708 Common and 2010 rare DISC1 locus variants identified in 1542 subjects

Citation for published version:
Thomson, PA, Parla, JS, McRae, AF, Kramer, M, Ramakrishnan, K, Yao, J, Soares, DC, McCarthy, S, Morris, SW, Cardone, L, Cass, S, Ghiban, E, Hennah, W, Evans, KL, Rebolini, D, Millar, JK, Harris, SE, Starr, JM, MacIntyre, DJ, Generation Scotland, McIntosh, AM, Watson, JD, Deary, IJ, Viesscher, PM, Blackwood, DH, McCombie, WR & Porteous, DJ 2014, '708 Common and 2010 rare DISC1 locus variants identified in 1542 subjects: analysis for association with psychiatric disorder and cognitive traits' Molecular Psychiatry, vol 19, no. 6, pp. 668-675. DOI: 10.1038/mp.2013.68

Digital Object Identifier (DOI):
10.1038/mp.2013.68

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Molecular Psychiatry

Publisher Rights Statement:
This work is licensed under a Creative Commons Attribution 3.0 Unported License.

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
708 Common and 2010 rare DISC1 locus variants identified in 1542 subjects: analysis for association with psychiatric disorder and cognitive traits

INTRODUCTION
Schizophrenia (SZ), bipolar disorder (BD) and recurrent major depressive disorder (rMDD) are common forms of serious mental illness, each with a strong and overlapping genetic component. A balanced t(1;11) translocation that transsects the Disrupted in schizophrenia 1 (DISC1) gene shows genome-wide significant linkage for schizophrenia and recurrent major depressive disorder (rMDD) in a single large Scottish family, but genome-wide and exome sequencing-based association studies have not supported a role for DISC1 in psychiatric illness. To explore DISC1 in more detail, we sequenced 528 kb of the DISC1 locus in 653 cases and 889 controls. We report 2718 validated single-nucleotide polymorphisms (SNPs) of which 2010 have a minor allele frequency of < 1%. Only 38% of these variants are reported in the 1000 Genomes Project European subset. This suggests that many DISC1 SNPs remain undiscovered and are essentially private. Rare coding variants identified exclusively in patients were found in likely functional protein domains. Significant region-wide association was observed between rs16856199 and rMDD (P = 0.026, unadjusted P = 6.3 × 10⁻⁵, OR = 3.48). This was not replicated in additional recurrent major depression samples (replication P = 0.11). Combined analysis of both the original and replication set supported the original association (P = 0.0058, OR = 1.46). Evidence for segregation of this variant with disease in families was limited to those of rMDD individuals referred from primary care. Burden analysis for coding and non-coding variants gave nominal associations with diagnosis and measures of mood and cognition. Together, these observations are likely to generalise to other candidate genes for major mental illness and may thus provide guidelines for the design of future studies.

Keywords: DISC1; recurrent major depressive disorder; sequencing

A balanced t(1;11) translocation that transsects the Disrupted in schizophrenia 1 (DISC1) gene shows genome-wide significant linkage for schizophrenia and recurrent major depressive disorder (rMDD) in a single large Scottish family, but genome-wide and exome sequencing-based association studies have not supported a role for DISC1 in psychiatric illness. To explore DISC1 in more detail, we sequenced 528 kb of the DISC1 locus in 653 cases and 889 controls. We report 2718 validated single-nucleotide polymorphisms (SNPs) of which 2010 have a minor allele frequency of < 1%. Only 38% of these variants are reported in the 1000 Genomes Project European subset. This suggests that many DISC1 SNPs remain undiscovered and are essentially private. Rare coding variants identified exclusively in patients were found in likely functional protein domains. Significant region-wide association was observed between rs16856199 and rMDD (P = 0.026, unadjusted P = 6.3 × 10⁻⁵, OR = 3.48). This was not replicated in additional recurrent major depression samples (replication P = 0.11). Combined analysis of both the original and replication set supported the original association (P = 0.0058, OR = 1.46). Evidence for segregation of this variant with disease in families was limited to those of rMDD individuals referred from primary care. Burden analysis for coding and non-coding variants gave nominal associations with diagnosis and measures of mood and cognition. Together, these observations are likely to generalise to other candidate genes for major mental illness and may thus provide guidelines for the design of future studies.

Keywords: DISC1; recurrent major depressive disorder; sequencing
in SZ cases and 40 in European controls with no evidence for a significantly increased burden of likely pathogenic variants.\textsuperscript{23} DISC1, however, continues to feature strongly in attempts to assess genome-wide association results in terms of networks\textsuperscript{24} or in combination with known biological function.\textsuperscript{25–27}

The biological functions of DISC1 fit well with current aetiological concepts in SZ-related major mental illness and cognition.\textsuperscript{28} DISC1 is a scaffold protein that interacts with, and modulates the activity of, multiple proteins with key roles in neurodevelopment, neurogenesis, neuronal migration, integration and signalling\textsuperscript{29–31} including the antidepressant and antipsychotic targets GSK3\textsuperscript{32} and PDE4.\textsuperscript{33} Several common and rare amino-acid polymorphisms of DISC1 have predicted deleterious effects on protein function and demonstrable biological effects in experimental settings.\textsuperscript{31} The 704C allele is associated with reduced activity of ERK1 and Akt kinases, altered binding affinities of DISC1 for NDE1 and NDEL1 and variation in DISC1 oligomeric status.\textsuperscript{34–37} The 607F allele results in (a) reduced binding and centrosomal localisation of PCM1, (b) reduced noradrenaline neurotransmitter release in SH-SY5Y cells,\textsuperscript{38} (c) altered mitochondrial trafficking\textsuperscript{39} and (d) a partial shift from neuronal to glial expression in the brain.\textsuperscript{40} Furthermore, Singh et al.\textsuperscript{41} reported that 607F impacts negatively on neural progenitor proliferation in E16 mouse brain, correlates with aberrant wnt signalling in human lymphoblasts, and is associated with a neurodevelopmental phenotype in morpholino mutant zebrafish. R37W lies within an arginine-rich nuclear localisation signal and a partially overlapping functional domain for PDE4\textsuperscript{42} and GSK3\textsuperscript{41,43} The 37W allele shows reduced nuclear DISC1 expression and altered DISC1 regulation of AFT4, a critical modulator of cAMP signalling and mediator of the stress response.\textsuperscript{44}

In summary, whereas the primary genetic evidence from the original Scottish t(1;11) family was significant at the genome-wide significance level for both SZ and rMDD and the experimental evidence and biological plausibility of DISC1 remains strong, the evidence from subsequent linkage and candidate gene association studies is however inconsistent and not supported by genome-wide association studies or meta-analysis.\textsuperscript{12} To explore these contrasting findings, we aimed here to establish the nature of DISC1 genomic sequence variants, identify rare variants in putative functional domains, and test for effects of these on cognitive traits and the risk of psychiatric illness. We comprehensively sequenced 528 kb covering the entire DISC1 locus, including TRAX (also known as TSNAX) for which there is evidence for intergenic splicing with DISC1\textsuperscript{45} and the intergenic region, which contains regulatory elements immediately 5' of DISC1.\textsuperscript{13,46–48}

**RESULTS**

Sequence analysis

We sequenced 1542 Caucasians from Scotland comprising 240 cases of SZ, 221 cases of BD, 192 cases of rMDD and, as controls, 889 members of the Lothian Birth Cohort 1936 (LBC1936), which have been extensively phenotyped.\textsuperscript{60} Each sample was sequenced to >80% coverage at a minimum of 30-fold read depth by long-range PCR and sequencing on either Illumina GAII or HiSeq 2000 sequencers. To ensure a robust data set, all variants within repetitive regions were removed. Final quality score thresholds for the data were derived from capillary sequence validation of 10% of the remaining variants. All variants with an MAF <1% were validated by ABI3730 sequencing. After quality control, there was no evidence for sequencing bias between cases and controls (Supplementary Figure S1). Allele frequencies from our sample showed strong concordance to those from the European subset of the 1000 Genomes Project\textsuperscript{61} (Supplementary Figure S2). We report 2718 SNPs in the 1542 samples analysed, 708 at >1% and 2010 at <1% MAF (Supplementary Table S1). Only 1027 of the 2718 SNPs (38%) were previously reported in the European subset of the 1000 Genomes Project.\textsuperscript{61}

As defined and annotated by the UCSC genome browser (http://genome.ucsc.edu), 489 SNPs mapped to regions of regulatory potential, 177 to non-coding exons (including DISC2) and 36 to coding regions of exons. Of these 36 variants, 12 were synonymous changes, 23 were non-synonymous changes, with one producing a stop codon consistent with the DISC1 E isoform (Figure 1; Supplementary Figure S3; Supplementary Tables S1 and S2). Supplementary Table S3 summarises the overlap between variants identified in this study and other DISC1 sequence studies and relevant association studies.\textsuperscript{10,11,13,14,16,17,23,46}

Association and segregation of common variants with psychiatric illness and related quantitative traits

Genome-wide association studies of SZ, BP and rMDD are most consistent with a polygenic liability for common variants, but they

**MATERIALS AND METHODS**

A full summary of the methods can be found in Supplementary Information. Briefly, all study participants gave signed consent for their data and samples to be used in studies that have been approved by the appropriate Research Ethics Committee or the GS access Committee. Genomic DNA from each individual was whole genome amplified in triplicate, the products pooled and amplified with primer pairs tiled across 528 kb of TRAX/DISC1 (hg18 chr1:1229723339-230521606; hg19 chr1:231656716-232184983). For each sample, the pooled products were sheared, converted into paired-end Illumina libraries and sequenced on an Illumina GAII or HiSeq 2000 sequencer to >80% coverage and >30-fold depth. Sequences were aligned to the UCSC hg18 reference sequence, variants called using MAQ software\textsuperscript{49} and the variants in repeats removed. Ten percent of all remaining variants were validated using Sanger sequencing chemistry on an ABI3730 sequencer, and the derived information used to optimise the quality control filters. After quality control screening, all exonic and low frequency (MAF <1%) variants were also validated by Sanger chemistry sequencing as above. The variants were functionally annotated using SNPnexus\textsuperscript{50} (http://www.snp-nexus.org).

Non-coding variants were annotated using the UCSC table browser for the following tracks: ‘RepeatMasker’,\textsuperscript{51} ‘CpG island’, ‘TFBS conserved’, ‘7x Reg Potential’ (which substantially overlaps with DNase hypersensitivity sites) and/or ‘28-Way Most conserved—PlacMammal’ (http://genome.ucsc.edu/). Sequence variants classified as coding were mapped to the DISC1 L isoform and potential pathogenicity ascribed using Pmut,\textsuperscript{52} Panther\textsuperscript{53} and PolyPhen-2\textsuperscript{54}. The coding sequence variants were mapped onto a list of known curated DISC1-interactor binding sites\textsuperscript{31} and with other functional elements (for example, phosphorylation sites\textsuperscript{55}). Case–control association was tested on the combined case samples as well as individually for SZ, BP and rMDD using Fisher’s exact test. Permutation was used to derive region-wide P-values and significance thresholds. Quantitative trait association analyses using LBC1936 samples were performed by linear regression of the trait residuals (adjusted for age and sex) on the number of minor alleles at each SNP, with empirical P-values estimated by permutation to avoid issues with the test statistic distribution caused by the combination of rare variants and slight deviations from normality in the phenotypes. All association analyses were performed using PLINK.\textsuperscript{46} Mark-recapture analysis followed the Lincoln-Peterson and Modified Petersen methods\textsuperscript{57,58} with 95% confidence intervals calculated following Chapman.\textsuperscript{59} Burden analysis was performed in PLINK/SEQ to implement BURDEN and VTTEST with empirical P-values estimated using permutation. Genotyping of the replication and familial samples was performed by the Edinburgh Wellcome Trust Clinical Research Facility Genetics Core using TaqMan SNP genotyping assay C__33950433_10 with concurrent genotyping of known heterozygotes.

Data access

The accession numbers for sequence data are NCBI ss472328925—ss472331023.
The association was seen in the combined analysis of both the rMDD and all three combined (best significant association was seen with any individual replication set). No drawn exclusively from Generation Scotland: Scottish Family Health Study (Supplementary additional text and Table S4). No asso-
ciation (P = 0.00065, OR = 1.76, 95% CI = 1.27–2.44). No association was seen in the combined analysis of both the rMDD and MDD population-based replication sets (P = 0.41). The risk allele for rs16856199 did not segregate with rMDD in 10 families of carriers identified from the Generation Scotland replication sample (Supplementary Figure S5). This suggests that there is increased evidence for association of rs16856199 in the more severely affected individuals.

SNP rs16856199 is on the Affymetrix 6.0 array, but the best tagging SNP on the Illumina 660W-Quad, Human Hap, Human1M-Duo arrays is rs16856189. SNP rs16856189 has an r² of 0.27 with rs16856199, which may explain in part why this association has not been reported previously in genome-wide association studies.32,62,63 SNP rs6678723, which lies 2.1 kb distal to this SNP in the recent mega-analysis of depression (P = 0.0092).20

The LBC1936 has quantitative measures of symptoms of anxiety, depression and the personality trait of neuroticism, plus psychometrically tested measures of cognitive ability (fluid (age sensitive) and crystallised (non-age sensitive)) and cognitive ageing,50,64 which have been shown to be highly heritable and polygenic.55,66 Association of these traits with DISC1 was tested by linear regression analyses, co-varied for age at testing and sex (Supplementary Figures S6–S8). There were no region-wide significant findings for any of these quantitative traits.

Estimating the net pool of DISC1 variants
To estimate the effective pool of common and rare sequence variants in the European population, we applied a ‘mark-recapture’ approach (see Supplementary methods) to our data and that of the 1000 Genomes Project (v3.20101123)27 after appropriate checks on read depth and Sanger sequence validation (Supplementary Table S5; Supplementary Figure 9).

Figure 1. TRAX/DISC1 (Disrupted in schizophrenia 1) genomic and exon structure: alignment of coding and regulatory variants. (a) Three-period moving average of all single-nucleotide polymorphisms (SNPs) identified per 5 kb across the region in this study with TRAX/DISC1 intron/exon structure given to scale. Total SNP number (black), those with a minor allele frequency (MAF) of <1% SNPs (blue), those ≥1% MAF (green), rs16856199 (arrow). For comparison, the number of SNPs identified in the 1000 genomes (red, dashed) and the number of bases repeat masked (top black) and 7x regulatory potential (top blue) are also shown. Exon and intron structure of TRAX and DISC1 are drawn to scale. (b) The position and diagnoses of exonic or regulatory SNPs. SNPs not seen previously (underlined), synonymous SNPs (black) and non-synonymous SNPs (red), stop or putative splice SNPs (green). Novel SNPs not previously reported in the European samples of the 1000 Genomes Project (v3.20101123) or the NHLBI GO Exome Sequencing Project (ESP6500) or relevant sequencing and association studies15,16,17,64,65 are underlined.

also imply that there is real ‘missing’ genetic variation, which is most likely due to risk variants having low frequency in the population. To test for evidence of DISC1 association, we applied the Fisher’s exact test across all variants and all diagnoses (Figure 2). There was no evidence for SNP association at genome-wide levels of significance for any diagnosis when considered separately or combined, nor was there evidence for locus-wide association of variants with SZ or BP. We did detect a novel, locus-wide empirical association P = 0.026 (OR = 3.48, 95% CI = 1.95–6.23, unadjusted P = 6.3 × 10⁻⁵) between intronic variant rs16856199 and rMDD. We speculated that individual risk alleles might be predicted to segregate with disease in families. Twelve additional family members were available for genotyping for four rs16856199 carriers. The rs16856199 risk allele segregated with rMDD in all four families (Supplementary Figure S4).

We next tested for association of rs1685199 with depression in three additional sample sets: a group of individuals referred from primary care to a hospital outpatient clinic (n = 467 rMDD patients), and two population-based samples drawn from primary care as part of the Generation Scotland: Scottish Family Health Study consisting of 645 cases with rMDD and 690 cases with single episode MDD. All three groups were compared with 4017 controls, and two population-based samples drawn from primary care to a hospital outpatient clinic (Supplementary Table S5; Supplementary Figure 9).
The total number of DISC1 SNPs $>1\%$ MAF was estimated at 905 (95%CI = 905 ± 5), of which 901 (99.5%) are known (Supplementary Table S6). The number of rare variants ($<1\%$ MAF) is less confidently predicted, but is likely to be substantially higher (95%CI = 3777 ± 252) (Supplementary Table S6). Thus, despite the ∼2500 European genomes in which the DISC1 locus has been completely sequenced and the 2305 rare DISC1 variants now known, ∼40% or more remain to be discovered, and will be essentially ‘private’.

Rare amino-acid substitutions

Of the 17 rare coding SNPs previously reported for DISC1,10,16,17,68,69 we identified 12 (70.6%) plus an additional 8 non-synonymous variants, of which 5 are also absent from the European samples of both the 1000 Genomes Project (v3.20101123)11 and the Exome Variant Server (NHBLI GO Exome Sequencing Project (ESP), Seattle, WA; http://evs.gs.washington.edu/EVS/; September, 2012), and previous DISC1 sequencing studies.10,11,13,14,16,17,23,46,70 (Figure 1; Supplementary Figure S3; Supplementary Tables S2 and S3).

Five variants, R37W, T453M, T603M, L607F and S704C, are predicted to be deleterious by all three applied prediction algorithms (PolyPhen-2, Pmut and Panther; Supplementary Table S2). From this set, the functional effects of the common variants L607F (rs6675281) and S704C (rs821616) have been well documented in the literature, as mentioned earlier. R37W lies within a defined nuclear localisation signal71 and PDE4B binding site72 and is seen in a single case of rMDD (discussed at the end of this section). T453M is present at low frequency in cases and control individuals, both in this study and others (Supplementary Table S3). T603M was only identified in a single control, but Song et al.11 reported a T603I variant in a schizophrenic individual that was absent in their set of control individuals (Supplementary Table S3).

Five variants were only observed in affected individuals from our study (R37W, A83V, W160L, R233K and R418H), but not in 889 control individuals. Four of these non-synonymous case-only singletons are located in the largest coding exon, DISC1 exon 2, and the remaining variant is in DISC1 exon 4 (Figure 1). A83V was seen in a single individual with BD; this variant is predicted to be deleterious by PolyPhen-2 and Pmut and has been shown to affect wnt signalling.41 It was however observed at low frequency in controls and in individuals with partial agenesis of the corpus callosum in previous studies (Supplementary Table S3).11,14,16 Apart from R37W and A83V, none of the three remaining non-synonymous variants, W160L, R233K and R418H, are consistently predicted by the three prediction algorithms to have functional effects (Supplementary Table S2). 160L and 418H were detected in single SCZ individuals and have also previously been reported in individuals with SCZ; but 160L has also been detected in control individuals.10 The variant 233K has not been previously reported, and was identified in an individual with rMDD. No non-synonymous variants in TRAX were found in cases only. A single
significant association was found in the LBC1936 data between conserved transcription factor binding sites and CpG islands SNPs, conserved regions, regions with regulatory potential, and Supplementary Methods). SNPs were classified on the basis of bioinformatic sequencing. There was no significant overall difference in the overall number of minor alleles by diagnosis (see Supplementary Table S8; Supplementary Methods). Details of nominally significant measures of cognitive ability (Supplementary Table S9). Summaries of the nominally significant results are given in Table 1.

DISCUSSION

Diagnoses of SZ, BP and rMDD were all present in the original Scottish family carrying the translocation that disrupts DISC1. All three DSMIV diagnoses have a strong and overlapping genetic component, but robust statistical analysis of gene-level contributions to risk are complicated by extensive genetic heterogeneity within and between diagnoses.4 We have provided the most comprehensive landscape of genetic variation at the DISC1 locus to date in patients with this spectrum of psychiatric illness and in healthy population controls with quantitative measures of mood and cognition. Comparison between our sequencing study and that of the 1000 Genomes Project confirms that current genome-wide association studies effectively captures the majority of common (but not rare) variants in the European population. Our sample size is large by current sequencing study standards, but we lack power to detect genome-wide significant P-values for either common or rare variants (see Supplementary Information and Supplementary Figure S11 for further details and also Kiezun et al.).78 Indeed, the predicted abundance of independent rare variants at this (and any other given) locus makes it highly improbable that any one will contribute to illness in the population at a frequency that will be statistically significant, given the numbers of patients we can afford to analyse by direct sequencing.74 We observed no evidence for association at the whole-genome level of statistical significance between individual rare or common variants and either psychiatric illness or cognition. This is consistent with recent findings,75 which suggest that much larger samples would be required to detect such associations. Burden analysis of multiple rare and/or deleterious putative functional variants also failed to show association with these traits. We do report both functional and putative regulatory variants that are both individually, and by functional classification, nominally associated with rMDD and/or cognitive ability at the locus-wide level of significance.

Our study identified a novel association between intronic SNP rs16856199 and rMDD in hospital-referral subjects. Segregation with diagnosis in the relatives of these probands corroborated the association, but further studies are required to understand the lack of replication in population-based cohorts with depression. This may be due to inherent differences between patients recruited from hospital-referral compared with those from population-based cohorts. Cohorts from primary care are more likely to have a family history of depression,76 and may have more physical and psychiatric comorbidity in general. Conversely, the population-based sample may have suffered, less severe episodic, or even none of the hospital-based cohorts.76 However, given the modestly significant P-value for rMDD in the discovery cohort, the number of psychiatric traits examined and the lack of replication, it is possible that the observed association is due to chance.

The nominal associations of the burden of common (threshold MAF = 30.7%) and rare potentially regulatory variants (threshold MAF = 0.060%) to measures of cognitive ability merit further study. A yet-to-be-defined subset of these is likely to have critical roles is spatial and temporal regulation of transcription and splicing.46,78 This highlights the need for annotation tools with improved predictive value for non-coding variants.79

More importantly, our study demonstrates that substantial coding and non-coding genetic variation at the DISC1 locus
remains undiscovered. Despite sequencing over 1500 subjects, we have probably captured only ~40% of the extant DISC1 variants in just the European population. Crowley et al. sequenced 2.7 kb of DISC1 exons and 5' and 3' regulatory sequence in 1460 samples of European or African origin. We observed 13/38 (34%) of the variants genotyped in the replication phase, supporting the argument for an abundance of rare variants.

The level of sequence variation identified in our study is unlikely to be exceptional, and indeed is consistent with evidence emerging from other genome sequencing studies. Consequently, it will be challenging to demonstrate robust (replicated) association by statistical evidence alone in case-control studies, exceptionally so with the numbers of patients that are currently affordable for sequencing. The original t(1;11) family illustrates the added issue of variable penetrance and cross-boundary diagnosis for a given mutation: ~70% of carriers had SZ, BP or rMDD, but ~30% had no formal psychiatric diagnosis, yet t(1;11) carriers, including both affected and unaffected, had ERP P300 measures in the range typical of individuals with SZ. The original identification of 37W in a case of SZ and here in a case of rMDD (and two offspring, one with rMDD, the other generalised anxiety disorder) may suggest variable penetrance of this biologically functional variant. Of note, the 37W variant was not observed in 10,000 control individuals, the 1000 Genomes project, the NHGRI GO Exome Sequencing Project (ESP), nor any of our 889 control samples. These findings on R37W reinforce the probable importance of this domain for DISC1 subcellular distribution and binding of interacting proteins and add to the weight of evidence for other functional studies of DISC1 amino-acid substitutions. Each observed amino-acid substitution provides a similar opportunity to tease out the relationships between genotype and phenotype and between structure and function. Overall, these results demonstrate a high level of sequence variation in DISC1, a subset of which may contribute to psychiatric disorder in some individuals who will be typically rare in the population precluding classical statistical analysis and requiring biological validation. This predicts a population-specific contribution of rare casual variants to risk. Our results indicate the potential value of sequencing non-coding regions of the genome, which may harbour disease-associated regulatory variants. Our findings of both functional and putative regulatory variants nominally associated with depression and cognitive ability merit replication in independent samples and biological exploration.

CONFLICT OF INTEREST

WRM has participated in Illumina sponsored meetings over the past four years and received travel reimbursement and an honorarium for presenting at these events.

ACKNOWLEDGEMENTS

We acknowledge the contribution of Prof Walter M Muir (deceased) for the collection of the clinical samples used in this study. We thank Gail Davies, Dave Liewald, Lee Murphy, Alan MacLean, Pamela Linksted, Archie Campbell, Shona Kerr, Andrew McQuillan and Stuart Macgregor for help in preparing samples, data and statistical advice and Generation Scotland for access to GSFHS samples and data. We thank the LBC1936 cohort members. We thank the LBC1936 study secretary Paula Davies and Alan Gov, Janie Corley and Caroline Brett, for cognitive data collection and entry data. We thank the nurses and staff at the WTCPF where the medical phenotypic data and DNA were collected. We thank the staff at the Lothian Health Board, and the staff at the SCRE Centre, University of Glasgow. The LBC1936 research was supported by a programme grant from Research Into Ageing and continues with programme grants from Help the Aged/Research Into Ageing (Disconnected Mind). The research was done within the University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, supported by the BBSRC, EPSRC, ESRC and MRC, as part of the cross-council Lifelong Health and Wellbeing Initiative (G0700704/84698). Generation Scotland receives core support from the Chief Scientist Office of the Scottish Government Health Directorates CZD/16/6/2 and the Scottish Funding Council. We are grateful to the GS Executive Committee Professors Andrew D Morris, Blair H Smith, Anna F Dominicaci, David J Porteous and Drs Lynne J Hocking and Sandosh Padmanabhan and all the families who took part, the general practitioners and the Scottish School of Primary Care for their help in recruiting them, and the whole Generation Scotland team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, health-care assistants and nurses. We thank Sharon Mitchell and Carolann Gundersen for help in the preparation of this manuscript. This work was supported by a gift from T and V Stanley to CSHL, the UK Medical Research Council (G0902166), the Wellcome Trust (WT 008179), the Chief Scientists Office of the Scottish Government Department of Health (CZD/16/6/1), a Scottish Funding Council Strategic Research Development Grant award (HR30006) and a Lifelong Health and Wellbeing (G0700704/84698). AMcR and PMV acknowledge support from the Australian National Health and Medical Research Council.
AUTHOR CONTRIBUTIONS
DJP, PT and WRM designed the study, interpreted the data and wrote the paper. WJM, DJM, AMM and DHB were responsible for the psychiatric evaluation, selection and provision of case samples. DJP, PT, UJD, DLM, AMM and DHB were responsible for the provision of GSSFH samples and associated phenotypic measures. SEH, JMS and UD were responsible for the provision of LBC1936 samples and associated phenotypic measures. JSP developed the sample processing methodology and edited the paper. MK developed the sequence analysis pipeline and conducted the quality control of the sequence data and edited the paper. PT, AFM and PMW conducted the statistical analysis. KR, DCS, SM, SC and PT conducted the bioinformatics analysis. All authors contributed to revise the paper.

REFERENCES
1 Gottesman II, Laurens TM, Bertelsen A, Mortensen PB. Severe mental disorders in infants with 2 psychiatrically ill parents. Arch Gen Psychiatry 2010; 67: 252–257.
2 Lichtenstein P, Yip BH, Björk C, Pawitan Y, Cannon TD, Sullivan PF et al. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. Lancet 2009; 373: 234–239.
3 Purcell S, Neale B, Todd-Brown KE, Thomas L, Ferreira MA, Bender D et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007; 81: 559–575.
4 Sullivan PF, Daly MJ, O’Donovan M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. Nat Rev Genet 2012; 13: 537–551.
5 Blackwood DH, Fordyce A, Walker MT, St Clair DM, Porteous DJ, Muir WJ. Schizophrenia and affective disorders–co-segregation with a translocation at chromosome 1q42 that directly disrupts brain-expressed genes: clinical and P300 findings in a family. Am J Hum Genet 2001; 69: 428–433.
6 Eykelenboom JE, Briggs GJ, Bradshaw NJ, Soares DC, Ogawa F, Christie S et al. A t(11;11) translocation linked to schizophrenia and affective disorders gives rise to aberrant chimeric DISC1 transcripts that encode structurally altered, deleterious mitochondrial proteins. Hum Mol Genet 2012; 21: 3374–3386.
7 Millar JK, Wilson-Annan JC, Andersson S, Christie S, Taylor MS, Semple CA et al. Disruption of two novel genes by a translocation co-segregating with schizophrenia. Hum Mol Genet 2000; 9: 1415–1423.
8 Zhou X, Chen Q, Schaukowitch K, Kelsoe JR, Geyer MA. Insoluble DISC1-Boymaw fusion proteins generated by DISC1 translocation. Mol Psychiatry 2010; 15: 669–672.
9 Johnstone M, Thomson PA, Hall J, McIntosh AM, Lawrie SM, Porteous DJ. DISC1 in schizophrenia: genetic mouse models and human genomic imaging. Schizophrenia Bull 2011; 37: 14–20.
10 Moens LN, De Rijk P, Reumers J, Van den Bossche MJ, Glassee W, De Zutter S et al. Identifying high risk DISC1 structural variants with a 2% attributable risk for schizophrenia. Mol Psychiatry 2011; 16: 1096–11041063.
11 Song W, Li W, Li W, Feng J, Heston LL, Scaringe WA, Sommer SS et al. Identification of high risk DISC1 protein structural variants in patients with bipolar spectrum disorder. Neurosci Lett 2010; 486: 136–140.
12 Williams JM, Beck TF, Pearson DM, Proud MB, Holmans PA et al. Genome-wide association study identifies five new schizophrenia loci. Nat Genet 2011; 43: 969–976.
13 Sklar P, Ripke S, Scout LJ, Andreassen OA, Cichon S, Craddock N et al. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near OD24. Nat Genet 2011; 43: 977–983.
14 Moody JJ, Hillard CE, Kim Y, Morgan MB, Lewis LR, Muzny DM et al. Deep resequencing and association analysis of schizophrenia candidate genes. Mol Psychiatry 2012; 18: 138–140.
15 Reimold AM, Porteous DJ, Millar JK, Brandon NJ, Sawa A. DISC1 variation on neuroanatomical and neurocognitive phenotypes. Mol Psychiatry 2012; 17: 887–905.
16 Le-Niculescu H, Porteous DJ, Sawa A. Linking neurodevelopmental and synaptic theories of mental illness through DISC1. Nat Rev Neurosci 2012; 13: 707–722.
17 McAuley WS, Porteous DJ, Brandt SD, Cole A, Sweeting M et al. Convergent functional genomics of genome-wide association data for bipolar disorder: comprehensive identification of candidate genes, pathways and mechanisms. Am J Med Genet B Neuropsychiatr Genet 2009; 150B: 155–181.
18 Tiwary KH. The severity of mental disorders is linked to interaction among candidate genes. Integr Biol (Camb) 2012; 4: 1096–1101.
19 Insel TR. Rethinking schizophrenia. Nature 2010; 468: 187–193.
20 Porteous DJ, Millar JK, Brandon NJ, Sawa A. DISC1 at 10: connecting psychiatric genetics and neuroscience. Front Neuroendocrinol 2011; 32: 599–706.
21 Brandon NJ, Sawa A. Linking neurodevelopmental and synaptic theories of mental illness through DISC1. Nat Rev Neurosci 2011; 12: 707–722.
22 Soares DC, Carlyle BC, Bradshaw NJ, Porteous DJ. DISC1: Structure, Function, and Therapeutic Potential for Major Mental Illness. ACS Chem Neurosci 2012; 3: 609–632.
23 unpublished 2011. Frank CL, Madison JM, Koehler AN, Doud MK et al. Disrupted in schizophrenia 1 regulates neuronal progenitor proliferation via modulation of GSK3beta/beta-catenin signaling. Cell 2009; 136: 1017–1031.
24 Millar JK, Pickard BS, Mackie S, James R, Christie S, Buchanan SR et al. DISC1 and PDE4B are interacting genetic factors in schizophrenia that regulate CAPMP signaling. Science 2005; 310: 1187–1191.
25 Hashimoto R, Numakawa T, Ohnishi T, Kumaumari E, Yagasaki Y, Isihomoto T et al. Impact of the DISC1 Ser704Cys polymorphism on risk for major depression, brain morphology and ERK signaling. Hum Mol Genet 2006; 15: 3024–3033.
26 Kamiya A, Tomoda T, Chang J, Takaki M, Zhan C, Morita M et al. DISC1-NDEL1/NUDEL protein interaction, an essential component for neurite outgrowth, is modulated by genetic variations of DISC1. Hum Mol Genet 2006; 15: 3313–3323.
27 Burdick KE, Kamiya A, Hodgkinson CA, Lancez T, DeRoose P, Ishizuka K et al. Elucidating the relationship between DISC1, NDE1, and NDE2 and the risk for schizophrenia: Evidence of epistasis and competitive binding. Hum Mol Genet 2008; 17: 2462–2473.
28 Leffeld SV, Hendrikx P, Michel M, Sajani N, Bader V, Trossbach S et al. Oligomer assembly of the C-terminal DISC1 domain (640–854) is controlled by self-association motifs and disease-associated polymorphisms 5704C. Biochemistry 2009; 48: 7746–7755.
29 Eastwood SL, Hodgkinson CA, Harrison PJ. DISC1-Leu607Phe alleles differentially affect centrosomal PC11 localization and neurotransmitter release. Mol Psychiatry 2009; 14: 556–557.
30 Atkin TA, MacAskill AF, Brandon NJ, Kittler JT. Disrupted in Schizophrenia-1 regulates neuronal progenitor proliferation via modulation of PDE4B and PDE4A8. PLoS ONE 2009; 4: e7746–7755.
31 Atkin TA, MacAskill AF, Brandon NJ, Kittler JT. Disrupted in Schizophrenia-1 regulates neuronal progenitor proliferation via modulation of PDE4B and PDE4A8. PLoS ONE 2009; 4: e7746–7755.
32 Atkin TA, MacAskill AF, Brandon NJ, Kittler JT. Disrupted in Schizophrenia-1 regulates neuronal progenitor proliferation via modulation of PDE4B and PDE4A8. PLoS ONE 2009; 4: e7746–7755.
33 Millar JK, Pickard BS, Mackie S, James R, Christie S, Buchanan SR et al. DISC1 and PDE4B are interacting genetic factors in schizophrenia that regulate CAPMP signaling. Science 2005; 310: 1187–1191.
by intergenic splicing to DISC1, a gene disrupted by a translocation segregating with schizophrenia. Genomics 2000; 67: 69–77.

46 Hennah W, Porteous D. The DISC1 pathway modulates expression of neuro-developmental, synapticogenic and sensory perception genes. PLoS ONE 2009; 4: e4906.

47 Thompson PA, Macintyre DJ, Hamilton G, Dominiczak A, Smith BH, Morris A et al. Association of DISC1 variants with age of onset in a population-based sample of recurrent major depression. Mol Psychiatry 2012; 18: 745–747.

48 Walker RM, Hill AE, Newman AC, Hamilton G, Torrance HS, Anderson SM et al. The DISC1 promoter: characterization and regulation by FOXP2. Hum Mol Genet 2012; 21: 2862–2872.

49 Li H, Ruan J, Durbin R. Mapping short DNA sequencing reads and calling variants by a web-based tool set for whole-genome association and population-based linkage analyses. Bioinformatics 2009; 25: 655–661.

50 Tempel S. Using and understanding RepeatMasker. Methods Mol Biol 2012; 859: 29–51.

51 Ferrer-Costa C, Gelpi JL, Zamakola L, Parraga I, de la Cruz X, Orozco M. PMUT: a web-based tool for the annotation of pathological mutations on proteins. Bioinformatics 2005; 21: 3176–3178.

52 Thomas PD, Campbell MJ, Kejarival A, Mi H, Karaklak B, Daverman R et al. PANTHER: a library of protein families and subfamilies indexed by function. Genome Res 2003; 13: 2129–2141.

53 Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P et al. A method and server for predicting damaging missense mutations. Nat Methods 7: 248–249.

54 Blom N, Gammeltoft S, Brunak S. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. J Mol Biol 1999; 294: 1351–1362.

55 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007; 81: 559–575.

56 Petersen CGJ. The yearly immigration of young place into the Limfjord from the German Sea. Danish Biol Station 1896; 6: 1–77.

57 Lincoln FC. Calculating Waterfowl Abundances on the Basis of Banding Returns, 1930, US Department of Agriculture Circular, 118, 1–4.

58 Chapman DG. Some properties of the hypergeometric distribution with applications using mapping quality scores. Genome Res 2008; 18: 1851–1858.

59 Chelala C, Khan A, Lemoine NR. SNPnexus: a web database for functional annotation of newly discovered and public single nucleotide polymorphisms. Bioinformatics 2009; 25: 635–641.

60 Purcell SM, Wray NR, Stone JL, Visscher PM, O’Donovan MC, Sullivan PF et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature 2009; 460: 748–752.

61 Consortium TGP. A map of human genome variation from population-scale sequencing. Nature 2010; 467: 1061–1073.

62 Davies G, Tenesa A, Payton A, Yang J, Harris SE, Liewald D et al. Genome-wide association studies establish that human intelligence is highly heritable and polygenic. Mol Psychiatry 2011; 16: 996–1005.

63 Hennah W, Porteous D, Thompson PA, Macintyre DJ, Hamilton G, Dominiczak A, Smith BH, Morris A et al. Association of DISC1 variants with age of onset in a population-based sample of recurrent major depression. Mol Psychiatry 2012; 18: 745–747.

64 Walker RM, Hill AE, Newman AC, Hamilton G, Torrance HS, Anderson SM et al. The DISC1 promoter: characterization and regulation by FOXP2. Hum Mol Genet 2012; 21: 2862–2872.

65 Li H, Ruan J, Durbin R. Mapping short DNA sequencing reads and calling variants by a web-based tool set for whole-genome association and population-based linkage analyses. Bioinformatics 2009; 25: 655–661.

66 Tempel S. Using and understanding RepeatMasker. Methods Mol Biol 2012; 859: 29–51.

67 Lin CH, Chen PR, Shih SH, Liu YH, Lin CH, Hsu YH et al. Polymorphisms disrupt Wnt/GSK3beta signaling and brain development. Neuron 2008; 54: 155–168.

68 Song W, Li W, Noltner K, Yan J, Green E, Grozeva D et al. Identification of high risk DISC1 structural variants in patients with bipolar spectrum disorder. Neurosci Lett 2010; 486: 136–140.

69 Hennah W, Thomson P, McQuillan A, Bass N, Loukola A, Anjorin A et al. DISC1 association, heterogeneity and interplay in schizophrenia and bipolar disorder. Mol Psychiatry 2009; 14: 865–873.

70 Sawamura N, Ando T, Maruyama Y, Fujimuro M, Mochizuki H, Honjo K et al. Nuclear DISC1 regulates CRE-mediated gene transcription and sleep homeostasis in the fruit fly. Mol Psychiatry 2008; 13: 1138–11481069.

71 Murdoch H, Mackie S, Collins DM, Hill EV, Bolger GB, Klussmann E et al. Isoform-selective susceptibility of DISC1 phosphodiesterase-4 complexes to dissociation by elevated intracellular cAMP levels. J Neurosci 2007; 27: 9513–9524.

72 Singh RK, De Rienzo G, Drane L, Mao Y, Flood Z, Madison J et al. Common DISC1 polymorphisms disrupt Wnt/GSK3beta signaling and brain development. Neuron 72: 545–558.

73 Kievz A, Garimella K, Do R, Stitziel NO, Neale BM, McLaren PJ et al. Exome sequencing and the genetic basis of complex traits. Nat Genet 2012; 44: 623–630.

74 Sullivan PF, Wells JE, Joyce PR, Bushnell JA, Mulder RT, Oakley-Browne MA. Family history of depression in clinic and community samples. J Affect Disord 1996; 40: 159–168.

75 Eaton WW, Anthony JC, Gallo J, Cai G, Tien A, Romanoski A et al. Natural history of Diagnostic Interview Schedule/DSM-IV major depression. The Baltimore Epidemiologic Catchment Area follow-up. Arch Gen Psychiatry 1997; 54: 993–999.

76 Colantuoni C, Lipska BK, Ye T, Hyde TM, Tao R, Leek JT et al. Temporal dynamics and genetic control of transcription in the human prefrontal cortex. Nature 2011; 478: 519–523.

77 Bernstein BE, Bioney E, Dunham I, Green ED, Gunter C, Snyder M. An integrated encyclopedia of DNA elements in the human genome. Nature 2012; 489: 57–74.

78 Nicolae LC, Le-Nicolae H. The P-value illusion: how to improve (psychiatric) genetic studies. Am J Med Genet B Neuropsychiatr Genet 2010; 153B: 847–849.

This work is licensed under a Creative Commons Attribution 3.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by/3.0/