

# 乳腺癌转移相关基因的文献计量分析

谢 菲, 黄东兰, 岑东芝, 张积仁<sup>△</sup>

(南方医科大学附属珠江医院肿瘤中心, 广州 510282)

**摘要:**目的 了解近 10 年来乳腺癌转移相关基因的研究概况, 寻找未来的研究方向。方法 文献检索数据库为 Embase、Medline/Pubmed 及 BIOSIS Preview 3 个数据库。文献检索年限为 2001 年 1 月至 2010 年 12 月, 对最终纳入的 363 篇文献的出版年、作者、国家、研究机构、发表期刊及基因进行计量学分析。结果 近 10 年乳腺癌转移基因相关文献在 2001~2007 年间呈缓慢增长趋势, 2008~2010 年开始呈现明显上升趋势。发文量最大的国家、期刊、研究机构分别为中国、Breast Cancer Research and Treatment、天津医科大学。论文中涉及的乳腺癌转移相关基因共计 567 个, 研究最多的基因为 ERBB2。结论 近 10 年间对乳腺癌转移相关基因的研究逐渐增多, 中国在该研究领域也有一定的优势。采用多基因模型评估乳腺癌转移风险将是未来研究的趋势。

**关键词:** 乳腺肿瘤; 转移相关基因; 计量学分析

doi:10.3969/j.issn.1671-8348.2012.05.002

文献标识码: A

文章编号: 1671-8348(2012)05-0420-03

## A bibliometric analysis of breast cancer metastatic genes

Xie Fei, Huang Donglan, Cen Dongzhi, Zhang Jiren<sup>△</sup>

(Oncology Center, Affiliated Zhujiang Hospital, Southern Medical University, Guangzhou 510282, China)

**Abstract: Objective** To comprehend the current study status of breast cancer metastatic genes in order to find the direction for further research. **Methods** We searched Embase, Medline/Pubmed and Biosis Preview database from January, 2001 to December, 2010. 363 included papers on breast cancer prognostic genes were analyzed on publication year, authors, country, research institute, reported journal and gene with bibliometric approaches. **Results** The identified essays grew slowly during 2001—2007, showed obviously increasing trend from 2008 to 2010. The country, institution and journal that published most essays were China, Tianjin Medical University and Breast Cancer Research and Treatment. The identified essays covered 567 genes, of which ERBB gene won the highest frequency. **Conclusion** The research essays on breast cancer prognostic genes are increasing. China in this field has some advantages. Using the multi-gene model to assess risk of breast cancer will be the trend of further research with the development of gene chip technology.

**Key words:** breast neoplasms; metastatic genes; bibliometrics

乳腺癌是女性最常见的恶性肿瘤, 其发病率占女性肿瘤的 20% 以上, 随着年龄的增加而增加, 而且早期即有远处转移, 其死亡原因几乎都是远处转移所致。乳腺癌的转移是一个较为复杂、由多基因参与及多步骤完成的过程。与肿瘤转移密切相关的基因被认为是抗转移干预治疗的潜在靶点, 深入阐述乳腺癌转移的机制将为预测和阻止乳腺癌转移提供坚实的基础。目前临床上病理和免疫组化指标不能准确判断乳腺癌患者转移的风险<sup>[1]</sup>, 严重影响肿瘤患者治疗的疗效和预后。为了了解目前乳腺癌转移相关基因的研究概况及发展趋势, 本文对国际上 2001~2010 年 10 年间乳腺癌转移相关基因研究的文献进行了计量学分析。

## 1 资料与方法

**1.1 一般资料** 数据库的选择以荷兰 Elsevier Science 出版公司建立的 Embase 数据库、美国国立医学图书馆建立的 Medline/Pubmed 数据库及美国生物科学信息服务社建立的世界最大的有关生命科学的文摘和索引数据库——BIOSIS Preview 数据库为数据源, 统计这 3 个数据库 2001 年 1 月至 2010 年 12 月收录的有关乳腺癌转移相关基因的文献, 选择这

3 个收录范围既有交叉又有区别的权威数据库来确保数据的完整性及准确性。

## 1.2 方法

**1.2.1 检索方法** 为提高文献的查准率, 对 Embase 数据库进行 Emtree 扩展检索, 辅以自由词的题目及摘要字段的检索; 对 Pubmed/Medline 数据库进行 MeSH 检索, 辅以自由词的题目及摘要字段的检索; 对 BIOSIS Preview 数据库进行主题检索。拟定的检索词为乳腺癌、转移及基因。乳腺癌的检索词包括: (1) MeSH 词 breast neoplasm; (2) Emtree 词 breast cancer; (3) 自由词检索使用 breast cancer、breast tumor \*、breast neoplasm \*、breast carcinoma \*、mammary neoplasm \*。转移的英文检索词包括: (1) MeSH 词 neoplasm metastasis; (2) Emtree 词检索使用 metastasis、cancer recurrence、cancer relapse 及 metastasis inhibition; (3) 自由词检索使用 metastasis、metastases。基因的英文检索词为 gene。

**1.2.2 文献纳入与排除** 文献形式主要为论著、会议论文及会议摘要, 纳入全部与乳腺癌转移相关基因及基因产物相关且以人类为研究对象的临床研究, 能提供足够信息的摘要或可以

<sup>△</sup> 通讯作者, Tel: (020) 61643200; E-mail: zhangjiren@126.com。

获取全文资料的文献。排除了综述、个案报道、回信、书籍、社论、无法获取摘要的文献、与转移相关的非基因研究、针对基因的结构和机制等研究、细胞株体外研究、老鼠及其他动物体内试验及重复文献与其他与研究目的不符的文献。纳入与排除的过程,由 2 名研究人员独立进行,最终结果互相比较,不一致的地方互相讨论,或向临床医生、实验研究人员(研究员)请教。

**1.2.3 文献计量分析** 采用 NoteExpress X3 软件对所检索文献进行管理,使用 Excel 2007 软件对最终纳入的文献的出版年、作者、国家、研究机构、发表期刊及基因进行计量学分析。

**2 结果**

**2.1 检索结果** Embase 数据库检出文献 2 246 篇,其中论著 1 573 篇,综述 353 篇,会议论文 105 篇,会议摘要 77 篇,信件 35 篇,笔记 33 篇,社论 31 篇。Pubmed/Medline 数据库检出文献 1 009 篇,其中综述 126 篇。BIOSIS Preview 数据库检出文献 1 441 篇,其中论著 974 篇,会议论文 440 篇,书籍目录 22 篇,信件 9 篇。三库共检出文献 4 696 篇。根据前述的文献纳入排除标准,最终纳入研究的文献共有 363 篇。

**2.2 文献发表年限分析** 对最终纳入的 363 篇文献进行发表年限分析,其年度发表数量及变化见图 1。从各年度的文献量分布上可以看出,近 10 年乳腺癌转移相关基因的临床研究在 2001~2007 年间呈缓慢增长趋势,研究处于萌芽阶段,2008 年总体开始呈现明显上升趋势,2009~2010 年间呈平稳发展状态,说明科研人员对该领域的研究越来越多,目前研究正处于稳定发展的阶段。由此推测,今后知识研究文献仍将会保持一定的增长速度,进一步达到研究的成熟阶段。

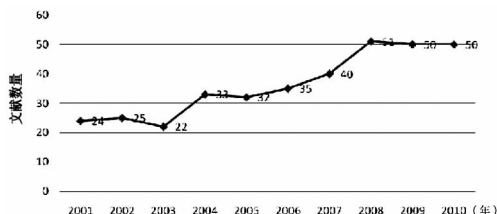


图 1 乳腺癌转移相关基因文献年份分布

**2.3 文献发文章国别分析** 通过对 363 篇文献的发表国家的统计分析得知,共有 45 个国家发表了关于乳腺癌转移相关基因方面的研究。发文章量前 20 位的国家见图 2。可以发现,中国发文章量为 84 篇,位列第 1 位,为发文章量最多的国家,占据了文献总数的 23.1%,遥遥领先于其他国家。美国继中国之后,发文章量为 65 篇,占据了文献总数的 17.9%,德国、日本、意大利、英国、法国、波兰、韩国及土耳其分列 3~10 位。以上表明,中国目前是研究乳腺癌转移相关基因最活跃的国家,体现了我国在该领域研究的先进地位和较高的科研能力,其他传统型生物学领域研究大国也是这一领域的主要力量。发文章量前 10 名的国家中,除中国外,均为高收入国家。除中国和美国外,在发文章量前 20 名的国家中,半数均为欧洲国家,说明欧洲在该领域研究中也占据了重要地位。从学术生产力方面分析<sup>[2]</sup>,由于我国人口基数大,人均发文章量与美、德、日等国相距甚远,因此我国的研究者应加大对该领域研究投入,吸取国外优秀科研经验,加强与国外权威科研机构合作,从而提高我国的学术生产力,增强我国的学术影响力。

**2.4 文献期刊分析** 对最终纳入的 363 篇论文进行统计分

析,共发表在 140 种生物医学期刊上,平均每刊发文 2.59 篇。发文章量达 8 篇的期刊有 9 种,根据 2009 年公布的 SCI 期刊目录,上述期刊均为 SCI 收录期刊,其发文章量及影响因子见表 1,其中 Cancer Research 的影响因子最高。根据布拉德福分散定律<sup>[4]</sup>,从发文章量的角度,Breast Cancer Research and Treatment 是该研究领域核心的英文期刊。通过查阅以上 9 种期刊可以较快地了解该领域的研究动态。

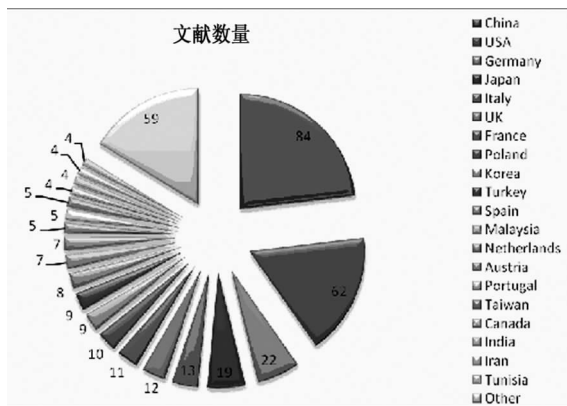


图 2 乳腺癌转移相关基因文献国家分布

表 1 发文章量达 8 篇及其以上的期刊

期刊	发文章量	影响因子
Breast Cancer Research and Treatment	22	4.696
Clinical Cancer Research	15	6.747
Anticancer Research	13	1.428
Cancer Research	13	7.543
British Journal of Cancer	11	4.346
International Journal of Cancer	9	4.722
European Journal of Cancer, Supplement	8	4.121
Journal of Cancer Research and Clinical Oncology	8	2.261
Oncology Reports	8	1.588

**2.5 文献发表的研究机构** 按第一作者所在单位进行统计,共统计出 280 个研究机构在进行乳腺癌转移相关基因方面的研究,发文章量达 4 篇的研究机构有 6 个,分别为 Tianjin Medical University (12 篇), The Netherlands Cancer Institute (6 篇), The University of Texas M. D. Anderson Cancer Center (4 篇), General Hospital of Chinese People's Liberation Army (4 篇), University of Lodz (4 篇), Fu Dan University (4 篇), University of Hong Kong (4 篇), 其中我国研究机构有 4 个,占据了主要地位,说明我国在乳腺癌转移相关基因方面实力较强。但总体来讲,在乳腺癌转移相关基因研究领域,无论国际上还是国内,均未形成一些核心研究机构。

**2.6 核心作者分析** 经调研统计得知, Netherlands 的 Mook、Malaysia 的 Naidu 在 10 年间以第一作者发表了乳腺癌转移相关基因方面的论文 3 篇,为该领域发文章量最大的作者。国外有学者主要研究 MammaPrint70 基因模型在预测乳腺癌预后及转移风险方面的作用<sup>[4-6]</sup>,也有很多学者研究方向为基因多态性与乳腺癌患病风险的关系<sup>[7-9]</sup>。根据洛特卡和普赖斯对科学家的生产率和活动规律的研究,发表论文为 N 篇以上的作者

为该学科的杰出科学家,即核心作者。计算公式为: $N=0.749 \times (\eta_{\max})^{1/2}$ 。式中  $\eta_{\max}$  为发文量最多的作者发表的论文数量<sup>[10]</sup>。根据上述公式,发表论文达 2 篇的作者为该研究领域的核心作者。经统计核心作者共有 20 位,对核心作者所在国家进行统计发现,美国、英国各有 3 位,日本、荷兰、爱尔兰、法国各有 2 位,马来西亚、加拿大、德国、韩国、意大利、瑞典各有 1 位,所有核心作者均在发达国家,特别是欧美地区,说明欧美的学者在该研究领域造诣较深。

**2.7 乳腺癌转移相关基因的文献量分析** 通过对最终纳入的 363 篇文献的详细阅读,共统计出 567 个乳腺癌转移相关基因。其中研究文献量达 6 篇的基因有 18 个,它们的基因名称及所涉及的文献量如下:ERBB2(52)、TP53(26)、ESR1(23)、NME1(16)、KRT19(14)、PGR(12)、VEGFA(12)、EGFR(10)、PTGS2(10)、BRCA1(8)、CD44(8)、CDH1(8)、MMP9(7)、BCL2(6)、CCND1(6)、GSTP1(6)、MKI67(6)、PTEN(6)。其中对 ERBB2 基因研究的文献量最多,共 52 篇,远远多于其他基因,TP53 基因作为最早发现的抑癌基因,紧随其后,有 26 篇文献进行研究。而在 2006~2010 年间文献量达 4 篇的基因有 19 个,ERBB2(33)、ESR1(14)、VEGFA(10)、TP53(10)、PTGS2(6)、PGR(6)、MMP9(6)、KRT19(6)、EGFR(6)、NME1(5)、BRCA1(5)、BCL2(5)、VEGFC(4)、TERT(4)、SCGB2A2(4)、GSTP1(4)、CXCL12(4)、CD44(4)、CCND1(4)。可以发现,近 10 年间与近 5 年间研究较多的基因稍有变化,ERBB2、TP53、ESR1、NME1、KRT19、PGR、VEGFA、EGFR、BRCA1、CD44、BCL2 等基因仍然是研究的热点,CDH1、MKI67、PTEN 基因在近 5 年的研究逐渐减少,对 CXCL12、TERT、VEGFC 基因的研究在近 5 年逐渐增多。

### 3 讨 论

通过对最终纳入的 363 篇乳腺癌转移相关基因研究文献的计量学分析,可以发现,乳腺癌转移相关基因方面的研究目前正处于稳定发展阶段。中国、美国、日本、英国、德国等国都已经开展了这方面的研究,我国在发文量方面遥遥领先于其他国家,但在人均发文量方面远远落后于其他国家。在发文量达 4 篇的 6 个研究机构中,有 4 个在中国,充分说明我国在乳腺癌转移相关基因方面研究方面已经形成了自己的优势,并有形成核心研究机构的趋势。各研究机构发文量均较少,目前尚未形成该研究领域的核心研究机构,迫切需要国内外研究机构间通力合作。各位学者发表文献数也较少,尚无该领域的领军人物。由此可见,目前的研究较为分散,多为单中心研究,样本量较小,迫切需要进行多中心、大样本的研究,因此研究人员之间甚至是国家之间合作是乳腺癌转移相关基因方面研究的趋势。

基因检测技术的发展,促使基因功能研究从基因产物表达检测逐渐发展为基因突变、启动子甲基化、单核苷酸多态性检测。而随着基因芯片技术的发展,乳腺癌转移相关基因的研究逐渐由单基因研究向多基因研究方向发展。多基因模型开始多用于评估乳腺癌预后,近年来开始有人把多基因模型用于评估乳腺癌转移风险,如国外已进入临床应用的 MammaPrint70 基因模型<sup>[7]</sup>,以及 Oncotype Dx21 基因模型<sup>[12]</sup>,Tutt 等<sup>[12]</sup>和 Yau 等<sup>[13]</sup>构建的 14 基因模型,Culhane 和 Quackenbush<sup>[14]</sup>及 Landemaine 等<sup>[15]</sup>建立的预测乳腺癌患者肺转移风险的 6 基因模型。由于乳腺癌转移是一个复杂的、由多基因参与及多步骤

完成的过程,应用单一基因预测乳腺癌的转移风险,其敏感度及精确度均欠佳,而应用多基因模型可显著提高其预测的敏感度及精确度。建立多基因预测模型的根本在于寻找预测价值高、性能稳定的候选基因。本文通过对文献的计量学分析,初步筛选了一些研究较成熟、预测性好的基因,为建立高效的多基因预测模型奠定了基础。多基因预测模型为乳腺癌复发转移风险预测、治疗疗效预测及个体化治疗方案的制订提供了可能,逐渐成为乳腺癌转移研究的热点,显示出巨大的潜力和广泛的应用前景。

### 参考文献:

- [1] Fisher B, Jeong JH, Bryanl J, et al. Treatment of lymph-node-negative, oestrogen\_receptor-positive breast cancer: long-term findings form National Surgical Adjuvant Breast and Bowel Project randomised clinical trials[J]. *Lancet*, 2004, 364(9437): 858-868.
- [2] Swaminathan M, Phillips-Bute BG, Grichnik KP. Bibliometric analysis of global clinical research by anesthesia departments[J]. *Anesth Analg*, 2007(105): 1741-1746.
- [3] 罗式胜. 文献计量学概论[M]. 广州: 中山大学出版社, 1994: 34-46.
- [4] Mook S, Schmidt MK, Weigelt B, et al. The 70-gene prognosis profile predicts early metastases in postmenopausal breast cancer patients[J]. *Cancer Res*, 2009, 69 Suppl 2: S124.
- [5] Mook S, Schmidt MK, Weigelt B, et al. The 70-gene prognosis signature predicts early metastasis in breast cancer patients between 55 and 70 years of age[J]. *Ann Oncol*, 2010, 21(4): 717-722.
- [6] Mook S, Knauer M, Bueno-de-Mesquita JM, et al. Metastatic potential of T1 breast cancer can be predicted by the 70-gene MammaPrint signature [J]. *Ann Surg Oncol*, 2010, 17(5): 1406-1413.
- [7] Naidu R, Yip CH, Taib NA. Polymorphisms of HER2 Ile655Val and cyclin D1 (CCND1) G870A are not associated with breast cancer risk but polymorphic allele of HER2 is associated with nodal metastases[J]. *Neoplasma (Bratislava)*, 2008, 55(2): 87-95.
- [8] Naidu R, Har YC, Taib NA. Polymorphism of FGFR4 Gly388Arg does not confer an increased risk to breast cancer development[J]. *Oncol Res*, 2009, 18(2/3): 65-71.
- [9] Naidu R, Har YC, Taib NA. Associations between hypoxia-inducible factor-1 (alpha) (HIF-1 (alpha)) gene polymorphisms and risk of developing breast cancer[J]. *Neoplasma*, 2009, 56(5): 441-447.
- [10] 邱均平. 信息计量学[M]. 武汉: 武汉大学出版社, 2007: 106-192.
- [11] Cobleigh MA, Tabesh B, Bitterman P, et al. Tumor gene expression and prognosis in breast cancer patients with 10 or more positive lymph nodes[J]. *Clin Cancer Res*, 2005, 11(24 Pt 1): 8623-8631. (下转第 425 页)

可通过侧枝出芽的方式取代损伤的轴突,代偿其功能<sup>[7-8]</sup>。NF 是构成神经元胞体和神经轴突细胞骨架的主要成分,在维护神经元的功能和轴浆转运等一系列与脊髓损伤修复相关的病理、生理变化中发挥着重要作用<sup>[9]</sup>。NF 依据相对分子质量的不同可分为 NF-68、NF-140 及 NF-200 三种。其中 NF-200 在正常情况下只存在于轴索中,而胞体不含或很少含有,故正常时 NF-200 胞体为阴性。有研究发现,在脊髓损伤区附近的神经元在创伤刺激及多种诱导因子的作用下,能大量合成及积存 NF-200 以适应神经再生的需要,此时 NF-200 可在神经元胞体内着色<sup>[10-11]</sup>。脊髓不完全损伤后 NF-200 阳性神经元的数量及神经元胞体着色的程度与后肢的功能恢复情况有密切的关系,说明 NF-200 染色不仅能显示伤后神经元的形态,也能反映其功能状态<sup>[12]</sup>。在缺血/再灌注后,NF-200 在皮质的表达明显减少,这与神经元缺血/再灌注损伤的发生是一致的。轴突的损伤及缺失,必然对神经细胞间的功能联系和传递有着不可避免的影响,因此大鼠的神经功能受损后恢复存在困难。这也与本实验结果是一致的。

CM 作为新一代的哌嗪类钙离子拮抗剂,除了能够抑制钙离子内流、扩张血管外,研究还发现它能够通过抑制腺苷脱氨酶的作用延迟腺苷的降解,减少肌酐和次黄嘌呤的生成,使腺苷在组织内蓄积。腺苷作为内源性保护因子在脑缺血损伤中表现出了显著的自稳态调节及神经调节作用<sup>[13]</sup>。本研究运用 CM 对脑缺血/再灌注大鼠进行急性期治疗,从减轻脑组织炎症反应、促进轴突再生及改善神经功能 3 个方面,探讨了 CM 对急性缺血性脑损伤的保护作用。CM 治疗组与缺血/再灌注组比较后发现:CM 能够显著抑制脑内 IL-1 $\beta$ 、IL-6 的过度表达,明显增加脑皮质 NF-200 的表达,还伴随神经功能缺损的明显改善。这些都提示,CM 可能从以上环节对缺血性脑损伤发挥了较好的保护作用,且这种保护作用随着用药时间的延长更加突出。

#### 参考文献:

- [1] Longa EZ, Weinstein PR, Carlson S, et al. Reversible middle cerebral artery occlusion without craniectomy in rats [J]. *Stroke*, 1989, 20(1): 84-91.
- [2] Nagasawa H, Kogure K. Correlation between cerebral blood flow and histologic changes in a new rat model of middle cerebral artery occlusion [J]. *Stroke*, 1989, 20(8): 1037-1043.
- [3] 王素香. 脑梗死的治疗 [J]. *中国全科医学*, 2004, 7(4): 215.
- [4] Suzuki S, Tanaka K, Nogawa S. Temporal profile and cellular localization of interleukin-6 protein after focal cerebral ischemia in rats [J]. *J Cereb Blood Flow Metab*, 1999, 19(11): 1256-1262.
- [5] Dong Y, Dobkin BH, Cen SY, et al. Motor cortex activation during treatment may predict therapeutic gains in paretic hand function after stroke [J]. *Stroke*, 2006, 37(6): 1552-1555.
- [6] Meintzschel F, Ziemann U. Modification of practice-dependent plasticity in human motor cortex by neuromodulators [J]. *Cereb Cortex*, 2006, 16(8): 1106-1115.
- [7] Dickson TC, Chung RS, McCormack GH. Acute reactive and regenerative changes in mature cortical axons following injury [J]. *Neuroreport*, 2007, 18(3): 283-288.
- [8] Fenrich KK, Skelton N, MacDermid VE, et al. Axonal regeneration and development of de novo axons from distal dendrites of adult feline commissural interneurons after a proximal axotomy [J]. *J Comp Neurol*, 2007, 502(6): 1079-1097.
- [9] Uchida K, Baba H, Maezawa Y, et al. Progressive changes in neurofilament proteins and growth-associated protein-43 immunoreactivities at the site of cervical spinal cord compression in spinal hyperostotic mice [J]. *Spine*, 2002, 27(5): 480-486.
- [10] Yabe JT, Wang FS, Chylinski T, et al. Selective accumulation of the high molecular weight neurofilament subunit within the distal region of growing axonal neurites [J]. *Cell Motil Cytoskeleton*, 2001, 50(1): 1-12.
- [11] Brook GA, Plate D, Franzen R, et al. Spontaneous longitudinally orientated axonal regeneration is associated with the Schwann cell framework within the lesion site following spinal cord compression injury of the rat [J]. *J Neurosci Res*, 1998, 53(1): 51-65.
- [12] 胥少汀, 郭世绂. 脊髓损伤基础与临床 [M]. 2 版. 北京: 人民卫生出版社, 2002.
- [13] Saransaari P, Oja SS. Mechanisms of adenosine release in the developing and adult mouse hippocampus [J]. *Neurochem Res*, 2002, 27(9): 911-918.

(收稿日期: 2011-09-22 修回日期: 2011-11-08)

(上接第 422 页)

- [12] Tutt A, Wang A, Rowland C, et al. Risk estimation of distant metastasis in node-negative, estrogen receptor-positive breast cancer patients using an RT-PCR based prognostic expression signature [J]. *BMC Cancer*, 2008(8): 339.
- [13] Yau C, Esserman L, Moore DH, et al. A multigene predictor of metastatic outcome in early stage hormone receptor-negative and triple-negative breast cancer [J]. *Breast Cancer Res*, 2010, 12(5): R85.
- [14] Culhane AC, Quackenbush J. Confounding effects in "A six-gene signature predicting breast cancer lung metastasis" [J]. *Cancer Res*, 2009, 69(18): 7480-7485.
- [15] Landemaine T, Jackson A, Bellahcene A, et al. A six-gene signature predicting breast cancer lung metastasis [J]. *Cancer Res*, 2008, 68(15): 6092-6099.

(收稿日期: 2011-06-09 修回日期: 2011-11-01)