

## 乳腺癌生存期相关长非编码MAPT-AS1及其共表达基因筛选与验证

许家瑞<sup>1</sup>, 刘冬冬<sup>1</sup>, 李贝贝<sup>1</sup>, 陈梦麟<sup>2</sup>, 汪湾<sup>1</sup>, 黄凯铃<sup>2</sup>, 郑周霞<sup>2</sup>, 徐建华<sup>3</sup>

(1. 广州中医药大学第二附属医院//广东省中医院检验医学部, 广东 广州 510120; 2. 广州万德基因医学科技有限公司, 广东 广州 510535; 3. 广州中医药大学顺德医院, 广东 佛山 528000)

**摘要:**【目的】筛选乳腺癌患者生存期相关差异表达长链非编码RNA(lncRNA)及其共表达基因,并验证其乳腺癌细胞中的表达情况。【方法】通过TCGA数据库筛选943例RNA-seq数据信息:(837例乳腺癌患者+106例正常对照),发现长非编码MAPT-AS1高表达,乳腺癌患者生存期更长。构建长非编码MAPT-AS1过表达和干扰质粒,将构建好的质粒转染乳腺癌细胞株T47D,并用嘌呤霉素筛选出稳定表达的T47D细胞株,通过RT-qPCR验证长非编码MAPT-AS1及其共表达基因的表达情况。【结果】荧光显微镜观察及RT-qPCR验证,成功构建长非编码MAPT-AS1过表达及干扰稳定转染乳腺癌细胞株T47D,并筛选出干扰效率最高的长非编码MAPT-AS1干扰片段shRNA3。验证其共表达基因得到MAPT、MAPT-IT1和NXNL2在转染了干扰片段shRNA3后表达量降低,与长非编码MAPT-AS1的表达趋势一致。【结论】成功构建长非编码MAPT-AS1过表达及干扰稳定转染乳腺癌细胞株T47D,验证其和共表达基因表达情况与数据库一致,为进一步研究长非编码MAPT-AS1基因在乳腺癌中的作用机制奠定基础。

**关键词:**乳腺癌;长链非编码RNA;共表达基因

**中图分类号:**R73 **文献标志码:**A **文章编号:**1672-3554(2019)03-0467-06

## Screening and Verification of Long Non-coding MAPT-AS1 and Their Co-expressed Genes in Breast Cancer Survival

XU Jia-rui<sup>1</sup>, LIU Dong-dong<sup>1</sup>, LI Bei-bei<sup>1</sup>, CHEN Meng-lin<sup>2</sup>, WANG Wan<sup>1</sup>,  
HUANG Kai-ling<sup>2</sup>, ZHENG Zhou-xia<sup>2</sup>, XU Jian-hua<sup>3</sup>

(1. Department of Laboratory Science, the Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou 510120, China; 2. Guangzhou Mendel Genomics and Medical Technology Co. Ltd, Guangzhou 510535, China; 3. Shunde Hospital of Guangzhou University of Chinese Medicine, Foshan 528000, China)

Correspondence to: XU Jian-hua, E-mail: jhXu1976@126.com

**Abstract:** 【Objective】To screen survival-related differential expression of long non-coding RNA (lncRNA) and its co-expressed genes in breast cancer patients and to verify their expression in breast cancer cells. 【Methods】RNA-seq data of 943 cases (837 breast cancer + 106 normal controls) by the TCGA database were screened, and found that long non-coding MAPT-AS1 highly expressed, and breast cancer patients had longer survival. The long non-coding MAPT-AS1 overexpression and interference plasmid was constructed, and the constructed plasmid was transfected into breast cancer cell line T47D, and the stably expressed T47D cell line was screened by puromycin. The expression of long non-coding MAPT-AS1 and its co-expressed genes was verified by the methods of RT-qPCR. 【Results】Fluorescence microscopy and RT-qPCR confirmed that the long non-coding MAPT-AS1 overexpression and interference-transfected breast cancer cell line T47D were successfully constructed, and the long non-coding MAPT-AS1 interference fragment shRNA3

收稿日期:2019-01-05

基金项目:广东省科技计划项目(2014A020212274);广州市科技计划项目(201510010044);广东省中医院科技专项(YN2016QJ10)

作者简介:许家瑞,在读研究生,研究方向:生物化学与分子生物学,E-mail:415474550@qq.com;徐建华,通信作者,博士研究生

导师,主任技师,研究方向:生物化学与分子生物学,E-mail:jhXu1976@126.com

with the highest interference efficiency was screened. The expression of *MAPT*, *MAPT-IT1* and *NXNL2* in the co-expressed gene was decreased after transfection of the shRNA3 interference fragment, which was consistent with the expression trend of the long non-coding *MAPT-AS1*. 【Conclusion】 The long non-coding *MAPT-AS1* overexpression and interference plasmid transfected breast cancer cell line T47D were successfully constructed, and the expression of the co-expressed gene was consistent with the database. The study laid the foundation for further study of the mechanism of action of long non-coding *MAPT-AS1* gene in breast cancer.

**Key words:** breast cancer; long non-coding RNA; co-expression gene

[J SUN Yat-sen Univ (Med Sci), 2019, 40(3):467-472]

乳腺癌是女性中最常见的一种恶性肿瘤,据美国癌症协会统计,乳腺癌发病率一直居女性各类恶性肿瘤之首,且呈逐年上升趋势,其死亡人数也在不断增加<sup>[1-3]</sup>,尽管手术、放化疗等治疗使乳腺癌患者的长期生存率有所提高,但乳腺癌患者的死亡率仍在升高。早期乳腺癌不具备典型症状和体征,不易引起病人重视,往往容易漏诊。因此,早期诊断乳腺癌对乳腺癌的治疗具有重要意义。近年来,随着分子生物学技术的不断发展,越来越多的特异性分子靶点被发现。长链非编码RNA (long non-coding RNA, lncRNA) 是一类不能编码蛋白质的RNA分子,转录本长度超过200个核苷酸。lncRNA 曾被认为不具有生物学功能,然而越来越多的研究发现,lncRNA 能在基因组印记、染色质修饰与重塑、转录激活与干扰和核内运输等生物学进程发挥重要调控作用<sup>[4]</sup>,并在各类癌症中起着重要的调节作用<sup>[5]</sup>。目前已有研究发现乳腺癌发生发展过程中,存在lncRNA 表达的异常改变,提示lncRNA 可作为乳腺癌潜在诊断标志物,这对乳腺癌药物选择以及治疗方法的寻找等方面具有重要意义<sup>[6-7]</sup>。本研究旨在发现乳腺癌生存期相关特异性表达lncRNA 及其共表达基因,并通过构建真核表达载体,使其在乳腺癌细胞T47D中稳定过表达和干扰,为进一步研究长非编码MAPT-AS1在乳腺癌发生发展中的作用机制奠定基础。

## 1 材料与方法

### 1.1 仪器与试剂

人乳腺癌细胞株T47D由广州医科大学附属肿瘤医院研究所邓敏教授实验室馈赠。RPMI1640培养基、胰蛋白酶购自美国Gibco公司;胎牛血

清购自Invitrogen公司;嘌呤霉素Puromycin购自Invivogen公司;逆转录试剂盒购自TAKARA公司;2X PowerUp SYBR Green Master Mix 购自Thermo公司;PCR引物由上海生工生物设计合成;过表达、干扰及对照质粒pCDH-CMV-MCS-EF1-copGFP、pLent-U6-GFP-puro 购自美国Vigene公司。Opti-MEM、转染试剂Lipofactamine™ 2000 购自Invitrogen公司;质粒提取试剂盒购自Biomiga公司。

### 1.2 数据库筛选差异表达lncRNA与共表达基因

课题组从癌症基因图谱(The Cancer Genome Atlas, TCGA)下载乳腺癌数据,包括转录组数据(RNA-seq)和临床生存数据;经单因素COX和Log Rank分析找到与生存期相关的lncRNA。通过lncRNA相关性分析,得到正相关的调控基因用于实验验证。引物设计见表1。

表1 lncRNA-MAPT-AS1及其靶基因引物序列  
Table 1 lncRNA-MAPT-AS1 and its target gene primer sequences

Primer	sequence(5'-3')
MAPT-AS1-Forward	CATGCTGGATTCTGAGCCACTCTC
MAPT-AS1-Reward	TGAAGAAGCCTGGACAGGAGGTC
MAPT-Forward	TGAAGAAGCAGGCATTGGAGACAC
MAPT-Reward	AGCCTTCCTGGACCACCTTACC
MAPT-IT1-Forward	CTCTTACATTCGAGCGAGTGACC
MAPT-IT1-Reward	TGAGCCTGGAGCCTGAGCAC
KDM4B-Forward	CGGAAGCGGATGAAGAAGGTGTC
KDM4B-Reward	AGGCAGGTGATGGAGACCACATAG
NXNL2-Forward	GCTGCTCTGCGACTTCTATACGG
NXNL2-Reward	GCTGTGACGTTGTACCTCTTCCTC
SPPL2C-Forward	GCAGAGTCTTCAAGCCATGAGAGG
SPPL2C-Reward	ACCACAATGTCACCGAAGCCAAG

### 1.3 细胞转染以及稳转细胞株的筛选

用含体积分数 10% 的胎牛血清的 RPMI1640 完全培养基于 37 °C、体积分数 5% 的 CO<sub>2</sub> 恒温培养箱中常规培养乳腺癌细胞 T47D; 当细胞处于对数生长期时用胰酶消化制成单细胞悬液, 按 2 mL/孔将细胞铺于 6 孔板中, 实验分为阴性对照(negative control, NC)组, 转染过表达空质粒 Empty vector 和干扰空质粒 shRNA-NC)组, 过表达组(转染重组过表达质粒 MAPT-AS1 vector)和 4 个干扰组(转染 shRNA1、shRNA2、shRNA3、shRNA4 干扰质粒)。质粒(2 μg)与转染试剂 Lipofactamine™ 2 000 (5 μL)进行转染, 并验证转染效率(质粒带绿色荧光), 转染后用最佳浓度嘌呤霉素(0.5 μg/mL)进

行筛选, 连续培养 1 周, 进行传代直至获得稳定细胞株, 此后用含 0.5 μg/mL 嘌呤霉素的完全培养基维持培养。

### 1.4 验证稳定转染细胞株内 lncRNA-MAPT-AS1 及其靶基因的表达水平

收集 7 组筛选后的稳转乳腺癌细胞, 用 Trizol 分别提取总 RNA, 按逆转录试剂盒说明书将总 RNA 逆转录为 cDNA, 然后以此 cDNA 为模板进行 PCR 扩增。PCR 反应体系为 20 μL, 反应条件设置如表 2。经罗氏 Light Cycler 480 荧光定量 PCR 仪检测, 以 GAPDH 为内参, 采用 2<sup>-ΔΔCt</sup> 法计算目的基因的相对表达量。每组样品重复 3 次, 试验重复 3 次。实验工作经伦理委员会批准。

表 2 PCR 反应条件  
Table 2 Reaction setting of PCR

Step	Temperature/°C	Time	Cycles
UDG enzyme activation	50	2 min	1
Dual-Lock™ DNA Polymerase Activation	95	2 min	1
Denaturation	95	15 s	40
Annealing and extension	60	1 min	
Melting curve analysis	95	15 s	1
	60	1 min	
	95	15 s	

### 1.5 统计学方法

数据分析采用 Graph Pad Prism 5 软件统计分析并作图, 计量资料采用均数±标准差( $\bar{x} \pm s$ )表示, RNA-seq 数据分析使用单因素 COX 和 Log Rank 分析, 各组之间比较用单因素方差分析, *P* 值均为双边检验, 以 *P* < 0.05 为差异有统计学意义的检验标准。

## 2 结果

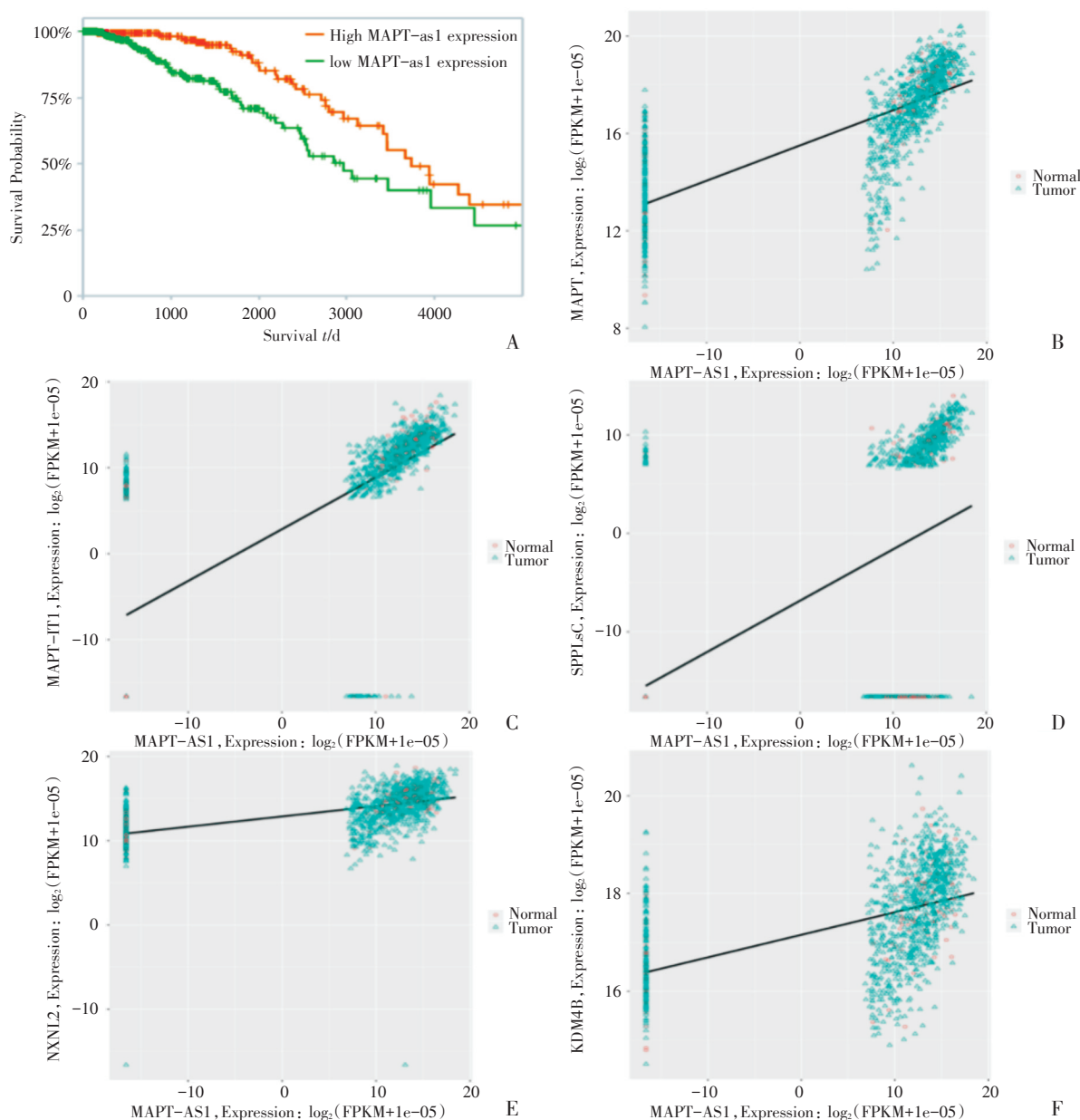
### 2.1 乳腺癌差异表达 lncRNA 与共表达基因筛选

我们分析了 943 例 RNA-seq 数据信息:(837 例乳腺癌患者 + 106 例正常对照), 经单因素 COX 和 Log Rank 分析找到与生存期相关的 lncRNA, 找到了显著性最高的长非编码 MAPT-AS1; 发现长非编码 MAPT-AS1 高表达, 预测乳腺癌患者生存期更长(图 1A)。通过 lncRNA 相关性分析, 找出正相关的调控基因共 14 个(表 3)。选择相关性最高的前 5 个基因(图 1B-F)用于实验验证。

表 3 MAPT-AS1 正相关调控基因  
Table 3 MAPT-AS1 positive correlation regulatory genes

Genes	Correlation	<i>P</i> value
MAPT	0.859	9.26E-239
MAPT-IT1	0.856	1.61E-235
SPPL2C	0.735	8.66E-140
NXNL2	0.654	1.10E-100
KDM4B	0.642	5.40E-96
PCR	0.634	3.15E-93
TMEM26	0.619	1.22E-87
PLAC9P1	0.613	1.31E-85
SUSD3	0.593	1.02E-78
GREB1	0.591	4.32E-78
PREX1	0.59	9.54E-78
FGD3	0.589	2.22E-77
STC2	0.58	1.73E-74
NEK10	0.579	1.99E-74

E-n in *P* value: 10<sup>-n</sup>



A: Survival analysis: MAPT-AS1 is highly expressed and has a long survival period; B:  $x = \text{MAPT-AS1}, y = \text{MAPT}$ ; C:  $x = \text{MAPT-AS1}, y = \text{MAPT-IT1}$  (a.k.a. LOC100130148); D:  $x = \text{MAPT-AS1}, y = \text{SPPL2C}$  (a.k.a. IMP5); E:  $x = \text{MAPT-AS1}, y = \text{NXNL2}$ ; F:  $x = \text{MAPT-AS1}, y = \text{KDM4B}$ .

图1 MAPT-AS1与共表达基因

Fig.1 MAPT-AS1 and co-expressed genes

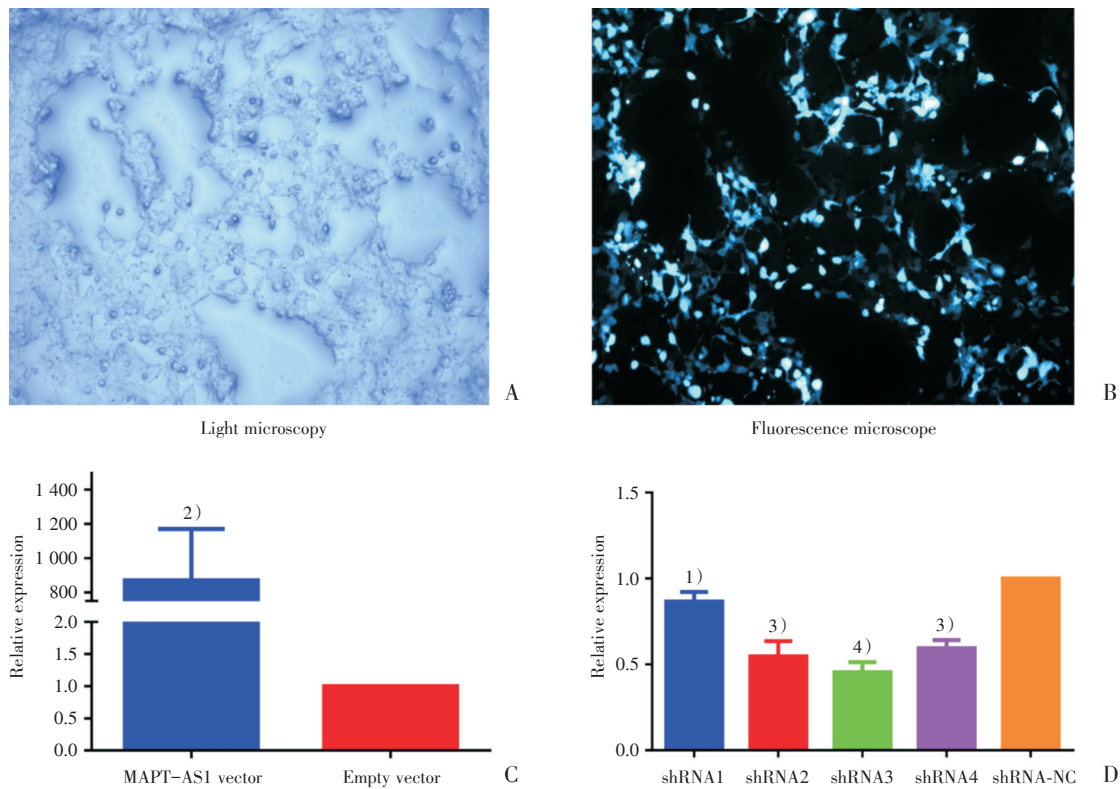
## 2.2 转染效率验证及长非编码MAPT-AS1干扰片段筛选

转染效率如图2所示,可见转染效率达90%以上,以此条件继续后续实验。转染过表达长非编码MAPT-AS1的稳转细胞株内长非编码MAPT-AS1的相对表达水平上升,差异有统计学意义( $P < 0.001$ ;图2C),表明筛选的细胞株为稳定过表达长非编码MAPT-AS1的乳腺癌细胞株T47D。转染不同干扰片段后长非编码MAPT-AS1

的相对表达水平下降,差异有统计学意义( $P$ 值均 $< 0.05$ ;图2D),结果显示shRNA3干扰效果最好。

## 2.3 稳定转染细胞株内长非编码MAPT-AS1共表达基因的表达水平

本研究选取预测结果中表达量最高的5个共表达基因: *MAPT* (microtubule associated protein tau)、*MAPT-IT1* (MAPT intronic transcript 1)、*SP-PL2C* (signal peptide peptidase like 2C)、*KDM4B* (lysine demethylase 4B)、*NXNL2* (nucleoredoxin like 2)



A: bright field of vision; B: fluorescence field of vision. Breast cancer cell line T47D transfected with lncRNA-MAPT-AS1 verification (100×). C: Transfection overexpression level of lncRNA-MAPT-AS1; D: Expression level of long non-coding MAPT-AS1 after transfection of different interference fragments. 1)  $P < 0.05$ , 2)  $P < 0.01$ , 3)  $P < 0.001$ , 4)  $P < 0.0001$  vs empty vector or shNNA-NC.

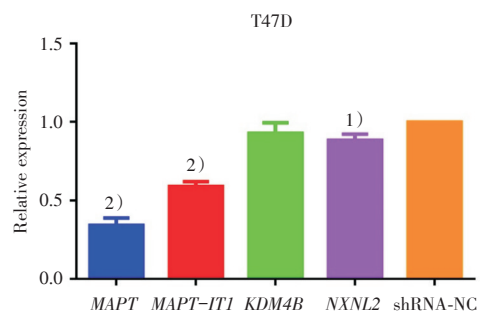
图2 乳腺癌细胞株 T47D 转染长非编码 MAPT-AS1 效率验证及其过表达和干扰的相对表达量

Fig.2 Efficiency verification and overexpression and interference relative expression of breast cancer cell line T47D transfected with lncRNA-MAPT-AS1

进行验证,在乳腺癌细胞 T47D 中预先转染 MAPT-AS1 干扰片段 shRNA3 和干扰对照 shRNA-NC,再进行 RT-qPCR 实验验证。结果显示: *MAPT*、*MAPT-IT1* 和 *NXNL2* 在转染了 shRNA3 干扰片段后相对表达量降低,差异有统计学意义( $P$  值均  $< 0.05$ ,图 3),与长非编码 MAPT-AS1 的表达趋势一致,*SPPL2C* 未见表达。

### 3 讨论

本研究通过 TCGA 数据库筛选乳腺癌患者差异表达 lncRNA,经单因素 COX 和 Log Rank 分析找到与生存期相关的 lncRNA,发现了显著性最高的长非编码 MAPT-AS1。通过 lncRNA 相关性分析,找出正相关的调控基因,选择相关性最高的前 5 个基因进行实验验证。荧光显微镜可见转染效率达 90% 以上,RT-qPCR 验证,成功构建长非编码 MAPT-AS1 过表达及干扰稳定转染乳腺癌细胞株



1)  $P < 0.01$ , 2)  $P < 0.0001$  vs. shNNA-NC.

图3 长非编码 MAPT-AS1 共表达基因表达水平  
Fig.3 lncRNA-MAPT-AS1 co-expression gene expression level

T47D,并筛选出干扰效率最高的长非编码 MAPT-AS1 干扰片段 sh-RNA3。RT-qPCR 验证长非编码 MAPT-AS1 共表达基因,得到 *MAPT*、*MAPT-IT1* 和 *NXNL2* 在转染了 shRNA3 干扰片段后表达量降低,与长非编码 MAPT-AS1 的表达趋势一致,结果显示,长非编码 MAPT-AS1 在乳腺癌患者中高表

达,可能具有致癌作用。

近年来,研究者发现 MAPT 表达增加可能与帕金森病(PD)疾病状态有关,而 MAPT-AS1 和 DNA 甲基转移酶(DNMT)为 MAPT 的潜在表观遗传调节因子<sup>[8]</sup>; MAPT-AS1、SPPL2C、MAPT、MAPT-IT1、CRHR1、MGC57346、CRHR1-IT1、STH 和 KANSLI 这些基因中的一个或多个的单倍不足可能是导致 Koolen de Vries 综合征(KDVS; MIM 610443)表型的原因<sup>[9]</sup>。MAPT-AS1 过表达可部分保护 MAPT 不被降解,而 MAPT-AS1 敲低可降低 MAPT 的表达,同时也降低了 MAPT 的稳定性。研究发现 MAPT-AS1 在年龄较小(<60),肿瘤较大( $\geq 2$  cm),转移淋巴结和分期(III-IV)的 ER 阴性患者中表达较高,并与细胞生长、侵袭性和紫杉醇抗性相关,可作为 ER 阴性乳腺癌的潜在治疗靶点<sup>[10]</sup>。MAPT 变体可能与家族性和散发性肌萎缩侧索硬化(ALS)

和额颞叶变性(FTLD)的发生相关<sup>[11]</sup>。系统性硬化症(SSc)相关性间质性肺病(ILD)与特发性间质性肺炎(IIP)中肺部受累的表型相似,预测用力肺活量百分比(FVC%)与 SPP2C 中的 SNP rs2076295、rs17690703 和 MAP19 中的 rs1981997 相关<sup>[12]</sup>。KDM4B 在 TGF- $\beta$  诱导的胰腺癌上皮-间质转化(EMT)过程中表观遗传地调节 ZEB1 的表达,是 EMT 过程中的关键介质。在胰腺癌患者的肿瘤组织中,KDM4B 的蛋白水平与 ZEB1 呈正相关,可作为人胰腺癌转移进展的重要预后标志物和治疗靶点<sup>[13]</sup>。

研究发现 MAPT-AS1 及其共表达基因 MAPT、MAPT-IT1、SPPL2C、KDM4B、NXNL2 在许多疾病的发生发展中均具有一定意义,但他们在乳腺癌分子水平上的作用机制研究仍不明确,进一步研究和证实他们在乳腺癌中的功能机制,或许能成为乳腺癌诊断治疗、生存预后的重要新靶点。

#### 参考文献

- [1] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018[J]. CA-Cancer J Clin, 2018, 68(1): 7-30.
- [2] Siegel RL, Miller KD, Dvm A J. Cancer statistics, 2017[J]. CA-Cancer J Clin, 2017, 67(1): 7-30.
- [3] Siegel RL, Miller KD, Jemal A. Cancer statistics in China, 2016[J]. CA-Cancer J Clin, 2016, 66(1): 7-30.
- [4] Bunch H. Gene regulation of mammalian long non-coding RNA [J]. Mol Genet Genomics, 2017, 89(4): 110-119.
- [5] Wei GH, Wang X. lncRNA MEG3 inhibit proliferation and metastasis of gastric cancer via p53 signaling pathway[J]. Eur Rev Med Pharmacol Sci, 2017, 21(17): 3850-3856.
- [6] Jiang YZ, Liu YR, Xu XE, et al. Transcriptome analysis of triple-negative breast cancer reveals an integrated mRNA-lncRNA signature with predictive and prognostic value[J]. Cancer Res, 2016, 76(8): 2105-2114.
- [7] Feng T, Shao F, Wu Q, et al. miR-124 downregulation leads to breast cancer progression via lncRNA-MALAT1 regulation and CDK4/E2F1 signal activation[J]. Oncotarget, 2016, 7(13): 16205-16216.
- [8] Coupland KG, Kim WS, Halliday GM, et al. Role of the long non-coding RNA MAPT-AS1 in regulation of microtubule associated protein Tau (MAPT) expression in Parkinson's disease [J]. PLoS one, 2016, 11(6): e0157924.
- [9] Nascimento GR, Pinto IP, De Melo AV, et al. Molecular characterization of koolen De Vries syndrome in two girls with idiopathic intellectual disability from central brazil[J]. Mol Syndromol, 2017, 8(3): 155-160.
- [10] Pan Y, Pan Y, Cheng Y, et al. Knockdown of lncRNA MAPT-AS1 inhibits proliferation and migration and sensitizes cancer cells to paclitaxel by regulating MAPT expression in ER-negative breast cancers[J]. Cell Biosci, 2018, 8(1): 7.
- [11] King A, Al-Sarraj S, Troakes C, et al. Mixed tau, T DP-43 and p62 pathology in FTLD associated with a C9ORF72 repeat expansion and p.Ala239 ThrMAPT(tau) variant[J]. Acta Neuropathologica, 2013, 125(2): 303-310.
- [12] Wu M, Assassi S, Salazar GA, et al. Genetic susceptibility loci of idiopathic interstitial pneumonia do not represent risk for systemic sclerosis: a case control study in Caucasian patients [J]. Arthr Res Ther, 2016, 18(1): 20.
- [13] Li S, Wu L, Wang Q, et al. KDM4B promotes epithelial-mesenchymal transition through up-regulation of ZEB1 in pancreatic cancer [J]. Acta Biochimicaet Biophysica Sinica, 2015, 47(12): 997-1004.

(编辑 刘清海)