Psychological Medicine (2011), **41**, 19–32. © Cambridge University Press 2010
doi:10.1017/S003329171000070X

REVIEW ARTICLE

Rethinking the genetic architecture of schizophrenia

K. J. Mitchell[1]* and D. J. Porteous[2]

[1] *Smurfit Institute of Genetics, Trinity College Dublin, Ireland*
[2] *Medical Genetics Section, University of Edinburgh Molecular Medicine Centre, Institute of Genetics and Molecular Medicine, Edinburgh, UK*

Background. For many years, the prevailing paradigm has stated that in each individual with schizophrenia (SZ) the genetic risk is due to a combination of many genetic variants, individually of small effect. Recent empirical data are prompting a re-evaluation of this polygenic, common disease–common variant (CDCV) model. Evidence includes a lack of the expected strong positive findings from genome-wide association studies and the concurrent discovery of many different mutations that individually strongly predispose to SZ and other psychiatric disorders. This has led some to adopt a mixed model wherein some cases are caused by polygenic mechanisms and some by single mutations. This model runs counter to a substantial body of theoretical literature that had supposedly conclusively rejected Mendelian inheritance with genetic heterogeneity. Here we ask how this discrepancy between theory and data arose and propose a rationalization of the recent evidence base.

Method. In light of recent empirical findings, we reconsider the methods and conclusions of early theoretical analyses and the explicit assumptions underlying them.

Results. We show that many of these assumptions can now be seen to be false and that the model of genetic heterogeneity is consistent with observed familial recurrence risks, endophenotype studies and other population-wide parameters.

Conclusions. We argue for a more biologically consilient mixed model that involves interactions between disease-causing and disease-modifying variants in each individual. We consider the implications of this model for moving SZ research beyond statistical associations to pathogenic mechanisms.

Received 21 December 2009; Revised 9 March 2010; Accepted 15 March 2010; First published online 12 April 2010

Key words: Heterogeneous, polygenic, schizophrenia.

'It ain't what you don't know that gets you into trouble. It's what you know for sure that just ain't so.' Mark Twain

Introduction

Schizophrenia (SZ) is a highly heritable and common disorder, with a population lifetime prevalence of 0.4–0.8% (Tandon *et al*. 2008). Although many cases of SZ are sporadic, a major genetic underpinning has been incontrovertibly demonstrated by findings from twin, family and adoption studies. The concordance rate between monozygotic (MZ) twins is approximately threefold higher than that of dizygotic (DZ) twins, the risk to first-degree relatives of people with SZ is about 10-fold higher than that of the overall population and the risk to adopted people is associated with the affected status of their biological but not their adoptive relatives (Riley *et al*. 2003). What has been far less clear and more contentious is the mode of inheritance or general genetic architecture of the disorder.

Two major models have been proposed that differ fundamentally in their conception of the disorder and its relationship to normal variation. The application of the common disease–common variant (CDCV) concept to SZ has led to a model that proposes that each case is caused by the inheritance of multiple genetic variants that are common in the population and that the disorder occurs when a threshold of genetic burden is passed; that is, inheritance of the disorder is *polygenic*. By contrast, the multiple rare variants model proposes that each case is caused by a single rare variant, but that these variants can occur in different genes in different families/individuals; that is, the disorder is genetically *heterogeneous*. In both models, interactions with non-genetic factors are most probably also involved.

For a variety of reasons, the polygenic CDCV model has had far wider acceptance and indeed has been taken by many as proven on theoretical grounds. Under this view, a few examples of single mutations

* Address for correspondence: Dr K. J. Mitchell, Smurfit Institute of Genetics, Trinity College Dublin, Ireland.
(Email: Kevin.Mitchell@tcd.ie)
(Email: David.Porteous@ed.ac.uk)

predisposing to SZ or related disorders with high penetrance have generally been considered as exceptional cases of arguable relevance to the generality of SZ or the genetic architecture of the disorder as a whole. Recent empirical data are, however, prompting a re-evaluation of the polygenic CDCV model. These include a lack of the expected number of strong positive findings from genome-wide association studies (GWAS) (Goldstein, 2009) and the concurrent discovery of many more single mutations that strongly predispose to SZ and other psychiatric disorders (Sebat *et al*. 2009). These findings, which contradict a substantial body of literature, prompted us to reconsider the theoretical foundations and underpinning assumptions of the polygenic model.

The foundations of the polygenic CDCV model of SZ

Early studies into the familiality of SZ proposed either single-gene recessive, single-gene dominant with incomplete penetrance or two-locus modes of inheritance (reviewed in O'Rourke *et al*. 1982; Riley *et al*. 2003). It is informative to reconsider why these models were rejected and replaced with a polygenic model. In many cases, the arguments presented rely on assumptions that seemed reasonable at the time, but that we now know do not hold.

Several observations have led to the proposal of a polygenic basis for SZ. First, there are few examples of pedigrees where SZ segregates in a clearly Mendelian manner with classical segregation ratios (Gregory, 1960; Gottesman & Shields, 1967). The observed segregation ratios, averaged across the population, are not consistent with a single-locus, simple mode of Mendelian inheritance. A polygenic model was also presumed to be more consistent with the continuing high prevalence of the disorder, supposedly by reducing the visibility of each risk allele to negative selection. Several studies that analysed the inheritance of SZ from familial relative risk data (O'Rourke *et al*. 1982; McGue *et al*. 1985; Risch, 1990*a*) or by segregation analysis (Tsuang *et al*. 1982; Risch & Baron, 1984) also clearly rejected the hypothesis that all cases of SZ are caused by mutation at a single locus. However, they could not distinguish between (*a*) genetic (locus and/or allelic) heterogeneity, (*b*) oligogenic inheritance, with a single gene of major effect and a few modifying loci, or (*c*) polygenic inheritance. Arguments against Mendelian inheritance at one or a few loci were reinforced by the subsequent failure of linkage studies across multiple, unrelated families to generate consistent results. The reasonable alternative explanation for this inconsistency – that the disorder is highly heterogeneous – was dismissed for several reasons.

In an influential paper, Risch (1990*a*) argued directly against heterogeneous Mendelian inheritance on the basis that, if that were the case, then distinct clinical presentations would be expected to segregate in different families: 'if it is allelic heterogeneity, then all related individuals carrying the same allele should show a similar clinical picture'. In fact, we now know that the opposite holds; individual mutations can result in very different phenotypes in different people, and even in distinct clinical diagnoses (e.g. autism, bipolar disorder, epilepsy, SZ) (Table 1). Risch also argued that, if modes of inheritance were different between pedigrees, this would be obvious. However, because of the large decrease in fitness associated with SZ, large multiplex pedigrees are rare and, as penetrance is incomplete, deduction and comparison of modes of inheritance across pedigrees is extremely difficult. Nevertheless, there are now numerous examples of multiplex pedigrees with a better fit to either recessive or dominant inheritance (reviewed by Riley *et al*. 2003). Finally, Risch suggested that the results from the linkage scans that had been conducted on SZ at that time were too inconsistent for heterogeneity to explain them; the lack of replication was taken as evidence instead that the findings were false positives. That a large number of loci might be mutable to cause SZ was thus dismissed as implausible, *a priori*.

A crucial assumption in that line of reasoning was that rates of mutation are far too low to continuously replenish highly penetrant alleles, which would be expected to be rapidly selected against in the population (Gottesman & Shields, 1967). That new mutations can be ignored is explicitly assumed in studies that have attempted to deduce the mode of inheritance from data on familial relative risks (O'Rourke *et al*. 1982; McGue *et al*. 1985; Risch, 1990*a*). In fact, we now know that the *de novo* mutation rate is much higher than previously expected (Crow, 2000). Whole-genome sequencing has revealed a remarkable plethora of rare and private mutations (Ng *et al*. 2008; Wheeler *et al*. 2008; Frazer *et al*. 2009) not detected by the HapMap project. In addition, copy number variants (CNVs) occur at a previously unrecognized and appreciable frequency; lower than point mutations, but affecting far more bases (Sebat *et al*. 2004; Lupski, 2007). *De novo* mutations in human sperm accumulate markedly with age (Crow, 2000) and numerous studies have shown a consistent link between increasing paternal age and risk of SZ in offspring (Malaspina *et al*. 2001; Wohl & Gorwood, 2007). New mutations may thus explain a significant fraction of sporadic SZ cases.

The most commonly cited evidence for a polygenic model of SZ is that the risk to relatives decreases by more than a factor of two with degree of relatedness

Table 1. *Mutations implicated in schizophrenia (SZ). Particular aberrations are included if they have been seen independently multiple times (assuming the possibility of broad phenotypic expression) or if they have strong supporting biological evidence. Only protein-coding genes are listed*

Single-gene mutations	Location	Associated phenotypes	Nature of mutation(s)	Gene length (kb)	References
ABCA13	7p12.3	BD, MD, SZ	Translocation, point mutations	449	1
CNTNAP2	7q35	ADHD, ASD, E, OCD, SZ, TS	CNV, point mutations	2305	2–6
DISC1	1q42.2	ASD, BD, MD, SZ,	Translocation, point mutations	414	7–9
ERBB4	2q34	SZ	CNV	1163	10
GRIK4	11q23.3	BD, MR, SZ	Translocation	326	11
NPAS3	14q13.1	ID, SZ	Translocation	865	12
NRXN1	2p16.3	ASD, SZ	CNV, point mutations	1112	3, 13
PCM1	8p22	SZ	Point mutations	111	14
PDE4B	1p31.3	SZ	Translocation	582	15
PINK1	1p36.12	ANX, MD, OCD, PD, SZ	Point mutations	18	16, 17
SYNGR1	22q13.1	BD, SZ	Point mutations	36	18, 19

Multigenic CNVs	Position (kb)	Associated phenotypes	Variant	Genes affected	References
1q21.1	144.9–146.3	ADHD, ASD, E, MD, MR, SZ	Del/Dup	NBPF11, HYDIN, PRKAB2, FMO5, CHD1L, BCL9, ACP6, GJA5, GJA8, GPR89B, PPIAL4, NBPF14	3, 10, 20–24
1q44	241.48–241.72	SZ	Dup	EFCAB2, KIF26B	20, 25
2p16.1–p15	61.10–61.29	SZ	Dup	AHSA2, KIAA1841, PEX13, USP34	26
2p21–p16.3	48.63–49.29	SZ	Dup	KLRAQ1, STON1–GTF2A1L, LHCGR, FSHR	10, 20
2p16.3	51.10–51.35	ASD, SZ	Del	NRXN1	3, 10, 20, 25, 27
2q12.2	107.3–109.2	SZ	Dup	10–15 genes	20, 22
2q34	211.79–212.19	SZ	Del (×1)	Affects ERBB4	10
3p26.3	1.55–1.63	SZ	Dup	3′ of CNTN6	26
3q29	197.23–198.58	ASD, MR, SZ	Del	20 genes, break in DLG1	3, 10, 28
4q32.1	160.10–160.82	SZ	Dup	C4orf45, RAPGEF2	26
4q35.2	189.86–190.50	SZ	Dup	Non-genic	20, 25, 29
7q35	146–147	ADHD, SZ		CNTNAP2	3, 4
7q36.3	157.40–157.49	SZ	Del	PTPRN2	
	157.62–158.80	ID, SZ	Del/Dup	PTPRN2, NCAPG2, FAM62B, WDR60, VIPR2	22, 30, 31
8p22	15.10–18.36	SZ	Del (×1)	16 genes, including PCM1	20
8p23.2	110	SZ	Dup	DLGAP2	32, 33
	3 Mb	E, MR	Del	DLGAP2…CSMD1, 22 others	32
	2.32–3.46	SZ	Dup	CSMD1	26
9q33.1	Not given	ASD, SZ	Del	ASTN2	20, 27, 34, 35
10q23.1	83.70–83.78	SZ	Dup	NRG3	26
11q14	83.6–83.94	SZ	Del	DLG2	10, 21
11q23.1	112.77–112.78	SZ	Del (×7)	ANKK1	20
12p11.23	Not reported	SZ	Del (×4)	Not reported	3
14q21.1	40.76–40.82	SZ	Del	5′ of LRFN5	26
15q11.2	20.45–20.85	SZ	Del	TUBGCP5, CYFIP1, NIPA2, NIPA1	3, 20, 22, 36
15q11.2	21.2–26.2	ASD, SZ	Dup	13 genes, including UBE3A, GABA-receptors	20, 22, 35, 37
15q13.1	27.0–28.4	SZ	Dup	APBA2, NDNL2, TJP1	3, 20, 25, 38
15q13.3	28.7–30.3	ASD, E, SZ	Del/Dup	MTMR15, MTMR10, TRPM1, KLF13, OTUD7A, CHRNA7	3, 20, 22, 36, 39, 40
16p11.2	29.56–30.08	ASD, BD, SZ	Del/Dup	29 genes	10, 35, 36, 37, 41
16p12.2	21.9–22.3	SZ	Del	UQCRC2, C16orf65, C16orf52, VWA3A, SDR42E2, EEF2K, POLR3E, CDR2	3, 25

[*continued overleaf*

Table 1 (*cont.*)

Multigenic CNVs	Position (kb)	Associated phenotypes	Variant	Genes affected	References
16p13.11	15.39–16.20	ASD, E, MR, SZ	Del/Dup	MPV17L, c16orf45, KIAA0430, NDE1, MYH11, KIAA0866, c16orf63, ABCC1, ABCC6	3, 20, 22, 30, 42, 43
17p12	14.05–15.36	SZ	Del	CDRT15, COX10, HS3ST3B1, PMP22, TEKT3, CDRT4, FAM18B2, CDRT1	3, 22, 30, 36
20p12.1	15.00–15.09	SZ	Del	MACROD2	26
21q11.2	13.69–13.99	SZ	Dup	A26B3, LOC441956	26
22q11.2	350	ASD, MR, SZ	Del	PRODH, DGCR6	32
	17.26–19.79	VCFS, SZ	Del/Dup	43 genes, including PRODH, DGCR6	3, 10, 20, 21, 22, 30, 38
Xp11.4	56	SZ, ASD	Dup	TSPAN7	32, 33

ADHD, Attention deficit hyperactivity disorder; ANX, anxiety disorder; ASD, autism spectrum disorder; BD, bipolar disorder; E, epilepsy; ID, intellectual disability; MD, major depression; MR, mental retardation; OCD, obsessive–compulsive disorder; PD, Parkinson's disease; TS, Gilles de la Tourette's syndrome; VCFS, velocardiofacial syndrome.

[1] Knight *et al*. (2009), [2] Friedman *et al*. (2008), [3] International Schizophrenia Consortium (2008), [4] Elia *et al*. (2009), [5] Bakkaloglu *et al*. (2008), [6] Verkerk *et al*. (2003), [7] Millar *et al*. (2000), [8] Song *et al*. (2008), [9] Chubb *et al*. (2008), [10] Walsh *et al*. (2008), [11] Pickard *et al*. (2006), [12] Pickard *et al*. (2005), [13] Rujescu *et al*. (2009), [14] Kamiya *et al*. (2008), [15] Millar *et al*. (2005), [16] Steinlechner *et al*. (2007), [17] Funayama *et al*. (2008), [18] Cheng & Chen (2007), [19] Verma *et al*. (2004), [20] Need *et al*. (2009), [21] Xu *et al*. (2008), [22] Kirov *et al*. (2009), [23] Brunetti-Pierri *et al*. (2008), [24] Mefford *et al*. (2008), [25] Kirov *et al*. (2008), [26] Xu *et al*. (2009), [27] Vrijenhoek *et al*. (2008), [28] Ballif *et al*. (2008), [29] Pickard *et al*. (2004), [30] Shi *et al*. (2008), [31] Tyson *et al*. (2005), [32] Guilmatre *et al*. (2009), [33] Marshall *et al*. (2008), [34] Park *et al*. 1991), [35] Glessner *et al*. (2009), [36] Stefansson *et al*. (2008), [37] Bucan *et al*. (2009), [38] Rodriguez-Santiago *et al*. (2009), [39] van Bon *et al*. (2009), [40] Miller *et al*. (2009), [41] Weiss *et al*. (2008), [42] Mefford *et al*. (2009), [43] Hannes *et al*. (2009).

(Gottesman & Shields, 1967; Risch, 1990*b*). A monotonic decrease in risk of a factor of two is expected if all cases follow a single-gene dominant mode of inheritance (even if different genes are involved in different families, and regardless of penetrance). It has been argued that this expected relationship is independent of the mode of inheritance (Risch, 1990*b*). In fact, under a standard recessive model, for any given mutation, risk to relatives decreases much more sharply and declines rapidly to zero beyond first-degree relatives if the risk allele is rare. If an appreciable proportion of cases are caused by *de novo* mutation, this will also dramatically reduce the familial relative risk rates. Observed familial risk rates are in fact consistent with expectation from a heterogeneous model, if varying proportions of SZ cases fall under *de novo*, dominant or recessive modes of inheritance (Table 2).

Thus, despite the commonly held view to the contrary, there are no valid theoretical arguments to reject a Mendelian model of inheritance of SZ, with locus and/or allelic heterogeneity and variable penetrance, that may or may not include a role for modifying alleles. As discussed later, the weight of the empirical evidence also favours such a rare variants model.

Empirical data on the influence of common variants

Advances in human genomics, particularly the HapMap project and associated array-based methods for low-cost, high-throughput genotyping, offered the hope that common variants contributing a modest increase in risk could be detected by GWAS (Risch & Merikangas, 1996; Reich & Lander, 2001). Several such studies have now been published (Mah *et al*. 2006; O'Donovan *et al*. 2008; Need *et al*. 2009), including recent ones with very large sample sizes, across which data could be pooled for meta-analysis (Purcell *et al*. 2009; Shi *et al*. 2009; Stefansson *et al*. 2009). Across these studies, only a few loci showed genome-wide significant association with increased risk, each with very small effect sizes [odds ratio (OR) $\sim$1.1–1.2]. With the exception of the long-established SZ association with the human leucocyte antigen (HLA) region, there was little evidence for concordance of 'top-ranked' associations between the three studies. Given the large sample sizes in each study, the uncontestable primary conclusion is that the predicated large number of common risk variants of even modest effect size do not exist.

To assess more generally the possibility of a polygenic influence on risk, the authors of one of these studies (Purcell *et al*. 2009) derived an aggregate score from the association results of the top 10–50% of over 74 000 single-nucleotide polymorphisms (SNPs). This score could predict about 3% of the variance in risk of SZ in a target group. At face value, these data indicate that 97% of the observed variance in risk is not due to polygenic common variants. The authors argue,

Table 2. *Familial recurrence risks with heterogeneous modes of Mendelian inheritance. Expected recurrence risks for various family members under a hypothetical equal distribution of cases into three different modes of inheritance:* de novo, *dominant and recessive*

Mode	MZ	DZ	FS	Offspring	Niece	Cousin
De novo	100	0	0	50	0	0
Dominant	100	50	50	50	25	12.5
Recessive	100	25	25	0	0	0
Average	100	25	25	33.3	8.3	4.2
50% penetrant	50	12.5	12.5	16.7	4.2	2.1
Observed	52.1	14.2	8.6	10.0	3.1	1.8

DZ, Dizygotic; FS, full sibling.

Average penetrance of 50% is assumed, based on monozygotic (MZ) twin concordance. A greater than twofold reduction in risk with increasing degrees of relatedness can readily be generated by such a distribution. This is not meant to suggest that these particular figures are correct, but rather the opposite, that frequencies in relatives cannot be used to disprove heterogeneous Mendelian inheritance. As originally pointed out by James (1971): there is 'an infinite number of parameter sets … which lead to the same frequencies in relatives'.

however, that the variants that are having a real effect are being drowned out by the noise of the majority of false positives and use simulations to attempt to estimate the 'true variance' that would be explained if only we knew which SNPs to look at. These simulations are, however, based explicitly on the liability-threshold model, which, we argue below, has no biological validity. Moreover, despite the claim of dramatic convergence, the resultant seven models (out of 560 tested) give wildly different estimates of the total true variance explained (from 34% to 98%) and the number of SNPs contributing (from 6% to 100%). By default, the necessary and striking conclusion from this study is thus that not common, lowly penetrant, but rare, highly penetrant variants explain at least two-thirds and possibly as much as 97% of the phenotypic variance across the population. This conclusion has sound empirical support.

Empirical support for the rare variants model

The most compelling evidence for rare variants is that many mutations have now been shown to segregate with SZ and are thus highly likely to be causal (McClellan *et al*. 2007; Sebat *et al*. 2009). These include single-gene mutations in addition to CNVs or other cytogenetic abnormalities (Table 1). Typically, such mutations show high, but incomplete, penetrance for SZ and may lead to other psychiatric disorders such as bipolar disorder, major depression, autism or attention deficit hyperactivity disorder, and also to epilepsy and mental retardation (Table 1). These findings are consistent with studies of increased co-morbidity of autism, bipolar disorder and epilepsy within families of schizophrenics (Cardno *et al*. 2002; Qin *et al*. 2005; Lichtenstein *et al*. 2009; Steinhausen *et al*. 2009) and provide additional evidence for an overlapping aetiology of these disorders.

It is important to note also several families where the appearance of a psychiatric diagnosis was not restricted to those members carrying the putative causative allele (Guilmatre *et al*. 2009; Xu *et al*. 2009). This apparent violation of the principle of causality could be explained by intra-familial heterogeneity in some families, where more than one highly penetrant mutation is segregating. Such a situation may be more common than expected because of assortative mating (Merikangas, 1982) and a high *de novo* mutation rate.

Thus, mutations in many genes affecting the development or function of the nervous system may lead to the same clinical phenotype, whereas mutations in the same gene can lead to different phenotypes in different individuals (Sebat *et al*. 2009). With this background, it is not surprising that traditional linkage studies using single diagnostic categories or combining multiple families segregating different mutations have had few successes in identifying such mutations. This contrasts with the remarkable return, by current proportion of effort and investment, from single family and single case studies.

At the moment, such Mendelian cases or specific CNVs constitute a small, but growing, fraction of SZ cases. There is every reason to think, however, that these represent only the tip of the iceberg and that CNV analyses with better genome coverage and deep resequencing or whole-genome sequencing will turn up many more such cases. For example, sequencing the exons of Disrupted-in-Schizophrenia 1 (DISC1) identified five ultra-rare missense mutations in six out of 288 unrelated cases of SZ (2% overall), which were not found in 10 000 control chromosomes (Song *et al*. 2008). Bioinformatic (Chubb *et al*. 2008) and biological analyses (D. Porteous, unpublished data) suggest that these rare variants alter DISC1 function. If a conservative 2% of cases can be explained by highly penetrant mutations in this single gene, and this case is not exceptional, then single mutations could clearly numerically account for all cases of SZ. Importantly, many of the implicated genes have biologically plausible roles in neural development (Guilmatre *et al*. 2009; Raychaudhuri *et al*. 2009; Sebat *et al*. 2009), and when mutated in mice model aspects of SZ with good construct, face and predictive validity (Waddington *et al*. 2007; Carpenter & Koenig, 2008; Porteous, 2008).

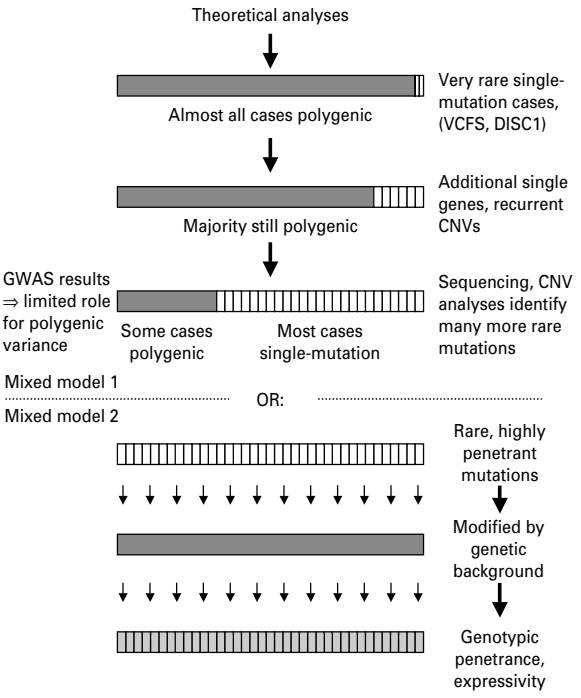

Fig. 1. Recent evolution of genetic models for schizophrenia. See text for details.

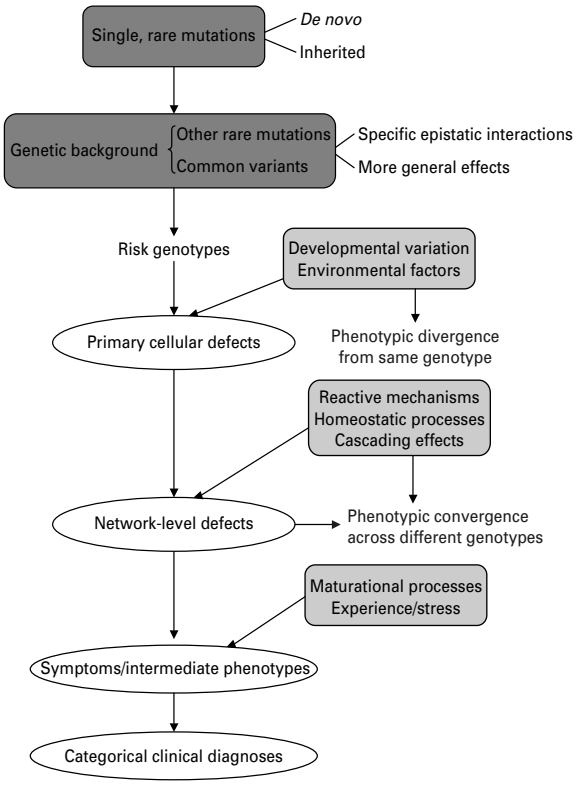

Fig. 2. A framework for the complex actiology of schizophrenia. Rare mutations in many different genes, modified by genetic background, will generate risk genotypes. Non-genetic factors will influence the trajectory of neurodevelopment and contribute to phenotypic divergence. Primary defects may trigger reactive mechanisms or have cascading effects on subsequent experience-dependent development. Phenotypic convergence may arise through overlapping primary defects or common secondary pathways. All these processes, possibly intersecting with ongoing maturation and/or stress, will ultimately result in a varied spectrum of intermediate phenotypes or symptoms in different individuals and thus in different clinical diagnoses, one category of which is schizophrenia.

The hypothesis that they are causative, although obviously requiring case-by-case confirmation, is thus highly parsimonious and well supported by the evidence to date.

Thus, there is now very strong evidence for a major contribution of single mutations of high penetrance to the genetic architecture of SZ. This role has been incorporated into mixed models that either apportion cases to those caused by single mutations *versus* polygenic influences (Psychiatric GWAS Consortium Steering Committee, 2009) or presume an interaction between these factors in individuals (Frazer *et al*. 2009).

Mixed models, mutations and modifiers

On the reasonable assumption that the rare variants already identified represent only a fraction of those causing disease, and if the model of Mendelian inheritance with genetic heterogeneity is broadly consistent with the distribution of familial risks across the population, is there any reason to invoke additional factors? Analyses of GWAS data do suggest some contribution from common alleles to phenotypic variance across the population. In addition, many rare variants show incomplete penetrance and variable phenotypic expressivity that could be partly explained by an interaction with other genetic factors. These and other considerations have led to the emergence of mixed models for the genetic architecture of SZ (Fig. 1). One such model apportions cases of SZ across the population to those caused by single mutations *versus* those caused by polygenic effects. A second model proposes that the phenotypic expression of single mutations of high penetrance will be modified by genetic background; that is, there is an epistatic interaction, within each individual, between disease-causing and disease-modifying variants. Under this model, the 'causal' variant may be necessary, but not always sufficient, to generate the phenotype.

We argue here that the second model is more parsimonious, more biologically plausible and has greater explanatory power. If we know that single mutations can lead to this phenotype, there is no obvious reason to invoke what is essentially a novel and highly

speculative genetic mechanism to explain the fraction of cases where a mutation of major influence has not yet been identified. It has been proposed that a cumulative polygenic effect could be manifest as a distinct phenotype under a liability-threshold model (Gottesman & Shields, 1967). Although this concept has been described mathematically for decades (Dempster & Lerner, 1950), it remains essentially metaphorical.

This model is founded on the notion of an underlying 'liability' to the disorder, which is normally distributed in the population (reflecting the distribution of many common alleles), but that the risk of developing the disease depends upon exceeding some 'threshold' of liability. This model posits a sharp increase in risk between someone with n liability alleles on one side of the threshold and someone with $n+1$ liability alleles on the other (e.g. Kendler & Kidd, 1986). Why this should be the case is not clear. Nor is there any empirical evidence to suggest that this mechanism ever applies, unless of course n is a very small number and segregation is essentially Mendelian.

The appeal of this model seems to arise from a conflation between a *threshold event*, in a developmental sense, which can explain incomplete penetrance of an all-or-none phenotype, such as cleft palate, and a *threshold of cumulative genetic burden*. In cleft palate, the threshold explains the bimodal phenotype distribution in people with the same underlying liability (even MZ twins). Whether or not the phenotype emerges depends on chance or other non-genetic factors, including epigenetic cascades or environmental triggers. A similar mechanism might very well be at play in the emergence of SZ (Mitchell, 2007) (Fig. 2). However, this mechanism does not relate in any way to the difference in the chance of the event happening in people with different genotypes, nor is there any evidence for a sudden change from very low to very high risk at some point along a genetic continuum.

In fact, biological systems are generally highly buffered and quite robust to the cumulative effects of weak mutations (Kitano, 2004). Most genetic systems incorporate substantial levels of noise, as a design feature, not a 'bug' (Kerszberg, 2004). Intrinsic variation in the amounts or activities of individual components serves to buffer the effects of extrinsic perturbations, and also has the effect of allowing the system to absorb small changes in gene expression of many individual components. These properties hold for both metabolic (Kacser & Burns, 1981) and developmental (Waddington, 1959; Siegal & Bergman, 2002) systems.

Several observations have been put forward as supporting a multifactorial threshold model in the case of SZ (Gottesman & Shields, 1967). These include: (i) increasing risk with increasing familial loading (number of affected relatives); (ii) increasing risk to siblings with the severity of the disorder in the affected proband; and (iii) increased risk with bilineal inheritance (where both parents are affected). In fact, all of these observations are consistent with a heterogeneous Mendelian mode of inheritance. If some alleles have higher penetrance than others, then some pedigrees will show more affecteds than others and this will correlate with statistical risk to individual members. Similarly, if more highly penetrant alleles also have higher expressivity (and there is no reason to expect that they would not), then individual risk would also be correlated with severity in relatives. Finally, if both parents are affected due to independent risk alleles of modest to high penetrance, then their offspring will obviously be at much higher risk, especially if there are also epistatic interactions between the two alleles (Gottesman *et al*. 2010).

A mixed model that apportions cases to either single-mutation or polygenic mechanisms also provides no explanation for the variability in phenotypic expression of particular mutations. The alternative mixed model incorporates a role for polygenic influences but proposes that these modify the phenotypic effects of highly penetrant mutations, rather than producing the phenotype themselves. This view is highly congruent with findings from experimental genetics, where genetic background effects are typical and can be large (Nadeau, 2001; Shao *et al*. 2008). For example, many phenotypes in mutant mice are affected by a change in genetic background, sometimes dramatically, although the mutant phenotype usually never occurs in either background without the major mutation. Indeed, given the complexities of the phenotypes concerned, it would be astonishing if such genetic background effects were not important in human psychiatric disease (e.g. Girirajan *et al*. 2010).

Of course, the genetic background will be made up of all other variants, rare and common, in the genotype. We might expect the rarer variants to make a larger contribution to phenotypic variance; indeed, this seems to be the case even for many quantitative traits (Cohen *et al*. 2004; Ji *et al*. 2008; Frazer *et al*. 2009; Goldstein, 2009). However, common variants (most obviously the presence or absence of a Y chromosome) may also play an important role, as outlined below.

Rare variants, common modifiers and endophenotypes

Particular common variants could directly modify the biochemical, cellular or physiological effects of specific rare variants. For example, common variants

in proline dehydrogenase (oxidase) 1 (PRODH) may modify the phenotypic effects of 22q11 deletions (Guilmatre *et al*. 2009). Such dominance or epistatic effects would most probably be specific to particular rare variants. Common variants might also influence risk indirectly, but more generally, for example through effects on cognitive reserve (discussed later).

A genetic background effect offers one possible explanation for what might be considered the strongest evidence for a polygenic influence on SZ risk, namely that unaffected relatives of SZ patients often show intermediate values on a range of endophenotypes. The endophenotype model is based firmly on the common variants hypothesis; specifically, it suggests that different variants may contribute to different aspects of the overall phenotype (such as defects in working memory or executive function, for example) and that the combination of many such defects results in the clinical phenotype (Gottesman & Gould, 2003; Braff *et al*. 2007). The prediction of this model is that relatives of affected individuals will carry some, but not all, of the multiple risk variants that in combination cause disease, and will thus show some of the endophenotypes of the disease, though normally at a less severe level. By contrast, the single-mutation hypothesis in its simplest form (without modifiers) predicts that unaffected relatives will not differ from the general population.

Many studies do indeed show that unaffected first-degree relatives of schizophrenics have endophenotypes that place them, on average, intermediate between cases and unrelated controls. These include studies of cognitive endophenotypes, such as working memory or executive function, and also physiological endophenotypes, such as the amplitude of various evoked potentials. How can the rare variants model explain these findings?

First, some endophenotypes, especially ones that measure cognitive abilities, may reflect a genetic background effect that modifies risk of expression of disease, rather than disease-causing genes *per se*. There is very good evidence that poor pre-morbid cognitive function, perhaps reflecting inefficient neural processing, is a strong independent risk factor for SZ (Woodberry *et al*. 2008). A less efficient brain (with lower 'cognitive reserve') may be less well buffered and more likely to be pushed into a pathophysiological state by the effects of rare variants (as seen in brain injury and dementia) (Barnett *et al*. 2006). SZ patients and their relatives might therefore be expected to have an underlying deficit in cognitive performance, compared to unrelated controls.

A second explanation for observed endophenotypes in 'unaffected' relatives is that many of them will also carry the risk variant, as the allelic penetrance for the clinical diagnosis is typically less than 50%. If the penetrance for a particular endophenotype is higher, then many first-degree relatives will show this endophenotype, and in fact may show it at the same level as affecteds. The average across relatives will thus be intermediate between cases and controls. There is good evidence that this situation applies to some endophenotypes. In the case of the DISC1 translocation, for example, relatives who carry the translocation, but who are not diagnosed with a psychiatric disorder, show the same P300 auditory evoked potential phenotype as affecteds (Blackwood & Muir, 2004). More generally, many endophenotypes are present only in relatives who have broader psychiatric symptoms or diagnosis (Saperstein *et al*. 2006; Turetsky *et al*. 2008; Prasad *et al*. 2009) or in relatives who later go on to develop full-blown SZ (Erlenmeyer-Kimling *et al*. 2000; Owens & Johnstone, 2006), again presumably differentiating carriers of a causal mutation from non-carriers.

Explaining the persistence of SZ

The strongest version of the rare variants model thus proposes that most, if not all, cases of SZ are dependent upon the presence of some highly penetrant mutation. This is consistent with general evidence that rare mutations are much more likely to affect protein function and to cause deleterious phenotypes (Kryukov *et al*. 2007; Ng *et al*. 2008). Conversely, mutations that have such effects are likely to be subject to strong negative selection and not rise to high allelic frequencies. To explain the high rate of the disorder thus requires both a high enough rate of new mutation and a large enough mutational target (Keller & Miller, 2006). For this model to be true, there must therefore be many genes that, when mutated, can give rise to SZ. More than 45 distinct loci have already been implicated (Table 1) and this number is very likely to increase. A mutation-selection balance model also predicts a preponderance of mutations in larger genes and in regions with a higher than average mutation rate. These predictions are borne out by the data, with the observed length of the implicated genes over 200-fold greater than the genomic average and a significant contribution from recurrent CNVs (Table 2).

This suggests a straightforward and plausible explanation for the existence and persistence of SZ in human populations. It simply takes a lot of genes to build and wire a human brain. This may seem to contradict the assertion made earlier that neurodevelopmental systems are robust to the cumulative effects of small changes. It is important to note, however, that such systems may nevertheless be sensitive to large perturbations of particular components, especially

highly connected nodes (e.g. DISC1) (Kitano, 2004). When early processes of brain wiring fail due to mutation in one or other of these genes, neurodevelopment may be channelled into a common, maladaptive state (Fig. 2). For now, we call this schizophrenia, but in time it may be usefully and appropriately redefined on the basis of the particular molecular pathology.

Rethinking the approach to SZ

In summary, there is every reason to expect that most, if not all, cases of SZ are dependent upon the presence of some highly penetrant mutation. The number of loci involved, the frequency of the mutations, the fractions that are dominant or recessive and their penetrance for either intermediate phenotypes or for the clinical diagnosis of SZ itself are all matters for empirical investigation. Where a polygenic effect is likely to be observed is in modifying the phenotypic expression of such mutations, rather than in generating a distinct proportion of cases. Specific epistatic interactions are very likely to be important in determining each individual's phenotype (e.g. Girirajan *et al*. 2010). This will undoubtedly complicate analyses of segregation, as strict Mendelian patterns of inheritance are unlikely to be the norm. In addition, there must also be a major contribution from modifying effects of environmental factors (Kinney *et al*. 2009) or stochastic developmental variation (Woolf, 1997; Mitchell, 2007) to explain phenotypic discordance of MZ twins (Fig. 2). However, none of these factors should distract from a focus on the primary, causative mutations. As in other areas of biology, mutations with the largest effects in individuals, regardless of their frequency in the population, will be the most informative as to the underlying pathogenic and pathophysiological mechanisms.

Acknowledgements

We thank M.-C. King and M. Slatkin for very helpful discussions and for sharing unpublished data and D. McConnell, D. Bradley, M. Ramaswami and S. Roche for their feedback on the manuscript.

Declaration of Interest

None.

References

Bakkaloglu B, O'Roak BJ, Louvi A, Gupta AR, Abelson JF, Morgan TM, Chawarska K, Klin A, Ercan-Sencicek AG, Stillman AA, Tanriover G, Abrahams BS, Duvall JA, Robbins EM, Geschwind DH, Biederer T, Gunel M, Lifton RP, State MW (2008). Molecular cytogenetic analysis and resequencing of contactin associated protein-like 2 in autism spectrum disorders. *American Journal of Human Genetics* **82**, 165–173.

Ballif BC, Theisen A, Coppinger J, Gowans GC, Hersh JH, Madan-Khetarpal S, Schmidt KR, Tervo R, Escobar LF, Friedrich CA, McDonald M, Campbell L, Ming JE, Zackai EH, Bejjani BA, Shaffer LG (2008). Expanding the clinical phenotype of the 3q29 microdeletion syndrome and characterization of the reciprocal microduplication. *Molecular Cytogenetics* **1**, 8.

Barnett JH, Salmond CH, Jones PB, Sahakian BJ (2006). Cognitive reserve in neuropsychiatry. *Psychological Medicine* **36**, 1053–1064.

Blackwood DH, Muir WJ (2004). Clinical phenotypes associated with DISC1, a candidate gene for schizophrenia. *Neurotoxicity Research* **6**, 35–41.

Braff DL, Freedman R, Schork NJ, Gottesman II (2007). Deconstructing schizophrenia: an overview of the use of endophenotypes in order to understand a complex disorder. *Schizophrenia Bulletin* **33**, 21–32.

Brunetti-Pierri N, Berg JS, Scaglia F, Belmont J, Bacino CA, Sahoo T, Lalani SR, Graham B, Lee B, Shinawi M, Shen J, Kang SH, Pursley A, Lotze T, Kennedy G, Lansky-Shafer S, Weaver C, Roeder ER, Grebe TA, Arnold GL, Hutchison T, Reimschisel T, Amato S, Geragthy MT, Innis JW, Obersztyn E, Nowakowska B, Rosengren SS, Bader PI, Grange DK, Naqvi S, Garnica AD, Bernes SM, Fong CT, Summers A, Walters WD, Lupski JR, Stankiewicz P, Cheung SW, Patel A (2008). Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities. *Nature Genetics* **40**, 1466–1471.

Bucan M, Abrahams BS, Wang K, Glessner JT, Herman EI, Sonnenblick LI, Alvarez Retuerto AI, Imielinski M, Hadley D, Bradfield JP, Kim C, Gidaya NB, Lindquist I, Hutman T, Sigman M, Kustanovich V, Lajonchere CM, Singleton A, Kim J, Wassink TH, McMahon WM, Owley T, Sweeney JA, Coon H, Nurnberger JI, Li M, Cantor RM, Minshew NJ, Sutcliffe JS, Cook EH, Dawson G, Buxbaum JD, Grant SF, Schellenberg GD, Geschwind DH, Hakonarson H (2009). Genome-wide analyses of exonic copy number variants in a family-based study point to novel autism susceptibility genes. *PLoS Genetics* **5**, e1000536.

Cardno AG, Rijsdijk FV, Sham PC, Murray RM, McGuffin P (2002). A twin study of genetic relationships between psychotic symptoms. *American Journal of Psychiatry* **159**, 539–545.

Carpenter WT, Koenig JI (2008). The evolution of drug development in schizophrenia: past issues and future opportunities. *Neuropsychopharmacology* **33**, 2061–2079.

Cheng MC, Chen CH (2007). Identification of rare mutations of synaptogyrin 1 gene in patients with schizophrenia. *Journal of Psychiatric Research* **41**, 1027–1031.

Chubb JE, Bradshaw NJ, Soares DC, Porteous DJ, Millar JK (2008). The DISC locus in psychiatric illness. *Molecular Psychiatry* **13**, 36–64.

Cohen JC, Kiss RS, Pertsemlidis A, Marcel YL, McPherson R, Hobbs HH (2004). Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science* **305**, 869–872.

Crow JF (2000). The origins, patterns and implications of human spontaneous mutation. *Nature Reviews Genetics* **1**, 40–47.

Dempster ER, Lerner IM (1950). Heritability of threshold characters. *Genetics* **35**, 212–236.

Elia J, Gai X, Xie HM, Perin JC, Geiger E, Glessner JT, D'Arcy M, Deberardinis R, Frackelton E, Kim C, Lantieri F, Muganga BM, Wang L, Takeda T, Rappaport EF, Grant SF, Berrettini W, Devoto M, Shaikh TH, Hakonarson H, White PS (2009). Rare structural variants found in attention-deficit hyperactivity disorder are preferentially associated with neurodevelopmental genes. *Molecular Psychiatry* Published online: 23 June 2009. doi:10.1038/mp.2009.57.

Erlenmeyer-Kimling L, Rock D, Roberts SA, Janal M, Kestenbaum C, Cornblatt B, Adamo UH, Gottesman II (2000). Attention, memory, and motor skills as childhood predictors of schizophrenia-related psychoses: the New York High-Risk Project. *American Journal of Psychiatry* **157**, 1416–1422.

Frazer KA, Murray SS, Schork NJ, Topol EJ (2009). Human genetic variation and its contribution to complex traits. *Nature Reviews Genetics* **10**, 241–251.

Friedman JI, Vrijenhoek T, Markx S, Janssen IM, van der Vliet WA, Faas BH, Knoers NV, Cahn W, Kahn RS, Edelmann L, Davis KL, Silverman JM, Brunner HG, van Kessel AG, Wijmenga C, Ophoff RA, Veltman JA (2008). CNTNAP2 gene dosage variation is associated with schizophrenia and epilepsy. *Molecular Psychiatry* **13**, 261–266.

Funayama M, Li Y, Tsoi TH, Lam CW, Ohi T, Yazawa S, Uyama E, Djaldetti R, Melamed E, Yoshino H, Imamichi Y, Takashima H, Nishioka K, Sato K, Tomiyama H, Kubo S, Mizuno Y, Hattori N (2008). Familial Parkinsonism with digenic parkin and PINK1 mutations. *Movement Disorders* **23**, 1461–1465.

Girirajan S, Rosenfeld JA, Cooper GM, Antonacci F, Siswara P, Itsara A, Vives L, Walsh T, McCarthy SE, Baker C, Mefford HC, Kidd JM, Browning SR, Browning BL, Dickel DE, Levy DL, Ballif BC, Platky K, Farber DM, Gowans GC, Wetherbee JJ, Asamoah A, Weaver DD, Mark PR, Dickerson J, Garg BP, Ellingwood SA, Smith R, Banks VC, Smith W, McDonald MT, Hoo JJ, French BN, Hudson C, Johnson JP, Ozmore JR, Moeschler JB, Surti U, Escobar LF, El-Khechen D, Gorski JL, Kussmann J, Salbert B, Lacassie Y, Biser A, McDonald-McGinn DM, Zackai EH, Deardorff MA, Shaikh TH, Haan E, Friend KL, Fichera M, Romano C, Gecz J, DeLisi LE, Sebat J, King MC, Shaffer LG, Eichler EE (2010). A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay. *Nature Genetics* **42**, 203–209.

Glessner JT, Wang K, Cai G, Korvatska O, Kim CE, Wood S, Zhang H, Estes A, Brune CW, Bradfield JP, Imielinski M, Frackelton EC, Reichert J, Crawford EL, Munson J, Sleiman PM, Chiavacci R, Annaiah K, Thomas K, Hou C, Glaberson W, Flory J, Otieno F, Garris M, Soorya L, Klei L, Piven J, Meyer KJ, Anagnostou E, Sakurai T, Game RM, Rudd DS, Zurawiecki D, McDougle CJ, Davis LK, Miller J, Posey DJ, Michaels S, Kolevzon A, Silverman JM, Bernier R, Levy SE, Schultz RT, Dawson G, Owley T, McMahon WM, Wassink TH, Sweeney JA, Nurnberger JI, Coon H, Sutcliffe JS, Minshew NJ, Grant SF, Bucan M, Cook EH, Buxbaum JD, Devlin B, Schellenberg GD, Hakonarson H (2009). Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature* **459**, 569–573.

Goldstein DB (2009). Common genetic variation and human traits. *New England Journal of Medicine* **360**, 1696–1698.

Gottesman II, Gould TD (2003). The endophenotype concept in psychiatry: etymology and strategic intentions. *American Journal of Psychiatry* **160**, 636–645.

Gottesman II, Laursen TM, Bertelsen A, Mortensen PB (2010). Severe mental disorders in offspring with 2 psychiatrically ill parents. *Archives of General Psychiatry* **67**, 252–257.

Gottesman II, Shields J (1967). A polygenic theory of schizophrenia. *Proceedings of the National Academy of Sciences USA* **58**, 199–205.

Gregory I (1960). Genetic factors in schizophrenia. *American Journal of Psychiatry* **116**, 961–972.

Guilmatre A, Dubourg C, Mosca AL, Legallic S, Goldenberg A, Drouin-Garraud V, Layet V, Rosier A, Briault S, Bonnet-Brilhault F, Laumonnier F, Odent S, Le Vacon G, Joly-Helas G, David V, Bendavid C, Pinoit JM, Henry C, Impallomeni C, Germano E, Tortorella G, Di Rosa G, Barthelemy C, Andres C, Faivre L, Frebourg T, Saugier Veber P, Campion D (2009). Recurrent rearrangements in synaptic and neurodevelopmental genes and shared biologic pathways in schizophrenia, autism, and mental retardation. *Archives of General Psychiatry* **66**, 947–956.

Hannes FD, Sharp AJ, Mefford HC, de Ravel T, Ruivenkamp CA, Breuning MH, Fryns JP, Devriendt K, Van Buggenhout G, Vogels A, Stewart H, Hennekam RC, Cooper GM, Regan R, Knight SJ, Eichler EE, Vermeesch JR (2009). Recurrent reciprocal deletions and duplications of 16p13.11: the deletion is a risk factor for MR/MCA while the duplication may be a rare benign variant. *Journal of Medical Genetics* **46**, 223–232.

International Schizophrenia Consortium (2008). Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* **455**, 237–241.

James JW (1971). Frequency in relatives for an all-or-none trait. *Annals of Human Genetics* **35**, 47–49.

Ji W, Foo JN, O'Roak BJ, Zhao H, Larson MG, Simon DB, Newton-Cheh C, State MW, Levy D, Lifton RP (2008). Rare independent mutations in renal salt handling genes contribute to blood pressure variation. *Nature Genetics* **40**, 592–599.

Kacser H, Burns JA (1981). The molecular basis of dominance. *Genetics* **97**, 639–666.

Kamiya A, Tan PL, Kubo K, Engelhard C, Ishizuka K, Kubo A, Tsukita S, Pulver AE, Nakajima K, Cascella NG, Katsanis N, Sawa A (2008). Recruitment of PCM1 to the

centrosome by the cooperative action of DISC1 and BBS4: a candidate for psychiatric illnesses. *Archives of General Psychiatry* **65**, 996–1006.

Keller MC, Miller G (2006). Resolving the paradox of common, harmful, heritable mental disorders: which evolutionary genetic models work best? *Behavioral and Brain Sciences* **29**, 385–404; discussion 405–452.

Kendler KS, Kidd KK (1986). Recurrence risks in an oligogenic threshold model: the effect of alterations in allele frequency. *Annals of Human Genetics* **50**, 83–91.

Kerszberg M (2004). Noise, delays, robustness, canalization and all that. *Current Opinion in Genetics and Development* **14**, 440–445.

Kinney DK, Teixeira P, Hsu D, Napoleon SC, Crowley DJ, Miller A, Hyman W, Huang E (2009). Relation of schizophrenia prevalence to latitude, climate, fish consumption, infant mortality, and skin color: a role for prenatal vitamin D deficiency and infections? *Schizophrenia Bulletin* **35**, 582–595.

Kirov G, Grozeva D, Norton N, Ivanov D, Mantripragada KK, Holmans P, Craddock N, Owen MJ, O'Donovan MC (2009). Support for the involvement of large copy number variants in the pathogenesis of schizophrenia. *Human Molecular Genetics* **18**, 1497–1503.

Kirov G, Gumus D, Chen W, Norton N, Georgieva L, Sari M, O'Donovan MC, Erdogan F, Owen MJ, Ropers HH, Ullmann R (2008). Comparative genome hybridization suggests a role for NRXN1 and APBA2 in schizophrenia. *Human Molecular Genetics* **17**, 458–465.

Kitano H (2004). Biological robustness. *Nature Reviews Genetics* **5**, 826–837.

Knight HM, Pickard BS, Maclean A, Malloy MP, Soares DC, McRae AF, Condie A, White A, Hawkins W, McGhee K, van Beck M, Macintyre DJ, Starr JM, Deary IJ, Visscher PM, Porteous DJ, Cannon RE, St Clair D, Muir WJ, Blackwood DH (2009). A cytogenetic abnormality and rare coding variants identify ABCA13 as a candidate gene in schizophrenia, bipolar disorder, and depression. *American Journal of Human Genetics* **85**, 833–846.

Kryukov GV, Pennacchio LA, Sunyaev SR (2007). Most rare missense alleles are deleterious in humans: implications for complex disease and association studies. *American Journal of Human Genetics* **80**, 727–739.

Lichtenstein P, Yip BH, Bjork C, Pawitan Y, Cannon TD, Sullivan PF, Hultman CM (2009). Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet* **373**, 234–239.

Lupski JR (2007). Genomic rearrangements and sporadic disease. *Nature Genetics* **39**, S43–S47.

Mah S, Nelson MR, Delisi LE, Reneland RH, Markward N, James MR, Nyholt DR, Hayward N, Handoko H, Mowry B, Kammerer S, Braun A (2006). Identification of the semaphorin receptor PLXNA2 as a candidate for susceptibility to schizophrenia. *Molecular Psychiatry* **11**, 471–478.

Malaspina D, Harlap S, Fennig S, Heiman D, Nahon D, Feldman D, Susser ES (2001). Advancing paternal age and the risk of schizophrenia. *Archives of General Psychiatry* **58**, 361–367.

Marshall CR, Noor A, Vincent JB, Lionel AC, Feuk L, Skaug J, Shago M, Moessner R, Pinto D, Ren Y, Thiruvahindrapduram B, Fiebig A, Schreiber S, Friedman J, Ketelaars CE, Vos YJ, Ficicioglu C, Kirkpatrick S, Nicolson R, Sloman L, Summers A, Gibbons CA, Teebi A, Chitayat D, Weksberg R, Thompson A, Vardy C, Crosbie V, Luscombe S, Baatjes R, Zwaigenbaum L, Roberts W, Fernandez B, Szatmari P, Scherer SW (2008). Structural variation of chromosomes in autism spectrum disorder. *American Journal of Human Genetics* **82**, 477–488.

McClellan JM, Susser E, King MC (2007). Schizophrenia: a common disease caused by multiple rare alleles. *British Journal of Psychiatry* **190**, 194–199.

McGue M, Gottesman II, Rao DC (1985). Resolving genetic models for the transmission of schizophrenia. *Genetic Epidemiology* **2**, 99–110.

Mefford HC, Cooper GM, Zerr T, Smith JD, Baker C, Shafer N, Thorland EC, Skinner C, Schwartz CE, Nickerson DA, Eichler EE (2009). A method for rapid, targeted CNV genotyping identifies rare variants associated with neurocognitive disease. *Genome Research* **19**, 1579–1585.

Mefford HC, Sharp AJ, Baker C, Itsara A, Jiang Z, Buysse K, Huang S, Maloney VK, Crolla JA, Baralle D, Collins A, Mercer C, Norga K, de Ravel T, Devriendt K, Bongers EM, de Leeuw N, Reardon W, Gimelli S, Bena F, Hennekam RC, Male A, Gaunt L, Clayton-Smith J, Simonic I, Park SM, Mehta SG, Nik-Zainal S, Woods CG, Firth HV, Parkin G, Fichera M, Reitano S, Lo Giudice M, Li KE, Casuga I, Broomer A, Conrad B, Schwerzmann M, Raber L, Gallati S, Striano P, Coppola A, Tolmie JL, Tobias ES, Lilley C, Armengol L, Spysschaert Y, Verloo P, De Coene A, Goossens L, Mortier G, Speleman F, van Binsbergen E, Nelen MR, Hochstenbach R, Poot M, Gallagher L, Gill M, McClellan J, King MC, Regan R, Skinner C, Stevenson RE, Antonarakis SE, Chen C, Estivill X, Menten B, Gimelli G, Gribble S, Schwartz S, Sutcliffe JS, Walsh T, Knight SJ, Sebat J, Romano C, Schwartz CE, Veltman JA, de Vries BB, Vermeesch JR, Barber JC, Willatt L, Tassabehji M, Eichler EE (2008). Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes. *New England Journal of Medicine* **359**, 1685–1699.

Merikangas KR (1982). Assortative mating for psychiatric disorders and psychological traits. *Archives of General Psychiatry* **39**, 1173–1180.

Millar JK, Pickard BS, Mackie S, James R, Christie S, Buchanan SR, Malloy MP, Chubb JE, Huston E, Baillie GS, Thomson PA, Hill EV, Brandon NJ, Rain JC, Camargo LM, Whiting PJ, Houslay MD, Blackwood DH, Muir WJ, Porteous DJ (2005). DISC1 and PDE4B are interacting genetic factors in schizophrenia that regulate cAMP signaling. *Science* **310**, 1187–1191.

Millar JK, Wilson-Annan JC, Anderson S, Christie S, Taylor MS, Semple CA, Devon RS, Clair DM, Muir WJ, Blackwood DH, Porteous DJ (2000). Disruption of two novel genes by a translocation co-segregating with schizophrenia. *Human Molecular Genetics* **9**, 1415–1423.

Miller DT, Shen Y, Weiss LA, Korn J, Anselm I, Bridgemohan C, Cox GF, Dickinson H, Gentile J, Harris DJ, Hegde V, Hundley R, Khwaja O, Kothare S, Luedke C, Nasir R, Poduri A, Prasad K, Raffalli P, Reinhard A, Smith SE, Sobeih MM, Soul JS, Stoler J, Takeoka M, Tan WH, Thakuria J, Wolff R, Yusupov R, Gusella JF, Daly MJ, Wu BL (2009). Microdeletion/duplication at 15q13.2q13.3 among individuals with features of autism and other neuropsychiatric disorders. *Journal of Medical Genetics* **46**, 242–248.

Mitchell KJ (2007). The genetics of brain wiring: from molecule to mind. *PLoS Biology* **5**, e113.

Nadeau JH (2001). Modifier genes in mice and humans. *Nature Reviews Genetics* **2**, 165–174.

Need AC, Ge D, Weale ME, Maia J, Feng S, Heinzen EL, Shianna KV, Yoon W, Kasperaviciute D, Gennarelli M, Strittmatter WJ, Bonvicini C, Rossi G, Jayathilake K, Cola PA, McEvoy JP, Keefe RS, Fisher EM, St Jean PL, Giegling I, Hartmann AM, Moller HJ, Ruppert A, Fraser G, Crombie C, Middleton LT, St Clair D, Roses AD, Muglia P, Francks C, Rujescu D, Meltzer HY, Goldstein DB (2009). A genome-wide investigation of SNPs and CNVs in schizophrenia. *PLoS Genetics* **5**, e1000373.

Ng PC, Levy S, Huang J, Stockwell TB, Walenz BP, Li K, Axelrod N, Busam DA, Strausberg RL, Venter JC (2008). Genetic variation in an individual human exome. *PLoS Genetics* **4**, e1000160.

O'Donovan MC, Craddock N, Norton N, Williams H, Peirce T, Moskvina V, Nikolov I, Hamshere M, Carroll L, Georgieva L, Dwyer S, Holmans P, Marchini JL, Spencer CC, Howie B, Leung HT, Hartmann AM, Moller HJ, Morris DW, Shi Y, Feng G, Hoffmann P, Propping P, Vasilescu C, Maier W, Rietschel M, Zammit S, Schumacher J, Quinn EM, Schulze TG, Williams NM, Giegling I, Iwata N, Ikeda M, Darvasi A, Shifman S, He L, Duan J, Sanders AR, Levinson DF, Gejman PV, Cichon S, Nothen MM, Gill M, Corvin A, Rujescu D, Kirov G, Owen MJ, Buccola NG, Mowry BJ, Freedman R, Amin F, Black DW, Silverman JM, Byerley WF, Cloninger CR (2008). Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nature Genetics* **40**, 1053–1055.

O'Rourke DH, Gottesman II, Suarez BK, Rice J, Reich T (1982). Refutation of the general single-locus model for the etiology of schizophrenia. *American Journal of Human Genetics* **34**, 630–649.

Owens DG, Johnstone EC (2006). Precursors and prodromata of schizophrenia: findings from the Edinburgh High Risk Study and their literature context. *Psychological Medicine* **36**, 1501–1514.

Park JP, Moeschler JB, Berg SZ, Wurster-Hill DH (1991). Schizophrenia and mental retardation in an adult male with a de novo interstitial deletion 9(q32q34.1). *Journal of Medical Genetics* **28**, 282–283.

Pickard BS, Hollox EJ, Malloy MP, Porteous DJ, Blackwood DH, Armour JA, Muir WJ (2004). A 4q35.2 subtelomeric deletion identified in a screen of patients with co-morbid psychiatric illness and mental retardation. *BMC Medical Genetics* **5**, 21.

Pickard BS, Malloy MP, Christoforou A, Thomson PA, Evans KL, Morris SW, Hampson M, Porteous DJ, Blackwood DH, Muir WJ (2006). Cytogenetic and genetic evidence supports a role for the kainate-type glutamate receptor gene, GRIK4, in schizophrenia and bipolar disorder. *Molecular Psychiatry* **11**, 847–857.

Pickard BS, Malloy MP, Porteous DJ, Blackwood DH, Muir WJ (2005). Disruption of a brain transcription factor, NPAS3, is associated with schizophrenia and learning disability. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics* **136B**, 26–32.

Porteous D (2008). Genetic causality in schizophrenia and bipolar disorder: out with the old and in with the new. *Current Opinion in Genetics and Development* **18**, 229–234.

Prasad KM, Sanders R, Sweeney J, Montrose D, Diwadkar V, Dworakowski D, Miewald J, Keshavan M (2009). Neurological abnormalities among offspring of persons with schizophrenia: relation to premorbid psychopathology. *Schizophrenia Research* **108**, 163–169.

Psychiatric GWAS Consortium Steering Committee (2009). A framework for interpreting genome-wide association studies of psychiatric disorders. *Molecular Psychiatry* **14**, 10–17.

Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, Sklar P (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748–752.

Qin P, Xu H, Laursen TM, Vestergaard M, Mortensen PB (2005). Risk for schizophrenia and schizophrenia-like psychosis among patients with epilepsy: population based cohort study. *British Medical Journal* **331**, 23.

Raychaudhuri S, Plenge RM, Rossin EJ, Ng AC, Purcell SM, Sklar P, Scolnick EM, Xavier RJ, Altshuler D, Daly MJ (2009). Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions. *PLoS Genetics* **5**, e1000534.

Reich DE, Lander ES (2001). On the allelic spectrum of human disease. *Trends in Genetics* **17**, 502–510.

Riley B, Asherson PJ, McGuffin P (2003). Genetics and schizophrenia. In *Schizophrenia* (ed. S. R. Hirsch and D. R. Weinberger), pp. 251–276. Wiley-Blackwell: Malden, MA.

Risch N (1990*a*). Genetic linkage and complex diseases, with special reference to psychiatric disorders. *Genetic Epidemiology* **7**, 3–16; discussion 17–45.

Risch N (1990*b*). Linkage strategies for genetically complex traits. I. Multilocus models. *American Journal of Human Genetics* **46**, 222–228.

Risch N, Baron M (1984). Segregation analysis of schizophrenia and related disorders. *American Journal of Human Genetics* **36**, 1039–1059.

Risch N, Merikangas K (1996). The future of genetic studies of complex human diseases. *Science* **273**, 1516–1517.

Rodriguez-Santiago B, Brunet A, Sobrino B, Serra-Juhe C, Flores R, Armengol L, Vilella E, Gabau E, Guitart M, Guillamat R, Martorell L, Valero J, Gutierrez-Zotes A, Labad A, Carracedo A, Estivill X, Perez-Jurado LA (2009). Association of common copy number variants at the glutathione S-transferase genes and rare novel genomic

changes with schizophrenia. *Molecular Psychiatry*. Published online: 16 June 2009. doi:10.1038/mp.2009.53.

Rujescu D, Ingason A, Cichon S, Pietilainen OP, Barnes MR, Toulopoulou T, Picchioni M, Vassos E, Ettinger U, Bramon E, Murray R, Ruggeri M, Tosato S, Bonetto C, Steinberg S, Sigurdsson E, Sigmundsson T, Petursson H, Gylfason A, Olason PI, Hardarsson G, Jonsdottir GA, Gustafsson O, Fossdal R, Giegling I, Moller HJ, Hartmann AM, Hoffmann P, Crombie C, Fraser G, Walker N, Lonnqvist J, Suvisaari J, Tuulio-Henriksson A, Djurovic S, Melle I, Andreassen OA, Hansen T, Werge T, Kiemeney LA, Franke B, Veltman J, Buizer-Voskamp JE, Sabatti C, Ophoff RA, Rietschel M, Nothen MM, Stefansson K, Peltonen L, St Clair D, Stefansson H, Collier DA (2009). Disruption of the neurexin 1 gene is associated with schizophrenia. *Human Molecular Genetics* **18**, 988–996.

Saperstein AM, Fuller RL, Avila MT, Adami H, McMahon RP, Thaker GK, Gold JM (2006). Spatial working memory as a cognitive endophenotype of schizophrenia: assessing risk for pathophysiological dysfunction. *Schizophrenia Bulletin* **32**, 498–506.

Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, Maner S, Massa H, Walker M, Chi M, Navin N, Lucito R, Healy J, Hicks J, Ye K, Reiner A, Gilliam TC, Trask B, Patterson N, Zetterberg A, Wigler M (2004). Large-scale copy number polymorphism in the human genome. *Science* **305**, 525–528.

Sebat J, Levy DL, McCarthy SE (2009). Rare structural variants in schizophrenia: one disorder, multiple mutations; one mutation, multiple disorders. *Trends in Genetics* **25**, 528–535.

Shao H, Burrage LC, Sinasac DS, Hill AE, Ernest SR, O'Brien W, Courtland HW, Jepsen KJ, Kirby A, Kulbokas EJ, Daly MJ, Broman KW, Lander ES, Nadeau JH (2008). Genetic architecture of complex traits: large phenotypic effects and pervasive epistasis. *Proceedings of the National Academy of Sciences USA* **105**, 19910–19914.

Shi J, Levinson DF, Duan J, Sanders AR, Zheng Y, Pe'er I, Dudbridge F, Holmans PA, Whittemore AS, Mowry BJ, Olincy A, Amin F, Cloninger CR, Silverman JM, Buccola NG, Byerley WF, Black DW, Crowe RR, Oksenberg JR, Mirel DB, Kendler KS, Freedman R, Gejman PV (2009). Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature* **460**, 753–757.

Shi YY, He G, Zhang Z, Tang W, Zhang Jr. J, Zhao Q, Zhang Sr. J, Li XW, Xi ZR, Fang C, Zhao XZ, Feng GY, He L (2008). A study of rare structural variants in schizophrenia patients and normal controls from Chinese Han population. *Molecular Psychiatry* **13**, 911–913.

Siegal ML, Bergman A (2002). Waddington's canalization revisited: developmental stability and evolution. *Proceedings of the National Academy of Sciences USA* **99**, 10528–10532.

Song W, Li W, Feng J, Heston LL, Scaringe WA, Sommer SS (2008). Identification of high risk DISC1 structural variants with a 2% attributable risk for schizophrenia. *Biochemical and Biophysical Research Communications* **367**, 700–706.

Stefansson H, Ophoff RA, Steinberg S, Andreassen OA, Cichon S, Rujescu D, Werge T, Pietilainen OP, Mors O, Mortensen PB, Sigurdsson E, Gustafsson O, Nyegaard M, Tuulio-Henriksson A, Ingason A, Hansen T, Suvisaari J, Lonnqvist J, Paunio T, Borglum AD, Hartmann A, Fink-Jensen A, Nordentoft M, Hougaard D, Norgaard-Pedersen B, Bottcher Y, Olesen J, Breuer R, Moller HJ, Giegling I, Rasmussen HB, Timm S, Mattheisen M, Bitter I, Rethelyi JM, Magnusdottir BB, Sigmundsson T, Olason P, Masson G, Gulcher JR, Haraldsson M, Fossdal R, Thorgeirsson TE, Thorsteinsdottir U, Ruggeri M, Tosato S, Franke B, Strengman E, Kiemeney LA, Melle I, Djurovic S, Abramova L, Kaleda V, Sanjuan J, de Frutos R, Bramon E, Vassos E, Fraser G, Ettinger U, Picchioni M, Walker N, Toulopoulou T, Need AC, Ge D, Yoon JL, Shianna KV, Freimer NB, Cantor RM, Murray R, Kong A, Golimbet V, Carracedo A, Arango C, Costas J, Jonsson EG, Terenius L, Agartz I, Petursson H, Nothen MM, Rietschel M, Matthews PM, Muglia P, Peltonen L, St Clair D, Goldstein DB, Stefansson K, Collier DA (2009). Common variants conferring risk of schizophrenia. *Nature* **460**, 744–747.

Stefansson H, Rujescu D, Cichon S, Pietilainen OP, Ingason A, Steinberg S, Fossdal R, Sigurdsson E, Sigmundsson T, Buizer-Voskamp JE, Hansen T, Jakobsen KD, Muglia P, Francks C, Matthews PM, Gylfason A, Halldorsson BV, Gudbjartsson D, Thorgeirsson TE, Sigurdsson A, Jonasdottir A, Bjornsson A, Mattiasdottir S, Blondal T, Haraldsson M, Magnusdottir BB, Giegling I, Moller HJ, Hartmann A, Shianna KV, Ge D, Need AC, Crombie C, Fraser G, Walker N, Lonnqvist J, Suvisaari J, Tuulio-Henriksson A, Paunio T, Toulopoulou T, Bramon E, Di Forti M, Murray R, Ruggeri M, Vassos E, Tosato S, Walshe M, Li T, Vasilescu C, Muhleisen TW, Wang AG, Ullum H, Djurovic S, Melle I, Olesen J, Kiemeney LA, Franke B, Sabatti C, Freimer NB, Gulcher JR, Thorsteinsdottir U, Kong A, Andreassen OA, Ophoff RA, Georgi A, Rietschel M, Werge T, Petursson H, Goldstein DB, Nothen MM, Peltonen L, Collier DA, St Clair D, Stefansson K (2008). Large recurrent microdeletions associated with schizophrenia. *Nature* **455**, 232–236.

Steinhausen HC, Foldager L, Perto G, Munk-Jorgensen P (2009). Family aggregation of mental disorders in the nationwide Danish three generation study. *European Archives of Psychiatry and Clinical Neuroscience* **259**, 270–277.

Steinlechner S, Stahlberg J, Volkel B, Djarmati A, Hagenah J, Hiller A, Hedrich K, Konig I, Klein C, Lencer R (2007). Co-occurrence of affective and schizophrenia spectrum disorders with PINK1 mutations. *Journal of Neurology, Neurosurgery, and Psychiatry* **78**, 532–535.

Tandon R, Keshavan MS, Nasrallah HA (2008). Schizophrenia, 'just the facts' what we know in 2008. 2. Epidemiology and etiology. *Schizophrenia Research* **102**, 1–18.

Tsuang MT, Bucher KD, Fleming JA (1982). Testing the monogenic theory of schizophrenia: an application of

segregation analysis to blind family study data. *British Journal of Psychiatry* **140**, 595–599.

Turetsky BI, Greenwood TA, Olincy A, Radant AD, Braff DL, Cadenhead KS, Dobie DJ, Freedman R, Green MF, Gur RE, Gur RC, Light GA, Mintz J, Nuechterlein KH, Schork NJ, Seidman LJ, Siever LJ, Silverman JM, Stone WS, Swerdlow NR, Tsuang DW, Tsuang MT, Calkins ME (2008). Abnormal auditory N100 amplitude: a heritable endophenotype in first-degree relatives of schizophrenia probands. *Biological Psychiatry* **64**, 1051–1059.

Tyson C, Harvard C, Locker R, Friedman JM, Langlois S, Lewis ME, Van Allen M, Somerville M, Arbour L, Clarke L, McGilivray B, Yong SL, Siegel-Bartel J, Rajcan-Separovic E (2005). Submicroscopic deletions and duplications in individuals with intellectual disability detected by array-CGH. *American Journal of Medical Genetics. Part A* **139**, 173–185.

van Bon BW, Mefford HC, Menten B, Koolen DA, Sharp AJ, Nillesen WM, Innis JW, de Ravel TJ, Mercer CL, Fichera M, Stewart H, Connell LE, Ounap K, Lachlan K, Castle B, Van der Aa N, van Ravenswaaij C, Nobrega MA, Serra-Juhe C, Simonic I, de Leeuw N, Pfundt R, Bongers EM, Baker C, Finnemore P, Huang S, Maloney VK, Crolla JA, van Kalmthout M, Elia M, Vandeweyer G, Fryns JP, Janssens S, Foulds N, Reitano S, Smith K, Parkel S, Loeys B, Woods CG, Oostra A, Speleman F, Pereira AC, Kurg A, Willatt L, Knight SJ, Vermeesch JR, Romano C, Barber JC, Mortier G, Perez-Jurado LA, Kooy F, Brunner HG, Eichler EE, Kleefstra T, de Vries BB (2009). Further delineation of the 15q13 microdeletion and duplication syndromes: a clinical spectrum varying from non-pathogenic to a severe outcome. *Journal of Medical Genetics* **46**, 511–523.

Verkerk AJ, Mathews CA, Joosse M, Eussen BH, Heutink P, Oostra BA (2003). CNTNAP2 is disrupted in a family with Gilles de la Tourette syndrome and obsessive compulsive disorder. *Genomics* **82**, 1–9.

Verma R, Chauhan C, Saleem Q, Gandhi C, Jain S, Brahmachari SK (2004). A nonsense mutation in the synaptogyrin 1 gene in a family with schizophrenia. *Biological Psychiatry* **55**, 196–199.

Vrijenhoek T, Buizer-Voskamp JE, van der Stelt I, Strengman E, Sabatti C, Geurts van Kessel A, Brunner HG, Ophoff RA, Veltman JA (2008). Recurrent CNVs disrupt three candidate genes in schizophrenia patients. *American Journal of Human Genetics* **83**, 504–510.

Waddington CH (1959). Canalization of development and genetic assimilation of acquired characters. *Nature* **183**, 1654–1655.

Waddington JL, Corvin AP, Donohoe G, O'Tuathaigh CM, Mitchell KJ, Gill M (2007). Functional genomics and schizophrenia: endophenotypes and mutant models. *Psychiatric Clinics of North America* **30**, 365–399.

Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM, Nord AS, Kusenda M, Malhotra D, Bhandari A, Stray SM, Rippey CF, Roccanova P, Makarov V, Lakshmi B, Findling RL, Sikich L, Stromberg T, Merriman B, Gogtay N, Butler P, Eckstrand K, Noory L, Gochman P, Long R, Chen Z, Davis S, Baker C, Eichler EE, Meltzer PS, Nelson SF, Singleton AB, Lee MK, Rapoport JL, King MC, Sebat J (2008). Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* **320**, 539–543.

Weiss LA, Shen Y, Korn JM, Arking DE, Miller DT, Fossdal R, Saemundsen E, Stefansson H, Ferreira MA, Green T, Platt OS, Ruderfer DM, Walsh CA, Altshuler D, Chakravarti A, Tanzi RE, Stefansson K, Santangelo SL, Gusella JF, Sklar P, Wu BL, Daly MJ (2008). Association between microdeletion and microduplication at 16p11.2 and autism. *New England Journal of Medicine* **358**, 667–675.

Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, He W, Chen YJ, Makhijani V, Roth GT, Gomes X, Tartaro K, Niazi F, Turcotte CL, Irzyk GP, Lupski JR, Chinault C, Song XZ, Liu Y, Yuan Y, Nazareth L, Qin X, Muzny DM, Margulies M, Weinstock GM, Gibbs RA, Rothberg JM (2008). The complete genome of an individual by massively parallel DNA sequencing. *Nature* **452**, 872–876.

Wohl M, Gorwood P (2007). Paternal ages below or above 35 years old are associated with a different risk of schizophrenia in the offspring. *European Psychiatry* **22**, 22–26.

Woodberry KA, Giuliano AJ, Seidman LJ (2008). Premorbid IQ in schizophrenia: a meta-analytic review. *American Journal of Psychiatry* **165**, 579–587.

Woolf CM (1997). Does the genotype for schizophrenia often remain unexpressed because of canalization and stochastic events during development? *Psychological Medicine* **27**, 659–668.

Xu B, Roos JL, Levy S, van Rensburg EJ, Gogos JA, Karayiorgou M (2008). Strong association of de novo copy number mutations with sporadic schizophrenia. *Nature Genetics* **40**, 880–885.

Xu B, Woodroffe A, Rodriguez-Murillo L, Roos JL, van Rensburg EJ, Abecasis GR, Gogos JA, Karayiorgou M (2009). Elucidating the genetic architecture of familial schizophrenia using rare copy number variant and linkage scans. *Proceedings of the National Academy of Sciences USA* **106**, 16746–16751.