

饮用水源致癌风险早期预警 SNPs 监测技术

孔 界,艾山江·亚琛,张孝林,赵大勇,吴 兵,张 宴,张徐祥,程树培
(污染控制与资源化研究国家重点实验室,南京大学环境学院,南京大学环境健康研究所,江苏 南京 210093)

摘 要:综述了饮用水源污染致癌早期预警的单核苷酸多态性(single nucleotide polymorphisms, SNPs)检测技术,内容包括 SNPs 技术及其应用研究领域,癌症敏感基因与 SNPs,饮用水源致癌早期预警 SNPs 监测技术等。该技术的应用对于早期诊断、预警、预防、控制饮用水源污染致癌,保护人体健康,保障后代生命安全,具有现实意义。
关键词:饮用水源;污染致癌;敏感基因;SNPs;早期预警
中图分类号:X835 **文献标识码:**A **文章编号:**1674-6732(2010)-01-0012-05

SNPs Based Early Warning Technology for Testing Carcinogenesis in Source Drinking Water

KONG Jie, YACHEN Ai-shanjiang, ZHANG Xiao-lin, ZHAO Da-yong, WU Bing, ZHANG Yan, ZHANG Xu-xiang, CHEN Shu-pei
(State Key laboratory of Pollution Control and Resource Reuse, School of the Environment at Nanjing University, Nanjing University Institute for Environmental Health, Nanjing, Jiangsu 210093, China)

ABSTRACT: The development of the single nucleotide polymorphisms (SNPs) technique for early warning of carcinogenesis in source drinking water was examined, including (1) SNPs technique and its application, (2) sensitive cancer genes and measurement of SNPs; (3) early warning of carcinogenesis in source drinking water with SNPs technique. The technique proved to be effective for early diagnosis and warning, prevention and control of cancers induced by the pollutants from source drinking water. It is of great significance to protection of human health and life safety for the future generations.
KEY WORDS: source drinking water; carcinogenesis; sensitive cancer genes; SNPs; early warning

1 SNPs 技术及其应用研究领域

1.1 SNPs 定义与范围

所有生命的基因,由4种不同的单核苷酸 SN (single nucleotide)组成。4种不同的 SN,就是4种不同的 DNA。4种不同的 DNA 含有相同的五碳糖和磷酸根,区别是分别含有 A、T、G、C 4种不同的碱基。在科学研究资料之中,4种不同的碱基通常可以作为4种 DNA 的简称。单核苷酸多态性(SNPs)是指基因分子中4种单核苷酸数量和排序发生变化的分子形态多样性。

自然存在的单核苷酸多态性(SNPs),决定了生命个体及生物种群的差异。已经用于法医鉴定,亲子鉴定,种族区划,分子药物设计,疾病个性化治疗,流行病调查,临床诊断等应用研究领域。由环境污染物引发的 SNPs,则可致使人类基因分子的单核苷酸发生转换、颠换、插入、缺失等多种变化,

导致正常基因的密码错乱,激活致癌基因,诱发致病基因,破坏生命基因系统的平衡,威胁人体的健康安全。

1.2 SNPs 数量与类型

不同人群的基因,有不同的 SNPs;不同个人的基因,也有不同的 SNPs。人类基因组的 SN 总数约30亿个。由美国126版 NCBI(National Center of Biotechnology Information)数据库定义的人类 SNPs 已接近3千万个。玉米 SNPs 频率更高,在3'非翻译区和编码区的 SNPs 频率,分别为1:48个碱基对和1:130个碱基对^[1-5]。根据 SNPs 在基因分子中位置的差异,可以将 SNPs 分为两类:编码蛋白

收稿日期:2009-05-12;修订日期:2009-05-31
基金项目:江苏省环境监测科研基金项目[0711]
作者简介:孔界(1986—),男,在读研究生,研究方向为环境健康生物技术。

SNPs、非编码蛋白 SNPs。

编码蛋白的 SNPs, 位于功能基因的外显子中。如果 SNPs 的变化不引起编码蛋白的氨基酸改变, 则称为同义 SNPs, 反之称为非同义 SNPs。非同义 SNPs 可改变氨基酸表达的数量位置和蛋白质的结构功能。

非编码蛋白 SNPs, 位于功能基因的内含子中, 或功能基因的间区, 不改变编码蛋白的氨基酸序列, 不影响编码蛋白的结构功能。在自然生命的个体中, 非编码蛋白 SNPs 的频率, 远远高出编码蛋白的 SNPs。

位于基因调节区的 SNPs, 称为调节 SNPs, 也称为基因边缘 SNPs。如果调节 SNPs, 影响到基因转录表达 RNA 产物的过程, 就会影响到蛋白氨基酸的表达, 从而影响到生命的活动功能^[6-9]。

1.3 SNPs 检测技术与功能

SNPs 检测技术, 有杂交法、电泳法、溶解法、化学法、测序法、酶学法、物理法以及它们的组合方法。其基本原理过程是, 提取、分离、纯化、扩增待测基因的 DNA, 检测 DNA 中的 A、T、G、C 排序, 发现 DNA 序列中存在的 SNPs, 应用信息软件分析 SNPs 的生物效应。

由于 SNPs 技术具有高遗传的稳定性, 高密度的可测位点, 易于判别分析多态性等诸多优点, 已经成为研究基因转录表达的一种先进技术, 并广泛用于生命科学、医学、畜牧业、药理学、法医学、环境健康学等应用研究领域。

在生命科学和医学领域, SNPs 用于定位致病基因、检出易感基因等。在人类基因组计划 (HGP) 的实施期间, 科学家们同时启动了环境基因组计划 (EGP), 研究环境因素 (包括药物因素) 与相关疾病敏感基因的相互关系。SNPs 技术, 既是实施环境基因组计划的有力工具, 也是实施环境基因组计划的突破口。解读致癌致病的基因学原理, 是环境基因组计划的中心目标。

可以认为, 人类基因组计划提供了 SNPs 技术的发展基础, 而环境基因组计划促进了 SNPs 技术的成熟应用。在药物基因组学中, 应用检测 SNPs 的多态性遗传标记, 揭示不同个体对同种药物敏感性差异的根本原因, 为开展个性化的用药和药物的开发, 提供科学导向及分子依据。在此基础上, SNPs 技术为人类攻克致癌机理, 提供了新的技术途径。

SNPs 技术已经被广泛应用在法医鉴定领域,

弥补了原有 DNA 分析方法的不足。分析现场遗留物证的 SNPs, 可构建出当事者的人种及地域特征, 有助于侦查破案。此外, SNPs 还可用于个体识别、起源预测、亲权鉴定等。

病原微生物进入人体后, 其基因组与人体基因组相互作用, 引发遗传病的发生。不同的人含有不同的易感基因, 所以不同的人对同种病原微生物有不同的致病敏感性。利用 SNPs 可以寻找这些易感基因, 从基因水平上了解传染病的发病机制, 为预防治疗病原微生物引发的遗传病, 提供充分的遗传学依据。

在群体遗传学中, 注重研究基因低重组区及功能基因外的 SNPs 标记, 用于继续探索人类起源、进化以及遗传多样性。如研究高加索人、西班牙人、美国黑人的 SNPs 遗传多样性, 得出美国黑人主要源于欧洲和非洲的推论。

应用 SNPs 技术, 研究果蝇、猪、犬、鸡、羊、牛等动物, 已经建立了相关基因的 SNPs 数据库。未来的重点目标是, 研究 SNPs 对生命基因产生有用影响的技术预警。在畜牧业中, 检测猪氟烷、雌激素受体、生长激素受体等基因的 SNPs, 将有利于家畜家禽疾病的预防治疗和生长繁殖。

SNPs 技术的诞生, 震撼了多项科学领域。SNPs 技术的开发和应用, 将促进人类对生命奥秘的解读, 促进人类对疾病的了解与控制, 加快致病基因或易感基因的搜寻过程, 最终揭示生命的本质和致病原理, 为人类的健康带来福音。据科学引文索引报道, 在生命科学及医学领域, 国际上研究 SNPs 起始于 1973 年, 而最近 5 年的相关研究报道占总量的 74%, 进入高峰阶段, 其中有相当数量的中国研究报道, SNPs 的理论与技术已进入预防医学的主流前沿。

2 癌症敏感基因污染物与 SNPs

2.1 水体致癌污染物

水体中检出的有毒有机污染物已达上万多之多, 饮用水中检出的达 769 种, 其中包括致癌、促癌、致突变物质^[10]。联合国 2003 年《斯德哥尔摩 POPs 公约》首次提出 12 种 POPs 名单, 其中农药有狄氏剂、艾氏剂、异狄氏剂、毒杀芬、灭蚁灵、七氯、氯丹、DDT 共 8 种; 化学品有 PCBs 和六氯苯 2 种; 工业副产物有二噁英和苯并呋喃 2 种。这类污染物的主要特征是浓度低、难降解, 通过肌体的吸收

富集和代谢转化,产生持久的分子毒性,诱发癌症等环境疾病,威胁后代的生命质量。

2009年5月9日,联合国环境规划署与全球160多个国家和地区的代表,在日内瓦达成共识,将《斯德哥尔摩 POPs 公约》提出的12种 POPs,增至21种。新增的9种是:林丹、6-溴联苯、 α -六氯环己烷、 β -六氯环己烷、四溴联苯醚/五溴联苯醚/六溴联苯醚/七溴联苯醚、十氯酮、五氯苯、全氟辛酸磺酸、全氟辛酸磺酸盐/全氟辛基磺酰氟。

2006年,中国政府修订颁布了《生活饮用水卫生标准》(GB 5749—2006),将原有的35项监测指标,增加至106项。2007年开始执行,2012年强制执行,2015年实现目标。在增加的71项监测指标中,51项为有毒有机污染物。显然,中国政府为了保护饮用水源的环境安全,对控制有毒有机污染物,特别提出了更高更严的标准。

中国近年来饮用水源水质监测中,累计检出有机污染物504种,确切定量的有213种,其中卤代烃42种、胺类8种、苯系物27种、醚类3种、单环芳烃30种、呋喃3种、脂类21种、酮类4种、多环芳烃类22种、酚类17种、亚硝胺及其他34种^[11]。

此外,水体中还含有致癌致突变的内分泌干扰物、藻毒素、消毒副产物等有毒有机污染物^[12]。彭华等检测河南主要城市饮用水源水,得知多环芳烃类是普遍存在的有机污染物,其中的强致癌物有苯并(a)芘、萘蒽、芘等,苯并(a)芘明显超标^[13]。高继军等对北京市饮用水源水重金属污染物存在的健康风险进行了初步评价,得出饮用水中的化学致癌物As所引起的健康风险比Cd大,As为主要污染物的结论^[14]。重金属As、Cd、Hg污染等广泛地分布于自然环境中^[15-17]。

黄磊等检测评价长江三角洲地下水的健康风险,指出其中化学致癌污染物的危险度高出非致癌污染物3~4个数量级,必须在人们饮用之前去除化学致癌污染物^[18]。杨全锁等检测评价青岛市黄岛区饮用水源健康风险,同样得出化学致癌污染物健康风险高于非致癌污染物的结论,其中铬的健康危害高于砷^[19]。邓熙等分析了1991—1998年广州市饮用水源中硝酸盐、亚硝酸盐以及癌症死亡率的历史数据,得出饮用水源中的硝酸盐和亚硝酸盐可能是重要的致癌因子的结论^[20]。

2.2 污染致癌敏感基因

已经完成的环境基因组计划,已建立起癌症敏

感基因及其他疾病敏感基因的基因库。癌症敏感基因,是对环境因素(包括药物)敏感的癌症基因,如结肠直肠癌基因、乳腺癌基因、骨髓癌基因、血癌/白血病基因、胃癌基因、生殖系统癌基因、胰腺癌基因、食道癌基因、肺癌基因、皮肤癌基因、肝癌基因、鼻咽癌基因等。

癌症发生是一个极其复杂的过程。一种癌症的发生,往往是多种敏感基因激活的结果;一个癌症敏感基因,也往往涉及多种癌症的诱发,还与其他诸多的因素和基因关联。环境因素,特别是有毒有机污染物进入肌体,可产生持久的分子毒性,激活癌症敏感基因,诱发癌症。

2.3 癌症敏感基因 SNPs

癌症敏感基因的 SNPs 与癌症的发生有着一定的联系。监测致癌敏感基因 SNPs,对于饮用水源污染致癌的早期预警,对于维护水源地居民的健康,有着极为重要的意义。国内研究癌症敏感基因的 SNPs,主要集中在基因与癌症的诊断治疗医学领域,直接对癌症敏感基因上的 SNPs 位点进行研究已见报道。

2.3.1 肝癌基因 SNPs

汪莉萍等应用 RT-PCR 方法,检测49例原发性肝细胞性肝癌(HCC)患者肝癌组织 MAGE-A1 基因,及其转录表达 mRNA 的结果;提取其中43例 HCC 患者的肝癌组织、癌旁组织、外周血细胞基因组 DNA,PCR 扩增 MAGE-A1 基因 DNA;对所得 PCR 产物进行测序,得出 MAGE-A1 肝癌敏感基因中存在3种 SNPs^[21]。王娟等定位肝癌高频缺失区的肿瘤相关基因,检索 SNPs 数据库,获得编码区 SNPs(cSNP)序列;设计引物,根据 SNPs 位点设计寡核苷酸探针,构建 SNPs 芯片;分别从 HBV(乙型肝炎病毒)患者和正常人血样中提取基因组为扩增模板;PCR 扩增标记含 SNPs 位点的序列;将地高辛标记的 PCR 产物和 SNPs 芯片杂交;检测到3个 cSNP 位点的基因频率在两组人群中有显著性差异;证明 HBV 患者中的高频多态位点可能与其肝癌易感性相关^[22]。王娟和倪虹等从肝癌高频缺失区的基因入手,从 NCBI dbSNP 数据库查询获得基因的 cSNP 序列;根据 SNPs 位点设计寡核苷酸探针,研制多种肝癌相关基因芯片,为肝癌多态性标记的检测奠定了基础。

2.3.2 肺癌基因 SNPs

姜毅等运用逆转录 RT-PCR 方法,和限制性

片段长度多态性技术,结合 DNA 测序法,探索 caspase 9 (CASP 9) 基因在非小细胞肺癌中的表达,及其 SNPs 位点在非小细胞肺癌患者和正常人中的分布情况,得出 CASP 9 的 rs1052576 多态位点 G 等位基因与肺癌发生的相关性^[23,24]。吕美霞等研究了 DNA 损伤、DDB2 基因单核苷酸多态性与肺癌易感性的关系,得出 DDB2 基因 SNPs 位点 rs830083 C/G 变异,可能与肺癌遗传易感性相关,吸烟可显著增加肺癌发生的危险性的结论^[25]。经剑颖等研究,中国汉族人群谷胱甘肽 S 转移酶 A4 基因调控区 - 1718 位 T/A 多态(rs182623) 与肺癌遗传的易感性,首次发现 GSTA 4 基因调控区 - 1718 位 T/A 多态性,可能与汉族人群肺癌遗传易感性相关^[26]。韦卫琴等研究了基质金属蛋白酶启动子区 - 1607 bp 处 SNPs 与贵州汉族人非小细胞肺癌遗传易感性的关系^[27]。刘林林等研究了 p53

基因单核苷酸多态性与肺癌易感性之间的关系,对肺癌患者的放疗疗效及其预测提供指导意义^[28]。

中国研究癌症敏感基因的 SNPs,已经取得了一定的进展。从癌症敏感基因 SNPs 入手,研究癌症发生机理,设计分子药物,用于临床诊断治疗,是当今医学发展突破的重大目标。而应用 SNPs 技术,研究早期预警预防污染致癌症的原理方法,也是对环境健康责无旁贷的一项重大任务。

3 饮用水源污染致癌早期预警 SNPs 检测技术

在饮用水源污染致癌的早期预警领域,应用 SNPs 技术的研究也已进入高峰阶段。本课题组探索应用 SNPs 检测技术,对长江南京段饮用水源污染致癌物质进行了早期预警,相关检测过程列于表 1。

表 1 长江南京段饮用水源水对小鼠结/直肠癌基因 SNPs 检测过程

	步骤	目标内容	操作方法	结果与说明
前期 检测 发现 致癌 敏感 基因	1	检测长江南京段饮用水源水质	有机 GC - MS, 金属 ICP 等	PAHs 超 WHO 标准
	2	长江饮用水源水饲喂受试小鼠	自动饮水,标准化饲料	军区总院动物中心,90 d
	3	基因芯片检测受试小鼠转录组	基因组抽提、RT - PCR、分子杂交、扫描图谱	基因公司操作
	4	基因芯片检测转录组信息处理	列出散点中毒强度参数	应用相关的信息软件
	5	从产物中发现转录异常基因数	判别转录中毒强度≥2	223 个基因转录异常
	6	从转录异常基因中发现致癌基因	前 8 种癌症敏感基因排序	结/直肠癌,乳腺癌,骨髓瘤,白血病,胃癌,胰腺癌,生殖系统癌,肝癌
应用 SNPs 检测 CYP7A1 肠癌 基因	7	结/直肠癌基因为 SNPs 检测基因	检索 NCBI 的 GenBank	确定目标基因 CYP7A1
	8	检索结/直肠癌基因 SNPs	检索 NCBI 的 GenBank	确定目标基因的 SNPs
	9	设计 CYP7A1 的 PCR 引物	Primer 5.0 信息软件	扩增 655 bp 基因片断
	10	抽提受试小鼠肝脏基因组模板	UNIQ 柱式试剂盒	紫外检测基因组 DNA
	11	PCR 过程	Eppendorf - PCR 扩增仪	94 ℃,1 min;49 ℃,1 min; 72 ℃,2 min;35 次循环
	12	PCR 扩增 CYP7A1 片断与纯化	扩增 655 bp,Kit 纯化	基因公司操作
	13	分子克隆 CYP7A1 片断 655 bp	M13 载体	基因公司操作
	14	测序 CYP7A1 片断 DNA 655 bp	ddXTP	基因公司操作
	15	分析直肠癌基因 SNPs 检测结果	BLAST 软件及 GenBank	8 个 SNPs 点,2 段缺失

上述探索应用 SNPs 技术,检测结/直肠癌基因(CYP7A1)的结果是,8 个单核苷酸位点发生突变,同时缺失了 2 段单核苷酸。表明长江南京段饮

用水源水可能存在引发小鼠 CYP7A1 肠癌敏感基因转录的可能。尽管检测的 CYP7A1 基因仅为 655 个碱基对(bp),不足以说明整个肠癌敏感基因

的致癌风险,但课题组应用基因芯片的检测结果,与 SNPs 的检测结果相符,可以定性说明长江南京饮用水源存在激活肠癌敏感基因的分子毒性,可能已经威胁到人民健康。

癌症发生是一个复杂的过程,仍然需要长期探索。但是,癌症敏感基因的 SNPs 是诱导癌症发生的一项重要标志这一结论不会改变。进一步的研究目标是:研究建立快速简易的 SNPs 检测技术,为早期预警饮用水源的致癌毒性,保护人体环境健康提供技术支持,从预防医学的角度,建立与国际前沿接轨的先进监测技术。

[参考文献]

- [1] 杨昭庆, 洪坤学. 单核苷酸多态性的研究进展[J]. 国外医学:遗传学分册, 2000, 23(1): 4-8.
- [2] COOPER D N, SMITH B A, COOKE H J, et al. An estimate of unique DNA sequence heterozygosity in the human genome [J]. Human Genetics, 1985, 69(3): 201-205.
- [3] REICH D E, SCHAFFNER S F, DALY M J, et al. Human genome sequence variation and the influence of gene history, mutation and recombination [J]. Nature Genetics, 2002, 32(1): 135-142.
- [4] TENAILLON M I, SAWKINS M C, LONG A D, et al. Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays ssp mays* L.) [J]. PNAS, 2001, 98(16): 9161-9166.
- [5] RAFALSKI A. Applications of single nucleotide polymorphisms in crop genetics [J]. Current Opinion in Plant Biology, 2002, 5(2): 94-100.
- [6] COLLINS F S, GUYER M S, CHAKRAVARTI A. Variations on a theme: Cataloging human DNA sequence variation [J]. Science, 1997, 278(5343): 1580-1581.
- [7] LOOTS G G, LOCKSLEY R M, BLANKESPOOR C M, et al. Identification of a coordinate regulator of interleukins 4, 13 and 5 by cross-species sequence comparisons [J]. Science, 2000, 288(5463): 136-140.
- [8] TAO H, COX D R, FRAZER K A. Allele-specific KRT1 expression is a complex trait [J]. PLoS Genetics, 2006, 2(6): 93.
- [9] YING S Y, LIN S L. Intron-derived microRNAs-fine tuning of gene functions [J]. Gene, 2004, 342(1): 25-28.
- [10] 董亮, 封跃鹏. 环境样品中 PCBs 的测定[J]. 中国环境监测, 2002, 18(1): 35-38.
- [11] 王金应, 刘国尧. 水污染对人体健康危害的现状及对策研究[J]. 环境科学与技术, 2006, 29(B08): 80-81.
- [12] 王芳. 健康饮用水的探讨[J]. 科技信息, 2008(16): 48-49.
- [13] 彭华, 李明, 王玲玲, 等. 河南省主要城市饮用水源水中多环芳烃污染状况的研究[J]. 中国环境监测, 2004, 20(3): 17-19.
- [14] 高继军, 张力平, 黄圣彪, 等. 北京市饮用水源水重金属污染物健康风险的初步评价[J]. 环境科学, 2004, 25(2): 47-50.
- [15] CHAPPELL W R, BECK B D, BROWN K G, et al. Inorganic arsenic: a need and an opportunity to improve risk assessment [J]. Environmental Health Perspect, 1997, 105(10): 1060-1067.
- [16] MISLIN H, RAVERA O. Cadmium in the Environment [M]. New York: Princeton Architectural Press, 1986.
- [17] GORDON C, HUTCHINSON T. Global Perspective on Lead, Mercury and Cadmium Cycling in the Environment [M]. Elkins Park, PA: Franklin Book Company Inc. 1994.
- [18] 黄磊, 李鹏程, 刘白薇. 长江三角洲地区地下水污染健康风险评价[J]. 安全与环境工程, 2008, 15(2): 26-29.
- [19] 杨全锁, 郑西来. 青岛市黄岛区饮用水源健康风险评价[J]. 安全与环境学报, 2008, 8(2): 83-86.
- [20] 邓熙, 林秋奇, 顾继光. 广州市饮用水源中硝酸盐亚硝酸盐含量与癌症死亡率联系[J]. 生态科学, 2004, 23(1): 38-41.
- [21] 汪莉萍, 陈红松, 魏来, 等. 黑色素瘤抗原-A1 基因多态性及其在肝癌组织中的表达[J]. 肝脏, 2006, 11(2): 92-94.
- [22] 王娟, 倪虹, 陈力, 等. 肝癌染色体高频缺失区肿瘤相关基因在患者和正常人群中的多态分布研究[J]. 南开大学学报:自然科学版, 2006, 39(3): 1-5.
- [23] 王娟, 倪虹, 陈力, 等. 肝癌相关基因单核苷酸多态性基因芯片制备与检测分析[J]. 中华检验医学杂志, 2004, 27(10): 649-652.
- [24] 娄毅, 方长清, 李建华. CASP 9 基因在非小细胞肺癌中的表达及其多态性研究[J]. 中华医学遗传学杂志, 2007, 24(1): 59-62.
- [25] 吕美霞, 杨晓波, 白云, 等. DNA 损伤结合蛋白(DDB 2)基因单核苷酸多态性与肺癌易感性研究[J]. 疾病控制杂志, 2007, 11(5): 444-447.
- [26] 经剑颖, 蒋跃明, 胡志斌, 等. GSTA 4 基因调控区 T/A 多态与肺癌易感性之间的关联分析[J]. 肿瘤, 2007, 27(1): 39-42.
- [27] 韦卫琴, 梁显泉, 闻心培. MMP-中单核苷酸多态性与贵州汉族人的非小细胞肺癌易感性[J]. 贵阳医学院学报, 2007, 32(3): 267-272.
- [28] 刘林林, 吕文天, 邹国光. p53 基因单核苷酸多态性与肺癌易感性[J]. 中国实验诊断学, 2003, 7(6): 556-557.