

人类单核苷酸多态性研究进展

苏 津 综述,包娜娜,李会来,陈占国 审校(河北省秦皇岛市抚宁县中医院 066300)

【关键词】 单核苷酸; 多态性; 基因组; 研究方法

DOI:10.3969/j.issn.1672-9455.2014.07.057 文献标志码:A 文章编号:1672-9455(2014)07-0977-02

根据“人类基因组计划”,人类对自身基因组中的核苷酸序列进行了深入的了解与认识,这个计划的关键目标在于寻找功能基因,并且说明特定基因组区域里序列产生的差异^[1]。单核苷酸多态性(SNP)在基因组中的序列差异最为普遍,在多个学科的领域中具有重要的意义,发挥着巨大的作用。SNP 指在基因组水平中由单核的苷酸变异引发 DNA 的序列产生多态的变化,在人类遗传变异中最为常见,也是人类基因组计划后才进行研究的最新遗传标记。作者针对 SNP 的概念与特性,阐述其研究的价值与意义,并对研究方法进行综述,旨在探讨 SNP 在遗传学、生物医学以及人类进化史中发挥的作用。

1 SNP 的概念

SNP 作为基因组 DNA 中单碱基的序列差异,在人群中的分布率占 1% 以上,但是基因组里其他序列差异导致的变异不属于 SNP 范围,比如重复、缺失以及插入序列复制数目发生变化^[2]。另外,SNP 所引起的人类遗传基因多态性在遗传信息中的本质表现占 90% 以上。

从生物遗传的角度来说,SNP 特点包括:(1)将两个随机染色体加以对比,每一千个碱基里面就会有一个 SNP。(2)在基因内部中有部分 SNP 将对蛋白结构,或是表达水平造成直接影响,以利于充分了解疾病的遗传机制。(3)在进化史中,SNP 仅发生过一次变异,因此其遗传特性更加稳定。(4)高通量 SNP 筛选随着 SNP 研究成本降低以及研究方法的发展,将对遗传学分析起到促进与推动的支持作用^[3]。

2 SNP 的研究价值

2.1 疾病病因 在大多数人的概念中,每一种疾病均可发现遗传因素的影子^[4]。SNP 在基因组中的变异能够起到两个作用,一个是作为遗传学的一种分子标记物,另一个至关重要的作用就是能表达出基因和疾病之间的关系,对疾病中存在的有关基因进行搜索。对于部分单基因的遗传性疾病,采用家系研究方法对疾病的病因进行搜索的应用,在目前临床上已经取得了一定的成果。但是,在多基因疾病中,因为基因数目和影响疾病发生发展存在的差异,且环境因素也起到了重要作用,在病因学研究方面多基因疾病增加了很多的困难。SNP 研究的重点为通过 SNP 图谱的构建,为寻求和疾病有关的基因,开展全基因方面的研究,在现阶段研究里面主要采取的研究方法就是直接法、间接法。直接法通过寻找 cSNPs 及基因编码,获取对疾病易感性产生影响的基因序列变异,并且根据对现在研究的分析,在人类基因序列编码区域中 cSNPs 的数量只有 2、3 个变异,其频率在 10% 以上,而编码区域内并不包括全部的功能性变异,因此不能被广泛应用。间接法所寻找的目标就是和致病性相关的中性多态位点,而且对于氨基酸编码和蛋白的结构不会被改变,且不会使表现型改变。但是,在染色体方面有些疾病有关的基因突变位点将可能发生连锁效应,甚至还会遗传至后代^[5]。

2.2 药物基因组学 人类基因组框架序列于 2001 年建立完成,与此同时,诞生了以个体遗传学为主,运用医源性药物原理,对患者临床的药效和不良反应加以评估的一门药物基因组学^[6]。此学科利用基因组学的研究技术,比如生物学信息、化学信息、基因图谱、高通量 DNA 的检测序列,使得研究人员可以针对个体和群体间产生的药效差异予以准确遗传学基础的定性,主要在于对特定患者进行诊疗,防止药物造成不良反应及对药物靶点加以设计。通过个体化医疗手段对患者用药起到正面作用,并按照个体基因型的特异性对患者进行药物的优化配置,使得药物的临床疗效更为显著,不良反应更小。在药物基因组研究中,由于 SNP 位点的差异,将会对基因的表达、蛋白质结构、氨基酸的剪切以及编码产生巨大的影响,进而改变药物与受体以及转运体之间的相互作用,影响药物代谢^[7]。在不断构建高密度 SNP 及相关药物代谢的 SNP 图谱的情况下,医生了解患者基因型后,认真对患者遗传档案加以分析,并对患者进行适宜药物的应用,以此避免药物产生不良反应。目前,临床所采用的药物应用条件需要根据患者的体质量、年龄等资料加以确定,而在未来发展过程中,此方法会逐渐由个体遗传信息取代。

2.3 人类进化史 由于 SNP 对进化水平的变异有所体现,因此也可以对种群间的 DNA 序列变异进行研究。大多数 SNP 在基因组在蛋白质编码区域之外,如分布方面是不会受到选择的干扰,而且会长期保留下来。一些对蛋白编码区域内的表现型 SNP 产生影响,虽然会受到自然选择施加的压力,但若是可以长时间被保留下来,就能够对个体成功繁衍后代发挥促进作用,充分说明这些变异会受到选择的干扰,而且还会作为代表成为进化史中的部分关键事件^[9]。所以,根据对不同物种相互之间的蛋白编码及非编码区和变异数目得出的比值予以计算,能够对进化树当中的分支点加以追溯,还能促进进化,并产生积极影响的变异,这样才可以在基因池中被固定保留。

3 研究方法与技术手段

在目前的 SNP 研究期间,很多技术会随生物学实验技术的进步发展而加以应用,具体表现为下面几方面。

3.1 微阵列杂交实验 在 SNP 研究中可以将杂交实验方法作为研究的基础理论^[10]。将 SNP 位点作为中心,再向两侧延伸 10~20 个碱基作为位点特异性探针,然后杂交包括该位点不同个体产生聚合酶链反应(PCR)的物质。该特异性探针具有较小的长度,所以可以对含有 SNP 位点碱基和靶 DNA 相匹配,从而对形成探针,也就是靶 DNA 复合物的稳定性产生关键作用,同时据此对不同个体位于此位点时表现的不同基因型加以辨别和对比。这种实验方法具体有两种模式,第一种适用于在规模较大的人群当中对少量 SNP 位点进行筛选,方法为将待测的 DNA 样品进行固定;第二种可于同一时间对大批量 SNP 位点进行统一分析,方法为采用微阵列将高密度探针在

基片上进行固定,并且要制备数万个和 ASO 杂交以后 PCR 产物。如果可以实现这样的操作,应该将每个 SNP 位点之间所产生的 PCR 尽可能缩短,保证能够实现多重 PCR。而且在 PCR 引物中应该存在通用尾巴,以利于二次扩增时,增加引物的使用效率,提高不同产物 PCR 的相似率^[11]。当扩增完成以后,将点有 ASO 的基因芯片与包含多个 SNP 位点或是多重混合物的不同样品实验杂交,同时利用共聚焦显微镜技术检测荧光信号并记录,根据样本间的对比及对照成像对杂交结果详尽分析。此方法的问题主要是多重 PCR 具有一定程度的局限性,还有就是很多 DNA 样本和 ASO 探针相杂交时存在的复杂性及设备与设计高密度芯片的问题。

3.2 分子灯塔 这项技术基于 PCR 反应对 SNP 位点进行区分,并需要由 3 个部分组成的杂交探针,两端为由荧光染料进行标记并可以参照 SNP 位点碱基的变化进行相应标记的 DNA 序列,中间为 SNP 位点的特异性序列。如果探针没有和序列实施特异性杂交的情况下,探针两端序列可以互补,从而产生发卡结构,致使两侧荧光材料因为过度紧密而发生荧光淬灭的现象^[12]。如果探针和序列在特异性杂交的情况下,两种染料各自分离,致使荧光强度变为杂交前的 900 倍以上。此种方法比较简单便捷,但是探针杂交的稳定性却不能满足实验的预期效果。

3.3 测序 这种方法为国际上应用最为广泛的 SNP 研究方法,具有直接、准确对序列差异进行反映、检测成本逐渐降低等优势。常用的手段就是 PCR 扩增产物的测序和随机性测序。

3.3.1 PCR 扩增产物测序 该种手段经常被应用于验证及发现 SNP 中,利用 PCR 把可能调控区域和编码区域里面的序列表现于不同个体中加以扩增以后,再对扩增后的产物予以序列或是测序分析。这种方法仅对已知序列组区域内的 SNP 开展研究,成本比较昂贵且操作复杂,也不适合对 SNP 进行大规模的筛选。

3.3.2 随机测序 这种策略在目前的研究中流行两种方法,基因组比与不完全鸟枪法。前者采取的方法为对全基因组进行随机测序,并将其与人类基因组序列进行对比,将新发现的 SNP 在全基因组的序列图谱中进行定位。缺点在于,在发现 SNP 的结果仅为 2 条染色体序列之间的比较,易导致数据分析时随机误差的产生和引入测序错误的出现,致使其产生假阳性的结果。不完全鸟枪法关键是把不同个体内的 DNA 样品加以等量混合以后再构建基因文库^[13]。随机性把全基因组核苷酸序列打断为基础,获取全基因组内的微量片段,然后对其采取随机测序,以至于其可以成倍的覆盖,从而在不同颜色体间相互重叠的区域对序列差异加以验证和发现,以利于搜索一些人群中的 SNP。通常情况下,若是想把基因组内的部分区域取得比较大的覆盖度,这就需要进行大量的测序反应达到目的,但是因为随机测序方法内 SNP 选自不同个体构建的群体样品,加强了结果的可信度。

4 人类单核苷酸多态性的研究现状

根据现阶段国际标准,SNP 的试验研究可以分为 3 个阶段,分别是发现、验证及筛选。而现在研究的重点倾向于第一个阶段,而这一阶段只能够获得与 SNP 有关的少量信息,例如多态位点在基因当中、染色体当中的定位,以及碱基的变异等。对于这个阶段得到的 SNP,还要在小规模人群中进行验证,根据 SNP 的分布频率对常见或稀有的类型进行判断,以此来避免数据、实验处理等方面产生假阳性的结果^[14]。下一步实验

研究将选择更广阔的人群对 SNP 进行扫描,进一步确定不同人群中 SNP 的分布情况,同时寻找更低频率的 SNP。对于不同人群之间的关系以及迁移情况,SNP 的分布频率尤其是稀有 SNP 能够起到重要的提示作用。按照进化的观点可以总结出一个结论,常见 SNP 表明了变异相当久远,在人群中长久的保存下来,而且以固定的频率存在着;稀有的 SNP 表明,人群会在近期内出现变化。

参考文献

- [1] 吉强,顾星,高春芳. 甲胎蛋白及其单核苷酸多态性与肝细胞癌易感性的研究[J]. 检验医学,2013,28(2):168-170.
- [2] 张孝平,菅子莹,陈宝安,等. 应用 MALDI-TOFMS 检测 5 种白血病细胞株 DCK 和 CDA 基因部分单核苷酸多态性[J]. 癌变·畸变·突变,2013,25(2):115-119.
- [3] 张利,郭雄,王磊. VDR 基因单核苷酸多态性与大骨节病关联分析[J]. 医学研究杂志,2013,42(3):21-24.
- [4] 彭茜,陈昌辉,吴青,等. CASP3 基因单核苷酸多态性与中国儿童川崎病的相关性研究[J]. 中华医学遗传学杂志,2013,30(2):180-184.
- [5] 朱雄,石毅,鲁芳,等. CDPG2 和 HSPG2 基因单核苷酸多态性与中国汉族人颅内动脉瘤的相关性研究[J]. 中华医学遗传学杂志,2013,30(2):218-221.
- [6] 张钰,魏威,李欣. SMAD7 基因单核苷酸多态性与散发性胃癌、肠癌发病风险相关性研究[J]. 现代预防医学,2013,40(7):1327-1329.
- [7] 何晓,王小华,邱振纲,等. IL-22 基因单核苷酸多态性与肝癌易感性的相关性研究[J]. 赣南医学院学报,2013,33(1):29-32.
- [8] 陈广,李云超,邱虹,等. 白介素 18 基因启动子区-137G/C 单核苷酸多态性与颅内动脉瘤发病和破裂的关系研究[J]. 中国全科医学,2013,16(7):629-632.
- [9] 毛翀,熊俊浩,宁湧,等. CCDC122 基因单核苷酸多态性与彝族麻风相关性研究[J]. 中国麻风皮肤病杂志,2013,29(2):85-87.
- [10] 张再伟,丁世芳,赵亚玲,等. 中国人脂联素基因+276G/T 单核苷酸多态性与冠心病相关性研究的 Meta 分析[J]. 循证医学,2013,13(1):61-64.
- [11] 李博,胡秋侠,钟贵芳,等. 系统性红斑狼疮患者信号转导及转录激活因子 4 基因单核苷酸多态性与其 mRNA 表达水平的相关性研究[J]. 海南医学,2013,24(4):471-473.
- [12] 刘小琦. 采用 SNaPshot 方法对中国老年黄斑变性 C2 和 C3 基因单核苷酸多态性进行相关性研究[J]. 国际检验医学杂志,2013,34(3):285-287.
- [13] 胡嘉乐,黄婉娴,高洋,等. FTO 基因单核苷酸多态性与 2 型糖尿病发病风险的相关性研究[J]. 分子诊断与治疗杂志,2013,1(1):36-40.
- [14] 周松,朱泽章,邱勇,等. 汉族人群白介素 17 受体 C 基因单核苷酸多态性与青少年特发性脊柱侧凸的相关性[J]. 中国脊柱脊髓杂志,2013,23(2):161-165.