

Genetic dissection of the common epilepsies

Nigel C.K. Tan^a, John C. Mulley^b and Ingrid E. Scheffer^c

Purpose of review

Only two functionally validated susceptibility genes, *CACNA1H* and *GABRD*, have so far been identified in the common epilepsies using a candidate gene approach. The difficulty with the alternative statistical approach, where none of the suggested candidates has been functionally validated, may partly be due to the posited genetic architecture of the common epilepsies, such as the idiopathic generalized epilepsies. A subset of both rare and common variants from a much larger pool of susceptibility genes may contribute to disease risk. We review methods and designs for the genetic dissection of common epilepsies.

Recent findings

Genetic association studies, though theoretically more powerful than linkage analysis, have not yet delivered validated susceptibility genes. Methodological flaws can undermine such studies but are correctable. Concerns remain, however, about the extent of underlying genetic heterogeneity in common epilepsies. Genome-wide association studies are increasingly feasible, but issues remain about their conduct and analysis. Metaanalysis may resolve conflicting association studies, facilitated by the establishment of databases of genetic association studies. Newer multi-locus and admixture mapping approaches are attractive alternatives to traditional association studies and may offer new insights into identifying epilepsy genes.

Summary

We conclude by emphasizing the importance of deeper endophenotyping using electroclinical, imaging, and molecular approaches to dissect the common epilepsies.

Keywords

complex disease, endophenotyping, genetic dissection, genetic heterogeneity

Abbreviations

CDCV	common disease-common variant
IGE	idiopathic generalized epilepsy
MRV	multiple rare variant
SNP	single nucleotide polymorphism
TLE	temporal lobe epilepsy

© 2006 Lippincott Williams & Wilkins
1350-7540

Introduction

Common epilepsies, the subject of this review, refer to the epilepsy syndromes that we regularly encounter in clinical practice. We focus on the idiopathic generalized epilepsies (IGEs) and temporal lobe epilepsy (TLE) as these are the common epilepsy syndromes in which the role of genetic factors has been investigated. We will not be examining the rare idiopathic epilepsies precipitated by single gene mutations of large effect, nor the genetics of febrile seizures, as these have recently been reviewed [1,2].

In reviewing the genetic evidence for common epilepsies, we begin by briefly describing the posited complex genetic architecture of common epilepsies. We examine the methods for genetic dissection for complex diseases, focusing on genetic association studies. We review newer methods, and conclude by offering a perspective on what is needed for future progress.

Monogenic versus complex diseases

Monogenic epilepsies are rare idiopathic epilepsy syndromes due to large effect mutations in a single gene, often ion channels [3], when both the mutation and the syndrome segregate together and follow simple Mendelian inheritance [1]. They are often characterized by genetic heterogeneity, incomplete penetrance and variable expressivity, with the last two characteristics suggestive of modulation by additional unidentified modifier genes.

In common epilepsy syndromes, however, the underlying genetic architecture is believed to be quite different [4••]. Instead of a rare mutation in a single gene resulting in epilepsy, common epilepsies are polygenic and result when an individual has susceptibility alleles of sufficient magnitude at enough genes to exceed a putative seizure threshold. An isolated susceptibility allele in a single gene does not, by itself, result in epilepsy. Extension of the concept of genetic heterogeneity to complex

Curr Opin Neurol 19:157–163. © 2006 Lippincott Williams & Wilkins.

^aDepartment of Neurology, National Neuroscience Institute, Singapore,

^bDepartment of Genetic Medicine, Women's and Children's Hospital, and School of Molecular and Biomedical Sciences, The University of Adelaide, Adelaide, South Australia, Australia and ^cDepartment of Paediatrics, Royal Children's Hospital, Parkville, and Epilepsy Research Centre and Department of Medicine (Neurology), University of Melbourne, Austin Health, Heidelberg West, Victoria, Australia

Correspondence to Prof. Ingrid E. Scheffer, Epilepsy Research Centre, Level 1, Neurosciences Building, Austin Health, Banksia Street, West Heidelberg, Victoria 3081, Australia

Tel: +613 9496 2737; fax: +613 9496 2291; e-mail: scheffer@unimelb.edu.au

Current Opinion in Neurology 2006, 19:157–163

epilepsies would imply ‘polygenic genetic heterogeneity’ with various subsets of susceptibility genes each capable of precipitating clinically identical seizure subsyndromes. The concepts of penetrance and expressivity disappear for complex genetic disorders as the effect of no single gene will dominate the additive or interactive effect of each polygenic set of susceptibility genes.

Common epilepsies are thus termed complex diseases, as the disease does not follow simple Mendelian inheritance, though each susceptibility allele does. Complexity in this ‘polygenic’ disease is further increased by incorporation of gene–environment interactions [5^{••}], the term ‘multifactorial’ taking the actions of several genes and the environment into account. The mystery of this multifactorial disease lies in our ignorance of the total number of susceptibility genes, distribution of their effect sizes and the magnitude of the required epilepsy precipitating ‘effect size’ in a given environment.

Susceptibility alleles: rare versus common variants

Much of the variation in the human genome is provided by over 9 million single nucleotide polymorphisms (SNPs) [6]. SNPs and rarer variants are what distinguish individuals within a population from one other; they are believed to represent susceptibility alleles for common diseases with complex inheritance [7[•]].

An ongoing debate, however, centres on the frequency of these susceptibility alleles [8]. Either multiple rare variants (with allele frequencies < 1%) [9] or a few common polymorphisms (SNPs) [10] (allele frequencies > 1%), or both, might play a causal role in complex disease. These models are recognized in complex genetic disorders as the multiple rare variant (MRV) and the common disease-common variant (CDCV) hypotheses.

Both hypotheses are not mutually incompatible, and evidence exists from candidate gene studies to support both hypotheses [8,11,12[•],13] in different complex diseases and within the same complex disease, such as epilepsy. Rare and common variants in *CACNA1H* [14] and *GABRD* [15] have been shown to change channel properties consistent with susceptibility to complex IGE but neither accounts for a high proportion of cases, suggesting that in common epilepsies, a mixture of both models is possible [4^{••}], and that there is considerable underlying polygenic heterogeneity.

Dissection and disappointment

Several strategies have been attempted for the genetic dissection of common epilepsies and other complex diseases [16^{••}]; their strengths and weaknesses have recently been summarized [17^{••}]. Linkage analysis using multiplex families has proven highly successful for

monogenic diseases, but not so in complex diseases and common epilepsies in particular. Affected sib-pair studies, a nonparametric method of linkage analysis, have also proven ineffective in genetic dissection, though these may be a result of the practicalities of ascertaining sample sizes with adequate power.

Genetic association studies

Association studies have been increasingly used as an alternative approach to analyse the common epilepsies. These are, in essence, genetic case–control studies, and represent an epidemiological approach to complex diseases [18[•]]. There are theoretical advantages of this strategy over traditional linkage methods in complex diseases [19]. Association studies, however, have been dogged by problems of false-positive results and non-replicability [20–22]. Association studies in epilepsy are equally disappointing [7[•],23[•]], with no clear susceptibility gene identified using this approach.

This might be explained by methodologic limitations [7[•],20,23[•],24,25,26^{••}]; potential shortcomings include inadequate statistical power [7[•]], inappropriate controls and population stratification [27], and insufficiently stringent significance thresholds [22,28]. Genotyping errors may also be contributory [26^{••},29,30[•]]. Some of these problems may be partially mitigated using designs based on family-based controls and larger sample sizes. Newer statistical techniques to control for bias (especially population stratification [27]) and multiple testing [31[•]] are being developed; careful consideration of significance thresholds [7[•]] and use of Bayesian methods [22,32] are also encouraging developments. Guidelines for epilepsy genetic association studies have been proposed [7[•]], incorporating these improvements.

More troubling, however, is the issue of heterogeneity. As sample sizes increase, it is increasingly likely that the disease population being examined will become more heterogeneous, without the development of greater specificity through endophenotyping. The rare versus common variant debate [8] then becomes of critical importance.

Under the MRV hypothesis, any population contains many rare mildly deleterious mutations that increase epilepsy susceptibility. An individual with epilepsy has several of these rare mutations; unfortunately, different combinations of genes may produce the same epilepsy phenotype. A sample of cases with the same epilepsy syndrome in an association study will thus include a mixture of different rare causative alleles spread across many genes. Such genetic heterogeneity will progressively nullify the statistical power of an association study [9]. In contrast, the most optimistic view is the CDCV hypothesis (in which all susceptibility polymorphisms are

common and genetic heterogeneity is minimal), for which association studies will have sufficient power; this assumption has yet to be confirmed for common epilepsies. In reality [4^{••}], both common and rare variants are likely to play a role; this will further increase the heterogeneity and reduce the power of association studies.

The rare versus common variant debate also affects sample sizes. Under the CDCV hypothesis, 500–750 case–control pairs would be needed to detect a common susceptibility allele [33[•]]. With a rare variant, however, required sample sizes would approach impracticality [9,34^{••}].

The sample size further increases with stringent significance thresholds. To complicate matters further, these calculations optimistically assume dominant inheritance, complete linkage disequilibrium between marker and causal allele, and no phenotyping or genotyping error [35[•]]. More pragmatic calculations [36] factoring in these considerations [33[•]] dictate even larger and more unrealistic sample sizes of thousands or tens of thousands.

Given such daunting sample sizes, multicentre studies will be needed [7[•],23[•]]. International League Against Epilepsy (ILAE) diagnostic criteria are a useful starting point for implementation of consistent phenotyping [7[•]]. IGE, however, will prove challenging given its clinical heterogeneity; for TLE, aetiologic heterogeneity presents difficulties. Selecting subgroups, for example TLE and hippocampal sclerosis, may be a solution [23[•]].

Genome-wide association studies

With falling costs and high-throughput genotyping methods [17^{••}], enthusiasm has been mounting for genome-wide association studies. Instead of testing one particular SNP at a time, hundreds of thousands of SNPs are tested throughout the genome for significance. The recent publication of HapMap data [37^{••}] should make this technically feasible. This method has been used for genetic dissection of myocardial infarction [38] but has yet to be applied to common epilepsies.

While such studies may initially appear attractive due to the ability to seemingly cover the whole genome in one fell swoop, many issues remain. There is no general consensus regarding population selection, marker selection (map-based versus sequence-based), appropriate statistical analyses, population stratification, statistical significance and genotyping errors. These problems and their potential solutions have recently been summarized [17^{••},34^{••},39,40^{••}]. One approach which appears to be gaining acceptance is a multistage approach [17^{••},40^{••}]. A small sample is tested for significance with a high-density set of markers; markers that cross the significance threshold are then re-tested in a second,

larger population. This results in increased efficiency and significantly reduced cost.

It is likely that genome-wide association studies will be applied to common epilepsies in the near future, if for no better reason than that it will be technically possible to do so. These studies will require meticulous attention to methodology to have any chance of success [17^{••}]. If candidate genes are identified via genome-wide association studies, biological plausibility is essential.

Given the likely number of false-positive results (despite evolving methods to control for these), each putative susceptibility gene should be shown to be linked to epilepsy in a biologically meaningful way [7^{••},41,42]. In epilepsy, for example, alterations in ion channel function [14,15] would provide functional evidence to validate any statistically significant association.

Metaanalyses of genetic association studies

Given the conflicting results of existing association studies, investigators have turned to metaanalysis to resolve inconsistency. Metaanalysis is an increasingly popular statistical tool for analysing and combining results across studies, producing a conclusion based on the existing body of evidence. Its attraction lies in its ability to combine results of many contradictory small studies, overcome their individual lack of power, and produce a final conclusion as to whether a putative susceptibility gene has a real effect [43^{••},44].

A good metaanalysis relies on robust study selection and statistical analysis. The adage of ‘garbage in, garbage out’ applies. A metaanalysis must state inclusion and exclusion criteria, as well as assess the methodologic strengths of each study considered [43^{••},44,45[•],46].

Studies to be included must fulfil quality control criteria [46]; for example, significant deviation from Hardy-Weinberg equilibrium may indicate sampling or genotyping error [47,48]. Inclusion of methodologically suspect studies would affect the conclusion of the metaanalysis [44].

A good metaanalysis must also determine and control for heterogeneity of the studies included [44,46]. Heterogeneity may arise because of differences in ascertainment methods, ethnic populations, case definitions and disease aetiology or severity.

For example, studies examining the relationship between the GABA(B) receptor 1 gene and TLE used two different cohorts – one with predominantly mild nonlesional TLE [49], the other with predominantly severe TLE with hippocampal sclerosis [50]. Case definitions of drug-resistant epilepsy may differ, as seen in recent epilepsy pharmacogenetic studies [51–53].

Existing evidence suggests that between-study heterogeneity is prevalent [54] but racial differences are unlikely to account for it [55^{*}]. There are existing methods for handling heterogeneity such as using a random-effects model or subgroup analysis, but these are not without their problems either [44,56].

Publication bias remains a vexing problem in meta-analysis [43^{**},57^{*}]. There is less incentive to submit negative studies for publication because these are less exciting; acceptance is thus more difficult. This naturally leads to a bias towards inclusion of positive studies in metaanalyses, potentially skewing the results towards a positive conclusion.

It is crucial for a robust metaanalysis to ascertain the extent of publication bias; numerous tests exist for this purpose [57^{*}]. Unpublished data should also be considered for inclusion, though these must still fulfil inclusion and quality control criteria. One way around this problem is to create an online registry of published and unpublished genetic association studies to facilitate metaanalysis [45]. The existing Genetic Association Database [58] is a laudable step in that direction, though its emphasis on published data is unlikely to solve the problem of publication bias. A similar database dedicated to epilepsy association studies, published and unpublished, is required.

Many published association studies in epilepsy are small and underpowered [7^{*}]. Can metaanalysis find us a small enough genomic region to enable identification of a susceptibility gene for common epilepsies? Its premise is intriguing and its methods are maturing quickly. Encouragingly, minimum standards have recently been proposed [44]. Increasing publication of negative studies is also reducing the problem of publication bias. Meta-analysis has already clarified the relationship between *UCHL1* and Parkinson's disease [59]; common epilepsies may soon benefit from insights provided by careful metaanalyses, though this is no substitute for an adequately powered and designed association study.

Newer methods: multi-locus approaches and admixture mapping

Inventive new approaches may provide fresh insight into epilepsy susceptibility genes.

Multi-locus approaches

Association studies currently examine a single locus at a time. If the CDCV theory is correct, each susceptibility allele only has a small effect, explaining only a small fraction of disease heritability. Finding such genes would require massive sample sizes.

The phenomenon of gene–gene interaction, however, would suggest that though individual effects of each

susceptibility gene are small, the joint effect of a pair of genetic variants may be greater than the sum of their individual effects, making identification easier with smaller sample sizes. This is the premise behind multi-locus approaches [6,60]. These aim to identify genetic variants that act in concert in order to affect disease risk. It is now possible to extend this approach genome wide using extant technology [6]. Initial methodologic concerns [60] have been eased with recent reports that this method is computationally feasible [61^{**}] and different approaches can be employed for varying models of gene–gene interaction [6].

Admixture mapping

An admixed population has genetic ancestry from two different ancestral populations. The possibility of using such populations to map genes was proposed over 50 years ago, but has come into prominence recently due to refinements in its theoretical and practical aspects [62^{**},63,64^{**}].

At its simplest, admixture mapping is a method of identifying susceptibility alleles that differ in frequency across populations [65]; it performs best when examining recently admixed populations that have ancestry from two different populations that were previously isolated from each other, and which have large differences in disease prevalence and susceptibility allele frequencies [62^{**},63,64^{**}].

Admixture mapping occupies a middle ground compared with the established methods of linkage analysis and association analysis for dissection of complex traits [17^{**},64^{**},65,66]. Its statistical power and mapping resolution falls between linkage analysis and association analysis [64^{**},65], though with caveats [62^{**}]. It is more robust to genetic heterogeneity and population stratification compared with association analysis [65]. Compared with genome-wide association studies, the number of markers required for a genome scan using admixture mapping is also much lower. This makes admixture mapping an attractive starting point for genome-wide searches.

Though the arguments in favour of admixture mapping are compelling and the statistical methods improving, the biggest obstacle facing this method is availability of suitably admixed populations (with differences in disease and allele frequency), and identifying genetic markers informative for ancestry [62^{**}]. A set of markers is available for African-Americans [67], who have 75% African and 25% European ancestry, but it is likely that different sets of markers will be needed for dissimilar populations [65].

Susceptibility loci for hypertension [68^{*}] and multiple sclerosis [69^{*}] have recently been reported using

admixture mapping in African-Americans. Admixture mapping is an intellectually appealing concept, but we must be mindful that stringent criteria have to be met [63,64**], and independent replication is pending for these early reports.

Conclusion

Linkage analysis has not yet identified susceptibility genes for common epilepsies, and association studies are now being heavily promoted as a more powerful alternative. Though methodological and statistical innovations have improved study quality, the approach is still dogged by the problem of false-positive reports. Concerns about the genetic architecture of common epilepsies (rare versus common variants) and heterogeneity (clinical and genetic) of epilepsy cases, however, still present fundamental difficulties to this approach. Demonstrating biological plausibility for statistical associations (validation) is rarely carried out, detracting from the impact of the findings. This may require collaboration with other researchers who have complementary skills, such as physiologists, substantiating the biological link between the gene and the phenotype.

Genome-wide association studies, metaanalyses and public databases of existing association studies hold promise for identifying susceptibility genes. Methodological and statistical problems are currently being solved and these studies now appear technically feasible and dependent only on funding. Newer multi-locus approaches and admixture mapping may also be used; multi-locus approaches may be considered for IGE, and admixture mapping for specific populations.

It is instructive to conclude with an overall view of the genetic epidemiology of common epilepsies. The first small steps have been made to identify a small number of epilepsy susceptibility genes by the candidate gene approach [4**], but the overwhelming majority await discovery. At the same time, major advances have been made in genotyping speed and costs, statistical analysis and bioinformatics as a complementary strategy to the application of functional genomic approaches to candidate genes. The major advantage of this alternative approach is that it should detect genes which fall outside our current notion of candidate genes.

A multidisciplinary approach involving clinical epileptology, epidemiology, genetics, statistics, bioinformatics and functional genomics is therefore needed [70**]. The clinician's key role is in assembling and endophenotyping cohorts of epilepsy patients. Large sample sizes must also be balanced against increasing clinical and genetic heterogeneity. The scientist's role is to pursue endophenotyping as close to the molecular level as possible in order to further refine existing clinical entities

to reduce the level of heterogeneity which will otherwise confuse attempts to unravel the genetic architecture by the association study approach.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

Additional references related to this topic can also be found in the Current World Literature section in this issue (pp. 221–222).

- 1 Gourfinkel-An I, Baulac S, Nabbout R, *et al.* Monogenic idiopathic epilepsies. *Lancet Neurol* 2004; 3:209–218.
- 2 Baulac S, Gourfinkel-An I, Nabbout R, *et al.* Fever, genes, and epilepsy. *Lancet Neurol* 2004; 3:421–430.
- 3 Mulley JC, Scheffer IE, Petrou S, Berkovic SF. Channelopathies as a genetic cause of epilepsy. *Curr Opin Neurol* 2003; 16:171–176.
- 4 Mulley JC, Scheffer IE, Harkin LA, *et al.* Susceptibility genes for complex epilepsy. *Hum Mol Genet* 2005; 14 (Suppl 2):R243–R249. This review summarizes the genetic architecture and extant evidence for susceptibility genes in idiopathic generalized epilepsy, emphasizing the possibility of extensive genetic heterogeneity and contribution of rare and common variants to disease.
- 5 Hunter DJ. Gene–environment interactions in human diseases. *Nat Rev Genet* 2005; 6:287–298. This review describes different models for gene–environment interactions and examines various study designs for identifying such interactions, and the challenges facing such studies.
- 6 Daly MJ, Altshuler D. Partners in crime. *Nat Genet* 2005; 37:337–338.
- 7 Tan NC, Mulley JC, Berkovic SF. Genetic association studies in epilepsy: 'the truth is out there'. *Epilepsia* 2004; 45:1429–1442. A critical review of over 50 genetic association studies in epilepsy, their methodological limitations, and a proposed set of guidelines for conducting and appraising such studies.
- 8 Smith DJ, Lusk AJ. The allelic structure of common disease. *Hum Mol Genet* 2002; 11:2455–2461.
- 9 Pritchard JK, Cox NJ. The allelic architecture of human disease genes: common disease-common variant ... or not? *Hum Mol Genet* 2002; 11:2417–2423.
- 10 Reich DE, Lander ES. On the allelic spectrum of human disease. *Trends Genet* 2001; 17:502–510.
- 11 Lohmueller KE, Pearce CL, Pike M, *et al.* Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet* 2003; 33:177–182.
- 12 Klein RJ, Zeiss C, Chew EY, *et al.* Complement factor H polymorphism in age-related macular degeneration. *Science* 2005; 308:385–389. This paper uses a limited genome-wide screen to identify a common susceptibility allele for age-related macular degeneration.
- 13 Cohen JC, Kiss RS, Pertsemidis A, *et al.* Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science* 2004; 305:869–872.
- 14 Khosravani H, Altier C, Simms B, *et al.* Gating effects of mutations in the Cav3.2 T-type calcium channel associated with childhood absence epilepsy. *J Biol Chem* 2004; 279:9681–9684.
- 15 Dibbens LM, Feng HJ, Richards MC, *et al.* GABRD encoding a protein for extra- or peri-synaptic GABAA receptors is a susceptibility locus for generalized epilepsies. *Hum Mol Genet* 2004; 13:1315–1319.
- 16 Freimer N, Sabatti C. The use of pedigree, sib-pair and association studies of common diseases for genetic mapping and epidemiology. *Nat Genet* 2004; 36:1045–1051. This paper reviews the principles behind the use of linkage and association analysis, emphasizing the need for different methods for varying circumstances, importance of significance thresholds, and careful selection of phenotypes.
- 17 Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 2005; 6:95–108. This paper, together with a companion review by Wang *et al.* [34**], provides a state-of-the-art summary of genome-wide association studies. The paper includes a useful table comparing linkage vs admixture mapping vs association analysis vs resequencing, discusses marker selection, significance thresholds, and sources of bias.
- 18 Cordell HJ, Clayton DG. Genetic association studies. *Lancet* 2005; 366:1121–1131. A well written summary of the basic premise and rationale behind association studies, part of a comprehensive seven-part series on genetic epidemiology.

- 19 Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science* 1996; 273:1516–1517.
- 20 Colhoun HM, McKeigue PM, Davey Smith G. Problems of reporting genetic associations with complex outcomes. *Lancet* 2003; 361:865–872.
- 21 Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K. A comprehensive review of genetic association studies. *Genet Med* 2002; 4:45–61.
- 22 Ott J. Association of genetic loci: replication or not, that is the question. *Neurology* 2004; 63:955–958.
- 23 Cavalleri GL, Lynch JM, Depondt C, *et al.* Failure to replicate previously reported genetic associations with sporadic temporal lobe epilepsy: where to from here? *Brain* 2005; 128:1832–1840.
- This study reports that seven previously published genetic associations for sporadic TLE could not be replicated, suggesting these were false positive results due to methodological issues in the original studies.
- 24 Cardon LR, Bell JI. Association study designs for complex diseases. *Nat Rev Genet* 2001; 2:91–99.
- 25 Zondervan KT, Cardon LR. The complex interplay among factors that influence allelic association. *Nat Rev Genet* 2004; 5:89–100.
- 26 Hattersley AT, McCarthy ML. What makes a good genetic association study? *Lancet* 2005; 366:1315–1323.
- This review describes characteristics of a well designed and conducted association study, emphasizing avoidance of genotyping error and biological plausibility for associated genes.
- 27 Cardon LR, Palmer LJ. Population stratification and spurious allelic association. *Lancet* 2003; 361:598–604.
- 28 Dahlman I, Eaves IA, Kosoy R, *et al.* Parameters for reliable results in genetic association studies in common disease. *Nat Genet* 2002; 30:149–150.
- 29 Makoff A, Asherson P, Nashef L. Authors' voluntary retraction: H. Wilkie, A. Osei-Lah, B. Chioza, L. Nashef, D. McCormick, P. Asherson, and A.J. Makoff: association of the mu-opioid receptor subunit gene and idiopathic generalized epilepsy. *Neurology* 2002; 59:724–728. *Neurology* 2005; 64:579.
- 30 Clayton DG, Walker NM, Smyth DJ, *et al.* Population structure, differential bias and genomic control in a large-scale, case-control association study. *Nat Genet* 2005; 37:1243–1246.
- This study demonstrates that population structure and genotyping errors can lead to false-positive results.
- 31 Newton-Cheh C, Hirschhorn JN. Genetic association studies of complex traits: design and analysis issues. *Mutat Res* 2005; 573:54–69.
- Critical issues in design and interpretation of association studies are detailed; the importance of carefully chosen cases and controls is pointed out, and statistical issues are discussed.
- 32 Wacholder S, Chanock S, Garcia-Closas M, *et al.* Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *J Natl Cancer Inst* 2004; 96:434–442.
- 33 Palmer LJ, Cardon LR. Shaking the tree: mapping complex disease genes with linkage disequilibrium. *Lancet* 2005; 366:1223–1234.
- Examines the use of linkage disequilibrium to construct haplotypes, the use of haplotypes in reducing genome-wide genotyping costs, and the HapMap project. Also discusses sample size and power.
- 34 Wang WY, Barratt BJ, Clayton DG, Todd JA. Genome-wide association studies: theoretical and practical concerns. *Nat Rev Genet* 2005; 6:109–118.
- This discusses marker selection, the impact of allele frequencies on statistical power, and epistasis. It should be read with the companion review of Hirschhorn *et al.* [17**].
- 35 Gordon D, Finch SJ. Factors affecting statistical power in the detection of genetic association. *J Clin Invest* 2005; 115:1408–1418.
- A detailed analysis of factors that need to be considered for an adequately powered study, and how bias and errors can erode statistical power.
- 36 Purcell S, Cherny SS, Sham PC. Genetic power calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 2003; 19:149–150.
- 37 Altshuler D, Brooks LD, Chakravarti A, *et al.* A haplotype map of the human genome. *Nature* 2005; 437:1299–1320.
- A landmark study describing how the human genome is structured into blocks of high linkage disequilibrium and low haplotype diversity, and the implications of these findings. An accompanying editorial sounds a cautionary note.
- 38 Ozaki K, Ohnishi Y, Iida A, *et al.* Functional SNPs in the lymphotoxin-alpha gene that are associated with susceptibility to myocardial infarction. *Nat Genet* 2002; 32:650–654.
- 39 Carlson CS, Eberle MA, Kruglyak L, Nickerson DA. Mapping complex disease loci in whole-genome association studies. *Nature* 2004; 429:446–452.
- 40 Thomas DC, Haile RW, Duggan D. Recent developments in genomewide association scans: a workshop summary and review. *Am J Hum Genet* 2005; 77:337–345.
- A detailed analysis of problems and possible solutions to the issues facing the genome-wide association study approach. It highlights the feasibility of a multi-stage approach to such studies.
- 41 Page GP, George V, Go RC, *et al.* 'Are we there yet?': deciding when one has demonstrated specific genetic causation in complex diseases and quantitative traits. *Am J Hum Genet* 2003; 73:711–719.
- 42 Rebbeck TR, Spitz M, Wu X. Assessing the function of genetic variants in candidate gene association studies. *Nat Rev Genet* 2004; 5:589–597.
- 43 Salanti G, Sanderson S, Higgins JP. Obstacles and opportunities in meta-analysis of genetic association studies. *Genet Med* 2005; 7:13–20.
- A balanced commentary detailing strengths and weaknesses of metaanalyses, highlighting the risk of publication bias and the need for careful study selection.
- 44 Munafo MR, Flint J. Meta-analysis of genetic association studies. *Trends Genet* 2004; 20:439–444.
- 45 Bracken MB. Genomic epidemiology of complex disease: the need for an electronic evidence-based approach to research synthesis. *Am J Epidemiol* 2005; 162:297–301.
- A call for use of online electronic databases, similar to the Cochrane Collection, to collect and systematically review genetic association studies.
- 46 Attia J, Thakkinstian A, D'Este C. Meta-analyses of molecular association studies: methodologic lessons for genetic epidemiology. *J Clin Epidemiol* 2003; 56:297–303.
- 47 Kocsis I, Vasarhelyi B, Gyorffy A, Gyorffy B. Reanalysis of genotype distributions published in 'Neurology' between 1999 and 2002. *Neurology* 2004; 63:357–358.
- 48 Little J, Bradley L, Bray MS, *et al.* Reporting, appraising, and integrating data on genotype prevalence and gene-disease associations. *Am J Epidemiol* 2002; 156:300–310.
- 49 Gambardella A, Manna I, Labate A, *et al.* GABA(B) receptor 1 polymorphism (G1465A) is associated with temporal lobe epilepsy. *Neurology* 2003; 60:560–563.
- 50 Tan NC, Heron SE, Scheffer IE, *et al.* Is variation in the GABA(B) receptor 1 gene associated with temporal lobe epilepsy? *Epilepsia* 2005; 46:778–780.
- 51 Siddiqui A, Kerb R, Weale ME, *et al.* Association of multidrug resistance in epilepsy with a polymorphism in the drug-transporter gene ABCB1. *N Engl J Med* 2003; 348:1442–1448.
- 52 Tan NC, Heron SE, Scheffer IE, *et al.* Failure to confirm association of a polymorphism in ABCB1 with multidrug-resistant epilepsy. *Neurology* 2004; 63:1090–1092.
- 53 Zimprich F, Sunder-Plassmann R, Stogmann E, *et al.* Association of an ABCB1 gene haplotype with pharmacoresistance in temporal lobe epilepsy. *Neurology* 2004; 63:1087–1089.
- 54 Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Replication validity of genetic association studies. *Nat Genet* 2001; 29:306–309.
- 55 Ioannidis JP, Ntzani EE, Trikalinos TA. 'Racial' differences in genetic effects for complex diseases. *Nat Genet* 2004; 36:1312–1318.
- The results of this study suggest that the effect of genetic variants may be relatively consistent across different ethnic groups.
- 56 Cohn LD, Becker BJ. How meta-analysis increases statistical power. *Psychol Methods* 2003; 8:243–253.
- 57 Munafo MR, Clark TG, Flint J. Assessing publication bias in genetic association studies: evidence from a recent meta-analysis. *Psychiatry Res* 2004; 129:39–44.
- This examines different methods for detecting publication bias in metaanalyses, and concludes that no single test is adequate, and that several tests are required.
- 58 Becker KG, Barnes KC, Bright TJ, Wang SA. The genetic association database. *Nat Genet* 2004; 36:431–432.
- 59 Maraganore DM, Lesnick TG, Elbaz A, *et al.* UCHL1 is a Parkinson's disease susceptibility gene. *Ann Neurol* 2004; 55:512–521.
- 60 Hoh J, Ott J. Mathematical multi-locus approaches to localizing complex human trait genes. *Nat Rev Genet* 2003; 4:701–709.
- 61 Marchini J, Donnelly P, Cardon LR. Genome-wide strategies for detecting multiple loci that influence complex diseases. *Nat Genet* 2005; 37:413–417.
- This study reports that multi-locus approaches are more powerful than single-locus testing, computationally feasible, and may be quite robust to different models of gene-gene interaction.
- 62 McKeigue PM. Prospects for admixture mapping of complex traits. *Am J Hum Genet* 2005; 76:1–7.
- A comprehensive review of the principle behind admixture mapping, and providing details about current obstacles and future approaches.

- 63** Reich D, Patterson N. Will admixture mapping work to find disease genes? *Philos Trans R Soc Lond B Biol Sci* 2005; 360:1605–1607.
- 64** Smith MW, O'Brien SJ. Mapping by admixture linkage disequilibrium: •• advances, limitations and guidelines. *Nat Rev Genet* 2005; 6:623–632. A useful set of guidelines, adapted from Reich and Patterson [63], on declaring statistical significance in admixture mapping studies, as well as an examination of feasibility and power of such studies.
- 65** Darvasi A, Shifman S. The beauty of admixture. *Nat Genet* 2005; 37:118–119.
- 66** Patterson N, Hattangadi N, Lane B, *et al.* Methods for high-density admixture mapping of disease genes. *Am J Hum Genet* 2004; 74:979–1000.
- 67** Smith MW, Patterson N, Lautenberger JA, *et al.* A high-density admixture map for disease gene discovery in African Americans. *Am J Hum Genet* 2004; 74:1001–1013.
- 68** Zhu X, Luke A, Cooper RS, *et al.* Admixture mapping for hypertension loci with • genome-scan markers. *Nat Genet* 2005; 37:177–181. This study identifies two putative susceptibility loci for hypertension using African-Americans as an admixed population.
- 69** Reich D, Patterson N, Jager PL, *et al.* A whole-genome admixture scan finds a • candidate locus for multiple sclerosis susceptibility. *Nat Genet* 2005; 37:1113–1118. A large study using admixture mapping in African-Americans to identify a possible susceptibility locus for multiple sclerosis.
- 70** Risch N. 2004 Curt Stern Award Address. The SNP endgame: a multi-•• disciplinary approach. *Am J Hum Genet* 2005; 76:221–226. A call for an integrated multidisciplinary approach to tackling complex diseases.