Molecular Psychiatry (2000) 5, 11–13

www.nature.com/mp

NEWS & VIEWS

Beware the chopsticks gene

Population stratification is a potential source of error in psychiatric genetics. New study designs and statistical methods can help guard against this problem. *Molecular Psychiatry* (2000) **5,** 11–13.

Once upon a time, an ethnogeneticist decided to figure out why some people eat with chopsticks and others do not. His experiment was simple. He rounded up several hundred students from a local university, asked them how often they used chopsticks, then collected buccal DNA samples and mapped them for a series of anonymous and candidate genes.

The results were astounding. One of the markers, located right in the middle of a region previously linked to several behavioral traits, showed a huge correlation to chopstick use, enough to account for nearly half of the observed variance. When the experiment was repeated with students from a different university, precisely the same marker lit up. Eureka! The delighted scientist popped a bottle of champagne and quickly submitted an article to *Molecular Psychiatry* heralding the discovery of the 'successful-use-of-selected-hand-instruments gene' (SUSHI).

It took another 2 years to discover that SUSHI is a histocompatibility antigen gene that has nothing to do with chopstick use but just happens to have different allele frequencies in Asians and Caucasians, who of course differ in chopstick use for purely cultural rather than biological reasons. Even though the association data were highly significant and readily replicated, they were biologically meaningless.

This well-known little story is an apocryphal example of the dangers of population stratification, also known as population admixture, which is the situation that arises when a study population contains two or more ethnic or racial subgroups that have different allele frequencies and, just coincidentally, different levels of a particular phenotype. The fatal flaw in the SUSHI gene story is obvious, but suppose that we replaced 'ethnogeneticist' with 'psychiatric geneticist', 'chopstick use' with 'schizophrenia', and 'Asian *vs* Caucasian' with 'Irish American *vs* Italian American', or 'French *vs* German'. We would then have a study that could very well be published in this very journal. Fortunately there are a variety of study designs and statistical methods, several of which have been described recently, that can be used to avoid the potential perils of population stratification.

The first two approaches use nuclear families. The basic idea is that if a trait is associated with a particular allele(s) of a candidate gene within the offspring of a nuclear family, then there must be genuine transmission since a sibship is, by definition, a homogeneous population subgroup in which all the members (children) have the same founders (parents). The transmission disequilibrium test (TDT), which involves analyzing parent-child trios, was one of the first methods to use families in this way.[1] For qualitative traits, such as psychiatric diagnoses, the TDT examines the frequency with which the allele of interest is transmitted from a heterozygous parent to an affected child; significant deviation from the Mendelian expectation of 50% transmission, as determined by a McNemar chi-square test, is taken as an indication of both association and linkage. For quantitative traits, such as personality scales, the degree of association between the allele of interest and the trait is measured only in those offspring known to have received the allele from a heterozygous parent, and the significance of the relationship is assessed by various methods such as a *t*-test for two independent samples[2] or multiple regression.[3]

The use of siblings to measure association within families is a more recent development.[4,5] For qualitative traits, one examines those sibships in which both the trait and the allele of interest are segregating; that is, sibships in which there is at least one affected and one unaffected child and not all the siblings have the same genotype. The significance of allelic association is then assessed by a permutation or *z* score test, which can conveniently be combined with child-parent trio data.[4] For quantitative traits, one also focuses on sibships in which the allele of interest is segregating. Benjamin *et al*,[6] who first described the use of siblings to control for population stratification in an association study, assessed the significance of the association within such genotype-discordant sib-pairs by a paired-samples *t*-test. Recently Allison *et al*[7] have shown that this type of sibling data can be analyzed by two additional methods: a mixed effects ANOVA in which phenotype is the dependent variable, genotype is the

Correspondence: Dr DH Hamer, NIH Building 37, Rm YA13, Bethesda, MD 20892, USA. E-mail: deanh@helix.nih.gov

fixed factor, and sibship number is the random factor; and a permutation-based procedure that generates a chi-square statistic. Fulker *et al*[8] have proposed a method to combine sib-pair association and linkage analysis by a maximum-likelihood variance-components method that simultaneously models the allelic means and sibling covariance matrix. A useful feature of this method is that it partitions the effect of the gene into between- and within-sibship components, thus providing a direct test of population stratification.

The key property of the TDT and sib-based tests is that they require *both* linkage and association to give a positive result. If chopstick use had been tested within families, instead of in unrelated individuals, it would have been obvious that the observed association was spurious since linkage is not present.

Pritchard and Rosenberg[9] have taken a quite different approach. They propose that if population stratification exists within a population of unrelated cases and controls, it should be possible to detect it by using unlinked markers. Their test consists simply of genotyping the experimental population with 15–20 random polymorphic markers. If there is no significant correlation between the trait and these control markers, then the observed association to the candidate gene is considered to be genuine; if a spurious correlation between the trait and the control markers is found, then the association to the candidate gene is written off as an artifact. Although Pritchard and Rosenberg deal only with qualitative traits, the extension of the method to quantitative phenotypes should be straightforward.

Which of these methods is 'best' for psychiatric genetics? In terms of validity, the two family-based methods provide a direct test of genuine transmission *vs* population stratification that should be valid even in ethnically mixed samples such as those typically collected in the United States. By contrast, the unlinked marker approach is only inferential; if the control and candidate loci do not have comparable degrees of population-specific variability, a spurious association could be accepted or a genuine association could be rejected. Another advantage of the family-based methods is that they can quantify the degree to which an observed association is due to genuine transmission *vs* population effects and therefore prove whether stratification actually exists; although population stratification is sometimes blamed for every failure to replicate, there are surprisingly few (if any) examples where this has been proven. Finally, the family methods are well suited to whole genome searches whereas the unlinked marker approach is problematic; how would one decide which markers are the controls and which are the candidates?

In terms of sample size requirements, the unlinked marker method is the big winner since it requires just the same number of subjects as a population-based association study. Family-based methods require a larger number of subjects because many of the parent-child trios or sib-pairs will not be usable since the trait or candidate gene allele is not segregating. Of the two family-based methods, the sib-pair test is generally less powerful than the TDT for qualitative traits, because of the problem of overmatching, but more powerful for quantitative traits because it takes advantage of the polygenic and environmental correlations between siblings, both of whom provide useful phenotypic information. (Although some investigators are reluctant to use siblings in association tests because of their nonindependence, the appropriate small corrections are easily made by using the ASSOC program[10] or estimating equations.[11])

As a practical example, consider a dominant biallelic locus that contributes to 5% of the variance in a quantitative trait. Assuming equal allele frequencies and a sibling correlation of 0.2, the number of individuals that would have to be genotyped (and phenotyped) to have 80% power at the $P = 0.05$ level are 651 (136) for the TDT, 432 (96) for sib-pairs, and 152 (152) for population association with or without unlinked markers. The corresponding numbers for the more robust significance level of $P = 0.0001$ are 1863 (388) for the TDT, 1246 (272) for sib-pairs, and 435 (435) for population association.

Several cautions about the use of these approaches should be mentioned:

- They do not magically decrease the sample size required to obtain a given level of statistical significance. The probability of type I error is not reduced by using families.
- Lack of significance in a family-based study is not a proof of population stratification. This requires evidence of a significant *difference* in effect sizes.
- Population stratification can cause false negatives as well as false positives. Reporting no association based only on a case-control or population-based experiment is just as problematic as reporting a positive result.
- The unlinked markers used for the Pritchard–Rosenberg method must be chosen in advance, not after the test has been performed; the idea is to determine whether *any* of the controls show association, not whether some of them don't.
- Although it is theoretically possible to increase the power of an association experiment by using subjects who are selected on the basis of extreme phenotypes, it is advisable to first examine the entire population since the effects of a gene may not be constant across the full distribution.[12]

So long as these caveats are kept in mind, psychiatric geneticists should have no problem distinguishing 'chopsticks genes' from the real thing.

D Hamer and L Sirota
Laboratory of Biochemistry
National Cancer Institute
National Institues of Health
Bethesda, MD, USA

References

1 Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 1993; **52**: 506–516.
2 Allison DB. Transmission-disequilibrium tests for quantitative traits. *Am J Hum Genet* 1997; **60**: 676–690.
3 George V, Tiwari HK, Zhu X, Elston RC. A test of transmission/disequilibrium for quantitative traits in pedigree data, by multiple regression. *Am J Hum Genet* 1999; **65**: 236–245.
4 Spielman RS, Ewens WJ. A sibship test for linkage in the presence of association: the sib transmission/disequilibrium test. *Am J Hum Genet* 1998; **62**: 450–458.
5 Boehnke M, Langefeld CD. Genetic association mapping based on discordant sib pairs: the discordant-alleles test. *Am J Hum Genet* 1998; **62**: 950–961.
6 Benjamin J, Li L, Patterson C, Greenberg BD, Murphy DL, Hamer DH. Population and familial association between the D4 dopamine receptor gene and measures of Novelty Seeking. *Nature Genet* 1996; **12**: 81–84.
7 Allison DB, Heo M, Kaplan N, Martin ER. Sibling-based tests of linkage and association for quantitative traits. *Am J Hum Genet* 1999; **64**: 1754–1763.
8 Fulker DW, Cherny SS, Sham PC, Hewitt JK. Combined linkage and association sib-pair analysis for quantitative traits. *Am J Hum Genet* 1999; **64**: 259–267.
9 Pritchard JK, Rosenberg NA. Use of unlinked genetic markers to detect population stratification in association studies. *Am J Hum Genet* 1999; **65**: 220–228.
10 SAGE. Statistical analysis for genetic epidemiology, release 3.1. Department of Epidemiology and Biostatistics: Case Western Reserve University, Cleveland, 1998.
11 Tregouet DA, Ducimetiere P, Tiret L. Testing association between candidate-gene markers and phenotype in related individuals, by use of estimating equations. *Am J Hum Genet* 1997; **61**: 189–199.
12 Sirota LA, Greenberg BD, Murphy DL, Hamer DH. Non-linear association between the serotonin transporter promoter polymorphism and neuroticism: a caution against using extreme samples to identify quantitative trait loci. *Psychiatr Genet* 1999; **9**: 35–38.