



哈特人群哮喘遗传学队列研究的启示

潘学栋

复旦大学现代人类学教育部重点实验室, 上海 200433

评论文献: Ober C, Tan Z, Sun Y, Possick JD, Pan L, Nicolae R, Radford S, Parry RR, Heinzmann A, Deichmann KA, Lester LA, Gern JE, Lemanske RF Jr., Nicolae DL, Elias JA, Chupp GL (2008) Effect of variation in CHI3L1 on serum YKL-40 level, risk of asthma, and lung function. *N Engl J Med* 358:1682-1691.

Ober C, Hoffjan S (2006) Asthma genetics 2006: the long and winding road to gene discovery. *Genes Immun* 7:95-100.

Ober C, Thompson EE (2005) Rethinking genetic models of asthma: the role of environmental modifiers. *Curr Opin Immunol* 17:670-678.

摘要: 芝加哥大学人类遗传学教授 Carole Ober 在她的哈特人群队列中所作的疾病遗传学研究被普遍赞赏, 哮喘研究是其中最出色的部分。本文以她的哮喘研究为例, 说明选择具有相对单纯背景的队列, 合适的疾病定义和分类, 必要的重复验证, 以及分析基因和环境的交互作用是 Ober 做好哮喘遗传学研究的关键。

关键词: 哈特派; 队列; 哮喘; 遗传学

Inspiration from the Asthma Genetics Study in the Hutterite Cohort

PAN Xuedong

MOE Key Laboratory of Contemporary Anthropology, School of Life Sciences, Fudan University, Shanghai 200433 China

ABSTRACT: The disease association studies by Professor Carole Ober from University of Chicago in the Hutterite cohort were widely applauded, especially for the asthma study. In this paper, the asthma study was reviewed to conclude the rules for cohort studies. A cohort with relatively pure background, a proper classification and definitions for the disease, necessary repeats, and the analyses on the joint effects of genes and environmental factors were all pivotal parts of the study.

Keywords: Hutterite; Cohort; Asthma; Genetics

Carole Ober 是芝加哥大学人类遗传学的教授, 她从上世纪七十年代开始跟踪调查一个祖先群体很小的人群——哈特派信徒。在随访的几十年中, Ober 用遗传学方法对很多儿科、妇科疾病进行了研究, 取得了很多成果[1-4]。其中, 最著名的要数她对哮喘的研究[5-7], 2008年, 新英格兰医学学报发表了她的关于哮喘全基因组关联分析的文章。我们希望能从她数年的积累中吸取一些经验, 即如何用遗传学方法研究人类疾病。

队列建设——人群的选择, 时间的积累

Ober 的研究成果得到公认, 与她所跟踪随访的人群队列的特性密不可分。哈特派信徒(Hutterites)起源于十六实际提洛尔人(Tyrol)

居住的阿尔卑斯山脉(Tyrolean Alps)。在三百多年的时间里人口从 120 增长到 1000。大约在 1870 年, 其中的 900 人移民到美国的南达科塔州并分三个公社型农场居住。由于该人群有高的自然受精率并且禁止避孕, 目前该人群人口已超过 35000 人, 他们分居于美国北部和加拿大西部 350 个以上公社型农场中。Steinberg 和他的学生在二十世纪五六十年代收集了哈特派居民的家系信息, 追溯现今的哈特派居民的祖先到十八世纪至十九世纪早期的 90 人。这些祖先之间的关系并不清楚, 其中有一些也许有亲缘关系。南达科他州的哈特派居民自 1910 年起形成了三个生殖隔离的亚人群, S-哈特派(Schmiedeleut)、D-哈特派(Dariusleut)、L-哈特派(Lehrerleut)。Ober 研究的人群是 S-哈特派居民, 他们有 64 个共同祖先。

收稿日期: 2009年4月14日 修回日期: 2009年4月24日 联系人: 潘学栋 andrewalwaystry@gmail.com

2009年5月5日 <http://COMonCA.org.cn/Abs/2009/007.htm>

68

©上海人类学学会 Shanghai Society of Anthropology

奠基者人群在绘制复杂疾病的遗传图谱方面很有优势。与通常的人群相比，祖先人群数目少导致一些位点上的致病等位基因数少，很可能也会使致病的位点数减少，从而减小寻找致病基因的难度[8]。同时，哈特人群的公社化生活习惯使他们受的环境影响相对均一。尤其是哈特人群在公共厨房和餐厅做饭和饮食。吸烟是被禁止的，抽烟的人极少。所以，主动吸烟和被动吸烟现象基本不存在，不会影响到哮喘和肺部疾病的诊断。当然，Ober 也强调，哈特人群有其自身特性，哈特队列的研究结果如果没有别的人群进行验证，是不可信的。

Ober 带领她的团队每三年对哈特人群做一次全面的随访。他们拥有 13000 人的数据库，并且有一个家系的 13 代人的数据，涉及 1000 多人。

我们知道，队列建设是从群体角度研究疾病的第一步，也是最关键的一步。首先，我们应当选取遗传背景相对简单的人群。然后，队列是一个庞大而复杂的系统工程，对科研团队的组织和协调能力是一种考验；队列建设也需要时间的积累。

定义疾病——并不容易

哮喘是一种慢性呼吸系统疾病，传统的模型是这样描述哮喘的：刺激因素导致免疫反应，慢性症状为发炎。刺激因素包括：病毒感染、运动、激素水平改变、精神压力、自然界过敏原、食物和用药。以自然界过敏原为例，免疫反应包括三步，过敏原被抗原呈递细胞(APC)捕获并呈递给 Th0(T help0)细胞，通常情况下，Th0 细胞忽略该抗原。但是在哮喘病患者中，由于未知的原因，Th0 细胞转变为 Th2 细胞，并激活激素反应(hormonal response)，导致炎症反应(气管阻塞，多痰，细胞介导的免疫反应)。

我们知道，迄今为止的医学遗传学研究中，对疾病状态的描述大多是一个二分类变量，即患病与不患病。虽然判定患病与否的标准随各个病种不一，但是每个病种的定义都必须明确的。如果不明确，则可能带来疾病状态的错误描述，导致关联结果偏倚；如果不明确，不同研究组间的数据则没有可

比性，不能进行汇总(meta)分析。所以，明确疾病的定义非常重要，对哮喘来说，这个定义并不简单[9,10]。Wenzel 等在 2006 年 Lancet 上发表了一篇文章，专门对哮喘的表型作了定义[11]。Wenzel 认为，哮喘不是一个单一的疾病，而是一个由多种互相重叠的症状组成的复合体。他提出了多种可能的对哮喘进行分类的方法，一是临床或者生理表型，如基于严重程度的，基于易恶化程度的；二是引起哮喘的刺激物，如阿司匹林、花粉；三是炎症表型，如嗜中性细胞、粒细胞。其中，基于严重程度的分类方法在关联分析中应用得最为广泛。Ober 在全基因组关联研究以及过去的研究中都采取了基于严重程度的分类方法(表 1)，将哮喘状态分为三类。有趣的是，Miller 等发现，如果用基于严重程度的分类方法，患者哮喘的严重程度与其治疗的花费相关，即患者治疗花费越大，他的症状越严重[12]。这个现象也许可以解释为患者患病程度越重，则越需要好的治疗，但是相关性似乎说明好的治疗并没有效果。我们并没有太多理由否定治疗的效果，所以我们倾向于认为基于严重程度的分类方法也许还不够合理。Wenzel 的文章中也提到，通过症状的控制程度对哮喘进行分类是一个不错的选择[11]。无论如何，疾病的定义很重要。虽然对于类似哮喘的复杂疾病，给出一个合理的定义是一件很困难的事情，但是我们还是尽可能寻找好的方法。

表 1 通过两个条件判定的三种疾病状态
Tab.1 Three disease situations judged from two symptoms

条件	医生或自己诊断 以下至少两种症状(咳嗽、呼吸困难、呼吸短促)	支气管过度应答(FEV1 值下降 20%以上)
哮喘	有	有
支气管过度应答	无	有
无症状	无	无

我们知道，对于绝大多数疾病来说，国际上都有通用的金标准。但是，我们在确定疾病定义的时候，不能只考虑金标准，还要

考虑我们所能获取的病人资料尽可能准确地、合理地在力所能及的范围内给病情一个精确的描述。

重复——朴实而有效的验证方法

重复验证，是指在不同的人群中对同一个已检测到阳性的位点进行关联研究。Ober指出，对关联分析来说，重复是验证阳性位点的金标准。即使是被重复许多次的阳性位点，例如被重复了超过十次的 IL-4，依然会在少部分实验中呈现阴性[13]。这可能是由于人群亚结构或实验操作误差引起的。所以，即使是刚开始没有检测到阳性的位点，也不代表它就不是真正的阳性位点；同时，如果一个位点有一两次被检测出阳性，也不足以说明它就是阳性位点。Ober 总结了一百个左右至少被检测到阳性一次的位点(图 1)。我们看到，被重复五次以上的位点只是少数，而 Ober 认为，这二十几个点才是相对可信的。另外，Neale 等人认为重复验证的对象应该是基因而不是一个位点或一个单倍型[14]。

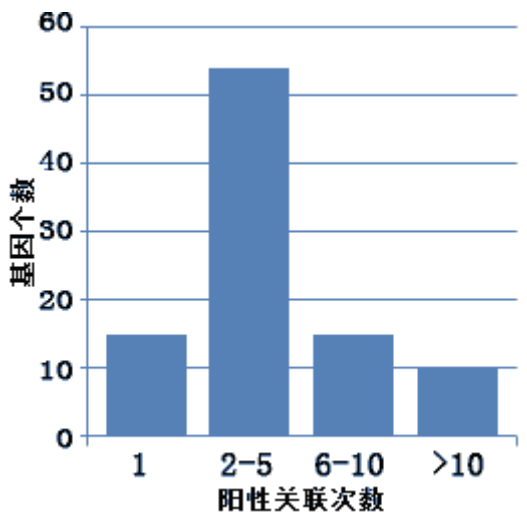


图 1 不同基因被检出与有关疾病正相关的次数分布
Fig.1 Frequencies of the positive correlation reports for various genes to the relevant diseases

Ober 在 2008 年新英格兰医学学报上发表了哮喘的全基因组关联结果[5]。在这篇文章中，为了增加结果的可信度，她在另外三个人群中进行了验证，一个是德国的，另一个是美国芝加哥的(不同于哈特人群)，还有一个是儿童队列。在这些队列中，她都检测到了 CHI3L1 基因上的 rs4950928(-131C→G)与

哮喘表型的关联。另外，这个位点的功能也在另一篇文章中被很好地验证[15]。

重复是验证阳性位点的不二之选，即使在全基因组关联分析中，被多次重复的位点或区域才真正可信。

基因与环境的相互作用——不能被忽视的因素

Ober 在另一篇文章中也指出，由于人群之间的差异，没有一个基因在所有人群中都是哮喘的致病基因[9]。并且大多数致病基因的致病效果并不高，而且基因-基因互作、基因-环境互作也有可能是重要的致病因素[13,16]。

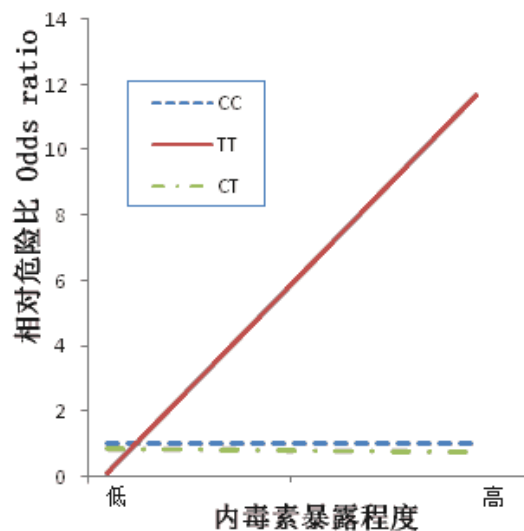


图 2 在不同的室内内毒素暴露程度下 CD14**C*-260T 不同基因型罹患巴巴多斯哮喘的相对危险比
Fig.2 Odds ratio of three genotypes of CD14**C*-260T for Barbados Asthma in different endotoxin exposure

交互指的是一个因素的改变会影响到另一个因素对结果的贡献。一个比较典型的例子如图 2，描述的是 CD14 基因与环境的交互作用对罹患巴巴多斯哮喘的影响。CD14 基因 C-260T 单核苷酸多态，在 TT 纯合子状态下，不同室内的内毒素浓度中有不同的相对危险比(OR)。该基因型在低内毒素暴露下，OR 值为 0.09，与该病没有关联；而在高内毒素暴露下，OR 值为 11.66，是危险因素。与之相对的是 CC/CT 基因型，与疾病的关联几乎不随内毒素浓度增加而改变。需要提及的是，这里的内毒素指的是家庭空气中漂浮的微生物

物进入人体呼吸系统后被裂解后释放的有毒物质。Zambelli 等人用定量的仪器测定了空气中内毒素的含量。虽然对于这种交互作用还没有很好的解释，但是，这至少提示我们，人体在应对内毒素的过程中可能用到了 CD14 基因的产物，而 TT 基因型导致了 CD14 产物的某些属性发生改变[17]。当然，如果这个交互作用能被重复验证，那么将大大提高其可信度，我们会有足够的理由检验以上提出的功能假设。

还有些因素本身就很简单，既有环境因素的属性，又有基因因素的属性，比方说性别[18]。Ober 在最近的一篇综述中专门讨论了疾病的性别特异性，因为性别差异与很多因素相关联，比如行为差异、心理差异、体内生化过程差异、基因调控差异等[19]。哮喘在男性和女性中的发病率也随年龄而不同，18 岁前男性高于女性，18 岁后女性高于男性。对这个现象并没有明确的解释。由于这类因素本身很复杂，我们在进行关联分析的时候需要格外注意，不能忽略其影响。目前的惯用做法是病例对照性别匹配。

我们可以展望一下未来，五年或十年以内，我们对基因的主效应已经有了较为深入的研究，我们的样本数量足够大，测序技术变得足够便宜，我们就可以来研究交互作用。也许，我们会对疾病有新的认识，交互作用才是导致疾病的主要因素[20]。

总结

Ober 能在人群疾病研究方面取得成功，与她长年累月的积累是分不开的。队列研究从来不是一件立竿见影的工程，研究者需要足够的耐心和持之以恒的态度来完成这项事业。同时，队列研究的质量更取决于我们的细心程度。需要尽可能详尽且准确的资料来判定疾病状态，做基因环境交互作用的分析；需要认真选取有代表性的样本人群。更进一步，还需要开展国际合作，找到用于重复验证的队列，使结论更加可靠。

参考文献

1. Newman DL, Abney M, Dytch H, Parry R, McPeck MS, Ober C (2003) Major loci influencing serum triglyceride levels on 2q14 and 9p21 localized by homozygosity-by-descent mapping in a

- large Hutterite pedigree. *Hum Mol Genet* 12:137-144.
2. Ober C, Elias S, Kostyu DD, Hauck WW (1992) Decreased fecundability in Hutterite couples sharing HLA-DR. *Am J Hum Genet* 50:6-14.
3. Ober C, Elias S, O'Brien E, Kostyu DD, Hauck WW, Bombard A (1988) HLA sharing and fertility in Hutterite couples: evidence for prenatal selection against compatible fetuses. *Am J Reprod Immunol Microbiol* 18:111-115.
4. Ober C, Bombard A, Dhaliwal R, Elias S, Fagan J, Laffler TG, Martin AO, Rosinsky B (1987) Studies of cystic fibrosis in Hutterite families by using linked DNA probes. *Am J Hum Genet* 41:1145-1151.
5. Ober C, Tan Z, Sun Y, Possick JD, Pan L, Nicolae R, Radford S, Parry RR, Heinzmann A, Deichmann KA, Lester LA, Gern JE, Lemanske RF, Jr, Nicolae DL, Elias JA, Chupp GL (2008) Effect of variation in CHI3L1 on serum YKL-40 level, risk of asthma, and lung function. *N Engl J Med* 358:1682-1691.
6. Ober C, Tsalenko A, Parry R, Cox NJ (2000) A second-generation genomewide screen for asthma-susceptibility alleles in a founder population. *Am J Hum Genet* 67:1154-1162.
7. Ober C, Leavitt SA, Tsalenko A, Howard TD, Hoki DM, Daniel R, Newman DL, Wu X, Parry R, Lester LA, Solway J, Blumenthal M, King RA, Xu J, Meyers DA, Bleeker ER, Cox NJ (2000) Variation in the interleukin 4-receptor alpha gene confers susceptibility to asthma and atopy in ethnically diverse populations. *Am J Hum Genet* 66:517-526.
8. Ober C, Abney M, McPeck MS (2001) The genetic dissection of complex traits in a founder population. *Am J Hum Genet* 69:1068-1079.
9. Hoffjan S, Nicolae D, Ober C (2003) Association studies for asthma and atopic diseases: a comprehensive review of the literature. *Respir Res* 4:14.
10. Ober C (2005) Perspectives on the past decade of asthma genetics. *J Allergy Clin Immunol* 116:274-278.
11. Wenzel SE (2006) Asthma: defining of the persistent adult phenotypes. *Lancet* 368:804-813.
12. Miller MK, Johnson C, Miller DP, Deniz Y, Bleeker ER, Wenzel SE (2005) Severity assessment in asthma: An evolving concept. *J Allergy Clin Immunol* 116:990-995.
13. Ober C, Hoffjan S (2006) Asthma genetics 2006: the long and winding road to gene discovery. *Genes Immun* 7:95-100.
14. Neale BM, Sham PC (2004) The future of association studies: gene-based analysis and replication. *Am J Hum Genet* 75:353-362.
15. Zhao X, Tang R, Gao B, Shi Y, Zhou J, Guo S, Zhang J, Wang Y, Tang W, Meng J, Li S, Wang H, Ma G, Lin C, Xiao Y, Feng G, Lin Z, Zhu S, Xing Y, Sang H, St Clair D, He L (2007) Functional variants in the promoter region of Chitinase 3-like 1 (CHI3L1) and susceptibility to schizophrenia. *Am J Hum Genet* 80:12-18.
16. Ober C, Thompson EE (2005) Rethinking genetic models of asthma: the role of environmental modifiers. *Curr Opin Immunol* 17:670-678.
17. Zambelli-Weiner A, Ehrlich E, Stockton ML, Grant AV, Zhang S, Levett PN, Beaty TH, Barnes KC (2005) Evaluation of the CD14/-260 polymorphism and house dust endotoxin exposure in the Barbados Asthma Genetics Study. *J Allergy Clin Immunol* 115:1203-1209.
18. Weiss LA, Pan L, Abney M, Ober C (2006) The sex-specific genetic architecture of quantitative traits in humans. *Nat Genet* 38:218-222.
19. Ober C, Loisel DA, Gilad Y (2008) Sex-specific genetic architecture of human disease. *Nat Rev Genet* 9:911-922.
20. Vineis P, Brennan P, Canzian F, Ioannidis JPA, Matullo G, Ritchie M, Stromberg U, Taioli E, Thompson J (2008) Expectations and challenges stemming from genome-wide association studies. *Mutagenesis* 23:439-444.