Identification of Issue Inhibitor of Metalloproteinase 1 (TIMP1) as a Potential Biomarker for the Diagnosis, Pathogenesis and Prognosis of Colorectal Cancer via Integrated Bioinformatic Analysis

Guona Li  
Shanghai University of Traditional Chinese Medicine Yueyang Hospital of Integrated Traditional Chinese Medicine and Western Medicine

Mengmeng Kang  
Shanghai University of Traditional Chinese Medicine Yueyang Hospital of Integrated Traditional Chinese Medicine and Western Medicine

Siyuan Sheng  
University of the Chinese Academy of Sciences

Ziyi Chen  
Shanghai University of Traditional Chinese Medicine Yueyang Hospital of Integrated Traditional Chinese Medicine and Western Medicine

Kunshan Li  
Shanghai University of Traditional Chinese Medicine Yueyang Hospital of Integrated Traditional Chinese Medicine and Western Medicine

Luyi Wu  
Shanghai University of Traditional Chinese Medicine Yueyang Hospital of Integrated Traditional Chinese Medicine and Western Medicine

Yan Huang  
Shanghai Research Institute of Acupuncture and Meridian

Muen Gu  
Shanghai Research Institute of Acupuncture and Meridian

Huangan Wu *(wuhuangan@126.com)*  
Shanghai Research Institute of Acupuncture and Meridian

---

**Research**

**Keywords:** TIMP1, prognostic signature, colorectal cancer, bioinformatic analysis,
Abstract

**Background:** Colorectal cancer (CRC) is a common malignant tumor of the digestive system. It is crucial to screen potential biomarkers for the diagnosis, pathogenesis, and prognosis of CRC because there are limited clinical symptoms associated with this cancer. Therefore, we attempted to identify biomarkers associated with the occurrence and progression of CRC by utilizing bioinformatic analysis and to elucidate a molecular mechanism for the diagnosis and treatment of CRC.

**Methods:** Two independent gene expression profile datasets of colonic neoplasms (GSE44076 and GSE37182) were collected from public GEO datasets, which included 182 tumor tissues and 236 normal tissues. Next, differentially expressed genes (DEGs) between CRC colonic samples and non-CRC colonic samples were obtained via GEO2R online tools. Subsequently, hub genes were selected by several analyses of DEGs, including GO pathway enrichment analysis, KEGG pathway enrichment analysis, and PPI network analysis. Finally, the correlation between the hub genes and the occurrence of CRC was tested by harnessing survival analysis and ROC curve analysis.

**Results:** Sixty-one shared DEGs were screened, including 44 high-expression genes and 17 low-expression genes, in CRC samples. Four genes (MYC, TIMP1, MMP7, and COL1A1) were considered to be hub genes because they exhibited higher connectivity degree scores through PPI network analysis. More importantly, there was a significant correlation between increased expression of TIMP1 and reduced survival time in patients with colorectal cancer.

**Conclusion:** By using bioinformatic analysis, this study suggested that Timp-1 may represent a potential biomarker for the diagnosis and prognosis of targeted molecular therapy for CRC.

Full Text

Due to technical limitations, full-text HTML conversion of this manuscript could not be completed. However, the latest manuscript can be downloaded and accessed as a PDF.

Figures
Figure 1
Heatmap of DEFs from GSE44076 (A) and GSE37182 (B). The green color indicates genes with low expression, while the red color indicates genes with high expression. Each row represents one gene probe, and each column represents one sample.

Figure 2
Venn diagrams of shared high-expression and lowly-expression DEGs in GSE44076 and GSE37182. (A) Forty-four shared upregulated DEGs in the two datasets and (B) seventeen shared downregulated DEGs in the two datasets were analyzed.
Figure 3

Go and KEGG pathway of the two datasets. (A) Biological process, (B) cellular components, (C) molecular function, and (D) KEGG pathway. The Y-axis reflects the GO and KEGG pathway terms, and the X-axis reflects the gene ratio. Different colors represent the degree of significance, and a redder color indicates higher significance. The size of the node indicates gene counts.
Figure 4

Highly-expression DEGs are shown in red, while the DEGs with low expression are shown in blue. The shape of the node indicates the difference in connectivity degree. MYC, IMP1, MMP7, and COL1A1 were considered to be hub genes due to a higher level of connectivity degree.
Figure 5

Overall survival analysis of TIMP1 (A), COL1A1 (B), MMP7 (C), and MYC (D). A P value < 0.05 was considered to be significant.
Figure 6

ROC curve for gene expression in the GSE44076 (A) and GSE37182 (B) datasets.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementarymaterial1.xlsx