·肿瘤专栏·

基于 TCGA 数据库分析 m6A 调节因子甲基化修饰对肺腺癌预后的影响

吴启飞，张 冬*
（内蒙古科技大学包头医学院第一附属医院呼吸及危重症科，内蒙古 包头 014010）

［摘要］ **目的** 通过生物信息学分析探究 m6A 甲基化调节因子在肺腺癌中的表达及预后。**方法** 通过 TCGA 数据库下载 516 例肺腺癌患者的转录组数据及其临床数据，使用 R 软件分析比较肺腺癌组织和癌旁组织中 20 个 m6A 甲基化调节因子的表达差异，并通过单变量 Cox 回归分析进行生存分析。通过 Consensus Cluster Plus R 非监督类的方法进行 m6A 聚类生存分析。**结果** 16 个 m6A 甲基化调节因子在肺腺癌中差异表达，其中有 5 个 m6A 甲基化调节因子（IGF2BP1、IGF2BP3、HNRNPC、RBMX、YTHDC2）与肿瘤分期、7 个 m6A 甲基化调节因子（IGF2BP1、IGF2BP3、HNRNPC、RBMX、METTL3、YTHDF2、METTL14）与 T 期和 2 个 m6A 甲基化调节因子（IGF2BP3、YTHDC2）与 N 期显著相关。Cox 回归分析结果表明 6 个 m6A 甲基化调节因子（IGF2BP1、IGF2BP2、IGF2BP3、HNRNPA2B1、HNRNPC、RBM15）是影响肺腺癌患者预后的独立危险因素。**结论** m6A 甲基化调节因子与肺腺癌进展有关，6 个 m6A 甲基化调节因子可以作为预测肺腺癌患者预后的潜在生物标志物。

［关键词］ 肺腺癌；m6A 甲基化；TCGA 数据库 doi:10.3969/j.issn.1007-3205.2024.03.004
［中图分类号］ R734.2 ［文献标志码］ A ［文章编号］ 1007-3205(2024)03-0271-07

Analysis of the effect of m6A methylation regulatory factors on the prognosis of lung adenocarcinoma based on TCGA database

WU Qi-fei, ZHANG Dong*
(*Department of Respiratory and Critical Care, the First Affiliated Hospital of Baotou Medical College, Inner Mongolia University of Science and Technology, Baotou* 014040, *China*)

［**Abstract**］ **Objective** To investigate the expression and prognosis of m6A methylation regulatory factors in lung adenocarcinoma by bioinformatics analysis. **Methods** The transcriptome data and clinical data of 516 cases of lung adenocarcinoma were downloaded from TCGA database. The expression differences of 20 m6A methylation regulatory factors in lung adenocarcinoma tissues and paracancerous tissues were compared using R software analysis, and survival analysis was performed by univariate Cox regression analysis. Consensus Cluster Plus R method was used to perform m6A cluster survival analysis. **Results** Sixteen m6A methylation regulatory factors were differentially expressed in lung adenocarcinoma, of which 5 m6A methylation regulatory factors (IGF2BP1, IGF2BP3, HNRNPC, RBMX, YTHDC2) were significantly associated with tumor stage, 7 m6A methylation regulatory factors (IGF2BP1, IGF2BP3, HNRNPC, RBMX, METTL3, YTHDF2, METTL14) were significantly associated with T stage, and 2 m6A methylation regulatory factors (IGF2BP3, YTHDC2) were

［收稿日期］2023－03－25
［基金项目］包头市卫生健康科技计划项目(wsjkkj009)
［作者简介］吴启飞(1995－)，男，河北张家口人，内蒙古科技大学包头医学院第一附属医院医学硕士研究生，从事呼吸系统疾病诊治研究。
* 通信作者。E-mail:zhangdonggh@163.com

significantly associated with N stage. Cox regression analysis showed that 6 m6A methylation regulatory factors (IGF2BP1, IGF2BP2, IGF2BP3, HNRNPA2B1, HNRNPC, RBM15) were independent prognostic factors of lung adenocarcinoma. **Conclusion** m6A methylation regulatory factors are associated with the progression of lung adenocarcinoma. In addition, the 6 m6A methylation regulatory factors can be used as potential biomarkers to predict the prognosis of lung adenocarcinoma patients.

[**Key words**] lung adenocarcinoma; m6A methylation; TCGA database

肺癌是世界上最常见的恶性肿瘤,据估计,全球每年约有2万新发病例,病死率为20%。在影响肺部系统的所有癌症类型中,非小细胞肺癌占所有肿瘤80%～85%,包括腺癌、鳞癌、大细胞癌等多种类型[1]。研究[2]表明,N6-甲基腺嘌呤(N6-methyladenosine,m6A)甲基化调节因子在癌症发展和预后预测中起着至关重要的作用,但是对m6A修饰与肺腺癌之间的关系却知之甚少,m6A甲基化调节与肺腺癌预后之间的关系仍需进一步验证。本研究通过生物信息学方法探究m6A甲基化调节因子在肺腺癌中的表达情况,并分析其与肺腺癌预后的相关性。

1 资料与方法

1.1 数据收集 对于癌症基因组图谱(The Cancer Genome Atlas,TCGA)数据库的516例肺腺癌肿瘤患者,可从Genomic Data Commons(GDC)数据门户网站下载肿瘤RNA-seq数据(TCGA)(https://portal.gdc.cancer.gov/),见表1。m6A相关基因来源于Li等[2]在33种癌症类型中m6A调节因子的分子表征和临床意义的研究。

1.2 m6A甲基化调节因子在肺腺癌中的表达 通过R(v4.0.3)软件包ggplot2和pheatmap分析肺腺癌中m6A甲基化调节因子的表达与临床病理变量之间的关系。采用Wilcox检验分析516个肿瘤组织和59个癌旁组织中相关m6A甲基化调节因子的表达水平。癌旁组织样本的所有数据均来TCGA数据库中的癌旁样本及GTEx V8版本(https://gtexportal.org/home/datasets)。采用Spearman相关性分析,探究不同m6A甲基化调节因子之间的相关性。

1.3 m6A甲基化调节因子与肺腺癌患者临床病理特征的关系 为了研究m6A甲基化调节因子与肺腺癌患者临床病理特征之间的关系,分别分析在肺腺癌中差异表达的m6A甲基化调节因子与肿瘤TNM分期、T分期、N分期之间的关系。

表1 肺腺癌患者的临床病理

Table 1 Clinical pathology of patients with lung adenocarcinoma

(n=516,例数,%)

临床特征	TCGA数据
年龄(岁)	
≤60	177(34.3)
>60	339(65.7)
性别	
女性	278(53.9)
男性	238(46.1)
TNM分期	
Ⅰ	276(53.5)
Ⅱ	122(23.7)
Ⅲ	92(17.8)
Ⅳ	26(5.0)
T分期	
T1	169(32.8)
T2	278(53.9)
T3	47(9.1)
T4	19(3.7)
TX	3(0.5)
M分期	
M0	347(67.3)
M1	29(5.6)
MX	140(27.1)
N分期	
N0	332(64.4)
N1	97(18.8)
N2	74(14.3)
N3	2(0.4)
NX	11(2.1)

1.4 m6A甲基化调节因子的聚类分析 为了确定肺腺癌患者中20个m6A甲基化调节因子的m6A修饰模式,使用Consensus Cluster Plus R软件包进行非监督类聚类分析(consensus clustering),根据累积分布函数(cumulative distribution function,CDF)曲线的聚类分数,并评估Delta面积确定最佳聚类数(k值),从而评估TCGA肺腺癌队列中m6A甲基化调节因子的m6A修饰特征。

1.5 统计学方法 应用R软件包(v4.0.3)分析数据。采用单因素Cox回归分析m6A甲基化调节因子和临床表型之间进行生存分析,采用Kaplan-

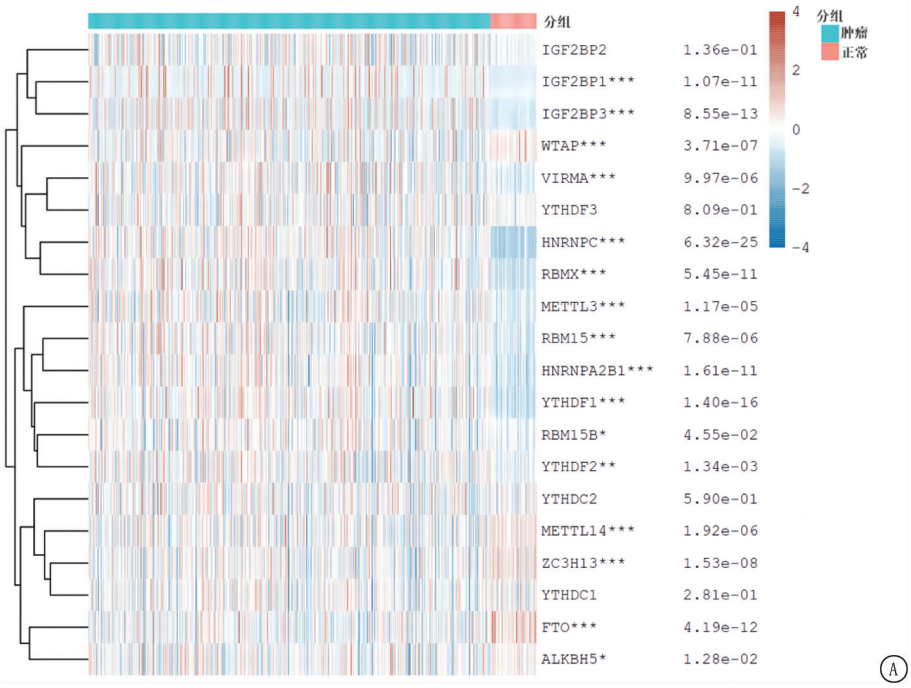
Meier法和log-rank检验分析高危组和低危组的OS；采用多变量Cox回归分析确定影响肺腺癌患者预后的独立危险因素。$P<0.05$为差异有统计学意义。

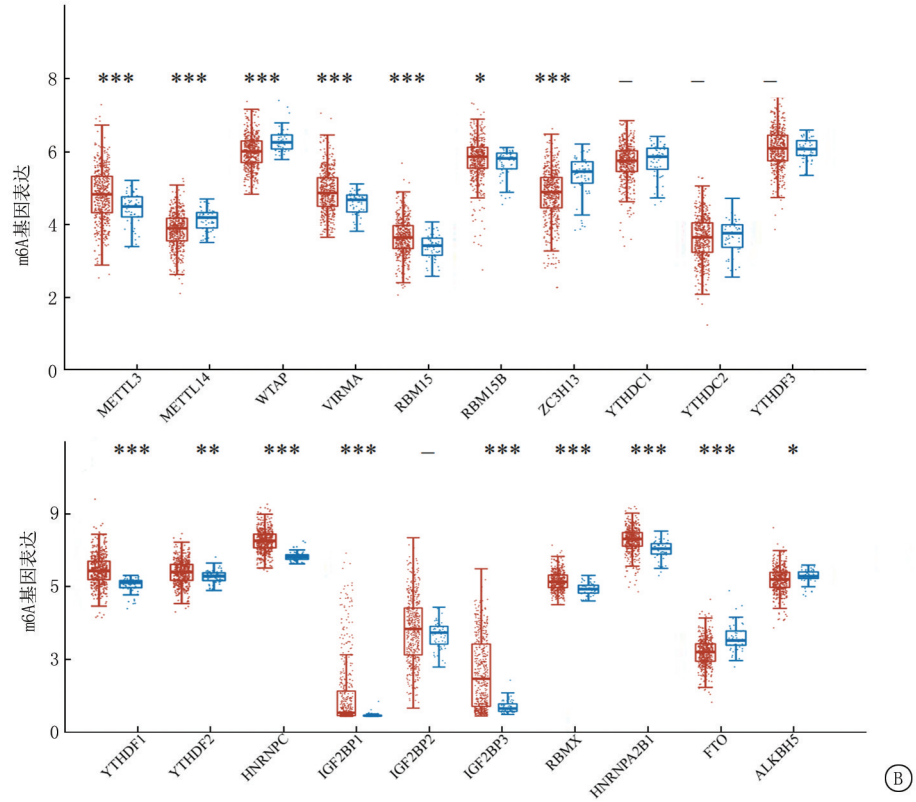
2 结　果

2.1 m6A RNA甲基化调节因子在肺腺癌中的表达水平 通过从TCGA数据库中获取的516例肺腺癌患者的RNA-seq转录组数据及其临床信息，分析比较了肿瘤组织和癌旁组织中20个m6A RNA甲基化调节因子的表达，显示了16个差异表达的调节因子。热图和小提琴图显示RBM15B（$P<0.05$）、ALKBH5（$P<0.05$）、YTHDF2（$P<0.01$）、IGF2BP1（$P<0.001$）、IGF2BP3（$P<0.001$）、VIRMA（$P<0.001$）、HNRNPC（$P<0.001$）、RBMX（$P<0.001$）、METTL3（$P<0.001$）、RBM15（$P<0.001$）、HNRNPA2B1（$P<0.001$）、YTHDF1（$P<0.001$）在癌组织中高表达；而WTAP（$P<0.001$）、METTL14（$P<0.001$）、ZC3H13（$P<0.001$）、FTO（$P<0.001$）在癌旁组织中高表达，见图1A，1B。Spearman相关性分析显示，在这16个m6A甲基化调节因子中，每两个调节因子之间相互比较，其中m6A甲基化调节因子之间相关性最强的是METTL14和ZC3H13，见图1C。

2.2 m6A甲基化调节因子与肺腺癌患者临床病理特征的关系 结果显示，在16个m6A甲基化调节因子中，IGF2BP1、IGF2BP3、HNRNPC、RBMX、YTHDC2在不同的TNM分期（Ⅰ、Ⅱ、Ⅲ、Ⅳ）中表达差异有统计学意义（$P<0.05$），见图2A。IGF2BP1、IGF2BP3、HNRNPC、RBMX、METTL3、YTHDF2、METTL14在不同的T分期（T1、T2、T3）中表达差异有统计学意义（$P<0.05$），见图2B。IGF2BP3、YTHDC2在不同的N分期（N0、N1、N2、N3）中表达差异有统计学意义（$P<0.05$），见图2C。

2.3 m6A甲基化调节因子的聚类分析 根据20个甲基化调节因子的表达水平进行非监督类聚类分析，结果显示聚类数为2时，CDF值增速平缓，没有样品特别小的分组，聚类稳定性最佳，见图3A、3B。此时分为2种不同的m6A修饰模式，称为m6A聚类C1、C2。对2组m6A聚类患者的预后分析，显示2组总体生存率比较差异无统计学意义（$P>0.05$），见图3C。然而当聚类数为6时，此时分为6中不同的m6A修饰模式，称为m6A聚类C1、C2、C3、C4、C5、C6，见图3D。对6组m6A聚类患者的预后分析，显示6组总体生存率比较差异有统计学意义（$P<0.05$），其中C3组预后最好。





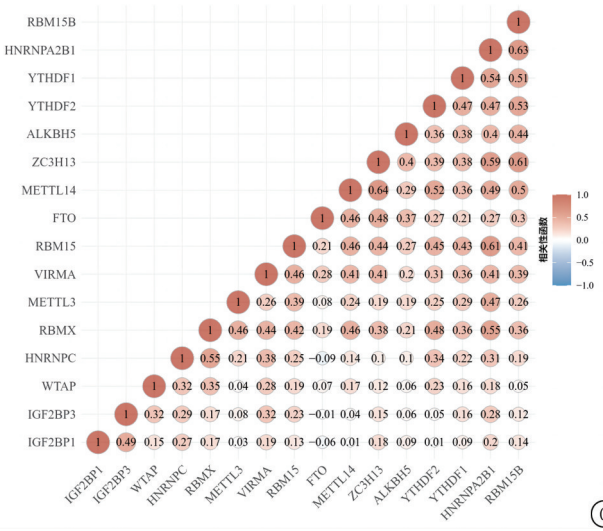


图 1　m6A 甲基化调节因子在肺腺癌中的表达

A.癌旁组织和肿瘤组织中 20 个 m6A 甲基化调节因子的表达热图；B.20 个 m6A 甲基化调节因子的小提琴图。癌旁组织标记为蓝色，肿瘤组织标记为红色；C.16 个 m6A RNA 甲基化调节因子的 Spearman 相关性分析

＊$P<0.05$　＊＊$P<0.01$　＊＊＊$P<0.001$

Figure 1　Expression of m6A methylation regulatory factors in lung adenocarcinoma

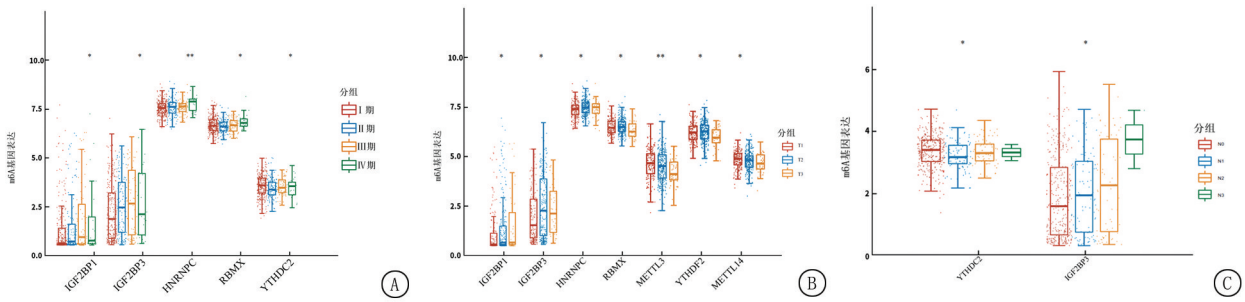


图 2　m6A 甲基化调节因子与肺腺癌患者临床病理特征的关系

A.m6A 甲基化调节因子与 TNM 分期的关系；B.m6A 甲基化调节因子与 T 分期的关系；C.m6A 甲基化调节因子与 N 分期的关系

＊$P<0.05$　＊＊$P<0.01$　＊＊＊$P<0.001$

Figure 2　Relationship between m6A methylation regulatory factors and clinicopathological features of lung adenocarcinoma

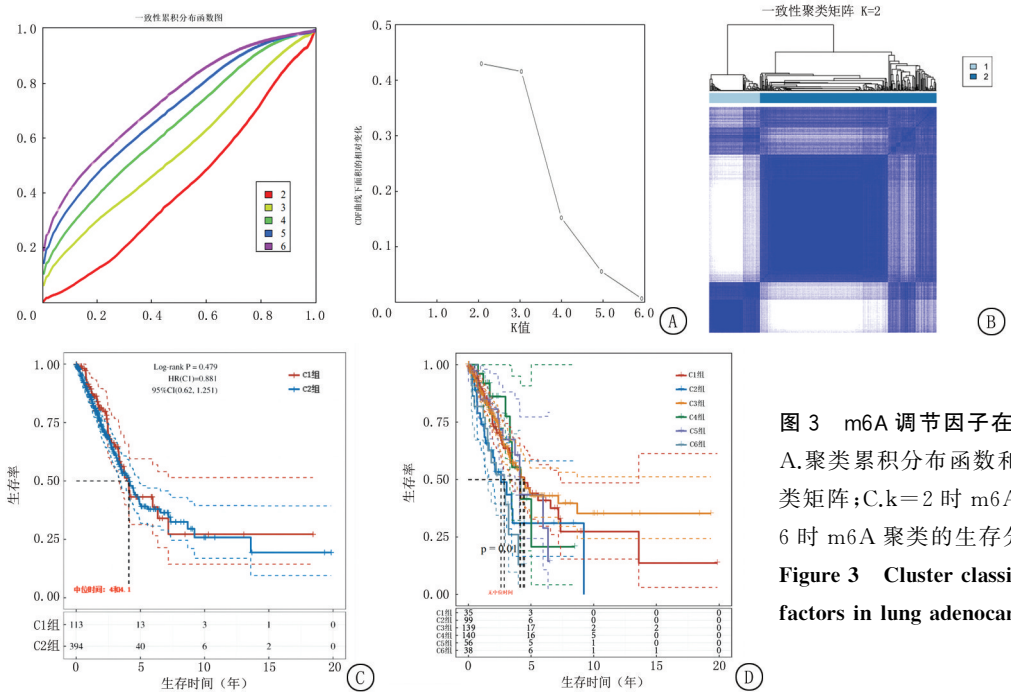


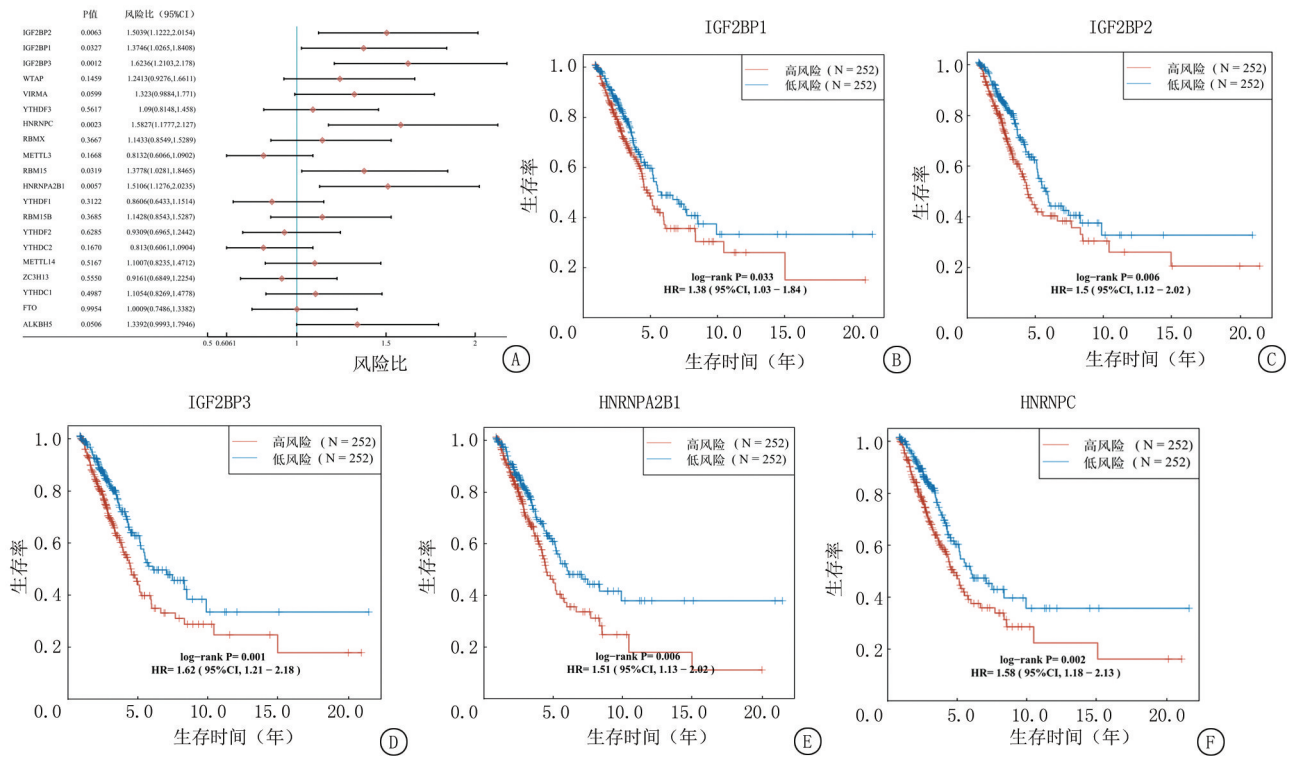
图3　m6A调节因子在肺腺癌中的聚类分类

A.聚类累积分布函数和Delta面积；B.k＝2时聚类矩阵；C.k＝2时m6A聚类的生存分析；D.k＝6时m6A聚类的生存分析

Figure 3　Cluster classification of m6A regulatory factors in lung adenocarcinoma

2.4　m6A甲基化调节因子对肺腺癌中预后的影响

通过单因素Cox回归分析，分析m6A甲基化调节因子在肺腺癌中的预后作用，见图4A。其中值得注意的是IGF2BP3的高表达导致了肺腺癌患者较差的生存[危险比(hazard ratio，HR)为1.624，95%可信区间(confidence interval，CI)为1.210～2.178]。Kaplan-Meier曲线结果显示IGF2BP1、IGF2BP2、IGF2BP3、HNRNPA2B1、HNRNPC和RBM15的高表达与不良预后显著相关($P<0.05$)，见图4B～4G。考虑到临床病理特征对患者生存率的影响，纳入患者年龄、性别和肿瘤TNM分期进行多因素Cox回归分析，与单因素Cox回归分析结果比较，显示IGF2BP1、IGF2BP2、IGF2BP3、HNRNPA2B1、HNRNPC和RBM15是影响肺腺癌患者预后的独立危险因素，见图4H。



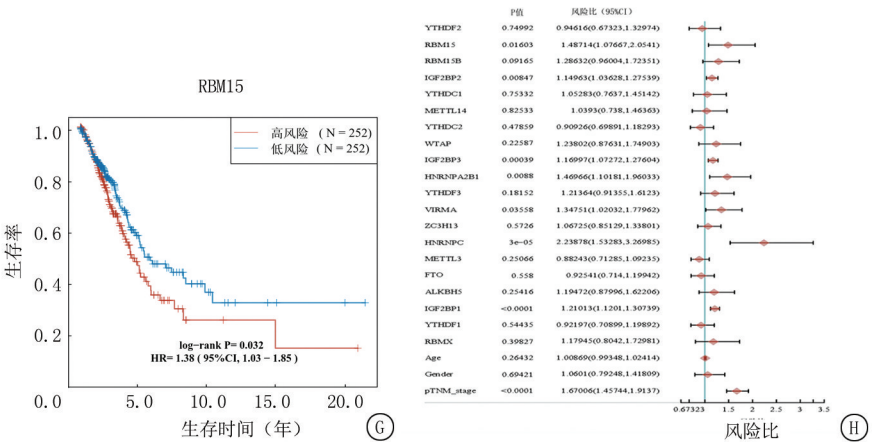


图4　m6A甲基化调节因子对肺腺癌预后的影响

A. 20个m6A甲基化调节因子的单变量Cox回归分析；B－G. IGF2BP1、IGF2BP2、IGF2BP3、HNRNPA2B1、HNRNPC、RBM15的Kaplan-Meier生存曲线；H.临床病理学特征和20个m6A甲基化调节因子的多变量Cox回归分析

Figure 4　Effect of m6A methylation regulatory factors on prognosis of lung adenocarcinoma

3　讨　论

N6甲基腺苷修饰，定义为在腺苷的N6位点添加甲基，是真核生物中最普遍的内部化学修饰[3]。由一组被称为“书写器”、“清除器”、“阅读器”的蛋白质组成，其可以影响RNA代谢的多个方面，包括mRNA的剪接、稳定性、定位和翻译。参与m6A甲基化的蛋白质主要有甲基转移酶、去甲基化酶和RNA结合蛋白3种类型[4]。m6A mRNA的甲基化主要由“书写器”蛋白完成，包括甲基转移酶样14（Methyltransferase-like 14，METTL14）、Wilms肿瘤1相关蛋白（Wilms′ tumor 1-associating protein，WTAP）、甲基转移酶类3（methyltransferase-like 3，METTL3）、RNA结合蛋白15（RNA-binding motif protein 15，RBM15）和含锌指CCCH结构域的蛋白13（zinc finger CCCH domain-containing protein 13，ZC3H13）。去甲基化过程主要由“清除器”蛋白进行，包括AlkB同系物5（Alk B homologous 5，ALKBH5）以及脂肪量和肥胖相关蛋白（fat mass and obesity associated protein，FTO）。“阅读器”蛋白主要是RNA降解和翻译，包括5个YT521-B（YT521-B homology，YTH）同源性YTH结构域家族成员、异质核核糖核蛋白（Heterogeneous nuclear ribonucleoprotein，hn RNP）和胰岛素样生长因子2 mRNA结合蛋白（Insulin like growth factor 2 mRNA binding protein，IGF2BP）[5]。由于“书写器”、“清除器”和“阅读器”基因协同控制可逆的m6A修饰，因此在生理活动和人类疾病，特别是癌症中发挥至关重要的作用[6]。

近年来，m6A修饰在免疫调节、代谢和细胞分化等许多生物过程中的作用已被证实。研究[7]显示RNA的m6A修饰通过调节癌蛋白表达，引起细胞增殖和肿瘤进展，在癌症发展中发挥着重要的调控作用。本研究结果表明RBM15B、ALKBH5、YTHDF2、IGF2BP1、IGF2BP3、VIRMA、HNRNPC、RBMX、METTL3、RBM15、HNRNPA2B1和YTHDF1等12个调节因子上调癌蛋白表达；而WTAP、METTL14、ZC3H13和FTO等4个调节因子下调癌蛋白表达。在大多数情况下，METTL3被认为是一种癌基因，促进多种癌症的发生和发展。例如在肠道中，微生物代谢物丁酸盐被认为下调METTL3的表达并抑制结直肠癌的发展[8]。Wang等[9]在胃癌中显示，P300介导组蛋白H3在赖氨酸27（H3K27ac）处乙酰化并促进METTL3转录。研究[10]报道了METTL3在乳腺癌中的过表达，并确定Bcl-2是METTL3的靶标，证明由于m6A mRNA的甲基化修饰，促进了癌细胞增殖。而METTL14被认为是一种抑癌基因，有研究[11]表明METTL14缺失与结直肠癌患者的不良预后相关，敲低METTL14显著增强了CRC细胞体外的增殖和侵袭能力，促进了体内致瘤性和转移性。METTL3和METLL14虽然分别作为促癌基因和抑癌基因，但在本研究中，METTL3和METTL14均在肺腺癌中差异表达，如何能针对两者设计出有效的靶点药物需要综合考虑。HNRNPC是一种RNA结合蛋白，Huang等[12]研究显示HNRNPC的过表达通过上皮－间充质转化（epithelial-mesenchymal transition，EMT）促进口腔鳞状细胞癌的致癌作用。此外，circBACH2（hsa_circ_0001625）被证实通过MAPK信号通路刺激HNRNPC的表达，从而促进乳腺癌细胞增殖[13]。含YT521-B同源结构域的蛋白质指导不同的复合物调节RNA信号通路，包括RNA折叠、RNA剪接、蛋白质翻译和RNA代谢[14]。研究[15]证实IGF1/IGF1R抑制剂林西替尼优先靶向表达YTHDF2的细胞，降低胶质母细胞瘤干细胞活力而不影响正常神经干细胞，抑制体内胶质母细胞瘤生

长。研究[16]表明除了原发性肝细胞癌中的 YTHDF2 以及在头颈部鳞状细胞癌中的 YTHDC2 以外，几乎所有的 YTH 家族蛋白在各种类型的癌症中都作为癌基因。IGF2BP 蛋白家族（包括 IGF2BP1/2/3）作为一个独特的 m6A 阅读器，已被证明与癌症进展密切相关，IGF2BP 蛋白家族识别经甲基化修饰的 mRNA，并在稳定的核糖核蛋白颗粒的背景下延长其生命周期，以促进癌症进展[17]。还有研究[18]证明 circNDUFB2 作为支架增强 TRIM25 和 IGF2BPs 之间的相互作用，促使 IGF2BPs 降解，从而参与非小细胞肺癌的进展。来自中国国家癌症中心的队列研究[19]证实了 IGF2BP3 高表达可能与性别、肿瘤长度、分化和 T 分期显著相关。本研究结果显示，IGF2BP 蛋白家族不仅参与调控肺腺癌的增殖、侵袭转移等多个生物过程，还导致肺腺癌患者的不良预后。肺腺癌中包含多个肿瘤位点，但是部分基因在不同的肿瘤类型中表现出相互矛盾的作用，因此生物信息学分析成为探索肺腺癌核心基因、为肿瘤治疗提供潜在靶点的有效途径。

综上所述，m6A 甲基化调节因子的表达与肺腺癌患者的临床病理特征密切相关。6 个 m6A 调节因子也是影响肺腺癌患者生存预后的独立危险因素，可以作为预测肺腺癌患者预后的潜在生物标志物。本研究基于 TCGA 数据库，是否适用于中国人还有待于进一步研究。由于临床分析有一定的局限性，接下来还需要开展基础实验来进一步验证结果的可靠性。

[参考文献]

[1] Suster DI, Mino-Kenudson M. Molecular pathology of primary non-small cell lung cancer[J]. Arch Med Res, 2020, 51(8): 784－798.

[2] Li Y, Xiao J, Bai J, et al. Molecular characterization and clinical relevance of m6A regulators across 33 cancer types [J]. Mol Cancer, 2019, 18(1): 137.

[3] Oerum S, Meynier V, Catala M, et al. A comprehensive review of m6A/m6Am RNA methyltransferase structures [J]. Nucleic Acids Res, 2021, 49(13): 7239－7255.

[4] Huang J, Shao Y, Gu W. Function and clinical significance of N6-methyladenosine in digestive system tumours[J]. Exp Hematol Oncol, 2021, 10(1): 40.

[5] Zhang N, Zuo Y, Peng Y, et al. Function of N6-methyladenosine modification in tumors[J]. J Oncol, 2021, 2021: 6461552.

[6] Deng LJ, Deng WQ, Fan SR, et al. m6A modification: recent advances, anticancer targeted drug discovery and beyond[J]. Mol Cancer, 2022, 21(1): 52.

[7] Huang H, Weng H, Chen J. m6A Modification in coding and non-coding RNAs: roles and therapeutic implications in cancer [J]. Cancer Cell, 2020, 37(3): 270－288.

[8] Zhu W, Si Y, Xu J, et al. Methyltransferase like 3 promotes colorectal cancer proliferation by stabilizing CCNE1 mRNA in an m6A-dependent manner[J]. J Cell Mol Med, 2020, 24(6): 3521－3533.

[9] Wang Q, Chen C, Ding Q, et al. METTL3-mediated m6A modification of HDGF mRNA promotes gastric cancer progression and has prognostic significance[J]. Gut, 2020, 69(7): 1193－1205.

[10] Wang H, Xu B, Shi J. N6-methyladenosine METTL3 promotes the breast cancer progression via targeting Bcl-2[J]. Gene, 2020, 722: 144076.

[11] Yang X, Zhang S, He C, et al. METTL14 suppresses proliferation and metastasis of colorectal cancer by down-regulating oncogenic long non-coding RNA XIST[J]. Mol Cancer, 2020, 19(1): 46.

[12] Huang GZ, Wu QQ, Zheng ZN, et al. M6A-related bioinformatics analysis reveals that HNRNPC facilitates progression of OSCC via EMT[J]. Aging (Albany NY), 2020, 12(12): 11667－11684.

[13] Lv W, Tan Y, Xiong M, et al. Analysis and validation of m6A regulatory network: a novel circBACH2/has-miR-944/HNRNPC axis in breast cancer progression[J]. J Transl Med, 2021, 19(1): 527.

[14] Liao J, Wei Y, Liang J, et al. Insight into the structure, physiological function, and role in cancer of m6A readers-YTH domain-containing proteins[J]. Cell Death Discov, 2022, 8(1): 137.

[15] Dixit D, Prager BC, Gimple RC, et al. The RNA m6A reader YTHDF2 maintains oncogene expression and is a targetable dependency in glioblastoma stem cells[J]. Cancer Discov, 2021, 11(2): 480－499.

[16] Liu S, Li G, Li Q, et al. The roles and mechanisms of YTH domain-containing proteins in cancer development and progression[J]. Am J Cancer Res, 2020, 10(4): 1068－1084.

[17] Ramesh-Kumar D, Guil S. The IGF2BP family of RNA binding proteins links epitranscriptomics to cancer[J]. Semin Cancer Biol, 2022, 86(Pt 3): 18－31.

[18] Li B, Zhu L, Lu C, et al. circNDUFB2 inhibits non-small cell lung cancer progression via destabilizing IGF2BPs and activating anti-tumor immunity[J]. Nat Commun, 2021, 12(1): 295.

[19] Guo W, Huai Q, Wan H, et al. Prognostic impact of IGF2BP3 expression in patients with surgically resected lung adenocarcinoma[J]. DNA Cell Biol, 2021, 40(2): 316－331.

（本文编辑：何祯）