Genetics of Schizophrenia and Bipolar Affective Disorder: Strategies to Identify Candidate Genes

D.J. PORTEOUS,* K.L. EVANS,* J.K. MILLAR ,* B.S. PICKARD,* P.A. THOMSON,* R. JAMES,* S. MACGREGOR,† N.R. WRAY,* P.M. VISSCHER,† W.J. MUIR,‡ AND D.H. BLACKWOOD‡

Medical Genetics Section, University of Edinburgh, Western General Hospital, Edinburgh, Scotland, EH4 2XU; †Institute of Cell, Animal and Population Biology, Ashworth Laboratories, School of Biology, The University of Edinburgh, The King's Buildings, Edinburgh, Scotland, EH9 3JT; ‡Division of Psychiatry, University of Edinburgh, Royal Edinburgh Hospital, Edinburgh, Scotland, EH10 5HF

Schizophrenia (SCZ) and bipolar affective disorder (BPAD) (formerly termed manic-depressive illness) are severe, disabling psychiatric illnesses that feature prominently in the top ten causes of disability worldwide (Lopez and Murray 1998). Each will affect about 1% of the population in their lifetime. The cost of providing treatment for mental illness in the UK National Health Service is estimated at 10% of total expenditure. The total costs (medical, social, economic) are estimated at £32 billion per annum for the population of 50 million in England (Bird 1999; see also www.mentalhealth.org.uk). World Health Organization predictions indicate that major depression will be second only to heart disease in terms of disability-adjusted life years (DALYs) by 2020 (Lopez and Murray 1998). Research into the causes of these devastating disorders and the development of improved interventions is a high scientific, social, individual, and public health priority. Unfortunately, these remain Cinderella disorders compared to cancer and heart disease, both in terms of governmental and societal recognition and national and international research support. Despite their high prevalence and, indeed, decades of neuroscience research, little is known with certainty about their cellular and molecular bases. Consequently, treatments remain largely empirical and palliative. This apparent impasse, however, justifies neither complacency nor despondency, because the one consistent, replicable finding is that family, twin, and adoption studies demonstrate a major genetic component in both SCZ and BPAD (Merikangas and Risch 2003). A decade ago only a handful of genes for monogenic disorders had been positionally cloned through linkage studies. Over the past five years, as the Human Genome Project and associated tools have developed, that number has soared to over 1000. As the Human Genome Project passes from formal completion to full development as a universal biological tool, we can expect accelerated progress in tackling the much more difficult task of genetically dissecting the common complex disorders including major mental illness.

EVIDENCE FOR A GENETIC COMPONENT TO SCZ AND BPAD

The risk to a first-degree relative of a person affected by SCZ or BPAD is an order of magnitude higher than that of the general population (from ~1% to 10–15% lifetime risk for each disorder) (Kendler and Diehl 1993; Merikangas and Risch 2003). Further evidence for the high heritability of SCZ and of BPAD comes from twin and adoption studies. Concordance rates for monozygotic twins are around 50% for SCZ and 60–80% for BPAD. Importantly, these concordance rates are much higher than for dizygotic twins (10–15%), and the biological risk is unaffected by adoption. Several chromosomal regions that may harbor susceptibility genes for SCZ and for BPAD have been identified by a range of genetic strategies (Owen et al. 2000; Potash and DePaulo 2000; Riley and McGuffin 2000). There is also growing evidence that relatives of sufferers are at higher risk of other psychiatric diagnoses within the schizophrenia–affective disorder spectrum (Kendler et al. 1998; Wildenauer et al. 1999; Berrettini 2000; Valles et al. 2000). This suggests that some genetic risk factors may contribute to a range of psychotic symptoms that cross the traditional diagnostic boundaries of SCZ and affective disorders. In part, this may reflect the fact that diagnosis of psychiatric illness is an imprecise science; psychiatric phenotypes are almost entirely based on profiles of behavioral indices and communication patterns. The use of standardized diagnostic criteria (such as the Diagnostic and Statistical Manual of Mental Disorders, DSM-IV, published by the American Psychiatric Association, and The ICD-10 Classification of Mental Health and Behavioural Disorders, published by the World Health Organization) has ensured good reproducibility of diagnoses between researchers, but there is wide overlap of symptoms between the diagnostic categories of SCZ, BPAD, and recurrent, major (unipolar) depression. In the absence of reliable biological or genetic markers specific for SCZ or BPAD, the validity of existing classification remains uncertain. It would be a major advance if genetic studies yielded molecular diagnostic methods that could not only resolve some of the present uncertainties in psychiatric diagnosis, but also inform treatment choice and disease prognosis. Until such time, we can only speculate that variability in individual diagnoses reflects a combination of genetic (and possibly allelic) heterogeneity, polygenic inheritance, and variation in individual life trajectories/environmental exposures. Nevertheless, post-genome science most certainly offers the best hope for determining the biological basis

of these devastating conditions and for developing effective, evidence-based treatments.

GENETICS OF COMPLEX DISEASE AND THE PARTICULAR CHALLENGES FOR PSYCHIATRIC GENETICS

The most appropriate way forward in dissecting the genetics of common disease is subject to much debate (Terwilliger and Weiss 1998; Evans et al. 2001a; Botstein and Risch 2003). The most pragmatic is to follow both main schools of thought as they each have strengths pertaining to different underlying genetic models. The first focuses on variants, which may be rare in the population, but which have detectable, individual effects. For these, the methods that worked well for single-gene disorders, a combination of molecular cytogenetics, family linkage studies, and candidate gene studies, can be transferred directly. It is, however, worth commenting on how problematic the last of these options is for the psychiatric geneticist. Unlike the common genetic disorders of the other major organs, such as cancer or heart disease, cellular pathology in the brain of SCZ and BPAD patients is poorly defined, and brain biopsies are rarely available. With perhaps a third of all known genes expressed in the brain and a plethora of neurodevelopmental and neurophysiological theoretical constructs to choose between, the number of possible candidates is perhaps greater than for any other class of organic disorder. Add to these considerable problems the more general issue of incomplete penetrance, and the full magnitude of the task will be appreciated. The age at first diagnosis for SCZ is typically in late adolescence/young adulthood, slightly later for BPAD, indicating a biological (and genetically influenced) developmental window for expression or, possibly and related to this, an age-dependent susceptibility to life changes and exposures. The severity and individual response to illness are also highly variable. The history of drug medication will likewise be variable and may result in pathological side effects. The field would benefit greatly from a set of validated biomarkers.

Linkage studies are dependent on the availability of informative families, and application to complex traits can be problematic, because it may be hard to define a genetic model that explains the observed inheritance pattern. Replication of linkage results is needed to provide credibility to initial linkage reports, which are likely to overestimate the size of effect, whereas subsequent studies should reflect the true size of effect (Lander and Kruglyak 1995; Terwilliger and Weiss 1998; Botstein and Risch 2003). Lander and Kruglyak (1995) provided a generally accepted yardstick by which to assess claims for genome-wide linkage. They recommended that a LOD of 3.3 be taken as primary evidence for significant linkage accounting for multiple testing issues. Where that value was matched or exceeded, they argued that additional studies with p values of 0.01 (or LOD = 1.2) can be taken as evidence for replication (as testing is of a prior hypothesis). Failure to replicate does not necessarily imply an initial report is a false positive; it may reflect lack of power in the replication study, population heterogeneity, diagnostic differences, or statistical fluctuation (Lander and Kruglyak, 1995). We explain below the extent to which these yardsticks have been matched.

A second school of thought argues from a theoretical and experimental perspective for a polygenic model to explain the genetic variation in complex traits. By virtue of their frequency in the human population, the argument goes that the origins of common genetic disease in humans most probably lie in common and ancient sequence variants and in multiple variants of additive action. As Risch and Merikangas (1996) pointed out, if the genotype relative risk (GRR) is small (as predicted by the common ancient variant hypothesis), then unachievable numbers of simplex families (sib pairs or trios) would be required to reach genome-wide significance (see also Botstein and Risch 2003). On the other hand, association studies, the formal comparison of allele frequencies in cases and matched controls, do have the power to detect even a very modest GRR with realistic numbers of unrelated samples. The association mapping strategy is powerful and attractive, yet problematic. It relies on the scored variant, typically a single-nucleotide polymorphism (SNP), of which there are some 3 million known examples to choose between in the public domain (http://www.ncbi.nlm.nih.gov/SNP; http://snp.cshl.org), being a functional (disease-causing) variant itself or, much more likely, in linkage disequilibrium with a causative variant. It is becoming apparent that linkage disequilibrium is unevenly and unpredictably distributed across the genome (varying by over two orders of magnitude) and that although some studies have shown that a limited number of common haplotypes occur across ethnically diverse populations (see, e.g., Gabriel et al. 2002), other studies have shown that there are a limited number of haplotypes in any one population, but that there is significant variation in haplotypes between populations (see, e.g., Kauppi et al. 2003). This emphasizes the importance of case-matching control strategies that avoid the confounder of population stratification (Clayton and McKeigue 2001). Of course, it is feasible to simply increase the number and density of SNPs tested, but this raises both cost and the statistical problem of multiple testing. The problems inherent in map-based association studies have prompted others to promote a sequence-based strategy that focuses on variation within coding elements as an empirical way to reduce cost, a rational way to reduce the problem of multiple testing, and a logical way to maximize the probability that causal SNPs will be discovered and typed (Botstein and Risch 2003). We and others would extend the logic to include regulatory elements, identified empirically or highlighted by virtue of comparative genome analysis. There is thus no doubt about the importance of association studies as part of the end game in positional cloning in complex trait genetics, but the value of extended family-based studies should not be underestimated.

As already alluded to, there are several alternative family-based linkage approaches, which typically test segregation in multiplex families (multiple affecteds in three- or more generation families), affected sib pairs or trios (one affected and both parents). To identify susceptibility loci for highly heritable, genetically complex diseases

such as SCZ or BPAD, appropriate study design is essential. This topic has been the subject of intense debate as to which of the two main family-based strategies should be employed: large numbers of small families, or smaller numbers of large extended families. The differential efficacy of the two depends on the distribution of genetic risk for the disease of interest. If the risk loci are common in the population and have a small effect on disease risk, the small-families strategy is preferred, largely because the overall sample size can be maximized. On the other hand, if the susceptibility loci are more genetically heterogeneous (multiple susceptibility loci), the best strategy is to identify larger, extended families. Such large families minimize heterogeneity by collecting large numbers of related individuals, most of whom share the same disease risk alleles. Although collecting large numbers of small families may seem a cost-effective strategy, if heterogeneity is expected (as in the case of SCZ and BAPD), the power to detect susceptibility loci will be limited. This point was made clearly by MacGregor et al. (2002) in response to Levinson et al. (2002); it is true that a very large ($n = 900$) sib pair collection has 80% power to detect linkage with a LOD >3 if 75% of families are segregating mutations at the gene of interest, but realistically, the power drops to 40% even if there is only 50% heterogeneity, and to just 5% power at a still modest level of 33% heterogeneity (MacGregor et al. 2002). Thus, unless selected from a single population isolate, the power to replicate linkage detected in multiplex families is quite limited, even with very large sib pair collections. This theoretical analysis justifies our decision to adopt a twin-pronged approach to the difficult task of identifying susceptibility genes for SCZ and BPAD, through genome-wide surveys for cytogenetic lesions in probands and through genome-wide linkage studies on multiplex families.

The usefulness of this approach is exemplified by the genetics of Alzheimer's disease (AD) (http://www.ncbi.nlm.nih.gov:80/entrez/dispomim.cgi?id=104300). The high incidence of AD in individuals with Down syndrome (DS) implicated a gene or genes on Chromosome 21 and, subsequently, a mutation was found in the Chromosome 21 amyloid precursor protein (APP) gene in an early-onset extended AD family. Linkage analysis in other early-onset families then also implicated the presenilin 1 (PS1) gene. Presenilin 2 (PS2) was then discovered by virtue of its homology with PS1. Linkage analysis identified a locus on 19q13.1-q13.3, and subsequent association studies identified the apolipoprotein E (ApoE) gene from this region. The ApoE4 allele has been shown to be a significant risk factor in both familial and sporadic AD. Downstream of these genetic discoveries, research has linked the presenilins to the γ-secretase activity that regulates APP processing (Hardy and Israel 1999) and has shown that the learning deficit in mice that overexpress human mutant APP can be rescued by immunization with APP antibodies (Chen et al. 2000; Janus et al. 2000). Thus, the genetics of AD illustrates the value of combining cytogenetics, linkage, and association to dissect late-onset, neurological disorders of complex behavioral phenotype and inheritance pattern and connecting these to the characteristic pathology through biological study.

GENOME-WIDE STRATEGIES FOR MAPPING SUSCEPTIBILITY GENES: MOLECULAR CYTOGENETICS

Genetic Evidence for the DISC1 Locus as a Risk Factor in Psychosis

We first reported evidence for linkage (LOD = 3.3) between a balanced reciprocal translocation between human Chromosomes 1 and 11 and major mental illness (SCZ, schizoaffective disorder, and recurrent major depression) in a single Scottish family (St Clair et al. 1990) as the Human Genome Project was formally launched. Follow-up investigations of the reciprocal translocation identified the two breakpoints as t(1;11) (1q42;q14) (Muir et al. 1995). Positional cloning of the breakpoint identified two novel genes disrupted by the translocation on Chromosome 1 (DISC1/DISC2, for *d*isrupted *i*n *sc*hizophrenia 1 and 2) (Fig. 1, top) (Millar et al. 2000b). A third gene, TRAX, undergoes intergenic splicing to DISC1 and may affect DISC1 expression (Fig. 1, top) (Millar et al. 2000a). There are no genes near the breakpoint on Chromosome 11. A clinical follow-up on the family (Blackwood et al. 2001), which described additional family members with psychosis, reported a LOD score of 7.1 when subjects with recurrent major depression, BPAD, or SCZ were classed as affected (Fig. 2).

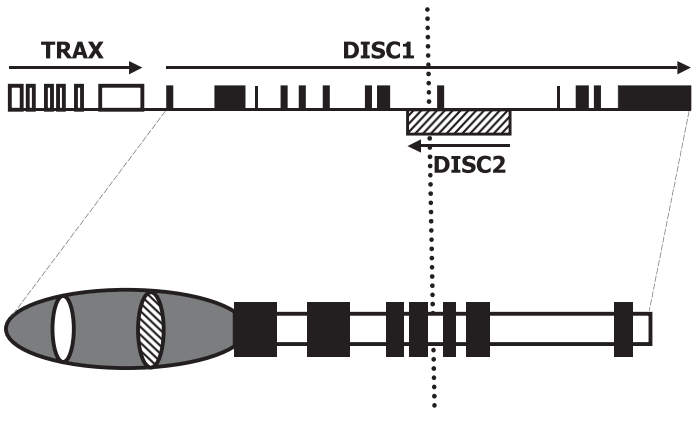
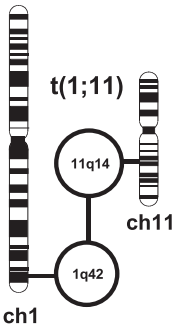


Figure 1. The genomic organization of the DISC1 locus and predicted protein structure of DISC1. (*Top*) Schematic of the transcriptional orientation and exon/intron structure of TRAX (*open boxes*) lying immediately centromeric (5′) of DISC1, the exon/intron structure of DISC1 (*solid black boxes*), the antiparallel and antisense position of DISC2 (*hatched box*) and, as a dotted vertical line, the position of the t(1;11) translocation breakpoint. (*Bottom*) Schematic view of the predicted protein structure of DISC1, indicating the amino-terminal globular head domain (*solid gray ellipse*), including putative nuclear localization signals (*open ellipse*), a serine/phenylalanine-rich domain (*hatched ellipse*), and seven domains (*solid black boxes*) with coiled-coil-forming potential within the carboxy-terminal tail domain (*open*).



	t(1;11) YES N=36	t(1;11) NO N=35
Schizophrenia	7	0
Bipolar affective disorder	1	0
Recurrent major depression	10	0
Conduct disorder with anxiety in adolescence	2	1
Minor depression	1	3
Alcoholism	0	1

Figure 2. The t(1:11) translocation and segregation with psychosis. Diagram of the t(1;11), which breaks at ch1q42 and ch11q14. The diagnosis of patients (ascertained blinded to karyotype status) is tabulated. The LOD score for SCZ alone is 3.4 and for SCZ plus BPAD and recurrent major depression is 7.1 (Blackwood et al. 2001).

The most parsimonious explanation for the correlation between the t(1;11) and psychosis is as a consequence of DISC1/DISC2 gene disruption (Millar et al. 2003b), but brief mention should be made of an alternative explanation proposed by Klar (2002), who seized on the observation that about half of the t(1:11) subjects developed a major psychotic diagnosis, and about half did not. Klar (2002) proposes an elegant, if elaborate, hypothesis that requires three hypothetical genes, *DOH1, SEG*, and *RGHT*. *DOH1* is required for brain laterality specification and is active on one chromatid, but switched off on the other by imprinting; *SEG* exists on the same chromosome, but is unlinked to DISC1; and a *trans*-acting factor, *RGHT* (*right* hander), utilizes *SEG* to govern nonrandom segregation of sister chromatids at the critical period of development when brain hemisphere asymmetry is determined. It is then hypothesized that the translocation separates *DOH1* from *SEG*. The net consequence of the t(1;11) according to this scheme is that 50% of individuals carrying the translocation will have abnormal brain laterality and develop psychiatric illness, and the other 50% will be normal. The clinical picture is not so simple. This five-generation family has been under clinical observation and care for over 30 years (Blackwood et al. 2001). Overall, 62% of t(1;11) carriers are affected, but if the youngest generation is set aside (many of the individuals not yet having reached the average age of onset at the time of ascertainment), the figure rises to 70% of translocation carriers being affected. When the P300 event-related potential (P300 ERP), a trait marker of risk, is taken into account, the proportion of affected translocation carriers increases still further (Blackwood et al. 2001). The P300 ERP is considered to be a measure of the pace and efficiency of information processing in the brain and is abnormal in translocation carriers, with or without a major psychiatric diagnosis. This indicates the presence of central nervous system abnormalities in *all* translocation carriers, including those who are clinically unaffected, and is thus inconsistent with the strand segregation hypothesis. In contrast, the proportion of clinically affected translocation carriers and the P300 ERP data are consistent with our model of dominant inheritance coupled with reduced penetrance and variable expressivity dependent on the action of modifiers (genetic and/or environmental). Indeed, we propose that disruption of *DISC1* and *DISC2*, plus the actions of genetic and environmental modifying factors, are sufficient to explain the psychiatric illness arising from inheritance of the t(1;11) chromosome (Millar et al. 2003b). In support of this argument, the relative risk of the t(1;11) is equivalent to that of the monozygotic co-twin of an affected individual.

Replication of the Genetic Evidence for the DISC1 Locus as a Risk Factor

The most convincing independent genetic evidence so far for the involvement of the TRAX/DISC1 region in mental illness has come from the Finnish population. Initial linkage evidence was found in an internal isolate of Finland (LOD = 3.7) and a common haplotype spanning 6.6 cM reported in 3 of the 20 families multiply affected with SCZ and schizoaffective disorder (Hovatta et al. 1999). A later analysis of 134 sib pairs collected from throughout Finland gave additional support for the same 1q32.2-q41 region on Chromosome 1 with a maximum LOD of 2.6 for a diagnostic class that included SCZ, schizophrenia spectrum disorders, and bipolar and unipolar disorders (Ekelund et al. 2000). Fine mapping of both Finnish populations resulted in maximum LOD scores for SCZ and schizoaffective disorder at D1S2709, a microsatellite located in intron 9 of DISC1, in both the combined sample ($Z_{max} = 2.71$) and outside the internal isolate ($Z_{max} = 3.21$) (Ekelund et al. 2001).

Other data implicating this region in psychosis have come from linkage analysis in families of diverse ethnic origin. Detera-Wadleigh et al. (1999) reported a maximum LOD of 2.67 in a 30-cM region spanning 1q25-1q42 in bipolar families, some of whose members were affected by SCZ or schizoaffective disorder. Gejman et al. (1993) reported a maximum LOD of 2.39 at D1S103 (1q42.2) in one North American family with bipolar disorder. Most recently, a suggestive nonparametric linkage score (NPL = 2.18), equating approximately to a LOD score of 1.2, has been reported in Taiwanese families with diagnoses of SCZ, schizoaffective disorder, and other non-affective psychotic disorders (Hwu et al. 2003). Our own studies have also found evidence for linkage in a subset of BPAD families in Scotland (S. MacGregor et al., unpubl.).

Finally, as discussed earlier in the context of the t(1;11) clinical follow-up study, the use of endophenotypes in molecular studies may overcome the effect of reduced penetrance on the ability to detect linkage and association and help to resolve the affected status of family members. In this context, it is worth mentioning that Gasperoni et al. (2003) undertook a QTL analysis of spatial working memory, an endophenotype for SCZ (Cannon et al. 2000; Glahn et al. 2003). Gasperoni et al. (2003) provide evi-

dence of linkage and association between D1S283 at ch1q42 and variation in spatial working memory between individuals affected by SCZ and their unaffected cotwins ($p = 0.007$ and $p = 0.003$, respectively). This marker was the most telomeric of the markers tested and lies in 1q42.2 approximately 1 Mb centromeric to DISC1.

In light of the growing evidence of independent linkage of 1q42 to major mental illness, we have reanalyzed data initially published in Devon et al. (2001). In that study, we described 15 SNP variants in the DISC1 gene, of which 4 were carried forward for an association study, but no evidence for statistically significant association with any one SNP, or pair of consecutive SNPs, was found (Devon et al. 2001). There is, however, weak evidence of association ($p = 0.0034$; $p = 0.044$, corrected for multiple testing) of a 3-SNP haplotype and BPAD. None of the other 3 or 4 SNP haplotypes showed significant association with BPAD or SCZ (data not shown). This preliminary finding provides tentative evidence of association in the Scottish population between BPAD and polymorphisms in DISC1/DISC2. Further higher resolution association studies stratified by clinical phenotype may allow us to define more specifically the nature and effect size of DISC1/DISC2 variants.

Biology and Predicted Function of DISC1

Full-length *DISC1* is composed of 13 exons and covers over 50 kb of genomic sequence and is subject to alternative splicing. The full-length or long (L) transcript utilizes all 13 exons. A commonly spliced variant of full-length *DISC1*, the long variant (Lv) transcript, utilizes a proximal splice site in exon 11, thereby reducing the transcript by 66 nucleotides. Additional *DISC1* transcripts have been identified by RT-PCR and Northern blotting, and protein isoforms consistent with these alternative splice forms have been detected by Western blotting. *DISC1* transcripts are detected in all tissues tested (Millar et al. 2000b), and similarly, DISC1 protein shows a widespread pattern of expression.

In silico analysis was performed on the protein sequence of the human L DISC1 isoform (Millar et al. 2000b; Taylor et al. 2003). DISC1 consists of amino- and carboxy-terminal domains (Fig. 1, bottom). The two termini approximate to exon boundaries, with the first two exons encoding the amino terminus and the remainder of the gene encoding the carboxyl terminus. The two termini can also be distinguished on the basis of secondary structure prediction and levels of conservation between species. The amino terminus is made up of one or more globular domains, with two nuclear localization signals (NLS) (Ma et al. 2002; Taylor et al. 2003). The amino terminus shares no homologies with any known proteins. Conversely, the carboxyl terminus consists of α-helical and looped structures, interspersed with regions of coiled-coil-forming potential. Similarities exist between the carboxyl terminus and structural proteins or proteins involved in transport and motility (Millar et al. 2000b). However, the similarities are within the coiled-coil regions and are unlikely to be functionally relevant. Due to the modular structure of the coiled coils, these regions are anticipated to represent the protein interaction domains of DISC1 (Taylor et al. 2003). In addition, three leucine zippers are present within the carboxyl terminus (Ma et al. 2002) and may also have a role in mediating DISC1 protein interactions (Fig. 1, bottom).

DISC1 orthologs have been identified in mouse, rat, and fish species (Ma et al. 2002; Ozeki et al. 2003; Taylor et al. 2003). The overall gene structure is maintained, and there is evidence for alternative splicing in mouse, rat, and pufferfish. DISC1 is poorly conserved across species, with the amino terminus being less well conserved than the carboxyl terminus. The amino terminus shows 52% identity and 63% similarity, and the carboxyl terminus 61% identity and 78% similarity, between human and mouse (Taylor et al. 2003). The amino and carboxyl termini are conserved, as are the amino-terminal NLS and carboxy-terminal coiled-coil regions, suggesting that these are functionally important.

Preliminary evidence for protein–protein interactions comes from Ozeki et al. (2003), who reported the results of yeast two-hybrid studies, using full-length and truncated DISC1 protein as bait. They identified NudE-like (NUDEL), a cytoskeletal protein expressed in the cortex, as a strong interactor with DISC1. Our own results support the NUDEL interaction and identify several other potential interactors of known neuronal function (Millar et al. 2003a).

To further investigate the function of DISC1, we have raised antibodies to both the amino and carboxyl termini of the protein. The subcellular distribution of DISC1 is cell-type-specific and most likely reflects the cell shape and concomitant organization of the cytoskeleton (data not shown). DISC1 is predominantly, but not exclusively, expressed in the mitochondrion (Ozeki et al. 2003; James et al. 2004). In differentiated neuroblastoma cells, DISC1 redistributes to the shafts and tips of developing neurites, suggesting potential involvement of DISC1 in neurite outgrowth (Fig. 3).

Other Cytogenetic Rearrangements Associated with Psychosis

The potential opportunities provided by rare chromosome aberrations, such as those used to pinpoint single-gene disorders in the first phase of the Human Genome Project (Collins 1992) and as described here for SCZ in the t(1;11) family, have led to the widespread use of karyotype screening of psychiatric patients. Many karyotypic rearrangements have been defined at the visual, "Giemsa-band" level, but relatively few as yet at the molecular level (for review, see MacIntyre et al. 2003). We have extended our studies to 12 additional cases and families in whom chromosome rearrangements are associated with diagnoses ranging from SCZ to BPAD as well as "comorbid" diagnoses where SCZ is coupled with learning disability (US: mental retardation) (B.S. Pickard et al., in prep.). Nearly all of the constituent chromosomal breakpoints from each of these abnormalities have now been mapped and their genomic environments searched for

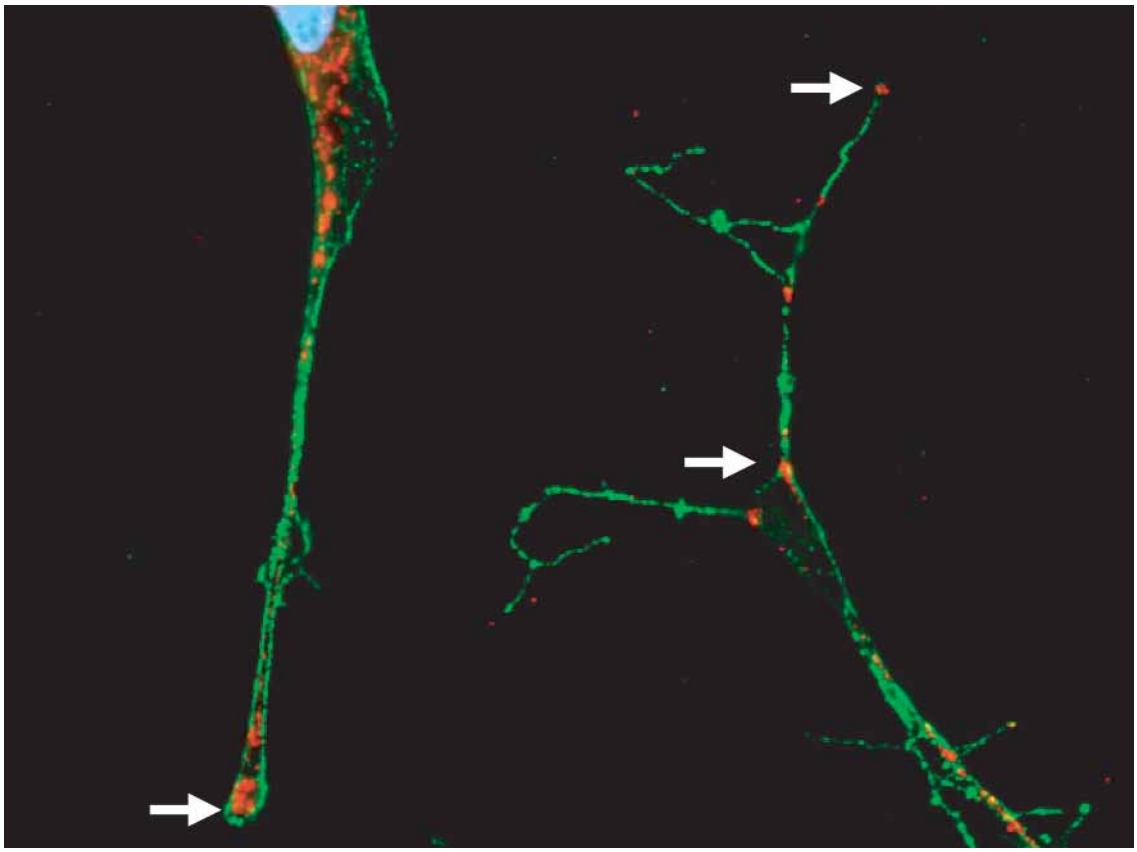


Figure 3. DISC1 protein expression in differentiated neuroblastoma-derived cells. DISC1 anti-peptide antibody staining (*red*) with anti-actin staining (*green*) and DAPI-stained nucleus (*blue*) shows that DISC1 protein is concentrated at the branch points (*arrowed*) and growing neurite tips (*arrowed*) of neuroblastoma-derived SH-SY5Y cells induced to differentiate in the presence of retinoic acid.

candidate genes. Emerging data from the Human Genome Project have been instrumental in this process in two ways. First, the physical map of BAC and PAC clones has provided cytogenetically cross-referenced molecular probes for fluorescence in situ hybridization (FISH) to patient chromosomes, positioning the breakpoints to ~200-kb windows (Fig. 4). Second, the collation, positioning, and annotation of expressed sequence tags (ESTs) and defined gene transcripts have allowed the rapid identification of potential candidate genes. Of 11 abnormalities where breakpoints have been experimentally defined, 8 directly disrupt or are predicted to positively influence at least one gene. The remaining 3 have breakpoints positioned where perturbations of nearby genes could be postulated. One such gene, *NPAS3*, encodes a transcription factor expressed in the central nervous system (Kamnasaran et al. 2003; B.S. Pickard et al., in prep.). Intriguingly, its close homolog, *NPAS2*, has been implicated in synaptic long-term potentiation (LTP) and the cellular energy-state-dependent modification of circadian rhythms (Garcia et al. 2000; Reick et al. 2001; Rutter et al. 2001; Dioum et al. 2002). Recently, *DIBD1* (for *d*isrupted *i*n *b*ipolar *d*isorder 1), encoding a component of the pathway responsible for adding carbohydrate moieties to membrane-bound and secreted proteins, has been cloned by a similar strategy from a family where BPAD was prevalent (Baysal et al. 2002). These suc-

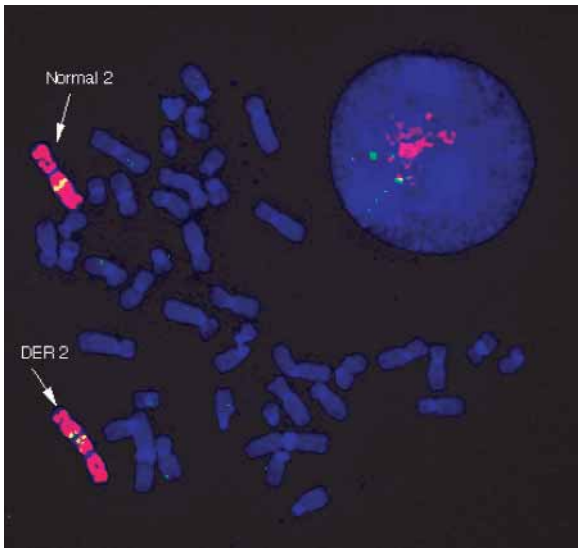


Figure 4. FISH of a metaphase chromosome spread from a patient with SCZ and mild learning disability (US: mental retardation). The red signal is obtained from a Chromosome-2 "paint" probe mix. The gap in the long arm of one Chromosome 2 represents the insertion of a portion of the short arm of Chromosome 5. This is highlighted by the yellow yeast artificial chromosome (YAC) probe which spans the point of insertion on Chromosome 2. Placing this YAC onto the Human Genome Project BAC/PAC physical map led to the identification of a Chromosome-2 gene disrupted by the insertion event.

cesses have been mirrored in the study of other complex neurological/psychiatric conditions such as autism (Vincent et al. 2000; Tentler et al. 2001; Sultana et al. 2002) and speech and language development (Fisher et al. 1998; Lai et al. 2001), where chromosome abnormalities have pinpointed individual susceptibility genes. Functional studies of these and other candidate genes resulting from this approach are certain to shed light on the developmental and molecular pathways and processes that go awry in mental illness, as well as suggesting rational choices for future population-based association and mutation detection studies.

GENOME-WIDE STRATEGIES FOR MAPPING SUSCEPTIBILITY GENES: LINKAGE STUDIES IN MULTIPLEX FAMILIES

Genome-wide linkage studies in multiplex families have been criticized because of inconsistent replications in genome scans of affected sib pairs or large collections of small families. As discussed above, failure to replicate could be due to lack of power of existing sample sets to detect alleles of small effect, and/or to the presence of substantial locus heterogeneity. Thus, it remains to be determined where the balance lies between the models of Mendelian inheritance with genetic heterogeneity and a fully quantitative model. In BPAD, significant linkage has been reported in extended pedigrees on Chromosomes 1q, 4p, 4q, 12q, 18q, and 21q (Potash and DePaulo 2000). The original 4p report came from our genome-wide linkage study in a large Scottish pedigree (F22) that segregates major affective disorder (Blackwood et al. 1996). A whole-genome scan of F22 found significant linkage to Chromosome 4p16 (LOD score = 4.8). Subsequent to that report, a number of other groups have also found evidence for linkage of major psychiatric illness to this region. Detera-Wadleigh et al. (1999) carried out a genome-wide scan of 22 BPAD families, and their largest family (F48) generated a LOD of 3.24. Asherson et al. (1998) found linkage (LOD = 1.96) in schizoaffective family CF50. Ewald et al. (1998) reported linkage in BPAD families (LOD = 2.0). Williams et al. (1999) found increased allele sharing (LOD = 1.73) in SCZ. Polymeropoulos and Schaffer (1996) described LODs between 1 and 2 in a BPAD family. In our follow-up studies of 57 Scottish families, we found evidence for linkage in a second Scottish BPAD family (F59, LOD = 1.15). The LOD score in this family comes very close to meeting the replication criteria proposed by Lander and Kruglyak (1995), but the maximum possible LOD score is limited by the small size of the family. It is not possible to estimate accurately the proportion of BPAD families linked to 4p, but a first approximation is 4%, based on the figures currently available; i.e., 2/58 in a Scottish sample, 1/24 in a Welsh sample (Asherson et al. 1998), and 1/22 in a US sample (Detera-Wadleigh et al. 1999).

The high LOD score of 4.8 generated by the linked markers in F22 indicates that the disease haplotype is very likely to contain a susceptibility locus for psychiatric illness. Further statistical evidence for this conclusion comes from variance component analysis of the same data, which found significant evidence (LOD = 3.7) for a BAPD locus in the region (Visscher et al. 1999). Family F48 is also of a size that by itself can generate a significant LOD score (Lander and Kruglyak 1995). The use of smaller families that do not have the ability to generate significant LOD scores is a higher risk strategy, due to the increased likelihood of false positives. However, if such families contain affected individuals with recombination events that are complementary to those of the larger families, this allows the division of the candidate region into subregions with different priority levels, as described for Chromosome 4p in Figure 5. The proximal and distal boundary of the candidate region is defined by the recombination breakpoints in F22 (see Fig. 5). Subregions have been prioritized for candidate gene analysis on the strength of evidence provided by the other linked families. On this basis, Minimal Region I is perhaps the most promising, being common to three out of four families (all of Celtic origin), followed by Minimal Region II (3/4, including F48 of Ashkenazi Jewish origin).

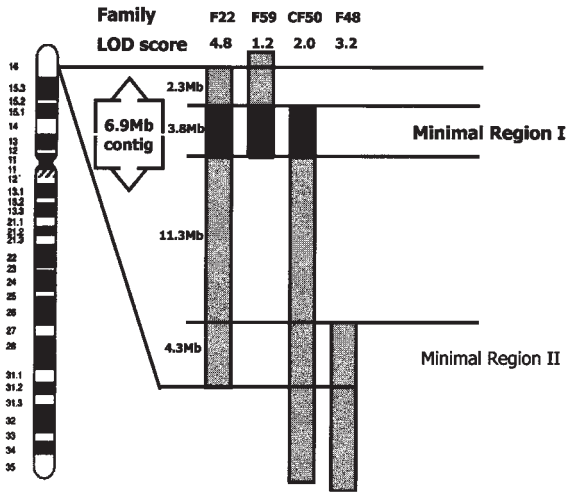


Figure 5. Genetic linkage to psychosis to Chromosome 4p16. The figure shows the LOD scores and a representation of the recombination intervals for each of four 4p16 linked families. Their diagnostic features are, respectively: F22 and F59, major affective disorder; F50, SCZ and schizoaffective disorder; F48, SCZ and major affective disorder. A BAC and PAC clone contig of 6.9 Mb with a single gap (estimated size 150–350 kb) covers all of Minimal Region I, plus proximal and distal sequence. A contig for the entire 20-Mbp region is curated in a custom version of AceDB.

As a platform for genome annotation and association mapping (collaboration with Dr. Mark Ross, Sanger Institute, UK), we combined large-insert genomic library screening with computational analysis to construct and curate in ACeDB (Eeckman and Durbin 1995) and SAM (Soderlund et al. 1997) a 6.9-Mb contig (consisting of 460 overlapping BAC and PAC clones) that encompasses and extends proximal and distal to Minimal Region I (Evans et al. 2001b). Advances in the Human Genome Project mean that these can now be largely constructed in silico. However, it was our experience that successive builds of the human genome (http://genome.ucsc.edu/

cgi-bin/hgGateway) resulted in quite major changes in the public domain view of the region, which were inconsistent with the physical map built "in-house" by detailed STS mapping. A small gap in the contig, estimated at 150–350 kb, still remains to be convincingly filled. This region is repeat-dense and underrepresented in YAC, PAC, and BAC libraries (K.L. Evans et al., unpubl.).

A clone contig of the region facilitates the effective use of genomic sequence to construct a transcript map and to conduct comparative genetic analyses that focus on the search for DNA sequence variants within coding or regulatory regions (Semple et al. 2002). We have set up our ACeDB database to display the automated results of exon and gene prediction programs and BLAST searches of EST databases and genomic sequence from human and other organisms, assisting recognition of both regulatory and coding regions, as well as prediction of biological function. The assembled evidence is being used to direct laboratory work to confirm these annotations and generate reagents for biological analyses, including accessing multiple, full-length cDNA libraries (in collaboration with Dr. Kate Rice and colleagues, Sanger Institute, UK). The physical and transcript maps form the basis for SNP discovery and association studies. The rest of the shared haplotype region has been built in silico from the Golden Path (http://genome.ucsc.edu/cgi-bin/hgGateway) and by in-house annotation and experimentation.

The most parsimonious explanation for the convergence of linkage evidence on Minimal Region I in the families of Celtic origin is a founder mutation. If there is a founder mutation common to more than one family, this will be flanked by a region of haplotype sharing that may vary in extent between families, but will include a functional variant that is common to and specific to all cases. With this in mind, we have looked for evidence of allele sharing between families. We have preliminary evidence for allele sharing in Minimal Region I, but definitive evidence awaits a much higher density of marker analysis. We have drawn upon the public domain SNP databases (dbSNP at NCBI, http://www.ncbi.nlm.nih.gov/SNP and the SNP Consortium, http://snp.cshl.org), but in practice, a significant proportion are uninformative in our samples. We are therefore undertaking SNP discovery, focusing on coding and putative regulatory regions in affected and control chromosomes from the linked families, and testing these SNPs for allele sharing between families prior to large-scale association analysis (Le Hellard et al. 2002). To enhance our capacity for SNP discovery and the power to detect associations with rare SNP variants (Thomas et al. 2004), we have adapted a somatic cell hybrid approach (Douglas et al. 2001) to derive haploid reagents for key probands and random cases.

A strength of our overall strategy is that the case and control samples, like F22 and F59, are drawn from the Scottish population, which has a low attrition rate (<2% per annum) and which, from HLA analyses, is thought to be one of the most genetically homogeneous populations in Europe (Cavalli-Sforza and Piazza 1993). This minimizes the possibility of admixture and increases the possibility of detecting a true association. The association study is being carried out in two phases. The first phase involves screening a subset of the sample in the form of parent offspring trios to allow selection of SNPs that represent haplotype blocks. The construction of this preliminary LD map also allows an examination of the extent to which all haplotype blocks in the Scottish population have been covered. Within haplotype blocks, SNP selection for phase 2 favors those that show preliminary evidence for association. In the second phase, association analysis is carried out on a much larger set of unrelated cases and controls, allowing replication of the first data set with high power to detect significant association. This two-stage strategy has a number of advantages: Replication is important in ruling out false-positive association results; the need for correction for multiple testing is reduced, as only a subset of markers are genotyped on the second set, and savings of both DNA and money are made. Our ACeDB database has been extended to facilitate the management and analysis of the data generated by association studies.

Once an associated region has been identified, candidate variants are prioritized by their predicted effect on protein function and the extent to which they are shared by linked families. We have invested in establishing lymphoblastoid cell lines for a large number of patients so that we have the capacity to undertake proteomic studies. If the region identified contains large numbers of variants, it may be necessary also to look for evidence for association with illness in other populations, particularly in populations that display lower levels of linkage disequilibrium, thus decreasing the number of associated variants. Prioritization of the variants will be followed by (1) assessment of individual and population attributable risk, (2) identification and analysis of gene(s) function, (3) study of how function might be affected by the variant(s), and (4) identification of biochemical pathways, which should also identify targets for the development of novel treatments and additional candidates for further genetic study.

CONCLUSIONS

There have been several false dawns in the early history of psychiatric genetics, but there is growing confidence, backed by persuasive statistical findings and biological associations, that there is at last cause for cautious optimism (Botstein and Risch 2003; Merikangas and Risch 2003). The recent evidence for linkage and association to the neuregulin (NRG1) gene by Stefansson and colleagues (2002, 2003) at deCODE Genetics both supports our position on the value of genome-wide studies of multiplex families and emphasizes the need for functional evaluation of gene discoveries.

Our own results provide clear statistical evidence for a BPAD locus of major effect in family F22 and persuasive evidence for replication in other families. The task of defining the locus in molecular terms is still challenging. It remains to be seen whether this locus represents a rare or common variant in the Scottish population and, indeed, other populations. If there is a founder effect, then alleles

sharing between (distantly related) families narrow the locus. On the other hand, allelic heterogeneity, as seen often in Mendelian disorders, may be valuable for confirming a candidate gene but poses additional, statistical challenges for discovery (Pritchard 2001). Nevertheless, within the overall framework of the Human Genome Project (Lander et al. 2001) and the steady impact of emergent genomic technologies, we can be confident that this is now a tractable problem.

The cytogenetic approach, which was so successful in the early stages of the Human Genome Project in pinpointing single-gene disorders, has, in our opinion, received insufficient attention as a strategy for dissecting complex genetic disorders. This is all the more surprising, given the fact that somatic rearrangement has been the single most productive approach to gene discovery in cancer and that phenotypically complex disorders of the nervous system such as Down syndrome, neurofibromatosis, AD, autism, and specific language disorder have all benefited from a cytogenetic approach. The explanation may reside in part in a passionate debate over the underlying genetic model for SCZ and for BPAD. In the absence of empirical evidence, however, it is unwise to settle on a preferred model or oblige the data to fit that model. Our empirical evidence from molecular cloning of a balanced t(1;11) breakpoint in a multigeneration family with a high loading of psychosis points to DISC1 as a very promising candidate gene. This finding, made through molecular cytogenetics, has been replicated by conventional family- and population-based studies in both Finland and Scotland. Bioinformatic analysis of DISC1 identified potential structural and interaction domains, but few clues as to function (Taylor et al. 2003). Protein–protein interaction studies (Millar et al. 2003a; Ozeki et al. 2003) will be important in establishing the biological pathway(s) in which DISC1 acts. Immunocytochemical studies point to a possible role in neurite outgrowth (Fig. 4) (James et al. 2004). The emergence of a growing list of equally interesting genes identified by molecular cytogenetic studies in other families further vindicates the general approach. Combining the availability of a BAC contig with high-resolution fluorescence in situ hybridization techniques to analyze metaphase chromosomes is a low-tech solution to genome-wide screens that, on our evidence, are likely to bring valuable benefits to other researchers and, we conjecture, are entirely applicable to other common, complex genetic disorders.

Figure 6 summarizes our broad hopes for where this research will eventually lead. We aim through family- and population-based studies to be able to nominate candidate genes for biological study and to feed a rational drug-development pipeline. The gene variants so discovered will also be valuable in the near term as molecular diagnostics, for disease surveillance, for monitoring response to treatment, and for treatment choice. Low compliance and adverse drug reaction are major problems in psychiatry. This is one specific and important area where the successful application of pharmacogenetic principles (Roses 2000) would be of immense value. If the gene-to-drug pathway is indeed successful, this opens the possibility for tailored drug prescription of rational medicines. The ethical, legal, and social issues that are fundamental to all applications of the new genetics are nowhere more sensitive than in relation to behavioral disorders. There is, in our view, a realistic possibility that genetic research will lead in time to completely novel insights into these perplexing and distressing disorders and that may in turn lead to completely novel therapeutic strategies which take account of both genetic *and* environmental risk factors. Scientifically sound and ethical application of this promised knowledge to the benefit of the many who suffer from mental illness is a major challenge in itself, but is the ultimate goal that drives the research.

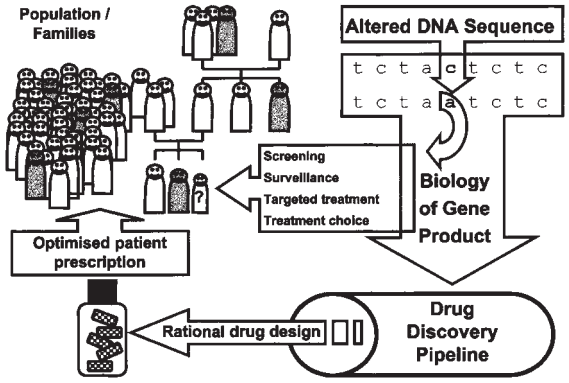


Figure 6. Putting it all together. Summary of the overall research objectives. Family- and population-based studies are used to nominate candidate genes for biological study and rational drug development. Gene variants can be used as molecular diagnostics, for disease surveillance, for monitoring response to treatment, and for treatment choice. If the gene-to-drug pathway is successful, this opens the possibility of tailored drug prescription of evidence-based medicines.

ACKNOWLEDGMENTS

This work was supported in part by the UK Medical Research Council, the Caledonian Research Foundation, the Chief Scientists Office in Scotland, the Scottish Hospitals Research Endowment Trust, the Stanley Medical Research Institute, the Cunningham Trust, the PPP Foundation, Organon NV, and Merck Sharpe Dohme. We acknowledge Professor Mike Owen for access to F50 and Dr. Sevilla Detera-Wadleigh for access to F48. We acknowledge Dr. Ole Mors and Dr. Henrik Ewald, Aarhus, Denmark, for access to materials from patients with suspected cytogenetic rearrangements. We thank the Genetics Core at the Wellcome Trust Clinical Research Facility, Edinburgh (www.wtcrf.med.ed.ac.uk), for genotyping support and Drs. David Bentley, Mark Ross, and Kate Rice, Sanger Institute, UK, for help with establishing ACeDB, SAM, and access to genomic and cDNA libraries. We thank colleagues at the Medical Genetics Section and Division of Psychiatry, Edinburgh University, for support and critical comment and the many families and individuals who have helped this research.

REFERENCES

Asherson P., Mant R., Williams N., Cardno A., Jones L., Murphy K., Collier D.A., Nanko S., Craddock N., Morris S., Muir W., Blackwood B., McGuffin P., and Owen M.J. 1998. A study of chromosome 4p markers and dopamine D5 receptor gene in schizophrenia and bipolar disorder. *Mol. Psychiatry* **3:** 310.

Baysal B.E., Willett-Brozick J.E., Badner J.A., Corona W., Ferrell R.E., Nimgaonkar V.L., and Detera-Wadleigh S.D. 2002. A mannosyltransferase gene at 11q23 is disrupted by a translocation breakpoint that co-segregates with bipolar affective disorder in a small family. *Neurogenetics* **4:** 43.

Berrettini W.H. 2000. Susceptibility loci for bipolar disorder: Overlap with inherited vulnerability to schizophrenia. *Biol. Psychiatry* **47:** 245.

Bird l. 1999. *The fundamental facts: All the latest facts and figures on mental illness*. The Mental Health Foundation, London, United Kingdom.

Blackwood D.H., Fordyce A., Walker M.T., St Clair D.M., Porteous D.J., and Muir W.J. 2001. Schizophrenia and affective disorders—Cosegregation with a translocation at chromosome 1q42 that directly disrupts brain-expressed genes: Clinical and P300 findings in a family. *Am. J. Hum. Genet*. **69:** 428.

Blackwood D.H., He L., Morris S.W., McLean A., Whitton C., Thomson M., Walker M.T., Woodburn K., Sharp C.M., Wright A.F., Shibasaki Y., St Clair D.M., Porteous D.J., and Muir W.J. 1996. A locus for bipolar affective disorder on chromosome 4p. *Nat. Genet.* **12:** 427.

Botstein D. and Risch N. 2003. Discovering genotypes underlying human phenotypes: Past successes for Mendelian disease, future approaches for complex disease. *Nat. Genet*. **33:** 228.

Cannon T.D., Huttunen M.O., Lonnqvist J., Tuulio-Henriksson A., Pirkola T., Glahn D., Finkelstein J., Hietanen M., Kaprio J., and Koskenvuo M. 2000. The inheritance of neuropsychological dysfunction in twins discordant for schizophrenia. *Am. J. Hum. Genet*. **67:** 369.

Cavalli-Sforza L.L. and Piazza A. 1993. Human genomic diversity in Europe: A summary of recent research and prospects for the future. *Eur. J. Hum. Genet.* **1:** 3.

Chen G., Chen K.S., Knox J., Inglis J., Bernard A., Martin S.J., Justice A., McConlogue L., Games D., Freedman S.B., and Morris R.G. 2000. A learning deficit related to age and beta-amyloid plaques in a mouse model of Alzheimer's disease. *Nature* **408:** 975.

Clayton D. and McKeigue P.M. 2001. Epidemiological methods for studying genes and environmental factors in complex diseases. *Lancet* **358:** 1356.

Collins F.S. 1992. Positional cloning: Let's not call it reverse anymore. *Nat. Genet.* **1:** 3.

Detera-Wadleigh S.D., Badner J.A., Berrettini W.H., Yoshikawa T., Goldin L.R., Turner G., Rollins D.Y., Moses T., Sanders A.R., Karkera J.D., Esterling L.E., Zeng J., Ferraro T.N., Guroff J.J., Kazuba D., Maxwell M.E., Nurnberger J.I., Jr., and Gershon E.S. 1999. A high-density genome scan detects evidence for a bipolar-disorder susceptibility locus on 13q32 and other potential loci on 1q32 and 18p11.2. *Proc. Natl. Acad. Sci.* **96:** 5604.

Devon R.S., Anderson S., Teague P.W., Burgess P., Kipari T.M., Semple C.A., Millar J.K., Muir W.J., Murray V., Pelosi A.J., Blackwood D.H., and Porteous D.J. 2001. Identification of polymorphisms within Disrupted in Schizophrenia 1 and Disrupted in Schizophrenia 2, and an investigation of their association with schizophrenia and bipolar affective disorder. *Psychiatr. Genet*. **11:** 71.

Dioum E.M., Rutter J., Tuckerman J.R., Gonzalez G., Gilles-Gonzalez M.A., and McKnight S.L. 2002. NPAS2: A gas-responsive transcription factor. *Science* **298:** 2385.

Douglas J.A., Boehnke M., Gillanders E., Trent J.M., and Gruber S.B. 2001. Experimentally-derived haplotypes substantially increase the efficiency of linkage disequilibrium studies. *Nat. Genet*. **28:** 361.

Eeckman F.H. and Durbin R. 1995. ACeDB and macace. *Methods Cell Biol.* **48:** 583.

Ekelund J., Lichtermann D., Hovatta I., Ellonen P., Suvisaari J., Terwilliger J.D., Juvonen H., Varilo T., Arajarvi R., Kokko-Sahin M.L., Lonnqvist J., and Peltonen L. 2000. Genome-wide scan for schizophrenia in the Finnish population: Evidence for a locus on chromosome 7q22. *Hum. Mol. Genet*. **9:** 1049.

Ekelund J., Hovatta I., Parker A., Paunio T., Varilo T., Martin R., Suhonen J., Ellonen P., Chan G., Sinsheimer J.S., Sobel E., Juvonen H., Arajarvi R., Partonen T., Suvisaari J., Lonnqvist J., Meyer J., and Peltonen L. 2001. Chromosome 1 loci in Finnish schizophrenia families. *Hum. Mol. Genet*. **10:** 1611.

Evans K.L., Muir W.J., Blackwood D.H.R., and Porteous D.J. 2001a. Nuts and bolts of psychiatric genetics: Building on the Human Genome Project. *Trends Genet*. **17:** 35.

Evans K.L., Le Hellard S., Morris S.W., Lawson D., Whitton C., Semple C.A., Fantes J.A., Torrance H.S., Malloy M.P., Maule J.C., Humphray S.J., Ross M.T., Bentley D.R., Muir W.J., Blackwood D.H., and Porteous D.J. 2001b. A 6.9-Mb high-resolution BAC/PAC contig of human 4p15.3-p16.1, a candidate region for bipolar affective disorder. *Genomics* **71:** 315.

Ewald H., Degn B., Mors O., and Kruse T.A. 1998. Support for the possible locus on chromosome 4p16 for bipolar affective disorder. *Mol. Psychiatry* **3:** 442.

Fisher S.E., Vargha-Khadem F., Watkins K.E., Monaco A.P., and Pembrey M.E. 1998. Localisation of a gene implicated in a severe speech and language disorder. *Nat. Genet.* **18:** 168.

Gabriel S.B., Schaffner S.F., Nguyen H., Moore J.M., Roy J., Blumenstiel B., Higgins J., DeFelice M., Lochner A., Faggart M., Liu-Cordero S.N., Rotimi C., Adeyemo A., Cooper R., Ward R., Lander E.S., Daly M.J., and Altshuler D. 2002. The structure of haplotype blocks in the human genome. *Science* **296:** 2225.

Garcia J.A., Zhang D., Estill S.J., Michnoff C., Rutter J., Reick M., Scott K., Diaz-Arrastia R., and McKnight S.L. 2000. Impaired cued and contextual memory in *Npas2*-deficient mice. *Science* **288:** 2226.

Gasperoni T.L., Ekelund J., Huttunen M., Palmer C.G., Tuulio-Henriksson A., Lonnqvist J., Kaprio J., Peltonen L., and Cannon T.D. 2003. Genetic linkage and association between chromosome 1q and working memory function in schizophrenia. *Am. J. Med. Genet*. **116B:** 8.

Gejman P.V., Martinez M., Cao Q., Friedman E., Berrettini W.H., Goldin L.R., Koroulakis P., Ames C., Lerman M.A., and Gershon E.S. 1993. Linkage analysis of fifty-seven microsatellite loci to bipolar disorder. *Neuropsychopharmacology* **9:** 31.

Glahn D.C., Therman S., Manninen M., Huttunen M., Kaprio J., Lonnqvist J., and Cannon T.D. 2003. Spatial working memory as an endophenotype for schizophrenia. *Biol. Psychiatry* **53:** 624.

Hardy I. and Israel A. 1999. Alzheimer's disease: In search of gamma-secretase. *Nature* **398:** 466.

Hovatta I., Varilo T., Suvisaari J., Terwilliger J.D., Ollikainen V., Arajarvi R., Juvonen H., Kokko-Sahin M.L., Vaisanen L., Mannila H., Lonnqvist J., and Peltonen L. 1999. A genomewide screen for schizophrenia genes in an isolated Finnish subpopulation, suggesting multiple susceptibility loci. *Am. J. Hum. Genet*. **65:** 1114.

Hwu H.G., Liu C.M., Fann C.S., Ou-Yang W.C., and Lee S.F. 2003. Linkage of schizophrenia with chromosome 1q loci in Taiwanese families. *Mol. Psychiatry* **8:** 445.

James R., Adams R.R., Christie S., Buchanan D., Porteous D.J., and Millar J.K. 2004. DISC1 is a multicompartmentalized protein that predominantly localises to mitochondia. *Mol. Cell. Neurosci*. (in press).

Janus C., Pearson J., McLaurin J., Mathews P.M., Jiang Y., Schmidt S.D., Chishti M.A., Horne P., Heslin D., French J., Mount H.T., Nixon R.A., Mercken M., Bergeron C., Fraser P.E., St George-Hyslop P., and Westaway D. 2000. A beta peptide immunization reduces behavioural impairment and plaques in a model of Alzheimer's disease. *Nature* **408:** 979.

Kamnasaran D., Muir W.J., Ferguson-Smith M.A., and Cox D.W. 2003 Disruption of the neuronal PAS3 gene in a family affected with schizophrenia. *J. Med. Genet*. **40:** 325.

Kauppi L., Sajantila A., and Jeffreys A.J. 2003. Recombination hotspots rather than population history dominate linkage disequilibrium in the MHC class II region. *Hum. Mol. Genet.* **12:** 33.

Kendler K.S. and Diehl S.R. 1993. The genetics of schizophrenia: A current, genetic-epidemiologic perspective. *Schizophr. Bull.* **19:** 261.

Kendler K.S., Karkowski L.M., and Walsh D. 1998. The structure of psychosis: Latent class analysis of probands from the Roscommon Family Study. *Arch. Gen. Psychiatry* **55:** 492.

Klar A.J. 2002. The chromosome 1;11 translocation provides the best evidence supporting genetic etiology for schizophrenia and bipolar affective disorders. *Genetics* **160:** 1745.

Lai C.S., Fisher S.E., Hurst J.A., Vargha-Khadem F., and Monaco A.P. 2001. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* **413:** 519.

Lander E. and Kruglyak L. 1995. Genetic dissection of complex traits: Guidelines for interpreting and reporting linkage results. *Nat. Genet.* **11:** 241.

Lander E.S., Linton L.M., Birren B., Nusbaum C., Zody M.C., Baldwin J., Devon K., Dewar K., Doyle M., FitzHugh W., Funke R., Gage D., Harris K., Heaford A., Howland J., Kann L., Lehoczky J., LeVine R., McEwan P., McKernan K., Meldrim J., Mesirov J.P., Miranda C., Morris W., and Naylor J., et al. (International Human Genome Sequencing Consortium). 2001. Initial sequencing and analysis of the human genome. *Nature* **409:** 860.

Le Hellard S., Ballereau S.J., Visscher P.M., Torrance H.S., Pinson J., Morris S.W., Thomson M.T., Semple C.A.M., Muir W.J., Blackwood D.H.R., Porteous D.J., and Evans K.L. 2002. SNP genotyping on pooled DNAs: Comparison of genotyping technologies and a semi automated method for data storage. *Nucleic Acids Res*. **30:** e74.

Levinson D.F., Holmans P.A., Laurent C., Riley B., Pulver A.E., Gejman P.V., Schwab S.G., Williams N.M., Owen M.J., Wildenauer D.B., Sanders A.R., Nestadt G., Mowry B.J., Wormley B., Bauche S., Soubigou S., Ribble R., Nertney D.A., Liang K.Y., Martinolich L., Maier W., Norton N., Williams H., Albus M., and Carpenter E.B., et al. 2002. No major schizophrenia locus detected on chromosome 1q in a large multicenter sample. *Science* **296:** 739.

Lopez A.D. and Murray C.C. 1998. The global burden of disease. *Nat. Med.* **4:** 1241.

Ma L., Liu Y., Ky B., Shughrue P.J., Austin C.P., and Morris J.A. 2002. Cloning and characterization of *Disc1*, the mouse ortholog of DISC1 (Disrupted-in-Schizophrenia 1). *Genomics* **80:** 662.

MacGregor S., Visscher P.M., Knott S., Porteous D., Muir W., Millar K., and Blackwood D. 2002. Is schizophrenia linked to chromosome 1q? *Science* **298:** 2277.

MacIntyre D.J., Blackwood D.H., Porteous D.J., Pickard B.S., and Muir W.J. 2003. Chromosomal abnormalities and mental illness. *Mol. Psychiatry* **8:** 275.

Merikangas K.R. and Risch N. 2003. Will the genomics revolution revolutionize psychiatry? *Am. J. Psychiatry* **160:** 625.

Millar J.K., Christie S., and Porteous D.J. 2003a. Yeast two-hybrid screens implicate DISC1 in brain development and function. *Biochem. Biophs. Res. Commun*. **311:** 1019.

Millar J.K., Christie S., Semple C.A., and Porteous D.J. 2000a. Chromosomal location and genomic structure of the human translin-associated factor X gene (TRAX; TSNAX) revealed by intergenic splicing to DISC1, a gene disrupted by a translocation segregating with schizophrenia. *Genomics* **67:** 69.

Millar J.K., Thomson P.A., Wray N.R., Muir W.J., Blackwood D.H., and Porteous D.J. 2003b. Response to Amar J. Klar: The chromosome 1;11 translocation provides the best evidence supporting genetic etiology for schizophrenia and bipolar affective disorders. *Genetics* **163**: 833.

Millar J.K., Wilson-Annan J.C., Anderson S., Christie S., Taylor M.S., Semple C.A., Devon R.S., Clair D.M., Muir W.J., Blackwood D.H., and Porteous D.J. 2000b. Disruption of two novel genes by a translocation co-segregating with schizophrenia. *Hum. Mol. Genet*. **9:** 1415.

Muir W.J., Gosden C.M., Brookes A.J., Fantes J., Evans K.L., Maguire S.M., Stevenson B., Boyle S., Blackwood D.H., St Clair D.M., Porteous D.J., and Weith A. 1995. Direct microdissection and microcloning of a translocation breakpoint region, t(1;11)(q42.2;q21), associated with schizophrenia. *Cytogenet. Cell Genet*. **70:** 35.

Owen M.J., Cardno A.G., and O'Donovan M.C. 2000. Psychiatric genetics: Back to the future. *Mol. Psychiatry* **5:** 22.

Ozeki Y., Tomoda T., Kleiderlein J., Kamiya A., Bord L., Fujii K., Okawa M., Yamada N., Hatten M.E., Snyder S.H., Ross C.A., and Sawa A. 2003. Disrupted-in-Schizophrenia-1 (DISC-1): Mutant truncation prevents binding to NudE-like (NUDEL) and inhibits neurite outgrowth. *Proc. Natl. Acad. Sci.* **100:** 289.

Polymeropoulos M.H. and Schaffer A.A. 1996. Scanning the genome with 1772 microsatellite markers in search of a bipolar disorder susceptibility gene. *Mol. Psychiatry* **1:** 404.

Potash J.B. and DePaulo J.R., Jr. 2000. Searching high and low: A review of the genetics of bipolar disorder. *Bipolar Disord.* **2:** 8.

Pritchard J.K. 2001. Are rare variants responsible for susceptibility to complex diseases? *Am. J. Hum. Genet.* **69:** 124.

Reick M., Garcia J.A., Dudley C., and McKnight S.L. 2001. NPAS2: An analog of clock operative in the mammalian forebrain. *Science* **293:** 506.

Riley B.P. and McGuffin P. 2000. Linkage and associated studies of schizophrenia. *Am. J. Med. Genet.* **97:** 23.

Roses A. 2000. Pharmacogenetics and the practice of medicine. *Nature* **405:** 857.

Rutter J., Reick M., Wu L.C., and McKnight S.L. 2001. Regulation of clock and NPAS2 DNA binding by the redox state of NAD cofactors. *Science* **293:** 510.

Semple C.A.M., Morris S.W., Porteous D.J., and Evans K.L. 2002. Computational comparison of human genomic sequence assemblies for a region of chromosome 4. *Genome Res*. **12:** 424.

Soderlund C., Longden I., and Mott R. 1997. FPC: A system for building contigs from restriction fingerprinted clones. *CABIOS* **13:** 523.

St Clair D., Blackwood D.H., Muir W.J., Carothers A., Walker M., Spowart G., Gosden C., and Evans H.J. 1990. Association within a family of a balanced autosomal translocation with major mental illness. *Lancet* **336:** 13.

Stefansson H., Sarginson J., Kong A., Yates P., Steinthorsdottir V., Gudfinnsson E., Gunnarsdottir S., Walker N., Petursson H., Crombie C., Ingason A., Gulcher J.R., Stefansson K., and St Clair D. 2003. Association of neuregulin 1 with schizophrenia confirmed in a Scottish population. *Am. J. Hum. Genet*. **72:** 83.

Stefansson H., Sigurdsson E., Steinthorsdottir V., Bjornsdottir S., Sigmundsson T., Ghosh S., Brynjolfsson J., Gunnarsdottir S., Ivarsson O., Chou T.T., Hjaltason O., Birgisdottir B., Jonsson H., Gudnadottir V.G., Gudmundsdottir E., Bjornsson A., Ingvarsson B., Ingason A., Sigfusson S., Hardardottir H., Harvey R.P., Lai D., Zhou M., Brunner D., and Mutel V., et al. 2002. Neuregulin 1 and susceptibility to schizophrenia. *Am. J. Hum. Genet*. **71:** 877.

Sultana R., Yu C.E., Yu J., Munson J., Chen D., Hua W., Estes A., Cortes F., de la Barra F., Yu D., Haider S.T., Trask B.J., Green E.D., Raskind W.H., Disteche C.M., Wijsman E., Dawson G., Storm D.R., Schellenberg G.D., and Villacres E.C. 2002. Identification of a novel gene on chromosome 7q11.2 interrupted by a translocation breakpoint in a pair of autistic twins. *Genomics* **80:** 129.

Taylor M.S., Devon R.S., Millar J.K., and Porteous D.J. 2003. Evolutionary constraints on the Disrupted in Schizophrenia locus. *Genomics* **81:** 67.

Tentler D., Brandberg G., Betancur C., Gillberg C., Anneren G., Orsmark C., Green E.D., Carlsson B., and Dahl N. 2001. A balanced reciprocal translocation t(5;7)(q14;q32) associated with autistic disorder: Molecular analysis of the chromosome 7 breakpoint. *Am. J. Med. Genet*. **105:** 729.

Terwilliger J.D. and Weiss K.M. 1998. Linkage disequilibrium mapping of complex disease: Fantasy or reality? *Curr. Opin. Biotechnol.* **9:** 578.

Thomas S., Porteous D.J., and Visscher P.M. 2004. Power of direct vs. indirect haplotyping in association studies. *Genet. Epidemiol.* **26:** 116.

Valles V., Van Os J., Guillamat R., Gutierrez B., Campillo M., Gento P., and Fananas L. 2000. Increased morbid risk for schizophrenia in families of in-patients with bipolar illness. *Schizophr. Res*. **42:** 83.

Vincent J.B., Herbrick J.A., Gurling H.M., Bolton P.F., Roberts W., and Scherer S.W. 2000. Identification of a novel gene on chromosome 7q31 that is interrupted by a translocation breakpoint in an autistic individual. *Am. J. Hum. Genet*. **67:** 510.

Visscher P.M., Haley C.S., Heath S.C., Muir W.J., and Blackwood D.H. 1999. Detecting QTLs for uni- and bipolar disorder using a variance component method. *Psychiatr. Genet.* **9:** 75.

Wildenauer D.B., Schwab S.G., Maier W., and Detera-Wadleigh S.D. 1999. Do schizophrenia and affective disorder share susceptibility genes? *Schizophr. Res.* **39:** 107.

Williams N.M., Rees M.I., Holmans P., Norton N., Cardno A.G., Jones L.A., Murphy K.C., Sanders R.D., McCarthy G., Gray M.Y., Fenton I., McGuffin P., and Owen M.J. 1999. A two-stage genome scan for schizophrenia susceptibility genes in 196 affected sibling pairs. *Hum. Mol. Genet.* **9:** 1729.

WEB SITES

www.mentalhealth.org.uk
http://www.ncbi.nlm.nih.gov:80/entrez/dispomim.cgi?id=104300
http://genome.ucsc.edu/cgi-bin/hgGateway
http://www.ncbi.nlm.nih.gov/SNP
http://snp.cshl.org
www.wtcrf.med.ed.ac.uk