

DOI: 10.3969/j.issn.1005-8982.2016.21.009

文章编号: 1005-8982(2016)21-0045-05

## 蛋白激酶基因在不同癌基因组中的变异频率及表达变化\*

郭冬梅,张蜀敏,易芳,唐珍

(川北医学院 生物化学教研室,四川 南充 637000)

**摘要:目的** 探讨蛋白激酶基因突变与癌症发生的关系。**方法** 本研究通过搜索癌基因组数据库,获得人类全部514个蛋白激酶基因在不同癌基因组中的变异频率及表达变化。**结果** 基因表达搜索结果显示,绝大多数基因的表达上调。基因变异搜索结果显示,512个成员发生变异,变异频率最高为90%。其中,基因扩增频率要高于突变频率和删除频率,且删除频率最低。12个成员的突变频率>30%,功能研究相关的文献资料显示,其中的8个成员可能与癌症产生密切相关。**结论** 笔者推测高频突变的蛋白激酶基因可能具有重要功能,其很可能是癌基因或者抑癌基因。本研究结果为进一步深入研究相关蛋白激酶基因的功能,尤其是功能未知的成员,及其与癌症发生的关系提供重要参考。

**关键词:** 蛋白激酶,基因突变,癌基因组学

**中图分类号:** R730

**文献标识码:** B

## Mutation frequency and expression variation of protein kinase genes in different types of cancers\*

Dong-mei Guo, Shu-min Zhang, Fang Yi, Zhen Tang

(Department of Biochemistry, North Sichuan Medical College,  
Nanchong, Sichuan 637000, China)

**Abstract: Objective** To study the relationship between mutation of protein kinase (PK) genes and cancer occurrence. **Methods** In this study, the frequency of sequence alteration and gene expression variation of 514 PK genes in different types of cancers were obtained by searching the databases of cancer genomes. **Results** Results of gene expression search indicated that most of the PK genes have overexpression. Results of gene alteration search indicated that 512 genes have mutations; among which the frequency of amplification is higher than mutation and deletion, and deletion has the lowest frequency. The highest alteration frequency is 90%. Mutation frequency of 12 genes is higher than 30%, and 8 of them are associated with cancer occurrence according to the functional studies from bibliographic databases. **Conclusions** We suggest that PK genes with high mutation frequency might have vital functions, and they could be oncogenes or tumor suppressor genes. This study will shed light on our understanding of the function of PK genes (especially those members without functional studies) and their roles in cancer occurrence.

**Keywords:** protein kinase; gene mutation frequency; cancer genomics

癌症已成为导致人类死亡的首要原因<sup>[1]</sup>。在过去几十年中,人们已经鉴定出许多与癌症发生密切相关

的基因,如肉瘤基因(sarcoma gene, Src)、骨髓细胞瘤病毒癌基因同源基因、大鼠肉瘤病毒癌基因等<sup>[2-4]</sup>。

收稿日期:2016-03-14

\* 基金项目:四川省教育厅自然科学基金(No:15ZA0213);川北医学院博士启动基金(No:CBY13-QD-10);川北医学院科研发展计划项目(No:CBY13-A-ZP04);川北医学院校级科研课题(No:CBY15-A-YB27)

[通信作者] 唐珍, E-mail: tzhbird@163.com; Tel: 18808171213

尽管导致癌症发生的遗传背景复杂多样,但是根据人类大规模癌基因组测序结果显示,肿瘤组织中突变频率最高的蛋白常常是蛋白激酶<sup>[5-7]</sup>。此外,许多原癌基因(突变后成为致癌基因)或抑癌基因(遏制肿瘤形成的基因)的编码产物也都是蛋白激酶<sup>[8]</sup>。因此,蛋白激酶突变与癌症发生之间有着千丝万缕的关系。

在真核细胞中,蛋白激酶参与绝大多数的细胞内及细胞间信号转导途径。蛋白激酶通过调节底物水平的磷酸化或者去磷酸化状态进而调控细胞内的许多发育过程,包括细胞的新陈代谢、细胞周期进程、细胞骨架重建,以及细胞的运动、凋亡和分化等。另外,蛋白激酶还参与生物发育、应激反应、激素调控、神经系统及免疫系统发挥作用时的细胞间信号转导途径<sup>[9-10]</sup>。真核生物中编码蛋白激酶的基因构成一个非常庞大的超基因家族,该超家族在人类中拥有 516 个成员,占人类所有基因数目的 1.7%<sup>[11]</sup>。已有研究结果显示,在某些情况下,首先是某些蛋白激酶成员突变,然后才导致人类疾病产生;而在另外一些情况下,首先是某些疾病已经发生,然后激酶基因再随之发生突变<sup>[12]</sup>。

由于蛋白激酶在细胞信号转导途径中的重要作用,该家族在最近几十年得到广泛研究。然而,目前该超基因家族中仍然有大量成员的功能不清楚。据 FEDOROV 等<sup>[8]</sup>统计,该家族中有 >100 个成员(约占总数的 25%)的功能不甚清楚,其中近 50%成员的功能完全没有被研究过。此外,LAHIRY 等<sup>[13]</sup>也对已有激酶蛋白的功能研究进行分类统计,结果显示,与疾病有关联的激酶突变类型高达近千种,其中 77%的突变类型为错义突变,而 >80%的突变发生在蛋白激酶的催化结构域及其周围,这意味着该突变类型很可能会影响相应蛋白的功能。已有的研究结果促使学者们思考以下 2 个问题:①蛋白激酶超家族中功能未知的成员是否具有重要功能呢?②如何从众多成员中挑选出合适的候选基因,作为进一步功能研究的目标基因呢?

为探讨上述 2 个问题,本研究通过搜索相关癌基因组数据库,获得人类 513 个蛋白激酶基因在不同癌基因组中的突变频率。在此基础上,通过广泛查阅有关蛋白激酶基因功能研究的文献,进而总结出在癌基因组中突变频率较高的基因多具有重要功能,很可能与癌症的产生密切相关。因此,笔者可以挑选一些在癌基因组中突变频率较高,但是目前缺乏清晰功能研究的成员,作为进一步功能研究的候选基因。该

研究结果将为进一步深入研究相关蛋白激酶基因,尤其是功能未知成员的功能打下坚实基础。

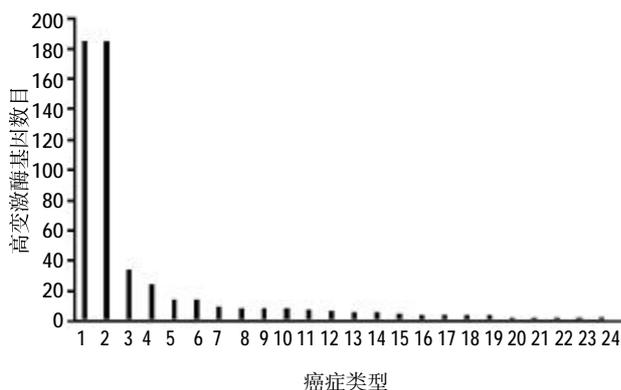
## 1 资料与方法

以人类全部蛋白激酶家族成员为模板<sup>[10]</sup>,利用 cBio Portal ([www.cbioportal.org](http://www.cbioportal.org)) 平台获取人类全部 516 个蛋白激酶基因在不同癌基因组数据库中的变异情况。该网站提供一个挖掘、查看以及分析不同层面的癌基因组数据库综合平台,其整合目前主要癌基因组数据库资源如癌症基因组图谱、癌症体细胞突变数据库、国际癌症基因组联盟等,对于检索基因在不同癌基因组中的变异情况提供便利条件。笔者对 516 个成员在不同癌基因组中的变异频率(包括突变、删除及扩增)进行由高到低的排序。依次从突变频率最高的基因开始,通过广泛的文献及基因数据库检索,如美国国家医学图书馆,美国国立生物技术信息中心等,获取相关基因的功能研究结果。另外,笔者还统计每一个基因的表达变化,包括基因表达是上升还是下降,又或者是两者兼有之。最后综合分析基因功能与基因突变频率的关系。

## 2 结果

MANNING 等<sup>[14]</sup>报道人类有 516 个蛋白激酶基因,但有 2 个成员信息有误,未检索到相关信息,因此本研究只分析 514 个成员。蛋白激酶基因在不同癌症中表达变化的统计发现,514 个基因中有 3 个基因的表达未检测到任何变化,1 个基因表达下降,157 个基因的表达既有上调也有下降,353 个基因的表达上调。癌症类型与最高变异频率激酶基因数的统计发现,514 个成员的最高变异率分布在 25 种癌症中。其中,乳腺癌和黑色素瘤是拥有最高变异率基因数最多的癌症,分别有 184 和 183 个激酶基因的最高变异率。笔者推测数据库中乳腺癌标本,黑色素瘤较多,因此得到的数据多或者激酶基因变异更易发生。见图 1。

蛋白激酶基因在不同癌症组织中的变异频率统计发现,514 个成员中有 512 个成员在相关癌症组织中发生变异,2 个成员在已有的癌基因组数据库中未检测到变异。在本研究中,基因变异包含以下 3 种情况:突变、删除及扩增。基因变异频率等于有基因变异的癌症组织样本数除以总癌症样品数目,而突变频率、删除频率及扩增频率的计算采用相似的计算公式。对于每一个基因,笔者分别统计其在各种



1:乳腺癌(184);2:神经性前列腺癌(183);3:结缔组织增生性黑色素瘤(32);4:胰腺癌(23);5:皮肤鳞状细胞癌(13);6:弥漫性大 B 细胞淋巴瘤(13);7:前列腺癌(8);8:子宫内膜癌(7);9:NCI-60 细胞系(7);10:肺癌(7);11:睾丸生殖细胞癌(6);12:子宫癌(5);13:原发性中枢神经系统淋巴瘤(4);14:皮肤黑色素瘤(4);15:肾透明细胞癌(3);16:腺样囊性癌(2);17:膀胱尿路上皮癌(2);18:多形性成胶质细胞瘤(2);19:恶性周边神经鞘瘤(2);20:急性髓细胞样白血病(1);21:总的癌细胞系(1);22:胆管癌(1);23:胰腺的腺泡细胞癌(1);24:肉瘤(1)。括号中数字代表该癌症拥有的高变异频率激酶基因的数目

图 1 不同癌症拥有的高变异频率激酶基因数目

癌症组织中最高变异率、最高突变率、最高删除频率,以及最高扩增频率。每个成员均在  $\geq 1$  种癌症组织中发生变异。基因最高变异率有可能同时包含 3 种变异,即突变、删除及扩增,或者只包含其中 1 或 2 种情况。在 3 种类型的变异中,频率的分布具有较大差异。总的说来,扩增频率要比其他两种变异频率高,有 231 个成员的最高扩增频率  $>20\%$ ,而只有 39 个成员的最高突变率及 4 个成员的最高删除率  $>20\%$ 。该结果与大量基因表达上调具有一致性,极有可能是基因扩增导致相应的基因表达上调。基因删除频率最低,451 个成员的最高删除率  $<10\%$ ,61 个成员在  $10.1\% \sim 20\%$ ,仅有 3 个成员在  $20.1\% \sim 30\%$ ,1 个成员为  $31\%$ 。突变频率中,504 个基因的最高突变率  $<30\%$ ,12 个成员的突变率  $>30\%$ ,其中 9 个成员最高突变率在  $30.1 \sim 40.0\%$ ,3 个成员是高突变率基因,分别是 DNA 依赖蛋白激酶(protein kinase,DNA-activated,catalytic polypeptide,PRKDC)(44.8%),

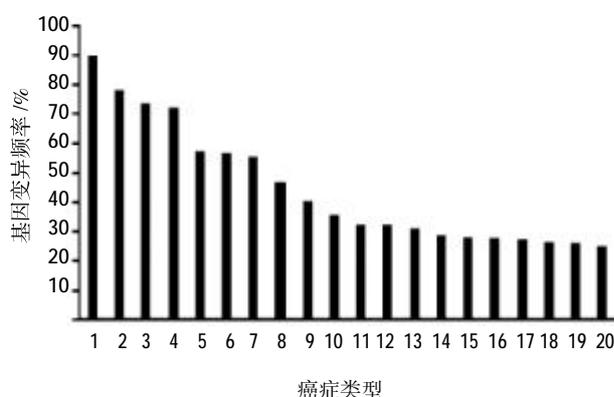
码丝氨酸 / 苏氨酸激酶的乳腺和卵巢原癌基因(B-Raf proto-oncogene,serine/threonine kinase,BRAF)(62.8%)和肌联蛋白(Titin,TTN)(90%)。在癌细胞中,染色体非整倍化的结果常常是染色体的加倍,由此引起大量基因拷贝数目的增加,即基因扩增<sup>[4]</sup>,笔者很难确定基因扩增到底是导致癌症的原因,还是癌症发生后的相应结果。因此,在本研究中,笔者着重探讨基因突变与癌症发生的关系。见图 1 和附表。

基因突变导致所编码蛋白质结构和功能的改变是原癌基因的活化的一种重要机制,越来越受到人们的重视。近年来,癌症研究的一个热门领域是癌基因组学。其研究思路是首先对同一个体的癌症组织与正常癌旁组织的基因组进行比较,从而找到那些与癌症产生密切相关的候选基因,再进一步对该候选基因进行功能验证。癌基因组学,将为研究癌症发生的原因和癌症治疗等提供重要线索。例如,在一项有关急性髓细胞白血病的癌基因组学研究中,研究人员通过对患者癌症组织和正常癌旁组织(皮肤)的基因组序列进行比较分析,从而发现一些重要的经常性突变的基因,进一步分析认为该突变极有可能与该癌症的发生有关<sup>[5]</sup>。因此在本研究中,笔者首先关注的对象是在癌基因组中具有较高突变频率的基因。突变频率最高的成员是 TTN 基因,在许多癌症中该基因都具有较高突变频率,其中在黑色素瘤的肿瘤组织中的突变频率高达 90%。通过检索文献,得知该基因对维持染色体的结构具有重要的功能<sup>[6]</sup>,这与最新有关癌症形成的新理论—癌细胞染色体的非整倍化是癌症的驱动者相吻合<sup>[6]</sup>。突变频率第二高的成员 BRAF 基因(62.8%)是一个研究较多并且已被确认的原癌基因,也被认为是癌症产生的驱动基因<sup>[7]</sup>。突变频率第三高的基因 PRKDC(44.8%)参与 DNA 的损伤与修复,在肿瘤组织中常常过表达,很可能是一个癌症相关基因<sup>[8]</sup>。突变频率为  $30\% \sim 40\%$  的 9 个成员中,有 5 个基因的功能研究结果显示,其可能与癌症发生相关;2 个基因功能研究结果显示,可能与癌症无关;剩下 2 个成员功能未知。以上 2 个功能未

附表 基因最高变异频率分布 个

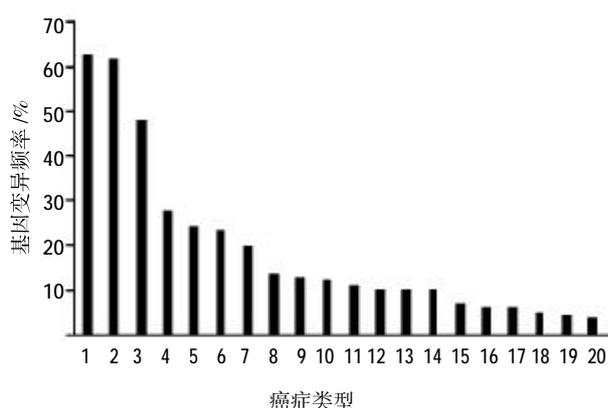
变异频率	0.0% ~ 10.0%	10.1% ~ 20.0%	20.1% ~ 30.0%	30.1% ~ 40.0%	40.1% ~ 50.0%	50.1% ~ 60.0%	60.1% ~ 70.0%	70.1% ~ 80.0%	80.1% ~ 90.0%	90.1% ~ 100.0%
突变率	326	151	27	9	1	0	1	0	1	0
删除率	451	61	3	1	0	0	0	0	0	0
扩增率	55	230	161	51	15	4	0	0	0	0
变异率	24	226	181	58	17	7	2	0	1	0

知成员,或者初步功能研究显示,与癌症有一定相关性的基因,都可作为进一步功能研究的候选基因。在突变率 <20%的基因中,也有成员是癌基因或者抑癌基因。例如,广为人知的原癌基因 *Src*,其最高突变频率仅为 4.5%。这与已有的功能研究结果相吻合—*Src* 基因致癌是通过表达量的上升完成<sup>[3,19-20]</sup>,因此,与该基因功能密切相关的区域在其调控区,而非编码区。此外,笔者通过挖掘 *Src* 基因在癌基因组数据库中的相关信息,也发现该基因的 mRNA 数量在许多癌组织中显著上调。见图 2、3。



1: 结缔组织增生性黑色素瘤; 2: 皮肤黑色素瘤 1; 3: 肺鳞状细胞癌; 4: 总的癌细胞系; 5: 膀胱尿路上皮癌; 6: 胃癌; 7: 食道腺癌; 8: 头颈部鳞癌; 9: 结直肠癌; 10: 食管鳞状细胞癌; 11: 子宫内膜癌; 12: 宫颈鳞癌; 13: 弥漫大 B 细胞淋巴瘤; 14: 儿童尤文肉瘤; 15: 肝细胞肝癌; 16: 前列腺癌; 17: 乳腺癌; 18: 恶性周围神经鞘膜瘤; 19: 胰腺癌; 20: 多形性成胶质细胞瘤

图 2 *TTN* 基因在不同癌症细胞中的变异频率



1: 皮肤黑色素瘤; 2: 乳头状甲状腺癌; 3: 黑色素瘤; 4: 乳腺癌; 5: NCI-60 细胞系; 6: 神经性前列腺癌; 7: 结缔组织增生性黑色素瘤; 8: 总的癌细胞系; 9: 胰腺的腺泡细胞癌; 10: 卵巢浆液性囊腺癌; 11: 肺癌; 12: 结直肠癌; 13: 鳞状上皮细胞癌; 14: 淋巴瘤; 15: 胃腺癌; 16: 骨髓瘤; 17: 弥漫大 B 细胞淋巴瘤; 18: 膀胱尿路上皮癌; 19: 子宫癌; 20: 脑胶质瘤 (该图只显示部分代表性癌症细胞)

图 3 *BRAF* 基因在不同癌组织中的变异频率

### 3 讨论

结合以上基因突变频率及基因功能研究的文献报道,笔者认为,具有较高突变频率的基因很可能行使重要的生理功能,也有可能是癌基因或抑癌基因。另外在肿瘤组织中的基因序列突变频率不高,但是表达量变化巨大的基因,也有可能是癌症相关基因。该基因都可作为癌症研究的候选基因,进而通过分子生物学实验技术验证其功能。本研究结果为进一步深入研究相关蛋白激酶基因,尤其是功能未知成员的功能打下坚实的基础,也将为研发相关蛋白激酶抑制剂,即为相关药物研究提供重要的理论指导。

#### 参 考 文 献:

- [1] International Agency for Research on Cancer, World Health Organization. GLOBOCAN 2012: estimated cancer incidence, mortality and prevalence worldwide in 2012 [EB/OL]. [2012-05-03]. <http://globocan.iarc.fr>.
- [2] BELL D W. Our changing view of the genomic landscape of cancer[J]. J Pathol, 2010, 202(2): 231-243.
- [3] BLUME-JENSEN P, HUNTER T. Oncogenic kinase signalling[J]. Nature, 2001, 411(6835): 355-365.
- [4] VOGT P K. Retroviral oncogenes: a historical primer[J]. Nat Rev Cancer, 2012, 12(9): 639-648.
- [5] GREENMAN C, WOOSTER R, FUTREAL P A, et al. Statistical analysis of pathogenicity of somatic mutations in cancer[J]. Genetics, 2006, 173(4): 2187-2198.
- [6] LIN J, GAN C M, ZHANG X S, et al. A multidimensional analysis of genes mutated in breast and colorectal cancers[J]. Genome Res, 2007, 17(9): 1304-1318.
- [7] WOOD L D, PARSONS D W, JONES S, et al. The genomic landscapes of human breast and colorectal cancers[J]. Sci, 2007, 318: 1108-1113.
- [8] FEDOROV O, MULLER S, KNAPP S. The (un) targeted cancer kinome[J]. Nat Chem Biol, 2010, 6: 166-169.
- [9] RAKSHAMBIKAI R, MANOHARAN M, GNANAVEL M, et al. Typical and atypical domain combinations in human protein kinases: functions, disease causing mutations and conservation in other primates[J]. RSC Adv, 2015, 5: 25132-25148.
- [10] LAI S, SAFAEI J, PELECH S. Evolutionary ancestry of eukaryotic protein kinases and choline kinases[J]. J Biol Chem, 2016, 291: 5199-5205.
- [11] MANNING G, WHYTE D B, MARTINEZ R, et al. The protein kinase complement of the human genome[J]. Science, 2002, 298 (5600): 1912-1934.
- [12] COHEN P. Protein kinases—the major drug targets of the twenty-first century[J]. Nat Rev Drug Discov, 2002, 1(4): 309-315.
- [13] LAHIRY P, TORKAMANI A, SCHORK N J, et al. Kinase mu-

- tations in human disease: interpreting genotype-phenotype relationships[J]. *Nat Rev Genet*, 2010, 11(1): 60-74.
- [14] HACKMAN J P, VIHOLA A K, UDD A B. The role of titin in muscular disorders[J]. *Ann Med*, 2003, 35(6): 434-441.
- [15] MARDIS E R, DING L, DOOLING D J, et al. Recurring mutations found by sequencing an acute myeloid leukemia genome[J]. *N Engl J Med*, 2009, 361(11): 1058-1066.
- [16] DAVOLI T, XU A W, MENGWASSER K E, et al. Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome[J]. *Cell*, 2013, 155(4): 948-962.
- [17] DAVIES H, BIGNELL G R, COX C, et al. Mutations of the BRAF gene in human cancer[J]. *Nature*, 2002, 417: 949-954.
- [18] KOTULA E, BERTHAULT N, AGRARIO C, et al. DNA-PKcs plays role in cancer metastasis through regulation of secreted proteins involved in migration and invasion[J]. *Cell Cycle*, 2015, 14(12): 1961-1972.
- [19] CHATZIZACHARIAS N A, KOURAKLIS G P, GIAGINIS C T, et al. Clinical significance of Src expression and activity in human neoplasia[J]. *Histol Histopathol*, 2012, 27(6): 677-692.
- [20] CHOI C W, KIM Y H, SOHN J H, et al. Focal adhesion kinase and Src expression in premalignant and malignant skin lesions [J]. *Exp Dermatol*, 2015, 24: 361-364.

(童颖丹 编辑)