Rare chromosomal deletions and duplications increase risk of schizophrenia

The International Schizophrenia Consortium*

Abstract

Schizophrenia is a severe mental disorder marked by hallucinations, delusions, cognitive deficits and apathy, with a heritability estimated at 73–90% (ref. 1). Inheritance patterns are complex, and the number and type of genetic variants involved are not understood. Copy number variants (CNVs) have been identified in individual patients with schizophrenia2, 3, 4, 5, 6, 7 and also in neurodevelopmental disorders8, 9, 10, 11, but large-scale genome-wide surveys have not been performed. Here we report a genome-wide survey of rare CNVs in 3,391 patients with schizophrenia and 3,181 ancestrally matched controls, using high-density microarrays. For CNVs that were observed in less than 1% of the sample and were more than 100 kilobases in length, the total burden is increased 1.15-fold in patients with schizophrenia in comparison with controls. This effect was more pronounced for rarer, single-occurrence CNVs and for those that involved genes as opposed to those that did not. As expected, deletions were found within the region critical for velo-cardio-facial syndrome, which includes psychotic symptoms in 30% of patients12. Associations with schizophrenia were also found for large deletions on chromosome 15q13.3 and 1q21.1. These associations have not previously been reported, and they remained significant after genome-wide correction. Our results provide strong support for a model of schizophrenia pathogenesis that includes the effects of multiple rare structural variants, both genome-wide and at specific loci.

The International Schizophrenia Consortium was established to promote rapid progress towards the identification of genetic causes underlying schizophrenia. The consortium is composed of investigators from the University of Aberdeen, Cardiff University, the University of Edinburgh, Karolinska Institutet, Massachusetts General Hospital, the University of North Carolina-Chapel Hill, the Queensland Institute of Medical Research, the University of Southern California, the Stanley Center for Psychiatric Research at the Broad Institute of Harvard and MIT, Trinity College Dublin and University College London.

We surveyed single nucleotide polymorphisms (SNPs) and CNVs using the Affymetrix Genome-Wide Human SNP 5.0 and 6.0 arrays in European cases of schizophrenia and in ancestrally matched controls (Table 1 and Supplementary Information)13. On the basis of the genome-wide SNP data there was no evidence of major population stratification within each site14 (data not shown). Intensity data from both SNP and CNV probes were used to identify autosomal deletions and duplications, based on a hidden Markov model15.

Table 1: Study sample characteristics and genotyping platforms

Sample	Ancestry	Case (*n*)	Control (*n*)	Genotyping platform
Figures are the numbers of cases and controls passing quality control and included in the final analyses. Case samples received a diagnosis of schizophrenia. 'Genotyping platform' indicates Affymetrix array type (5.0 or 6.0).				
***Cases were excluded if IQ was less than 70.**				
†Controls were screened for psychiatric disorders.				
‡University College London control samples genotyped with the Affymetrix 500K two-chip genotyping platform were excluded because CNV data were not available.				
§Swedish cases and controls matched for array type for all analyses.				
University of Aberdeen	Scottish	727	694[†]	5.0
University College London	British	547	n/a[‡]	5.0
Portuguese Island Collection	Portuguese	333	200[†]	5.0
Karolinska Institutet	Swedish	622	437	5.0/6.0[§]
Cardiff University	Bulgarian	479[*]	646	6.0
Trinity College Dublin	Irish	280	914	6.0
University of Edinburgh	Scottish	403[*]	290	6.0

This study focused on rare but highly penetrant structural variation in schizophrenia, following a natural extension of the classical medical genetic approach. Common CNVs are better identified with different algorithms and are better tested for association separately13, 15. Considering CNVs that were present in less than 1% of our total sample, there were 6,753 larger than 100 kilobases (kb) that passed sample and CNV quality filtering (see Supplementary Information and Supplementary Table 1). The median size was 182.1 kb (166.3 kb for deletions, 194.4 kb for duplications), 39% were deletions and the median number per individual was 1. We assessed the impact of rare structural variation on the risk for schizophrenia in two ways: first in terms of an individual's genome-wide burden, and second by searching for specific loci that were significantly associated with disease.

Structural variants have been identified for severe neurodevelopmental disorders9, 10, 11, 16, 17. Because it has been postulated that schizophrenia might, at least in part, have a developmental aetiology18, we posited a role for CNVs in schizophrenia, as have others2, 3, 4, 5, 6. Several loci have been identified, including variants containing genes with neurodevelopmental roles2, 3, 4, 5. However, a critical question is the extent to which this is a general mechanism for producing schizophrenia in typical clinical populations rather than in cases selected for atypical phenotypic features such as very early onset or mental retardation. This motivated our primary hypothesis: that individuals with schizophrenia have a greater genome-wide burden of CNVs. Considering all CNVs, we observed that cases had a greater average burden than controls (one-sided, empirical P = 3 times 10-5 controlling for array type; Table 2). Controls on average had 0.99 CNVs per person, whereas cases showed a 1.15-fold higher rate.

Table 2: Global CNV burden analysis: event type and frequency

Table 2 | Global CNV burden analysis: event type and frequency

			CNV burden (number)			CNV burden (gene count)		
CNV type	Frequency	CNV (*n*)	*P*	Case/control ratio	Baseline rate (controls)	*P*	Case/control ratio	Baseline rate (controls)
Deletions and duplications	All	6,753	3×10^{-5}	1.15	0.99	2×10^{-6}	1.41	2.01
	Single occurrence	890	5×10^{-6}	1.45	0.11	0.0057	1.67	0.32
	2–6 occurrences	2,465	0.0013	1.16	0.35	5×10^{-4}	1.36	0.80
Deletions only	All	2,652	0.11	1.08	0.40	3×10^{-5}	1.55	0.72
	Single occurrence	470	0.011	1.29	0.06	0.005	1.77	0.12
	2–6 occurrences	994	0.048	1.15	0.15	0.13	1.38	0.21
Duplications only	All	4,101	2×10^{-5}	1.20	0.59	10^{-4}	1.28	1.94
	Single occurrence	734	8×10^{-6}	1.58	0.09	0.015	1.60	0.30
	2–6 occurrences	1,532	0.011	1.16	0.22	0.012	1.30	0.69

The table shows an analysis of global CNV burden in cases versus controls. As described in the text, CNVs have previously been filtered for a maximum of about 1% sample frequency. These analyses were further stratified according to type (deletions versus duplications) and frequency (single occurrences and CNVs observed two to six times). Empirical *P* values (one-sided, controlling for array type) are given for two measures of CNV burden (number of CNVs and number of genes affected by CNVs). The average rate in controls (baseline, number of CNVs per person) and the fold increase in cases (case/control ratio) are shown for each analysis. Note that the 'Deletions only' and 'Duplications only' counts are not expected to sum to the 'Deletions and duplications' count for the two lower-frequency groups (see Supplementary Information).

We next explored this subtle, but highly statistically significant, observation of increased burden. We defined burden in two ways: as the number of CNVs carried by an individual (as above), and also as the number of genes spanned by those CNVs. This second metric (the 'gene count') in fact showed a stronger association with schizophrenia (1.41-fold increase, empirical P = 2 times 10-6) than burden defined simply as the number of CNVs. Characteristics of CNV subgroups studied here are their frequency, type, size, and proximity to a gene (Tables 2 and 3, and Supplementary Table 2). We observed an increased burden across multiple independent subgroups of CNVs, a finding that was more pronounced for rarer CNVs and those involving genes. Deletions and duplications also had somewhat different profiles: the association of deletions varied more noticeably with respect to CNV size and proximity to a gene, whereas duplications showed a more uniform pattern.

Table 3: Global CNV burden analysis: event type and size

Table 3 | Global CNV burden analysis: event type and size

			CNV burden (number)			CNV burden (gene count)		
CNV type	Size range (kb)	CNV (*n*)	*P*	Case/control ratio	Baseline rate (controls)	*P*	Case/control ratio	Baseline rate (controls)
Deletions and duplications	100 – 200	3,725	0.0017	1.15	0.55	8×10^{-6}	1.35	0.73
	200 – 500	2,156	0.028	1.11	0.32	0.0088	1.25	0.66
	≥500	872	0.0013	1.32	0.12	8×10^{-4}	1.79	0.62
Deletions only	100 – 200	1,612	0.54	1.02	0.25	0.28	1.07	0.26
	200 – 500	755	0.39	1.04	0.12	0.059	1.27	0.14
	≥500	285	3×10^{-4}	1.67	0.03	2×10^{-5}	3.57	0.14
Duplications only	100 – 200	2,113	10^{-4}	1.26	0.30	2×10^{-6}	1.50	0.47
	200 – 500	1,401	0.017	1.14	0.20	0.026	1.24	0.52
	≥500	587	0.11	1.17	0.09	0.17	1.29	0.48

The table shows an analysis of global CNV burden in cases versus controls. CNVs were stratified into three size categories (100–200 kb, 200–500 kb, and 500 kb or more). See Table 2 for further details.

A total of 890 CNVs were observed in either a case or a control as a single occurrence. This rarest subset of CNVs would be expected to show enrichment under the model that genetic causes of schizophrenia are individually unique in some proportion of patients. Indeed, this set of CNVs showed a 1.45-fold increase in cases (empirical P = 5 times 10-6). On average, 13.1% of cases of schizophrenia possessed a deletion or duplication observed only once in the sample, in contrast to 10.4% of controls. Under a model in which very rare (occurring in under 1/1,000 individuals) inherited or recurrently de novo events increase risk, we would expect to observe a greater overall burden in schizophrenia. Although our study was statistically underpowered to identify the actual loci involved, such variants could in theory be mapped in extremely large

samples. In this intermediate group, we observed 2,465 CNVs occurring between two and six times in the total sample, for which there was an increased burden, both for number of CNVs (empirical P = 0.0013) and gene count (empirical P = 5 times 10-4).

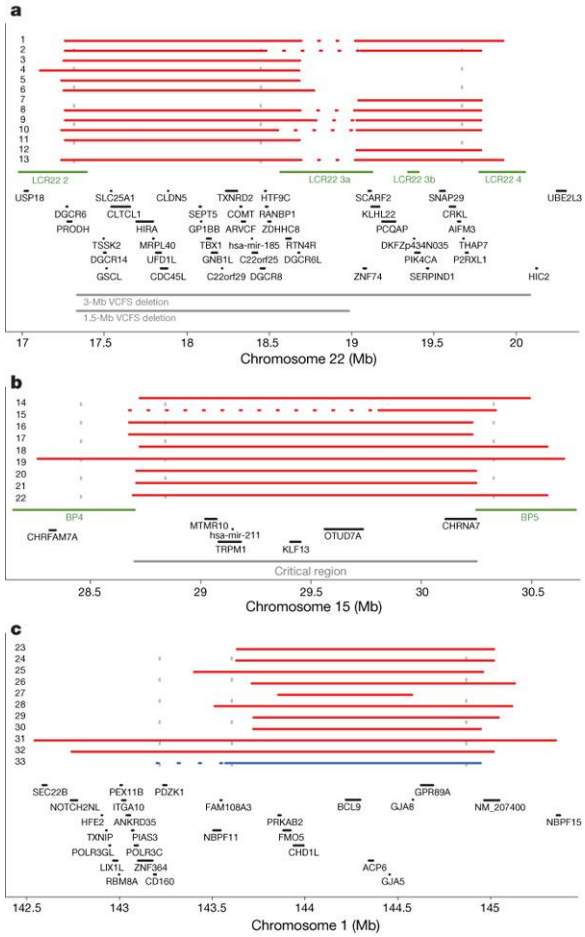
Because several known genomic disorders of the nervous system result from large CNVs, which are often many hundreds of kilobases long11, we additionally stratified by size of event (Table 3). Of deletions, only larger (more than 500 kb) variants were enriched (empirical P = 3 times 10-4) despite being the least frequent set of CNVs (n = 285), displaying a 3.57-fold increase in gene count between cases and controls (empirical P = 2 times 10-5). In contrast, shorter duplications showed a stronger association with disease than longer duplications, albeit with a smaller fold increase than deletions (Table 3).

In general, the gene count definition of CNV burden yielded stronger results, particularly for deletions (gene count P = 3 times 10-5 versus number P = 0.11; Table 2). In fact, dividing all CNVs into two sets, of those that intersect at least one gene and those that do not, we saw an increased burden only in the number of 'genic' CNVs (P = 5 times 10-6; Supplementary Table 2) and not for non-genic CNVs (P = 0.16). There was a similar trend for CNVs seen two to six times when comparing enrichment in genic and non-genic CNVs (P = 7 times 10-4 and 0.19, respectively) but not single-occurrence CNVs (P = 6 times 10-4 and 6 times 10-4, respectively). These results may reflect biological distinctions, although they may to some extent also reflect variable performance in CNV detection for different classes of variant.

We conducted a set of analyses to rule out several sources of bias and confounding in the primary genome-wide burden analysis (Supplementary Tables 3–6). Although, in general, low specificity and sensitivity decrease power, of concern here is potential measurement error that varied systematically between cases and controls, leading to spurious results. In this respect, an obvious concern is that both Affymetrix 5.0 and 6.0 arrays were used; as a consequence, we performed all analyses controlling for array type. As described in Supplementary Information, the primary result was also robust to the following. First, in addition to array type, we controlled for sample collection site, genotyping plate and average probe variance. Second, sensitivity analyses showed that no single sample collection site accounted for the observations. Third, we restricted analysis to the most homogeneous 90% of the sample with respect to intra-individual probe variance. Fourth, if differences in CNV burden between cases and controls were purely due to unmeasured confounders, we would not expect an enriched gene count after controlling for the overall extent and rate of CNVs. However, after controlling for overall (genic and non-genic) CNV burden there remained a significantly enriched gene-count burden in patients with schizophrenia.

Our large sample size further enabled us to search for specific CNV regions associated with schizophrenia. One locus previously reported to increase risk for schizophrenia is 22q11.2 (17–21 megabases (Mb)), at which hemizygosity occurs in 1 in every 4,000 live births12. These deletions produce a range of clinically heterogeneous phenotypes, including velo-cardio-facial syndrome and DiGeorge syndrome, that together are known as 22q11.2 deletion syndrome (22q11.2DS)12; about 30% of carriers develop psychosis12. Previous studies estimated the frequency of 22q11.2 deletions to be 0.6–2.0% in cases of schizophrenia, although many of these studies had technically incomplete characterization of this region19. We therefore expected to find examples of 22q11.2 deletions in our sample of 6,572 individuals. The most common form of 22q11.2DS is a 3-Mb loss (about 90% frequency), although a nested 1.5-Mb deletion is also observed (about 7%) along with infrequent (about 3%) atypical deletions20. We identified 13 large deletions (more than 500 kb) in cases of schizophrenia within this interval, and none in controls (Supplementary Table 7). Of these, six were consistent with the larger deletion, five were consistent with the shorter deletion, and two were atypical. The 11 samples with typical deletions defined an interval with the strongest association (empirical P = 0.0017; genome-wide corrected P = 0.0046; odds ratio = 21.6) (Fig. 1a). Controlling for sample collection site or genotyping plate instead of array type did not change the results (Supplementary Table 10). The two other atypical deletions in this region overlap the distal end of the 3-Mb variant. Deletion events within the region were confirmed in all 13 patients by quantitative polymerase chain reaction (qPCR) with three individual assays (Supplementary Fig. 1 and Supplementary Tables 11 and 12). Our findings provide additional evidence that hemizygosity in 22q11.2 is a rare but powerful risk factor for schizophrenia.

Figure 1: Regions with excess large deletions in cases.



a, The positions of CNVs of more than 500 kb across the chromosome 22q11.2 region. Red lines, case deletions; horizontal dashed sections, qPCR or visual inspection of array intensity data suggest an extended deletion; green lines, locations of low-copy repeats (LCR22-2–LCR22-4).; grey lines, recurrent 3-Mb and 1.5-Mb velo-cardio-facial syndrome (VCFS) deletions. qPCR primers are marked by vertical dashed lines. **b**, Chromosome 15q13.3 region, as above, except for the locations of breakpoint regions (BP4 and BP5; green) and the critical region defined previously[17] (grey line). **c**, Chromosome 1q21.1, as above, except that a single deletion identified in a control subject is marked by blue line. Genes based on build 35 UCSC browser (Supplementary Table 8) .

a, The positions of CNVs of more than 500 kb across the chromosome 22q11.2 region. Red lines, case deletions; horizontal dashed sections, qPCR or visual inspection of array intensity data suggest an extended deletion; green lines, locations of low-copy repeats (LCR22-2–LCR22-4).; grey lines, recurrent 3-Mb and 1.5-Mb velo-cardio-facial syndrome (VCFS) deletions. qPCR primers are marked by vertical dashed lines. b, Chromosome 15q13.3 region, as above, except for the locations of breakpoint regions (BP4 and BP5; green) and the critical region defined previously17 (grey line). c, Chromosome 1q21.1, as above, except that a single deletion identified in a control subject is marked by blue line. Genes based on build 35 UCSC browser (Supplementary Table 8).

High resolution image and legend (186K)Download Power Point slide (722K)

The larger 22q11.2 deletion harbours 43 genes (Supplementary Table 8). Despite the efforts of many groups, the psychiatric symptoms observed in 22q11.2DS have not been ascribed to a reduced copy number of any individual gene12. Variants within COMT, which encodes catechol-O-methyltransferase, an enzyme responsible for degrading catecholamines, including dopamine, have been implicated in a wide variety of phenotypes, but with inconsistent results12.

Removing the thirteen 22q11.2DS individuals, we observed a further 271 deletions of more than 500 kb (161 in cases and 110 in controls). Two additional regions (15q13.3 and 1q21.1) were identified that harboured a significant excess of deletions in cases of schizophrenia after correction for multiple testing (Fig. 1b, c; see Supplementary Table 7 for case descriptions). On chromosome 15 (28–31 Mb) there were deletions in nine cases and no controls (empirical $P = 0.0029$; genome-wide corrected $P = 0.046$; odds ratio 17.9). On chromosome 1 (142.5–145.5 Mb) there were ten deletions in cases and one in controls (empirical $P = 0.0076$; genome-wide corrected $P = 0.046$; odds ratio 6.6). All 20 large deletions at 15q13.3 and 1q21.1 were validated by one or more qPCR reactions (Supplementary Fig. 1 and Supplementary Tables 11 and 12). The multiple test correction factors were small as a consequence of our having restricted attention to this small class of rare variants. We did not observe any regions with a corrected $P < 0.05$ for either duplications or smaller (less than 500 kb) deletions. In addition, the primary CNV burden tests remained significant after removing individuals with a deletion at one of these three loci (number P = 10-4 and gene count P = 3 times 10-5); for deletions of more than 500 kb specifically, the burden test remained significant for number ($P = 0.02$) but not for gene count ($P = 0.11$).

The large deletions on chromosome 15q13.3 have not previously been associated with schizophrenia. This region does not include the nearby critical region for Prader–Willi/Angelman syndrome21 but is consistent with the critical region defined by recurrent deletion in cases of mental retardation with seizures that have been reported recently17. Furthermore, our estimated breakpoints fall within the segmental duplications reported (BP4 and BP5). In the present study, evidence consistent with mildly impaired cognition was seen in five of the nine patients with deletions, and one individual also had a history of epilepsy (Supplementary Table 7). This broad region has been the focus of previous genetic studies in schizophrenia. The gene CHRNA7, encoding the alpha7 subunit of the nicotinic acetylcholine receptor, is a candidate based on an initial identification from linkage analysis of auditory evoked potential deficits observed in patients with schizophrenia22, 23.

The deleted region on 1q21.1 is consistent with a previously reported de novo deletion in a patient with learning disability and seizures24 and two patients with autism (one de novo and one inherited)10. In the present study, three cases had mild cognitive abnormalities and one had a history of epilepsy (Supplementary Table 7). The region contains 27 known genes, most of which are expressed in the brain (Supplementary Table 8), and previous reports have shown linkage25, 25 but there have been no previous reports of CNVs associated with schizophrenia.

Regions of highly homologous segmental duplication flank the deletions we report at 22q11.2, 15q13.3 and 1q21.1. A prominent mechanism for CNV genesis is non-allelic homologous recombination mediated by segmental duplications, resulting in deletions and reciprocal duplications of the interval between segmental duplications16, 27. Neurodevelopmental and psychiatric syndromes have been associated with deletions and duplications flanked by segmental duplications, many of which occur de novo10, 11, 17. Segmental duplications and non-allelic homologous recombination mediate CNVs at 22q11.2 (ref. 28) and may be involved in the genesis of CNVs at 15q13.3 (ref. 17) and 1q21.1, although other mechanisms may be involved29.

While this work was under review, Walsh et al.2 reported a higher frequency of cases with CNVs (15%, based on 23 CNVs in 150 patients with schizophrenia) than in controls (5%). Of the 21 autosomal case CNVs identified in that report, we observed overlapping control CNVs at six loci (for example DLG2 and PTPRM; Supplementary Table 9), illustrating that large sample numbers are needed to conclude that any one particular CNV or implicated gene can cause schizophrenia. Our global burden analysis demonstrated that, on aggregate, single-occurrence and very rare (under about 1 in 1,000) CNVs have increased rates in cases of schizophrenia in comparison with controls, in line with Walsh et al.2. This indicates that at least some of these rare CNVs seen in cases but not in controls are probably risk factors for schizophrenia, although like Walsh et al. we

are unable to identify which. Some examples of possible risk CNVs that were observed multiple times only in cases include deletions at 12p11.23 (four cases) and 16p12.2–12.1 (four cases). These deletions were more than 500 kb, flanked by segmental duplications and spanning several brain-expressed genes. In addition, duplications in two genes relevant to neural development and growth (NOTCH1 and p21-activated kinase 7, PAK7) were found in five and six cases, respectively, and no controls. Furthermore, we identified CNVs at two recently reported loci, NRXN1 and CNTNAP2 (refs 3, 5) (Supplementary Fig. 2).

The aetiology of schizophrenia has been vigorously debated. We now have strong and replicated2 evidence that individuals with schizophrenia have a greater burden of structural variation across their genomes. Our data show that CNVs in at least three loci act as strong risk factors for schizophrenia in a minority of individuals. We can therefore now posit that some cases of schizophrenia are 'genomic disorders'16 although we do not yet know whether the risk is specific for schizophrenia as opposed to a more general risk factor for neuropsychiatric or central nervous system illness.

Exactly how a subtle, 1.15-fold increase in CNV burden translates mechanistically into illness in a given patient is currently unknown. We also do not know whether common genetic variants of more subtle effect are components of the aetiology of schizophrenia, an empirical question that we and others are addressing. Similarly, we do not know how environmental risk or protective factors might act in concert with specific CNVs or with the overall burden of CNVs.

A critically important goal will be to determine the full clinical and phenotypic spectrum in carriers of these deletions. Our data provide preliminary evidence of a variable phenotype in patients with schizophrenia who would otherwise be regarded as clinically typical. Examining the role of these variants in related psychotic disorders, such as bipolar disorder, is imperative. Further work explicating the epidemiology and mechanism of these variants in schizophrenia may ultimately lead to a role for them in genetic counselling and understanding disease biology.

Note added in proof: Samples from the University of Aberdeen were genotyped independently by the SGENE Consortium30.

Top of page

Methods Summary

Cases satisfied DSM-IV31 or ICD-10 (ref. 32) criteria for schizophrenia and were broadly representative of clinical cases in contact with psychiatric services. DNA was extracted from whole blood, with approval from institutional review boards. CNVs were identified with the Birdseye package15 and analysed with PLINK v1.03 (ref. 14). See Supplementary Information for details. A list of all CNVs passing quality control is available (http://pngu.mgh.harvard.edu/isc/).

References

Sullivan, P. F., Kendler, K. S. & Neale, M. C. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. Arch. Gen. Psychiatry 60, 1187–1192 (2003) | Article | PubMed | ISI |

Walsh, T. et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. Science 320, 539–543 (2008) | Article | PubMed | ChemPort |

Kirov, G. et al. Comparative genome hybridization suggests a role for NRXN1 and APBA2 in schizophrenia. Hum. Mol. Genet. 17, 458–465 (2008) | Article | PubMed | ChemPort |

Flomen, R. H. et al. Association study of CHRFAM7A copy number and 2 bp deletion polymorphisms with schizophrenia and bipolar affective disorder. Am. J. Med. Genet. B. Neuropsychiatr. Genet. 141, 571–575 (2006)

Friedman, J. I. et al. CNTNAP2 gene dosage variation is associated with schizophrenia and epilepsy. Mol. Psychiatry 13, 261–266 (2008) | Article | PubMed | ChemPort |

Wilson, G. M. et al. DNA copy-number analysis in bipolar disorder and schizophrenia reveals aberrations in genes involved in glutamate signaling. Hum. Mol. Genet. 15, 743–749 (2006) | Article | PubMed | ISI | ChemPort |

Moon, H. J. et al. Identification of DNA copy-number aberrations by array-comparative genomic hybridization in patients with schizophrenia. Biochem. Biophys. Res. Commun. 344, 531–539 (2006) | Article | PubMed | ChemPort |

Marshall, C. R. et al. Structural variation of chromosomes in autism spectrum disorder. Am. J. Hum. Genet. 82, 477–488 (2008) | Article | PubMed | ChemPort |

Sebat, J. et al. Strong association of de novo copy number mutations with autism. Science 316, 445–449 (2007) | Article | PubMed | ISI | ChemPort |

Weiss, L. A. et al. Association between microdeletion and microduplication at 16p11.2 and autism. N. Engl. J. Med. 358, 667–675 (2008) | Article | PubMed | ChemPort |

Lee, J. A. & Lupski, J. R. Genomic rearrangements and gene copy-number alterations as a cause of nervous system disorders. Neuron 52, 103–121 (2006) | Article | PubMed | ISI | ChemPort |

Williams, N. M., O'Donovan, M. C. & Owen, M. J. Chromosome 22 deletion syndrome and schizophrenia. Int. Rev. Neurobiol. 73, 1–27 (2006) | Article | PubMed | ChemPort |

McCarroll, S. et al. Integrated detection and population-genetic analysis of SNPs and copy number variation. Nature Genet. (in the press)

Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575 (2007) | Article | PubMed | ISI | ChemPort |

Korn, J. et al. Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms, and rare CNVs. Nature Genet. (in the press)

Lupski, J. R. Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. Trends Genet. 14, 417–422 (1998) | Article | PubMed | ISI | ChemPort |

Sharp, A. J. et al. A recurrent 15q13.3 microdeletion syndrome associated with mental retardation and seizures. Nature Genet. 40, 322–328 (2008) | Article |

Ross, C. A. & Pearlson, G. D. Schizophrenia, the heteromodal association neocortex and development: potential for a neurogenetic approach. Trends Neurosci. 19, 171–176 (1996) | Article | PubMed | ChemPort |

Karayiorgou, M. et al. Schizophrenia susceptibility associated with interstitial deletions of chromosome 22q11. Proc. Natl Acad. Sci. USA 92, 7612–7616 (1995) | Article | PubMed | ChemPort |

Shaikh, T. H. et al. Chromosome 22-specific low copy repeats and the 22q11.2 deletion syndrome: genomic organization and deletion endpoint analysis. Hum. Mol. Genet. 9, 489–501 (2000) | Article | PubMed | ISI | ChemPort |

Butler, M. G., Fischer, W., Kibiryeva, N. & Bittel, D. C. Array comparative genomic hybridization (aCGH) analysis in Prader–Willi syndrome. Am. J. Med. Genet. A 146, 854–860 (2008) | PubMed |

Freedman, R. et al. Linkage of a neurophysiological deficit in schizophrenia to a chromosome 15 locus. Proc. Natl Acad. Sci. USA 94, 587–592 (1997) | Article | PubMed | ChemPort |

Xu, J. et al. Evidence for linkage disequilibrium between the alpha7-nicotinic receptor gene (CHRNA7) locus and schizophrenia in Azorean families. Am. J. Med. Genet. 105, 669–674 (2001) | Article | PubMed | ChemPort |

Sharp, A. J. et al. Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome. Nature Genet. 38, 1038–1042 (2006) | Article |

Brzustowicz, L. M., Hodgkinson, K. A., Chow, E. W., Honer, W. G. & Bassett, A. S. Location of a major susceptibility locus for familial schizophrenia on chromosome 1q21–q22. Science 288, 678–682 (2000) | Article | PubMed | ISI | ChemPort |

Gurling, H. M. et al. Genomewide genetic linkage analysis confirms the presence of susceptibility loci for schizophrenia, on chromosomes 1q32.2, 5q33.2, and 8p21–22 and provides support for linkage to schizophrenia, on chromosomes 11q23.3–24 and 20q12.1–11.23. Am. J. Hum. Genet. 68, 661–673 (2001) | Article | PubMed | ISI | ChemPort |

Stankiewicz, P. & Lupski, J. R. Genome architecture, rearrangements and genomic disorders. Trends Genet. 18, 74–82 (2002) | Article | PubMed | ISI | ChemPort |

Shaikh, T. H. et al. Low copy repeats mediate distal chromosome 22q11.2 deletions: sequence analysis predicts breakpoint mechanisms. Genome Res. 17, 482–491 (2007) | Article | PubMed | ISI | ChemPort |

Lee, J. A., Carvalho, C. M. & Lupski, J. R. A. DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. Cell 131, 1235–1247 (2007) | Article | PubMed | ChemPort |

Stefansson, H. et al. Large recurrent microdeletions associated with schizophrenia. Nature doi: doi: 10.1038/nature07229 (this issue) | Article | PubMed | ChemPort |

DSM-IV. Diagnostic and Statistical Manual of Mental Disorders 4th edn (American Psychiatric Association, 2000)

ICD-10. International Statistical Classification of Diseases and Related Health Problems 10th revision (World Health Organization, 2007)

Acknowledgements

We thank the patients and families who contributed their time and DNA to these studies, and also D. Altshuler and members of the Medical and Population Genetics group at the Broad Institute of Harvard and Massachusetts Institute of Technology for valuable discussion. The group at the Stanley Center for Psychiatric Research at the Broad Institute was supported by the Stanley Medical Research Institute (E.M.S.), the Sylvan C. Herman Foundation (E.M.S.), and MH071681 (P.S.). The Cardiff University group was supported by a Medical Research Council (UK) Programme grant and the National Institutes of Mental Health (USA) (CONTE: 2 P50 MH066392-05A1). The group at Karolinska Institutet was supported by the Swedish Council for Working Life and Social Research (FO 184/2000; 2001-2368). The Massachusetts General Hospital group was supported by the Stanley Medical Research Institute (P.S.), MH071681 (P.S.) and a Narsad Young Investigator Award (S.P.). The group at the Queensland

Institute of Medical Research was supported by the Australian National Health and Medical Research Council. The Trinity College Dublin group was supported by Science Foundation Ireland, the Health Research Board (Ireland), the Stanley Medical Research Institute and the Wellcome Trust; Irish controls were supplied by J. McPartlin from the Trinity College Biobank. The work at the University of Aberdeen was partly funded by GlaxoSmithKline and Generation Scotland, Genetics Health Initiative. The University College London clinical and control samples were collected with support from the Neuroscience Research Charitable Trust, the Camden and Islington Mental Health and Social Care Trust, East London and City Mental Heath Trust, the West Berkshire NHS Trust, the West London Mental Health Trust, Oxfordshire and Buckinghamshire Mental Health Partnership NHS Trust, South Essex Partnership NHS Foundation Trust, Gloucestershire Partnership NHS Foundation Trust, Mersey Care NHS Trust, Hampshire Partnership NHS Trust and the North East London Mental Health Trust. The collection of the University of Edinburgh cohort was supported by grants from the Wellcome Trust, London, and the Chief Scientist Office of the Scottish Executive. The group at the University of North Carolina, Chapel Hill, was supported by MH074027, MH077139 and MH080403, the Sylvan C. Herman Foundation (P.F.S.) and the Stanley Medical Research Institute (P.F.S.) The group at the University of Southern California thanks the patients and their families for their collaboration, and acknowledges the support of the National Institutes of Mental Health and the Department of Veterans Affairs.

The International Schizophrenia Consortium

Manuscript preparation

Jennifer L. Stone1,2,3,4, Michael C. O'Donovan5, Hugh Gurling6, George K. Kirov5, Douglas H. R. Blackwood7, Aiden Corvin8, Nick J. Craddock5, Michael Gill8, Christina M. Hultman9,10, Paul Lichtenstein9, Andrew McQuillin6, Carlos N. Pato11, Douglas M. Ruderfer1,2,3,4, Michael J. Owen5, David St Clair12, Patrick F. Sullivan13, Pamela Sklar1,2,3,4 & Shaun M. Purcell (Leader)1,2,3,4

Data analysis

Jennifer L. Stone1,2,3,4, Douglas M. Ruderfer1,2,3,4, Joshua Korn3,4, George K. Kirov5, Stuart Macgregor14, Andrew McQuillin6, Derek W. Morris8, Colm T. O'Dushlaine8, Mark J. Daly2,3,4, Peter M. Visscher14, Peter A. Holmans5, Michael C. O'Donovan5, Patrick F. Sullivan13, Pamela Sklar1,2,3,4 & Shaun M. Purcell (Leader)1,2,3,4

Management committee

Hugh Gurling6, Aiden Corvin8, Douglas H. R. Blackwood7, Nick J. Craddock5, Michael Gill8, Christina M. Hultman9,10, George K. Kirov5, Paul Lichtenstein9, Andrew McQuillin6, Michael C. O'Donovan5, Michael J. Owen5, Carlos N. Pato11, Shaun M. Purcell1,2,3,4, Edward M. Scolnick2,3, David St Clair12, Jennifer L. Stone1,2,3,4, Patrick F. Sullivan13 & Pamela Sklar (Leader)1,2,3,4

Cardiff University

Michael C. O'Donovan5, George K. Kirov5, Nick J. Craddock5, Peter A. Holmans5, Nigel M. Williams5, Lucy Georgieva5, Ivan Nikolov5, N. Norton5, H. Williams5, Draga Toncheva15, Vihra Milanova16 & Michael J. Owen5

Karolinska Institutet/University of North Carolina at Chapel Hill

Christina M. Hultman9,10, Paul Lichtenstein9, Emma F. Thelander9 & Patrick Sullivan13

Trinity College Dublin

Derek W. Morris8, Colm T. O'Dushlaine8, Elaine Kenny8, John L. Waddington17, Michael Gill8 & Aiden Corvin8

University College London

Andrew McQuillin6, Khalid Choudhury6, Susmita Datta6, Jonathan Pimm6, Srinivasa Thirumalai18, Vinay Puri6, Robert Krasucki6, Jacob Lawrence6, Digby Quested19,29, Nicholas Bass6, David Curtis20 & Hugh Gurling6

University of Aberdeen

Caroline Crombie21, Gillian Fraser21, Soh Leh Kwan12, Nicholas Walker22 & David St Clair12

University of Edinburgh

Douglas H. R. Blackwood7, Walter J. Muir7, Kevin A. McGhee7, Ben Pickard7, Pat Malloy7, Alan W. Maclean7 & Margaret Van Beck7

Queensland Institute of Medical Research

Peter M. Visscher14 & Stuart Macgregor14

University of Southern California

Michele T. Pato11, Helena Medeiros11, Frank Middleton23, Celia Carvalho11, Christopher Morley23, Ayman Fanous11, David Conti11, James A. Knowles11, Carlos Paz Ferreira27, Antonio Macedo28, M. Helena Azevedo28 & Carlos N. Pato11

Massachusetts General Hospital

Jennifer L. Stone1,2,3,4, Douglas M. Ruderfer1,2,3,4, Joshua Korn3,4, Steve A. McCarroll3,4, Mark Daly2,3,4, Shaun M. Purcell1,2,3,4 & Pamela Sklar1,2,3,4

Stanley Center for Psychiatric Research and Broad Institute of MIT and Harvard

Shaun M. Purcell1,2,3,4, Jennifer L. Stone1,2,3,4, Kimberly Chambert2,3, Douglas M. Ruderfer1,2,3,4, Joshua Korn3,4, Steve A. McCarroll3,4, Casey Gates3, Stacey B. Gabriel3, Scott Mahon3, Kristen Ardlie3, Mark J. Daly2,3,4, Edward M. Scolnick2,3 & Pamela Sklar1,2,3,4

The International Schizophrenia Consortium

Department of Psychiatry, Massachusetts General Hospital and Harvard Medical School, 185 Cambridge Street, Boston, Massachusetts 02114, USA.

Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA.

Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, USA.

Center for Human Genetic Research, Massachusetts General Hospital, 185 Cambridge Street, Boston, Massachusetts 02114, USA.

School of Medicine, Department of Psychological Medicine, School of Medicine, Cardiff University, Cardiff C14 4XN, UK.

Molecular Psychiatry Laboratory, Department of Mental Health Sciences, University College London Medical School, Windeyer Institute of Medical Sciences, 46 Cleveland Street, London W1T 4JF, UK.

Division of Psychiatry, School of Molecular and Clinical Medicine, University of Edinburgh, Edinburgh EH10 5HF, UK.

Neuropsychiatric Genetics Research Group, Department of Psychiatry and Institute of Molecular Medicine, Trinity College Dublin, Dublin 2, Ireland.

Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, SE-171 77 Stockholm, Sweden.

Department of Neuroscience, Psychiatry, Ulleråker, Uppsala University, SE-750 17 Uppsala, Sweden.

Center for Genomic Psychiatry, University of Southern California, Los Angeles, California 90033, USA.

Institute of Medical Sciences, University of Aberdeen, Foresterhill, Aberdeen AB25 2ZD, UK.

Departments of Genetics, Psychiatry, and Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA.

Queensland Institute of Medical Research, 300 Herston Road, Brisbane, Queensland 4006, Australia.

Department of Medical Genetics, University Hospital Maichin Dom, Sofia 1431, Bulgaria.

Department of Psychiatry, First Psychiatric Clinic, Alexander University Hospital, Sofia 1431, Bulgaria.

Molecular and Cellular Therapeutics and RCSI Research Institute, Royal College of Surgeons in Ireland, Dublin 2, Ireland.

West Berkshire NHS Trust, 25 Erleigh Road, Reading RG3 5LR, UK.

West London Mental Health Trust, Hammersmith and Fulham Mental Health Unit and St Bernard's Hospital, London W6 8RF, UK.

Queen Mary College, University of London and East London and City Mental Health Trust, Royal London Hospital, Whitechapel, London E1 1BB, UK.

Department of Mental Health, University of Aberdeen, Aberdeen AB25 2ZD, UK.

Ravenscraig Hospital, Inverkip Road, Greenock PA16 9HA, UK.

State University of New York – Upstate Medical University, Syracuse, New York 13210, USA.

Washington VA Medical Center, Washington, DC 20422, USA.

Department of Psychiatry, Georgetown University School of Medicine, Washington, DC 20057, USA.

Department of Psychiatry, Virginia Commonwealth University, Richmond, VA 23298, USA.

Department of Psychiatry, Sao Miguel, 9500-310 Azores, Portugal.

Department of Psychiatry, University of Coimbra, 3004-504 Coimbra, Portugal.

Present address: Department of Psychiatry, University of Oxford, Warneford Hospital, Headington, Oxford OX3 7JX, UK.

Lists of members and affiliations appear at the end of the paper.

Correspondence to: Correspondence and requests for materials should be addressed to P.S. (Email: sklar@chgr.mgh.harvard.edu).