

基于 GEO 数据库的宫颈鳞癌预后分析及 NEK2 对糖酵解的调控

于盼盼

哈尔滨医科大学附属第四医院 黑龙江 哈尔滨 150000

【摘要】目的：利用公共生物信息学数据库筛选出与宫颈鳞状细胞癌（cervical squamous cell carcinoma, CSCC）相关的关键基因，分析这些基因在肿瘤发生发展过程中所起的作用。方法：从 Gene Expression Omnibus (GEO) 数据库下载 CSCC 相关基因表达数据集 GSE29570 和 GSE63514，利用 GEO2R 在线分析工具筛选差异表达基因 (DEGs)，以 $P_{adj} < 0.05$ 且 $|\text{LogFC}| > 1$ 为筛选条件，并通过韦恩图获得两个数据集的交集基因。结果：GEO 数据库分析得到 340 个交集差异表达基因，上调基因 231 个，下调基因 103 个。结论：NEK2 可以促进糖酵解途径来提高肿瘤细胞的增殖能力，进而影响患者的预后。NEK2 可以成为 CSCC 的预后生物标志物和分子治疗靶点，给宫颈鳞癌的机制研究以及靶向治疗提供新的研究思路。

【关键词】：宫颈鳞癌；NEK2；GEO 数据库；糖酵解；生物信息学分析

DOI:10.12417/2705-098X.26.13.077

引言

NEK2 (NIMA-related kinase 2) 属于细胞周期调控的丝氨酸/苏氨酸蛋白激酶，它参与中心体分离、纺锤体形成以及细胞有丝分裂。已有研究显示 NEK2 在多种恶性肿瘤中存在高表达现象，它同肿瘤细胞的增殖，染色体不稳定以及化疗耐药等过程有着联系。但是 NEK2 在宫颈鳞癌中起的作用还不清楚，尤其是它是否参与肿瘤代谢重编程还没有系统的研究。因此，本研究从 GEO、TCGA 数据库中筛选出 CSCC 相关的关键基因，主要对 NEK2 在宫颈鳞癌中的表达情况以及它所参与的糖酵解通路进行分析，希望给 CSCC 的分子机制研究以及靶向治疗提供新的理论基础。

1 材料与方法

1.1 GEO 数据库分析关键基因

本研究利用基因综合表达 (gene expression omnibus, GEO) 数据库的 CSCC 患者芯片，筛选出差异基因并进行生物信息学分析，从而发现与 CSCC 相关的关键基因和信号通路。

1.2 数据收集

以 “cervical cancer”、“cervical squamous cell carcinoma” 为检索词，以 “Homo Sapiens” 为研究对象进行检索，从 GEO 数据库 (<https://www.ncbi.nlm.nih.gov/geo/>) 中下载 CSCC 相关的基因表达数据集 GSE29570 和 GSE63514 的基因表达谱，采用 GEO2R 在线分析软件分析，以 $P_{adj} < 0.05$ 且 $|\text{LogFC}| > 1$ 为筛选条件，得到与 CSCC 相关的差异表达基因 (differentially expressed gene, DEGs)。绘制火山图、韦恩图、热图。

1.3 GO 功能注释和 KEGG 富集分析

使用 DAVID 线数据进行基因本体 (gene ontology, GO)

分析，包括：生物过程 (biological process, BP)、分子功能 (molecular function, MF) 和细胞组成 (cellular component, CC)。并进行 KEGG 的富集分析。

1.4 蛋白质互作网络(PPI)的构建

使用 STRING 数据库对交集的差异基因获取靶点蛋白之间的相互作用信息。蛋白互作分析结果导入 Cytoscape 软件，绘制蛋白质互作网络图。

1.5 关键基因构建

使用 CytoHubba 基于四种拓扑算法，对差异基因连通性进行计算，最大聚集中心 (MCC)、最大邻域组件 (MNC)、度 (Degree)、边缘渗透组件 (EPC)，分别计算出 PPI 网络中前二十个重要节点，将上述四种算法所获得的结果取交集，将重叠的 DEGs 作为关键枢纽基因。

1.6 显著模块基因簇构建

使用 Cytoscape 中 MCODE 插件，以 $k_{core}=2$, $degree\ cutoff=2$, $node\ score\ cutoff=0.2$ 与 $max.depth=100$ 为标准，筛选出显著模块。

1.7 生存曲线和关键基因的预后分析

应用 GEPIA2 对关键基因进行预后评估分析，此平台共有 292 例宫颈癌患者可用于分析总生存率，对于每个基因 cut off 值设置为 50%，将患者分为高表达和低表达两组进行比较， $P < 0.05$ 提示关键基因对患者总体生存率有影响。

2 TCGA 数据库 GSEA 富集分析糖酵解通路

2.1 数据下载与整理

从 TCGA 数据库中下载 CSCC 基因转录组数据及临床数据，并删除临床数据中的未知及缺失数据。

2.2 基因富集分析

利用 GSEA 探索 NEK2 可能参与的信号通路,发现涉及到一些糖代谢相关通路,其中有关糖代谢的改变就是有氧糖酵解,即肿瘤细胞在有氧的情况下,也优先选择糖酵解供能,以此来满足肿瘤细胞快速生长的需要。将 CSCC 患者数据按 NEK2 的高低分为两组,进一步对 NEK2 进行糖酵解基因富集分析。分析结果以标准化富集分数(NES)>1.5 和 P 值(P<0.05)设定临界值。

3 结果

3.1 GEO 数据库分析结果

3.1.1 DEGs 基因

GSE29570 数据集得到 881 个 DEGs,其中上调基因 348 个,下调基因 533 个,火山图,见图 1A。GSE63514 数据集得到 1553 个 DEGs,其中上调基因 1028 个,下调基因 525 个,火山图,见图 1B。韦恩图分析得到交集 340 个 DEGs,其中上调交集基因 231 个,下调交集基因 103 个,见图 1C。

3.1.2 DEGs 的 GO 功能注释和 KEGG 富集分析

表明 CSCC 可能是通过参与这些生物学过程和通路发挥调控作用。各组排名前 10 的条目,见图 1D 和 1E。这些差异基因对细胞周期、细胞分裂和 DNA 复制至关重要。不受控制的细胞增殖和 DNA 复制被认为是癌症的标志之一。

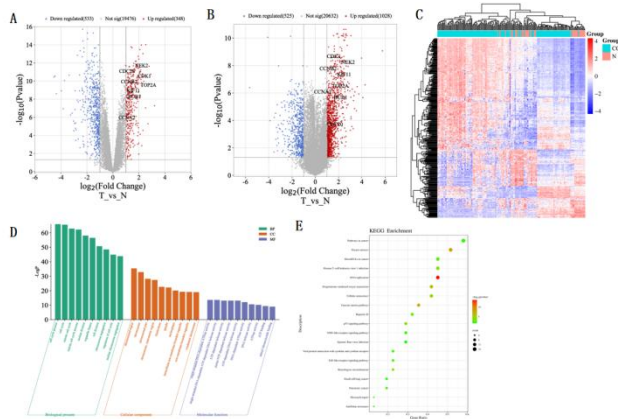


图 1 GEO 数据库分析结果

A.GSE29570 数据集火山图 B.GSE63514 数据集火山图 C.交集 DEGs 热图 D.GO 功能注释 E.KEGG 富集分析

3.1.3 关键基因构建

构建 PPI 网络,利用 CytoHubba 基于 MCC、MNC、Degree 和 EPC 四种拓扑算法筛选出排名前二十的关键基因,确定了四种算法的重叠基因作为关键枢纽基因,结果显示蛋白质相互作用网络中存在 1 个比较重要的聚类。发现关键基因与显著模块基因簇的基因完全重合。

3.1.4 NEK2 生存曲线和预后分析

在 GEPIA 上 8 个关键基因表达均上调,有作为早期诊断的潜在可能,见图 2A。其中 NEK2 没有 CSCC 方面的报道。生存分析显示 NEK2 的 Disease Free Survival 和 Overall Survival,见图 2B 和 2C, P 值 0.052 和 0.16 虽然不是严格小于 0.05,但由于 GEPIA 中样本量有限,且 NEK2 高表达的患者生存期也有低于低表达患者的趋势结果,也能推测 NEK2 可能是影响预后的潜在靶点。

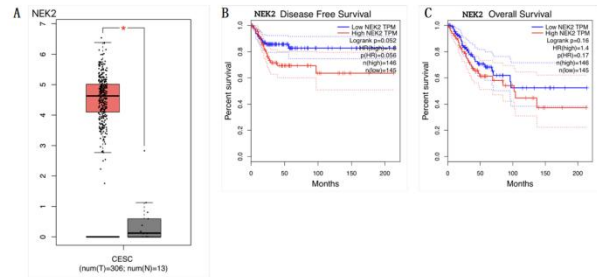


图 2 NEK2 的表达和生存曲线

A.NEK2 的 mRNA 表达值 B.NEK2 的 RFS 生存曲线 C.NEK2 的 OS 生存曲线

3.2 TCGA 数据库分析结果

CSCC 患者表达谱 GSEA 富集分析富集到糖酵解通路,说明 CSCC 的发生伴随着糖酵解途径的激活,见图 3A、3B、3C、3D。进一步对 NEK2 单基因 GSEA 富集分析也富集到糖酵解通路,见图 3E 和 3F。说明 NEK2 基因可能通过影响糖酵解途径影响肿瘤的增殖。

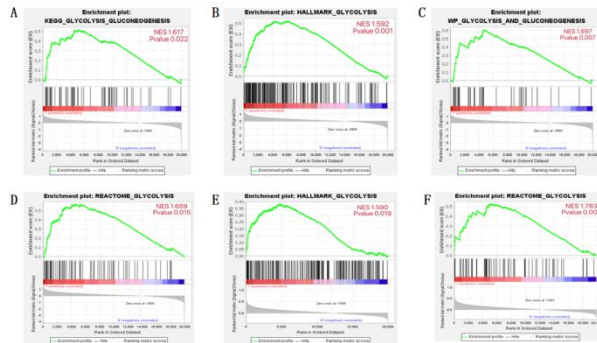


图 3 CSCC 患者 GSEA 糖酵解通路富集分析

A.CSCC 表达谱 KEGG 分析富集糖酵解通路 B.CSCC 表达谱 KEGG 分析富集糖酵解通路 C.CSCC 表达谱 KEGG 分析富集糖酵解通路 D.CSCC 表达谱 KEGG 分析富集糖酵解通路 E.NEK2 单基因分析富集糖酵解通路 F.NEK2 单基因分析富集糖酵解通路

4 讨论

本研究从两个独立的 GEO 数据集中筛选出 341 个宫颈癌共有差异基因,在其中 NEK2 上调表达最明显,提示它在宫

颈鳞癌的发生中可能起到原癌基因的作用^[5]。GO 功能富集分析显示, 差异基因主要集中在细胞周期过程、有丝分裂、染色体分离等生物学过程中, 这与 NEK2 作为细胞周期调节的核心

因子所知的生理功能高度一致, NEK2 主要参与中心体分裂、染色体分离等细胞周期的重要步骤, 它的异常高水平表达会导致细胞周期的紊乱, 促进肿瘤细胞的无限增殖。

参考文献:

- [1] Souid M,Ghedira R,Souissi S,Bouzgarrou N,Gabbouj S,Shini-Hadhri S,et al.Arginase is involved in cervical lesions progression and severity.Immunobiology.2022;227:152189.
- [2] Hernandez F.Glycolysis and gluconeogenesis:A teaching view.J Biol Chem.2021;296:100016.
- [3] Ma D,Huang Y,Song S.Inhibiting the HPV16 oncogene-mediated glycolysis sensitizes human cervical carcinoma cells to 5-fluorouracil. Onco Targets Ther.2019;12:6711-20.
- [4] Zhang Y,Zhao L,Yang S,Cen Y,Zhu T,Wang L,et al.CircCDKN2B-AS1 interacts with IMP3 to stabilize hexokinase 2 mRNA and facilitate cervical squamous cell carcinoma aerobic glycolysis progression.J Exp Clin Cancer Res.2020;39:281.
- [5] Zhang G,Zhao X,Liu W.NEDD4L inhibits glycolysis and proliferation of cancer cells in oral squamous cell carcinoma by inducing ENO1 ubiquitination and degradation.Cancer Biol Ther.2022;23:243-53.