

基因多态性与疾病易感性的分子流行病学研究:问题与对策

沈洪兵

传统流行病学研究表明,在相同的环境暴露下,不同个体对疾病的易感性存在差异。研究表明,单核苷酸多态性(single nucleotide polymorphisms, SNPs)在阐明疾病的发生发展过程,个体对环境危险因素易感性与耐受性等方面都起着重要的作用。与基因多态性相关的分子流行病学研究将促进在分子水平上探讨环境-基因交互作用在疾病发生发展中的影响。有关基因多态性与疾病易感性的分子流行病学研究近年来在我国也已逐步展开并渐深入,虽然在数量和质量上均无法与国外相比,但也有自己的特色和优势,即中国人群不同的环境暴露特点和中华民族有别于其他种族的遗传背景特征。本文主要讨论和分析目前基因多态性在疾病分子流行病学研究中存在的主要问题,并希望能对将来的研究有所裨益。

1. 基因选择的问题:据估计,人类 30 亿碱基中大约有数百万个 SNPs 位点^[1]。因此,在疾病分子流行病学研究中如何进行基因多态性研究?在如此众多的基因及其 SNPs 位点中选择哪些基因?在同一基因中选择哪些 SNPs 位点进行的研究?这些已成为许多分子流行病学研究工作者共同关注的问题。其中的原则是:

(1)应着重研究参与疾病发生发展的具有共性的基因。以肿瘤为例,应关注癌基因、抑癌基因、调控细胞周期、细胞增殖和分化的基因等,因为这些基因在不同种类肿瘤的发生发展中均起着较为重要的作用,如 p53^[2]、cyclin D1^[3]、interleukin-1^[4]等。

(2)候选基因的选择要具有明确的病理生理学或生物化学基础,即要有明确的科研假设。例如,肺癌的主要危险因素为吸烟,烟雾中的多环芳烃等致癌物质可引起 DNA 损伤,因此,选择代谢酶基因

GSTM1、GSTT1、细胞色素 P450 家族基因、DNA 修复基因等研究其多态性与肺癌易感性的关系就比较科学合理。

(3)候选基因多态性具有潜在的功能影响,即所谓的“功能性 SNPs”。一个基因序列中往往有多个 SNPs 位点,但并不是所有的 SNPs 位点都会对该基因所编码的蛋白质或 mRNA 产生影响。因此,应优先选择那些能对其编码的蛋白质或 mRNA 产生影响的 SNPs 进行研究。例如优先选择在外显子和启动子区域的 SNPs,在外显子区域应优先选择能引起氨基酸改变的 SNPs 等。

(4)如所研究的目标基因 SNPs 位点较多时,可以以单倍型区域(haplotype block)为单位,每个区域选择 1 个有代表性的频率相对较高的 SNP 位点,每个基因约 5~8 个位点,且可以覆盖整个基因。

(5)对于多基因疾病,应根据不同疾病可能的发病机理,选择同一通路的多个候选基因或相互关联的不同通路的候选基因进行联合分析,既要研究某一基因某一 SNPs 位点的作用,更重要的是要对同一基因不同位点以及不同的候选基因之间的相互作用(即基因-基因交互作用及单倍型分析等)的整体影响进行分析^[5]。

2. 样本量的问题:关于基因多态性和疾病可能关联的文献每年均有大量发表,但研究结果中缺乏一致性甚至矛盾的现象较为普遍,其中一个重要原因是到目前为止大多数的研究样本量都太小。以病例对照研究为例,目前国内关于基因多态性的病例对照研究的平均样本量在 150~300 对。这样的样本量不足以判断人群中存在较罕见的 SNP(如频率 < 5%)或虽频率较高但其与疾病的联系较弱的 SNP($OR < 2$),但对于能增加 2 倍以上危险度的一般频率 SNP(如 20%)已足够了。然而,Meta 分析表明,即使某一 SNPs 确实能增加疾病的危险度,它的作用通常很缓和(为低共显性基因)。例如,最近

我们进行了一项包括 3725 例肺癌病例和 4152 例对照的 9 个病例对照研究的 Meta 分析,表明 DNA 修复基因 XPD 751CC 和 312AA 变异基因型与其野生型相比仅分别增加 21% 和 27% 的肺癌危险性,其 95% CI 分别为 2%~43% 和 4%~56%。因此,将来关于 SNPs 的研究,预计它确实增加疾病危险度的一般范围应该是 20%~60%,而不是从前研究中假设的 100%~200%,同时,为了判断如此小的危险度,将来的样本含量也应该足够大。

因此,为了进行可能由 SNPs 起主效应的研究,开始时一般建议要有 300 对以上的对子。如要进行相加或相乘交互作用的分析,常常需要 1000 对对子以上,在这样的样本下才可估计在某一亚群中这些 SNPs 的作用(即分层分析)。目前国外的一些研究已经开始这样的设计并已有论文发表,但这样的大样本对发病率较低的疾病进行起来较为困难,可以通过多中心联合研究来克服。

3. 研究设计问题:目前基因多态性与疾病易感性的分子流行病学研究设计大多采用传统的病例对照研究(包括巢式病例对照研究)方案,其原则有诸多相同之处,包括选择合适、可比的、具有人群代表性的对照和控制可能的选择偏倚。因此,在选择对照时应注意避免人群分层的影响,注意使病例和对照在种族特征、地区分布上的可比。在病例的选择时,最好选择新发病例而不是现患病例,以避免生存偏倚的影响。不合适的无代表性的研究人群可以直接导致错误的结论。例如,研究妇女乳腺癌时,如果选取的病例均来自乳腺癌高发的家庭,所得到的关于 BRCA1 和 BRCA2 基因与乳腺癌的关联就可能被夸大了,不能完全代表一般人群和散发病例的情况^[6,7]。

在低共显性基因 SNPs 和复杂性状疾病关联的分子流行病学研究中,对于混杂因素的控制十分重要。在这样的研究中,年龄、性别、环境暴露因素、疾病性状和病理组织类型等都有可能成为混杂因素,应在研究设计时加以控制。在统计分析阶段,通常可采用调整或分层分析的方法来控制。

因研究目的不同,近年来也出现了一些非传统病例对照研究,如单纯病例研究(case-only study)、病例同胞对研究(case-sibling study)等。因此,在进行分子流行病学研究设计时,应根据不同的研究目的,选用不同的研究设计方案。如以估计基因环境交互作用为主要目的,可推荐选用单纯病例研究方

法,其优点是所需样本量较小,精确度较高,可避免由于选择对照而可能产生的选择偏倚,缺点是不能估计主效应因素(基因、环境)单独与疾病的关系。

值得提出的是,在分子流行病学研究中不应过分强调基因 SNPs 或其他分子标志物的作用而忽略环境因素的影响,应同样注意收集准确可靠的暴露因素信息和宏观流行病学资料,宏观和微观并举,以综合评价基因-环境交互作用在疾病发生发展中的意义。

4. 基因型的错误分类问题:实验室基因型的分类错误(misclassification)也是一个引起错误结论的可能原因。在实验室进行基因型检测时,采用不同灵敏度和特异度的方法(如 PCR-RFLP、PCR-SSCP 或直接测序等)均有可能导致基因型不同程度的误分类,导致假阳性或假阴性结果^[8]。因此,准确的实验方法和实验室质量控制至关重要。一方面,可以通过盲法分型和一定比例(如 10%)样本的重复实验来控制,另一方面,可采用已明确基因型的样本作为内对照以减少实验室的系统误差。其他如固定实验技术人员、采用统一判断标准、病例和对照样本混合检测等避免信息偏倚的方法均可以减少潜在的基因型错误分类。当采用高通量基因分型技术时,在分型过程中应优先确认和使用平行样本或金标准对照。

5. 结果的解释问题:许多研究报道过分强调了来自分层分析的结果,即使未发现显著的主效应时也是如此,而这样的结果往往缺乏普遍意义,从而导致了来自不同研究的结果不一致。例如仅限于妇女、非暴露人群或限于某一病理类型(如鳞癌)的患者等。但这些报道忽略了一点,即观察的并非是一个整体的作用。显然,在一个未发现显著主效应的研究中,其中一个亚人群危险度的增加意味着在其余人群中危险度的相应降低,这样的研究结果一般在随后的报道中重复性较差,且由于分层亚群的样本数量更少,相关结论很可能是由于机会所致^[9]。因此,当缺乏很强的先验假设时,在鉴定某一亚人群是否真的增加了危险度之前,基因多态性研究结论的解释和推论应持谨慎态度。

6. 当前基因多态性研究的策略:作为总结,这里提出当前基因多态性分子流行病学研究的一些策略以供参考:

(1)科学的分子流行病学研究设计,足够的样本量,鼓励多中心联合研究。

(2) 选择有价值的基因及其多态性位点进行研 究,最好能筛选和发现新的未发表过的具有功能意义 的多态性,尤其是中国人群特有的基因多态性。

(3) 研究一个基因内多个 SNPs 和/或一段染色 体区域内多个 SNPs,关注 tagSNP 和单倍型区段 (haplotype block) 构建的研究进展,参与全基因组范 围内遗传标记和疾病的关联分析。

(4) 同一通路中或相关通路的多基因联合研究, 克服单一基因多态性研究的缺陷,有助于研究基因- 基因交互作用。

(5) 注重宏观与微观相结合,研究基因-环境交 互作用(gene-environment interaction)、基因型-表型 相互关系(genotype-phenotype correlation)。

(6) 评价并运用方便而经济的高通量技术进行 大样本的基因多态性研究。

参 考 文 献

1 Sachidanandam R, Weissman D, Schmidt SC, et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature*, 2001, 409:928-933.

2 Shen H, Zheng Y, Sturgis EM, et al. P53 codon 72 polymorphism and risk of squamous cell carcinoma of the head and neck: a case-control study. *Cancer Lett*, 2002, 183:123-130.

3 Zheng Y, Shen H, Sturgis EM, et al. Cyclin D1 polymorphism and risk for squamous cell carcinoma of the head and neck: a case-control study. *Carcinogenesis*, 2001, 22:1195-1199.

4 El-Omar EM, Carrington M, Chow WH, et al. Interleukin-1 polymorphisms associated with increased risk of gastric cancer. *Nature*, 2000, 404:398-402.

5 Fu YP, Yu JC, Cheng TC, et al. Breast cancer risk associated with genotypic polymorphism of the nonhomologous end-joining genes: a multigenic study on cancer susceptibility. *Cancer Res*, 2003, 63: 2440-2446.

6 Hopper JL, Southey MC, Dite GS, et al. Population-based estimate of the average age-specific cumulative risk of breast cancer for a defined set of protein-truncating mutations in BRCA1 and BRCA2. Australian Breast Cancer Family Study. *Cancer Epidemiol Biomarkers Prev*, 1999, 8:741-747.

7 Struwing JP, Hartge P, Wacholder S, et al. The risk of cancer associated with specific mutations of BRCA1 and BRCA2 among Ashkenazi Jews. *N Engl J Med*, 1997, 336:1401-1408.

8 Rothman N, Garcia-Closas M, Stewart WT, et al. The impact of misclassification in case-control studies of gene-environment interactions. *IARC Sci Publ*, 1999, 148:89-96.

9 Green J, Banks E, Berrington A, et al. N-acetyltransferase 2 and bladder cancer: an overview and consideration of the evidence for gene-environment interaction. *Br J Cancer*, 2000, 83:412-417.

(收稿日期 2003-12-25)

(本文编辑:张林东)

· 疾病控制 ·

广西灵川县两例全程接种狂犬病疫苗后发病死亡的调查分析

叶源彪

广西灵川县于 2003~2004 年间共发生狂犬病 9 例,其 中有 2 例全程接种狂犬病疫苗后发病死亡。

【病例 1】 67 岁男性,于 2003 年 9 月 10 日被犬咬伤手 臂,伤情不严重,当时到村医生处就诊,未进行伤口处理,只 接种狂犬病疫苗 5 支(疫苗来源不明),接种程序按照 0,3,7, 14,30 天进行,于同年 10 月 9 日全程接种完,10 月 13 日发 病,16 日死亡,病程 4 天。

【病例 2】 2 岁女性,于 2004 年 4 月 5 日被自家犬咬伤 面部,伤情有牙痕,破皮三处,出血。立即到卫生院接种狂犬 病疫苗 5 支,未进行伤口处理,免疫程序按照 0,3,7,14,30 天进行,于同年 5 月 4 日全程接种完,5 月 11 日发病,15 日 死亡,病程 5 天。

通过以上 2 例分析认为:凡参与疫苗接种的人员必须经

作者单位 541200 广西壮族自治区灵川县卫生防疫站
过培训后才能上岗。未经免疫接种培训的人员,对生物制品 的管理和接种技术均不规范,使疫苗未按冷藏的条件保存和 接种前疫苗未充分摇匀,影响疫苗的效价,达不到免疫效果。 乡村两级接种狂犬病疫苗人员均不是防疫专业人员,对基本 技能不了解。建议接种预防用生物制品最好是经防疫专业 人员。所有预防用生物制品必须按正规渠道购进。目前一 些疫苗经销商既无冷藏设备,也不懂冷藏知识,在疫苗运输、 保存方面不能达到冷藏条件,使疫苗效价降低,影响其疫苗 效果。乡村级就诊人员未进行过伤口处理。狂犬咬伤后伤 口处理是重要的环节,经过伤口处理后可以增加 40% 以上的 保护效果,因此在接种疫苗前应用 20% 肥皂水或新洁尔灭冲 洗伤口,冲洗后用 75% 酒精或碘酊涂擦伤口。

(收稿日期 2004-06-01)

(本文编辑:张林东)