

肺腺癌中 HPRT1 基因表达对患者总生存的影响

杨红秀, 朱中山

(湖南省第二人民医院肿瘤科, 湖南 长沙 410007)

[摘要] **目的** 探索 HPRT1 基因在肺腺癌中的表达特征、总生存率、功能激活以及免疫浸润中的影响。**方法** 通过对 TCGA 肺腺癌数据以及多个 GEO 数据库中的肺腺癌数据进行挖掘分析, 对比并验证 HPRT1 表达量与预后总生存(overall survival, OS)的关系, 通过 clusterProfiler 分析 HPRT1 基因高表达组中上调基因的功能富集情况。且通过 TIMER 以及 CIBERSORT 算法计算肺腺癌中不同免疫细胞的浸润水平并对比在 HPRT1 高低表达组间的浸润程度差异。**结果** HPRT1 在肺腺癌组织中表达量显著上调, 且在 TCGA 中证明 HPRT1 基因高表达的患者预后 OS 更差($P < 0.01$)。经过 2 个 GEO 数据集的验证同样发现 HPRT1 基因高表达表现为预后 OS 更差(GSE13213, $P < 0.01$; GSE67639, $P < 0.001$)。差异分析显示高表达的患者中有 683 个基因的表达量显著上调且这些上调基因的功能主要富集与 p53 以及细胞周期等癌症相关的信号通路。TIMER 以及 CIBERSORT 算法进行的免疫细胞浸润程度分析发现在高表达差预后人群中 B 细胞以及 CD4T 细胞含量均更低($P < 0.05$)。**结论** HPRT1 基因表达量越高肺腺癌患者的总生存越差, 且高表达患者的 p53 信号通路上调, B 细胞以及 CD4T 细胞浸润程度显著下降。

[关键词] 肺腺癌; HPRT1 基因; 总生存; 数据挖掘

[中图分类号] R734.2 **[文献标志码]** A **[文章编号]** 2095-610X(2024)08-0017-07

Effect of HPRT1 Gene Expression on Overall Survival of Patients with Lung Adenocarcinoma

YANG Hongxiu, ZHU Zhongshan

(Dept. of Oncology, The 2nd People's Hospital of Hunan Province,
Changsha Hunan 410007, China)

[Abstract] **Objective** To explore the expression characteristics, prognostic effect, functional activation and immune infiltration of HPRT1 gene in lung adenocarcinoma. **Methods** The data of adenocarcinoma of the lung from TCGA and multiple GEO databases were collected and analyzed. The relationship between HPRT1 expression level and overall survival (OS) was compared and validated, and the functional enrichment of upregulated genes in the HPRT1 high-expression group was analyzed using ClusterProfiler. Furthermore, the infiltration levels of immune cells in lung adenocarcinoma were calculated using TIMER and CIBERSORT algorithms and compared between the HPRT1 high- and low-expression groups. **Results** The expression of HPRT1 was significantly upregulated in lung adenocarcinoma tissues, and TCGA dataset analysis showed the patients with high expression of HPRT1 gene had a poorer overall survival (OS) prognosis ($P < 0.01$). Validation through two GEO datasets also shows that high expression of the HPRT1 gene is associated with a worse OS prognosis (GSE13213, $P < 0.01$, GSE67639, $P < 0.001$). Differential analysis indicates that among patients with high expression, the expression levels of 683 genes are significantly upregulated, with these upregulated genes primarily enriched in signaling pathways related to cancer, such as p53 and the cell cycle. Analysis of immune cell infiltration levels using TIMER and CIBERSORT

[收稿日期] 2024-04-02

[基金项目] 湖南省自然科学基金科卫联合基金资助项目(2020JJ8080)

[作者简介] 杨红秀(1975~), 女, 湖南怀化人, 医学学士, 副主任护师, 主要从事肿瘤专业护理与科研工作。

[通信作者] 朱中山, E-mail: 84059616@qq.com

algorithms reveals that in the group with high expression and poor prognosis, the levels of B cells and CD4 T cells are lower ($P < 0.05$). **Conclusion** The higher the expression level of the HPRT1 gene, the significantly worse the overall survival of patients with lung adenocarcinoma, and in patients with high expression levels, the p53 signaling pathway is upregulated, and the infiltration level of B cells and CD4T cells is significantly decreased.

[**Key words**] Lung Adenocarcinoma; HPRT1 gene; Overall Survival; Data Mining

肺癌始终是 1 个重要的世界性公共卫生及健康问题, 基于全球癌种数据统计其中非小细胞肺癌类型占比接近 80%~85%, 非小细胞肺癌基于病理学特征分为肺腺癌以及肺鳞癌^[1-2]。经过多年的精准医学转化研究以及基因组学研究, 肺腺癌和肺鳞癌在临床诊疗中俨然成为 2 种非常不同的癌症类型。

肺腺癌已典型的激酶激活为主要驱动因素, 其中包括 EGFR、ALK、KRAS、BRAF、MET、ERBB2、RET、ROS1 等重要基因, 并在不同基因上均发现了多种热点变异并开发了相应的小分子抑制剂^[3-5]。虽然临床中针对肺腺癌已经取得了非常显著的进展和突破, 但是以小分子抑制剂为首的治疗方式较为依赖驱动变异的检出情况, 且依然存在较高比例的患者未能检测到常见基因的驱动变异^[6]。因此, 针对肺腺癌患者影响治疗以及影响预后的分子特征仍需要进一步挖掘。

HPRT1 基因可转录翻译成 1 种蛋白酶, 主要参与细胞周期信号通路, 而且是 1 种重要的蛋白质可为细胞生长提供必需的成分, 在多种组织中高表达^[7-8]。传统认为 HPRT1 为管家基因, 但是越来越多研究表明该基因和多种类型的癌种存在关联, 在头颈癌患者中过表达与不良预后相关^[9], 还被发现该基因与癌细胞的化疗耐药相关并且被提议为化疗药物的治疗靶点^[10]。因此本研究将通过多个数据集对 HPRT1 基因在肺腺癌中的预后关联、生存率、调控通路、免疫浸润特征等进行分析。

1 资料与方法

1.1 TCGA 以及 GEO 中肺腺癌数据

本研究数据下载自 TCGA 数据库 (<https://xenaproject.ucsc.edu/>)^[11], 其中包含共计 493 例肺腺癌原发灶的 RNAseq 表达量数据, 表达量信息包含 2 种定量方式。测序得到的原始序列计数 (Reads count) 用于样本之间的差异分析, 每百万份转录数 (transcript per million, TPM) 用于进行表达量高低的对比, 该 493 例患者均具有完整的预后 OS 数据用于生存分析。另外纳入 TCGA 数据集中共计

49 例癌旁对照样本的表达量数据。基因表达综合数据库 (gene expression omnibus, GEO) 中提取了 2 个研究的数据包含 GSE13213^[12] 以及 GSE67639^[13], 分别有 116 例及 1024 例肺腺癌样本的表达数据以及预后 OS 数据。

1.2 研究方法

1.2.1 数据集定义 训练集: TCGA 的 493 例肺腺癌数据用于本研究的训练数据集来探索 HPRT1 基因的表达与预后关联以及各类生物学特征的分析。验证集: GSE13213 以及 GSE67639 2 个独立数据集主要用于验证 HPRT1 基因和预后的关联性。

1.2.2 基因表达量分析 分析 TCGA 数据集中的 493 例肺癌样本与 49 例癌旁样本之间的 HPRT1 基因表达量差异, 表达量的定量方式为 TPM。

1.2.3 HPA 分析肺腺癌中 HPRT1 蛋白的表达量 在 The Human Protein Atlas (HPA)^[14] 数据库中检索 HPRT1 基因对应的抗体, 并分别统计 2 种抗体在肺腺癌组织中的表达水平。

1.2.4 基因与患者的预后生存分析 训练集 TCGA 数据以及 GEO 的独立验证集中均按照相同的标准对患者进行高表达以及低表达分组, 然后在 R 语言软件 (版本号为 4.0.3) 中使用 survival 算法包 (版本号为 3.2.13) 对 2 组对高表达以及低表达患者进行比较分析。其中高表达定义为 HPRT1 基因表达量最高的 30% 患者, 低表达定义为 HPRT1 基因表达量最低的 70% 的患者。多因素校正分析使用 survival 算法包中的 Cox 比例风险回归模型 (cox proportional hazards regression model) 进行计算。

1.2.5 OSluca 进行基因预后的验证分析 OSluca^[15] 提供了汇总之后的 GEO 预后数据, 通过该数据库进行了 HPRT1 在多个 GEO 数据集中与预后 OS 的关联性验证。

1.2.6 差异表达以及功能富集分析 通过 R 语言软件中的 DESeq2 包 (版本号为 1.30.1) 对 TCGA 数据集中 HPRT1 基因高表达组与低表达组之间的样本进行差异表达分析。其中 Fold change 大于 2 且 adjust P value < 0.05 即为表达量上调。并通过 ggplot2 包对差异表达结果进行火山图绘制。基于上

述差异表达基因的定义条件提取上调基因之后通过 clusterProfiler 包^[16](版本号为3.18.1)进行KEGG富集分析。

1.2.7 免疫浸润分析 通过 TIMER^[17] 以及 CIBERSORT 算法^[18] 对 493 例 TCGA 中肺腺癌患者的全部免疫细胞类型进行定量分析, 并按照 HPRT1 基因的高表达与低表达分组进行各个细胞类型的含量差异分析。

1.3 统计学处理

使用 Wilcoxon test 检验方法进行 HPRT1 基因表达量分组差异, 免疫细胞含量分组差异比较, Log-rank test 检验计算预后总生存(overall survival, OS)差异, $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 HPRT1 基因在肺腺癌样本中表达情况

在 TCGA 数据集中共计有 493 例肺腺癌组织的转录组测序以及 59 例正常癌旁组织用于对照分析, 通过对比肺癌组织与癌旁组织的 HPRT1 基因

表达量(使用TPM进行定量, Transcript per million)发现该基因在肺癌组织中呈现出显著的高表达现象($P < 0.001$), 见图 1A。但将所有肺癌样本按照患者的肿瘤分期进行表达量分析发现, 该基因的表达量并未随着分期增加而显著增加表达量, 见图 1B。为了进一步确认该基因在蛋白水平的表达情况, 通过 Human Protein Atlas 数据库检索了 HPRT1 的蛋白表达结果发现, HPRT1 基因抗体(HPA006360)IHC 定量为高表达, 见图 1C。

2.2 HPRT1 基因表达与肺腺癌患者预后生存关系

将提供了预后 OS 的 TCGA 中肺腺癌患者按照 HPRT1 基因表达量从高到低排序, 其中前 30% 的患者定义为 HPRT1 基因高表达组, 后 70% 患者定义为 HPRT1 基因低表达组。比较 2 组患者的预后 OS 发现 HPRT1 基因高表达组患者的预后显著更差($P < 0.05$), 见图 2A。为了进一步验证该基因可能独立影响到患者预后, 联合患者的年龄、性别以及病理分期进行预后校正发现, 经过多因素校正的 Cox 预后模型中 HPRT1 基因的高表

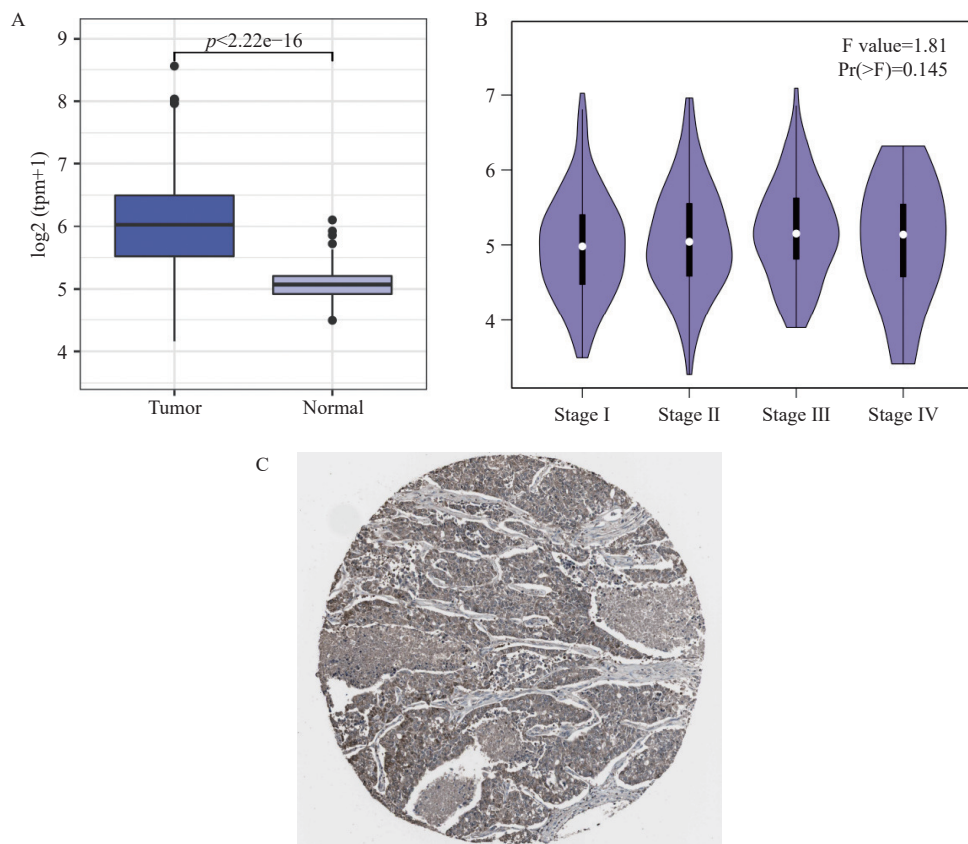


图 1 HPRT1 基因在肺腺癌样本中表达情况

Fig. 1 HPRT1 gene expression in lung adenocarcinoma samples

A: HPRT1 基因在肺腺癌组织与癌旁组织的差异; B: HPRT1 基因表达量在患者肺腺癌不同分期样本中的对比; C: 肺腺癌组织中 HPRT1 基因抗体(HPA006360)IHC 定量表达。

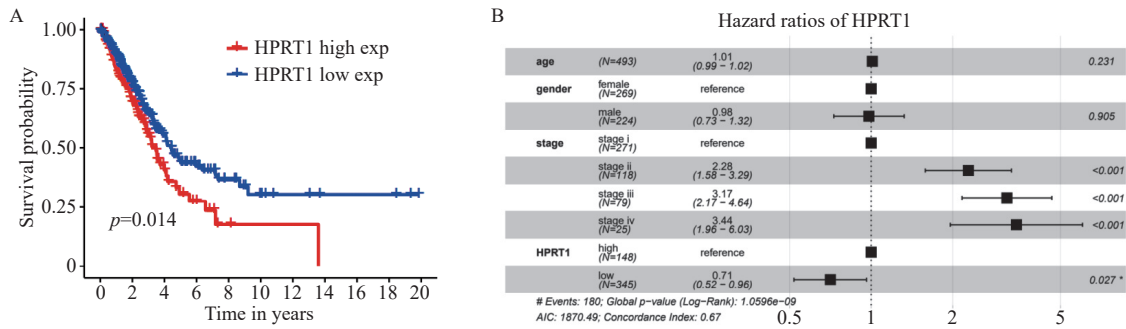


图 2 HPRT1 基因表达与肺腺癌患者预后生存关系

Fig. 2 Relationship between HPRT1 gene expression and prognostic survival in patients with lung adenocarcinoma

A: TCGA 肺腺癌数据集中 HPRT1 基因高表达与低表达生存曲线; B: 多因素校正分析。

达依然为显著预后因素 ($P < 0.05$), 见图 2B。

2.3 HPRT1 基因表达与肺腺癌患者预后关系

为进一步确定 HPRT1 基因与肺腺癌患者的预后关系是否存在样本集偶然因素, 通过检索 GEO 数据库中 HPRT1 基因表达谱以及预后 OS 发现在多个独立数据集中发现 HPRT1 基因的高表达患者预后更差所示不同数据集中 $HR > 1$, 见图 3A。数据集 GSE13213 以及 GSE67639 结果显示在各自 HPRT1 基因高表达的 30% 患者中其预后 OS 均更差 ($P < 0.01$, $P < 0.001$), 见图 3B。

2.4 HPRT1 基因高表达、低表达患者组之间基因功能分析

按照高表达的 30% 患者即为预后较差的患者定义为高风险组, 其余低表达的 70% 患者定义为低风险组。比较 2 组患者中的差异表达基因发现在高风险组中有 683 基因发生了显著上调, 见图 4A。通过 KEGG 富集分析发现该组上调的基因其功能显著富集于细胞周期以及 p53 信号通路, 见图 4B, 提示在高风险组的患者中细胞周期以及 p53 相关功能通路处于更高的激活状态。

2.5 HPRT1 高表达组患者的免疫微环境的特征分析

基于 TIMER 算法分析高、低风险组不同的免疫细胞类型浸润程度。发现 B 细胞以及 CD4T 细胞在高风险组中富集的含量更少 ($P < 0.001$), 见图 5A、5B。通过 CIBERSORT 算法的分析验证了该结果, 发现多类 B 细胞, 见图 5C。以及 CD4T 细胞, 见图 5D, 在高风险组中更低 ($P < 0.001$), 提示 HPRT1 高表达所代表的高风险患者存在显著更低的免疫细胞浸润特点。

3 讨论

本研究通过对 TCGA 肺腺癌数据中的表达谱

以及临床信息进行广泛的数据挖掘分析发现 HPRT1 基因在肺腺癌中显著上调, 且表达量越高的患者其临床中表现为预后总生存越差, 该结果提示肺腺癌患者中 HPRT1 基因的表达量会显著的影响到患者的生存情况。且该研究发现经过多个临床因素的校正证明其不受到常见临床特征的影响, 是可独立影响到肺腺癌患者预后的关键指标。本研究在多个数据队列中也进一步验证了该基因的发现, 证明了该结果的稳定性以及可靠性。

既往已有研究表明 HPRT1 基因在头颈癌以及卵巢癌中均为预后差的关键因素, 且该基因编码的蛋白是参与细胞周期活动的重要成本^[19]。本研究进一步分析发现 HPRT1 基因高表达的患者会显著上调 p53 以及细胞周期信号通路, 与既往研究结论一致^[9]。因此推测该基因会通过上调 p53 以及细胞周期等癌症高度相关的细胞通路活性, 促进肿瘤细胞的生长和存活从而影响到治疗效果。因此在临床实践中, 抑制该基因的表达可能成为改善 HPRT1 基因高表达肺腺癌患者临床疗效的一个有效措施。

肿瘤的生长永远处在与患者免疫功能斗争的过程中, 且肿瘤中免疫细胞浸润越强则杀伤肿瘤细胞的潜力越大对肿瘤细胞的抑制作用也更为明显^[20]。相反如果肿瘤环境中较少有免疫细胞的浸润, 则表现为免疫功能的活性较弱进而其肿瘤抑制的程度越弱。本研究也针对 HPRT1 基因高表达的患者和低表达的患者进行了多种免疫细胞的浸润程度对比。通常患者的免疫微环境会与患者的整体生存有着较强的关联性, 免疫细胞活跃的肿瘤中患者的整体预后会更优。为了进一步解释 HPRT1 基因高表达呈现出来预后更差的可能因素, 故笔者使用 TIMER 算法分析, 发现 B 细胞以及 CD4T 细胞均在 HPRT1 基因高表达这类预后差的患者中呈现了显著更低的特点。即表明 HPRT1 高

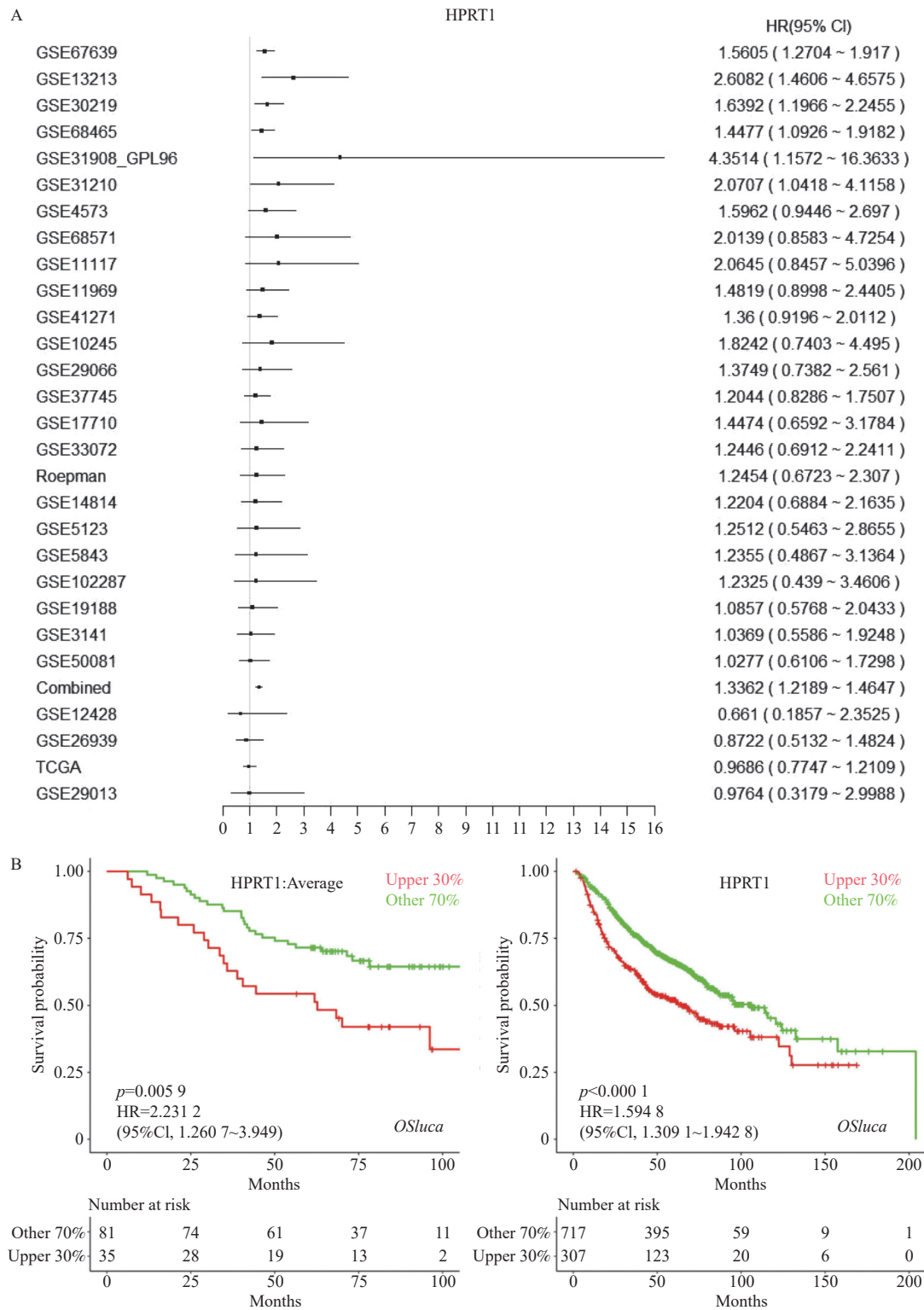


图 3 HPRT1 基因表达与肺腺癌患者预后关系

Fig. 3 Relationship between HPRT1 gene expression and prognosis of lung adenocarcinoma patients

A: GEO 中多组数据集分析 HPRT1 基因在肺腺癌中的预后特征; B: GSE13213(左)以及 GSE67639(右)验证 HPRT1 基因与预后 OS 的关系。

表达所表征的差预后人群呈现出了一定程度的免疫浸润更差的现象, 通过 CIBERSORT 算法分析提示了完全一致的结果。因此笔者为 HPRT1 高表达多表征的差预后人群找到了另 1 条差预后的可

能原因, 进而为改善临床疗效提供了 1 个备选思路。

综上所述, 本研究发现了肺腺癌中存在 1 个显著上调的基因即 HPRT1 且该基因的表达上调会

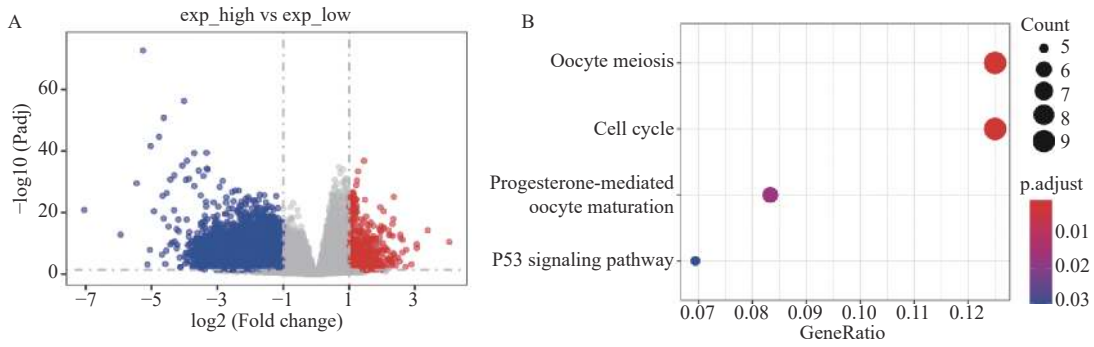


图 4 高表达、低表达 HPRT1 基因之间功能分析

Fig. 4 Functional enrichment analysis between high and low HPRT1 expression groups

A: 差异基因火山图; B: 高表达的 KEGG 富集分析。

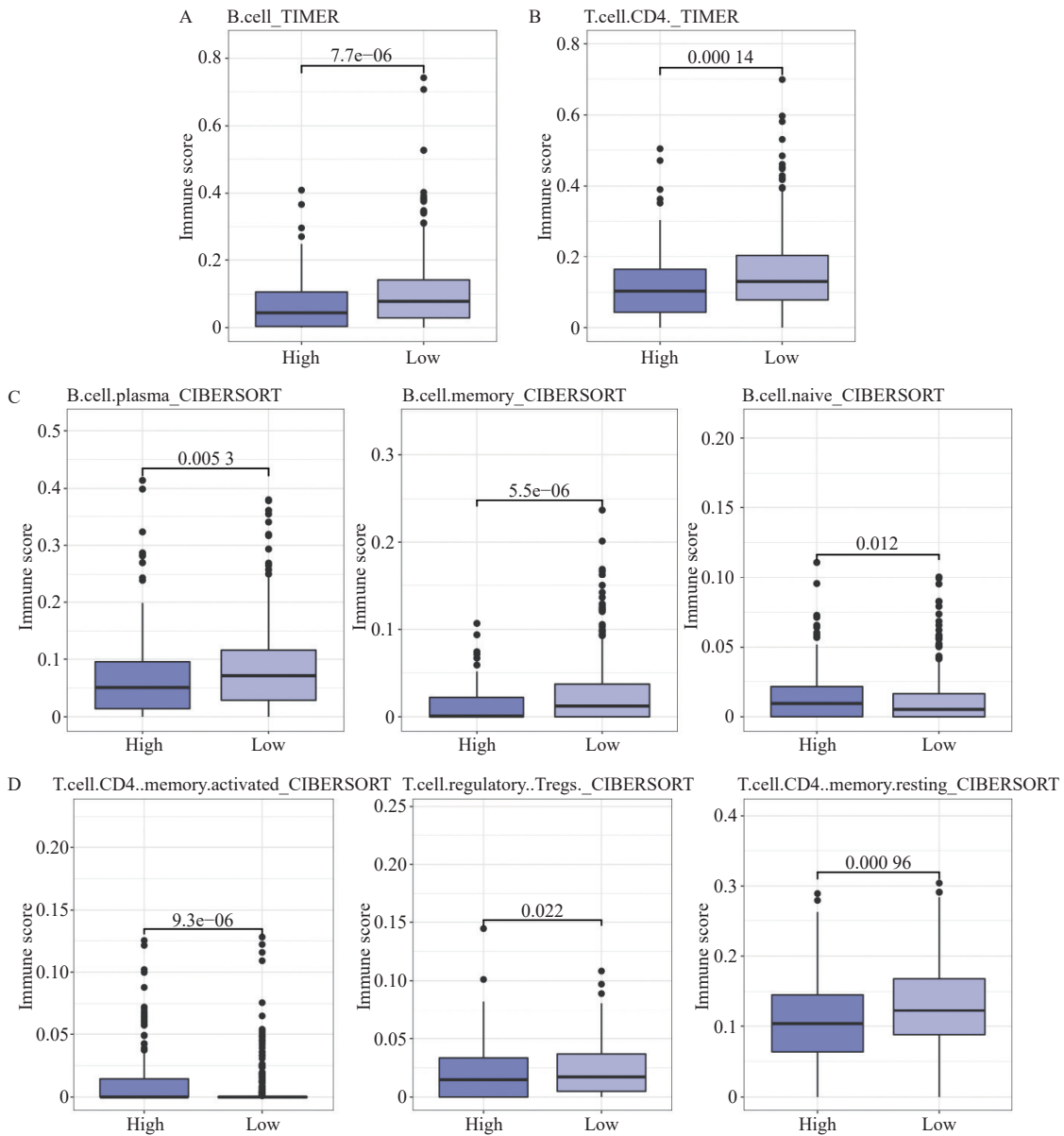


图 5 HPRT1 高表达组免疫微环境特征分析

Fig. 5 Characterization of the immune microenvironment in the HPRT1 high expression group

A: IMER 算法分析 HPRT1 基因不同表达组间 B 细胞含量对比; B: TIMER 算法分析 HPRT1 基因不同表达组间 CD4T 细胞含量对比; C: CIBERSORT 算法验证 HPRT1 基因不同表达组间不同 B 细胞含量; D: CIBERSORT 算法验证 HPRT1 基因不同表达组间不同 T 细胞含量。

在肿瘤患者中表现为 p53 以及细胞周期信报通路的过度激活, B 细胞以及 CD4T 细胞也会呈现出浸润程度的下降。患者的整体预后也会随着 HPRT1 基因的表达量升高而显著下降。

[参考文献]

- [1] Tewari K S, Colombo N, Monk B J, et al. Pembrolizumab or placebo plus chemotherapy with or without bevacizumab for persistent, recurrent, or metastatic cervical cancer: Subgroup analyses from the KEYNOTE-826 randomized clinical trial[J]. *JAMA Oncol*, 2024, 10(2): 185-192.
- [2] Ma W, Zeng J, Montoya D J, et al. Labyrinthin expression is associated with poor prognosis in patients with non-small cell lung cancer[J]. *Cancers (Basel)*, 2023, 15(3): 924.
- [3] Wang X S, Bai Y F, Verma V, et al. Randomized trial of first-line tyrosine kinase inhibitor with or without radiotherapy for synchronous oligometastatic EGFR-mutated non-small cell lung cancer[J]. *J Natl Cancer Inst*, 2023, 115(6): 742-748.
- [4] Chen F, Byrd A L, Liu J, et al. Polycomb deficiency drives a FOXP2-high aggressive state targetable by epigenetic inhibitors[J]. *Nat Commun*, 2023, 14(1): 336.
- [5] John P, Sudandiradoss C. A comprehensive integrated gene network construction to explore the essential role of Notch 1 in lung adenocarcinoma (LUAD) [J]. *J Biomol Struct Dyn*, 2024, 1-13.
- [6] Xu H, Hu Y, Peng X, et al. Prediction of prognostic and immune therapy response in lung adenocarcinoma based on MHC-I-related genes[J]. *Immunopharmacol Immunotoxicol*, 2024, 46(1): 93-106.
- [7] Gurioli G, Conteduca V, Brighi N, et al. Circulating tumor cell gene expression and plasma AR gene copy number as biomarkers for castration-resistant prostate cancer patients treated with cabazitaxel[J]. *BMC Med*, 2022, 20(1): 48.
- [8] Miao J, Kang L, Lan T, et al. Identification of optimal reference genes in golden Syrian hamster with ethanol- and palmitoleic acid-induced acute pancreatitis using quantitative real-time polymerase chain reaction[J]. *Animal Model Exp Med*, 2023, 6(6): 609-618.
- [9] Chen L, He X, Yi S, et al. Six glycolysis-related genes as prognostic risk markers can predict the prognosis of patients with head and neck squamous cell carcinoma[J]. *Biomed Res Int*, 2021, 2021: 8824195.
- [10] Gurioli G, Conteduca V, Brighi N, et al. Circulating tumor cell gene expression and plasma AR gene copy number as biomarkers for castration-resistant prostate cancer patients treated with cabazitaxel[J]. *BMC Med*, 2022, 20(1): 48.
- [11] Goldman M J, Craft B, Hastie M, et al. Visualizing and interpreting cancer genomics data via the Xena platform[J]. *Nat Biotechnol*, 2020, 38(6): 675-678.
- [12] Tomida S, Takeuchi T, Shimada Y, et al. Relapse-related molecular signature in lung adenocarcinomas identifies patients with dismal prognosis[J]. *J Clin Oncol*, 2009, 27(17): 2793-2799.
- [13] Gentles A J, Bratman S V, Diehn M, et al. Integrating tumor and stromal gene expression signatures with clinical indices for survival stratification of early-stage non-small cell lung cancer[J]. *J Natl Cancer Inst*, 2015, 107(10): djv211.
- [14] Uhlen M, Zhang C, Lee S, et al. A pathology atlas of the human cancer transcriptome[J]. *Science*, 2017, 357(6352): eaan2507.
- [15] Yan Z, Wang Q, Lu Z, et al. OSLuca: An interactive web server to evaluate prognostic biomarkers for lung cancer[J]. *Front Genet*, 2020, 11: 420.
- [16] Yu G, Wang L G, Han Y, et al. clusterProfiler: An R package for comparing biological themes among gene clusters[J]. *OMICS*, 2012, 16(5): 284-287.
- [17] Wu M, Jiang L. Hydroxysteroid 17-beta dehydrogenase 6 is a prognostic biomarker and correlates with immune infiltrates in hepatocellular carcinoma[J]. *Dig Dis Sci*, 2022, 67(1): 146-158.
- [18] Chen B, Khodadoust M S, Liu C L, et al. Profiling tumor infiltrating immune cells with CIBERSORT[J]. *Methods Mol Biol*, 2018, 1711: 243-259.
- [19] Mavridis K, Stravodimos K, Scorilas A. Quantified KLK15 gene expression levels discriminate prostate cancer from benign tumors and constitute a novel independent predictor of disease progression[J]. *Prostate*, 2013, 73(11): 1191-201.
- [20] 朱中山, 杨洲, 江承川, 等. 肺腺癌患者 PLA2G1B 表达情况与预后的相关性 [J]. *昆明医科大学学报*, 2022, 43(9): 70-76.