

Schizophrenia: a common disease caused by multiple rare alleles[†]

JON M. McCLELLAN, EZRA SUSSER and MARY-CLAIRE KING

Summary Schizophrenia is widely held to stem from the combined effects of multiple common polymorphisms, each with a small impact on disease risk. We suggest an alternative view: that schizophrenia is highly heterogeneous genetically and that many predisposing mutations are highly penetrant and individually rare, even specific to single cases or families. This ‘common disease – rare alleles’ hypothesis is supported by recent findings in human genomics and by allelic and locus heterogeneity for other complex traits. We review the implications of this model for gene discovery research in schizophrenia.

Declaration of interest None. Funding detailed in Acknowledgements.

Current research in the genetics of schizophrenia is guided primarily by the ‘common disease – common alleles’ model (Chakravarti, 1999). This model originated from the hypothesis that the illness results from the cumulative impact of multiple common small-effect, genetic variants, interacting with environmental exposures to exceed a biological threshold (Gottesman & Shields, 1982). The ‘common disease – common alleles’ model for schizophrenia is heuristically appealing. The illness is relatively frequent and is found worldwide. Thus common susceptibility alleles shared across populations are plausible. The ‘common disease – common alleles’ model has also been posited to explain the variable and inconsistent results of linkage studies devoted to finding genes of large effect responsible for schizophrenia, and the weak associations of various candidate genes with schizophrenia. Furthermore, mathematical

modelling suggested that the observed decline in recurrence risk of disease with increased genetic distance from affected individuals is inconsistent with monogenic inheritance of large-effect alleles (Risch, 1990). As pointed out by Risch, these models were based on the assumption that the illness was genetically homogeneous throughout the population for which the recurrence risks were calculated (Risch, 1990). Increasing evidence suggests that schizophrenia is genetically heterogeneous (Fanous & Kendler, 2005). If so, then recurrence risk data are also consistent with monogenic inheritance of large-effect alleles in a proportion of people with schizophrenia, with different alleles for different families.

We suggest that the ‘common disease – rare alleles’ model explains many cases of schizophrenia. Our hypothesis is that many mutations predisposing to schizophrenia are highly penetrant and individually rare, even specific to single patients or families. In this model, different families harbour different mutations, either in the same gene or in different genes, but any one family carries only one or two mutations. Many different disease-associated mutations may occur in the same gene.

The ‘common disease – common alleles’ and ‘common disease – rare alleles’ models are not mutually exclusive (Goldstein & Chikhi, 2002). Rare severe mutations may occur in genes that also harbour more common variants with modest effects on disease risk. However, the two models have distinctly different implications for gene-finding strategies. Most current psychiatric genetic research is designed to identify common alleles or haplotypes associated with increased risk of disease and shared by large numbers of patients compared with appropriate controls (Merikangas & Risch, 2003). If many cases of schizophrenia stem from individually rare large-effect alleles, current approaches – even if executed perfectly – will fail to identify critical genes.

We argue that current observations from epidemiology and genetics of schizophrenia are consistent with the influence of a large number of individually rare deleterious mutations, many of which have occurred in the present or recent generations. Several features of schizophrenia support this view:

- (a) Schizophrenia is familial, i.e. close relatives of affected persons are at increased risk of the illness. Occasional families are very severely affected (Gottesman & Shields, 1982), but most patients have no close affected relative. Taken together, these observations are consistent with a subset of families harbouring high-penetrance, recently occurring alleles predisposing to schizophrenia, with different alleles present in different families. Unless caused by detectable chromosomal alterations (as in the case of *DISC1*), such alleles have heretofore been difficult to find because individual families are not sufficiently informative for single-family linkage analyses.
- (b) Paternal age is consistently associated with increased risk of schizophrenia (Brown *et al*, 2002; Dalman & Allebeck, 2002; Malaspina *et al*, 2002; Byrne *et al*, 2003; El-Saadi *et al*, 2004; Sipos *et al*, 2004; Tsuchiya *et al*, 2005). Paternal age is also associated with increased rates of several types of *de novo* germline mutations (Crow, 2003).
- (c) The illness is associated with decreased fertility (Nimgaonkar, 1998; Haukka *et al*, 2003). If this had been the case over long periods, then the frequencies of any ancient common alleles associated with schizophrenia would be reduced. An ongoing contribution of new and therefore individually rare risk alleles could explain the persistence of the disorder.

To explain our reasoning, we first describe recent findings in genomics and genetics of other complex human disorders. Then we show that the results of schizophrenia research are consistent with the existence of multiple individually rare alleles of large effect. Finally, we consider the implications of this model for future schizophrenia research.

THE HUMAN GENOME

The dynamics of the human genome are proving more complex than anticipated, revealing more mechanisms by which

[†]See invited commentary, pp. 200–203, this issue

genetic changes may lead to human disease. Particularly relevant to our argument are the following observations (International Human Genome Sequencing Consortium, 2001a,b, 2004).

- (a) Only about 2% of the genome consists of protein-coding genes. There are approximately 20 000 protein-coding genes, far fewer than the 80 000–100 000 hypothesised a decade ago. However, the human proteome is enormously complex. At most genes, variable transcription leads to multiple transcripts, and thus multiple proteins, derived from the same locus but with different amino acid sequences. Variable transcription is frequently tissue-specific. As a result, the consequences of a mutation may also be tissue-specific.
- (b) Germ-line mutations occur more commonly than previously thought. Potentially deleterious new mutations may occur at a rate as high as three per zygote (Eyre-Walker & Keightley, 1999; Crow, 2000). Rates of occurrence of different classes of *de novo* mutations are differently influenced by parent of origin and parental age (Crow, 2000, 2003). The increased mutation rate associated with greater paternal age is particularly relevant, given that risk of schizophrenia is also associated with paternal age.
- (c) Epigenetic alterations – stable changes in gene expression that do not depend on changes in DNA sequence (Jaenisch & Bird, 2003) – may play an important part in human disease, including psychiatric disorders. Recent intriguing observations of possible epigenetic effects related to development include phenotypic variability in monozygotic twins and imprinting effects on neurodevelopmental disorders (Fraga *et al*, 2005; Wong *et al*, 2005).

GENETIC HETEROGENEITY IN HUMAN DISEASE

The genetic heterogeneity of complex illnesses is the natural result of the origins of human genetic variation. The oldest human alleles originated in Africa millions of years before people first migrated out of Africa some 50 000 years ago (Cavalli-Sforza *et al*, 1994). These ancient variants are found in all human populations, are the most common worldwide, and account for approximately 95% of human variation. Yet the exponential growth of the

human population has resulted in many new alleles, each individually rare and each specific to one population (or even one family). Most alleles are of this sort. Thus the paradox: most human variation is ancient and shared; most alleles are recent and individually rare. Given the size of the present human population and the rate of occurrence of new mutations, all mutations compatible with life have probably already occurred and will occur again. However, mutations with deleterious effects before or during the reproductive years will be less frequently transmitted to subsequent generations owing to their adverse impact on fertility or viability. Therefore, mutations with large effects on disease may be disproportionately of recent origin and individually very rare.

To the extent that any class of mutation – point mutations, copy number errors or abnormalities of chromosome number – occur spontaneously, they appear at similar rates in all human populations. All humans share the same basic genomic architecture, including the same genomic regions vulnerable to mutations. The incidence of schizophrenia does not appear to vary substantially across populations by virtue of genetic ancestry. This pattern is consistent with a disease due to multiple, independent *de novo* mutations that arise in many different vulnerable genes and genomic regions. Furthermore, environmental exposures with mutagenic consequences may lead to high rates of new mutations among exposed individuals. For example, environmental factors such as maternal starvation that are associated with disease (Cannon *et al*, 2003) may mediate their effects through *de novo* genetic or epigenetic mutations. We explore this theme in more detail below.

Complex illnesses are almost universally characterised by allelic heterogeneity (multiple different mutations in the same gene leading to disease) and locus heterogeneity (mutations in multiple different genes leading to the same disease) (Botstein & Risch, 2003; Goldstein *et al*, 2003). We propose that both are characteristic of schizophrenia. To understand the potential implications of genetic heterogeneity for schizophrenia, we briefly consider other complex disorders for which genes have been identified.

Deafness

To date, nearly a hundred genes have been identified that harbour inherited mutations

leading to hearing loss (Petit *et al*, 2001; Friedman & Griffith, 2003). All mutations are recent and all but one are individually rare. The one frequent mutation, 30delG in connexin 26, is the exception that proves the rule, in that the same mutation has occurred independently numerous times in a mutational hot-spot.

Epilepsy

The inherited forms of epilepsy are characterised by allelic and locus heterogeneity (Meisler *et al*, 2001). Mutations in any of several genes involved with neuronal signalling can lead to broadly defined epilepsy. Rare mutations in three different sodium channel genes lead to one more narrowly defined form of epilepsy (generalised epilepsy with febrile seizures plus).

Alzheimer's disease

Alzheimer's disease illustrates that the 'common disease – common allele' and 'common disease – rare allele' models need not be mutually exclusive. The common $\epsilon 4$ allele of *APOE* (apolipoprotein E) is associated with a threefold to fourfold increased risk in individuals of European descent of developing common, late-onset Alzheimer's disease (Bird, 2005). On the other hand, multiple rare mutations in genes encoding amyloid precursor protein (*APP*), presenilin 1 and 2 (*PS1* and *PS2*) and ubiquitin 1 (*UBQLN1*) are responsible for familial early-onset Alzheimer's disease. Therefore, both common modest-effect alleles and rare large-effect alleles have a role in Alzheimer's disease. The role of *APOE4* is an excellent example of the 'common disease – common allele' model. However, the effect of *APOE4* on Alzheimer's disease risk is substantially larger than the effect sizes of 2 or less that are typically estimated for schizophrenia susceptibility genes.

Inherited predisposition to cancer

In each of the two major genes for inherited breast and ovarian cancer, *BRCA1* and *BRCA2*, more than a thousand different pathogenic mutations have been found (Walsh *et al*, 2006). Large genomic rearrangements account for about 10% of these mutations. All inherited *BRCA1* and *BRCA2* mutations are individually rare. Both locus and allelic heterogeneity are also characteristic of inherited colon cancer and the rarer cancer syndromes (Vogelstein & Kinzler, 2004).

Lipid metabolic pathways

Rare variants in genes related to lipid metabolism are associated with low levels of high-density lipoprotein cholesterol (HDL-C) (Cohen *et al*, 2004) and low-density lipoprotein cholesterol (LDL-C) (Cohen *et al*, 2006). Although each variant is individually rare, in total these variants are found in a substantial portion of individuals at the far end of the spectrum in terms of levels of HDL-C or LDL-C respectively.

These examples illustrate two ways in which mutations of large effect are important for understanding human disease. First, the collective effect of individually rare mutations in the same gene may explain a considerable proportion of an illness. Second, rare mutations in genes of large effect can reveal pathways critical to disease development.

IMPLICATIONS OF GENETIC HETEROGENEITY FOR GENE DISCOVERY

All complex illnesses evaluated thus far are characterised by locus and allelic heterogeneity. Disease genes for these illnesses have been identified primarily by positional cloning in large kindreds. Although subsequent association studies confirmed their role, the original gene discoveries were dependent upon individual highly informative families. Such large informative kindreds are extremely rare in schizophrenia. Linkage studies of schizophrenia based on single gene models have not been successful at identifying causal mutations (Owen *et al*, 2004). Because pedigrees with schizophrenia have not been large enough to be individually informative, studies generally pool data from different families. If many different genes were responsible for the illness in different families, pooling results across families would preclude identification of any of them.

Currently, most gene-discovery strategies for schizophrenia research – case–parent triad studies, candidate gene studies and haplotype association studies – are designed to identify alleles or haplotypes that appear more frequently among affected individuals than among appropriate controls (Cannon *et al*, 2003). These designs are not robust to either allelic heterogeneity or locus heterogeneity. Sib-pair linkage analyses are designed to detect genomic regions consistently shared by affected siblings. Sib-pair analyses are robust to allelic

heterogeneity but not to locus heterogeneity. For each of these designs, hundreds or thousands of rigorously diagnosed cases of unrelated affected and unaffected individuals are examined. If a substantial portion of schizophrenia stems from different individually rare alleles, increasing the number of cases also increases the number of different disease risk mutations represented among them. As a result, increasing sample size does not confer a corresponding increase in statistical power. In the most extreme scenario, in which every case results from a different mutation, an increase in sample size would not lead to any increase in the power to detect any one mutation. Consequently, even very large studies may fail to detect individually rare disease risk mutations.

In addition, genetic analyses that focus only on single nucleotide polymorphisms (SNPs), either individually or in haplotypes, rather than fully sequenced DNA, will inevitably miss rare disease alleles and thus fail to detect critical genes harbouring such alleles. Association and linkage studies generally assume that individuals sharing the same SNP-defined haplotype share the entire region, including any hypothetical embedded disease alleles. This assumption is reasonable for ancient alleles and nearly always true for related individuals for whom there is direct inheritance of the haplotype. However, this assumption is not reasonable for a study of unrelated individuals who carry disease alleles of recent origin. Rare recent mutations causing schizophrenia within the same haplotype will differ among unrelated individuals, diluting any association.

TWO CANDIDATE GENES FOR SCHIZOPHRENIA: DYSBINDIN AND DISC1

The familial nature of schizophrenia is well established. Large collaborative linkage studies have suggested multiple candidate chromosomal regions that may harbour genes associated with the illness. Regions best supported by genome-wide scans include 6p22–p24 (Straub *et al*, 1995), 1q21–q22 (Brustowicz *et al*, 2000) and 13q32–q34 (Blouin *et al*, 1998). Other regions with positive linkage findings include 1q42, 5q21–q33, 6q21–q25, 8p21–p22, 10p15–p11 and 22q11–q12 (Owen *et al*, 2004). These regions combined represent a substantial portion of the genome.

Candidate genes have been suggested in several of these regions, including dysbindin on 6p22, neuregulin on 8p22, *G72* on 13q34, *COMT* on 22q11, *RGS4* on 1q21 and *GRM3* on 7q21 (see reviews by Blouin *et al*, 1998; Harrison & Weinberger, 2005). Each of these genes is biologically plausible (Owen *et al*, 2004). However, for each candidate gene, both positive and negative associations have been reported with the same SNPs; strengths of effects are generally weak; the specific allele or haplotype associated with the illness varies across studies; and definitive causative mutations have not been identified.

To illustrate the implications of the common allele *v.* rare allele models, we will review two promising susceptibility genes, dysbindin (*DTNBP1*) and *DISC1*. Very different study designs revealed these genes, with correspondingly different results to date. Linkage, association and functional studies all support some role for dysbindin in schizophrenia. Several studies involving different populations have found positive associations of dysbindin alleles or haplotypes with schizophrenia (Straub *et al*, 2002; Schwab *et al*, 2003; Funke *et al*, 2004; Kirov *et al*, 2004; Kohn *et al*, 2004; Numakawa *et al*, 2004; Williams *et al*, 2004; Bray *et al*, 2005; Gornick *et al*, 2005). Dysbindin is widely expressed in brain, and appears to play a part in cognitive functioning and capacity (Owen *et al*, 2004). Post-mortem studies suggest that brain levels of dysbindin may be reduced in individuals with schizophrenia (Weickert *et al*, 2004). However, no variant of dysbindin has been specifically linked to schizophrenia. Across different studies, the risk conferred by any dysbindin variant is small, with effect sizes of about 2.0 or less. Among positive association studies, the specific alleles associated with the disease differ. Moreover, an allele may be associated with increased disease risk in some studies and decreased risk in others (Owen *et al*, 2004). In general, the variants of interest (defined by SNPs) are common and without known functional significance. An exception is SNP rs1047631, which has been associated with differences in the expression of dysbindin in brain (Funke *et al*, 2004). However, the frequency of the haplotype with this SNP was similar between cases (45.6%) and controls (40.4%). Thus far, resequencing efforts have not revealed any coding sequence mutations in dysbindin among individuals with schizophrenia (Liao & Chen, 2004).

Therefore, at present the evidence supporting dysbindin is mixed. Variable associations with different alleles have been attributed to allelic heterogeneity; yet allelic heterogeneity refers to different disease-causing mutations in the same gene, not to the same allele reducing risk in some cases and increasing risk in others. There are at least three possible interpretations of these data. The most favoured in the literature is that dysbindin variants mediate disease risk as part of a complex interaction with other genes and environmental factors. This is possible, in principle, although difficult to test given the challenge of establishing the role of a mediating factor of small effect on a complex illness of unknown cause.

A second possibility is that relatively rare, as yet unidentified variants in the dysbindin locus are embedded in illness-associated haplotypes in some, but not all, cases (including those potentially located in non-coding regulatory regions). Such alleles could have substantial effects on the phenotype, but would be masked by studying only the common haplotype. The third possibility is that many, most, or all of the various positive associations with dysbindin are false positives. The number of different positive association studies (albeit with different variants) is taken as *prima facie* evidence that the gene must be involved with the disorder. However, dysbindin, like many genes involved with brain development, is large (>140 kb). The dysbindin locus includes at least 363 SNPs (Hinrichs *et al*, 2006), from which various candidates are selected for association studies. Incorporating linkage disequilibrium across the locus, many thousands of SNP and haplotype combinations appear in different populations. The potential for false positives is enormous. Unless negative and positive studies were published with equal frequency, this possibility is also nearly impossible to test.

In contrast, *DISC1* and its associated non-coding antisense RNA *DISC2* were originally identified by a balanced translocation involving chromosome 1q42 which segregated with schizophrenia (and other major psychiatric disorders) over four generations in a large Scottish kindred (St Clair *et al*, 1990; Millar *et al*, 1998, 2000). Sachs *et al* (2005) found that a frameshift mutation that abnormally truncates *DISC1* co-segregated with schizophrenia in three siblings. However, this mutation is also rarely found in controls with unknown diagnostic status (Green *et al*, 2006). The gene *PDE4B*

(phosphodiesterase 4B), which interacts with *DISC1* in the neuronal cyclic adenosine monophosphate (AMP) pathway, was disrupted by a balanced translocation in two related individuals with chronic psychotic illnesses (Millar *et al*, 2005). Finally, mice with a deletion variant that disrupts the *DISC1* protein have working memory deficits (Koike *et al*, 2006). These findings suggest a role of rare large-effect mutations of *DISC1* and of genes involved in *DISC1* pathways in the development of schizophrenia.

Not surprisingly, *DISC1* became the subject of association studies. Haplotypes of *DISC1* were associated with schizophrenia and other mental illness in European and North American populations (Ekelund *et al*, 2001; Hennah *et al*, 2003; Hodgkinson *et al*, 2004; Callicott *et al*, 2005; Cannon *et al*, 2005; Hamshere *et al*, 2005) but not in Japanese or Scottish populations (Devon *et al*, 2001; Kockelkorn *et al*, 2004; Zhang *et al*, 2005). Within populations with positive associations, *DISC1* haplotypes were also associated with putative endophenotypes, including neuroanatomical and/or neurocognitive profiles (Hodgkinson *et al*, 2004; Burdick *et al*, 2005; Cannon *et al*, 2005). However, disease-risk haplotypes vary across populations and effect sizes are small. Therefore, although there is compelling evidence that rare large-effect mutations in *DISC1* are associated with schizophrenia, it is not clear whether common polymorphisms play a part.

IMPLICATIONS FOR GENETIC RESEARCH IN SCHIZOPHRENIA

The 'common disease – rare allele' model has important implications for gene-finding strategies. A current mantra in schizophrenia genetic research is the need for ever-larger sample sizes in order to obtain adequate statistical power to detect common small-effect variants (Devon *et al*, 2001). These designs are dependent upon the existence of disease-risk alleles that are shared across large numbers of unrelated cases. These strategies will be inadequate if schizophrenia in large part stems from individually rare disease-risk mutations in a large number of different genes.

We propose an alternative strategy in selecting cases for study. Rare cases of schizophrenia with mutations that can be individually detected using current genomic technologies are extremely valuable. It is worthwhile devoting resources to finding

them. It has been recognised for decades that such cases would include any large kindred with a number of well-diagnosed individuals or cases with identifiable genomic events of recent origin (e.g. balanced translocations). The emerging story of *DISC1* highlights this strategy, since the gene was originally identified by a balanced translocation on chromosome 1q42 co-inherited with schizophrenia (Millar *et al*, 1998). Similar promising findings have been noted for specific genes in the 22q11 region (Maynard *et al*, 2002), stemming in part from the recognised association between deletions at 22q11 (i.e. the velo-cardio-facial syndrome) and schizophrenia.

Current genomic technology now enables the identification of an increasingly large number of classes of mutations. Here-fore, identifiable genomic events have been limited to chromosomal abnormalities such as translocations or deletions. However, as the resolution of genome-wide mutation screening technologies improves, smaller genomic events in informative cases or families can be detected. Once identified, a gene altered by a single chromosomal event becomes a candidate to be screened for other (typically smaller) mutations in other cases. Advanced resequencing technology allows for more rapid identification of mutations in candidate genes. The occurrence of multiple deleterious mutations among unrelated cases provides both biological evidence and epidemiological support for the causal role of the gene, using gene-based hypothesis testing strategies (Chen *et al*, 2006).

Individuals who develop schizophrenia following a known environmental exposure are also potentially informative. Such exposures may focus gene discovery in two ways. First, genomic approaches (e.g. resequencing efforts) can focus on candidate genetic pathways relevant to the suspected exposure, screening for otherwise benign variants that are deleterious given the exposure. For example, associations between schizophrenia and *in utero* exposure to maternal starvation (Susser *et al*, 1996; St Clair *et al*, 2005), and associations between schizophrenia and genes in the folate metabolic pathway (Lewis *et al*, 2005; Picker & Coyle, 2005) suggest that mutations in genes in the folate metabolism network could be linked to the illness. Second, the mutagenic effects of the environmental exposure can be evaluated. Following the same example, gestational folate deficiency may be mutagenic, in that

it leads to an increase in the rate of mutation genome-wide (McClellan *et al*, 2006). Among such cases, genome-wide mutational screening may detect *de novo* mutations. If disease-associated alleles are identified, other mutations within the same gene may confer some risk for other individuals with the same exposure. Furthermore, severe mutations in the same gene may lead to the disorder without the exposure.

As informative cases are identified, genomic technologies are needed for efficient screening for potential disease-associated mutations. Effective transcriptome- and proteome-based tools are needed to characterise such variants. These technologies are under rapid development. It is already possible to detect deletions, duplications and other chromosomal aberrations of multiple kilobases anywhere in the genome, and the sensitivity of these methods for detecting smaller mutations is improving (Sebat *et al*, 2004; Sharp *et al*, 2005). Resequencing tools are increasingly efficient, so large numbers can be screened for rare events in candidate genes first identified by rare, individually detectable events. User-friendly bioinformatics resources now exist to help characterise the structure and function of potential candidate genes. It is increasingly possible to characterise not only mutations in protein coding sequence but also mutations and epigenetic changes in regulatory regions, in non-coding RNA and in transposable elements.

To summarise, we propose that individually rare alleles with large effect, many of which are recent in origin, have a substantial role in causing schizophrenia. Current research designs that focus on collecting large samples of unrelated individuals for analysis of shared alleles or haplotypes are not suitable for detecting such disease alleles. In contrast, rare disease mutations may be revealed by studies of individuals and families that harbour informative genomic events, and by studies of exposed cohorts. A gene harbouring one mutation predisposing to schizophrenia is likely to harbour more than one, with frequencies ranging from relatively common to rare, and effects ranging from modest to severe. To see one is not to see them all.

ACKNOWLEDGEMENTS

Support for this work was provided by grants from the National Institutes of Health (MH01120) and

the Stanley Medical Research Foundation to J.M.M., the Lieber Center for Schizophrenia Research to E.S. and a National Association for Research on Schizophrenia and Affective Disorders Distinguished Investigator Award to M.-C.K.

REFERENCES

- Bird, T. D. (2005)** Genetic factors in Alzheimer's disease. *New England Journal of Medicine*, **352**, 862–864.
- Blouin, J. L., Dombroski, B. A., Nath, S. K., et al (1998)** Schizophrenia susceptibility loci on chromosomes 13q32 and 8p21. *Nature Genetics*, **20**, 70–73.
- Botstein, D. & Risch, N. (2003)** Discovering genotypes underlying human phenotypes: past successes for Mendelian disease, future approaches for complex disease. *Nature Genetics*, **33**, 228–237.
- Bray, N. J., Preece, A., Williams, N. M., et al (2005)** Haplotypes at the dystrobrevin binding protein 1 (DTNBP1) gene locus mediate risk for schizophrenia through reduced DTNBP1 expression. *Human Molecular Genetics*, **14**, 1947–1954.
- Brown, A. S., Schaefer, C. A., Wyatt, R. J., et al (2002)** Paternal age and risk of schizophrenia in adult offspring. *American Journal of Psychiatry*, **159**, 1528–1533.
- Brustowicz, L. M., Hodgkinson, K. A., Chow, E. W., et al (2000)** Location of a major susceptibility locus for familial schizophrenia on chromosome 1q21–q22. *Science*, **288**, 678–682.
- Burdick, K. E., Hodgkinson, C. A., Szeszko, P. R., et al (2005)** DISC1 and neurocognitive function in schizophrenia. *Neuroreport*, **16**, 1399–1402.
- Byrne, M., Agerbo, E., Ewald, H., et al (2003)** Parental age and risk of schizophrenia: a case-control study. *Archives of General Psychiatry*, **60**, 673–678.
- Callicott, J. H., Straub, R. E., Pezawas, L., et al (2005)** Variation in DISC1 affects hippocampal structure and function and increases risk for schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 8627–8632.
- Cannon, M., Kendell, R., Susser, E., et al (2003)** Prenatal and perinatal risk factors for schizophrenia. In *The Epidemiology of Schizophrenia* (eds R. Murray, P. Jones, E. Susser, et al), pp. 74–100. Cambridge University Press.
- Cannon, T. D., Hennah, W., van Erp, T. G., et al (2005)** Association of DISC1/TRAX haplotypes with schizophrenia, reduced prefrontal gray matter, and impaired short- and long-term memory. *Archives of General Psychiatry*, **62**, 1205–1213.
- Cavalli-Sforza, L. L., Menozzi, P. & Piazza, A. (1994)** *History and Geography of Human Genes*. Princeton University Press.
- Chakravarti, A. (1999)** Population genetics – making sense out of sequence. *Nature Genetics*, **21** (suppl.), 56–60.
- Chen, B. E., Sakoda, L. C., Hsing, A. W. et al (2006)** Resampling-based multiple hypothesis testing procedures for genetic case-control association studies. *Genetic Epidemiology*, **30**, 495–507.
- Cohen, J. C., Kiss, R. S., Pertsemliadis, A., et al (2004)** Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science*, **305**, 869–872.
- Cohen, J. C., Pertsemliadis, A., Fahmi, S., et al (2006)** Multiple rare variants in NPC1L1 associated with reduced sterol absorption and plasma low-density lipoprotein levels. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 1810–1815.
- Crow, J. F. (2000)** The origins, patterns and implications of human spontaneous mutation. *Nature Reviews Genetics*, **1**, 40–47.
- Crow, J. F. (2003)** Development: there's something curious about paternal-age effects. *Science*, **301**, 606–607.
- Dalman, C. & Allebeck, P. (2002)** Paternal age and schizophrenia: further support for an association. *American Journal of Psychiatry*, **159**, 1591–1592.
- Devon, R. S., Anderson, S., Teague, P. W., et al (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. *Psychiatric Genetics*, **11**, 71–78.
- Ekelund, J., Hovatta, I., Parker, A., et al (2001)** Chromosome 1 loci in Finnish schizophrenia families. *Human Molecular Genetics*, **10**, 1611–1617.
- El-Saadi, O., Pedersen, C. B., McNeil, T. F., et al (2004)** Paternal and maternal age as risk factors for psychosis: findings from Denmark, Sweden and Australia. *Schizophrenia Research*, **67**, 227–236.
- Eyre-Walker, A. & Keightley, P. D. (1999)** High genomic deleterious mutation rates in hominids. *Nature*, **397**, 344–347.
- Fanous, A. H. & Kendler, K. S. (2005)** Genetic heterogeneity, modifier genes, and quantitative phenotypes in psychiatric illness: searching for a framework. *Molecular Psychiatry*, **10**, 6–13.
- Fraga, M. F., Ballestar, E., Paz, M. F., et al (2005)** Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 10604–10609.
- Friedman, T. B. & Griffith, A. J. (2003)** Human nonsyndromic sensorineural deafness. *Annual Review of Genomics and Human Genetics*, **4**, 341–402.
- Funke, B., Finn, C. T., Plocik, A. M., et al (2004)** Association of the DTNBP1 locus with schizophrenia in a U.S. population. *American Journal of Human Genetics*, **75**, 891–898.
- Goldstein, D. B. & Chikhi, L. (2002)** Human migrations and population structure: what we know and why it matters. *Annual Review of Genomics and Human Genetics*, **3**, 129–152.
- Goldstein, D. B., Cavalleri, G. L. & Ahmadi, K. R. (2003)** The genetics of common diseases: 10 million times as hard. *Cold Spring Harbor Symposium on Quantitative Biology*, **68**, 395–402.
- Gornick, M. C., Addington, A. M., Sporn, A., et al (2005)** Dysbindin (DTNBP1, 6p22.3) is associated with childhood-onset psychosis and endophenotypes measured by the Premorbid adjustment scale (PAS). *Journal of Autism and Developmental Disorders*, **10**, 1–8.
- Gottesman, I. I. & Shields, J. (1982)** *Schizophrenia: The Epigenetic Puzzle*. Cambridge University Press.
- Green, E. K., Norton, N., Peirce, T. et al (2006)** Evidence that a DISC1 frame-shift deletion associated with psychosis in a single family may not be a pathogenic mutation. *Molecular Psychiatry*, **11**, 798–799.
- Hamshere, M. L., Bennett, P., Williams, N. et al (2005)** Genomewide linkage scan in schizoaffective disorder: significant evidence for linkage at 1q42 close to DISC1, and suggestive evidence at 22q11 and 19p13. *Archives of General Psychiatry*, **62**, 1081–1088.
- Harrison, P. J. & Weinberger, D. R. (2005)** Schizophrenia genes, gene expression, and neuropathology: on the matter of their convergence. *Molecular Psychiatry*, **10**, 40–68.

- Haukka, J., Suvisaari, J. & Lonnqvist, J. (2003)** Fertility of patients with schizophrenia, their siblings, and the general population: a cohort study from 1950 to 1959 in Finland. *American Journal of Psychiatry*, **160**, 460–463.
- Hennah, W., Varilo, T., Kestila, M., et al (2003)** Haplotype transmission analysis provides evidence of association for *DISC1* to schizophrenia and suggests sex-dependent effects. *Human Molecular Genetics*, **12**, 3151–3159.
- Hinrichs, A. S., Karolchik, D., Baertsch, R., et al (2006)** The UCSC Genome Browser Database: update 2006. *Nucleic Acids Research*, **34**, D590–598.
- Hodgkinson, C. A., Goldman, D., Jaeger, J., et al (2004)** Disrupted in schizophrenia 1 (*DISC1*) association with schizophrenia, schizoaffective disorder, and bipolar disorder. *American Journal of Human Genetics*, **75**, 862–872.
- International Human Genome Sequencing Consortium (2001a)** Initial sequencing and analysis of the human genome. *Nature*, **409**, 860–921.
- International Human Genome Sequencing Consortium (2001b)** A physical map of the human genome. *Nature*, **409**, 934–941.
- International Human Genome Sequencing Consortium (2004)** Finishing the euchromatic sequence of the human genome. *Nature*, **431**, 931–945.
- Jaenisch, R. & Bird, A. (2003)** Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature Genetics*, **33** (suppl), 245–254.
- Kirov, G., Ivanov, D., Williams, N. M., et al (2004)** Strong evidence for association between the dystrobrevin binding protein 1 gene (*DTNBPI*) and schizophrenia in 488 parent-offspring trios from Bulgaria. *Biological Psychiatry*, **55**, 971–975.
- Kockelkorn, T. T., Arai, M., Matsumoto, H., et al (2004)** Association study of polymorphisms in the 5' upstream region of human *DISC1* gene with schizophrenia. *Neuroscience Letters*, **368**, 41–45.
- Kohn, Y., Danilovich, E., Filon, D., et al (2004)** Linkage disequilibrium in the *DTNBPI* (dysbindin) gene region and on chromosome 1p36 among psychotic patients from a genetic isolate in Israel: findings from identity by descent haplotype sharing analysis. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics*, **128**, 65–70.
- Koike, H., Arguello, P. A., Kvaajo, M., et al (2006)** *Disc1* is mutated in the I29S6/SvEv strain and modulates working memory in mice. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 3693–3697.
- Lewis, S. J., Zammit, S., Gunnell, D., et al (2005)** A meta-analysis of the *MTHFR C677T* polymorphism and schizophrenia risk. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics*, **135**, 2–4.
- Liao, H. M. & Chen, C. H. (2004)** Mutation analysis of the human dystrobrevin-binding protein 1 gene in schizophrenic patients. *Schizophrenia Research*, **71**, 185–189.
- Malaspina, D., Corcoran, C. & Fahim, C., et al (2002)** Paternal age and sporadic schizophrenia: evidence for *de novo* mutations. *American Journal of Medical Genetics*, **114**, 299–303.
- Maynard, T. M., Haskell, G. T., Lieberman, J. A., et al (2002)** 22q11 DS: genomic mechanisms and gene function in DiGeorge/velocardiofacial syndrome. *International Journal of Developmental Neuroscience*, **20**, 407–419.
- McClellan, J. M., Susser, E. & King, M. C. (2006)** Maternal famine, *de novo* mutations, and schizophrenia. *JAMA*, **296**, 582–584.
- JON M. McCLELLAN, MD, Department of Psychiatry, University of Washington, Seattle; EZRA SUSSER, MD, DrPH, Department of Epidemiology, Mailman School of Public Health, Department of Psychiatry, Columbia University, and New York State Psychiatric Institute, New York; MARY-CLAIRE KING, PhD, Department of Genome Sciences and Department of Medicine (Medical Genetics), University of Washington, Seattle, Washington, USA
- Correspondence: Dr Jon McClellan, Department of Psychiatry, Box 356560, University of Washington, Seattle, WA 98195, USA. Email: drjack@u.washington.edu
- (First received 21 April 2006, final revision 18 August 2006, accepted 4 October 2006)
- Meisler, M. H., Kearney, J., Ottman, R., et al (2001)** Identification of epilepsy genes in human and mouse. *Annual Review of Genetics*, **35**, 567–588.
- Merikangas, K. R. & Risch, N. (2003)** Will the genomics revolution revolutionize psychiatry? *American Journal of Psychiatry*, **160**, 625–635.
- Millar, J. K., Brown, J., Maule, J. C., et al (1998)** A long-range restriction map across 3 Mb of the chromosome 11 breakpoint region of a translocation linked to schizophrenia: localization of the breakpoint and the search for neighbouring genes. *Psychiatric Genetics*, **8**, 175–181.
- Millar, J. K., Wilson-Annan, J. C., Anderson, S., et al (2000)** Disruption of two novel genes by a translocation co-segregating with schizophrenia. *Human Molecular Genetics*, **9**, 1415–1423.
- Millar, J. K., Pickard, B. S., Mackie, S., et al (2005)** *DISC1* and *PDE4B* are interacting genetic factors in schizophrenia that regulate cAMP signaling. *Science*, **310**, 1187–1191.
- Nimgaonkar, V. L. (1998)** Reduced fertility in schizophrenia: here to stay? *Acta Psychiatrica Scandinavica*, **98**, 348–353.
- Numakawa, T., Yagasaki, Y., Ishimoto, T., et al (2004)** Evidence of novel neuronal functions of dysbindin, a susceptibility gene for schizophrenia. *Human Molecular Genetics*, **13**, 2699–2708.
- Owen, M. J., Williams, N. M. & O'Donovan, M. C. (2004)** The molecular genetics of schizophrenia new findings promise new insights. *Molecular Psychiatry*, **9**, 14–27.
- Petit, C., Levilliers, J. & Hardelin, J. P. (2001)** Molecular genetics of hearing loss. *Annual Review of Genetics*, **35**, 589–646.
- Picker, J. D. & Coyle, J. T. (2005)** Do maternal folate and homocysteine levels play a role in neurodevelopmental processes that increase risk for schizophrenia? *Harvard Review of Psychiatry*, **13**, 197–205.
- Risch, N. (1990)** Linkage strategies for genetically complex traits. I. Multilocus models. *American Journal of Human Genetics*, **46**, 222–228.
- Sachs, N. A., Sawa, A., Holmes, S. E., et al (2005)** A frameshift mutation in Disrupted in Schizophrenia 1 in an American family with schizophrenia and schizoaffective disorder. *Molecular Psychiatry*, **10**, 758–764.
- Schwab, S. G., Knapp, M., Mondabon, S., et al (2003)** Support for association of schizophrenia with genetic variation in the 6p22.3 gene, dysbindin, in sib-pair families with linkage and in an additional sample of triad families. *American Journal of Human Genetics*, **72**, 185–190.
- Sebat, J., Lakshmi, B., Troge, J., et al (2004)** Large-scale copy number polymorphism in the human genome. *Science*, **305**, 525–528.
- Sharp, A. J., Locke, D. P., McGrath, S. D., et al (2005)** Segmental duplications and copy-number variation in the human genome. *American Journal of Human Genetics*, **77**, 78–88.
- Sipos, A., Rasmussen, F., Harrison, G., et al (2004)** Paternal age and schizophrenia: a population based cohort study. *BMJ*, **329**, 1070.
- St Clair, D., Blackwood, D., Muir, W., et al (1990)** Association within a family of a balanced autosomal translocation with major mental illness. *Lancet*, **336**, 13–16.
- St Clair, D., Xu, M., Wang, P., et al (2005)** Rates of adult schizophrenia following prenatal exposure to the Chinese famine of 1959–1961. *JAMA*, **294**, 557–562.
- Straub, R. E., MacLean, C. J., O'Neill, F. A., et al (1995)** A potential vulnerability locus for schizophrenia on chromosome 6p24–22: evidence for genetic heterogeneity. *Nature Genetics*, **11**, 287–293.
- Straub, R. E., Jiang, Y., MacLean, C. J., et al (2002)** Genetic variation in the 6p22.3 gene *DTNBPI*, the human ortholog of the mouse dysbindin gene, is associated with schizophrenia. *American Journal of Human Genetics*, **71**, 337–348.
- Susser, E., Neugebauer, R., Hoek, H. W., et al (1996)** Schizophrenia after prenatal famine. Further evidence. *Archives of General Psychiatry*, **53**, 25–31.
- Tsuchiya, K. J., Takagai, S., Kawai, M., et al (2005)** Advanced paternal age associated with an elevated risk for schizophrenia in offspring in a Japanese population. *Schizophrenia Research*, **76**, 337–342.
- Vogelstein, B. & Kinzler, K. W. (2004)** Cancer genes and the pathways they control. *Nature Medicine*, **10**, 789–799.
- Walsh, T., Casadei, S., Coats, K. H., et al (2006)** Spectrum of mutations in *BRCAl*, *BRCA2*, *CHEK2*, and *TP53* in families at high risk of breast cancer. *JAMA*, **295**, 1379–1388.
- Weickert, C. S., Straub, R. E., McClintock, B. W., et al (2004)** Human dysbindin (*DTNBPI*) gene expression in normal brain and in schizophrenic prefrontal cortex and midbrain. *Archives of General Psychiatry*, **61**, 544–555.
- Williams, N. M., Preece, A., Morris, D. W., et al (2004)** Identification in 2 independent samples of a novel schizophrenia risk haplotype of the dystrobrevin binding protein gene (*DTNBPI*). *Archives of General Psychiatry*, **61**, 336–344.
- Wong, A. H., Gottesman, I. I. & Petronis, A. (2005)** Phenotypic differences in genetically identical organisms: the epigenetic perspective. *Human Molecular Genetics*, **14**, R11–R18.
- Zhang, X., Tochigi, M., Ohashi, J., et al (2005)** Association study of the *DISC1/TRAX* locus with schizophrenia in a Japanese population. *Schizophrenia Research*, **79**, 175–180.