

前列腺癌易感位点 rs2252004 的 eQTL 和 meQTL 研究

葛雨秋, 金 经, 马高祥, 邵 卫, 储海燕, 王美林, 张正东*

南京医科大学公共卫生学院职业医学与环境卫生学系, 江苏 南京 211166

[摘要] 目的:全基因组关联性研究(GWAS)发现 rs2252004 与前列腺癌的易感性相关,本研究拟进一步探讨 rs2252004 参与前列腺癌发生的可能分子生物学机制。方法:运用TCGA公共数据库中的基因型、甲基化、基因表达数据,采用线性回归分析 rs2252004 基因型与其上下游1 Mb区域 CpG 位点的甲基化水平及相关基因表达的关联性。结果:rs2252004 与 FGFR2 基因上的3个 CpG 位点的甲基化水平显著相关(cg25294906, $P=0.031$; cg03552039, $P=0.042$; cg16499947, $P=0.049$)。而 FGFR2 在50例配对的前列腺癌组织和癌旁组织中表达差异有统计学意义($P=9.3\times 10^{-11}$)。与 GG 基因型相比,rs2252004 GT 基因型的前列腺癌患者的 FGFR2 基因表达量显著降低($P=0.007$)。此外,GT/TT 基因型与 GG 基因型患者的 FGFR2 表达量差异有统计学意义($P=0.006$)。结论:rs2252004 可能通过影响 FGFR2 甲基化水平,进而调控 FGFR2 基因的表达,从而参与前列腺癌的发生。

[关键词] rs2252004; 甲基化; FGFR2; 前列腺癌

[中图分类号] R737.25

[文献标志码] A

[文章编号] 1007-4368(2018)01-0031-03

doi: 10.7655/NYDXBNS20180107

eQTL and meQTL analysis of prostate cancer susceptibility locus rs2252004

Ge Yuqiu, Jin Jing, Ma Gaoxiang, Shao Wei, Chu Haiyan, Wang Meilin, Zhang Zhengdong*

Department of Occupational Medicine Environmental Health, School of Public Health, NMU, Nanjing 211166, China

[Abstract] **Objective:** Previous genome-wide association study (GWAS) has identified that rs2252004 was associated with the susceptibility of prostate cancer. This study was aimed to investigate the molecular biological mechanisms underlying the effect of rs2252004 on tumorigenesis of prostate cancer. **Methods:** We obtained the genotype, methylation and gene expression data from The Cancer Genome Atlas (TCGA) database. The relationships between rs2252004 genotype (0, 1, or 2) and the β -value or proportion of methylation at each CpG site (within 500 kb of rs2252004) and the expression level of involved genes were evaluated via linear regression. **Results:** The results showed that rs2252004 was significantly associated with the methylation level of three CpG sites located in FGFR2 (cg25294906, $P=0.031$; cg03552039, $P=0.042$; cg16499947, $P=0.049$, respectively). Furthermore, the expression of FGFR2 remarkably differed between 50 paired prostate tumors and adjacent normal tissues ($P=9.3\times 10^{-11}$). Besides, individuals carrying rs2252004 GT genotypes were significantly associated with decreased expression of FGFR2 than those with GG genotypes ($P=0.007$). We also observed a notable different expression of FGFR2 between GT/TT genotype and GG genotype individuals ($P=0.006$). **Conclusion:** Altogether, our results suggest that rs2252004 may regulate the methylation level of FGFR2, which contributes to influence FGFR2 expression, thus be involved in the tumorigenesis of prostate cancer.

[Key words] rs2252004; methylation; FGFR2; prostate cancer

[Acta Univ Med Nanjing, 2018, 38(01):0031-0033, 87]

在全球范围内,前列腺癌发病率和死亡率分别位居男性肿瘤的第2位和第6位。据统计在2012年全球约有110多万例前列腺癌新发病例及30多万

[基金项目] 国家自然科学基金(81473050),江苏省公共卫生与预防医学优势学科

*通信作者(Corresponding author), E-mail: drzdzhang@gmail.com

例死亡病例。前列腺癌的发病率和死亡率存在着明显的地区差异,西方发达国家发病率最高,而加勒比海地区以及非洲中部和南部是前列腺癌死亡率最高的地区。虽然亚洲人群发病率和死亡率较低,但是近年来中国、日本等亚洲国家的前列腺癌的发病率呈现明显上升趋势^[1-2]。年龄、种族、家族

史是目前公认的前列腺癌的危险因素,既往流行病学研究以及双生子研究表明除了环境危险因素,遗传因素在前列腺癌的发生过程中发挥了重要作用^[3]。全基因组关联性研究(genome-wide association studies, GWASs)是近年来研究复杂性疾病常见变异强有力的手段,目前已经发现多种疾病的多个易感性位点^[4]。至今, GWAS已经发现超过100个前列腺癌易感性位点,而其中大部分是在欧美人群中发现的,只有10个是亚洲人群中报道的易感性位点^[5-8]。

然而GWAS发现的易感性位点主要是位于非编码区及基因间区的常见变异,多态性位点自身并未改变基因的编码序列,肿瘤的发生机制及疾病的真正致病位点和致病基因有待进一步研究。有研究表明易感性位点可能具有eQTL(expression quantitative trait loci)作用,即多态性位点可能位于调控原件,通过改变基因的表达从而影响肿瘤的发生^[9]。多项研究发现DNA多态性位点可能会影响CpG岛附近的DNA甲基化水平,进而在不改变DNA序列的情况下调控基因表达,即meQTL(methylation quantitative trait loci)。meQTL可以作为桥梁连接基因型与表型,有助于对非编码危害性位点(risk allele)的解释,以及发现新的疾病相关候选基因^[10]。

2012年, Akamatsu等^[7]在日本人群中进行了一项多中心、多阶段、大样本的前列腺癌GWAS,发现10q26区的单核苷酸多态性位点(SNP)rs2252004与前列腺癌易感性显著相关。此外Gu等^[11]研究结果显示rs2252004 A等位基因能够提高中国人群前列腺增生的发病风险。然而rs2252004位于基因荒漠区,其相关的10q26高连锁不平衡区域(LD blocks)不包含任何编码基因。本研究旨在探讨10q26区域的rs2252004是否作为meQTL发挥作用从而影响前列腺癌的易感性。

1 材料和方法

1.1 材料

TCGA数据库由美国国家癌症研究所(National Cancer Institute, NCI)和国家人类基因组研究所(National Human Genome Research Institute, NHGRI)共同维护,通过对多达30余种癌症患者进行高通量测序分析以深入了解疾病的遗传基础,为疾病的诊断、治疗和预防提供新的线索。TCGA数据库中样本主要来自欧美人群,每个样本包含多种基因组学信息,有基因表达谱数据(gene expression profiling)、基因分型数据(SNP genotyping)、全基因组DNA甲

基化表达数据(DNA methylation profiling)、microRNA表达数据(microRNA profiling)等。本研究从TCGA数据库中下载了前列腺癌的基因型数据、基因表达数据以及甲基化表达数据,其中meQTL分析使用了503例前列腺癌患者的基因型和甲基化数据。eQTL包含了495例前列腺癌患者的基因型和基因表达数据。同时还使用了50例基因表达数据比较配对癌组织与癌旁组织的基因表达差异。

1.2 方法

在rs2252004上下游各扩500 kb,提取该区域内所有的CpG位点。采用线性回归分析CpG位点甲基化水平(β -value)与基因型(0, 1, 2)的关联性。对gene expression数据中的基因表达量进行了lg转换后,应用线性回归分析基因表达与基因型之间的关联性。配对t检验用于分析前列腺癌组织与癌旁组织的基因表达差异。rs2252004不同基因型患者基因表达差异,所采用的统计学方法是方差分析。所有的统计检验均为双侧概率检验, $P \leq 0.05$ 为差异有统计学意义。使用的统计学软件为R软件。

2 结果

2.1 rs2252004的meQTL分析

在rs2252004周围1 Mb范围内(上下游各扩500 kb)共提取了72个CpG岛的甲基化数据。如图1所示,通过线性回归分析发现,rs2252004与FGFR2基因上的3个CpG位点的甲基化水平显著相关(cg25294906, $P=0.031$; cg03552039, $P=0.042$; cg16499947, $P=0.049$)。这3个CpG位点均位于FGFR2的内含子区,距rs2252004约400 kb范围。由于rs2252004所在的10q26区是基因荒漠区,故我们推测rs2252004可能通过远端调控关键基因的表达,进而参与前列腺癌的发生。

2.2 rs2252004的eQTL分析

本研究通过meQTL发现rs2252004能够改变FGFR2基因上的3个CpG位点的甲基化状态,为此我们进一步探讨rs2252004能否影响前列腺癌中FGFR2基因的表达。首先分析TCGA数据库中配对的前列腺癌组织和癌旁组织FGFR2基因的表达,结果发现癌组织中FGFR2显著低于癌旁组织(图2, $P=9.29 \times 10^{-11}$)。eQTL结果显示(表1),与GG基因型相比,rs2252004 GT基因型前列腺癌患者的FGFR2基因表达量显著降低($P=0.007$)。此外,将GT、TT基因型合并,与GG基因型患者的FGFR2表达量相比差异有统计学意义($P=0.006$)。而TT与GG基因型之间

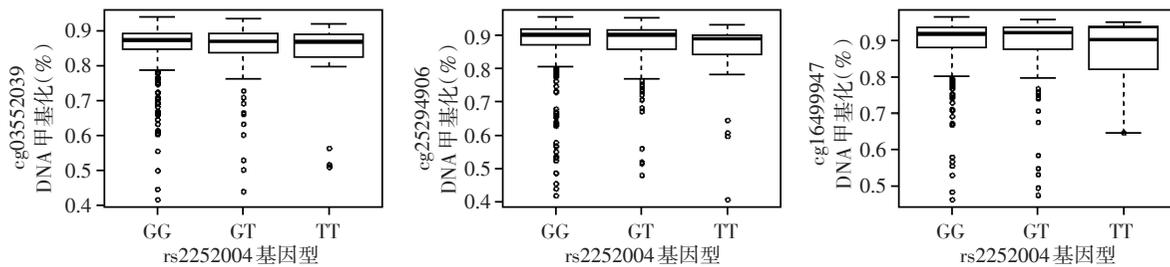


图1 rs2252004 不同基因型在 FGFR2 3 个 CpG 位点的甲基化表达量

Figure 1 Genotypes of rs2252004 correlate with the methylation levels of three CpG sites in FGFR2 gene

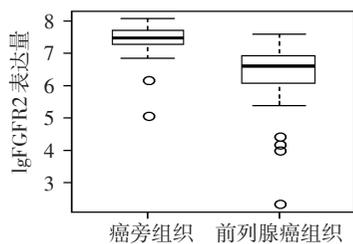


图2 前列腺癌组织与癌旁组织中 FGFR2 基因的表达

Figure 2 FGFR2 expression in paired prostate tumors and adjacent normal tissues

表1 rs2252004 不同基因型的 FGFR2 基因表达差异分析

Table 1 Association between rs2252004 genotype and FGFR2 gene expression

| 基因型 | 例数 | lgFGFR2 | P 值 |
|-------|-----|-------------|-------|
| GG | 363 | 6.12 ± 1.02 | - |
| GT | 109 | 5.81 ± 1.15 | 0.007 |
| TT | 23 | 5.89 ± 1.30 | 0.304 |
| GT/TT | 132 | 5.82 ± 1.17 | 0.006 |

的基因表达差异无统计学意义($P=0.304$)。

3 讨论

近年来, GWAS 已经发现了多种复杂性疾病的多个易感性位点, 而这些易感性位点主要位于内含子区及基因间区。多数 GWAS 位点并不是显著的功能性位点, 不能完全阐明该位点作用的独立致病基因, 故疾病的易感基因和发病机制尚不完全清楚^[12]。有研究表明, GWAS 发现的复杂性疾病相关性的 SNPs 相较其他 SNPs 更有可能通过影响基因的表达参与疾病进展。这些研究显示 GWAS 报道的 SNP 主要是通过影响蛋白产物的表达量, 而不是改变蛋白产物的类型^[13]。众所周知, 甲基化能够影响基因的表达, SNPs 可以通过影响甲基化状态进而调控关键基因的表达, 改变疾病的发生风险^[14]。所以通过对易感性位点的功能注释, 可以进一步了解 SNP 影响疾病易感性的潜在生物学机制。

2012年, 在日本人群进行的 GWAS 发现了 10q26 区的 1 个新的前列腺癌易感性位点 rs2252004。而在 rs2252004 的连锁不平衡分析发现 10q26 不包含任何的蛋白编码基因^[7]。既往的多项研究表明, 10q26 区域在包括前列腺癌的多种肿瘤中是一个杂合性缺失区域^[15-16]。为了探讨 rs2252004 影响前列腺癌易感性的生物学机制, 本文旨在研究 rs2252004 能否作为 meQTL 调控附近基因的甲基化状态, 从而影响关键基因的表达参与前列腺癌的发生。

分析结果显示, rs2252004 影响位于 FGFR2 基因内含子区的 3 个 CpG 位点的甲基化水平。位于 10q26 区域的 FGFR2 能够通过选择性剪切编码两种亚型 FGFR2b 和 FGFR2c, 两者都作为纤维母细胞生长因子受体而发挥作用^[17]。FGFR2b 和 FGFR2c 在不同组织表达有所不同, 且具有配体差异性。FGFR2b 主要作用于上皮细胞, 作为 FGF1、FGF3 等的受体, 而 FGFR2c 与间质细胞中 FGF1、FGF2 具有较高的亲和力^[18]。多项研究表明 FGFR2b 作为抑癌基因参与了前列腺癌和膀胱癌的肿瘤形成过程^[19-20]。我们对 TCGA 中 50 例配对的前列腺癌组织和癌旁组织的 FGFR2 基因进行差异性表达分析, 结果发现前列腺癌组织的 FGFR2 水平显著低于癌旁组织。进一步探讨 rs2252004 不同基因型的 FGFR2 基因表达差异, 发现与 GG 基因型相比, GT 基因型的前列腺癌患者的 FGFR2 基因表达量显著降低。GG 基因型与 GT/TT 基因型患者的 FGFR2 表达量差异有统计学意义。而 TT 与 GG 基因型之间的基因表达差异无统计学意义, 可能由于 TT 基因型的前列腺癌样本太少, 后期还需要更大样本进一步验证结果。

综上所述, GWAS 报道的前列腺癌易感性位点 rs2252004 可能通过影响 FGFR2 甲基化水平, 进而影响 FGFR2 基因的表达, 从而参与前列腺癌的发生。

【参考文献】

[1] Jemal A, Bray F, Center MM, et al. (下转第 87 页)

- 54(9):935-944
- [9] 帕力达·阿不力孜,周晓辉.代谢综合征与良性前列腺增生的相关性[J].中国动脉硬化杂志,2011,19(11):937-940
- [10] 罗文娟.代谢综合征与良性前列腺增生的关系[J].医学综述,2013,19(9):1650-1652
- [11] Gacci M, Corona G, Vignozzi L, et al. Metabolic syndrome and benign prostatic enlargement: a systematic review and meta-analysis[J]. *BJU Int*, 2015, 115(1):24-31
- [12] Pashootan P, Ploussard G, Cocaul A, et al. Association between metabolic syndrome and severity of lower urinary tract symptoms (LUTS): an observational study in a 4666 European men cohort [J]. *BJU Int*, 2015, 116(1):124-130
- [13] Agrawal A. Metabolic syndrome and BPH: What do we know?[J]. *Med J Armed Forces India*, 2017, 73(1):102-103
- [收稿日期] 2017-02-23

(上接第33页)

- Global cancer statistics [J]. *CA Cancer J Clin*, 2011, 61(2):69-90
- [2] Torre LA, Bray F, Siegel RL, et al. Global cancer statistics, 2012 [J]. *CA Cancer J Clin*, 2015, 65(2):87-108
- [3] Lichtenstein P, Holm NV, Verkasalo PK, et al. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland [J]. *N Engl J Med*, 2000, 343(2):78-85
- [4] Hindorf LA, Sethupathy P, Junkins HA, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits [J]. *Proc Natl Acad Sci U S A*, 2009, 106(23):9362-9367
- [5] Takata R, Akamatsu S, Kubo M, et al. Genome-wide association study identifies five new susceptibility loci for prostate cancer in the Japanese population [J]. *Nat Genet*, 2010, 42(9):751-754
- [6] Xu J, Mo Z, Ye D, et al. Genome-wide association study in Chinese men identifies two new prostate cancer risk loci at 9q31.2 and 19q13.4 [J]. *Nat Genet*, 2012, 44(11):1231-1235
- [7] Akamatsu S, Takata R, Haiman CA, et al. Common variants at 11q12, 10q26 and 3p11.2 are associated with prostate cancer susceptibility in Japanese [J]. *Nat Genet*, 2012, 44(4):426-429
- [8] Wang ML, Takahashi A, Liu F, et al. Large-scale association analysis in Asians identifies new susceptibility loci for prostate cancer [J]. *Nat Commun*, 2015, 6:8469
- [9] Rockman MV, Kruglyak L. Genetics of global gene expression [J]. *Nat Rev Genet*, 2006, 7(11):862-872
- [10] Rushton MD, Reynard LN, Young DA, et al. Methylation quantitative trait locus analysis of osteoarthritis links epigenetics with genetic risk [J]. *Hum Mol Genet*, 2015, 24(25):7432-7444
- [11] Gu X, Huang T, Xu D, et al. Association of a common variant at 10q26 and benign prostatic hyperplasia aggressive-ness in han Chinese descent [J]. *Biochem Res Int*, 2013, 2013:820849
- [12] Rivas MA, Beaudoin M, Gardet A, et al. Deep resequencing of GWAS loci identifies Independent rare variants associated with inflammatory bowel disease [J]. *Nat Genet*, 2011, 43(11):1066-1073
- [13] Nicolae DL, Gamazon E, Zhang W, et al. Trait-Associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS [J]. *PLoS Genet*, 2010, 6(4):e1000888
- [14] Hancock DB, Wang JC, Gaddis NC, et al. A multiancestry study identifies novel genetic associations with CHRNA5 methylation in human brain and risk of nicotine dependence [J]. *Hum Mol Genet*, 2015, 24(20):5940-5954
- [15] Katoh M, Katoh M. FGFR2 and WDR11 are neighboring oncogene and tumor suppressor gene on human chromosome 10q26 [J]. *Int J Oncol*, 2003, 22(5):1155-1159
- [16] Ittmann M. Allelic loss on chromosome 10 in prostate adenocarcinoma [J]. *Cancer Res*, 1996, 56(9):2143-2147
- [17] Dionne CA, Crumley G, Bellot F, et al. Cloning and expression of two distinct high-affinity receptors cross-reacting with acidic and basic fibroblast growth factors [J]. *EMBO J*, 1990, 9(9):2685-2692
- [18] Zhang X, Ibrahimi OA, Olsen SK, et al. Receptor specificity of the fibroblast growth factor family. The complete mammalian FGF family [J]. *J Biol Chem*, 2006, 281(23):15694-15700
- [19] Matsubara A, Kan M, Feng SJ, et al. Inhibition of growth of malignant rat prostate tumor cells by restoration of fibroblast growth factor receptor 2 [J]. *Cancer Res*, 1998, 58(7):1509-1514
- [20] Ricol D, Cappellen D, El Marjou A, et al. Tumour suppressive properties of fibroblast growth factor receptor 2-IIIb in human bladder cancer [J]. *Oncogene*, 1999, 18(51):7234-7243
- [收稿日期] 2016-01-27