

肝细胞癌内源竞争 RNA 调控网络的构建

王瑶雯¹, 刘紫薇^{1*}

(1. 武汉工程大学, 湖北省武汉市, 430205; *通讯作者, liuzw04@wit.edu.cn)

摘 要: 肝细胞癌 (Hepatocellular Carccinoma, HCC) 是一种常见的恶性肿瘤, 因其发病隐蔽, 发展迅速, 探索其调控机制并寻找分子靶点对 HCC 的治疗至关重要。内源竞争 RNA (competing endogenous RNAs, ceRNA) 假说表明, mRNA 与非编码 RNA 可竞争性结合 miRNA 应答元件, 参与转录后调控, 形成复杂的调控网络。然而, 目前 HCC 的 ceRNA 调控网络研究还不够深入。本文通过 R 筛选 TCGA 数据库中 HCC 差异表达的 mRNA lncRNA 和 miRNA, 并借助 cBioPortal、Metascape、Tarbase、miRDB 和 Lncbase 等工具构建 HCC ceRNA 调控网络, 评估 ceRNA 网络中 RNA 分子的表达差异与预后价值。最终筛选出纳入 HCC ceRNA 调控网络的 RNA 分子包括两种 mRNA、三种 miRNA 和两种 lncRNA。

关键词: 肝细胞癌; ceRNA; 调控网络

引言

肝细胞癌 (Hepatocellular Carcinoma, HCC) 约占原发性肝癌的 90%[1]。根据《2020 年全球癌症发病率、死亡率和患病率》, 在全球范围内, HCC 是第六大常见癌症, 也是癌症相关死亡的第三大常见原因[2]。由于早期无症状阶段的诊断延迟, HCC 往往发展到严重的侵袭性阶段, 严重影响患者的生存。此外, HCC 对化疗和放疗等常规癌症治疗表现出显著耐药性[3]。因此, 研究 HCC 发生和发展中的分子机制, 寻找 HCC 治疗的合适靶点, 是有效治疗 HCC 的关键策略。内源竞争 RNA (competing endogenous RNAs, ceRNA) 属于非编码 RNA (Non-coding RNA, ncRNA)。ceRNA 假说提出, 微小 RNA (microRNA, miRNA) 可以通过结合 mRNA 导致基因沉默, 而 ceRNA 可以通过竞争性结合 miRNA 来调节基因表达[4]。ceRNA 可以通过应答元件 (microRNA response elements, MREs) 与 miRNA 结合从而使 miRNA 失效。ceRNA 调控网络包括 mRNA、编码基因的假基因、长链 ncRNA 和 miRNA 等[5]。研究表明, ceRNA 网络在各种恶性肿瘤的发展中起着至关重要的作用。例如, HOXD-AS1 与 miR-130a-3p 竞争性结合, 阻止 miRNA 介导的 SOX4 降解, 从而激活 EZH2 和 MMP2 的表达并促进 HCC 转移[6]。DLEU2L 与 BRCA2 竞争结合 miR-210-3, 抑制胰腺癌细胞的增殖、侵袭和迁移[7]。因此, 找到新的 ceRNA 目标节点, 不仅能够丰富 RNA 介导的 HCC 等癌症发病机制的知识库, 也有助于开发新的癌症治疗策略。本研究通过构建 HCC ceRNA 调控网络, 探究了其中重要的 ceRNA。

1 材料与方法

1.1 TCGA 数据的下载与处理

从 TCGA 数据库下载 HCC 的 mRNA 表达数据、肿瘤突变数据及患者临床生存数据等, 使用 R (V4.4.3) 对数据进行归一化处理。本研究的 RNA-seq 数据包含了 371 例 HCC 肿瘤样本与 50 例正常样本。

1.2 差异基因的筛选

使用 R 整理 HCC RNA-seq 矩阵文件, 使用程序包 “edgeR” 对 HCC 肿瘤样本与正常样本之间的 mRNA 表达差异进行分析。设置筛选条件: $\log_2FC > 1$, $p < 0.05$ 。

1.3 PPI 网络的构建与关键基因的筛选

选取差异表达显著的 mRNA，利用 STRING 在线分析工具对它们进行分析。设置：条件为多基因研究、物种为人源，置信分数（confidence）为 0.4，去除游离基因，构建蛋白质互作网络（Protein-Protein Interaction Networks, PPI），以探究其蛋白质相互作用关系。在 STRING 网站中将 PPI 网络以“tsv”格式下载，并导入 Cytoscape（V3.10.2），利用 Cytoscape 对 PPI 网络进行可视化，建立可视化网络后，在 CytoHubba 插件中选择不同算法对 PPI 网络进行节点分析，评估各个基因在 PPI 网络中占有的位置，结合十种算法分析结果取交集选取差异表达基因中节点评分位于前 15 位的关键基因用于预测下游 miRNA。

1.4 KEGG 富集分析

使用 KEGG 数据库对 PPI 网络中筛选出的差异表达基因中节点评分前 15 个关键基因进行分析，得到这些 mRNA 在 HCC 中显著富集的信号通路。

1.5 GO 富集分析

利用 R 程序包“clusterProfiler”对关键 mRNA 进行 GO 功能富集，分析它们的生物学功能相关生物学过程（biological process, BP）、细胞组成成分（Cellular Components, CC）、分子生物功能（Molecular Function, MF），从而深入了解其机制和功能。

1.6 HPA 数据库分析

利用 HPA 数据库（Human Protein Atlas）分析 KARS1 在 HCC 肿瘤组织与正常组织的免疫组化染色情况。HPA 数据库是一个基于蛋白组学、转录组学以及系统生物学数据的数据库，可以绘制组织、细胞、器官等图谱。

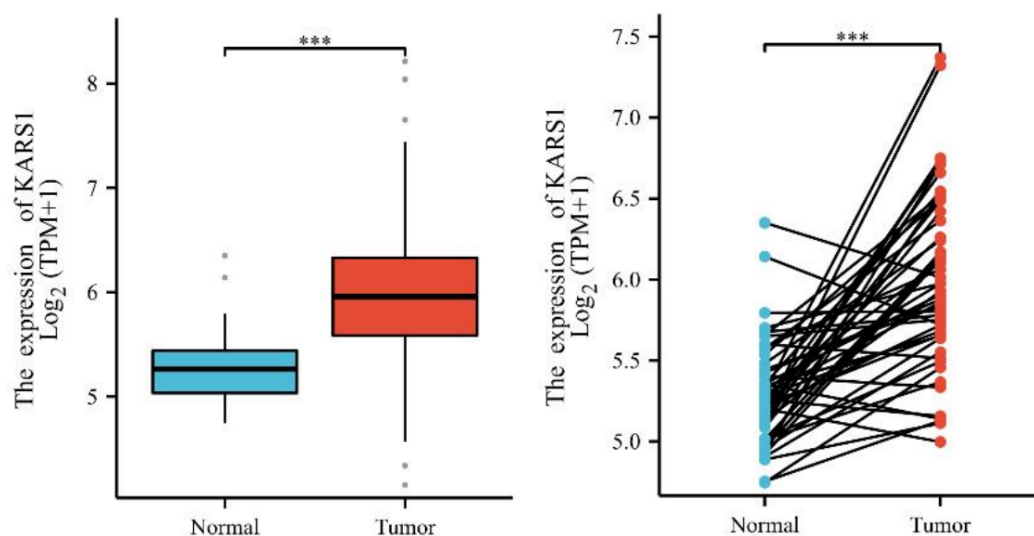
1.7 ceRNA 网络的构建

筛选出显著差异表达的 lncRNA、mRNA 和 miRNA 用于构建 HCC 的 ceRNA 调控网络。利用 miRcode 数据库预测 miRNA-lncRNA 之间的靶点关系。利用 miRDB 数据库、miRTarBase 数据库和 TargetScan 数据库预测 miRNA-mRNA 之间的靶点关系。为了提高结果的可靠性，只有在这三个数据库中均能查到的 miRNA-mRNA 靶点关系才可作为构建 ceRNA 调控网络的候选基因。基于 lncRNA-miRNA 之间以及 miRNA-mRNA 之间相互的作用关系完成 HCC 初始 ceRNA 调控网络的构建，并利用 Cytoscape 将其可视化。

2 结果

2.1 HCC KARS1 差异表达分析

本研究以 KARS1 为 HCC 基因差异表达分析的基准参考基因，首先通过对比分析 KARS1 在 HCC 肿瘤样本与正常样本中的表达水平差异，说明 KARS1 在 HCC 中的潜在促癌作用。差异表达分析样本分为非配对样本与配对样本，其中非配对样本包含 371 例肿瘤样本与 50 例正常样本。结果显示，如图 1（A）所示，非配对样本的 HCC 肿瘤组织中 KARS1 表达水平显著高于正常组织中 KARS1 表达水平，表明 KARS1 有潜在促进 HCC 发生发展的作用。配对样本包含 50 组配对的 HCC 肿瘤样本与正常样本。分析结果显示，如图 1（B）所示，配对样本的 HCC 肿瘤组织中 KARS1 表达水平也显著高于正常组织中 KARS1 表达水平，同样说明 KARS1 在 HCC 中有潜在的促癌作用。



(A) 非配对样本的 KRAS1 表达差异图 (B) 配对样本的 KRAS1 表达差异图

图 1 HCC KRAS1 差异表达分析

2.2 KARS1 在 HCC 中的促癌作用及预后价值

为了进一步探究 KARS1 在 HCC 中的潜在促癌作用，基于 HPA 数据库对比了 HCC 肿瘤组织和正常肝组织的 KARS1 表达与免疫组化结果。结果表明，如图 2 (A) 和图 2 (B) 所示，与正常肝组织相比，KARS1 在 HCC 肿瘤组织中呈表达上调。如图 2 (C) 所示，HCC 肿瘤标本的免疫组化染色结果显示，与正常肝组织相比，KARS1 在 HCC 肿瘤组织的染色范围更广，颜色更深。如图 2 (D) 所示，Kaplan-Meier 生存曲线显示，HCC 患者的 KARS1 高表达与较差的总生存期显著相关。

2.3 HCC 差异表达 ceRNA 的筛选

为了构建与 KARS1 相关的 HCC ceRNA 网络，以作为 HCC 患者的潜在预后模型，将 TCGA 数据库的 HCC 测序数据分为 KARS1 高表达(KARS1^{high})和 KARS1 低表达(KARS1^{low})两组，根据三种 ceRNA——lncRNA、miRNA 和 mRNA 各自的筛选标准，分别筛选其差异表达基因 (Differentially expressed genes, DEGs)。如图 3 (A) 所示，筛选得到了 852 个差异表达 lncRNA (DElncRNAs)，其中 624 个表达上调，228 个表达下调。如图 3 (B) 所示，筛选得到了 47 个差异表达 miRNA (DEmiRNAs)，其中 11 个表达上调，36 个表达下调。如图 3 (C) 所示，筛选得到了 1129 个差异表达 mRNA (DEmRNAs)，其中 879 个表达上调，250 个表达下调。此外，HCC 肿瘤样本和正常样本的基因差异表达分析分别鉴定出了 3650 个 DElncRNAs，其中 1769 个表达上调，1881 个表达下调；523 个 DEmiRNAs，其中 57 个表达上调，466 个表达下调；10418 个 DEmRNAs，其中 5104 个表达上调，5314 个表达下调。如图 3 (D) 所示，该热图显示了前 15 个显著 DElncRNAs 与每例 HCC 肿瘤样本中 KARS1 表达水平的相关性。如图 3 (E) 所示，该热图显示了前 15 个显著 DEmiRNAs 与每例 HCC 肿瘤样本中 KARS1 表达水平的相关性。如图 3 (F) 所示，该热图显示了前 15 个显著 DEmRNAs 与每例 HCC 肿瘤样本中 KARS1 表达水平的相关性。

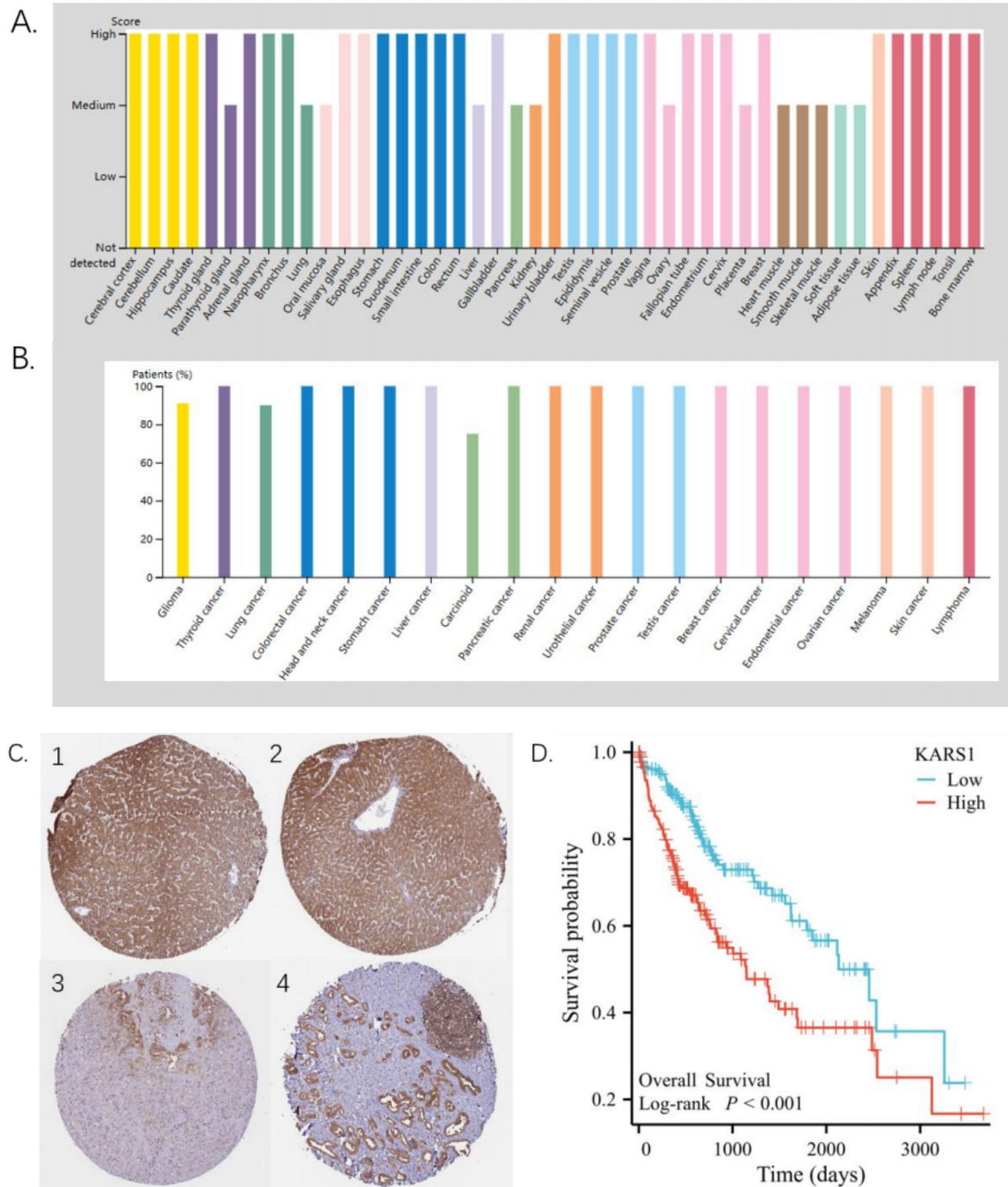


图2 KARS1 在 HCC 中的促癌作用及预后价值 (A) KARS1 在正常组织中的表达水平 (B) KARS1 在泛癌组织中的表达水平 (C) HCC 肿瘤组织和正常肝组织的免疫组化染色结果图 (D) KARS1 的 HCC Kaplan-Meier 生存曲线, 低表达 ($n=185$), 高表达 ($n=184$)

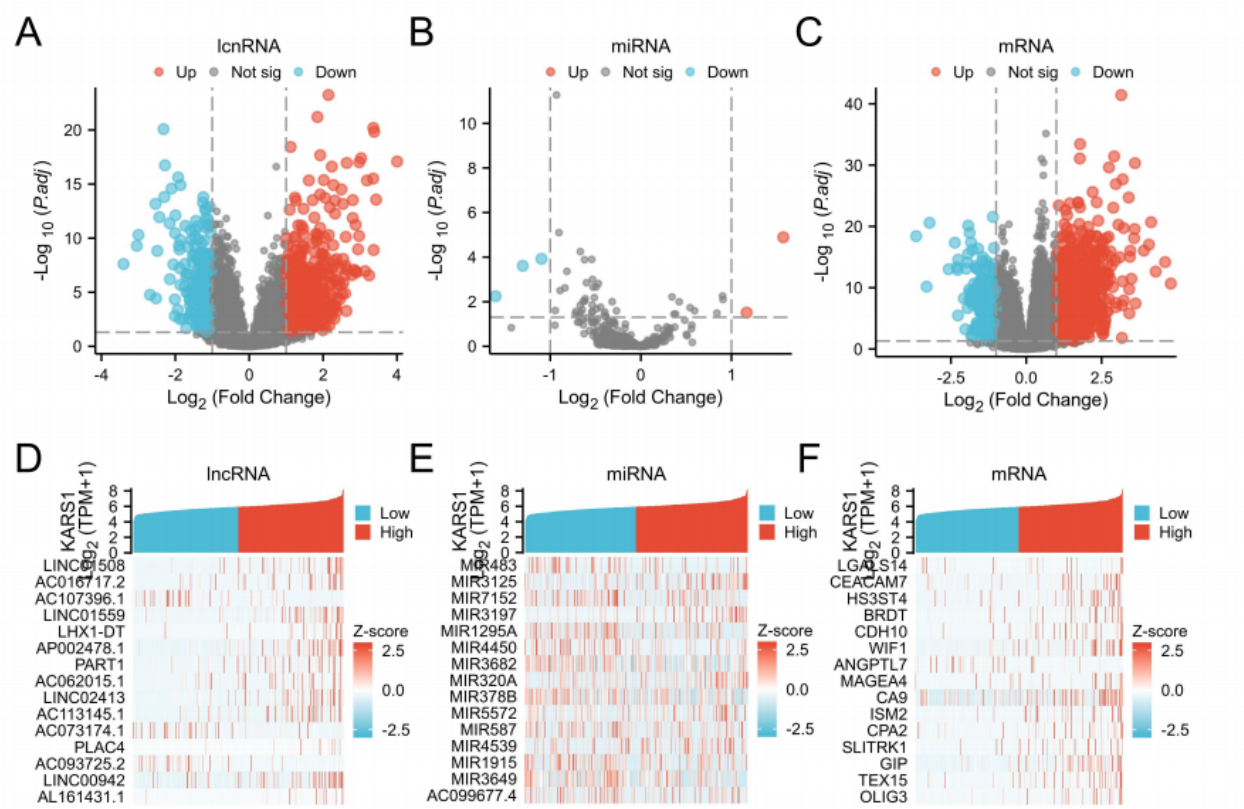


图 3 HCC $KARS1^{\text{high}}$ 组与 $KARS1^{\text{low}}$ 组之间的 lncRNA、miRNA 和 mRNA 差异表达分析 (A) lncRNA 的差异表达分析火山图 ($|\log_2\text{FC}| > 0.5$, $p < 0.05$) (B) miRNA 的差异表达分析火山图 ($|\log_2\text{FC}| > 0.3$, $p < 0.05$) (C) mRNA 的差异表达分析火山图 ($|\log_2\text{FC}| > 1$, $p < 0.05$) (D) 15 个显著 DElncRNAs 与 HCC 样本中 KARS1 的相关性热图 (E) 15 个显著 DEmiRNAs 与 HCC 样本中 KARS1 的相关性热图 (F) 15 个显著 DEmRNAs 与 HCC 样本中 KARS1 的相关性热图

2.4 HCC lncRNA-miRNA-mRNA 三重调控网络的构建及富集分析

为了构建 HCC lncRNA-miRNA-mRNA 三重调控网络，接下来对 $KARS1^{\text{high}}$ 组、 $KARS1^{\text{low}}$ 组及 HCC 肿瘤样本组、正常样本组进行了联合分析。将 DElncRNAs 放入 Metascape 数据库中，以识别潜在的靶向这些 DElncRNAs 的 miRNA，在与 47 个 DEmiRNAs 交叉后，从预测的 miRNA 中选择了 5 个。然后分别使用 miRDB 和 TargetScan 数据库预测这 5 个 miRNA 下游的靶 mRNA，为提高预测的准确性，选择两个数据库均识别到的 mRNA 为候选基因，从 10418 个 DEmRNAs 中鉴定出 519 个候选 mRNA。如图 4 (A) 所示，lncRNA 调控网络中包括 2 个 lncRNA，其中 1 个高表达，1 个低表达。如图 4 (B) 所示，miRNA 调控网络中包括 5 个 miRNA，其中 4 个高表达，1 个低表达。如图 4 (C) 所示，mRNA 调控网络中包括 519 个 mRNA，其中 289 个高表达，230 个低表达。如图 4 (D) 所示，该图显示了 HCC 的 lncRNA-miRNA-mRNA 三重调控网络。如图 4 (E) 所示，mRNA 调控网络的 GO 和 KEGG 富集分析结果表明，这 519 个 mRNA 显著富集于与 NABA 基质体相关、上皮细胞分化及细胞连接组织等相关的通路。

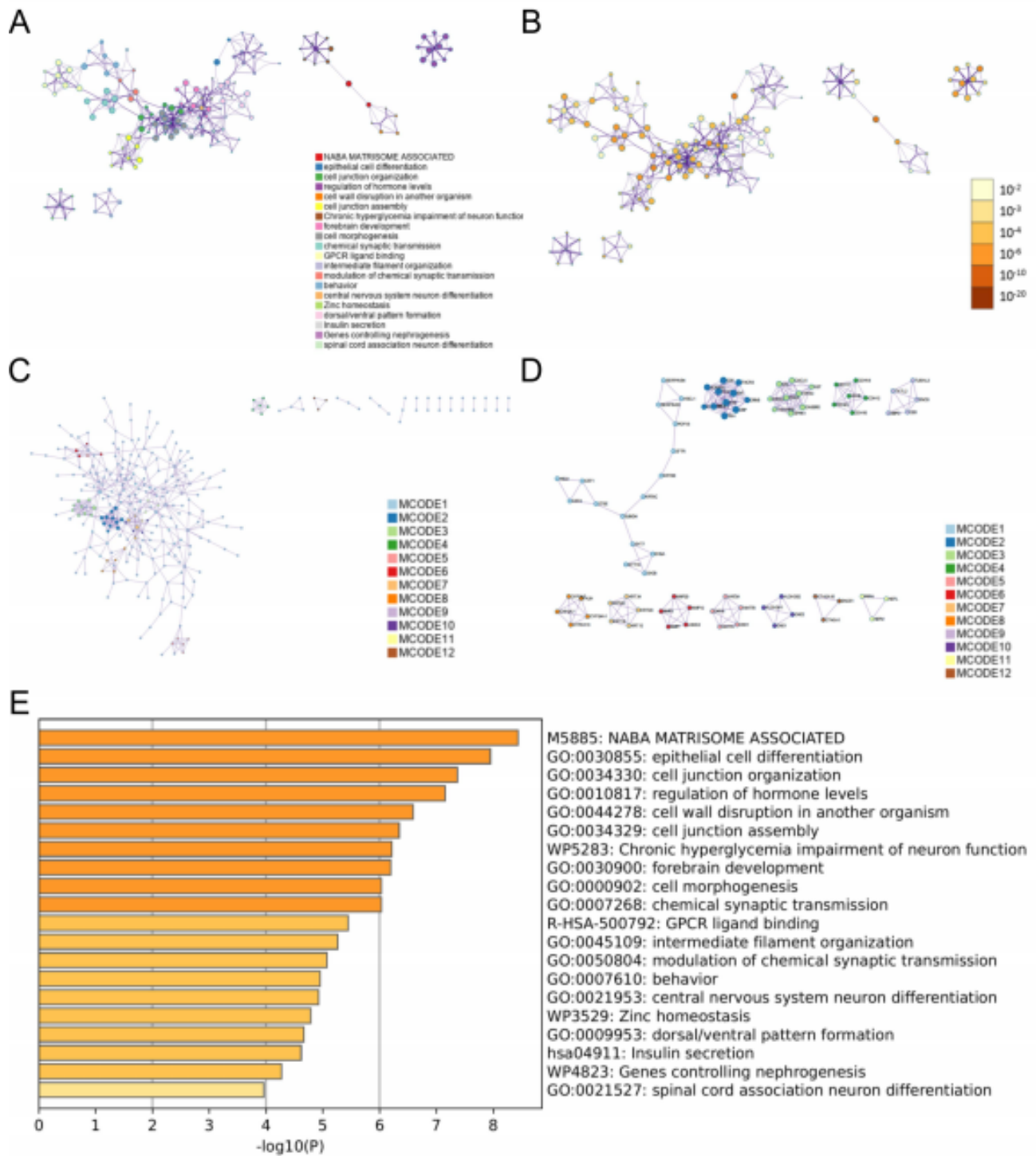


图 4 HCC lncRNA-miRNA-mRNA 三重调控网络的构建及 GO、KEGG 富集分析 (A) KARS1 相关的 lncRNA 调控网络 (B) KARS1 相关的 miRNA 调控网络 (C) KARS1 相关的 mRNA 调控网络 (D) KARS1 相关的 lncRNA-miRNA-mRNA 三重调控网络 (E) mRNA 网络的富集分析条形图

2.5 ceRNA 网络的构建和验证及具有 HCC 特异性预后价值模型的选择

为了构建具有重要预后价值的 KARS1 相关 HCC ceRNA 调控网络, 接下来对比分析了 lncRNA-miRNA-mRNA 三重调控网络中的 hub RNA 分别在 HCC 肿瘤样本和正常样本中的表达水平。结果表明, 这些 RNA 在 HCC KARS1^{high} 和 KARS1^{low} 样本组对比下的表达水平不同于 HCC 肿瘤样本和正常样本组对比下的表达水平, 如图 5 (A) 和 (B) 所示, LINC-ROR 和 GAS5 两个 lncRNA 呈表达上调; 如图 5 (E) - (H) 所示, miR-324-3p、miR-126-5p、miR-145-5p、miR-193a-3p4 四个 miRNA 呈表达上调, miR-210-3p 一个 miRNA 呈表达下调; 如图 5 (C)、(D) 和 (J) - (N) 所示, RFX6 和 DLX1 两个 mRNA 呈表达上调, IL13、SGPP2、LAMC2、ADGRF1、HS3ST4 五个 mRNA 表达变化不显著。此外, 对这些 RNA 在 TCGA 中 50 对 HCC 配对样本中的表达水平进行了验证, 结果与上述分析几乎一致。

为了确定这些 RNA 与 HCC 预后的相关性, 使用 Kaplan-Meier 生存曲线和 log-rank 检验对 HCC 患者进行了总生存期分析。结果显示, 如图 6 所示, 有 1 个 lncRNA (LINC-ROR)、3 个 miRNA (miR-324-p、miR-126-5p 和 miR-145-5p) 和 4 个 mRNA (IL13、SGPP2、RFX6 和 LAMC2) 与 HCC 预后密切相关, 其中 LINC-ROR、IL13 和 SGPP2 均 $p < 0.1$, 可能为潜在的 KARS1 相关 HCC ceRNA 网络关键调控因子。

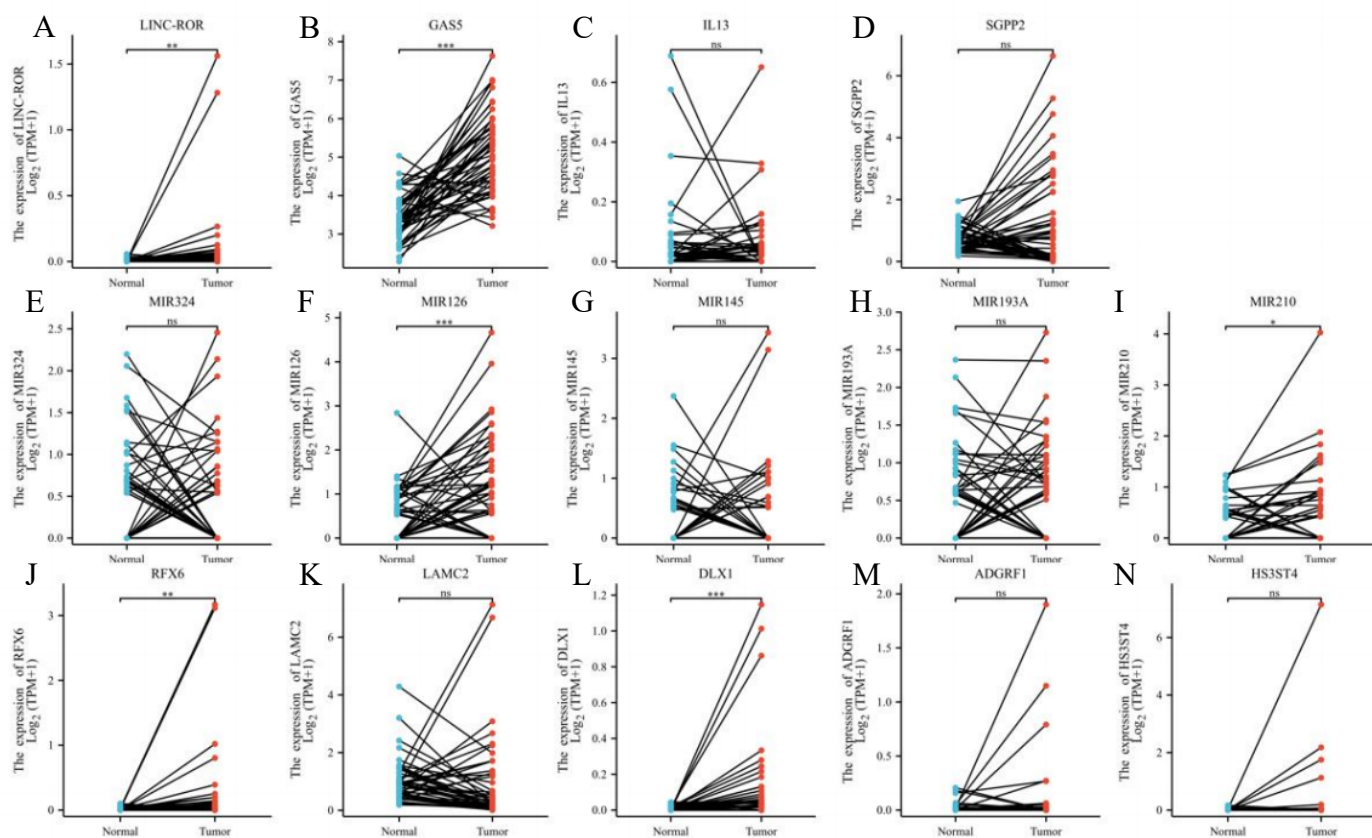
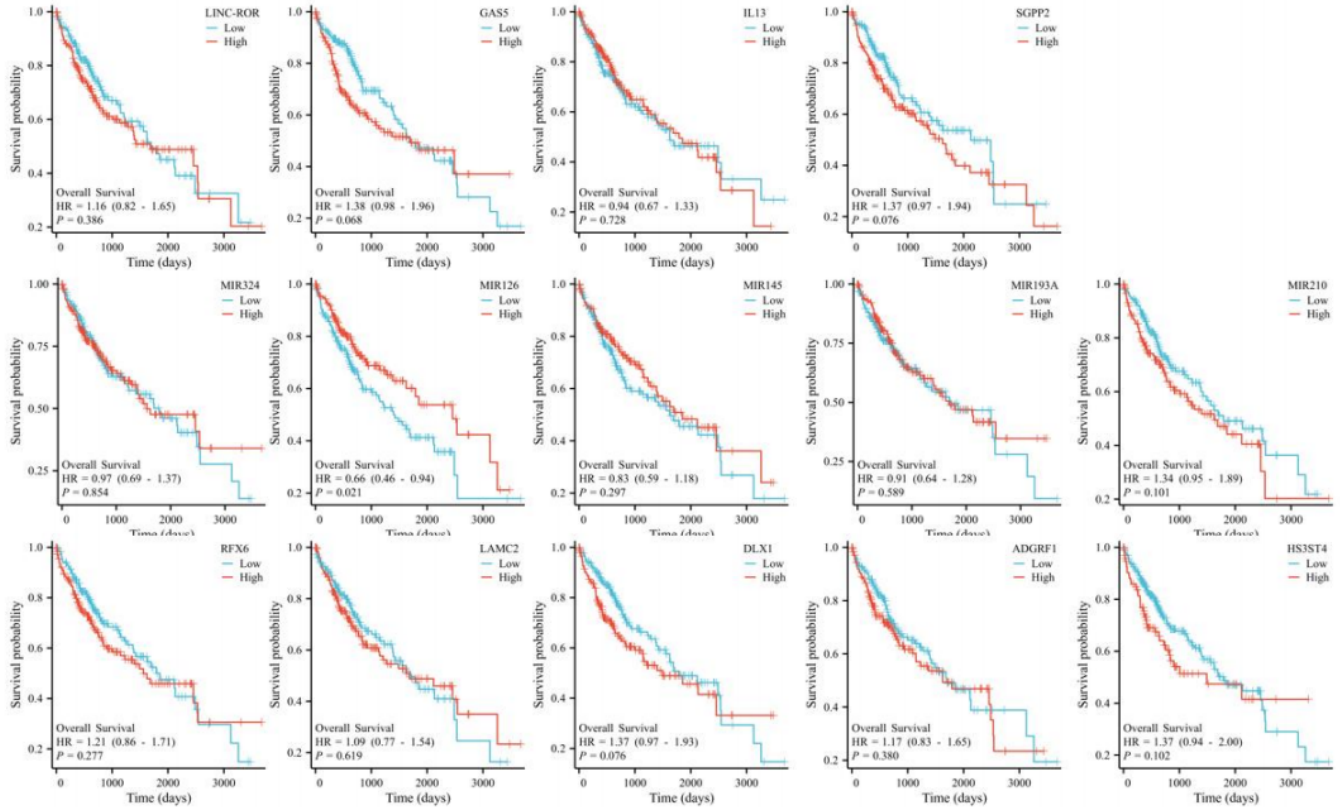


图5 关键 DEGs 分别在 HCC 肿瘤样本和正常样本中的表达水平

图6 关键 DEGs 的 KM 生存分析 ($p < 0.05$)

2.6 HCC ceRNA 调控网络的构建及各通路的碱基配对关系

为了构建与 HCC 预后相关的 ceRNA 调控网络，从前述 2 个关键 lncRNA、5 个关键 miRNA 和 7 个关键 mRNA 中选取了高表达与 HCC 预后不良相关的 lncRNA，靶向这些 lncRNA 但与预后良好相关的 miRNA，以及这些 miRNA 靶向的高表达与 HCC 预后不良相关的 mRNA。如图 7 (A) 所示，在 ceRNA 调控网络中，包含了“LINC-ROR-Hsa-miR-145-5p”、“GAS5-Hsa-miR-126-5p”和“GAS5-miR-324-3p”3 条 ceRNA 调控机制通路。如图 7 (B) 所示，该图显示了 LINC-ROR (lncRNA) 与 Hsa-miR-145-5p (miRNA) 的碱基配对关系。如图 7 (C) 所示，该图显示了 GAS5 (lncRNA) 与 Hsa-miR-126-5p (miRNA) 的碱基配对关系。如图 7 (D) 所示，该图显示了 GAS5 (lncRNA) 与 Hsa-miR-324-3p (miRNA) 的碱基配对关系。

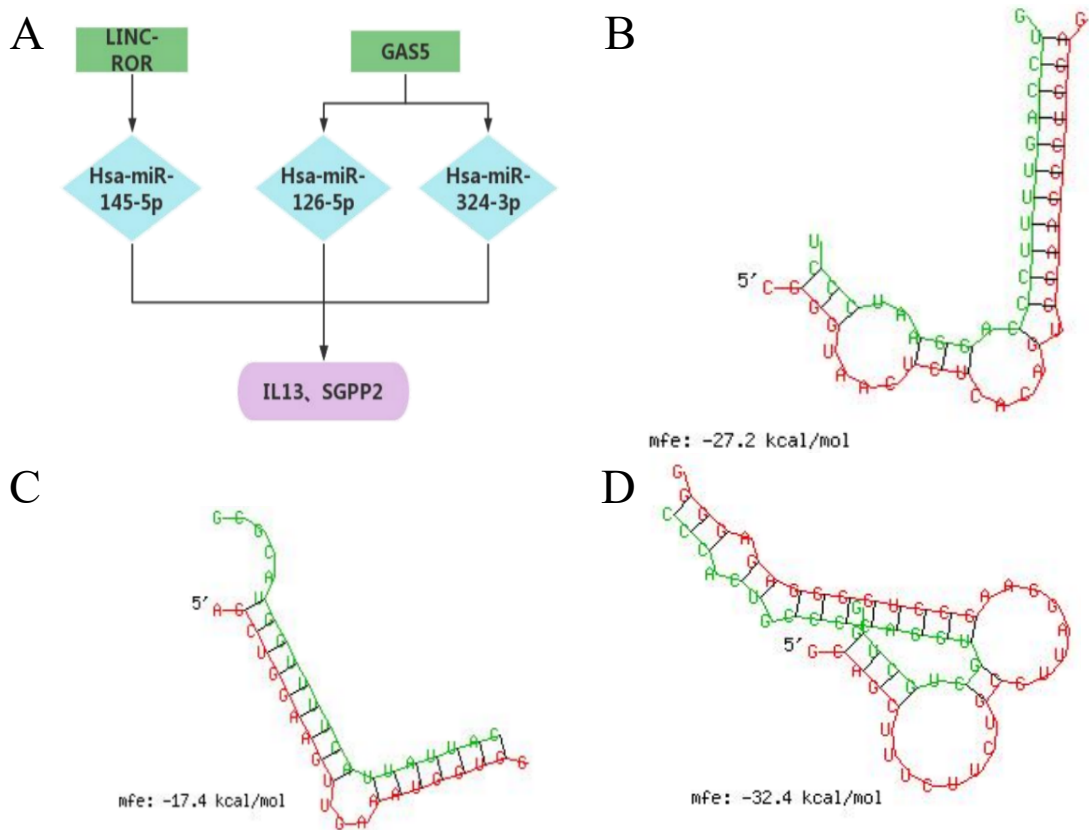


图 7 HCC ceRNA 调控网络的构建及各通路的碱基配对关系 (A) 与 HCC 预后相关的 ceRNA 调控网络，其中绿色模块代表 lncRNA，蓝色模块代表 miRNA，紫色模块代表 mRNA (B) LINC-ROR 与 miR-145-5p 的碱基配对关系图 (C) GAS5 与 miR-126-5p 的碱基配对关系图 (D) GAS5 与 miR-324-3p 的碱基配对关系图

3 讨论

肝细胞癌 (Hepatocellular Carcinoma, HCC) 是最常见的原发性肝癌类型，也是全球最致命的癌症之一。与 HCC 相关的主要已知风险因素是病毒 (慢性乙型和丙型肝炎)、代谢 (糖尿病和非酒精性脂肪肝)、毒性 (酒精和黄曲霉毒素) 和免疫系统相关疾病[8]。当早期诊断且肿瘤大小<5cm 时，肝移植和手术切除是有效的 HCC 治疗选择。然而，由于 HCC 在早期大多没有症状，大多数患者在晚期就诊。晚期 HCC 采用射频消融、肝动脉化疗栓塞、酪氨酸激酶抑制剂和免疫疗法治疗，但是，随着治疗耐药性和疾病复发，这些治疗方式不能显著延长 HCC 患者寿命[9]。到目前为止，尽管人们已对 HCC 进行了广泛的研究，治疗 HCC 的新途径也越来越多样化，但对于 HCC 发生和发展的生物学机制仍未完全了解清楚，特别是与 ceRNA 相关的机制。ceRNA 假说表明，参与 ceRNA 网络的可以是所有类型的 RNA 转录本，它们通过竞争性结合 miRNA 应答元件，形成跨转录组的大规模调控网络，同时破坏 ceRNA 和 miRNA 之间的动态平衡，会极大的影响机体的生物学功能，进而导致癌症发生[10]。多项研究表明，ceRNA 网络在许多肿瘤的发生和发展中具有重要作用[11]。但是，关于 ceRNA 网络在 HCC 中的研究还不够全面，仍有许多待挖掘的 ceRNA 分子。寻找与 HCC 预后相关的 ceRNA 分子，构建 ceRNA 调控网络，对于了解 HCC 的分子调控机制及发现新的药物治疗靶点具有重要意义。

首先预测并得到了 HCC 中靶向 7 个关键 mRNA 的 300 个 miRNA，这 7 个关键 mRNA 在 HCC 中均表现为表达水平上调。通过对 300 个 miRNA 进行 KM 生存分析，得到了与 HCC 患者总体生存率显著相关的 5 个 miRNA，其中 3 个关键 miRNA (miR-145、miR-126、miR-324) 在 HCC 中均表现为表达水平下调，且它们的高表达均与 HCC 预后良好相关，将它们纳入 ceRNA 网络中。研究表明，miR-145-5p 能够通过调节前列腺癌的 SOX11/MYCN 轴抑制神经内分泌分化和肿瘤生长[12]。血小板衍生的 miR-126-3p 能够直接靶向 AKT2 并在乳腺癌细胞中发挥抗肿瘤作用[13]。假基因 PLGLA 能够通过调节 miR-324-3p/GLYATL1 轴对 HCC 发挥抗肿

瘤作用[14]。然后,预测了关键 miRNA 的上游 lncRNA,并对这些 lncRNA 进行了表达水平分析和 KM 生存分析,鉴定出 2 个关键 lncRNA (LINC-ROR 和 GAS5),它们在 HCC 中均表现为表达水平上调。研究表明,LINC-ROR 能够通过激活上皮-间充质转化促进胃癌细胞的侵袭和转移[15]。GAS5 能够通过激活人肝母细胞瘤 HepG2 细胞中的 CHOP 依赖性内质网应激通路诱导细胞凋亡并抑制肿瘤生长[16]。随后构建出 HCC ceRNA 调控网络,其中包含 3 条与 HCC 预后相关的调控机制通路,分别为“LINC-ROR-Hsa-miR-145-5p”、“GAS5-Hsa-miR-126-5p”和“GAS5-miR-324-3p”。

综上所述,本研究利用多种生物信息学方法构建了 HCC ceRNA 调控网络,揭示了 HCC 分子机制上的一些特点,并寻找到了网络中与患者生存率相关的预后关键基因,这些基因是潜在的 HCC 的预后分子标志物,这些关键基因有望为后续深入研究提供线索。

参考文献

- [1] KOSHY A. Evolving Global Etiology of Hepatocellular Carcinoma (HCC): Insights and Trends for 2024[J]. *Journal of Clinical and Experimental Hepatology*, 2025, 15(1): 102406.
- [2] BICER F, KURE C, OZLUK A A, et al. Advances in Immunotherapy for Hepatocellular Carcinoma (HCC)[J]. *Current Oncology*, 2023, 30(11): 9789-9812.
- [3] SHAH M, SARKAR D. HCC-Related lncRNAs: Roles and Mechanisms[J]. *International Journal of Molecular Sciences*, 2024, 25(1): 39.
- [4] CAI S, DU R, ZHANG Y, et al. Construction and Comprehensive Analysis of ceRNA Networks and Tumor-Infiltrating Immune Cells in Hepatocellular Carcinoma With Vascular Invasion[J]. *Frontiers in Bioinformatics*, 2022, 2: 836981.
- [5] LI X, PENG C, LIU H, et al. Constructing methylation-driven ceRNA networks unveil tumor heterogeneity and predict patient prognosis[J]. *Human Molecular Genetics*, 2025, 34(3): 251-264.
- [6] WANG H, HUO X, YANG X R, et al. STAT3-mediated upregulation of lncRNA HOXD-AS1 as a ceRNA facilitates liver cancer metastasis by regulating SOX4[J]. *Molecular Cancer*, 2017, 16(1): 136.
- [7] XU F, WU H, XIONG J, et al. Long Non-coding RNA DLEU2L Targets miR-210-3p to Suppress Gemcitabine Resistance in Pancreatic Cancer Cells via BRCA2 Regulation[J]. *Frontiers in Molecular Biosciences*, 2021, 8: 645365.
- [8] FANG J, CELTON-MORIZUR S, DESDOUETS C. NAFLD-Related HCC: Focus on the Latest Relevant Preclinical Models[J]. *Cancers*, 2023, 15(14): 72.
- [9] KIMURA T, FUJIWARA T, KAMEOKA T, et al. The Current Role of Stereotactic Body Radiation Therapy (SBRT) in Hepatocellular Carcinoma (HCC)[J]. *Cancers*, 2022, 14(18): 87.
- [10] WU S, ZHONG B, YANG Y, et al. ceRNA networks in gynecological cancers progression and resistance[J]. *Journal of Drug Targeting*, 2023, 31(9): 920-930.
- [11] LUO Z, RONG Z, ZHANG J, et al. Circular RNA circCCDC9 acts as a miR-6792-3p sponge to suppress the progression of gastric cancer through regulating CAV1 expression[J]. *Molecular Cancer*, 2020, 19(1): 86.
- [12] JI S, SHI Y, YANG L, et al. miR-145-5p Inhibits Neuroendocrine Differentiation and Tumor Growth by Regulating the SOX11/MYCN Axis in Prostate cancer[J]. *Frontiers in Genetics*, 2022, 13: 790621.
- [13] SIBILANO M, TULLIO V, ADORNO G, et al. Platelet-Derived miR-126-3p Directly Targets AKT2 and Exerts Anti-Tumor Effects in Breast Cancer Cells: Further Insights in Platelet-Cancer Interplay[J]. *International Journal of Molecular Sciences*, 2022, 23(10): 67.
- [14] BAO L, LI P, ZHAO H, et al. Pseudogene PLGLA exerts anti-tumor effects on hepatocellular carcinoma through modulating miR-324-3p/GLYATL1 axis[J]. *Digestive and Liver Disease*, 2022, 54(7): 918-926.
- [15] LIU M, ZHANG M, YIN H. Linc-ROR promotes invasion and metastasis of gastric cancer by activating epithelial-mesenchymal transition[J]. *Indian Journal of Pathology and Microbiology*, 2022, 65(3): 545-550.
- [16] ZHANG W Y, ZHAN H L, LI M K, et al. Long noncoding RNA Gas5 induces cell apoptosis and inhibits tumor growth via activating the CHOP-dependent endoplasmic reticulum stress pathway in human hepatoblastoma HepG2 cells[J]. *Journal of Cellular Biochemistry*, 2022, 123(2): 231-247.