Published in final edited form as:

Cogn Neuropsychiatry. 2009 ; 14(4-5): 391–418. doi:10.1080/13546800903059829.

Genome-wide Strategies for Discovering Genetic Influences on Cognition and Cognitive Disorders: Methodological Considerations

Steven G. Potkin¹, Jessica A. Turner¹, Guia Guffanti^{1,2}, Anita Lakatos¹, Federica Torri², David B. Keator¹, and Fabio Macciardi^{1,2}

¹Dept of Psychiatry and Human Behavior, University California of Irvine, 5251 California Ave, Irvine, CA

²Dept of Science & Biomedical Technologies - Fondazione Filarete and School of Medicine, Università degli Studi di Milano, viale Ortles 22/4 - 20100 Milan – Italy

Abstract

Introduction—Genes play a well-documented role in determining normal cognitive function. This paper focuses on reviewing strategies for the identification of common genetic variation in genes that modulate normal and abnormal cognition with a genome-wide association scan (GWAS). GWASs make it possible to survey the entire genome to discover important but unanticipated genetic influences.

Methods—The use of a quantitative phenotype in combination with a GWAS provides many advantages over a case-control design, both in power and in physiological understanding of the underlying cognitive processes. We review the major features of this approach, and show how, using a General Linear Model method, the contribution of each Single Nucleotide Polymorphism (SNP) to the phenotype is determined, and adjustments then made for multiple tests. An example of the strategy is presented, in which fMRI measures of cortical inefficiency while performing a working memory task is used as the quantitative phenotype. We estimate power under different effect sizes (10 to 30%) and variations in allelic frequency for a quantitative trait (10 to 20%), and compare them to a case-control design with an Odds Ratio (OR) of 1.5, showing how a QT approach is superior to a traditional case-control. In the presented example, this method identifies putative susceptibility genes for schizophrenia which affect prefrontal efficiency and have functions related to cell migration, forebrain development and stress response,

Conclusion—The use of Quantitative Traits (QT) as phenotypes provide increased statistical power over categorical association approaches and when combined with a GWAS creates a strategy for identification of unanticipated genes that modulate cognitive processes and cognitive disorders.

Keywords

cognition; imaging phenotype; endophenotype; quantitative trait; GWAS; power; permutation; replication

Corresponding Authors: Steven G. Potkin, Dept of Psychiatry and Human Behavior, 5251 California Ave, Suite 240, Irvine, CA 92617, Telephone: 949-824-8040, Fax: 949-824-3324, sgpotkin@uci.edu. Fabio Macciardi, Dept of Science & Biomedical Technologies, Università degli Studi di Milano, via F.lli Cervi, 93 - 20090 Segrate (MI), Italy, Telephone: +39 02.56660 212, fabio.macciardi@unimi.it.

Introduction

Genes play a well-documented role in determining normal cognitive abilities as well as the development of cognitive disorders. This paper focuses on strategies for the identification of normal genetic variation in genes that modulate normal and abnormal cognition by combining brain images acquired during cognitive performance with genetic data. This approach is part of an emerging field referred to as Imaging Genetics. There are several excellent comprehensive reviews on Genome-Wide Association Study (GWAS) design methodology (Cardon & Bell, 2001; Hirschhorn & Daly, 2005; McCarthy et al., 2008; D. C. Thomas, Haile, & Duggan, 2005; Zondervan & Cardon, 2004) and corresponding statistical issues (Balding, 2006; Ziegler, Konig, & Thompson, 2008), as well as stratification (Cardon & Palmer, 2003; Freedman et al., 2004; Price et al., 2006; Serre et al., 2008), replication (Chanock et al., 2007) and epistasis (Keedwell et al., 2008 {Phillips, 2008 #6532}) that this paper refers to and builds upon. The reader is referred to these publications for greater depth on these subjects.

Imaging genetics leverages the strengths of both neuroimaging and genetic studies, visualizing the brain activation patterns in the context of genetic background. Imaging as an intermediate phenotype can clarify the functional link between genes and cognition. While imaging studies reveal many functional (and structural) aspects of function and dysfunction in cognition, their explanatory power may be limited by not considering the genetic bases of cognition, as many aspects of cognition are clearly heritable (Kennedy, Farrer, Andreasen, Mayeux, & St George-Hyslop, 2003). Cognitive performance may be more distant from the biological mechanisms underlying cognition than are brain images of functional circuitry (Basile, Masellis, Potkin, & Kennedy, 2002; Potkin, Anand, Alphs, & Fleming, 2003; Potkin, Kennedy, & Basile, 2002a, 2002b) as cognitive performance is a complex expression of function of several brain regions working as coordinated brain circuits. Given the known importance of both genetics and environment in brain function, and the role of neuroimaging in revealing brain dysfunction, the synergy of integrating genetics with brain imaging carries clear advantages:

- the function of the genes of interest can be revealed in neuroimaging data, an intermediate phenotype which has relevance to normal cognition, as well as to many diseases in which cognition is abnormal (e.g. (Meyer-Lindenberg & Weinberger, 2006);
- the unwanted variation in imaging studies can be reduced through controlling for genetic background;
- neuroimaging may identify dimensions of cognitive functions that are more closely related to susceptibility genes than are current clinical cognitive standardized assessments (e.g. (Gottesman & Gould, 2003);
- gene discovery - identification of new “candidate” genes related to normal and abnormal cognition that would not be discovered by traditional candidate gene approaches (e.g. (Papassotiropoulos et al., 2006; S. G. Potkin, Turner, Fallon et al., 2009; S. G. Potkin, Turner, Guffanti et al., 2009).

The availability of high-throughput genotyping technologies and genomic resources such as HapMap (www.hapmap.org) has made it possible to survey the entire genome and increase the probability of discovering important unanticipated genetic influences. In this paper, we argue that the use of Quantitative Traits (QT) as phenotypes provide increased statistical power over categorical association approaches and when combined with a Genome-Wide Association Study (GWAS) creates a strategy for identification of genes that modulate cognitive processes.

The outline of this paper is as follows: first we review the motivation for performing a GWAS, followed by a short review of what constitutes a GWAS and the major statistical issues involved in its analysis. Then we discuss the use of cognitive measures as phenotypes in GWAS, the increase in power involved with such use, and present our General Linear Model (GLM) approach, with a fuller discussion of the statistical considerations and the resulting interpretation that follows from this approach, using an imaging genetics example from previously published work.

Motivation for performing a GWAS

Much has been learned from candidate gene approaches, in which *a priori* hypotheses are tested. These hypotheses may be based on linkage studies that have identified a key region of the genome as linked to a cognitive function, or based on current physiological or pathophysiological knowledge. This approach is very attractive as it takes current knowledge, and extends it. However, a limitation of this approach is precisely that it starts with our current understanding of physiology or pathophysiology, which we know is incomplete.

For example, both acetylcholine and dopamine are involved in cognitive function (Gray & Roth, 2007; Robbins & Roberts, 2007; Tamminga, 2006), and investigation of the synthesis/release/re-uptake and metabolism of these neurotransmitters leads to a number of obvious candidate genes targets. What is less clear is whether and how these monoamines are modulated or regulated by various transcription factors, or other regulatory elements or growth factors, such as the Brain Derived Neurotrophic Factor (BDNF). BDNF could then be considered as a candidate gene. Such regulatory influences can be determined to some extent by investigating *in silico* pathway and epigenetic databases still there are many important interactions and functions not yet identified. Additionally, this process typically leads to one study per candidate gene, and with thousands of genes expressed in the brain to consider, is inefficient. While such interactions are continuing to be discovered; we do not yet know how to place the dopaminergic system in context within the whole brain or genetic milieu.

It is also important to keep in mind that most of the current cognitive candidate genes have only become candidates through *post hoc* genetic annotation and selection (Abdolmaleky, Faraone, Glatt, & Tsuang, 2004; Egan et al., 2001; Lohmueller, Pearce, Pike, Lander, & Hirschhorn, 2003; Shifman et al., 2002). Plausible hypotheses build upon current knowledge will extend that knowledge (Meyer-Lindenberg et al., 2006; Roffman, Weiss, Goff, Rauch, & Weinberger, 2006), but are unlikely to result in a fundamental shift in the conceptualization of cognition.

An alternative approach is to screen the entire genome (GWAS) for genes that influence these cognitive processes or disorders, without being limited to potentially less important chromosomal regions by our current state of incomplete knowledge. This strategy applied to Imaging Genetics can identify new, unanticipated candidate genes for cognition. These insights have the potential to fundamentally alter our current understanding of cognitive processes, since this application of Imaging Genetics can visualize the brain activation patterns in the context of genetic background, thereby synergizing the strengths of each individual approach (S. G. Potkin, Turner, Fallon et al., 2009; S. G. Potkin, Turner, Guffanti et al., 2009) and ultimately representing a strategy for gene discovery.

What constitutes a genome wide association study

A Genome-Wide Association Study (GWAS) surveys most of the genome for potentially causally-related Single Nucleotide Polymorphisms (SNPs). The selection criterion of the

markers to include in a GWAS is based on the hypothesis and methods of Linkage Disequilibrium (LD) (Lewontin, 1958) between genetic marker alleles and the true causative alleles (Durrant et al., 2004). The two most common statistical measures used to summarize LD between two markers are D' and r^2 . Both measures are built on the difference between the probabilities of observing two markers (e.g. SNPs) alleles on the same haplotype compared to observing them independently in the population. The D' statistics is equal to 1 unless recombination across the two SNPs has occurred, therefore the higher the D' the greater the LD; however, the allele frequency of the SNPs is not considered. The r^2 measure directly depends on allele frequency; a r^2 value of 1 can be achieved only when the marker loci have identical allele frequencies and every occurrence of each allele marker perfectly predicts the allele at the other locus (i.e. an identical proxy). An r^2 of 1 thus indicates two SNPs are in perfect LD so that only one SNP needs to be genotyped to know the genotype of the other. This "tag" SNP approach can provide a high level of genomic coverage as a large number of untyped SNPs will be identifiable with this reference set of tagging-SNPs genotyped. To generate a comprehensive LD map of the human genome, the HapMap Project has already genotyped the most common DNA variants and is now completing the identification of less frequent variants (Kruglyak, 2008). The general consensus is that an r^2 of 0.8 or greater is sufficient for a tag SNP to obtain a good coverage of LD-dependent untyped SNPs. Thus tagging SNPs allow genotyping of far fewer marker SNPs with relatively small losses in power (Anand et al., 2005). The tagging SNP method is effective even with alleles frequency $< 5\%$ and therefore even rare variants can be detectable (Fallin et al., 2001).

The completion of the human genome sequence and advances in SNP genotyping have made this kind of coverage both feasible and affordable (Botstein & Risch, 2003; Hirschhorn & Daly, 2005; Stranger et al., 2005). Current microarray development, like for example the Illumina Human1M-Duo BeadChip (Di Piero, Ferracuti, Sabatini, & Pantano, 1994) currently covers 96% (in CEU) of the genome with 1 million SNPs at average LD-related distance corresponding to an r^2 of 0.8. The costs of GWAS have also decreased to the point that they are being almost routinely conducted for many disorders even in medium size laboratories.

Statistical challenges in GWAS

GWAS techniques produce anywhere from 100,000 to more than 1,000,000 genotypes, e.g. data points, per subject. Classical statistical analytical techniques are not designed for situations where the number of variables so grossly outnumbers the number of subjects, and such analyses have rarely been attempted in biological studies. The classical approach of considering each association between a SNP and the phenotype as a "repeated" test, and thus correcting for the number of tests performed (under the assumption that they are multiple tests), is out of its depth when faced with a million or more tests. The potential for detecting genetic influences on phenotypes is lost if the sample size required for study by conventional measures literally verges on a significant proportion of the world's population. The technical developments in genotyping have apparently outstripped the ability to analyze the resulting information, if we keep using statistical methods and strategies that have been developed for much simpler tasks. In fact, the traditional methods to correct for multiple tests, for example the well-known Bonferroni's method is not ideal in this application, since it assumes that all the factors, the SNPs in our case, are independent from each other. However, this is not the reality of the genome, where SNPs are somewhat correlated as function of their relative distance independently from their status as "taggers" or not. While new methods have been developed (Benjamini & Hochberg, 1995) to address this problem, a more compelling theoretical question is whether we are dealing with a problem of multiple testing or rather with an issue of multiple hypotheses. In fact, if we consider each single

SNP a possible "cause" of the disease that we are investigating, then testing one million SNPs reverts to testing one million different hypotheses, each one of them with a high degree of uncertainty due to the overall complexity of the genome in addition to the number of hypotheses tested. As one attempts to deal with the "one million hypotheses" question, the proposed solutions span a wide range of alternatives: at one extreme, theoretical epidemiologists (Rothman, 1990) state that there is no need to correct and the investigator should simply report any significant finding together with the number of hypotheses tested, leaving the burden to a subsequent focused experiment to prove or disprove the previous reported positive findings. With a more conservative view, others focus on the risk of too many potential false positives and attempt to balance and trade-off between the complementary risks of false positives and false negatives. They propose to minimize the risk of false positives with techniques such as permutations (Westfall, 1993), FDR (Benjamini & Hochberg, 1995) and others. This latter approach is more or less conceptually based on *a priori* establishing a threshold for a "genome-wide" significant finding, almost in the way R.A. Fisher originally set up the 5% significant threshold for classical statistical tests (Fisher, 1935). It is important to note that false negatives can be potentially more harmful than false positives, since may be discarded forever, while a false positive can always be disproved or "falsifiable" albeit at some cost. Any established threshold for genome-wide significance is a subjective decision that cannot be universally proven or disproven, regardless of the arguments in support of a given threshold.

In addition to the problem of multiple testing / multiple hypotheses, there are several other important issues that are the current focus of interest in statistical genetics. Other topics that are beyond the scope of the present paper, include research related to how to statistically identify "genes" (or chromosomal regions) rather than SNPs, given that our primary interest is in mapping putative functional elements of the genome rather than simple DNA point variations. Addressing gene*gene interactions (Brzustowicz, 2008; Moore, 2008), hence epistasis (Chapman & Clayton, 2007; Evans, Marchini, Morris, & Cardon, 2006; Jiang, Tang, Wu, & Fu, 2009) for broad, or gene*environment (Clayton & McKeigue, 2001; Glazier, Nadeau, & Aitman, 2002; Hoffmann, Lange, Vansteelandt, & Laird, 2009; Lander & Kruglyak, 1995) interactions or even how to validate a causative or regulatory network (Barabasi, 2007; Hidalgo, Blumm, Barabasi, & Christakis, 2009) are areas of active statistical research.

It is in fact becoming increasingly evident that multivariate analysis methods can be applied to extract meaningful results from a GWAS, including the complex interdependence that exists across the SNPs within the genome (see e.g., {Meng, 2009 #6690}). These multivariate approaches not only address a gap in our current analytical strategies, but can reach new and sometimes unexpected information in the way our genome is organized and how it can be deranged in complex disease (Slawski, Daumer, & Boulesteix, 2008). The application of these methods such as Random Forest, Support Vector Machines, etc. is finding increased popularity as classification techniques in gene expression datasets for identifying risk genes, or to case-control GWS analyses (for an introduction to their applications in gene expression analyses see, {Boulesteix, 2008 #6687}). While application of some of these methods to quantitative regressions are possible, they have been applied to quantitative traits with less frequency (though see {Calhoun, Liu, & Adali, 2009}) for ICA analyses of imaging and genetics data, and (S. S. Lee, Sun, Kustra, & Bull, 2008) for an example of applying Random Forest techniques with the quantitative trait of blood pressure in a linkage study). Their application to quantitative traits such as cognitive phenotypes in GWS is still very much in development. Some of the issues that are challenges for the univariate analyses do not apply in the multivariate analyses, e.g repeated testing of hypotheses, but the larger issues of interpretation and confirmation of the results apply in either analytical approach.

We focus in this paper on the application of massively univariate statistical techniques in imaging genetics analyses.

Cognition as a phenotype

GWAS have also found genetic influences on various cognitive measures. One of the most studied traits is human memory, with a heritability estimate of approximately 50% for various measures (McClearn et al., 1997). Performance on a memory task has been used successfully as a quantitative trait (QT) in a GWAS study of young healthy adults, leading to the identification of KIBRA as a memory-related gene (Papassotiropoulos et al., 2006). A common variant of the KIBRA gene, identified by the SNP rs17070145, was found to be significantly associated with memory performances, based on quartile ranking in verbal episodic memory, in 3 independent samples. The genome significance threshold for the first GWAS was established based on a strategy that combined two different statistical approaches: a single-point method using Bonferroni's correction for independent comparisons and a sliding-window approach to identify chromosomal clusters harboring the most significant SNPs at various window sizes. The positive findings selected for high "statistical confidence" were further genotyped in independent samples (Papassotiropoulos et al., 2006). Subsequently non-carriers of the T allele (rs17070145) in KIBRA are associated with an increased risk of developing late-onset Alzheimer's disease (AD) (Corneveaux et al., 2008). KIBRA was over-expressed and three of its four known binding partners under-expressed in hippocampal, posterior cingulate, and temporal cortex regions of AD subjects studied post mortem. Interestingly, cognitively normal, late-middle-aged persons who were non-carriers of the T allele exhibited lower glucose metabolism on PET scanning in posterior cingulate and precuneus, brain regions sensitive to detecting AD, than did carriers. These findings suggest that KIBRA is associated with individual variation in episodic memory in healthy normals as well as in the predisposition to AD. It is noteworthy that the case-control study required far more subjects to achieve the same statistical power as that found in the much smaller study using cognition as a quantitative trait (QT).

Almasy and her colleagues (Almasy et al., 2008a) used a quantitative trait design to examine multiple cognitive domains in families with schizophrenia, using 386 autosomal microsatellite markers (selected from the Genethon human linkage map) in a genome-wide linkage study. In this study they identified evidence for genetic loci on chromosome 19 and chromosome 5q.31–34 associated with cognitive performance, as well as with risk for schizophrenia. The Papassotiropoulos and Almasy studies indicate the value of employing Quantitative Traits, where unanticipated genes are identified and become candidates for further molecular and functional exploration. Their work expands the potential mechanisms and circuitry underlying cognitive function and provides new scientific models of brain function, as well as new targets for therapeutic interventions.

Power of a Quantitative Trait in a GWAS

The power of an association study depends on the specific design of the study, such as whether replication or joint analyses are to be accommodated. Several methods have been developed to minimize the financial costs of genotyping, such as a two-stage design (Ennis et al., 2008; Li, 2007; Skol, Scott, Abecasis, & Boehnke, 2007). These designs recognize the need to balance the risk of false positives with the greater need to minimize false negative results. False negative results may be more of a problem, as false positive results will be identified in joint analysis studies or properly powered replication studies. In any event, an appropriate correction for multiple comparisons / multiple hypotheses must be used, but determining the appropriate correction is challenging for GWAS studies as we and others have discussed (Balding, 2006; D. C. Thomas, 2006; Ziegler et al., 2008).

Our approach to power analysis is different from that assumed in standard association studies since we will focus on quantitative phenotypes as contrasted with case controls. We estimate power under different assumptions regarding the effect size and the variation in allelic frequency, using the approach proposed by Purcell et al. (Purcell, Cherny, & Sham, 2003) for a quantitative complex trait. The results are presented in Figure 1 and Figure 2. For comparison, power curves using a categorical, case-control design are also shown with an odds ratio (OR) of 1.5. The usual cautions apply here: for example, the putative allelic frequency of the “cognitive or disease” allele should not be too different from the “marker” allele frequency, with the risk of missing the association (Durrant et al., 2004). With the following conservative assumptions, an effect size of 10%, a low minor allele frequency of 20% in figure 1 and 10% in figure 2, a sample size of 800 and 500, respectively, will provide 80% power for a QT phenotype to be detected even when the estimated p-value is very conservative ($p \leq 10^{-7}$), for both a multiplicative and an additive model. In contrast, a much larger sample of many thousands is required to obtain 80% power in a case-control design, given an OR of 1.5 or even greater than that.

Table 1 shows, for a constant 80% power, how the needed sample size to detect a related locus with a QT phenotype depends on the amount of variance explained in the QT and in the SNP allele frequencies, over a 10 to 30% range. A range of minor allele frequencies (MAFs) for the tagging SNPs is also indicated, assuming a stable and constant LD (measured as D') between the tagging SNP and the causative SNP of .95. As the variance of the QT, explained by a given SNP, increases, the sample size dramatically decreases for all frequencies of the tagging SNPs. The curve distributions in Figure 1 and Figure 2 again indicate that the closer the match between the allelic frequency of the QT locus and the tagging SNP, the greater the power and the smaller the needed sample size, while the relationship is very different in the case-control approaches. It is well-known that our ability to detect a “causative” SNP in a case-control design is also a function of how close the allelic frequencies are for the unobserved SNP and the marker or tagging-SNP even when the allelic frequency for the causative SNP is as high as 30%. If we cannot match the causative SNP frequency with the “tagging” SNPs that we are experimentally testing, we risk missing an important association unless we use an unrealistically large sample size. Figure 3 shows how the power can vary as a function of the mismatch between the two allelic frequencies of the causative and tagging SNPs, even under the assumption of high LD between them.

Statistical Model for a QT phenotype

All SNPs that pass quality control checks (Teo, 2008) are included in the GWAS analysis. The simplified statistical model compares the differential effects of SNP alleles or genotypes on the quantitative trait.

The simplest General Linear Model (GLM) is:

$$Y = \mu + \beta_1 * SNP + \beta_i * V_i + \varepsilon \quad (1)$$

in which Y is the quantitative trait, μ is the population mean, β_i are the coefficients being estimated, SNP represents the marker being tested, V_i are other variables in the model (e.g. the age of the subject, diagnostic group, gender, etc.), and ε is the error term. The SNP term represents the main effect of the “gene” on the QT, independent of any class grouping or other variable.

The beauty of the GLM approach lies in its extensibility. In the study of a cognitive QT in the presence of a cognitive disorder (e.g. looking for genetic effects modulating both the

development of autism and specific cognitive dysfunction within the disease), the GLM can be extended to:

$$Y = \mu + \beta_1 * SNP + \beta_2 * diagnosis + \beta_3 * SNP * diagnosis + \beta_i * V_i + \varepsilon \quad (2)$$

In which the $\beta_2 * diagnosis$ term indicates how the cognitive phenotype Y differs across the diagnostic variable. The SNP term represents the overall effect of genotype on the QT; the $SNP * diagnosis$ interaction term in this model is an integrative term reflecting how the genotype differentially affects cognitive phenotype in the disorder. This approach can determine the effects of genes on cognition as measured by brain activation or behaviorally measured performance, and further, it can determine if these genetic effects differ by diagnosis.

Depending on the study design, of course, the GLM can also be extended to explicitly model effects of known, candidate genes, and/or gene-gene interactions. In the example below, Y could be a quantitative measure of attention in a study of Alzheimer's disease (AD) and healthy subjects with different APOE alleles, and the research question is looking for genetic effects on cognitive function that interact with the APOE genotype:

$$Y = \mu + \beta_1 * SNP + \beta_2 * diagnosis + \beta_3 * SNP * diagnosis + \beta_4 * APOE + \beta_5 * APOE * SNP + \beta_i * V_i + \varepsilon \quad (3)$$

How complicated the analysis model can be will depend, of course, on other design considerations such as expected effect sizes and sample sizes, etc.

Covariates in the model

When there is an established strong genetic influence on the phenotype, it is important to eliminate or control for this known genetic effect to enhance the strategy to identify unanticipated genes. For example and using the example before, APOE4 is a well documented risk factor for AD, and thus should be controlled or covaried for, as the effects of APOE4 can otherwise obscure the identification of other risk genes. Other relevant variables can also be included in the model. In the case of AD, we can include age and gender as covariates in the regression model as they are associated with the risk of developing AD.

Genetic Models

The genetics underlying complex traits is characterized by a sophisticated interplay between multiple genes. When applied to common diseases, genome wide association studies must take into account the degree of etiopathological complexity using the most appropriate model. Classical statistical genetics offers four possible models to code the SNP term in the GLM above: *additive*, *codominant*, *dominant* and *recessive*. The *additive* model is generally preferred as it reflects the additive contribution to risks for complex diseases, and additive models also can detect strong non-additive effects (Gianola & de los Campos, 2008; Hill, Goddard, & Visscher, 2008; S. H. Lee, van der Werf, Hayes, Goddard, & Visscher, 2008; Yamada & Okada, 2009). When the hypothesized risk allele (B) is rare, the dominant model can be used as it pools the homozygotes (BB) and heterozygotes (AB) genotypes together in the analysis. Such a model tests the hypothesis that carrying even one copy of that particular allele increases the risk of disease.

This approach may not be exhaustive when applied to complex traits, as these models were developed to study single major loci. For complex traits, functional pathways should be

considered because multiple genes are involved and interact. Since additive, dominant, codominant and recessive models refer to the single SNP as the unit of the analysis new statistical approaches that model interaction effects, either gene-environment (e.g. gene x diagnosis) and gene-gene (i.e. epistasis) are needed. There is relatively little work on how to model these interaction relationships, and it is an area of active research (Moore, 2003; Phillips, 2008; Wang, Comaniciu, & Fasulo, 2006).

How to calculate a GWAS threshold

A major challenge to the interpretation of results from a GWAS is setting the appropriate statistical threshold. In classical statistics, as the number of tests against a single null hypothesis increases, the statistical threshold (p-value) has to account for the probability of a false positive occurring by chance. This is typically done using a Family-Wise Error Rate (FWER) approach such as a Bonferroni correction; if a GWAS is going to test for 1 million SNPs (and not considering here the complexities potentially related to testing 1 million hypotheses that we have previously discussed), the set significance level is 10^{-8} for a global significance level of 5%. However, the Bonferroni's correction is not appropriate for observational studies such as GWAS (Perneger, 1998) because it does not account for the dependencies of SNPs that are close to each other across chromosomes, thus leading to an overcorrection. Arking et al. proposed less stringent significance levels that accounted for LD across SNPs (Arking et al., 2006). Zondervan and Cardon (2007) provide a method to adjust for the actual number of independent tests (Zondervan & Cardon, 2007). The HapMap consortium proposed a local significance threshold of 5.5×10^{-8} based on re-sampling from empirical data under the null hypothesis (Altshuler et al., 2005), reaching a similar result to the WTCCC consortium (5×10^{-7}) (Dudbridge & Gusnanto, 2008; WTCCC, 2007).

Various combinations of a classical method with other methods have been suggested, including multi-stage experimental designs, ranking the results and identifying a top subset of likely genes, and permutation testing to determine an empirical p-value. Also, expanding the work by Benjamini and Hochberg (Benjamini & Hochberg, 1995), Efron and Tibshirani (Abi-Dargham et al., 2002) proposed controlling for the risk of false positives over all the positive results rather than over all the possible tests (the False Discovery Rate approach – FDR), thus partially compensating for the overcorrection of the more traditional FWER methods.

It is important to re-emphasize that all these proposed methods address the evaluation of a single SNP at a time. If two or more SNPs are in (strong) LD to each other and show a similar pattern of significant association with the trait of interest, it is appropriate to use a less stringent threshold (Ziegler et al., 2008). Alternatively one can avoid putting forth a prespecified formal significance threshold, and present all results ordered from lowest to highest p-values (Almasy et al., 2008b; Helgadottir et al., 2007).

In GWASs, a false positive results in an additional cost of following up the initial finding, always a necessity considering the exploratory nature of GWAS. The consequences of false positives in a GWAS are less than finding a false positive in a pivotal clinical trial that could lead to potentially dangerous treatment (Rothman, 1998). Importantly, any correction for multiple testing to avoid increasing false positives also affects the false negative rate, reducing power to detect a true significant finding (Samani et al., 2007). The latter may be a more serious error, in that an important finding is prematurely dismissed.

Several other approaches are being developed to determine which GWAS significance threshold is most appropriate to a specific research question. It is important to keep in mind that in a GWAS each SNP is its own hypothesis; a GWAS involves testing of hundreds of

thousands of hypotheses. This is a fundamentally different question than testing the same hypothesis a million times. Bonferroni corrections are more suited to the latter case but are not well suited to the testing of many different hypotheses (WTCCC, 2007). Perneger (1998) made the case that Bonferroni adjustments are concerned with the wrong hypothesis (i.e. that all null hypotheses are true simultaneously which is not of interest) and increases the likelihood of type II errors concluding that the Bonferroni method may create more problems than it solves. Appropriate statistical correction for multiple statistical tests is an area of intense statistical research (Dudoit S, 2008). Despite a lack of consensus various practical approaches are in use to address the problem of multiple testing and determining the appropriate GWAS threshold, several of which are briefly summarized.

Permutation tests

Permutation methods (e.g.(Manly, 1997)) offer the possibility of using the dataset collected to empirically determine the statistical threshold, both for case-control and QT phenotypes. Although computationally expensive, current technology is robust enough to perform permutation testing.

While permutation methods are a broad class of techniques, in GWAS applications they are used to determine the proportion of cases in which the F or chi-square statistics would arise under the null hypothesis of no genetic effect. With a permutation approach, the “labels” identifying cases and controls (or a given measure for a quantitative phenotype) are randomly reassigned, the original analysis re-done and the subsequent statistics is noted; when this is done many thousands or hundreds of thousands of times, the distribution of the test statistic under the null hypothesis for that specific dataset is known. This data derived distribution can deviate significantly from the *a priori* distribution. The probability of the original statistics in the current sample arising by chance, then, can be empirically estimated.

In a dichotomous case-control study, when testing the main genetic effect on case-control status, usually the case-control designation is permuted across subjects; these methods are now standard with GWAS software (e.g., PLINK release v1.03 (Purcell et al., 2007)). However, for quantitative trait interactions it is not clear if the QT measure (*Y*) should be permuted, and what level of randomization (full, reduced, or constrained) is required. If the significance of an interaction term between SNP and diagnosis is being assessed, for example, the method may permute residuals for that term rather than the original raw data. Depending on the precise factorial design, full permutation may not be used, but rather permutation within exchangeable units within the design (for more detail on these issues, see for example (Manly, 1997); (Anderson & Ter Braak, 2002; Jung, Jhun, & Song, 2006).

With the increasing interest about the identification of improved multiple testing procedures for GWAS, we suggest a possible strategy to modify the existing non-parametric inferential procedures in order to fit the interaction event statistical assumptions.. For illustrative purposes we propose to randomly select a sample of SNPs (e.g. say 1% or 5,000 of the 500,000 SNPs in the GWAS) and perform the permutation tests 1,000 times for each of the 5,000 SNPs creating a set of 5 million interaction p values, representing a null distribution for a random sample of SNPs in the GWAS. From this null distribution, the top 5% or 1% empirical threshold p-value can be set as an empirically determined, genome-wide threshold. The number of SNPs and permutations required to achieve the desire power will depend on sample size. This combines the strength of the permutation testing method with the GWAS data to determine an underlying null distribution for the full GWAS (see Weinberger presentation to Psych Gen NYC 2008).

Additional SNPs in the same region

Combinations of SNPs can also be considered, through various “haplotype-based” approaches (Seasholtz et al., 2006), though that has not been applied to a GWAS with a QT. The greater the number of significant tagging SNPs in proximity to one another, the smaller is the likelihood of a finding being a false positive. As chromosomal regions rather than individual SNPs represent the inherited units into which an inherited genome can be partitioned, finding several significant tagging SNPs in physical proximity provide additional support for the locus being a true risk factor. The focus on tagging SNPs is critical, since we already know that adjacent SNPs are more likely to be in greater linkage disequilibrium and hence finding many close SNPs significant may not be particularly relevant since they share the same meaning. Haplotype approaches may also be able to capture effects in SNPs which are rare, in the 1–5% range (de Bakker et al., 2005; Kamatani et al., 2004; Lin, Chakravarti, & Cutler, 2004). The haplotype approach known as the “sliding window” (Fallin & Schork, 2000) can be used to detect the effects of rare alleles, using the extended block method of Purcell (Purcell et al., 2007). A complete catalogue of all the low frequency SNPs and CNVs discovered will improve pinpointing genes, structural variants in chromosomes and other individual genomic variations that are associated with disease.

Results of an imaging genetics analysis of cognitive function

In our own studies, we have used functional neuroimaging as a quantitative phenotype in the context of GWAS to identify genes related to cognition in schizophrenia, and genes responsible for cognitive differences between subjects with schizophrenia and controls. We measured Blood Oxygenation Level Dependent (BOLD) signal in the dorsal lateral prefrontal cortex (DLPFC) during a the Sternberg Item Recognition Paradigm (SIRP), a working memory task, in a group of schizophrenia and healthy control subjects (see Potkin et al., 2009a and 2009b). In a similar fMRI study of schizophrenia and healthy control twins, DLPFC activation during the SIRP as well as the behavioral measures was found to be heritable (Karlsgodt et al., 2007).

DLPFC has a well documented role in memory performance in normal cognition as well as being a locus of dysfunction in schizophrenia; differences in the DLPFC between schizophrenic and control subjects have been described in local gene expression (Dean, Keriakous, Scarr, & Thomas, 2007; Vawter et al., 2004), cell morphometry, structural circuitry (Lewis & Gonzalez-Burgos, 2008; Shenton, Dickey, Frumin, & McCarley, 2001), and both local and distributed functional activation (Manoach et al., 1999; Menon, Anagnoson, Mathalon, Glover, & Pfefferbaum, 2001; Weinberger, Berman, & Zec, 1986).

In these data, schizophrenics show more (BOLD) activation in the DLPFC than do healthy controls when matched for accuracy performance on the Sternberg Item Recognition Paradigm (SIRP), a working memory task, consistent with cortical inefficiency (S. G. Potkin, Turner, Brown et al., 2009). Subjects performed the Sternberg Item Recognition Paradigm (SIRP) with memory loads of 1, 3, or 5 items. During the “encode” condition, subjects memorized a set of target digits (memory load 1, 3 and 5 digits). They were then presented with probes (single digits) and responded by indicating whether the probe was a target (a member of the memorized set). The associated fMRI BOLD measure from the DLPFC while performing the probe condition while holding 3 items in memory was used as a quantitative phenotype in a GWAS.

Subjects were genotyped using the Illumina HumanHap370-Duo, providing 370404 SNPs suitable for analysis with the fMRI QT. Using the GLM analyses presented and statistical thresholds described above, two genes with functions related to cell migration to forebrain

structures were identified, that had not been previously associated with cognition or schizophrenia (S. G. Potkin, Turner, Fallon et al., 2009). Six additional genes (or chromosomal regions) related to forebrain development and stress response, affecting prefrontal efficiency were also identified (S. G. Potkin, Turner, Guffanti et al., 2009). These studies highlight the ability of brain imaging as a QT to identify unanticipated genes related to cognition.

Interpretation of the results of GWAS

The complexity of human psychiatric diseases and the questionable accuracy of the psychiatric diagnoses require a partitioning process of the potential underlying pathophysiological mechanisms (Meyer-Lindenberg & Weinberger, 2006). Imaging offers a tool that reflects underlying quantitative traits that may be more directly linked to the neurobiology of normal and abnormal function, serving as a proxy for the underlying causal mechanism. Therefore imaging phenotypes have great potential for gene discovery by increasing the power of a study and decreasing phenotypic heterogeneity (Walters & Owen, 2007). However, in choosing the imaging phenotype, choice of the task is an important consideration. A complex cognitive task, which may reflect important cognitive differences between groups, may be so complex that it is difficult to deconstruct into the many components and processes involved. On the other hand, a simple cognitive task, which is easy to interpret, may be so simple or narrow as to no longer reflect the essence or complexity of the disorder.

In addition, the combination of the quantitative trait with a whole genome wide scan can discover unanticipated genes previously not associated with the trait or the disorder. The discovery of novel associations between genes and risk for cognitive or neuropsychiatric disorder offers a powerful impetus to postulate new biological mechanisms. Although this approach is focused on discovery of novel genetic associations with cognitive function, the biological meaning of the associations emerging from the analyses must be considered.

The strategy leverages the collective biological insights of ancillary studies, published results from other research programs, databases, and genetic annotation. The consistency of findings across the different studies, and the correspondence to any *in vitro* functional experimental results based on the findings will determine how compelling the results are.

The annotation and comparative analysis of genes of interest can be a challenging and time consuming task. Fortunately, numerous databases and tools are available for assessing the evolutionary, structural and functional significance of the genes. The genes' and their products' structure, their structural changes and evolutionary significance can be assessed by a homology-based approach or simple information searching by taking advantage of the publically available genomic and protein databases (e.g., NCBI Genomic, NCBI conserved domains, NCBI GenBank Proteins, UniProtKB-Swiss-Prot, UniprotKB-PIR, and UniProtKB-TrEMBL, Pfam, Map Viewer). The genes can be functionally classified by using the Gene Ontologies (GO) and establishing pathway associations by using, for example, the Kyoto Encyclopedia of Genes and Genomes (KEGG). The function of the genes can be analyzed not only by an individual component but also by systematic approaches in the context of complex cellular processes (Hecker, Lambeck, Toepfer, van Someren, & Guthke, 2009). *In silico* pathway and network analyses Ingenuity (<http://www.ingenuity.com>), Ariadne (<http://www.ariadnegenomics.com/>), <http://Metacorehttp://www.genego.com/metacore.php>, PANTHER (P. D. Thomas et al., 2003), The Database for Annotation, Visualization and Integrated Discovery (DAVID) (Dennis et al., 2003) - can illustrate novel potential underlying neurobiological mechanisms and delineate neuronal circuitries that may translate genetic effects into behavior.

Due to the complexity of human brain function, it has been very difficult to generate animal models that mimic human cognition and cognitive disorders. The hypothesized role of novel genes in the underlying physiology of cognition or diseases involving cognition can be further tested with in vitro and in vivo experiments such as establishing aspects of gene function in cell culture experiments or transgenic animals. One of the examples of this top-down approach is the generation of a dominant-negative DISC1 transgenic mice, as DISC1 was found first in human studies in relation to schizophrenia, then targeted in the mouse, with measurable translatability as a model of schizophrenia (Hikida et al., 2007). This approach evidently can facilitate animal models relevant to cognition and potentially the development of novel therapeutics and gene-based therapies. Other approaches can consider the gene regulators network that the gene participates in rather than the gene itself. These networks can be identified by computational biology and then these statistical models validated by animal experiments.

Replication

The problem of failure to replicate a GWAS finding is common. Replication in the context of GWAS must be carefully considered. Based on current knowledge, the genetic bases of cognitive processes are not Mendelian; rather, multiple genetic influences can affect the same cognitive process (or an underlying mechanism such as signaling pathway). There are at least two levels at which replication of a study can be considered depending on the unit of analysis. At the most abstract level the replication can implicate the same pathway in a cognitive process or disorder. Replication at the molecular level, at a more conservative level, can include the identical sequence or SNP or genes depending on the research question. Most SNPs for GWASs are typically chosen as tagging SNPs (htSNPs), i.e. proxies for “blocks” of DNA (Balding, 2006). Tagging SNPs identify loci on a chromosomal region associated with the trait of interest. That region can harbor the “true” causal SNP which typically does not coincide with the htSNP, therefore the “true” SNP is not genotyped. Consequently, it is the block of DNA tagged by the htSNP(s) that becomes the focus for gene sequencing and subsequent studies of the causal molecular mechanisms.

The implications for replication of a GWAS are considerable. The most stringent replication requirement is finding the same allele of the original SNP in the second study. However, this requirement can be misleading if the SNP to be replicated is a “tagging” SNP and the second population is not genetically homogeneous to the initial population (a likely event). In such cases, the same SNP may be implicated but the directionality may be different, or a different SNP that is located in the same chromosomal (“block”) region may be implicated. This is more likely if the allelic frequency of the tagging SNP is not identical in the two populations. On the other hand, if the identified SNP in the first study is hypothesized to be a causal SNP, then failure to identify the same causal SNP in the second study, provided there is sufficient power (see Figure 1 and Figure 2), is a failure of replication.

As an example of the complexity of the replication, even in a “simple” association, we can consider the case of the phenotype of cystic fibrosis, which is known to be caused by an homozygous recessive genotype at the CFTR gene (cystic fibrosis transmembrane conductance regulator) (Zielenski, 2000). However, only 66% of all cases have a particular mutation, a 3-bp deletion called $\Delta F508$, and more than a thousand mutations have been described in the remaining 33%. The frequency distribution of the mutations differs by ethnic groups. In fact, failure to find the particular $\Delta F508$ mutation in cases with cystic fibrosis led to sequencing the entire gene, identifying the other mutations. The problems in replicating the non- $\Delta F508$ mutations are obvious. The complexity of replication is likely to be much greater in non-Mendelian diseases or cognitive processes with multiple genetic influences.

Population stratification (PS) is also an important consideration in attempts to replicate a GWAS finding. Failure to recognize and correct for population stratifications is a major reason for failure in replicating genetic findings. PS occurs when allele frequencies differ between cases and controls due to ancestry differences, ethnic background or even "hidden" stratification, thus leading to spurious findings between a phenotype and unlinked candidate loci with either false positive or false negative results. To overcome this problem, different correction strategies have been proposed, however finding the most appropriate method to control for PS is not an easy task (Salvi et al., Submitted). Most of the current software used to analyze GWAS includes statistical controls and procedures to correct for PS, whenever present. With the most popular method, investigators calculate the inflation score due to PS with a statistic called *lambda* that represent the deviation from the median of a central chi-square distribution (Devlin, Bacanu, & Roeder, 2004), with PS considered to be present for values of *lambda* higher than 1. A method called Genomic Control (GC) (Bacanu, Devlin, & Roeder, 2000; Devlin & Roeder, 1999) is typically used to correct association tests for the estimated inflation score. Algorithms in the various software packages differ from each other in the way *lambda* is calculated as well as how GC is applied, leading to slightly, but sometimes consistently, different results (Salvi et al., Submitted). Investigators should be aware of the complexities related to PS, especially with the current trend in GWAS studies of considering large samples that rarely derive from a single population.

Summary

Genes play a well documented role in cognition. Brain imaging obtained during performance of a cognitive task is a sensitive and quantitative reflection of cognition. Memory performance and regional brain activation during various cognitive tasks as measured in neuroimaging studies are examples of objective and quantitative measures that are usually heritable, and well suited as quantitative phenotypes (Blokland et al., 2008).

Brain imaging and genetic data can be integrated to enhance our understanding of cognition, and to discover unanticipated genetic influences on cognition. Using a summary measure of brain activation from a neuroimaging study as a cognitive phenotype in the context of a GWAS provides a strategy with increased statistical power over non-quantitative trait approaches. This QT approach can identify both known and unanticipated cognitive-related genes.

There are several caveats related to findings with a QT GWAS approach. There is no consensus on the appropriate adjustments for multiple testing or how to calculate a GWAS threshold. The identified genes are merely candidates that require confirmation in an independent sample. The SNPs that identify these genes are primarily tagging SNPs, and therefore they do not necessarily convey any special meaning rather than "flagging" that particular gene or region. Study replication can be viewed at several levels and is dependent on the composition of the sample (population stratification). The choice of the quantitative trait (QT) in a GWAS is crucial to an interpretable result. The QT should represent an objective, quantitative and reliable measure that is relevant to the cognitive process. To the extent that the chosen trait is proximal to the genetic mechanism, the expected variance would be less, and increase the likelihood of identifying meaningful genes associated with the cognitive trait. The role of these newly identified genes require further exploration by *in silico* annotation, sequencing and molecular studies.

References

- Abdolmaleky HM, Faraone SV, Glatt SJ, Tsuang MT. Meta-analysis of association between the T102C polymorphism of the 5HT2a receptor gene and schizophrenia. *Schizophrenia Research* 2004;67(1):53–62. [PubMed: 14741324]

- Abi-Dargham A, Mawlawi O, Lombardo I, Gil R, Martinez D, Huang Y, et al. Prefrontal dopamine D1 receptors and working memory in schizophrenia. *Journal of Neuroscience* 2002;22(9):3708–3719. [PubMed: 11978847]
- Almasy L, Gur RC, Haack K, Cole SA, Calkins ME, Peralta JM, et al. A genome screen for quantitative trait loci influencing schizophrenia and neurocognitive phenotypes. *American Journal of Psychiatry* 2008a;165(9):1185–1192. [PubMed: 18628350]
- Almasy L, Gur RC, Haack K, Cole SA, Calkins ME, Peralta JM, et al. A Genome Screen for Quantitative Trait Loci Influencing Schizophrenia and Neurocognitive Phenotypes. *American Journal of Psychiatry*. 2008b
- Altshuler D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P. A haplotype map of the human genome. *Nature* 2005;437(7063):1299–1320. [PubMed: 16255080]
- Anand A, Li Y, Wang Y, Wu J, Gao S, Bukhari L, et al. Antidepressant effect on connectivity of the mood-regulating circuit: an fMRI study. *Neuropsychopharmacology* 2005;30(7):1334–1344. [PubMed: 15856081]
- Anderson MT, Ter Braak C. Permutation test for multi-factorial analysis of variance. *Journal of Statistical Computation and Simulation* 2002;73(2):85–113.
- Arking DE, Pfeufer A, Post W, Kao WH, Newton-Cheh C, Ikeda M, et al. A common genetic variant in the NOS1 regulator NOS1AP modulates cardiac repolarization. *Nature Genetics* 2006;38(6):644–651. [PubMed: 16648850]
- Bacanu SA, Devlin B, Roeder K. The power of genomic control. *American Journal of Human Genetics* 2000;66(6):1933–1944. [PubMed: 10801388]
- Balding DJ. A tutorial on statistical methods for population association studies. *Nature Reviews Genetics* 2006;7(10):781–791.
- Barabasi AL. Network medicine—from obesity to the "diseasome". *New England Journal of Medicine* 2007;357(4):404–407. [PubMed: 17652657]
- Basile VS, Masellis M, Potkin SG, Kennedy JL. Pharmacogenomics in schizophrenia: the quest for individualized therapy. *Human Molecular Genetics* 2002;11(20):2517–2530. [PubMed: 12351588]
- Benjamini, Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B* 1995;57:289–300.
- Blokland GA, McMahon KL, Hoffman J, Zhu G, Meredith M, Martin NG, et al. Quantifying the heritability of task-related brain activation and performance during the N-back working memory task: a twin fMRI study. *Biological Psychology* 2008;79(1):70–79. [PubMed: 18423837]
- Botstein D, Risch N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nature Genetics* 2003;33 Suppl:228–237. [PubMed: 12610532]
- Brzustowicz LM. NOS1AP in schizophrenia. *Current Psychiatry Reports* 2008;10(2):158–163. [PubMed: 18474209]
- Calhoun VD, Liu J, Adali T. A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. *Neuroimage* 2009;45(1 Suppl):S163–S172. [PubMed: 19059344]
- Cardon LR, Bell JI. Association study designs for complex diseases. *Nature Reviews Genetics* 2001;2(2):91–99.
- Cardon LR, Palmer LJ. Population stratification and spurious allelic association. *Lancet* 2003;361(9357):598–604. [PubMed: 12598158]
- Chanock SJ, Manolio T, Boehnke M, Boerwinkle E, Hunter DJ, Thomas G, et al. Replicating genotype-phenotype associations. *Nature* 2007;447(7145):655–660. [PubMed: 17554299]
- Chapman J, Clayton D. Detecting association using epistatic information. *Genetic Epidemiology* 2007;31(8):894–909. [PubMed: 17654599]
- Clayton D, McKeigue PM. Epidemiological methods for studying genes and environmental factors in complex diseases. *Lancet* 2001;358(9290):1356–1360. [PubMed: 11684236]
- Corneveaux JJ, Liang WS, Reiman EM, Webster JA, Myers AJ, Zismann VL, et al. Evidence for an association between KIBRA and late-onset Alzheimer's disease. *Neurobiology of Aging*. 2008
- de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and power in genetic association studies. *Nature Genetics* 2005;37(11):1217–1223. [PubMed: 16244653]

- Dean B, Keriakous D, Scarr E, Thomas EA. Gene expression profiling in Brodmann's area 46 from subjects with schizophrenia. *Australian and New Zealand Journal of Psychiatry* 2007;41(4):308–320. [PubMed: 17464717]
- Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biology* 2003;4(5):P3. [PubMed: 12734009]
- Devlin B, Bacanu SA, Roeder K. Genomic Control to the extreme. *Nature Genetics* 2004;36(11):1129–1130. author reply 1131. [PubMed: 15514657]
- Devlin B, Roeder K. Genomic control for association studies. *Biometrics* 1999;55(4):997–1004. [PubMed: 11315092]
- Di Piero V, Ferracuti S, Sabatini U, Pantano P. A cerebral blood flow study on tonic pain activation in man. *Pain* 1994;56(2):167–173. [PubMed: 8008407]
- Dudbridge F, Gusnanto A. Estimation of significance thresholds for genomewide association scans. *Genetics Epidemiology* 2008;32(3):227–234.
- Dudoit, ScdLK. *Multiple Testing Procedures with application to Genomics*. New York: Springer Verlag; 2008.
- Durrant C, Zondervan KT, Cardon LR, Hunt S, Deloukas P, Morris AP. Linkage disequilibrium mapping via cladistic analysis of single-nucleotide polymorphism haplotypes. *American Journal of Human Genetics* 2004;75(1):35–43. [PubMed: 15148658]
- Egan MF, Goldberg TE, Kolachana BS, Callicott JH, Mazzanti CM, Straub RE, et al. Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America* 2001;98(12):6917–6922. [PubMed: 11381111]
- Ennis S, Jomary C, Mullins R, Cree A, Chen X, Macleod A, et al. Association between the SERPING1 gene and age-related macular degeneration: a two-stage case-control study. *Lancet*. 2008
- Evans DM, Marchini J, Morris AP, Cardon LR. Two-stage two-locus models in genome-wide association. *PLoS Genetics* 2006;2(9):e157. [PubMed: 17002500]
- Fallin D, Cohen A, Essioux L, Chumakov I, Blumenfeld M, Cohen D, et al. Genetic analysis of case/control data using estimated haplotype frequencies: application to APOE locus variation and Alzheimer's disease. *Genome Research* 2001;11(1):143–151. [PubMed: 11156623]
- Fallin D, Schork NJ. Accuracy of haplotype frequency estimation for biallelic loci, via the expectation-maximization algorithm for unphased diploid genotype data. *American Journal of Human Genetics* 2000;67(4):947–959. [PubMed: 10954684]
- Fisher, RA. *Design of Experiments*. Edinburgh: Oliver & Boyd; 1935.
- Freedman ML, Reich D, Penney KL, McDonald GJ, Mignault AA, Patterson N, et al. Assessing the impact of population stratification on genetic association studies. *Nature Genetics* 2004;36(4):388–393. [PubMed: 15052270]
- Gianola D, de los Campos G. Inferring genetic values for quantitative traits non-parametrically. *Genetical Research* 2008;90(6):525–540.
- Glazier AM, Nadeau JH, Aitman TJ. Finding genes that underlie complex traits. *Science* 2002;298(5602):2345–2349. [PubMed: 12493905]
- Gottesman II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. *American Journal of Psychiatry* 2003;160(4):636–645. [PubMed: 12668349]
- Gray JA, Roth BL. Molecular targets for treating cognitive dysfunction in schizophrenia. *Schizophrenia Bulletin* 2007;33(5):1100–1119. [PubMed: 17617664]
- Hecker M, Lambeck S, Toepfer S, van Someren E, Guthke R. Gene regulatory network inference: data integration in dynamic models-a review. *Biosystems* 2009;96(1):86–103. [PubMed: 19150482]
- Helgadottir A, Thorleifsson G, Manolescu A, Gretarsdottir S, Blondal T, Jonasdottir A, et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* 2007;316(5830):1491–1493. [PubMed: 17478679]
- Hidalgo CA, Blumm N, Barabasi AL, Christakis NA. A dynamic network approach for the study of human phenotypes. *PLoS Computational Biology* 2009;5(4):e1000353. [PubMed: 19360091]

- Hikida T, Jaaro-Peled H, Seshadri S, Oishi K, Hookway C, Kong S, et al. Dominant-negative DISC1 transgenic mice display schizophrenia-associated phenotypes detected by measures translatable to humans. *Proceedings of the National Academy of Sciences of the United States of America* 2007;104(36):14501–14506. [PubMed: 17675407]
- Hill WG, Goddard ME, Visscher PM. Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genetics* 2008;4(2):e1000008. [PubMed: 18454194]
- Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nature Reviews Genetics* 2005;6(2):95–108.
- Hoffmann TJ, Lange C, Vansteelandt S, Laird NM. Gene-environment interaction tests for dichotomous traits in trios and sibships. *Genetic Epidemiology*. 2009
- Jiang R, Tang W, Wu X, Fu W. A random forest approach to the detection of epistatic interactions in case-control studies. *BMC Bioinformatics* 2009;10 Suppl 1:S65. [PubMed: 19208169]
- Jung BC, Jhun M, Song SH. A new random permutation test in ANOVA models. *Statistical Paper Springer-Verlag* 2006;48:47–62.
- Kamatani N, Sekine A, Kitamoto T, Iida A, Saito S, Kogame A, et al. Large-scale single-nucleotide polymorphism (SNP) and haplotype analyses, using dense SNP Maps, of 199 drug-related genes in 752 subjects: the analysis of the association between uncommon SNPs within haplotype blocks and the haplotypes constructed with haplotype-tagging SNPs. *American Journal of Human Genetics* 2004;75(2):190–203. [PubMed: 15202072]
- Karlsgodt KH, Glahn DC, van Erp TG, Therman S, Huttunen M, Manninen M, et al. The relationship between performance and fMRI signal during working memory in patients with schizophrenia, unaffected co-twins, and control subjects. *Schizophrenia Research* 2007;89(1–3):191–197. [PubMed: 17029749]
- Keedwell P, Drapier D, Surguladze S, Giampietro V, Brammer M, Phillips M. Neural markers of symptomatic improvement during antidepressant therapy in severe depression: subgenual cingulate and visual cortical responses to sad, but not happy, facial stimuli are correlated with changes in symptom score. *Journal of Psychopharmacology*. 2008
- Kennedy JL, Farrer LA, Andreasen NC, Mayeux R, St George-Hyslop P. The genetics of adult-onset neuropsychiatric disease: complexities and conundra? *Science* 2003;302(5646):822–826. [PubMed: 14593167]
- Kruglyak L. The road to genome-wide association studies. *Nature Reviews Genetics* 2008;9(4):314–318.
- Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nature Genetics* 1995;11(3):241–247. [PubMed: 7581446]
- Lee SH, van der Werf JH, Hayes BJ, Goddard ME, Visscher PM. Predicting unobserved phenotypes for complex traits from whole-genome SNP data. *PLoS Genetics* 2008;4(10):e1000231. [PubMed: 18949033]
- Lee SS, Sun L, Kustra R, Bull SB. EM-random forest and new measures of variable importance for multi-locus quantitative trait linkage analysis. *Bioinformatics* 2008;24(14):1603–1610. [PubMed: 18499695]
- Lewis DA, Gonzalez-Burgos G. Neuroplasticity of neocortical circuits in schizophrenia. *Neuropsychopharmacology* 2008;33(1):141–165. [PubMed: 17805309]
- Lewontin RC. A General Method for Investigating the Equilibrium of Gene Frequency in a Population. *Genetics* 1958;43(3):419–434. [PubMed: 17247767]
- Li J. Marker selection for whole-genome association studies with two-stage designs using dense single-nucleotide polymorphisms. *BMC Proceedings* 2007;1 Suppl 1:S136. [PubMed: 18466479]
- Lin S, Chakravarti A, Cutler DJ. Exhaustive allelic transmission disequilibrium tests as a new approach to genome-wide association studies. *Nature Genetics* 2004;36(11):1181–1188. [PubMed: 15502828]
- Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nature Genetics* 2003;33(2):177–182. [PubMed: 12524541]
- Manly. *Randomization, Bootstrap, and Monte Carlo Methods in Biology*. (2nd ed. ed.). London: Chapman and Hall; 1997.

- Manoach DS, Press DZ, Thangaraj V, Searl MM, Goff DC, Halpern E, et al. Schizophrenic subjects activate dorsolateral prefrontal cortex during a working memory task, as measured by fMRI. *Biological Psychiatry* 1999;45(9):1128–1137. [PubMed: 10331104]
- McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nature Reviews Genetics* 2008;9(5):356–369.
- McClearn GE, Johansson B, Berg S, Pedersen NL, Ahern F, Petrill SA, et al. Substantial genetic influence on cognitive abilities in twins 80 or more years old. *Science* 1997;276(5318):1560–1563. [PubMed: 9171059]
- Menon V, Anagnoson RT, Mathalon DH, Glover GH, Pfefferbaum A. Functional neuroanatomy of auditory working memory in schizophrenia: relation to positive and negative symptoms. *Neuroimage* 2001;13(3):433–446. [PubMed: 11170809]
- Meyer-Lindenberg A, Nichols T, Callicott JH, Ding J, Kolachana B, Buckholtz J, et al. Impact of complex genetic variation in COMT on human brain function. *Molecular Psychiatry* 2006;11(9):867–877. 797. [PubMed: 16786032]
- Meyer-Lindenberg A, Weinberger DR. Intermediate phenotypes and genetic mechanisms of psychiatric disorders. *Nature Reviews Neuroscience* 2006;7(10):818–827.
- Moore JH. The ubiquitous nature of epistasis in determining susceptibility to common human diseases. *Human Heredity* 2003;56(1–3):73–82. [PubMed: 14614241]
- Moore JH. Chapter 1, Analysis of gene-gene interactions. *Current Protocols in Human Genetics* 2008;14. Unit 1. [PubMed: 18972370]
- Papassotiropoulos A, Stephan DA, Huentelman MJ, Hoerdli FJ, Craig DW, Pearson JV, et al. Common Kibra alleles are associated with human memory performance. *Science* 2006;314(5798):475–478. [PubMed: 17053149]
- Perneger TV. What's wrong with Bonferroni adjustments. *British Medical Journal* 1998;316(7139):1236–1238. [PubMed: 9553006]
- Phillips PC. Epistasis--the essential role of gene interactions in the structure and evolution of genetic systems. *Nature Reviews Genetics* 2008;9(11):855–867.
- Potkin, Anand R.; Alphs, L.; Fleming, K. Neurocognitive performance does not correlate with suicidality in schizophrenic and schizoaffective patients at risk for suicide. *Schizophrenia Research* 2003;59(1):59–66. [PubMed: 12413643]
- Potkin, Kennedy; Basile. Brain imaging and pharmacogenetics in Alzheimer's disease and schizophrenia. In: Lerer, B., editor. *Pharmacogenetics of Psychotropic Drugs*. Cambridge: Cambridge University Press; 2002a.
- Potkin, Kennedy; Basile. Combining brain imaging and pharmacogenetics in understanding clinical response in Alzheimer's disease and schizophrenia. In: Lerer, B., editor. *Pharmacogenetics of Psychotropic Drugs*. Cambridge: Cambridge University Press; 2002b. p. 391-400.
- Potkin SG, Turner JA, Brown GG, McCarthy G, Greve DN, Glover GH, et al. Working memory and DLPFC inefficiency in schizophrenia: the FBIRN study. *Schizophrenia Bulletin* 2009;35(1):19–31. [PubMed: 19042912]
- Potkin SG, Turner JA, Fallon JA, Lakatos A, Keator DB, Guffanti G, et al. Gene discovery through imaging genetics: identification of two novel genes associated with schizophrenia. *Molecular Psychiatry* 2009;14(4):416–428. [PubMed: 19065146]
- Potkin SG, Turner JA, Guffanti G, Lakatos A, Fallon JH, Nguyen DD, et al. A genome-wide association study of schizophrenia using brain activation as a quantitative phenotype. *Schizophrenia Bulletin* 2009;35(1):96–108. [PubMed: 19023125]
- Price A, Patterson N, Plenge R, Weinblatt M, Shadick N, DA R. Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics* 2006;38(8):904–909. [PubMed: 16862161]
- Purcell S, Cherny SS, Sham PC. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 2003;19(1):149–150. [PubMed: 12499305]
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* 2007;81(3):559–575. [PubMed: 17701901]

- Robbins TW, Roberts AC. Differential regulation of fronto-executive function by the monoamines and acetylcholine. *Cerebral Cortex* 2007;17 Suppl 1:i151–i160. [PubMed: 17725997]
- Roffman JL, Weiss AP, Goff DC, Rauch SL, Weinberger DR. Neuroimaging-genetic paradigms: a new approach to investigate the pathophysiology and treatment of cognitive deficits in schizophrenia. *Harvard Review of Psychiatry* 2006;14(2):78–91. [PubMed: 16603474]
- Rothman KJ. Statistics in nonrandomized studies. *Epidemiology* 1990;1(6):417–418. [PubMed: 2090277]
- Rothman, KJ. *Modern Epidemiology*. (2 ed.). Lippincott Williams & Wilkins; 1998.
- Salvi E, Guffanti G, Orro A, Lupoli S, Torri F, Potkin S, et al. Ancestry correction in genome-wide association studies: Comparison of different methods to control for population stratification. (Submitted).
- Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, et al. Genomewide association analysis of coronary artery disease. *New England Journal of Medicine* 2007;357(5):443–453. [PubMed: 17634449]
- Seasholtz TM, Wessel J, Rao F, Rana BK, Khandrika S, Kennedy BP, et al. Rho Kinase Polymorphism Influences Blood Pressure and Systemic Vascular Resistance in Human Twins. *Role of Heredity. Hypertension*. 2006
- Serre D, Montpetit A, Pare G, Engert JC, Yusuf S, Keavney B, et al. Correction of population stratification in large multi-ethnic association studies. *PLoS ONE* 2008;3(1):e1382. [PubMed: 18196181]
- Shenton ME, Dickey CC, Frumin M, McCarley RW. A review of MRI findings in schizophrenia. *Schizophrenia Research* 2001;49(1–2):1–52. [PubMed: 11343862]
- Shifman S, Bronstein M, Sternfeld M, Pisante-Shalom A, Lev-Lehman E, Weizman A, et al. A highly significant association between a COMT haplotype and schizophrenia. *American Journal of Human Genetics* 2002;71(6):1296–1302. [PubMed: 12402217]
- Skol AD, Scott LJ, Abecasis GR, Boehnke M. Optimal designs for two-stage genome-wide association studies. *Genetic Epidemiology* 2007;31(7):776–788. [PubMed: 17549752]
- Slawski M, Daumer M, Boulesteix AL. CMA: a comprehensive Bioconductor package for supervised classification with high dimensional data. *BMC Bioinformatics* 2008;9:439. [PubMed: 18925941]
- Stranger BE, Forrest MS, Clark AG, Minichiello MJ, Deutsch S, Lyle R, et al. Genome-wide associations of gene expression variation in humans. *PLoS Genetics* 2005;1(6):e78. [PubMed: 16362079]
- Tamminga CA. The neurobiology of cognition in schizophrenia. *Journal of Clinical Psychiatry* 2006;67 Suppl 9:9–13. discussion 36–42. [PubMed: 16965183]
- Teo YY. Common statistical issues in genome-wide association studies: a review on power, data quality control, genotype calling and population structure. *Current Opinion in Lipidology* 2008;19(2):133–143. [PubMed: 18388693]
- Thomas DC. Are we ready for genome-wide association studies? *Cancer Epidemiology Biomarkers Prevention* 2006;15(4):595–598.
- Thomas DC, Haile RW, Duggan D. Recent developments in genomewide association scans: a workshop summary and review. *American Journal of Human Genetics* 2005;77(3):337–345. [PubMed: 16080110]
- Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, et al. PANTHER: a library of protein families and subfamilies indexed by function. *Genome Research* 2003;13(9):2129–2141. [PubMed: 12952881]
- Vawter MP, Shannon Weickert C, Ferran E, Matsumoto M, Overman K, Hyde TM, et al. Gene expression of metabolic enzymes and a protease inhibitor in the prefrontal cortex are decreased in schizophrenia. *Neurochemical Research* 2004;29(6):1245–1255. [PubMed: 15176481]
- Walters JT, Owen MJ. Endophenotypes in psychiatric genetics. *Molecular Psychiatry* 2007;12(10):886–890. [PubMed: 17895920]
- Wang LY, Comaniciu D, Fasulo D. Exploiting interactions among polymorphisms contributing to complex disease traits with boosted generative modeling. *Journal of Computational Biology* 2006;13(10):1673–1684. [PubMed: 17238838]

- Weinberger DR, Berman KF, Zec RF. Physiologic dysfunction of dorsolateral prefrontal cortex in schizophrenia. I. Regional cerebral blood flow evidence. *Archives of General Psychiatry* 1986;43(2):114–124. [PubMed: 3947207]
- Westfall, PH.; Young, SS. Resampling-based multiple testing: examples and methods for p-value adjustment. New York: John Wiley & Sons; 1993.
- WTCCC. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447(7145):661–678. [PubMed: 17554300]
- Yamada R, Okada Y. An optimal dose-effect mode trend test for SNP genotype tables. *Genetic Epidemiology* 2009;33(2):114–127. [PubMed: 18688846]
- Ziegler A, König IR, Thompson JR. Biostatistical aspects of genome-wide association studies. *Biometrical Journal* 2008;50(1):8–28. [PubMed: 18217698]
- Zielenski J. Genotype and phenotype in cystic fibrosis. *Respiration* 2000;67(2):117–133. [PubMed: 10773783]
- Zondervan KT, Cardon LR. The complex interplay among factors that influence allelic association. *Nature Reviews Genetics* 2004;5(2):89–100.
- Zondervan KT, Cardon LR. Designing candidate gene and genome-wide case-control association studies. *Nature Protocols* 2007;2(10):2492–2501.

Figure 1.

The graph shows the power distribution curves for QT analysis contrasted with a case-control design at $p < .01$ and 10^{-7} ($OR = 1.5$). The x axis portrays the sample sizes and on the y axis the power at each value of the sample size for a 10% percentage variance explained for the QT, a QT MAF of 10% and a marker SNP MAF at 20%. The results for QT and case-control are displayed in black and grey respectively. Compared to QT the case-control curves are shifted to the right indicating that much larger sample sizes are required to reach the same power.

Figure 2.

The graph shows the power distribution curves for QT analysis, contrasted with a case control design using the same parameters as in Fig 1 but with the marker SNP MAF at 10%. Comparing Fig 1 and Fig 2 highlight the effects of the match between tagging and causal SNP MAFs shown in more detail in Fig 3.

Figure 3.

The graph depicts the sample size required for a power = 80 % for a QT phenotype with tagging SNP frequencies from .05 to 0.5 given a 10%, 20%, and 30% MAF for the QT and when the total variance explained by the QT is 10, 20 and 30%. The curves show the sample size required for any combination of the QT values: the lowest sample size is $N=132$ and is reached when there is agreement between the allelic frequencies of the tagging SNPs with that of the actual QT allele, which is unknown. This is true independent of the total amount of variance explained by the QT.

Table 1

Sample sizes needed to reach 80% power to detect a locus responsible for the QT.

SNP MAF	QT MAF = .1			QT MAF = .2			QT MAF = .3		
	QTL .10	QTL .20	QTL .30	QTL .10	QTL .20	QTL .30	QTL .10	QTL .20	QTL .30
0.05	961	470	146	1064	528	349	3957	1968	1305
0.1	439	209	132	1041	510	333	1847	913	602
0.2	1026	502	328	440	210	132	793	386	250
0.3	1781	880	580	780	380	247	441	210	132
0.4	2787	1383	915	1241	610	400	714	346	224
0.5	4195	2087	1385	1882	931	613	1097	538	351