Meta-analysis in Psychiatric Genetics

Douglas F. Levinson, MD

Address
Department of Psychiatry, University of Pennsylvania School of Medicine, 3535 Market Street, Room 4006, Philadelphia, PA 19104, USA.
E-mail: DFL@mail.med.upenn.edu

Current Psychiatry Reports 2005, **7:**143–151
Current Science Inc. ISSN 1523-3812

The article reviews literature on methods for meta-analysis of genetic linkage and association studies, and summarizes and comments on specific meta-analysis findings for psychiatric disorders. The Genome Scan Meta-Analysis and Multiple Scan Probability methods assess the evidence for linkage across studies. Multiple Scan Probability analysis suggested linkage of two chromosomal regions (13q and 22q) to schizophrenia and bipolar disorder, whereas Genome Scan Meta-Analysis on a larger sample identified at least 10 schizophrenia linkage regions, but none for bipolar disorder. Meta-analyses of pooled ORs support association of schizophrenia to the Ser311Cys polymorphism in DRD2 and the T102C polymorphism in HTR2A, and of attention deficit hyperactivity disorder to the 48-bp repeat in DRD4. The 5-HTTLPR polymorphism in the serotonin transporter gene (SLC6A4) may contribute to the risk of bipolar disorder, suicidal behavior, and neuroticism, but association to the lifetime risk of major depression has not been shown. Meta-analyses support linkage of schizophrenia to regions where replicable associations to candidate genes have been identified through positional cloning methods. There are additional supported regions where susceptibility genes are likely to be identified. Linkage meta-analysis has had less clear success for bipolar disorder based on a smaller dataset. Meta-analysis can guide the prioritization of regions for study, but proof of association requires biological confirmation of hypotheses about gene actions. Elucidation of causal mechanisms will require more comprehensive study of sequence variation in candidate genes, better statistical and meta-analytic methods to take all variation into account, and biological strategies for testing etiologic hypotheses.

Introduction

Meta-analysis refers to a set of statistical methods to combine data from different datasets. Meta-analysis has become important in psychiatric genetics because of rapid increases in the number and size of datasets. When modern molecular genetic studies first were initiated approximately 15 years ago, sample sizes typically were less than 100 families for linkage studies, or up to a few hundred cases for association studies. Studies have become larger (500 to 1000 families, or thousands of cases and control subjects, with denser marker maps), because genotyping is cheaper and faster, linkage analysis software (using more powerful hardware) is faster and more robust, and multicenter strategies have evolved for recruiting larger samples. Meta-analytic methods are useful to extract more information from older studies and to combine them with newer, larger studies.

Numerous possible sources of bias should be considered when reading meta-analyses [1•]. Some investigators decide to combine datasets without regard to previous findings and others do so because of their positive results. Publication bias (failure to publish negative results) can produce spuriously positive meta-analysis results. Neither the reader nor the analyst knows for sure whether some of the studies had high genotyping error rates. Usually, these errors produce false negative results, but an error-prone assay also could produce false positive "association." Diagnoses are never completely comparable across studies. "Narrow" (severe) categories (schizophrenia, bipolar I) are most reliable across sites [2,3], but there often is variation across studies in interviewing styles, collection of collateral data, and subjective weighting of clinical features in assigning diagnoses. Therefore, combining datasets will produce a different range of clinical features than that found in any one study. Fortunately, for each major psychiatric disorder there is a spectrum of genetically-related phenotypes, so that (for example) there is likely to be a genetic relationship between cases that "just make" a diagnosis of schizophrenia at one center but "just miss" at another. Lastly, every meta-analyst must make a set of decisions about which hypotheses to test and which studies to include. Therefore, the availability of more than one meta-analysis of the same hypothesis at a similar point in time can be more instructive than a single analysis, because each analysis sheds light on the strengths and limitations of the other.

Genetic linkage

Genetic linkage analysis tests whether there is correlation between diagnosis and inheritance of DNA marker alleles within families with two or more ill relatives (other than parent-child pairs). Linkage to continuous trait measurements also can be studied. There now are many genome scan studies of markers on all chromosomes. The failure of

early psychiatric genetic studies to find consistently replicable, significant linkage suggests an etiologic role for the combined actions of multiple genes of smaller effect rather than "major locus" effects. Meta-analysis can facilitate the search for these genes by increasing sample size. However, like any linkage analysis of a complex disorder, meta-analysis can provide support for linkage, but can never disprove that there could be undetected linkage (very weak, or strong, but only in a small minority of families). An interesting meta-analytic study looked at linkage studies of different complex disorders, and found that highly significant linkage was more likely to be observed when more families were studied, and when they were drawn from a reasonable homogeneous population ethnically [4•].

Methods

If genotypes are available for each sample, the families can be combined into a single linkage analysis, although it may be wise to apply statistical methods to determine whether results differ across sites [5,6]. Two formal meta-analysis methods have been applied to psychiatric data. One is multiple scan probability (MSP) [7•], which uses a variation of Fisher's method [8] to combine probability values from different scans while correcting for the genetic distance between peaks in different studies. The second method, genome scan meta-analysis (GSMA) [9,10•], has been more widely applied [11–17]. GSMA is a rank-ordered method that starts by dividing the genome into segments (120 "bins" of ~ 30 centiMorgans). For each scan, each bin is assigned the best linkage score observed within that segment, and these scores are used to rank each bin. The ranks can be weighted for sample size. For each bin, the ranks are summed across studies. Probability values can be determined for unweighted tanks from a theoretical distribution, or for weighted ranks by permutation test, with Bonferroni's correction to determine genome-wide significance. A simulation study also showed that, when there are many "linked" bins, they tend to have probability values less than 0.05 (without correction) for the summed/averaged ranks (P_{AvgRnk}), and for a statistic (P_{ord}) which determines whether (for example) the summed rank of the fifth-place bin is higher than expected by chance for fifth-place bins [10]. When sets of genome scans were simulated with no linkage present, less than 5% of GSMAs detected five or more bins with both probability values less than 0.05, but this threshold was frequently exceeded when linkage was simulated in multiple bins. True linkage also more often produced significant results in adjacent bins. This threshold often was met even when the genetic effect was too weak to produce a significant linkage score. Therefore, GSMA may be more sensitive to weak linkage, although at the expense of poor localization.

Results

Several collaborations have combined samples for linkage analyses of candidate regions (chromosomal regions that had produced positive results in a reported study, but not in all samples). For example, two collaborative studies (using partially overlapping samples) provided early support for the linkage to schizophrenia on chromosome 6p [18,19], and the second of these provided stronger support for linkage of schizophrenia to chromosome 8p [19]. A multicenter study also supported linkage on chromosome 6q [20]. A combined analysis of bipolar disorder in nine extended pedigrees from three studies failed to find a common location for the linkage peaks that had been observed in these studies [21].

Genome scan meta-analyses [22••,23•] and MSP [24•] analyses have been published for schizophrenia and bipolar disorder (see Tables 1 and 2 for a summary of the GSMA results). The MSP analysis also combined the bipolar and schizophrenia scans, because it has been hypothesized that some loci predispose to both disorders [25]. The samples used for the GSMA studies are described in Tables 1 and 2. The MSP analyses used 18 schizophrenia genome scan datasets (681 pedigrees, 1929 patients with schizophrenia or schizoaffective disorder), 11 bipolar datasets (353 pedigrees, 1228 patients with bipolar I or II disorder, corresponding to the Intermediate model of the GSMA shown in Table 2), and both sets of scans combined. Therefore, the MSP analysis considered a smaller number of families in each diagnostic category, drawn only from published data, whereas the GSMAs added data provided by investigators. The results also were different. For schizophrenia, both methods produced positive results for the same regions of chromosomes 8p and 22q. However, the MSP analysis detected significant linkage on chromosome 13q and 22q for both disorders, and for the two disorders combined, whereas the GSMA studies did not detect linkage on 13q for either disorder, nor for 22q for bipolar disorder. Therefore, the MSP analysis concluded that there is a genetic relationship between these two disorders, whereas the GSMA of a larger dataset does not support that conclusion. Because the MSP sample was not a simple subset of the GSMA sample, the results cannot be directly compared, but it is likely that the main factor responsible for the differing results was the differences in the datasets, given that R. Segurado reported on a GSMA of published data comparable to that used in the MSP, with similar conclusions (oral presentation, World Congress on Psychiatric Genetics, St. Louis, October, 2001).

The results of the schizophrenia GSMA are highly encouraging. There were 12 bins in 10 chromosomal regions with P_{AvgRnk} and P_{ord} less than 0.05. This would be extremely rare in the absence of linkage in many of these bins [10•]. There were additional regions with P_{ord} less than 0.05, and simulation data suggest that some of these will ultimately prove to contain linked loci. These linkage regions include the locations of several genes for which association evidence has now been reported in several samples each, and which were originally identified through systematic association studies of dense

Table 1. Genome scan meta-analysis of 20 schizophrenia genome scans [11,27]

Chromosome	Bin	Begin-end cM	Cytogenetic loc	P_{AvgRnk}	P_{ord}	Both<0.05
1	6	142.2–170.8	1p13.3–q23.3	*	*	†
1	7	170.8–201.6	1q23.3–q31.1	‡	*	
2	5	101.6–128.4	2p12–q22.1	§	*	†
2	6	128.4–154.5	2q22.1–q23.3	*	*	†
3	2	32.4–63.1	3p25.3–p22.1	*	*	†
5	5	131.5–164.2	5q23.2–q34	*	*	†
6	1	0–32.6	6pter–p22.3	*	*	†
6	2	32.6–65.1	6p22.3–p21.1	*	*	†
6	4	99.0–131.1	6q15–q23.2	‡	*	
8	2	27.4–55.0	8p22–p21.1	*	*	†
10	1	0–29.2	10pter–p14	‡	*	
11	5	99.0–123.0	11q22.3–q24.1	*	*	†
14	1	0–40.1	14pter–q13.1	*	*	†
15	3	52.3–85.6	15q21.3–q26.1	‡	*	
16	2	32.1–67.6	16p13–q12.2	‡	*	
17	3	63.6–94.0	17q21.33–q24.3	‡	*	
18	4	96.5–126.0	18q22.1–qter	‡	*	
20	2	21.2–47.5	20p12.3–p11	*	*	†
22	1	0–33.8	22pter–q12.3	*	*	†

*Uncorrected P_{AvgRnk} or P_{ord} <0.05
†Both *P*-values were <0.05
‡Uncorrected P_{AvgRnk} or P_{ord} <0.10
§*P*-value was significant after correcting for 120 tests
P_{AvgRnk}—*P*-value for the average weighted rank achieved by that bin across studies, obtained by a permutation test; P_{ord}–probability of observing a given average weighted rank by chance considering the order of the bins.

Table 2. Genome scan meta-analysis of 18 bipolar disorder genome scans [28]

				Diagnostic model		
Chromosome	Bin	Begin-end cM	Cytogenetic loc	BP-I/SAB	+BP-II	+MDDR
1	4	83.07–113.69	1p32.1–q31.1	*		
2	6	128.41–154.48	2q22.1–q23.3		*	†
7	6	122.48–148.11	7q34–qter		†	*
8	1	0–27.4	8pter–p22		*	
8	6	137.92–167.9	8q24.21–qter		*	*
9	3	53.6–84.9	9p21.1–q21.32	*	†	†
10	3	62.23–91.13	10q11.21–q22.1	*		
14	3	74.96–105	14q24.1–q32.12	‡	*	*
17	2	25.14–63.62	17p12–q21.33	*		
18	1	0–24.08	18pter–p11	‡		
18	2	24.08–62.84	18p11–q12.3	*	*	*
18	3	62.84–96.48	18q12.3–q22.1	†	†	*
19	4	75.41–105.0	19q13.33–qter	*		

*Uncorrected P_{AvgRnk} or P_{ord} <0.05
†Uncorrected P_{AvgRnk} or P_{ord} <0.10
‡Both *P*-values were <0.05

maps of markers in linkage regions: DTNBP1 (dysbindin, bin 6.1; bin 6.2 was also positive in the GSMA) [26–28], and catechol-O-methyltransferase (COMT) in the region of the microdeletion responsible for velo-cardio facial syndrome (bin 22.1) [29] (PRODH2 [30], in the same region, is reported to be associated to schizophrenia but this has not yet been replicated beyond the original report). A third promising candidate gene, NRG1 (neuregulin-1, bin 8.3), was detected by systematic association mapping and then replicated in additional samples [31–35]. However, the adjacent bin on chromosome 8p (8.2) is the one implicated by most of the linkage evidence. One wonders whether some of the more common linkage regions contain more than one susceptibility gene, and this may prove to be the case for 6p [36] and 8p [37].

Genome scan meta-analysis supported linkage in several other regions, including the proximal part of chromosome 1q, which has produced evidence for linkage in some samples [38,39], but not others [40]. Bin 2.5 achieved the single most significant result (a corrected genome-wide significant P_{AvgRnk}), with P_{AvgRnk} and P_{ord} less than 0.05 in bins 2.5 and 2.6. Linkage peaks in this region have varied in location and significance levels generally have been modest, as also has been the case for chromosome 3p, which has never produced striking evidence for linkage in a sample. It probably will require systematic association mapping of these regions to determine whether there are one or more susceptibility genes in each of them. MSP supported schizophrenia linkage on chromosome 13q, but GSMA (using a larger dataset) did not. This region produced significant linkage results in two samples [41,42], near the G72 gene for which there is evidence for association to schizophrenia and bipolar disorder [43–49]. This pattern of results offers no simple explanation.

The bipolar disorder GSMA was more disappointing: there was no significant result by any metric. There were far fewer bipolar-I families (approximately 400) available for this GSMA than there were schizophrenia families (approximately 1200), so it is possible that future meta-analyses will be more successful when newer studies are included [50–52]. It was surprising to some investigators that broader diagnostic models did not produce more significant GSMA results despite the larger sample sizes.

Multiple scan probability analysis of four autism genome scans supported linkage on chromosome 7q [7•]. Additional collaborative autism linkage data are expected in the near future.

Genetic Association

Linkage analysis identifies a wide chromosomal region likely to contain susceptibility gene(s). Association analysis identifies a specific DNA "polymorphism" that increases disease risk or that is close to such a sequence (a "polymorphism" can be a single-base change, or an insertion, deletion, inversion or repetition of a sequence of bases). Linkage is a correlation between disease and marker alleles within families, whereas association is a correlation across families. In linkage analysis, most subjects may carry the "1" marker allele in one family, and the "2" in another. Meiotic recombination is more and more likely to re-shuffle these relationships in succeeding generations as the distance increases between the disease locus and the nearest marker. Association is a correlation between an allele and disease in the entire sample (one also can test a haplotype or alleles at several locations on one chromosome). Association analysis essentially treats subjects as members of one huge family, testing for correlations dating back thousands or millions of years. An association can be observed by chance (false positive), or if the associated allele alters disease risk, or if it is so close to a susceptibility polymorphism that meiotic recombination seldom separates them, so that the marker allele that was next to the "ancestral" disease mutation is still next to it much of the time (linkage disequilibrium). Numerous meta-analyses have been reported of psychiatric association studies, mostly for genes relevant to dopaminergic and serotonergic neurotransmission.

Methods

Association can be studied with case-control and family-based analyses. The former tests allele, genotype, or haplotype frequency differences between ill individuals and control subjects (well individuals or a random population sample). The classic, family-based (transmission disequilibrium) analysis determines, for each case-parent trio, whether a parent "transmits" the allele of interest to the patient significantly more or less than the expected 50% of the time [53]. Other family constellations can be analyzed [54–56]. Case-control studies require careful matching of the two groups, because allele frequencies vary in world populations because of chance events in their reproductive and migratory histories. Family-based analysis eliminates this issue, but for adult diseases, it is often difficult to recruit both parents. Meta-analyses have considered case-control and family-based studies separately or together.

Meta-analysis requires estimating a pooled OR. If the ratio of the two allele frequencies is 3:1 in cases and 2:1 in control subjects, then the OR is 3/2 or 1.5. Depending on sample size, an OR has a CI with a 95% probability of including the population's true OR. If this CI includes 1.0, then the groups do not differ (P>0.05). One cannot compute an OR from frequencies summed across studies. This would be comparing apples (cases from one population) with oranges (a large control group from a second population with different allele frequencies). ORs are statistically combined using either fixed effects meta-analysis (ORs are pooled in a way that assumes a uniform underlying effect across studies) or random effects meta-analysis which assumes some variability across populations and usually results in a wider CI (more conservative test) [57•,58]. Random effects meta-analysis generally is preferred, but the two methods produce similar results [59••]. Well-conducted meta-analyses select studies based on clear criteria and an exhaustive search of English and non-English journals, with resolution of sample overlap in different reports. Ideally, there should be statistical tests of heterogeneity among samples (subgroups of samples with different genetic effects), of sensitivity (the effect of dropping each study in turn), and of publication bias (whether negative studies seem to be missing from the literature) – smaller studies produce a wide range of ORs above and below the pooled result, and one can test whether this range is truncated (absence of the expected less significant ORs) [60,61]. The sensitivity of this test to small biases is unclear. Therefore, CIs that barely exclude 1.0 must be regarded with caution. Multiple testing can be a problem if

Table 3. Meta-analyses of psychiatric disorders and traits for which at least one analysis showed significant association

Disorder, polymorphism	Study	Samples	*n*	Effect	OR	CI	*P*	Comments
Schizophrenia								
DRD2 Ser311Cys	73	24	3733	G(Cys):	1.3	1.1–1.6	0.007	
	74	26	3506	G(Cys):	1.43	1.16–1.78	<0.001	
DRD3 Ser9Gly c-c:	75	44	5430	Hom:	1.08	1–1.16	ns	
				1-1:	1.05	0.97–1.13	ns	
				1:	1	0.95–1.06	ns	
TDT:	76	8	463	Hom:	1.24	0.91–1.70	ns	
				1-1:	1.23	0.90–1.67	ns	
				1:	10.9	0.89–1.35	ns	
c-c and TDT:	69	48	nr	1-1:	1.13	1.02–1.25	<0.05	a; pub bias, *P*=0.04
DRD4 prom~521C/T	77	3	592	Alleles:	1.22	1.04–1.4	<0.02	
HTR2A T102C c-c:	78	31	4632	C:	1.129*	1.02–1.25	0.015	*P*<0.001 Eur
TDT:	78	5	473	C:	1.3	0.9–1.8	<0.05	a
Mood/anxiety								
SLC6A4 5-HTTLPR								
BP c-c:	79	15	2774	S:	1.13	1.05–1.22	0.001	
BP c-c:	80	12	1356	S; GTs:	nr		ns	
BP TDT:	80	6	nr	S:	nr		ns	
UP c-c:	79	14	1961	S:	1.05	0.96–1.14	ns	
UP c-c:	80	10	910	S; GTs:			ns	P<0.05 Eur, S/S, sensitive to one study
SP vs CS:	81	17	1521	SS+SL:	1.21	0.94–1.57	ns	NS for S
SP vs CS:	82	12	1168	S:	1.17	1.04–1.32	0.009	+ for SS, SS+SL, -for completers
SP vs NSP:	81	8	511	SS+SL:	1.55	1.15–2.10	0.004	+ alcoholism only; - for S
VSP vs CS:	81	5	190	SS+SL:	3.32	1.51–7.31	0.003	+ for violent vs nonviolent
NVSP vs CS:	81	5	375	SS+SL:	0.94	0.71–1.25	ns	- for S
Neuroticism/harm avoidance:	83	23	5629	All:			0.087	T scores; mean 50, SD 10
				NEO:			0.000016	For 10 studies using the NEO/Neuroticism
ADHD								
DRD4 48-bp rpt c-c:	84	8	1266	7-rpt:	1.9	1.5–2.2	0.001	overlap with TDT samples
TDT:	84	13	1665	7-rpt:	1.4	1.1–1.6	0.02	

a–omitted the first positive report; ADHD—attention deficit hyperactivity disorder; BP—bipolar; CS—control subjects; c-c—case-control studies; Eur—European ancestry; GT—genotype; Hom—homozygosity; NEO—NEO personality inventory; NSP—non-suicidal patient; NVSP—patients with non-violent suicidal behavior ; rpt—repeats; S—short; SP—suicidal patient; TDT—family-based study; UP—unipolar; VSP—patients with violent suicidal behavior

*More precise OR provided by S. Glatt (personal communication)

several effects are tested (alleles, genotypes, phenotypes) – and 5% of meta-analyses will be "significant" by chance. Lastly, the current author advocates contacting authors when necessary to inquire about sample overlap and to obtain unpublished details to permit comparison with other studies. This produces larger and more accurate meta-analyses.

Results

Meta-analysis has supported significant association of four polymorphisms to schizophrenia, one to certain clinical features related to mood and anxiety disorders, and one to attention deficit disorder (Table 3) [62–73]. The many negative meta-analyses will not be reviewed here (details available from the author on request), nor will pharmacogenomic studies be considered.

For schizophrenia, genes related to dopaminergic neurotransmission have been most widely studied, particularly DRD2, DRD3 and DRD4 which encode receptors in the D2-like receptor family (inhibitory effect on cyclic-AMP production and high affinity for antipsychotic drugs) [74]. In DRD2 (chromosome 11q23), a Ser311Cys amino acid change (T-to-G substitution) lowers the receptor's DA affinity [75,76]. Two recent meta-analyses found significant association of schizophrenia to the G/Cys variant. Each of the two available recent meta-analyses [62,63] has some flaws (undetected sample overlap [62] or omitted studies [63], details available on request), but the consistency in the results suggests a likely significant association. In DRD3 (3q13), an A-to-G substitution produces a Ser9Gly amino acid change [77]. Recent meta-analyses of the many studies of this polymorphism produced contradictory conclusions. Separate meta-analyses of case-control and family-based studies found no significant association for alleles, genotypes or homozygosity [64,65]. A meta-analysis that combined the two types of studies (omitting the first positive report) found a modest, significant association for the 1-1 genotype, but also significant publication bias (P=0.04), and several studies seem to have been omitted from that analysis [59]. Although a very modest association is possible, it has not been convincingly established. In DRD4 (11p15), an analysis of only three studies [66] found modestly significant association between schizophrenia and a ~ 521C/T variant in the promoter region that influences transcription [78]. This finding requires additional study. Lastly, in HTR2A (13q14) (type 2A serotonin receptor), the C allele of an anonymous T102C substitution [79] in the 3′-UTR region is apparently associated with schizophrenia [67].

The possible association of schizophrenia to COMT (22q11) is unclear. Meta-analyses [59,80] found no association to the well-known Val158/108Met functional polymorphism.[81] Subsequently, Shifman *et al.* [29] reported very weak association of the G allele (P=0.024) in a large sample, but addition of these data to previous studies does not produce a significant result by random effects meta-analysis (unpublished analysis by the present author, results available on request). However, Shifman *et al.* [29] reported much stronger association (P<0.001 and P<0.0001) for two other COMT single nucleotide polymorphisms (SNPs), and another study [82] suggests that the association could be with either COMT or an adjacent gene, ARVCF. This region deserves more comprehensive study in larger samples.

For mood disorders, the only significant meta-analyses have been for the 44-bp insertion/deletion (5-HTTLPR) in the promoter region of SLC6A4 (17q11), the gene encoding the serotonin transporter (re-uptake site) which is a target of antidepressant drugs. The "short" allele reduces transcription and functional activity [83]. The larger [68] (but not the smaller [69]) of two recent meta-analyses of case-control data found a significant association between the short allele and bipolar disorder; neither analysis found an association with major depression. The short allele was found to be associated with suicidal behavior and ideation only in the smaller [71] of two recent meta-analyses, but this was not observed in the larger meta-analysis [70], which found an excess of SS+SL genotypes for suicidal versus non-suicidal patients (but in alcoholic and not mood disorder patients), and for those with violent suicidal behavior versus control subjects or versus patients with non-violent suicidal behavior. Short allele carriers have been reported to have an increased risk of major depression onsets after life stress [84], an effect that was recently replicated (Wilhelm *et al.*, poster presentation at the World Congress of Psychiatric Genetics, Dublin, October, 2004), but not an increased overall risk of major depression. Thus 5-HTTLPR could be more closely related to emotional reactivity, or to personality characteristics relevant to reactivity and depression, as suggested by a large meta-analysis that found a trend toward association of 5-HTTLPR genotypes and neuroticism, harm avoidance, or related personality traits, with a small (0.106 SD) but highly significant effect of the S/S genotype in the subset of studies using the NEO inventory to measure neuroticism [72].

Lastly, a meta-analysis found an association of the DRD4 48-bp repeat polymorphism and attention deficit hyperactivity disorder in family-based analyses, and in a subset of studies that also carried out case-control analyses (in cases overlapping with the family samples) [73].

Discussion

Modern meta-analysis methods can derive valuable information from large numbers of psychiatric genetic linkage and association studies. These methods do have limitations. For example, they depend on the quality of each study. The analyst cannot know whether some of the apparently eligible studies are introducing undetected sources of error or other variation. Each meta-analysis also involves decisions about whether to test one hypothesis to

preserve power, or multiple hypotheses (genetic effects, diagnostic models, ethnic subgroups, pedigree types, methodological differences) with more chances to detect effects, but with possible loss of power. A meta-analysis of all data cannot be replicated in a second sample. Despite these limitations, the results summarized above show that significant effects can be observed, thereby providing guidance to investigators prioritizing future studies.

The positional cloning strategy has achieved dramatic successes in the study of schizophrenia: numerous linkage regions have shown some consistency of results across studies, and highly plausible candidate genes have been identified in some of these regions with association mapping ("linkage disequilibrium" mapping)— genotyping very dense maps of DNA markers (separated by thousands of base pairs rather than millions or tens of millions in linkage studies) across a region to determine in which genes there might be evidence of association to disease.

The reader will note the absence of meta-analyses of these "new schizophrenia genes." The findings in these genes are not single associated polymorphisms, but rather various haplotypes of SNPs or other markers that seem to be associated in different samples (although a single associated NRG1 haplotype was seen in two samples [36,37]). In such cases, there is no single hypothesis for a meta-analysis. Only in the past few years has it been feasible to plan studies in which large numbers of SNPs are tested in each gene of interest. Someday it may be possible to study nearly all polymorphic sites in a candidate gene and to analyze their effects jointly, permitting each study to test the hypothesis of "association in this gene [85,86,87•]. Anyone who has attempted a gene-based analysis in a reasonably long gene can confirm that the power of these methods is not yet clear, but it seems likely that as they evolve, they could provide a basis for meta-analyses.

No psychiatric disorder candidate gene has yet been studied sufficiently comprehensively in enough subjects to arrive at definitive conclusions about the genetic mechanisms by which sequence variation is contributing to pathology. The ultimate test of a hypothesis of genetic association is biological, not statistical. The actions and interactions of a gene and the effects of sequence variation in the gene must be sufficiently well understood to provide a plausible and testable explanation for its role in disease susceptibility. This requires convergence of multiple sources of biological evidence, supported by statistical association.

Acknowledgments

This work was supported by grants K24-MH64197, R01-MH062276 and R01-MH61675 from the US National Institute of Mental Health. The assistance of Drs. Pablo V. Gejman, Kenneth S. Kendler and Claudine Laurent in providing suggestions about this manuscript is gratefully acknowledged.

References and Recommended Reading

Papers of particular interest, published recently, have been highlighted as:

- • Of importance
- •• Of major importance

1.• Egger M, Smith GD, Sterne JA: **Uses and abuses of meta-analysis.** *Clin Med* 2001, **1:**478–484.

This is an excellent review on meta-analysis, with a focus on the importance of providing a critical review of the literature and not simply a statistical pooling of results.

2. Pulver AE, Karayiorgou M, Lasseter VK, *et al.*: **Follow-up of a report of a potential linkage for schizophrenia on chromosome 22q12-q13.1: Part 2.** *Am J Med Genet* 1994, **54:**44–50.
3. Nurnberger Jr JI, Blehar MC, Kaufmann CA, *et al.*: **Diagnostic interview for genetic studies. Rationale, unique features, and training. NIMH Genetics Initiative.** *Arch Gen Psychiatry* 1994, **51:**849–859.
4.• Altmüller J, Palmer LJ, Fischer G, *et al.*: **Genomewide scans of complex human diseases: true linkage is hard to find.** *Am J Hum Genet* 2001, **69:**936–950.

This is an important report of a meta-analysis, not of linkage studies of one disease, but of genome scans for many complex disorders, to attempt to identify factors that predict which studies obtain significant results. The two most important predictors proved to be larger sample size, and restriction of a sample to a single reasonably homogeneous ethnic group (i.e., "European" or "Chinese" families).

5. Rice JP: **The role of meta-analysis in linkage studies of complex traits.** *Am J Med Genet* 1997, **74:**112–114.
6. Dorr DA, Rice JP, Armstrong C, *et al.*: **Meta-analysis of chromosome 18 linkage data for bipolar illness.** *Genet Epidemiol* 1997, **14:**617–622.
7.• Badner JA, Gershon ES: **Regional meta-analysis of published data supports linkage of autism with markers on chromosome.** *Mol Psychiatry* 2002, **7:**56–66.

This paper includes the description of Multiple Scan Probability analysis, a method of meta-analysis based on Fisher's formula for combining p-values, modified to take into account the genetic distance between the locations of peak p-values in different studies. The method was also applied to schizophrenia and bipolar disorder (reference 29).

8. Province MA: **The significance of not finding a gene.** *Am J Hum Genet* 2001, **69:**660–663.
9. Wise LH, Lanchbury JS, Lewis CM: **Meta-analysis of genome searches.** *Ann Hum Genet* 1999, **63:**263–72.
10.• Levinson DF, Levinson MD, Segurado R, Lewis CM: **Genome scan meta-analysis of schizophrenia and bipolar disorder, part I: Methods and power analysis.** *Am J Hum Genet* 2003, **73:**17–33.

This is a report on a comprehensive simulation study to explore the properties of the Genome Scan Meta-Analysis (GSMA) method (reference 10) in datasets modelled after the schizophrenia and bipolar disorder analyses reported in references 27 and 28. Not all GSMA studies (see references 12-18) have utilized the empirical significance thresholds derived in this study, but they appear to make GSMA more sensitive to weak linkage effects.

11. Chiodini BD, Lewis CM: **Meta-analysis of 4 coronary heart disease genome-wide linkage studies confirms a susceptibility locus on chromosome 3q.** *Arterioscler Thromb Vasc Biol* 2003, **23:**1863–1868.
12. Demenais F, Kanninen T, Lindgren CM, *et al.*: **A meta-analysis of four European genome screens (GIFT Consortium) shows evidence for a novel region on chromosome 17p11.2-q22 linked to type 2 diabetes.** *Hum Mol Genet* 2003, **12:**1865–1873.
13. Liu W, Zhao W, Chase GA: **Genome scan meta-analysis for hypertension.** *Am J Hypertens* 2004, **17:**1100–1106.
14. Marazita ML, Murray JC, Lidral AC, *et al.*: **Meta-analysis of 13 genome scans reveals multiple cleft lip/palate genes with novel loci on 9q21 and 2q32-35.** *Am J Hum Genet* 2004, **75:**161–173.

15. Sagoo GS, Tazi-Ahnini R, Barker JW, *et al.*: **Meta-analysis of genome-wide studies of psoriasis susceptibility reveals linkage to chromosomes 6p21 and 4q28-q31 in Caucasian and Chinese Hans population.** *J Invest Dermatol* 2004, **122:**1401–1405.

16. van Heel DA, Fisher SA, Kirby A, *et al.*: **Genome Scan Meta-Analysis Group of the IBD International Genetics Consortium. Inflammatory bowel disease susceptibility loci defined by genome scan meta-analysis of 1952 affected relative pairs.** *Hum Mol Genet* 2004, **13:**763–770.

17. Williams CN, Kocher K, Lander ES, *et al.*: **Using a genome-wide scan and meta-analysis to identify a novel IBD locus and confirm previously identified IBD loci.** *Inflamm Bowel Dis* 2002, **8:**375–381.

18. Moises HW, Yang L, Kristbjarnarson H, *et al.*: **An international two-stage genome-wide search for schizophrenia susceptibility genes.** *Nat Genet* 1995, **11:**321–324.

19. Schizophrenia Linkage Collaborative Group for Chromosomes 3, 6 and 8: **Additional support for schizophrenia linkage on chromosomes 6, 8: a multicenter study.** *Am J Med Genet* 1996, **67:**580–594.

20. Levinson DF, Holmans P, Straub RE, *et al.*: **Multicenter linkage study of schizophrenia candidate regions on chromosomes 5q, 6q, 10p, and 13q: schizophrenia linkage collaborative group III.** *Am J Hum Genet* 2000, **67:**652–663.

21. Visscher PM, Haley CS, Ewald H, *et al.*: **Joint multi-population analysis for genetic linkage of bipolar disorder or "wellness" to chromosome 4p.** *Am J Med Genet* 2004

22.•• Lewis CM, Levinson DF, Wise LH, *et al.*: **Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: Schizophrenia.** *Am J Hum Genet* 2003, **73:**34–48.

This meta-analysis of schizophrenia genome scan projects provides evidence supporting at least 10 (and possibly more) chromosomal regions likely to contain schizophrenia linkage. Several of the best-supported schizophrenia candidate genes are in some of these regions and were discovered as a result of systematic association mapping of linkage findings. These results demonstrate the utility of systematic positional cloning as a strategy to identify multiple genes that influence risk for common, genetically complex disorder.

23.• Segurado R, Detera-Wadleigh SD, Levinson DF, *et al.*: **Genome scan meta-analysis of schizophrenia and bipolar disorder, part III: Bipolar disorder.** *Am J Hum Genet* 2003, **73:**49–62.

This GSMA of bipolar disorder was carried out using the same methods as the schizophrenia analysis reported in reference 22, but no significant findings were observed. The dataset was larger than that in reference 24, and differed in some other ways in the selection of studies and in the data provided by authors for analysis beyond the published data in some cases. It was striking to see how few families (under 400) had been collected worldwide for linkage analysis using a "narrow" diagnostic model (bipolar-I and schizoaffective disorder-bipolar type). Fortunately, additional large datasets (references 50-52)have since become available and can be analyzed in the future.

24.• Badner JA, Gershon ES: **Meta-analysis of whole-genome linkage scans of bipolar disorder and schizophrenia.** *Mol Psychiatry* 2002, **7:**405–411.

This analysis, using the MSP method [7], produced different results than the two GSMA studies of these disorders [22,23], as discussed in the text. Although it is likely that the main contributing factor was the difference in the datasets that were studied, more work will need to be done to understand whether MSP has advantages over GSMA in some situations.

25. Berrettini W: **Bipolar disorder and schizophrenia: convergent molecular data.** *Neuromolecular Med* 2004, **5:**109–17.

26. Straub RE, Jiang Y, MacLean CJ, *et al.*: **Genetic variation in the 6p22.3 gene DTNBP1, the human ortholog of the mouse dysbindin gene, is associated with schizophrenia.** *Am J Hum Genet* 2002, **71:**337–348.

27. Schwab SG, Knapp M, Mondabon S, *et al.*: **Support for association of schizophrenia with genetic variation in the 6p22.3 gene, dysbindin, in sib-pair families with linkage and in an additional sample of triad families.** *Am J Hum Genet* 2003, **72:**185–190.

28. Van Den Bogaert A, Schumacher J, Schulze TG, *et al.*: **The DTNBP1 (dysbindin) gene contributes to schizophrenia, depending on family history of the disease.** *Am J Hum Genet* 2003, **73:**1438-43.

29. Shifman S, Bronstein M, Sternfeld M, *et al.*: **A highly significant association between a COMT haplotype and schizophrenia.** *Am J Hum Genet* 2002, **71:**1296–302.

30. Liu H, Heath SC, Sobin C, *et al.*: **Genetic variation at the 22q11 PRODH2/DGCR6 locus presents an unusual pattern and increases susceptibility to schizophrenia.** *Proceedings of the Nat Acad Sci U S A* 2002, **99:**3717–3722.

31. Stefansson H, Sigurdsson E, Steinthorsdottir V, *et al.*: **Neuregulin 1 and susceptibility to schizophrenia.** *Am J Hum Genet* 2002, **71:**877–892.

32. Stefansson H, Sarginson J, Kong A, *et al.*: **Association of neuregulin 1 with schizophrenia confirmed in a Scottish population.** *Am J Hum Genet* 2003, **72:**83–87.

33. Tang JX, Chen WY, He G, *et al.*: **Polymorphisms within 5′ end of the Neuregulin 1 gene are genetically associated with schizophrenia in the Chinese population.** *Mol Psychiatry* 2004, **9:**11–12.

34. Yang JZ, Si TM, Ruan Y, *et al.*: **Association study of neuregulin 1 gene with schizophrenia.** *Mol Psychiatry* 2003, **8:**706–709.

35. Williams NM, Preece A, Spurlock G, *et al.*: **Support for genetic variation in neuregulin 1 and susceptibility to schizophrenia.** *Mol Psychiatry* 2003, **8:**485–487.

36. Straub RE, Matsumoto M, Egan MF, *et al.*: **MRDS1 (6p24.3) is associated with schizophrenia in both adult onset and childhood onset schizophrenia families.** *Am J Med Genet Part B (Neuropsychiatric Genetics)* 2003, **122B:**18.

37. Datta SR, McQuillin A, Rizig MA, *et al.*: **Further genetic analyses of the pcm1 gene association with schizophrenia on chromosome 8p21 and tests of the G72, dysbindin, RGS4, calcineurin, COMT, frizzled 3, MRDS1, AKT1 and CAPON associations.** *Am J Med Genet Part B (Neuropsychiatric Genetics)* 2004, **130B:**86.

38. Brzustowicz LM, Hodgkinson KA, Chow EW, *et al.*: **Location of a major susceptibility locus for familial schizophrenia on chromosome 1q21-q22.** *Science* 2000, **288:**678–682.

39. Gurling HM, Kalsi G, Brynjolfson J, *et al.*: **Genomewide genetic linkage analysis confirms the presence of susceptibility loci for schizophrenia, on chromosomes 1q32.2, 5q33.2,, 8p21-22, provides support for linkage to schizophrenia, on chromosomes 11q23.3-24, 20q12.1-11.23.** *Am J Hum Genet* 2001, **68:**661–673.

40. Levinson DF, Holmans PA, Laurent C, *et al.*: **No major schizophrenia locus detected on chromosome 1q in a large multicenter sample.** *Science* 2002, **296:**739–741.

41. Blouin JL, Dombroski BA, Nath SK, *et al.*: **Schizophrenia susceptibility loci on chromosomes 13q32, 8p21.** *Nat Genet* 1998, **20:**70-73.

42. Brzustowicz LM, Honer WG, Chow EW, *et al.*: **Linkage of familial schizophrenia to chromosome 13q32.** *Am J Hum Genet* 1999, **65:**1096–103.

43. Chumakov I, Blumenfeld M, Guerassimenko O, *et al.*: **The family based association test method: strategies for studying general genotype-phenotype associations.** *Eur J Hum Genet* 2001, **9:**301–306.

44. Zou F, Li C, Duan S, *et al.*: **A family-based study of the association between the G72/G30 genes and schizophrenia in the Chinese population.** *Schizophr Res* 2005, **73:**257–261.

45. Wang X, He G, Gu N, *et al.*: **Association of G72/G30 with schizophrenia in the Chinese population.** *Biochem Biophys Res Commun* 2004, **319:**1281–1286.

46. Addington AM, Gornick M, Sporn AL, *et al.*: **Polymorphisms in the 13q33.2 gene G72/G30 are associated with childhood-onset schizophrenia and psychosis not otherwise specified.** *Biol Psychiatry* 2004, **55:**976–980.

47. Schumacher J, Jamra RA, Freudenberg J, *et al.*: **Examination of G72 and D-amino-acid oxidase as genetic risk factors for schizophrenia and bipolar affective disorder.** *Mol Psychiatry* 2004, **9:**203–207.

48. Chen YS, Akula N, Detera-Wadleigh SD, *et al.*: **Findings in an independent sample support an association between bipolar affective disorder and the G72/G30 locus on chromosome 13q33.** *Mol Psychiatry* 2004, **9:**87–92.
49. Hattori E, Liu C, Badner JA, *et al.*: **Polymorphisms at the G72/G30 gene locus, on 13q33, are associated with bipolar disorder in two independent pedigree series.** *Am J Hum Genet* 2003, **72:**1131–1140.
50. Middleton FA, Pato MT, Gentile KL, *et al.*: **Genomewide linkage analysis of bipolar disorder by use of a high-density single-nucleotide-polymorphism (SNP) genotyping assay: a comparison with microsatellite marker assays and finding of significant linkage to chromosome 6q22.** *Am J Hum Genet* 2004, **74:**886–897.
51. Fallin MD, Lasseter VK, Wolyniec PS, *et al.*: **Genomewide linkage scan for bipolar-disorder susceptibility loci among Ashkenazi Jewish families.** *Am J Hum Genet* 2004, **75:**204–219.
52. McInnis MG, Dick DM, Willour VL, *et al.*: **Genome-wide scan and conditional analysis in bipolar disorder: evidence for genomic interaction in the National Institute of Mental Health genetics initiative bipolar pedigrees.** *Biol Psychiatry* 2003, **54:**1265–1273.
53. 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.
54. Clayton D: **A generalization of the transmission/disequilibrium test for uncertain-haplotype transmission.** *Am J Hum Genet* 1999, **65:**1170–1177.
55. Horvath S, Xu X, Laird NM: **The family based association test method: strategies for studying general genotype-phenotype associations.** *Eur J Hum Genet* 2001, **9:**301–306.
56. Martin ER, Monks SA, Warren LL, Kaplan NL: **A test for linkage and association in general pedigrees: the pedigree disequilibrium test.** *Am J Hum Genet* 2000,**67:**146–154.
57.• Egger M, Smith GD, Phillips AN: **Meta-analysis: principles and procedures.** *BMJ* 1997, **315:**1533–1537.

This is not a very recent paper, but it provides a highly readable review of the meta-analysis methods currently being used for pooling of odds ratios of genetic association studies, and references for the various methods.

58. Sutton AJ, Abrams KR, Jones DR, *et al.*: *Methods for Meta-Analysis in Medical Research.* New York: Wiley; 2000.
59.•• 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.

This paper includes meta-analyses of 25 reported associations of specific polymorphisms to common disorders, including psychiatric studies as listed in Tables 4 and 5. The authors make the valuable point that association results will tend to be quite variable in small samples due to lack of power, but that large samples or meta-analyses not infrequently detect significant association for polymorphisms that have produced positive results in some and negative results in other (mostly small) samples. The value of the individual meta-analyses reported in the paper is somewhat diminished by the failure of the journal to provide the details of each analysis (the studies included and the allele or genotype counts from each study) either in the text or online.

60. Egger M, Smith GD: **Bias in location and selection of studies.** *BMJ* 1998, **316:**61–66.
61. Egger M, Davey Smith G, Schneider M, Minder C: **Bias in meta-analysis detected by a simple, graphical test.** *BMJ* 1997, **315:**629–634.
62. Glatt SJ, Faraone SV, Tsuang MT: **Meta-analysis identifies an association between the dopamine D2 receptor gene and schizophrenia.** *Mol Psychiatry* 2003, **8:**911–915.
63. Jonsson EG, Sillen A, Vares M, *et al.*: **Dopamine D2 receptor gene Ser311Cys variant and schizophrenia: association study and meta-analysis.** *Am J Med Genet B Neuropsychiatr Genet* 2003, **119:**28–34.
64. Jonsson EG, Kaiser R, Brockmoller J, *et al.*: **Meta-analysis of the dopamine D3 receptor gene (DRD3) Ser9Gly variant and schizophrenia.** *Psychiatr Genet* 2004, **14:**9–12.
65. Jonsson EG, Flyckt L, Burgert E, *et al.*: **Dopamine D3 receptor gene Ser9Gly variant and schizophrenia: association study and meta-analysis.** *Psychiatr Genet* 2003, **13:**1–12.
66. Jonsson EG, Sedvall GC, Nothen MM, Cichon S: **Dopamine D4 receptor gene (DRD4) variants and schizophrenia: meta-analyses.** *Schizophr Res* 2003, **61:**111–119.
67. Abdolmaleky HM, Faraone SV, Glatt SJ, Tsuang MT: **Meta-analysis of association between the T102C polymorphism of the 5HT2a receptor gene and schizophrenia.** *Schizophr Res* 2004, **67:**53–62.
68. Lasky-Su JA, Faraone SV, Glatt SJ, Tsuang MT: **Meta-analysis of the association between two polymorphisms in the serotonin transporter gene and affective disorders.** *Am J Med Genet* 2004
69. Lotrich FE, Pollock BG: **Meta-analysis of serotonin transporter polymorphisms and affective disorders.** *Psychiatr Genet* 2004, **14:**121–129.
70. Lin PY, Tsai G: **Association between serotonin transporter gene promoter polymorphism and suicide: results of a meta-analysis.** *Biol Psychiatry* 2004, **55:**1023–1030.
71. Anguelova M, Benkelfat C, Turecki G: **A systematic review of association studies investigating genes coding for serotonin receptors and the serotonin transporter: II. Suicidal behavior.** *Mol Psychiatry* 2003, **8:**646–653.
72. Sen S, Burmeister M, Ghosh D: **Meta-analysis of the association between a serotonin transporter promoter polymorphism (5-HTTLPR) and anxiety-related personality traits.** *Am J Med Genet B Neuropsychiatr Genet* 2004, **127:**85–89.
73. Faraone SV, Doyle AE, Mick E, Biederman J: **Meta-analysis of the association between the 7-repeat allele of the dopamine D(4) receptor gene and attention deficit hyperactivity disorder.** *Am J Psychiatry* 2001, **158:**1052–1057.
74. Girault JA, Greengard P: **The neurobiology of dopamine signaling.** *Arch Neurol* 2004, **61:**641–644.
75. Itokawa M, Arinami T, Futamura N, *et al.*: **A structural polymorphism of human dopamine D2 receptor, D2(Ser311Cys).** *Biochem Biophys Res Commun* 1993, **196:**1369–1375.
76. Cravchik A, Sibley DR, Gejman PV: **Functional analysis of the human D2 dopamine receptor missense variants.** *J Biol Chem* 1996, **271:**26013–26017.
77. Lannfelt L, Sokoloff P, Martres M, *et al.*: **Amino-acid substitution in the dopamine D3 receptor as a useful polymorphism for investigating psychiatric disorders.** *Psychiatr Genet* 1992, **2:**249–256.
78. Lichter JB, Barr CL, Kennedy JL, *et al.*: **A hypervariable segment in the human dopamine receptor D4 (DRD4) gene.** *Hum Mol Genet* 1993, **2:**767–773.
79. Warren Jr JT, Peacock ML, Rodriguez LC, Fink JK: **An MspI polymorphism in the hyman serotonin receptor gene (HTR2): detection by DGGE and RFLP analysis.** *Hum Mol Genet* 1993, **2:**338.
80. Glatt SJ, Faraone SV, Tsuang MT: **Association between a functional catechol O-methyltransferase gene polymorphism and schizophrenia: meta-analysis of case-control and family-based studies.** *Am J Psychiatry* 2003, **160:**469–476.
81. Lotta T, Vidgren J, Tilgmann C, *et al.*: **Kinetics of human soluble and membrane-bound catechol O-methyltransferase: a revised mechanism and description of the thermolabile variant of the enzyme.** *Biochemistry* 1995, **34:**4202–4210.
82. Sanders AR, Rusu I, Duan J, *et al.*: **Haplotypic association spanning the 22q11.21 genes COMT and ARVCF with schizophrenia.** *Mol Psychiatry* 2004
83. Lesch KP, Balling U, Gross J, *et al.*: **Organisation of the human serotonin transporter gene.** *J Neural Trans* 1994, **95:**157–162.
84. Caspi A, Sugden K, Moffitt TE, *et al.*: **Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene.** *Science* 2003, **301:**386–389.
85. Clayton D, Chapman J, Cooper J: **Use of unphased multilocus genotype data in indirect association studies.** *Genet Epidemiol* 2004, **27:**415–428.

86. Seltman H, Roeder K, Devlin B: **Evolutionary-based association analysis using haplotype data.** *Genet Epidemiol* 2003, **25:**48–58.

87.• Neale BM, Sham PC: **The future of association studies: gene-based analysis and replication.** *Am J Hum Genet* 2004, **75:**353–362.

This paper reviews prospects for more systematic genetic association studies based on methods for combining the effects of all polymorphisms in a gene into a single analysis (see also references 127 and 128), which would permit direct comparison and meta-analysis of studies when no single functional polymorphism is responsible for an association. Although there are problems with current methods, this is an important short-term goal in the field of genetic association studies.