microarray datasets was independently performed, and the 6 datasets were integrated by the robust rank aggregation method to detect consistent differentially expressed genes (DEGs). Aberrant expression patterns of these genes were further validated in RNA sequencing data. Then, gene set enrichment analysis (GSEA) was performed to investigate significantly dysregulated biological functions in CRC. Finally, univariate, LASSO and multivariate Cox regression models were built to identify key prognostic genes in CRC patients.

Results: A total of 990 DEGs (495 downregulated and 495 upregulated genes) were acquired after integratedly analyzing the 6 microarray datasets, and 4131 DEGs (2050 downregulated and 2081 upregulated genes) were obtained from the RNA sequencing dataset. Subsequently, these DEGs were intersected and 885 consistent DEGs were finally identified, including 458 downregulated and 427 upregulated genes. Two risky prognostic genes (*TIMP1* and *LZTS3*) and 5 protective prognostic genes (*AXIN2*, *CXCL1*, *ITLN1*, *CPT2* and *CLDN23*) were identified, which were significantly associated with the prognosis of CRC.

Conclusions: The 7 genes that we identified would provide more evidence for further applying novel diagnostic and prognostic biomarkers in clinical practice to facilitate personalized treatment of CRC.

MeSH Keywords: Biological Markers • Colorectal Neoplasms • Gene Expression Profiling • Prognosis

Full-text PDF: https://www.medscimonit.com/abstract/index/idArt/918906

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

Background

Colorectal cancer (CRC) is the third most commonly diagnosed malignancy and the second leading cause of cancer death in a global context [1]. The past decades have witnessed a remarkable decline in CRC incidence and mortality overall, and a dramatic rise in the median overall survival (OS) of metastatic colorectal cancer patients [2–9]. The exciting fact is ascribed to advances in comprehensive medical options, such as laparoscopic surgery, radiotherapy, neoadjuvant and palliative chemotherapies and targeted therapies, along with a deeper understanding of epidemiology, pathology and molecular mechanisms related to CRC [2,10]. Despite that, CRC, which accounts for almost one-tenth of cancer cases and deaths (with an estimated 1.8 million new cases and 881 000 deaths in 2018), contributes to high medical burden worldwide [1]. It has been well-known that many of CRC patients present discrepant treatment response and prognosis despite having histologically identical tumors, and thus personalized treatment based on biomarkers is likely to generate great clinical efficacy and public health significance, which not only enhances therapeutic effectiveness but also decreases treatment-related injury and costs [10,11]. Therefore, although the numerous molecular characterization, biological markers and therapeutic targets of CRC formerly discovered have greatly contributed to the diagnosis and treatment of this malignancy, more evidence on predictive and prognostic biomarkers is meaningful and urgently demanded in view of the biological complexity, worse outcome and high metastasis of this deadly disease [2,10,12].

Striking advancements in microarray and high-throughput sequencing technologies have facilitated the discovery of not only the crucial genetic or epigenetic alternations in carcinogenesis, tumor growth, metastasis and recurrence but also the promising cancer biomarkers for diagnosis, prognosis and treatment prediction [12–14]. Nevertheless, inconsistent results often occur due to sample heterogeneity in individual experiments or discrepancy in technological platforms [15]. Furthermore, application of relatively small sample size decreases statistical power, which blocks informative and useful findings [16–18]. To overcome the limitations and obtain convictive outcomes, integrated bioinformatics analysis, a comprehensive strategy to increase sample size, unify cross-platform standardization of expression profiles and discard invalid raw data, has been widely adopted to identify differentially expressed genes (DEGs) at mRNA and non-coding RNA level in CRC [16,19].

This study performed the integrative analysis for the gene expression patterns of 6 microarray datasets in Gene Expression Omnibus (GEO) via using the robust rank aggregation (RRA) method, aiming at discovering the consistent DEGs between human CRC and paired adjacent normal tissue samples. We further validated the aberrant expression patterns of these genes in the RNA sequencing data of the CRC patients from The Cancer Genome Atlas (TCGA). Additionally, we conducted gene set enrichment analysis (GSEA) to investigate significantly dysregulated biological functions in CRC. Finally, we constructed a gene signature with prognostic value in CRC patients through implementing univariate, LASSO and multivariate Cox regression analyses.

Material and Methods

Data collection and preprocessing

Six microarray-based gene expression data (GSE21510, GSE22598, GSE37182, GSE39582, GSE44076 and GSE89076) were accessed from Gene Expression Omnibus [20,21] (GEO; *http://www.ncbi.nlm.nih.gov/geo/*). All the included datasets met the following inclusion criteria: 1) they used colorectal tissues of CRC patients; 2) they included paired tumor and adjacent normal tissue samples and 3) the sample size of each dataset was at least 30. The samples included in this study came from these datasets and only the paired tumor and adjacent normal samples from colon tissues were used. The microarray data of GSE21510 (23 paired samples), GSE22598 (17 paired) and GSE39582 (17 paired) implemented the platform of Affymetrix Human Genome U133 Plus 2.0 Array. The platforms for GSE44076 (98 paired), GSE89076 (24 paired) and GSE37182 (82 paired) were Affymetrix Human Genome U219 Array, Agilent-039494 SurePrint G3 Human GE v2 8x60K Microarray 039381 and Illumina HumanHT-12 V3.0 expression beadchip, respectively. In total, 261 CRC and 261 matched normal cases were chosen for integrated analysis. The RNA sequencing data containing 398 colon adenocarcinoma and 39 normal samples were downloaded from The Cancer Genome Atlas (TCGA) (up to December 18, 2018, *http://cancergenome.nih.gov/*). Only protein-coding genes were eventually reserved for further study, and corresponding annotation information was accessed from Ensembl (*http://www.ensembl.org/index.html*). The AnnotationDbi [22] and org.Hs.eg.db [23] packages were applied to achieve conversion among gene symbol, Entrez ID and Ensembl ID. Furthermore, clinical information of 385 colon adenocarcinoma patients in TCGA was also downloaded, among whom 349 patients were reserved for further study. Thirty-six patients were excluded from our study for 3 reasons: 1) 12 patients for not having overall survival (OS) time, survival status, or pathological stage; 2) 20 patients for having an overall survival time shorter than 30 days and 3) 4 patients for lacking corresponding mRNA expression data.

Background correction, normalization, and expression calculation for the raw data (.cel format) of GSE21510, GSE22598, GSE39582 and GSE44076 (based on the Affymetrix platform) were conducted by the Robust Multi-array Average (RMA)

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

method [24,25] in the affy package [26]. The marray package [27] and the neqc function in the limma package [28,29] were used for preprocessing the raw data of the Agilent (GSE89076) and Illumina (GSE37182) microarray platforms, respectively. Annotation files for probes in the different datasets were downloaded from the GEO database. If multiple probes were mapped to one same gene, the average expression value of the different probes represented the final expression level of this gene. Moreover, conversion among gene symbol, Entrez ID and Ensembl ID was also achieved by the AnnotationDbi and org.Hs.eg.db packages.

Differentially expressed genes (DEGs) screening

For each of the 6 microarray datasets, gene expression difference between the tumor and adjacent noncancerous tissues were calculated by the limma package. Then, the integration for the genes in every list was conducted by the RobustRankAggreg package [30], which was based on the robust rank aggregation (RRA) method. This rank aggregation approach detects genes that are ranked consistently better than expected under null hypothesis of randomly ordered input lists and assigns a *P* value for each gene. Bonferroni correction was also employed in case of false positive results, and genes meeting the criterion of $|\log_2$ fold change (FC)| >1 and adjust $P<0.05$ were taken as DEGs.

For the mRNA sequencing data from TCGA, protein-coding genes with counts >1 in more than 75% samples were retained, and duplicate gene expression values were averaged. Expression calculation, normalization and DEGs screening were carried out by edgeR [31], with $|\log_2$FC| >1 and false discovery rate (FDR) <0.05 as the threshold. The impute package [32] was used to fill missing values of the normalized expression data. The consistent DEGs in the 6 microarray profiles were intersected with the DEGs in the TCGA dataset by Entrez ID, and the eventually consistent DEGs between the microarray and sequencing data were reserved for further study. Moreover, the expression values of the eventually consistent DEGs in the TCGA colon adenocarcinoma dataset were log2 transformed before the following analysis.

Gene set enrichment analysis (GSEA)

To identify significantly dysregulated biological pathways in CRC, the GSEA [33] was performed by clusterProfiler [34], under functional annotations of the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (*https://www.genome.jp/kegg/*). Entrez IDs and corresponding $\log_2$FC values of all the genes in each dataset were submitted to clusterProfiler, with the permutation number and the minimum gene set size set as 100 000 and 120, respectively. Activated and suppressed pathways with adjust $P<0.05$ in each dataset were merged, and ones with higher frequency (found in ≥3 datasets) were identified as dramatically changed biological functions in CRC.

Survival analysis

For the included 349 CRC patients, survival time, status, and mRNA expression levels of the consistent DEGs were applied for survival analysis. Firstly, a univariate Cox proportional hazards regression model was built for preliminarily screening OS-related genes, and the genes with $P<0.05$ were considered statistically significant. Secondly, a least absolute shrinkage and selection operator (LASSO) Cox regression model was adopted to further select key genes from significant ones in univariate analysis. The glmnet package [35] was utilized to perform the LASSO Cox analysis. The maximum number of replacements was set as 100 000 times, and a sequence of tuning parameters (lambdas, λs) were returned according to the expected generalization error estimated from 10-fold cross-validation. The lambda with minimum mean cross-validated error (lambda.min) was employed. Finally, a multivariate Cox proportional hazards regression model was established to estimate the contribution of a gene as an independent prognostic factor for patient survival. The optimal model was selected by the Akaike information criterion (AIC) method, and thereby a prognostic gene signature was established. The univariate and multivariate Cox regression analyses were all conducted by the survival package [36]. A prognosis risk score was calculated based on a linear combination of the expression value of the gene in this prognostic signature multiplied by its regression coefficient derived from the multivariate Cox model. The formula is as follows:

$$\text{Risk score} = \sum_{i=1}^{n} (exp_i \times coef_i)$$

Where n is the number of genes, *expi* is the expression value of the *i*th variable and *coefi* is the regression coefficient of the *i*th variable. These 349 patients were categorized into either low-risk or high-risk group based on the median prognostic risk score. The Kaplan-Meier method with the log-rank test was used to assess the correlation between the risk and OS, and the survival curve was generated by the survminer package [37]. The time-dependent receiver operating characteristic (ROC) curve analysis was conducted by the survivalROC package [38], and the area under the curve (AUC) was calculated to measure the predictive accuracy of this prognostic signature for time-dependent cancer death. All the statistical analyses were performed with R (version 3.5.2, *https://www.r-project.org/*) in this study.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

Table 1. Clinical information for the included 349 patients.

Characteristics	Number of cases (%)	
Gender		
Male	189	(54.2)
Female	160	(45.8)
Age		
≤60	105	(30.1)
>60	244	(69.9)
TNM stage		
Stage I	62	(17.8)
Stage II	138	(39.5)
Stage III	99	(28.4)
Stage IV	50	(14.3)
T stage		
Tis	1	(0.3)
T1	7	(2.0)
T2	62	(17.8)
T3	242	(69.3)
T4	37	(10.6)
M stage		
M0	266	(76.2)
M1	50	(14.3)
MX	31	(8.9)
Not reported	2	(0.6)
N stage		
N0	207	(59.3)
N1	83	(23.8)
N2	59	(16.9)
Vital status		
Alive	279	(79.9)
Dead	70	(20.1)

Results

Identification of DEGs

The clinical information for the CRC patients included in the present study is shown in Table 1, Supplementary Tables 1–3. We obtained 990 DEGs (495 downregulated and 495 upregulated genes) after performing the integrated analysis of the 6 microarray datasets (Figures 1A, 2A–2F, Supplementary Tables 4, 5), and we also acquired 4131 DEGs (2050 downregulated and 2081 upregulated genes) from the TCGA colon cancer dataset (Figures 1B, 2G, and Supplementary Table 6). Subsequently, we intersected these DEGs and finally identified 885 consistent DEGs, including 458 downregulated and 427 upregulated genes (Figure 1C, 1D, and Supplementary Table 7).

Identification of dysregulated pathways

According to the results of GSEA (Figures 3, 4 and Supplementary Table 8, 9), 24 pathways (including 5 activated and 19 suppressed) found in more than or equal to 4 datasets were identified as significantly dysregulated biological pathways in CRC. Eight suppressed pathways existed in all the 7 datasets, namely, adrenergic signaling in cardiomyocytes, apelin signaling pathway, calcium signaling pathway, cAMP signaling pathway, cGMP-PKG signaling pathway, neuroactive ligand-receptor interaction, Rap1 signaling pathway, and regulation of actin cytoskeleton. The top 3 activated pathways were cell cycle, RNA transport, and Wnt signaling pathway, which were respectively found in 7, 7, and 5 datasets, respectively. Among the 24 significantly changed pathways, it has long been known that the cell cycle, Ras signaling pathway and Wnt signaling pathway play important roles in the initiation and progression of CRC [12,39].

Survival analysis

We performed the univariate Cox regression to investigate the correlation of the DEGs with OS of CRC patients, and identified 101 OS-related genes with P was <0.05 (Supplementary Table 10). Then, in order to further narrow genes, we employed the LASSO Cox model with 10-fold cross-validation and 100 000 repetitions to acquire optimal penalty parameters. As a result, 22 genes were identified when we chose the minimum criteria where the log (λ)=–3.52 with λ=0.02957 (Figure 5). Finally, we developed a 7-gene prognostic signature after performing the multivariate Cox analysis, which was composed of TIMP metallopeptidase inhibitor 1 (*TIMP1*), Axin 2 (*AXIN2*), C-X-C motif chemokine ligand 1 (*CXCL1*), leucine zipper tumor suppressor family member 3 (*LZTS3*), intelectin 1 (*ITLN1*), carnitine palmitoyltransferase 2 (*CPT2*) and claudin 23 (*CLDN23*) (Figures 6A, 7A). As shown in Figure 7B, *TIMP1*, *AXIN2*, *CXCL1* and *LZTS3* were upregulated, whereas *ITLN1*, *CPT2* and *CLDN23* were downregulated in CRC compared with normal groups. Moreover, lower expression of *CXCL1* and *CPT2* was shown to be associated with advanced tumor stage (Kruskal-Wallis test P<0.05, Figure 7C, 7D), while the correlation of the other 5 genes with pathological stage was not statistically significant. Among these 7 genes, *AXIN2*, *CXCL1*, *ITLN1*, *CPT2*, and *CLDN23* with HR<1 were identified as protective prognostic genes, whereas *TIMP1* and *LZTS3* with HR>1 were identified as risky prognostic genes. The regression coefficient for each gene was also generated, and the survival risk score was calculated as follows: risk score=(0.3259×expression level of *TIMP1*)+(–0.2607×expression level of *AXIN2*)+(–0.1289×expression level of *CXCL1*)+(0.4504×expression level of *LZTS3*)+(–0.0619×expression level of *ITLN1*)+(–0.7526×expression

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

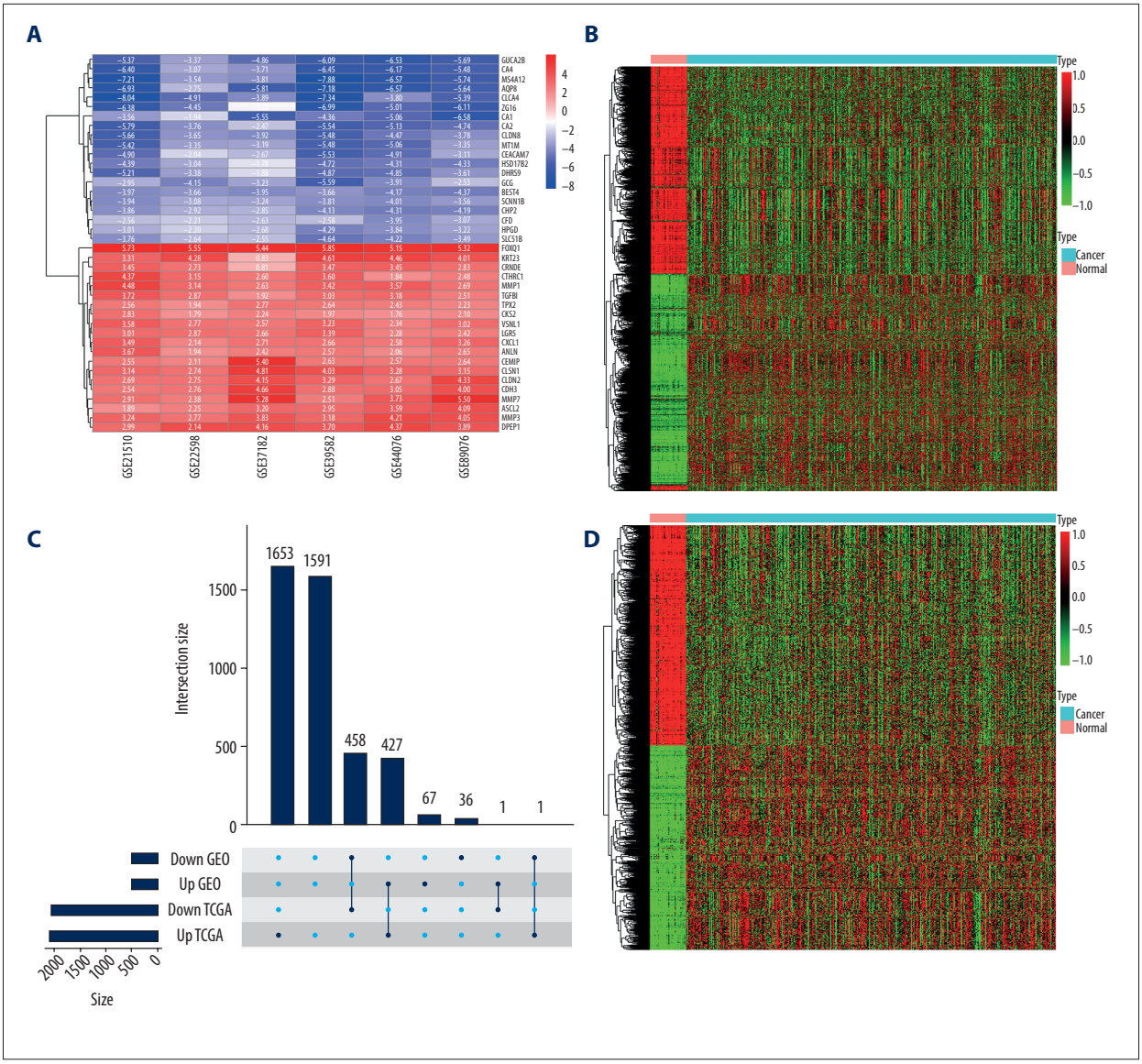


Figure 1. Identification of differentially expressed genes (DEGs). (**A**) The heatmap of top 20 downregulated and upregulated DEGs identified by the integrated analysis of the 6 microarray datasets. Each column represents 1 dataset and each row represents 1 gene. The number in each rectangle represents the value of log_2FC. The gradual color ranging from blue to red represents the changing process from downregulation to upregulation. (**B**) The heatmap of the 4131 DEGs in The Cancer Genome Atlas (TCGA) colorectal cancer (CRC) dataset. Each column represents 1 sample and each row represents 1 gene. The gradual color ranging from green to red represents the changing process from downregulation to upregulation. (**C**) The Venn diagram of the DEGs between the integrated Gene Expression Omnibus (GEO) dataset and the TCGA CRC dataset. (**D**) The heatmap of the 885 consistent DEGs (using the TCGA dataset). Each column represents 1 sample and each row represents 1 gene. The gradual color ranging from green to red represents the changing process from downregulation to upregulation.

level of *CPT2*)+(–0.4304×expression level of *CLDN23*). The 174 patients with risk scores higher than the median risk score (1.0048) were included into the high-risk group, whereas the rest 175 patients were included into the low-risk group (Figure 6B). The Kaplan-Meier survival analysis showed that patients in the high-risk group had shorter survival time and more deaths compared with patients in the low-risk group (Log-rank test $P<0.0001$), suggesting expression levels of these 7 genes could effectively distinguish the high-risk and low-risk of these colon cancer patients (Figure 6C). The AUC of the time-dependent ROC curve was 0.738, 0.769, and 0.851 for 1-year, 3-year, and 5-year OS, respectively, confirming the good prediction accuracy of this prognostic gene signature (Figure 6D). The nomogram for survival time prediction of CRC patients is shown in ure 8.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

Discussion

Integrated bioinformatics analysis of CRC gene expression profiles and construction of gene signatures associated with CRC prognosis have aroused extensive attention recently. For example, Sun et al. identified 352 overlapping DEGs in 5 GEO datasets which totally included 207 CRC and matched normal samples and proposed a 5-gene prognostic signature using Cox regression models [40]. Chen et al. detected a 7-gene signature that can predict OS of CRC patients by employing Cox regression analysis combined with a robust likelihood-based survival modeling approach [11]. Xiong et al. extracted expression data of mRNAs, miRNAs, and lncRNAs in TCGA, and built a multi-RNA-based classifier for CRC patient stratification by Cox survival analysis and Lasso regression [41]. Dai et al. also used Lasso Cox regression modeling and developed a robust 15-mRNA prognostic signature from GSE39582 for predicting early relapse in stage I–III colon cancer patients [42]. As for the

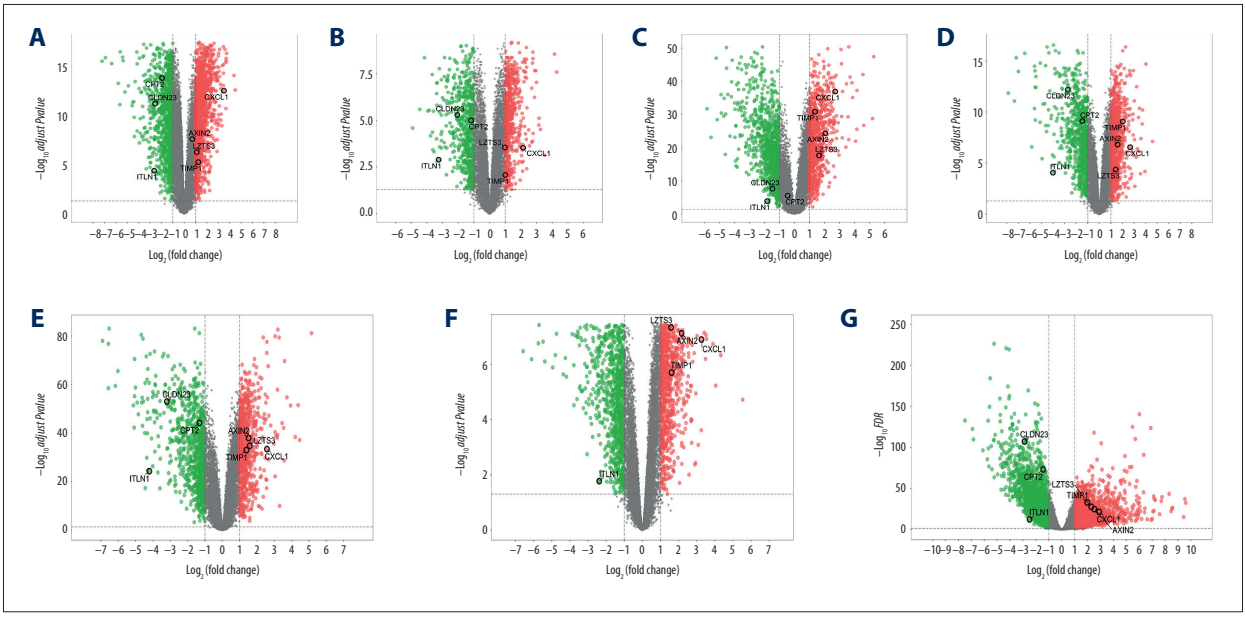
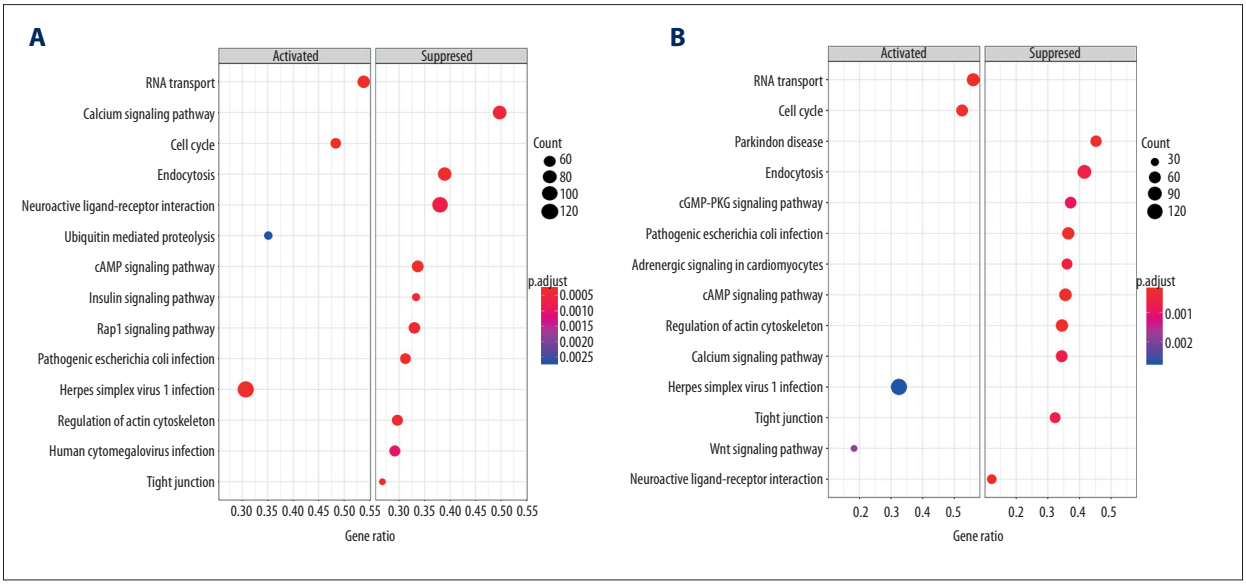


Figure 2. The volcano plot of the genes in the 7 datasets. (**A**) The volcano plot of the genes in GSE21510. (**B**) The volcano plot of the genes in GSE22598. (**C**) The volcano plot of the genes in GSE37182. (**D**) The volcano plot of the genes in GSE39582. (**E**) The volcano plot of the genes in GSE44076. (**F**) The volcano plot of the genes in GSE89076. (**G**) The volcano plot of the genes in The Cancer Genome Atlas (TCGA) dataset. The red dot represents the genes with adjust $P<0.05$ and $\log_2FC >1$, and the green dot represents the genes with adjust $P<0.05$ and $\log_2FC <-1$.



Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

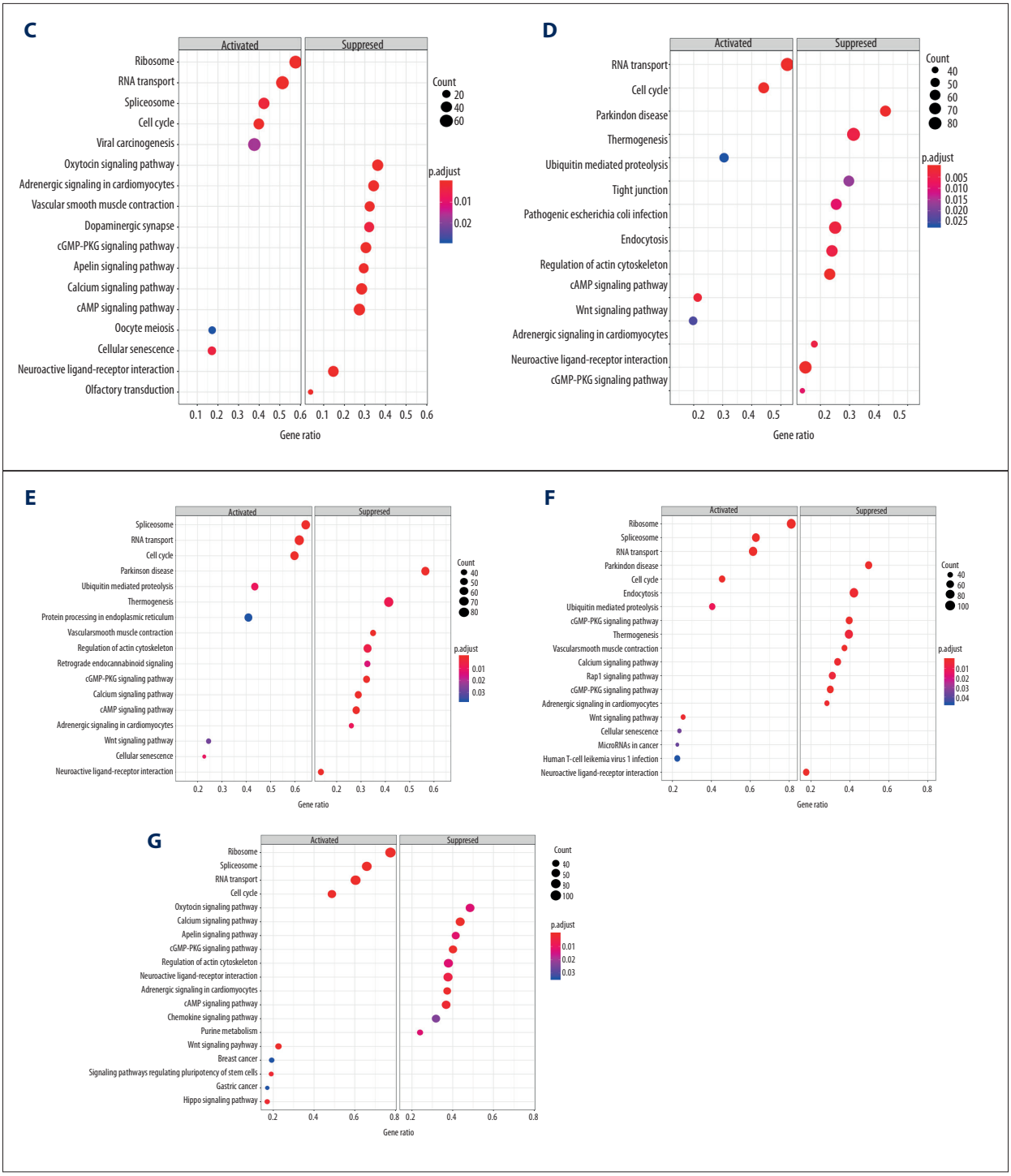


Figure 3. The gene set enrichment analysis (GSEA) for the 7 colorectal cancer (CRC) datasets. (**A**) The GSEA for GSE21510. (**B**) The GSEA for GSE22598. (**C**) The GSEA for GSE37182. (**D**) The GSEA for GSE39582. (**E**) The GSEA for GSE44076. (**F**) The GSEA for GSE89076. (**G**) The GSEA for The Cancer Genome Atlas (TCGA) dataset. The top 20 suppressed and activated Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways in each dataset were shown. The y-axis shows the KEGG pathway terms, and the x-axis shows the gene ratio of each term.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

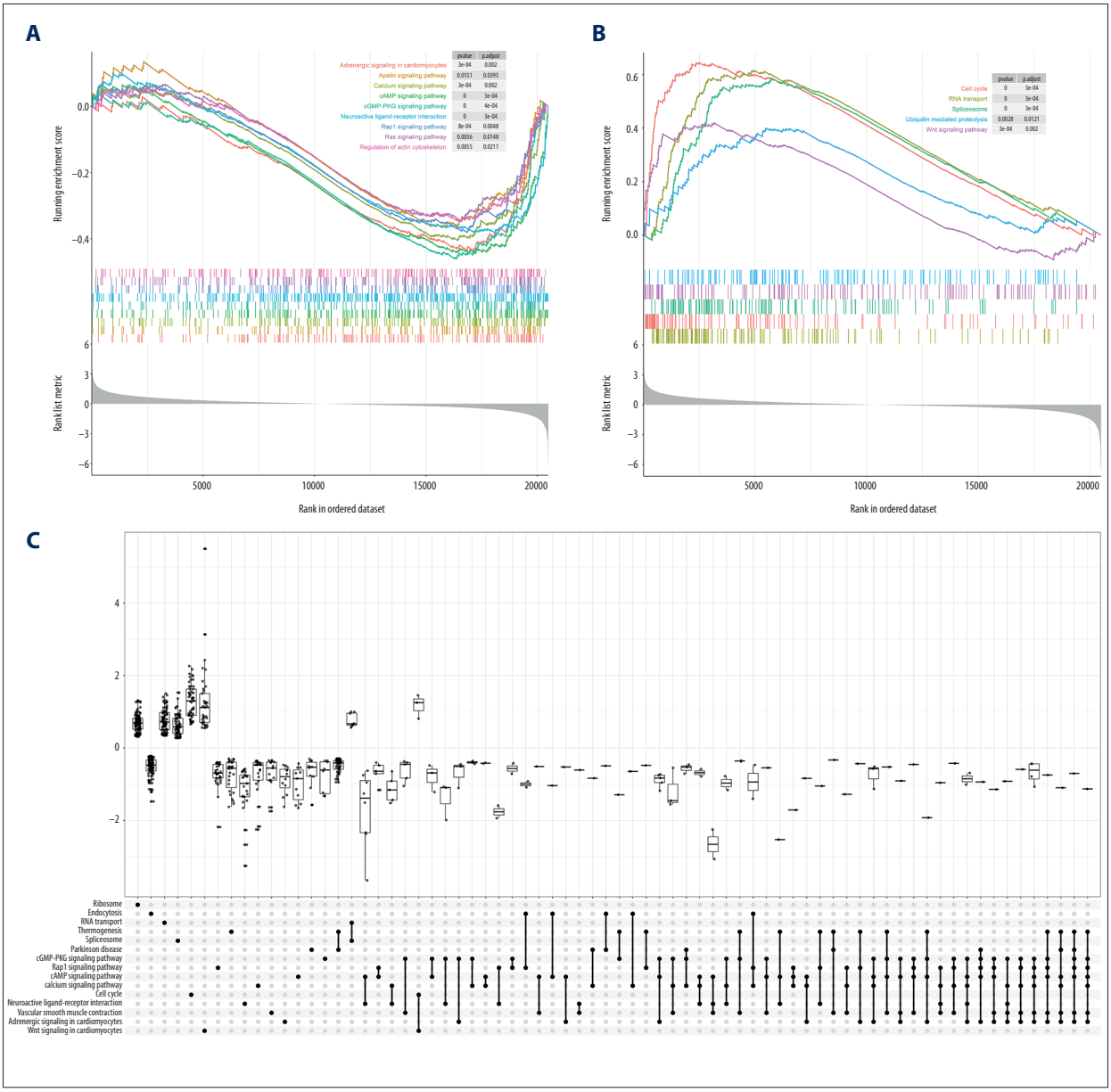


Figure 4. The enrichment plot of the gene set enrichment analysis (GSEA) (using GSE89076). (**A**) The enrichment plot of 9 suppressed pathways. (**B**) The enrichment plot of 5 activated pathway. (**C**) The upSet plot for the GSEA.

present study, we used the raw data of 6 whole genome platform-based microarray datasets with paired tumor and non-cancerous samples and conducted corresponding normalization for them to make these data more comparable. Meanwhile, we applied the RRA approach to integrate the shared DEGs across the 6 datasets, making the results more reliable than only intersecting DEGs of different expression profiles. Moreover, to detect significantly changed biological functions in CRC, we performed GSEA for each CRC dataset and the pathways found in more than 4 datasets were taken into consideration. Ultimately, we integrated univariate, LASSO and multivariate Cox regression models to identify key prognostic genes in CRC patients.

In the current study, we detected 990 common DEGs between 261 CRC and matched normal tissues in 6 microarray datasets, 885 of which were validated thorough TCGA. When conducting the GSEA, we identified 22 significantly dysregulated biological pathways in CRC. The univariate and LASSO Cox regression models selected 22 survival-related genes, and a 7-gene signature with prognostic value in CRC was finally established by the multivariate Cox analysis. The 7-gene prognostic signature consisted of 2 risky prognostic genes (*TIMP1* and *LZTS3*) and 5 protective prognostic genes (*AXIN2*, *CXCL1*, *ITLN1*, *CPT2*, and *CLDN23*). Among them *TIMP1*, *AXIN2*, *CXCL1* and *LZTS3* were upregulated, whereas *ITLN1*, *CPT2*, and *CLDN23* were

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

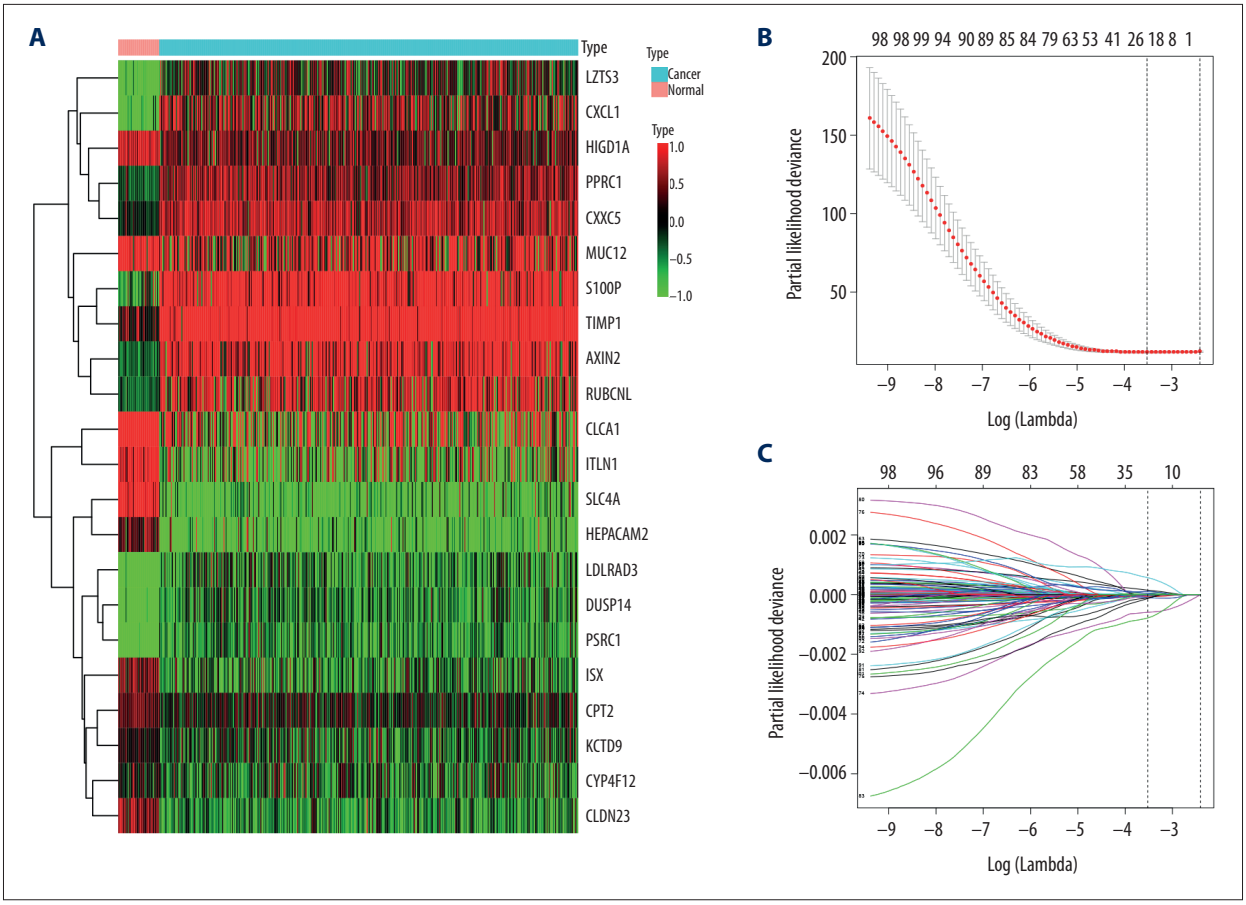


Figure 5. Gene selection through the least absolute shrinkage and selection operator (LASSO) Cox regression model. (**A**) The heatmap of the 22 differentially expressed genes (DEGs) identified by the LASSO Cox regression model. (**B**) Ten-fold cross-validation for tuning parameter (λ) selection in the LASSO Cox regression model. The vertical lines were drawn at the optimal values by the minimum criteria and the 1-SE criteria. (**C**) The LASSO coefficient profiles of the 101 DEGs.

downregulated in CRC compared with normal groups according to our bioinformatics analysis. For the 2 risky prognostic genes, the prognostic value of *TIMP1* in CRC has been confirmed in former works, while that of *LZTS3* has not. TIMP-1 is among human natural endogenous inhibitors of matrix metalloproteinases (MMPs). It has been acknowledged that MMPs, a group of proteolytic enzymes, play an important role in the degradation of extracellular matrix (ECM) components, which is critical for tumor growth, invasion and metastasis [43]. In addition to its function as an inhibitor of MMPs, TIMP-1 can stimulate cell proliferation, induce anti-apoptotic signaling and influence angiogenesis in an MMP-independent manner [44–47]. Increasing evidence, especially from meta-analysis, has shown that TIMP-1 has potential diagnostic and prognostic value in CRC, and elevated TIMP-1 may predicts shorter OS among patients with no systemic inflammatory response [48–54]. Consistent with these reports, our study also found that *TIMP1* is upregulated in CRC patients and severs as a risky prognostic gene. Members of the leucine zipper tumor suppressor (LZTS) protein family are thought to play roles in cell growth modulation [55]. A past *in silico* work presented that LZTS3, a member of this protein family, served as a potential tumor suppressor [55]. A latest study showed that highly expressed miR-1275 could promote proliferation and metastasis of non-small cell lung cancer thorough targeting LZTS3 [56]. However, much less is known about the function of *LZTS3* in CRC.

Regarding the 5 protective prognostic genes, the prognostic value of *CXCL1*, *ITLN1*, *CPT2* and *CLDN23* in CRC have been reported, while that of *AXIN2* has not been totally elucidated. CXCL1, a chemotactic cytokine, involves in cancer progression and invasion [57]. Highly elevated CXCL1 expression is found in CRC, promoting tumorigenicity, progression and metastasis [57–61], and higher CXCL1 expression is correlated with larger tumor size and later tumor stage [57]. Recent researches showed that CXCL1 serves as an independent adverse prognostic biomarker in CRC patients, and it might be a novel biomarker and potential therapeutic target for CRC treatment [57,61]. In contrast, our results showed the

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

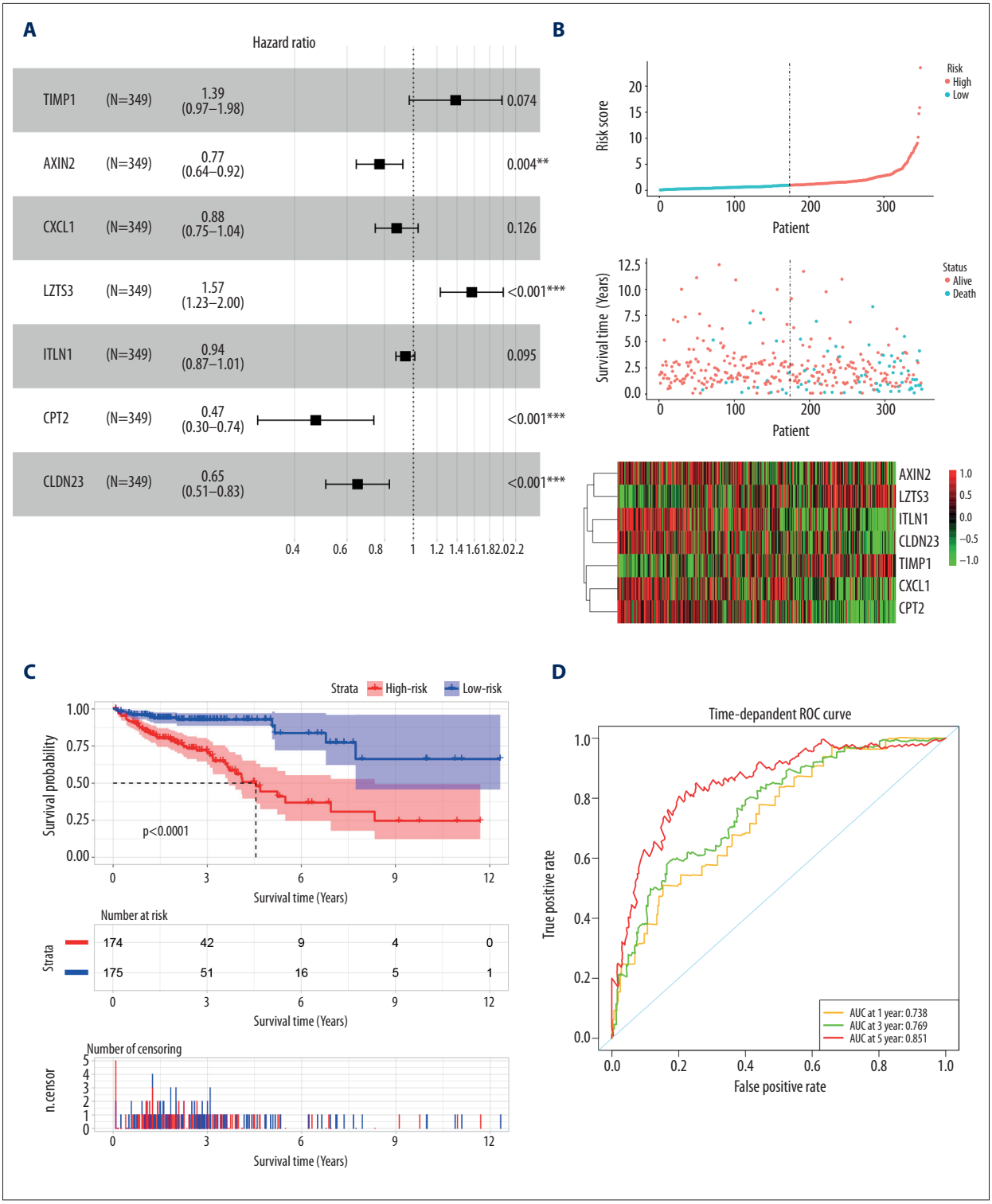


Figure 6. Construction of the 7-gene signature with prognostic value. (**A**) The forest plot of the 7 genes identified by the multivariate Cox regression analysis. (**B**) The characteristics of the patients order by their risk score. Dotted line: the median risk score (1.0048). From top to bottom is the risk score, patients' survival status distribution and heatmap of the 7 genes for patients in the low- and high-risk groups. (**C**) The Kaplan-Meier survival curve for patients in the low- and high- risk groups. (**D**) The time-dependent receiver operating characteristic (ROC) curve for predicting overall survival (OS) in colorectal cancer (CRC) patients by the risk score.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

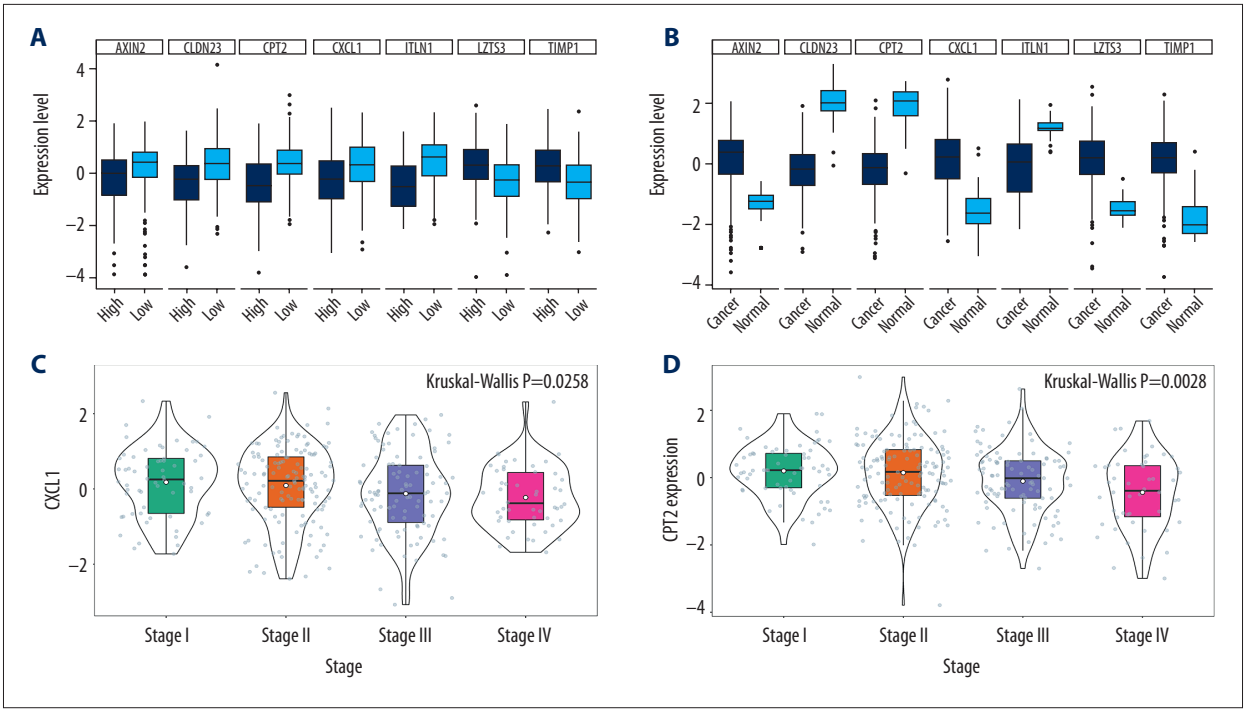


Figure 7. The expression level distribution of the 7 genes. (**A**) The expression level of the 7 genes between the low-risk and high-risk groups. (**B**) The expression level of the 7 genes between the normal and tumor groups. (**C**) The correlation of *CXCL1* expression with pathological stage. (**D**) The correlation of *CPT2* expression with pathological stage.

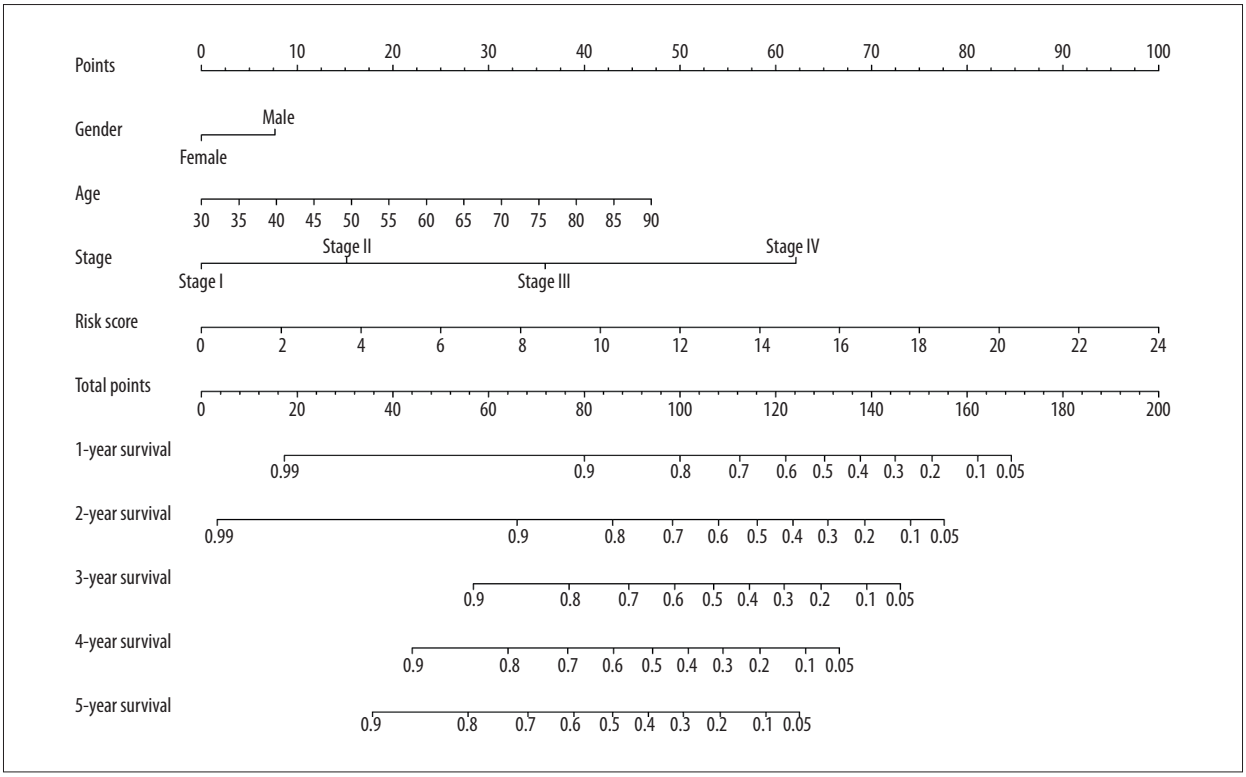


Figure 8. The nomogram for the survival time prediction of the colorectal cancer (CRC) patients.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

correlation of higher *CXCL1* expression with lower tumor stage in CRC and that high level of *CXCL1* predicts better outcome in CRC. The difference may derive from population heterogeneity and small sample size, and thus large-scale multi-center clinical research studies are needed due to limited evidence on the prognostic value of *CXCL1* in patients with CRC. Intelectin-1 (also known as omentin-1), encoded by the *ITLN1* gene, is reported as a protein that possess metabolic, inflammatory, and immune-related properties, and thereby might be correlated with CRC risk [62–66]. A previous research presented that high intelectin-1 expression is closely associated with favorable prognosis in gastric cancer patients [67]. As for CRC, our findings identified *ITLN1* as a protective prognostic gene. Likewise, Kim et al. reported that intelectin-1 predicts better prognosis in stage IV CRC [68]. These findings support the functions of *ITLN1* as a potential tumor suppressor in gastrointestinal cancers. Conversely, a prospective cohort study presented that higher circulating intelectin-1 concentrations were related to a higher CRC risk [62]. Since whether *ITLN1* is a tumor suppressor or promoter in colorectal carcinogenesis has not been absolutely clarified [69], the prognostic value of *ITLN1* should be highly valued and deserves deeper investigation. CPT2, the key enzyme in fatty acid oxidation, locating on the mitochondrial membrane [70]. Consistent with our study, decreased expression of CPT2 was detected in CRC tissue [70,71], and higher expression of CPT2 in cancer tissue as an independent prognostic factor predicts better prognosis in CRC patients [70]. The *CLDN23* gene encodes a member of the claudin family, and claudins are known to be crucial in cancer growth and progression [72,73]. It has been reported that *CLDN23* expression is significantly reduced in CRC tissue and lower expression of this gene correlates with shorter OS rates in CRC patients [74–76], which is consistent with our finding that *CLDN23* could serve as a protective prognostic factor. Furthermore, *CLDN23* expression is shown to be epigenetically regulated, and disruption of bivalent histone modifications at the CLDN23 locus probably result in remarkably reduced *CLDN23* expression in CRC tissue [74]. As for the *AXIN2* gene, both germline and somatic mutations in this gene were found in CRC [77]. The AXIN2 protein, acting as an essential scaffold to help assemble the β-catenin destruction complex, negatively regulates β-catenin-dependent Wnt signaling, the well-known pathway that is critical in initiation and progression of CRC and is featured by accumulation of genetic and epigenetic changes [12,77,78]. Meanwhile, AXIN2 is a transcriptional target of β-catenin-dependent Wnt signaling [79–81], and highly expressed AXIN2 is found in malignancies with activating Wnt pathway mutations [77]. Give that AXIN2 is not only a β-catenin downstream target but also a key negative feedback regulator of Wnt signaling with induction of β-catenin degradation, AXIN2 has long been hypothesized as a potential tumor suppressor [77,82]. However, the prognostic value of *AXIN2* in CRC has hardly been reported.

In the current study, the gene expression data we used for the integrative analysis were generated from different institutions and accessed from publicly available databases, so we cannot guarantee the quality of these data. Furthermore, the influence of the detailed features such as age, gender, race, tumor grade and stage on gene expression patterns was not considered because our study solely focused on genes consistently identified as significantly altered ones in different researches, which makes some biological information overlooked in our study. Ultimately, given that our findings came from the comprehensive *in silico* research, additional results from biological experiments and large-scale multi-center clinical research studies will be pivotal for supporting our findings.

Conclusions

In conclusion, we identified 7 potential prognostic biomarkers for CRC by performing the integrative analysis of the gene expression profiles of microarray and RNA sequencing. Our findings would provide more evidence for further applying novel diagnostic and prognostic biomarkers in clinical practice to facilitate the personalized treatment of CRC. Meanwhile, further biological experiments and large-scale multi-center clinical research studies are required to validate our results since our study was conducted based on data analysis.

Supplementary Data

Supplementary Table 1. The information for the samples in the 6 Gene Expression Omnibus (GEO) datasets.

Supplementary Table 2. The information for the colon adenocarcinoma patients in The Cancer Genome Atlas (TCGA).

Supplementary Table 3. The information for the 349 included colon adenocarcinoma patients in The Cancer Genome Atlas (TCGA).

Supplementary Table 4. The information for the 6 gene lists of the 6 Gene Expression Omnibus (GEO) datasets generated by the limma package.

Supplementary Table 5. The information for the 990 differentially expressed genes (DEGs) identified by the integrated analysis of the 6 Gene Expression Omnibus (GEO) datasets ($|\log_2 FC| > 1$ and adjust $P<0.05$).

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

Supplementary Table 6. The information for the 4131 differentially expressed genes (DEGs) of The Cancer Genome Atlas (TCGA) colon adenocarcinoma dataset ($|\log_2 FC| > 1$ and adjust $P<0.05$).

Supplementary Table 7. The information for the 885 consistent differentially expressed genes (DEGs).

Supplementary Table 8. The results of the gene set enrichment analysis (GSEA) for the 7 colorectal cancer (CRC) datasets (adjust $P<0.05$).

Supplementary Table 9. The 24 dysregulated biological pathways in colorectal cancer (CRC).

Supplementary Table 10. The information for the 101 overall survival (OS)-related genes identified by the univariate Cox regression analysis ($P<0.05$).

Supplementary/raw data available from the corresponding author on request.

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. Cancer J Clin, 2018; 68(6): 394–424
2. Kuipers EJ, Grady WM, Lieberman D et al: Colorectal cancer. Nat Rev Dis Primers, 2015; 1: 15065
3. Cremolini C, Loupakis F, Antoniotti C et al: FOLFOXIRI plus bevacizumab versus FOLFIRI plus bevacizumab as first-line treatment of patients with metastatic colorectal cancer: Updated overall survival and molecular subgroup analyses of the open-label, phase 3 TRIBE study. Lancet Oncol, 2015; 16(13): 1306–15
4. Giantonio BJ, Catalano PJ, Meropol NJ et al: Bevacizumab in combination with oxaliplatin, fluorouracil, and leucovorin (FOLFOX4) for previously treated metastatic colorectal cancer: Results from the Eastern Cooperative Oncology Group Study E3200. J Clin Oncol, 2007; 25(12): 1539–44
5. Hoff PM, Ansari R, Batist G et al: Comparison of oral capecitabine versus intravenous fluorouracil plus leucovorin as first-line treatment in 605 patients with metastatic colorectal cancer: results of a randomized phase III study. J Clin Oncol, 2001; 19(8): 2282–92
6. Hurwitz H, Fehrenbacher L, Novotny W et al: Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. N Engl J Med, 2004; 350(23): 2335–42
7. Van Cutsem E, Cervantes A, Nordlinger B, Arnold D, ESMO Guidelines Working Group: Metastatic colorectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol, 2014; 25(Suppl. 3): iii1–9
8. Venook A, Niedzwiecki D, Lenz HJ et al: CALGB/SWOG 80405: Phase III trial of irinotecan/5-FU/leucovorin (FOLFIRI) or oxaliplatin/5-FU/leucovorin (mFOLFOX6) with bevacizumab (BV) or cetuximab (CET) for patients (pts) with KRAS wild-type (wt) untreated metastatic adenocarcinoma of the colon or rectum (MCRC). J Clin Oncol, 2014; 32(18 Suppl.): LBA3
9. Volker H, Ludwig Fischer VW, Thomas D et al: FOLFIRI plus cetuximab versus FOLFIRI plus bevacizumab as first-line treatment for patients with metastatic colorectal cancer (FIRE-3): A randomised, open-label, phase 3 trial. Lancet Oncol, 2014; 15(10): 1065–75
10. Schirripa M, Lenz HJ: Biomarker in colorectal cancer. Cancer J, 2016; 22(3): 156–64
11. Chen H, Sun X, Ge W et al: A seven-gene signature predicts overall survival of patients with colorectal cancer. Oncotarget, 2016; 8(56): 95054–65
12. Cancer Genome Atlas Network: Comprehensive molecular characterization of human colon and rectal cancer. Nature, 2012; 487(7407): 330–37
13. Kulasingam V, Diamandis EP: Strategies for discovering novel cancer biomarkers through utilization of emerging technologies. Nat Clin Pract Oncol, 2008; 5(10): 588–99
14. Nannini M, Pantaleo MA, Maleddu A et al: Gene expression profiling in colorectal cancer using microarray technologies: Results and perspectives. Cancer Treat Rev, 2009; 35(3): 201–9
15. Cardoso J, Boer J, Morreau H, Fodde R: Expression and genomic profiling of colorectal cancer. Biochim Biophys Acta, 2007; 1775(1): 103–37
16. Chan SK, Griffith OL, Tai IT, Jones SJ: Meta-analysis of colorectal cancer gene expression profiling studies identifies consistently reported candidate biomarkers. Cancer Epidemiol Biomarkers Prev, 2008; 17(3): 543–52
17. Shangkuan WC, Lin HC, Chang YT et al: Risk analysis of colorectal cancer incidence by gene expression analysis. Peer J, 2017; 5: e3003
18. Yang J, Han S, Huang W et al: A meta-analysis of microRNA expression in liver cancer. PLoS One, 2014; 9(12): e114533
19. Sun M, Song H, Wang S et al: Integrated analysis identifies microRNA-195 as a suppressor of Hippo-YAP pathway in colorectal cancer. J Hematol Oncol, 2017; 10(1): 79
20. Ron E, Michael D, Alex EL: Gene expression omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res, 2002; 30(1): 207–10
21. Barrett T, Wilhite SE, Ledoux P et al: NCBI GEO: Archive for functional genomics data sets – update. Nucleic Acids Res, 2013; 41(Database issue): D991–95
22. AnnotationDbi: Annotation Database Interface [computer program]. R package version 1.44.0. *http://bioconductor.org/packages/release/bioc/html/AnnotationDbi.html2018*
23. org.Hs.eg.db: Genome wide annotation for Human [computer program]. R package version 3.7.0. *http://www.bioconductor.org/packages/release/data/annotation/html/org.Hs.eg.db.html2018*
24. Bolstad BM, Irizarry RA, Astrand M, Speed TP: A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics, 2003; 19(2): 185–93
25. Irizarry RA, Bridget H, Francois C et al: Exploration, normalization, and summaries of high-density oligonucleotide array probe level data. Biostatistics, 2003; 4(2): 249–64
26. Gautier L, Cope L, Bolstad BM, Irizarry RA: affy – analysis of Affymetrix GeneChip data at the probe level. Bioinformatics, 2004; 20(3): 307–15
27. marray: Exploratory analysis for two-color spotted microarray data [computer program]. R package version 1.60.0. *http://www.maths.usyd.edu.au/u/jeany/2018*
28. Smyth GK: limma: Linear models for microarray data. Bioinformatics & Computational Biology Solutions Using R & Bioconductor, 2011; 397–420
29. Ritchie ME, Phipson B, Wu D et al: limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res, 2015; 43(7): e47
30. Kolde R, Laur S, Adler P, Vilo J: Robust rank aggregation for gene list integration and meta-analysis. Bioinformatics, 2012; 28(4): 573–80
31. Robinson MD, McCarthy DJ, Smyth GK: edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics, 2010; 26(1): 139–40
32. impute: impute: Imputation for microarray data [computer program]. R package version 1.56.0. *http://bioconductor.org/packages/release/bioc/html/impute.html2018*
33. Subramanian A, Tamayo P, Mootha VK et al: Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA, 2005; 102(43): 15545–50
34. Yu G, Wang LG, Han Y, He QY. clusterProfiler: An R package for comparing biological themes among gene clusters. OMICS, 2012; 16(5): 284–87
35. Friedman J, Hastie T, Tibshirani R: Regularization paths for generalized linear models via coordinate descent. J Stat Softw, 2010; 33(1): 1–22
36. A Package for Survival Analysis in S [computer program]. version 2.38. *https://CRAN.R-project.org/package=survival2015*
37. survminer: Drawing Survival Curves using 'ggplot2' [computer program]. Version 0.4.3. *https://cran.r-project.org/web/packages/survminer/index.html2018*
38. survivalROC: Time-dependent ROC curve estimation from censored survival data [computer program]. Version 1.0.3. *https://cran.r-project.org/web/packages/survivalROC/index.html2013*

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

39. Fearon ER. Molecular genetics of colorectal cancer. Annu Rev Pathol, 2011; 6: 479–507
40. Sun G, Li Y, Peng Y et al: Identification of a five-gene signature with prognostic value in colorectal cancer. J Cell Physiol, 2019; 234(4): 3829–36
41. Xiong Y, Wang R, Peng L et al: An integrated lncRNA, microRNA and mRNA signature to improve prognosis prediction of colorectal cancer. Oncotarget, 2017; 8(49): 85463–78
42. Dai W, Li Y, Mo S et al: A robust gene signature for the prediction of early relapse in stage I–III colon cancer. Mol Oncol, 2018; 12(4): 463–75
43. Herszenyi L, Hritz I, Lakatos G et al: The behavior of matrix metalloproteinases and their inhibitors in colorectal cancer. Int J Mol Sci, 2012; 13(10): 13240–63
44. Moller Sorensen N, Vejgaard Sorensen I, Ornbjerg Wurtz S et al: Biology and potential clinical implications of tissue inhibitor of metalloproteinases-1 in colorectal cancer treatment. Scand J Gastroenterol, 2008; 43(7): 774–86
45. Wurtz SO, Schrohl AS, Mouridsen H, Brunner N: TIMP-1 as a tumor marker in breast cancer – an update. Acta Oncol, 2008; 47(4): 580–90
46. Jiang Y, Goldberg ID, Shi YE.: Complex roles of tissue inhibitors of metalloproteinases in cancer. Oncogene, 2002; 21(14): 2245–52
47. Chirco R, Liu XW, Jung KK, Kim HR: Novel functions of TIMPs in cell signaling. Cancer Metastasis Rev, 2006; 25(1): 99–113
48. Spindler KL, Christensen IJ, Nielsen HJ et al: TIMP-1 and CEA as biomarkers in third-line treatment with irinotecan and cetuximab for metastatic colorectal cancer. Tumour Biol, 2015; 36(6): 4301–8
49. Bockelman C, Beilmann-Lehtonen I, Kaprio T et al: Serum MMP-8 and TIMP-1 predict prognosis in colorectal cancer. BMC Cancer, 2018; 18(1): 679
50. Meng C, Yin X, Liu J et al: TIMP-1 is a novel serum biomarker for the diagnosis of colorectal cancer: A meta-analysis. PLoS One, 2018; 13(11): e0207039
51. Lee JH, Choi JW, Kim YS: Plasma or serum TIMP-1 is a predictor of survival outcomes in colorectal cancer: A meta-analysis. J Gastrointestin Liver Dis, 2011; 20(3): 287–91
52. Birgisson H, Nielsen HJ, Christensen IJ et al: Preoperative plasma TIMP-1 is an independent prognostic indicator in patients with primary colorectal cancer: A prospective validation study. Eur J Cancer, 2010; 46(18): 3323–31
53. Song G, Xu S, Zhang H et al: TIMP1 is a prognostic marker for the progression and metastasis of colon cancer through FAK-PI3K/AKT and MAPK pathway. J Exp Clin Cancer Res, 2016; 35(1): 148
54. Niewiarowska K, Pryczynicz A, Dymicka-Piekarska V et al: Diagnostic significance of TIMP-1 level in serum and its immunohistochemical expression in colorectal cancer patients. Pol J Pathol, 2014; 65(4): 296–304
55. Teufel A, Weinmann A, Galle P, Lohse A: In silico characterization of LZTS3, a potential tumor suppressor. Oncol Rep, 2005; 14(2): 547–51
56. He J, Yu L, Wang CM, Zhou XF: MiR-1275 promotes non-small cell lung cancer cell proliferation and metastasis by regulating LZTS3 expression. Eur Rev Med Pharmacol Sci, 2018; 22(9): 2680–87
57. Zhuo C, Wu X, Li J et al: Chemokine (C-X-C motif) ligand 1 is associated with tumor progression and poor prognosis in patients with colorectal cancer. Biosci Rep, 2018; 38(4): pii: BSR20180580
58. Divella R, Daniele A, De Luca R et al: Circulating levels of VEGF and CXCL1 are predictive of metastatic organotropismin in patients with colorectal cancer. Anticancer Res, 2017; 37(9): 4867–71
59. Triner D, Xue X, Schwartz AJ et al: Epithelial hypoxia-inducible factor 2alpha facilitates the progression of colon tumors through recruiting neutrophils. Mol Cell Biol, 2017; 37(5): pii: e00481-16
60. Wang D, Sun H, Wei J et al: CXCL1 is critical for premetastatic niche formation and metastasis in colorectal cancer. Cancer Res, 2017; 77(13): 3655–65
61. le Rolle AF, Chiu TK, Fara M et al: The prognostic significance of CXCL1 hypersecretion by human colorectal cancer epithelia and myofibroblasts. J Transl Med, 2015; 13: 199
62. Aleksandrova K, di Giuseppe R, Isermann B et al: Circulating omentin as a novel biomarker for colorectal cancer risk: Data from the epic-potsdam cohort study. Cancer Res, 2016; 76(13): 3862–71
63. Jaikanth C, Gurumurthy P, Cherian KM, Indhumathi T: Emergence of omentin as a pleiotropic adipocytokine. Exp Clin Endocrinol Diabetes, 2013; 121(07): 377–83
64. Suzuki YA, Shin K, Lönnerdal B: Molecular cloning and functional expression of a human intestinal lactoferrin receptor. Biochemistry, 2001; 40(51): 15771–79
65. Yang R, Lee M, Hu H et al: Identification of omentin as a novel depot-specific adipokine in human adipose tissue: Possible role in modulating insulin action. Am J Physiol Endocrinol Metab, 2006; 290(6): 1253–61
66. Tsuji S, Uehori J, Matsumoto M et al: Human intelectin is a novel soluble lectin that recognizes galactofuranose in carbohydrate chains of bacterial cell wall. J Biol Chem, 2001; 276(26): 23456–63
67. Dan L, Xiang Z, Yong X et al: Intelectin 1 suppresses tumor progression and is associated with improved survival in gastric cancer. Oncotarget, 2015; 6(18): 16168–82
68. Kim HJ, Kang UB, Lee H et al: Profiling of differentially expressed proteins in stage IV colorectal cancers with good and poor outcomes. J Proteomics, 2012; 75(10): 2983–97
69. Kawashima K, Maeda K, Saigo C et al: Adiponectin and intelectin-1: Important adipokine players in obesity-related colorectal carcinogenesis. Int J Mol Sci, 2017; 18(4): pii: E866
70. Guo H, Zeng W, Feng L et al: Integrated transcriptomic analysis of distance-related field cancerization in rectal cancer patients. Oncotarget, 2017; 8(37): 61107
71. Zhang S, Jin J, Tian X, Wu L: hsa-miR-29c-3p regulates biological function of colorectal cancer by targeting SPARC. Oncotarget, 2017; 8(61): 104508–24
72. Ding L, Lu Z, Lu Q, Chen YH: The claudin family of proteins in human malignancy: A clinical perspective. Cancer Manag Res, 2013; 5: 367–75
73. Turksen K, Troy TC: Junctions gone bad: Claudins and loss of the barrier in cancer. Biochim Biophys Acta, 2011; 1816(1): 73–79
74. Maryan N, Statkiewicz M, Mikula M et al: Regulation of the expression of claudin 23 by the enhancer of zeste 2 polycomb group protein in colorectal cancer. Mol Med Rep ,2015; 12(1): 728–36
75. Bujko M, Kober P, Mikula M et al: Expression changes of cell-cell adhesion-related genes in colorectal tumors. Oncol Lett, 2015; 9(6): 2463–70
76. Pavel P, Ondrej V, Jan B et al: Differential expression and prognostic role of selected genes in colorectal cancer patients. Anticancer Res, 2013; 33(11): 4855–65
77. Mazzoni SM, Fearon ER: AXIN1 and AXIN2 variants in gastrointestinal cancers. Cancer Lett, 2014; 355(1): 1–8
78. Galamb O, Kalmar A, Peterfia B et al: Aberrant DNA methylation of WNT pathway genes in the development and progression of CIMP-negative colorectal cancer. Epigenetics, 2016; 11(8): 588–602
79. Aulehla A, Wehrle C, Brand-Saberi B et al: Wnt3a plays a major role in the segmentation clock controlling somitogenesis. Dev Cell, 2003; 4(3): 395–406
80. Jho Eh, Zhang T, Domon C et al: Wnt/-Catenin/Tcf signaling induces the transcription of Axin2, a negative regulator of the signaling pathway. Mol Cell Biol, 2002; 22(4): 1172–83
81. Leung JY, Kolligs FT, Wu R et al: Activation of AXIN2 expression by beta-catenin-T cell factor. A feedback repressor pathway regulating Wnt signaling. J Biol Chem, 2002; 277(24): 21657–65
82. Klaus A, Birchmeier W: Wnt signalling and its impact on development and cancer. Nat Rev Cancer, 2008; 8: 387–98

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]