Genetic association analysis and candidate genes

[On the use of specific 'candidate' genes to measure association with a trait, some faltering efforts, and some strategies for improving reliability]

Academic life contains many torments. One is public speaking when you have to cram months of work that you still only partly understand into 15 minutes with an audience half asleep from the combined effects of lunch and three prior grueling hours of such talks by others. Then there is the agony of receiving the funding decision on a grant application, a grant that took weeks of anxiety-filled effort to complete and a decision that will impact on whether your post-docs and lab technicians will have a pay check or not. We should not forget the agony of having to attend strategy review meetings organized by funding agencies whose pointlessness you cannot comment on as the agency pays for your science. However, as bad as all these torments are, a thing you do not wish even on your competitors (well, at least not all of them) is to have to review psychiatric genetic association studies.

What makes them so dreary? Why is it after reviewing dozens of them over years, one looks with dread at seeing another invitation from a well-respected journal in your inbox? Reviewing these papers should be a simple decision: have the investigators shown that the DNA sequence variant they have studied is commoner in the group with the disorder than in the unaffected controls; that is, is it correct to conclude that possessing this sequence variant increases the chances of becoming psychotic, violent, depressed, autistic, or whatever disease they have studied? But it is not simple, and understanding why this is so takes us closer to many of the issues that surround the deeper question of how genes influence behavior.

The attraction of genetic association studies is their power to indicate cause. A functional DNA sequence change—a mutation that alters a protein—that changes a site at which the RNA molecule is cleaved or alters a promoter so as to discourage the start of RNA transcription will have been present since birth. Finding that everyone with the same change has a disease, and that no one without the disease has the variant, comes very close to proving a causal relation—the sort of finding all scientists strive for and dream about. This is true because of a simple principle: the genome that we are saving to make the next generation, the DNA in our eggs or sperm, is inviolable. The environment can't change it.

This means that, in claiming that a particular genetic variant is associated with schizophrenia, we do not worry that schizophrenia could

cause the DNA to mutate. Genetic association studies are quite unlike studies that look for associations between sugar intake and hyperactivity, marital conflict and depression, watching violent films and committing crimes, studies where it is difficult, often impossible, to establish what is causing what.

The difficulty in interpreting genetic association studies lies not in confusing association with cause, but with assuming from the genetic data that 'gene X really is associated with schizophrenia'. There are hidden traps, a problem of unacknowledged assumptions that surface later to confound any simple interpretation of the findings. We need to explain how this could happen and we begin by contrasting the two ways psychiatric geneticists have used to find susceptibility genes: linkage and association.

Linkage versus association: 'candidate' genes

As we outlined in Chapter 3, linkage analysis has the major virtue of interrogating the entire genome all at once. In doing a genome-wide linkage study—when you place some 300 evenly spaced markers over all the chromosomes—you have a chance of detecting a signal from any gene, anywhere in the genome. A gene of large effect (meaning one that confers a more than fourfold risk of developing a disorder) has no place to hide from a well-designed linkage study. You don't need to know anything about that gene, where it is, or what it does. Because we are so ignorant about the causes of psychiatric disorders, linkage analysis is an attractive method.

However, linkage analysis has two important drawbacks. Firstly, it has low power. It is only good at detecting relatively large genetic signals—gene regions that contain variants that quite substantially alter the risk for a disorder. Secondly, even when you find positive results, linkage signals are very broad, typically smeared over tens of millions of base pairs, a region large enough to contain hundreds, if not thousands, of genes.

Association analysis has the opposite combination of strengths and weaknesses. Firstly, signals detected by association are much more focused, typically stretching over tens of thousands rather than tens of millions of DNA base pairs. Secondly, association analysis can detect genes with modest effects on disease risk. The main drawback to association is that, until recently, it could not screen the genome, as linkage can. Instead, you could use a few (typically less than 20) markers in what is called a 'candidate' gene. You gather affected individuals—your cases—and a group of matched unaffected individuals—your controls—and see whether the frequency of the marker variants differs between the two groups.

Firstly, a sequence polymorphism is needed, preferably in the coding region of a candidate gene (shown at the top of Figure 5.1). In this example, the polymorphism has two alleles (*a* and *g*), which form three genotypes (the two homozygotes, *aa* and *gg*, and the heterozygote, *ag*). Secondly, you need a couple of hundred patients with a psychiatric illness (such as schizophrenia) and the same number of controls (unaffected people who are the same in every way to the cases (same age, sex, and so on)). Thirdly, cases and controls must be genotyped and classified as *aa*, *ag* or, *gg*, as shown in the lower right of the figure. Finally, a statistical test is needed to determine whether the two columns of figures are significantly different (in this case they are) or not.

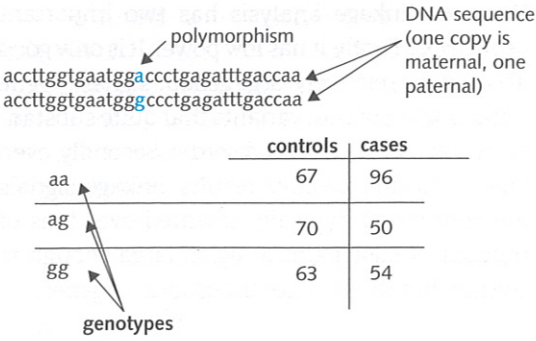

Figure 5.1 Ingredients of a genetic association test.

The candidate gene approach itself comes in two flavors. In the more common approach, the gene is picked because you think it might have something to do with the physiology of the illness. Thus, these genes are called 'physiological candidate' genes. In the second approach, genes are picked because of where they lie in the genome; in particular because they are under linkage peaks. Such genes are called 'positional candidates'.

Physiological candidate genes at first seemed appealing, and there were high hopes that their analysis would lead to major breakthroughs in the genetics of schizophrenia. There was, however, one deep problem with this method: you had to be able to pick good candidate genes, genes that had a plausible chance of being involved in the physiology of the disorder. To pick good physiological candidate genes, you needed to know something about what caused the disorder, which of course we did not when it came to schizophrenia. Just to be clear about what we mean here—two examples of good physiological candidate genes would be the insulin gene if you were studying diabetes or a gene for a cholesterol receptor if you were studying heart disease. While there is no guarantee that variants in these genes would influence risk for the disease, at least you know that the products of these genes are directly involved in the disease process.

Now, one of the great appeals of applying genetic approaches to psychiatric illness was the chance it offered to understand a disorder's origins, its etiology, without recourse to prior knowledge. The old-fashioned top-down approaches had not worked—boiling down urine, blood, or cerebrospinal fluid (the fluid in which your brain and spinal cord float) to look for key biological differences between the psychiatrically well and ill had not worked. Molecular geneticists arrived and said, 'We will solve your problem from the bottom-up. We know from all those good family, twin, and adoption studies that genes must play a role in the etiology of psychiatric illness. Just build us some good labs and give us lots of grant money and we will find those genes that will lead us to the cause.' But when it came to genetic association studies, to get genes they needed to know the cause of the illness—which is what we hired them to find out in the first place.

The fact that the choice of candidate genes might be ad hoc, a question of the scientist's particular interests and prejudices, should not matter if the way of validating their involvement in the disease is robust. And, on the face of it, the validation is straight forward. We ask whether the frequency of a sequence variant in the candidate gene is higher in those with the disease than those without. The question can be answered by well-established tests, tests used in all areas of science, the sort of tests that tell us whether drugs are effective or not, or what risk factors contribute to heart disease and cancer. Entire university departments are staffed with people who do this sort of thing: they are called statisticians. Surely when I carry out or review a genetic association study, if I don't understand how the

statistical test works, all I need do is ring up the Department of Statistics and get a statistician to help. Admittedly statisticians may not be fun to talk to ('What's the difference between a statistician and an accountant? The statistician looks at the other person's shoe'), but statistics is a well-established discipline with a lot of quantitative methodology, lots of equations, and computers; in short pretty reliable stuff.

Late on a Friday evening, you get a request from the *Journal of Uninteresting Studies* to review their latest submission, entitled 'Evidence that *DUP25* contributes to the risk of developing anxiety'. It's the first time this prestigious magazine has asked for your views, and it's good for your career to develop a close, friendly relationship with the editors (for after all, you hope soon to submit your own paper, entitled 'Evidence that *DUP30* contributes to the risk of developing anxiety'). How are you going to decide whether the results are right or not?

Fleeting candidates: *DUP25*

In 2001, *Cell* published a paper about a chromosomal duplication, something you can see down a microscope, that was associated with anxiety (Gratacos *et al.*, 2001). *Cell* has a thoroughbred pedigree as a high-profile, well-respected science journal. The authors, a group from Barcelona, wrote: '*We have identified an interstitial duplication of human chromosome 15q24–26 (named DUP25) which is significantly associated with panic/agoraphobia/social phobia/joint laxity in families and with panic disorder in nonfamilial cases.*' Figure 5.2 shows a picture from the paper showing what they found.

The dots are two DNA probes labeled with different-colored fluorescent dyes and hybridized to a patient's chromosomes (a pair of chromosome 15 homologs). You can see that there appear to be four, not two, pale dots. This indicates that the DNA homologous to the labeled probe is duplicated (*DUP25*). '*Ninety percent of patients diagnosed with one or several anxiety disorders had the duplication. Remarkably, all patients with panic disorder with or without agoraphobia and all patients with social phobia carried DUP25.*' An impressive finding, and even if '*20% of patients with DUP25 did not have any anxiety phenotypes*', it was still enough to get the paper through the review process and out into the world. Was the result true?

One way to find out is to replicate: is *DUP25* also more prevalent in the anxious patients in your clinical sample, just as it is in that of the authors' sample? When other groups tried this, the answer was straightforward: no *DUP25* occurred in either cases or controls. So was anxiety in Barcelona due, in part, to a rare sequence variant found only in Catalonia? No, not that either, for when some geneticists in the UK (Tabiner *et al.*, 2003) tested the Spanish samples that had been used in the original study they reported:

> **There was no evidence of a duplication of signals in distal 15q that would be indicative of *DUP25* in any of the 16 patient samples or in any of the 40 control samples... It is difficult to think of any logical scientific or technical explanation for the differences between the two laboratories in scoring the positive control cultures. However, we were unable to detect any *DUP25*-positive cells, either in the positive control samples from CEPH or in our patient or control samples. Furthermore, we have never had a report of such a duplication in any of the thousands of diagnostic samples that have been scored on high-resolution chromosomes in our laboratory.**

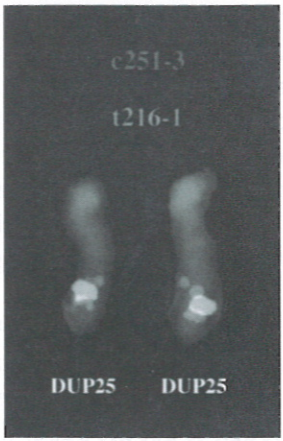

Figure 5.2 Duplication of *DUP25* on chromosome 15.
Source: Gratacos *et al.* (2001).

Coquettish candidates: *COMT*

DUP25 is unusual in a number of ways: it was not a candidate gene, it could not be replicated, and it probably didn't exist in the first place—it was a technical artifact. So in the end, making a decision as to its importance was easy: it's not important. The more typical situation is more complicated. For example, in genetic association studies of schizophrenia, authors pick a candidate gene implicated from one of the etiologic theories of the psychiatric disease, itself based on a large and forbidding neuroscience literature.

The history of neuroscience can briefly be summarized as progressing from the discovery that a specific type of cell (the neuron) is the functional unit of the brain (the alternative view was that there was a continuous connecting membrane arranged like a set of wires) to the discovery of specialized connections between cells (the synapse) and the chemicals that carried information from one side of the synapse to the other (neurotransmitters) (see Chapter 10 for a description of neurons and neurotransmitters).

Neurotransmitters have always fascinated psychiatrists. Many of the drugs used to treat psychiatric and neurological illnesses act by mimicking, blocking, prolonging the life of, or otherwise interfering with neurotransmitters. Many illicit psychoactive drugs act in the same way: the fact that you can make yourself excited, calm, happy, paranoid, depressed, transcendental, or traumatized by interfering with neurotransmitter function has spurred psychiatrists to investigate whether abnormalities of neurotransmitter function are a cause of psychiatric disease.

The neurotransmitter dopamine has been a magnet for the attention of psychosis researchers in large part because, with stunning consistency, drugs that treat the symptoms of schizophrenia block one particular class of dopamine receptors in the brain. Furthermore, their potency in treating symptoms and blocking the receptor are uncannily correlated. These two sets of observations—which are widely accepted—form the basis of the dopamine hypothesis of schizophrenia—far and away the most influential theory from its first proposal in the 1970s to today. This theory posits that, in schizophrenics, the dopamine system is somehow and somewhere hyperactive.

Dopamine itself is not encoded by a gene: the body makes it from raw ingredients using a series of enzymes that are encoded by genes. Once released, the chemical can be degraded enzymatically or by re-uptake into neurons via a transporter in the cell membrane (Figure 5.3).

Theoretically, alterations in any of these steps could lead to functional hyperactivity: excess production, excess release, excess stimulation,

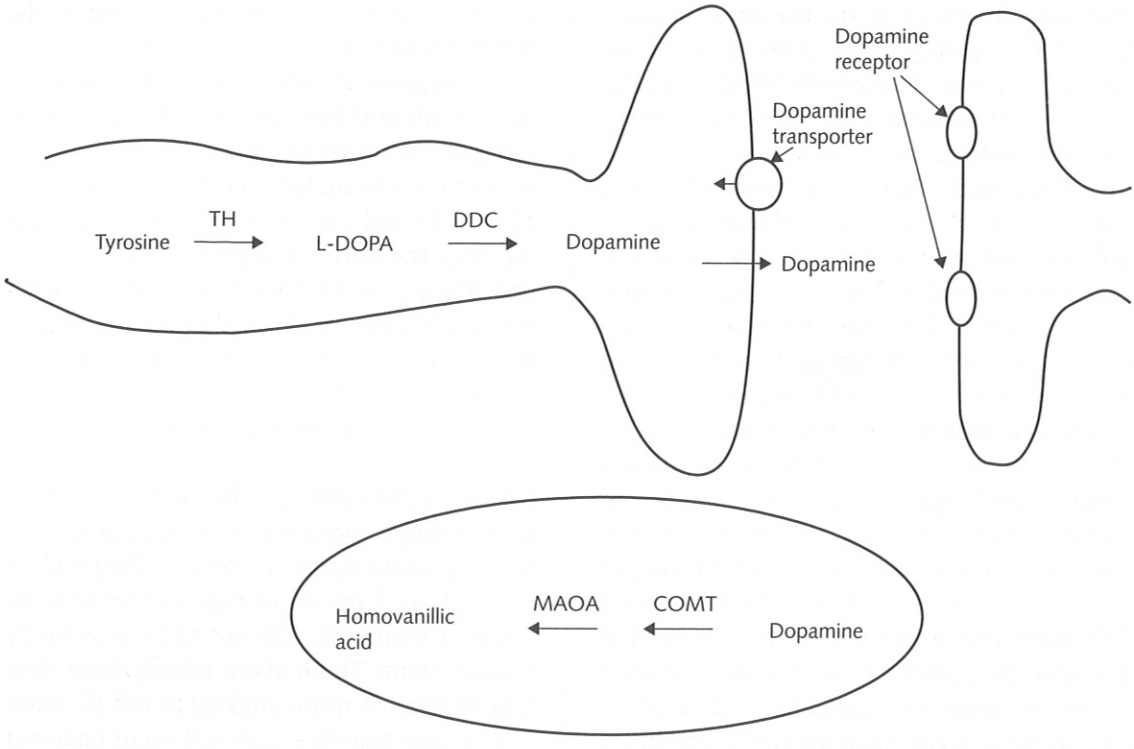

Figure 5.3 Metabolism of dopamine. Dopamine is manufactured from the amino acid tyrosine through the action of two enzymes, tyrosine hydroxylase (TH) and DOPA decarboxylase (DDC), shown here acting within a neuron. Dopamine is released from the neuron and can then act on dopamine receptors, and can then be taken back into the neuron via a dopamine transporter or into another cell type where it is degraded by the action of two enzymes, catechol-*O*-methyl-transferase (COMT) and monoamine oxidase (MAO—there are two variants known as MAOA and MAOB).

excess receptor number, or reduced removal. Despite the accumulated biochemical and neuroanatomical knowledge about dopamine, after 30 years of research no one had been able to prove whether the dopamine hypothesis is true or false.

Genetic association studies, it was widely hoped, would settle this issue once and for all. Dopamine is degraded in several ways, but one of the most important is by an enzyme called catechol-O-methyltransferase or COMT for short (Figure 5.3). As you might gather from the name, it sticks a methyl group (one carbon and three hydrogens) onto the dopamine molecule. Unlike dopamine, COMT is a protein encoded by a gene and that gene contains a sequence variant that alters the activity of the enzyme: one form is less efficient at breaking down dopamine. Consequently, a test of the dopamine hypothesis is that individuals with the inefficient *COMT* gene variant will be at greater risk of developing schizophrenia. This is because more dopamine will hang around longer in the synapse in a person with the inefficient form of COMT. Therefore, the question can be put this way: is the inefficient *COMT* variant more common in schizophrenics than in controls? If it is, then here is evidence in favor of the dopamine hypothesis.

The experiment is straightforward: I determine the frequency of the *COMT* gene variant in 100 cases and 100 controls, which might be 23% in schizophrenics and 30% in controls. I assume that in reality there is no difference in the frequency of the genetic variant between patients and controls, so any difference I do see is due to chance. I then work out the probability that the observed difference is due to chance. This is usually done mathematically, but an easier way to see what's going on (and a method we also use) is to simulate the experiment many times on a computer, on the assumption that there is indeed no real difference. By using a random number generator, I have access to infinitely large populations of cases and controls (of the sick and the well), in which the frequency of the *COMT* variant is the same. I repeatedly select 100 individuals at random from each population and calculate the frequency of the variant in the two samples. I might find that a difference as big, or bigger, than the one I found in my real experiment occurs in a fifth of the simulated experiments. This means the result of the real experiment could easily have occurred by chance alone. Conventionally, when the probability drops to less than 1 in 20 (a 5% chance), I conclude that my results cast doubt on my first assumption, namely that there is no difference in the frequency between the two groups.

In 1996, genetic association was used in exactly this way. '*This study investigated this [COMT] polymorphism in 78 unrelated schizophrenic patients and 78 comparison subjects matched for age and ethnicity. The frequency of the polymorphism [the key variant] was 0.51 in the schizophrenic patients and 0.53 in the comparison subjects, and no significant allelic or genotypic associations were observed*' (Daniels *et al.*, 1996). That should do it. No association. The dopamine hypothesis has failed the test.

But by 2003, there were 18 further publications that reported the results of the same kind of study; by 2006, the number had grown to 44. Nevertheless, 11 years after the first study, we still have no systematic evidence that the presence of an inefficient form of the COMT protein at synapses increases the risk for schizophrenia. A careful meta-analysis was performed by one of us (J.F.) and concluded that when including only the better-designed studies, no evidence was found for any association between the *COMT* gene and schizophrenia (Munafò *et al.*, 2005). (See Box 5.1 for a description of meta-analyses.) Why, you might ask, did scientists keep testing this hypothesis?

One important reason why studies of *COMT* kept coming is that a few of the individual studies did produce significant results. The problem here is that if you do enough statistical tests, a few of them will turn out to be positive by chance alone. Think about rolling three dice. It is, by chance, quite unlikely to roll all 'ones' (if you have true dice it should occur once out of every 6^3 times or once in 216 rolls). But if you sit there for a while and roll the dice enough times, you are guaranteed, if you are patient enough, to get three ones.

Gauging probabilities and biases

When we use genetic association to find genes involved in psychiatric illness, the answer we get back is usually a *P* value, or some similar statistic, whose interpretation is not simple. The result of the genetic association test, in a rather counter-intuitive fashion, is defined by the likelihood that it is wrong. The *P* value that is conventionally used is 0.05, 1 in 20, or 5%, a figure that has a sacrosanct place in medical research, not always well deserved. It is particularly inappropriate for association studies of complex diseases and we can explain why.

Let's assume, as a first approximation, that we have 20,000 human genes of which 20 are involved in the etiology of schizophrenia. If I pick a gene at random (which is almost what we are doing when we study physiological candidate genes because we have so little idea of which genes are actually involved), that gene has 20/20,000 or a 1/1,000 chances of being a real schizophrenia susceptibility gene. Imagine I am a successful scientist with a grant to study, by genetic association, 1,000 of these genes using the case–control design. Applying the 5% *P* value, we would expect that around 50 out of our 1,000 genes will be 'significant' by chance alone—random results that we call by the inappropriately benign phrase 'false positives.' (These are anything but benign because other research groups will often spend months of time and thousands of dollars trying to replicate such results.) Of the 1,000 genes, one is likely to be a true positive. So, to a rough first approximation, our well-funded project would be expected to produce 51 significant results from our 1,000 genes tested of which 50 (roughly 98%) are false.

Clearly, we need to use much more stringent *P* values (that is, much lower than 5%) to give us a more reasonable proportion of true positive findings. In one of the deeper ironies of the field of psychiatric genetics, this problem was much better worked out for linkage studies. Due to the influence of one statistical geneticist—Newton Morton—it was very early imprinted on the field that an LOD ($\log_{10}$ of odds) score of 3.0 (which depending on some technical issues equals a *P* value of between 0.001 and 0.0001) was needed to declare significant linkage. This is about right in that most linkage studies would involve something like 300 different tests. With stunning inconsistency, association studies, when they began to appear in the literature, utilized a *P* value of 0.05, even though the multiple testing problem was far greater than in linkage studies. This fateful decision allowed many scientists to publish 'significant' association results and was therefore very beneficial to their Curriculum Vitae. But it was certainly not so helpful to the science.

A second reason why people kept testing the *COMT* gene and schizophrenia returns us to the realm of the sociology of science. Put simply, it made such a good story, it just had to be true. In defending a truly powerful theory, this position is not as illogical as it sounds. Astronomers in the 19th century noted abnormalities in the orbit of Uranus that could not be explained by the principles of Newtonian theories. As a result of these findings, they could have concluded that there was something wrong with Newton's theories but they didn't; the theory had been so good in other ways, leading them to suspect that there might be some way of reconciling the findings with Newtonian mechanics. So they kept looking for other explanations, assuming the theory was right. Their perseverance was rewarded when they discovered the new planet of Neptune whose gravitational influences explained the abnormalities in Uranus's orbit.

To put it politely, the dopamine hypothesis of schizophrenia is not in the same league as Newton—not even close. But scientists have strong affiliations to their theories, not unlike many people's affiliation to their political parties. *COMT* was not the only dopamine gene repeatedly tested in schizophrenia. For example, there have been multiple case–control association studies of the dopamine transporter gene (the product of which 'sucks up' dopamine back into the pre-synaptic cell), and dozens and dozens of studies of the dopamine receptors of which there are five, each with their own gene. The results have not all been negative. Indeed, the gene for one type of dopamine receptor—called *DRD2*—has a variant that

alters an amino acid in the protein. In 2008, a meta-analysis suggested that this variant did modestly influence the risk for schizophrenia (Allen *et al.*, 2008).

The *COMT* gene story demonstrates a pattern that is repeated with other genetic association studies: a physiological candidate gene is picked based on quite weak evidence of disease involvement. Lots of studies are done using liberal statistical criteria for significance. Some positive studies emerge that then generate more attempts to replicate them. Meta-analyses then begin to accumulate, typically showing that the combined effect is either small or non-existent. This pattern is common enough to warrant a couple more examples, which also demonstrate some of the pitfalls of interpretation and the difficulties in assessing the robustness of the evidence.

Groundswell candidates: *DRD2*

Dopamine is also blamed for alcoholism. There is good evidence that part of the pleasurable effect obtained when one drinks alcohol is mediated by dopamine. Not surprisingly attempts have been made to nail the problem of alcoholism to variation in dopamine receptors. In 1990, Kenneth Blum at San Antonio and Ernest Noble at UCLA wrote in the *Journal of the American Medical Association* (Blum *et al.*, 1990):

> **We report the first allelic association of the dopamine D2 receptor gene in alcoholism... In the present samples, the presence of A1 allele of the dopamine D2 receptor gene correctly classified 77% of alcoholics, and its absence classified 72% of nonalcoholics. The polymorphic pattern of this receptor gene suggests that a gene that confers susceptibility to at least one form of alcoholism is located on the q22–q23 region of chromosome 11.**

In other words, Blum and Noble had found a genetic variant that predicted susceptibility to alcoholism, which was big news indeed.

Four years later, the story deserved a commentary in *Science*—two major failures to confirm their findings out of three studies didn't look good for the Blum and Noble work. But the bad news wasn't over. In 1992, a study from a group at Washington University in St Louis initially appeared to support the A1 connection—but Washington University psychiatrist Robert Cloninger says that when the group expanded the sample it found, to his 'chagrin', that the association between the D2 receptor and alcoholism 'faded out.' In all, the article concluded that '*attempts to replicate the finding have been largely unsuccessful.*' Joel Gelernter had complained, '*It is now four years since the paper came out and we still don't have a mutation or anything that could explain the effects that this A1 allele is supposed to convey.*'

Blum and Noble didn't take this lying down. They wrote to *Science*, criticizing the article for sending 'the wrong message' and creating '*embarrassment for scientists who are pioneering at the forefront of research in the genetics of addictive–compulsive disorder.*' In their view, '*We are witnessing the birth of a new paradigm in our understanding of the genetic basis of addictive–compulsive behaviors, and from the total evidence available it should be clear that the* DRD2 *gene will continue to play an important role in these behaviors*' (Blum and Noble, 1994). Blum and Noble argued against Cloninger's negative result as follows: '*Careful scrutiny of their follow up paper revealed that the sample of alcoholics in the second study was heterogeneous, including both severe and*

less severe alcoholics. The inclusion of less severe alcoholics diluted the sample.' In other words, the genetic effect will vary depending on whether it is measured in those with mild or severe alcohol dependence.

Back in 1993, an analysis of all of the available studies of the A1 allele of the *DRD2* gene concluded: '*The findings to date can best be explained by more conservative interpretations than a confirmed physiologically important allelic association between* DRD2 *alleles and alcoholism*' (Gelernter *et al.*, 1993). By 2007, over 40 studies of the *DRD2* gene and alcoholism had been published. Again, part of the reason for the continuing interest is the inherent attraction of the idea, its biological plausibility.

Another reason, still sociological, is that scientific journals afford more importance to positive than negative results. This is usually difficult to observe, as a lot of studies are needed to detect its effect, but sufficient were available in the *DRD2* alcoholism literature for us to be certain of its existence (Munafò *et al.*, 2007). Figure 5.4 shows the effect size of each study (y-axis) against the year of publication (x-axis). The correlation is clear and highly significant: studies with the largest effects were published first, while negative studies only appeared later.

As with the *COMT* story, there are also biological reasons for the continuing interest in *DRD2*. Firstly, it might be that the effect was just too small too detect. We can estimate just how small by putting together data from all of the published samples: combined analyses of all *DRD2* studies indicated that, if the effect was there, it accounted for as little as 0.2% of the variation in alcoholism. This is so small that no single study had enough subjects to detect the effect reliably. Small effect size could explain the inconsistent results.

Secondly, conflicting results might arise because samples had been taken from different populations that happened to differ at the *DRD2* locus for reasons other than alcoholism. A key assumption of genetic association is that the two groups—cases and controls—are equivalent in all respects other than the disease being tested. Consider what would be found if that assumption was relaxed.

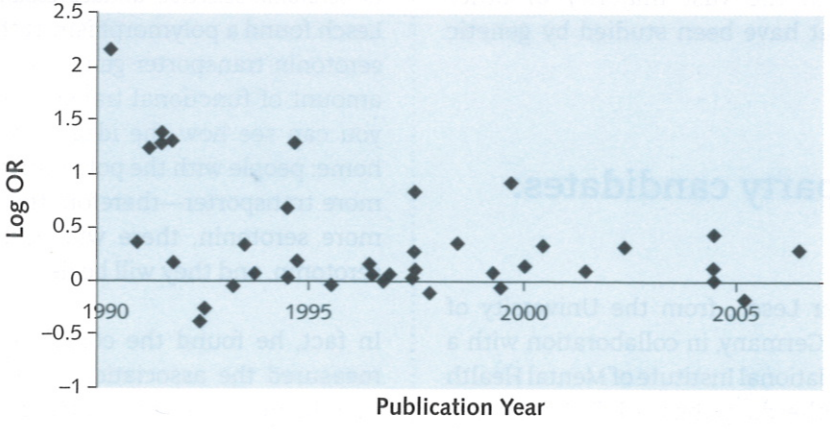

Figure 5.4 Effect size of each study on *DRD2* and alcoholism against the year of publication (OR, odds ratio). Source: Munafo *et al.* (2007).

The frequencies of alleles at many genes differ among ethnic groups, so that an association study comparing alcoholics of one ethnicity with non-alcoholics of another would identify a large number of significant differences, none related to alcoholism. For example, if you carried out a study of the genetic basis of religion and compared gene frequencies in groups who practice Buddhism and those who do not, you would be very likely to find many significant differences, solely due to the fact that Buddhism is commoner in East Asians than it is in Europeans, and there are many genetic differences between East Asians and Europeans. Even a slight degree of mismatching in the ethnicity of the cases and controls could introduce bias into the results.

This problem of genetic admixture is often blamed for the reporting of spurious associations, although, to be fair, it's not that easy to find confirmed examples. In fact, a joint analysis of *DRD2* studies in different ethnic populations gives the same answer as analyzing the studies separately (again the large number of *DRD2* papers makes this possible) (Munafò *et al.*, 2007), indicating that ethnic heterogeneity is probably not an issue in this case, nor indeed in the vast majority of other disorders that have been studied by genetic association.

Prozac party candidates: 5-*HTT*

In 1996, Peter Lesch, from the University of Wurzburg in Germany, in collaboration with a group at the National Institute of Mental Health (NIMH) in Bethesda, published, in *Science*, the results of a genetic association study between a personality trait and a polymorphism in the serotonin (or 5-hydroxytryptamine, 5-HT) transporter gene (5-*HTT*) (Lesch *et al.*, 1996).

In the same way that the dopamine theory has dominated neurobiological theories of the cause of schizophrenia, 'something wrong with serotonin' has been a standard theory of the origin of depression for some time, again despite the lack of conclusive evidence. Serotonin, like dopamine, is a neurotransmitter in the brain, released by neurons and degraded or re-used by re-uptake via a transporter in the cell membrane, the serotonin transporter. There's evidence that the amount of serotonin available to neurons correlates with mood: low levels correlate with the occurrence and the lethality of suicide attempts and can be used to predict future suicide attempts. Prozac and other drugs that block the re-uptake of serotonin (serotonin reuptake inhibitors or SSRIs) are effective antidepressants and anxiolytics, presumably because of the increased availability of serotonin at the synapse. Finding any evidence that genetic variants in the serotonin system are associated with mood disorders could, according to NIMH Director Tom Insel, in 2005, '*lead to a genetic test for vulnerability to depression and a way to predict which patients might respond best to serotonin-selective antidepressants*.' So when Lesch found a polymorphism at the start of the serotonin transporter gene that increased the amount of functional transporter (Figure 5.5), you can see how the idea must have struck home: people with the polymorphism will have more transporter—therefore they will take up more serotonin, there will be less available serotonin, and they will be depressed.

In fact, he found the complete opposite. He measured the association between the polymorphism and a personality trait called neuroticism, which is genetically related to anxiety and depression (that is, a large proportion of the

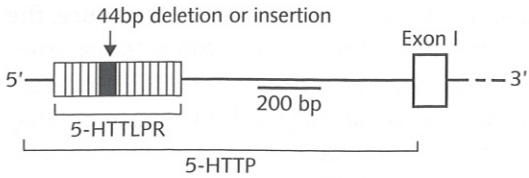

Figure 5.5 Serotonin is also known as 5-hydroxytryptamine (5-HT), so the promoter is denoted as 5-HTTP. On the left is the region of the gene called the promoter, where transcription starts, containing a repeated sequence motif (shown as a box with bars and marked 5-HTTLPR—the serotonin-transporter-linked polymorphic region). Subsequent work has shown that there is considerable sequence diversity in the promoter, in addition to the length polymorphism.
Source: Lesch *et al.* (1996).

genetic variants that contribute to variation in neuroticism also contribute to susceptibility to anxiety and depression) (see Chapter 4, p. 64). As personality traits are, by definition, stable and enduring characteristics and relatively easy to measure, they could be a better phenotype for genetic studies than mood itself, which is transient and difficult to assess.

Lesch found differences in the DNA sequence at the start of the serotonin transporter gene, differences that altered the amount of gene product. There were two alleles, which he called *s* (for short) and *l* (for long) at the transporter (later work has established that the situation is considerably more complicated and that there are multiple alleles). Individuals with an *s* allele were more neurotic and more prone to anxiety and depression than individuals with an *l* allele. That ought to mean that the *s* allele increases the amount of transporter. However, '*the basal activity of the* l *variant was more than twice that of the* s *form of the serotonin transporter gene promoter.*' Lesch argued his way out of this unexpected finding on the following grounds: '*The therapeutic effects of the SSRIs have primarily been demonstrated in neuropsychiatric patients, who may have some primary serotonin or other neurotransmitter dysfunction that is ameliorated by the SSRIs, whereas our findings are in a sample of the general population.*' In other words, serotonin behaves differently in patients than it does in you or me.

The subsequent decade has seen literally hundreds of papers investigating the role of the transporter gene in psychiatric disorders. As with *DRD2* and schizophrenia, many papers report an association and many papers don't. In an attempt to resolve the issue, in 2002, one of us (J.F.) and a colleague, Marcus Munafò, reviewed the data from 20 papers and concluded that there was no effect (Munafò *et al.*, 2003). We also carried out a large study, enrolling the 2,000 most and least neurotic people in the south west of England, selected from a community sample of more than 88,000 people. We tested whether there was an association with the serotonin transporter polymorphism. There wasn't (Willis-Owen *et al.*, 2005). Surely the story would end there?

It did not. A number of objections to our findings were raised: the assessment of neuroticism we used was different from that in the Lesch study, the study was in a different population, the genetic effects at the tails of the distribution might be different from those in the middle, and the environmental effects might somehow be different. A lot of interest was then being paid to gene–environment interactions: might there be a differential effect in people with the *s* allele who had had a stressful life event? It was not enough to have the genetic predisposition: there had to be an environmental stressor for its effect to be manifest.

5-*HTT* gene–environment interactions

Richie Poulton and colleagues in New Zealand have for many years been collecting information on the same 1,000 people, contacting them every year and flying them back home if necessary, to find out what has happened to them and assess how they are. The cohort were in their 20s when Poulton, together with Terrie Moffitt at the Institute of Psychiatry in London and her partner Avshalom Caspi, carried out a genetic analysis of the 5-*HTT* gene. They reported in *Science*: '*The effect of life events on informant reports of depression was stronger among individuals carrying an s allele than among l/l homozygotes. These analyses attest that the* 5-HTT *gene interacts with life events to predict depression symptoms, an increase in symptoms, depression diagnoses, new-onset diagnoses, suicidality, and an informant's report of depressed behavior*' (Caspi *et al.*, 2003).

Psychiatrists and psychologists around the world have loved this piece of work; it's inventive and interesting and suits our belief that genes act in complicated ways, in combination with the environment, to work their effects. Genetic tests for this gene variant are currently being marketed on the internet (for those who can afford them). There are, naturally, attempts to replicate the result—14 independent studies (by 2009) including the original report (Munafò *et al.*, 2009). But only one reported a statistically significant interaction apparently identical to that observed in the original report. Three studies reported no evidence of a statistically significant interaction, one interpreted their results as offering 'modest support' based on subgroup analyses, and six reported a significant interaction, which was different from that observed in the original report.

Overall, when we reviewed the literature, the positive results for the serotonin-transporter-linked polymorphic region (5-HTTLPR) interactions are still compatible with chance findings (Munafò *et al.*, 2009). Moreover, as the main effect of 5-HTTLPR genotype and the interaction effect between 5-HTTLPR and environment on risk of depression are negligible, given reasonable assumptions regarding likely genetic and environmental effect sizes, the published studies are underpowered. And a significant finding from an underpowered study is a false positive.

Why reviewing association studies induces headaches

We hope by now you understand and maybe even sympathize with our suffering at yet another request to review yet another genetic association study. The studies all begin the same way. Some rather unconvincing story is told trying to relate a physiological candidate gene to schizophrenia, alcoholism, depression, personality disorder, or autism. The methods section describes a sample size of about 200 cases and 200 controls (too small to have power to detect any but the largest of effects). Many markers are tested against a range of diagnostic definitions and, sure enough, a few modest *P* values emerge, allowing the authors to claim in their conclusion that this gene is associated with disease. Perhaps like a lapsed Catholic forced back to mass, you knew what was coming but the sense of belief was gone.

'*The great tragedy of science—the slaying of a beautiful hypothesis by an ugly fact*' (Thomas Huxley,

1825–1895) has not stopped the seemingly endless production of genetic association studies, nor dampened enthusiastic endorsement of claims to have identified genes contributing to psychiatric disease, as well as to personality, sexual orientation, intelligence, even empathy. It is true that there are cases where things have simply gone wrong (as the *DUP25* example shows). And there are statistical problems: the significance threshold (the *P* value to use to decide whether a result is significant or not) has been set too high so that the result is likely to be a false positive. But these factors are not enough to hold back the tide of publications.

We've emphasized the role that the sociology of science plays, in wedding scientists to received opinion, so that there is entrenched opposition to accepting a negative result (which is anyway more difficult to publish than a positive result). The pattern of the *COMT* and *5-HTT* stories is typical: a physiological candidate gene is picked based on quite weak evidence of disease involvement. Lots of studies are done using liberal statistical criteria for significance. Some positive studies emerge that generate more attempts to replicate. And so on, ad infinitum.

One other reason why negative association studies do not slay the theories that gave birth to them is that the experimental design makes it impossible for a negative result to do so. All a negative result tells you is that the effect could not be detected. It might be there, just too small to see. A common thread running through all of the genetic association studies is that the effects are small, contributing to less than 1% of the variation in the phenotype, or increasing the risk of the disorder by a small amount. This realization has led people to consider whether there might be circumstances in which the effects could be increased.

For example, could we study populations in which the total genetic variation is less, so that the contribution of any one variant is relatively increased? This genetic simplification is possible with model organisms but is difficult to do in humans. Results from studying island populations (in Sardinia, Croatia, and the Shetlands, for instance), where the degree of relatedness between individuals is higher than in other parts of the world, has not radically improved things (Eaves *et al.*, 2000): finding the molecular basis of personality variation, such as neuroticism and other traits, has been no easier, implying the existence of many loci of very small effect, as is the case in more outbred, genetically diverse populations.

'Endophenotypes': getting closer to *5-HTT*

Could there be phenotypes related to psychiatric disease or behavior that have a different genetic architecture, composed of genetic loci with larger effect sizes? Working with such phenotypes gives the candidate-gene approach a better chance of success. A number of investigators have pursued this idea, among whom Danny Weinberg at NIMH has made the strongest claims for the success of the approach (Meyer-Lindenberg and Weinberger, 2006). In his words:

> **Genes do not encode for psychiatric phenomena (for example, hallucinations and panic attacks), and so, almost by definition, the more behavioral the phenotype, the less directly it will be predicted by a genotype. This leads to the strategy of studying underlying quantitative traits that more directly**

index biology, analogous to moving from the study of cardiac insufficiency or stroke (complex diseases) to ventricular hypertrophy and cholesterol metabolism. This strategy offers several advantages for behavioral disorders: biological traits are expected to be closer to the genetic substrate, enhancing penetrance; the traits should be observable in genetically at risk but behaviorally unaffected individuals; and, if the traits are sufficiently causally upstream to index a biological process that makes a separable contribution to disease, the genetic architecture should be simplified.

In this context, 'enhancing penetrance' means increasing the effect size. The closer the phenotype is to the site of genetic action, the more direct, the more immediate it is, the larger the effect should be. From this standpoint, working with behavior to find genes is like trying to work out a river's course by standing on its bank and looking back to the mountains from which it descends, standing at such a great distance that the snow peaks appear to be no more than wrinkles on the horizon. Travel up the river closer to the foot hills, so that the mountains are easier to discern, and the route from glacier to river creek emerges.

Figure 5.6 illustrates the concept of the endophenotype. On the left of the figure are the

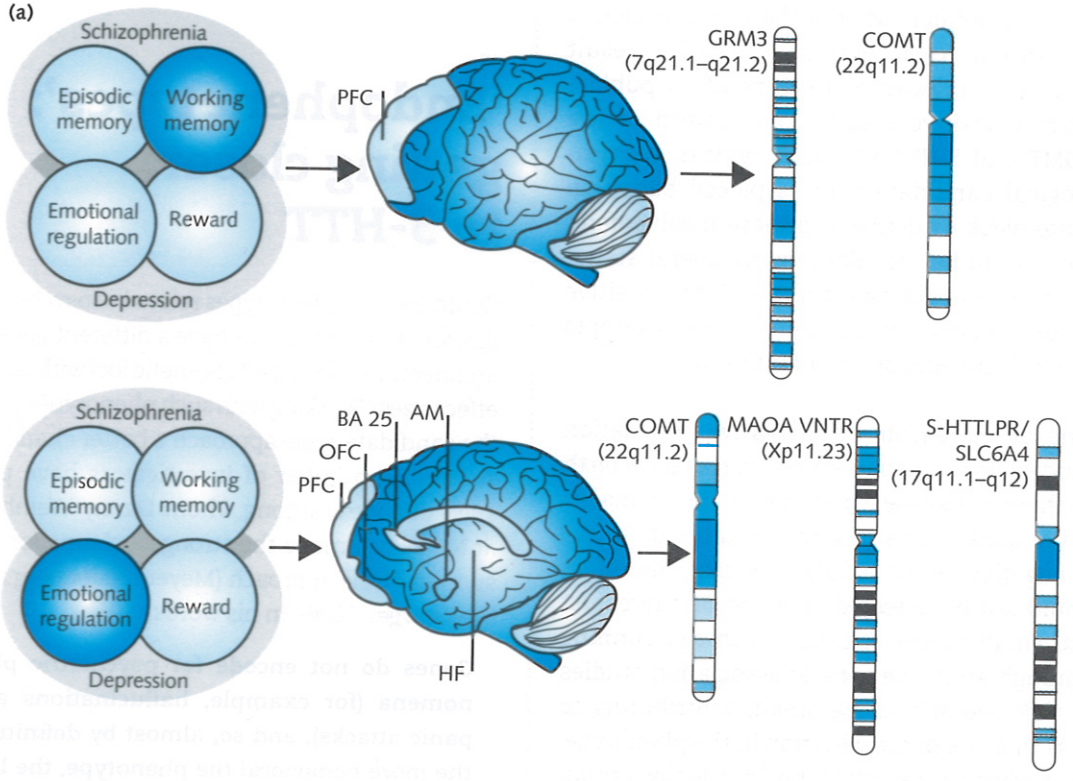

Figure 5.6 Concept of the endophenotype (AM, amygdala; BA, Brodmann's area; HF, hippocampal formation; OFC, orbitofrontal cortex; PFC, prefrontal cortex).
Source: Meyer-Lindenberg and Weinberger (2006).

phenotypes that psychiatrists work with, the diffuse, poorly defined diagnostic entities of schizophrenia and depression. Within these are better but still diffuse psychological constructs such as working memory and emotional regulation. In the middle are the brain regions, the neural circuitry whose activity gives rise to the psychology and psychiatric disease. And on the right are the molecules of DNA from which all of this pathology arises. The closer the researcher is to the DNA, the better the chance of observing its effect. Endophenotypes are measures of brain function that lie between DNA and psychiatric illness.

For example, as shown in Figure 5.6 from Meyer-Lindenberg and Weinberger (2006), cortical dysfunction is genetically analyzed instead of schizophrenia, and emotional regulation instead of depression. Cortical dysfunction is linked to variation in the *COMT* and *GRM3* genes. Emotional regulation is linked to variation in *COMT*, the monoamine oxidase A gene (*MAOA*) and the serotonin transporter polymorphism. Genetic association is said to be easier to find when analyzing these neuronal phenotypes because they are closer to the genetic action.

Do the published data bear this out? It's still not clear. In one influential study, Weinberger and colleagues gave people a task in which they had to match the emotion (angry or afraid) of one of two faces to that of a third (Hariri *et al.*, 2002). People were asked to do this emotional matching test while the activity of brain regions was monitored using magnetic resonance imaging, or MRI. The amygdala, a region known to be involved in emotional processing, was found to be more active in people with an *s* allele (the same allele whose possession Peter Lesch had identified as increasing neuroticism).

The remarkable thing about this study is that the effect was found with just 28 individuals. The genetic effect attributable to the 5-*HTT* locus explained about 40% of the variation in brain activity. This is almost two orders of magnitude larger than the effects we have been discussing in the genetic association of psychiatric disease, consistent with the idea that intermediate, or endophenotypes, do indeed 'enhance genetic penetrance.' But this is, in design, just another genetic association study, prone to all of the problems that genetic association analyses face, and we should take the same critical stance in assessing its value as we did with the other work. In other words, is it true?

The pattern of results emerging from the genetic analysis of intermediate phenotypes looks very similar to what we saw for case–control studies of psychiatric disease: a high-profile publication reports a large effect size, with a small sample, and is followed by other studies using larger samples that report smaller effects, or non-replication. The difference with the case–control studies we have discussed above is that the number of intermediate phenotypes studies is relatively small. This is because imaging brains is much more expensive (say at least US$400 for each subject) and time consuming than assessing phenotypes by asking people to fill in a questionnaire (costs less than 50 cents and takes about 5 minutes).

By 2007, 14 studies had been published that looked at the relationship between 5-HTT and amygdala activation, far fewer than the hundreds that analyzed the relationship between 5-HTT and personality, but enough to carry out a meta-analysis of the results (Munafò *et al.*, 2008). This showed a significant result, but with a greatly reduced effect size: down to 10%. In fact, Marcus Munafò at Bristol University who carried out the meta-analysis

plotted the downward trend in the estimated effect and predicted that, in 2008, the first study showing an effect in the opposite direction would be published. That prediction was fulfilled, raising the possibility that there may after all be no true effect attributable to the 5-HTT locus, or indicating that the effect is small, just as small as in the classical psychiatric genetic association studies. Similar conclusions have been reached in meta-analyses of other intermediate phenotypes (Flint and Munafò, 2007).

Is there a way out of the quagmire?

What else could be done to improve success rates? One suggestion is to give up candidate genes and find something better, but that means testing all genes, an idea that for technical reasons it has been difficult to realize. In the next chapter, we'll describe how that became possible. Secondly, large sample sizes, much larger than countenanced, could be collected. During the time that the genetic association studies of psychiatric disorders were being carried out, other diseases whose origins were also obscure were subject to the same genetic analysis: cancer, diabetes, hypertension, stroke, arthritis, asthma, and other common illnesses. Researchers in all of these areas were facing the same difficulties. For example, here is David Allison, a statistical geneticist, summarizing progress in obesity genetics (Redden and Allison, 2003):

> **Over the past decade, numerous research projects have reported associations between nutritional phenotypes (obesity, type 1 and 2 diabetes mellitus and energy expenditure) and regions of the human chromosomes. Unfortunately, many of the reported associations have not been replicated in independent research. The nonreplication of these association findings is a concern and has caused some researchers to question the utility of association methodology in genetic studies.**

John Ioannidis has been a particularly outspoken critic of genetic association studies. Here he is in an editorial in the *Journal of the National Cancer Institute* lambasting studies of cancer (Ioannidis, 2006):

> **In 2005 alone, 194 original research articles were published that probed gene–disease associations for breast cancer; I selected every 10th article ($n = 19$) for perusal. Fifteen of these articles claimed associations overall, in subgroups, or for specific outcomes. The parade of claimed associated genes in this tiny sample is already impressive: *HER2*, *IL10*, *NCOA3*, *TGFBR1*, *TGFB1*, *ESR2*, *HFE*, *IGF-I*, *ESR1*, *AR*, *CHEK2*, *PAI-1*, *XRCC1*, *HSMH2*, *SULT1A1*, and *IFNG*. If all these claimed associations are real, a 10% sample of the published genetic association research in a single year alone seemingly suffices to explain all that causes breast cancer: the total attributable fraction from this small sample of associations already reaches close to 100%.**
>
> **Is this an apotheosis of data dredging? Even in my small sample of 19 articles, one comes across an association that is statistically significant only in the sub-sub-subgroup of postmenopausal women who have at least three pregnancies and also have no wild-type allele; a polymorphism with statistically significantly decreased risk for early-stage breast cancer but increased risk for advanced-stage disease; another increasing the risk especially for grade 3 tumors; a marker with 13 variants, of which one shows a statistically significant association versus**

all others combined, while hundreds of different groupings are conceivable; polymorphisms that have no statistically significant associations on their own but do in one of their many constructed haplotypes; joint effects of polymorphisms of different genes acting in obvious or not-so-obvious pathways; associations that are not even tested statistically and so forth.

This quote exemplifies the frustration that many of us working on the genetics of common disease, not just psychiatric illness, felt. The job ought to be easy: we knew that the disorders had a heritable component and molecular genetics gave us the power to investigate them at a molecular level. Why could we not find robust results? A few papers were pointing to the importance of sample size. For example, an analysis of over 3,000 people in a study of a gene thought to be involved in type 2 diabetes (*PPARG*) gave a clear indication that the investigators had found the correct variant (Altshuler *et al.*, 2000). Studies with similar sample sizes also detected a signal of roughly the same effect size, and combined analyses improved the significance of the result (rather than weakening it as we found in meta-analyses of psychiatric association studies).

Studies that analyzed thousands of cases and controls were extremely rare in psychiatry. It was becoming clear that they were needed. But we also needed a way to interrogate something other than candidate genes. Candidate gene studies in psychiatry were not proving to be productive. Of the 20,000 or so genes in the human genome, about 10,000 are expressed in the brain. How could we test their involvement? The next chapter explains the technological and methodological developments that made it possible.

BOX 5.1 Meta-analysis

Meta-analysis is a quantitative method that combines results from a body of evidence—typically a number of published studies—in order to arrive at a consensus conclusion. The basic elements of meta-analytic techniques can be traced back to Fisher (1925), but Glass introduced the term meta-analysis in 1976 (Glass, 1976). In genetics, the method has been applied to linkage and genetic association studies as a way of determining whether the literature supports claims for a relationship between genetic variants and disease.

The advantage of meta-analysis over simply counting the number of studies that report a positive or negative finding is that it takes into account the power of each study, so that a small study with a highly significant result does not outweigh a much larger, and therefore more powerful, study with a non-significant result.

A meta-analysis produces two key results. First, it tells us if the results from the various studies are statistically homogeneous. That is, can they be seen as replications of one another? If the

answer to this is yes, then a meta-analysis gives an aggregate estimate of the effect under study that is properly averaged over all the studies. This statistic reflects the best possible aggregate estimate of the effect based on all the available information.

A meta-analysis is only as good as its constituent parts, and its success is in part determined by the diligence of the investigators identifying suitable studies and extracting the correct information.

Summary

1. Association analysis, the correlation of variants in specific genes with a trait, offers an alternative to linkage analysis that has greater power and focus, but casts a narrower net and requires knowing enough about the disorder to nominate 'candidate' genes.

2. Results of association studies for behavioral and psychiatric phenotypes have often been inconsistent, difficult to replicate, and influenced by the inevitable preconceptions inherent in choosing candidate genes.

3. 'Endophenotypes' that are considered to be closer to the site of genetic action, such as brain area activation, are a strategy for enhancing effect size in association studies.

4. Larger sample sizes improve the robustness and reliability of results from association studies.

References

Allen, N.C., Bagade, S., McQueen, M.B., Ioannidis, J.P., Kavvoura, F.K., Khoury, M.J., Tanzi, R.E., and Bertram, L. (2008). Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database. *Nat Genet* **40**:827–834.

Altshuler, D., Hirschhorn, J.N., Klannemark, M., Lindgren, C.M., Vohl, M.C., Nemesh, J., Lane, C.R., Schaffner, S.F., Bolk, S. *et al.* (2000). The common PPARγ Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat Genet* **26**:76–80.

Blum, K. and Noble, E.P. (1994). The sobering D2 story. *Science* **265**:1346–1347.

Blum, K., Noble, E. P., Sheridan, P.J., Montgomery, A., Ritchie, T., Jagadeeswaran, P., Nogami, H., Briggs, A.H., and Cohn, J.B. (1990). Allelic association of human dopamine D2 receptor gene in alcoholism. *JAMA* **263**:2055–2060.

Caspi, A., Sugden, K., Moffitt, T.E., Taylor, A., Craig, I.W., Harrington, H., McClay, J., Mill, J., Martin, J. *et al.* (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* **301**:386–389.

Daniels, J.K., Williams, N.M., Williams, J., Jones, L.A., Cardno, A.G., Murphy, K.C., Spurlock, G., Riley, B., Scambler, P. *et al.* (1996). No evidence for allelic association between schizophrenia and a polymorphism determining high or low catechol O-methyltransferase activity. *Am J Psychiatry* **153**:268–270.

Eaves, I.A., Merriman, T.R., Barber, R.A., Nutland, S., Tuomilehto-Wolf, E., Tuomilehto, J., Cucca, F., and Todd, J.A. (2000). The genetically isolated populations of Finland and Sardinia may not be a panacea for linkage disequilibrium mapping of common disease genes. *Nat Genet* **25**:320–323.

Fisher, R.A. (1925). *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd.

Flint, J. and Munafò, M.R. (2007). The endophenotype concept in psychiatric genetics. *Psychol Med* **37**:163–180.

Gelernter, J., Goldman, D., and Risch, N. (1993). The A1 allele at the D2 dopamine receptor gene and alcoholism. A reappraisal. *JAMA* **269**:1673–1677.

Glass, G.V. (1976). Primary, secondary and meta-analysis. *Educat Res* **5**:3–8.

Gratacos, M., Nadal, M., Martin-Santos, R., Pujana, M.A., Gago, J., Peral, B., Armengol, L., Ponsa, I., Miro, R. *et al.* (2001). A polymorphic genomic duplication on human chromosome 15 is a susceptibility factor for panic and phobic disorders. *Cell* **106**:367–379.

Hariri, A.R., Mattay, V.S., Tessitore, A., Kolachana, B., Fera, F., Goldman, D., Egan, M.F., and Weinberger, D.R. (2002). Serotonin transporter genetic variation and the response of the human amygdala. *Science* **297**:400–403.

Ioannidis, J.P. (2006). Common genetic variants for breast cancer: 32 largely refuted candidates and larger prospects. *J Natl Cancer Inst* **98**:1350–1353.

Lesch, K.-P., Bengel, D., Heils, A., Sabol, S.Z., Greenberg, B.D., Petri, S., Benjamin, J., Muller, C.R., Hamer, D.H., and Murphy, D.L. (1996). Association of anxiety related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science* **274**:1527–1530.

Meyer-Lindenberg, A. and Weinberger, D. R. (2006). Intermediate phenotypes and genetic mechanisms of psychiatric disorders. *Nat Rev Neurosci* **7**:818–827.

Munafò, M.R., Clark, T.G., Moore, L.R., Payne, E., Walton, R., and Flint, J. (2003). Genetic polymorphisms and personality in healthy adults: a systematic review and meta-analysis. *Mol Psychiatry* **8**:471–484.

Munafò, M.R., Bowes, L., Clark, T.G., and Flint, J. (2005). Lack of association of the *COMT* ($Val^{158/108}$Met) gene and schizophrenia: a meta-analysis of case–control studies. *Mol Psychiatry* **10**:765–770.

Munafò, M.R., Matheson, I.J., and Flint, J. (2007). Association of the *DRD2* gene Taq1A polymorphism and alcoholism: a meta-analysis of case–control studies and evidence of publication bias. *Mol Psychiatry* **12**:454–461.

Munafò, M.R., Brown, S.M., and Hariri, A.R. (2008). Serotonin transporter (5-HTTLPR) genotype and amygdala activation: a meta-analysis. *Biol Psychiatry* **63**:852–857.

Munafò, M.R., Durrant, C., Lewis, G., and Flint, J. (2009). Gene × environment interactions at the serotonin transporter locus. *Biol Psychiatry* **65**:211–219.

Redden, D.T. and Allison, D.B. (2003). Nonreplication in genetic association studies of obesity and diabetes research. *J Nutr* **133**:3323–3326.

Tabiner, M., Youings, S., Dennis, N., Baldwin, D., Buis, C., Mayers, A., Jacobs, P.A., and Crolla, J.A. (2003). Failure to find DUP25 in patients with anxiety disorders, in control individuals, or in previously reported positive control cell lines. *Am J Hum Genet* **72**:535–538.

Willis-Owen, S.A., Turri, M.G., Munafò, M.R., Surtees, P.G., Wainwright, N.W., Brixey, R.D., and Flint, J. (2005). The serotonin transporter length polymorphism, neuroticism, and depression: a comprehensive assessment of association. *Biol Psychiatry* **58**:451–456.