

泛癌分析长链非编码RNA MIR22HG的表达特征

王惠¹, 黎雯雯², 张冬艳¹

南方医科大学南方医院¹放疗科,²肝病中心, 广东 广州 510515

摘要:目的 泛癌分析长链非编码RNA(lncRNA)MIR22HG的表达及其与临床特征的关系。方法 R语言分析癌症基因组图谱(TCGA)中MIR22HG在不同肿瘤组织中的表达及其与临床分期、淋巴结转移、肿瘤突变负荷(TMB)及微卫星不稳定(MSI)的关系;应用TIMER比较MIR22HG表达与免疫浸润的关系;使用cBioPortal分析MIR22HG的突变频率及其与预后的关系;使用GDSC在线数据库分析MIR22HG与化疗药物敏感性的关系。利用GEO数据库肝癌数据集和12对肝癌标本验证MIR22HG在肝癌中的表达及其与索拉非尼治疗反应的关系;分析MIR22HG表达与索拉非尼治疗预后的关系;在肝癌细胞株HCC-LM3中过表达MIR22HG(NC组, MIR22HG过表达组),在肝癌细胞株MHCC-97H中敲除MIR22HG(NC组, sh-MIR22HG组),采用CCK-8检测肝癌细胞中MIR22HG对索拉非尼IC₅₀的影响。结果 MIR22HG在大多数肿瘤中低表达($P<0.05$),且多数肿瘤中MIR22HG基因存在缺失突变,其突变与肿瘤不良预后相关($P<0.05$)。MIR22HG的表达水平与多种肿瘤的临床分期及淋巴结转移相关($P<0.05$)。在多种肿瘤中MIR22HG的表达水平与TMB、MSI、免疫评分、免疫检查点相关基因表达及化疗药的药物敏感性显著相关($P<0.05$)。6种常见免疫细胞中,中性粒细胞浸润水平与MIR22HG的表达水平相关性最强,尤其在乳腺癌、直肠癌及肾乳头状细胞癌中明显。多个数据集的分析结果验证MIR22HG在肝癌中低表达,且其低表达与肝癌进展相关($P<0.05$)。MIR22HG低表达的患者经索拉非尼治疗后预后较好($HR=2.94, P=0.075$),其低表达能有效预测肝癌患者索拉非尼治疗效果($AUC=0.8095$)。过表达MIR22HG能降低肝癌细胞对索拉非尼的敏感性($IC_{50\text{ NC}} \text{ vs } IC_{50\text{ MIR22HG}}=7.731 \text{ vs } 15.61$);而敲除MIR22HG的表达则能增加肝癌细胞对索拉非尼的敏感性($IC_{50\text{ NC}} \text{ vs } IC_{50\text{ sh-MIR22HG}}=7.986 \text{ vs } 5.085$)。结论 MIR22HG的表达与多种肿瘤的分期、有否淋巴结转移、肿瘤突变负荷、微卫星不稳定、免疫细胞浸润及化疗药物敏感性相关。

关键词:长链非编码RNA;MIR22HG;泛癌分析;肝癌;索拉非尼

Pan-cancer analysis of the expression pattern of long non-coding RNA MIR22HG

WANG Hui¹, LI Wenwen², ZHANG Dongyan¹

¹Department of Radiation Oncology, ²Hepatology Unit and Department of Infectious Diseases, Nanfang Hospital, Southern Medical University, Guangzhou 510515, China

Abstract: Objective To conduct a pan-cancer analysis of the expression of long non-coding RNA (lncRNA) MIR22HG and explore its association with clinical characteristics. **Methods** We analyzed the expression of MIR22HG in different tumors and its association with clinical staging, lymph node metastasis, tumor mutation burden (TMB) and microsatellite instability (MSI) using R package based on the Cancer Genome Atlas (TCGA) datasets. The relationship between MIR22HG expression and infiltrating immune cells was analyzed using TIMER algorithm. The association of MIR22HG gene alteration frequency with the clinical outcomes was examined using cBioPortal online software. Data from Genomics of Drug Sensitivity in Cancer (GDSC) were used to analyze the relationship between MIR22HG and the sensitivity of chemotherapy drugs. We specifically analyzed MIR22HG expression in hepatocellular carcinoma (HCC) and its correlation with sorafenib treatment using GEO database and verified the results in 12 pairs of HCC specimens. Kaplan-Meier analysis was performed to analyze the correlation of MIR22HG with the outcomes of sorafenib treatment. We also tested the effects of MIR22HG overexpression and knockdown on IC₅₀ of sorafenib in HCC cells. **Results** MIR22HG was downregulated in most tumors ($P<0.05$), where its deletion mutations were frequent, and associated with a poor prognosis ($P<0.05$). In many tumors, MIR22HG expression level was correlated with clinical stage, lymph node metastasis, TMB, MSI, immune cell infiltration, immune checkpoint-related genes, and sensitivity to common chemotherapeutic drugs ($P<0.05$). Among the 6 common infiltrating immune cells in cancers, neutrophil infiltration had the strongest correlation with MIR22HG expression level, especially in breast cancer, rectal cancer and kidney renal papillary cell carcinoma ($P<0.05$). MIR22HG was downregulated in HCC in association with HCC progression ($P<0.05$). In HCC patients, a low MIR22HG expression was associated with a favorable outcome after sorafenib treatment ($HR=2.94, P=0.075$) and was capable of predicting the response to sorafenib treatment ($AUC=0.8095$). Compared with the negative control, MIR22HG overexpression obviously reduced sorafenib sensitivity (with IC₅₀ of 7.731 vs 15.61) while

MIR22HG knockdown increased sorafenib sensitivity of HCC cells (with IC₅₀ of 7.986 vs 5.085). **Conclusion** MIR22HG expression level is correlated with clinical stage, lymph node metastasis, TMB, MSI, immune cell infiltration, and chemosensitivity in most cancer, suggesting its potential as an immunotherapeutic target and also a prognostic biomarker for tumors.

Keywords: long non-coding RNA; MIR22HG; pan-cancer analysis; hepatocellular carcinoma; sorafenib

收稿日期:2021-11-29

基金项目:国家自然科学基金青年项目(81902478);中国博士后科学基金(2019M652975);广东省自然科学基金(2019A1515011436)

Supported by Youth Program of National Natural Science Foundation of China (81902478).

作者简介:王惠,主管护师,E-mail: 635137884@qq.com

通信作者:张冬艳,博士,副研究员,医师,E-mail: 15013091401@163.com

近年来,随着疾病谱的变化,癌症已成为人类健康的重大威胁和挑战,给全球公共卫生带来沉重负担^[1-3]。预计到2040年将发生超过2840万例新癌症病例,比2020年增加47%^[4]。癌症是复杂的基因组疾病,根据其位置和细胞来源具有多种形式^[5]。泛癌研究的意义在于通过跨肿瘤相似性将诊断和治疗应用于更多的肿瘤^[6,7]。因此,识别不同癌症类型之间的关键基因有助于癌症的诊断和治疗。

随着基础研究及转化研究的不断深入,免疫治疗的应用开启了肿瘤治疗的新纪元^[8],其在肺癌、黑色素瘤、肝癌等多种肿瘤治疗中取得令人鼓舞的治疗效果,极大程度改善肿瘤患者的预后^[9]。目前的免疫疗法,包括程序性死亡1(PD-1)^[10]、程序性死亡配体1(PD-L1)^[11]和细胞毒性T淋巴细胞相关抗原4(CTLA-4)^[12]抑制剂,但在大部分肿瘤中其临床疗效有限^[13]。鉴于制约免疫治疗疗效的复杂性,迫切需要可以预测免疫治疗疗效和预后的标志物。在癌症基因组图谱(TCGA)的帮助下,学者们可以通过进行泛癌表达分析来寻找新的免疫治疗靶点,并研究它们与临床病理特征、免疫浸润和与癌症相关信号通路的相关性。

长链非编码RNA(lncRNA)是长度大于200核苷酸的调节性RNA^[14]。越来越多的证据表明lncRNAs在转录、翻译和表观遗传水平上调节肿瘤的生物学过程^[15-17]。lncRNA MIR22HG位于17p13.3,这是一个杂合性缺失或高甲基化染色体区域^[18]。现有研究表明,MIR22HG在胃癌^[19]、肝癌^[18]、肺癌^[20]及直肠癌^[21]中发挥抑癌基因作用进而抑制肿瘤发生和发展,而在食管癌及胶质瘤中则能促进肿瘤进展^[22]。鉴于MIR22HG在不同肿瘤中的差异表达及生物学功能,其可能对肿瘤的诊断、预后和治疗具有重要的价值。尽管已有文献报道MIR22HG与肿瘤进展的关系,但其在大部分肿瘤中的表达、功能及其与预后的关系尚不清楚;尚未有文献报道MIR22HG与肿瘤免疫微环境的关系。本研究借助癌症基因组图谱数据库,通过生物信息学的方法泛癌分析MIR22HG的表达及其与临床特征的关系,并通过体外实验对结果进行验证。

1 资料和方法

1.1 材料

1.1.1 细胞 人肝癌细胞株HCC-LM3和MHCC-97H购自中国科学院上海细胞库。

1.1.2 药品和试剂 DMEM培养基(Gibco)、胎牛血清(Gibco);MIR22HG慢病毒过表达及干扰系统由上海吉凯基因构建并包装;TRIzol(Invitrogen)、逆转录试剂盒(TAKARA)、SYBR荧光定量PCR试剂盒(TAKARA)、CCK-8试剂盒(日本同仁)。

1.1.3 肝癌标本收集 本研究通过了南方医院伦理委员会的审批,批准伦理号:NFEC-201208-K3。所有标本均来源于就诊南方医院的肝癌初诊患者,患者均未接受过抗肿瘤治疗;标本收集前与患者签署知情同意书,标本离体后,分别取癌组织和癌旁组织(与癌组织的距离>2 cm);并立即放入液氮保存。

1.2 方法

1.2.1 数据集下载及数据处理 从UCSC Xena (<https://xena.ucsc.edu>)数据库下载癌症基因组图谱(TCGA),该数据库包括11069例、共33种不同肿瘤类型的表达谱。从GTEx (<https://commonfund.nih.gov/GTEx>)网站下载31个正常组织的基因表达谱。利用Strawberry Perl (Version 5.32.0, <http://strawberryperl.com/>)软件从上述数据集中提取MIR22HG的表达谱,并分析上述肿瘤组织及正常组织中MIR22HG的表达情况。对MIR22HG的表达水平进行log₂转化后行统计分析, $P < 0.05$ 表示有统计学意义。利用R语言包“ggpubr”绘制差异表达的小提琴图。UALCAN portal (<http://ualcan.path.uab.edu/analysis-prot.html>)是一种开放资源,可用于分析不同肿瘤基因表达与临床特征的关系,我们使用该网站分析不同肿瘤中MIR22HG表达与临床特征的关系。从GEO (<https://www.ncbi.nlm.nih.gov/geo/>)及ICGC (<https://dcc.icgc.org/projects/LIRI-JP>)数据库下载肝癌及相应癌旁组织的基因表达谱(GSE22058、GSE25097、STORM trial、ICGC-LIRI-JP),分析正常肝组织及癌旁组织中MIR22HG的表达情况,并分析索拉非尼治疗有反应及无反应肝癌患者中MIR22HG的表达情况。

1.2.2 MIR22HG拷贝数变异和突变的泛癌分析 cBioPortal癌症基因组学门户网站(<http://cbioportal.org>)是开放式资源,可用于分析各种肿瘤样本的基因组变化,我们应用cBioPortal来识别MIR22HG拷贝数变化和突变的频率,并行生存分析。

1.2.3 肿瘤突变负荷(TMB)估算 使用R语言包“TCGAbiolinks”下载TCGA中突变注释文件。每兆碱基组的体细胞变异数被定义为肿瘤突变负荷。

1.2.4 微卫星不稳定(MSI)分析 MANTIS是用于识别肿瘤样本MSI的常用工具。我们使用MANTIS从TCGA数据分析了33中肿瘤类型的MSI,平均距离阈值=0.4用于区分微卫星稳定(MSS)和不稳定(MSI-H)肿瘤。利用R语言包“reshape2”和“RColorBrewer”绘制不同肿瘤中MIR22HG表达与微卫星不稳定相关基因MLH1、MSH2、MSH6、PMS2和EPCAM的相关性。

1.2.5 MIR22HG表达与免疫细胞的相关性 TIMER是系统分析不同肿瘤类型免疫浸润的理想工具。TIMER采用统计反褶积方法从基因表达谱中推断肿瘤浸润免

疫细胞的丰度。本研究利用TIMER分析MIR22HG表达与总共6种免疫浸润细胞(B细胞、CD4⁺T细胞、CD8⁺T细胞、中性粒细胞、巨噬细胞及树突状细胞)丰度之间的关系。ESTIMATE是一种可以预测肿瘤纯度及肿瘤组织中浸润的基质/免疫细胞的工具,本研究利用ESTIMATE算法来推断不同肿瘤类型中每个肿瘤样本的免疫评分。

1.2.6 MIR22HG表达与化疗药物敏感性的关系 GDSC(<https://www.cancerxgene.org>)是公共的肿瘤药物敏感性基因组学数据库,从GDSC网站中下载肿瘤药物敏感性基因组学数据,利用R语言包“PharmacoGx”分析MIR22HG表达与化疗药物敏感性的关系。

1.2.7 Kaplan-Meier生存分析 Kaplan-Meier Plotter是用于评估21种癌症类型中54 000个基因对生存影响的线上工具。本研究利用Kaplan-Meier Plotter工具分析MIR22HG的表达与索拉非尼治疗肝癌患者预后的关系,并计算了log-rank P值和95%置信区间(CI)的风险比率(HRs)。

1.2.8 荧光定量PCR(qRT-PCR)检测MIR22HG的表达 利用TRIzol抽提肝癌细胞RNA,纯化后用紫外分光光度仪检测RNA浓度。用逆转录试剂盒将RNA逆转录成cDNA。利用染料法(SYBR Green I)将反应体系置于荧光定量PCR系统进行荧光定量PCR分析,U6作为内参。本实验所用引物由上海生工公司设计并合成(表1)。

表1 qRT-PCR检测差异表达基因的引物序列

Tab.1 Primer sequence of qRT-PCR for differentially expressed genes

Gene	Primer sequence 5'-3'
U6	F: CTCGCTTCGGCAGCACACA
	R: AACGCTTCACGAATTTGCGT
MIR22HG	F: ATCCAAAGCAGGACAGCA
	R: TGGCAGGTTTACTCACT

F: Forward primer; R: Reverse primer.

1.2.9 细胞培养 用含10%血清的DMEM培养基培养人肝癌细胞株HCC-LM3和MHCC-97H,并置于含5% CO₂,37℃的恒温箱内培养,待细胞密度达到85%左右进行常规传代和后续实验,1~2 d/次进行换液。

1.2.10 CCK-8法检测MHCC-97H和HCC-LM3细胞的IC₅₀值 将相应处理后的MHCC-97H和HCC-LM3细胞分别制成1×10³/mL的细胞混悬液,接种于96孔板中,加入200 μL用无血清培养基配制的10% CCK-8,培养

2 h,用Bio-Rad酶联免疫检测仪检测A_{450 nm}。每组重复3个复孔,取平均值,根据 $[(1-A_{450 \text{ nm}} \text{实验组})/A_{450 \text{ nm}} \text{对照组}] \times 100$ 计算细胞抑制率(%),以抑制50%细胞生长的药物浓度定义为IC₅₀。

1.2.11 统计学方法 Wilcoxon检测用于评估正常组织和肿瘤组织间MIR22HG的表达差异。Kaplan-Meier方法用于计算总生存期(OS),并使用对数秩检验比较生存曲线。Pearson相关性分析用于评估MIR22HG表达水平与免疫检查点相关基因的相关性。用R软件(3.6.3版本)行统计分析,P<0.05为差异有统计学意义。

2 结果

2.1 泛癌中MIR22HG的表达水平及突变频率和拷贝数变化

TCGA数据集的泛癌表达谱分析显示,与癌旁组织相比,MIR22HG在16种癌症组织中低表达(P<0.05),包括膀胱尿路上皮癌(BLCA)、乳腺癌(BRCA)、结肠癌(COAD)、头颈部鳞状细胞癌(HNSC)、肾嫌色细胞癌(KICH)、肾透明细胞癌(KIRP)、肾乳头状细胞癌(KIRP)、低级别胶质瘤(LGG)、肺腺癌(LUAD)、肺鳞癌(LUSC)、胰腺癌(PAAD)、前列腺癌(PRAD)、直肠腺癌(READ)、甲状腺癌(THCA)及子宫内膜癌(UCEC)(图1A)。因TCGA数据集中部分癌症类型无相应的正常组织表达谱,我们结合GTEx数据库中31种正常组织的表达谱和TCGA数据集表达谱进行综合分析,结果显示,除上述癌症类型外,MIR22HG在肾上腺皮质癌(ACC)、食管癌(ESCA)、肝细胞肝癌(LIHC)、卵巢癌(OV)、皮肤黑色素瘤(SKCM)、胃癌(STAD)、睾丸癌(TGCT)和子宫肉瘤(UCS)中的表达均较正常组织的表达明显降低(P<0.05);而在胶质瘤(GBM)和急性髓细胞样白血病(LAML)中MIR22HG的表达较正常组织明显增高(P<0.05,图1B)。MIR22HG基因突变和拷贝数变异分析显示,在多种肿瘤类型中,MIR22HG基因存在不同程度基因突变,包括缺失突变、扩增突变及融合突变,其中缺失突变最常见;而STAD中MIR22HG突变频率最高。缺失突变占拷贝数变化的0.59%(63/10 953)(图1C)。生存分析显示MIR22HG基因突变患者无病生存及无进展生存时间较短(P<0.05,图1D)。

2.2 MIR22HG表达与临床特征的关系

在BLCA、BRCA、KIRC、KIRP、LIHC、睾丸癌(TGCT)和THCA中,MIR22HG的表达随临床分期增加而逐渐降低(P<0.05,图2A)。同时,在KIRC、KIRP、THCA和LUSC中,有淋巴结转移较无淋巴结转移肿瘤MIR22HG的表达进一步降低(P<0.05,图2B)。

2.3 MIR22HG表达与肿瘤突变负荷和微卫星不稳定之间的相关性

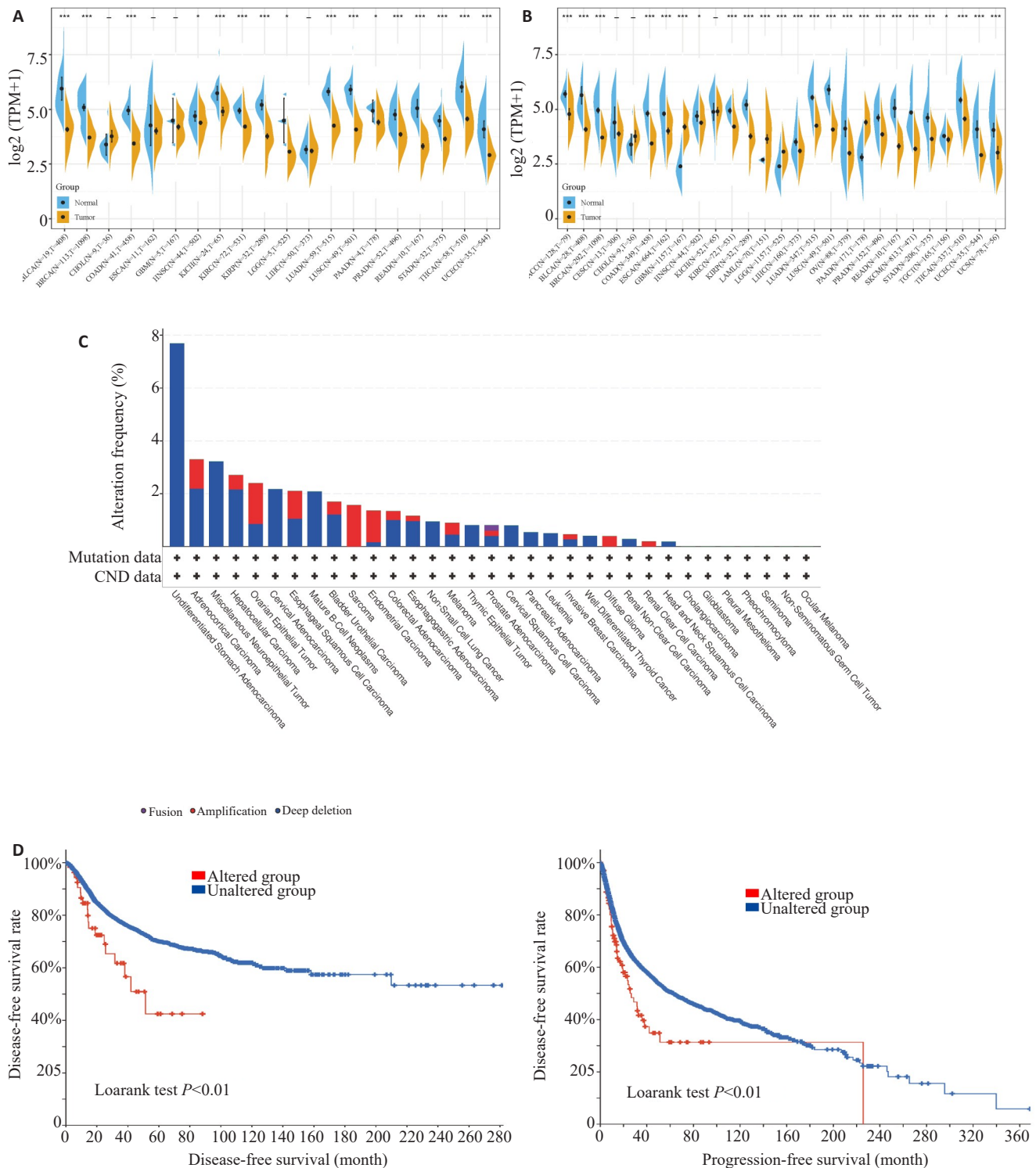


图1 泛癌中MIR22HG的表达水平及突变频率和拷贝数变化

Fig.1 Pan-cancer analysis of MIR22HG expression level, genomic mutation frequency and copy number variation. **A:** Expression profile of MIR22HG gene in paired tumor and non-tumor tissues across 22 cancer types from TCGA. **B:** MIR22HG expression in 33 cancer types with corresponding normal tissues of GTEx dataset as controls. **C:** Alteration frequency of MIR22HG gene analyzed using cBioPortal online software. **D:** Disease-free rate and progression-free rate of patients with alterations in MIR22HG gene analyzed by cBioPortal online software.

雷达图显示MIR22HG的表达与COAD、OV、肉瘤(SARC)及胸腺瘤(THYM)的肿瘤突变负荷呈正相关,而与BRCA、KIRC、KIRP、LIHC、LUAD、LUSC和PAAD等的肿瘤突变负荷呈负相关($P < 0.05$);在其它肿

瘤中未发现明显相关性(图3A)。MIR22HG的表达与弥漫大B细胞淋巴瘤(DLBC)、LIHC、LUSC及STAD的微卫星不稳定呈负相关,而与低级别胶质瘤(LGG)的微卫星不稳定呈正相关($P < 0.05$);而在其它肿瘤中未发现

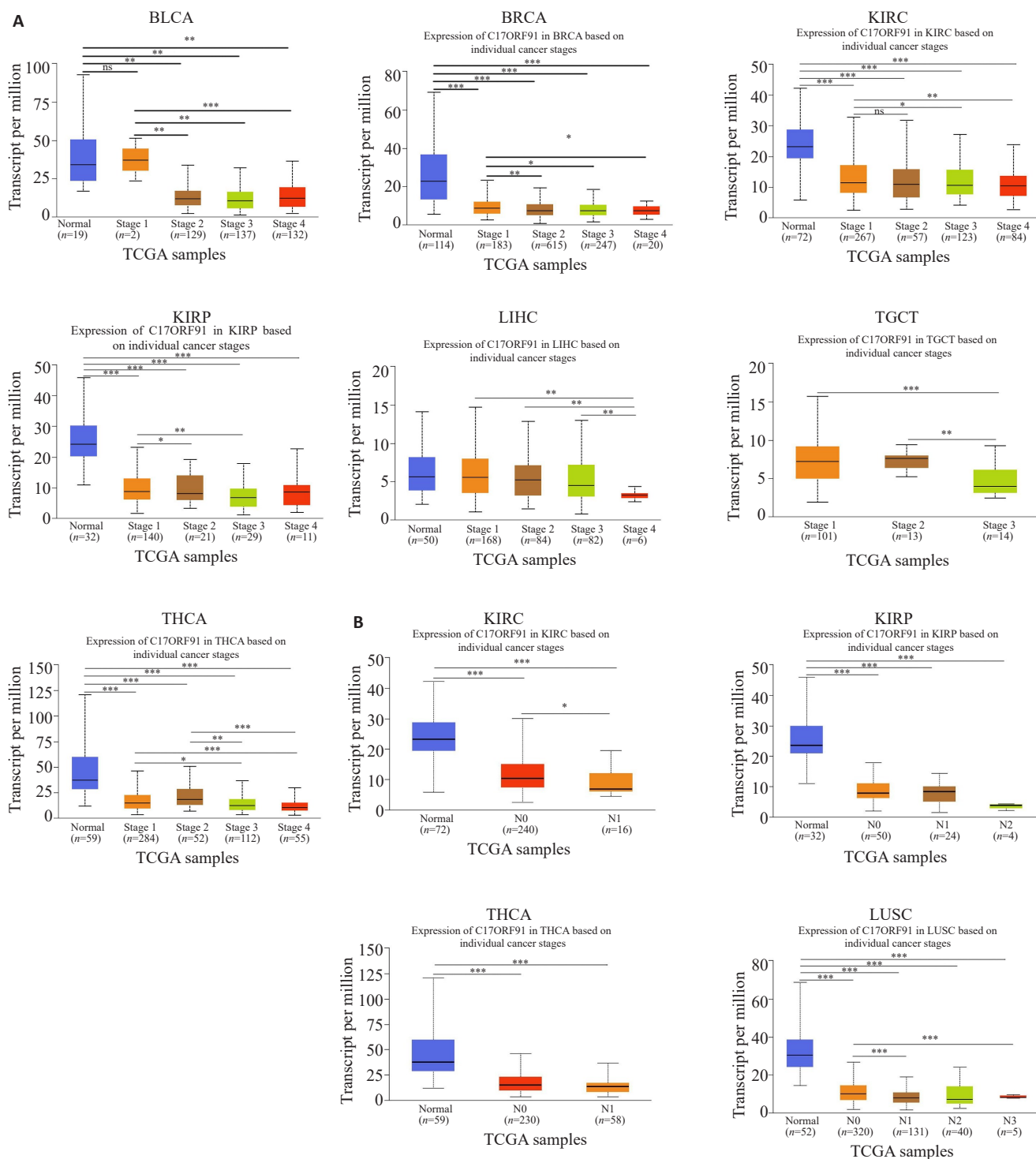


图2 MIR22HG表达与临床特征的关系

Fig.2 Relationship between MIR22HG expression and clinical characteristics. A: Expression level of MIR22HG gene analyzed by pathological stage of BLCA, BRCA, KIRC, KIRP, LIHC, TGCT, and THCA. B: Expression level of MIR22HG gene analyzed by status of lymphatic metastasis in KIRC, KIRP, THCA, and LUSC. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

任何相关性(图3B)。同时,在包括BRCA、LIHC和KIRP在内的25种肿瘤中,MIR22HG的表达与微卫星不稳定相关基因的表达相关($P < 0.05$,图3C)。

2.4 MIR22HG与免疫评分的关系

ESTIMST算法免疫评分结果显示,在包括BRCA、COAD及LIHC在内的20中肿瘤中,MIR22HG的表达

与免疫细胞浸润呈正相关,而在肾上腺皮质癌(ACC)和THCA中,MIR22HG的表达与免疫细胞浸润呈负相关($P < 0.05$,图4)。

2.5 MIR22HG与免疫浸润的相关性

图5A列举了与免疫浸润相关性最强的3种癌症。在BRCA中,MIR22HG表达水平与B细胞($R = 0.113$,

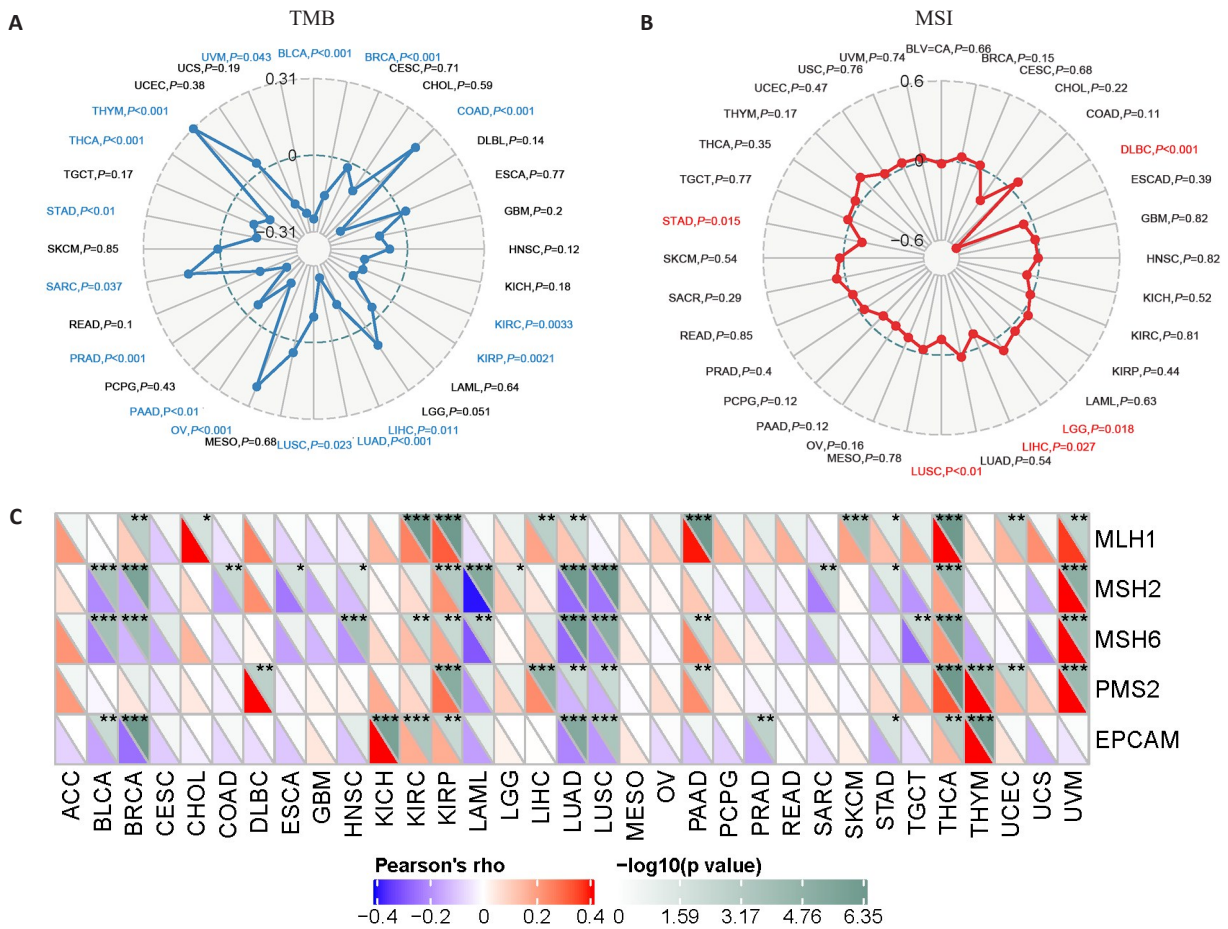


图3 MIR22HG表达与肿瘤突变负荷和微卫星不稳定之间的相关性

Fig.3 Correlation of MIR22HG expression with tumor mutation burden (TMB), microsatellite instability (MSI), and mismatch repair (MMR) gene. A, B: Radar image illustrating the association of MIR22HG expression with TMB (A) and MSI (B). C: Heatmap demonstrating the relationship between MIR22HG expression and MMR genes. For each pair, the bottom left triangle represents the correlation coefficient, and the top right triangle represents the P value. *P<0.05, **P<0.01, ***P<0.001.

P<0.001)、CD4⁺T细胞(R=0.371, P<0.001)、CD8⁺T细胞(R=0.344, P<0.001)、树突状细胞(R=0.383, P<0.001)、巨噬细胞(R=0.439, P<0.001)和中性粒细胞(R=0.41, P<0.001)的浸润呈正相关;在结肠癌中, MIR22HG表达水平与B细胞(R=0.272, P<0.001)、CD4⁺T细胞(R=0.355, P<0.001)、CD8⁺T细胞(R=0.363, P<0.001)、树突状细胞(R=0.57, P<0.001)、巨噬细胞(R=0.32, P<0.001)和中性粒细胞(R=0.584, P<0.001)的浸润呈正相关;在KIRP中, MIR22HG表达水平与B细胞(R=0.162, P<0.01)、CD4⁺T细胞(R=0.19, P<0.01)、CD8⁺T细胞(R=0.229, P<0.001)、树突状细胞(R=0.349, P<0.001)、巨噬细胞(R=0.243, P<0.001)和中性粒细胞(R=0.32, P<0.001)的浸润呈正相关。而在HNSC中, MIR22HG表达水平与B细胞(R=-0.128, P<0.01)和CD8⁺T细胞(R=-0.155, P<0.001)的浸润呈负相关;在THYM中, MIR22HG表达水平与B细胞(R=-0.22, P=0.017)、CD4⁺T细胞(R=-0.356, P<0.001)、CD8⁺T细胞(R=-0.204, P=0.027)和树突状细胞(R=-0.315, P<0.001)的浸润呈负相关

(图5B)。

2.6 MIR22HG表达与免疫检查点相关基因的相关性

MIR22HG表达在包括BRCA、COAD和LIHC在内的28中肿瘤中与大多数免疫检查点相关基因相关, 其中相关性最强的为COAD、LGG和UVM(P<0.05); 而在ACC、BLCA、CESC、ESCA和SARC中, 未发现MIR22HG与免疫检查点基因相关(图6)。

2.7 MIR22HG表达与化疗药物敏感性的相关性

相关性分析显示MIR22HG的表达与164种化疗药物IC50呈正相关(P<0.05), 图7A呈现正相关且相关性最强的9种药物, 包括AZD1208(Pim激酶抑制剂, r=0.45, P<0.05)、AZD5991(Mcl-1抑制剂, r=0.44, P<0.05)、Zoledronate(唑来膦酸, r=0.43, P<0.05)、Entinostat(恩替诺特, r=0.44, P<0.05)、JQ1(BET抑制剂, r=0.47, P<0.05)、MIRA-1(p53激动剂, r=0.43, P<0.05)、sorafenib(索拉非尼, r=0.43, P<0.05)、Venetoclax(BCL-2抑制剂, r=0.45, P<0.05)和Vorinostat(HDAC抑制剂, r=0.52, P<0.05); 而MIR22HG表达水平与Dasatinib(达沙替

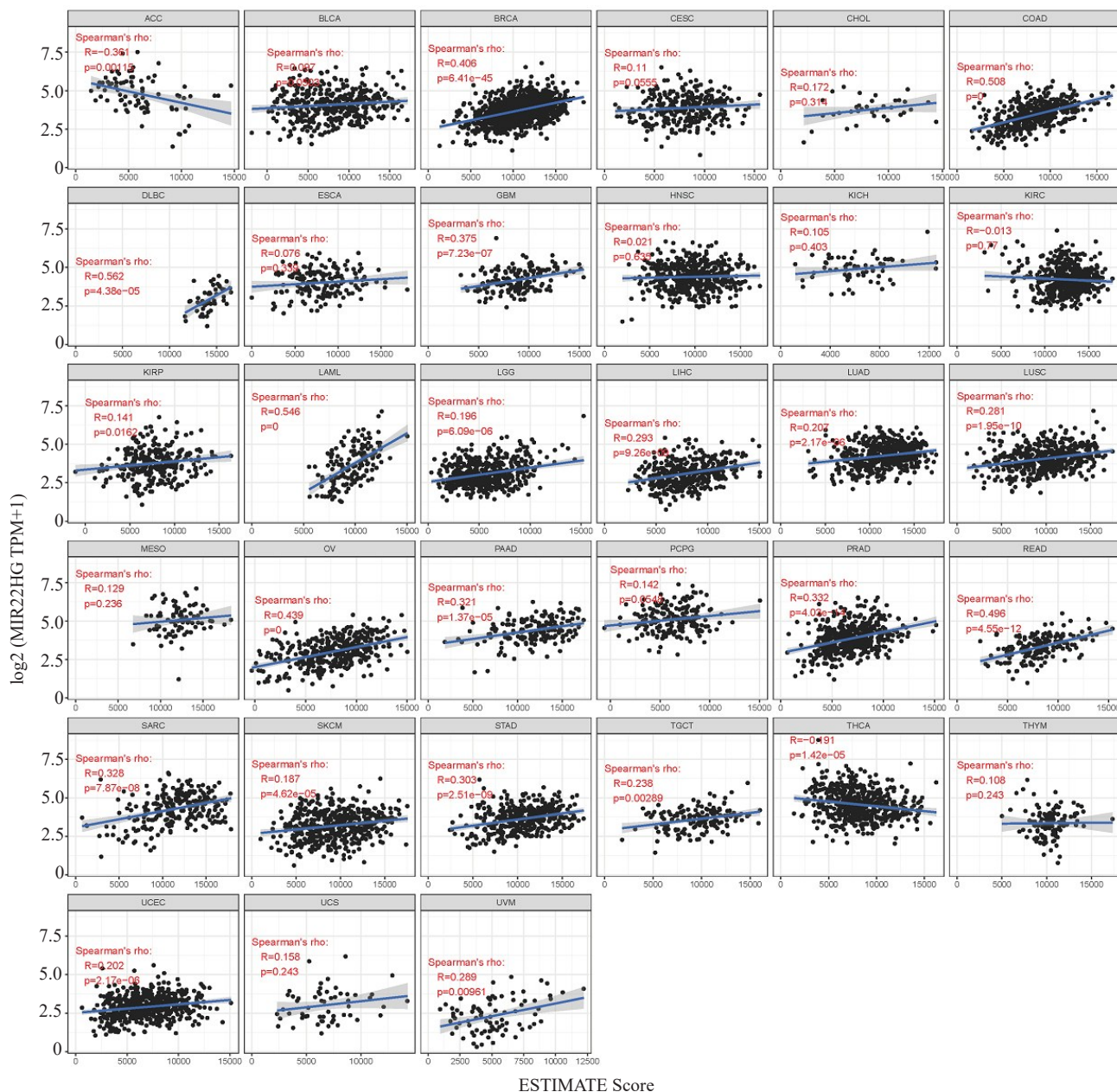


图4 MIR22HG与免疫评分的关系

Fig.4 Correlations between MIR22HG expression and ESTIMATE Score across 33 different types cancers from TCGA datasets.

尼, $r=-0.24, P<0.05$)和Trametinib(曲美替尼, $r=-0.16, P<0.05$)的 IC_{50} 呈负相关(图7B)。

2.8 MIR22HG在肝癌中的表达及其与索拉非尼耐药的关系

多数据集分析结果证实,与癌旁组织相比, MIR22HG在肝癌组织中低表达($P<0.05$,图8A)。qRT-PCR结果进一步证实MIR22HG在肝癌中低表达($P<0.05$,图8B)。MIR22HG的表达随着疾病进展逐渐降低($P<0.05$,图8C、D)。MIR22HG与临床特征的相关性分析显示MIR22HG表达与肝癌TNM分期相关($P=0.036$,表2)。相关性分析显示MIR22HG的表达与索拉非尼 IC_{50} 呈正相关,因此,本研究探讨MIR22HG表达与索拉非尼治疗肝癌患者预后的关系,生存分析显示

MIR22HG低表达的肝癌患者经索拉非尼治疗后总生存期明显延长($HR=2.94, P=0.075$,图8E)。相似地,与索拉非尼治疗无效患者相比,索拉非尼治疗有效的患者MIR22HG的表达明显降低($P<0.05$,图8F),MIR22HG低表达能有效预测肝癌患者索拉非尼治疗效果($AUC=0.8095$,图8G)。与对照组相比(HCC-LM3-NC),过表达MIR22HG组(HCC-LM3-MIR22HG)细胞的对索拉非尼的 IC_{50} 明显增高($IC_{50\ NC}$ vs $IC_{50\ MIR22HG}=7.731$ vs 15.61 ,图8H);反之,与对照组相比(MHCC-97H-NC),敲除MIR22HG组细胞(MHCC-97H-sh-MIR22HG)对索拉非尼的 IC_{50} 能明显降低($IC_{50\ NC}$ vs $IC_{50\ sh-MIR22HG}=7.986$ vs 5.085 ,图8I)。

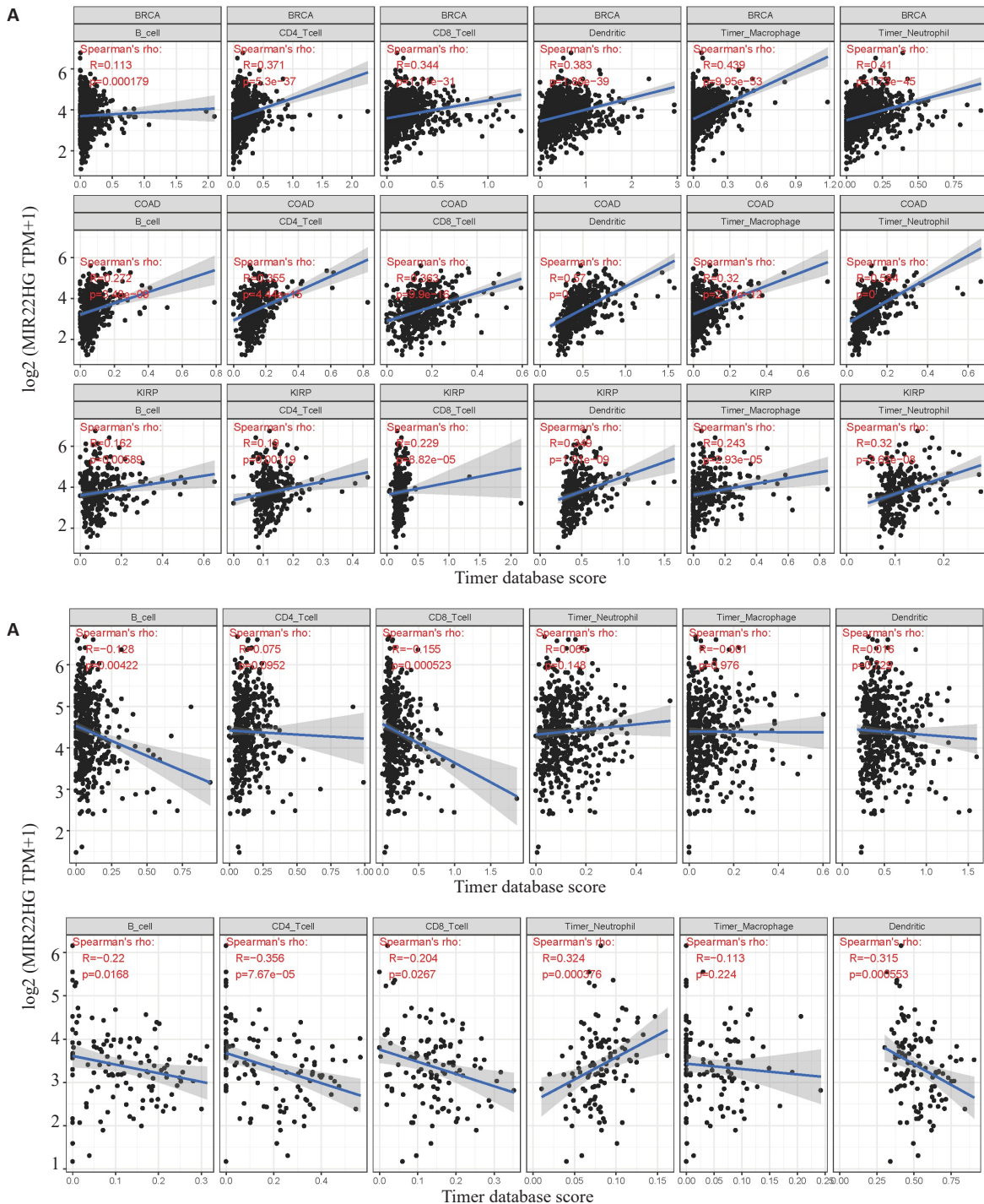


图5 MIR22HG与免疫浸润的相关性

Fig.5 Correlation between MIR22HG and immune infiltration levels. A: Three tumors with the highest correlation coefficients between MIR22HG expression and tumor infiltration level of different immune cells. B: Correlation between MIR22HG and immune infiltration levels in HNSC and THYM.

3 讨论

尽管越来越多的文献报道MIR22HG在肿瘤进展中发挥重要的作用,尚无任何从整体角度对MIR22HG进行泛癌分析的研究。因此,本研究基于TCGA和GEO数据库,通过基因表达、突变频率和拷贝数变化等角度对不同肿瘤中的MIR22HG基因进行综合分析。我们旨在探讨MIR22HG在肿瘤组织和正常组织的表达高低,而TCGA数据库中某些癌症组织缺少正常组织

及癌旁组织,或者相应癌旁组织样本量较少,比如肾上腺皮质癌(ACC)、食管癌(ESCA)和卵巢癌(OV)等,因此我们将TCGA与含有31个正常组织表达谱的GTEx数据库数据进行联合分析。我们观察到MIR22HG在大多数肿瘤中低表达,包括BRCA、COAD、LUAD、LUSC、PRAD、READ和LIHC等;而在脑胶质瘤中MIR22HG的表达较正常组织明显增高。我们的分析结果与已有文献报道相一致,MIR22HG在不同肿瘤中

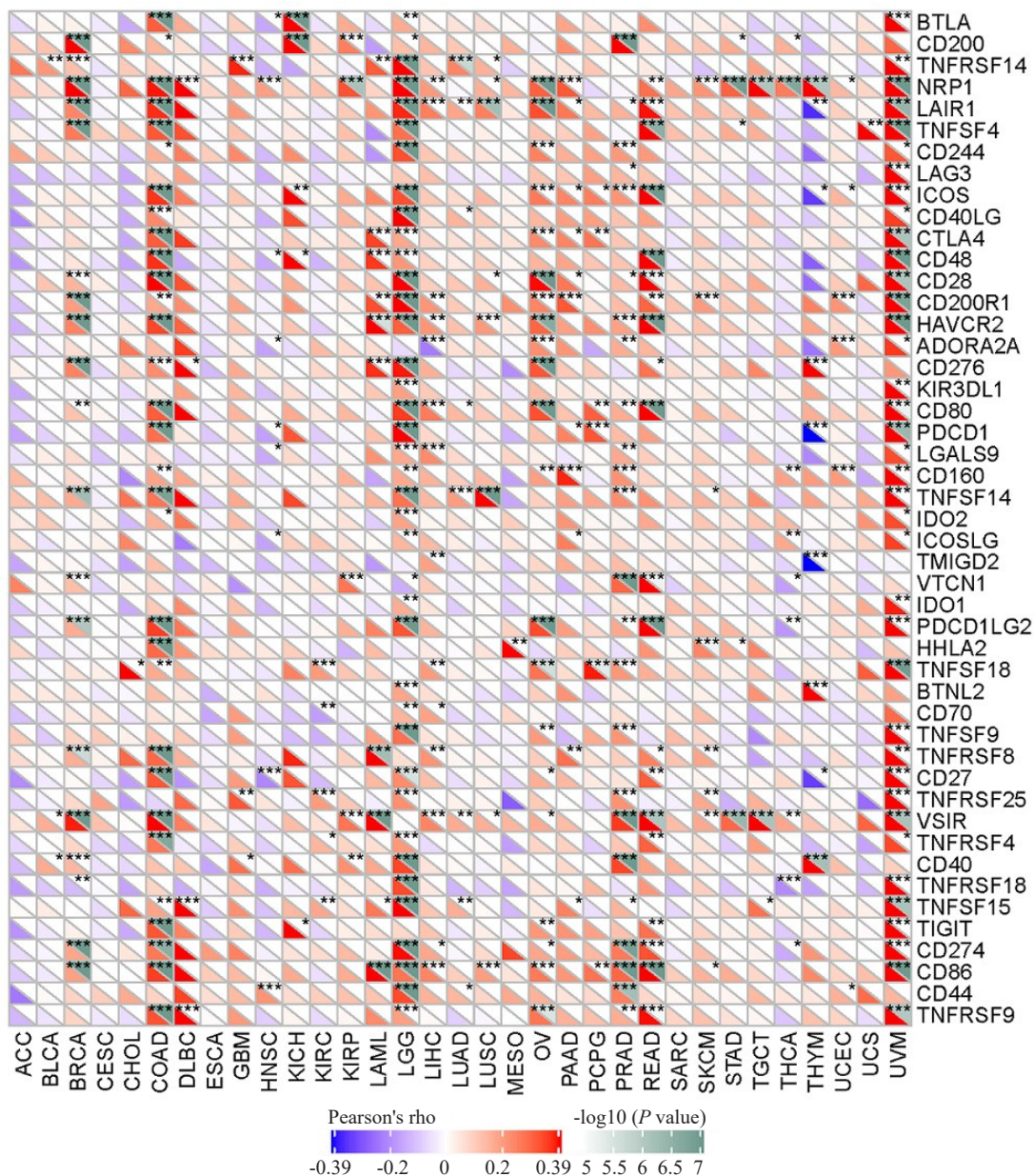


图6 MIR22HG表达与免疫检查点相关基因的相关性
 Fig.6 Correlation between MIR22HG expression and immune checkpoint-related genes. * $P<0.05$, ** $P<0.01$, *** $P<0.001$.

的表达水平及生物学功能不同。已有报道 MIR22HG 在结直肠癌中显著低表达,其作为抑癌基因能通过调控 TGF- β /SMAD 信号通路增强免疫治疗疗效^[21]。在肺癌中,抑制 MIR22HG 表达能激活 YBX1/MET/p21 信号通路进而促进肺癌进展^[20]。而脑胶质瘤中,高表达的 MIR22HG 能通过 wnt/beta-catenin 信号通路促进胶质瘤进展^[23]。这些争议的结果强调了 MIR22HG 在不同肿瘤类型发挥特定作用的可能性。基因组突变及拷贝数变异在一定程度上影响基因的表达水平^[24],我们观察到 MIR22HG 在多种肿瘤中均存在不同程度缺失突变,因此,我们推测基因组缺失突变可能是 MIR22HG 在不同肿瘤中低表达的原因之一,而 MIR22HG 在不同肿瘤中差异表达的原因仍需进一步探讨。

异常表达的 lncRNA 可以通过多种途径调控肿瘤的发生和发展,除了调控肿瘤细胞本身的之外,lncRNA 还能通过调控不同信号通路影响肿瘤微环境^[25]。肿瘤微环境由肿瘤细胞、成纤维细胞、免疫细胞、内皮细胞等细胞成分以及非细胞成分组成^[26]。免疫细胞可塑性极强,容易受到肿瘤微环境的影响,可以被激活并促进肿瘤发生和发展^[27]。针对免疫细胞和肿瘤细胞相互作用的免疫治疗已成为近年的研究热点,并已成功应用于临床^[28]。然而,只有在某些特定肿瘤如黑色素瘤中,有限比例的患者对免疫治疗有效^[29]。因此,寻找新的免疫治疗靶点至关重要。越来越多的研究表明,肿瘤免疫细胞浸润水平和免疫治疗疗效及患者预后密切相关^[30]。为了阐明 MIR22HG 与免疫浸润的关系,我们首先利用

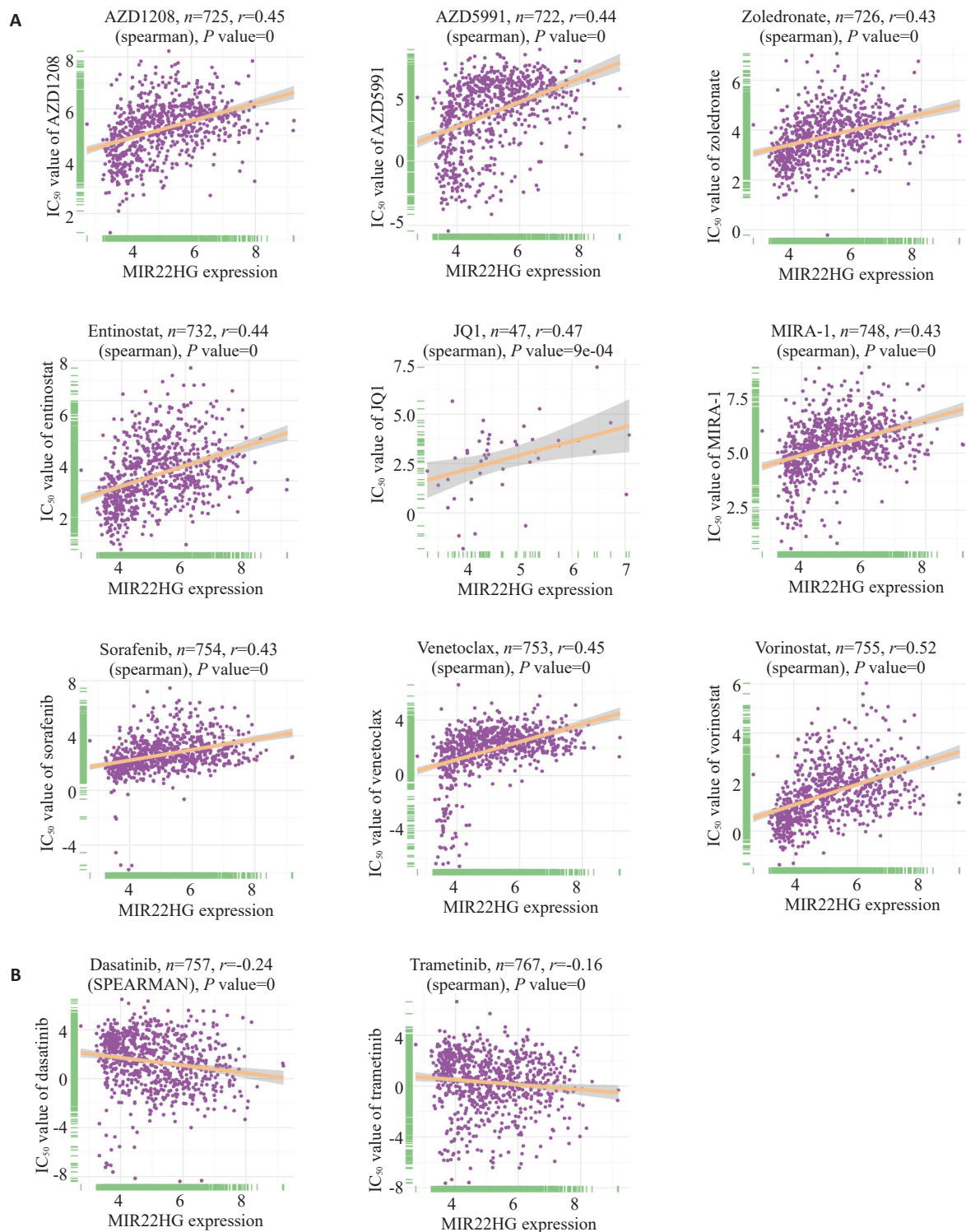


图7 MIR22HG表达与化疗药物敏感性的相关性

Fig.7 Association of MIR22HG expression level with cancer drug resistance and sensitivity. **A:** High MIR22HG expression correlated with high IC_{50} of each drug in different cancer cell lines from the Genomics of Drug Sensitivity in Cancer (GDSC) dataset. **B:** High MIR22HG expression correlated with low IC_{50} of each drug in different cancer cell lines from the Genomics of Drug Sensitivity in Cancer (GDSC) dataset.

ESTIMAT算法计算TCGA队列中MIR22HG表达水平与免疫评分的关系,结果显示在包括BRCA、COAD及LIHC在内的20中肿瘤中,MIR22HG的表达与免疫细胞浸润呈正相关,而在ACC和THCA中,MIR22HG的表达与免疫细胞浸润呈负相关,表明在大多数肿瘤中,

MIR22HG的表达水平与高免疫细胞浸润相关。利用TIMER分析我们发现,对于不同肿瘤类型,MIR22HG表达与肿瘤浸润免疫细胞之间的相关性不同。进一步探讨了MIR22HG表达与免疫检查点相关基因的相关性,发现MIR22HG表达在包括BRCA、COAD和LIHC

表2 203例肝癌患者中MIR22HG表达水平与临床特征的关系 (ICGC-LIRI-JP)

Tab.2 Correlation between MIR22HG expression and clinicopathologic features in 203 patients with HCC (ICGC-LIRI-JP)

Characteristics	MIR22HG high expression (n=102)	MIR22HG low expression (n=1101)	P
Gender			0.968
Male	77	76	
Female	25	25	
Age (year)			0.647
≤55	11	13	
>55	91	88	
Virus (HBV)			0.412
With	26	31	
without	76	70	
TNM stage			0.036
I + II	72	57	
III + IV	30	44	
Portal vein invasion			0.131
Yes	20	29	
No	82	72	
Vein invasion			0.193
Yes	25	22	
No	87	79	
Artery invasion			0.316
Yes	0	1	
No	103	100	
Bile duct invasion			0.382
Yes	5	8	
No	97	93	

HBV: Hepatitis B virus; HCV: Hepatitis C virus.

在内的28中肿瘤中与大多数免疫检查点相关基因相关;而在ACC、膀胱癌BLCA、CESC、ESCA和SARC中,未发现MIR22HG与免疫检查点基因相关。这些结果提示MIR22HG的表达水平与肿瘤免疫细胞浸润及免疫细胞的功能密切相关,MIR22HG有可能成为免疫治疗的有效靶点。

GDSC是公共的肿瘤药物敏感性基因组学数据库,利用该数据库可分析基因表达与常用化疗药物敏感性的关系。我们发现,MIR22HG的表达水平与MIR22HG的表达与包括索拉非尼在内的164种化疗药物IC₅₀成正相关;而与2种化疗药物(达沙替尼和曲美替尼)的IC₅₀呈负相关。以上结果提示尽管MIR22HG

在多种肿瘤中显著低表达,但是低表达MIR22HG的肿瘤患者可能对包括索拉非尼在内的化疗药物更敏感。

尽管在TCGA、GTEx和GDSC数据库中整合了信息,单纯生物信息学分析仍有局限性。为了克服这一问题,我们利用基础实验对泛癌分析结果进行验证。本课题组长期致力于肝癌发生发展机制研究,索拉非尼是晚期肝癌一线治疗手段,泛癌分析结果提示MIR22HG在肝癌显著低表达,且尚无文献报道其表达水平与索拉非尼治疗的关系,因此,我们在肝癌中验证MIR22HG的表达及其与索拉非尼治疗的相关性。我们重新整合GEO(GSE22058、GSE25097和STORM trial)及ICGC-LIRI-JP数据库信息初步验证MIR22HG在肝癌中低表达,接着,利用qRT-PCR法检测12对新鲜肝癌标本及对应癌旁组织中MIR22HG的表达情况,结果再次验证MIR22HG在肝癌中低表达。MIR22HG表达与药物敏感性分析提示MIR22HG高表达与索拉非尼耐药相关,提取TCGA数据集中经索拉非尼治疗肝癌患者行生存分析,我们观察到,MIR22HG低表达的患者经索拉非尼治疗后生存期明显延长,差异未达到统计学意义可能与样本例数不足相关(n=29)。利用肝癌索拉非尼治疗的经验临床研究(STORM trial)的基因谱数据进行分析,也提示相较于索拉非尼治疗无效的患者,索拉非尼治疗有效的患者MIR22HG表达水平较低,且MIR22HG能有效预测索拉非尼治疗的有效性。这些结果初步验证MIR22HG表达与索拉非尼治疗敏感性相关。利用体外实验进行验证,结果显示过表达MIR22HG能降低肝癌细胞对索拉非尼的敏感性;而敲除MIR22HG的表达则能增加肝癌细胞对索拉非尼的敏感性。综上所述,MIR22HG在肝癌中低表达,且其低表达与肝癌进展密切相关,而MIR22HG低表达的患者更能从索拉非尼治疗中获益。因此,治疗前检测MIR22HG的表达水平有助于筛选索拉非尼治疗的潜在获益人群。

总之,本研究阐明了MIR22HG在多种癌症类型中显著低表达,其表达水平与多种肿瘤的分期、有否淋巴结转移、肿瘤突变负荷、微卫星不稳定、免疫细胞浸润及化疗药物敏感性相关,因此,MIR22HG有望成为免疫疗法的潜在靶点和筛选化疗获益人群的生物标志物。

参考文献:

[1] Miller KD, Ortiz AP, Pinheiro PS, et al. Cancer statistics for the US hispanic/Latino population, 2021 [J]. CA Cancer J Clin, 2021, 71 (6): 466-87.
 [2] Siegel RL, Miller KD, Fuchs HE, et al. Cancer statistics, 2022 [J]. CA Cancer J Clin, 2022, 72(1): 7-33.
 [3] Byrd DR, Brierley JD, Baker TP, et al. Current and future cancer staging after neoadjuvant treatment for solid tumors [J]. CA Cancer

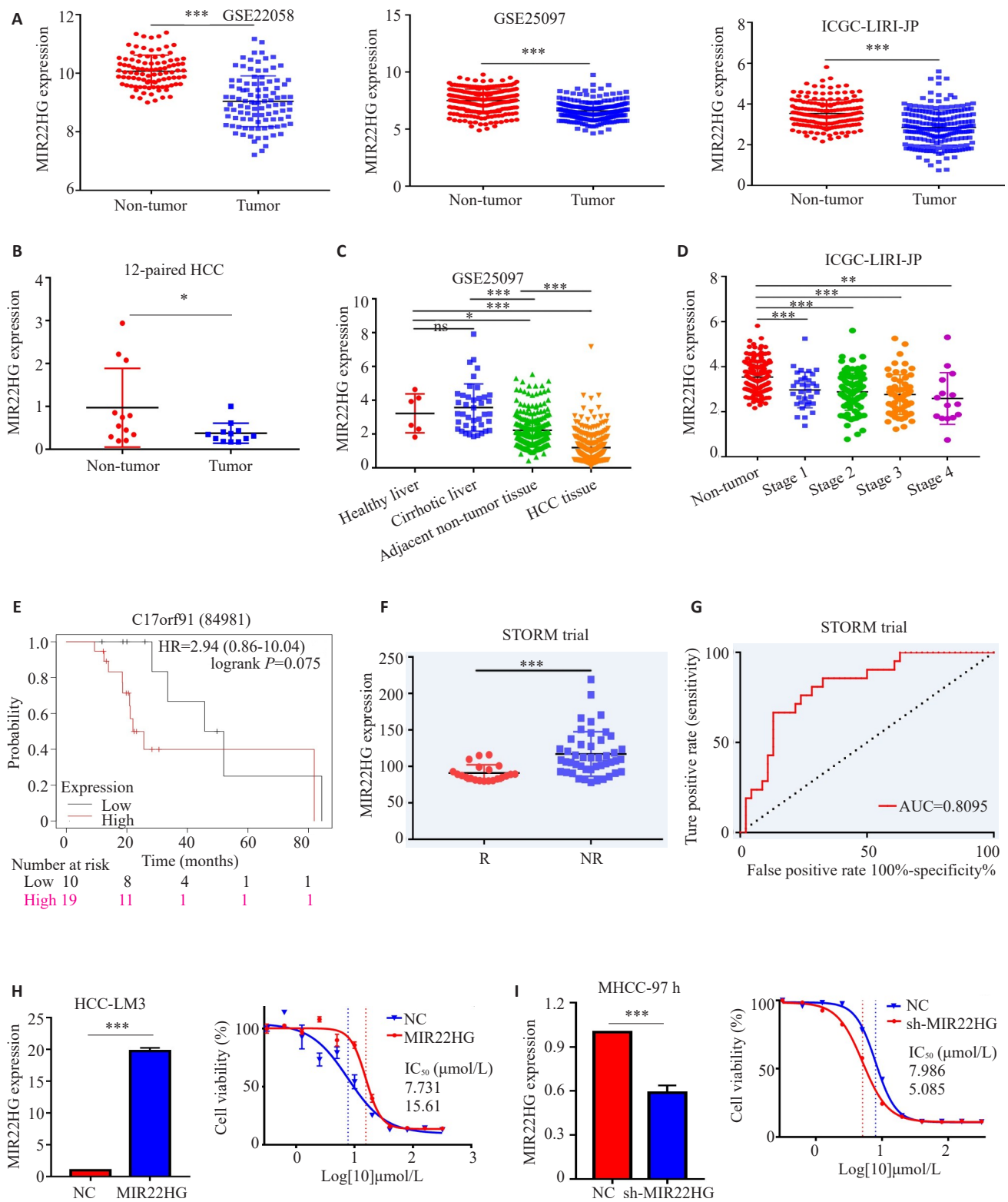


图8 MIR22HG在肝癌中的表达与索拉非尼耐药的关系

Fig.8 Expression level of MIR22HG in HCC and its correlation with sorafenib resistance. A: Expression of MIR22HG in HCC from the indicated datasets. B: Expression level of MIR22HG from 12 paired fresh HCC tissues and matched non-tumor tissues. C, D: MIR22HG expression in indicated groups from GSE25057 and ICGC-LIRI-JP. ns: not significant. E: Overall survival of sorafenib-treated patients with HCC. F: Differences of MIR22HG expression between sorafenib responders and non-responders with HCC from DTROM trial. R: Responders; NR: Non-responders. G: Receiver operating characteristic curves for prediction of sorafenib sensitivity in HCC using MIR22HG probe as the predictor. H, I: IC_{50} values of sorafenib in HCC cells transfected with MIR22HG-overexpressing (H) or sh-MIR22HG (I) lentivirus analyzed by CCK-8 assay. * $P<0.05$, ** $P<0.01$, *** $P<0.001$.

- J Clin, 2021, 71(2): 140-8.
- [4] Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. *CA Cancer J Clin*, 2021, 71(3): 209-49.
- [5] Liu XH, Yu SN, Feng C, et al. In situ analysis of cancer cells based on DNA signal amplification and DNA nanodevices[J]. *Crit Rev Anal Chem*, 2021, 51(1): 8-19.
- [6] Urbanek-Trzeciak MO, Galka-Marciniak P, Nawrocka PM, et al. Pan-cancer analysis of somatic mutations in miRNA genes [J]. *EBioMedicine*, 2020, 61: 103051.
- [7] O'Donnell JS, Teng MWL, Smyth MJ. Cancer immunoediting and resistance to T cell-based immunotherapy[J]. *Nat Rev Clin Oncol*, 2019, 16(3): 151-67.
- [8] Ju MY, Bi J, Wei Q, et al. Pan-cancer analysis of NLRP3 inflammasome with potential implications in prognosis and immunotherapy in human cancer[J]. *Brief Bioinform*, 2021, 22(4): bbaa345.
- [9] Abbott M, Ustoyev Y. Cancer and the immune system: the history and background of immunotherapy[J]. *Semin Oncol Nurs*, 2019, 35(5): 150923.
- [10] Herbst RS, Giaccone G, de Marinis F, et al. Atezolizumab for first-line treatment of PD-L1-selected patients with NSCLC[J]. *N Engl J Med*, 2020, 383(14): 1328-39.
- [11] Paz-Ares L, Vicente D, Tafreshi A, et al. A randomized, placebo-controlled trial of pembrolizumab plus chemotherapy in patients with metastatic squamous NSCLC: protocol-specified final analysis of KEYNOTE-407[J]. *J Thorac Oncol*, 2020, 15(10): 1657-69.
- [12] Kennedy LB, Salama AKS. A review of cancer immunotherapy toxicity[J]. *CA Cancer J Clin*, 2020, 70(2): 86-104.
- [13] Baas P, Scherpereel A, Nowak AK, et al. First-line nivolumab plus ipilimumab in unresectable malignant pleural mesothelioma (CheckMate 743): a multicentre, randomised, open-label, phase 3 trial[J]. *Lancet*, 2021, 397(10272): 375-86.
- [14] Tan YT, Lin JF, Li T, et al. LncRNA-mediated posttranslational modifications and reprogramming of energy metabolism in cancer [J]. *Cancer Commun (Lond)*, 2021, 41(2): 109-20.
- [15] Zhang DY, Sun QC, Zou XJ, et al. Long noncoding RNA UPK1A-AS1 indicates poor prognosis of hepatocellular carcinoma and promotes cell proliferation through interaction with EZH2[J]. *J Exp Clin Cancer Res*, 2020, 39(1): 229.
- [16] Yang Z, Jiang S, Shang JJ, et al. LncRNA: Shedding light on mechanisms and opportunities in fibrosis and aging[J]. *Ageing Res Rev*, 2019, 52: 17-31.
- [17] Bridges MC, Daulagala AC, Kourtidis A. LNCcation: lncRNA localization and function[J]. *J Cell Biol*, 2021, 220(2): e202009045.
- [18] Zhang DY, Zou XJ, Cao CH, et al. Identification and functional characterization of long non-coding RNA MIR22HG as a tumor suppressor for hepatocellular carcinoma [J]. *Theranostics*, 2018, 8(14): 3751-65.
- [19] Zhang XQ, Jiang YY, Xie Y, et al. Comprehensive analysis of lncRNAs associated with the pathogenesis and prognosis of gastric cancer[J]. *DNA Cell Biol*, 2020, 39(2): 299-309.
- [20] Su WM, Feng SM, Chen XY, et al. Silencing of long noncoding RNA MIR22HG triggers cell survival/death signaling via oncogenes YBX1, MET, and p21 in lung cancer [J]. *Cancer Res*, 2018, 78(12): 3207-19.
- [21] Xu J, Shao TT, Song MX, et al. MIR22HG acts as a tumor suppressor via TGFβ/SMAD signaling and facilitates immunotherapy in colorectal cancer [J]. *Mol Cancer*, 2020, 19(1): 51.
- [22] Zhang L, Li CX, Su XL. Emerging impact of the long noncoding RNA MIR22HG on proliferation and apoptosis in multiple human cancers[J]. *J Exp Clin Cancer Res*, 2020, 39(1): 271.
- [23] Han MZ, Wang S, Fritah S, et al. Interfering with long non-coding RNA MIR22HG processing inhibits glioblastoma progression through suppression of Wnt/β-catenin signalling [J]. *Brain*, 2020, 143(2): 512-30.
- [24] Song Y, Wang H, Zou XJ, et al. Reciprocal regulation of HIF-1α and Uroplakin 1A promotes glycolysis and proliferation in Hepatocellular Carcinoma[J]. *J Cancer*, 2020, 11(22): 6737-47.
- [25] Hinshaw DC, Shevde LA. The tumor microenvironment innately modulates cancer progression[J]. *Cancer Res*, 2019, 79(18): 4557-66.
- [26] Vitale I, Manic G, Coussens LM, et al. Macrophages and metabolism in the tumor microenvironment[J]. *Cell Metab*, 2019, 30(1): 36-50.
- [27] DeBerardinis RJ. Tumor microenvironment, metabolism, and immunotherapy[J]. *N Engl J Med*, 2020, 382(9): 869-71.
- [28] Jarosz-Biej M, Smolarczyk R, Cichoń T, et al. Tumor microenvironment as a "game changer" in cancer radiotherapy[J]. *Int J Mol Sci*, 2019, 20(13): 3212.
- [29] Soekojó CY, Ooi M, de Mel S, et al. Immunotherapy in multiple myeloma[J]. *Cells*, 2020, 9(3): 601.
- [30] Bai X, Wu DH, Ma SC, et al. Development and validation of a genomic mutation signature to predict response to PD-1 inhibitors in non-squamous NSCLC: a multicohort study [J]. *J Immunother Cancer*, 2020, 8(1): e000381.

(编辑:余诗诗)