Title	Cognitive analysis of schizophrenia risk genes that function as epigenetic regulators of gene expression
Author(s)	Whitton, Laura; Cosgrove, Donna; Clarkson, Christopher; Harold, Denise; Kendall, Kimberley; Richards, Alex; Mantripragada, Kiran; Owen, Michael J.; O'Donovan, Michael C.; Walters, James; Hartmann, Annette; Konte, Betina; Rujescu, Dan; Gill, Michael; Corvin, Aiden; Rea, Stephen; Donohoe, Gary; Morris, Derek W.
Publication Date	2016-10-20
Publication Information	Whitton, Laura, Cosgrove, Donna, Clarkson, Christopher, Harold, Denise, Kendall, Kimberley, Richards, Alex, Mantripragada, Kiran, Owen, Michael J., O'Donovan, Michael C., Walters, James, Hartmann, Annette, Konte, Betina, Rujescu, Dan, Gill, Michael, Corvin, Aiden, Rea, Stephen, Donohoe, Gary, Morris, Derek W. (2016). Cognitive analysis of schizophrenia risk genes that function as epigenetic regulators of gene expression. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 171(8), 1170-1179. doi:https://doi.org/10.1002/ajmg.b.32503
Publisher	Wiley
Link to publisher's version	https://doi.org/10.1002/ajmg.b.32503
Item record	http://hdl.handle.net/10379/16277
DOI	http://dx.doi.org/10.1002/ajmg.b.32503


BY NC ND

Cognitive analysis of schizophrenia risk genes that function as epigenetic regulators of gene expression

Laura Whitton[1], Donna Cosgrove [1], Christopher Clarkson[2], Denise Harold[3,4], Kimberley Kendall[5], Alex Richards[5], Kiran Mantripragada[5], Michael J Owen[5], Michael C. O'Donovan[5], James Walters[5], Annette Hartmann[6], Betina Konte[6], Dan Rujescu[6], WTCCC2, Michael Gill[3], Aiden Corvin[3], Stephen Rea[2], Gary Donohoe [1] & Derek W Morris [1]

[1] Cognitive Genetics and Cognitive Therapy Group, Neuroimaging, Cognition & Genomics (NICOG) Centre & NCBES Galway Neuroscience Centre, School of Psychology and Discipline of Biochemistry, National University of Ireland Galway, Ireland.

[2] Centre for Chromosome Biology, Discipline of Biochemistry, National University of Ireland Galway, Ireland.

[3] Neuropsychiatric Genetics Research Group, Institute of Molecular Medicine and Discipline of Psychiatry, Trinity College Dublin, Ireland.

[4] School of Biotechnology, Dublin City University, Dublin, Ireland.

[5] MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University School of Medicine, Cardiff, UK.

[6] Department of Psychiatry, University of Halle, Halle, Germany.

Corresponding author:

Dr Derek Morris, Room 106, Discipline of Biochemistry, National University of Ireland Galway,

University Road, Galway, Ireland.

Tel: + 353 91 494439

Email: derek.morris@nuigalway.ie

Abstract

Epigenetic mechanisms are an important heritable and dynamic means of regulating various genomic functions, including gene expression, to orchestrate brain development, adult neurogenesis and synaptic plasticity. These processes when perturbed are thought to contribute to schizophrenia pathophysiology. A core feature of schizophrenia is cognitive dysfunction. For genetic disorders where cognitive impairment is more severe such as intellectual disability, there are a disproportionally high number of genes involved in the epigenetic regulation of gene transcription. Evidence now supports some shared genetic aetiology between schizophrenia and intellectual disability. GWAS have identified 108 chromosomal regions associated with schizophrenia risk that span 350 genes. This study identified genes mapping to those loci that have epigenetic functions, and tested the risk alleles defining those loci for association with cognitive deficits. We developed a list of 350 genes with epigenetic functions and cross-referenced this with the GWAS loci. This identified 8 candidate genes: *BCL11B, CHD7, EP300, EPC2, GATAD2A, KDM3B, RERE, SATB2*. Using a dataset of Irish psychosis cases and controls(n=1235), the schizophrenia risk SNPs at these loci were tested for effects on IQ, working memory, episodic memory and attention. Strongest associations were for rs6984242 with both measures of IQ(p=0.001) and episodic memory(p=0.007). We link rs6984242 to *CHD7* via a long range eQTL. These associations were not replicated in independent samples. Our study highlights that a number of genes mapping to risk loci for schizophrenia may function as epigenetic regulators of gene expression but further studies are required to establish a role for these genes in cognition.

Introduction

Epigenetic mechanisms are heritable and have an important role in the dynamic regulation of various genomic functions, including gene expression. This is largely achieved through the covalent modification of DNA and histones. Epigenetic chromatin modifications establish and maintain gene expression programs to help orchestrate brain development, adult neurogenesis and synaptic plasticity (Qureshi and Mehler, 2013), processes that when perturbed are thought to contribute to the pathophysiology of schizophrenia (Wockner et al., 2014, Liu et al., 2014). Recent exome sequencing (McCarthy et al., 2014, Takata et al., 2014) has identified rare mutations in genes associated with epigenetic chromatin remodeling in schizophrenia. Pathway analysis of genome-wide association study (GWAS) data has supported a role for epigenetic mechanisms (histone methylation) in major psychiatric disorders (O'Dushlaine, 2015).

A core feature of schizophrenia is cognitive dysfunction in the form of impairments in memory, attention and IQ. This has a genetic basis and decline in cognitive performance occurs years prior to onset of illness as the brain is still developing (Kahn and Keefe, 2013). For genetic disorders where cognitive impairment is more severe than in schizophrenia such as intellectual disability, there are a disproportionally high number of genes involved in the epigenetic regulation of gene transcription (Kleefstra et al., 2014). Loss-of-function (LoF) mutations in the *MECP2* gene (encoding a methyl-DNA binding protein) causing Rett syndrome is the classic example of an "epigenetic gene" being mutated to cause a disorder with severe cognitive deficits. There are currently >50 genes identified that function in different epigenetic mechanisms and are mutated in cognitive disorders (Kleefstra et al., 2014). Evidence now supports some shared genetic aetiology between schizophrenia and intellectual disability (Fromer et al., 2014). Therefore, it is pertinent to ask if an overlapping phenotypic characteristic of schizophrenia and intellectual disability (cognitive deficits) could be due to genetic variability in a shared pathobiology (epigenetic mechanisms).

A meta-analysis of over 50 GWAS studies performed by the Psychiatric Genomics Consortium (PGC) has culminated in the identification of 108 risk loci for schizophrenia that contain some 350 genes (Schizophrenia Working Group of the Psychiatric Genomics, 2014). Our study asks: do any genes at these new schizophrenia risk loci regulate epigenetic mechanisms, specifically as chromatin modulators of gene expression? For those loci containing those genes, we next ask whether they independently associate with cognitive functions? Our overall hypothesis is that GWAS-identified risk genes for schizophrenia, which regulate epigenetic mechanisms, are associated with deficits in cognitive function.

Materials & Methods:

Chromatin modulators

We prepared a list of genes encoding the major classes of chromatin modulators which encompasses epigenetic regulators of gene expression (Table 1). Relevant genes were collected from two of the main online gene databases; Gene Cards (http://www.genecards.org/) and the Gene Ontology Consortium (http://geneontology.org/) as follows: Searches of these databases were conducted using specific keywords to distinguish the type of chromatin modulation that was being investigated. For example, in the case of Genecards, a human gene database, we used the advanced search option (http://www.genecards.org/Search) using terms such as "histone AND acetyltransferase" to search the summaries section. In the case of Gene Ontology we searched the ontology database (http://amigo2.berkeleybop.org/amigo/search/ontology) using terms such as "histone acetylation" giving the branch "GO:0016573 histone acetylation" under inferred tree view, following the "link to all genes and gene products annotated", filtering with the term "Homo sapiens" returns a list of relevant genes. In addition to using these two databases, in the case of histone ubiquitination we have included the full list of E1 ubiquitin activating enzymes, and E2 ubiquitin conjugating enzymes as it is not always known which of these coordinates with the E3 ubiquitin ligase enzyme to mediate the modification. This list was selected from the list compiled by van Wijk and Timmers (van Wijk and Timmers, 2010). The two lists for the same chromatin modulation were combined, duplicate entries were removed and each gene was then manually queried for legitimate inclusion. At this point an associated list was appended onto the list of known modulators, this includes factors that exist in complex with the enzyme in question, and are required for its activity (Supplementary Table 1). Also shown in the list are the target residues/histone of the enzyme where known, and a brief statement of evidence as taken from Genecards' summary section, which summarises gene function from Entrez Gene, UniProtKB (UniprotKB / Swiss-Prot / UniprotKB / TrEMBL), Tocris Bioscience, PharmGKB, and Gene Wiki. These genes were assembled into a final list of 350 unique chromatin-modulating genes (Supplementary Table 1), which from here on we refer to as "epigenetic genes". The epigenetic gene list (n = 350) was cross-referenced with genes located in 108 chromosomal regions associated with schizophrenia risk (n = 350) (Schizophrenia Working Group of the Psychiatric Genomics, 2014) to identify candidate genes for this study.

<Table 1 here>

Study Participants and Neuropsychological Testing

This study used an Irish dataset of broad psychosis cases (n =890) and controls (n =330) who had completed tests in five main areas of cognition and for whom GWAS data were available. Cases were clinically stable patients with a DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, 4th Edition) diagnosis of schizophrenia (n = 559), schizoaffective disorder (n = 111), bipolar disorder (n = 148), major depressive disorder with psychotic features (n = 38), or psychosis not otherwise specified (n = 34) recruited from five sites across Ireland (Table 2). Inclusion criteria required that participants were clinically stable at the time of neuropsychological testing, were aged 18 to 65 years, had no history of comorbid psychiatric disorder, no reported substance abuse in the preceding 6 months, no prior head injury with loss of consciousness, and no history of seizures. Diagnosis was confirmed by trained psychiatrists using the Structured Clinical Interview for DSM-IV-TR (First M, 2002). A more detailed description of sample ascertainment has previously been published (Cummings et al., 2013). Healthy control participants were recruited via online and poster

advertising. They were aged 18 to 65 years, with no reported history of substance abuse in the preceding 6 months, no prior head injury with loss of consciousness, no history of seizures, and no personal history of psychosis or in their first-degree relatives. All assessments were conducted in accordance with the relevant ethics committees' approval from each participating site. All participants had four grandparents born in Ireland and provided written informed consent.

<Table 2 here>

Participants completed a neuropsychological assessment battery designed to target the cognitive deficits observed in schizophrenia, this battery included general cognitive function, episodic memory, working memory and attentional control. General cognitive functioning (IQ) was measured using selected subtests (Vocabulary, Similarities, Block Design, and Matrix Reasoning) from the Wechsler Adult Intelligence Scale (WAIS; (Weschler, 1997a) which assessed verbal, performance and full-scale IQ as well as the Wechsler Test of Adult Reading (WTAR) (Wechsler, 2001). Episodic memory was assessed using the logical memory (LM) 1 & 2 tasks as well as the "Faces 1 & 2" task subtests from the Wechsler Memory Scale (WMS) III (Weschler, 1997b) as well as the Paired-Associates Learning task (PAL). Working memory was assessed using the spatial working memory task (SWM) from the Cambridge Neuropsychological Test Automated Battery (CANTAB) (Robbins TW, 1994) and letter-number sequencing from the WMS III (Weschler, 1997b). Attentional control was assessed using the Continuous Performance Task (CPT)–Identical Pairs version, and the Sustained Attention to Response Task (Cornblatt BA, 1988, Robertson, 1994).

Genotyping and Statistical Analysis

Genotyping was conducted on DNA extracted from blood or saliva from patient and control participants. SNP data were obtained from two different sites; a recent GWAS using the Affymetrix SNP Array 6.0 platform, conducted as part of the Wellcome Trust Case Control Consortium 2 (Donnelly P, 2012) and a collaborative GWAS with Cardiff University using an Illumina HumanCoreExome (+custom) SNP array. Full GWAS data on these samples included recently completed imputation using 1000 Genomes Phase I data and IMPUTE2 (Howie et al., 2009) to give ~10 million SNPs genome-wide per sample. Only samples that had passed the quality control filtering were imputed using the 1000 Genomes reference panel. Imputed data were converted to best guess genotypes using PLINK (Purcell et al., 2007), which was also used to extract data on the index SNP at each gene from GWAS data and combine it in Statistical Package for the Social Sciences (SPSS) 21, IBM) with data from an extensive neuropsychological battery. Statistical analysis was carried out in SPSS using a linear based model of regression. Because of differences in cognitive scores between male and female participants and between older and younger individuals, gender and the age at the time of assessment were used as co-variate measures to determine a general genotype effect across all cases and controls. Cognitive test scores for each genotype group were compared using a three group analysis, i.e. homozygous non-risk versus heterozygous versus homozygous risk genotype groups. R^2 is a measure of how close the data fits to the regression line. R^2 change values were used in this analysis which identifies changes in the original R^2 based on the linear contribution of variables added into the regression model. Regarding power in this sample, which in itself is too small for GWAS, to detect e.g. differences in Performance IQ for a SNP where the mean difference in scores between risk allele carriers and non-carriers was 0.5 standard deviations (SD), our sample had

87% power at α=0.05. For a mean difference of 0.33 SD or 0.75 SD, the power was 59% and 99% respectively for the same type I error rate.

Interaction Analysis

Online tools and resources including STRING (http://string-db.org/) and the ingenuity pathway analysis (IPA) database (http://www.ingenuity.com/) were used to identify direct (physical binding) and indirect (functional) interacting partners of each candidate gene based on genomic and co-expression data as well as experimental data from literature searches . Gene IDs from each interaction network were compiled into lists and cross-referenced with the epigenetic gene list (n = 350) and the GWAS schizophrenia gene list (n =350) to determine if any interacting partners had epigenetic functions and were associated with schizophrenia risk. Among the candidate genes, three pairs of genes were identified as interacting partners: *SATB2* and *BCL11B*, *EP300* and *BCL11B* and *CHD7* and *GATAD2A*. All 3 combinations were taken forward for neuropsychological interaction analysis using the same Irish dataset. Interaction analysis was carried out in SPSS based on a linear model of regression using up to five genotype groups where study participants were assigned to groups based on the number of schizophrenia risk alleles they carried at the two SNPs from a pair of interacting genes. Gender and age at time of assessment were used as co-variate measures in this analysis.

UK Replication Sample

772 individuals with schizophrenia or schizoaffective disorder - depressive subtype were recruited from community mental health teams in Wales and England. The Schedule for Clinical Assessment in Neuropsychiatry (SCAN) interview was conducted followed by case note review and consensus diagnosis according to DSM-IV criteria. Individuals were assessed using the Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) Consensus Cognitive Battery (MCCB). The MCCB is a 10 test cognitive battery consisting of seven cognitive domains: speed of processing, attention and vigilance, working memory, verbal learning, visual learning, reasoning and problem solving and social cognition (Nuechterlein et al., 2008). These domains were chosen as primary cognitive outcomes given that the MATRICS tests were selected on the basis of the domains they measured, and the fact that this approach minimised measurement error as some domains are assessed by more than one test. All interviewers were trained in the use of the SCAN interview and MCCB. The UK Multicentre Research Ethics Committee (MREC) approved the study and all participants provided valid informed consent. These samples were genotyped at the Broad Institute, Massachusetts on the Illumina HumanOmniExpressExome-8v1. QC procedures followed those of the PGC (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). Imputation was carried out by the PGC using IMPUTE2 and for this analysis we converted imputed to best guess genotypes using PLINK (Howie et al., 2009, Purcell et al., 2007).

German Replication Sample

Patients with a DSM-IV diagnosis of schizophrenia were ascertained from mental health services in the Munich area. Exclusion criteria were a history of head injury or neurological disease. Patients, ages 18–80 years, were interviewed with the Structured Clinical Interview for DSM-IV (SCID), and the

interviews were rated by psychiatrists or psychologists. Healthy comparison subjects of German descent, ages 18–80 years, were randomly selected from population registers from the Munich area. Participants underwent an extensive screening process to exclude those with neurological or psychotic disorders and those who had first-degree relatives with psychotic disorders. In the case of participants older than 60 years, the Mini-Mental State Examination was employed to exclude individuals with possible cognitive impairment (Walters et al., 2013). In this sample, IQ was measured by the full German Wechsler Adult Intelligence Scale–Revised (WAIS-R). Episodic memory was assessed using the immediate and delayed logical memory tests from the German Wechsler Memory Scale–Revised (WMS-R)(Härting, 2000). For Full Scale IQ and Verbal IQ, data were available for 355 patients and 2,147 healthy comparison subjects. For Episodic Memory, data were available for 362 patients and 558 healthy comparison subjects. GWAS data were available for these samples. Individuals were genotyped on different platforms and imputed in six batches. Quality control and imputation of batches 1 (Human610-Quadv1_B, Human660W-Quad_v1_A, Human610-Quadv1_B), 2 (HumanHap 300 v1.0.0) and 3 (Affymetrix 6.0) were performed in the framework of the PGC2 schizophrenia meta-analysis (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). Batches 4 (HumanHap 300 v1.0.0), 5 (Illumina Human OmniExpress 12v1.0) and 6 (Illumina Omni1-Quad) were imputed following the same QC and imputation protocols with minor adaptations. Sample call-rate was adjusted to account for smaller sample sizes with a minimum threshold of 96%. After imputation the batches were combined and all markers that differed more than 0.1 in minor allele frequency between any of the batches were excluded. Association analysis was done with PLINK using a linear regression model corrected for age, sex, education, affection status and batch.

Results

The PGC GWAS identified SNPs associated with SZ in 108 loci that contain 350 genes (Schizophrenia Working Group of the Psychiatric Genomics, 2014). Cross referencing our list of chromatin modulators with the genes in the GWAS schizophrenia risk loci identified a shortlist of 17 epigenetic genes in schizophrenia risk regions. As some of the risk loci include large numbers of genes, we systematically reviewed each GWAS results plot to conservatively reduce our gene list to seven genes (*BCL11B, EP300, EPC2, GATAD2A, KDM3B, RERE* and *SATB2*), where the associated SNP is located within or close to the gene location. Three of these genes, *BCL11B, EPC2* and *SATB2*, are the only gene within their associated locus. For the remaining four genes, positional and linkage disequilibrium evidence from the relevant plots was strong enough to suggest that they could be affected by the risk SNPs at these loci.

We also considered the possibility that a risk SNP may affect a gene outside its risk region. In the schizophrenia GWAS it is worth noting that there are 21 loci that do not contain any genes. To identify such distantly affected genes, we took advantage of a recent genome-wide brain expression quantitative trait loci (eQTL) study (Kim et al., 2014). This study reports results for eQTL analysis between SNPs and the expression of genes within 1 Mb of the SNP (i.e. a 2 Mb window). We extracted the eQTL results for all 128 schizophrenia index risk SNPs reported by the PGC and found that 59 SNPs had an eQTL at nominal $p < 0.05$. For each of these SNPs we calculated the number of genes within the 2 Mb window and Bonferroni corrected for that number of genes at each SNP (Dunn, 1961). This identified 16 unique SNPs with significant eQTL results that totalled 21 genes (Supplementary Table 2). We cross-referenced these 21 genes with our epigenetic gene list and identified one extra gene for study, *CHD7*. For each candidate gene, the highest associated SNP (the index SNP from the GWAS region containing that gene) was used for analysis of neuropsychological variables.

The final eight genes (with corresponding index SNPs) for analysis are listed here including brief details of function and descriptions of any cognitive disorders caused by these genes:

(1) *BCL11B* (rs2693698) encodes B-Cell CLL/Lymphoma 11B, a protein associated with the Mi-2/nucleosome remodelling and deacetylase complex (NuRD) (Cismasiu et al., 2005).

(2) *CHD7* (rs6984242) encoding chromodomain helicase DNA binding protein 7, is an ATP-dependent chromatin remodeller. *De novo* mutations in *CHD7* cause Charge syndrome, a disorder characterised by abnormal ear formation, hearing loss, cleft palate, coloboma, genital hypoplasia, delayed development and intellectual disability (Lalani et al., 2006). *De novo* mutations in *CHD7* have also been found in cases of autism (Jiang et al., 2013, O'Roak et al., 2012b) and *CHD7* has been shown to interact directly and indirectly with *CHD8*, a confirmed risk gene for autism (Batsukh et al., 2010, Wilkinson et al., 2015).

(3) *EP300* (rs9607782) encodes the histone acetyltransferase, E1A Binding Protein P300. *De novo* mutations within the *EP300* gene cause rare cases of Rubinstein-Taybi syndrome, a disorder characterized by intellectual disability and developmental abnormalities (Negri et al., 2014).

(4) *EPC2* (rs200327371) encodes Enhancer Of Polycomb Homolog 2, protein associated with a chromatin repressive complex. Microdeletions within the 2q23.1 region containing *EPC2* are

associated with intellectual disability, seizures, microcephaly, development delay, hypotonia and behavioural features similar to autism or Angelman syndrome (van Bon et al., 2010).

(5) *KDM3B* (rs10043984), Lysine (K)-Specific Demethylase 3B is required for normal spermatogenesis and sexual behaviours in male mice (Liu et al., 2015) and encodes an enzyme that removes a key transcriptional repressive modification from chromatin (Kim et al., 2012). Levels of this histone modification correlates with symptom severity in schizophrenia in a sex-specific manner (Chase et al., 2015).

(6) *GATAD2A* (rs2905426) is a transcriptional repressor and subunit of the NuRD complex (Wade et al., 1999, Brackertz et al., 2002).

(7) *RERE* (rs34269918) is a transcriptional co-repressor that binds chromatin and in mouse is involved in cerebellar development (Kim and Scott, 2014).

(8) *SATB2* (rs6704641) encodes a protein involved in chromatin remodelling. *De novo* structural and point mutations in *SATB2* result in *SATB2* haploinsufficiency and a defined *SATB2* phenotype with developmental delay, mild to severe intellectual disability, behavioural problems and abnormal craniofacial features, specifically a cleft palate (Docker et al., 2014).

Each of the eight schizophrenia risk variants was analysed against four areas of cognitive function (IQ, working memory, episodic memory and attentional control). To maximize power, this was done in the combined sample of psychosis cases and controls. Table 3 highlights all nominally significant associations detected. The strongest results were for rs6984242 (*CHD7*) where one association survives Bonferroni correction for the four cognitive domains tested and the eight variants tested (corrected p value threshold = 0.002). For that association, the schizophrenia risk allele (G) was associated with lower Verbal IQ ($p = 0.001$) where an R^2 change value of 0.010 indicates that 1.0% of the variance in test performance is explained by the genotype. The same variant was associated with two other measures of IQ (Full Scale IQ, $p = 0.007$, and WTAR, $p = 0.003$), as well as four measures of episodic memory function with the strongest result for the Faces task, delayed condition ($p= 0.007$).

<Table 3 here>

Of the remaining seven variants, four variants (at *EP300, GATAD2A, KDM3B* and *RERE*) had nominally significant associations with at least one cognitive task (Table 3). Of note, the risk allele for rs2905426 (*GATAD2A*, which interacts with *CHD7*) was also associated with lower Full Scale IQ ($p = 0.016$). To follow this up we performed an interaction analysis using the data for the variants at *CHD7* and *GATAD2A*. In this regression analysis participants were categorized into four groups on the basis of the number of risk alleles they carried (0, 1, 2 or 3/4 (few participants carried 4 risk alleles so these individuals were grouped with those that carried 3 risk alleles)). Association was found with four measures of IQ (Table 4): Full Scale IQ ($p = 0.001$), Verbal IQ ($p = 0.002$), WTAR ($p = 0.005$) and Performance IQ ($p = 0.016$). For Full Scale IQ, the R^2 change value for the 2-SNP analysis was 0.013 whereas the R^2 change value for rs6984242 (*CHD7*) on its own was 0.008. For Verbal IQ, the R^2 change value for both the 2-SNP and single SNP analysis is 0.010 and the range of mean scores for the 2-SNP and single SNP analysis largely overlap. For other nominally significant associations from the 2-SNP analysis, the results do not differ much from the single SNP analysis. Thus, there is no strong evidence from this analysis that the effect from a genetic interaction between these two

variants is contributing more significantly to poorer cognitive performance than the individual variants on their own. Interaction analysis using the two other pairs of interacting partners identified (*SATB2* and *BCL11B*, and *EP300* and *BCL11B*) did not detect any significant associations.

<Table 4 here>

We sought independent replication of our Bonferroni-corrected significant association between Verbal IQ and rs6984242 (*CHD7*). In doing these analyses, we also sought independent support for the associations we detected between this SNP and Full Scale IQ and with Episodic Memory. We used two independent datasets. The UK dataset is based on the MATRICS battery and so composite scores were used for cognitive domains. Similar to the Irish dataset, the German dataset used the WAIS to measure IQ and the WMS to measure Episodic Memory. rs6984242 did not show association with IQ or Episodic Memory in either dataset (Tables 5 and 6).

<Table 5 here>

<Table 6 here>

Discussion

This study developed a list of 350 genes that function as chromatin modulators of gene expression. These genes were investigated for a role in cognitive deficits in schizophrenia by first determining which of these genes are present at loci that are genome-wide significant for schizophrenia from the largest GWAS to date. The eight candidate schizophrenia risk genes identified had various epigenetic functions and a number are the site of causative mutations for rare cognitive disorders, e.g. Charge syndrome (*CHD7*), Rubinstein-Taybi syndrome (*EP300*) and *SATB2*-associated syndrome. Five of the eight risk SNPs in question (rs6984242, rs2905426, rs9607782, rs10043984 and rs34269918) were found to be associated with poorer performance in tests from three of the four cognitive domains; IQ, episodic memory and attention. For seven of the eight candidate genes, the schizophrenia risk SNP is located within or near the gene of interest. The impact of these variants on gene function is not certain. For three genes, there is brain eQTL data that identifies the potential functional impact of SNPs on gene expression. Significant eQTLs for rs9607782 and *EP300*, for rs6984242 and CHD7 and for rs2905426 and GATAD2A are reported in the Kim et al. (2014) brain eQTL data that included 424 brain samples. Finally, our extended analysis of this eQTL dataset identifies *CHD7* as a putative schizophrenia risk gene for the first time. The index SNP in this instance is rs6984242 and the association signal from the GWAS maps to chr8:60,475,469-60,954,469(hg19), with the SNP located 891Kb upstream of the first exon of *CHD7*. There are only four genes within the 2Mb region that centres on rs6984242 and *CHD7* is the only gene with a significant eQTL involving this SNP in the Kim et al. dataset. In an independent but much smaller brain eQTL dataset (n=134; www. braineac.org) the rs6984242-*CHD7* eQTL is not supported by data from the ten different brain regions analysed.

CHD7 belongs to the chromodomain helicase DNA (CHD) binding family of proteins, where it is a member of subgroup III (along with CHD6, CHD8 and CHD9) based on sequence similarities and additional functional domains (Marfella and Imbalzano, 2007). *CHD* genes are members of the SNF2 superfamily of ATP-dependent chromatin remodelers (Flaus et al., 2006, Marfella and Imbalzano, 2007) and function as part of large multisubunit complexes to regulate gene transcription, DNA repair, replication and recombination. *CHD7* interacts directly and indirectly with *CHD8 (Batsukh et al., 2010)*, a gene that is recurrently mutated in autism (O'Roak et al., 2012a). As brain eQTL and related functional datasets increase in size and power, we will seek more definite evidence to associate *CHD7* with schizophrenia risk via rs6984242.

It is rs6984242 that provides the strongest evidence of association where the schizophrenia risk allele (G) was associated with lower Verbal IQ ($p = 0.001$) and explains 1.0% of the variance in test performance in our sample. This variant was also associated with other measures of IQ and episodic memory ($p < 0.01$). A combined analysis of rs6984242 with rs2905426 (at *GATAD2A*, known to interact with *CHD7*) also revealed a number of associations with IQ but the effect sizes did not point to an epistatic interaction between the two variants. Our analyses combined psychosis cases and healthy participants together to maximise sample size and power. When, in *post hoc* analyses, these samples were considered separately, the association between the risk allele at rs6984242 was not significant in the healthy sample, but was marginally stronger in the case sample than for the sample as a whole (for example, WTAR ($p = 0.004$; $R^2 = 0.013$), Verbal IQ ($p = 0.002$; $R^2 = 0.012$), Full Scale IQ (($p = 0.003$; $R^2 = 0.014$), Faces 2 ($p = 0.009$; $R^2 = 0.013$)).

Our attempts to replicate significant results were conservatively restricted to following up the associations between rs6984242 and IQ and episodic memory. We did not uncover independent data to support our associations in UK and German samples or in a combined meta-analysis (data not shown). Similarly, restricting the replication analyses to case-only samples did not uncover significant results. There are a number of caveats to consider that could explain the non-replication. For the UK sample, comparable measures of general cognitive ability and memory were not available. In the German sample, where the same measures were available, the samples consisted mainly of healthy participants, for whom the association between rs6984242 and cognition was weaker in the Irish data. Finally, the modest though significant amount of variance associated with rs6984242, which was estimated as being around 1%, may also suggest that the samples included here, though large for these types of studies, lacked power to detect association.

Overall our study finds up to eight genes that function as chromatin modulators of gene expression and that map to loci that are associated with schizophrenia in the largest GWAS published to date. Rare mutations in a number of genes cause rare cognitive disorders. We found strong association between rs6984242, which we link to *CHD7* via an eQTL, and measures of IQ and episodic memory in our dataset. However, we did not find replicating evidence of association between this variant in independent datasets. Further eQTL, functional and behavioural studies will be required both to confirm a link between rs6984242 and *CHD7*, thereby implicating this gene in schizophrenia biology, and to establish an association between this variant and cognition.

Acknowledgements

We have no conflict of interest to declare. We wish to thank all patients and their support staff, and all healthy volunteers for participating in the data collection on which this manuscript is based. Recruitment, genotyping and analysis was supported by Science Foundation Ireland (grants 12/IP/1670, 12/IP/1359 and 08/IN.1/B1916) and the Wellcome Trust Case Control Consortium 2 project (grants 085475/B/08/Z and 085475/Z/08/Z) and the Wellcome Trust (grants 072894/Z/03/Z, 090532/Z/09/Z and 075491/Z/04/B). We thank Lucinda Hopkins and Lesley Bates for assistance in generating genotype data at Cardiff University.

The Wellcome Trust Case Control Consortium 2 investigators include: Peter Donnelly, Lesley Bates, Ines Barroso, Jenefer M. Blackwell, Elvira Bramon, Matthew A. Brown, Juan P. Casas, Aiden Corvin, Panos Deloukas, Audrey Duncanson, Janusz Jankowski, Hugh S. Markus, Christopher G. Mathew, Colin N. A. Palmer, Robert Plomin, Anna Rautanen, Stephen J. Sawcer, Richard C. Trembath, Ananth C. Viswanathan, Nicholas W.Wood, Chris C. A. Spencer, Gavin Band, Céline Bellenguez, Colin Freeman, Garrett Hellenthal, Eleni Giannoulatou, Lucinda Hopkins, Matti Pirinen, Richard Pearson, Amy Strange, Zhan Su,Damjan Vukcevic, Cordelia Langford, Sarah E. Hunt, Sarah Edkins, Rhian Gwilliam, Hannah Blackburn, Suzannah J. Bumpstead, Serge Dronov, Matthew Gillman, Emma Gray, Naomi Hammond, Alagurevathi Jayakumar, Owen T.McCann, Jennifer Liddle, Simon C. Potter, Radhi Ravindrarajah, Michelle Ricketts, Matthew Waller, Paul Weston, Sara Widaa, and Pamela Whittaker.

References

BATSUKH, T., PIEPER, L., KOSZUCKA, A. M., VON VELSEN, N., HOYER-FENDER, S., ELBRACHT, M., BERGMAN, J. E., HOEFSLOOT, L. H. & PAULI, S. 2010. CHD8 interacts with CHD7, a protein which is mutated in CHARGE syndrome. *Hum Mol Genet,* 19**,** 2858-66.

BRACKERTZ, M., BOEKE, J., ZHANG, R. & RENKAWITZ, R. 2002. Two highly related p66 proteins comprise a new family of potent transcriptional repressors interacting with MBD2 and MBD3. *J Biol Chem,* 277**,** 40958-66.

CHASE, K. A., ROSEN, C., RUBIN, L. H., FEINER, B., BODAPATI, A. S., GIN, H., HU, E. & SHARMA, R. P. 2015. Evidence of a sex-dependent restrictive epigenome in schizophrenia. *J Psychiatr Res,* 65**,** 87-94.

CISMASIU, V. B., ADAMO, K., GECEWICZ, J., DUQUE, J., LIN, Q. & AVRAM, D. 2005. BCL11B functionally associates with the NuRD complex in T lymphocytes to repress targeted promoter. *Oncogene,* 24**,** 6753-6764.

CORNBLATT BA, R. N., FARIS G, FRIEDMAN D, ERLENMEYER-KIMLING L. 1988. Performance Test, identical pairs version (CPT-IP), I: new findings about sustained attention in normal families. *Psychiatry Res,* 26(2):223-238.

CUMMINGS, E., DONOHOE, G., HARGREAVES, A., MOORE, S., FAHEY, C., DINAN, T. G., MCDONALD, C., O'CALLAGHAN, E., O'NEILL, F. A., WADDINGTON, J. L., MURPHY, K. C., MORRIS, D. W., GILL, M. & CORVIN, A. 2013. Mood congruent psychotic symptoms and specific cognitive deficits in carriers of the novel schizophrenia risk variant at MIR-137. *Neurosci Lett,* 532**,** 33-8.

DOCKER, D., SCHUBACH, M., MENZEL, M., MUNZ, M., SPAICH, C., BISKUP, S. & BARTHOLDI, D. 2014. Further delineation of the SATB2 phenotype. *Eur J Hum Genet,* 22**,** 1034-9.

DONNELLY P, B. I., BLACKWELL JM, BRAMON E, BROWN MA, CASAS JP, CORVIN A, DELOUKAS P, DUNCANSON A, JANKOWSKI J, MARKUS HS, MATHEW CG, PALMER CN, PLOMIN R, RAUTANEN A, SAWCER SJ, TREMBATH RC, VISWANATHAN AC, WOOD NW, SPENCER CC, BAND G, BELLENGUEZ C, DILTHEY A, FREEMAN C, HELLENTHAL G, GIANNOULATOU E, MCVEAN G, MOUTSIANAS L, PIRINEN M, PEARSON R, STRANGE A, SU Z, TASHAKKORI-GHANBARIA A, VUKCEVIC D, DONNELLY P, LANGFORD C, HUNT SE, EDKINS S, GWILLIAM R, BLACKBURN H, BUMPSTEAD SJ, DRONOV S, GILLMAN M, GRAY E, HAMMOND N, JAYAKUMAR A, MCCANN OT, LIDDLE J, POTTER SC, RAVINDRARAJAH R, RICKETTS M, WALLER M, WESTON P, WIDAA S, WHITTAKER P, BARROSO I, DELOUKAS P, RILEY BP, MORRIS DW, O'DUSHLAINE T, MAHER BS, CORMICAN P, KENNY EM, WORMLEY B, DONOHOE G, QUINN E, JUDGE R, COLEMAN K, TROPEA D, ROCHE S, CUMMINGS L, KELLEHER E, MCKEON P, DINAN T, MCDONALD C, MURPHY KC, O'CALLAGHAN E, O'NEILL FA, WADDINGTON JL, WALSH D, STONE J, SCOLNICK E, PURCELL S, SKLAR P, RIPKE S, WALTERS J, OWEN MJ, O'DONOVAN MC, KENDLER KS, CORVIN A, GILL M. 2012. Genome-wide association study implicates HLA-C*01:02 as a risk factor at the major histocompatibility complex locus in schizophrenia. *Biol Psychiatry,* 72**,** 620-8.

DUNN, O. J. 1961. "Multiple Comparisons Among Means". *Journal of the American Statistical Association*

FIRST M, S. R., GIBBON M, WILLIAMS J. 2002. Structured Clinical Interview for DSM-IV-TR Axis I Disorders. *Biometrics Res. New York, NY:New York State Psychiatric Institute*.

FLAUS, A., MARTIN, D. M., BARTON, G. J. & OWEN-HUGHES, T. 2006. Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Res,* 34**,** 2887-905.

FROMER, M., POCKLINGTON, A. J., KAVANAGH, D. H., WILLIAMS, H. J., DWYER, S., GORMLEY, P., GEORGIEVA, L., REES, E., PALTA, P., RUDERFER, D. M., CARRERA, N., HUMPHREYS, I., JOHNSON, J. S., ROUSSOS, P., BARKER, D. D., BANKS, E., MILANOVA, V., GRANT, S. G., HANNON, E., ROSE, S. A., CHAMBERT, K., MAHAJAN, M., SCOLNICK, E. M., MORAN, J. L., KIROV, G., PALOTIE, A., MCCARROLL, S. A., HOLMANS, P., SKLAR, P., OWEN, M. J., PURCELL,

S. M. & O/'DONOVAN, M. C. 2014. De novo mutations in schizophrenia implicate synaptic networks. *Nature,* 506**,** 179-184.

HÄRTING, C., MARKOWITSCH, H. J., NEUFELD, H., CALABRESE, P., DEISINGER, K., & KESSLER, J. 2000. Wechsler Gedächtnis Test-Revidierte Fassung (WMS-R). *Huber, Bern*.

HOWIE, B. N., DONNELLY, P. & MARCHINI, J. 2009. A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. *PLoS Genet,* 5**,** e1000529.

JIANG, Y.-H., YUEN, RYAN K., JIN, X., WANG, M., CHEN, N., WU, X., JU, J., MEI, J., SHI, Y., HE, M., WANG, G., LIANG, J., WANG, Z., CAO, D., CARTER, MELISSA T., CHRYSLER, C., DRMIC, IRENE E., HOWE, JENNIFER L., LAU, L., MARSHALL, CHRISTIAN R., MERICO, D., NALPATHAMKALAM, T., THIRUVAHINDRAPURAM, B., THOMPSON, A., UDDIN, M., WALKER, S., LUO, J., ANAGNOSTOU, E., ZWAIGENBAUM, L., RING, ROBERT H., WANG, J., LAJONCHERE, C., WANG, J., SHIH, A., SZATMARI, P., YANG, H., DAWSON, G., LI, Y. & SCHERER, STEPHEN W. 2013. Detection of Clinically Relevant Genetic Variants in Autism Spectrum Disorder by Whole-Genome Sequencing. *American Journal of Human Genetics,* 93**,** 249-263.

KAHN, R. S. & KEEFE, R. S. 2013. Schizophrenia is a cognitive illness: time for a change in focus. *JAMA Psychiatry,* 70**,** 1107-12.

KIM, B. J. & SCOTT, D. A. 2014. Mouse model reveals the role of RERE in cerebellar foliation and the migration and maturation of Purkinje cells. *PLoS One,* 9**,** e87518.

KIM, J. Y., KIM, K. B., EOM, G. H., CHOE, N., KEE, H. J., SON, H. J., OH, S. T., KIM, D. W., PAK, J. H., BAEK, H. J., KOOK, H., HAHN, Y., KOOK, H., CHAKRAVARTI, D. & SEO, S. B. 2012. KDM3B is the H3K9 demethylase involved in transcriptional activation of lmo2 in leukemia. *Mol Cell Biol,* 32**,** 2917-33.

KIM, Y., XIA, K., TAO, R., GIUSTI-RODRIGUEZ, P., VLADIMIROV, V., VAN DEN OORD, E. & SULLIVAN, P. F. 2014. A meta-analysis of gene expression quantitative trait loci in brain. *Transl Psychiatry,* 4**,** e459.

KLEEFSTRA, T., SCHENCK, A., KRAMER, J. M. & VAN BOKHOVEN, H. 2014. The genetics of cognitive epigenetics. *Neuropharmacology,* 80**,** 83-94.

LALANI, S. R., SAFIULLAH, A. M., FERNBACH, S. D., HARUTYUNYAN, K. G., THALLER, C., PETERSON, L. E., MCPHERSON, J. D., GIBBS, R. A., WHITE, L. D., HEFNER, M., DAVENPORT, S. L., GRAHAM, J. M., BACINO, C. A., GLASS, N. L., TOWBIN, J. A., CRAIGEN, W. J., NEISH, S. R., LIN, A. E. & BELMONT, J. W. 2006. Spectrum of CHD7 mutations in 110 individuals with CHARGE syndrome and genotype-phenotype correlation. *Am J Hum Genet,* 78**,** 303-14.

LIU, J., CHEN, J., EHRLICH, S., WALTON, E., WHITE, T., PERRONE-BIZZOZERO, N., BUSTILLO, J., TURNER, J. A. & CALHOUN, V. D. 2014. Methylation patterns in whole blood correlate with symptoms in schizophrenia patients. *Schizophr Bull,* 40**,** 769-76.

LIU, Z., CHEN, X., ZHOU, S., LIAO, L., JIANG, R. & XU, J. 2015. The Histone H3K9 Demethylase Kdm3b Is Required for Somatic Growth and Female Reproductive Function. *International Journal of Biological Sciences,* 11**,** 494-507.

MARFELLA, C. G. & IMBALZANO, A. N. 2007. The Chd family of chromatin remodelers. *Mutat Res,* 618**,** 30-40.

MCCARTHY, S. E., GILLIS, J., KRAMER, M., LIHM, J., YOON, S., BERSTEIN, Y., MISTRY, M., PAVLIDIS, P., SOLOMON, R., GHIBAN, E., ANTONIOU, E., KELLEHER, E., O'BRIEN, C., DONOHOE, G., GILL, M., MORRIS, D. W., MCCOMBIE, W. R. & CORVIN, A. 2014. De novo mutations in schizophrenia implicate chromatin remodeling and support a genetic overlap with autism and intellectual disability. *Mol Psychiatry,* 19**,** 652-8.

NEGRI, G., MILANI, D., COLAPIETRO, P., FORZANO, F., DELLA MONICA, M., RUSCONI, D., CONSONNI, L., CAFFI, L. G., FINELLI, P., SCARANO, G., MAGNANI, C., SELICORNI, A., SPENA, S., LARIZZA, L. & GERVASINI, C. 2014. Clinical and molecular characterization of Rubinstein-Taybi syndrome patients carrying distinct novel mutations of the EP300 gene. *Clin Genet*.

NUECHTERLEIN, K. H., GREEN, M. F., KERN, R. S., BAADE, L. E., BARCH, D. M., COHEN, J. D., ESSOCK, S., FENTON, W. S., FRESE, F. J., 3RD, GOLD, J. M., GOLDBERG, T., HEATON, R. K., KEEFE, R. S., KRAEMER, H., MESHOLAM-GATELY, R., SEIDMAN, L. J., STOVER, E., WEINBERGER, D. R., YOUNG, A. S., ZALCMAN, S. & MARDER, S. R. 2008. The MATRICS Consensus Cognitive Battery, part 1: test selection, reliability, and validity. *Am J Psychiatry,* 165**,** 203-13.

O'ROAK, B. J., VIVES, L., FU, W., EGERTSON, J. D., STANAWAY, I. B., PHELPS, I. G., CARVILL, G., KUMAR, A., LEE, C., ANKENMAN, K., MUNSON, J., HIATT, J. B., TURNER, E. H., LEVY, R., O'DAY, D. R., KRUMM, N., COE, B. P., MARTIN, B. K., BORENSTEIN, E., NICKERSON, D. A., MEFFORD, H. C., DOHERTY, D., AKEY, J. M., BERNIER, R., EICHLER, E. E. & SHENDURE, J. 2012a. Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science,* 338**,** 1619-22.

O'ROAK, B. J., VIVES, L., GIRIRAJAN, S., KARAKOC, E., KRUMM, N., COE, B. P., LEVY, R., KO, A., LEE, C., SMITH, J. D., TURNER, E. H., STANAWAY, I. B., VERNOT, B., MALIG, M., BAKER, C., REILLY, B., AKEY, J. M., BORENSTEIN, E., RIEDER, M. J., NICKERSON, D. A., BERNIER, R., SHENDURE, J. & EICHLER, E. E. 2012b. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature,* 485**,** 246-50.

O'DUSHLAINE 2015. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nat Neurosci,* 18**,** 199-209.

PURCELL, S., NEALE, B., TODD-BROWN, K., THOMAS, L., FERREIRA, M. A., BENDER, D., MALLER, J., SKLAR, P., DE BAKKER, P. I., DALY, M. J. & SHAM, P. C. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet,* 81**,** 559-75.

QURESHI, I. A. & MEHLER, M. F. 2013. Developing epigenetic diagnostics and therapeutics for brain disorders. *Trends Mol Med,* 19**,** 732-41.

ROBBINS TW, J. M., OWEN AM, SAHAKIAN BJ, MCINNES L, RABBITT P. 1994. Cambridge Neuropsychological Test Automated Battery (CANTAB): a factor analytic study of a large sample of normal elderly volunteers. *Dementia*.

ROBERTSON, I. 1994. Sustained Attention to Response Task (SART). *Dublin, Ireland: Trinity College Dublin*.

SCHIZOPHRENIA WORKING GROUP OF THE PSYCHIATRIC GENOMICS, C. 2014. Biological insights from 108 schizophrenia-associated genetic loci. *Nature,* 511**,** 421-7.

TAKATA, A., XU, B., IONITA-LAZA, I., ROOS, J. L., GOGOS, J. A. & KARAYIORGOU, M. 2014. Loss-of-function variants in schizophrenia risk and SETD1A as a candidate susceptibility gene. *Neuron,* 82**,** 773-80.

VAN BON, B. W., KOOLEN, D. A., BRUETON, L., MCMULLAN, D., LICHTENBELT, K. D., ADES, L. C., PETERS, G., GIBSON, K., MOLONEY, S., NOVARA, F., PRAMPARO, T., DALLA BERNARDINA, B., ZOCCANTE, L., BALOTTIN, U., PIAZZA, F., PECILE, V., GASPARINI, P., GUERCI, V., KETS, M., PFUNDT, R., DE BROUWER, A. P., VELTMAN, J. A., DE LEEUW, N., WILSON, M., ANTONY, J., REITANO, S., LUCIANO, D., FICHERA, M., ROMANO, C., BRUNNER, H. G., ZUFFARDI, O. & DE VRIES, B. B. 2010. The 2q23.1 microdeletion syndrome: clinical and behavioural phenotype. *Eur J Hum Genet,* 18**,** 163-70.

VAN WIJK, S. J. & TIMMERS, H. T. 2010. The family of ubiquitin-conjugating enzymes (E2s): deciding between life and death of proteins. *FASEB J,* 24**,** 981-93.

WADE, P. A., GEGONNE, A., JONES, P. L., BALLESTAR, E., AUBRY, F. & WOLFFE, A. P. 1999. Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. *Nat Genet,* 23**,** 62-6.

WALTERS, J. T., RUJESCU, D., FRANKE, B., GIEGLING, I., VASQUEZ, A. A., HARGREAVES, A., RUSSO, G., MORRIS, D. W., HOOGMAN, M., DA COSTA, A., MOSKVINA, V., FERNANDEZ, G., GILL, M., CORVIN, A., O'DONOVAN, M. C., DONOHOE, G. & OWEN, M. J. 2013. The role of the major histocompatibility complex region in cognition and brain structure: a schizophrenia GWAS follow-up. *Am J Psychiatry,* 170**,** 877-85.

WECHSLER, D. 2001. Wechsler Test of Adult Reading: WTAR: . *Psychological Corporation.*

WESCHLER, D. 1997a. Weschler Adult Intelligence Test, Third Edition (WAIS-III). *San Antonio, TX: Harcourt Assessment*.

WESCHLER, D. 1997b. Weschler Memory Scale, Third Edition (WAIS-III). *San Antonio, TX: Harcourt Assessment*.

WILKINSON, B., GREPO, N., THOMPSON, B. L., KIM, J., WANG, K., EVGRAFOV, O. V., LU, W., KNOWLES, J. A. & CAMPBELL, D. B. 2015. The autism-associated gene chromodomain helicase DNA-binding protein 8 (CHD8) regulates noncoding RNAs and autism-related genes. *Transl Psychiatry,* 5**,** e568.

WOCKNER, L. F., NOBLE, E. P., LAWFORD, B. R., YOUNG, R. M., MORRIS, C. P., WHITEHALL, V. L. J. & VOISEY, J. 2014. Genome-wide DNA methylation analysis of human brain tissue from schizophrenia patients. *Transl Psychiatry,* 4**,** e339.

Tables

Table 1. Post-translational modification categories for epigenetic gene list	
Writers	**Erasers**
DNA Methyltransferases	DNA Demethylases
Histone Methyltransferases	Histone Demethylases
Histone Ubiquityltransferases	Histone Deubiqutinases
Ubiquitin-like histone modifiers	
Histone Acetyltransferases	Histone Deacetylases
Histone Kinases	Histone Phosphatases
Chromatin Remodelers	
Histone Chaperones	

Table 2. Participant demographics and neurocognitive and clinical measures			
Characteristic	**Patients**		**Healthy Participants** (n = 330)
	Psychosis Broad (n = 890)	*Psychosis Narrow* (n =670)	
Psychosis Subtype (n)			**NA**
Schizophrenia	559	559	
Schizoaffective disorder	111	111	
Bipolar Disorder	148	**NA**	
Major Depressive Disorder	34	**NA**	
Psychosis not otherwise specified	38	**NA**	
Male: Female Ratio	1.94 : 1.0	2.4 : 1.0	1.0 : 1.2
Age, mean (S.D.)	42.93 (12.42)	42.28 (12.48)	35.81 (12.66)
Age at onset, mean (S.D.)	23.78 (8.28)	23.23 (7.68)	**NA**
Cognitive Full Scale IQ, mean (S.D.)	92.04 (19.21)	90.29 (18.08)	119.81 (15.57)

Table 3. Regression analysis of neuropsychological variable test scores in all broad psychosis cases and controls					
Gene/Variant (risk allele)	**Cognitive Domain**	**Neuropsychological Variable**	**Genotype (sample N)**	**Mean Score (SD)**	**p-value (r2 change)**
CHD7/rs6984242 (G)	IQ	WTAR	A/A (290)	36.67 (10.38)	0.003 (.011)
			G/A (385)	35.57 (10.25)	
			G/G (134)	33.91 (11.19)	
CHD7/rs6984242 (G)	IQ	Verbal IQ	A/A (334)	99.40 (22.29)	0.001 (.010)
			G/A (441)	98.43 (22.34)	
			G/G (164)	93.65 (21.87)	
CHD7/rs6984242 (G)	IQ	Full Scale IQ	A/A (277)	100.69 (22.38)	0.007 (.008)
			G/A (373)	99.06 (22.17)	
			G/G (126)	96.25 (22.69)	
CHD7/rs6984242 (G)	Episodic Memory	"Faces 2" Task	A/A (232)	36.74 (5.50)	0.007 (.010)
			G/A (312)	36.19 (5.21)	
			G/G (97)	35.44 (5.91)	
CHD7/rs6984242 (G)	Episodic Memory	Logical Memory 1	A/A (343)	32.66 (15.17)	0.029 (.005)
			G/A (454)	32.39 (14.71)	
			G/G (170)	29.74 (14.92)	
CHD7/rs6984242 (G)	Episodic Memory	Logical Memory 2	A/A (341)	19.03 (11.41)	0.023 (.005)
			G/A (452)	18.67 (10.75)	
			G/G (167)	17.06 (10.71)	
CHD7/rs6984242 (G)	Episodic Memory	PAL (stages completed)	A/A (52)	- 0.53 (1.84)	0.034 (.03)
			G/A (76)	-0.72 (2.65)	
			G/G (16)	- 2.22 (3.97)	
EP300/rs9607782 (A)	Episodic Memory	Logical Memory 2	T/T (465)	19.33 (11.11)	0.038 (.004)
			A/T (379)	18.03 (10.93)	
			A/A (64)	17.19 (10.92)	
GATAD2A/rs2905426 (G)	IQ	Performance	T/T (318)	99.00 (22.36)	0.042 (.005)
			G/T (329)	97.45 (22.59)	
			G/G (113)	95.25 (22.299)	
GATAD2A/rs2905426 (G)	IQ	Full Scale	T/T (315)	100.44 (22.38)	0.016 (.007)
			G/T (325)	99.94 (22.19)	
			G/G (113)	95.49 (22.12)	
KDM3B/rs10043984 (T)	Attention	3-digit D'Prime	T/T (208)	1.16 (.98)	0.048 (.008)
			T/C (159)	1.17 (.85)	
			C/C (37)	1.26 (.99)	
RERE/rs34269918 (InsA)	Attention	3-digit D'Prime	G/G (150)	1.29 (0.96)	0.03 (.011)
			GA/G (188)	1.21 (0.93)	
			GA/GA (59)	1.08 (0.88)	

Table 4. Regression analysis of rs6984242 (CHD7) - rs2905426 (GATAD2A) interaction				
Cognitive Domain	**Neuropsychological Variable**	**Genotype Group (N)**	**Mean (SD)**	**P-value (r2 change)**
IQ	WTAR	0 (114)	36.38 (10.15)	0.005 (0.01)
		1 (284)	36.11 (10.91)	
		2 (261)	36.36 (9.89)	
		3/4 (119)	32.48 (10.34)	
IQ	Verbal	0 (133)	99.96 (22.22)	0.002 (0.01)
		1 (316)	99.22 (22.37)	
		2 (314)	97.58 (22.81)	
		3/4 (142)	94.18 (20.83)	
IQ	Performance	0 (108)	101.07 (23.14)	0.016 (0.007)
		1 (276)	97.71 (22.22)	
		2 (258)	97.06 (22.78)	
		3/4 (117)	96.19 (22.22)	
IQ	Full Scale	0 (108)	102.83 (22.81)	0.001 (0.013)
		1 (271)	99.63 (21.93)	
		2 (256)	98.83 (22.34)	
		3/4 (117)	95.56 (22.08)	
Episodic Memory	LM1	0 (138)	32.85 (15.00)	0.049 (0.004)
		1 (323)	32.39 (15.15)	
		2 (319)	32.35 (15.15)	
		3/4 (145)	30.13 (14.20)	
Episodic Memory	LM2	0 (137)	19.32 (11.23)	0.018 (0.005)
		1 (322)	19.07 (11.31)	
		2 (315)	18.45 (10.92)	
		3/4 (145)	17.30 (10.32)	
Attention	CPT-28	0 (64)	2.10 (1.20)	0.009 (0.015)
		1 (141)	2.12 (1.13)	
		2 (150)	1.93 (1.18)	
		3/4 (73)	1.78 (1.09)	

Table 5: UK Replication Data for IQ and Episodic Memory							
Gene	**Variant (risk allele)**	**Cognitive Domain**	**Neuropsychological Variable**	**Genotype**	**Mean (SD)**	**Beta (95% C.I)**	**p-value**
CHD7	rs6984242 (G)	IQ	ZCOMPOSITE Z	AA	-2.30 (1.55)	-0.023 (-0.159 to 0.114)	0.745
				GA	-2.24 (1.58)		
				GG	-2.48 (1.63)		
CHD7	rs6984242 (G)	Episodic Memory	ZSCORE HVLTTOTAL	AA	-2.15 (1.27)	-0.026 (-0.191 to 0.138)	0.752
				GA	-2.16 (1.32)		
				GG	-2.23 (1.53)		

Table 6: German Replication Data for IQ and Episodic Memory					
Gene	**Variant (risk allele)**	**Cognitive Domain**	**Neuropsychological Variable**	**Beta (Standard Error)**	**p-value**
CHD7	rs6984242 (G)	Full Scale IQ	WAIS IQ	-0.38450 (.3474)	0.2685
CHD7	rs6984242 (G)	Verbal IQ	WAIS Verbal	-0.3495 (0.3308)	0.2908
CHD7	rs6984242 (G)	Episodic Memory	WMS_VISUAL_MEMORY	0.0005 (0.0126)	0.9675