Research Article

Identification and verification of three key genes associated with survival and prognosis of COAD patients via integrated bioinformatics analysis

Yong Liu1,* , Chao Li2,* , Lijin Dong3, Xuewei Chen2 and Rong Fan2,4
1Department of General Surgery, Tianjin Xiqing Hospital, Tianjin 300380, P.R. China; 2Department of Operational Medicine, Tianjin Institute of Environmental and Operational Medicine, Tianjin 300050, P.R. China; 3Editorial Department of Education and Research Security Centre, Logistic University of Chinese People’s Armed Police Force, Tianjin 300309, P.R. China; 4Central Laboratory, Tianjin Xiqing Hospital, Tianjin 300380, P.R. China

Correspondence: Xuewei Chen (chenxuewei11@sina.com) or Rong Fan (prosperity_39@sina.cn)

Background: Colorectal cancer (CRC) is the third most lethal malignancy in the world, wherein colon adenocarcinoma (COAD) is the most prevalent type of CRC. Exploring biomarkers is important for the diagnosis, treatment, and prevention of COAD. Methods: We used GEO2R and Venn online software for differential gene screening analysis. Hub genes were screened via Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) and Cytoscape, following Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis. Finally, survival analysis and RNA expression validation were performed via UALCAN online software and real-time PCR. Immunohistochemistry (IHC) was performed to verify the protein expression level of hub genes from tissues of COAD patients. Results: In the present study, we screened 323 common differentially expressed genes (DEGs) from four GSE datasets. Furthermore, four hub genes were selected for survival correlation analysis and expression level verification, three of which were shown to be statistically significant. Conclusion: Our study suggests that Serpin Family E Member 1 (SERPINE1), secreted phosphoprotein 1 (SPP1) and tissue inhibitor of metalloproteinase 1 (TIMP1) may be biomarkers closely related to the prognosis of CRC patients.

Introduction

Colorectal cancer (CRC) originates in the colon or rectum [1], and typical symptoms include alternating diarrhea, constipation, mucus, bloody stools, as well as weight loss [2]. According to annual reports on global cancer statistics, nearly 1.4 million newly diagnosed cases and 700000 deaths occur worldwide due to CRC [3]. The 5-year survival rate of patients with CRC at its early stage is 90%, while those at later stages show a rate of no more than 12% [4,5]. Colon adenocarcinoma (COAD) is one of the most prevalent types of CRC, the incidence and mortality of which was 10.2 and 9.2%, respectively, in 2018 [6,7]. Therefore, it is very important to select and identify the specific biomarkers of COAD for its early diagnosis, development of an efficient treatment strategy, and assessment of patient prognosis.

Gene microarray technology has been extensively applied in several fields of medicine and biology. This high-throughput technology can effortlessly acquire gene expression profile information, which is very useful for exploring various molecular mechanisms [8]. Recently, integrated bioinformatics analysis has also been widely applied; it includes mathematics and biology, which makes it possible to analyze large-scale microarray data. In addition, it is advantageous as an efficient method for the systematic screening of tumor-associated genes [9,10].
In the present study, we screened the common differentially expressed genes (DEGs) in four GSE datasets from the Gene Expression Omnibus (GEO) repository and constructed a protein–protein interaction (PPI) network. Via integrated bioinformatics analysis, we identified the four hub genes associated with survival of COAD patients and verified the mRNA expression levels in CRC cells. Finally, our results showed that secreted phosphoprotein 1 (SPP1), Serpin Family E Member 1 (SERPINE1) and tissue inhibitor of metalloproteinase 1 (TIMP1) could be potential biomarkers for prognostics of COAD in patients.

Materials and methods
Download of microarray data
The high-throughput gene expression profiles of normal colorectal tissues and COAD tissues were extracted from the GEO database. The independent datasets of GSE37364, GSE41328, GSE81558, and GSE110225 were selected which included 14 COAD tissues and 37 normal tissues, 10 COAD tissues and 10 normal tissues, 23 COAD tissues and 9 normal tissues, and 17 COAD tissues and 17 normal tissues, respectively.

Identification of DEGs
The DEGs between normal colorectal tissues and COAD tissues were screened via the GEO2R tool. We limited the cut-off criterion of $|\log FC| \geq 1$ to be statistically significant. Then, the DEGs in the four datasets were screened out using online Venn software. The DEGs with $\log FC \geq 1$ were considered as up-regulated genes, while DEGs with $\log FC \leq -1$ were considered as down-regulated genes.

Enrichment analysis via Gene Ontology and Kyoto Encyclopedia of Genes and Genomes pathways
In order to characterize the functional roles of the DEGs, we used Database for Annotation, Visualization and Integrated Discovery (DAVID; version 6.8 [11]) database for Gene Ontology (GO) and signal pathway enrichment analysis, which comprised biological process (BP), molecular function (MF) and cellular component (CC), and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways analysis.

Construction of PPI network and analysis of module
The PPI network was built via the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) [12] database to uncover the relationships of DEGs, based on a minimum required interaction score of 0.4. We used Cytoscape (version 3.7.1) [13] software to complete the analysis and visualization of the PPI network. In addition, the remarkable modules of PPI network were selected via the Molecular Complex Detection (MCODE) app (node score cutoff = 0.2, degree cutoff = 2, max. depth = 100, and k-score = 2).

Survival and RNA expression analysis of hub genes
UALCAN [14] was used for analyzing the relationship between hub genes expression and survival of patients with COAD, as it is a good tool for analyzing transcriptome data of cancers from The Cancer Genome Atlas (TCGA). Simultaneously, UALCAN was used to analyze RNA expression data based on TCGA. $P$-value $<0.05$ was considered statistically significant.

Cell lines and culture
The normal colon epithelial cell line, NCM460, and COAD cell lines, HCT116, HCA-7, HT-29, and SW480 were maintained by our laboratory. And all cells were cultured in a 5% CO$_2$ humidified incubator in Dulbecco’s modified Eagle’s medium (HyClone) containing 10% fetal bovine serum (FBS, Gibco) at 37°C.

Quantitative real-time polymerase chain reaction (PCR)
Total RNA from cell cultures was collected by using Total RNA Extraction Kit (Solarbio Life Sciences, China). The specific primers were designed by using Primer Express 3.0 software. Each sample was used in triplicate for the qPCR using PowerUp™ SYBR™ Green Master Mix (Thermo Fisher Scientific, U.S.A.) on Applied Biosystems™ StepOnePlus instruments, and analyzed with the StepOne software, using GAPDH as a control. The primers used in the present study are designed as follows: TIMP1: Forward: 5’-GGGACACCCAGTCAACCCAG-3’; Reverse: 5’-CAATGAGAAGTCITCGCTG-3’; C-X-C motif chemokine ligand 2 (CXCL2): Forward: 5’-CCAAACCGTAAGTCCACAG-3’; Reverse: 5’-TGCCATTTTCAGCATCTTTTC-3’; SPP1: Forward:
and 212 down-regulated genes in COAD tissues compared with normal colorectal tissues.

Results
Identification of DEGs in COAD

In the present study, we selected and downloaded four GEO datasets and extracted DEGs on the basis of the cut-off criteria. The present study covered 64 COAD and 73 normal colorectal tissues in total. The results showed that 2729, 957, 1864, and 617 genes were down-regulated while 1917, 770, 1282, and 469 genes were up-regulated in GSE37364, 41328, 81585, and 110250, respectively. By using the Venn diagram software, we detected a total of 323 common DEGs in the COAD samples, composed of 111 up-regulated genes and 212 down-regulated genes (Table 1 and Figure 1).

Tissue samples and immunohistochemistry

Sixty paired primary cancer tissues and adjacent tissues from colon cancer patients who underwent surgery between September 2011 and April 2020, were collected at the Tianjin Xiqing Hospital. No chemotherapy and radiotherapy was given to patients before surgery. All the pathology materials were reviewed independently by two pathologists. Consent was obtained from each patient, and the collection of tissue specimens was approved by the Institutional Review and Ethics Boards of the Tianjin Xiqing Hospital. The two-step EnVision method was used to perform immunohistochemical experiments, along with different primary antibodies against CXCL2 (1:50), SERPINE1 (1:50), TIMP1 (1:50) and SPP1 (1:50). The slides were analyzed using the Image-Pro Plus software program (Media Cybernetics, Inc. U.S.A.).

Statistical analysis

Statistical analysis was performed using SPSS 22.0 and GraphPad Prism 8.0 software. Results were presented as the mean ± standard deviation. The statistically significant changes were indicated with asterisks and the P-values were calculated via Student’s t test, wherein *, **, and *** represented P < 0.05, P < 0.01 and P < 0.001, respectively.

Table 1 A total of 323 commonly DEGs were screened from four profile datasets, including 111 up-regulated and 212 down-regulated genes in COAD tissues compared with normal colorectal tissues.

<table>
<thead>
<tr>
<th>DEGs</th>
<th>Gene symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up-regulated</td>
<td>CD44, ADAM12, CXCL3, PLAU, GDF15, MMP7, TRIB3, SERPINB5, FOXQ1, MET, SFTA2, RFC3, RGS1, CD258, COL1A1, ANLN, DUSP4, SLC6A8, TIMG6, CLDN2, VSN1L, PHD1A, CHD5L, NEBL, CXCL11, CXCL8, MS22, PIM1H, SOX9, STC1, GRHPE1, CADPS1, THY1, CBP4L3, HSST2, RAD54B, SPP1, CA9, BMP17, OLR1, EPF4X, NK4X4, FPR4, SLC7A11, RNF38E, TIMP1, BGN, TNFRSF1A, IBSP, MMP1, S100A11, CKB2, COL8A1, TNFDE, TEAD4, CD43, SLC4A1, MMP3, UBE2S, KRIT1, FAP, INHBA, SLC7A5, CXCL2, GALNT6, UBE2C, MYEOV, FXYDS, SRPX2, HOMER1, PSAT1, TMVRSS3, LRP8, NFE2L3, CEMIP, AZG1P1, TACSTD2, TNS4, CERS1, PMAIP1, S100A2, CX1R1, SORL1, DPEP1, FABP6, PPAR, SERPINE1, MSLN, KLK10, PLEKHS1, EVA1A, SIM2, AJUBA, HILPD1, SH2B2, CLDN1, ENC1, STC2, COL10A1, KRIT1B, SP5, PMEP1, TNFSF19, COL11A1, CTHRC1, TESC, TGFBI, COL1A2, ZAK, C2D4A4, KRIT2, FABP6, PPAR, SERPINE1, MSLN, KLK10, PLEKHS1, EVA1A, SIM2, AJUBA, HILPD1, SH2B2, CLDN1, ENC1, STC2, COL10A1, KRIT1B, SP5, PMEP1, TNFSF19, COL11A1, CTHRC1, TESC, TGFBI, COL1A2, ZAK, C2D4A4, KRIT2, FABP6, PPAR, SERPINE1, MSLN, KLK10, PLEKHS1, EVA1A, SIM2, AJUBA, HILPD1, SH2B2, CLDN1, ENC1, STC2, COL10A1, KRIT1B, SP5, PMEP1, TNFSF19, COL11A1, CTHRC1, TESC, TGFBI, COL1A2, ZAK, C2D4A4, KRIT2, FABP6, PPAR, SERPINE1, MSLN, KLK10, PLEKHS1, EVA1A, SIM2, AJUBA, HILPD1, SH2B2, CLDN1, ENC1, STC2, COL10A1, KRIT1B, SP5, PMEP1, TNFSF19, COL11A1, CTHRC1, TESC, TGFBI, COL1A2, ZAK, C2D4A4, KRIT2, FABP6, PPAR, SERPINE1, MSLN, KLK10, PLEKHS1, EVA1A, SIM2, AJUBA, HILPD1, SH2B2, CLDN1, ENC1, STC2, COL10A1, KRIT1B, SP5, PMEP1, TNFSF19, COL11A1, CTHRC1, TESC, TGFBI, COL1A2, ZAK, C2D4A4, KRIT2, FABP6, PPAR, SERPINE1, MSLN, KLK10, PLEKHS1, EVA1A, SIM2, AJUBA, HILPD1, SH2B2, CLDN1, ENC1, STC2, COL10A1, KRIT1B, SP5, PMEP1, TNFSF19, COL11A1, CTHRC1, TESC, TGFBI, COL1A2, ZAK, C2D4A4, KRIT2, FABP6, PPAR, SERPINE1, MSLN, KLK10, PLEKHS1, EVA1A, SIM2, AJUBA, HILPD1, SH2B2, CLDN1, ENC1, STC2, COL10A1, KRIT1B, SP5, PMEP1, TNFSF19, COL11A1, CTHRC1, TESC, TGFBI, COL1A2, ZAK, C2D4A4, KRIT2</td>
</tr>
<tr>
<td>Down-regulated</td>
<td>PLCE1, XHD, HD17B2, NAP1L2, EPB41L4B, ZG16, TMCC3, DHFRS11, HS1D1B2, TRPM6, IGH, CES2, PYH, WWA5A, MAOA, BOCHE, THRIP, TSPN7, HHLA2, ACYLI2, ILBR, PKB, CDHS, RBLD, FHL1, PTTP4, SRBP2, KLFA, TCFC1, NEDD4L, SLC51B, AOP8, LYPD8, TNFRSF17, CD177, C1orf115, SLC25A34, HPGDS, ADHC1, SGK2, TMEM27, SLC17A4, NPY1R, UGT2A3, GCNT2, LRRC19, SST, SNN1BB, DENND2A, CLIC5, PRKACB, THIDE, EBP41L3, LAMA1, CVH43, NRX11, MPE1A, PDK4, GUC2A2, ANGPTL1, SLC22A5, GREN2, LGALS2, NR3C2, CHGB, OGN, SLC25A23, CA4, NQO3-3, ARPP2, PR1MA, NAAA, RGS13, SEMA6D, LINC01082, FUC1A1, LRMP, SLC26A2, ADTRP, VSTM4A, MTIF, PLPP1, SOSTDC1, ACA8B, MSA412, COH19, GIS92, ABC13, CEACAM1, C2orf88, PLP1, BEST4, STMN2, PPP2R3A, TIMM6, TEX1, SLC4A4, C40MEP18, MTPB, TMEET, SI, TMEM37, EVA2, PDE6A, CHRD1L, LDHD, PHLPK2, AKR1B10, TUBAL3, ARL14, RMDN2, PTGDR, S1P, POU2AF1, TFCP2L1, DNAE3L1, MIER3, SCIN, CKB, SLC26A3, ACOX1, PIATOR, MYOT, RCAN2, CHGA, ITGA5, MTIM, TPS3, CNP1, RNR1, PDE9A, ABC8, SLC13A2, SLC16A9, CLDN23, EDN3, ANK2, HSD11A1, SGE1, PKC1, CPM, SMPD3A4, ABI3BP, WDRT7B, C1orf210, RNF238, USP2, HPGD, FKB18B, C7, MOGA72, EFHC2, EC2, EPHF2, SEMA6A, TIMM100, DHR95, CA7, ENTPD5, SPINK2, CPX83, CNTN3, ABCG2, MMP28, GUC2A8, CHPP2, SCARAS5, CAPN13, CLCA4, AT1PA2, MOYD1, NKPE4, EFTDH, MB, LIFR, MA01, GDPD3, CXCL12, NR5A2, C16orf89, FM05, PAQR5, CCL28, CA2, B2CS1, SMIM14, ANPEP, KSNIP4, CEACAM7, PAD2, METTL7A, ITMD2, ARFIP44A, NR1H4, TMEM220, BEST2, ADH1B, ZNF575, GCO, GBA3, PGM6B, ANK3, TCEA3, NXPE1, LPAR1, MAMDC2, CA1, SULT1B1, DPI8, SRPR1, SELENBP1, GPA73, HAPLN1</td>
</tr>
</tbody>
</table>
Figure 1. Selection of 323 common DEGs from the four datasets (GSE37364, GSE41328, GSE81558, and GSE110225)

(A) Volcano plot of DEGs from the four datasets. (B) On the left, 136 DEGs are down-regulated (logFC ≤ 1). On the right, 198 DEGs are up-regulated (logFC ≥ 1).

GO and signaling pathway enrichment analysis

We analyzed the DEGs via the DAVID software and identified 125 significant enrichment terms, including BP (71), MF (24), and CC (26). With regard to BP, DEGs were seen to be significantly enriched during cell adhesion, negative regulation of cell proliferation, angiogenesis, cytoskeleton organization, steroid hormone-mediated signaling pathway, inflammatory response, negative regulation of protein kinase activity, and so on. Whereas, for MF, DEGs were significantly enriched during chemokine activity, hormone activity, CXCR chemokine receptor binding, receptor binding, calcium ion binding, structural constituent of cytoskeleton, protein kinase inhibitor activity, oxidoreductase activity etc. With regard to CC, DEGs were significantly enriched in extracellular space, apical plasma membrane, extracellular exosome, extracellular region, cell surface, plasma membrane, basal plasma membrane, and so forth. The top 20 GO terms are depicted in Figure 2A–C. Furthermore, KEGG pathway analysis indicated that the DEGs were enriched in 11 pathways including nitrogen metabolism, bile secretion, proximal tubule bicarbonate reclamation, cytokine–cytokine receptor interaction, aldosterone-regulated sodium reabsorption, mineral absorption, phosphoinositide-3-kinase (PI3K)-Akt signaling pathway, PPAR signaling pathway etc. (Figure 2D).

Construction of PPI network and module analysis

Via the STRING and Cytoscape analysis, we drew the PPI network complex constructed with 268 nodes and 709 edges, consisting of 97 down-regulated and 171 up-regulated genes. Then we conducted further analyses by applying Cytoscape MCODE plus, and found 32 central nodes that included 20 up-regulated genes and 12 down-regulated genes (Figures 3 and 4 and Table 2).
Table 2 A total of 32 central genes were selected by STRING and Cytoscape software from PPI network including 20 up-regulated and 12 down-regulated genes

<table>
<thead>
<tr>
<th>DEGs</th>
<th>Gene symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up-regulated</td>
<td>BGN, BMP7, COL1A1, COL1A2, CXCL1, CXCL11, CXCL2, CXCL3, CXCL8, EVA1A, IBSP, MMP1, MMP3, MMP7, MSLN, SERPINE1, SPP1, STC2, THY1, TIMP</td>
</tr>
<tr>
<td>Down-regulated</td>
<td>CCL28, CD44, CHGB, CHRD1, CNR1, CXCL12, LPAR1, NPY1R, PLAU, PYY, SOX9, SST</td>
</tr>
</tbody>
</table>

Selection of hub genes and validation of the expression levels

We used UALCAN to identify the survival data of 32 central genes and found that the expression levels of four hub genes, namely CXCL2, SERPINE1, SPP1, and TIMP1, were remarkably related to the survival of COAD patients (Figure 5). Furthermore, we validated the expression level of these four hub genes in COAD and normal tissues via UALCAN. Results showed that expression levels of the four genes were significantly increased in the COAD samples as compared with the normal samples (Figure 6). Furthermore, we verified the mRNA levels of the four hub genes in NCM460, HCT116, HCA-7, HT-29, and SW480 cells. The real-time PCR results showed that the expression levels of SPP1, TIMP1, and SERPINE1 mRNAs were significantly increased in the HCT116, HCA-7, HT-29, and SW480 cell lines as compared with the NCM460 cell line (Figure 7). Moreover, CXCL2 expression was higher in the HCA-7, HT-29, and SW480 cells, but not in the HCT116 cells when compared with NCM460. In addition, we also analyzed the protein expressions of the four hub genes by using tissue microarray. The results indicated that SPP1, TIMP1, and SERPINE1 expression appeared to be remarkably higher in CRC tissues than in the adjacent tissues, while the expression of CXCL2 was not statistically different between cancer tissues and adjacent tissues (Figure 8).
CRC is the third main cause of deaths from malignant cancers worldwide, and the fourth most common cancer in China [15,16]. One of the most common types of CRC is COAD, which originates in the superficial glandular epithelial cells [6]. Previous studies have indicated that the mechanism behind the tumorigenesis and development of COAD is a complicated process, including chromosomal instability, dysregulation of oncogenes and tumor suppressor, epigenetic alterations, metabolic alterations, abnormal immune response etc. [17].
Figure 4. Module analysis by MCODE in the Cytoscape software
(A) Module1; (B) module2 (blue circles signify up-regulated DEGs and red circles denoted down-regulated DEGs).

Figure 5. Correlation between the four hub genes expression and the overall survival of COAD patients, via UALCAN analysis
Figure 6. Analysis of the four hub genes expression levels via UALCAN analysis.

Figure 7. Relative mRNA levels of the four hub genes in HCT116, HCA-7, HT-29, and SW480 cells compared with NCM460 cells, via real-time PCR.

** Represents $P < 0.01$, *** represents $P < 0.001$, ns represents no significance.
In the present study, we screened DEGs via GEO2R, and selected 111 up-regulated genes and 212 down-regulated genes found in COAD tissues. Then we analyzed their enrichment in GO and signaling pathways and found that the DEGs were enriched in 11 pathways, including the PI3K-Akt signaling pathway. Previous studies showed that several signal pathways could be involved in the progress and development of CRC [18–23]. PI3K/AKT is one of the classic pathways responsible for multiple BPs, such as cell proliferation, differentiation, and migration. As a key effector of PI3K, Akt phosphorylation is closely related to CRC cell proliferation and apoptosis inhibition [18]. The specific inhibition of the PI3K/AKT signaling pathway could result in a decrease in proliferation and increase in apoptosis of the cells [24].

Furthermore, we constructed a PPI network and performed survival analysis along with verification of the RNA expression levels via the GEO database, DAVID and STRING online tools, Cytoscape software, and UALCAN online
analysis. After the above series of operations, four hub genes, CXCL2, SERPINE1, SPP1, and TIMP1, were selected. Furthermore, we tested the expression of these four genes at the mRNA as well as protein levels and found that SPP1, TIMP1, and SERPINE1 had higher expression levels in COAD cell lines than the colon epithelial cell line as well as in CRC tissue compared with para-cancerous tissue.

SPP1, named as osteopontin, is a phosphoprotein secreted by malignant cells. Cheng et al. [25] found that SPP1 has a higher expression level in CRC tissues as compared with normal tissue, wherein its expression is closely associated with lymph node metastasis and the TNM stage. Inhibition of SPP1 in the HCT116 cells consequently resulted in the suppression of cellular growth, migration, and invasion. SPP1 overexpression could thus enhance the proliferation, migration, and invasion of CRC cells as well as the activation of the PI3K pathway, accompanied by the increased expression of Snail and matrix metalloproteinase 9 (MMP9) [26]. As one of the transcription factors of epithelial–mesenchymal transition (EMT), snail can inhibit the expression of E-cadherin, thereby promotes the activation of EMT in tumor cells [27]. MMP9 is one of the members in the MMP family which is a class of zinc-dependent proteases with the main function of participating in the degradation of extracellular matrix. MMP9 is often overexpressed in malignant tumors and is associated with aggressive malignant phenotypes and poor prognosis in cancer patients [28]. Similarly, SPP1 could not only promote CRC cell growth and migration, but it also inhibited autophagy, which may be thus achieved by activating the p38 mitogen-activated protein kinase (MAPK) signaling pathway [29]. In addition, a study by Choe et al. [30] also confirmed higher SPP1 expression along with shorter survival times in CRC patients.

TIMP1 is one of the members of the TIMP family which can bind to MMP substrates and mediate MMP degradation that is necessary for matrix renewal [31]. TIMP1 has been shown to be highly expressed in many types of cancer, consequently correlated to their poor prognosis [32–36]. TIMP1 may also promote liver metastasis in CRC [37]. After chemotherapy, high expression levels of TIMP1 in the serum have been closely related to the decline over survival (OS) [38]. Moreover, knockdown of TIMP1 leads to decreased proliferation of HT29 cells [39]. Huang et al. [40], Zeng et al. [41], and Zuo et al. [42] demonstrated that the expression levels of TIMP1 may be negatively correlated with the prognosis of COAD patients. Song et al. [43] found that TIMP1 knockdown could result in decreased proliferation, migration, and invasion as well as increased apoptosis in vitro, whereas there is weakened tumorigenesis and metastasis in vivo via the FAK-PI3K/Akt and MAPK pathways. The elevated TIMP1 levels in serum were related to lower tumor grade and shorter overall survival times of CRC patients [44].

SERPINE1, also called PAI-1, can regulate plasmin formation from plasminogen [45]. The expression level of SERPINE1 was seen to increase in CRC tissue and colon cancer cells lines showing active proliferation [46]. Moreover, overexpression of SERPINE1 led to decreased apoptosis and caspase-3 amidolytic activity in vascular smooth muscle cells (VSMCs) compared with control [47]. LC-MS/MS analysis has shown that SERPINE1 may be a peripheral participant in the apoptosis of CRC cells [48]. Shun et al. found that the expression level of SERPINE1 was remarkably up-regulated in tumor stroma and associated with the relapse of CRC stage II–III patients [49]. Several recent studies have suggested that SERPINE1 may be associated with prognosis of CRC patients [41,50,51].

CXCL2 belongs to the chemokine superfamily, and is responsible for the attraction of leukocytes to inflammatory sites apart from the regulation of angiogenesis, tumorigenesis [52]. Dietrich et al. found that there was significantly increased expression of CXCL2 in CRC, in contrast with normal tissue [53]. Further, there was increased expression of mRNA and protein levels of CXCL2 in colon cancer stem cells, whereas CXCL2 knockdown led to the decreased expression of EMT markers, MMPs, and Gai-2, which consequently promoted tumor initiation and development [54]. Down-regulation of CXCL2 destroyed the proliferation and colony formation in CRC cells [55]. Moreover, CXCL2 remarkably promoted the proliferation and migration of human umbilical vein endothelial cells (HUVECs) [56]. But the qPCR results showed no significant difference in CXCL2 expression between HCT116 and NCM460 cell lines. This may be due to the source of the cell lines; more cell lines should thus be selected for verification in future studies.

In summary, in the present study, we screened 323 DEGs from four COAD GEO datasets and identified SERPINE1, SPP1, and TIMP1 as key genes which may be closely related to the prognosis of COAD in patients. Further investigations need to be performed to explore the mechanisms of these genes. These results would consequently provide candidate targets for the diagnosis and prognosis of CRC in patients, as well as new ideas for treatment.

Competing Interests
The authors declare that there are no competing interests associated with the manuscript.

Funding
This work was supported by the National Natural Science Foundation of China [grant number 81201757]; and the Basic Research Projects of Logistic University of Chinese People’s Armed Police Force [grant numbers WHJ201702, BWS17J025, AWS16J004].

© 2020 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).
The visualization of GO analysis, including BP, MF and CC of DEGs, was performed via the online software ImageGP (http://www.ebihio.com/ImageGP/index.php/Home/Index/index).

Author Contribution
Yong Liu conceived the idea behind the present paper. Chao Li and Xuwei Chen collected and analyzed the data. Lijin Dong and Ping Li drafted the paper. Rong Fan revised the final paper.

Ethics Approval
Clinical colon cancer samples were obtained from the Tianjin Xiqing Hospital, with informed written consents from the patients or their guardians and the study protocol was approved by the Institutional Review Board of Tianjin Xiqing Hospital.

Abbreviations
BP, biological process; CC, cellular component; COAD, colon adenocarcinoma; CRC, colorectal cancer; CXCL2, C-X-C motif chemokine ligand 2; DAVID, Database for Annotation, Visualization and Integrated Discovery; DEG, differentially expressed gene; EMT, epithelial-mesenchymal transition; FAK, focal adhesion kinase-related nonkinase; Gai-2, Gxi-2; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; GEO, Gene Expression Omnibus; GO, Gene Ontology; GSE, Gene Expression Omnibus Series; HUVEC, human umbilical vein endothelial cell; KEGG, Kyoto Encyclopedia of Genes and Genomes; MAPK, mitogen-activated protein kinase; MCODE, Molecular Complex Detection; MF, molecular function; MMP, matrix metalloproteinase; PAI-1, plasminogen activator inhibitor 1; P13K, phosphoinositide-3-kinase; PPAR, peroxisome proliferator activated receptor; SERPINE1, Serpin Family E Member 1; SP1, secreted phosphoprotein 1; STRING, Search Tool for the Retrieval of Interacting Genes/Proteins; TCGA, The Cancer Genome Atlas; TIMP1, tissue inhibitor of metalloproteinase 1; TNM, tumor node metastasis.

References

© 2020 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).


41 Zeng, C. and Chen, Y. (2019) HTR1D, TIMP1, SERPINE1, MMP3 and CNR2 affect the survival of patients with colon adenocarcinoma. Oncol. Lett. 18, 2448–2454


