

· 基础研究 ·

同源序列对CYP2A13基因SNP研究的影响

滑峰 万海粟 梅朝蓉 郑德杰 孙琳琳 陈军 刘红雨 周清华

【摘要】背景与目的 已有研究表明：细胞色素P450酶2A13 (cytochrome P450 2A13, CYP2A13) 在单核苷酸多态性 (single nucleotide polymorphisms, SNP) 与疾病关联中起重要作用。细胞色素P450酶是一组同工酶，基因序列间有高度的同源性，可能对SNP分析产生影响。本研究初步探讨同源序列对CYP2A13基因SNP研究的影响。方法 应用Taqman探针检测573例人群rs8192789位点的分布，采用BLAST方法分析引物结合序列。对60例样本的CYP2A13序列进行测序，进一步进行了TA克隆测序，应用BLAST方法分析了克隆测序结果。结果 rs8192789位点在573例人群中只有3例为TT纯合型，其余均为CT杂合型，BLAST分析为同源序列导致。60例人群CYP2A13部分序列测序结果完全一致，有大量套峰，101氨基酸位点处没有dbSNP数据库中报道的SNP位点。克隆测序为247 bp、235 bp两个片段。结论 CYP2A13的同源序列对SNP研究造成了干扰，部分SNP位点可能是不存在的。

【关键词】 细胞色素P450代谢酶2A13；单核苷酸多态性；同源序列

【中图分类号】 R734.2 DOI: 10.3779/j.issn.1009-3419.2010.02.02

Interference of Homologous Sequences on the SNP Study of CYP2A13 Gene

Feng HUA, Haisu WAN, Chaorong MEI, Dejie ZHENG, Linlin SUN, Jun CHEN, Hongyu LIU, Qinghua ZHOU

Tianjin Key Laboratory of Lung Cancer Metastasis and Tumor Microenvironment, Tianjin Lung Cancer Institute,

Tianjin Medical University General Hospital, Tianjin 300052, China

Corresponding author: Qinghua ZHOU, E-mail: zhouqh1016@yahoo.com.cn

【Abstract】 Background and objective It has been proven that cytochrome P450 enzyme 2A13 (CYP2A13) played an important role in the association between single nucleotide polymorphisms (SNP) and human diseases. Cytochrome P450 enzymes are a group of isoenzymes, whose sequence homology may interfere with the study for SNP. The aim of this study is to explore the interference on the SNP study of CYP2A13 caused by homologous sequences. **Methods** Taqman probe was applied to detect distribution of rs8192789 sites in 573 subjects, and BLAST method was used to analyze the amplified sequences. Partial sequences of CYP2A13 were amplified by PCR from 60 cases. The amplified sequences were TA cloned and sequenced. **Results** For rs8192789 loci in 573 cases, only 3 cases were TT, while the rest were CT heterozygotes, which was caused by homologous sequences. There are a large number of overlapping peaks in identical sequences of 60 cases, and the SNP of 101 amino acid site reported in the SNP database is not found. The cloned sequences are 247 bp, 235 bp fragments. **Conclusion** The homologous sequences may interfere the study for SNP of CYP2A13, and some SNP may not exist.

【Key words】 Cytochrome P450 2A13; Single nucleotide polymorphisms; Homologous sequences

This study was partly supported by grants from the National Eleventh-Five-Year Key Task Project of China (to Qinghua ZHOU)(No.2006BAI02A01), National 863 Project (to Qinghua ZHOU)(No.2006AA02401) and Tianjin Scientific Supporting Project, China-Sweden Cooperative Foundation (to Qinghua ZHOU)(No.09ZCZDSF04100).

单核苷酸多态性是新一代遗传标记，其与人类疾病的关系是遗传学研究的重要内容^[1-4]。细胞色素P450酶2A13 (cytochrome P450 2A13, CYP2A13) 是人呼吸系统中

最常见的细胞色素P450酶，可代谢烟草中前致癌物，因此被认为与肺癌等肿瘤的发生有关^[5,6]。近年有较多研究者探讨了CYP2A13不同单核苷酸多态性位点与肺癌的关系，得出了有意义的结论，使得关于CYP2A13单核苷酸多态性的研究成为一个活跃的领域^[7-10]，从相关核酸序列数据库中已经可以查到大量不同研究者报道的SNP数据。我们在研究CYP2A13单核苷酸多态性与肺癌关联性实验中发现了同源序列的干扰，查阅比对了不同数据库 (dbSNP、EGPsnp、HAPMAP) 中的信息，发现部分信息相互矛盾，于是进一步做了克隆测序，以观察同源序

本研究受国家“十一五”科技攻关项目 (No.2006BAI02A01)、国家“863”项目 (No.2006AA02401)、天津市科技支撑计划中瑞合作项目 (No.09ZCZDSF04100) 项目资助

作者单位：300052 天津，天津医科大学总医院，天津市肺癌研究所，天津市肺癌转移与肿瘤微环境重点实验室 (通讯作者：周清华，E-mail: zhouqh1016@yahoo.com.cn)

列是否影响SNP研究，以期引起研究者的注意，并能正确参考应用数据库中的数据。

1 材料与方法

1.1 研究对象 原发性肺癌266例，正常体检人群307例，来源于天津医科大学总医院，均为汉族人群。收集样本人群外周静脉血2 mL，采用Axygen外周血DNA提取试剂盒提取基因组DNA备用，检测DNA浓度为40 μg/mL左右。

1.2 CYP2A13基因rs8192789位点 (Arg257Cys) 基因分型及序列比对 探针及引物由ABI公司设计合成，上游引物：GGAGGACTTCATCGCCAAGAA，下游引物：CGGATGAGAAAGGAGTTCGATGA，VIC标记探针：AGCGTGCGCTGGTT，FAM标记探针：AGCGTGCACTGGTT。实时定量PCR反应体系为10 μL，含探针0.25 μL，mix 5 μL，H₂O 2.75 μL，基因组DNA 2 μL。反应条件为：94 °C预变性10 min，92 °C、15 s，60 °C、1 min，共40个循环。结果判断：根据PCR曲线，如为VIC标记探针S形曲线，则基因型为CC型，如为FAM标记探针S形曲线，则基因型为TT型，如为双重S形曲线则为CT型。检测仪器为ABI 7500 Real Time PCR system，分

析软件为SDS V1.4。每个检测单元均设立阴性对照及两个以上重复样本，选择10%样本复测以进行实验质量控制。根据ABI公司设计引物应用BLAST进行序列比对。

1.3 CYP2A13基因序列测序分析及比对 设计上游引物：CGGACATGATGCCGTCAA，下游引物：GAGGCCACGATGAAGGGAG AT (由赛百盛公司合成)，PCR扩增反应体系：Taq (5 U/μL) 0.2 μL，10× PCR Buffer 5 μL，dNTP 5 μL，模板DNA 2 μL，引物各1 μL，加灭菌蒸馏水至50 μL体系。PCR反应条件：95 °C预变性7 min，95 °C变性45 s，58 °C退火45 s，72 °C延伸50 s，72 °C再延伸7 min，共30个循环。采用TA克隆试剂盒 p-EASY-T1 Cloning Kit (购自TransGen Biotech)，将扩增产物克隆于p-EASY-T1载体中，转染至大肠杆菌感受态细胞，利用蓝白斑及抗生素双重筛选，选择30个白斑过夜摇菌后取菌液送测序。

2 结果

2.1 rs8192789位点 (Arg257Cys) 基因分型结果 rs8192789位点在573例人群中只有3例为TT纯合型 (0.5%)，其余均为CT杂合型 (图1)。复测结果完全一致，计算分型结果等位基因C 0.5，等位基因T 0.5。应用BLAST比对了

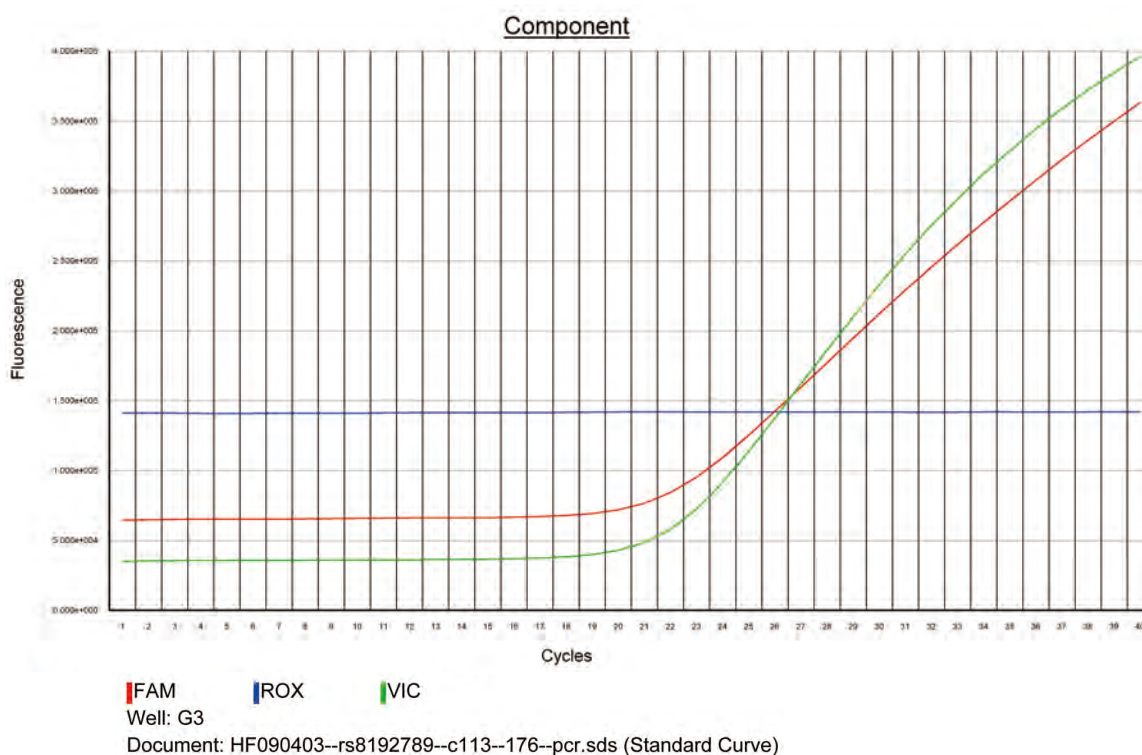


图1 CYP2A13基因rs8192789位点实时荧光PCR曲线

Fig 1 The real time fluorescence PCR curve for the genotype of CYP2A13 rs8192789

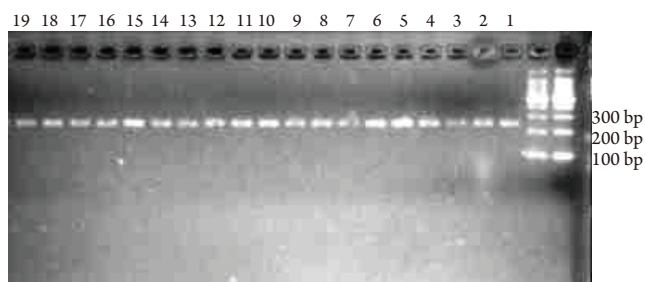


图2 PCR扩增目的片段的琼脂糖电泳图 (247 bp)

Fig 2 Agarose gel electrophoresis of PCR products (247 bp)

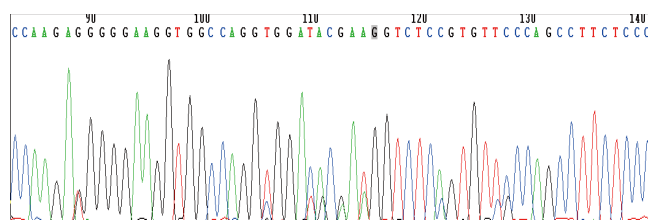


图3 PCR产物测序图

Fig 3 Sequence of PCR products

引物结合区序列,发现在CYP2A13基因3' 64 kb处有同源序列,同源性97%。

2.2 CYP2A13基因序列分析结果 PCR扩增目的片段247 bp (图2),60个样本测序完全一致,有大量套峰,其中101氨基酸位点处没有dbSNP数据库中报道的SNP位点(图3)。将目的片段克隆,随机挑取30个白斑,摇菌后菌液送测序,结果示247 bp、235 bp两个序列,经BLAST分析分别为来源于CYP2A13、CYP2A6基因的序列(图4A, B)。

3 讨论

CYP2A13是P450 II家族最新发现的一个酶,被认为是人呼吸系统中最常见的细胞色素P450酶,也是最主要的代谢烟草中前致癌物以及黄曲霉素的P450酶^[5,11-14]。关于CYP2A13多态性与肺癌的关联研究是当前一个活跃的领域,从报道的结果看,结论一致而且OR值达到2倍以上,这表明CYP2A13多态性在肺癌遗传研究中非常值得关注^[7-10]。CYP2A13基因位于人19号染色体长臂的CYP2A-T基因簇,全长7.3 kb,包括9个外显子、8个内含子,与CYP2A6、CYP2A7序列具有高度同源性,编码蛋白均为494个氨基酸。由于色素氧化酶P450是一大类参与内源性和外源性化合物代谢的酶,因此CYP2A13与其它P450酶也起码具有40%以上的序列同源性^[15]。关于CYP2A13单核苷酸公布的数据,在dbSNP数据库中上传了139个SNP位点,Hapmap数据库于2009年2月公布的

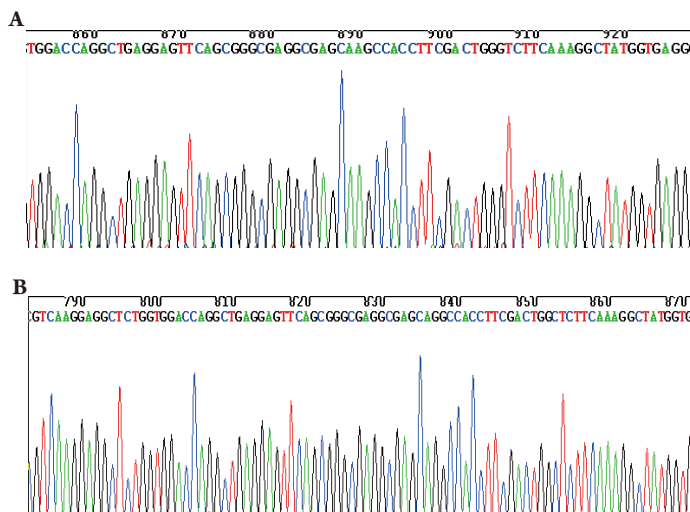


图4 PCR产物克隆测序图

A: CYP2A13序列; B: CYP2A6序列。

Fig 4 Sequence of cloned products

A: CYP2A13 sequence; B: CYP2A6 sequence.

最新数据报告了6个SNP位点,全部包括在dbSNP中,EGPSNP数据库暂未收录该基因数据。

我们选择rs8192789位点定制了Taqman探针,该位点位于257位氨基酸位点,密码子的第一个碱基为T/C,对应氨基酸分别为色氨酸、精氨酸。该位点在dbSNP数据库中由Hapmap上传的数据为单一的CT杂合子,而该位点在Hapmap数据库2009年2月公布的最新数据中已经删除。Wang^[9]利用限制性酶切的方法研究了791例中国汉族正常人群rs8192789位点的基因型频率,其中CC、CT、TT频率分别为82.4%、16.4%、1.2%。我们分型的结果是单一的CT杂合子,同dbSNP数据库中的频率是一致的,从遗传的角度来讲,这种基因型分布实际上是不可能的。利用引物结合区域序列进行了BLAST比对,发觉引物结合序列同位于该基因簇的一段假基因有97%同源性,我们认为该同源序列导致的。虽然Hapmap更新的数据库删除了rs8192789位点,但我们认为这一位点SNP是应该存在的。60例样本人群的CYP2A13序列测序完全一致,有大量套峰集中分布,包括了dbSNP数据中的rs3826712位点,进一步将序列进行了克隆测序,结果表明为CYP2A6干扰所致,我们认为rs3826712位点可能是不存在的,另外dbSNP数据库中也没有发现rs72552266位点。

造成SNP假阳性的原因除了分型技术误差外,另一个原因就是由于同源序列导致的干扰,这也可能是导致推断人类SNP出现频率差异的主要原因。通过对CYP2A13序列的初步分析,我们认为在dbSNP数据库中的139个

SNP位点部分是由于同源序列导致的假阳性的结果。Hapmap数据库中包含了中国汉族人群的SNP分布,是研究SNP重点参考的数据库^[16],但由于工作量庞大,其发布的数据可能也会有疏漏,因此我们在参考这些相关数据时要重点考虑到同源序列导致的对结果的干扰,以便正确参考应用相关数据。

致谢 天津医科大学总医院肺癌研究所白云、孙丽亚等对本研究给予的帮助,谨致谢意。

参 考 文 献

- Li J, Pan YC, Li YX, *et al.* Analysis and application of SNP and haplotype in the human genome. *Yichuan Xue Bao*, 2005, 32(8): 879-889.
- Giacomini KM, Brett CM, Altman RB, *et al.* The pharmacogenetics research network: from SNP discovery to clinical drug response. *Clin Pharmacol Ther*, 2007, 81(3): 328-345.
- Sobrinho B, Brion M, Carracedo A. SNPs in forensic genetics: a review on SNP typing methodologies. *Forensic Sci Int*, 2005, 154(2-3): 181-194.
- Suh Y, Vijg J. SNP discovery in associating genetic variation with human disease phenotypes. *Mutat Res*, 2005, 573(1-2): 41-53.
- Zhang X, D'Agostino J, Wu H, *et al.* CYP2A13: variable expression and role in human lung microsomal metabolic activation of the tobacco-specific carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. *J Pharmacol Exp Ther*, 2007, 323(2): 570-578.
- Su T, Bao Z, Zhang QY, *et al.* Human cytochrome P450 CYP2A13: predominant expression in the respiratory tract and its high efficiency metabolic activation of a tobacco-specific carcinogen, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. *Cancer Res*, 2000, 60(18): 5074-5079.
- D'Agostino J, Zhang X, Wu H, *et al.* Characterization of CYP2A13*2, a variant cytochrome P450 allele previously found to be associated with decreased incidences of lung adenocarcinoma in smokers. *Drug Metab Dispos*, 2008, 36(11): 2316-2323.
- Timofeeva MN, Kropp S, Sauter W, *et al.* CYP450 polymorphisms as risk factors for early-onset lung cancer: gender-specific differences. *Carcinogenesis*, 2009, 30(7): 1161-1169.
- Wang H, Tan W, Hao B, *et al.* Substantial reduction in risk of lung adenocarcinoma associated with genetic polymorphism in CYP2A13, the most active cytochrome P450 for the metabolic activation of tobacco-specific carcinogen NNK. *Cancer Res*, 2003, 63(22): 8057-8061.
- Cauffiez C, Lo-Guidice JM, Quaranta S, *et al.* Genetic polymorphism of the human cytochrome CYP2A13 in a French population: implication in lung cancer susceptibility. *Biochem Biophys Res Commun*, 2004, 317(2): 662-669.
- He XY, Tang L, Wang SL, *et al.* Efficient activation of aflatoxin B1 by cytochrome P450 2A13, an enzyme predominantly expressed in human respiratory tract. *Int J Cancer*, 2006, 118(11): 2665-2671.
- von Weyarn LB, Chun JA, Hollenberg PF. Effects of benzyl and phenethyl isothiocyanate on P450s 2A6 and 2A13: potential for chemoprevention in smokers. *Carcinogenesis*, 2006, 27(4): 782-790.
- Zhu LR, Thomas PE, Lu G, *et al.* CYP2A13 in human respiratory tissues and lung cancers: an immunohistochemical study with a new peptide-specific antibody. *Drug Metab Dispos*, 2006, 34(10): 1672-1676.
- Ling G, Wei Y, Ding X. Transcriptional regulation of human CYP2A13 expression in the respiratory tract by CCAAT/enhancer binding protein and epigenetic modulation. *Mol Pharmacol*, 2007, 71(3): 807-816.
- Smith BD, Sanders JL, Porubsky PR, *et al.* Structure of the human lung cytochrome P450 2A13. *J Biol Chem*, 2007, 282(23): 17306-17313.
- Phillips C. Online resources for SNP analysis: a review and route map. *Mol Biotechnol*, 2007, 35(1): 65-97.

(收稿: 2009-12-10 修回: 2010-01-05)

(本文编辑 南娟)

· 启 事 ·

解基严、周清华主译《心胸外科学精要》已出版

该书由国际最权威的心胸外科专家所著。英文原版书一经问世,即受到广大心胸外科医师的喜爱,拥有众多读者。

作者在编写上独具匠心,使得本书秉承了经典教科书的传统和临床手术学的风格,取二者之精华。全书分为3篇,共101章,全面而系统地论述了胸部外科、成人心脏病的外科治疗和先天性心脏病的外科治疗等内容。

此次引进的是该书的最新版本。在第一版的基础上,作者又添加了近几年心胸外科领域的最新理论、临床决策及手术方式。包括各种内窥镜手术、微创手术、肺移植、心肺移植、机器人手术、小儿心脏移植、小儿瓣膜手术以及临床数据库建立等。此外,每章结尾都配有编者评述,对该章所涉及的各种学术观点、最新进展及临床实际问题做出精辟的分析及点评,并附有权威性参考文献,以拓展读者的知识面。

本书对于承担临床繁忙工作的一线心胸外科医师而言,不失为一本图文并茂、条理清晰、实用性强的案头书,对初涉心胸外科领域的青年医师,也是一部极好的专业参考书。它可以帮助读者在较短时间内理解心胸外科相关的基础理论、各种学术观点、实用技术及发展趋势。

本书由北京大学人民医院心脏中心副主任解基严教授和天津医科大学总医院周清华教授担任主译,由国内多家医院的心胸外科界同仁共同翻译。相信这部著作的引进出版,会使我国广大心胸外科医师从中获益,并推动我国心胸外科技术向更先进的水平发展。