

论 著

· 临床研究 ·

# 基于数据挖掘分析GALNT2基因在肺腺癌中的表达及临床意义

陈凯锐<sup>1</sup>, 童华生<sup>2</sup>, 周泉<sup>1</sup>, 孙朝晖<sup>1\*</sup><sup>1</sup>南部战区总医院检验科, 广州 510010; <sup>2</sup>南部战区总医院急诊科, 广州 510010

**[摘要]** **目的** 基于数据挖掘探索N-乙酰氨基半乳糖转移酶2(GALNT2)基因在肺腺癌中的表达及临床意义。**方法** 利用Oncomine数据库对GALNT2基因在肺腺癌中的表达进行荟萃分析; 利用UALCAN平台分析TCGA及CPTAC数据库中肺腺癌和正常组织中GALNT2基因的表达情况及其与临床特征的关系; 采用KM-plotter网站对肺腺癌中GALNT2表达量与预后的关系进行分析; 利用cBioportal在线分析工具分析GALNT2基因在肺腺癌中的突变情况; 使用String数据库构建相关蛋白网络并进行功能富集分析。**结果** Oncomine数据库荟萃分析结果显示, 肺腺癌组织中GALNT2的mRNA表达及DNA拷贝数变异(CNV)明显高于正常组织( $P<0.001$ )。UALCAN平台分析结果显示, 肺腺癌患者肿瘤组织中GALNT2的mRNA及蛋白表达水平明显高于正常组织, 但其表达水平与患者临床病理特征不相关。cBioportal在线分析工具结果显示, 在503例肺腺癌样本中共有52例样本发生GALNT2基因变异, 变异率约为10%。KM-plotter生存曲线分析结果显示, GALNT2高表达组及GALNT2变异组的总体生存期短于GALNT2低表达组及无变异组( $P<0.001$ )。多因素Cox回归分析结果显示, GALNT2高表达是影响肺腺癌患者预后的独立危险因素( $HR=1.48$ ,  $P=0.01$ )。String数据库分析结果表明, GALNT1、RAB6A、RAB6B、DYNC1H1、DYNC1LI1、DCTNs、CAPZAs及多种黏蛋白与GALNT2具有明显的相互作用, 主要参与的生物学过程有蛋白质丝氨酸/苏氨酸O型糖基化、细胞器微管运输及白细胞介素-17信号通路。**结论** 数据挖掘结果表明GALNT2在肺腺癌组织中呈高表达, 具有较高的基因变异频率, 且与患者预后不良相关。

**[关键词]** 肺腺癌; 基因, GALNT2; 数据挖掘; 预后**[中图分类号]** R734.2**[文献标志码]** A**[文章编号]** 0577-7402(2021)10-0975-09**[DOI]** 10.11855/j.issn.0577-7402.2021.10.04

## Expression of GALNT2 in lung adenocarcinoma and its prognostic significance based on data mining

Chen Kai-Rui<sup>1</sup>, Tong Hua-Sheng<sup>2</sup>, Zhou Quan<sup>1</sup>, Sun Zhao-Hui<sup>1\*</sup><sup>1</sup>Department of Clinical Laboratory, <sup>2</sup>Department of Critical Care Medicine, General Hospital of Southern Theatre Command of Chinese PLA, Guangzhou 510010, China

\*Corresponding author, E-mail: zhaohui3@126.com

This work was supported by the National Natural Science Foundation of China (31901035), the Natural Science Foundation of Guangdong Province (2020A1515010951), and the Science and Technology Program of Guangzhou (202102021265)

**[Abstract]** **Objective** To explore the expression and clinical significance of GALNT2 protein in lung adenocarcinoma (LAC) based on data mining. **Methods** GALNT2 gene expression in LAC was analyzed using Oncomine database. UALCAN analysis platform was used to analyze the relationship between GALNT2 gene expression in The Cancer Genome Atlas (TCGA) and Clinical Proteomic Tumor Analysis Consortium (CPTAC) datasets and clinical features, and KM-plotter website was used to analyze the relationship between GALNT2 gene expression in LAC and prognosis. The mutations of GALNT2 gene in LAC from TCGA dataset were explored using cBioportal online analysis tools. The related protein networks were constructed for function enrichment analysis by using String database. **Results** The data in the Oncomine database showed that the mRNA expression and DNA copy number variation (CNV) of GALNT2 in LAC tissues were significantly higher than those in normal tissues ( $P<0.01$ ). The transcription and protein level of GALNT2 in tumor tissues in LAC patients were significantly higher than that in marching tissues, but its expression level had no correlation with the clinical pathological characteristics of patients. A total of 52 samples of 503 LAC samples had GALNT2 gene mutation, and the mutation rate was about 10%. Survival curve analysis showed that the overall survival

**[基金项目]** 国家自然科学基金(31901035); 广东省自然科学基金(2020A1515010951); 广州市科技计划项目(202102021265)**[作者简介]** 陈凯锐, 主管技师, 主要从事临床医学检验方面的研究**[通信作者]** 孙朝晖, E-mail: zhaohui3@126.com

periods of GALNT2 high expression group and GALNT2 mutation group were shorter than those of GALNT2 low expression group and no mutation group ( $P<0.01$ ). Multivariate Cox regression analysis showed that the high expression of GALNT2 was an independent risk factor of prognosis in LAC ( $HR=1.48, P=0.01$ ). String database analysis revealed that GALNT1, RAB6A, RAB6B, DYNC1H1, DYNC1L1, DCTNs, CAPZAs and mucins had obvious interactions with GALNT2, and these proteins were mainly involved in minus-end-directed organelle transport along microtubule, protein O-linked glycosylation *via* serine/threonine, O-glycan processing and IL-17 signaling pathway. **Conclusions** GALNT2 can be highly expressed in LAC tissues, has a high frequency of gene mutation, and may be related to poor prognosis of patients using data mining methods, thus providing a theoretical basis for further research on the role of GALNT2 in the occurrence and progression of LAC.

**[Key words]** lung adenocarcinoma; gene, GALNT2; data mining; prognosis

肺癌是全世界癌症相关死亡人数最多的癌症,根据2018年的全球癌症统计数据,肺癌的病死人数占癌症总死亡人数的18.4%<sup>[1]</sup>。肺癌中约有85%被归类为非小细胞肺癌(non-small cell lung cancer, NSCLC),并可细分为鳞状细胞癌、肺腺癌及大细胞癌<sup>[2-3]</sup>。肺腺癌是最常见的肺癌亚类,占NSCLC的65%,占有所有肺癌的40%以上<sup>[4-5]</sup>。2015年,中国新发肺癌约有78.7万例<sup>[6]</sup>。肺癌的发生发展是一个复杂、多步骤的过程,与各种基因突变、细胞信号转导通路激活及异常新生血管形成密切相关,但其分子机制仍不完全清楚<sup>[4]</sup>。通过筛选与肿瘤相关的基因网络,有可能确定肺腺癌新的药物靶标并发现用于治疗及诊断的生物标志物。糖蛋白发挥着多种生物学功能,如细胞间信息传递、细胞内信号转导及维持蛋白质稳定性<sup>[7]</sup>。黏蛋白型或N-乙酰基-D-半乳糖胺(GalNAc)型O-糖基化是由一大类多肽GalNAc-转移酶(GALNT,也称为GalNAc-T)同工酶催化的O-蛋白质的糖基化。N-乙酰氨基半乳糖转移酶2(GALNT2)是GALNT家族的第2个成员<sup>[8]</sup>,广泛分布于肾脏、胰腺、胃、肠等多种器官中<sup>[9]</sup>。

全基因组关联研究(GWAS)确定GALNT2参与了人类脂蛋白的代谢调控<sup>[10]</sup>,其催化底物包括脂蛋白脂肪酶抑制剂类血管生成素3(ANGPTL3)、载脂蛋白C-III(apoC-III)<sup>[11]</sup>及磷脂转移蛋白(PLTP)等<sup>[12]</sup>。后续研究显示,该基因在多种恶性肿瘤,如胃癌、胶质瘤、肝细胞癌及口腔鳞状细胞癌中表达失调<sup>[13-15]</sup>。然而,GALNT2在肺癌尤其是肺腺癌发生、发展中的作用仍不清楚。本研究基于Oncomine数据库、UALCAN数据库、cBioportal数据库及String数据库进行数据挖掘,进而分析GALNT2基因在肺腺癌中的表达情况及其预后意义,旨在为后续研究GALNT2基因在肺腺癌发生发展中的作用奠定基础。

## 1 材料与方法

**1.1 利用Oncomine数据库分析GALNT2在肺腺癌中的表达** 收集分析Oncomine数据库肿瘤样本的基因表达谱芯片数据,同时提供转录组数据,

筛选条件如下。(1)Gene: GALNT2; (2)Analysis type: cancer *vs.* normal analysis; (3)Cancer type: lung cancer; (4)Date type: mRNA and DNA; (5)Sample type: clinical specimen; (6)临界值设定条件( $P<0.05$ ; gene rank: top 10%)。利用Oncomine数据库对来自(The Cancer Genome Atlas, TCGA)和Gene Expression Omnibus(GEO)中评估了GALNT2转录水平的多个恶性肿瘤研究进行分析。将所有结果汇总后进行综合分析比较。

**1.2 利用UALCAN数据库分析GALNT2表达量与临床特征及预后的关系** 通过基于TCGA搭建的数据可视化平台UALCAN,在UALCAN数据库中的TCGA及CPTAC(Clinical Proteomic Tumor Analysis Consortium)数据集中检索GALNT2,于数据库选项中选择肺腺癌。分别对年龄、性别、TNM分期、淋巴结转移等临床特征分组及预后情况进行分析。

**1.3 利用cBioportal工具分析TCGA数据库中GALNT2基因在肺腺癌中的突变情况** 利用cBioportal在线工具,在数据库类型中选择“TCGA database”,肿瘤类型选择肺腺癌,基因选择“GALNT2”,分析GALNT2在肺腺癌中的突变情况。应用survival选项分析GALNT2变异与患者生存时间的Kaplan-Meier曲线。

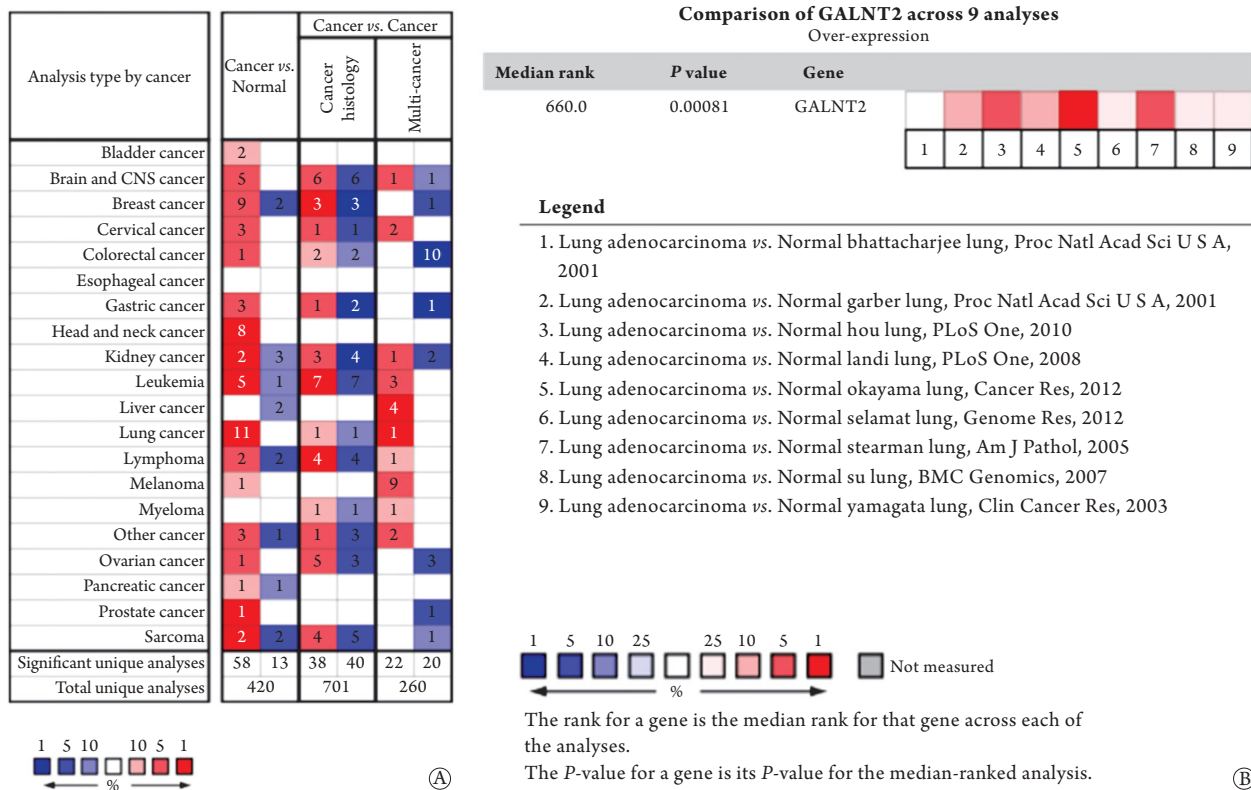
**1.4 利用String数据库构建GALNT2蛋白相互作用网络** String用于构建蛋白质-蛋白质相互作用(PPI)网络,可生成有关基因功能的假设,分析基因列表并进行功能分析。本研究中设置蛋白相互作用网络来源为:物理相互作用、基因共表达、基因共定位、基因富集分析及网站预测。

**1.5 统计学处理** 采用SPSS 26.0软件进行统计分析。正常组织与肺腺癌组织中GALNT2基因表达的差异采用两独立样本非参数Mann-Whitney U检验,定量资料的多组间两两比较采用P值BH校正法分析。GALNT2表达与患者预后的关系采用Kaplan-Meier模型进行生存分析。在校正了相应的协变量后,采用多因素Cox比例风险回归模型分析GALNT2表达对总体生存率的影响。 $P<0.05$ 为差异有统计学意义。

## 2 结 果

**2.1 GALNT2在肺腺癌中的表达情况** Oncomine 4.5 数据库数据中共有420项肿瘤组织与正常组织的比较研究，其中GALNT2在肿瘤中高表达的研究有58项，低表达的研究有13项；肺癌研究共33项，其中11项研究显示GLANT2在肺癌组织中的表达明显高于肺部正常组织，而无一项研究显示GALNT2

在肺癌中的表达低于正常组织(表达水平变化>1.5倍,  $P<0.05$ , Gene rank: Top 10%)(图1A)。肺腺癌中GALNT2表达情况分析结果显示，共有9项研究比较了肺腺癌与正常组织中GALNT2的表达差异，其中8项研究数据表明，GALNT2在肺腺癌组织中呈高表达(图1B)。肺腺癌组织中GALNT2的mRNA表达及DNA拷贝数变异(CNV)明显高于正常组织( $P<0.01$ , 图2)。



**图1** Oncomine数据库中GALNT2基因在所有主要肿瘤(A)及肺腺癌(B)中与相应正常组织的差异表达情况  
**Fig.1** Differential expression of GALNT2 in all major cancers (A) and lung adenocarcinomas (B) and their corresponding normal tissues in Oncomine database

红色方格示mRNA表达水平高于正常组织均值，蓝色方格示mRNA水平低于正常组织均值，方格颜色深浅由分析的基因秩百分位数决定，颜色越深，百分位数越小，即差异倍数越大。

### 2.2 GALNT2表达量与肺腺癌患者临床特征的关系

利用UALCAN数据库对TCGA中的574个肺腺癌样本的多种临床病理特征进行亚组分析，结果显示，肺腺癌中GALNT2 mRNA的表达水平明显高于正常组织( $P<0.001$ )，但其表达水平与患者临床病理特征如肿瘤分期、淋巴结转移、性别、吸烟史及TP53突变均无明显相关性(图3)。利用UALCAN数据库进一步分析GALNT2蛋白表达水平与临床特征的关系，结果显示，GALNT2蛋白表达水平在肺腺癌组织中明显升高( $P<0.001$ )，但与上述临床特征仍无相关性(图4)。

### 2.3 GALNT2表达水平与肺腺癌患者预后的关系

由于GALNT2在肺腺癌中表达水平明显升高，为进一步明确GALNT2表达与肺腺癌预后的关系，从

KM-plotter数据库中收集肺腺癌的临床病理参数，按GALNT2的表达水平分为高表达组与低表达组并进行Kaplan-Meier生存分析。结果显示，GALNT2表达水平与肺腺癌患者的总生存时间(overall survival, OS)及首次进展时间(fires progression, FP)均明显相关( $P<0.001$ , 图5)。

本研究通过多因素Cox比例风险模型进一步分析GALNT2表达水平与肺腺癌总生存时间的关系，并对患者的年龄、临床分期、T分期、M分期和N分期进行了协变量校正，结果显示，GALNT2的表达状态(分类变量)与总生存时间明显相关( $HR=1.48$ , 95%CI 1.098~1.997,  $P=0.01$ )。

### 2.4 GALNT2基因在肺腺癌中的变异情况

采用

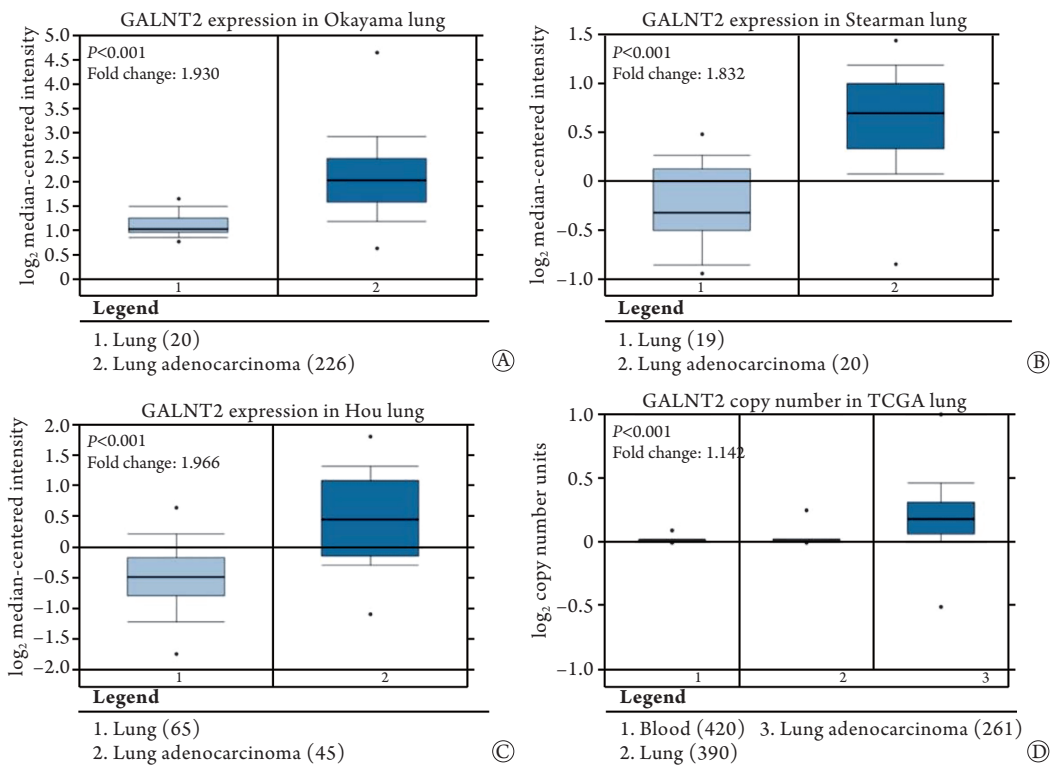


图2 Oncomine数据库中GALNT2基因在肺腺癌与正常组织中的差异表达情况

Fig.2 Differential expression of GALNT2 in lung adenocarcinoma and corresponding normal tissues in Oncomine database  
A-C. 分别为3个不同数据集中的GALNT2 mRNA水平; D. 癌基因组图谱(TCGA)数据库中的GALNT2拷贝数差异

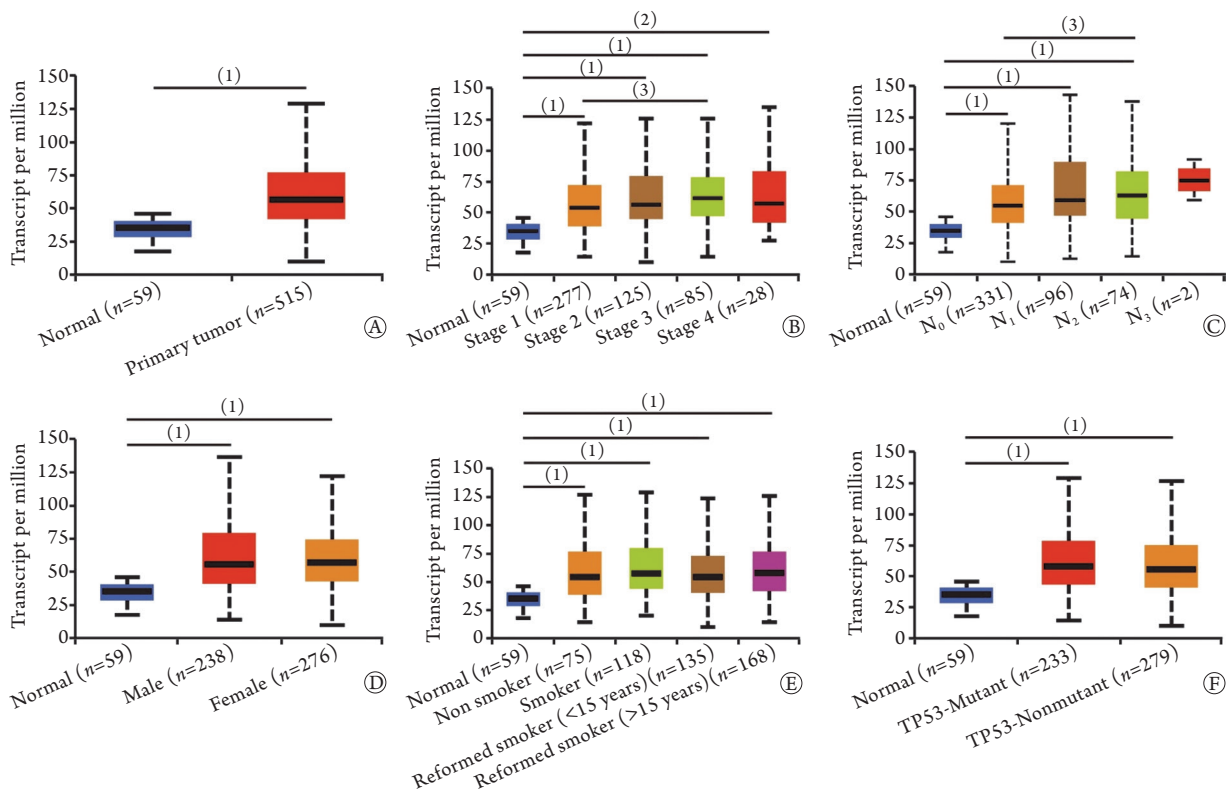


图3 TCGA数据库中根据不同临床特征分组的GALNT2 mRNA表达水平(UALCAN)

Fig.3 Expression levels of GALNT2 mRNA in TCGA database grouped according to different clinical characteristics (UALCAN)

A. GALNT2在正常组织与肺腺癌样品中的相对表达水平; B. 不同分期肺腺癌组织与正常组织中GALNT2的表达水平; C. GALNT2在不同淋巴结转移肺腺癌组织与正常组织的相对表达水平; D. GALNT2在正常组织与不同性别肺腺癌患者中的相对表达水平; E. GALNT2在正常组织与不同吸烟习惯肺腺癌患者中的相对表达水平; F. GALNT2在正常组织与不同TP53突变情况肺腺癌患者中的相对表达量。箱线图使用 $M(Q_1, Q_3)$ 表示, 上下边缘为最大值与最小值。(1) $P<0.001$ ; (2) $P<0.01$ ; (3) $P<0.05$ 。

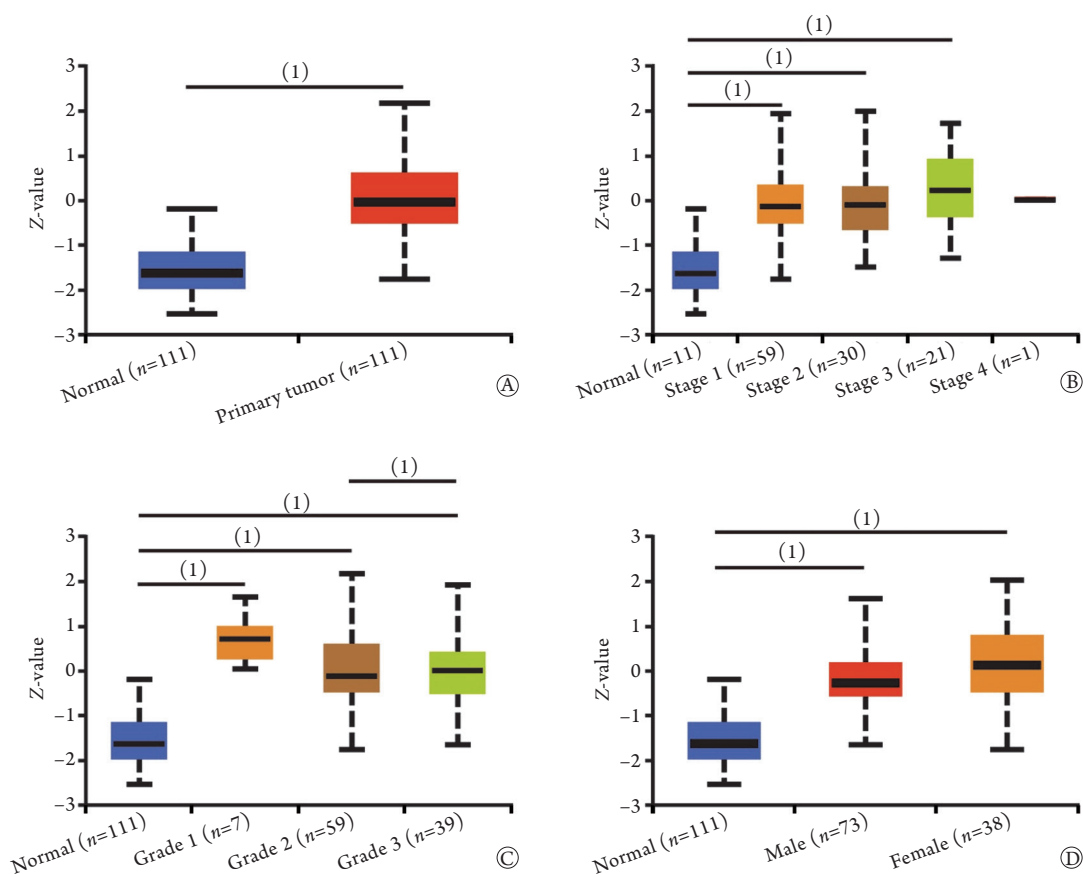


图4 CPTAC数据库中根据不同临床特征分组的GALNT2 蛋白表达水平(UALCAN)

**Fig.4** Expression levels of GALNT2 protein in CPTAC database grouped according to different clinical characteristics (UALCAN)  
 A. GALNT2在正常组织与肺腺癌样本中的蛋白表达水平; B. 不同分期肺腺癌组织与正常组织中GALNT2的蛋白表达水平;  
 C. GALNT2在不同肿瘤分级肺腺癌患者与正常组织中的蛋白表达水平; D. GALNT2在正常组织与不同性别肺腺癌患者中的蛋白表达水平; (1) $P < 0.001$ .

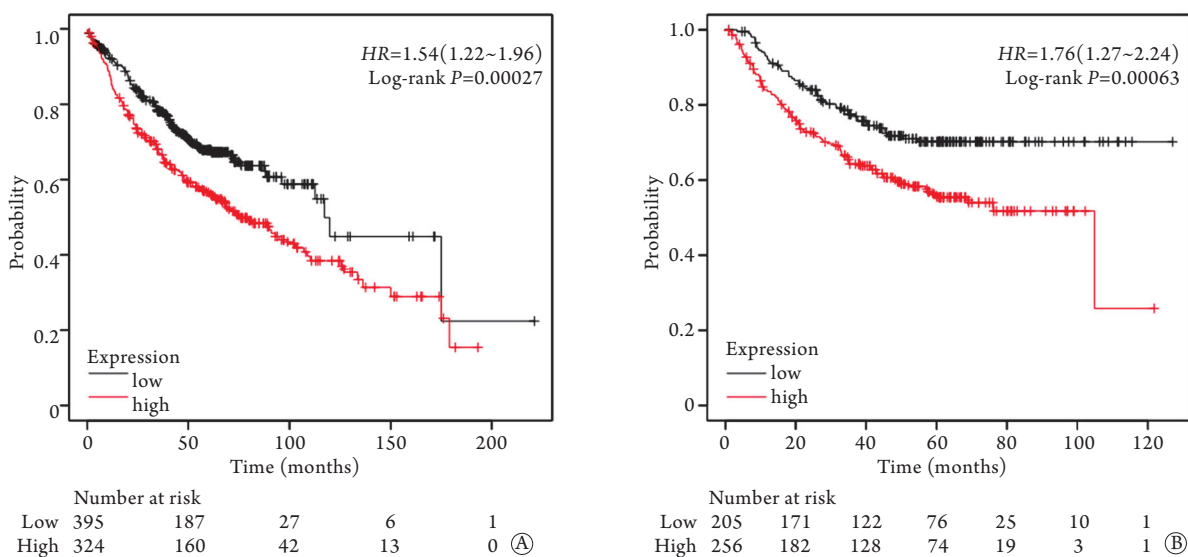


图5 GALNT2表达水平与肺腺癌患者预后的关系(KM-plotter)

**Fig.5** Relationship between GALNT2 expression status and the prognosis of lung adenocarcinoma (KM-plotter)

A. 总生存率; B. 首次进展时间

cBioportal网站对TCGA中503例资料完整的肺腺癌样本进行分析, 如图6A所示, 其中共有52例样本发生GALNT2基因变异, 变异率约为10%, 这些变

异包括扩增15例(3%), mRNA上调31例(6.1%), 突变9例(1.8%)。GALNT2基因共有10个突变位点, 均为Missense突变(图6B)。比较GALNT2基因变异组

与未变异组患者的总生存率，结果表明，GALNT2发生变异的患者总生存时间较短，同样GALNT2高表达组的生存率明显降低(图6C、D)。然而基因扩

增的15例样本中，仅2例出现了mRNA高表达现象(图6A)，表明GALNT2在转录层面的表达异常并非由基因组层面的扩增变异引起。

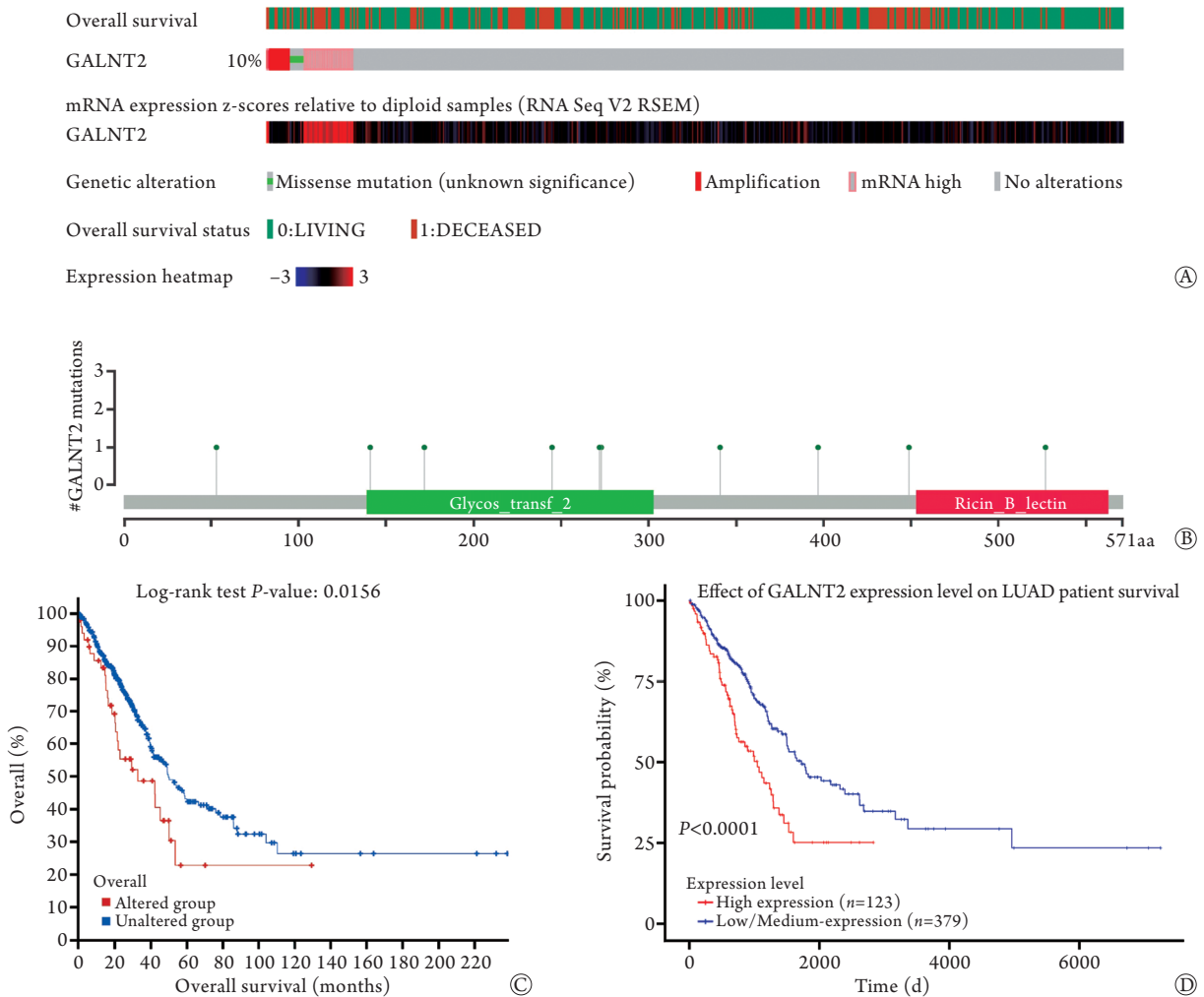


图6 肺腺癌中GALNT2基因变异与预后的关系(cBioportal)

Fig.6 Relationship between GALNT2 gene alteration and the prognosis in lung adenocarcinoma (cBioportal)

A. 肺腺癌GALNT2基因在TCGA数据库中的变异情况(不同类型的遗传改变以不同的颜色突出显示); B. GALNT2基因突变位点示意图; C. GALNT2基因变异对总生存时间的影响; D. TCGA数据库中GALNT2表达量对总生存时间的影响

**2.5 与GALNT2蛋白相互作用的蛋白网络** 利用String构建GALNT2蛋白相互作用网络，PPI富集P值(PPI enrichment  $P$ -value) $<0.001$ ，其中节点数目为26个，主要包括GALNT1、RAB6A、RAB6B、DYNC1H1、DYNC1LI1、DCTNs、CAPZAs及多种黏蛋白等，主要参与的生物学过程有蛋白质丝氨酸/苏氨酸O型糖基化、细胞器微管运输及白介素-17信号通路等(图7)。

### 3 讨论

肺腺癌是全球最常见的恶性肿瘤之一<sup>[1,16]</sup>。早期发现的肺腺癌患者经过手术切除等治疗后，预后较好，但大多数肺腺癌患者确诊时已处于中晚期，且多发生了肿瘤转移。虽然目前已经开发出诸如靶

向间变性淋巴瘤激酶(ALK)、表皮生长因子受体-酪氨酸激酶抑制剂(EGFR-TKI)等药物治疗携带此类基因突变的肺腺癌患者，但是相当一部分患者并未携带靶向药物基因突变，或对靶向治疗产生耐药，预后较差<sup>[17-18]</sup>。同时，肺腺癌发生发展的分子机制仍不完全明了，需要进一步研究。相关研究表明，肺腺癌的发生和发展是受自身遗传及环境因素共同影响的复杂过程，在病变过程中，有多种基因的参与及调控<sup>[19]</sup>。因此，进一步研究肺腺癌发病的分子机制及其相应的治疗靶点及检测标志物，是目前亟需解决的问题。

糖基化是蛋白质最常见的翻译后修饰之一。黏蛋白型O-糖基化(Mucin-type O-糖基化)为O-糖基化的最常见类型。Mucin-type O-糖基化的起始步骤

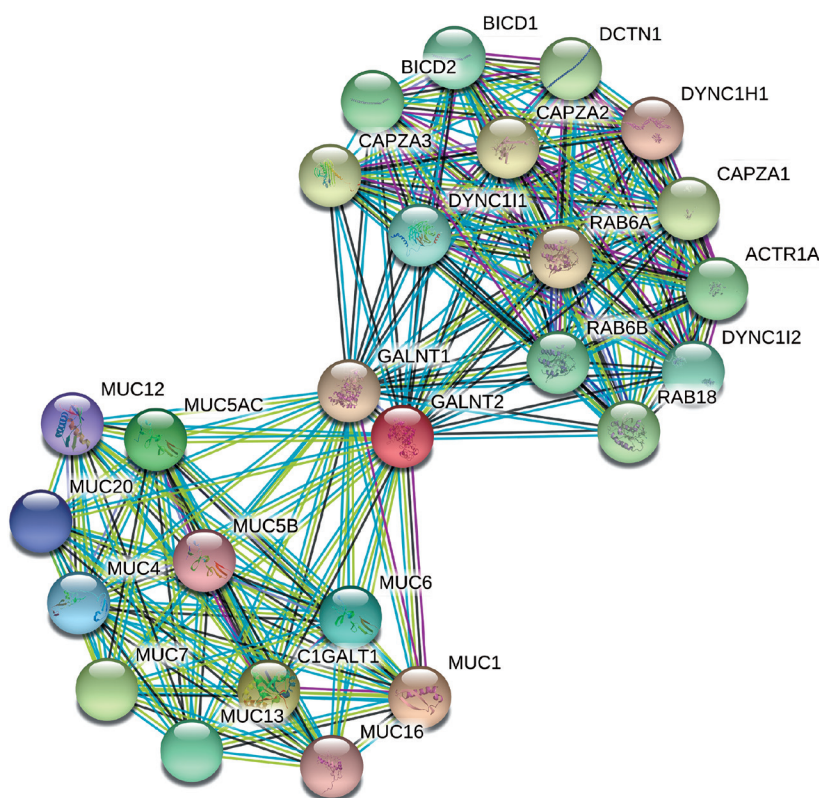


图7 GALNT2在肺腺癌中的蛋白质-蛋白质相互作用网络(String)

Fig.7 The protein-protein interaction network of GALNT2 in lung adenocarcinoma (String)

由多肽N-乙酰氨基半乳糖转移酶催化,将GalNAc连接至糖蛋白的丝/苏氨酸残基上<sup>[20]</sup>。此修饰过程发生于高尔基复合体,受GALNT家族成员的调控<sup>[21]</sup>。大量研究表明,异常调节的Mucin-type O-糖基化蛋白在细胞增殖、分化、凋亡、迁移、侵袭及黏附等过程中发挥了重要的调节作用,进而影响肿瘤的发生发展及复发转移情况<sup>[22-24]</sup>,且已有研究证实,GALNTs的异常表达与多种恶性肿瘤的进展程度及预后明显相关<sup>[25-26]</sup>。

GALNT2是N-乙酰氨基半乳糖转移酶家族的第2个成员。近年来,多个GWAS研究发现GALNT2基因多态性与高密度脂蛋白胆固醇(HDL)及三酰甘油(TG)水平存在相关性,且为2型糖尿病及心血管疾病的潜在调控因子<sup>[10,27]</sup>。也有研究发现,GALNT2基因启动子的甲基化与急性脑梗死密切相关<sup>[28]</sup>。GALNT2可能通过调节靶蛋白如APOC3、低密度脂蛋白受体及极低密度脂蛋白的O-糖基化,影响了机体内脂质的代谢调节功能。还有研究发现,GALNT2还参与了胰岛素信号与脂肪形成过程<sup>[29]</sup>。此外,GALNT2在恶性肿瘤中的作用研究同样值得关注。GALNT2可通过降低胃癌患者EGFR、MET的磷酸化修饰活性抑制肿瘤的生长及转移<sup>[15]</sup>。肝细胞癌中GALNT2的mRNA表达下调,且与肿瘤的血管浸润及复发相关;在肝细胞癌细胞系中过表达

GALNT2可抑制EGF诱导的肿瘤细胞生长、迁移及侵袭,原因是GALNT2催化了EGFR的O-糖基化修饰,从而改变了EGF与EGFR的结合反应<sup>[30]</sup>。然而在其他一些肿瘤中,GALNT2具有相反的生物学功能,即能够促进肿瘤的发生发展。在神经胶质瘤中GALNT2表达水平升高,且与预后不良及肿瘤分级升高有关<sup>[13]</sup>;口腔鳞状细胞癌中GALNT2呈过表达,且同样与高侵袭性及高肿瘤级别相关,可通过修饰EGFR的O-糖基化调节受体活性,增强肿瘤细胞的侵袭能力<sup>[14]</sup>。有研究通过WGCNA和LASSO约束的Cox回归模型进行分析,结果显示GALNT2可作为肺癌风险预测模型中的关键基因<sup>[31]</sup>。综上所述,GALNT2在多种肿瘤中表达异常,功能各不相同,可作为潜在的肿瘤诊断标志物,但是在肺癌中GALNT2的研究十分有限。

本研究首先利用Oncomine数据库验证GALNT2是否在肺腺癌中有表达量和拷贝数的变化。Oncomine数据库是目前世界上最大的基因芯片数据平台,主要针对大多数癌症类型和相应的正常组织及各种癌症亚型进行差异表达分析。利用Oncomine数据库中多组数据集进行分析,可以有效避免由于样本量和取样问题导致的误差,增加结论的可信度。结果显示,在多组肺腺癌的研究中,GALNT2在肿瘤组均有表达量上调和拷贝数变异,可能在

肺腺癌的进展中发挥重要作用。随后的研究选取了TCGA作为数据来源, TCGA数据库是最大的癌症基因信息数据库, 为肿瘤研究提供了大量的基因组数据及临床相关数据, 可以使研究者进行进一步的数据挖掘和验证。本研究利用UALCAN和cBioportal平台对TCGA数据进行了可视化处理, 结果发现, GALNT2在肺腺癌中呈高表达, 但与肿瘤恶性程度及淋巴结转移等临床特征无相关性。在基因组层面, GALNT2基因变异率(10%)较高, 但在基因扩增的15例样本中, 仅2例出现了mRNA高表达, 说明GALNT2在转录层面的表达异常并非由基因组层面的扩增变异引起的, 其表达异常升高的原因仍需进一步探究。此外, 通过生存分析发现, 与GALNT2低表达组相比, GALNT2高表达组肺腺癌患者的OS和FP较短, 差异有统计学意义, 表明GALNT2的表达水平可以一定程度上预测肺腺癌患者的预后情况, 且其高表达是潜在的危险因素, 同时通过多因素Cox比例风险模型进行协变量校正后分析GALNT2表达水平与肺腺癌总生存时间的关系发现, GALNT2的表达水平增高是肺腺癌患者预后不佳的危险因素。此外, 本研究根据GALNT2在肺腺癌中的相互作用蛋白分析了其可能影响的信号通路及生物学过程, 结果发现, GALNT2调控网络参与的生物学过程有蛋白质丝氨酸/苏氨酸O型糖基化、细胞器微管运输及白介素-17信号通路等。多项研究结果表明, GALNT2可通过糖基化EGFR调控下游信号通路<sup>[13-15]</sup>, 但其在肺腺癌中的具体分子机制仍需进一步验证。

综上所述, 本研究通过对多个数据库中肺腺癌的数据进行深入挖掘, 发现GALNT2基因在肺腺癌组织中高表达, 且与肺腺癌的预后不良相关, 这为进一步研究GALNT2基因在肺腺癌中的作用和分子机制提供了依据和线索, 为肺腺癌的临床基因治疗及新型抗肿瘤药物的研发提供了理论基础。

### 【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] Wang LP. Progress in targeting and immunotherapy of non-small cell lung cancer[J]. *J Zhengzhou Univ (Med Sci)*, 2020, 55(2): 176-182. [王丽萍. 非小细胞肺癌的靶向和免疫治疗进展[J]. 郑州大学学报(医学版), 2020, 55(2): 176-182.]
- [3] Liu YL, Yu YF, Lu S. Research progress of positive ALK fusion gene non-small cell lung cancer[J]. *Med J Chin PLA*, 2019, 44(6): 527-535. [刘颖蕾, 虞永峰, 陆舜. ALK融合基因阳性非小细胞肺癌的临床研究进展[J]. 解放军医学杂志, 2019, 44(6): 527-535.]
- [4] Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer[J]. *Nature*, 2018, 553(7689): 446-454.
- [5] Chen Z, Fillmore CM, Hammerman PS, *et al.* Non-small-cell lung cancers: a heterogeneous set of diseases[J]. *Nat Rev Cancer*, 2014, 14(8): 535-546.
- [6] Gao S, Li N, Wang S, *et al.* Lung cancer in People's republic of China[J]. *J Thorac Oncol*, 2020, 15(10): 1567-1576.
- [7] Pinho SS, Reis CA. Glycosylation in cancer: mechanisms and clinical implications[J]. *Nat Rev Cancer*, 2015, 15(9): 540-555.
- [8] White T, Bennett EP, Takio K, *et al.* Purification and cDNA cloning of a human UDP-N-acetyl-alpha-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase[J]. *J Biol Chem*, 1995, 270(41): 24156-24165.
- [9] Tian E, Ten Hagen KG. Recent insights into the biological roles of mucin-type O-glycosylation[J]. *Glycoconj J*, 2009, 26(3): 325-334.
- [10] Kathiresan S, Melander O, Guiducci C, *et al.* Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans[J]. *Nat Genet*, 2008, 40(2): 189-197.
- [11] Schjoldager KT, Vakhrushev SY, Kong Y, *et al.* Probing isoform-specific functions of polypeptide GalNAc-transferases using zinc finger nuclease glycoengineered SimpleCells[J]. *Proc Natl Acad Sci U S A*, 2012, 109(25): 9893-9898.
- [12] Khetarpal SA, Schjoldager KT, Christoffersen C, *et al.* Loss of function of GALNT2 lowers high-density lipoproteins in humans, nonhuman primates, and rodents[J]. *Cell Metab*, 2016, 24(2): 234-245.
- [13] Sun Z, Xue H, Wei Y, *et al.* Mucin O-glycosylating enzyme GALNT2 facilitates the malignant character of glioma by activating the EGFR/PI3K/Akt/mTOR axis[J]. *Clin Sci (Lond)*, 2019, 133(10): 1167-1184.
- [14] Lin MC, Huang MJ, Liu CH, *et al.* GALNT2 enhances migration and invasion of oral squamous cell carcinoma by regulating EGFR glycosylation and activity[J]. *Oral Oncol*, 2014, 50(5): 478-484.
- [15] Hu WT, Yeh CC, Liu SY, *et al.* The O-glycosylating enzyme GALNT2 suppresses the malignancy of gastric adenocarcinoma by reducing EGFR activities[J]. *Am J Cancer Res*, 2018, 8(9): 1739-1751.
- [16] Yang SN, Li N, Wang LP, *et al.* Efficacy of icotinib combined with chemotherapy on advanced non-small cell lung cancer patients with EGFR mutations[J]. *J Zhengzhou Univ (Med Sci)*, 2020, 55(1): 53-56. [杨双宁, 李楠, 王丽萍, 等. 埃克替尼联合化疗治疗EGFR突变型晚期非小细胞肺癌的临床效果[J]. 郑州大学学报(医学版), 2020, 55(1): 53-56.]
- [17] Swanton C, Govindan R. Clinical implications of genomic discoveries in lung cancer[J]. *N Engl J Med*, 2016, 374(19): 1864-1873.
- [18] Dagogo-Jack I, Shaw AT. Crizotinib resistance: implications for therapeutic strategies[J]. *Ann Oncol*, 2016, 27(Suppl 3): iii42-iii50.
- [19] Devarakonda S, Morgensztern D, Govindan R. Genomic alterations in lung adenocarcinoma[J]. *Lancet Oncol*, 2015, 16(7): e342-e351.
- [20] Tran DT, Ten Hagen KG. Mucin-type O-glycosylation during development[J]. *J Biol Chem*, 2013, 288(10): 6921-6929.
- [21] Brooks SA, Carter TM, Bennett EP, *et al.* Immunolocalisation of

- members of the polypeptide N-acetylgalactosaminyl transferase (ppGalNAc-T) family is consistent with biologically relevant altered cell surface glycosylation in breast cancer[J]. *Acta Histochem*, 2007, 109(4): 273-284.
- [22] Burchell JM, Beatson R, Graham R, *et al.* O-linked mucin-type glycosylation in breast cancer[J]. *Biochem Soc Trans*, 2018, 46(4): 779-788.
- [23] Kudelka MR, Ju TZ, Heimburg-Molinaro J, *et al.* Simple sugars to complex disease: mucin-type O-glycans in cancer[J]. *Adv Cancer Res*, 2015, 126: 53-135.
- [24] Fuster MM, Esko JD. The sweet and sour of cancer: glycans as novel therapeutic targets[J]. *Nat Rev Cancer*, 2005, 5(7): 526-542.
- [25] He H, Shen Z, Zhang H, *et al.* Clinical significance of polypeptide N-acetylgalactosaminyl transferase-5 (GalNAc-T5) expression in patients with gastric cancer[J]. *Br J Cancer*, 2014, 110(8): 2021-2029.
- [26] Li Z, Yamada S, Wu Y, *et al.* Polypeptide N-acetylgalactosaminyltransferase-6 expression independently predicts poor overall survival in patients with lung adenocarcinoma after curative resection[J]. *Oncotarget*, 2016, 7(34): 54463-54473.
- [27] Willer CJ, Sanna S, Jackson AU, *et al.* Newly identified loci that influence lipid concentrations and risk of coronary artery disease[J]. *Nat Genet*, 2008, 40(2): 161-169.
- [28] Guo D, Liu D, Zhang GW. Correlation analysis of GALNT2 gene promoter region methylation and acute cerebral infarction[J]. *J Apoplexy Nerv Dis*, 2019, 36(2):136-139. [郭丹, 刘丹, 张广炜. GALNT2基因启动子区甲基化与急性脑梗死的相关分析[J]. *中风与神经疾病杂志*, 2019, 36(2): 136-139.]
- [29] Marucci A, Antonucci A, de Bonis C, *et al.* GALNT2 as a novel modulator of adipogenesis and adipocyte insulin signaling[J]. *Int J Obes (Lond)*, 2019, 43(12): 2448-2457.
- [30] Wu YM, Liu CH, Hu RH, *et al.* Mucin glycosylating enzyme GALNT2 regulates the malignant character of hepatocellular carcinoma by modifying the EGF receptor[J]. *Cancer Res*, 2011, 71(23): 7270-7279.
- [31] Wang H, Lu D, Liu XG, *et al.* Survival-related risk score of lung adenocarcinoma identified by weight gene co-expression network analysis[J]. *Oncol Lett*, 2019, 18(5): 4441-4448.

(收稿日期: 2021-01-25; 修回日期: 2021-05-17)

(责任编辑: 熊晓然)