·基础研究·

可变剪接在肺腺癌预后中的价值*

张世超[1]，欧阳燕[1]，兰应武[1]，唐福州[1]，胡祖权[1]，曾柱[2**]
(1. 贵州医科大学 生物与工程学院，贵州 贵阳 550025；2. 贵州医科大学 基础医学院，贵州 贵阳 550025)

[摘 要] **目的**：探讨可变剪接在肺腺癌预后中的作用。**方法**：在 TCGA 数据库中收集 474 例肺腺癌组织的 RNA-seq 数据，绘制可变剪接图谱；通过单因素 *Cox* 回归分析确定肺腺癌中存活相关的剪接事件，再利用最小绝对收缩模型筛选用于监测存活的预后因子；收集 5 例肺腺癌患者的癌组织，采用 PCR 扩增及琼脂糖凝胶电泳检测 *CHEK*1、*KIF*23、*MCM*7 和 *FN*1 的 mRNA 亚型条带。**结果**：在肺腺癌中存在着大量的可变剪接事件并与预后相关，包括 *CHEK*1、*KIF*23、*MCM*7 和 *FN*1，通过生物信息学利用可变剪接事件构建了高效的风险预测模型；KEGG 途径分析表明，主要富集的基因与“ “regulation of autophagy” 、“central carbon metabolism in cancer” 和“cell cycle”有关；5 例肺腺癌患者的癌组织 PCR 扩增结果显示，可变剪接核心基因 *CHEK*1、*KIF*23、*FN*1 确实存在 CHEK1-AP-19309、KIF23-AP-31390、FN1-ES-57398 等剪接体。**结论**：可变剪接事件可作为预测肺腺癌患者临床结果的指标，可为肺腺癌患者制定个体化治疗方案提供依据。

[关键词] 肺肿瘤；预后；肺腺癌；可变剪接；临床结果；核心基因
[中图分类号] R734. 2；R34 [文献标识码] A [文章编号] 1000-2707(2019)11-1254-08
DOI：10. 19367/j. cnki. 1000-2707. 2019. 11. 003

The Role of Alternative Splicing in the Prognosis of Lung Adenocarcinoma

ZHANG Shichao[1], OUYANG Yan[1], LAN Yingwu[1], TANG Fuzhou[1], HU Zuquan[1], ZENG Zhu[2]
(*1. School of Biology and Engineering, Guizhou Medical University, Guiyang 550025, Guizhou, China;*
2. School of Basic Medical Science, Guizhou Medical University, Guiyang 550025, Guizhou, China)

[**Abstract**] **Objective**: To explore the relationship between alternative splicing and prognosis of lung adenocarcinoma. **Methods**: Integrative analysis of RNA sequencing and clinical information were utilized to identify survival associated splicing events in lung adenocarcinoma. Then, several prognosis-related splicing events were submitted to develop moderate predictors for survival monitoring by using least absolute shrinkage and selection operator model. **Results**: There are a large number of alternative splicing events in lung adenocarcinoma, which are related to prognosis, including *CHEK*1, *KIF*23, *MCM*7 and *FN*1. We constructed an efficient risk prediction model using alternative splicing events. KEGG pathway analysis showed that the major enriched genes were "regulation of autophagy", "carbon metabolism in cancer" and "cell cycle". **Conclusion**: Alternative splicing events could be used as a prognostic indicator for predicting the clinical outcomes of patients with lung adenocarcinoma, providing individualized treatments with important clinical significance for patients with lung adenocarcinoma.

[**Key words**] lung neoplasms; prognosis; lung adenocarcinoma; alternative splicing; clinical outcome; core genes

*[基金项目]国家自然科学基金项目(31771014，11762006，31660258)；贵州省科技计划项目(2019－1275，2019－2787)；贵州医科大学博士启动基金(YJ2017－10)
**通信作者 E-mail：zengzhu100@ sina. com
网络出版时间：2019－11－21 网络出版地址：http://kns. cnki. net/kcms/detail/52. 1164. R. 20191120. 2125. 003. html

mRNA 可变剪接是指特定的剪切外显子被选择性去除或保留的过程[1]。*pre*-mRNA 的选择性加工是产生 mRNA 亚型的重要机制,是蛋白多样性的主要来源[1-2],同时还是调节 mRNA 翻译效率的主要途径之一[3]。因此,选择性剪接是必不可少的生物学过程,剪接模式的改变与蛋白质功能密切相关。基因组功能研究结果显示,剪接缺陷和特定类型的剪接加工是癌症的驱动因素[4]。肺癌是常见的恶性肿瘤,也是导致死亡的主要癌症,主要包括肺腺癌(lung adenocarcinoma,LUAD) 和肺鳞状细胞癌(lung squamous cell carcinoma,LUSC)[5-6]。肺腺癌属于非小细胞癌,与肺鳞状细胞癌不同,多发生于女性及不吸烟者,基因组研究表明,肺腺癌基因的拷贝数、DNA 甲基化和基因表达均与其正常细胞不同[7];而且,癌症特异性的 mRNA 可变剪接在肺腺癌与其正常细胞中也呈现明显的差异[7-8],但目前肺腺癌中基于可变剪接事件的系统生存分析尚缺乏。以 mRNA 可变剪接事件构建癌症患者的预后模型,对为患者提供个体化治疗具有重要的临床意义。因此,本研究基于肿瘤基因组图谱(the cancer genome atlas, TCGA)数据库,系统分析肺腺癌患者的生存率与 mRNA 可变剪接之间的关系,构建预后模型和预后相关可变剪接基因的网络互作图,并预测、分析预后相关可变剪接基因的信号通路,现将结果报道如下。

1　材料与方法

1.1　数据获取

从 TCGA 数据库(https://tcga data. nci. nih. gov/tcga/)下载 LUAD 队列中 474 例肺腺癌组织的 RNA-seq 数据,并使用 Java 应用程序 SpliceSeq v3. 2 拼接工具分析样本转录剪接模式,得到 7 种类型可变剪接事件的"选择剪接百分比值"(percent spliced in,*PSI*),7 种可变剪接事件包括 Mutually Exclusive Exons(ME)、Retained Intron(RI)、Alternate Donor site(AD)、Alternate Acceptor site(AA)、Alternate Terminator(AT)、Alternate Promoter(AP) 和 Exon Skip(ES)[9-11],同时从 TCGA 的泛癌图谱数据库中提取患者相应的临床数据,包括年龄、性别和总生存期。

1.2　可变剪接事件检测

收集 5 例肺腺癌患者的癌组织,提取其 RNA 进行反转录。进一步通过 PCR 扩增及琼脂糖凝胶电泳,检测 *CHEK*1、*KIF*23、*MCM*7 和 *FN*1 的 mRNA 亚型条带。

1.3　统计学分析

数据分析采用 R/Bioconductor。单因素 *Cox* 回归分析评价可变剪接事件 *PSI* 值与肺腺癌患者总生存期之间的关系并构建预后模型,"survival ROC"绘制受试者工作特征曲线(receiver operating characteristic curve,*ROC*),*ROC* 曲线下面积(area under ROC curve,*AUC*)评价预后模型的有效性;随后绘制 *Kaplane-Meier* 曲线,并用 *Log-Rank* 检验分析高风险组和低风险组间的差异。利用 cytoscape 的 reactome fi 插件构建基因网络图,探索与预后相关的可变剪接核心基因;利用 Kyoto Encyclopedia of Genes and Genomes (KEGG) 和 Gene Ontology(GO) 对预后相关的可变剪接基因进行功能注释和信号通路分析。$P<0.05$ 差异有统计学意义。

2　结果

2.1　总体可变剪接事件

474 名患者的 13 684 个基因中检测到 43 948 个 mRNA 可变剪接事件,其中 6 324 个基因中检测到 16 793 个 ES 事件,3 816 个基因中检测到 8 992 个 AP 事件,3 901 个基因中检测到 8 546 个 AT 事件,3 291 个基因中检测到 3 559 个 AA 事件,2 889 个基因中检测到 3 057 个 AD 事件,2 677 个基因中检测到 2 781 个 RI 事件,216 个基因中检测到 220 个 ME 事件。见图 1。

2.2　预后相关的可变剪接事件

通过单因素 *Cox* 回归分析发现,共有 1 908 个预后相关的可变剪接事件,其中包括 430 个 ES 事件, 704 个 AP 事件,432 个 AT 事件,104 个 AA 事件,126 个 AD 事件,104 个 RI 事件,8 个 ME 事件;UpSet plot 图表明一个基因可以有多达几种与预后相关的可变剪接事件。见图 2。

2.3　肺腺癌患者的预后信号

通过预后相关的可变剪接事件构建预后模型,结果表明可变剪接事件可作为预后因子能高效地预测肺腺癌患者的临床结果,见图 3;*ROC* 曲线也证实了该预后模型具有较高的可靠性,*AUC* 值为 0. 748,见图 4;每位患者基于可变剪接事件得到的风险值显示,患者风险值越高,存活率越低,存活期越短,同时可见用于构建预后模型的可变剪接事件在每位患者中的 *PSI* 值,见图 5。

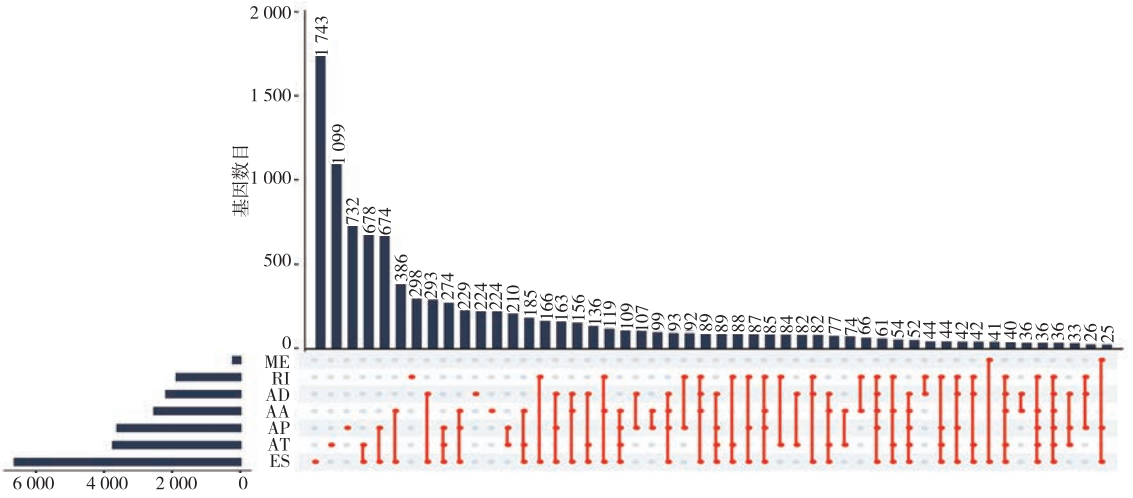


图 1　肺腺癌中总体可变剪接事件的 UpSet 结果

Fig. 1　UpSet plots in lung adenocarcinoma showing the interactions among the seven types of alternative splicing events

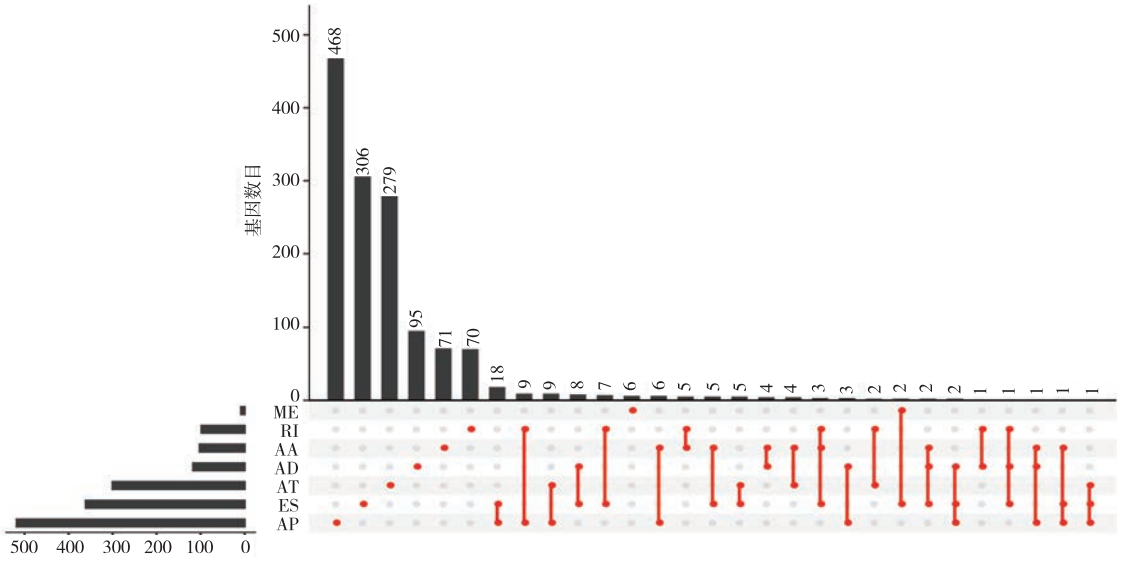


图 2　肺腺癌中预后相关的可变剪接事件的 UpSet 结果

Fig. 2　UpSet plots of prognosis – related alternative splicing events in lung adenocarcinoma

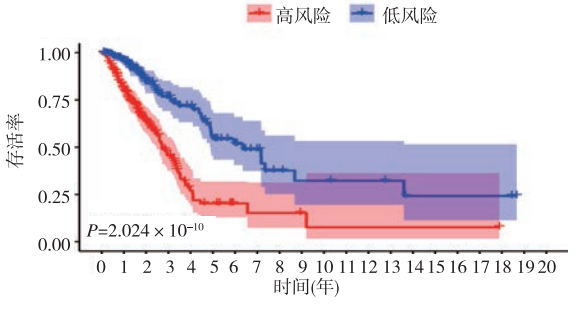


图 3　可变剪接预测肺腺癌预后的 *Kaplan-Meier* 曲线

Fig. 3　*Kaplan-Meier* curves of prognostic predictors for patients with lung adenocarcinoma

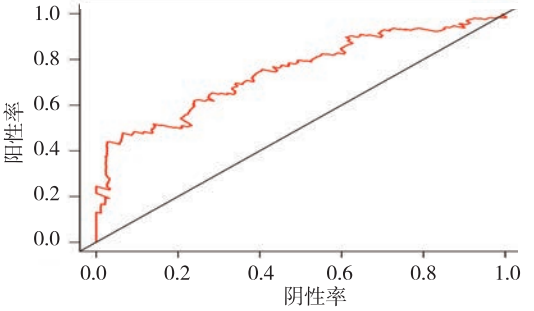
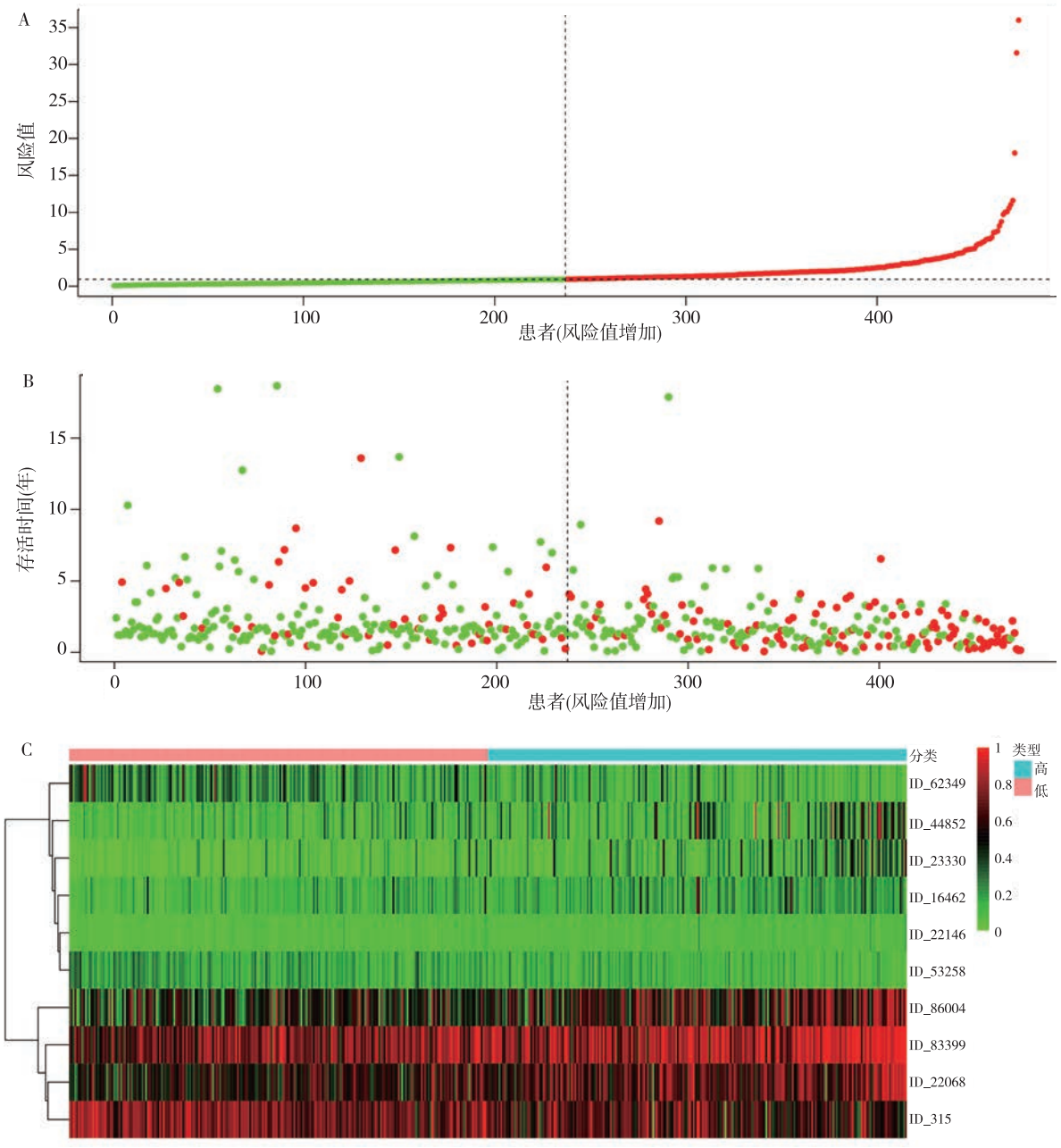


图 4　肺腺癌预后模型的 *ROC* 曲线

Fig. 4　*ROC* curves of prognostic predictors for patients with lung adenocarcinoma



注：A 为 476 名患者通过风险值分组，分为低风险（绿色点）和高风险组（红色点）；B 为肺腺癌患者的生存时间，左侧和右侧分别显示低风险和高风险患者，绿点表示存活患者，红点表示死亡患者；C 为用于构建预后模型的可变剪接事件的 *PSI* 值。

图 5　预后模型对肺腺癌患者预后的识别

Fig. 5　The recognition capability of prognostic signature for dividing patients with lung adenocarcinoma into low-and high-risk groups

2. 4　预后相关可变剪接事件的分子特征

基因功能注释表明“cell-cell adhesion”、“macroautophagy”和“regulation of transcription from RNA polymerase II promoter”是 3 个最有效的“biological process”，“nucleoplasm”、“cytosol”和“cytoplasm”是 3 个最有效的“cellular component”，“protein binding”、“cadherin binding involved in cell-cell adhesion”和“kinase activity”是 3 个最重要的“molecular

function”,见表 1、2、3;KEGG 途径分析表明,主要富集的基因与“ “regulation of autophagy”,“central carbon metabolism in cancer”和“cell cycle”有关,见表 4;基因网络互作图结果表明 *CHEK*1、*KIF*23、*MCM*7 和 *FN*1 基因位于互作图的核心位点,见图 6。

表 1　预后相关可变剪接基因功能注释(生物学过程)

Tab. 1　Gene ontology analysis of genes with survival-associated alternative splicing events: biological process

ID	生物学过程	*P*	*n*
GO:0098609	cell-cell adhesion	4.57×10^{-5}	39
GO:0016236	macroautophagy	2.57×10^{-4}	16
GO:0006357	regulation of transcription from RNA polymerase II promoter	4.09×10^{-4}	52
GO:0010719	negative regulation of epithelial to mesenchymal transition	5.99×10^{-4}	8
GO:0007062	sister chromatid cohesion	9.21×10^{-4}	18
GO:0018107	peptidyl-threonine phosphorylation	1.13×10^{-3}	10
GO:0043161	proteasome-mediated ubiquitin-dependent protein catabolic process	1.23×10^{-3}	28

表 2　预后相关可变剪接基因功能注释(细胞成分)

Tab. 2　Gene ontology analysis of genes with survival-associated alternative splicing events: cellular component

ID	细胞成分	*P*	*n*
GO:0005654	nucleoplasm	6.24×10^{-18}	307
GO:0005829	cytosol	2.22×10^{-16}	346
GO:0005737	cytoplasm	2.87×10^{-15}	491
GO:0005634	nucleus	9.75×10^{-11}	482
GO:0005925	focal adhesion	2.20×10^{-7}	57
GO:0016020	membrane	2.91×10^{-6}	208
GO:0005815	microtubule organizing center	5.79×10^{-5}	26
GO:0005913	cell-cell adherens junction	1.48×10^{-4}	42

表 3　预后相关可变剪接基因功能注释(分子功能)

Tab. 3　Gene ontology analysis of genes with survival-associated alternative splicing events: molecular function

ID	分子功能	*P*	*n*
GO:0005515	protein binding	3.46×10^{-24}	800
GO:0098641	cadherin binding involved in cell-cell adhesion	1.16×10^{-4}	40
GO:0016301	kinase activity	1.23×10^{-4}	35
GO:0044822	poly(A) RNA binding	6.29×10^{-4}	111
GO:0008134	transcription factor binding	6.54×10^{-4}	37
GO:0003779	actin binding	8.74×10^{-4}	36
GO:0005524	ATP binding	8.78×10^{-4}	140
GO:0003700	transcription factor activity, sequence-specific DNA binding	8.99×10^{-4}	96

表 4　预后相关可变剪接基因 KEGG 信号通路分析

Tab. 4　KEGG pathway analysis of genes with survival-associated alternative splicing events

Hsa	通路	*P*	*n*
04140	regulation of autophagy	1.38×10^{-3}	8
05230	central carbon metabolism in canceradhesion	5.04×10^{-3}	12
04110	cell cycle	6.57×10^{-3}	18
04512	ECM-receptor interaction	1.97×10^{-2}	13
00051	fructose and mannose metabolism	2.34×10^{-2}	7
04510	focal adhesion	3.57×10^{-3}	23

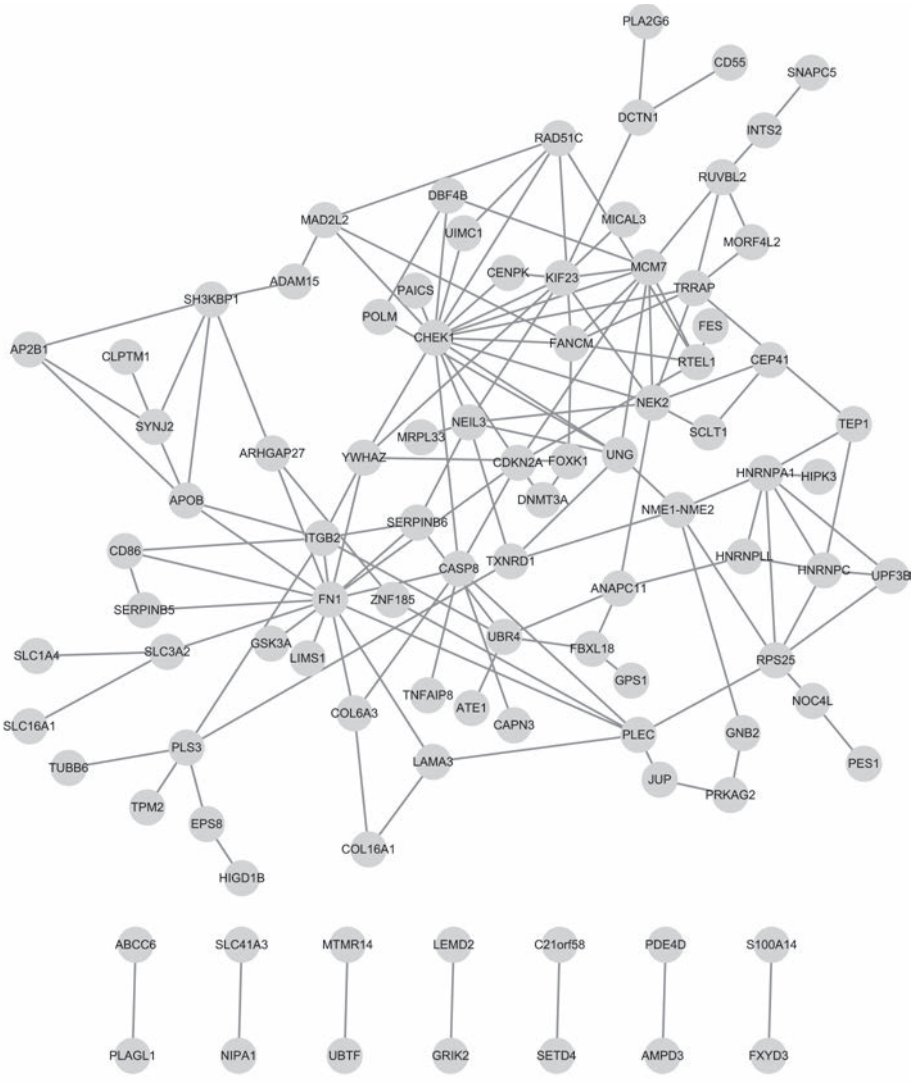


图 6　预后相关基因的网络互作

Fig. 6　Prognosis-associated gene networks

2.5　肺腺癌组织验证

5 例肺腺癌患者的癌组织 PCR 扩增结果显示，可变剪接核心基因 *CHEK*1、*KIF*23、*FN*1 确实存在 CHEK1-AP-19309、KIF23-AP-31390、FN1-ES-57398 等剪接体。

3　讨论

前体 mRNA 的可变剪接作为基因修饰的转录后过程产生了大量的 mRNA，形成蛋白质组和转录组的多样性。特定的剪接亚型是肿瘤发生和发展的驱动因子，其与肿瘤的发生、增殖、转移、生存和耐药性有关[1-3]。然而，由于肺腺癌细胞分子功能的复杂性和多样性以及缺乏可获得的大样本公共选择性剪接图谱和对选择性剪接的系统分析，选择性剪接事件在肺腺癌中的作用仍有许多未解之谜。

本文采用多种生物信息学计算方法，结合可变剪接事件和肺腺癌患者的临床结果，对选择性剪接事件进行综合分析。根据预后相关的剪接事件和存活时间，将肺腺癌患者分为好和坏的预后的两个亚组，本研究筛选了与预后相关的选择性剪接的核心基因 *CHEK*1、*KIF*23、*MCM*7 及 *FN*1，也分析了预后相关可变剪接基因的信号通路，如“regulation of autophagy”，“central carbon metabolism in cancer”和“cell cycle”；通过单因素回归分析在 474 例肺腺癌患者得到 1 908 个预后相关的可变剪接事件，整合 7 种可变剪接事件的最终模型呈现较高的可靠效率值（*AUC* 值为 0.748），并且能显著区分好的或坏

的临床结果,表明可变剪接可作为预测肺腺癌患者临床预后的指标。预后相关基因的基因网络互作图表明 *CHEK*1、*KIF*23、*MCM*7 及 *FN*1 微管靶向药物对非小细胞肺癌的疗效[12],*KIF*23 是治疗肺癌的潜在分子靶点[13],*MCM*7 可作为非小细胞肺癌的预后标志物[14],*FN*1 与肺癌的免疫治疗效果相关[15],因而这些核心基因的可变剪接可能在肺腺癌患者的预后起着重要的作用;进一步 KEGG 对预后相关的可变剪接基因进行信号通路分析,呈现 3 个重要的信号通路,如:"regulation of autophagy","central carbon metabolism in cancer"和"cell cycle"。先前的研究也报道了这些信号通路与非小细胞癌的关联,如 *lncrna-blacat*1 通过调节自噬参与非小细胞肺癌细胞的耐药性[16-20], *Microrna*-21 通过 ampk/ulk1 信号通路调节自噬活性从而促进非小细胞肺癌 a549 细胞的增殖、迁移和侵袭[21-24],*FASN* 通过调节糖代谢和 AKT/ERK 通路抑制非小细胞肺癌的恶化[25-27],*NME*4 调节细胞周期阻滞从而促进非小细胞肺癌的增殖和扩散[28-30]。这些信号通路也可能与肺腺癌患者的预后相关。

综上所述,本研究系统描述了肺腺癌中的可变剪接事件,结果显示可变剪接与肺腺癌患者临床预后显著相关,而且可变剪接可作为肺腺癌患者的临床预后指标,对为肺腺癌患者提供个体化治疗具有重要的临床意义。

4 参考文献

[1] GRIFFITH M, GRIFFITH O L, MWENIFUMBOM J, et al. Alternative expression analysis by RNA sequencing [J]. Nature Methods, 2010, 7(10):843 - 847.

[2] LU M, ZHAN X. The crucial role of multiomic approach in cancer research and clinically relevant outcomes [J]. The EPMA Journal, 2017, 9(1):77 - 102.

[3] ÁNGELES A, CONESA A. Single-cell RNAseq for the study of isoforms-how is that possible[J]. Genome Biology, 2016,19(1):110 - 115.

[4] 从美丽, 周蓓, 吕晓敏, 等. 大鼠酒精性肝损伤模型 RNA 可变剪接变化分析[J]. 热带医学杂志, 2018, 18(8):41 - 44.

[5] LI Y, SUN N, LU Z, et al. Prognostic alternative mRNA splicing signature in non-small cell lung cancer [J]. Cancer Letters, 2017, 393:40 - 51.

[6] CHEN J, HU L, ZHANG F, et al. Downregulation of carbonic anhydrase IV contributes to promotion of cell proliferation and is associated with poor prognosis in non-small cell lung cancer [J]. Oncology Letters, 2017, 14(4):5046 - 5050.

[7] WABG E T, SANDBERG R , LUO S, et al. Alternative isoform regulation in human tissue transcriptomes [J]. Nature, 2008, 456(7221):470 - 476.

[8] ZHENG Z, WEI X, HILDEBRANDT A, et al. A computational method for studying the relation between alternative splicing and DNA methylation [J]. Nucleic Acids Research, 2016, 44(2):19 - 19.

[9] TOMCZAK K, CZERWISKA P, WIZNEROWICZ M. The cancer genome atlas (TCGA): an immeasurable source of knowledge [J]. Contemporary Oncology, 2015, 19(1A):68 - 77.

[10] WU Z, PENG Z, HE R, et al. Prognostic index of aberrant mRNA splicing profiling acts as a predictive indicator for hepatocellular carcinoma based on TCGA SpliceSeq data [J]. International Journal of Oncology, 2019, 55(2):425 - 438.

[11] LIN P, HE R, HUANG Z, et al. Role of global aberrant alternative splicing events in papillary thyroid cancer prognosis [J]. Aging, 2017,11(7):2082 - 2097.

[12] YU X, ZHANG Y, MA X, et al. miR-195 Potentiates the efficacy of Microtubule-targeting agents in Non-small cell lung cancer [J]. Cancer Letters, 2018, 427(28): 85 - 93.

[13] VIKBERG A L, VOODER T, LOKK K, et al. Mutation analysis and copy number alterations of KIF23 in non-small-cell lung cancer exhibiting KIF23 over-expression [J]. Onco Targets and Therapy, 2017,10:4969 - 4979.

[14] LIU Y Z, WANG B S, JIANG Y Y, et al. MCMs expression in lung cancer: implication of prognostic significance [J]. Journal of Cancer, 2017, 8(18):3641 - 3647.

[15] HARAO M, MITTENDORF E A, RADVANYI L G. Peptide-based vaccination and induction of CD^{8+} T-cell responses against tumor antigens in breast cancer [J]. Biodrugs, 2015, 29(1):15 - 30.

[16] NIU Y, FENG M, HUANG W, et al. Long non-coding RNA TUG1 is involved in cell growth and chemoresistance of small cell lung cancer by regulating LIMK2b via EZH2 [J]. Molecular Cancer, 2017, 16(1):5 - 10.

[17] SUC C, HRYDZIUSZKO O, LEE D, et al. Integration of somatic mutation, expression and functional data reveals potential driver genes predictive of breast cancer survival [J]. Bioinformatics, 2015, 31(16):2607 - 2613.

[18] MC R, CLELAND J, KIM R, et al. SpliceSeq: a resource for analysis and visualization of RNA-Seq data on

alternative splicing and its functional impacts [J]. Bioinformatics, 2012, 28: 2385 – 2387.

[19] SOAVE D M, STRUG L J. Testing calibration of cox survival models at extremes of event risk [J]. Frontiers in Genetics, 2018, 9:177 – 182.

[20] FU Y T, ZHANG D Q, ZHOU L, et al. Has-MiR-196a-2 is up-regulated and acts as an independent unfavorable prognostic factor in thyroid carcinoma [J]. European Review for Medical & Pharmacological Sciences, 2018, 22(9):2707 – 2712.

[21] QIU X, MAO Q, TANG Y, et al. Reversed graph embedding resolves complex single-cell trajectories [J]. Nature Methods, 2017, 14(10):979 – 982.

[22] NEELAMRAJU Y, GONZALEZ P A, BHAT N P, et al. Mutational landscape of RNA-binding proteins in human cancers [J]. RNA Biology, 2018, 15(1):115 – 129.

[23] KAMTCHUENG C, STEBENNE M, DELANNOY A, et al. Alternative splicing of TAF6: downstream transcriptome impacts and upstream RNA splice control elements [J] Plos One, 2014, 9(7):e102399.

[24] CHANG L, FANG S, CHEN Y, et al. Inhibition of FASN suppresses the malignant biological behavior of non-small cell lung cancer cells via deregulating glucose metabolism and AKT/ERK pathway [J]. Lipids in Health and Disease, 2019, 18(1):118 – 123.

[25] GUO W, FRENTEET P S. Alternative CD44 splicing in intestinal stem cells and tumorigenesis [J]. Oncogene, 2014, 33: 537 – 538.

[26] MIGUEL F D, PAJARES M J, MARTINEZ-TERROBA E, et al. A large-scale analysis of alternative splicing reveals a key role of QKI in lung cancer [J]. Molecular Oncology, 2016, 10(9):1437 – 1449.

[27] PAIK P K, DRILIN A, FAN P D, et al. Response to MET inhibitors in patients with stage IV lung adenocarcinomas harboring MET mutations causing exon 14 skipping [J]. Cancer Discovery, 2015, 5(8): 842 – 849.

[28] ZHU D Y, LI X N, QI Y, et al. MiR-454 promotes the progression of human non-small cell lung cancer and directly targets PTEN [J]. Biomedicine & Pharmacotherapy, 2016, 81:79 – 85.

[29] MENG L, XU Y, XU C, et al. Biomarker discovery to improve prediction of breast cancer survival: using gene expression profiling, meta-analysis, and tissue validation [J]. Oncotargets & Therapy, 2016, 9:6177 – 6185.

[30] MARTINEZ-MONTIEL N, ROSAS-MURRITEA N H, MARTINEZ-MONTIEL M, et al. Microbial and natural metabolites that inhibit splicing: a powerful alternative for cancer treatment [J]. Biomed Research International, 2016, 2016(23):1 – 14.

（2019-09-02 收稿，2019-10-17 修回）

中文编辑：严　征；英文编辑：张启芳