

基于生物信息学途径预测 *MELK* 基因在肺腺癌及肺鳞癌中的功能机制及临床意义

杨云英, 王雪涔, 彭振维, 陈勇
(中山大学附属第一医院放射治疗科, 广东 广州 510080)

摘要:【目的】通过生物信息学方法,预测肺腺癌及肺鳞癌中高表达的母体胚胎亮氨酸拉链激酶(*MELK*) 在疾病发生发展中的功能机制及临床意义。【方法】从基因表达数据(GEO)获取芯片数据集GSE7670,结合肿瘤基因组图谱(TCGA)肺腺癌及肺鳞癌数据库进行差异分析,筛选差异激酶($P < 0.05$)。对这些激酶进行GO生物学功能富集分析及KEGG信号通路分析。通过String数据库绘制蛋白相互作用网络,并利用Cytoscape的MCODE插件找到关键节点蛋白,挑选出*MELK*激酶。通过Oncomine数据库分析*MELK*表达,从TCGA数据库获取临床信息,分析*MELK*表达与肺腺癌及肺鳞癌临床病理特征及预后的关系。采用GeneMANIA预测*MELK*功能,进而用GEPIA2验证相关共表达基因。【结果】共筛选出159个与细胞增殖、凋亡及各信号通路密切相关的差异激酶。蛋白质相互作用网络分析发现21个关键基因。*MELK*在男性、年龄较小、高分期的肺腺癌及肺鳞癌中表达较高。在肺腺癌中,*MELK*高表达与较短的总生存期(OS)及无进展生存期(PFS)显著相关,而在肺鳞癌中则无显著差异。*MELK*可能参与G2/M期转化,且在肺腺癌及肺鳞癌中均与增殖相关基因*MKI67*、*PCNA*、*CCNB1*、*MCM2*和*TOP2A*共表达。【结论】*MELK*在肺腺癌及肺鳞癌中高表达,可能通过促进细胞有丝分裂及细胞周期G2/M期转化参与肺腺癌的发生发展,有望成为肺腺癌新的生物标记物。

关键词:肺腺癌;肺鳞癌;生物信息学;*MELK*

中图分类号:R73 文献标志码:A 文章编号:1672-3554(2020)06-0891-11

Prediction of Functional Mechanism and Clinical Significance of *MELK* in Lung Adenocarcinoma and Lung Squamous Cell Carcinoma Based on Bioinformatics Analysis

YANG Yun-ying, WANG Xue-cen, PENG Zhen-wei, CHEN Yong

(Department of Radiation Oncology, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou 510080, China)

Correspondence to: CHEN Yong; E-mail: chenying@mail.sysu.edu.cn

Abstract: 【Objective】 To predict the functional mechanism and clinical significance of highly expressed maternal embryonic leucine zipper kinase (*MELK*) in lung adenocarcinoma and lung squamous cell carcinoma by bioinformatics. 【Methods】 The chip data set GSE7670 was obtained from the Gene Expression Omnibus (GEO) database and combined with the TCGA lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) database for differential analysis, and to screen out differential kinases ($P < 0.05$). GO biological function enrichment and KEGG pathway analysis of these kinases were performed. Based on String, a protein interaction network was obtained. MCODE plug-in of Cytoscape was applied to find key node proteins and then *MELK* kinase protein was picked out. *MELK* expression was analyzed through Oncomine database. Clinical information from TCGA database was obtained, and the relationship between *MELK* expression and clinicopathological characteristics, prognosis of LUAD and LUSC was analyzed. GeneMANIA was used to predict the function of *MELK*, and then to verify related co-expressed genes with GEPIA2. 【Results】 A total of 159 differ-

收稿日期:2020-07-15

基金项目:国家自然科学基金(81770608)

作者简介:杨云英,主管护师,研究方向:肿瘤学,E-mail: yangyy9@mail.sysu.edu.cn;陈勇,通信作者,教授,研究方向:肿瘤学、头颈部放疗,E-mail: chenying@mail.sysu.edu.cn

ential kinases were screened out, which were closely related to cell proliferation, apoptosis and various signal pathways. Twenty-one important genes were identified by analyzing protein interaction networks. *MELK* level was higher in male, young or advanced clinical stage lung adenocarcinoma and lung squamous cell carcinoma. The patients with high *MELK* levels had a shorter progression-free survival (PFS) and overall survival (OS) in LUAD, but there was no significant difference in LUSC. *MELK* was involved in G2/M transition possibly and was co-expressed with proliferation-related genes (*MKI67*, *PCNA*, *CCNB1*, *MCM2*, and *TOP2A*) in LUAD and LUSC. 【Conclusions】 *MELK* is highly expressed in LUAD and LUSC. *MELK* may be involved in the occurrence and development of LUAD by promoting cell mitosis and G2/M transition. *MELK* could be a new biomarker for LUAD.

Key words: lung adenocarcinoma; lung squamous cell carcinoma; bioinformatics; *MELK*

[J SUN Yat-sen Univ (Med Sci), 2020, 41(6): 891-901]

肺癌是我国发生率和死亡率最高的恶性肿瘤,非小细胞肺癌(non-small cell cancer, NSCLC)占肺癌的85%^[1],主要病理类型为肺腺癌和肺鳞癌。肺癌是生存率最低的恶性肿瘤之一,由于难以及时发现,大多数患者确诊时即为晚期,而晚期肺癌5年生存率仅为5%^[2]。蛋白激酶参与调控各类细胞信号转导途径,与细胞凋亡、增殖、细胞代谢及恶性肿瘤的发生发展密切相关^[3],因此,激酶在恶性肿瘤的治疗中成为最重要的靶点之一,截至2019年,美国FDA已批准48种小分子蛋白激酶抑制剂进行肿瘤靶向治疗^[4]。大约30%非小细胞肺癌患者存在预测治疗反应的分子改变,如靶向表皮生长因子受体(epidermal growth factor receptor, EGFR)突变,靶向间变性淋巴瘤激酶(anaplastic lymphoma kinase, ALK)重排、c-ros原癌基因1受体酪氨酸激酶(c-ros oncogene 1 receptor tyrosine kinase, ROS1)重排等等,因此,靶向EGFR酪氨酸激酶抑制剂(tyrosine kinase inhibitors, TKI)、ALK、ROS1、间质表皮转化因子(mesenchymal to epithelial transition factor, MET)、血管内皮细胞生长因子受体(vascular endothelial growth factor receptor, VEGFR)等激酶抑制剂在晚期非小细胞肺癌的治疗中发挥着重要作用^[5],但仍然存在靶向药适用患者少、治疗无效、肿瘤耐药复发等情况。因此,在NSCLC中挖掘新的激酶,探究相关功能与机制、明确相关临床意义很有必要,可能为NSCLC的治疗提供新的靶点。随着生物信息学快速发展,基因组学、转录组学、蛋白组学数据库越来越完善,可以为我们发现功能基因、预测相关机制、阐明临床意义提供重要线索。本研究利用高通量基因表达数据库(Gene Expression Omnibus, GEO)和肿瘤基因组图谱(The Cancer Genome Atlas,

TCGA)数据库筛选得到159个差异激酶基因,预测相关功能和机制,寻找关键节点蛋白,筛选出母体胚胎亮氨酸拉链激酶(maternal embryonic leucine zipper kinase, *MELK*)基因进行后续分析。*MELK*参与多种细胞过程,与细胞周期调控、细胞增殖、凋亡、基因表达、肿瘤发生密切相关^[6]。预测*MELK*功能与机制、明确其临床意义,可能为其成为肺腺癌、肺鳞癌的治疗新靶点提供重要依据。

1 材料与方法

1.1 基因芯片数据

从美国国家生物信息中心(National Center for Biotechnology Information, NCBI)中的GEO数据库里查找同时存在癌组织与相应癌旁组织的肺癌芯片,获取GSE7670芯片(来自Affymetrix Human Genome U133A Array平台),得到27个病人配对的肺腺癌和癌旁组织的mRNA表达数据。

1.2 差异基因筛选

采用GEO数据库中的GEO2R工具对GSE7670芯片中的癌和癌旁组织mRNA表达数据进行差异分析($P < 0.05$),并绘制火山图。接着,在TCGA肺腺癌和肺鳞癌数据库中获取GEO2R分析出的差异激酶基因的mRNA表达数据,使用R软件包“Limma”进一步筛选癌和相应正常组织的差异激酶基因($P < 0.05$),最终筛选出候选的差异激酶基因。

1.3 基因富集分析

利用“DAVID Bioinformatics Resources”工具(<https://david.ncifcrf.gov/>)对筛选出的159个差异激酶基因进行GO富集分析:生物过程(biological process, BP)、细胞成分(cellular component)、分子

功能 (molecular function, MF) 以及 KEGG 通路富集分析。绘制气泡图:GO 分析采用 FDR 为横坐标, KEGG 采用 P 值为横坐标, 各信号通路为纵坐标。

1.4 蛋白质相互作用网络构建及分析

使用 String 数据库 (<http://string-db.org/cgi/input.pl>) 绘制 159 个激酶蛋白质相互作用网络 (Protein-protein Interaction Network, PPIN) 图, 导入 Cytoscape3.8.0 软件。利用软件中的 MCODE (Molecular Complex Detection) 插件, 根据关联分值, 在蛋白网络中聚类, 构建关键功能模块, 找到关键节点蛋白, 并将其标注于 (图 1) 的火山图中。MCODE 插件筛选核心基因的标准: Degree Cutoff=2, Node Score Cutoff=0.2, K-core=2, Max Depth=100。

1.5 *MELK* 基因表达及与肺腺癌及肺鳞癌临床病理特征的关系

在各关键节点蛋白中选择研究相对较少的 *MELK* 进行后续分析。在 Oncomine 数据库 (<https://www.oncomine.org/resource/main.html>) 中明确 *MELK* 在不同肿瘤中的表达概况; 比较 *MELK* 在正常肺组织、肺腺癌、肺鳞癌组织的 mRNA 水平 (Hou Lung Dataset Summary, 来源于 GEO 数据库 <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE19188>)。同时, 使用 SPSS 20.0, 以表达值、总生存时间 (overall survival, OS)、OS 生存状态作 ROC 曲线, 曲线中敏感性+特异性最大时所对应的 *MELK* 表达值, 即为划分 *MELK* 高表达和低表达的临界值。利用该临界值, 分析 TCGA 数据库中 *MELK* 与肺腺癌及肺鳞癌各临床病理特征的关系。TCGA 数据库中有肺鳞癌 502 例, 癌旁组织 49 例以及肺腺癌 535 例, 癌旁组织 59 例, 部分临床病理指标存在缺失值。

1.6 *MELK* 基因表达及与肺腺癌及肺鳞癌预后的关系

在 TCGA 数据库中获得各患者的生存状态、总生存时间 OS 以及肿瘤进展情况、无进展生存期 (progression-free survival, PFS) 数据, 计算 *MELK* 高表达组和低表达组生存期平均值, 并进行单因素生存分析。部分数据存在缺失值, 删除掉有缺失值的样本, 样本量如下: 肺腺癌 OS 分析中, *MELK* 低表达样本 193 例, *MELK* 高表达样本 316 例; PFS 分析中, *MELK* 低表达样本 184 例, 高表达样本 301 例。肺鳞癌 OS 分析中, *MELK* 低表达样本 198 例, *MELK* 高表达样本 299 例; PFS 分析中,

MELK 低表达样本 194 例, 高表达样本 288 例。之后, 进一步对各临床参数进行单因素、多因素的生存分析。

1.7 *MELK* 功能预测及共表达分析

利用 GeneMANIA (<http://genemania.org/>) 工具对目标基因 *MELK* 进行相关基因网络构建并预测其功能, 进而采用 GEPIA2 (<http://gepia2.cancer-pku.cn/#correlation>) 工具检验在 TCGA 数据库中, *MELK* 在肺腺癌和肺鳞癌中是否与增殖相关基因 (*MKI67*、*PCNA*、*CCNB1*、*MCM2* 和 *TOP2A*) 共表达 (相关系数采用 Pearson 分析)。

1.8 数据分析和统计

采用 SPSS 20.0 软件进行统计学分析, 采用 GraphPad8.0 软件作图。中位生存期采用中位数和四分位数, 即 $M(P_{25} \sim P_{75})$ 表示。*MELK* 与临床参数相关性使用卡方检验。三组比较采用单因素方差分析, 有统计学意义时, 组间多重比较采用 Dunnett T3 法。单因素生存分析使用 Kaplan-Meier (log-rank), 多因素生存分析使用 COX 比例风险回归模型。 $P < 0.05$ 时差异有统计学意义。

2 结果

2.1 差异表达基因分析

在 GSE7670 芯片中, 将肺腺癌组织和相应癌旁组织的 mRNA 表达量进行差异分析并绘制火山图 (图 1): 筛选出 6 380 个差异基因, 其中, 3 394 个基因上调 (红色), 2 986 个基因下调 (绿色)。其中包含 303 个差异激酶基因。接着, 在 TCGA 肺腺癌和肺鳞癌数据库中查找这 303 个激酶的 mRNA 表达数据, 再次进行差异基因分析, 最终筛选出 159 个激酶。

2.2 基因富集分析

通过 GO 分析和 KEGG 分析, 发现这 159 个激酶与各磷酸化级联反应相关, 参与调控细胞周期、细胞增殖、细胞分化以及细胞黏附等生物学过程。同时参与调控各信号传导通路, 如 MAPK、VEGF、mTOR、P53、细胞凋亡等, 与肿瘤的发生发展密切相关 (图 2)。

2.3 蛋白质相互作用网络

利用 Cytoscape3.8.0 的 MCODE 插件, 在 159 个激酶中筛选出 21 个关键节点激酶蛋白 (图 3): TK1、*MELK*、NEK2、CDK4、PLK1、BUB1、AURKB、

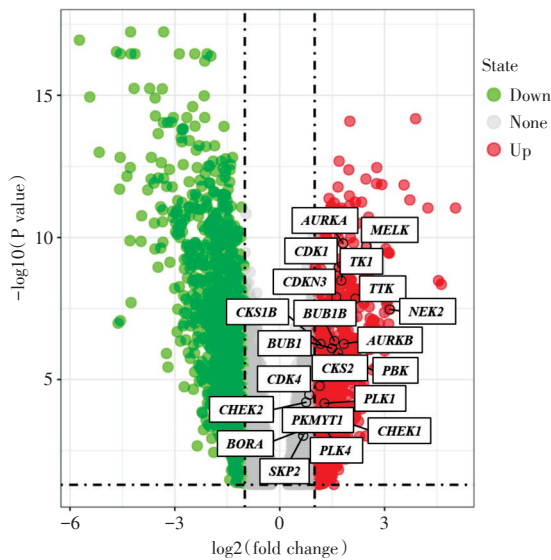


图1 GSE7670芯片中肺腺癌和癌旁组织差异表达基因火山图

Fig.1 Volcano plots of differentially expressed genes in lung adenocarcinoma and adjacent tissues in GSE7670 chip

PKMYT1、BUB1B、BORA、PBK、CKS2、CDKN3、CHEK2、CKS1B、CHEK1、PLK4、TTK、CDK1、SKP2和AURKA。预测评分MCODE-score=12~14之间,颜色越深,代表预测评分越高。将这21个蛋白标注于GSE7670芯片中肺腺癌和癌旁组织差异表达基因火山图中(图1),可以看到,绝大部分基因在癌组织中高表达。

2.4 MELK的表达情况

在Oncomine数据库中可以看到,*MELK*在包括肺癌在内的多种恶性肿瘤中高表达(图4A)。将正常肺组织、肺鳞癌(LUSC)、肺腺癌(LUAD)3组*MELK*基因的mRNA进行比较(图4B),整体而言组间均数有统计学差异($F = 148.024, P < 0.001$),进一步进行组间多重比较:LUSC和LUAD *MELK*基因的mRNA水平较正常肺组织均显著增高($P < 0.001$),同时,LUSC与LUAD比较,其*MELK*基因的mRNA水平显著增高($P = 0.001$)。

2.5 MELK与肺腺癌及肺鳞癌的临床病理特征关系

根据ROC曲线,选择肺腺癌*MELK*表达值分界点为2.71,肺鳞癌*MELK*表达值分界点为7.767,按照分界点将*MELK*分为高表达、低表达组进行后续分析。在肺腺癌中,*MELK*高表达与男性、年龄 ≤ 66 岁(中位数)、吸烟、较高的临床N分期、较高的临床TNM分期、患者已死亡(OS状态)、肿瘤

未进展(PFS阴性事件)具有显著相关性($P < 0.05$;表1);在肺鳞癌中,*MELK*高表达与性别女性、年龄 ≤ 68 (中位数)、较高的临床N分期、较高的临床TNM分期显著相关($P < 0.05$),但与OS、PFS均无相关性(表2)。

2.6 MELK与肺腺癌及肺鳞癌预后的关系

进行Kaplan-Meier单因素生存分析(图5),在肺腺癌中*MELK*低表达组OS中位生存期为26.87(18.00~42.97)月,高于*MELK*高表达组OS中位生存期22.32(11.4~37.4)月,差异有统计学意义($\chi^2 = 16.744, P < 0.001$);*MELK*低表达组PFS中位生存期为20.3(13.8~36.3)月,高于高表达组PFS中位生存期15.9(7.68~26.43)月,差异有统计学意义($\chi^2 = 12.511, P < 0.001$)。在肺鳞癌中*MELK*低表达组OS中位生存期为21.33(11.65~37.47)月,低于*MELK*高表达组OS中位生存期23.30(10.23~44.60)月,差异无统计学意义($\chi^2 = 0.625, P = 0.429$);*MELK*低表达组PFS中位生存期为19.28(8.72~35.3)月,高于*MELK*高表达组PFS中位生存期18.97(7.86~39.3)月,差异无统计学意义($\chi^2 = 0.053, P = 0.817$)。利用Cox模型对肺腺癌进行单因素和多因素生存分析显示(表3~4),在单因素生存分析中,PFS与临床T分期、N分期、肿瘤TNM分期、*MELK*表达水平显著相关($P < 0.05$),且HR均大于1;OS与年龄、临床T分期、临床N分期、临床M分期及TNM分期、*MELK*表达水平显著相关($P < 0.05$)。而在多因素生存分析中,PFS仅与临床T分期、*MELK*表达水平显著相关($P < 0.05$);OS仅与年龄、临床T分期、N分期及*MELK*表达水平显著相关($P < 0.05$)。

2.7 MELK功能预测及共表达分析

我们通过GeneMANIA预测可能与*MELK*具有相同功能且相互作用的基因,并且预测了*MELK*的潜在功能(图6)。在功能上提示*MELK*可能促进G2/M细胞周期的转化,推测*MELK*可能与细胞增殖相关。因而,进一步检测*MELK*与增殖相关基因的共表达(图7),结果表明,在LUAD中,*MELK*均与*MKI67*、*PCNA*、*CCNB1*、*MCM2*和*TOP2A*表达呈明显正相关($P < 0.05$,相关系数R分别为:0.75,0.65,0.81,0.7,0.66);在LUSC中,*MELK*也与其呈显著正相关($P < 0.05$,相关系数R分别为:0.36,0.46,0.36,0.37,0.38),但相关性较LUAD较低。



A-C:Go analysis:Biological Process(A), Cellular Component(B), Molecular function(C);D:KEGG analysis

图2 159个激酶的GO分析与KEGG分析
Fig.2 GO and KEGG analysis of 159 kinases

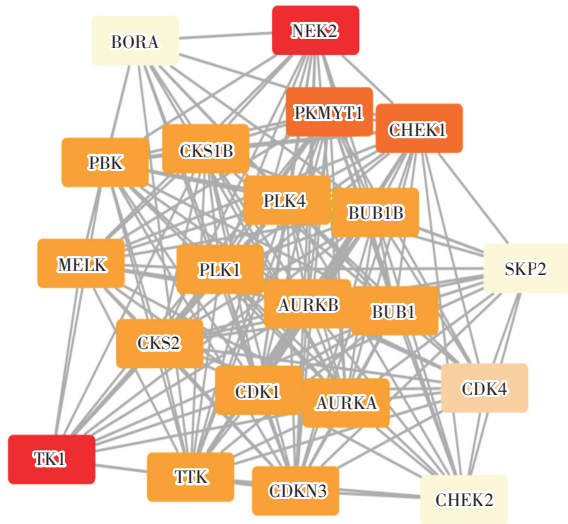
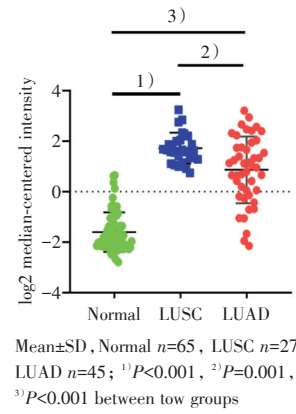


图3 159个激酶中关键蛋白的蛋白质相互作用网络
Fig.3 Protein-protein interaction network of key genes in 159 kinases

Disease Summary for MELK

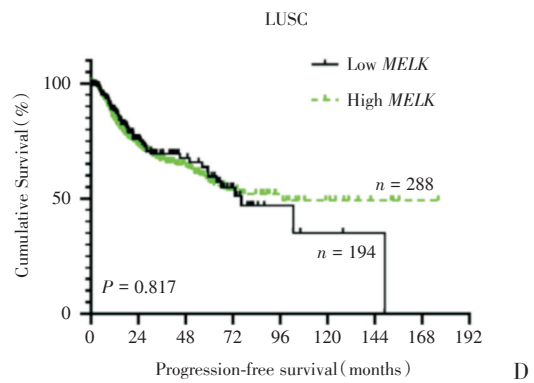
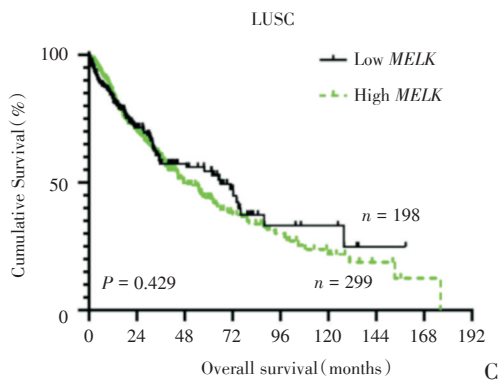
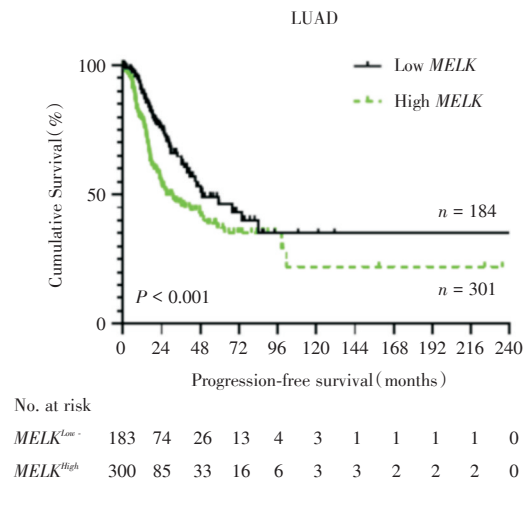
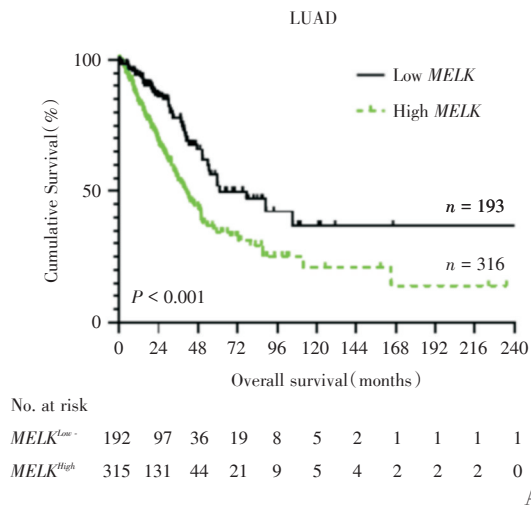
Analysis Type by Cancer	Cancer vs. Normal
Bladder Cancer	4
Brain and CNS Cancer	6 1
Breast Cancer	15 1
Cervical Cancer	4
Colorectal Cancer	12
Esophageal Cancer	2
Gastric Cancer	4
Head and Neck Cancer	5
Kidney Cancer	1
Leukemia	3 3
Liver Cancer	3
Lung Cancer	9
Lymphoma	6
Melanoma	1
Myeloma	1
Other Cancer	3 1
Ovarian Cancer	1
Pancreatic Cancer	4
Prostate Cancer	1
Sarcoma	11
Significant Unique Analyses	90 6
Total Unique Analyses	427

Legend: 1 5 10 10 5 1 %



A: In different tumors; B: From GEO database GSE19188 chip, in normal lung tissue, LUSC and LUAD

图4 Oncomine 网站中 MELK 基因表达概况
Fig.4 Overview of MELK gene expression on Oncomine



Overall Survival of MELK low/high expression group in LUAD (A) and LUSC (C); Progression-free survival of MELK low/high expression group in LUAD (B) and LUSC (D).

图5 肺腺癌和肺鳞癌中 MELK 表达与预后的相关性
Fig.5 Correlation of MELK expression with prognosis in LUAD and LUSC patients

表1 肺腺癌 *MELK* 表达量与临床病理参数的相关性Table 1 Relationship between the expression of *MELK* and the clinical characteristics of lung adenocarcinoma

Variables	n	<i>MELK</i> expression		χ^2	P
		Low	High		
Gender		n = 196	n = 326	6.316	0.012 ¹⁾
Female	280(53.6)	119(60.7)	161(49.4)		
Male	242(43.4)	77(39.3)	165(50.6)		
Age/year ⁴⁾		n = 196	n = 326	10.035	0.002 ²⁾
≤ 66	273(52.3)	85(43.4)	188(57.7)		
> 66	249(47.7)	111(56.6)	138(42.3)		
Smoking		n = 191	n = 317	9.285	0.002 ²⁾
No	75(14.8)	40(20.4)	35(11.0)		
Yes	433(85.2)	151(77.0)	282(89.0)		
cT		n = 195	n = 324	0.003	0.956
T1/T2	353(84.2)	170(87.2)	283(82.7)		
T3/T4	66(15.8)	25(9.2)	41(12.0)		
cN		n = 196	n = 326	12.718	0.005 ²⁾
N0	347(66.5)	147(75.0)	200(61.3)		
N1	98(18.8)	32(16.3)	66(20.2)		
N2	75(14.4)	17(8.7)	58(17.8)		
N3	2(0.4)	0	2(0.6)		
cM		n = 139	n = 232	1.876	0.171
M0	353(95.1)	135(97.1)	218(94.0)		
M1	18(4.9)	4(2.9)	14(6.0)		
cTNM Stage		n = 192	n = 321	11.650	0.009 ²⁾
I	279(54.4)	121(63.0)	158(49.2)		
II	123(24.0)	43(22.4)	80(24.9)		
III	85(16.6)	22(11.5)	63(19.6)		
IV	26(5.1)	6(3.1)	20(6.2)		
OS Status		n = 194	n = 319	15.267	< 0.001 ³⁾
Alive	329(64.1)	145(74.7)	184(57.7)		
Death	184(35.9)	49(25.3)	135(42.3)		
PFS Status		n = 196	n = 326	7.147	0.008 ²⁾
Negative	220(42.1)	68(34.7)	152(46.6)		
Positive	302(57.9)	128(68)	174(53.4)		

cT, clinical tumor stage; cN, clinical lymph nodes stage; cM, clinical distant metastasis stage; 1) $P < 0.05$, 2) $P < 0.01$, 3) $P < 0.001$, 4) Median age

3 讨论

激酶是一种催化磷酸基团从高能磷酸供体分子转移到特定底物的酶,底物获得一个磷酸基,而高能 ATP 分子提供一个磷酸基,这个过程被称为

磷酸化。蛋白激酶主要通过磷酸化丝氨酸、苏氨酸、酪氨酸或组氨酸残基,启动或沉默激酶与其他蛋白的相互作用,参与信号传导(信号级联放大)、调节细胞生命活动等过程^[7]。多种蛋白激酶参与肿瘤发生发展的各个阶段,并成为肿瘤治疗的重要靶点,如细胞周期蛋白依赖激酶家族(cyclin de-

表2 肺鳞癌 *MELK* 表达量与临床病理参数的相关性Table 2 Relationship between the expression of *MELK* and the clinical characteristics of lung squamous carcinoma

Variables	<i>n</i>	<i>MELK</i> expression		χ^2	<i>P</i>
		Low	High		
Gender		<i>n</i> = 201	<i>n</i> = 303	10.680	0.001 ²⁾
Female	131 (26.0)	68 (33.8)	63 (49.4)		
Male	373 (74.0)	133 (66.2)	240 (50.6)		
Age/year ⁴⁾		<i>n</i> = 201	<i>n</i> = 303	15.032	< 0.001 ³⁾
≤ 68	264 (52.4)	84 (41.8)	180 (59.4)		
> 68	240 (47.6)	117 (58.2)	123 (40.6)		
Smoking		<i>n</i> = 198	<i>n</i> = 294	1.821	0.177
No	18 (3.7)	10 (5.1)	8 (2.7)		
Yes	474 (96.3)	188 (94.9)	286 (97.3)		
cT		<i>n</i> = 201	<i>n</i> = 303	0.817	0.366
T1/T2	409 (81.1)	167 (83.0)	242 (79.9)		
T3/T4	95 (18.8)	34 (16.9)	61 (20.1)		
cN		<i>n</i> = 198	<i>n</i> = 300	10.695	0.013 ¹⁾
N0	320 (64.3)	143 (72.2)	177 (59.0)		
N1	98 (26.7)	44 (22.2)	89 (29.7)		
N2	75 (8.0)	9 (4.5)	31 (10.3)		
N3	1 (1.0)	2 (1.0)	3 (1.0)		
cM		<i>n</i> = 158	<i>n</i> = 261	0.676	0.411
M0	414 (98.8)	157 (99.4)	257 (98.5)		
M1	5 (1.2)	1 (0.6)	4 (1.5)		
TNM Stage		<i>n</i> = 200	<i>n</i> = 300	13.390	0.004 ²⁾
I	245 (49.0)	118 (59.0)	127 (42.3)		
II	163 (32.6)	53 (26.5)	110 (36.7)		
III	85 (17.0)	27 (13.5)	58 (19.3)		
IV	7 (1.4)	2 (1.0)	5 (1.7)		
OS Status		<i>n</i> = 198	<i>n</i> = 300	3.214	0.073
Alive	285 (57.2)	123 (62.1)	162 (54.0)		
Death	213 (42.8)	75 (37.9)	138 (46.0)		
PFS Status		<i>n</i> = 201	<i>n</i> = 303	0.269	0.604
Negative	152 (30.2)	58 (28.9)	94 (31.0)		
Positive	352 (69.8)	143 (71.1)	209 (69.0)		

cT, clinical tumor stage; cN, clinical lymph nodes stage; cM, clinical distant metastasis stage; 1) $P < 0.05$, 2) $P < 0.01$, 3) $P < 0.001$, 4) Median age

pendent kinases, CDKs)^[8]、丝裂原激活蛋白激酶类 (mitogen-activated protein kinases, MAPKs) 等^[9]。因此,探究激酶在恶性肿瘤中的功能、机制及临床意义对于其诊断治疗有着重要意义。

本文利用 GEO 芯片数据及 TCGA 数据库筛选出肺腺癌、肺鳞癌中的 21 个关键激酶蛋白:

TK1、*MELK*、*NEK2*、*CDK4*、*PLK1*、*BUB1*、*AURKB*、*PKMYT1*、*BUB1B*、*BORA*、*PBK*、*CKS2*、*CDKN3*、*CHEK2*、*CKS1B*、*CHEK1*、*PLK4*、*TTK*、*CDK1*、*SKP2* 和 *AURKA*。已有不少报道表明,这些激酶蛋白参与肿瘤发生发展、细胞调控、耐药等过程中,与多种临床指标密切相关,是多种肿瘤的预后因子。

表3 肺腺癌患者中影响OS预后因素的单因素和多因素Cox回归分析

Table 3 Univariate and Multivariate Cox regression analysis of prognostic factors affecting OS in LUAD patients

Variables	Overall survival			
	Univariate		Multivariate	
	HR(95% CI)	P	HR(95% CI)	P
Gender (Male vs. Female)	1.033(0.776, 1.376)	0.824		
Age/year(> 66 y vs. \geq 66 y)	1.072(0.803, 1.432)	0.638		
Smoking (Yes vs. No)	0.939(0.623, 1.415)	0.764		
cT (T3/T4 vs. T1/T2)	1.848(1.227, 2.785)	0.003 ²⁾	1.711(1.087, 2.693)	0.020 ¹⁾
cN (N3 vs. N2 vs. N1 vs. N0)	1.322(1.106, 1.579)	0.002 ²⁾		
cM (M1 vs. M0)	1.873(0.949, 3.697)	0.070		
TNM Stage (IV vs. III vs. II vs. I)	1.681(1.188, 2.379)	0.003 ¹⁾		
MELK level (High vs. Low)	1.735(1.273, 2.363)	< 0.001 ³⁾	1.538(1.118, 2.116)	0.008 ²⁾

HR, hazard ratio; CI, confidence interval; cT, clinical tumor stage; cN, clinical lymph nodes stage; cM, clinical distant metastasisstage; 1) $P < 0.05$, 2) $P < 0.01$, 3) $P < 0.001$

表4 肺腺癌患者中影响PFS预后因素的单因素和多因素Cox回归分析

Table 4 Univariate and Multivariate Cox regression analysis of prognostic factors affecting PFS in LUAD patients

Variables	Overall survival			
	Univariate		Multivariate	
	HR(95% CI)	P	HR(95% CI)	P
Gender (Male vs. Female)	1.063(0.794, 1.421)	0.683		
Age/year(> 66 y vs. \geq 66 y)	1.013(1.002, 1.024)	0.019 ¹⁾	1.014(1.001, 1.026)	0.028 ¹⁾
Smoking (Yes vs. No)	0.903(0.598, 1.363)	0.626		
cT (T3/T4 vs. T1/T2)	2.343(1.600, 3.432)	< 0.001 ³⁾	1.89(1.190, 3.000)	0.007 ²⁾
cN (N3 vs. N2 vs. N1 vs. N0)	1.691(1.425, 2.007)	< 0.001 ³⁾	1.469(1.070, 2.019)	0.018 ¹⁾
cM (M1 vs. M0)	2.188(1.277, 3.748)	0.004 ²⁾		
TNM Stage (IV vs. III vs. II vs. I)	2.624(1.922, 3.582)	< 0.001 ³⁾		
MELK level (High vs. Low)	2.005(1.445, 2.783)	< 0.001 ³⁾	1.885(1.266, 2.808)	0.002 ^{**}

HR, hazard ratio; CI, confidence interval; cT, clinical tumor stage; cN, clinical lymph nodes stage; cM, clinical distant metastasisstage; 1) $P < 0.05$, 2) $P < 0.01$, 3) $P < 0.001$

MELK 在多种肿瘤中高表达,且与不良预后显著相关,如黑色素瘤、结直肠癌、卵巢癌、乳腺癌、肝癌、前列腺癌等^[10-13]。在神经母细胞瘤中,*MELK*抑制剂OST167能抑制肿瘤体内外的增殖,且能抑制喜树碱或放射诱导后的复制叉折叠形成^[14]。在宫颈癌中,敲降*MELK*可能通过同源重组修复途径加重宫颈癌细胞的DNA损伤,且能促进P53、caspase3信号通路的激活^[15]。虽然已有文献表明*MELK*是肺腺癌不良预后的生物标志物,但是并没有具体分析*MELK*与各个临床指标的相关性,也没有分析各临床指标与预后的

关系^[16]。

本文利用Oncomine数据库,证实*MELK*在包括肺腺癌、肺鳞癌中的多种恶性肿瘤中高表达,同时肺鳞癌较肺腺癌表达更高。在LUAD与LUSC中,年龄较小、肿瘤TNM高分期的患者*MELK*表达较高;在LUAD中,男性、吸烟患者*MELK*表达高,而*MELK*高表达与OS、PFS较短显著相关。但是,在LUSC中,女性患者*MELK*表达较高,而*MELK*表达与OS、PFS均无相关性。在生存分析中,肺腺癌*MELK*高表达与不良预后(短OS、PFS)显著相关,但*MELK*表达量在肺鳞癌中与OS、PFS无显著

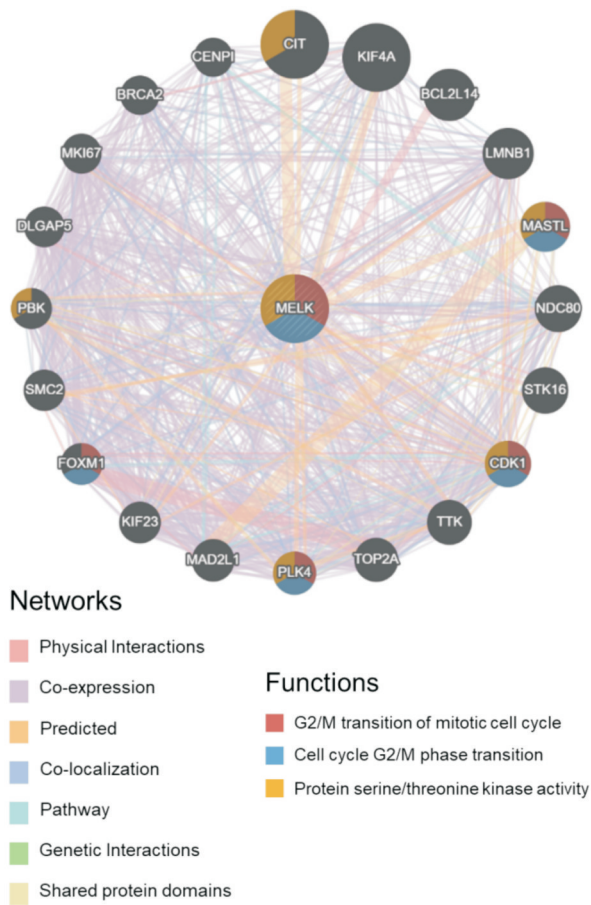


图6 MELK 相关信号网络及功能预测

Fig.6 MELK related signal network and function prediction

差别。在LUAD的多因素生存分析中提示,患者预后不良主要与高临床TNM分期、高表达 *MELK* 显著相关。这表明, *MELK* 是LUAD的独立预后因子, 这为肺腺癌的精准确治疗提供了可能的治疗靶点。虽然肺鳞癌 *MELK* 的 mRNA 表达水平相比正常肺组织与肺腺癌都显著增高, 但 *MELK* 与肺鳞癌的预后并没有显著相关性, 这提示 *MELK* 在肺腺癌和肺鳞癌中可能存在不同的功能和机制, 其在肺鳞癌中是否有其他临床意义尚需进一步探究。

MELK 对基底样乳腺癌、结直肠癌、黑色素瘤的增殖并不是必需的^[17-18], *MELK* 除了在肿瘤细胞中高表达之外, 同时也在正在分裂的细胞中高表达, 当细胞结束有丝分裂时, *MELK* 蛋白水平下降^[19]。使用 GeneMANIA 预测 *MELK* 功能, 提示 *MELK* 与 G2/M 期转化有关, 这说明 *MELK* 依赖于细胞分裂及细胞周期发挥功能。有研究证实, *MELK* 可以通过与 CDC25B 相互作用使其磷酸化, 激活 CDK1, 进而促进 G2/M 期转化^[6, 20]。在肺腺癌和肺鳞癌中发现 *MELK* 与细胞增殖相关基因 (*MKI67*、*PCNA*、*CCNB1*、*MCM2* 和 *TOP2A*) 共表达, 在肺腺癌中相关性较肺鳞癌中更强, 这表明, *MELK* 可能通过促进细胞有丝分裂、促进细胞周期 G2/M 期转化, 进而促进肺腺癌的发生发展及不良预后的产生。本文为 *MELK* 能成为肺腺癌的

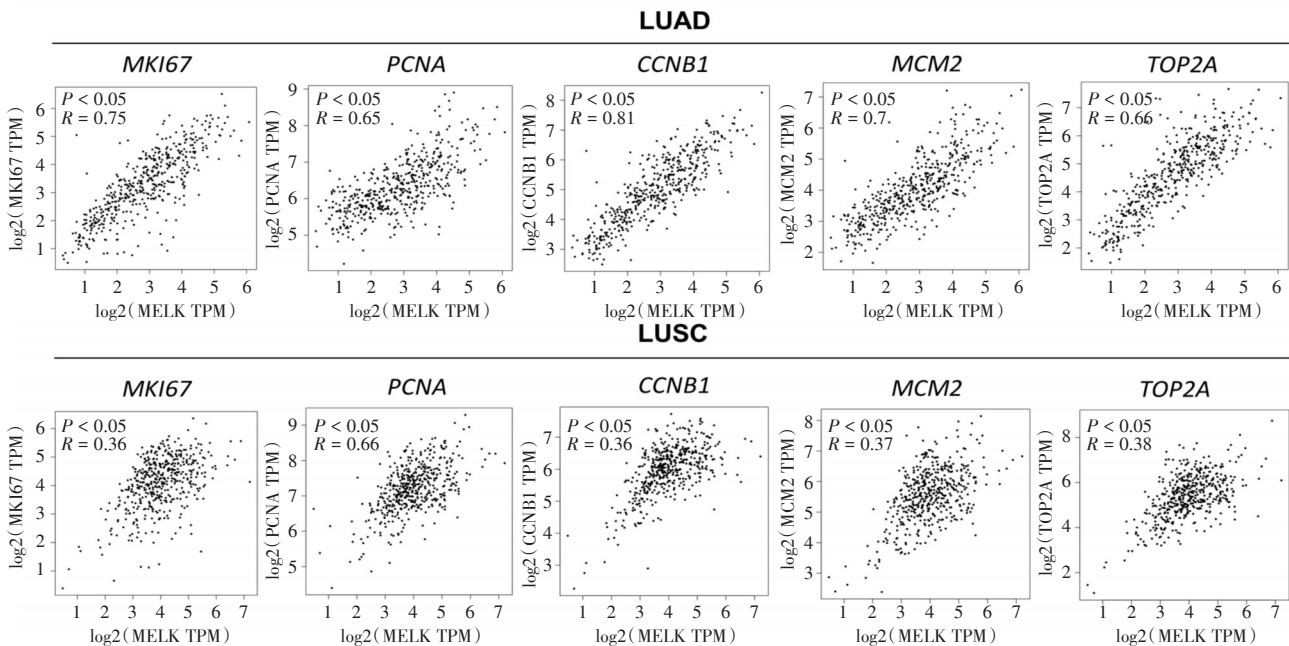


图7 MELK 与增殖相关基因的共表达分析

Fig.7 Co-expression analysis of MELK and proliferation-related genes

治疗靶点提供了理论依据,但是, *MELK* 如何在肺腺癌中发挥功能机制及如何使 *MELK* 成为其

治疗靶点,进而为精准治疗提供依据尚需进一步研究。

参考文献:

- [1] Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. *CA Cancer J Clin*, 2018, 68(6): 394-424.
- [2] De Groot PM, Wu CC, Carter BW, et al. The epidemiology of lung cancer [J]. *Transl Lung Cancer Res*, 2018, 7(3): 220-233.
- [3] Czubaty A, Piekiełko-Witkowska A. Protein kinases that phosphorylate splicing factors: Roles in cancer development, progression and possible therapeutic options [J]. *Int J Biochem Cell Biol*, 2017, 91 (Pt B): 102-115.
- [4] Roskoski R, Jr. Properties of FDA-approved small molecule protein kinase inhibitors [J]. *Pharmacol Res*, 2019, 144(2019): 19-50.
- [5] Arbour KC, Riely GJ. Systemic therapy for locally advanced and metastatic non-small cell lung cancer: a review [J]. *Jama*, 2019, 322(8): 764-774.
- [6] Jiang P, Zhang D. Maternal embryonic leucine zipper kinase (MELK): a novel regulator in cell cycle control, embryonic development, and cancer [J]. *Int J Mol Sci*, 2013, 14(11): 21551-21560.
- [7] Manning G, Whyte DB, Martinez R, et al. The protein kinase complement of the human genome [J]. *Science*, 2002, 298(5600): 1912-1934.
- [8] Roskoski R, Jr. Cyclin-dependent protein serine/threonine kinase inhibitors as anticancer drugs [J]. *Pharmacol Res*, 2019, 139 (January 2019): 471-488.
- [9] Najafi M, Ahmadi A, Mortezaee K. Extracellular-signal-regulated kinase/mitogen-activated protein kinase signaling as a target for cancer therapy: an updated review [J]. *Cell Biol Int*, 2019, 43(11): 1206-1222.
- [10] Gray D, Jubb AM, Hogue D, et al. Maternal embryonic leucine zipper kinase/murine protein serine-threonine kinase 38 is a promising therapeutic target for multiple cancers [J]. *Cancer Res*, 2005, 65 (21): 9751-9761.
- [11] Speers C, Zhao SG, Kothari V, et al. Maternal embryonic leucine zipper kinase (MELK) as a novel mediator and biomarker of radioresistance in human breast cancer [J]. *Clin Cancer Res*, 2016, 22(23): 5864-5875.
- [12] Hiwatashi K, Ueno S, Sakoda M, et al. Expression of maternal embryonic leucine zipper kinase (MELK) correlates to malignant potentials in hepatocellular carcinoma [J]. *Anticancer Res*, 2016, 36 (10): 5183-5188.
- [13] Kuner R, Fälth M, Pressinotti NC, et al. The maternal embryonic leucine zipper kinase (MELK) is up-regulated in high-grade prostate cancer [J]. *J Mol Med (Berl)*, 2013, 91(2): 237-248.
- [14] Chlenski A, Park C, Dobratic M, et al. Maternal embryonic leucine zipper kinase (MELK), a potential therapeutic target for neuroblastoma [J]. *Mol Cancer Ther*, 2019, 18(3): 507-516.
- [15] Wang J, Wang Y, Shen F, et al. Maternal embryonic leucine zipper kinase: A novel biomarker and a potential therapeutic target of cervical cancer [J]. *Cancer Med*, 2018, 7(11): 5665-7568.
- [16] Zang X, Qian C, Ruan Y, et al. Higher maternal embryonic leucine zipper kinase mRNA expression level is a poor prognostic factor in non-small-cell lung carcinoma patients [J]. *Biomark Med*, 2019, 13(16): 1349-1361.
- [17] Giuliano CJ, Lin A, Smith JC, et al. MELK expression correlates with tumor mitotic activity but is not required for cancer growth [J]. *Elife*, 2018, 7 (e3283), doi: 10.7554/eLife.32838.
- [18] Huang HT, Seo HS, Zhang T, et al. MELK is not necessary for the proliferation of basal-like breast cancer cells [J]. *Elife*, 2017, 6 (e26693), doi: 10.7554/eLife.26693.
- [19] Badouel C, Chartrain I, Blot J, et al. Maternal embryonic leucine zipper kinase is stabilized in mitosis by phosphorylation and is partially degraded upon mitotic exit [J]. *Exp Cell Res*, 2010, 316 (13): 2166-2173.
- [20] Davezac N, Baldin V, Blot J, et al. Human pEg3 kinase associates with and phosphorylates CDC25B phosphatase: a potential role for pEg3 in cell cycle regulation [J]. *Oncogene*, 2002, 21(50): 7630-7641.

(编辑 孙慧兰)