

· 论著 ·

长链非编码 RNA 在乳腺癌患者外周血与癌组织中的表达

吕明明¹ 许楠^{1,2} 王凤良¹ 陈飞¹ 陆迅³ 陆澄^{1,2}

【摘要】 目的 比较分析长链非编码 RNA(lncRNA)在乳腺癌组织与外周血中的表达情况,筛选出参与乳腺癌发生、复发及转移过程中重要的 lncRNA。**方法** 利用芯片技术分别检测并筛选乳腺癌组织与外周血差异表达的 lncRNA 及 mRNA,应用生物信息学方法,包括基因本体论(gene ontology, GO)和 Pathway 分析,分析乳腺癌组织与外周血中 lncRNA 及 mRNA 的表达谱。分析得到乳腺癌外周血与癌组织中共同表达与特异表达的 lncRNA,并且进一步分析差异基因可能调控的相关通路,进而发现参与乳腺癌发生发展和转移过程的 lncRNAs。**结果** 笔者发现在乳腺癌组织和外周血中共同上调的 lncRNA 有 7 条(包括 RP11-707G14.7、RP11-771K4.1、XLOC_013658、RP11-201A3.1、RP11-298H24.1、BM151951、XLOC_011403),共同下调的 lncRNA 有 12 条(LOC100506035、PWRN1、MAGI2-AS3、RP11-67L3.2、CTC-454M9.1、AC017006.2、AC002456.2、RP11-430B1.2、RP11-945C19.4、RP11-839D17.3、CTA-407F11.8、LINC00420)。GO 和 Pathway 分析结果表明在乳腺癌患者外周血和组织中均差异表达的基因主要参与调控免疫反应与 TGF- β 通路等。**结论** lncRNA 在乳腺癌组织和外周血中呈差异表达,可能参与到乳腺癌的发生、发展以及复发转移中,有可能成为临床诊断与治疗乳腺癌的新靶点。

【关键词】 乳腺肿瘤; 分子诊断技术; 芯片分析技术; 基因表达谱; 微 RNAs

【中图分类号】 R737.9

【文献标志码】 A

Expression of long non-coding RNA in peripheral blood and carcinoma tissue of breast cancer patients

Lyu Mingming¹, Xu Nan^{1,2}, Wang Fengliang¹, Chen Fei¹, Lu Xun³, Lu Cheng^{1,2}. ¹ Department of Breast Surgery, Nanjing Maternity and Child Health Care Hospital Affiliated to Nanjing Medical University, Nanjing 210004, China; ² First Clinical Medicine College, Nanjing University of Chinese Medicine, Nanjing 210046, China; ³ Jinling High School, Nanjing 210005, China

Corresponding author: Lu Cheng, Email: lucheng66@126.com

【Abstract】 Objective To compare the expression of long non-coding RNA (lncRNA) in both peripheral blood and carcinoma tissues of breast cancer patients and reveal the important lncRNAs that may participating in the development, recurrence and metastasis of breast cancer. **Methods** Microarray chip was used to detect and screen the differentially expressed lncRNA and mRNA in carcinoma tissue and peripheral blood of breast cancer patients. lncRNA and mRNA expression profiles were analyzed by bioinformatics tools, including gene ontology (GO) and Pathway analysis. The differentially expressed and commonly expressed lncRNAs both in peripheral blood and carcinoma tissues were detected. Moreover, we further analyzed the relevant pathways regulated by differentially expressed genes, as well as lncRNAs participating in the development and metastasis of breast cancer. **Results** There were 7 up-regulated lncRNAs (including RP11-707G14.7, RP11-771K4.1, XLOC_013658, RP11-201A3.1, RP11-298H24.1, BM151951, XLOC_011403) and 12 down-regulated lncRNAs in both carcinoma tissue and peripheral blood (including LOC100506035, PWRN1, MAGI2-AS3, RP11-67L3.2, CTC-454M9.1, AC017006.2, AC002456.2, RP11-430B1.2, RP11-945C19.4, RP11-839D17.3, CTA-407F11.8, LINC00420). GO and Pathway analysis indicated that

DOI:10.3877/cma.j.issn.1674-0807.2015.04.004

基金项目:国家自然科学基金资助项目(81172501);南京市医学科技发展项目(YKK12103)

作者单位:210004 南京医科大学附属南京市妇幼保健院乳腺病科¹;210046 南京中医药大学第一临床医学院²;210005 南京金陵中学³

通信作者:陆澄,Email:lucheng66@126.com

differentially expressed genes in both peripheral blood and carcinoma tissues of breast cancer patients were mainly involved in the regulation of immune response and TGF- β pathway. **Conclusion** lncRNAs were differentially expressed genes in both peripheral blood and carcinoma tissues of breast cancer patients, which may be involved in the development, recurrence and metastasis of breast cancer, regarded as possible potential targets for clinical diagnosis and treatment.

【Key words】 Breast neoplasms; Molecular diagnostic techniques; Microchip analytical procedures; Gene expression profiling; MicroRNAs

乳腺癌是中国女性最常见的恶性肿瘤,在全世界新诊断乳腺癌病例约占 12.2%,约占全世界癌症死亡的 9.6%^[1]。自 1990 年以来,中国乳腺癌的发病率增长为过去的两倍,而乳腺癌检出率相对较低,中晚期患者的治愈率低^[2]。因此,寻求有效的诊断和预后指标为近年来研究的重点^[3]。乳腺癌治疗后复发转移是患者死亡的主要原因,且远处转移患者的病死率显著高于局部复发的患者。

近些年来,对肿瘤发生、发展机制的认识越来越深入,已经从功能基因层面(编码蛋白质)逐渐拓展到非编码 RNA^[4]。非编码 RNA 领域中的 microRNA 已经被证明在乳腺癌中发挥重要作用^[5-6]。长链非编码 RNA(long non-coding RNA, lncRNA)一类转录本长度超过 200 nt 的长链非编码 RNA,因其重要的调控功能以及复杂的作用机制,成为近年肿瘤研究的新热点,已被证实其与多种肿瘤的发生、侵袭转移及耐药有关,有望成为肿瘤诊疗的新靶标^[7-8]。

癌细胞的脱落、侵袭并进入血液循环是实现肿瘤转移的重要一步,乳腺癌外周血微转移与远处转移存在相关性^[9]。笔者前期研究发现 lncRNA 在乳腺癌患者的组织中异常表达^[10],随后在外周血中也发现了游离的 lncRNA,其表达水平有异于正常人。乳腺癌组织与外周血中异常表达的 lncRNA,是否存在关联?是否存在特异的 lncRNA 在乳腺癌组织与外周血中均异常表达,从而对乳腺癌的发生发展、复发转移发挥重要的调控作用?笔者利用生物信息学方法分析比较了乳腺癌组织与外周血中异常表达的 lncRNAs,对于筛选出的差异基因进行生物信息学分析,从而为深入研究 lncRNA 在乳腺癌中的作用提供理论依据。

1 资料和方法

1.1 样本的选取

所入选的乳腺癌病例为 2013 年 9 月至 2014 年

3 月期间南京市妇幼保健院收治的 112 例患者,均为女性,且有完整的临床超声乳腺 X 线摄影检查资料和病理资料。随机数字表法选取 3 例乳腺癌患者的癌组织与癌旁组织,用芯片对组织中 lncRNA 的表达进行检测。以同期体检的正常人外周血为对照组,同时用芯片对乳腺癌患者外周血中 lncRNA 的表达进行检测,正常对照组与乳腺癌组年龄差异无统计学意义[(46.3 \pm 10.4)岁比(48.0 \pm 7.5)岁, $t=0.225$, $P=0.833$]。

1.2 差异 lncRNA 的筛选及交集分析

应用 t 检验筛选两个实验组之间差异表达的 lncRNA,计算 lncRNA 差异的显著性水平(P -value)和误判率(false discovery rate, FDR)从而得到差异表达的 lncRNA,并以邻近的 mRNA 进行注释。分别筛选出乳腺癌组织与外周血中的差异基因后,对组织和外周血中差异倍数 ≥ 3 的差异基因取交集。对取交集得到的 lncRNA 进行生物信息学分析,主要包括基因本体论(gene ontology, GO)和 Pathway 分析。

1.3 GO 分析

GO 分析是基因功能国际标准分类体系^[11],分为分子功能(molecular function)、生物学过程(biological process)和细胞组分(cellular component)三部分。使用生物学软件 TopGO 进行差异基因的 GO 分析。

1.4 Pathway 分析

根据挑选出的差异基因的 Pathway 分析^[12],可以找到富集差异基因的 Pathway 条目,寻找不同样品的差异基因可能和哪些细胞通路的改变有关。Pathway 分析对实验结果有提示的作用,计算这些差异基因同 Pathway 的关系,Pathway 分析会对每个有差异基因存在的通路返回一个 P 值, P 值越小表示差异基因在该通路中出现的富集程度越高。

1.5 统计分析

结果用 $\bar{x} \pm s$ 表示,采用 SPSS 20.0 统计软件对

数据进行 t 检验。 $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 乳腺癌组织与外周血中差异表达 lncRNA 的交集分析

分别筛选出乳腺癌组织和血液中差异倍数 ≥ 3 的 lncRNA, 取交集分析, 并将交集后的基因进行注释和功能分析。在乳腺癌组织中, 上调的 lncRNA (1 637 条) 多于下调的 lncRNA (1 143 条); 而在外周血中, 表达量下调的 lncRNA (330 条) 远多于上调的 lncRNA (83 条)。通过分析, 笔者发现乳腺癌组织与外周血中表达共同上调的 lncRNA 有 7 条, 而表达共同下调的 lncRNA 有 12 条, 见表 1~2。

2.2 乳腺癌组织与外周血中差异表达 mRNA 的交集分析

分别筛选出乳腺癌组织和血液中差异倍数 ≥ 3 倍以上的 mRNA, 取交集分析, 并将交集后的基因进行注释和功能分析。在乳腺癌组织中, 上调的 mRNA (1 379 条) 多于下调的 mRNA (851 条); 而在外周血中, 表达量下调的 mRNA (725 条) 远

多于上调的 mRNA (216 条)。乳腺癌组织与外周血中表达量共同上调的 mRNA 有 10 条, 而表达量共同下调的 mRNA 有 20 条。

2.3 差异基因的 GO 分析

组织和外周血中差异基因交集的 GO 分析结果如图 1。乳腺癌组织与外周血中共同的差异基因主要参与的生物过程包括分泌过程、受体分解代谢过程、细胞分泌过程以及 TGF- β 通路等。如图 1 所示, 乳腺癌组织与外周血中同时上调或下调的差异基因涉及的细胞组分与分子功能较多, 主要参与的生物过程部分集中在免疫学方面, 例如 IL-17 以及炎症反应等方面; 差异基因涉及的细胞组分以及分子功能还包括核糖体以及受体(死亡受体和肿瘤坏死因子受体超家族)的结合。

2.4 差异基因的 Pathway 分析

组织和外周血中共同变化的差异基因主要集中在参与免疫反应、TGF- β 、病毒致癌以及癌症相关的蛋白多糖等通路中(图 2)。系统性红斑狼疮信号通路主要是 T 和 B 细胞所参与的细胞与体液免疫, 两者在乳腺癌发生发展过程起重要作用。

表 1 乳腺癌组织与外周血中共同上调的长链非编码 RNA

序列名	长链非编码 RNA	RNA 长度(nt)	染色体部位	正/反义链(+/-)	序列起始位点	序列结束位点
ENST00000535936	RP11-707G14.7	383	chr11	+	73639577	73641417
ENST00000541749	RP11-771K4.1	868	chr12	-	31516414	31522235
TCONS_00028623	XLOC_013658	652	chr20	-	12915531	12916654
ENST00000569873	RP11-201A3.1	2 277	chr1	+	188187530	188189807
ENST00000424701	RP11-298H24.1	551	chr10	+	107900048	107917903
BM151951	BM151951	330	chr3	+	72792571	72792897
TCONS_00024045	XLOC_011403	211	chr15	-	24587892	24588655

表 2 乳腺癌组织与外周血中共同下调的长链非编码 RNA

序列名	长链非编码 RNA	RNA 长度(nt)	染色体部位	正/反义链(+/-)	序列起始位点	序列结束位点
NR_038826	LOC100506035	2 054	chr4	+	80413746	80497614
ENST00000569908	PWRN1	1 435	chr15	+	24803304	24804739
NR_038346	MAGI2-AS3	901	chr7	+	79082272	79100524
ENST00000450503	RP11-67L3.2	343	chr1	+	55258557	55260578
ENST00000514571	CTC-454M9.1	594	chr5	+	88261691	88464485
ENST00000444077	AC017006.2	464	chr2	-	46305153	46305967
ENST00000412669	AC002456.2	449	chr7	-	90219936	90226667
ENST00000560518	RP11-430B1.2	1 972	chr15	+	52472413	52498071
ENST00000582939	RP11-945C19.4	620	chr18	+	5887485	5889462
ENST00000533504	RP11-839D17.3	517	chr11	-	113150104	113185159
ENST00000453811	CTA-407F11.8	451	chr22	-	25959067	25960226
ENST00000443554	LINC00420	1 059	chr13	-	19982181	19983669



注:a 图所示上调基因参与的生物学过程分析;b 图所示上调基因涉及细胞组分分析;c 图所示上调基因参与的分子功能分析;d 图所示下调基因参与的生物学过程分析;e 图所示下调基因涉及细胞组分分析;f 图所示下调基因参与的分子功能分析

图 1 乳腺癌差异基因 GO 分析

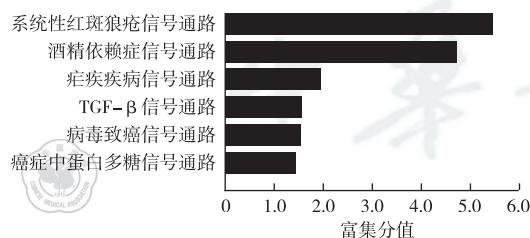


图 2 乳腺癌差异基因 Pathway 分析

3 讨论

在中国乳腺癌发病率一直在攀升,统计数据显示在过去 10 年,发病率增加了将近 1 倍,且发病趋势年轻化,成为目前威胁女性健康和生命的主要疾病^[13]。乳腺癌的早期诊断是提高患者生存率的一项重要因素。研究乳腺癌的分子机制,探索新型肿瘤标志物是目前乳腺癌研究领域中的热点之一^[14]。

研究证实,以 miRNA 和 lncRNA 为代表的非编码 RNA 在癌症发生、发展中扮演着重要角色^[15]。目前,miRNA 作为肿瘤标志物已经得到较为广泛的认同,然而,lncRNA 作为新型肿瘤标志物的研究尚处于起步阶段。lncRNA 异常表达于多种肿瘤中,起着致癌或抑癌作用,是调控肿瘤发生发展过程的一类重要因素。目前已在组织和体液中检测到与肿瘤相关的 lncRNA,如结肠癌中异常表达的 lncRNA-422^[16],前列腺癌中异常表达的基因 PCAT-1^[17],基因 HULC 在胰腺癌中异常表达并且可以作为诊断标志物^[18],胃癌患者的癌组织与血浆中均可检测到基因 FER1L4 的异常表达。而在乳腺癌中,很少有研究系统地比较分析乳腺癌组织与外周血中 lncRNA 的表达。本研究分别筛选了乳腺癌组织与外周血中差异表达的 lncRNA,对筛选出的差异 lncRNA 取交集并进一

步对其进行了 GO 和 Pathway 的生物信息学分析。

笔者发现在乳腺癌组织与外周血中有 7 条 lncRNA 表达上调, 12 条 lncRNA 表达下调。GO 分析和 Pathway 分析显示差异基因主要参与了免疫反应、TGF- β 通路等。TGF- β 是一种多向性、多效应的细胞因子, 以自分泌或旁分泌的方式通过细胞表面的受体信号转导途径调节细胞的增殖、分化和凋亡, 其信号通路在胚胎的早期发育、免疫炎症反应、肿瘤发生以及机体代谢平衡中都扮演着非常重要的角色^[19]。研究 TGF- β 信号通路在肿瘤中调控的分子机制具有重要的意义。笔者发现部分 lncRNA 在乳腺癌组织与外周血中均异常表达, 但其在乳腺癌发生发展中具体的作用机制仍需要进一步研究。

早期诊断、手术后放疗、化疗、内分泌治疗和分子靶向治疗等综合手段的应用, 使乳腺癌病死率呈下降趋势, 而中国乳腺癌发病率逐年上升, 寻找乳腺癌早期诊断的分子标志物十分必要。相对于肿瘤组织而言, 外周血中的 lncRNA 作为诊断标志物更有优势。外周血作为乳腺癌远处转移的重要一站, 组织与外周血中共同差异表达的 lncRNA, 对乳腺癌的转移可能发挥重要的调控作用。因此, lncRNA 作为新型肿瘤标志物, 在肿瘤诊断与治疗中有较好的应用前景, 本研究可能为乳腺癌的治疗和诊断提供理论依据。

参 考 文 献

- [1] Ferlay J, Shin HR, Bray F, et al. Cancer incidence and mortality worldwide; IARC Cancer Base No. 10. GLOBOCAN 2008 [M]. Lyon: International Agency for Research on Cancer, 2010.
- [2] Fan L, Strasser-Weippl K, Li JJ, et al. Breast cancer in China [J]. Lancet Oncol, 2014, 15(7):e279-289.
- [3] Dos Anjos Pultz B, da Luz FA, de Faria PR, et al. Far beyond the usual biomarkers in breast cancer: a review[J]. J Cancer, 2014, 5(7):559-571.
- [4] ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome[J]. Nature, 2012, 489(7414), 57-74.
- [5] Tang J, Ahmad A, Sarkar FH. The role of microRNAs in breast cancer migration, invasion and metastasis[J]. Int J Mol Sci, 2012, 13(10):13414-13437.
- [6] Bergamaschi A, Katzenellenbogen BS. Tamoxifen downregulation of miR-451 increases 14-3-3 ζ and promotes breast cancer cell survival and endocrine resistance [J]. Oncogene, 2012, 31(1):39-47.
- [7] Spizzo R, Almeida MI, Colombatti A, et al. Long non-coding RNAs and cancer: a new frontier of translational research? [J]. Oncogene, 2012, 31(43):4577-4587.
- [8] Gupta RA, Shah N, Wang KC, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis[J]. Nature, 2010, 464(7291):1071-1076.
- [9] 高纪东, 王军, 张保宁, 等. 乳腺癌外周血微转移与远处转移的相关性[J]. 癌症, 2007, 26(12):1385-1387.
- [10] Xu N, Wang F, Lv M, et al. Microarray expression profile analysis of long non-coding RNAs in human breast cancer: a study of Chinese women[J]. Biomed Pharmacother, 2015, 69: 221-227.
- [11] Gene Ontology Consortium. The Gene Ontology (GO) project in 2006[J]. Nucleic Acids Res, 2006, 34(Database issue): D322-326.
- [12] Curtis RK, Oresic M, Vidal-Puig A. Pathways to the analysis of microarrays data[J]. Trends Biotechnol, 2005, 23(8):429-435.
- [13] Dong C, Chen L. Second malignancies after breast cancer: the impact of adjuvant therapy[J]. Mol Clin Oncol, 2014, 2(3): 331-336.
- [14] Kalia M. Biomarkers for personalized oncology: recent advances and future challenges[J]. Metabolism, 2015, 64(3 Suppl 1): S16-21.
- [15] Ye S, Yang L, Zhao X, et al. Bioinformatics method to predict two regulation mechanism: TF-miRNA-mRNA and lncRNA-miRNA-mRNA in pancreatic cancer [J]. Cell Biochem Biophys, 2014, 70(3):1849-1858.
- [16] Xue Y, Ma G, Gu D, et al. Genome-wide analysis of long noncoding RNA signature in human colorectal cancer [J]. Gene, 2015, 556(2):227-234.
- [17] Prensner JR, Chen W, Han S, et al. The long non-coding RNA PCAT-1 promotes prostate cancer cell proliferation through cMyc[J]. Neoplasia, 2014, 16(11):900-908.
- [18] Peng W, Gao W, Feng J. Long noncoding RNA HULC is a novel biomarker of poor prognosis in patients with pancreatic cancer[J]. Med Oncol, 2014, 31(12):346.
- [19] Principe DR, Doll JA, Bauer J, et al. TGF- β : duality of function between tumor prevention and carcinogenesis [J]. J Natl Cancer Inst, 2014, 106(2):djt369.

(收稿日期:2015-01-22)

(本文编辑:刘军兰)

吕明明, 许楠, 王凤良, 等. 长链非编码 RNA 在乳腺癌患者外周血与癌组织中的表达[J/CD]. 中华乳腺病杂志: 电子版, 2015, 9(4):242-246.