[文章编号] 1671-587X(2024)04-1076-11 DOI:10.13481/j.1671-587X.202404022

基于SSP1和TGFB1与食管腺癌发生、预后和免疫浸润关系的生物信息学分析

王元国，张 鹏

（天津医科大学总医院心胸外科，天津 300052）

[摘 要] **目的：**分析基因表达综合（GEO）数据库和癌症基因组图谱（TCGA）数据库中食管腺癌（EAC）基因表达数据，阐明EAC发病的潜在核心基因与肿瘤淋巴细胞浸润的关系，为EAC的诊断和治疗提供分子靶标。**方法：**在GEO数据库检索“esophageal adenocarcinoma”，下载包括EAC和食管正常组织的高通量芯片数据集GSE13898、GSE26886、GSE74553和GSE92396。采用R软件的limma包筛选EAC组织和食管正常组织的差异表达基因（DEGs），并通过韦恩图获取共同DEGs，采用STRING数据库分析后导入Cytoscape软件筛选核心基因并构建蛋白-蛋白互作（PPI）网络，采用基因表达谱交互分析（GEPIA）数据库验证核心基因表达水平，采用阿拉巴马大学伯明翰分校癌症数据分析门户（UALCAN）和Kaplan-Meier Plotter数据库分析核心基因与EAC患者预后和临床资料的关联性，采用肿瘤免疫评价资源（TIMER）数据库分析核心基因与肿瘤免疫浸润的关系，采用基因本体论（GO）和京都基因与基因组百科全书（KEGG）对LinkedOmics数据库获得的核心基因中正相关表达基因进行功能和信号通路富集分析。**结果：**对GEO获得的4个数据集的DEGs取交集，共获得340个DEGs，其中上调基因127个，下调基因213个。经STRING数据库和Cytoscape软件筛选后，最终获得评分最高的关键核心基因分泌型磷蛋白1（SPP1）和转化生长因子β1（TGFB1）。GEPIA数据库分析，与食管正常组织比较，癌组织中SPP1和TGFB1 mRNA表达水平明显升高（$P<0.01$）；SPP1低表达组EAC患者1、3和5年总体生存期均高于SPP1高表达组（HR=10.1，$P<0.05$；HR=3.09，$P<0.05$；HR=2.32，$P<0.05$），TGFB1低表达组EAC患者5年总体生存期高于TGFB1高表达组（HR=2.36，$P<0.05$）。UALCAN数据库分析，与食管正常组织比较，Ⅱ-Ⅲ期及N1-N2期淋巴结转移的EAC患者癌组织中SPP1和TGFB1 mRNA表达水平明显升高（$P<0.01$）。TIMER分析，SPP1和TGFB1 mRNA表达水平与EAC患者癌组织中巨噬细胞（$r=0.353$，$P<0.01$；$r=0.187$，$P<0.05$）和树突状细胞（$r=0.236$，$P<0.01$；$r=0.221$，$P<0.01$）浸润呈正相关关系。GO功能和KEGG信号通路富集分析，SPP1和TGFB1及其排名前50位正相关基因主要参与细胞迁移、细胞活性和血管发育等生物学过程及肿瘤蛋白多糖、细胞外基质（ECM）-受体互作和磷脂酰肌醇3-激酶（PI3K）/蛋白激酶B（AKT）等信号通路。**结论：**SPP1和TGFB1与EAC患者临床分期、淋巴结转移和总体生存期有密切关联。SPP1和TGFB1高表达可能导致巨噬细胞和树突状细胞浸润，从而改变肿瘤微环境。SPP1和TGFB1可能成为EAC诊断和治疗的新靶点。

[关键词] 食管腺癌；生物信息学；生存分析；预后；肿瘤免疫浸润

[中图分类号] R735.1；Q811.4 [文献标志码] A

[收稿日期] 2023-03-22

[基金项目] 天津市科技局京津冀基础研究合作专项项目（20JCZXJC00190）

[作者简介] 王元国（1984－），男，山东省平阴县人，主管技师，工程硕士，主要从事胸部肿瘤分子机制方面的研究。

[通信作者] 张 鹏，教授，主任医师，博士研究生导师（E-mail：zhangfapeng@aliyun.com）

Bioinformatics analysis based on relationship between SSP1 and TGFB1 and occurrence, prognosis, and immune invasion of esophageal adenocarcinoma

WANG Yuanguo, ZHANG Peng
(Department of Cardio-Thoracic Surgery, General Hospital, Tianjin Medical University, Tianjin 300052, China)

ABSTRACT Objective: To analyze gene expression data of esophageal adenocarcinoma (EAC) in the Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) databases and clarify the relationship between the potential core genes and tumor lymphocyte infiltration in the EAC, and to provide the molecular targets for the diagnosis and treatment of EAC. Methods: The high-throughput chip datasets GSE13898, GSE26886, GSE74553, and GSE92396, including EAC and normal esophageal tissues, were downloaded from the GEO database by searching for "esophageal adenocarcinoma". The limma package of R software was used to screen the differentially expressed genes (DEGs) in EAC tissue and esophageal normal tissue, and the common DEGs were obtained through Venn diagram. After the DEGs were analyzed by STRING database, the results were imported into Cytoscape software to screen the core genes and construct the protein-protein interaction (PPI) network. The Gene Expression Profiling Interactive Analysis (GEPIA) database was used to verify the expression levels of core genes. The UALCAN and Kaplan-Meier Plotter databases were used to analyze the correlations between the core genes and prognosis and clinical data of the EAC patients. The Tumor Immune Estimation Resource (TIMER) database was used to analyze the relationship between core genes and tumor immune infiltration. Gene Ontology (GO) functional and Kyoto Encyclopedia of Genes and Genomes (KEGG) signaling pathway enrichment analysis were performed to analyze the positively correlated genes of core genes obtained from the LinkedOmics database. Results: A total of 340 DEGs were obtained from the intersection of DEGs from the four GEO datasets, including 127 upregulated genes and 213 downregulated genes. After screening with the STRING database and Cytoscape software, the key core genes with the highest scores were secreted phosphoprotein 1 (SPP1) and transforming growth factor beta 1 (TGFB1). The GEPIA database analysis results showed that compared with esophageal normal tissue, the expression levels of SPP1 and TGFB1 mRNA in cancer tissue were significantly increased ($P<0.01$). The 1-year, 3-year, and 5-year overall survival of the EAC patients in SPP1 low expression group was higher than those in SPP1 high expression group (HR=10.1, $P<0.05$; HR=3.09, $P<0.05$; HR=2.32, $P<0.05$), and the 5-year overall survival of the EAC patients in TGFB1 low expression group was higher than that in TGFB1 high expression group (HR=2.36, $P<0.05$). The UALCAN database analysis results showed that compared with esophageal normal tissue, the expression levels of SPP1 and TGFB1 mRNA in cancer tissue of the EAC patients with stage Ⅱ-Ⅲ and N1-N2 lymph node metastasis were significantly increased ($P<0.01$). The TIMER analysis results showed that the expression levels of SPP1 and TGFB1 mRNA in cancer tissue of the EAC patients were positively correlated with the infiltration of macrophages ($r=0.353$, $P<0.01$; $r=0.187$, $P<0.05$) and dendritic cells ($r=0.236$, $P<0.01$; $r=0.221$, $P<0.01$). The GO and KEGG pathway enrichment analysis results showed that SPP1, TGFB1, and their top 50 positively correlated genes mainly participated in the biological processes such as cell migration, cell activity, and angiogenesis, and signaling pathways such as tumor proteoglycans, extracellular matrix (ECM)-receptor interaction, and phosphatidylinositol 3-kinase (PI3K)/protein kinase B (AKT). Conclusion: SPP1 and

TGFB1 are closely associated with clinical staging, lymph node metastasis, and overall survival of the EAC patients. High expressions of SPP1 and TGFB1 may lead to the infiltration of the macrophages and dendritic cells, and change the tumor microenvironment. SPP1 and TGFB1 may become new targets for the diagnosis and treatment of EAC.

KEYWORDS Esophageal adenocarcinoma; Bioinformatics; Survival analysis; Prognosis; Tumor immune infiltration

食管腺癌（esophageal adenocarcinoma，EAC）是常见的上消化道恶性肿瘤，在流行病学和发病危险因素方面与食管鳞癌有较大差异，EAC患者早期症状不明显，由于确诊较晚和晚期肿瘤转移，导致其5年生存率仅为17%，中位生存期不到1年[1]。EAC病因和发病机制尚未明确，且在临床上难以早期发现和诊断。因此，急需探索用于EAC早期诊断和预后评估的生物学标志物。分泌型磷蛋白1（secreted phosphoprotein 1，SPP1）是一种趋化因子样钙化的细胞外基质（extracellular matrix，ECM）相关蛋白，其主要功能为促进肿瘤细胞增殖和抑制细胞凋亡[2]，在评估宫颈癌和胃癌等多种肿瘤的转移潜力中具有重要的价值[3-4]。转化生长因子β1（transforming growth factor-beta 1，TGFB1）在免疫反应和肿瘤发生中发挥重要作用，在肿瘤进展阶段，TGFB1可以促进肿瘤生长，同时诱导肿瘤细胞发生上皮间质转化，促进肿瘤转移[5-6]；在肿瘤浸润淋巴细胞中，TGFB1可以通过减少白细胞介素2（interleukin-2，IL-2）的分泌抑制T淋巴细胞增殖活性[7]。虽然SPP1和TGFB1与多种肿瘤的发生发展有关，但其与EAC及肿瘤免疫浸润的关系尚不明确。本研究旨在应用基因表达综合（Gene Expression Omnibus，GEO）数据库分析SPP1和TGFB1在EAC组织及食管正常组织中表达水平的差异，并通过癌症基因图谱（The Cancer Genome Atlas，TCGA）数据库验证其表达水平，探讨SPP1和TGFB1与EAC患者临床特征和免疫浸润的关系，为EAC的诊断和治疗提供候选分子靶标。

1 资料与方法

1.1 GEO数据库芯片信息获取 采用GEO数据库（http：//www. ncbi. nlm. nih. gov/geo/）以关键词“esophageal Adenocarcinoma”进行检索。筛选条件：①mRNA表达谱数据集；②数据集同时包含肿瘤组织和正常组织；③单个数据集样本数量不低于20例。最终筛选到4个基因芯片数据集：GSE13898包含64例EAC组织和28例食管正常组织，GSE26886包含21例EAC组织和19例食管正常组织，GSE74553包含52例EAC组织和8例食管正常组织，GSE92396包含12例EAC组织和9例食管正常组织，其中GSE13898数据表达谱芯片平台为Illumina芯片，GSE26886、GSE74553和GSE92396数据表达谱平台为Affymetrix芯片。

1.2 差异表达基因（differentially expressed genes，DEGs）的提取和分析 在GEO网站下载筛选的Series数据文件和平台文件，应用Perl软件（5.38.0版本）将芯片的探针名注释为官方基因名，多个探针对应一个基因的数据取探针表达水平的平均值，并删除未检测到数据（not available，NA）和空值，应用R软件的limma包（3.56.2版本）进行差异分析。筛选标准为$P<0.05$，且$|\log_2FC|\geq1.5$，DEGs筛选后，以$-\log_{10}P$为纵坐标，以$\log_2FC$为横坐标，采用R软件中的ggplot2包（3.3.3版本）绘制火山图。采用在线Venn工具（http://bioinformatics. psb. ugent. be/webtools/Venn/）绘制韦恩（Venn）图，确定4个数据集的共同DEGs用于进一步分析。

1.3 蛋白-蛋白互作（protein-protein interaction，PPI）网络构建和核心基因筛选 为了评估DEGs蛋白的互作关系，将上述获得的340个DEGs导入STRING在线数据库（https://cn. string-db. org/，版本号：11.5），互作评分>0.4，将计算得到的数据导入Cytoscape软件（3.9.1版本），进行PPI网络构建和数据可视化，并计算节点和边，采用Cytoscape软件中的Cytohubba模块计算每个蛋白质节点的连接度，并应用MCODE模块对蛋白进行排序，构建PPI，网络中排序靠前且评分较高的基因即为筛选出的核心基因。

1.4 核心基因差异表达分析和EAC患者生存分析 采用基因表达谱交互分析（Gene Expression Profiling Interactive Analysis，GEPIA）数据库（http://gepia. cancer-pku. cn/）对核心基因SPP1和TGFB1 mRNA表达差异进行验证，采用

Kaplan-Meier Plotter数据库（https://kmplot.com/analysis/）分析核心基因与EAC患者预后的关系，以$P<0.05$为差异有统计学意义。

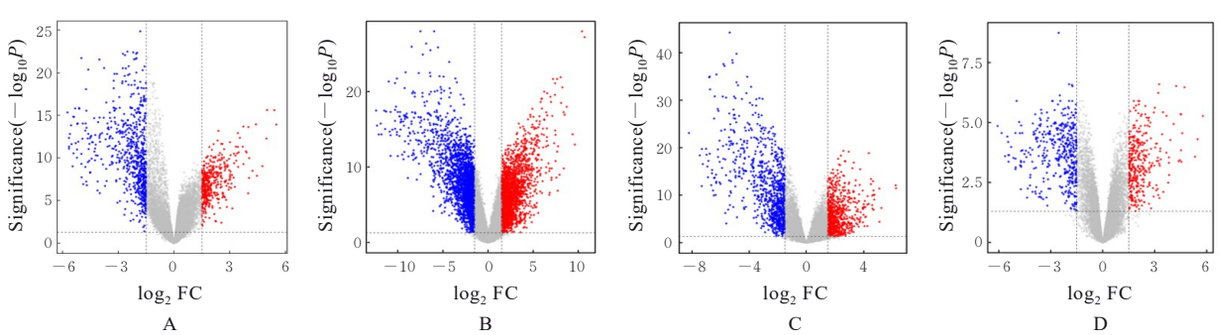
1.5 不同临床病理特征EAC患者癌组织中核心基因表达情况 采用阿拉巴马大学伯明翰分校癌症数据分析门户（University of Alabama at Birmingham Cancer Data Analysis Portal，UALCAN）（https://ualcan.path.uab.edu/）数据库分析SPP1和TGFB1表达情况与EAC患者临床分期和淋巴结转移等的关系。条件如下：①登录网站后选择“TCGA基因”模块；②网页搜索框中输入基因名SPP1和TGFB1；③在TCGA数据集中选择EAC；④在表达谱标签中选择肿瘤分析、淋巴结转移情况和患者性别；⑤以$P<0.05$为差异有统计学意义。

1.6 核心基因与免疫细胞浸润的相关性分析 采用肿瘤免疫评价资源（Tumor Immune Estimation Resource，TIMER）2.0在线数据库（https://cistrome.shinyapps.io/timer/）分析SPP1和TGFB1与EAC患者癌组织中多种免疫细胞浸润水平的关系。筛选条件如下：①登录网站后选择“基因”模块；②在基因符号中输入SPP1和TGFB1；③在肿瘤类型中选择食管癌；④在免疫浸润标签中选择B淋巴细胞、CD8+T淋巴细胞、CD4+T淋巴细胞、巨噬细胞、中性粒细胞和树突状细胞；⑤分析结果，相关系数（r）>0为正相关关系，$r<0$为负相关关系，以$P<0.05$为差异有统计学意义。

1.7 核心基因参与的生物学过程和信号通路 采用LinkedOmics数据库（https://www.linkedomics.org/）分析核心基因的正相关表达基因及其参与的生物学过程和信号通路。筛选条件：①在肿瘤组标签中选择食管癌；②在搜索数据集中选择RNA测序数据集；③数据集基因属性为SPP1和TGFB1；④统计学方法选择皮尔森（Pearson）相关性检验；⑤下载与核心基因呈正相关的基因数据集，筛选$r>0$且$P<0.01$排名前50位基因。将筛选的50个基因导入DAVID数据库（https://david.ncifcrf.gov/）进行基因本体论（Gene Ontology，GO）功能和京都基因与基因组百科全书（Kyoto Encyclopedia of Genes and Genomes，KEGG）信号通路富集分析，分析呈正相关基因参与的生物学过程和信号通路，以$P<0.05$为差异有统计学意义，采用R软件ggplot2包绘制气泡图。

2 结 果

2.1 DEGs筛选 按照筛选条件，分别从GSE13898、GSE26886、GSE74553和GSE92396数据集中获得1 094个（上调基因407个，下调基因687个）、4 961个（上调基因2 463个，下调基因2 498个）、1 566个（上调基因710个，下调基因856个）和789个（上调基因352个，下调基因437个）DEGs（图1）。通过Venn在线分析网站，进一步筛选4个数据集的共同DEGs，共获得340个DEGs，其中上调基因127个，下调基因213个（图2）。

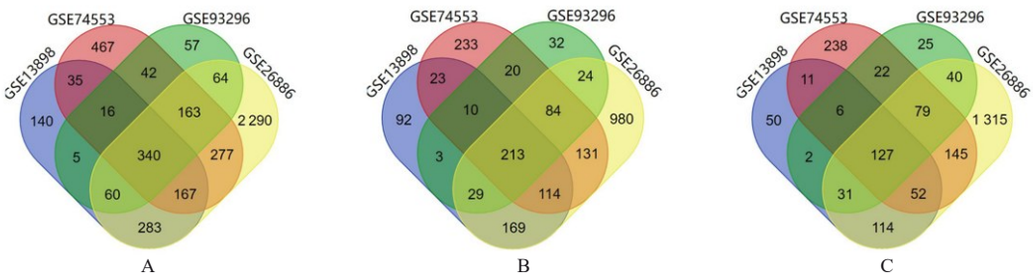


A：GSE13898 dataset DEGs volcano map；B：GSE26886 dataset DEGs volcano map；C：GSE74553 dataset DEGs volcano map；D：GSE92396 dataset DEGs volcano map. Red dots represented significantly upregulated DEGs，and blue dots represented significantly downregulated DEGs.

图1 4个数据集EAC组织中DEGs表达情况

Fig. 1 Expressions of DEGs in EAC tissue in four datasets

2.2 PPI网络和核心基因 将筛选的340个DEGs导入STRING数据库中进行PPI网络分析，并将分析结果导入Cytoscape软件，获得272个节点和870条边，通过Cytohubba和MCODE模块分



A：Venn diagram of DEGs in GSE dataset；B：Venn diagram of downregulated DEGs in GSE dataset；C：Venn diagram of upregulated DEGs in GSE dataset.

图2 4个数据集EAC组织共表达的DEGs Venn图

Fig. 2 Venn diagram of co-expressed DEGs in EAC tissue in four datasets

析后取交集，构建PPI网络图（图3），获得18个核心基因：SPP1，TGFB1，Ⅳ型胶原α1链（collagen type Ⅳ alpha 1 chain，COL4A1），多功能蛋白聚糖（versican，VCAN），基质金属蛋白酶1（matrix metallopeptidase 1，MMP1），Ⅴ型胶原α2链（collagen type Ⅴ alpha 2 chain，COL5A2），Ⅰ型胶原α1链（collagen type Ⅰ alpha 1 chain，COL1A1），巢蛋白2（nidogen 2，NID2），Ⅻ型胶原α1链（collagen type Ⅻ alpha 1 chain，COL12A1），双糖链蛋白聚糖（biglycan，BGN），Serpin家族H成员1（serpin family H member 1，SERPINH1），Ⅲ型胶原α1链（collagen type Ⅲ alpha 1 chain，COL3A1），Lumican蛋白多糖（Lumican proteoglycan，LUM），Ⅵ型胶原α3链（collagen type Ⅵ alpha 3 chain，COL6A3），金属蛋白酶组织抑制剂1（tissue inhibitor of metalloproteinases 1，TIMP1），凝血酶敏感蛋白1（thrombospondin 1，THBS1），钙黏着蛋白11（cadherin 11，CDH11），骨膜蛋白（periostin，POSTN）。将18个核心基因分别与127个上调基因和213个下调基因取交集，18个核心基因均属于上调基因。

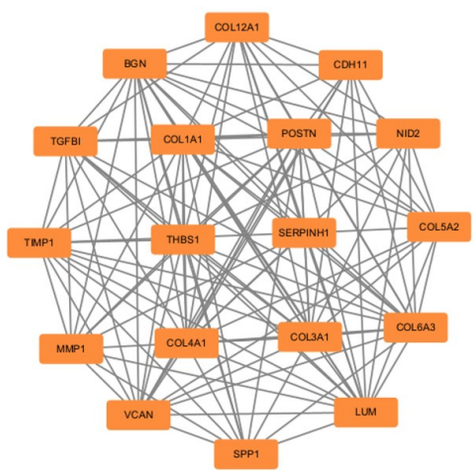
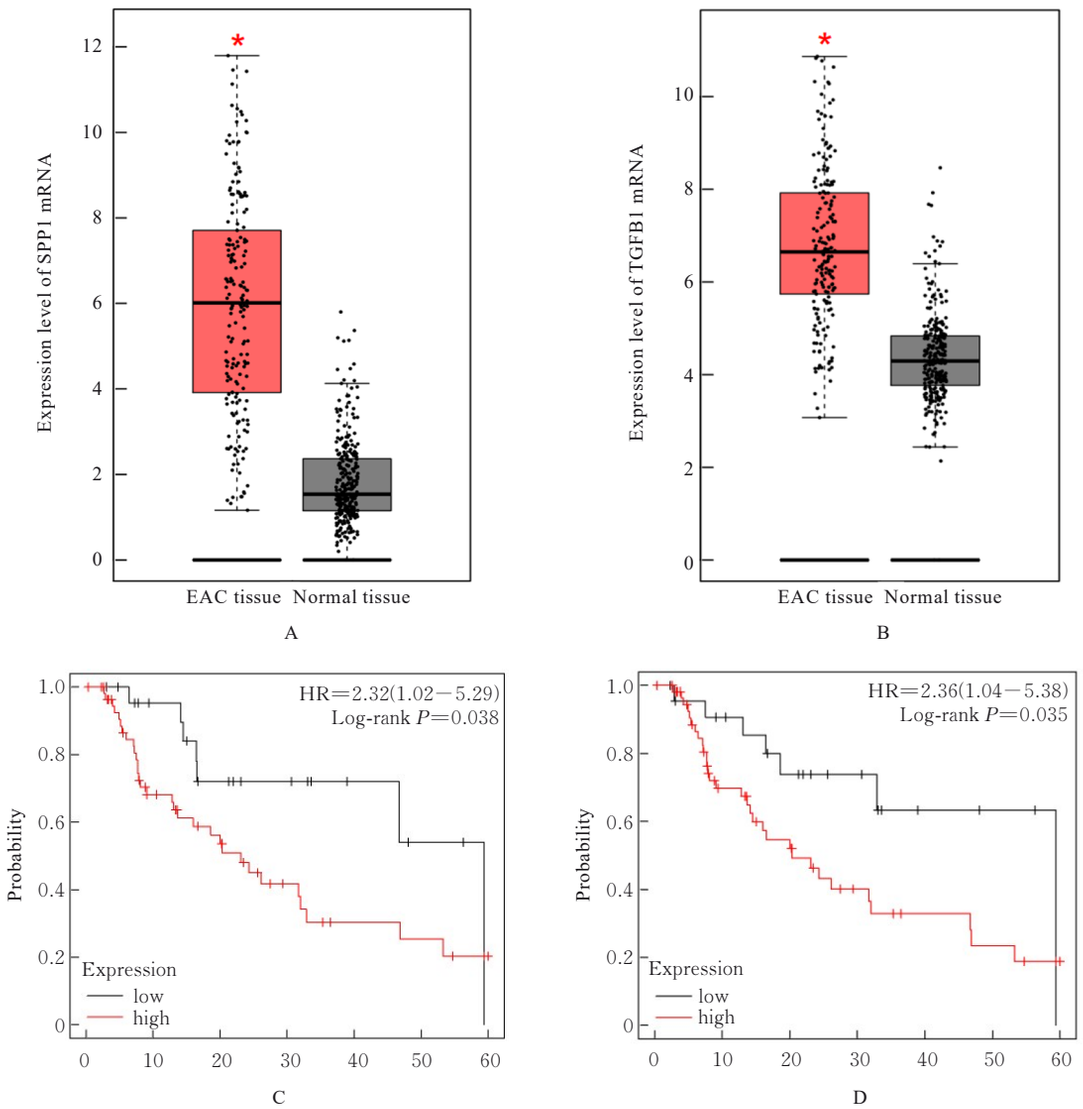


图3 PPI网络分析图

Fig. 3 PPI network analysis diagram

2.3 EAC患者癌组织中SPP1和TGFB1 mRNA表达水平及EAC患者预后分析 选择MCODE模块评分最高且在EAC中研究较少的2个基因SPP1和TGFB1为研究对象，通过GEPIA数据库验证其mRNA表达水平，非配对数据分析结果显示：EAC患者癌组织（n=182）中SPP1 mRNA表达水平高于食管正常组织（n=286）（$\log_2$FC=4.473，P<0.01）；EAC患者癌组织（n=182）中TGFB1 mRNA表达水平高于食管正常组织（n=286）（$\log_2$FC=1.846，P<0.01）（图4A和4B）。采用Kaplan-Meier Plotter数据库分析SPP1和TGFB1 mRNA表达水平与EAC患者预后的关系，结果显示：SPP1 mRNA低表达组患者1、3和5年总体生存期均高于SPP1 mRNA高表达组（HR=10.1，P=0.006；HR=3.09，P=0.014；HR=2.32，P=0.038），TGFB1 mRNA低表达组患者5年总体生存期高于TGFB1 mRNA高表达组（HR=2.36，P=0.035），而TGFB1 mRNA低表达组患者1和3年总体生存期与TGFB1 mRNA高表达组比较差异无统计学意义（HR=0.55，P=0.28；HR=0.65，P=0.28）（图4C和4D）。

2.4 不同临床病理特征EAC患者癌组织中SPP1和TGFB1 mRNA表达水平 通过UALCAN网站分析来源于TCGA数据库的EAC患者的临床资料，分别按照临床分期、淋巴结转移和性别等资料分析EAC患者癌组织中SPP1和TGFB1 mRNA表达情况。与食管正常组织比较，临床分期Ⅱ和Ⅲ期EAC患者癌组织中SPP1和TGFB1 mRNA表达水平升高（P<0.01），Ⅰ和Ⅳ期EAC患者癌组织中



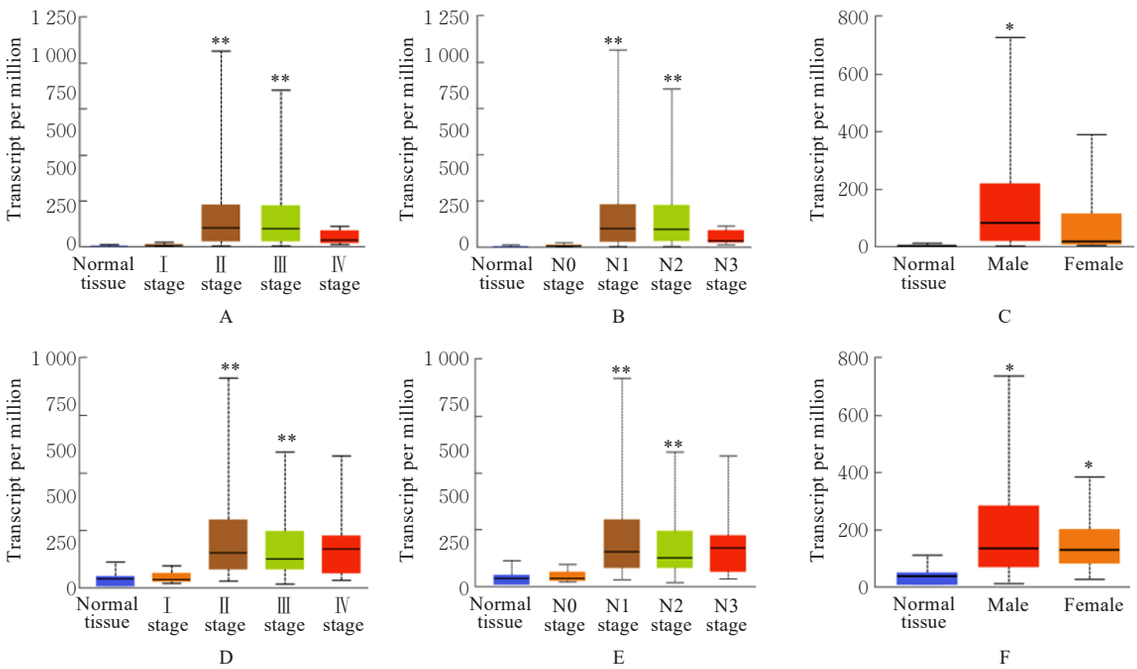
A：Expression of SPP1 mRNA in EAC and normal tissues（TCGA）；B：Expression of TGFB1 mRNA in EAC and normal tissues（TCGA）；C：Expression level of SPP1 mRNA；D：Expression level of TGFB1 mRNA.*$P<0.01$ compared with normal tissue.

图4 EAC患者癌组织中SPP1 mRNA和TGFB1 mRNA表达水平及EAC患者预后分析

Fig. 4 Expression levels of SPP1 mRNA and TGFB1 mRNA in cancer tissue and prognosis analysis on EAC patients

SPP1和TGFB1 mRNA表达水平差异无统计学意义（$P>0.05$）；与食管正常组织比较，N1和N2期淋巴结转移EAC患者癌组织中SPP1和TGFB1 mRNA表达水平升高（$P<0.01$），N0和N3期淋巴结转移EAC患者癌组织中SPP1和TGFB1 mRNA表达水平差异无统计学意义（$P>0.05$）；与食管正常组织比较，男性EAC患者癌组织中SPP1 mRNA表达水平升高（$P<0.05$），女性EAC患者癌组织中SPP1 mRNA表达水平差异无统计学意义（$P>0.05$），男性和女性EAC患者癌组织中TGFB1 mRNA表达水平升高（$P<0.05$）。见图5。

2.5 EAC患者癌组织中SPP1和TGFB1 mRNA表达水平与免疫细胞浸润的相关性 采用TIMER 2.0数据库在线分析EAC患者癌组织中SPP1和TGFB1 mRNA表达水平与多种免疫细胞浸润的关系，结果显示：SPP1 mRNA表达水平与巨噬细胞和树突状细胞浸润呈正相关关系（$r=0.353$，$P<0.001$；$r=0.236$，$P<0.01$），与细胞纯度（$r=0.019$，$P>0.05$）、B淋巴细胞（$r=-0.138$，$P>0.05$）、CD8+T淋巴细胞（$r=-0.128$，$P>0.05$）、CD4+T淋巴细胞（$r=0.011$，$P>0.05$）和中性粒细胞（$r=0.017$，$P>0.05$）均无相关性；TGFB1 mRNA表达水平与巨噬细胞和树突状细胞



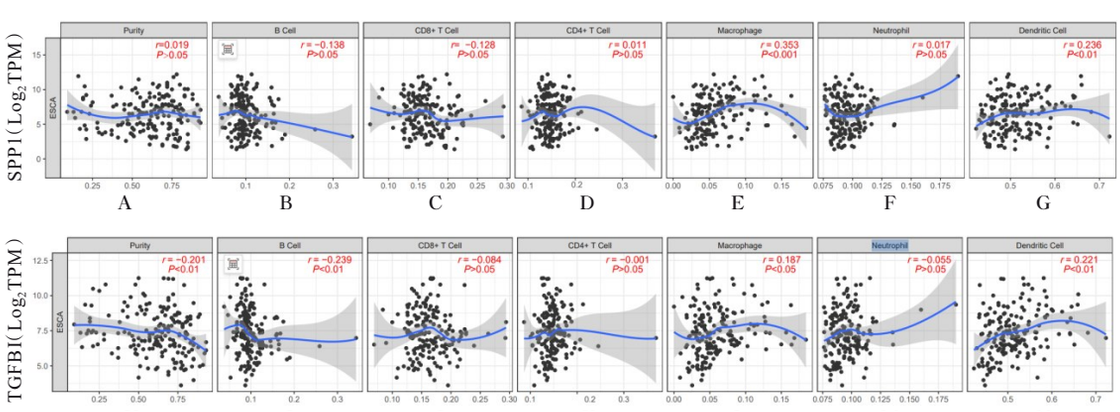
A: Correlation between SPP1 mRNA and cancer stages; B: Correlation between SPP1 mRNA and nodal metastasis status; C: Correlation between SPP1 mRNA and gender; D: Correlation between TGFB1 mRNA and cancer stages; E: Correlation between TGFB1 mRNA and nodal metastasis status; F: Correlation between TGFB1 and gender. $^{*}P<0.05$, $^{**}P<0.01$ compared with normal tissue.

图 5　不同临床病理特征 EAC 患者癌组织中 SPP1 和 TGFB1 mRNA 表达水平

Fig. 5　Expression levels of SPP1 and TGFB1 mRNA in cancer tissue of EAC patients with different clinicopathological characteristics

呈正相关关系（$r=0.187$，$P<0.05$；$r=0.221$，$P<0.01$），与细胞纯度、B 淋巴细胞呈负相关关系（$r=-0.201$，$P<0.01$；$r=-0.239$，$P<0.01$），与 CD8＋T 淋巴细胞（$r=-0.084$，$P>0.05$）、CD4＋T 淋巴细胞（$r=-0.001$，$P>0.05$）和中性粒细胞（$r=-0.055$，$P>0.05$）均无相关性。见图 6。

2.6　SPP1 和 TGFB1 相关基因参与调控的信号通路　为进一步分析预测 SPP1 和 TGFB1 相关基因在 EAC 中可能参与的生物学过程和信号通路，

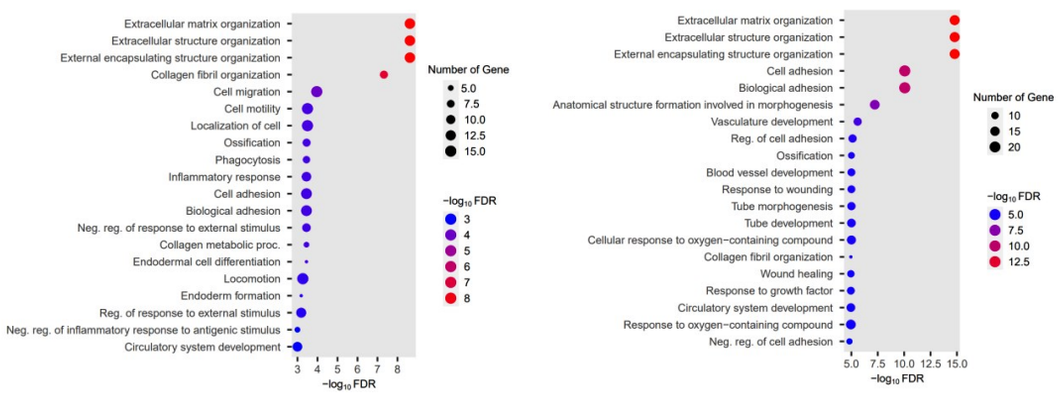


A－G: Expression level of SPP1 mRNA and immune cell infiltration; H－N: Expression level of TGFB1 and immune cell infiltration. A,H: Cell purities; B,I: B lymphocytes; C,J: CD8＋T lymphocytes; D,K: CD4＋T lymphocytes; E,L: Macrophages; F,M: Neutrophils; G,N: Dendritic cells.

图 6　EAC 患者癌组织中 SPP1 和 TGFB1 mRNA 表达水平与 EAC 免疫细胞浸润的相关性

Fig. 6　Correlations between expression levels of SPP1 and TGFB1 mRNA and immune cell infiltration in cancer tissue of EAC patients

采用LinkedOmics数据库获得SPP1和TGFB1的正相关基因，将差异显著的排名前50位正相关基因导出，导入DAVID数据库进行GO功能和KEGG信号通路分析。GO功能富集分析结果显示：SPP1和TGFB1均参与细胞外基质（extracellular matrix，ECM）、细胞外结构、外部封装结构和细胞黏附的生物学过程，SPP1还参与细胞迁移、细胞活性、细胞定位、吞噬作用和炎症反应等生物学过程，TGFB1参与细胞黏附、血管发育、血管形态发生、细胞对含氧化合物反应和伤口愈合等生物学过程（图7）。KEGG信号通路分析结果显示：SPP1主要参与蛋白质消化吸收、破骨细胞分化、ECM-受体互作、肿瘤蛋白多糖和吞噬体等信号通路（表1），TGFB1主要参与黏着斑、肿瘤蛋白多糖、蛋白质消化吸收、磷脂酰肌醇3-激酶（phosphatidylinositol 3-kinase，PI3K）/蛋白激酶（protein kinase B，AKT）信号通路和细菌侵入上皮细胞等信号通路（表2）。



A B

A: SPP1 positively associated genes involved in biological processes balloon map; B: TGFB1 positively associated genes involved in biological processes balloon map.

图7 SPP1和TGFB1正相关基因的GO功能富集分析

Fig. 7 GO functional enrichment analysis on SPP1 and TGFB1 positively related genes

表1 SPP1正相关基因的KEGG信号通路富集分析

Tab. 1 KEGG signaling pathway enrichment analysis on SPP1 positively related genes

KEGG pathway	Count	*P*	Gene
Protein digestion and absorption	8	<0.01	COL1A1, COL3A1, COL1A2, COL5A1, COL24A1, COL11A1, COL7A1, COL10A1
AGE-RAGE signaling pathway in diabetic complications	6	<0.01	COL1A1, COL3A1, COL1A2, FN1, NOX4, MAPK12
Osteoclast differentiation	6	<0.01	FCGR3A, FCGR2A, NCF2, TREM2, FCGR1A, MAPK12
Leishmaniasis	5	<0.01	FCGR3A, FCGR2A, NCF2, FCGR1A, MAPK12
Neutrophil extracellular trap formation	6	<0.01	FCGR3A, FCGR2A, SIGLEC9, NCF2, FCGR1A, MAPK12
Platelet activation	5	<0.01	COL1A1, COL3A1, FCGR2A, COL1A2, MAPK12
Phagosome	5	<0.01	MSR1, FCGR3A, FCGR2A, NCF2, FCGR1A
ECM-receptor interaction	4	<0.01	COL1A1, COL1A2, SPP1, FN1
Fc gamma R-mediated phagocytosis	4	<0.01	HCK, FCGR3A, FCGR2A, FCGR1A
Diabetic cardiomyopathy	5	<0.01	COL1A1, COL3A1, COL1A2, NCF2, MAPK12
Proteoglycans in cancer	5	<0.01	COL1A1, COL1A2, PLAU, FN1, MAPK12
Amoebiasis	4	<0.01	COL1A1, COL3A1, COL1A2, FN1
Relaxin signaling pathway	4	<0.01	COL1A1, COL3A1, COL1A2, MAPK12
Tuberculosis	4	<0.05	FCGR3A, FCGR2A, FCGR1A, MAPK12
Focal adhesion	4	<0.05	COL1A1, COL1A2, SPP1, FN1
Staphylococcus aureus infection	3	<0.05	FCGR3A, FCGR2A, FCGR1A

表2 TGFB1正相关基因的KEGG信号通路富集分析

Tab. 2 KEGG signaling pathway enrichment analysis on TGFB1 positively related genes

KEGG pathway	Count	*P*	Gene
Focal adhesion	8	＜0.01	COL1A1, CAV2, ACTN1, CAV1, COL6A2, TNC, ITGA5, PGF
Proteoglycans in cancer	6	＜0.01	COL1A1, TGFB1, PLAU, CAV2, CAV1, ITGA5
ECM-receptor interaction	4	＜0.01	COL1A1, COL6A2, TNC, ITGA5
Protein digestion and absorption	4	＜0.01	COL1A1, COL6A2, COL5A3, COL7A1
PI3K-AKT signaling pathway	5	＜0.05	COL1A1, COL6A2, TNC, ITGA5, PGF
Bacterial invasion of epithelial cells	3	＜0.05	CAV2, CAV1, ITGA5
AGE-RAGE signaling pathway in diabetic complications	3	＜0.05	COL1A1, TGFB1, SERPINE1
Amoebiasis	3	＜0.05	COL1A1, TGFB1, ACTN1

3 讨 论

食管癌是全球最常见的恶性肿瘤之一，包括食管鳞状细胞癌和EAC。与亚洲地区不同的是，在西方国家以EAC为主，占食管癌发病的80%，虽然我国以食管鳞状细胞癌为主，但EAC的发病率也逐年上升[8]。目前已明确的EAC危险因素为巴雷特食管、胃食管返流病、肥胖和吸烟等[9]，研究[10-11]显示：每周出现胃食管返流病症状可使EAC的发生增加5倍。虽然靶向药物在其他肿瘤的治疗中取得较好疗效，但对EAC的治疗效果不佳[12-14]。

肿瘤的发生发展是多基因参与的过程，基因的表达失衡或基因突变均可导致肿瘤发生，通过公共数据库基因的检索，可筛选肿瘤相关驱动基因，为肿瘤发生发展机制的研究提供理论依据[9]。随着生物信息学的发展，更多的潜在肿瘤基因被公开报道，但实际应用于临床的较少，其主要原因是生物信息学分析依赖高质量的生物学数据，不同单一数据质量不同，可导致结果的不准确和不可靠[15]。本研究纳入GEO数据库中的4个EAC芯片数据集，分析获取每个数据集的DEGs后，再对4个数据集的DEGs取交集，最终筛选出340个共同DEGs，有效地保证了数据的准确性。对所有DEGs进行PPI分析，并将分析结果导入Cytoscape软件筛选核心基因构建核心调控网络，最终筛选到18个核心基因，根据MCODE模块评分，挑选评分最高且在食管癌中研究较少的SPP1和TGFB1基因，分析其表达水平与EAC患者临床资料、肿瘤免疫细胞浸润和信号通路调控的关系。

SPP1属于分泌磷酸蛋白的小整合素结合配体N-连接糖蛋白家族，其编码的蛋白参与破骨细胞和矿化骨基质的附着，在细胞-基质相互作用中具有重要作用，同时作为一种细胞因子，可促进干扰素γ（interferon-γ，IFN-γ）和白细胞介素12（interleukin-12，IL-12）的产生，降低白细胞介素10（interleukin-10，IL-10）的产生并参与Ⅰ型免疫反应。SPP1参与前列腺癌、乳腺癌、肺腺癌和卵巢癌等多种肿瘤的发生，并在肿瘤组织中呈高表达，促进肿瘤的生长、迁移和侵袭[16-19]。在卵巢癌中，沉默SPP1可通过抑制整合素β1（integrin β1）/黏着斑激酶（focal adhesion kinase，FAK）/AKT信号通路阻止卵巢癌细胞的增殖[20]。一项针对肝癌的单细胞RNA测序中的受体配体分析[21]结果显示：SPP1可以通过SPP1-CD44和SPP1-前列腺素E受体4（prostaglandin E receptor 4，PTGER4）结合介导肝癌细胞和巨噬细胞之间串扰，促进巨噬细胞向M2型肿瘤相关巨噬细胞极化。SPP1的过表达可维持PI3K/AKT和细胞外调节蛋白激酶1/2（extracellular regulated protein kinases 1/2，ERK1/2）信号通路的活化，与前列腺癌的细胞侵袭和转移有密切的关联，抑制SPP1可以减弱前列腺癌的上皮-间质转化，SPP1可作为前列腺癌的潜在靶点[22]。本研究结果显示：与食管正常组织比较，ECA患者癌组织中SPP1 mRNA表达水平升高，且与ECA患者Ⅱ、Ⅲ期临床分期和N1、N2期淋巴结转移有关联；生存分析结果显示：SPP1 mRNA高表达EAC患者与总生存期呈负相关关系，与其他肿瘤相关研究[23-25]结果一致。肿瘤微环境中的免疫细胞浸润在肿瘤进展中起关键作用，并影响肿瘤患者的生存期。巨噬细胞可为肿瘤细胞提供营养和支持，促进肿瘤侵袭、转移和耐药的发生，一项单细胞测序研究[26]表明：巨噬细胞来源的SPP1在肝癌进展中具有重要作用。本研究中，SPP1 mRNA表达水平与巨噬细胞和树突状细胞浸润呈正相关关系，与

GAO等[27]研究结果一致。

TGFB1可编码一种相对分子质量为66 000的细胞外基质蛋白，主要表达于树突状细胞和巨噬细胞。TGFB1是许多生物学过程的关键调节因子，参与细胞分化、血管生成、免疫稳态和肿瘤发生等。TGFB1在肿瘤细胞中发挥双重作用，在癌症早期阶段，TGFB1通过抑制细胞周期进程和促进细胞凋亡而表现出抑制肿瘤的作用；在晚期阶段TGFB1发挥促瘤作用，增加肿瘤侵袭性和转移性，其主要机制为在癌症进展过程中，TGFB1信号传导部分功能丧失，导致TGFB1从抑制肿瘤转变为促进肿瘤进展，同时TGFB1的改变可激活Ras-丝裂原活化蛋白激酶（mitogen-activated protein kinase，MAPK）信号通路等致癌信号，进一步促进肿瘤进展[28-29]。近年来，多项研究[30-31]表明：TGFB1表达失调与肿瘤发生有关，TGFB1表达升高可促进肾癌细胞的黏附、迁移和侵袭。乳腺癌动物模型研究[32]显示：维持TGFB1的低表达水平可以改善血管灌注，减少缺氧，抑制肿瘤干细胞的生长和肿瘤转移。TGFB1过表达的正常肝细胞更易在小鼠体内形成肝肿瘤，同时TGFB1可通过激活FAK/AKT/蛋白激酶B底物1（protein kinase B substrate 1，AKT1S1）/核糖体蛋白S6（ribosomal protein S6，RPS6）/真核翻译起始因子4E结合蛋白（eukaryotic translation initiation factor 4E binding protein，EIF4EBP）通路促进射线损伤的细胞存活[33]。研究[34]显示：TGFB1可以激活肺癌细胞和乳腺癌细胞中黏附相关信号通路，重塑ECM蛋白，导致细胞运动减少和克隆形成能力减弱，说明其有一定的抑癌作用。本研究结果显示：TGFB1在EAC患者癌组织中高表达，且与EAC患者临床分期和淋巴结转移有关联，TGFB1高表达EAC患者的5年总体生存率降低；TGFB1与EAC肿瘤微环境中巨噬细胞和树突状细胞免疫浸润呈正相关关系，与B淋巴细胞浸润呈负相关关系，目前有关巨噬细胞和树突状细胞免疫浸润在食管腺癌中作用的研究较少，需要进一步完善相关实验，分析其在EAC发生发展中的作用机制。

综上所述，SPP1和TGFB1在EAC患者癌组织中高表达，其表达水平与患者的总体生存率、肿瘤分期和淋巴结转移有密切关联，可作为EAC患者诊断、预后判断和药物靶点的潜在生物标志物，SPP1和TGFB1参与了巨噬细胞和树突状细胞浸润，为免疫细胞浸润在EAC中作用的研究提供了新的方向。

利益冲突声明：

所有作者声明不存在利益冲突。

作者贡献声明：

王元国参与研究设计、数据整理和论文撰写，张鹏参与研究设计、论文撰写指导和论文审阅。

[参考文献]

[1] COLEMAN H G, XIE S H, LAGERGREN J. The epidemiology of esophageal adenocarcinoma [J]. Gastroenterology, 2018, 154(2): 390-405.

[2] POLEBOYINA P K, ALAGUMUTHU M, PASHA A, et al. Entrectinib a plausible inhibitor for osteopontin (SPP1) in cervical cancer-integrated bioinformatic approach[J]. Appl Biochem Biotechnol, 2023, 195(12): 7766-7795.

[3] ZHAO K D, MA Z, ZHANG W. Comprehensive analysis to identify *SPP1* as a prognostic biomarker in cervical cancer[J]. Front Genet, 2022, 12: 732822.

[4] XIE W, CHENG J, HONG Z J, et al. Multi-transcriptomic analysis reveals the heterogeneity and tumor-promoting role of SPP1/CD44-mediated intratumoral crosstalk in gastric cancer [J]. Cancers, 2022, 15(1): 164.

[5] STREEL G D, LUCAS S. Targeting immunosuppression by TGF- β1 for cancer immunotherapy [J]. Biochem Pharmacol, 2021, 192: 114697.

[6] SOUKUPOVA J, MALFETTONE A, BERTRAN E, et al. Epithelial-mesenchymal transition (EMT) induced by TGF-β in hepatocellular carcinoma cells reprograms lipid metabolism[J]. Int J Mol Sci, 2021, 22(11): 5543.

[7] MOREAU J M, VELEGRAKI M, BOLYARD C, et al. Transforming growth factor-β1 in regulatory T cell biology[J]. Sci Immunol, 2022, 7(69): eabi4613.

[8] 田 东，付茂勇，赵泽良，等. 食管腺癌临床病理特征及影响预后的因素[J]. 现代肿瘤医学，2013，21(8)：1748-1753.

[9] RUSTGI A K, EL-SERAG H B. Esophageal carcinoma[J]. N Engl J Med, 2014, 371(26): 2499-2509.

[10] RUBENSTEIN J H, TAYLOR J B. Meta-analysis: the association of oesophageal adenocarcinoma with symptoms of gastro-oesophageal reflux [J]. Aliment Pharmacol Ther, 2010, 32(10): 1222-1227.

[11] SNIDER E J, FREEDBERG D E, ABRAMS J A. Potential role of the microbiome in barrett's esophagus and esophageal adenocarcinoma[J]. Dig Dis Sci, 2016, 61(8): 2217-2225.

[12] KHALAFI S, LOCKHART A C, LIVINGSTONE A S, et al. Targeted molecular therapies in the treatment of esophageal adenocarcinoma, are we there yet? [J]. Cancers, 2020, 12(11): 3077.

[13] ZHANG X, WANG Y X, MENG L H. Comparative genomic analysis of esophageal squamous cell carcinoma and adenocarcinoma: new opportunities towards molecularly targeted therapy[J]. Acta Pharm Sin B, 2022, 12(3): 1054-1067.

[14] FATEHI HASSANABAD A, CHEHADE R, BREADNER D, et al. Esophageal carcinoma: towards targeted therapies[J]. Cell Oncol, 2020, 43(2): 195-209.

[15] 刘 迁, 祁国萍, 于华裔, 等. 结肠癌核心基因和独立预后因子筛选的生物信息学分析[J]. 吉林大学学报(医学版), 2022, 48(3): 755-765.

[16] MESSEX J K, BYRD C J, THOMAS M U, et al. Macrophages cytokine Spp1 increases growth of prostate intraepithelial neoplasia to promote prostate tumor progression[J]. Int J Mol Sci, 2022, 23(8): 4247.

[17] GÖTHLIN EREMO A, LAGERGREN K, OTHMAN L, et al. Evaluation of SPP1/osteopontin expression as predictor of recurrence in tamoxifen treated breast cancer[J]. Sci Rep, 2020, 10(1): 1451.

[18] ZHANG W G, FAN J L, CHEN Q, et al. SPP1 and AGER as potential prognostic biomarkers for lung adenocarcinoma[J]. Oncol Lett, 2018, 15(5): 7028-7036.

[19] PERIYASAMY A, GOPISETTY G, SUBRAMANIUM M J, et al. Identification and validation of differential plasma proteins levels in epithelial ovarian cancer[J]. J Proteomics, 2020, 226: 103893.

[20] ZENG B, ZHOU M, WU H, et al. SPP1 promotes ovarian cancer progression via Integrin β1/FAK/AKT signaling pathway[J]. Onco Targets Ther, 2018, 11: 1333-1343.

[21] LIU L L, ZHANG R Y, DENG J W, et al. Construction of TME and Identification of crosstalk between malignant cells and macrophages by SPP1 in hepatocellular carcinoma [J]. Cancer Immunol Immunother, 2022, 71(1): 121-136.

[22] PANG X C, ZHANG J L, HE X, et al. SPP1 promotes enzalutamide resistance and epithelial-mesenchymal-transition activation in castration-resistant prostate cancer via PI3K/AKT and ERK1/2 pathways[J]. Oxid Med Cell Longev, 2021, 2021: 5806602.

[23] RABJERG M, BJERREGAARD H, HALEKOH U, et al. Molecular characterization of clear cell renal cell carcinoma identifies CSNK2A1, SPP1 and DEFB1 as promising novel prognostic markers[J]. APMIS, 2016, 124(5): 372-383.

[24] CHEN J W, HOU C X, ZHENG Z T, et al. Identification of secreted phosphoprotein 1 (SPP1) as a prognostic factor in lower-grade gliomas [J]. World Neurosurg, 2019, 130: e775-e785.

[25] BIE T W, ZHANG X W. Higher expression of SPP1 predicts poorer survival outcomes in head and neck cancer[J]. J Immunol Res, 2021, 2021: 8569575.

[26] LIU Y H, ZHANG L, JU X Y, et al. Single-cell transcriptomic analysis reveals macrophage-tumor crosstalk in hepatocellular carcinoma [J]. Front Immunol, 2022, 13: 955390.

[27] GAO W, LIU D L, SUN H Y, et al. SPP1 isa prognostic related biomarker and correlated with tumor-infiltrating immune cells in ovarian cancer [J]. BMC Cancer, 2022, 22(1): 1367.

[28] GUASCH G, SCHOBER M, PASOLLI H A, et al. Loss of TGFbeta signaling destabilizes homeostasis and promotes squamous cell carcinomas in stratified epithelia [J]. Cancer Cell, 2007, 12(4): 313-327.

[29] SYED V. TGF-β signaling in cancer [J]. J Cell Biochem, 2016, 117(6): 1279-1287.

[30] IVANOV S V, IVANOVA A V, SALNIKOW K, et al. Two novel VHL targets, TGFBI (BIGH3) and its transactivator KLF10, are up-regulated in renal clear cell carcinoma and other tumors[J]. Biochem Biophys Res Commun, 2008, 370(4): 536-540.

[31] SHANG D H, LIU Y T, YANG P Q, et al. TGFBI-promoted adhesion, migration and invasion of human renal cell carcinoma depends on inactivation of von Hippel-Lindau tumor suppressor[J]. Urology, 2012, 79(4): 966.e1-966.e7.

[32] FICO F, SANTAMARIA-MARTÍNEZ A. TGFBI modulates tumour hypoxia and promotes breast cancer metastasis[J]. Mol Oncol, 2020, 14(12): 3198-3210.

[33] HAN B, CAI H L, CHEN Y, et al. The role of TGFBI (βig-H3) in gastrointestinal tract tumorigenesis[J]. Mol Cancer, 2015, 14: 64.

[34] WEN G Y, PARTRIDGE M A, LI B Y, et al. TGFBI expression reduces *in vitro* and *in vivo* metastatic potential of lung and breast tumor cells[J]. Cancer Lett, 2011, 308(1): 23-32.