**多组学生物信息学分析宫颈癌中YTHDF2功能及其关键靶基因预测**

**冉朝霞2、王金桃2、郭翀宇2、李德铖2、彭煜舒2、宋丽2、赵卫红3、**

**田志强1、丁玲2**

**1. 山西省白求恩医院**

**2. 山西医科大学公共卫生学院**

**3. 山西医科大学第二医院妇产科**

**目的：**探索m6A识别蛋白YTHDF2在宫颈癌发生、进展中的功能与作用并预测其关键靶基因。

**方法：**从Gene Expression Omnibus（GEO）的数据集GSE7803、GSE9750获取包含正常宫颈和宫颈癌组织的基因表达数据，使用limma包分析基因差异表达和YTHDF2表达情况。同时，从GEO数据库获取敲除YTHDF2前后hela细胞差异基因表达数据集GSE49339，对差异基因进行基因本体论（GO）、京都基因与基因组百科全书（KEGG）通路分析。GSE49339与ENCORI数据库中YTHDF2 CLIP-seq数据、GEO数据库中hela细胞的meRIP-seq数据集GSE165690取交集获得潜在靶基因数据；进一步将获得的潜在靶基因数据与人群基因数据集（GSE7803、GSE9750）取交集筛选出YTHDF2关键靶基因。从GEPIA数据库获取宫颈癌生存数据，采用在线工具Kaplan-meier plotter进行关键靶基因的生存分析。从UCSC数据库获取宫颈癌患者基因表达数据，并利用ESTIMATE和IOBR包计算B细胞、CD4+T细胞、CD8+T细胞等6种免疫细胞浸润评分，利用软件包psych计算关键靶基因表达量与各类免疫浸润评分的Pearson系数。

**结果：**GSE7803和GSE9750共筛选出255个差异表达基因，其中YTHDF2在宫颈癌组织中呈现高表达。敲除YTHDF2前后hela细胞的GSE49339数据集中筛选出3763个差异表达基因，GO分析显示，差异表达基因主要发挥RNA聚合酶II启动子的转录调控、DNA模板转录调控和RNA聚合酶II核心启动子近端区序列特异性DNA结合的功能；KEGG分析表明，差异基因主要参与HPV感染、MAPK、mTOR、Wnt等细胞信号通路。多组学数据集互作分析鉴定出12个关键靶基因，其中DDIT4、MEIS2、RHOB、IRS1、KLF4、SOWAHC、MFHAS1、PLAGL1与生存率有关；KLF4与免疫浸润评分相关，Timer分析显示KLF4与B 细胞、中性粒细胞、巨噬细胞相关。

**结论：**YTHDF2在宫颈癌中高表达，可参与细胞周期、凋亡的信号通路，其关键靶基因KLF4可能同时影响宫颈癌患者的预后和免疫环境，特别发现YTHDF2可作用于HPV感染细胞信号通路，提示YTHDF2与宫颈癌的发生、进展密切相关。