

# 乳腺癌化疗紫杉类药物敏感性相关基因的生物信息学分析

An Analysis of Chemotherapy Sensitivity of Taxane-based Drugs Related Genes in Breast Cancer//PENG Li-ming, CEN Dong-zhi, WANG Yi-nuo, et al.

彭黎明<sup>1</sup>, 岑东芝<sup>1</sup>, 王贻锬<sup>2</sup>, 鲁衍强<sup>2</sup>, 张积仁<sup>1</sup>  
(1.南方医科大学附属珠江医院肿瘤中心, 广东 广州 510282;  
2.上海市靶向分子医学研究所, 上海 200433)

**摘要:** [目的] 评价乳腺癌化疗紫杉类药物敏感性相关基因的研究状况, 并对未来研究方向进行预测。[方法] 选用 Embase、Medline/Pubmed、BIOSIS Preview 数据库为文献来源, 对 2001~2010 年期间与乳腺癌化疗紫杉类药物敏感性相关基因的文献进行计量学分析。[结果] 相关文献 93 篇, 文献中涉及 429 个基因。生物信息学分析说明所涉及的基因及其产物处于相互联系的复杂网络之中, 同时为关键基因和瓶颈基因主要有 PLK1、CCND1、ESR1、FEN1、CDKN1A、CCNA2、TP53、SRC、EGFR、PIK3CA、MDM2、CCNB1、VEGFA、TOP2A、BRCA1、Bcl-2、FOXM1、EP300、CTNNB1、BIRC5 等。[结论] 乳腺癌化疗紫杉类药物敏感性相关基因的研究可能在未来研究中更受关注, 我们预测的关键基因可能成为未来研究的热点。  
**关键词:** 乳腺癌; 紫杉类药物; 基因; 文献计量学; 生物信息学  
**中图分类号:** R737.9 **文献标识码:** A **文章编号:** 1004-0242(2012)02-0111-04

化疗在乳腺癌的治疗中占有非常重要的地位。在新一代乳腺癌化疗药物中, 应用最广的是紫杉类药物, 如紫杉醇和多西他赛都是通过与微管蛋白结合干扰有丝分裂, 进而抑制肿瘤增殖<sup>[1]</sup>。临床发现在乳腺癌中紫杉类单药有效率 5%~15%<sup>[2]</sup>, 与其他化疗药联合的有效率 8%~31%<sup>[3]</sup>。因此有必要对紫杉类药物有效性的预测指标进行研究, 对乳腺癌患者实行个体化治疗, 现对紫杉类药物敏感性基因进行分析。

## 1 材料与方法

### 1.1 文献检索

在 Embase、Medline/Pubmed 及 BIOSIS Preview 这 3 个权威数据库中进行相关文献检索。

为提高文献的查准率, 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; 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; BIOSIS Preview 数据库进行主题检索。乳腺癌的英文检索词包括: 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\*”。紫杉类药物的英文检索词包括: ① MeSH 词“taxoids”, ② EMtree 词“terpenoid”, ③ 自由词“taxane”、“paclitaxel”、“taxol”、“docetaxel”、“doxtaxol”及“taxotere”。基因的英文检索词为“gene”。物种限定为人类, 时间限定为 2001~2010 年, 语种不限。

### 1.2 文献纳入与排除

纳入全部与乳腺癌患者紫杉类药物敏感性基因及基因产物相关的临床研究。纳入与排除标准: ① 文献形式包括论著、会议论文及会议摘要, 排除综述、个案报道、书信、书籍及社论等; ② 纳入文献需包括能提供足够信息的摘要, 或可以获取全文资料, 排除无法获取摘要的文献; ③ 排除与药物代谢等相关基因的研究; ④ 排除与紫杉类药物敏感性相关的非基

收稿日期: 2011-05-18; 修回日期: 2011-07-24  
通讯作者: 张积仁, E-mail: zhangjiren@126.com

因研究;⑤纳入临床研究文献,排除针对基因的结构、机理等研究;⑥排除细胞株体外研究及动物体内实验;⑦排除重复文献,不同数据库来源,作者、题目及出版年相同的文献按一篇统计;⑧排除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(副主任医师)、实验研究人员(研究员)请教。

### 1.3 文献计量学分析

对最终纳入文献的出版年、第一作者、第一作者国家、研究机构、发表期刊分布进行计量学分析。

### 1.4 紫杉类药物敏感相关基因生物信息学分析

对涉及的基因及其产物进行功能分类及通路分析。利用在线 STRING 对乳腺癌化疗紫杉类药物敏感性基因表达产物构建蛋白质相互作用网络,在线软件 GATHER (<http://gather.genome.duke.edu/>) 按照 Gene Ontology (GO,即基因本体论)分类的标准对其进行功能分类。GO 数据库是一个较大的公开的生物分类学网络资源的一部分,并在每一个分类中,都提供一个描述功能信息的分级结构。这样,GO 中每一个分类术语都以一种被称为定向非循环图表 (DAGs) 的结构组织起来。研究者可以通过 GO 分类号和各种 GO 数据库相关分析工具将分类与具体基因联系起来,从而对这个基因的功能进行描述。将 STRING 软件构建的紫杉类药物敏感性相关基因相互作用网络图导入 Cytoscape 软件可视化,利用其插件 CentiScape 计算网络及各个节点的拓扑特性。

## 2 结果

### 2.1 检索结果

Embase 数据库检出文献 1 063 篇, Pubmed/Medline 数据库检出文献 266 篇, BIOSIS Preview 数据库检出文献 877 篇,共检出文献 2 206 篇。符合要求的文献共 93 篇。

### 2.2 文献计量学分析

近 10 年,紫杉类药物敏感性基因相关文献在 2001~2005 年间缓慢增多,2006~2010 年总体呈现明显上升趋势。93 篇论文发表在 36 种生物医学期刊上,其中 *Breast Cancer Research and Treatment* 发文 14 篇高居第一。

共有 18 个国家发表了关于紫杉类药物敏感性相关基因方面的研究,其中美国(26 篇)最多,日本发表的文献数量也较多,为 15 篇,其次为德国、韩国、法国、希腊、意大利、中国、西班牙、比利时,文献量分别为 9、7、6、5、5、3、3、2 篇。

共 68 个研究机构进行了紫杉类药物敏感性相关基因方面研究。其中美国的 University of Texas MD Anderson Cancer Center 发文量达 8 篇排名第一,其次为美国 Baylor College of Medicine、日本 Osaka University Graduate School of Medicine、日本 National Cancer Center Hospital、日本 Department of Surgery, Shinshu University School of Medicine、美国 Baylor College of Medicine、比利时 Jules Bordet Institute、韩国 Seoul National University College of Medicine、德国 J.W.Goethe-University、韩国 Seoul National University Hospital、法国 Laboratory of Molecular Genetics, Centre Georges Francois Leclerc, 发文量分别为 4、4、3、2、2、2、2、2、2、2 篇。

10 年间在紫杉类药物敏感性相关基因方面有 5 位作者发文量  $\geq 2$  篇,美国 Baylor College of Medicine 的 Chang JC 和英国 Jules Bordet Institute 的 Di Leo A 发文量为 3 篇,法国 Breast Cancer Unit and Translational Research Unit 的 Andre F、日本 National Cancer Center Hospital 的 Park S 和德国 JW Goethe-University 的 Rody A 发文量达 2 篇。

### 2.3 紫杉类药物敏感性相关基因及产物相互作用

STRING 是由 EMBL、SIB 和 UZH 共同开发管理的数据库,主要包含已有文献报道的相互作用蛋白信息以及通过共表达研究、高通量实验和基因组信息分析预测的结果。93 篇文献中涉及相关基因共 429 个,采用 STRING 软件对这 429 个紫杉类药物敏感性相关基因表达产物构建蛋白质相互作用网络(图 1)。由图 1 可知,部分基因处于网络的中心,这些基因之间以及与网络周边的基因之间存在着密切的相互作用关系。

### 2.4 基因的功能分类

利用在线软件 GATHER 分析,429 个紫杉类药物敏感性相关基因共涉及不同分级的 GO 分类 744 类,其中贝叶斯 (Byes) 因子大于 0 的 GO 分类 90 类,Byes 因子大于 3 的 GO 分类 60 类。利用 GODAG 软件对 60 类 GO 分类进行可视化,了解其

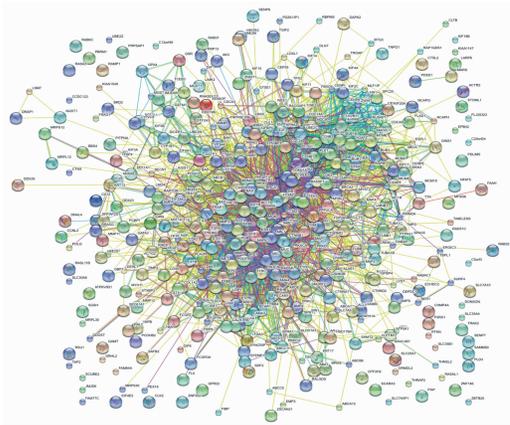


图1 紫杉类药物敏感性基因表达产物蛋白质相互作用网络图

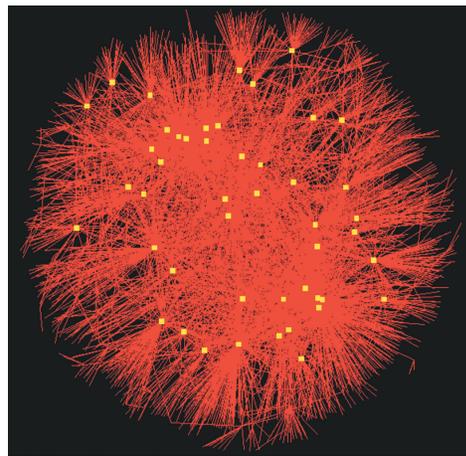


图2 紫杉类药物敏感性相关基因的相互作用图

之间层次相互包含关系,去除过大或过细分类。具体如下:细胞周期包括 85 个基因,细胞增殖包括 99 个基因,细胞有丝分裂 M 期包括 33 个基因,核分裂包括 31 个基因,细胞周期调控点包括 13 个基因,细胞浆移动包括 19 个基因,微管包括 21 个基因,有丝分裂调控点包括 7 个基因,G 蛋白耦联受体信号传导包括 4 个基因,DNA 复制和染色体周期包括 22 个基因,DNA 损伤应激反应包括 22 个基因,凋亡包括 34 个基因,DNA 修复包括 20 个基因,DNA 复制包括 17 个基因,磷酸化包括 40 个基因,细胞骨架内的细胞内转移包括 12 个基因,DNA 再结合包括 10 个基因,有丝分裂周期中 G<sub>2</sub> 期向 M 期的过渡包括 8 个基因,细胞组织和生源论包括 43 个基因,酶相关的受体蛋白信号转导通路包括 14 个基因。

### 2.5 关键基因及瓶颈基因分析

HPRD 数据库是研究蛋白质相互作用的数据库。将 STRING 软件构建的紫杉类药物敏感性相关基因相互作用网络图导入 Cytoscape 软件可视化(图 2),利用其插件 CentiScape 计算网络及各个节点的拓扑特性。网络是由 369 个节点 9 743 条边组成,网络直径为 7,网络的平均距离为 2.72。度为节点连接的边的总数目,最大度值为 282,最小为 2,平均为 52.80±61.25,网络中连接度高的节点对网络的稳定性起到重要作用,本研究度数≥均数+1SD(标准差)的蛋白所对应的基因为关键基因(Hub 基因),共有 67 个,它们的功能改变可能对于整个网络来说是至

关重要的。Betweenness 为网络中所有的最短路径中经过该节点的数量比例,最大 Betweenness 值为 11 672.76,最小为 0,平均为 633.51±1 176.35。相应的节点或者边在整个网络中的作用和影响力,本研究规定 Betweenness 大于等于均数+1SD 的蛋白所对应的基因为瓶颈基因(Bottleneck 基因),共有 29 个。我们认为同时为 Hub 基因和 Bottleneck 基因的对紫杉类药物敏感性相关基因网络至关重要,主要有 PLK1、CCND1、ESR1、FEN1、CDKN1A、CCNA2、TP53、SRC、EGFR、PIK3CA、MDM2、CCNB1、VEGFA、TOP2A、BRCA1、Bcl-2、FOXM1、EP300、CTNNB1、BIRC5 等。

## 3 讨论

我们对 2001~2010 年期间与紫杉类药物敏感性相关基因的文献进行计量学分析,从中可以发现该领域的研究历史、目前的研究现状以及未来的研究趋势。

目前指导乳腺癌临床用药的标准主要有 TNM 分期、病理学特征等,为了避免过度治疗,有必要从分子生物学方面进行预测,建立预测紫杉类药物敏感性相关的基因,实现紫杉类药物个体化治疗,指导临床实践,最大程度地使乳腺癌患者获益。

这 10 年中研究最多是 p53、HER-2。综合近 10 年的研究可以看出 p53 基因缺失或突变的乳腺癌患者对紫杉类获益较大<sup>[4]</sup>。而对 HER-2 的研究得出

同的结论,一部分学者认为 HER-2 阴性患者对紫杉类临床获益大<sup>[5,6]</sup>,另外一部分学者考虑 HER-2 阳性的乳腺癌患者对紫杉类为主的化疗有更好的获益<sup>[7]</sup>。

通过对关键基因和瓶颈基因分析可以看出该领域未来的研究方向、热点基因,利用 Cytoscape 软件将相互作用网络可视化,利用其插件 CentiScape 计算网络及各个节点的拓扑特性可以看出 p53、HER-2、BRCA1、Bcl-2、ESR1、EGFR 等可能是未来研究的热点。

### 参考文献:

[1] Noguchi S. Predictive factors for response to docetaxel in human breast cancers [J]. *Cancer Sci*, 2006, 97 (9):813-820.

[2] Sledge GW, Neuberg D, Bernardo P, et al. Phase III trial of doxorubicin, paclitaxel, and the combination of doxorubicin and paclitaxel as front-line chemotherapy for metastatic breast cancer [J]. *J Clin Oncol*, 2003, 21 (4): 588-592.

[3] Hess KR, Anderson K, Symmans WF, et al. Pharmacogenomic predictor of sensitivity to preoperative chemothera-

py with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer [J]. *J Clin Oncol*, 2006, 24(26): 4236-4244.

[4] Di Leo A, Tanner M, Desmedt C, et al. p-53 gene mutations as a predictive marker in a population of advanced breast cancer patients randomly treated with doxorubicin or docetaxel in the context of a phase III clinical trial[J]. *Ann Oncol*, 2007, 18(6): 997-1003.

[5] Hayes DF, Thor AD, Dressler LG, et al. HER2 and response to paclitaxel in node-positive breast cancer [J]. *N Engl J Med*, 2007, 357(15): 1496-1506.

[6] Wang X, Yao F, Liu N, et al. Clinical implications of HER-2 and p53 in taxane-based and anthracycline-based neoadjuvant chemotherapy in breast Cancer [J]. *Chinese Journal of Clinical Oncology*, 2008, 5(6): 424-428.

[7] Liu F, Jiang ZF, Song ST, et al. HER-2 and ER expression in prediction of chemo-sensitivity of taxane for advanced breast cancer[J]. *Zhonghua Zhong Liu Zazhi*, 2006, 28(6): 449-451.

[8] Sekine I, Shimizu C, Nishio K, et al. A literature review of molecular markers predictive of clinical response to cytotoxic chemotherapy in patients with breast cancer[J]. *Int J Clin Oncol*, 2009, 14(2): 112-119.