

# 生物信息学辅助研究乳腺癌转移相关 lncRNA 进展

柴慈曼 综述, 杨绍时 审校

(天津医科大学第二医院甲状腺乳腺外科, 天津 300211)

**摘要** 长链非编码 RNA 在肿瘤进展中的研究日益增多, 尤其随着二代测序技术的出现和广泛应用, 大量数据涌现推动了生物信息学的迅猛发展, 为进一步阐明乳腺癌相关 lncRNA 作用机制提供了很大的帮助。阐述乳腺癌转移相关 lncRNA 研究现状, 可为进一步研究 lncRNA 在乳腺癌转移中的作用机制、寻找治疗新靶点提供思路。

**关键词** 长链非编码 RNA; 乳腺癌; 生物信息学; 转移

**中图分类号** R737.9

**文献标志码** A

根据国家癌症中心最新公布的数据, 中国每年新发乳腺癌患者约 26.86 万人, 死亡 6.95 万人<sup>[1]</sup>。绝大部分乳腺癌患者死亡是由肿瘤转移导致重要脏器功能衰竭造成的, 如何攻克这个难题关系到乳腺癌患者的生存质量和生存时间。因此针对乳腺癌的转移问题, 寻求有效的诊疗和预后指标也成为近年来该领域的研究热点。随着对长链非编码 RNA(lncRNA)作用的研究深入, 其在肿瘤发生、发展中的调节作用也日趋明确。目前已经有实验证明, 很多 lncRNA 与乳腺癌的转移关系密切, 也已经涌现出一批用于 lncRNA 的鉴定、定量、结构分析以及功能预测的生物信息学工具和数据库。本文将利用这些生物信息学工具总结乳腺癌进展相关 lncRNA 的研究现状。

## 1 lncRNA 概述

随着人类基因组计划的完成, 人们发现人类基因组中仅有 1.5% 负责编码蛋白的基因, 其余 98.5% 的非编码调控元件能转录产生大量非编码 RNA(ncRNA)<sup>[2]</sup>。依据 ncRNA 所含碱基的数量, 可分为短链非编码 RNA(sncRNA)、中链非编码 RNA(mncRNA)及 lncRNA。lncRNA 被定义为长度大于 200 个核苷酸的不具有编码蛋白质功能的 RNA 序列。有大量实验证明, lncRNA 的异常表达与多种肿瘤的发生、发展密切相关<sup>[3-5]</sup>。lncRNA 能够在转录、转录后和染色体修饰等多个层面调控胚胎发育、细胞增殖、凋亡、转移和分化等各种生命活动, 故而近年来 lncRNA 一直是肿瘤领域研究的明星分子。

在目前各类研究中, lncRNAs 有很多种分类方法, 多种名称重叠可能会造成交流上的误解, 故而 2015 年 St. Laurent 等<sup>[6]</sup>对 lncRNA 各种分类方式进行了全面综述。

目前对于 lncRNA 的研究多采用染色质免疫共沉淀技术收集 lncRNA, 用微阵列(microarrays)、平铺阵列(tiling arrays)、转录组测序技术(RNA-seq)发现差异表达并鉴定<sup>[3]</sup>。随着生物信息学的迅猛发展, 将新一代高通量测序技术与生物计算方法联合使用, 出现了大量 lncRNA 测序信息并因此建立了很多相关生物信息数据库, 使得对 lncRNA 的研究更加便捷。目前已有的 lncRNA 数据库包括 LncRNAdb、NONCODE、Rfam、LncRNAdisease、CuiLab、UCSC、GENCODE、NRED、lncipedia 等, 这些数据库收录了 lncRNA 序列及其相关信息, 各常用 lncRNA 数据库特点简要介绍见表 1。

## 2 乳腺癌转移相关 lncRNA

研究表明, 很多 lncRNA 广泛参与了正常组织和疾病组织的生物学功能。目前已在乳腺癌细胞及组织中发现多种异常表达 lncRNA, 它们在乳腺癌细胞增殖、凋亡、侵袭、转移及药物敏感性等方面可能起了重要作用<sup>[7-9]</sup>。目前有很多生物数据库储存了与疾病相关的 lncRNA, 如 CuiLab 数据库或 LncRNAdisease 数据库等。

临床上有大量患者死于乳腺癌的复发和转移, 因此明确乳腺癌进展相关机制, 对于提高患者生存率和生活质量非常有价值。肿瘤转移是一个多阶段、多步骤的复杂过程: 肿瘤细胞脱离原发灶组织进入循环系统, 沿循环系统到达靶器官, 最终形成远端转移。在这个过程中, 上皮-EMT 被认为是整个肿瘤转移的起始步骤, 肿瘤细胞发生 EMT, 脱离原发灶组织, 获得运动能力, 然后通过 EMT 的可逆过程——MET 在靶器官定植并形成远端转移灶<sup>[10-11]</sup>。乳腺癌转移还与其他多种因素相关, 涉及肿瘤细胞自身生长优势的获得<sup>[12]</sup>、细胞间黏附能力的降低<sup>[13]</sup>、细胞因子诱导的癌细胞器官特异性定向归巢<sup>[14]</sup>、肿瘤干细胞<sup>[15]</sup>以及肿瘤血管和淋巴管生成<sup>[16]</sup>、癌细胞逃

**基金项目** 天津医科大学第二医院青年基金项目(2017ydey05)

**作者简介** 柴慈曼(1985~), 女, 主治医师, 博士, 研究方向: 乳腺癌内分泌耐药机制; 通信作者: 杨绍时, E-mail: m15122901226@163.com。

逸凋亡和免疫杀伤<sup>[17-18]</sup>等诸多因素。目前已有实验证实 lncRNA 通过表观修饰调节、转录及转录后水平调节、翻译水平调节等多种途径调控乳腺癌转移,但其作用机制仍不明确。生物信息数据库中存储着大量可能与乳腺癌进展相关的候选 lncRNA 信息,目前已经有不少研究在使用这些信息。

通过 lncRNA 生物数据库(表 1)获得乳腺癌转移相关 lncRNA 见表 2。

在 lncRNA 数据库中选择乳腺癌相关芯片信息,经过数据归一化后,在线统计分析软件或者使用 R 软件分析下载后的 lncRNA 信息,对比乳腺癌转移组与未转移组中差异基因,筛选出与乳腺癌转

移相关的 lncRNA 并制表。

### 3 生物信息学推动研究 lncRNA 在乳腺癌转移中的作用机制

lncRNA 几乎参与了肿瘤转移相关基因表达和调控的全过程,包括染色质重构、转录调控、转录后加工,见图 1<sup>[9]</sup>。目前的研究多为 DNA 甲基化、组蛋白修饰、基因组印迹、随机染色体失活及染色体重构、EMT 等方向<sup>[20-22]</sup>。已有研究发现,lncRNA 可与蛋白质结合,引起染色体重构、影响转录因子功能,还可在线性水平与 microRNA 结合,间接影响 mRNA 表达,也可以直接与 mRNA 结合,影响 mRNA 翻译、剪切、降解等过程<sup>[3]</sup>。

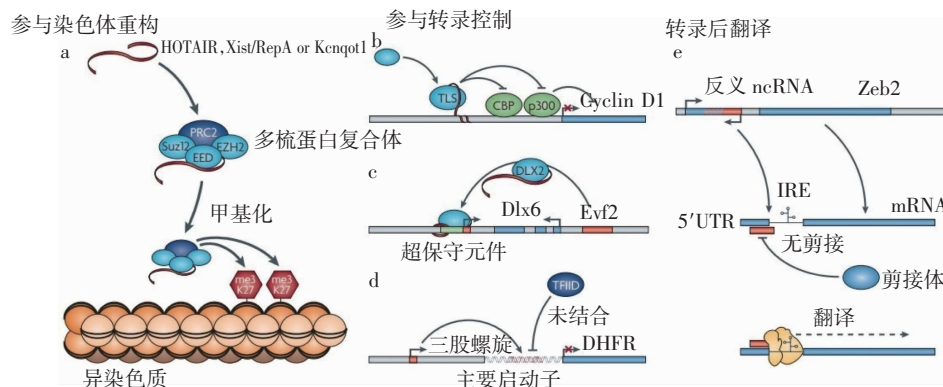
表 1 lncRNA 数据库汇总

数据库名称	简要特点	网址
ChiPBase	来自 ChiP-seq 数据的转录因子对 lncRNA 和 miRNA 的调控信息	http://Deeppbase.sysu.edu.cn/chipbase
deepBase	利用 RNA-seq 数据鉴定 lncRNA,提供表达谱及功能预测及进化保守性分析	http://Biocenter.sysu.edu.cn/deeppbase
LncRNAdisease	与人类疾病相关的 lncRNA	http://210.73.221.6/LncRNAdisease
NONCODE	16 个物种中 lncRNA 的位置与序列、表达谱、进化保守性、功能注释、相关疾病等信息	http://www.noncode.org
Lncipedia	除人类 lncRNA 一级序列外,还提供二级结构信息,用核糖体测序数据对 lncRNA 的蛋白编码潜能进行评估	http://www.lncipedia.org
UCSC	基于人类基因组计划,包括脊椎动物和模型生物的 lncRNA 信息	http://Genome.ucsc.edu/cgi-bin/hgGateway
starbase	Clip-seq 数据中获得高通量的 miRNA-lncRNA 靶向关系和 lncRNA-RNA 结合蛋白的互作关系,TCGA 的 lncRNA 表达谱数据及 ceRNA 调控网络	http://Starbase.sysu.edu.cn
DIANA-LncBase	提供基于 2 个研究的 CLIP-Seq 数据 miRNA 调控 lncRNA 的信息	http://www.microna.gr/LncBase
LncRNAdb <sup>[5]</sup>	提供有生物学功能的 lncRNA 的全面注释	http://LncRNAdb.org
Cuilab	提供疾病相关 lncRNA 及序列结果	http://www.cuilab.cn/lncmadisease

表 2 乳腺癌转移相关 lncRNA

名称	Genbank 号	作用机制
BCAR4	NR_024049.1	BCAR4 高水平表达与低 MFS(远端无转移生存)及总生存率相关,是肿瘤侵袭性的反应 <sup>[7]</sup>
BCYRN1	NR_001568.1	在原位导管癌中,BC200(BCYRN1)显著高表达与高级别相关,可以用作肿瘤进展的预后指示物
H19	NR_002196.42	H19 的下调可以抑制乳腺癌和肺癌细胞的集落和附壁生长;c-Myc 和 H19 在原发性乳腺癌和肺癌中关系密切
HOTAIR	NR_047517.1	通过 LSD1/CoREST&PRC2 沉默基因表达;转移相关
Loc554202	-	调节乳腺癌细胞的增殖和迁移
MALAT1	NR_002819.6	选择性剪接的调节剂
SRA1	NM_001035235.3	类固醇受体和其他转录因子的共激活因子;与转移相关

注:BCAR4:乳腺癌抗雌激素抗性基因 4;BCYRN1:脑胞质 RNA1;HOTAIR:HOX 转录反义基因间 RNA;MALAT1:肺腺癌转移相关转录本 1;SRA1:类固醇受体 RNA 激活因子



注:HOTAIR:HOX 转录反义基因间 RNA;Xist:智人 X 无效特异性转录本;PRC2:多克隆抑制复合物 2;me3k27:三甲基化赖氨酸;TLS:RNA 结合蛋白;CBP:CREB 结合蛋白;Evf2:Dlx5/6 超保守增强子;TFHD:转录因子;DHFR:二氢叶酸还原酶基因;Zeb2:锌指同源盒 mRNA;IRE:内含子内部核糖体进入位点

图 1 lncRNA 的功能

在乳腺癌疾病进展中 lncRNA 起重要作用,但目前为止其具体作用机制和途径尚未知晓。乳腺癌晚期转移患者迫切需要明确 lncRNA 的作用机制,从而对转移性乳腺癌的治疗靶点提供帮助。本文从基因表达的几个重要阶段总结了 lncRNA 的作用机制。

**3.1 染色体重构** lncRNA 可以通过招募染色质重构复合体到特定位置进而介导相关基因的表达沉默,如 HOTAIR。在原发性乳腺癌与转移性乳腺癌中,HOTAIR 的过表达预示了不良预后。HOTAIR 过表达将引起 PRC2 复合体在全基因组范围内的重新定位,使若干抑癌基因沉默,终将促进乳腺癌的转移。HOTAIR 通过隔离 miRNA-449b-5p,促进 HSPA1A 表达,从而增强了乳腺癌的放射抵抗力<sup>[23]</sup>,也通过介导雌激素受体+乳腺癌细胞中的 HOXA1 表达,促进他莫昔芬耐药<sup>[24]</sup>。在转移性乳腺癌患者中,HOTAIR 的表达量可作为预测乳腺癌转移及预后的有效指标。

**3.2 转录调控** 一些 lncRNA 分子被转录后,可以调控下游基因转录。如启动子甲基化影响下游基因。miRNA-31 可通过作用于某些促转移基因,从而抑制乳腺癌转移<sup>[25-26]</sup>。miRNA-31 在转移能力较强的三阴性乳腺癌细胞系中明显低表达<sup>[25]</sup>。miRNA-31 的转录活性受 LOC554202 的调节<sup>[27]</sup>。LOC554202 在三阴性乳腺癌细胞系中的表达明显降低;而在 luminal 细胞系中,其表达明显升高。加入甲基化抑制剂 5Aza2Cd 后,LOC554202 及 miRNA-31 在三阴性乳腺癌中的表达明显上调。利用甲基化特异性 PCR 及亚硫酸氢盐测序法等发现,LOC554202 启动子相关的 CpG 岛在三阴性乳腺癌和 luminal 型乳腺癌细胞系中具有两种完全不同的状态,在三阴性细胞系中被高度甲基化,而在 luminal 亚型中被低甲基化。

HOTAIR 通过募集 PRC2 连接至同源异型盒位点 *hoxD*,引起 H3 组蛋白的第 27 位赖氨酸甲基化,再通过促进一系列靶蛋白,如连接黏附分子 2(JAM2)、细胞黏附分子原钙黏素(PCDH)、转录因子同源盒蛋白 D10(HOXD10)等,从而促进细胞迁移、侵袭和肿瘤转移<sup>[28]</sup>。

类固醇受体 RNA 激活因子(SRA)的表达与乳腺癌相关<sup>[29-30]</sup>。SRA 可以调节类固醇受体和其他转录因子的 RNA 表达水平,以及调节 SRA 的表达,因此可能参与了性激素的致癌作用。已发现 SRA 与其他基因通过协同作用增强细胞侵袭能力,抑制 SRA 的表达后,细胞侵袭能力明显减弱。

**3.3 影响 EMT** lncRNA 在 EMT 中的作用已部分阐明,研究发现大量 lncRNA 在 EMT 中表达上调、

下调,但与 EMT 具体作用机制尚未明确。

lncRNA ATB 可以促进曲妥珠单抗克隆抗体(单抗)耐药产生以及乳腺癌侵袭转移的发生。曲妥珠单抗耐药是造成 HER2+乳腺癌患者死亡率增加的原因,其作为转化生长因子,在曲妥珠单抗耐药机制中起到诱导 EMT 的作用<sup>[31]</sup>。但是其中牵涉的 lncRNA 仍未知。研究发现,曲妥珠单抗耐药的乳腺癌细胞与亲代相比,侵袭力明显增强,而且在这些细胞中转化生长因子- $\beta$  信号表达显著增加。在 TR SKBR-3 细胞系和 TR 乳腺癌组织中,大部分的 lncRNA 显著高表达。研究发现,lncRNA ATB 通过与 miRNA-200c 结合上调 ZEB1 和 ZNF-217 的表达,从而导致 EMT,促进曲妥珠单抗耐药及乳腺癌侵袭转移。而且研究还发现,lncRNA ATB 高表达与乳腺癌患者曲妥珠单抗耐药相互关联。因此,这些发现表明,lncRNA ATB 是一种转化生长因子- $\beta$  的媒介,使乳腺癌患者容易发生 EMT 和曲妥珠单抗耐药<sup>[31]</sup>,提示预后不良<sup>[32]</sup>。

HOTAIR、XIST、MALAT、H19 已被证实与肿瘤干细胞可塑性相关<sup>[33]</sup>。肿瘤干细胞通路包含 HOX 基因,而且已经证实一些 lncRNA 与 HOX 基因表达相关。SOX2OT 是一种已知的 lncRNA,可以诱导 SOX2 的表达,后者是干细胞分化多样性的要素<sup>[34]</sup>。乳腺癌与较高的 SOX2OT 表达有关,后者可作为乳腺癌的一种新的检测指标,未来可能成为乳腺癌患者的潜在治疗靶点。

## 4 结语

目前,lncRNA 的研究还需要大量的数据和有效分析来支持,lncRNA 在乳腺癌等相关恶性肿瘤的进程中发挥重要作用。随着高质量的数据库的建立,继续完善研究手段及系统分析步骤,可有助于对乳腺癌的预测、早期诊断及靶向药物的研发。

## 参考文献:

- [1] Chen W, Zheng R, Baade P D, et al. Cancer statistics in China, 2015[J]. CA Cancer J Clin, 2016, 66(2): 115
- [2] Balakirev E S, Ayala F J. Pseudogenes: are they "junk" or functional DNA[J]. Annu Rev Genet, 2003, 37: 123
- [3] Wu P, Mo Y, Peng M, et al. Emerging role of tumor-related functional peptides encoded by lncRNA and circRNA[J]. Mol Cancer, 2020, 19(1): 22
- [4] Wu M Z, Fu T, Chen J X, et al. lncRNA GOLGA2P10 is induced by PERK/ATF4/CHOP signaling and protects tumor cells from ER stress-induced apoptosis by regulating Bcl-2 family members[J]. Cell Death Dis, 2020, 11(4): 276
- [5] Huang D, Chen J, Yang L, et al. NKILA lncRNA promotes tumor immune evasion by sensitizing T cells to activation-induced cell death[J]. Nat Immunol, 2018, 19(10): 1112
- [6] St Laurent G, Wahlestedt C, Kapranov P. The Landscape of long



- noncoding RNA classification[J]. Trends Genet, 2015, 31(5): 239
- [7] Jin X, Ge L P, Li D Q, et al. LncRNA TROJAN promotes proliferation and resistance to CDK4/6 inhibitor via CDK2 transcriptional activation in ER+ breast cancer[J]. Mol Cancer, 2020, 19(1): 87
- [8] Liang Y, Song X, Li Y, et al. LncRNA BCRT1 promotes breast cancer progression by targeting miR-1303/PTBP3 axis[J]. Mol Cancer, 2020, 19(1): 85
- [9] Luo L, Zhang J, Tang H, et al. LncRNA SNORD3A specifically sensitizes breast cancer cells to 5-FU by sponging miR-185-5p to enhance UMPS expression[J]. Cell Death Dis, 2020, 11(5): 329
- [10] Wu H T, Zhong H T, Li G W, et al. Oncogenic functions of the EMT-related transcription factor ZEB1 in breast cancer[J]. J Transl Med, 2020, 18(1): 51
- [11] Huang W, Liu C, Liu F, et al. Hinokiflavone induces apoptosis and inhibits migration of breast cancer cells via EMT signalling pathway[J]. Cell Biochem Funct, 2020, 38(3): 249
- [12] Tiedemann K, Hussein O, Komarova S V. Role of altered metabolic microenvironment in osteolytic metastasis[J]. Front Cell Dev Biol, 2020, 8: 435
- [13] Mendonsa A M, Bandyopadhyay C, Gumbiner B M. p120-catenin phosphorylation status alters E-cadherin mediated cell adhesion and ability of tumor cells to metastasize[J]. PLoS One, 2020, 15(6): e0235337
- [14] Meyer-Schaller N, Tiede S, Ivanek R, et al. A dual role of Irfl in maintaining epithelial identity but also enabling EMT and metastasis formation of breast cancer cells[J]. Oncogene, 2020, 39(24): 4728
- [15] Yi T, Kabha E, Papadopoulos E, et al. 4EGI-1 targets breast cancer stem cells by selective inhibition of translation that persists in CSC maintenance, proliferation and metastasis[J]. Oncotarget, 2014, 5(15): 6028
- [16] Huang R, Li J, Pan F, et al. The activation of GPER inhibits cells proliferation, invasion and EMT of triple-negative breast cancer via CD151/miR-199a-3p bio-axis[J]. Am J Transl Res, 2020, 12(1): 32
- [17] Li J T, Wang L F, Zhao Y L, et al. Nuclear factor of activated T cells 5 maintained by Hotair suppression of miR-568 upregulates S100 calcium binding protein A4 to promote breast cancer metastasis[J]. Breast Cancer Res, 2014, 16(5): 454
- [18] Lin C W, Xie J, Zhang D, et al. Immunity against cancer cells may promote their proliferation and metastasis[J]. Proc Natl Acad Sci U S A, 2020, 117(1): 426
- [19] Li X, Wu Z, Fu X, et al. LncRNAs: insights into their function and mechanics in underlying disorders[J]. Mutat Res Rev Mutat Res, 2014, 762: 1
- [20] Qu T, Zhao Y, Chen Y, et al. Down-regulated MAC30 expression inhibits breast cancer cell invasion and EMT by suppressing Wnt/ $\beta$ -catenin and PI3K/Akt signaling pathways[J]. Int J Clin Exp Pathol, 2019, 12(5): 1888
- [21] Deng F, Weng Y, Li X, et al. Overexpression of IL-8 promotes cell migration via PI3K-Akt signaling pathway and EMT in triple-negative breast cancer[J]. Pathol Res Pract, 2020, 216(4): 152902
- [22] Xu M, Xu X, Pan B, et al. LncRNA SATB2-AS1 inhibits tumor metastasis and affects the tumor immune cell microenvironment in colorectal cancer by regulating SATB2[J]. Mol Cancer, 2019, 18(1): 135
- [23] Zhang S, Wang B, Xiao H, et al. LncRNA HOTAIR enhances breast cancer radioresistance through facilitating HSPA1A expression via sequestering miR-449b-5p[J]. Thorac Cancer, 2020, 11(7): 18016
- [24] Kim C Y, Oh J H, Lee J Y, et al. The LncRNA HOTAIRM1 promotes tamoxifen resistance by mediating HOXA1 expression in ER+ breast cancer cells[J]. J Cancer, 2020, 11(12): 341
- [25] Luo L J, Yang F, Ding J J, et al. MiR-31 inhibits migration and invasion by targeting SATB2 in triple negative breast cancer[J]. Gene, 2016, 594(1): 47
- [26] Soheilyfar S, Velashjerdi Z, Sayed Hajizadeh Y, et al. *In vivo* and *in vitro* impact of miR-31 and miR-143 on the suppression of metastasis and invasion in breast cancer[J]. J BUON, 2018, 23(5): 1290
- [27] Augoff K, McCue B, Plow E F, et al. miR-31 and its host gene lncRNA LOC554202 are regulated by promoter hypermethylation in triple-negative breast cancer[J]. Mol Cancer, 2012, 11: 5
- [28] Lu L, Zhu G, Zhang C, et al. Association of large noncoding RNA HOTAIR expression and its downstream intergenic CpG island methylation with survival in breast cancer[J]. Breast Cancer Res Treat, 2012, 136(3): 875
- [29] Cooper C, Guo J, Yan Y, et al. Increasing the relative expression of endogenous non-coding Steroid Receptor RNA Activator (SRA) in human breast cancer cells using modified oligonucleotides[J]. Nucleic Acids Res, 2009, 37(13): 4518
- [30] Arshi A, Sharifi F S, Khorramian Ghahfarokhi M, et al. Expression analysis of MALAT1, GAS5, SRA, and NEAT1 lncRNAs in breast cancer tissues from young women and women over 45 years of age[J]. Mol Ther Nucleic Acids, 2018, 12: 751
- [31] Shi S J, Wang L J, Yu B, et al. LncRNA-ATB promotes trastuzumab resistance and invasion-metastasis cascade in breast cancer[J]. Oncotarget, 2015, 6(13): 11652
- [32] Li R H, Chen M, Liu J, et al. Long noncoding RNA ATB promotes the epithelial-mesenchymal transition by upregulating the miR-200c/Twist1 axis and predicts poor prognosis in breast cancer[J]. Cell Death Dis, 2018, 9(12): 1171
- [33] Deng J, Yang M, Jiang R, et al. Long Non-coding RNA HOTAIR regulates the proliferation, self-renewal capacity, tumor formation and migration of the cancer stem-like cell (CSC) subpopulation enriched from breast cancer cells[J]. PLoS One, 2017, 12(1): e0170860
- [34] Dehghani-Samani M, Hassanzadeh N, Kabiri H, et al. Correlations between overexpression of SOX2OT long non-coding RNA and susceptibility to breast cancer[J]. Comb Chem High Throughput Screen, 2020, [Online ahead of print]

(2020-05-16 收稿)