

## 基于网络靶点收敛算法预测治疗晚期肺腺癌的候选药物

刘溪, 关双, 俞成诚, 王忠

中国中医科学院中医临床基础医学研究所, 北京 100700

**【摘要】目的:**筛选晚期肺腺癌调控网络的收敛基因集,借助关联性图谱(CMap)预测治疗晚期肺腺癌的候选药物。**方法:**利用TCGA数据库检索肺腺癌转录组与临床数据,使用R4.0.3软件筛选肺腺癌早晚期患者的差异基因,使用Kaplan-Meier与log秩检验识别预后基因。DAVID和KEGG数据库对预后基因进行富集分析。依据背景网络构建差异预后基因调控网络,集体影响(CI)算法计算网络收敛基因集,将基因集导入CMap获得治疗晚期肺腺癌的候选药物,进一步对候选药物进行查找和分析。**结果:**共获得差异表达基因3409个,其中1981个与生存显著相关。富集分析结果显示,预后基因主要与细胞分裂、染色体分离、有丝分裂细胞周期、DNA复制、B细胞激活、T细胞激活等生物学过程相关;CI方法筛选得到晚期肺腺癌收敛预后基因96个,通过CMap连接图计算得到排名前20的候选化合物,其中,thapsigargin和nutlin-3通过文献验证对晚期肺腺癌有潜在的治疗作用。**结论:**借助生物信息学、网络靶点收敛算法和CMap数据库挖掘对晚期肺腺癌具有治疗作用的药物,为发现疾病的候选治疗靶点与药物开辟了新途径和思路。

**【关键词】**晚期肺腺癌;预后基因;复杂网络;靶点收敛;抗致癌药

**【中图分类号】**R318

**【文献标志码】**A

**【文章编号】**1005-202X(2024)04-0504-08

## Prediction of drug candidates for the treatment of advanced lung adenocarcinoma based on network target convergence algorithm

LIU Xi, GUAN Shuang, YU Chengcheng, WANG Zhong

Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, Beijing 100700, China

**Abstract: Objective** To identify the convergent gene sets in the regulatory network of advanced lung adenocarcinoma, and predict drug candidates for the treatment of advanced lung adenocarcinoma using connectivity map (CMap). **Methods** The TCGA database was used to search for the transcriptome and clinical data of lung adenocarcinoma, and R4.0.3 software to screen the differential genes of early- and advanced-stage patients, and Kaplan-Meier and log-rank tests to identify prognostic genes. The enrichment analysis of prognostic genes was carried out in DAVID and KEGG databases. The differential prognostic gene regulatory network was constructed based on the background network, and the collective influence algorithm was employed to calculate the convergent gene set which was then imported into CMap to obtain drug candidates for the treatment of advanced lung adenocarcinoma. Further investigation and analysis were conducted on the drug candidates. **Results** A total of 3 409 differentially expressed genes were obtained, with 1 981 genes significantly associated with survival. Enrichment analysis showed that prognostic genes were mainly related to biological processes such as cell division, chromosome segregation, mitotic cell cycle, DNA replication, B-cell activation, T-cell activation, etc. The collective influence method identified 96 convergent prognostic genes in advanced lung adenocarcinoma. The top 20 candidate compounds were obtained through CMap linkage map calculation, of which thapsigargin and nutlin-3 had been proven to have potential therapeutic effects on advanced lung adenocarcinoma through literature review. **Conclusion** The study leverages bioinformatics, network target convergence algorithm and CMap database to explore drugs with therapeutic effects on advanced lung adenocarcinoma, which opens up new ways and ideas for discovering candidate therapeutic targets and drugs for diseases.

**Keywords:** advanced lung adenocarcinoma; prognostic gene; complex network; target convergence; anticancer agent

**【收稿日期】**2024-01-16

**【基金项目】**国家科技重大专项-重大新药创制项目(2017ZX09301059)

**【作者简介】**刘溪, 博士生, 研究方向: 中药临床药理, E-mail: 736146265@qq.com

**【通信作者】**王忠, 博士, 研究员, 研究方向: 中药临床药理, E-mail: zhonw@vip.sina.com

前言

肺癌,包括小细胞肺癌(Small Cell Lung Cancer, SCLC)和非小细胞肺癌(Non-Small Cell Lung Cancer, NSCLC),是人类癌症死亡的主要原因<sup>[1-3]</sup>。肺腺癌是NSCLC的主要组织学亚型,在<45岁原发性肺癌病例中占比20%~30%,与吸烟史无关<sup>[4]</sup>。临床上,手术切除仍然是早期NSCLC(I-II期)患者最有效的治疗方法,干预后5年生存率为30%~60%<sup>[5]</sup>。然而,对于大多数NSCLC患者来说,由于诊断过晚,当肿瘤不能切除时,患者5年生存率大约下降到10%~15%。使用顺铂(DDP)与其他抗肿瘤剂(例如紫杉醇、吉西他滨、长春瑞滨等)仍然是晚期NSCLC的首选治疗方案。近年来,一些针对癌细胞特异性酪氨酸激酶的小分子药物的使用显示了良好的效果,但这种改善对于延长NSCLC患者的生命往往是微不足道的<sup>[6]</sup>。因此,需要寻找新的并且有效的肺腺癌治疗药。

将网络理论应用于疾病研究可有助于从分子层面认识疾病机理,为重大疾病的预防、诊疗研究提供了方法<sup>[7]</sup>。目前研究肺腺癌的范式侧重于构建疾病网络,通过研究和评估网络的拓扑结构及其多种中心度量,探索具有生物学意义的分子靶标<sup>[8-9]</sup>。然而,这些方法无法定位网络收敛的节点集合。最近, Morone等<sup>[10]</sup>介绍了一种优化方法来确定一组被称为集体影响者的收敛节点集合。旨在确定网络以最有效的方式分裂的最小节点数量。根据渗流理论,随机移除节点将导致网络以临界比例崩溃,其中巨大连接组件的概率为 $G=0$ 。最优渗流的目标是发现最小数量的影响者 $q$ ,以便获得 $G(q)=0$ <sup>[11]</sup>。

基因表达谱分析是阐明特定疾病分子机制、生物通路<sup>[12]</sup>以及预测药物反应或耐药性<sup>[13]</sup>、疾病预后<sup>[14]</sup>和发现新靶点<sup>[15]</sup>的有力工具。Lamb等<sup>[16]</sup>创建了一个可搜索数据库(关联性图谱, CMap),其中包含数千个暴露于大量小分子药物的各种培养癌细胞的基因表达特征。据推测,如果分子诱导的特征与CMap中发现的疾病诱导特征显著负相关,小药物分子可能会逆转疾病特征<sup>[17]</sup>。事实上,已经通过CMap成功确定了潜在的几种癌症的新疗法,包括结肠癌、肝细胞癌和肾细胞癌<sup>[18-20]</sup>。

本研究从TCGA数据库中收集了肺腺癌的转录组和临床数据,通过计算与预后相关的肺腺癌早期与晚期差异基因,基于背景的人类基因调控网络构建肺腺癌晚期基因调控网络,使用集体影响(CI)方法计算网络的收敛基因集合,通过CMap计算得到晚期肺腺癌相关的候选药物。最后利用文本挖掘的方法对这些预后差异基因和候选药物进行分析。

1 材料与方法

1.1 数据收集

使用癌症基因组图谱 TCGA 数据库 (<https://portal.gdc.cancer.gov/>) 下载肺腺癌转录组和临床数据进行整理,共得到496例肺腺癌患者的转录组数据和临床数据。根据临床病理分期分为386例早期(I、II阶段)患者与110例(III、IV阶段)晚期患者。

1.2 差异表达基因的筛选

使用R软件中的Limma软件包(版本3.40.6)对检索得到的基因表达数据集根据肺腺癌早期与晚期患者的不同分组进行基因表达差异分析,以“Adjusted  $P<0.05$ ”为标准阈值进行差异表达基因的筛选。差异分析结果通过生信 (<https://www.bioinformatics.com.cn/>) 在线网页工具对结果进行可视化处理。

1.3 肺腺癌生存预后分析

筛选预后相关基因,使用R软件中的survival软件包(版本3.40.6),对临床数据中的生存时间、生存状态和基因表达数据,利用单因素Cox回归方法评估了每个基因的预后显著性并筛选影响总生存期的风险基因。 $P<0.05$ 为差异有统计学意义。

1.4 基因的通路富集分析

基因本体(Gene Ontology, GO)富集分析是通过注释可视化和集成发现数据库DAVID (<https://david.ncifcrf.gov>) 工具进行的<sup>[21-22]</sup>。包括了基因的生物学过程(Biological Process, BP)、细胞组分(Cellular Component, CC)和分子功能(Molecular Function, MF)的注释。设定物种为人源,选取排名前10条信号通路,绘制热图。京都基因与基因组百科全书(Kyoto Encyclopedia of Genes and Genomes, KEGG)分析工具对于差异基因的功能和相关通路进行富集分析<sup>[23]</sup>。使用R软件的“ggplot2”包进行富集图形绘制。

1.5 基因调控网络的构建

通过将筛选出的与预后相关的差异表达基因集映射到人类基因调控网络上,构建出疾病调控网络。人类基因调控网络是通过整合Hu等<sup>[24]</sup>编制的“Reactome, KEGG, and NCI-Nature Pathway Interaction Database”的注释而构建的。

1.6 使用CI方法识别网络的收敛节点集合

CI算法的主要思想是去除每次迭代中CI值最大的节点,快速崩溃网络。每个节点的CI值可以通过式(1)计算:

$$CI_i = (d_i - 1) \sum_{j \in \delta B(i, I)} (d_j - 1)$$

(1)

其中,  $d_i$ 、 $d_j$  分别表示节点  $i, j$  的度。  $\delta B(i, l)$  是到节点  $i$  的最短路的长度为  $l$  的节点组成的球边界。 Morone 等<sup>[25]</sup> 已经证明, 较大的  $l$  将提供更好的结果, 但将导致更差的计算性能。 一般设置  $l=1$ 。

移除过程将在大多数巨型组件消失后停止, 即由式(2)计算的  $\lambda$  值决定, 当  $\lambda \leq 1$  时停止移除<sup>[26]</sup>。

$$\lambda(l)=\left(\frac{\sum_i CI_i}{N\bar{d}}\right)^{\frac{1}{l+1}}$$

(2)

1.7 CMap 分析

CMap(build02, <http://www.broad.mit.edu/CMap/>) 包含 7 000 多个表达特征, 代表 1 309 种药物。 将计算得到的网络收敛集合中上调和下调基因组同时提交给 CMap 进行分析。 使用基因组富集分析算法计算了数据库中每个药物的富集得分<sup>[16]</sup>。 具有负连接性

得分的药物被记录为肺腺癌的潜在治疗药物, 负连接性得分意味着匹配的药物具有逆转肺腺癌中查询基因表达方向的作用模式。 将候选化合物输入 DrugBank 数据库(<https://www Drugbank.ca/>) 和 stitch 数据库(<http://stitch.E embl.de/>) 进行查找验证。

2 结果

2.1 肺腺癌预后差异基因的鉴定

从癌症基因组图谱(The Cancer Genome Atlas, TCGA) 共获得了 386 个早期与 110 个晚期肺腺癌患者 RNAseq 数据和相应的临床信息。 经计算获得 3 409 个差异基因, 用火山图及热图分析表示(图 1)。 基于差异基因表达量和患者的生存时间(Overall Survival, OS) 进行单因素 Cox 生存分析。 确定了 1 981 个与患者生存结果相关的差异基因。

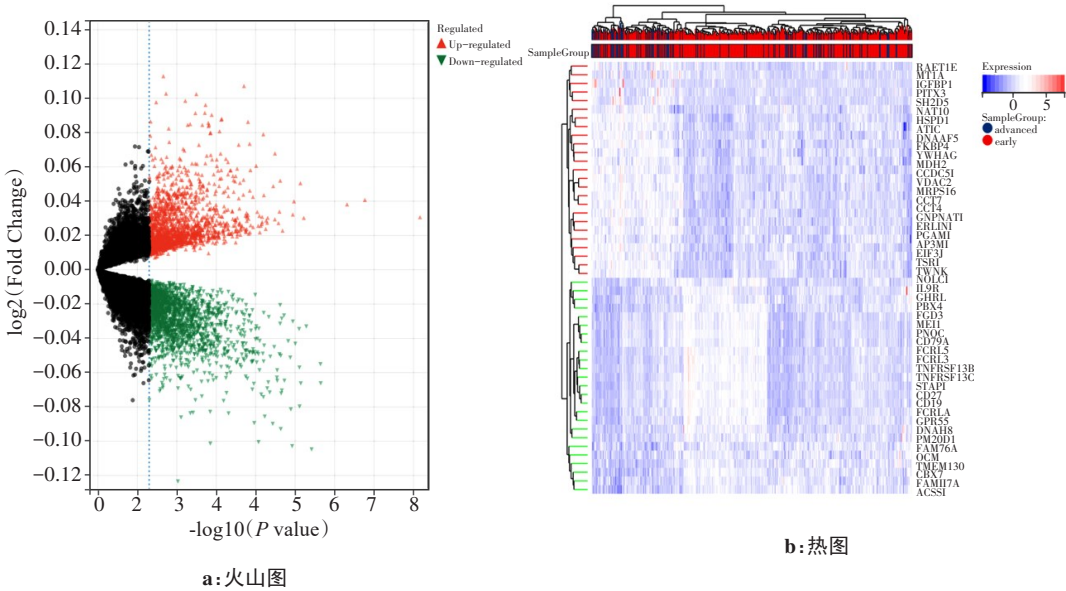


图 1 肺腺癌早期和晚期患者差异表达基因的火山图及热图分析  
Figure 1 Volcano and heat map analysis of differentially expressed genes between early- and advanced-stage lung adenocarcinoma

2.2 预后差异基因的潜在功能分析

使用 DAVID 数据库对共同差异基因进行 GO 富集分析后, 共获得 344 个富集结果。 其中包括 BP 190 项, 主要涉及细胞分裂、染色体分离、有丝分裂细胞周期、DNA 复制、细胞周期、B 细胞激活、T 细胞激活等。 MF 93 项, 主要涉及胞液、着丝粒、细胞质、核、核质等。 CC 61 项, 主要涉及蛋白质结合、ATP 结合、ATP 酶活性、跨膜信号受体活性、受损的 DNA 结合、RNA 结合等。 KEGG 富集出 24 条通路, 其中富集靶点较多的主要相关信号通路为: 爱泼斯坦-巴尔病毒(EBV)感染、人类 T 细胞白血病病毒感染、细胞周期信号通路、造血细胞谱系、细胞粘附分子、吞噬体、

Th17 细胞分化、Th1 和 Th2 细胞分化、用于 IgA 产生的肠道免疫网络、抗原加工和呈递(图 2)。

2.3 预后差异基因调控网络的构建与收敛基因集合的识别

将 1 981 个预后差异基因映射到人类调控疾病网络中, 最终获得一个 446 个节点, 1 867 条边的肺腺癌预后差异基因调控网络(图 3), 网络的具体参数见表 1。 采用 CI 算法进行网络节点收敛集合的计算, 最终得到 96 个基因集合, 包括 56 个上调基因和 40 个下调基因, 具体的基因见表 1。

2.4 CMap 数据筛选晚期肺腺癌候选治疗药物

将计算的晚期肺腺癌收敛基因集中的差异表达



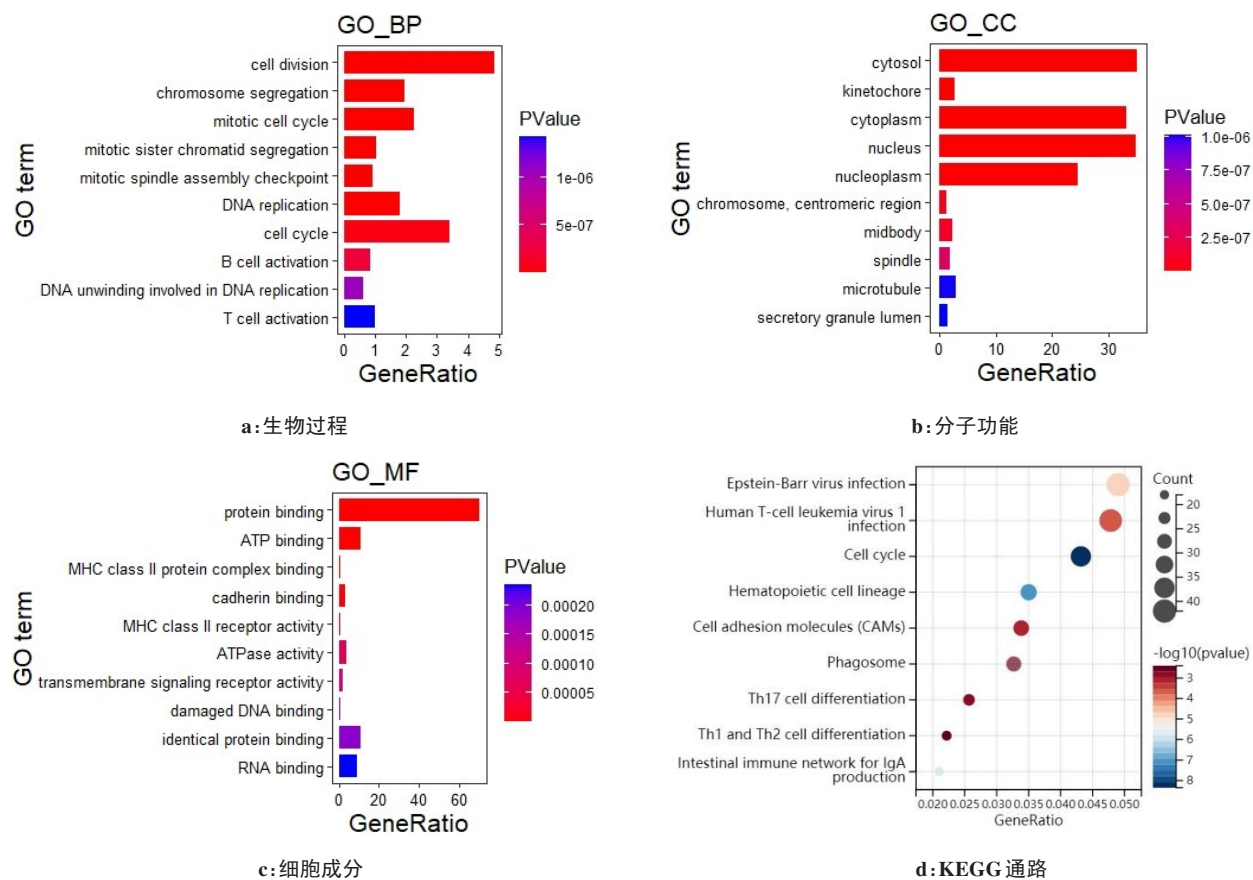


图2 肺腺癌预后差异基因的GO和KEGG通路富集分析

Figure 2 GO and KEGG pathway enrichment analysis of prognostic genes in lung adenocarcinoma

上调和下调基因导入CMap数据库,筛选出关联度最大的20种负相关小分子药物,见表2。其中的thapsigargin(毒胡萝卜素)和nutlin-3与晚期肺腺癌的治疗相关。由stitch数据库获得与thapsigargin和nutlin-3相关蛋白,如图4所示。HSPA5、XBP1、CASP12、CASP3、ATP2A1、CALR、ATP2A2、FREM2、EIF2AK3、FOS蛋白与thapsigargin相关;Mdm2、Tp53、Lgals3、Abcb1a、Foxm1、H2afx、Ndn、Bcl2a1、Tcl1a、Ceacam1蛋白与nutlin-3相关。

### 3 讨论

近年来,高通量测序技术的发展为理解表观遗传变化和破译肿瘤发生和发展中的关键遗传变化带来了希望,同时生物信息学及复杂网络分析在发现疾病预后和诊断生物标志物以及治疗靶点中发挥着重要作用<sup>[27-28]</sup>,这种技术已被广泛用于识别与肺腺癌的诊断、预后和治疗相关的潜在新生物标志物<sup>[29-30]</sup>。本研究通过整合数据库中的转录组数据和临床数据,确定早期和晚期肺腺癌组织中的差异基因,筛选1981个与生存相关的预后基因,依靠背景调控网络构建包含446个节点,1867条边的肺腺癌预后差异

基因调控网络,用CI算法计算得到包含96个基因的集合,CMap对这些收敛的差异表达基因进行计算得到与晚期肺腺癌相关的候选化合物。其中thapsigargin(负相关性最强)和nutlin-3(出现次数最多)可能对晚期肺腺癌有潜在的治疗作用。

肺腺癌晚期收敛基因集即肺腺癌患者晚期的疾病靶点数量存在上限,当超过这个上限,即使靶点个数增加,但靶点的影响力趋于饱和。其中周期蛋白依赖性激酶(Cyclin-Dependent Kinase, CDK),是蛋白激酶家族重要成员之一,CDK1对细胞周期的调控、检查点的激活、DNA损伤修复发挥着关键性的作用。故凭借信号通路的重要位点,CDK1抑制剂的开发受到了广泛的关注,尤其是其临床研究更多指向胰腺癌和胶质瘤,使得研究者对该靶点潜力充满憧憬<sup>[31]</sup>。研究表明淋巴细胞特异性蛋白酪氨酸激酶(Lymphocyte Specific Protein Tyrosine Kinase, Lck)是一种在T细胞和自然杀伤细胞中表达的Src家族细胞质酪氨酸激酶,在其他细胞中的异位表达可以诱发细胞发生癌变,因此抑制Lck还能起到抗癌作用<sup>[32]</sup>。蛋白酪氨酸磷酸酶,非受体型蛋白酪氨酸磷酸酶11(Protein Tyrosine Phosphatase Nonreceptor 11,

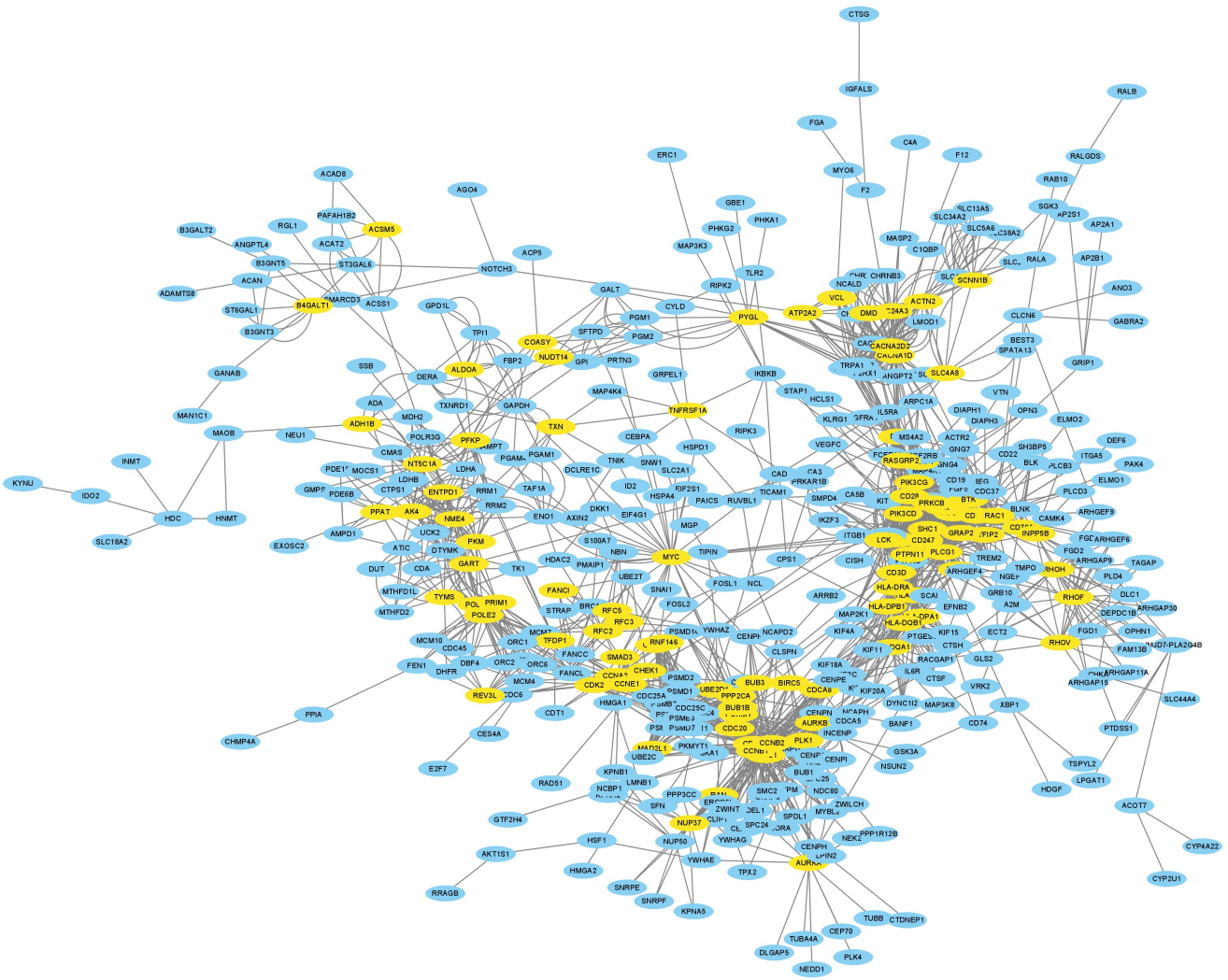


图 3 肺腺癌预后差异基因调控网络(黄色节点为收敛基因集合)

Figure 3 Differentially expressed prognostic gene regulatory network in lung adenocarcinoma (yellow nodes are the set of convergent genes)

表 1 网络收敛基因集合

Table 1 Convergent gene sets

类型	基因
上调	CCNB1、PLK1、RHOV、GART、BUB3、FOXM1、B4GALT1、CHEK1、CCNB2、BUB1B、POLD2、PXN、RAN、RFC、AURKB、PKM、FGF5、TNFR、SF1A、RFC3、FANCI、CDK1、RHOF、AK4、CCNA2、SHC1、NUP37、CDCA8、PYGL、ATP2A2、VCL、MYC、TFDP1、CDC20、MAD2L1、BIRC5、ALDOA、PRIM1、PTPN11、PFKP、CCNE1、SMAD3、TYMS、NCKAP1、RAC1、ESPL1、RFC5、NME4、CDK2、UBE2D1、COASY、TXN、PPP2CA、CKS1B、POLE2、PPAT、AURKA
下调	CD79A、CD79B、GRAP2、ZAP70、RAS、GRP2、ACSM5、CD247、RHOH、BTK、ITK、INPP5J、PRKCB、LCK、RNF146、HLA-DPB1、SLC24A3、PIK3CG、CACNA1D、ENTPD1、INPP5B、PTPN6、SLC4A8、HLA-DRB1、CD3D、HLA-DQB1、PLCG1、NT5C1A、PIK3CD、ADH1B、CYFIP2、SCNN1B、HLA-DPA1、HLA-DRA、RASGRF1、CACNA2D2、HLA-DQA1、REV3L、ACTN2、GALT

PTPN11)是一种多功能酪氨酸磷酸酶,在许多类型的肿瘤中具有重要作用<sup>[33]</sup>。但是目前这些基因在晚期肺腺癌中的作用尚不清楚,需要进一步研究。

晚期肺腺癌预后基因与富集生物学过程中的细胞分裂以及细胞周期等密切相关。KEGG分析显示

晚期肺腺癌的预后基因主要涉及EBV感染、人类T细胞白血病病毒感染、细胞周期、造血细胞谱系、细胞粘附分子、吞噬体、Th17细胞分化、Th1和Th2细胞分化等信号通路。研究表明EBV感染可能与肺癌相关,因为与健康对照个体相比,肺癌患者中观察到更

表 2 CMap 数据库筛选出的关联度最大的 20 种负相关小分子药物

Table 2 Twenty most negatively correlated small molecule drugs screened for association by CMap database

排序	药物名称	剂量/ $\mu\text{mol}\cdot\text{L}^{-1}$	细胞	评分	作用机理
1	thapsigargin	0.04	PC3	-0.850 7	ATPase inhibitor
2	Ro-4987655	2.22	MCF10A	-0.850 5	MEK inhibitor
3	MGCD-265	10	HUVEC	-0.842 0	VEGFR inhibitor
4	sunitinib	10	HT29	-0.839 9	FLT3 inhibitor KIT inhibitor PDGFR inhibitor RET inhibitor VEGFR inhibitor
5	C-646	10	HUVEC	-0.839 8	Histone acetyltransferase inhibitor
6	NNC-55-0396	10	HCC515	-0.839 6	Calcium channel blocker
7	rociletinib	10	A375	-0.838 9	EGFR inhibitor
8	mepacrine	10	MCF10A	-0.835 9	NFKB inhibitor Cytokine production inhibitor p53 activator
9	nutlin-3	10	A549	-0.835 8	MDM inhibitor
10	DL-PDMP	60	HT29	-0.832 8	Glucosyltransferase inhibitor
11	PD-184352	10	M0A	-0.831 3	MEK inhibitor
12	AZD-9291	10	A549	-0.831 1	EGFR inhibitor
13	fingolimod	10	HCC515	-0.830 7	Immunosuppressant Sphingosine 1 phosphate receptor agonist
14	PCI-29732	10	TMD8	-0.830 5	BTK inhibitor
15	nutlin-3	10	HEPG2	-0.829 0	MDM inhibitor
16	BMS-754807	10	HT29	-0.826 5	IGF-1 inhibitor
17	palbociclib	2.22	THP1	-0.824 7	CDK inhibitor
18	etoposide	2.22	HUVEC	-0.824 2	Topoisomerase inhibitor
19	ellagic-acid	10	HCT116	-0.823 1	Glutathione transferase inhibitor Aldose reductase inhibitor Beta secretase inhibitor HCV inhibitor HIV inhibitor Quorum sensing signaling modulator Reverse transcriptase inhibitor
20	refametinib	2.22	MCF10A	-0.823 1	MEK inhibitor

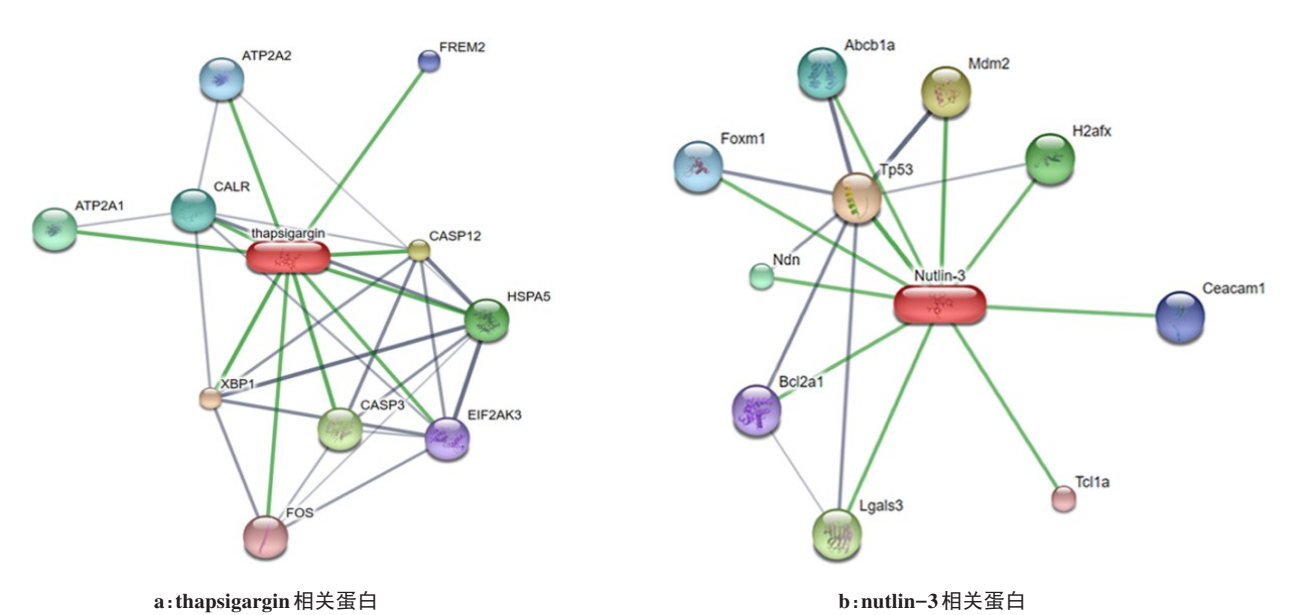


图4 化合物与相关蛋白

Figure 4 Compounds and related proteins

高的EBV血清阳性<sup>[34]</sup>。从肺癌患者收集的支气管肺泡液中检测到EBV,这表明肺组织可能充当潜在的EBV储存库<sup>[35]</sup>。多项研究也报告了EBV在肺鳞状细胞癌和肺腺癌中的存在<sup>[36]</sup>。进一步表明治疗肺腺癌



药物可能通过多靶点、多通路作用于肺腺癌。

通过比较表达特征,CMap可作为搜索所有被研究药物新适应症的方法,并在药物再发现方面取得了成功。有学者利用CMap发现雷帕霉素是一种潜在的糖皮质激素抗性逆转剂<sup>[37]</sup>。利用这种方法发现了两种新的hsp90抑制剂 celastrol 和 gedunin<sup>[38]</sup>。基于CMap预测潜在的治疗晚期肺腺癌小分子药物,这些小分子药物的分子机制大多为ATP酶抑制剂、MEK抑制剂、VEGFR抑制剂、钙通道阻滞剂、EGFR抑制剂、MDM抑制剂、葡萄糖基转移酶抑制剂等。其中thapsigargin是一种天然产物,对多种肿瘤细胞系具有潜在的抗癌活性<sup>[39]</sup>。研究表明其通过损害人肺腺癌细胞的细胞骨架动力学来诱导凋亡从而发挥抗肿瘤功效<sup>[40]</sup>。Stitch蛋白质结果表明thapsigargin通过参与内质网应激的反应,内质网中蛋白质加工响应,内质网应激的内在凋亡信号通路来发挥作用。nutlin-3被确定为抑制p53-Mdm2相互作用的第一个有效和特异性的小分子Mdm2拮抗剂<sup>[41]</sup>,研究表明同时引起p53激活和NF- $\kappa$ B抑制,nutlin-3具有双靶向抗癌的潜力<sup>[42]</sup>。Stitch蛋白质结果表明nutlin-3通过参与微小核糖核酸的转录失调和p53、PI3K-Akt信号通路等来发挥作用。这些小分子扰动因素可能对晚期肺腺癌患者靶基因高表达的患者具有有利的治疗效果。这将为肺腺癌的个体治疗提供有效的治疗方案。

本研究表明,基于基因表达特征的药物发现对于识别现有药物的新适应症具有潜在价值,这对于转化研究和临床应用至关重要。这种方法的一个主要优势是上市时间比新药开发短得多,成本节约也很重要,因为CMap中检测的许多药物都得到了食品和药品监督管理局的批准。来自这种平台的任何有希望的药物都可能对没有得到有效治疗的患者有益。

## 4 结论

综上,在本研究中,通过综合生物信息学和网络分析,对基因转录组数据进行了筛选、整合、挖掘和分析,探索与晚期肺腺癌发生发展相关的差异表达基因集合、信号通路,候选化合物和作用机制等,从多个角度分析晚期肺腺癌的表达特征。揭示了与肺腺癌预后相关的关键信号通路。thapsigargin和nutlin-3有望成为晚期肺腺癌治疗的候选药物。这些结果为进一步生物学验证和研究提供了依据,也为今后肺腺癌的治疗研究提供了新方向。本研究挖掘到关于肺腺癌的基因和与治疗相关的候选化合物,但也存在部分局限,例如未对候选化合物进行试验验证和临床研究,后续可做更深入的研究。

## 【参考文献】

- [1] Leiter A, Veluswamy RR, Wisnivesky JP. The global burden of lung cancer: current status and future trends[J]. *Nat Rev Clin Oncol*, 2023, 20(9): 624-639.
- [2] Feng RM, Su QL, Huang XY, et al. Cancer situation in China: what does the China cancer map indicate from the first national death survey to the latest cancer registration?[J]. *Cancer Commun (Lond)*, 2023, 43(1): 75-86.
- [3] Wu CC, Li MN, Meng HB, et al. Analysis of status and countermeasures of cancer incidence and mortality in China[J]. *Sci China Life Sci*, 2019, 62(5): 640-647.
- [4] Chen PX, Liu YH, Wen YK, et al. Non-small cell lung cancer in China[J]. *Cancer Commun (Lond)*, 2022, 42(10): 937-970.
- [5] Galetta D. Advances in lung cancer therapy[J]. *Cancers (Basel)*, 2023, 15(10): 2671.
- [6] Piotrowska Z, Hata AN. Resistance to first-line osimertinib in EGFR-mutant NSCLC: tissue is the issue[J]. *Clin Cancer Res*, 2020, 26(11): 2441-2443.
- [7] 王欢. 复杂网络理论在疾病研究中的应用[J]. *高教学刊*, 2015(20): 36-38.  
Wang H. Application of complex network theory in disease research[J]. *Journal of Higher Education*, 2015(20): 36-38.
- [8] Wu Q, Zhang B, Sun YD, et al. Identification of novel biomarkers and candidate small molecule drugs in non-small-cell lung cancer by integrated microarray analysis[J]. *Oncotargets Ther*, 2019, 12: 3545-3563.
- [9] Mosharaf MP, Reza MS, Gov E, et al. Disclosing potential key genes, therapeutic targets and agents for non-small cell lung cancer: evidence from integrative bioinformatics analysis[J]. *Vaccines (Basel)*, 2022, 10(5): 771.
- [10] Morone F, Makse HA. Influence maximization in complex networks through optimal percolation[J]. *Nature*, 2015, 524(7563): 65-68.
- [11] Bollobás B, Riordan O. Percolation[M]. Cambridge: Cambridge University Press, 2006.
- [12] Alam MS, Sultana A, Wang GH, et al. Gene expression profile analysis to discover molecular signatures for early diagnosis and therapies of triple-negative breast cancer[J]. *Front Mol Biosci*, 2022, 9: 1049741.
- [13] Tang HM, Cheung PC. Gene expression profile analysis of gallic acid-induced cell death process[J]. *Sci Rep*, 2021, 11(1): 16743.
- [14] Yousefi H, Bahramy A, Zafari N, et al. Notch signaling pathway: a comprehensive prognostic and gene expression profile analysis in breast cancer[J]. *BMC Cancer*, 2022, 22(1): 1282.
- [15] Pan X, Liu JH. Identification of four key biomarkers and small molecule drugs in nasopharyngeal carcinoma by weighted gene co-expression network analysis[J]. *Bioengineered*, 2021, 12(1): 3647-3661.
- [16] Lamb J, Crawford ED, Peck D, et al. The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease[J]. *Science*, 2006, 313(5795): 1929-1935.
- [17] Lan MY, Chen CL, Lin KT, et al. From NPC therapeutic target identification to potential treatment strategy[J]. *Mol Cancer Ther*, 2010, 9(9): 2511-2523.
- [18] Liu C, Liu DW, Wang FF, et al. Construction of a novel choline metabolism-related signature to predict prognosis, immune landscape, and chemotherapy response in colon adenocarcinoma[J]. *Front Immunol*, 2022, 13: 1038927.
- [19] Han CF, Chen JR, Huang J, et al. Single-cell transcriptome analysis reveals the metabolic changes and the prognostic value of malignant hepatocyte subpopulations and predict new therapeutic agents for hepatocellular carcinoma[J]. *Front Oncol*, 2023, 13: 1104262.
- [20] Bai SH, Wu YY, Yan YL, et al. Construct a circRNA/miRNA/mRNA regulatory network to explore potential pathogenesis and therapy options of clear cell renal cell carcinoma[J]. *Sci Rep*, 2020, 10(1): 13659.
- [21] Dennis GJ, Sherman BT, Hosack DA, et al. DAVID: database for annotation, visualization, and integrated discovery[J]. *Genome Biol*, 2003, 4(5): P3.
- [22] Gene Ontology Consortium. The gene ontology (GO) project in 2006[J]. *Nucleic Acids Res*, 2006, 34(S1): D322-D326.
- [23] Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes[J]. *Nucleic Acids Res*, 2000, 28(1): 27-30.

- [24] Hu YX, Chen CH, Ding YY, et al. Optimal control nodes in disease-perturbed networks as targets for combination therapy [J]. Nat Commun, 2019, 10(1): 2180.
- [25] Morone F, Min B, Bo L, et al. Collective influence algorithm to find influencers *via* optimal percolation in massively large social media[J]. Sci Rep, 2016, 6: 30062.
- [26] Zhang HW, Zhang H, Wu C. Identification of essential proteins based on centrality methods using improved collective influence algorithm [C]//2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP). Piscataway, NJ, USA: IEEE, 2019: 1-5.
- [27] Zhang B, Wu Q, Xu R, et al. The promising novel biomarkers and candidate small molecule drugs in lower-grade glioma: evidence from bioinformatics analysis of high-throughput data[J]. J Cell Biochem, 2019, 120(9): 15106-15118.
- [28] Shen HY, Wang ZH, Ren SQ, et al. Prognostic biomarker MITD1 and its correlation with immune infiltrates in hepatocellular carcinoma (HCC)[J]. Int Immunopharmacol, 2020, 81: 106222.
- [29] Ling B, Liao XJ, Huang YH, et al. Identification of prognostic markers of lung cancer through bioinformatics analysis and *in vitro* experiments [J]. Int J Oncol, 2020, 56(1): 193-205.
- [30] Cuccu A, Francescangeli F, De Angelis ML, et al. Analysis of dormancy-associated transcriptional networks reveals a shared quiescence signature in lung and colorectal cancer[J]. Int J Mol Sci, 2022, 23(17): 9869.
- [31] Wang QS, Bode AM, Zhang TS. Targeting CDK1 in cancer: mechanisms and implications[J]. NPJ Precis Oncol, 2023, 7(1): 58.
- [32] Gocho Y, Liu JJ, Hu JZ, et al. Network-based systems pharmacology reveals heterogeneity in LCK and BCL2 signaling and therapeutic sensitivity of T-cell acute lymphoblastic leukemia[J]. Nat Cancer, 2021, 2(3): 284-299.
- [33] Cao YP, Duan HX, Su AL, et al. A pan-cancer analysis confirms PTPN11's potential as a prognostic and immunological biomarker[J]. Aging (Albany NY), 2022, 14(13): 5590-5610.
- [34] Roy A, Dey S, Chatterjee R. Prevalence of serum IgG and IgM antibodies against Epstein-Barr virus capsid antigen in Indian patients with respiratory tract carcinomas[J]. Neoplasma, 1994, 41(1): 29-33.
- [35] Lung ML, Lam WK, So SY, et al. Evidence that respiratory tract is major reservoir for Epstein-Barr virus [J]. Lancet, 1985, 1(8434): 889-892.
- [36] Kheir F, Zhao MM, Strong MJ, et al. Detection of Epstein-Barr virus infection in non-small cell lung cancer[J]. Cancers (Basel), 2019, 11(6): 759.
- [37] Wei G, Twomey D, Lamb J, et al. Gene expression-based chemical genomics identifies rapamycin as a modulator of MCL1 and glucocorticoid resistance[J]. Cancer Cell, 2006, 10(4): 331-342.
- [38] Hieronymus H, Lamb J, Ross KN, et al. Gene expression signature-based chemical genomic prediction identifies a novel class of HSP90 pathway modulators[J]. Cancer Cell, 2006, 10(4): 321-330.
- [39] Wu LL, Huang XM, Kuang YQ, et al. Thapsigargin induces apoptosis in adrenocortical carcinoma by activating endoplasmic reticulum stress and the JNK signaling pathway: an *in vitro* and *in vivo* study[J]. Drug Des Devel Ther, 2019, 13: 2787-2798.
- [40] Wang F, Liu DZ, Xu H, et al. Thapsigargin induces apoptosis by impairing cytoskeleton dynamics in human lung adenocarcinoma cells [J]. Scientific World J 2014, 2014: 619050.
- [41] Vassilev LT, Vu BT, Graves B, et al. *In vivo* activation of the p53 pathway by small-molecule antagonists of MDM2[J]. Science, 2004, 303(5659): 844-848.
- [42] Dey A, Wong ET, Bist P, et al. Nutlin-3 inhibits the NFkappaB pathway in a p53-dependent manner: implications in lung cancer therapy[J]. Cell Cycle, 2007, 6(17): 2178-2185.

(编辑:薛泽玲)