Common and Rare Genetic Risk Factors for Schizophrenia and their Associations with Cognition

Leon Hubbard

2014

Thesis submitted for the degree of Doctor of Philosophy at Cardiff University

Supervisors: Dr James Walters and Dr Andrew Pocklington

MRC Centre for Neuropsychiatric Genetics & Genomics

DECLARATION

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed …… (candidate) Date ……

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD

Signed …… (candidate) Date ……

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated.Other sources are acknowledged by explicit references.

Signed …… (candidate) Date ……

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed …… (candidate) Date ……

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loans **after expiry of a bar on access previously approved by the Graduate Development Committee.**

Signed …… (candidate) Date ……

Acknowledgements

The completion of this thesis would not have been possible without the support of a large number of people. I would first like to thank my supervisors Dr James Walters and Dr Andrew Pocklington for their invaluable expertise and guidance. They always made time for me despite their busy schedules and I appreciated their constant support over the years. As James' first PhD student, I am sure this experience will not be forgotten any time soon!

I would also like to acknowledge Dr James Walters and Dr Stan Zammit who co-led on the work outlined in Chapter 2 relating to the polygenic overlap of schizophrenia and cognition in ALSPAC. I worked closely with Dr Katherine Tansey who performed other complimentary analyses as part of the project. Our combined work on this project has been submitted for publication. I would like to thank the families who took part in ALSPAC, and for the permission to use the data. In addition, Dr Dan Rujescu kindly allowed us permission to investigate the polygenic overlap between schizophrenia and cognition, using an adult cognition sample he recruited in Munich. His data were used in analyses presented in Chapters 2 and 3.

I would like to extend thanks to the patients and families who were recruited into Cardiff COGS. In addition, I would like to acknowledge the schizophrenia field team who were responsible for the collection and input of clinical and cognitive data from patients. Cognitive data from Cardiff COGS were used in Chapters 3 and 4.

I would like to acknowledge Dr Andrew Pocklington for his work relating to pathway analyses in Chapters 3 and 4. He grouped gene-sets from the Mouse Genome Informatics database into broader biological categories, and also provided gene-sets derived from proteomic studies in rodents and humans relating to pre and post-synaptic functions.

I would like to acknowledge Professor George Kirov and Dr Elliott Rees for their contribution into the work outlined in Chapter 4. They called CNVs and performed quality control on individuals in Cardiff COGS upon which subsequent analyses were based.

I would also like to thank everyone in the schizophrenia group and BBU for informal discussions relating to the PhD. Dr Davy Kavanagh who previously worked in the BBU helped me enormously regarding coding in R.

,

Finally, the completion of this thesis would not have been possible my parents Lynn and Mick. They have always been there for me during life's ups and downs, and their love and support has enabled me to achieve my ambitions. For that, I will be eternally grateful.

Summary

Individuals with schizophrenia have severe cognitive impairments that impact upon their ability to function within society. Better understanding the genetic mechanisms underlying schizophrenia and cognition provides an opportunity for targeted pharmacological intervention.

This thesis investigates common and rare genetic variation in schizophrenia and their associations with cognitive ability in schizophrenia cases and healthy controls.

Polygenic risk of schizophrenia and bipolar disorder predicted ability on tests of cognitive domains affected in schizophrenia, performance, verbal and full scale IQ in healthy controls. Increased polygenic risk of schizophrenia was robustly associated with lower performance IQ at different training thresholds in two independent cognition samples. There was no consistent association between bipolar polygenic risk and cognition. Common genetic differences between schizophrenia and bipolar disorder were associated with verbal and full scale IQ.

I investigated the hypothesis that 155 gene-sets across six biological categories relating to cognition, brain function and structure were enriched for SNPs influencing general cognitive ability. Schizophrenia polygenic pathway scores for gene-sets were not associated with general cognitive ability in schizophrenia patients, or performance IQ in healthy individuals. Separately, neither gene-sets

nor general categories were enriched for common SNPs showing association with general cognitive ability in schizophrenia cases.

Associations between rare CNVs and general cognitive ability were tested in schizophrenia cases. Cases with a known pathogenic CNV performed approximately one standard deviation below other schizophrenia cases on the MATRICS composite score. In addition, increases in the number of genes hit by large (>100kb) and rare (frequency <1%) CNVs were associated with lower general cognitive ability. However, the number of genes hit in gene-sets previously mentioned was not associated with the MATRICS composite score.

These findings indicate genetic variation in schizophrenia is associated with cognitive ability in schizophrenia cases and healthy controls, providing direction for future research.

Table of Contents

List of tables

List of figures

1 Chapter 1 - Introduction

1.1 Impact of schizophrenia

Schizophrenia is a common psychiatric disorder characterized by severe and heterogeneous symptomatology and deficits in cognitive ability. It has a lifetime risk of approximately 0.5-1% (Jablensky, 2000), affecting individuals across countries and cultures. Patients have a low quality of life characterized by low employment prospects (Rosenheck et al., 2006), low adherence to medication (Valenstein et al., 2004) and increased mortality (Saha et al., 2007). Furthermore, only 1/7 patients have satisfactory improvements regarding both clinical symptoms and functional outcome (Jaaskelainen et al., 2013). Life expectancy is 19 years below the population mean (Laursen, 2011), which may be attributable to increased rates of smoking and their secondary diseases (Lasser et al., 2000), neuroleptic medication (Joukamaa et al., 2006) and increased suicidality (Hor & Taylor, 2010). The societal burden is large, with direct (health/social care) and indirect (reduced productivity, social welfare, judicial) costs estimated at £6.7 billion in England alone in 2007 (Mangalore & Knapp, 2007).

1.2 History of schizophrenia

The first descriptions of schizophrenia were characterised by Emil Kraeplin under the term “dementia praecox”, meaning early dementia. Kraeplin differentiated between the psychotic symptoms of dementia praecox and “manic depressive psychosis”, now classified as bipolar disorder.

The term “schizophrenia”, meaning “splitting of the mind”, was first used by Bluer, who considered schizophrenia to be categorised by a number of smaller related subtypes. In the following decades a distinction was made between psychotic and non-psychotic symptoms of schizophrenia, with Crow defining the modern terms of positive (irrational beliefs/delusions/hallucinations/paranoia) and negative symptoms (alogia, blunted affect, anhedonia) (Crow, 1981).

A third model incorporating disorganised symptoms is now more widely accepted (Arndt et al., 1991). However, other studies have shown between 5-10 symptom dimensions can be identified (Cuesta & Peralta, 2001; McGrath et al., 2004a). This number varies based on the type of analytic methodology used (for example factor analysis, latent class analysis or grade of membership), and the number of symptom variables entered into the model (Jablensky, 2006).

1.3 Schizophrenia diagnosis

The Diagnostic Statistical Manual of Mental Disorders (DSM) V diagnosis of schizophrenia requires individuals to present with delusions, hallucinations or disorganised speech. Furthermore, individuals must present at least two core symptoms of hallucinations, delusions, disorganised speech, catatonic/disorganised behaviour or the presence of negative symptoms. Other diagnostic factors may include impairment through social or occupational dysfunction, signs of persistent disturbance over a minimum of 6 months period.

Furthermore, symptoms should not be attributable to recreational drugs or prescription medication (American Psychiatric Association, 2013).

Nevertheless, the dimensional structure of schizophrenia remains poorly defined (Jablensky, 2006), and from a diagnostic perspective the criteria for diagnosing schizophrenia has remained largely unchanged across the DSM, and their numerous iterations.

1.4 Schizophrenia environmental risk factors

A number of environmental risk factors are associated with increased risk of schizophrenia (van Os et al., 2010). Living in urbanised areas increases the risk of schizophrenia by 2.4 times compared to rural areas (Vassos et al., 2012). Other factors include recreational drug use, for example frequent cannabis users have a near fourfold increase for developing schizophrenia compared to non-users (Manrique-Garcia et al., 2012), and on average develop psychosis approximately 3 years earlier than non-cannabis users (Large et al., 2011). Environmental factors may contribute during gestation and include maternal smoking (Stathopoulou et al., 2013), dietary malnutrition (Brown & Susser, 2008) and maternal infection (Canetta & Brown, 2012).

1.5 Genetics of schizophrenia

Significant progress has been made over the past decade with respect to identifying both common and rare genetic variation contributing to

schizophrenia pathogenesis. This section will describe schizophrenia heritability estimates, linkage studies of schizophrenia, positional candidate genes, common variants identified through genome wide association studies (GWAS), rare variation from copy number variations (CNVs) and rare single nucleotide variations (SNVs).

1.5.1 Familial schizophrenia risk

Schizophrenia risk aggregates in families of affected individuals. Relative to an affected individual, monozygotic MZ twins have the highest lifetime risk at around 50%, dizygotic twins have a lower lifetime risk of 17%, with risk decreasing as the genetic distance of relatives increases (Gottesman, 1991). Increased risk of developing schizophrenia is also found in adopted children whose biological parents have the disorder (Ingraham & Kety, 2000), further supporting a genetic as opposed to environmental aetiology.

1.5.2 Schizophrenia heritability

1.5.2.1 Deriving heritability estimates

An individual's phenotype is the product of their genotype, environment and their respective interactions. The classic twin design is the most popular method for quantifying the variance of a trait attributable to genetic and environmental factors. MZ twins share approximately 100% of their genetic information, whereas DZ twins share on average 50%. For traits under strong genetic influence, their correlation should be higher for MZ twins than for DZ

twins. A basic model of heritability investigates the contribution of additive genetic effect (V_A), common environmental effects (V_C) and unique environmental effects (V_E) upon the phenotypic variation for a trait (V_P) (Visscher et al., 2008):

$V_p = V_A + V_C + V_E$

Heritability refers to the proportion of phenotypic variance (V_p) attributable to genetic factors (V_G). Narrow-sense heritability (h^2) refers to the proportion of phenotypic variance explained by additive genetic factors (V_A): $h^2 = V_A / V_p$. Finally, broad-sense heritability refers to the proportion of genetic variance explained by all genetic factors, which may include interactions between alleles at the same (dominance) or different (epistasis) loci: $H^2 = V_G / V_p$

Other methods are available for calculating narrow-sense heritability in large number of unrelated individuals, for example genome-wide complex trait analysis (GCTA) (Yang et al., 2011). GCTA estimates the variation of a particular trait explained by common SNPs across the genome by comparing the genotypic and phenotypic similarity of each pair of individuals within the sample. Furthermore, GCTA does not rely on twin pairs, permitting the use of large samples that have already been genotyped.

1.5.2.2 Schizophrenia heritability from twin studies & GCTA

Twin study heritability estimates show additive genetic contribution to schizophrenia is approximately 80% (Sullivan et al., 2003). GCTA estimates have shown approximately 25% of schizophrenia liability is explained by common SNPs (Lee et al, 2012). Whilst this is substantially less than twin estimates, this is likely attributable to low linkage disequilibrium (LD) between causal SNPs and those tagged by genotyping platforms.

Whilst heritability estimates can stipulate the genetic contribution to a disease or trait, they cannot identify the location or effect size of genetic variants contributing to them.

1.5.3 Linkage associations with schizophrenia

Linkage analysis uses the principle that one or more markers/genes in the same region of a chromosome have a greater probability of being inherited together. Linkage can therefore be applied to identifying traits or diseases that segregate amongst members of a family. Linkage analysis is a useful method for identifying pathogenic genes in diseases following simple Mendelian rules of inheritance, typically monogenic disorders (Elston, 1998). Early linkage studies of schizophrenia identified several regions harbouring genes of possible pathogenicity (Riley & Kendler, 2006), with a later meta-analysis of 32 schizophrenia linkage studies finding association on chromosome 1, 2q, 3q, 4q, 5q, 8p, and 10q (Ng et al., 2009). However, no single bin reached genome wide-

significance. A region at 8p showed the strongest association in European data ($p=5x10^{-4}$), and across all samples at 5q ($p= 0.004$) (Ng et al., 2009).

1.5.4 Positional candidate gene studies

In response to findings implicating several regions through linkage, further studies tried to identify plausible susceptibility genes for schizophrenia.

Catechol-O-methyltransferase (COMT)

COMT is located within the 22q11.2 locus, and has a crucial role in the regulation of dopamine in the prefrontal cortex (Gogos et al., 1998). *COMT* has two main isoforms: soluble *COMT* (S-COMT) is commonly expressed in non-brain tissues, whereas membrane-bound *COMT* (MB-COMT) is primarily expressed in the brain, particularly in the prefrontal cortex (Lachman et al., 1996).

The Val (G) /Met (A) polymorphism at rs4680 is posited as the functional SNP in *COMT*. Changes in the amino acid from valine to methionine result in increased (val) or decreased (met) levels of both S-COMT and MB-COMT (Williams et al., 2007). Whilst COMT has attractive biological functions potentially supporting the dopamine hypothesis of schizophrenia (Williams et al., 2007), a meta-analysis (Munafo et al., 2005) and other large studies (Williams et al., 2005) have failed to support an association of common variants in *COMT* with schizophrenia.

Disrupted in Schizophrenia 1 (DISC1)

DISC1 is located on the 1q42 locus and was identified through an inherited chromosomal translocation linkage in a Scottish pedigree (Stclair et al., 1990). *DISC1*, and is arguably the most polarising candidate gene for association with schizophrenia (Sullivan, 2013; Porteous et al., 2014). *DISC1* has appealing neurobiological functions (Brandon & Sawa, 2011), specifically it localises at the post-synaptic density in glutamatergic synapses which are strongly implicated in schizophrenia pathogenesis. However, no robust common or rare genetic variants in DISC1 have been identified in large genetic studies over the past decade.

Dysbindin (DTNBP1)

DTNBP1 maps to a region on chromosome 6p, which was identified as a putative schizophrenia region through linkage analysis of 265 Irish pedigrees (Straub et al., 1995). *DTNBP1* forms part of the dystrophin-associated protein complex (Benson et al., 2001) and is strongly expressed in axon bundles and mossy fiber synaptic terminals within the cerebellar and hippocampal structures (Benson et al., 2001). *DTNBP1* is localised on the pre-synapse, and changes in expression result in increased or decreased glutamate synthesis and release via exocytosis (Numakawa et al., 2004). In addition, *DTNBP1* also contributes to changes in synaptic transmutation and plasticity via α-Amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptors in the hippocampus (Orozco et al., 2014).

1.5.4.1 Common variants in schizophrenia candidate genes

Subsequent studies have not identified significant associations with common SNPs and schizophrenia in 14 candidate genes (*RGS4, DISC1, DTNBP1, STX7, TAAR6, PPP3CC, NRG1, DRD2, HTR2A, DAOA, AKT1, CHRNA7, COMT*, and *ARVCF*) (Sanders, 2008). Thus, whilst linkage studies may be useful where a disease is highly penetrant within families, or harbour rare variants, their use as a tool for gene discovery in highly polygenic disorders such as schizophrenia has proven to be limited.

Technological advances now allow for the investigation of millions of single nucleotide variants (SNPs) and structural chromosomal changes including CNVs, single nucleotide variants (SNVs) and indels. These methods have revolutionised the approaches aimed at identifying genetic variants associated with schizophrenia.

1.5.5 Schizophrenia genome-wide association studies

1.5.5.1 Overview of GWAS

Genome wide association studies (GWAS) are a hypothesis free method of testing the common variant, common disease hypothesis (Lander, 1996) by assessing variation across the genome using single nucleotide polymorphisms (SNPs). Each SNP is tested for association with a dichotomous or continuous phenotype. The ability to identify risk variants depends upon several factors including effect size, population frequency and sample size (Bergen & Petryshen, 2012). SNPs may have a direct causal relationship with the phenotype, or be in linkage

disequilibrium with the causal variant. SNP genotyping platforms have hundreds of thousands of individual markers, which can be supplemented by imputation procedures, which ascertain the probability of an ungenotyped marker having a specific allele based upon linkage disequilibrium with a tagged variant.

1.5.5.2 Early schizophrenia GWAS findings

Early schizophrenia GWAS failed to identify common genetic variants reaching genome wide significance (Lencz et al., 2007; O'Donovan et al., 2008a; Sullivan et al., 2008). However, with the inclusion of bipolar cases representing a broader psychosis phenotype, rs1344706, an intronic SNP in *Zinc Finger Protein 804a* (*ZNF804A*) reached this threshold (O'Donovan et al., 2008a); the first such observation in a psychiatric disorder.

To identify whether rs1344706 was the true functional variant in *ZNF804A*, this region was extensively mapped using de-novo polymorphism and high-density LD mapping (Williams et al., 2011a). The strongest association remained at rs1344706, providing evidence this was the susceptibility variant in the region. Furthermore, a meta-analysis of multiple schizophrenia GWAS showed rs1344706 reached genome wide significance, becoming stronger with the inclusion of bipolar cases (Williams et al., 2011a).

Functionally, the risk T allele at rs1344706 reduces *ZNF804A* expression during neurodevelopment in the second trimester (Hill & Bray, 2012). A *ZNF804A* knockdown study showed altered gene expression for *A2M*, *C2ORF80*, *CRYAB*,

FRZB and *STMN3* (Hill et al., 2012). The strongest evidence was observed in *C2ORF80* and *STMN3*. Whilst *C2ORF80* remains uncharacterised, *STMN3* contributes to branching of dendrites and axons (Poulain & Sobel, 2007).

1.5.5.3 International Schizophrenia Consortium GWAS (2009)

In 2009, using combined data from three schizophrenia consortia consisting of 12945 cases and 43591 controls (Stefansson et al., 2009), seven genome wide significant loci emerged from this analysis. Five variants were within the major histocompatibility complex (MHC) region; one was an intronic SNP in *TCF4*, while the other was proximal to *NRGN* (Stefansson et al., 2009).

The MHC is on chromosome 6, and is strongly associated with immune functioning and autoimmunity (Fernando et al., 2008). However, identifying the true functional variants are challenging due to high linkage disequilibrium (LD). A degree of scepticism has previously surrounded the role of immune response in schizophrenia (DeLisi, 1996), in part attributable to inadequate study design in human studies and unclear mechanisms of action (Strous & Shoenfeld, 2006). However, evidence is accumulating that shows MHC class 1 proteins also have non-immune related functions, including synaptic pruning and roles in synaptic plasticity (Oliveira et al., 2004; Lee et al., 2014).

TCF4 is part of the helix-loop-helix protein group, which has largely been studied in respect of its role in the development of B and T lymphocytes (Murre, 2005). A fine mapping study of *TCF4* found no non-synonymous variants, or cis-regulated

variants altering mRNA expression (Williams et al., 2011b). However, knockdown of *TCF4* in neuroblastoma cells has substantial effects upon genes in signalling pathways that regulate cell differentiation and survival (Forrest et al., 2013). Furthermore, *TCF4* also alters the expression of other mental retardation genes including *UBE3A* and *ZEB2*, which are associated with Angelman Sybdrome and Mowat-Wilson syndrome respectively (Forrest et al., 2013).

TCF4 does not appear to influence brain structure. A large structural MRI study in 1300 healthy individuals found no evidence of association between common SNPs in *TCF4* and total volume of the brain, or grey or white matter volumes (Cousijn et al., In Press).

1.5.5.4 Psychiatric Genetics Consortium (2010-Present)

By combining samples across groups, the increase in power to detect variants increases substantially. This principle was important for the formation of the Psychiatric Genetics Consortium (PGC) (Psychiatric GWAS Consortium Steering Committee, 2009; Sullivan, 2010). In 2011, a GWAS found SNPs in *TCF4*, *MIR137, TRIM26*, *CSMD1*, *CNNM2*, *NT5C2*, and proximal to *PGCEM1*, *MMP16*, *STT3A* and *CCDC68* reached genome wide significance for association with schizophrenia (Schizophrenia PGC, 2011). Furthermore, variants in *CACNA1C*, *ANK3* and *ITIH3-ITIH4* surpassed genome wide significance when individuals with schizophrenia and bipolar disorder were combined, providing further evidence for their genetic relatedness.

The PGC recently reported the largest GWAS of any psychiatric disorder to date, using a discovery sample of 34241 schizophrenia cases and 45604 controls (Schizophrenia Working Group of the Psychiatric Genetics Consortium, 2014). They found 128 loci reaching genome wide significance representing 108 independent genomic regions. SNPs reaching genome-wide significance were found in genes whose biological functions were previously implicated in schizophrenia aetiology including; calcium channel (*CACNA1C*, *CACNB2* and *CACNA1I*), dopamine receptor (*DRD2*) and glutamatergic (*GRM3*, *GRIN2A*, *SRR*, *GRIA1*) genes.

Calcium channel genes are implicated in several psychiatric disorders including schizophrenia, bipolar disorder, autism, major depressive disorder and attention deficit hyperactivity disorder (Smoller et al., 2013). Genes encoding for voltage-gated calcium channels have important roles regarding the activation of intracellular signalling pathways, altering both neuronal development and functionality, as well as affecting gene expression (Dolmetsch, 2003). A de-novo missense mutation G406R in *CACNA1C* is the cause of Timothy Syndrome, a disorder characterised by an autistic phenotype with cognitive abnormalities, as well as other cardiac and immune irregularities (Splawski et al., 2004). Further work is required to identify their role in schizophrenia.

Glutamate is the most abundant neurotransmitter acting on excitatory neurons across the brain (Erecinska & Silver, 1990). NMDA receptor-mediated glutamate transmission within the post-synaptic density is a prime candidate for dysfunction in schizophrenia (Harrison & Weinberger, 2005) Furthermore,

glutamate binds to NMDARs, which are responsible for changes in synaptic firing by altering long term potentiation that may cause structural changes at the synapse. These processes are considered to underpin synaptic plasticity and cognitive ability (Hunt & Castillo, 2012).

Immune processes outside of the MHC were also implicated. In particular, genes contributing to the lineage of B-lymphocytes were enriched for schizophrenia associations (Schizophrenia Working Group of the Psychiatric Genetics Consortium, 2014). These findings, along with consistent associations within the MHC provide support that genes regulating the immune response may contribute to schizophrenia pathogenesis.

Despite the successes of GWAS, they do not yet explain the majority of heritability identified through twin studies of schizophrenia (and other diseases/traits). Often labelled as "missing heritability", a number of theories have suggested possible explanations including the role of SNP (Zuk et al., 2012) and gene-environment interactions (Kaprio, 2012), although other studies have suggested this is attributable to incomplete linkage disequilibrium between SNPs tagged on genotyping platforms and their true causal counterparts (Yang et al., 2010a).

1.5.6 Polygenic risk scoring

When testing for association with a phenotype, the majority of SNPs will not meet the threshold for genome wide significance. Nonetheless, these variants

may contribute substantially to the genetic architecture of the phenotype in question. Polygenic risk scoring (PRS) is a method designed to identify whether en masse, common genetic variants associated with a particular phenotype in a discovery sample can predict the same, or different phenotype in an independent target sample. SNPs are selected based upon the strength of their association with the discovery phenotype using pre-specified p-value thresholds. Polygenic scores are calculated for each individual in the target sample by summing the number of susceptibility alleles of the reference SNP weighted by the log of the SNP odds ratios (or beta coefficient is the discovery phenotype is on a continuous scale).

PRS analyses have shown common variants contribute substantially to schizophrenia. The first study to apply PRS used a training sample of 3322 schizophrenia cases and 3587 controls (International Schizophrenia Consortium, 2009). Using the Molecular Genetics of Schizophrenia - European American (MGS-EA) independent schizophrenia case/control sample, schizophrenia polygenic risk predicted approximately 3% of the variance for schizophrenia liability and achieved high levels of statistical significance ($p=2x10^{-28}$).

A larger study used a training set comprised of 6458 cases and 8971 controls (Schizophrenia PGC, 2011). A target sample of 2936 cases and 3492 controls showed that the variance of schizophrenia liability explained by schizophrenia polygenic risk, increased to 6%.

This demonstrates that the power of PRS analyses is more dependent on the size of the initial discovery sample, rather than the target set (Chatterjee et al., 2013; Dudbridge, 2013). The explanation for this lays in the inaccuracy of the estimates of SNP effect sizes in the discovery sample, where error decreases as sample size increases.

The largest PRS analysis for schizophrenia was performed in the recent publication by the PGC (Schizophrenia Working Group of the Psychiatric Genetics Consortium, 2014). Using a discovery sample comprising of 32838 cases and 44357 controls, it predicted 18.4% of the variance of schizophrenia liability in MGS-EA. The variance explained is substantially larger than previous studies, supporting previously simulated projections (International Schizophrenia Consortium, 2009).

PRS have also been used to investigate the common genetic architecture across neuropsychiatric disorders. Schizophrenia polygenic risk explains between 2-2.5% of variance of the liability of bipolar disorder (International Schizophrenia Consortium, 2009; Smoller et al., 2013), 0.75% of the liability for major depressive disorder and approximately 0.1% for liability of autistic spectrum disorders (Smoller et al., 2013).

Common genetic risk of schizophrenia is not associated with non-psychiatric disorders including coronary heart disease, rheumatoid arthritis and types 1 & 2 diabetes mellitus (International Schizophrenia Consortium, 2009). These findings show psychiatric disorders have a degree of common genetic overlap

that is not shared with unrelated disorders, with implications for classification, disease models and treatment (Smoller et al., 2013).

1.5.7 Copy Number Variations (CNVs)

CNVs are a source of structural variation in the genome, whereby a chromosomal section can be deleted or duplicated. Although CNVs constitute normal genetic variation within healthy populations (Sebat et al., 2004), enrichment of large (>100kb) and rare (population frequency < 1%) CNVs at multiple loci are observed in schizophrenia probands (Rees et al., 2014b) and other neuropsychiatric disorders including autism and intellectual disability (Malhotra & Sebat, 2012).

A landmark paper by the ISC reported associations of CNV burden in 3391 schizophrenia cases and 3181 controls (International Schizophrenia Consortium, 2008). They found a 1.15 fold increase in rare (<1%) and large (>100kb) CNVs in schizophrenia cases, with particular enrichment of genic regions. Furthermore, the number of genes hit by CNVs was 1.41 times higher in schizophrenia cases compared to controls. CNVs at 1q21.1, 15q13.2 and 22q11.2 were also enriched in schizophrenia cases and conferred substantial increases in schizophrenia risk (odds ratios between 6.6-21.6).

The latest meta-analysis of 15 loci with prior association to schizophrenia implicated eleven regions (1q21.1, *NRXN1*, 3q29, 15q.11.2, 15q13.3 and 22q11.2, and duplications at 1q21.2, WBS, Angelman/Prader-Willie Syndrome on

chromosome 15, 16p13.11 and 16p11.2) with odds ratios typically ranging between 2 and 50 (Rees et al., 2014b).

Deletions at 22q11.2 show the strongest associations with schizophrenia, in addition to other physical, psychiatric and neurocognitive phenotypes collectively known as 22q11.2 Deletion Syndrome (22qDS) (Bassett & Chow, 2008). 22q11.2 deletions typically span over 3 megabases covering 43 genes (International Schizophrenia Consortium, 2008). Approximately 25% of individuals with 22qDS will develop schizophrenia (Murphy et al., 1999), making it the largest genetic risk factor, second only to monozygotic twins (Gottesman, 1991). Furthermore, the 22qDS schizophrenia phenotype is not clinically different compared to schizophrenia cases without the deletion (Bassett et al., 2003).

Interestingly, the reciprocal 22q11.2 duplication is associated with a protective effect against schizophrenia (Rees et al., 2014a) . They showed the rate of duplications at 22q11.2 in healthy controls (0.085%) was significantly higher compared to schizophrenia probands (0.014%). Conversely, carriers of the 22q11.2 deletion are at greatest risk of developing schizophrenia, suggesting dosage effects for genes at this locus may pertain to both protective and damaging effects (Rees et al., 2014a).

Neuropsychiatric CNVs typically span large genomic regions affecting multiple genes, therefore identifying single or multiple pathogenic genes is challenging. There are a number of exceptions where CNVs affect only one gene, for example,

deletions hitting exons in *NRXN1*, which are strongly associated with schizophrenia (Kirov et al., 2009b; Rujescu et al., 2009) and autism (Yan et al., 2008).

Functionally, neurexins and their binding partners neuroligins, have important roles in the organisation of GABAergic and glutamatergic synapses (Graf et al., 2004). Knockdown models of *NRXN1* using human induced pluripotent and embryonic stem cells demonstrate its role in biological networks needed for neurodevelopment (Zeng et al., 2013). Specifically, lowering the expression of *NRXN1* caused significant changes in genes associated with cell adhesion and neuron differentiation (Zeng et al., 2013).

Another exception is *VIPR2*. Duplication rates between 0.25-0.35% were observed in schizophrenia cases compared to 0.03-0.05% in controls (Levinson et al., 2011a; Vacic et al., 2011). However, both studies used the MGS and ISC datasets as secondary cohorts, thus their findings are not independent. Doubt over the validity of these findings came from a recent meta-analysis showing duplication rates in cases (0.015%) were lower than controls (0.095%), and after meta-analysis was no longer significant (Rees et al., 2014b).

De-novo CNVs are observed in offspring, but not in either parent. Stefansson et al (2008) were the first to show an enrichment of deleterious de-novo CNVs at 1q21.2, 15q11.2 and 15q13.3 in schizophrenia cases. More generally, de-novo CNVs are more frequent in schizophrenia compared to healthy populations (Xu et al., 2008), and are a possible explanation for the persistence of schizophrenia

despite cases' reduced fecundity (Rees et al., 2011). Furthermore, pathway analyses of de-novo CNVs also show enrichment for genes in the post-synaptic density, and more specifically, neuronal activity-regulated cytoskeleton-associated protein (ARC) and N-methyl-D-aspartate receptor (NMDAR) complexes in schizophrenia cases (Kirov et al., 2012).

1.5.8 Exome sequencing

Exome sequencing is a method that identifies variation in exons, which are protein-coding regions in the genome. For each codon, SNVs and indels can be identified. SNVs are called as either synonymous (no amino acid change) or non-synonymous (amino acid change) variants. Non-synonymous SNVs and indels can be further categorised according to "damaging" or "non-damaging" by their predicted effect on the resulting protein structure.

To date no single rare exonic variants have been robustly associated with schizophrenia (Need et al., 2012). In addition, the largest exonic de-novo mutation study to date found no evidence of enrichment of de-novo mutations in schizophrenia cases (Fromer et al., 2014).

Recently, rare disruptive variants within 2546 genes previously implicated with schizophrenia are found at a higher rate in cases compared to controls (Purcell et al., 2014). Furthermore, whilst significance for individual variants was modest, significant enrichment was observed for ARC, NMDAR and calcium channel pathways. The enrichment of ARC pathways shows consistency with other

findings from schizophrenia de-novo CNVs (Kirov et al., 2012) and SNVs (Fromer et al., 2014), whilst calcium channel SNPs show enrichment from the largest schizophrenia GWAS (Schizophrenia Working Group of the Psychiatric Genetics Consortium, 2014) and across other neuropsychiatric disorders (Smoller et al., 2013).

Some studies have indicated de-novo SNVs are enriched for genes that are expressed during gestation (Xu et al., 2012; Gulsuner et al., 2013). These genes are expressed in the developing dorsolateral and ventrolateral prefrontal cortex (Gulsuner et al., 2013). Furthermore, they regulate migration of neurons, synaptic transmission, and neuronal signaling/transportation (Gulsuner et al., 2013). However, other studies have found no evidence for enrichment of prenatally expressed genes (Fromer et al., 2014). Thus the contribution of de-novo SNVs upon the developmental aetiology of schizophrenia remains unclear.

1.5.9 Summary

The hypothesis that schizophrenia is highly polygenic disease was first posited nearly 50 years ago (Gottesman & Shields, 1967). This polygenic model is now substantiated through evidence showing schizophrenia risk is conferred through common, rare and structural genetic variation. This polygenic model has also shown that psychiatric disorders including schizophrenia, bipolar disorder, autism and attention deficit hyperactivity disorder share a degree of common (Smoller et al., 2013) and rare (Malhotra & Sebat, 2012) genetic risk across psychotic and autistic disorders (Craddock & Owen, 2010). With improved understanding of the complex genetic and molecular basis of schizophrenia, the

opportunity presents to develop novel therapeutic targets that look beyond dopaminergic antagonists that have underpinned schizophrenia treatment for the past 60 years (Lieberman et al., 2005).

1.6 Genetic architecture of cognition

1.6.1 Defining cognitive ability

Intelligence is a broad term encapsulating mental processes responsible for many aspects of cognitive functioning including abstract thinking, problem solving, comprehending and applying complex ideas, learning and using acquired knowledge for future problems (Gottfredson, 1997). However, definitions and conceptualisations of cognition vary widely, and have included measures of emotional or social intelligence (Mayer et al., 2001) and creativity (Barron & Harrington, 1981).

A number of theories regarding the structure of cognition have been posited, however perhaps the most widely accepted are those that stress the centrality of generalised cognitive ability ("g"). Originally conceptualised by Galton (Galton, 1883), and mathematically conceived by Spearman, this theory proposes that a single general cognitive factor is a substantial mediator of performance across lower order cognitive ability (Spearman, 1904). "g" is the name given to the first unrotated principle component derived across tests measuring different cognitive abilities, and consistently explains approximately 40% of inter-test variability (Deary et al., 2010). Other measurements of general cognitive ability

include Wechsler Adult Intelligence Scale (WAIS) IQ or Raven's Progressive Matrices (Sheppard & Vernon, 2008).

1.6.1.1 Fluid & Crystallised Intelligence

General cognitive ability can be subdivided into fluid and crystallised intelligence (Cattell, 1963). Fluid ability is a general term representing the ability to solve novel problems by utilising attention, working memory, processing speed and executive functioning (Blair, 2006). Conversely, crystallised intelligence represents the ability to use acquired skills relating to basic numeracy and general knowledge (Cattell, 1963). Both fluid and crystallised constructs are conceptually similar to Wechsler performance and verbal IQ respectively (Kline, 1991; Deary, 1993).

1.6.2 Heritability of cognitive ability

1.6.2.1 Heritability of general cognitive ability across the lifespan

General cognitive ability is under strong genetic influence. Heritability estimates derived from twin studies have identified intriguing patterns with respect to genetic effects and their influence upon general cognitive ability over time. A large study of 11000 twin pairs found heritability estimates for "g" increases linearly across childhood (41%), early adolescence (55%) and early adulthood (66%) (Haworth et al., 2010). Heritability estimates derived from GCTA are similar to that of twin studies. Childhood "g" is approximately 26% (Trzaskowski

et al., 2013), increasing to between 24-45% in early adolescence (Deary et al., 2012; Plomin et al., 2013; Trzaskowski et al., 2013).

Heritability estimates continue to increase across the lifespan to approximately 80% in old age (Pedersen et al., 1992). This counterintuitive finding shows general cognitive ability is under stronger genetic influence in later life compared to childhood. One argument has suggested genetic effects on cognition during childhood results in environmental choices throughout adulthood that compliments their cognitive genetic propensity (Haworth et al., 2010).

1.6.2.2 GCTA Heritability of fluid & crystallised intelligence

In adults, the heritability estimates for individual cognitive tests measuring fluid and crystallised abilities are generally lower compared to that of general cognitive ability. However, GCTA estimates show that both fluid (51%) and crystallised (40%) are still highly heritable (Davies et al., 2011).

1.6.2.3 Heritability of WAIS performance and verbal tests

Using twin studies, heritability of individual cognitive tests are substantially lower than general cognitive ability, and have been calculated separately for some performance and verbal subtests of the WAIS. Heritability estimates for performance IQ vary between 43-83% (Posthuma et al., 2001; Wright et al., 2001; Posthuma et al., 2003; Luciano et al., 2005). However, estimates for performance subtests are generally lower. Block design varies between 24-31% (Finkel et al., 1995; Rijsdijk et al., 2002), whereas digit symbol coding heritability

estimates are typically higher at 65% (Plomin et al., 1994; Posthuma et al., 2003). Other heritability estimates for digit symbol coding are lower (24-41%) and show substantial differences across countries (Finkel et al., 1995).

WAIS verbal IQ scores show higher heritability estimates between 55-85% (Posthuma et al., 2001; Wright et al., 2001; Rijsdijk et al., 2002; Posthuma et al., 2003; Luciano et al., 2005). WAIS verbal tests have shown moderate heritability: similarities (40%), vocabulary (52%), information (44%) and arithmetic (53%) (Rijsdijk et al., 2002).

1.6.2.4 Heritability of childhood IQ and socioeconomic status

Some studies have questioned these findings, using socioeconomic status as a mediating factor for assessing the genetic contribution of general cognitive ability (Nisbett et al., 2012). Twin heritability estimates of general cognitive ability in children from affluent families are approximately 60%, whereas almost no genetic contribution to variance in cognition was reported in children from poor backgrounds (Turkheimer et al., 2003). These findings are supported by other studies reporting a large effect of shared environment in lower socioeconomic families (Hanscombe et al., 2012), Other results suggest common genetic variation contributes to the correlation between socioeconomic status and childhood IQ (Trzaskowski et al., 2014), and the literature remains inconclusive.

1.6.3 Linkage studies and cognition

The first genome-wide linkage study for general cognitive ability was performed in 329 Australian and 100 Dutch families (Posthuma et al., 2005). 2q24.1-31.1 showed linkage with performance IQ, and 6p25.3-22.3 showed suggestive linkage with full scale IQ. Other studies have reported suggestive linkage at 2q24-31 with performance IQ and Cambridge reading tests (Luciano et al., 2006). Suggestive linkage at chromosome 6p to full scale IQ in individuals with alcohol dependence has also been identified (Dick et al., 2006); however, the functional variants contributing to cognition at these loci are unknown.

1.6.4 Genome wide association studies of cognition

1.6.4.1 Episodic & Working Memory

In 2006, a GWAS of episodic memory in 333 individuals identified a significant association at rs17070145 in *KIBRA*, and was replicated in two independent samples (Papassotiropoulos et al., 2006). Furthermore, this association was stronger for immediate word recall ($p=4x10^{-6}$) compared to delayed recall ($p=0.002$). Since its publication, a number of studies have supported this association (Almeida et al., 2008; Schaper et al., 2008; Hayashi et al., 2010; Vassos et al., 2010), whilst others have not (Need et al., 2008).

A meta-analysis tested the association between rs17070145 and episodic and working memory in 8909 and 4696 individuals respectively (Milnik et al., 2012). rs17070145 was significantly associated with episodic ($p=0.001$) and working

memory (p=0.018), explaining approximately 0.5% and 0.1% of their variances respectively, although effect sizes were low (d=0.14 and 0.07).

KIBRA binds to the protein PICK1 and forms complexes with AMPA receptors (Makuch et al., 2011). Functionally, *KIBRA* knockout mice have altered long-term potentiation and depression in the hippocampus, and perform poorly on contextual fear memory paradigms (Makuch et al., 2011).

1.6.4.2 Processing Speed

A small number of GWAS for processing speed are reported in the literature. One study performed a GWAS on seven measures of processing speed and a generalised speed factor in a combined sample of 2958 individuals (Luciano et al., 2011). No variant achieved genome wide significance for association with any phenotype. These findings are concordant with a smaller GWAS for digit symbol coding in 1086 individuals, which failed to identify SNPs reaching genome-wide significance (Cirulli et al., 2010).

1.6.4.3 Fluid/Crystalised Intelligence

A GWAS of fluid and crystalised intelligence was performed in 3511 healthy individuals (Davies et al., 2011). No SNP reached genome-wide significance. Using a gene-based test, formin-binding protein 1-like (*FNBP1L*) reached genome wide significance for association with fluid intelligence ($P = 9.2x10^{-7}$), although this failed to replicate at a nominally significant level in an independent sample.

1.6.4.4 General cognitive ability

A GWAS of "g" in 7900 individuals did not identify any SNP reaching genome wide-significance (Davis et al., 2010). A handful of studies have performed GWAS for educational attainment, measuring school test results (Martin et al., 2011) and number of years spent in education (Rietveld et al., 2013). No SNPs reached genome wide significance for association with school test results (Martin et al., 2011). However Rietveld and colleagues identified two SNPs reaching genome wide significance for association with a binary variable (individuals who completed college or not) at rs11584700 in *MDM4* and rs4851266 proximal to *BC105019*. An intergenic SNP rs9320913 also reached genome wide significance for association with number of years in education; however, the number of identified loci was modest, and effect sizes were small.

Other GWAS investigated change in general cognitive ability across the lifespan. Several studies have implicated *APOE* in normal cognitive aging (De Jager et al., 2012) (Davies et al., 2014). Furthermore, the e4 isoform of *APOE* is the largest genetic risk factor for Alzheimer's disease (Corder et al., 1993; Harold et al., 2009) (Harold et al., 2009), which is traditionally characterised by severe cognitive decline in elderly populations. Healthy carriers of the e4 isoform have lower generalised cognitive ability, executive functioning, episodic memory and processing speed compared to non-carriers (Wisdom et al., 2011).

1.6.5 CNVs and Cognition

The contribution of rare CNVs towards general cognitive ability in healthy populations appears limited, and will be discussed at greater length in Chapter 4. Briefly, studies have investigated various measures of rare CNV burden including: total CNV number, total CNV length and number of genes hit with respect to general cognitive ability (MacLeod et al., 2012; McRae et al., 2013; Kirkpatrick et al., 2014). Kirkpatrick *et al* (2014) found no association between total CNV number or length and IQ in over 6000 individuals, replicating findings from a smaller study of 800 controls (McRae et al., 2013). No association was observed between rare CNV burden and fluid/crystallised intelligence in over 3000 individuals (MacLeod et al., 2012).

1.6.6 Summary

Despite the substantial genetic contribution to cognition identified through heritability studies, the identification of genetic variants associated with either general or specific cognitive measures in the general population remain elusive. No robust associations between SNPs and cognition have emerged, either in candidate genes, or through GWAS. Furthermore, rare CNVs appear to have little influence on general cognitive ability in healthy individuals. Substantially larger sample sizes are required to identify genetic variants influencing cognitive ability.

1.7 Cognitive Phenotype in Schizophrenia

1.7.1 Overview of cognitive impairment in schizophrenia

Deficits in cognitive functioning are present across psychotic disorders (Green, 2006), however schizophrenia cases show the greatest impairment. An influential review showed schizophrenia cases performed poorly on tasks measuring processing speed, executive functioning, working memory, verbal learning, and general cognitive ability (Heinrichs & Zakzanis, 1998), and are typically 1-2 standard deviations below healthy controls (Reichenberg, 2005). Nonetheless, there is substantial heterogeneity across patients, with some individuals functioning at a level indistinguishable from the general population (MacCabe et al., 2012).

Representing the structure of cognitive deficits in schizophrenia cases has proven difficult. The abundance of different cognitive batteries of varying length, type and methods of administration makes interpretation of results across studies challenging (Keefe et al., 2004). Furthermore, the length of some cognitive batteries is arduous and impractical for patients with schizophrenia, and considered a hindrance when applied during clinical trials (Harvey & Keefe, 2001).

1.7.2 Development of the Measurement to Improve Cognition In Schizophrenia Initiative

The Measurement to Improve Cognition In Schizophrenia (MATRICS) cognitive battery was designed to standardise cognitive testing in schizophrenia cases for use in research and clinical trial settings (Green et al., 2004). Their aims were to identify specific cognitive domains showing the greatest impairment in schizophrenia. Furthermore, they aimed to promote the use of pharmacological approaches by improving cognition in patients, leading to improvements in functional outcome (Marder, 2006).

The MATRICS committee initially performed a comprehensive review of principle component/factor analytic studies of cognition in schizophrenia to identify recurring domains (Nuechterlein et al., 2004), with additional input from a committee of experts (Kern et al., 2004). They identified 7 domains: attention/vigilance, processing speed, verbal learning, visual learning, reasoning/problem solving, social cognition and working memory as being core cognitive features of schizophrenia (Nuechterlein et al., 2004).

Tests for each domain were required to fulfil several criteria including: good test-retest reliability with low practice effects, show associations with functional outcome, be tolerated by patients, have low inter-test correlations, and be suitable for repeated testing for research purposes and clinical trials (Green et al., 2004). The final version of the MATRICS comprised of 10 individual tests

measuring 7 cognitive domains, including a standardised composite score measuring general cognitive ability (Table 1.1).

MATRICS Domain	Test
Attention/Vigilance	Continuous Performance Test: Identical Pairs (CPT-IP)
Reasoning/Problem Solving	Neuropsychological Assessment Battery (NAB) Mazes
Speed of Processing	Brief Assessment of Cognition in Schizophrenia (BACS): Symbol-Coding Category Fluency: Animal Naming Trail Making Test (TMT): Part A
Verbal Learning/Memory	Hopkins Verbal Learning Test—Revised (HVLT-R)
Visual Learning/Memory	Brief Visuospatial Memory Test—Revised (BVMT-R)
Social Cognition	Mayer-Salovey-Caruso Emotional Intellgence Test (MSCEIT): Managing Emotions
Working Memory	Wechsler Memory Scale: 3rd Ed. (WMS-III): Spatial Span Letter-Number Span

Table 1-1 - MATRICS domains and their respective cognitive tests

The majority of MATRICS domains and the composite score are also correlated with IQ (see Table 1.2), although social cognition is not (Mohn et al., 2014).

MATRICS Domain	Correlation with IQ
Attention/Vigilance	0.26 ($p<0.01$)
Reasoning/Problem Solving	0.37 ($p<0.01$)
Speed of Processing	0.39 ($p<0.01$)
Verbal Learning/Memory	0.43 ($p<0.01$)
Visual Learning/Memory	0.54 ($p<0.01$)
Social Cognition	0.09 ($p>0.05$)
Working Memory	0.51 ($p<0.01$)
Composite	0.60 ($p<0.01$)

Table 1-2 - Correlation with MATRICS domains and IQ (adapted from Mohn et al, 2014)

1.7.3 MATRICS Domains

1.7.3.1 Attention/Vigilance

Attention is the ability to selectively focus upon a stimulus, often within the context of suppressing information from distractive stimuli (vigilance), and are commonly measured using various versions of the CPT. Patients show large attentional deficits (Heinrichs & Zakzanis, 1998; Liu et al., 2002) that may result from a reduction in capacity to store information (Nuechterlein & Dawson, 1984) or poor encoding of information (Elvevag et al., 2000).

1.7.3.2 Reasoning/Problem Solving (Executive Functioning)

Schizophrenia cases perform poorly on the Tower of London (Morice & Delahunty, 1996) and Wisconsin Card Sorting Test (WCST) (Heinrichs & Zakzanis, 1998), which utilise higher-level cognitive abilities such as advanced planning, abstract thinking and other lower order cognitive abilities (Nuechterlein et al., 2004).

1.7.3.3 Speed of Processing

Processing speed represents the speed at which different cognitive functions can be performed. These tasks are typically measured using digit symbol coding, trail making test part A, or through verbal fluency tasks. Digit symbol coding measures the speed at which individuals can correctly draw a symbol under the corresponding number from a given key. The TMT-A measures the speed at which lines are drawn between sequentially increasing numbers. Verbal fluency

measures the number of words an individual can recall in a set time period in response to an overarching category such as naming animals.

Schizophrenia cases perform particularly poorly on processing speed tasks when tested in tandem with other cognitive domains (Dickinson et al., 2007). A meta-analysis of individual tests across cognitive domains measuring processing speed, episodic memory, executive functioning, working memory, attention, problem solving, motor speed and general cognitive ability showed that the largest deficit was present for the digit symbol coding task with an effect size of -1.57 (Dickinson et al., 2007).

Another meta-analysis showed digit symbol coding had an effect size of -1.40, second only to verbal episodic measures of story learning and memory (Dickinson et al., 2013a).

Processing speed contributes to performance of other cognitive domains including working memory, verbal IQ, sustained attention, problem solving, verbal learning and memory, visual learning and memory, reaction time, executive functioning and motor dexterity in schizophrenia cases (Rodriguez-Sanchez et al., 2007; Andersen et al., 2013). When processing speed was added as a covariate, differences between schizophrenia cases and controls for most cognitive domains (except general cognitive ability and verbal episodic memory) were not significant. Furthermore, a large study of around 7000 healthy adolescents between the ages of 13-17 showed that processing speed explained approximately 90% of the effect of "g" (Coyle et al., 2011). Other findings have

shown that processing speed is both mediated by, and independent of "g" across schizophrenia cases and controls (Dickinson et al., 2008).

A meta-analysis in 2010 showed several variables moderate performance on digit symbol coding tasks (Knowles et al., 2010). First, more recent studies were likely to report smaller effect sizes. Second, substantial differences in IQ between patient and control populations were more likely to inflate effect sizes. Third, lower doses of the antipsychotic chlorpromazine were associated with smaller effect sizes.

1.7.3.4 Social Cognition

Social cognition refers to the ability to comprehend interpersonal interaction and to respond to social cues in an appropriate manner. Unlike more traditional cognitive domains, studies of social cognition in schizophrenia have become more prominent over the past decade (Green & Horan, 2010).

In individuals who later develop schizophrenia, abnormal social interaction is observed during early childhood and into adolescence (Jones et al., 1994). Furthermore, children who subsequently develop schizophrenia show abnormal facial expressions in social situations, indicative of aberrant social behaviour (Walker et al., 1993; Schiffman et al., 2004).

In adults, general cognitive ability contributes substantially to social cognition, which in turn mediates functional outcome (Schmidt et al., 2011). In particular,

tests measuring emotional perception and social knowledge explained a significant proportion of the relationship between cognition and functional outcome. In addition, a recent meta-analysis showed emotion perception was significantly impaired in schizophrenia cases, with a large effect size of 0.91 (Kohler et al., 2010), which may result from abnormalities in the anterior and posterior cingulate cortex (Reske et al., 2009).

1.7.3.5 Working Memory

Factor analytic studies of cognition in schizophrenia commonly identify a working memory factor and comprised of tests including WAIS arithmetic, letter number sequencing and digit span and mental control (Nuechterlein et al., 2004). Schizophrenia cases have severe deficits in working memory (Lee & Park, 2005; Forbes et al., 2009) encompassing several stages of mental processing (Baddeley, 1981), that are linked to aberrant connectivity in the prefrontal cortex (Barch et al., 2002).

Patients experience difficulties in almost all aspects of working memory, from initial encoding of information into short term memory (Hartman et al., 2003), to storage (Gold et al., 2010) and subsequent recalling of information (Aleman et al., 1999).

1.7.3.6 Verbal Memory/Learning

A recent meta-analysis by Dickinson et al (2013) showed the largest deficits in schizophrenia cases were for story memory and story learning tasks, both

measures of verbal episodic memory. However, it is unclear what cognitive battery they were part of. Nonetheless, these findings are concordant with previous a previous meta-analysis where the largest deficits were present for global verbal memory (Heinrichs & Zakzanis, 1998).

The cognitive mechanisms underlying impairments in verbal learning suggest patients have initial problems encoding information (Cirillo & Seidman, 2003), which may be partly attributable to aberrant attention and processing speed (Brebion et al., 2000), rather than problems in retaining or retrieving information (Gold et al., 2000).

1.7.3.7 Visual Memory/Learning

Schizophrenia cases show reductions in performance on spatial memory tasks (Fleming et al., 1997; Park et al., 1999). Worse performance may be attributable to problems with the organisational processing and retention of visual information (Seidman et al., 2003), possibly caused by widespread and inefficient processing of information across both hemispheres in frontal, parietal and temporal brain regions (Lee et al., 2008).

1.7.4 The generalised cognitive deficit in schizophrenia

Schizophrenia cases show cognitive impairment across multiple cognitive domains. However, these domains are not necessarily independent (Dickinson & Harvey, 2009), and show substantial inter-correlation (Keefe et al., 2006). Furthermore, individual domains load strongly upon a single generalised factor.

63% of schizophrenia case/control variance associated with cognitive performance was attributable to "g" (Dickinson et al., 2008), however some additional case/control variance was also mediated through verbal memory (13.8%) and processing speed domains (9.1%). The proportion of cognitive variance explained by "g" is also more substantial in schizophrenia cases compared to healthy controls (Dickinson et al., 2008; Deary et al., 2010; Dickinson et al., 2011). This suggests general cognitive ability in schizophrenia is more unified compared to healthy individuals. Furthermore, given "g" is substantial in schizophrenia cases, identifying neurobiological mechanisms underlying the generalised deficit may be more informative in contrast to individual cognitive domains.

Other ways of assessing the generalised deficit include the MATRICS composite score, or IQ. The MATRICS cognitive battery derives a composite score reflective of the generalised deficit using scores across each of the 7 individual domains. Furthermore, composite scores are strongly correlated with IQ ($r \sim 0.7$) (Kern et al., 2008; August et al., 2012). Other studies have found the MATRICS composite score is approximately 1 standard deviation lower compared to their projected composite score based upon premorbid IQ estimates (Gray et al., 2013), suggesting differences between current "g" and premorbid IQ .

Performance and verbal IQ scores are also affected in schizophrenia cases. Whilst there are not substantial differences between their premorbid scores (Woodberry et al., 2008), longitudinal studies have shown performance IQ has the largest decline between childhood and adulthood compared to verbal and

full scale IQ (Meier et al., 2014). This is in concordance with several meta-analyses showing performance IQ is the most impaired measure of general cognitive ability in schizophrenia cases (Heinrichs & Zakzanis, 1998; Dickinson et al., 2013a).

1.7.5 The developmental course of cognitive deficits in schizophrenia

Individuals with sub-clinical psychotic experiences during childhood are more likely to develop psychotic disorders in adulthood (Poulton et al., 2000). A meta-analysis of 1188 individuals with high risk of psychosis (HR) showed significant impairment on measures of general intelligence, attention, executive functioning, processing speed (digit symbol coding), social cognition, verbal fluency, verbal and visual memory (Fusar-Poli et al., 2012).

Niarchou et al (2013) investigated the association between self-reported psychotic experiences and cognitive ability based on the MATRICS domains in children at ages 8 and 12. Children with lower processing speeds at age 8 showed association with increased risk of psychosis- like symptoms at age 12 (Niarchou et al., 2013). Although not all individuals with HR will develop a psychotic disorder, these findings nonetheless suggest early psychotic experiences are accompanied by lower cognitive ability, which are reliable precursors to schizophrenia.

Schizophrenia cases exhibit signs of milder cognitive impairment preceding the onset of psychosis. Infants and children who take longer to develop motor

coordination and language capabilities are at greater risk of developing schizophrenia (Jones et al., 1994; Isohanni et al., 2001). In addition, children with low IQ measured as early as age 4 have increased risk of developing schizophrenia in later life (David et al., 1997; Cannon et al., 2000; Zammit et al., 2004). Furthermore, childhood general cognitive ability is associated with psychotic experiences in the adult general population (Barnett et al., 2012).

1.7.6 Longitudinal studies of cognitive ability in schizophrenia during childhood/adolescence

Several longitudinal studies have been conducted on independent birth cohorts providing insights into cognitive ability across childhood/adolescence in individuals who later develop schizophrenia. The New Zealand Dunedin cohort reported children aged 7-13 who later developed schizophrenia performed between 0.25-0.45 standard deviations lower than controls for IQ, performance IQ, block design, picture completion, verbal IQ, information, similarities, vocabulary and arithmetic tasks (Meier et al., 2014).

MacCabe and colleagues investigated the change in verbal, spatial and inductive performance in Swedish adolescents between the ages 13 and 18 (MacCabe et al., 2013). Individuals who developed schizophrenia showed marked deterioration on verbal tests, with similar deterioration observed for individuals with affective and non-affective psychosis. Spatial ability was also affected in schizophrenia cases, although this decline was less pronounced.

A separate study showed that prospective schizophrenia cases in the USA experienced a decline in reading, verbal, executive functioning and general cognitive ability between the ages of 13-16 (Fuller, 2002).

A large meta-analysis compared premorbid IQ in individuals who later developed schizophrenia against age-matched controls (Woodberry et al, 2008). Individuals who developed schizophrenia had significantly lower IQ. Furthermore, impairments in both verbal and performance IQ were present, although no evidence of decline was observed across childhood and adolescence.

Cognitive impairment is a key prodromal feature of schizophrenia with evidence suggestive of neurodevelopmental aetiology. These findings show individuals who develop schizophrenia in adulthood have cognitive impairments from early stages of development, which persists through childhood and adolescence. However, it is unclear if these deficits represent a causal influence on schizophrenia development, or whether neurobiological mechanisms contributing to schizophrenia also influence cognitive ability.

1.7.7 Cognitive ability and schizophrenia symptom dimensions

Understanding the relationship between symptom dimensions and cognitive performance may provide an opportunity for treatment intervention with reciprocal benefits for patients.

Dominguez and colleagues performed the largest review of cognition and symptom dimensions in over 5000 individuals (Dominguez et al., 2009). They looked at the association between 9 cognitive domains: attention, executive control, reasoning/problem solving, processing speed, verbal fluency, verbal learning/memory, verbal working memory, visual learning/memory, and full scale IQ), and 4 symptom clusters: depressive, disorganised, positive and negative.

Verbal fluency, verbal learning and full scale IQ were significantly associated with negative symptoms, although effect sizes were small (-0.21 to -0.29)). Disorganised symptoms significantly correlated with attention, visual learning and full scale IQ, again with small effect sizes -0.21 to -0.28. Positive symptoms showed a weak association with processing speed, however no significant correlations were observed with the other cognitive domains. Depressive symptoms showed no association with any cognitive domain.

The low correlation between positive symptoms and cognitive deficits could be attributable to selection bias, whereby patients with severe psychotic episodes are less likely to be included for selection for studies (Keefe & Harvey, 2012). Alternatively, this may simply reflect that severity of positive symptoms is not related to cognitive ability.

1.7.8 Treating Cognitive Symptoms In Schizophrenia

1.7.8.1 Pharmacological Targets

Cognitive deficits are not significantly treated by current medications or other treatment approaches. A large meta-analysis investigating the effects of anti-psychotic medication upon cognitive functioning was performed in 1513 schizophrenia cases across 14 studies (Woodward et al., 2005). They found atypical anti-psychotics (clozapine, olanzapine, quetiapine & risperidone) improved general cognitive ability, processing speed, verbal learning, fluency and motor skills. However, these improvements were approximately 1/3 of one standard deviation, suggesting for the majority of patients this will not remediate their cognitive functioning to within a normal range.

Several potential confounders are present for studies investigating anti-psychotic treatment and cognitive improvements, including practice and placebo effects (Goldberg et al., 2010). Where studies administer the same, or similar cognitive test over a specific time period, individuals undergoing testing have greater familiarity with the task, and may show improvement due to practice effects. Thus, better cognitive performance may be attributed to anti-psychotic medication, however practice, rather than medication effects may contribute to this improvement.

Anti-psychotic medication may modestly, but not substantially improve cognitive functioning in schizophrenia cases. Thus, the development of

pharmacological targets that improves cognitive functioning would be of substantial benefit to patients (Gold, 2004; Buchanan et al., 2005).

A number of potential molecular compounds targeting cholinergic, muscarinic, glutamatergic, dopaminergic, serotonergic, adrenergic and GABAergic receptors may benefit cognitive functioning in schizophrenia (Gray & Roth, 2007).

NMDA receptors are a promising target for treatment due to strong evidence from genetic studies in schizophrenia (Kirov et al., 2012; Fromer et al., 2014) and their association with cognition through synaptic plasticity (Hunt & Castillo, 2012).

Buchanan et al (2007) reported the results of a 16-week double blind trial in 157 schizophrenia cases investigating the clinical efficacy of glycine, a co-transmitter with glutamate for NMDARs, and D-cycloserine, which is a partial NMDAR agonist. Changes in negative symptoms and performance on neurocognitive domains (speed of processing, verbal fluency, motor speed, verbal memory, visual memory, auditory memory, visuo-spatial ability, attention and executive functioning, composite) were recorded at baseline and at the end of the trial. Unequivocally, there was no improvement across any neurocognitive measure or in the reduction of negative symptoms (Buchanan et al., 2007). This null finding was replicated in a second study of 195 patients (Weiser et al., 2012), however they noted that this could partially be explained by a higher than expected placebo response in the control group.

Other studies have investigated the effects of nicotinergic agents. High rates of smoking are found in schizophrenia populations relative to the general population, suggesting patients use nicotine as a form of self- medication (Kumari & Postma, 2005). Whilst some studies have shown nicotinic agonists may improve attention (Freedman et al., 2008; Quisenaerts et al., 2014), working memory and negative symptom severity (Freedman et al., 2008), possibly through the enhancement of glutamatergic synapses in the dorsolateral prefrontal cortex (Yang et al., 2013), their efficacy for targeted intervention is currently unknown.

In general, pharmacological trials have produced disappointing results regarding their impact on cognitive functioning. A review of 118 studies documenting change in cognitive functioning resulting from pharmacological intervention found many clinical trials used designs that are not optimal for detecting improvements in cognition (Keefe et al., 2013). Specifically, studies have small sample sizes and thus low power, whilst demographics are commonly limited to older males with chronic schizophrenia. These individuals may respond more poorly to treatment compared to recent onset patients who are generally younger.

1.7.8.2 Cognitive Remediation Behavioural Therapy (CRBT)

Non-pharmacological treatments have also been used to improve cognitive ability in schizophrenia cases, the most popular being CRBT. A recent meta-analysis reported a moderate effect size (d=0.45) for cognitive remediation

improving global cognitive ability in patients, although effect sizes for individual cognitive domains are typically lower (Wykes et al., 2011). However, it is unclear whether CRBT improves cognitive functioning sufficiently to produce meaningful long-term benefits relating to functional outcome.

1.8 Cognition as a schizophrenia endophenotype

The heterogeneous phenotype of schizophrenia, combined with inconclusive genetic results through linkage lead to a resurgence of the endophenotype approach (Gottesman, 2003). Endophenotypes are unobservable traits associated with disease under stronger genetic rather than environmental influence, and considered to be less complex and heterogeneous than the overarching disease. Gottesman and Shields proposed several criteria for defining an endophenotype (Gottesman & Shields, 1973) :

1) They show association with disease.
2) They are heritable.
3) They are independent of disease state and show consistency across time.
4) They are closely aligned with disease in families.
5) Within families, endophenotypes show greater association with unaffected family members compared to the general population.

Identifying genes underlying endophenotypic traits may provide an easier method for revealing genes or biological pathways in disease populations.

Furthermore, genes associated with endophenotypes are ideal candidates for modelling their biological function within animals (Gould & Gottesman, 2006).

1.8.1 Heritability of neurocognitive traits in schizophrenia families

The Consortium on the Genetics of Schizophrenia (COGS) is an on-going family study with the purpose of identifying reliable endophenotypes in schizophrenia. The study uses families where at least one proband is affected, and information on both parents and at least one unaffected sibling is required. Cognition is measured using a shortened version of the computerised neurocognitive battery (CNB) and other tests measuring attention, working memory, verbal memory, spatial memory/processing and emotional recognition (Calkins et al., 2007; Gur et al., 2007).

Heritability estimates were between 0.24 and 0.55 for all cognitive tasks, demonstrating a consistent low to moderate effect (Greenwood et al., 2007). These findings are in concordance with estimates from other schizophrenia family studies (Glahn et al., 2007).

1.8.2 State independence

State independence refers to the stability of an endophenotypic trait over time, regardless of other disease processes. Longitudinal studies measuring change in cognitive ability over time have showed performance is stable across time for both specific and generalised cognitive ability, and not influenced by changes in

symptom severity (Heaton, 2001) for at least a decade after initial hospitalisation (Hoff et al., 2005).

1.8.3 Aligned with disease in families

For a trait to be considered an endophenotype it must be observed in unaffected relatives of the proband, demonstrating probable genetic influence. A number of studies have investigated the cognitive profiles of biologically related individuals with affected schizophrenia probands relative to the population.

Unaffected siblings have lower performance on general cognitive ability (Wisner et al., 2011), processing speed (Wisner et al., 2011) and working memory (Conklin et al., 2005; Wisner et al., 2011). One meta-analysis used results from 58 studies encompassing: attention/working memory, verbal memory, visual memory, executive functioning, spatial ability, motor function, language function and general cognitive ability in affected first degree relatives (Snitz et al., 2006). They identified moderate effect sizes ($d > 0.5$) for tasks measuring attention and executive functioning, with lower effect sizes ($d > 0.4$) for verbal and spatial tests. These findings suggest unaffected family members also exhibit cognitive impairments relative to population controls, but not to the extent of affected schizophrenia probands.

1.8.4 Other endophenotype approaches

An alternative approach is to test the association between genetic variants associated with disease and endophenotypic candidates in healthy individuals

(Walters & Owen, 2007). Walters & Owen (2007) described five viewpoints on the relationship between genotype, phenotype and endophenotypic candidates:

1) Genetic factors are causal to endophenotypes on the disease pathway, which in turn influences phenotypic presentation of a symptom.
2) The endophenotype is the resulting factor of the disease symptomatology, or arise through treatment intervention.
3) Genetic factors demonstrate pleiotropy for both a symptom and the endophenotype
4) Non-overlapping genetic factors contribute independently to symptoms and the endophenotype.
5) The endophenotype influencing symptoms is the product of environmental factors.

In these scenarios, only the first and third scenarios may be useful for investigating either disease mechanisms or shared genetic aetiology (Walters & Owen, 2007).

In the absence of robust genetic variants associated with cognition, the approach of testing the association between genetic variants associated with schizophrenia and cognitive ability is currently more widely used.

1.9 Genetic Overlap of Schizophrenia and Cognition

Quantifying the genetic effects underlying schizophrenia and cognitive ability is important for assessing their suitability as endophenotypes. Several studies have used genetic modelling of twin pairs to answer this question.

Using a mixture of 267 monozygotic and dizygotic schizophrenia and healthy twin pairs, a strong phenotypic correlation between schizophrenia and IQ was observed (r=-0.61), with 92% of this correlation attributable to shared genetic influences (Toulopoulou et al., 2007).

A larger study using 657 schizophrenia cases, 674 first-degree relatives and 725 controls found approximately 89% of the correlation between schizophrenia and general cognitive ability was attributable to genetic factors (Toulopoulou et al., 2010). In addition, they also showed that genetic factors contributing to the phenotypic correlation between schizophrenia, and immediate and delayed recall were 72% and 86% respectively (Toulopoulou et al., 2010).

A separate study using 1986 MZ and 2253 DZ Swedish twins showed greatly reduced phenotypic correlation between psychosis and premorbid IQ (r= -0.11) (Fowler T, 2012), although the proportion of this correlation attributable to genetic factors was 91%, similar to the findings by Toulopoulou and colleagues. However, this study provided additional details showing genetic factors associated with psychosis are largely independent of IQ; only 6.8% was shared.

The authors conclude that using IQ as an endophenotype is likely to be uninformative in finding genes for schizophrenia.

1.10 Genetic Risk of Schizophrenia and Its Associations with Cognition

A number of studies have investigated the genetic basis of cognitive impairment in schizophrenia at the level of single SNPs, overall polygenic risk, and to a lesser extent CNVs.

1.10.1 Candidate schizophrenia SNPs and their association with cognition

This section will briefly describe studies that have investigated SNPs in schizophrenia candidate genes, and those discovered through schizophrenia GWAS with respect to their associations with cognition.

BDNF

BDNF was a schizophrenia functional candidate gene (Muglia et al., 2003), and resides within the neurotrophin gene family. These genes regulate neuronal differentiation and proliferation and may contribute to synaptic plasticity in the hippocampus (An et al., 2008).

Over 20 studies have been performed investigating the Val66Met genotype at rs6265 in BDNF and cognitive functioning. A large meta-analysis in 2012 investigated this association in 7095 cases and controls collectively (Mandelman

& Grigorenko, 2012). They found no association between the Val66Met genotype and general cognitive ability, executive functioning, verbal fluency, visual ability, or memory. Although this study pooled cases and controls, thus removing possible associations with cognition that are case specific, the findings nonetheless suggest rs6265 has no role upon cognitive functioning across healthy and schizophrenia populations.

COMT

The Val158/108Met polymorphism in COMT is a regulator of dopamine in the prefrontal cortex. A large meta-analysis of 46 studies investigated associations between the val/met genotype and 6 cognitive phenotypes; n-back task, trail making task, verbal recall, verbal fluency, Wisconsin card sorting and IQ (Barnett et al., 2008). The val/met genotype had a low effect size (d=0.06) and explained approximately 0.1% of the variance of IQ. No significant associations were observed between the val/met genotype and other cognitive phenotypes.

APOE

The *APOE* locus is the most strongly associated risk factor for developing Alzheimer's Disease (AD) (Harold et al., 2009), with the e4 allele conveying the largest disease risk (Corder et al., 1993). AD is largely characterised by severe cognitive decline in later life. Furthermore, *APOE* also contributes to a cognitive decline in later life within healthy individuals (Davies et al., 2014). The role of *APOE* in cognitive decline makes this a plausible candidate gene for the aetiology of cognitive deficits in schizophrenia.

No association was observed between *APOE-e4* and schizophrenia through meta-analysis (Xu et al., 2006). Furthermore, schizophrenia cases carrying the risk *APOE-e4* genotype carriers do not have lower IQ scores compared to non=carriers (Thibaut et al., 1998). This suggests *APOE-e4* does not contribute towards cognitive impairment in schizophrenia. However, its effect upon longitudinal changes in cognition in schizophrenia is unknown.

1.10.1.1 Schizophrenia GWAS associated SNPs their association with cognition

Nitric oxide synthase 1 (NOS1)

rs6490121 in *NOS1* was one of the earliest SNPs to show association at a level just below genome wide significance with schizophrenia (O'Donovan et al., 2008a). In hippocampal neurons, nitric oxide is released from the post-synaptic membrane and acts upon the presynaptic terminal, altering long term potentiation (Arancio et al., 1996). Furthermore, NOS1 couples with the N-methyl-D-aspartate receptors (NMDAR) in the post-synaptic density (PSD), which is strongly enriched for *de novo* CNVs in schizophrenia (Kirov et al., 2012) and linked to synaptic plasticity (Malenka & Nicoll, 1993) . In mice, knock out effects of nitric oxide synthase reduces memory and learning capacity using performance on the Morris water maze (Weitzdoerfer et al., 2004). One study investigated the effect of rs6490121 genotypes on general and specific cognitive ability in both schizophrenia cases and healthy controls (Donohoe et al., 2009). They showed homozygous carriers of the risk "G" allele at rs6490121 had lower verbal IQ and working memory, but not attention impairments.

TCF4

Several studies have investigated the association between rs9960767 genotype and cognition. One study of 173 schizophrenia cases showed carriers of the "C" risk allele had lower performance on the MATRICS reasoning/problem solving domain, with no other significant associations observed for the remaining cognitive domains (excluding social cognition) and general cognitive ability (Albanna et al., 2014). Other studies have found the "C" allele is associated with less impaired verbal memory in 401 schizophrenia cases, but not delayed recall or IQ (Lennertz et al., 2011).

ZNF804A

rs1344706, in the gene *ZNF804A* , is a SNP reaching genome wide significance for schizophrenia (Schizophrenia Working Group of the Psychiatric Genetics Consortium, 2014) and broader psychosis (O'Donovan et al., 2008a). Several studies have investigated the association between rs1344706 genotype and cognition.

Walters and colleagues looked at the association between the "A" risk allele at rs1344706 and IQ, verbal working memory, episodic memory and attention (Walters et al., 2010). Counterintuitively, proband carriers of the "A" allele were less cognitively impaired on verbal and spatial working memory tasks. Post hoc tests showed the homozygous AA group was significantly less impaired compared to the CC group. However, these results are not suggestive that risk carriers display preserved cognitive functioning. Interestingly, these results were not observed within healthy controls.

A separate study has shown differential patterns of working memory and executive functioning in schizophrenia cases with high (>100) and low (<100) IQ when grouping by rs1344706 genotype (Chen et al., 2012). Patients in the high IQ group with the homozygous AA genotype had significantly lower performance on both working memory and executive functioning tasks, whereas those in the low IQ group had more spared cognitive ability.

The counterintuitive finding that homozygous AA cases have less impaired cognitive functioning in working memory tasks may be partially explained by relatively preserved grey matter in the hippocampus (Donohoe et al., 2011a).

rs1344706 also shows association with differential activation in the right dorsolateral prefrontal cortex using a facial working memory paradigm (Linden et al., 2013). Separately Esslinger and colleagues measured working memory and emotional recognition with respect to the rs1344706 genotype using fMRI (Esslinger et al., 2011). They observed no phenotypic differences grouping by genotype, however risk allele carriers showed decreased connectivity across the prefrontal cortex and hippocampal regions.

NRGN

Several neurobiological functions relating to *NRGN* and cognition have been identified. Specifically, neurogranin is strongly enriched in hippocampal pyramidal neurons (Huang et al., 2007), and particularly concentrated in dendritic spines (Zhabotinsky et al., 2006). Furthermore, mouse models have

been used to show *NRGN* is important for hippocampal synaptic plasticity and subsequent performance on spatial learning tasks (Pak et al., 2000). However, its role in cognition in humans seems to be limited (Donohoe et al., 2011b; Walters et al., 2013).

MHC

Whilst SNPs in the MHC shows substantial association with schizophrenia, its association with cognition is not well defined. A recent study found rs6904071 was most robustly associated with delayed episodic memory, and the schizophrenia risk allele was associated with poorer cognitive performance (Walters et al., 2013). Furthermore, the risk allele was also associated with lower hippocampal volume, which itself has links to episodic memory (Chadwick et al., 2010). It is unknown whether this association is attributable to the role immune system or via other pathways. Evidence is now immerging showing MHC class I proteins may influence synaptic functioning (Lee et al., 2014), possibly through localisation with postsynaptic density protein 95 in hippocampal dendrites (Goddard et al., 2007).

CACNA1C

A number of studies have investigated the association between rs1006737 in *CACNA1C* and cognitive ability in schizophrenia patents and healthy controls. One study reported homozygous and heterozygous schizophrenia cases carrying the "A" risk allele were significantly impaired on a logical memory task (Hori et al., 2012). One study in over 500 healthy males found no association between

rs1006737 genotype and general cognitive ability, executive functioning, verbal learning and memory (Roussos et al., 2011).

1.10.2 Summary of candidate/GWAS SNPs and their association with cognition

Studies investigating schizophrenia or candidate SNPs with respect to effects of genotype on cognition have typically had little success. Although associations may be identified in small samples, effect sizes are typically small, not always in the expected direction of effect, or have inconsistent effects across studies. Furthermore, for well-studied variants such as *COMT* and *BDNF*, findings from large meta-analyses have failed to show strong associations between the risk genotype and cognition, despite initial successes of smaller studies. For SNPs recently identified through GWAS, large sample sizes are required to show meaningful associations to prevent false positive associations found in earlier smaller studies.

1.11 GWAS of Cognition in Schizophrenia

There have been a limited number of GWAS of cognition in schizophrenia reported in the literature, which is most likely due to issues with small sample sizes. However, Dickinson and colleagues recently performed a GWAS for association with "g" (Dickinson et al., 2014). Using a discovery cohort of 339 schizophrenia probands and 363 controls, they identified two SNPs (rs10174400 and rs10182570) reaching genome wide significance within *SCN2A*, with the

former SNP explaining over 10% of the variance of "g". rs10174400 was nominally significant in unaffected siblings (p=0.03) explaining 3.4% of the variance of "g". In healthy controls, this association was not significant, and the direction of effect was reversed. Furthermore, this effect was substantially lower in their replication sample, where rs10174400 was less significant (p=0.02) and explained approximately 1% of the variance of "g", suggesting a "winners curse" effect (Kraft, 2008). Larger studies are required to replicate this finding.

1.12 Polygenic risk of schizophrenia and its association with cognition

Recently, common polygenic risk of schizophrenia has been applied to test associations with potential cognitive endophenotypes in the general population. Increased polygenic risk of schizophrenia was associated with lower performance on general cognitive ability (Lencz et al., 2014) and a decline in general cognitive ability across the lifespan (McIntosh et al., 2013). Whilst these findings support the hypothesis that increased genetic risk of schizophrenia is associated with poorer cognition, it explains only modest proportions (%$r^2 \approx 1$) of general cognitive ability, and have not consistently replicated (van Scheltinga et al., 2013). Conversely, a reverse approach showed polygenic scores for "g" was associated with schizophrenia status, albeit with a very limited degree of variance explained (%$r^2 < 0.5$) (Lencz et al., 2014).

1.13 SNVs and their association with cognition in schizophrenia

The contribution of SNVs and their association with cognitive ability in schizophrenia is largely unknown. A single study reported higher rates of rare de-novo loss of function mutations in individuals with low academic achievement (Fromer et al., 2014). Individuals with the lowest grades (D & C) had approximately double the number of loss of function mutations compared to those achieving higher grades (A & B).

1.14 Neuropsychiatric CNVs and their association with cognition

Neuropsychiatric CNVs are structural genetic variants implicated in schizophrenia, autism and intellectual disability (Stefansson et al., 2014). These psychiatric disorders share a common theme of cognitive impairment, thus their risk CNVs may contribute to cognitive ability. A recent study by Stefansson and colleagues showed healthy carriers of neuropsychiatric CNVs had cognitive scores between that of schizophrenia cases and other healthy controls, which was largely driven by general cognitive ability. This study was the first to show neuropsychiatric CNVs have a detrimental impact upon cognition in the absence of a psychiatric diagnosis.

1.15 Summary

Identifying genetic variants contributing to both schizophrenia and cognitive ability may substantially improve our understanding of their underlying biology, with the prospect of potential therapeutic targets that could be of benefit to

patients. However they are both highly polygenic traits with a complex genetic architecture, and dissecting their shared biology has been arduous. To date, no genetic variant has been robustly associated schizophrenia and cognition. Whilst the MATRICS cognitive battery has led to improvements regarding the assessment and standardisation of cognitive ability, pharmacological targets are still elusive for the treatment of cognitive symptoms. At least part of this failure can be attributed to a poor understanding of genetic factors that underpin both cognitive ability and schizophrenia.

1.16 Thesis aims and Hypotheses

This thesis investigated the effect of common and rare schizophrenia genetic variation, and their associations with specific and general cognitive ability in schizophrenia cases and healthy controls.

Chapter 2 investigated the common polygenic risk of schizophrenia and bipolar disorder and their association with specific and generalised cognitive ability within the general population. Previous studies have identified schizophrenia polygenic risk is associated with lower general cognitive ability in healthy populations. This chapter expanded upon these findings by using three GWAS discovery datasets for schizophrenia, bipolar disorder and schizophrenia versus bipolar. Secondly, polygenic risk scores were calculated in two independent cognition datasets using tests measuring individual cognitive domains and general cognitive ability. This chapter addresses the following hypotheses:

1) Increased polygenic risk of schizophrenia would be associated with lower cognitive ability.
2) Increased polygenic risk of bipolar disorder would be associated with lower cognitive ability.
3) Polygenic risk scores derived from SNPs associated specifically with schizophrenia (and not bipolar disorder) would be associated with lower cognitive ability.

Chapter 3 investigated the hypothesis that gene sets related to brain function, development, behaviour and cognition are *a priori* more likely to be enriched for SNPs influencing general cognitive ability. 155 gene-sets were used and grouped into six biological categories relating to behaviour, cellular physiology, cellular morphology, development, region tract morphology and subcellular neuronal. Expanding on findings from chapter 2, it is unclear whether schizophrenia polygenic risk is associated with general cognitive ability using polymorphisms within specific biological pathways. In addition, it is unknown what pathways are important for the generalised cognitive deficit in schizophrenia cases. This chapter examines the following hypotheses:

1. Do schizophrenia polygenic risk scores derived from common SNPs in candidate pathways predict general cognitive ability measured through performance IQ in healthy controls, and the MATRICS composite score a schizophrenia patient sample? In addition, are there differences in association between the different pathway categories?

2. Using Brown's method, is there an enrichment of SNPs in candidate pathways that show association with general cognitive ability in a schizophrenia patient sample? In addition, are there differences in association between the different pathway categories?

Chapter 4 examined the association between rare CNVs and general cognitive ability in schizophrenia cases. Healthy carriers of well-supported neuropsychiatric CNVs have lower cognitive ability compared to non-carriers, however no study to our knowledge has investigated this within a schizophrenia cohort. In addition, the impact of CNV burden in schizophrenia cases is poorly defined, along with their impact upon biological pathways relating to brain function, development, behaviour and cognition. This chapter addresses the following hypotheses:

1. Do schizophrenia carriers with well-supported neuropsychiatric CNVs have lower general cognitive ability compared to non-carriers?

2. Does total rare (frequency <1%) CNV burden of small (15-100kb) and large (>100kb) CNVs, their total length or number of genes hit show association with general cognitive ability? Furthermore, does the type of CNV (deletions or duplications) show the same, or differential associations?

3. Does the number of genes hit by CNVs within 155 candidate pathways show association with general cognitive ability in schizophrenia cases?

2 Polygenic Risk of Schizophrenia and Bipolar Disorder and their Associations with Cognition

2.1 Summary

Schizophrenia and bipolar disorder are severe psychiatric conditions with a substantial genetic aetiology. This chapter used established polygenic risk score methodology to assess whether large numbers of common genetic variants associated with schizophrenia, bipolar disorder and their differences are associated with cognition in healthy adults and children at age 8.

Increased polygenic risk of schizophrenia was consistently associated with lower performance IQ using three different schizophrenia discovery datasets, across multiple training thresholds in two independent cognition samples. Schizophrenia polygenic risk predicted up to 0.59% of the performance IQ variance in 936 adult controls, and 0.34% in over 5500 controls in ALSPAC.

Increased polygenic risk of bipolar disorder was associated with lower processing speed at the 0.1 training threshold (%r^2=0.1), and better social cognition at the 0.0001 training threshold (%r^2=0.08).

Polygenic risk scores of schizophrenia relative to bipolar disorder were associated with lower performance on full scale IQ (%r^2=0.3), verbal IQ (%r^2= 0.3) and performance IQ (r^2=0.1%).

2.2 Introduction

Schizophrenia and bipolar disorder are common psychiatric disorders with lifetime prevalence estimated to be around 1% (McGrath et al., 2004b; Merikangas et al., 2007). Schizophrenia is characterised by positive (irrational beliefs/ delusions/hallucinations/paranoia) and negative symptoms (alogia, blunted affect, anhedonia) (van Os & Kapur, 2009). Bipolar disorder is predominantly a mood disorder characterised by extreme fluctuations in mental state, specifically mania or hypomania and depression (Phillips & Kupfer, 2013). Manic episodes are characterised by elevated or irritable mood, impulsivity, lowered inhibition, and at times psychosis. Conversely, depressive symptoms mirror those of major depressive disorder including anhedonia, low motivation and sleep disturbances. Bipolar disorder is categorised by two subtypes; bipolar I, and bipolar II. The differentiating factor between bipolar diagnoses is the severity of mania, which is lower in bipolar II cases (Phillips & Kupfer, 2013). There are no psychotic symptoms restricted to either schizophrenia or bipolar disorder, however, the frequency and severity of psychosis typically varies between the two disorders, for example schizophrenia cases are more likely to report persistent and severe hallucinations compared to bipolar cases (Baethge et al., 2005).

2.2.1 Common genetic architecture of schizophrenia and bipolar disorder

Schizophrenia and bipolar disorder have a strong genetic aetiology. Heritability estimates from twin studies are approximately 80% for schizophrenia (Sullivan

et al., 2003) and bipolar disorder (Cardno et al., 1999). Common genetic variants of small effect contribute substantially to the genetic architecture of schizophrenia (International Schizophrenia Consortium, 2009). The most recent analysis by the schizophrenia PGC used a discovery sample comprising of 32838 cases and 44357 controls, predicting 18.4% of the variance of schizophrenia liability in MGS-EA (Schizophrenia Working Group of the Psychiatric Genetics Consortium, 2014). The variance explained is substantially larger than the 3% identified using ISC discovery sample (International Schizophrenia Consortium, 2009). An equivalent analysis for bipolar disorder showed common genetic variants explained approximately 3% of variance of bipolar liability (PGC Bipolar Disorder Working Group, 2011).

For analyses relating to schizophrenia, increases in the variance of schizophrenia liability explained by common genetic variation over time is attributable to increasingly large schizophrenia discovery samples. Statistical modelling of polygenic risk score analyses show that increases in discovery sample size lead to more reliable effect sizes for individual SNPs, resulting in better predictive power to detect association with a phenotype in an independent sample (Chatterjee et al., 2013; Dudbridge, 2013)

Common genetic architecture of schizophrenia and bipolar disorder has overlapping features. Schizophrenia polygenic risk predicts between 2-2.5% of variance for bipolar disorder liability (International Schizophrenia Consortium, 2009; Smoller et al., 2013). In addition, schizophrenia polygenic risk is a better predictor of bipolar cases with psychotic traits (Hamshere et al., 2011),

suggesting the genetic risk of psychosis is dimensional rather than discrete (Craddock et al., 2009).

Individual loci also show association for schizophrenia and bipolar disorder. Genome wide association studies (GWAS) have shown individual risk loci for schizophrenia such as *ZNF804A* become strengthened with the inclusion of bipolar individuals (O'Donovan et al., 2008a). Conversely, *CACNA1C* originally associated with bipolar disorder (PGC Bipolar Disorder Working Group, 2011) has also been reported as genome wide significant in schizophrenia (Schizophrenia Working Group of the Psychiatric Genetics Consortium, 2014). In addition, a GWAS combining schizophrenia and bipolar cases into a unitary phenotype identified a novel locus within *PIK3C2A* reaching genome wide significance (Ruderfer et al., 2014). This locus was not previously identified using individual GWAS of the two disorders, showing common genetic variants can be identified using a broader psychosis phenotype despite increased phenotypic heterogeneity,

The role of copy number variations (CNVs) in schizophrenia will be discussed at greater length in Chapter 4. Briefly, specific large, rare (both inherited and de-novo) CNVs are highly enriched in schizophrenia probands (Rees et al., 2011; Rees et al., 2014b). De-novo CNVs are also enriched in bipolar probands (Malhotra et al., 2011), although large CNVs are less robustly associated with bipolar disorder compared to schizophrenia (Grozeva et al., 2010). Furthermore, neuropsychiatric CNVs enriched in schizophrenia probands are not more common in bipolar probands relative to controls (Grozeva et al., 2010). In

summary, schizophrenia and bipolar have aspects of shared genetic architecture, however their genetic differences may contribute to their separable phenotypes.

2.2.2 Cognitive Deficits in Schizophrenia and Bipolar Disorder

Schizophrenia cases show greater cognitive impairment compared to individuals with bipolar disorder (Murray et al., 2004; Green, 2006). Differences in premorbid impairment are also observed, for example general cognitive ability is associated with increased risk of schizophrenia but not bipolar disorder (Zammit et al., 2004; Sorensen et al., 2012).

Schizophrenia is characterised by premorbid deficits across multiple cognitive domains from performance (block design, picture completion) and verbal (information, similarities, vocabulary and arithmetic) tasks (Meier et al., 2014). In contrast, less is known regarding premorbid cognitive ability in bipolar disorder. Individuals with both inferior and superior school ability have an increased risk of developing bipolar disorder (MacCabe et al., 2010), and particular impairments on visuospatial reasoning have been identified (Tiihonen et al., 2005).

After the onset of psychosis, schizophrenia cases show the greatest deficits on general cognitive ability (Dickinson et al., 2008) and specific domains of attention/vigilance, reasoning/problem solving, speed of processing, social cognition, verbal learning/memory, visual learning/ working memory (Nuechterlein et al., 2004). Euthymic bipolar cases exhibit cognitive impairment

in several domains affected in schizophrenia (Green, 2006), and a meta-analysis of cognitive impairment in bipolar disorder have shown the largest impairments are in domains of verbal memory and executive functioning (Robinson et al., 2006; Arts et al., 2008; Bora et al., 2009).

Other studies have tested cognition in bipolar individuals using the MATRICS cognitive battery making results more easily comparable with schizophrenia. Burdick and colleagues found bipolar I cases could be separated into 3 subgroups using cognitive performance (Burdick et al., 2014). The first showed normal performance across all domains, but cases have superior social cognition. The second showed selective deficits on speed of processing, attention, verbal learning and social cognition, whilst the third demonstrated severe deficits across all 7 domains. However, this study was performed in a small sample of 136 bipolar cases, and larger studies are required to replicate these findings. Another study in 50 bipolar cases found deficits in processing speed, working memory, verbal and visual learning (Van Rheenen & Rossell, 2014). However, studies using the MATRICS cognitive battery are performed in substantially smaller samples when compared to meta-analyses, and it is unclear whether these domains accurately measure impairments in cognitive domains that impact cases with bipolar disorder.

The association between symptom domains and cognition also varies across disorders. Cognitive impairment shows no association with the positive symptoms of schizophrenia (Dominguez et al., 2009). In contrast, the level of

cognitive impairment in bipolar cases may be mediated by the severity of psychosis (Simonsen et al., 2011).

Differences in bipolar type I and II have also been observed. A meta-analysis (Bora et al., 2011) of cognitive impairment in bipolar I and II cases showed the biggest difference was in the domain of verbal learning, although significant impairments were also observed for semantic fluency and visual memory.

In summary. these findings indicate that cognitive deficits in schizophrenia are substantial, and effect multiple cognitive domains and general cognitive ability. The findings from bipolar disorder are suggestive of less substantial impairment across fewer cognitive domains. In addition, symptom dimensions have no association with of schizophrenia, although differences may be present across bipolar subtypes.

2.2.3 Heritability estimates of cognition in schizophrenia and bipolar disorder

The genetic basis to cognitive deficits in schizophrenia and bipolar disorder has been investigated using twin and sibling study designs. The Consortium on the Genetics of Schizophrenia (COGS) (Calkins et al., 2007) is an on-going family study with the purpose of identifying reliable endophenotypes in schizophrenia. The study uses families with at least one affected proband, and information on both parents and at least one unaffected sibling is required. Cognition was measured using a shortened version of the CNB and other tests measuring

attention, working memory, verbal memory, spatial memory/processing and emotional recognition (Calkins et al., 2007). Heritability estimates were between 0.24 and 0.55 for all cognitive tasks (Greenwood et al., 2007), demonstrating a low to moderate effect. Other studies have shown IQ is more heritable compared to individual cognitive traits in schizophrenia cases and unaffected family members (Husted et al., 2009; Owens et al., 2011). Unaffected siblings of schizophrenia probands also show lower cognitive ability on "g" (Wisner et al., 2011), processing speed (Wisner et al., 2011) and working memory (Conklin et al., 2005; Wisner et al., 2011) compared to healthy controls.

A meta-analysis of cognitive traits in bipolar cases and their first-degree relatives showed both had deficits on attention, verbal memory, executive functioning, inhibition and set shifting (Bora et al., 2009). Deficits in speed of processing, verbal fluency and visual memory were seen exclusively in cases. Heritability studies of cognition in bipolar families are suggestive of genetic influence on executive functioning ($h^2 \sim 0.6$), processing speed ($h^2 = 0.72$) and verbal ability ($h^2 = 0.96$) (Antila et al., 2007), suggesting a genetic basis for cognitive deficits in bipolar disorder.

En masse, cognitive deficits are heritable and present in unaffected siblings, thus making them potential endophenotypic candidates. Endophenotypes are unobservable traits associated with disease under stronger genetic than environmental influence, and considered to be less complex and heterogeneous than the overarching disease (Gottesman, 2003). One approach is to test the association between genetic variants associated with disease and

endophenotypic candidates (Walters & Owen, 2007). The following section discusses the use of polygenic risk scoring to test this association.

2.2.4 Polygenic Scores and Cognition

Given that schizophrenia is characterised by severe deficits in cognitive functioning, several studies have used polygenic risk score analysis to investigate the association between schizophrenia polygenic risk and general cognitive ability in healthy individuals. By testing cognition in healthy controls, this removes potentially confounding variables present within schizophrenia cohorts including illness duration, symptom severity and medication effects. Three previous publications have investigated the relationship between cognition in the general population and schizophrenia using polygenic score analysis. Increased polygenic risk of schizophrenia was associated with lower general cognitive ability (Lencz et al., 2014) and a decline in general cognitive ability across the lifespan (McIntosh et al., 2013). However, these studies only explain modest proportions (%$r^2 \approx 1$) of cognitive variance, and have not consistently replicated (van Scheltinga et al., 2013). Conversely, a direct assessment of the endophenotype approach showed polygenic scores for "g" were associated with schizophrenia caseness, however only a small proportion of variance for schizophrenia liability (%$r^2 < 0.5$) was found (Lencz et al., 2014).

A limitation of these studies was that they all limited their investigation to global measures of IQ or g, which neglects possible associations between schizophrenia polygenic risk and individual cognitive domains, or across verbal or non-verbal ability. Furthermore, the identification of cognitive domains most closely aligned

with genetic risk of schizophrenia would be useful for future endophenotype studies.

2.2.5 Aims/Hypotheses

This chapter builds upon previous research by investigating polygenic risk of schizophrenia, bipolar disorder and their differences, to assess their association with measures of generalised and specific cognitive ability. The following hypotheses are addressed:

1) Is increased polygenic risk of schizophrenia associated with lower cognitive ability using various measures of generalised cognition and specific cognitive domains? This hypothesis is tested by training on three different schizophrenia datasets: PGC1 (Schizophrenia PGC, 2011), CLOZUK (Hamshere et al., 2013) and PGC2 (Schizophrenia Working Group of the Psychiatric Genetics Consortium, 2014). Polygenic risk of schizophrenia was tested for association with cognitive ability in an adult German cohort (Rujescu et al., 2003) and in the ALSPAC sample(Golding et al., 2001), a UK longitudinal cohort with results of cognitive testing available at age 8.

The following analyses were performed:

i) Schizophrenia PGC1 to German Cognition

ii) Schizophrenia PGC1 to ALSPAC

ii) CLOZUK to ALSPAC

iii) PGC2 to ALSPAC

2) Is polygenic risk of bipolar disorder associated with lower cognitive ability using various measures of generalised cognition and specific cognitive domains? This hypothesis was tested using one bipolar dataset (PGC Bipolar Disorder Working Group, 2011), with polygenic risk of bipolar disorder tested for association with cognition in ALSPAC.

3) Given differences in the cognitive profiles of schizophrenia and bipolar cases, are common variants associated with increased risk of schizophrenia relative to bipolar disorder associated with performance on the cognitive domains outlined previously? This hypothesis was tested using a discovery GWAS of schizophrenia versus bipolar (case/control) (Ruderfer et al., 2014). Polygenic risk of schizophrenia relative to bipolar disorder was tested for association with cognition in ALSPAC.

2.3 Methods

2.3.1 Schizophrenia/Bipolar Samples

2.3.1.1 PGC1

The stage 1 discovery sample of the PGC1 was comprised of 9394 schizophrenia cases and 12462 controls across 17 separate studies. Genotyping was performed separately for each sample, but combined imputation and quality control was performed by the PGC Bipolar Disorder Working Group. As described in PGC Bipolar Disorder Working Group, 2011, SNPs with <5% missing data, individuals were retained if the missing genotype rate per individual was less than 2%.

Subsequently, SNPs were retained if the missing genotype rate per SNP was less than 2%, the missing genotype rate between cases and controls per SNP was less than 2%, Hardy-Weinberg equilibrium (controls) $P > 1 \times 10^{-6}$ and the frequency difference to the HapMap reference was <0.15. Data was imputed using BEAGLE 3.0.4, with phased HapMap phase 3 data as a reference panel.

Using the publically available PGC1 schizophrenia GWAS summary statistics, we subsequently included SNPs in the present study if they had an INFO score > 0.9 and a minor allele frequency > 0.01.

2.3.1.2 CLOZUK

CLOZUK is a UK sample comprising of 5,554 schizophrenia cases and 6299 controls (Schizophrenia Working Group of the Psychiatric Genetics Consortium, 2014). Cases were ascertained through collaboration with Norvartis, a manufacturer of the atypical antipsychotic Clozapine (Clozaril) in the UK. Cases with a diagnosis of treatment-resistant schizophrenia (at least two unsuccessful courses of previous anti-psychotic treatment) were included based upon clozapine registration forms completed by treating psychiatrists. The controls were from the Wellcome Trust Case Control Consortium (WTCCC), National Blood Service and 1958 British Birth Cohort. Samples were genotyped on the Illumina OmniExpress (Illumina Inc.).

CLOZUK samples were genotyped in two waves on the Illumina Omni Express and Illumnia Combo. Quality control measures for SNPs were performed by (Schizophrenia Working Group of the Psychiatric Genetics Consortium, 2014)

and outlined below. The quality control parameters for retaining SNPs and subjects were: SNP missingness < 0.05 (before sample removal); subject missingness < 0.02; autosomal heterozygosity deviation (| Fhet | < 0.2); SNP missingness < 0.02 (after sample removal); difference in SNP missingness between cases and controls < 0.02; and SNP Hardy-Weinberg equilibrium (P > 10−6 in controls or P > 10−10 in cases). Genotype imputation was performed using the pre-phasing/imputation stepwise approach implemented in IMPUTE2 / SHAPEIT (chunk size of 3 Mb and default parameters). The imputation reference set consisted of 2,186 phased haplotypes from the full 1000 Genomes Project dataset (August 2012, 30,069,288 variants, release "v3.macGT1"). We retained SNPs that were imputed with high quality (imputation information score INFO) > 0.9) and a minor allele frequency of > 1% for subsequent analysis.

2.3.1.3 PGC2

The PGC2 discovery dataset was comprised of 52 separate studies. QC of the individual datasets was described above. They tested all 52 GWAS datasets separately for association with schizophrenia using an additive logistic regression model in PLINK using population principal components as covariates (Schizophrenia Working Group of the Psychiatric Genetics Consortium, 2014). They performed a meta-analysis of the 52 sets of results using an inverse-weighted fixed effects model. Summary statistics were available for 35476 schizophrenia cases and 46839 controls. Only SNPs with high confidence (imputation information score > 0.9) and minor allele frequency greater than 1% were used for further analysis.

2.3.1.4 PGC1 Bipolar

Publically available data from the PGC bipolar disorder primary sample was used. This comprised of 7841 individuals with bipolar disorder and 9250 controls from eleven samples collected across Europe and America (PGC Bipolar Disorder Working Group, 2011). Bipolar cases had a diagnosis of either Bipolar Disorder type I or II, schizoaffective bipolar or "other" bipolar illness. Genotyping was performed separately for each sample, but combined imputation and quality control was performed by the PGC Bipolar Disorder Working Group. As described in PGC Bipolar Disorder Working Group, 2011), SNPs with <5% missing data, individuals were retained if the missing genotype rate per individual was less than 2%. Subsequently, SNPs were retained if the missing genotype rate per SNP was less than 2%, the missing genotype rate between cases and controls per SNP was less than 2%, Hardy-Weinberg equilibrium (controls) $P > 1 \times 10^{-6}$ and the frequency difference to the HapMap reference was <0.15. Data was imputed using BEAGLE 3.0, with phased HapMap phase 2 data as a reference.

Using the publically available PGC bipolar GWAS summary statistics, we subsequently included SNPs in the present study if they had an INFO score > 0.9 and a minor allele frequency > 0.01.

2.3.1.5 PGC Schizophrenia vs Bipolar Disorder

A GWAS was performed comparing 7129 schizophrenia cases (cases) against 9252 bipolar cases (controls) (Ruderfer et al., 2014) using data from the PGC. All

quality control steps are described from (Ruderfer et al., 2014). Quality control was performed separately on the on the 31 datasets. To generate reliable SNP data from across the different datasets, individuals (low genotyping rates < 0.98, IBD < 0.1, abnormal heterozygosity F > 0.15) and SNPs (minor allele frequency < 0.01, genotyping < 0.98, Hardy-Weinberg equilibrium P < 1E-6, and missing data > 0.02 between cases and controls) underwent additional quality control. Samples were imputed using BEAGLE and HapMap Phase 3+TSI was used as the reference panel.

Using GWAS summary statistics from this analysis, we included SNPs with an INFO score > 0.9 and minor allele frequency > 1% for the current analysis.

2.3.2 Cognition Samples

2.3.2.1 Sample description/Genotyping

2.3.2.1.1 German sample

Individuals of German decent were recruited at random via mail invitation in the city of Munich (Rujescu et al., 2003). Individuals and their first-degree relatives were screened for psychiatric diagnoses. The sample was genotyped on an Affymetrics 5 custom chip and underwent standard quality control. Specifically, individuals were removed for ambiguous sex coding, abnormal heterozygosity, high identity by decent (IBD) and genotyping rates less than 98%. SNPs were filtered by minor allele frequency (>0.01), a genotyping rate of 99% and Hardy–Weinberg equilibrium (p>1E-06). After quality control, 955 unrelated individuals and 698,308 genotyped SNPs were available for analysis.

2.3.2.1.2 ALSPAC

The Avon Longitudinal Study of Parents and Children (ALSPAC) is an on-going longitudinal study of over 14000 mothers and their offspring born in Avon between April in 1991 and the end of December in 1992 (Golding et al., 2001; Boyd et al., 2013). They were recruited via "expression of interest" through media communication, community centres and health services. Mothers interested in partaking in the study were invited to contact ALSPAC for further information. In total, 14676 women enrolled, in phase 1. A large amount of behavioural, cognitive, health and environmental data has been collected in children across 68 time points between time of birth and age 18. During this time period, both mothers and children were followed up via 59 questionnaires measuring health, social behaviour and physical development. Nine clinical assessments between the ages of 7-17 were used to obtain physiological information, cognitive information and psychological/social wellbeing. Attrition is problematic for any longitudinal study. Approximately 3000 families have partaken in every assessment, whilst approximately 5800 families have completed at least 75% of assessments. Furthermore, whilst some individuals have been lost due to mortality, have withdrawn from the study or become untraceable, currently nearly 12000 children are eligible for follow up assessments.

In addition to phenotypic collection, a large number of individuals have also been genotyped as part of the study. To date, 9912 children were genotyped using the Illumina HumanHap550. Samples underwent routine quality control. Specifically, individuals were removed for ambiguous sex coding, abnormal heterozyogisty,

high IBD, genotyping rates less than 97%, and population outliers were removed using multidimensional scaling analysis. SNPs were filtered by minor allele frequency (>1%), a 95% genotyping rate, and Hardy–Weinberg equilibrium (p>5E-07). After quality control, 8,365 unrelated individuals and 500,527 genotyped SNPs were available for analysis. EIGENSTRAT analysis revealed no additional obvious population stratification (Price et al., 2006). Data was imputed using MACH 1.0.16 Markov Chain Haplotype software (Li et al., 2010) and HapMap phase 2 CEPH population (HG18, release 22) was used as a reference panel.

2.3.2.2 Cognition testing

2.3.2.2.1 Cognitive assessment in the German sample

936 individuals had complete data on the HAWIE-R (Tewes, 1991), a German version of the WAIS-R. This measures five performance tests, six verbal tests, and provides performance, verbal and full scale IQ scores. Brief descriptions of the individual tests are given below.

Verbal tests:

Arithmetic - individuals are told stories containing mathematical information and are asked questions relating where they are required to use mental arithmetic.

Comprehension - individuals are asked questions relating to real life situations requiring social judgement and awareness of social norms.

Digit span - individuals are asked to repeat a set a numbers forwards and backwards. The length of the number string increases during the test.

Information - individuals are asked general knowledge questions.

Similarities - individuals hear word pairings, and asked to say in what way the two words are alike.

Vocabulary - individuals are asked to explain the meaning of words that increase in difficulty during the test.

Performance tests:
Block design - individuals manipulate cubes coloured with red and white patterns, and arrange them to match a shown pattern.

Digit Symbol Coding - A printed key that contains 9 matching pairs of numbers and symbols. Individuals are given a sheet containing numbers and are asked to draw the corresponding symbol in a blank space underneath.

Object Assembly - individuals are asked to make complete objects using individual jigsaw style pieces.

Picture Arrangement - individuals are required to arrange a scrambled set of cards telling a story into a logical order.

Picture Completion - Individuals are presented with pictures containing a missing section. They are asked to identify what is missing from each picture.

Scaled test scores for performance and verbal tests are summed separately, and converted to performance and verbal IQ scores respectively using standardised tables, which adjusts for sex and age. The scaled scores for all tests are summed and converted to full-scale IQ scores using standardised tables, adjusting for sex and age.

We chose to include a subset of the WAIS tests for further analysis based upon their similarity with the MATRICS domains (Nuechterlein et al., 2004) as cognitive domains most affected in schizophrenia cases may be under stronger influence from schizophrenia genetic risk factors. Three WAIS tests were considered to reliably test the following MATRICS domains: verbal working memory (digit span), speed of processing (digit symbol coding), reasoning/problem solving (block design). We also used performance and verbal IQ as composite measures of non-verbal and verbal ability, and full scale IQ as a measure of general cognitive ability.

2.3.2.2.2 Cognitive assessment in ALSPAC

Phenotype data was taken from the ALSPAC collection undertaken when the child was 8 years old. Individuals were administered the short form Wechsler Intelligence Scale for Children (WISC-III)(Wechsler et al., 1992) (alternate items used for all subtests except the coding subtest) as well as two additional items taken from the Test of Everyday Attention for Children (TEACh)(Robertson et al.,

1996) and the Diagnostic Analysis of Nonverbal Accuracy (DANVA).(Nowicki & Duke, 1994).

Cognitive tests were selected a priori based on their similarity to the MATRICS domains based on our previous work (Niarchou et al., 2013). The following tasks were selected from the WISC-III, with their representative cognitive constructs in parenthesises: coding (processing speed), digit span backward (working memory), and block design (reasoning and problem solving). As a measure of verbal learning, we used the total number of non-words correctly recalled from an adapted version of the Nonword Repetition Test.(Gathercole et al., 1994) From the TEACh, the Sky Search task was selected as our measure of attention., after adjusting for motor speed. The DANVA was used to create a social cognition variable by creating a total number of errors (incorrect assignment of emotions) across all four emotional domains. WISC IQ scores for verbal, performance and total IQ were used as measures of verbal, nonverbal and general cognitive ability.

2.3.3 Polygenic analyses

GWAS summary statistics were publically available for PGC1 (Schizophrenia PGC, 2011), whilst CLOZUK and PGC2 summary statistics were obtained through collaboration with the PGC (Schizophrenia Working Group of the Psychiatric Genetics Consortium, 2014). SNPs were filtered by a minor allele frequency of less than 1% and imputation (INFO) score < 0.9. To maximise the number of variants for each individual analysis, SNPs were selected if they appeared in both the psychiatric discovery sample and cognition target sample under investigation.

Linkage disequilibrium (LD) based clumping selects the most significant SNP within a sliding LD block. This was performed in Plink (Purcell et al., 2007)across SNPs in common across datasets for each analysis (500kb sliding window, r^2 =0.25, minimum association p=0.5). For the analysis using the German cognition sample, clumped SNPs were extracted from schizophrenia datasets and filtered by 4 p-value thresholds ($p < 0.001$, 0.01, 0.1 and 0.5). For analyses using ALSPAC, Genetic clumped SNPs were extracted from schizophrenia datasets and filtered by 5 p-value thresholds ($p < 0.0001$, 0.01, 0.1, 0.3, 0.5). The 'score' command in PLINK was used to calculate polygenic scores (Purcell et al., 2007). Polygenic scores are calculated by summing the number of susceptibility alleles of the reference SNP weighted by the log of the SNP odds ratios. Schizophrenia polygenic scores were calculated for each individual in ALSPAC and German cognition samples.

Linear regressions were performed in R, where schizophrenia polygenic scores were used as predictors of performance on the cognitive tasks. Analyses using the German cognition sample covaried for age, sex and the first principle component from Eigenstrat. Analyses performed in ALSPAC used no covariates. There was no population stratification in ALSPAC, and individuals completed cognitive tests at the same developmental time point. All p-values are uncorrected for multiple comparisons.

For the German cognition sample, polygenic scores, WAIS tests and covariate data was available in 936 individuals. For ALSPAC the number of individuals with polygenic scores and cognitive data can be found in Table 2.1.

Cognitive Domain	N Individuals
Attention	5319
Problem Solving	5500
Processing Speed	5557
Social Cognition	5110
Verbal Learning	5553
Working Memory	5421
Performance IQ	5536
Verbal IQ	5541
Full scale IQ	5518

Table 2-1 - Number of individuals in ALSPAC with polygenic scores and cognition

2.4 Results

2.4.1 PGC1 to cognition in German sample

Schizophrenia polygenic risk (at 4 training thresholds; 0.5, 0.1, 0.01 and 0.001) was tested for association with cognition in in 936 adult controls. Full results can be found in Table 2.2 and Figure 2.1.

Increased polygenic risk of schizophrenia was associated with performance IQ at 0.5 and 0.01 training thresholds (%r^2~0.5-0.6). This association was in the predicted direction of effect (increased schizophrenia polygenic risk was associated with lower performance IQ). These findings were replicated at a trend level for 0.1 and 0.001 training thresholds in the same direction of effect

Significant associations were also observed between schizophrenia polygenic risk and processing speed at 0.5, 0.1 and 0.01 training thresholds (%r^2~0.5-1.1). Again, this association was in the predicted direction of effect (increased schizophrenia polygenic risk was associated with lower processing speed). This association was observed at a trend level for the 0.001 training threshold in the same direction of effect.

No significant associations were observed between schizophrenia polygenic risk and reasoning/problem solving, working memory, verbal IQ or full scale IQ.

Training Dataset	Training Threshold	Target Cognition	Uncorrected P (2-Tailed)	%r2	Direction of Coefficient
PGC1	0.01	Processing Speed	0.005	1.06	-
PGC1	0.1	Processing Speed	0.016	0.76	-
PGC1	0.01	Performance IQ	0.030	0.59	-
PGC1	0.5	Processing Speed	0.043	0.52	-
PGC1	0.5	Performance IQ	0.048	0.53	-
PGC1	0.001	Processing Speed	0.058	0.66	-
PGC1	0.001	Performance IQ	0.064	0.48	-
PGC1	0.1	Performance IQ	0.089	0.39	-
PGC1	0.1	Verbal IQ	0.271	0.08	+
PGC1	0.01	Full scale IQ	0.293	0.20	-
PGC1	0.01	Problem Solving	0.310	0.34	-
PGC1	0.001	Problem Solving	0.314	0.40	-
PGC1	0.5	Verbal IQ	0.317	0.04	+
PGC1	0.5	Working Memory	0.321	0.03	+
PGC1	0.001	Full scale IQ	0.351	0.22	-
PGC1	0.1	Problem Solving	0.558	0.21	-
PGC1	0.1	Working Memory	0.588	< 0.01	+
PGC1	0.5	Full scale IQ	0.741	0.07	-
PGC1	0.5	Problem Solving	0.854	0.14	-
PGC1	0.1	Full scale IQ	0.903	0.02	-
PGC1	0.001	Verbal IQ	0.903	0.01	+
PGC1	0.01	Verbal IQ	0.964	< 0.01	+
PGC1	0.001	Working Memory	0.978	< 0.01	-
PGC1	0.01	Working Memory	0.985	< 0.01	+

Table 2-2 - Regression analyses for schizophrenia polygenic risk (PGC1) predicting cognition in the adult German sample

Training threshold refers to the discovery sample SNP p-threshold. Direction of coefficient refers to whether increased polygenic risk of schizophrenia is associated with lower cognition (-), or better cognition (+)

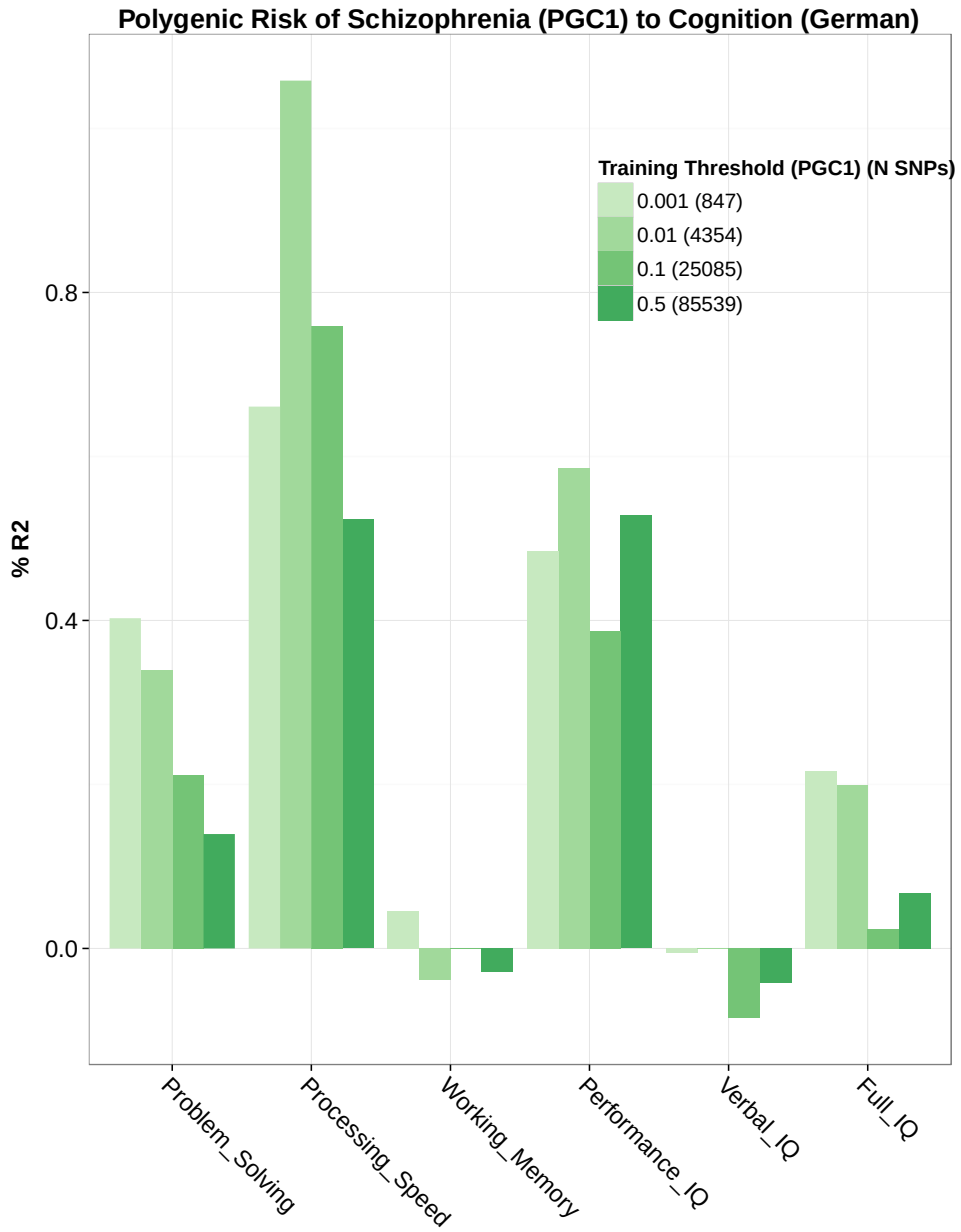

Figure 2.1 - Polygenic risk of schizophrenia (PGC1) and its association with cognition in German sample

%r^2 refers to the percentage of cognitive variance explained by schizophrenia polygenic risk at the four training thresholds. The 4 bars for each cognitive domain represent each of the different training threshold p-values. The number inside parentheses refers to the number of SNPs used for that training set. Bars with a positive %r^2 show associations are in the predicted direction of effect, whereby increased polygenic risk is associated with poorer performance.

2.4.2 CLOZUK to cognition in ALSPAC

This analysis derived schizophrenia polygenic risk from CLOZUK (Table 2.), and tested for association in ALSPAC on attention, problem solving, processing speed, social cognition, verbal learning, working memory., performance IQ, verbal IQ and full IQ. Full results can be found in Table 2.3 and Figure 2.2.

Polygenic risk of schizophrenia was significantly associated with verbal learning and performance IQ. The association between schizophrenia polygenic risk and performance IQ was found at 0.5, 0.3, 0.1 and 0.01 training thresholds (%r^2~0.07-0.09), and in the predicted direction of effect (increased schizophrenia polygenic risk was associated with lower performance IQ).

The association with verbal learning was observed for 0.5, 0.3 and 0.1 training thresholds (%r^2~0.09-0.13). However, this association was against the predicted direction of effect (increased schizophrenia polygenic risk was associated with better verbal learning). These findings replicated in the same direction of effect at a trend level for 0.01 and 0.0001 training thresholds.

Schizophrenia polygenic risk was not significantly associated with attention, problem solving, processing speed, social cognition, working memory, verbal IQ or full scale IQ even at a trend level.

Training Dataset	Training Threshold	Target Cognition	Uncorrected P (2-Tailed)	%r^2	Direction of Coefficient
CLOZUK	0.3	Verbal Learning	0.007	0.13	+
CLOZUK	0.5	Verbal Learning	0.011	0.11	+
CLOZUK	0.01	Performance IQ	0.022	0.09	-
CLOZUK	0.1	Verbal Learning	0.026	0.09	+
CLOZUK	0.1	Performance IQ	0.035	0.08	-
CLOZUK	0.3	Performance IQ	0.037	0.08	-
CLOZUK	0.5	Performance IQ	0.045	0.07	-
CLOZUK	0.01	Verbal Learning	0.061	0.06	+
CLOZUK	0.0001	Verbal Learning	0.068	0.06	+
CLOZUK	0.3	Processing Speed	0.120	0.04	-
CLOZUK	0.5	Social Cognition	0.173	0.04	-
CLOZUK	0.1	Social Cognition	0.189	0.03	-
CLOZUK	0.1	Processing Speed	0.214	0.03	-
CLOZUK	0.0001	Verbal IQ	0.230	0.03	-
CLOZUK	0.5	Processing Speed	0.238	0.02	-
CLOZUK	0.3	Verbal IQ	0.271	0.02	+
CLOZUK	0.01	Full scale IQ	0.273	0.02	-
CLOZUK	0.1	Verbal IQ	0.278	0.02	+
CLOZUK	0.5	Verbal IQ	0.284	0.02	+
CLOZUK	0.1	Attention	0.298	0.02	+
CLOZUK	0.3	Social Cognition	0.338	0.02	-
CLOZUK	0.3	Attention	0.349	0.02	+
CLOZUK	0.5	Attention	0.361	0.02	+
CLOZUK	0.01	Problem Solving	0.386	0.01	-
CLOZUK	0.01	Social Cognition	0.426	0.01	+
CLOZUK	0.0001	Full scale IQ	0.430	0.01	-
CLOZUK	0.0001	Working Memory	0.485	0.01	+
CLOZUK	0.1	Working Memory	0.487	0.01	+
CLOZUK	0.1	Full scale IQ	0.544	0.01	-

CLOZUK	0.01	Attention	0.574	0.01	+
CLOZUK	0.3	Full scale IQ	0.590	0.01	-
CLOZUK	0.5	Full scale IQ	0.599	< 0.01	-
CLOZUK	0.0001	Problem Solving	0.648	< 0.01	-
CLOZUK	0.3	Working Memory	0.675	< 0.01	+
CLOZUK	0.01	Processing Speed	0.676	< 0.01	-
CLOZUK	0.0001	Processing Speed	0.700	< 0.01	+
CLOZUK	0.5	Working Memory	0.736	< 0.01	+
CLOZUK	0.5	Problem Solving	0.791	< 0.01	+
CLOZUK	0.01	Verbal IQ	0.795	< 0.01	+
CLOZUK	0.0001	Performance IQ	0.822	< 0.01	-
CLOZUK	0.01	Working Memory	0.849	< 0.01	+
CLOZUK	0.3	Problem Solving	0.886	< 0.01	+
CLOZUK	0.1	Problem Solving	0.914	< 0.01	-
CLOZUK	0.0001	Social Cognition	0.917	< 0.01	-
CLOZUK	0.0001	Attention	0.966	< 0.01	+

Table 2-3 - Regression analyses for schizophrenia polygenic risk (CLOZUK) predicting cognition in ALSPAC

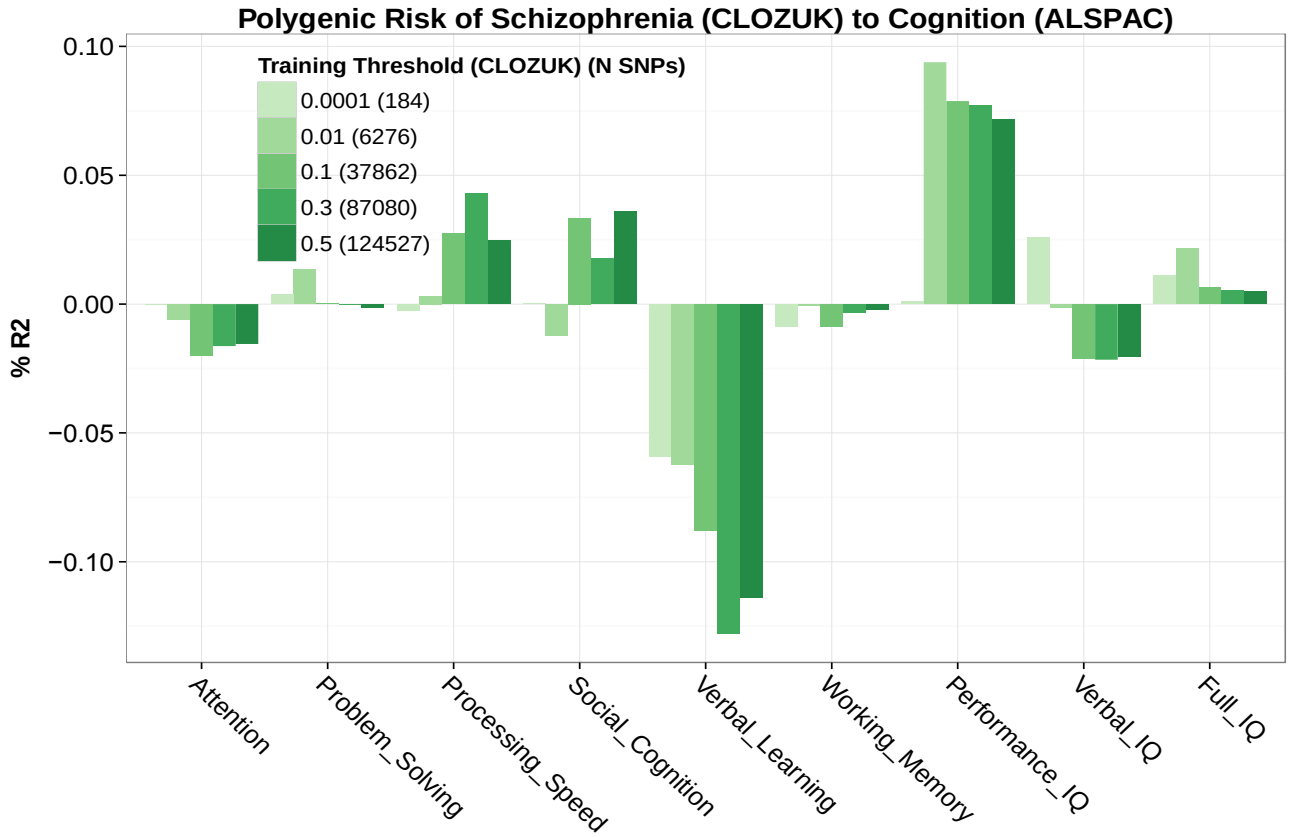

Figure 2.2 - Polygenic risk of schizophrenia (CLOZUK) and its association with cognition in ALSPAC

2.4.3 PGC1 to cognition in ALSPAC

This analysis derived schizophrenia polygenic risk from PGC1, and tested for association in ALSPAC on attention, problem solving, processing speed, social cognition, verbal learning, working memory, performance IQ, verbal IQ and full scale IQ. Full results can be found in Table 2.4 & Figure 2.3.

Polygenic risk of schizophrenia was significantly associated with processing speed, performance IQ, verbal IQ and full scale IQ. The strongest associations were for schizophrenia polygenic risk and performance IQ at 0.5, 0.3, 0.1 and 0.01 training thresholds (%r^2~0.21-0.28), and in the predicted direction of effect (increased schizophrenia polygenic risk was associated with lower performance IQ).

Increased polygenic risk of schizophrenia was also significantly associated with lower full scale IQ at 0.5, 0.3, 0.1 and 0.01 training thresholds (%r2~0.08-0.8). Significant associations in the predicted direction of effect were also observed with verbal IQ at 0.1 and 0.01 training thresholds (%r2~0.08-0.09).

Regarding individual cognitive domains, increased polygenic risk of schizophrenia was associated with lower processing speed at 0.5 and 0.3 training thresholds (%r2~0.09-0.12).

Associations between schizophrenia polygenic risk and both attention and working memory were also observed at a trend level against the predicted

direction of effect (increased polygenic risk was associated with better performance).

Training Dataset	Training Threshold	Target Cognition	Uncorrected P (2-Tailed)	%r^2	Direction of Coefficient
PGC1	0.5	Performance IQ	8.35E-05	0.2794	-
PGC1	0.3	Performance IQ	1.05E-04	0.2715	-
PGC1	0.1	Performance IQ	7.25E-04	0.2063	-
PGC1	0.3	Full scale IQ	0.002	0.1767	-
PGC1	0.5	Full scale IQ	0.002	0.1700	-
PGC1	0.1	Full scale IQ	0.003	0.1571	-
PGC1	0.5	Processing Speed	0.011	0.1150	-
PGC1	0.3	Processing Speed	0.022	0.0949	-
PGC1	0.01	Verbal IQ	0.027	0.0888	-
PGC1	0.01	Full scale IQ	0.030	0.0855	-
PGC1	0.1	Verbal IQ	0.036	0.0790	-
PGC1	0.3	Verbal IQ	0.057	0.0653	-
PGC1	0.5	Attention	0.076	0.0593	+
PGC1	0.5	Verbal IQ	0.076	0.0567	-
PGC1	0.5	Working Memory	0.076	0.0579	+
PGC1	0.1	Processing Speed	0.077	0.0564	-
PGC1	0.1	Working Memory	0.082	0.0558	+
PGC1	0.01	Performance IQ	0.082	0.0546	-
PGC1	0.3	Attention	0.093	0.0531	+
PGC1	0.3	Working Memory	0.094	0.0516	+
PGC1	0.1	Attention	0.115	0.0466	+
PGC1	0.01	Verbal Learning	0.131	0.0412	+
PGC1	0.3	Verbal Learning	0.216	0.0276	+
PGC1	0.5	Verbal Learning	0.249	0.0239	+
PGC1	0.01	Working Memory	0.275	0.0220	+

PGC1	0.01	Processing Speed	0.277	0.0213	-
PGC1	0.01	Attention	0.304	0.0199	+
PGC1	0.5	Social Cognition	0.332	0.0184	+
PGC1	0.1	Social Cognition	0.333	0.0184	+
PGC1	0.3	Social Cognition	0.355	0.0168	+
PGC1	0.1	Verbal Learning	0.363	0.0149	+
PGC1	0.0001	Processing Speed	0.371	0.0144	+
PGC1	0.01	Problem Solving	0.393	0.0133	+
PGC1	0.0001	Performance IQ	0.411	0.0122	-
PGC1	0.0001	Working Memory	0.518	0.0077	-
PGC1	0.01	Social Cognition	0.645	0.0042	+
PGC1	0.0001	Full scale IQ	0.649	0.0038	-
PGC1	0.0001	Attention	0.780	0.0015	+
PGC1	0.0001	Verbal IQ	0.878	0.0004	-
PGC1	0.0001	Verbal Learning	0.907	0.0002	+
PGC1	0.0001	Problem Solving	0.920	0.0002	-
PGC1	0.1	Problem Solving	0.924	0.0002	+
PGC1	0.5	Problem Solving	0.959	0.0000	+
PGC1	0.0001	Social Cognition	0.986	0.0000	-
PGC1	0.3	Problem Solving	0.987	0.0000	-

Table 2-4 - Regression analyses for schizophrenia polygenic risk (PGC1) predicting cognition in ALSPAC

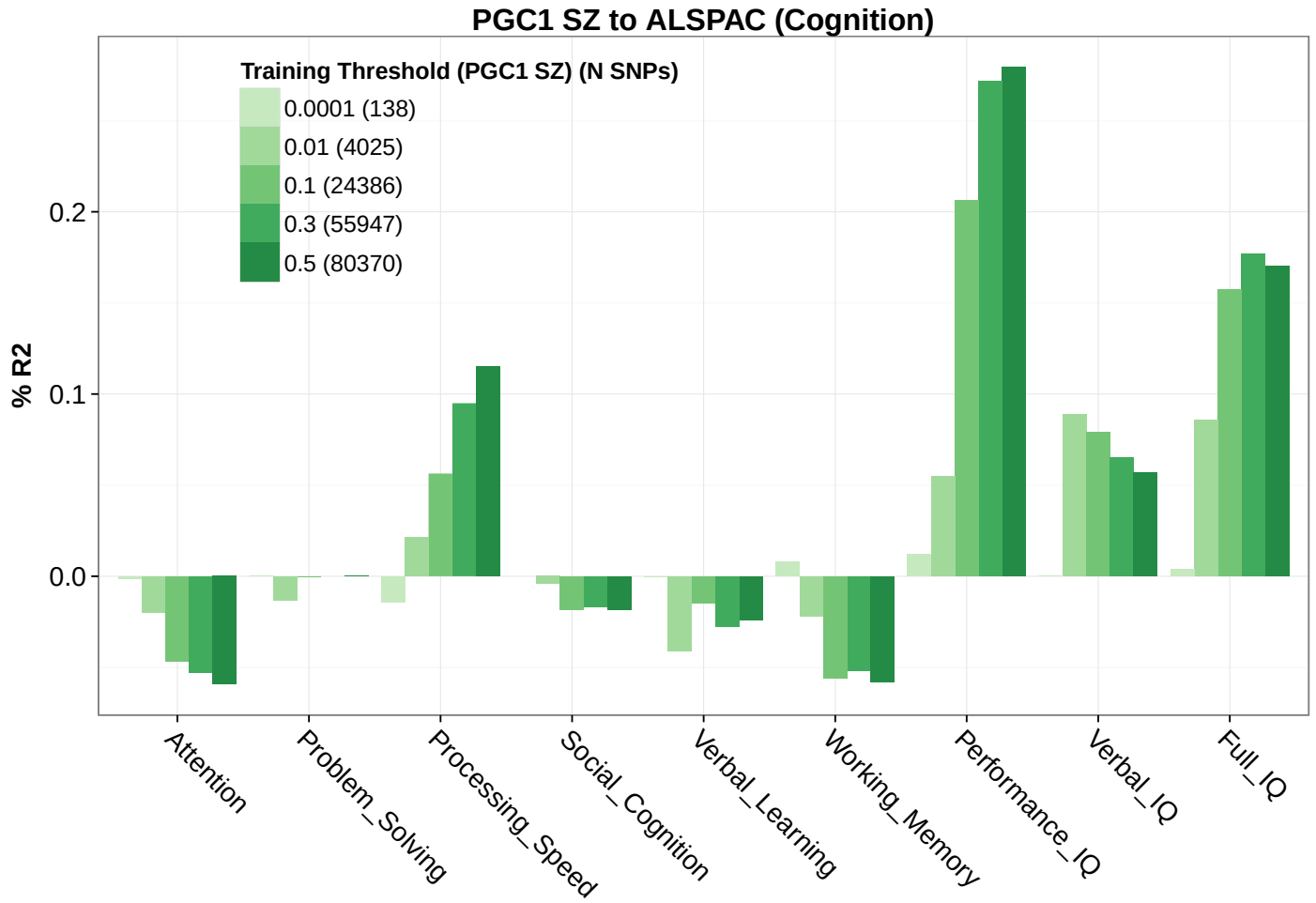

Figure 2.3 - Polygenic risk of schizophrenia (PGC1) and its association with cognition in ALSPAC

2.4.4 PGC2 to cognition in ALSPAC

This analysis derived schizophrenia polygenic risk from PGC2, the largest of the schizophrenia discovery samples, and was tested for association with cognition in ALSPAC (Table 2.5 & Figure 2.4).

Polygenic risk of schizophrenia was robustly associated with performance IQ, and nominally significant for processing speed and full scale IQ. Highly significant associations between schizophrenia polygenic risk and performance IQ were found at 0.5, 0.3, 0.1 and 0.01 training thresholds (%r^2~0.22-0.34), and in the predicted direction of effect (increased schizophrenia polygenic risk was associated with lower performance IQ).

Significant associations were also observed between schizophrenia polygenic risk and full scale IQ at 0.1 and 0.01 training thresholds (%r^2~0.1-0.13), and processing speed at 0.5, 0.3 and 0.1 training thresholds (%r^2~0.12-0.17) in the predicted direction of effect.

Training Dataset	Training Threshold	Target Cognition	Uncorrected P (2-Tailed)	% R2	Direction of Coefficient
PGC2	0.1	Performance IQ	1.23E-05	0.34	-
PGC2	0.5	Performance IQ	2.75E-04	0.24	-
PGC2	0.01	Performance IQ	3.76E-04	0.23	-
PGC2	0.3	Performance IQ	5.34E-04	0.22	-
PGC2	0.1	Processing Speed	0.002	0.17	-
PGC2	0.5	Processing Speed	0.005	0.14	-
PGC2	0.1	Full scale IQ	0.008	0.13	-
PGC2	0.3	Processing Speed	0.009	0.12	-
PGC2	0.01	Full scale IQ	0.019	0.10	-
PGC2	0.5	Verbal Learning	0.064	0.06	+
PGC2	0.5	Full scale IQ	0.069	0.06	-
PGC2	0.0001	Performance IQ	0.070	0.06	-
PGC2	0.3	Full scale IQ	0.074	0.06	-
PGC2	0.01	Processing Speed	0.081	0.05	-
PGC2	0.0001	Attention	0.082	0.06	+
PGC2	0.3	Verbal Learning	0.084	0.05	+
PGC2	0.5	Attention	0.090	0.05	+
PGC2	0.01	Working Memory	0.109	0.05	+
PGC2	0.01	Attention	0.119	0.05	+
PGC2	0.0001	Full scale IQ	0.126	0.04	-
PGC2	0.3	Attention	0.146	0.04	+
PGC2	0.1	Attention	0.177	0.03	+
PGC2	0.0001	Verbal IQ	0.248	0.02	-
PGC2	0.01	Problem Solving	0.298	0.02	-
PGC2	0.1	Problem Solving	0.308	0.02	-
PGC2	0.0001	Working Memory	0.317	0.02	+
PGC2	0.1	Verbal Learning	0.370	0.01	+
PGC2	0.01	Social Cognition	0.373	0.02	+
PGC2	0.01	Verbal IQ	0.443	0.01	-

PGC2	0.0001	Social Cognition	0.459	0.01	+
PGC2	0.01	Verbal Learning	0.511	0.01	+
PGC2	0.1	Verbal IQ	0.572	0.01	-
PGC2	0.0001	Processing Speed	0.572	0.01	-
PGC2	0.1	Social Cognition	0.576	0.01	+
PGC2	0.5	Problem Solving	0.590	0.01	-
PGC2	0.3	Problem Solving	0.633	< 0.01	-
PGC2	0.5	Social Cognition	0.647	< 0.01	+
PGC2	0.3	Social Cognition	0.709	< 0.01	+
PGC2	0.0001	Problem Solving	0.762	< 0.01	+
PGC2	0.0001	Verbal Learning	0.805	< 0.01	-
PGC2	0.3	Working Memory	0.840	< 0.01	+
PGC2	0.5	Working Memory	0.842	< 0.01	+
PGC2	0.1	Working Memory	0.858	< 0.01	+
PGC2	0.5	Verbal IQ	0.941	< 0.01	+
PGC2	0.3	Verbal IQ	0.978	< 0.01	-

Table 2-5 - Regression analyses for schizophrenia polygenic risk (PGC2) predicting cognition in ALSPAC

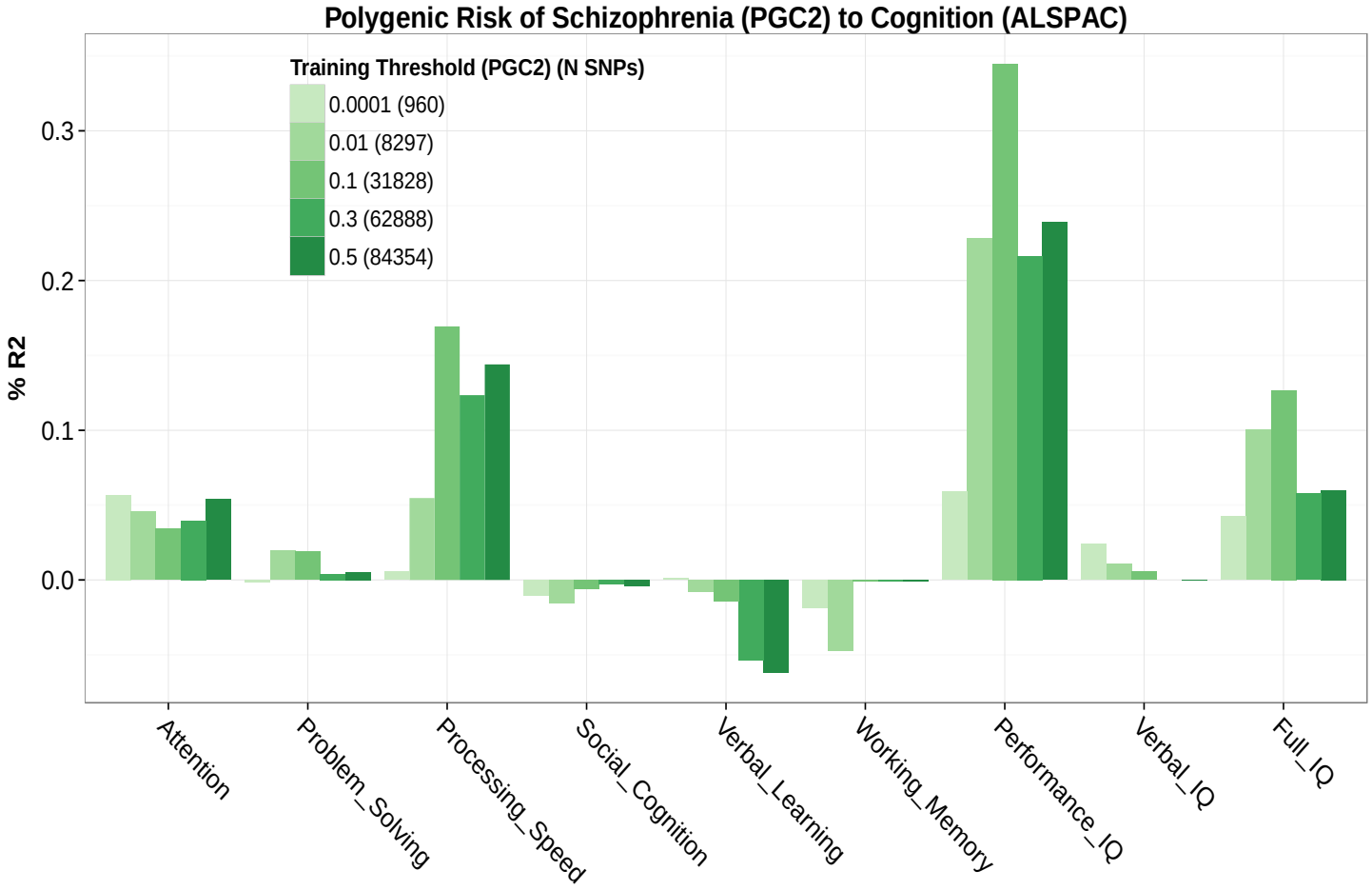

Figure 2.4 - Polygenic risk of schizophrenia (PGC2) and its association with cognition in ALSPAC

2.4.5 PGC1 Bipolar to cognition in ALSPAC

This analysis derived bipolar polygenic risk from the bipolar PGC1 GWAS, and tested for association in ALSPAC on attention, problem solving, processing speed, social cognition, verbal learning, working memory, performance IQ, verbal IQ and full scale IQ (Table 2.6 & Figure 2.5).

Polygenic risk of bipolar disorder was significantly associated with processing speed using the 0.1 training threshold (%r^2=0.11) in the predicted direction of effect (increased bipolar polygenic risk was associated with lower processing speed). However, bipolar polygenic risk was not significantly associated with processing speed at 0.5, 0.3, 0.01 and 0.0001 training thresholds even at the trend level.

Significant associations were observed between bipolar polygenic risk and social cognition at the 0.0001 training threshold (%r^2=0.08) against the predicted direction of effect (increased polygenic risk was associated with better social cognition). This trend was also observed for 0.5 and 0.3 training thresholds. Increased bipolar polygenic risk was also associated with better working memory at a trend level for 0.5, 0.3, 0.01, although none reached nominal levels of significance.

No robust association was observed between bipolar polygenic risk and attention, problem solving, processing speed, performance IQ, verbal IQ or full scale IQ.

Training Dataset	Training Threshold	Target Cognition	Uncorrected P (2-Tailed)	% R2	Direction of Coefficient
PGC1 Bipolar	0.1	Processing Speed	0.012	0.1116	-
PGC1 Bipolar	0.0001	Social Cognition	0.039	0.0823	+
PGC1 Bipolar	0.3	Social Cognition	0.051	0.0736	+
PGC1 Bipolar	0.5	Social Cognition	0.064	0.0663	+
PGC1 Bipolar	0.3	Working Memory	0.068	0.0607	+
PGC1 Bipolar	0.1	Performance IQ	0.068	0.0593	-
PGC1 Bipolar	0.5	Working Memory	0.073	0.0587	+
PGC1 Bipolar	0.01	Verbal IQ	0.074	0.0568	+
PGC1 Bipolar	0.01	Working Memory	0.080	0.0557	+
PGC1 Bipolar	0.1	Attention	0.100	0.0503	+
PGC1 Bipolar	0.01	Problem Solving	0.102	0.0480	+
PGC1 Bipolar	0.0001	Problem Solving	0.102	0.0480	+
PGC1 Bipolar	0.3	Attention	0.107	0.0483	+
PGC1 Bipolar	0.1	Working Memory	0.135	0.0407	+
PGC1 Bipolar	0.1	Social Cognition	0.155	0.0390	+
PGC1 Bipolar	0.0001	Processing Speed	0.180	0.0319	+
PGC1 Bipolar	0.3	Processing Speed	0.183	0.0315	-
PGC1 Bipolar	0.01	Social Cognition	0.196	0.0322	+
PGC1 Bipolar	0.0001	Working Memory	0.208	0.0289	+
PGC1 Bipolar	0.0001	Verbal IQ	0.227	0.0260	+
PGC1 Bipolar	0.0001	Full scale IQ	0.234	0.0253	+
PGC1 Bipolar	0.01	Full scale IQ	0.260	0.0226	+
PGC1 Bipolar	0.5	Processing Speed	0.285	0.0203	-
PGC1 Bipolar	0.5	Attention	0.294	0.0204	+
PGC1 Bipolar	0.5	Problem Solving	0.309	0.0186	+
PGC1 Bipolar	0.1	Full scale IQ	0.321	0.0176	-
PGC1 Bipolar	0.0001	Performance IQ	0.337	0.0165	+
PGC1 Bipolar	0.3	Problem Solving	0.448	0.0103	+

PGC1 Bipolar	0.5	Verbal Learning	0.465	0.0095	+
PGC1 Bipolar	0.3	Verbal Learning	0.503	0.0080	+
PGC1 Bipolar	0.3	Full scale IQ	0.523	0.0073	-
PGC1 Bipolar	0.3	Performance IQ	0.536	0.0068	-
PGC1 Bipolar	0.01	Attention	0.662	0.0035	+
PGC1 Bipolar	0.5	Full scale IQ	0.663	0.0034	-
PGC1 Bipolar	0.0001	Verbal Learning	0.702	0.0026	+
PGC1 Bipolar	0.5	Performance IQ	0.703	0.0026	-
PGC1 Bipolar	0.3	Verbal IQ	0.764	0.0016	-
PGC1 Bipolar	0.01	Processing Speed	0.800	0.0011	+
PGC1 Bipolar	0.0001	Attention	0.820	0.0010	-
PGC1 Bipolar	0.5	Verbal IQ	0.831	0.0008	-
PGC1 Bipolar	0.1	Verbal Learning	0.858	0.0006	-
PGC1 Bipolar	0.01	Performance IQ	0.907	0.0002	-
PGC1 Bipolar	0.1	Problem Solving	0.929	0.0001	-
PGC1 Bipolar	0.01	Verbal Learning	0.935	0.0001	+
PGC1 Bipolar	0.1	Verbal IQ	0.998	0.0000	-

Table 2-6 - Regression analyses for bipolar polygenic risk predicting cognition in ALSPAC.

P-values are two-tailed and uncorrected

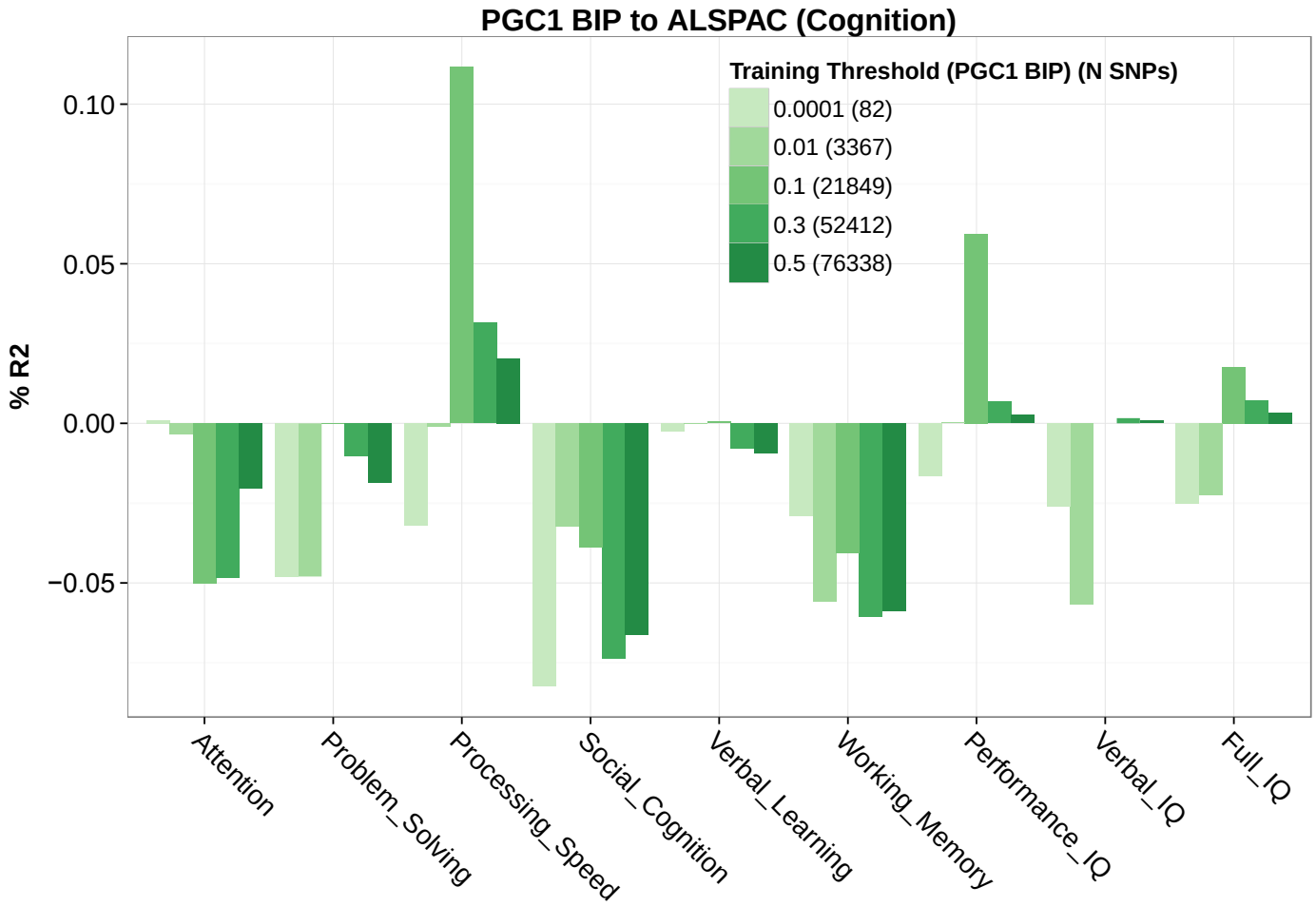

Figure 2.5 - Polygenic risk of bipolar disorder and its association with cognition in ALSPAC

2.4.6 Schizophrenia vs Bipolar to cognition in ALSPAC

This analysis derived schizophrenia versus bipolar disorder polygenic risk, and tested for association in ALSPAC on attention, problem solving, processing speed, social cognition, verbal learning, working memory, performance IQ, verbal IQ and full scale IQ (Table 2.7 & Figure 2.6).

Increased polygenic risk of schizophrenia relative to bipolar disorder was significantly associated with full IQ, verbal IQ and performance IQ at 0.5, 0.3, 0.1 and 0.01 training thresholds (full IQ $\%r^2$=0.13-0.27; verbal IQ $\%r^2$ =0.09-0.27; performance IQ $\%r^2$ =0.10-0.17). Polygenic risk scores were also significantly associated with problem solving at 0.3 and 0.1 training thresholds ($\%r^2$ =0.07-0.08).

No robust association was observed between schizophrenia vs bipolar polygenic risk and attention, processing speed, social cognition, verbal learning or working memory.

Training Dataset	Training Threshold	Target Cognition	Uncorrected P (2-Tailed)	% R2	Direction of Coefficient
PGC1 SCZ vs Bipolar	Full scale IQ	0.1	1.03E-04	0.2732	-
PGC1 SCZ vs Bipolar	Verbal IQ	0.1	1.21E-04	0.2666	-
PGC1 SCZ vs Bipolar	Verbal IQ	0.3	1.65E-04	0.2559	-
PGC1 SCZ vs Bipolar	Full scale IQ	0.3	1.72E-04	0.2556	-
PGC1 SCZ vs Bipolar	Full scale IQ	0.5	1.96E-04	0.2512	-
PGC1 SCZ vs Bipolar	Verbal IQ	0.5	2.31E-04	0.2446	-
PGC1 SCZ vs Bipolar	Performance IQ	0.1	0.002	0.1693	-
PGC1 SCZ vs Bipolar	Performance IQ	0.3	0.004	0.1510	-
PGC1 SCZ vs Bipolar	Performance IQ	0.5	0.004	0.1492	-
PGC1 SCZ vs Bipolar	Full scale IQ	0.01	0.006	0.1380	-
PGC1 SCZ vs Bipolar	Performance IQ	0.01	0.019	0.0999	-
PGC1 SCZ vs Bipolar	Verbal IQ	0.01	0.019	0.0994	-
PGC1 SCZ vs Bipolar	Problem Solving	0.01	0.034	0.0822	-
PGC1 SCZ vs Bipolar	Problem Solving	0.3	0.048	0.0713	-
PGC1 SCZ vs Bipolar	Problem Solving	0.5	0.058	0.0653	-
PGC1 SCZ vs Bipolar	Problem Solving	0.1	0.080	0.0559	-
PGC1 SCZ vs Bipolar	Social Cognition	0.0001	0.081	0.0598	+
PGC1 SCZ vs Bipolar	Attention	0.01	0.162	0.0367	+
PGC1 SCZ vs Bipolar	Verbal IQ	0.0001	0.163	0.0351	-
PGC1 SCZ vs Bipolar	Verbal Learning	0.01	0.178	0.0327	-
PGC1 SCZ vs Bipolar	Verbal Learning	0.1	0.211	0.0282	-
PGC1 SCZ vs Bipolar	Processing Speed	0.5	0.233	0.0256	-
PGC1 SCZ vs Bipolar	Full scale IQ	0.0001	0.250	0.0240	-
PGC1 SCZ vs Bipolar	Verbal Learning	0.3	0.268	0.0221	-
PGC1 SCZ vs Bipolar	Processing Speed	0.3	0.308	0.0187	-
PGC1 SCZ vs Bipolar	Verbal Learning	0.5	0.330	0.0171	-
PGC1 SCZ vs Bipolar	Performance IQ	0.0001	0.426	0.0114	-
PGC1 SCZ vs Bipolar	Problem Solving	0.0001	0.447	0.0105	-
PGC1 SCZ vs Bipolar	Working Memory	0.1	0.465	0.0098	-

PGC1 SCZ vs Bipolar	Processing Speed	0.1	0.496	0.0084	-
PGC1 SCZ vs Bipolar	Working Memory	0.0001	0.530	0.0073	+
PGC1 SCZ vs Bipolar	Processing Speed	0.0001	0.538	0.0068	+
PGC1 SCZ vs Bipolar	Social Cognition	0.3	0.573	0.0062	+
PGC1 SCZ vs Bipolar	Working Memory	0.01	0.609	0.0048	+
PGC1 SCZ vs Bipolar	Attention	0.5	0.642	0.0041	+
PGC1 SCZ vs Bipolar	Attention	0.0001	0.667	0.0035	+
PGC1 SCZ vs Bipolar	Social Cognition	0.01	0.697	0.0030	-
PGC1 SCZ vs Bipolar	Social Cognition	0.5	0.755	0.0019	+
PGC1 SCZ vs Bipolar	Working Memory	0.3	0.803	0.0011	-
PGC1 SCZ vs Bipolar	Attention	0.3	0.814	0.0010	+
PGC1 SCZ vs Bipolar	Verbal Learning	0.0001	0.836	0.0008	-
PGC1 SCZ vs Bipolar	Processing Speed	0.01	0.848	0.0007	-
PGC1 SCZ vs Bipolar	Social Cognition	0.1	0.854	0.0007	+
PGC1 SCZ vs Bipolar	Attention	0.1	0.882	0.0004	+
PGC1 SCZ vs Bipolar	Working Memory	0.5	0.887	0.0004	-

Table 2-7 - Polygenic risk of schizophrenia versus bipolar disorder and its association with cognition

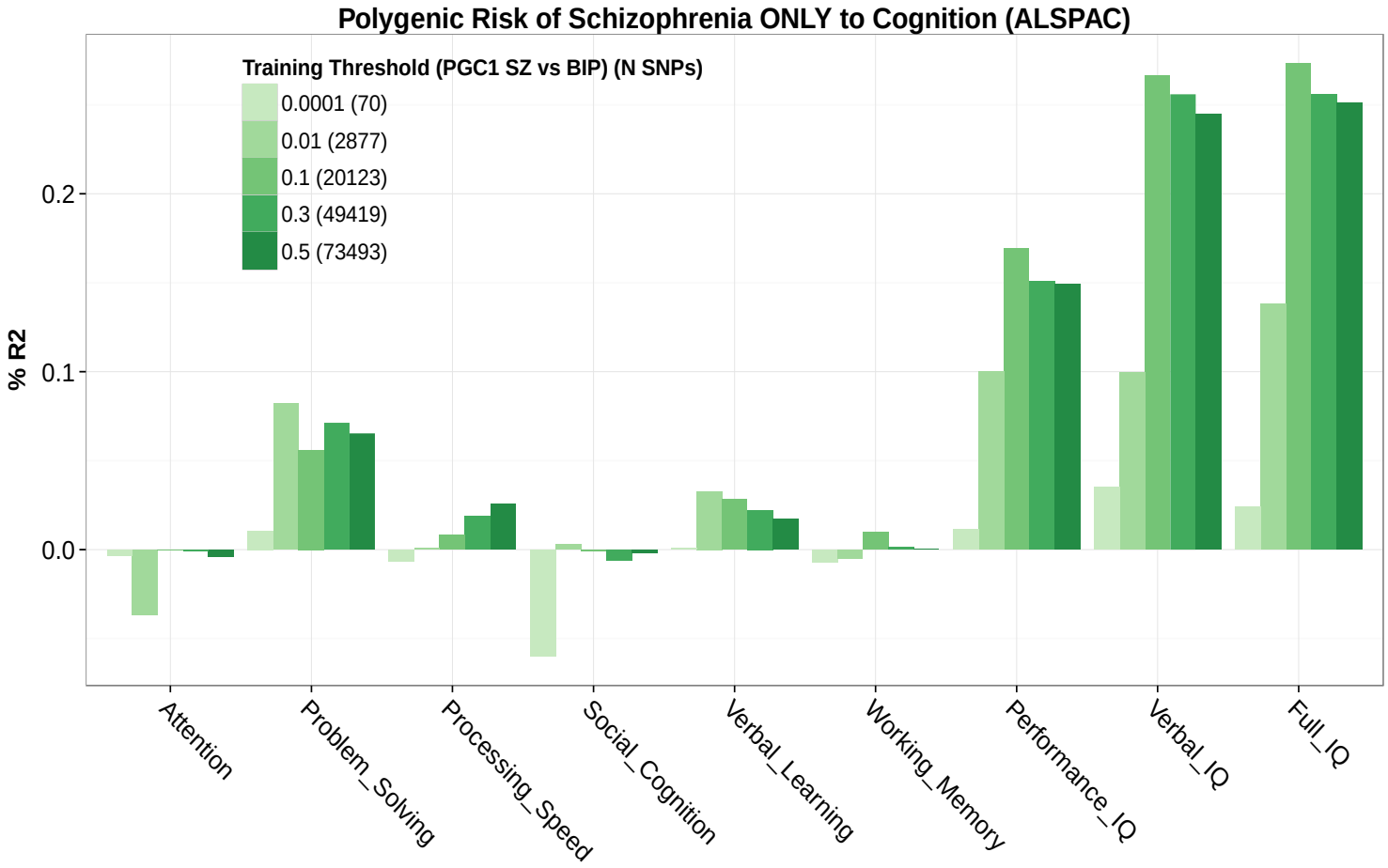

Figure 2.6 - Polygenic risk of schizophrenia versus bipolar disorder and its association with cognition in ALSPAC

2.5 Discussion

2.5.1 Polygenic risk of schizophrenia and its association with cognition

The first hypothesis in this chapter postulated that increased polygenic risk of schizophrenia was associated with lower performance across various measures of generalised and specific cognition. This was tested using three schizophrenia training datasets, and two independent cognition samples.

We observed an association between schizophrenia polygenic risk and performance IQ. This result was consistent across all analyses, and in the predicted direction of effect whereby increased polygenic risk was associated with lower performance IQ. When training on PGC1 where the target sample was comprised of 936 healthy adults, the proportion of performance IQ variance explained by schizophrenia polygenic risk was between 0.5-0.6% in regression analyses reaching nominal levels of significance. This is substantially higher than the variance of performance IQ explained within ALSPAC. The variance of performance IQ in ALPSAC explained by CLOZUK was 0.1%, PGC1 was 0.28%, increasing to 0.34% using PGC2.

These findings are also supported using a traditional endophenotype approach with regards to polygenic risk scoring (Hubbard et al., submitted). Polygenic scores for performance IQ explain between 0.04-0.09% of the variance of schizophrenia liability when compared against individual cognitive domains of attention, problem solving, processing speed, social cognition, working memory, verbal learning, verbal IQ and full scale IQ. Furthermore, within this study,

bivariate GCTA analyses showed a substantial genetic correlation between schizophrenia and performance IQ in ALSPAC (r_g~ -0.22 to –0.38); the largest of all the cognitive phenotypes tested. Collectively, these findings provide strong evidence showing that schizophrenia and performance IQ show the strongest genetic relationship with respect to common genetic variation. However, validation of these results outside of our group would be beneficial.

The relatively low increase in variance between PGC1 and PGC2 suggests further increasing the size of the schizophrenia training set is unlikely to explain substantially greater amounts of performance IQ variance. Given the severity of cognitive deficits in schizophrenia cases, the modest proportion of the variance explained across cognitive domains, even amongst the most strongly associated tests, is surprising given the large sample sizes available.

Processing speed is one of the most impaired cognitive domains in schizophrenia cases (Dickinson et al., 2007). The results from the present study found increased polygenic risk of schizophrenia was associated with lower processing speed, providing evidence for a degree of genetic relatedness. However, there were notable differences using different schizophrenia training sets and cognition samples. The largest processing speed variance (>1%) was observed at one training threshold in the German cognition sample. However, within ALSPAC, processing speed was not significant using CLOZUK (the smallest of the schizophrenia training sets), although 4/5 training thresholds were in the predicted direction of effect. Furthermore, processing speed showed greater

association training on PGC1 and PGC2 datasets, although the proportion of the variance explained (0.2-0.3%) was substantially lower compared to the German sample. However, the German sample contained over 4000 fewer individuals compared to ALSPAC. Possible inflation of results may have occurred due to sample bias in the German sample caused by low numbers, thus these results may have overestimated the variance explained of processing speed and performance IQ in relation to schizophrenia polygenic risk.

2.5.2 Polygenic risk of bipolar disorder and its association with cognition

The second hypothesis in this chapter asked whether polygenic risk of bipolar disorder was associated with lower cognitive ability. The findings from both specific and general cognitive domains generally show no such association. A weak association was observed showing increased polygenic risk of bipolar disorder was associated with lower processing speed at a single training threshold (0.1). A nominally significant association was observed showing increased bipolar polygenic risk was associated with better social cognition at a single training threshold (0.0001), continuing at a trend level for the other training thresholds.

Several studies have reported upon social cognition in bipolar cases. A tentative explanation comes from one study that investigated performance on the MATRICS cognitive battery in 136 bipolar cases and 148 controls (Burdick et al., 2014). They identified cases were categorised by three clusters based upon MATRICS performance. One cluster was comprised of cases with relatively preserved cognitive functioning relative to controls, but with superior social

cognition. However, only 43 bipolar cases were assigned to this cluster, and thus are not characteristic of the majority of bipolar cases.

Heritability studies of cognition in bipolar families are suggestive of genetic influence on executive functioning, processing speed and verbal ability (Antila et al., 2007), which generally have large effect sizes in bipolar cases (Bora et al., 2009). However, our results suggest polygenic risk of bipolar disorder does not contribute to these cognitive domains in healthy individuals.

2.5.3 Polygenic risk of schizophrenia versus bipolar disorder and its association with cognition

The third hypothesis in this chapter investigated whether common genetic variants associated with increased risk of schizophrenia relative to bipolar disorder were associated with cognitive ability. This analysis used a schizophrenia (case) / bipolar (control) GWAS, where genetic variants showing stronger association had a higher allelic frequency in schizophrenia cases compared to bipolar cases. Conversely, SNPs with weak association meant no significant differences in the allelic frequency. Differences in the common genetic architecture of schizophrenia and bipolar disorder may explain differences their respective phenotypes of these disorders, and within the context of this study, cognition.

The results of this analysis showed differences in the common genetic architecture of schizophrenia and bipolar disorder were robustly associated with verbal and full scale IQ. One interpretation is that genetic differences between

schizophrenia and bipolar disorder may explain premorbid differences in cognitive ability. Low premorbid IQ is associated with increased risk of schizophrenia but not bipolar disorder (Zammit et al., 2004). Common genetic variation between these two disorders may reflect neurodevelopmental differences resulting in earlier, and more severe cognitive impairment in schizophrenia relative to bipolar disorder.

2.5.4 Strengths/Limitations

This study has several strengths and limitations. First, this study has used the largest schizophrenia training set to date (PGC2) for investigating polygenic overlap with cognition, which is almost 4 times larger than previous studies (McIntosh et al., 2013; Lencz et al., 2014). This is particularly important for polygenic analyses, as the power to detect associations is largely attributable to the size of the training set (Dudbridge, 2013). Second, the number of individuals with cognitive data in ALSPAC is larger than previous studies (McIntosh et al., 2013; van Scheltinga et al., 2013; Lencz et al., 2014) and homogeneous regarding cognitive domains investigated, collection of cognitive data at a fixed time point (thus reducing the confounding effects of age) and low impact of population stratification. A third strength was the availability of two cognition samples, allowing replication of results in an independent dataset. However, it is unclear how comparable the results are for polygenic risk of schizophrenia and cognition between child and adult samples. Specifically, heritability of general cognitive ability becomes stronger across the lifespan, thus genetic factors influencing cognitive ability differ in children and adults. Furthermore, late childhood and early adolescence is when prospective schizophrenia cases begin to notably lag

behind their peers in cognitive tests (Reichenberg et al., 2010), meaning schizophrenia genetic risk variants may contribute more strongly to cognition before adulthood.

However, there are several limitations. Firstly, the schizophrenia datasets are not independent. Whilst PGC1 and CLOZUK samples do not overlap, both samples are included within the PGC2.

Secondly, it is unclear how comparable the results are for polygenic risk of schizophrenia and cognition between child and adult samples. Specifically, genetic factors exert stronger influence on cognition across the lifespan (Deary et al., 2012). Genetic factors influencing schizophrenia also modestly overlap with cognition (Fowler T, 2012). The differences in variance explained between processing speed and performance IQ in the adult German sample compared ALSPAC may therefore be attributable to a greater genetic influence over cognition in adults.

Third, although we used SNPs that were imputed with high quality, different imputation procedures were performed for ALSPAC, PGC1 and CLOZUK/PGC2 datasets that could contribute to differences in the results (Marchini & Howie, 2010)

2.5.5 Further work

Schizophrenia polygenic risk is associated with a negative change in general cognitive ability over the adult lifespan (McIntosh et al., 2013). However, relative

to healthy peers, cognitive lag during childhood and adolescence in prospective schizophrenia cases (Reichenberg et al., 2010) may be indicative of schizophrenia genetic factors exerting influence over cognitive ability during these developmental stages. Associations between schizophrenia polygenic risk and change in cognitive ability between childhood and adolescence may therefore be stronger in comparison to cognitive ability at age 8 alone. It may be change in cognition that the schizophrenia polygenic risk is associated with. Should go onto to say that this would explain the lesser degree of variance explained in our study compared to those with target datasets in adult populations as well as McIntosh results.

Although polygenic risk of bipolar disorder could not reliably predict any single cognitive phenotype, the PGC bipolar discovery dataset use a combination of type I and II cases. Cognitive deficits in bipolar disorder may be more prominent in type I relative to type II cases (Schenkel et al., 2012). If SNPs associated with cognitive ability was used as the training phenotype, cognition polygenic scores could be used to predict bipolar type I and type II individuals. This method may be more successful at identifying cognitive endophenotypes within bipolar subtypes.

2.5.6 Conclusion

We investigated whether increased polygenic risk of schizophrenia, bipolar disorder and their differences were associated with specific cognitive domains most affected in schizophrenia,

Common schizophrenia genetic variation en masse did not contribute substantially to cognitive ability within the general population. However, our evidence consistently demonstrated increased polygenic risk of schizophrenia was associated with lower performance IQ. These findings have important implications for future endophenotypic studies of neurocognition in schizophrenia and bipolar disorder. Previous studies have focused on schizophrenia and full scale IQ or "g" for deriving estimates of shared genetic variance (Toulopoulou 2007;Fowler et al, 2012) and common polygenic risk (Lencz et al., 2014). However, of the cognitive domains tested in the present study, this shows schizophrenia is most genetically related, at least in the context of common alleles, to performance IQ.

Polygenic risk of bipolar disorder was not strongly associated with any of the cognitive domains tested.

An analysis using a schizophrenia versus bipolar case control GWAS showed schizophrenia variants were associated with full and verbal IQ. Further interrogation of these variants may be useful for identifying genes associated with the generalised cognitive deficit in schizophrenia and provide clues pertaining to neurodevelopmental processes.

Selecting cognitive measures with the strongest genetic association with schizophrenia and bipolar disorder are required to maximise the benefit from future endophenotypic studies, thus providing greater insight into the neurobiology of mental illness.

3 Functional Pathways Underlying Cognitive Phenotypes in Schizophrenia cases and Healthy Controls

3.1 Summary

This chapter investigated the hypothesis that gene sets related to brain function, development and behaviour are *a priori* more likely to be enriched for SNPs influencing general cognitive ability. 155 gene-sets were used and grouped into six overarching categories: behaviour, cellular physiology, cellular morphology, development, region tract morphology and subcellular neuronal.

Two hypotheses were addressed. Firstly, does schizophrenia polygenic risk from common SNPs in brain related gene-sets predict general cognitive ability? It was found that schizophrenia polygenic scores for individual gene-sets were not predictive of performance IQ in 936 controls, or the MATRICS composite score in 496 cases after correction for multiple testing. Furthermore, there was no evidence that polygenic scores for gene-sets in specific categories showed greater association with general cognitive ability.

Secondly, are brain related gene-sets enriched for common SNPs associated with general cognitive ability in 496 schizophrenia cases? Using Brown's method implemented in set-screen there was no evidence of enrichment for any of the gene-sets tested. A comparison of the broader gene-set categories using the Mann-Whitney U-Test showed that gene-sets related to abnormal brain region and fibre tract morphology were ranked more highly.

3.2 Introduction

Over the past decade, it has emerged that schizophrenia has a complex polygenic architecture comprised of common polygenic risk (Schizophrenia Working Group of the Psychiatric Genetics Consortium, 2014), rare CNVs (Rees et al., 2014b), and rare *de novo* CNVs (Malhotra et al., 2011).

Progress in schizophrenia genetics can be attributed to increasingly large sample sizes and advances in genotyping technology allowing greater coverage of the genome, thus increasing the ability to detect both common and rare genetic variation. However, using this mass of information to further understand biological mechanisms contributing to disease poses several challenges. When considering findings from GWAS, the majority of SNPs show either weak, or no association with disease (Balding, 2006). In the largest schizophrenia GWAS to date, odds ratios for SNPs reaching genome wide-significance were typically less than 1.2 (Schizophrenia Working Group of the Psychiatric Genetics Consortium, 2014). This suggests that individual polymorphisms do not contribute substantially to schizophrenia liability. Modelling the functional effects of single SNPs or genes through cellular or animal based techniques is time consuming and expensive. Furthermore, the effects may be small, or change, depending on interactions with other genetic variants or environmental influences (Purcell, 2002).

To overcome these limitations, increasingly gene set-based approaches are utilised to identify the convergence of genetic signals upon specific biological

processes (Civelek & Lusis, 2014). Biological systems, pathways or gene-sets are comprised of genes with functional relatedness (Schadt, 2009), and thus are potentially more informative and more powerful for detecting association with a trait (Evangelou et al., 2012). An increasing number of databases with functional annotations are now available.

Gene ontology (GO) is one of the largest publically available databases. Pathways are organised in in a hierarchical structure based upon their associations with molecular functions, cellular components or biological processes (Ashburner et al., 2000). Molecular functions represent any biochemical reaction that results from a gene; cellular components document the cellular location where these biochemical reactions may occur, and biological processes represents the function of these biochemical reactions (Ashburner et al., 2000). These annotations may be derived experimentally, curated using non-experimentally derived data, electronic curation, or come from unknown sources (Rhee et al., 2008). Whilst the majority of annotations are derived from electronic curation, the proportion of manually curated entries is substantially increasing (Gene Ontology Consortium et al., 2013).

Other databases provide additional information by linking genes with quantitative or anatomical traits observed through the study of animal models. The mouse genome database is curated partially via automated processing of peer reviewed papers, and partially from groups who provide data directly to the database (Bult et al., 2013). This database links over 10,000 phenotypes to genes from experimental mouse models, and is a powerful translational resource that

can be applied to human models of disease or traits (Bult et al., 2013). Gene sets manually curated by experts are increasing in popularity because of their higher quality compared to automated methods (Wang et al., 2010).

A number of methods have been devised to analyse the association of SNPs in gene sets with a trait. Whilst it is recognised that different gene-set analytic methods are unlikely to produce identical results due to differences in their approaches and assumptions when modelling the data (Duncan et al., 2014), they are nonetheless valuable for extending results from GWAS.

Over-representation methods use genes containing an associated SNP below a specific p-threshold. The number of genes that reach significance are counted for each gene-set, and compared against the number of significant genes that are not within the gene-set (Holmans, 2010; Evangelou et al., 2012). Gene-sets are enriched when the proportion of significant genes associated with a trait is significantly higher than those not in the gene set (Holmans, 2010). Gene set enrichment analysis (GSEA) is an alternative method that ranks genes using the most significant SNP or gene-wide p-value (Subramanian et al., 2005). The Kolmogorov-Smirnov test compares whether genes within specified pathways are ranked significantly higher than genes in random pathways. This method negates the need to specify a p-value threshold for association used in overrepresentation analyses (Holmans, 2010).

Self-contained tests use a different approach by taking the association results for a set of SNPs within a gene set. A statistic is derived for all SNPs collectively to

test for association with a disease or trait. Set-based statistics may use the mean of SNP p-values (Purcell et al., 2007), kernel logistic regression (Wu et al., 2010) or Brown's approximation of Fisher's statistics (Moskvina et al., 2011). Brown's approximation utilises Fisher's approach for combining probabilities for non-independent tests (Brown, 1975), and is suitable for gene-set based analyses where SNPs within genes are unlikely to be independent due to LD.

Set-based tests are advantageous over overrepresentation analyses when multiple SNPs within the same gene contribute independently or semi-independently with a disease or trait. However, in large genes where only one causal variant is present, taking the combined p-value of all SNPs within the gene is likely to deflate the association signal (Duncan et al., 2014).

Several general methodological considerations exist for gene set analyses. There is little consensus regarding the contribution of SNPs within gene windows, referring to the area outside of the transcriptional start and end position of a gene. Some studies have used no gene window (Moskvina et al., 2011), whilst others have used between 5-500kb (Torkamani et al., 2008; Wang et al., 2010). SNPs within gene windows may have regulatory effects up to 20kb away (Veyrieras et al., 2008), although others will have no functional role and have a confounding effect (Holmans, 2010; Ramanan et al., 2012). This issue is partly an artefact of older genotyping technology before the implementation of genotype imputation, whereby the number of SNPs available for analysis was limited to genotyped SNPs exclusively. Imputation has resulted in a more dense distribution of SNPs within genes; meaning genic SNPs may also capture

regulatory signals through linkage disequilibrium. Thus, the usefulness of allocating gene windows is currently unclear (Holmans, 2010).

Second, if an associated SNP is located within more than one gene, and these genes are members of the same gene-set, the gene-set may show artificial association with the trait under investigation (Ramanan et al., 2012). This would have a greater impact upon overrepresentation analyses where the most significant SNP within a gene is used, however the use of bootstrapping would partly mitigate against this. Set-based tests are less likely to be affected as all available SNPs within the gene set are used, although spurious results could arise if SNP coverage in a gene is low. Third, if linkage disequilibrium is not accounted for between SNPs, this may result in false positive associations within, or across genes in an LD window because multiple SNPs may tag the true causal variant. Finally, larger genes are likely to contain more SNPs. Thus, gene-sets containing larger genes are more likely to show association with a trait due to chance.

Little is known regarding gene-sets underling cognitive ability in either healthy individuals or schizophrenia cases, or those that overlap between schizophrenia and general cognitive ability. Their identification could provide an additional opportunity for treatment intervention for the cognitive symptoms of schizophrenia.

A number of studies have investigated candidate gene-sets and their associations with cognition. One study using GSEA showed polymorphisms within genes

encoding for proteins in the mouse NMDA receptor complex were associated with fluid intelligence in 3511 healthy controls, but not crystallised intelligence, working memory or processing speed (Hill et al., 2014a). However, it is unclear how much variation in fluid intelligence NMDA genes explained. Other post-synaptic density gene-sets including mGlu5, AMPA and ARC were not associated with cognition, which may be indicative of a lack of power, or rather NMDA genes have specific roles in fluid intelligence, whereas other glutamatergic post-synaptic complexes do not.

Common SNPs in genes coding for synaptic G-protein receptors have been shown to be associated with general cognitive ability (Ruano et al., 2010), although this finding has not been replicated elsewhere (Hill et al., 2014b). Synaptic G-protein receptors contribute to synaptic transmission (Klose et al., 2010), synaptic plasticity and abnormal learning/memory in mice (Cooper et al., 2012). However, their role in human cognition remains uncertain.

A recent paper highlighted differential association of common SNPs in pathways relating to fluid and crystallised intelligence using Ingenuity (Christoforou et al., 2014). Pathways related to long-term synaptic depression were associated with crystallised intelligence. Synaptic depression is one form of synaptic plasticity (Collingridge et al., 2010) and may influence cognitive abilities such as long-term memory (Ge et al., 2010). In addition, gene-sets influencing neuronal density, morphology and integrity were associated with fluid intelligence. Gene-sets that overlapped between fluid and crystallised intelligence were related to structural properties including dendritic development, organisation of microtubles and

migration of GABAergic neurons (Christoforou et al., 2014). However, after correction for gene length, these associations were noted to be weaker, and it is unclear whether they reached nominal levels of significance.

Gene-sets underlying cognitive impairment in schizophrenia have received less attention. One group has applied polygenic risk scores to common SNPs in candidate pathways to test associations with cognition in cases with psychosis. (Nicodemus et al., 2014). Whilst autosomal polygenic risk scores can be used to broadly assess the common genetic overlap of the same, or different traits, this method lacks specificity when looking at biological processes that may be shared between traits. Nicodemus and colleagues investigated SNPs within *ZNF804A* and their functionally related genes (Hill et al., 2012). *ZNF804A* polygenic pathway scores showed association with performance IQ, spatial working memory and social cognition in 424 patients with psychosis, predicting between 1-3% of cognitive variance. However, significant associations were not consistently observed across all training thresholds.

Although these studies provide some insight into possible biological mechanisms underlying cognitive ability, many findings have not been externally replicated, or show weak levels of association.

3.3 Aims & Hypotheses

It is unclear whether common genetic risk factors contributing to schizophrenia are associated with general cognitive ability when restricted to polymorphisms

within specific biological pathways. In addition, it is unknown what pathways are specifically involved in the general cognitive deficit in schizophrenia cases. Using two separate methodologies, the following hypotheses were investigated:

3. Do schizophrenia polygenic risk scores derived from common SNPs in candidate pathways predict general cognitive ability measured through performance IQ in healthy controls and the MATRICS composite score in a schizophrenia patient sample? In addition, are there differences in association between the different pathway categories?

4. Using Brown's method, is there an enrichment of SNPs in candidate pathways that show association with general cognitive ability in a schizophrenia patient sample? In addition, are there differences in association between the different pathway categories?

3.4 Methods

3.4.1 Samples

3.4.1.1 Schizophrenia Discovery Samples (Polygenic Pathways)

Information for the individual sample recruitment, ascertainment and diagnoses for all samples used in PGC2 can be found in the supplementary data in the original paper (Schizophrenia Working Group of the Psychiatric Genetics Consortium, 2014). Briefly, all cases included in the analysis had a diagnosis of schizophrenia or schizoaffective disorder. Population matched controls were

available for all samples. Summary statistics were available for 35476 schizophrenia cases and 46839 controls. SNPs with high confidence imputation scores (INFO > 0.9), and a minor allele frequency greater than 0.01 were used to generate polygenic risk scores predicting performance IQ in the German cognition sample.

A second schizophrenia discovery sample used all individuals in the schizophrenia PGC sample described above, but without CLOZUK and Cardiff COGS samples. The number of individuals in this dataset was 29415 schizophrenia cases and 40101 controls. Schizophrenia polygenic risk was used to predict general cognitive ability in the Cardiff COGS sample, meaning both discovery and target datasets were independent.

3.4.1.2 Cognition Samples

3.4.1.2.1 German cognition sample

See section 2.3.2.2.1 for sample and genotype information regarding the German cognition sample. 936 individuals had available data for performance IQ, which were used in the present study. Analyses using controls were restricted to performance IQ because this cognitive measure showed the strongest association with schizophrenia polygenic risk in Chapter 2.

Individuals within the German cognition sample were used as controls within the PGC2 schizophrenia case/control analysis, thus these datasets are not independent. However, the aims these studies were different; specifically the

target variable in the present study is performance IQ rather than association with schizophrenia.

3.4.1.2.2 Cardiff COGS

CardiffCOGS is comprised of patients diagnosed with schizophrenia from across the UK, recruited from in-patient, community and voluntary mental health sectors. They were interviewed with the Schedules for Clinical Assessment in Neuropsychiatry (Wing et al., 1990), and best estimate lifetime diagnosis was based upon a review of case notes and their concordance with DSM-IV criteria. 496 schizophrenia cases were used in the present study.

Cardiff COGS were genotyped on the Illumnina HumanOmniExpressExome-8v1. Genotype quality control and imputation for Cardiff COGS was performed by the PGC Statistical Analysis Group. The quality control parameters for retaining SNPs and subjects were: SNP missingness < 0.05 (before sample removal); subject missingness < 0.02; autosomal heterozygosity deviation (| Fhet | < 0.2); SNP missingness < 0.02 (after sample removal); difference in SNP missingness between cases and controls < 0.02; and SNP Hardy-Weinberg equilibrium (P > 10−6 in controls or P > 10−10 in cases). Genotype imputation was performed using IMPUTE2 / SHAPEIT (chunk size of 3 Mb) using phased haplotypes from the full 1000 Genomes Project dataset (August 2012).

Dosage data was obtained for Cardiff COGS through the Schizophrenia PGC. Dosage files were converted into Plink bed/bim/fam format using

"gprobs2beagle" (http://faculty.washington.edu/browning/beagle utilities/utilities.html#gprobs2beagle) and "beagle2linkage" (http://faculty.washington.edu/browning/beagle utilities/utilities.html#beagle2linkage) java applications. SNPs with an INFO score > 0.9 and minor allele frequency greater than 1% were retained for further analysis.

3.4.2 Cardiff COGS cognition data

Patients completed the MATRICS cognitive battery (Kern et al., 2008; Nuechterlein et al., 2008), which measures ability on 10 separate tests encompassing 7 cognitive domains (attention/vigilance, reasoning/problem solving, speed of processing, social cognition, verbal learning/memory, visual learning/memory and working memory).

All subtests were normally distributed apart from the Trail Making Test , which underwent log transformation. The direction of the sign was reversed indicating longer task duration was associated with worse performance. Standardised domain z-scores were calculated in cases relative to performance of 103 controls. A composite z-score was derived using the standardized sum of the domain z-scores relative to controls.

Domain score calculations and imputation of missing individual test/domain scores were performed according to the instructions within the MATRICS manual (Nuechterlein & Green, 2006) (see Appendix A for additional details).

All further analyses used the MATRICS composite score only. The rationale behind this decision was firstly because previous studies observed the largest effects for general cognitive ability (Stefansson et al., 2014). Secondly, general cognitive ability is highly correlated with individual cognitive domains (Dickinson et al., 2008), and finally this strategy minimises the burden of multiple testing.

3.4.3 Candidate Gene-Sets

Gene-sets from the MGI Mammalian Phenotype database were previously categorised according to 5 overarching categories: behaviour, cellular physiology, cellular morphology, development and region tract morphology. The Mammalian Phenotype (MP) ontology and gene annotations were downloaded from the Mouse Genome Informatics (MGI) database (ftp://ftp.informatics.jax.org/pub/reports/index.html). These represent data from pharmacological and functional genetic studies in mice. Gene annotations arising from transgene and multi-gene manipulations were removed. Parent terms were identified for each MP term and assigned to all genes annotated with that child term. Genes were mapped from mouse to human using the mapping file HMD Human5.rpt, also downloaded from MGI.

Behavioural pathways pertain to abnormal behavioural or cognitive traits. Cognitive traits include abnormal temporal memory, spatial learning, motor learning and object recognition. Other behavioural traits include emotional affect, aggressive behaviour, and other abnormal social interactions.

Cellular morphology gene-sets are functionally related to abnormal structural properties of neurons and synapses. Several neuronal types are represented including dopaminergic, GABAergic, pyramidal and glial cells. Furthermore, axon, dendritic and synaptic malformations are also included.

Cellular physiology pathways pertain to abnormal synaptic processes. This includes abnormal synaptic transmission, excitatory and inhibitory potentials. Other synaptic properties include abnormal synaptic plasticity, long-term potentiation and depression.

Developmental gene-sets categorise malformations of different brain regions during neurodevelopment. This includes abnormal hippocampal, cerebellum, and forebrain amongst others. Other gene-sets include synaptic malformations between the axon and other neural or central nervous system tissues, and abnormal neuron differentiation.

Region tract morphology gene-sets are associated with abnormal structural properties within, and across brain regions. They include, but are not limited to enlarged ventricles, abnormal brain size, abnormal white matter morphology, abnormal temporal and parietal lobes, and other structures, including the hippocampus and hypothalamus.

Separately, subcellular neuronal gene-sets were derived from proteomic studies in rodents and humans that are predominantly associated with pre and

postsynaptic components. A subset of these pathways (ARC and NMDA receptors) previously showed enrichment for *de novo* CNVs in schizophrenia cases (Kirov et al., 2012). Definitions of all gene-sets and their respective pathway membership can be found in Appendix B.

3.4.4 Polygenic Pathway Analysis

Polygenic risk scores were derived for each individual pathway within the 6 overarching categories described above, and tested for association with performance IQ in controls and the MATRICS composite score in schizophrenia cases. The analysis proceeded as follows::

1) Overlapping SNPs between schizophrenia PGC and German cognition sample were identified. Separately, overlapping SNPs between schizophrenia PGC minus CLOZUK/Cardiff COGS and Cardiff COGS were identified. Where differences in strand alignment differences were found, strand flipping was performed in Plink where appropriate.
2) For each pathway, genomic coordinates of genes were identified according to the build of the target set (human genome (HG) assembly 18 for the German cognition sample, and HG19 assembly for Cardiff COGS). Only SNPs that were located within the transcriptional start and end position of the genes were taken forward for further analysis.
3) Schizophrenia SNPs within each pathway were filtered by INFO score > 0.9, and assigned to one of two training sets based upon association-

p values below 0.5 and 0.05. Polygenic pathway scores were derived using the "score" function in Plink.

4) For the German cognition sample, performance IQ was regressed against polygenic risk for each pathway separately, and the first principle component from Eigenstrat was used as a covariate. Performance IQ was already corrected for age and sex, and these variables were not used as covariates. For Cardiff COGS, the MATRICS composite score was regressed against polygenic risk for each pathway separately. Age and sex were used as covariates.
5) To correct for multiple testing, stages 4 and 5 were repeated 10,000 times by randomising the performance IQ score within the German sample, and the MATRICS composite score in Cardiff COGS. Permuted regression p-values are 2-tailed, reflecting the original analysis and do not account for direction of effect. The corrected regression p-value represents the number of times the original regression p-value for a pathway was less than the lowest permuted p-value across all pathways.
6) To assess whether association between the pathways and general cognitive ability differentiated across pathway categories, Mann-Whitney U-Tests were performed. This tested whether pathways belonging to a specific category were ranked more highly (based upon their two-tailed regression p-value) compared to the remaining 5 categories.
7) We used permutations to generate empirical p-values for each category that would correct for overlap between gene-sets. Using the

10,000 simulated datasets created in step 5, step 6 was repeated by randomising the performance IQ score within the German sample, and the MATRICS composite score in Cardiff COGS. The corrected Mann-Whitney p-value represents the number of times the original Mann-Whitney p-value for a pathway was less than the lowest permuted p-value across all pathways.

3.4.5 Brown's Method

The enrichment of SNPs in gene-sets influencing general cognitive ability in 488 schizophrenia cases was performed using the following method:

1) A linear GWAS was performed for the MATRICS composite score in Plink with age and sex included as covariates. SNP p-values were adjusted for genomic inflation using the "--adjust" flag in Plink.
2) For each gene, only SNPs within the transcriptional start and end were used.
3) Brown's method (Morton, 1975) was implemented in the Plink set-screen test (Purcell et al., 2007; Moskvina et al., 2011). The 1000 Genomes Project "Phase1 integrated release version3" (released in April 2012, ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/20110521/) was used to infer LD structure. A p-value for enrichment of common SNPs associated with the MATRICS composite score was produced for each of the 155 pathways.

4) A permutation approach was used to correct for the number of pathways tested. First, the MATRICS composite score was permuted across all individuals 1000 times. Stages 2-4 were repeated. Brown's corrected p-value represents the number of times the original brown p-value for a pathway was less than the lowest permuted p-value across all pathways.
5) To investigate differences amongst pathway categories, Brown's p-values for individual pathways were grouped by membership of their specific pathway category. Mann-Whitney U Tests were performed to assess whether Brown's p-values for pathways in a particular category were ranked more highly compared to the other 5 categories.
6) Permutations were used to correct for differences amongst the pathway categories by comparing the original Mann-Whitney p-value against that of 1000 simulations. The category corrected p-value represents the number of times the original Mann-Whitney p-value for a category was less than the lowest permuted p-value across all pathways.

3.5 Results

3.5.1 Polygenic Pathways (Controls)

Using alleles weighted by the log odds ratio in the full PGC dataset, polygenic pathway scores based upon all SNPs within each of the 6 categories were regressed against performance IQ at two training thresholds (0.5 & 0.05). Table 3.1 contains results for regressions of individual pathways reaching nominal significance. Appendix C contains results for all pathways.

Increased polygenic risk of schizophrenia in the pathway "abnormal hypothalamus morphology" was associated with lower performance IQ (training threshold = 0.05, corrected p=0.049, $\%r^2$=0.807). Polygenic risk of other pathways was not significantly associated with performance IQ after correction for multiple testing.

Individual pathways belonging to each category were ranked by their regression p-values separately for 0.5 and 0.05 training thresholds, and combined.
Each pathway category was compared against the remaining categories using the Mann-Whitney U-Test (Table 3.2). After correction for multiple testing, there was no evidence to suggest specific gene-set categories showed greater association with performance IQ.

Pathway	Pathway Category	Training Threshold	P	P (Corrected)	%R2	Direction of Effect
abnormal hypothalamus morphology	region tract morphology	0.05	0.006	**0.049**	0.807	-
abnormal neuron morphology	cellular morphology	0.5	0.007	0.063	0.797	+
abnormal basal ganglion morphology	region tract morphology	0.5	0.010	0.088	0.727	+
abnormal neurite morphology	cellular morphology	0.5	0.011	0.099	0.662	-
abnormal spinal cord morphology	region tract morphology	0.05	0.013	0.118	0.662	+
abnormal social conspecific interaction	behaviour	0.05	0.014	0.124	0.651	+
abnormal neuron morphology	cellular morphology	0.05	0.015	0.131	0.658	+
abnormal social investigation	behaviour	0.5	0.015	0.131	0.635	+
abnormal nervous system development	development	0.05	0.015	0.134	0.647	+
abnormal long term potentiation	cellular physiology	0.05	0.018	0.157	0.607	-
abnormal telencephalon development	development	0.5	0.020	0.178	0.582	+
abnormal nervous system morphology	region tract morphology	0.05	0.021	0.186	0.592	+
abnormal learning memory	behaviour	0.05	0.021	0.190	0.571	+
abnormal emotion affect behaviour	behaviour	0.5	0.022	0.194	0.565	-
abnormal midbrain morphology	region tract morphology	0.5	0.023	0.207	0.575	-
abnormal nervous system morphology	region tract morphology	0.5	0.026	0.226	0.558	+
abnormal brain morphology	region tract morphology	0.05	0.026	0.229	0.550	+
abnormal associative learning	behaviour	0.5	0.027	0.243	0.522	+
abnormal glial cell morphology	cellular morphology	0.5	0.030	0.262	0.535	+
abnormal forebrain morphology	region tract morphology	0.5	0.031	0.273	0.528	+
abnormal synaptic depression	cellular physiology	0.05	0.031	0.274	0.494	+
abnormal spinal cord morphology	region tract morphology	0.5	0.036	0.317	0.492	+

abnormal parental behaviour	behaviour	0.05	0.037	0.324	0.475	+
abnormal eating drinking behaviour	behaviour	0.05	0.038	0.340	0.474	+
abnormal paired pulse facilitation	cellular physiology	0.5	0.040	0.354	0.456	+
abnormal postnatal subventricular zone morphology	region tract morphology	0.05	0.041	0.364	0.466	-
abnormal inhibitory postsynaptic currents	cellular physiology	0.5	0.042	0.368	0.432	+
abnormal behavioural response to xenobiotic	behaviour	0.5	0.042	0.368	0.450	+
abnormal behavioural response to xenobiotic	behaviour	0.05	0.043	0.376	0.441	+
abnormal social conspecific interaction	behaviour	0.5	0.045	0.395	0.441	+
abnormal forebrain morphology	region tract morphology	0.05	0.045	0.397	0.457	+
abnormal brain morphology	region tract morphology	0.5	0.047	0.418	0.447	+
abnormal learning memory	behaviour	0.5	0.048	0.429	0.422	+
abnormal basal ganglion morphology	region tract morphology	0.05	0.051	0.448	0.415	+

Table 3-1 – ***Polygenic Pathways predicting performance IQ in healthy controls with regression p<0.05***

Direction of effect refers to the sign of the regression coefficient. "+" means increased schizophrenia polygenic risk of that pathway was associated with better performance IQ. "–" means increased polygenic risk of that pathway was associated with worse performance IQ.

Pathway Category	Mann Whitney (Corrected P)	Mann Whitney (Corrected P) (0.5)	Mann Whitney (Corrected P) (0.05)
Behaviour	1	0.8721	1
Cellular Morphology	0.7645	1	0.1089
Cellular Physiology	0.5623	1	0.5241
Development	1	1	1
Region Tract Morphology	0.0873	0.1432	0.1563
Subcellular Neuronal	1	1	0.9234

Table 3-2 - Mann-Whitney P-values for pathway categories (healthy controls)

Mann-Whitney U-Test p–values comparing each pathway category against the 5 remaining categories for combined, 0.5 and 0.05 training thresholds

3.5.2 Polygenic Pathways (Schizophrenia Cases)

Polygenic pathway scores based upon all SNPs within each of the 6 categories were regressed against the MATRICS composite score at two training thresholds (0.5 & 0.05) (See Table 3.3 for results reaching uncorrected $p<0.05$, and Appendix D for a complete list of results).

No schizophrenia polygenic pathway scores were associated with the MATRICS composite score after correction for multiple testing.

Individual pathways belonging to each category were ranked by their regression p-values separately for 0.5 and 0.05 training thresholds, and combined.
Each pathway category was compared against the remaining categories using the Mann-Whitney U-Test (Table 3.4). After correction for multiple testing, there was no evidence to suggest specific gene-set categories showed greater association with the MATRICS composite score.

Gene-set	Gene-set category	Training threshold	P	P (Corrected)	%R2	Direction of Effect
Abnormal cerebellum development	Development	0.5	0.0105	0.0854	1.3312	+
Abnormal depression related behaviour	Behaviour	0.05	0.0119	0.0967	1.2864	-
Abnormal hippocampus development	Development	0.05	0.0194	0.1576	1.1126	+
Abnormal oligodendrocyte morphology	Cellular morphology	0.05	0.0217	0.1762	1.0731	+
Abnormal depression related behaviour	Behaviour	0.5	0.0246	0.1996	1.0293	-
Dilated third ventricle	Region tract morphology	0.5	0.0252	0.2044	1.0100	-
Abnormal temporal memory	Behaviour	0.5	0.0277	0.2254	0.9866	-
Abnormal hippocampus development	Development	0.5	0.0307	0.2496	0.9511	+
Abnormal glial cell morphology	Cellular morphology	0.05	0.0366	0.2976	0.8900	+
Abnormal fear anxiety related behaviour	Behaviour	0.5	0.0395	0.3207	0.8643	-
Abnormal seizure response to inducing agent	Behaviour	0.5	0.0451	0.3660	0.8189	-
Abnormal contextual conditioning behaviour	Behaviour	0.5	0.0456	0.3703	0.8149	-

Table 3-3 - Polygenic Pathways predicting the MATRICS composite score in schizophrenia cases with regression p<0.05

Pathway Category	Mann Whitney (Corrected P)	Mann Whitney (Corrected P) (0.5)	Mann Whitney (Corrected P) (0.05)
Behaviour	0.0610	0.4441	0.1365
Cellular Morphology	1	1	1
Cellular Physiology	1	1	1
Development	0.6686	0.4759	1
Region Tract Morphology	1	1	1
Subcellular Neuronal	1	1	1

Table 3-4 – Mann-Whitney P-values for pathway categories (schizophrenia cases)

3.5.3 Brown's Test

Brown's test was used to investigate the hypothesis that SNPs en masse in the pathways tested would show association with the MATRICS composite score. Table 3.5 reports the top 10 most significant associations (complete results can be found in Appendix E). No single pathway showed nominal levels of significance ($p>0.05$) with the MATRICS composite score before correction for multiple testing.

To assess whether association could be localised to a particular pathway category, a Mann-Whitney U Test was performed comparing the ranks of p-values for a particular category against all other categories. After permutation corrections for gene-set overlap and multiple testing, gene sets associated with region tract morphology ($p=0.002$) had en masse higher p-values than those in other pathway categories (see Table 3.6 for full results).

Pathway Category	Pathway	N SNPs	P	P (Corrected)
Behaviour	Abnormal motor learning	12793	0.056	0.0951
Behaviour	Abnormal discrimination learning	5607	0.073	0.1240
Development	Abnormal CNS synapse formation	5657	0.075	0.1274
Region Tract Morphology	Abnormal hindbrain morphology	78393	0.083	0.1409
Cellular Physiology	Abnormal GABA mediated receptor currents	5186	0.094	0.1596
Cellular Physiology	Abnormal miniature inhibitory postsynaptic currents	5658	0.119	0.2021
Region Tract Morphology	Abnormal cerebrum morphology	87344	0.121	0.2055
Region Tract Morphology	Abnormal lateral ventricle morphology	16676	0.121	0.2055
Region Tract Morphology	Abnormal telencephalon morphology	109098	0.134	0.2275
Region Tract Morphology	Abnormal forebrain morphology	132002	0.136	0.2309

Table 3-5 – Top 10 most significant Brown p-values for enrichment of MATRICS SNPs in tested pathways (schizophrenia cases)

Pathway Category	Mann-Whitney P (1-Tailed)	Mann-Whitney Corrected P (1-Tailed)
Behaviour	0.997	1
Cellular Morphology	0.464	0.784
Cellular Physiology	0.339	0.543
Development	0.011	0.241
Region Tract Morphology	3.9×10^{-5}	**0.002**
Subcellular Neuronal	1.000	1

Table 3-6 - Mann-Whitney P-values ranks of p-values for a particular category against all other categories (schizophrenia cases).

3.6 Discussion

3.6.1 Polygenic risk scores in candidate pathways and their association with cognition

155 candidate pathways within 6 functional categories were tested for association with general cognitive ability using a polygenic risk score approach in a schizophrenia sample, and independently in 936 healthy controls.

In healthy controls, at the 0.05 training threshold, increased schizophrenia polygenic pathway risk for "abnormal hypothalamus morphology" was associated with lower performance IQ after correction for multiple testing. However. given the corrected p-value is only marginally significant, the size of the cognition sample is small, and as the equivalent analysis at the less stringent threshold of 0.5 showed no association, this is unlikely to represent a true association.

There was also no evidence to suggest that polygenic scores from gene-sets in particular categories were preferentially associated with performance IQ.

Schizophrenia polygenic pathway scores were not associated with the MATRICS composite score in schizophrenia cases after correction for multiple testing in either individual pathways, or across the 6 pathway categories.

These results suggest the application of polygenic risk scores to SNPs across large numbers of gene-sets is unlikely to be informative in small cognition

samples due to low power and multiple testing burdens. In addition, the findings from Chapter 2 using SNPs from across the autosomes showed schizophrenia polygenic risk only predicted a small proportion of performance IQ variance in a substantially larger cognition sample. Effect sizes for individual pathways influenced by common schizophrenia genetic risk that may also contribute to performance IQ in healthy controls are likely to be low. In addition to large schizophrenia discovery samples, large target cognition samples are also likely to be required to identify pathways contributing to schizophrenia liability and cognition.

3.6.2 GWAS of cognition in schizophrenia / Brown's method

A linear GWAS for association with the MATRICS composite score was performed in 490 schizophrenia cases. No SNP reached genome wide significance for association with the MATRICS composite score. Other large GWAS of general cognitive ability have yet to identify genome-wide associations between single polymorphisms and general cognitive ability. Thus, it is unsurprising we failed to detect any genome-wide significant loci. Although one study recently reported a genome-wide significant association at SCN2A with general cognitive ability in schizophrenia, this finding has yet to be replicated (Dickinson et al., 2014).

Brown's method was used to test for associations between SNPs in the 155 candidate pathways and general cognitive ability in schizophrenia cases. Brown's method was preferred over overrepresentation methods because

general cognitive ability is highly polygenic, with no single SNP showing genome-wide association in healthy individuals (Davis et al., 2010). Thus, we reasoned that multiple semi-independent SNPs within brain genes (which are typically larger and more likely to contain multiple semi-independent SNPs) were more likely to contribute towards general cognitive ability as opposed to only one causal SNP.

No single pathway reached nominal levels of significance before correction for multiple testing, suggesting common SNPs in these pathways are not associated with general cognitive ability in schizophrenia cases.

Brown's p-values for individual brain region/fibre tract morphology pathways were significantly higher than those in other pathway categories. The MGI definitions of region tract morphology gene-sets are based upon structural abnormalities of the mouse brain. Approximately 1653 genes are present across 44 region tract morphology gene-sets. They share on average 5 genes, thus the observed association cannot be explained by a substantial overlap in genes, which were also accounted for using permutations.

To relate these findings to existing literature, structural MRI has revealed differences in brain structure between schizophrenia cases and healthy controls. The most robust findings include decreases in grey and white matter (Olabi et al., 2011) and enlarged ventricles (Lawrie & Abukmeil, 1998) in patients. How structural changes correspond with cognitive ability may show particular brain regions that are involved with cognitive symptoms in schizophrenia (Pantelis et

al., 2005). One longitudinal study showed a reduction in white matter within the frontal lobe as associated with decreased ability on attention, verbal learning and working memory tasks (Andreasen et al., 2011). Furthermore, decreased white matter in the temporal lobes was also associated with decreased ability on attention, problem solving, fluency, verbal learning and working memory (Andreasen et al., 2011). However, this study had several limitations. First, no result would have survived correction for multiple testing. Second, structural MRI findings are subject to confounding effects from a variety of sources. Changes in grey and white matter may partly be mediated by anti-psychotic medication (Fusar-Poli et al., 2013), cannabis (Cousijn et al., 2012) and smoking (Yu et al., 2011). These confounders were not addressed in Andreasen *et al's* (2011) study.

The relationship between longitudinal brain changes and cognition in adult patients may not be particularly informative regarding the developmental aetiology of the general cognitive deficit. Specifically, cognitive impairment is observed through childhood (Meier et al., 2014) and further declines around the onset of psychosis (Wood et al., 2007). However, cognitive ability remains relatively stable thereafter, thus continuing structural changes within the brain are unlikely to be the primary cause of cognitive symptoms in adult schizophrenia cases.

Support from structural MRI for brain structure affecting cognition in schizophrenia is tentative. Whilst the results of the present study provide limited evidence that common genetic variation in genes affecting brain structure is

associated with general cognitive ability in schizophrenia, replication in a large independent schizophrenia cognition sample is required.

3.6.3 Strengths/Limitations

A strength of the present study was its use of gene-sets stratified by their functional relevance to brain structure, function, behaviour and cognition. These pathways are *a priori* more likely to be enriched for SNPs influencing general cognitive ability. By choosing a more narrow set of pathways in comparison with those of the full MGI database, or gene ontology, this reduced multiple testing burdens considerably, allowing greater power to detect associations between the 155 gene-sets and general cognitive ability. However, the results indicate we were still underpowered.

This study has several limitations. A major limitation is that of sample size, as both control and schizophrenia cognition samples were small. Furthermore, when testing pathways with SNP data using association p-values from small samples, caution is warranted over the interpretation of results. Specifically, negative results may result from poorly defined pathways or low power (Duncan et al., 2014). Further studies may benefit from analysing a smaller number of candidate pathways until samples are adequately sized and powered to detect robust associations that can withstand multiple correction burdens.

Secondly, SNPs were assigned to genes only if they were within the transcriptional start and end sites. Widening the gene window size around each

gene may capture regulatory elements that are currently missed in the present analysis. Conversely, their addition may simply add noise. The balance between including SNPs that may have functional relevance to the gene with those that do not is not easily solved. A possible solution is to use functional annotations provided by ENCODE (Encode Project Consortium, 2012) to identify SNPs with functional relatedness to gene outside of its transcriptional boundaries.

Third, no pathways were used as negative controls. However, as the analyses performed on the 155 candidate pathways did not yield robust associations with general cognitive ability in schizophrenia cases or controls across the two methods, the addition of negative control pathways would not change the interpretation of the present findings.

Fourth, set-screen and polygenic risk scores are complementary but distinct methodologies making it difficult to directly compare results. Set-based tests have some advantages over polygenic risk scoring. For example, polygenic risk scores are typically calculated using SNPs that are in low linkage disequilibrium (LD) using either LD based clumping. However, when applied to gene sets, the number of SNPs remaining in genes after clumping may be low, and thus there is insufficient variability in the polygenic risk score to show association with the phenotype under investigation. In addition, in genes where there are multiple independent signals, these may also be removed through LD-based clumping. Whilst we used an alternative approach that omitted clumping procedures (Nicodemus et al., 2014), this means results may potentially be inflated due to the inclusion of SNPs in LD. Brown's method has the advantage of using p-values

from all available SNPs within the gene set whilst accounting for the underlying LD structure (Moskvina et al., 2011).

3.7 Conclusion

This chapter used two methods to investigate the association of candidate gene-sets and their overarching biological functions, with cognitive ability in schizophrenia cases and healthy controls.

First, schizophrenia polygenic pathway scores were used to predict the MATRICS composite score in schizophrenia cases, and performance IQ in healthy individuals. No robust evidence was found associating polygenic pathway scores with general cognitive ability in patients or controls after correction for multiple testing.

Brown's test used SNPs associated with the MATRICS composite score to assess whether they collectively showed association with specific gene-sets and their respective categories. No single gene-set was shown to be significantly associated with the MATRICS composite score, however a comparison of the pathway categories showed region tract morphology pathways were ranked more highly compared to the remaining five.

Future analyses will require larger samples in order to provide greater power to detect whether the biological pathways involved with schizophrenia and general cognitive ability overlap.

4 The contribution of rare copy number variation to general cognitive ability in schizophrenia

4.1 Summary

Structural genetic variation in the form of copy number variants (CNVs) has emerged as an important contributor to many psychiatric and non-psychiatric diseases. This chapter investigates three aspects of rare CNVs with respect to general cognitive ability in schizophrenia cases.

First, do those with schizophrenia who are carriers of well- supported 'neuropsychiatric CNVs with prior association to schizophrenia, autism and intellectual disability have lower general cognitive ability?
Second, does CNV burden (as measured by size and number of genes hit) show association with general cognitive ability in those with schizophrenia?
Third, does the number of genes hit by CNVs in candidate pathways (previously outlined in Chapter 3) show association with general cognitive ability in schizophrenia cases?

We found schizophrenia cases carrying neuropsychiatric CNVs perform nearly one standard deviation lower on a measure of general cognitive ability compared to other schizophrenia cases. Regarding CNV burden, an increase in the numbers of genes hit by large (>100kb) rare CNVs was associated with lower general cognitive ability in schizophrenia cases. Finally, no evidence was observed for associations with a number of candidate pathways with links to brain structure, function, behaviour and cognition.

These results provide evidence showing rare CNVs contribute to generalised cognitive ability in schizophrenia cases, although larger samples are required to definitively assess whether disruptions in particular pathways contribute to these observations.

4.2 Introduction

4.2.1 Overview of CNVs associated with schizophrenia

Schizophrenia is a severe psychiatric disorder with a complex polygenic architecture. Common polygenic risk contributes nearly 20% of schizophrenia variance, and over 100 independent loci have surpassed genome-wide significance for association with schizophrenia (Schizophrenia Working Group of the Psychiatric Genetics Consortium, 2014). Nonetheless, even the most significant SNPs individually have small effect sizes with odds ratios for risk alleles typically below 1.2 (Schizophrenia Working Group of the Psychiatric Genetics Consortium, 2014).

Over the past decade advances in genotyping technology have allowed for the study of structural genetic variation, particularly CNVs. CNVs refer to changes in the structure of a chromosome when a section on one or both alleles are deleted, duplicated or rearranged. Approximately 12% of the human genome is comprised of chromosomal regions enriched for deletions or duplications across ethnically diverse populations (Redon et al., 2006). Whilst CNVs constitute normal genetic variation in healthy populations (Sebat et al., 2004; Redon et al., 2006). a small number of rare CNVs are high risk factors for developing

schizophrenia (Kirov et al., 2009a; Rees et al., 2014b). A recent meta-analysis confirmed 11 rare CNVs as associated with schizophrenia and that their odds ratios are between ~2-50 (Rees et al., 2014b), substantially larger than for established common variants. CNVs showing strong association with schizophrenia include deletions at 1q21.1, *NRXN1*, 3q29, 15q.11.2, 15q13.3 and 22q11.2, and duplications at 1q21.2, WBS, Angelman/Prader-Willi Syndrome on chromosome 15, 16p13.11 and 16p11.2 (Rees et al., 2014b). However, the proportion of schizophrenia cases carrying a neuropsychiatric CNV is low (~2%), with rates for individual neuropsychiatric CNVs typically between 0.1-0.5% (Rees et al., 2014b).

Identifying single or multiple genes within neuropsychiatric CNVs that contribute to pathogenesis is challenging. Neuropsychiatric CNVs are typically large, and may span several megabases incorporating a large number of genes. In comparison, a small number of CNVs associated with schizophrenia affect single genes, making biological inferences regarding pathogenesis somewhat easier.

One example is *NRXN1* (Kirov et al., 2009b; Rujescu et al., 2009), which is part of the neurexin family of genes. These genes, along with their binding partners neuroligins, organize GABAergic and glutamatergic synapses (Graf et al., 2004). Moreover, they have important roles in neurodevelopment, particularly in the expression of genes that regulate cell adhesion and neuron differentiation (Zeng et al., 2013).

CNVs encompassing *VIPR2* have also been associated with schizophrenia in several studies (Levinson et al., 2011b; Vacic et al., 2011). *VIPR2* encodes the G-

protein-coupled receptor VPAC2, which is expressed in several brain regions including the suprachiasmatic nucleus, amygdala, thalamus and pyramidal cells in the CA1-C3 layer of the hippocampus (Sheward et al., 1995). Furthermore, VPAC2 receptors may have a regulatory role on hippocampal NMDA receptors through the AMP/PAK pathway (Yang et al., 2010b), which may be important in the pathogenesis of schizophrenia (Sanderson et al., 2012). However, a recent meta-analysis showed VIPR2 was not significantly associated with schizophrenia (Rees et al., 2014b), casting doubt over its role in pathogenesis.

De-novo mutations (that are not transmitted from either parent) are found at a higher rate in schizophrenia cases compared to controls (Xu et al., 2008; Rees et al., 2011). This partly explains the stable prevalence of schizophrenia despite reduced fecundity (Rees et al., 2012). Functionally, de-novo schizophrenia CNVs are enriched for genes affecting the post-synaptic density (PSD), specifically ARC and NMDA receptors (Kirov et al., 2012).

4.2.2 CNVs conferring risk across psychiatric disorders

Some neuropsychiatric CNVs are associated with increased risk of multiple psychiatric and cognitive phenotypes (Sebat et al., 2009), in particular schizophrenia, autism and intellectual disability (Malhotra & Sebat, 2012). CNVs associated with both schizophrenia and autism include 1q21.1 (Brunetti-Pierri et al., 2008), 3q29 (Ballif et al., 2008; Levinson et al., 2011b; Levinson et al., 2011a), 16p11.2 (Weiss et al., 2008; McCarthy et al., 2009), 17q12 (Moreno-De-Luca et al., 2010) and 22q11.2 (Murphy et al., 1999; Kobrynski & Sullivan, 2007).

Deletions at 22q11.2 are one of the most extensively studied neuropsychiatric CNVs, and show the most robust association with schizophrenia (Rees et al., 2014b), and cause a wide range of phenotypes including cardiac abnormalities, autism and cognitive impairment, collectively known as 22q11.2 Deletion Syndrome (22qDS) (Kobrynski & Sullivan, 2007). Individuals with 22qDS have a substantial increase in risk for developing schizophrenia. Approximately 25% of individuals with 22qDS develop schizophrenia (Murphy et al., 1999), which is substantially higher than the general population risk of approximately 0.5-1%. Furthermore, schizophrenia cases with a 22q11.2 deletion are largely indistinguishable from non-22qDS patients (Bassett et al., 2003).

4.2.3 The association between neuropsychiatric CNVs and general cognitive ability

The overlap of neuropsychiatric CNVs across neuropsychiatric disorders with differing levels of cognitive impairment has led to questions over whether they contribute to cognitive impairment independently of or in tandem with disease (O'Donovan et al., 2008b). This question was addressed in a recent study investigating differences in cognitive ability in healthy individuals who carry a known neuropsychiatric CNV, compared to those who do not (Stefansson et al., 2014). They showed neuropsychiatric CNV carriers displayed cognitive profiles between that of schizophrenia cases and healthy individuals who did not carry a neuropsychiatric CNV. These findings were consistent across a broad range of generalized and specific cognitive domains including verbal IQ, performance IQ, working memory and processing speed. However, after correction for full scale

IQ, associations between neuropsychiatric CNV status and individual cognitive tests were weaker. This finding is unsurprising given performance on individual cognitive domains is largely mediated through generalized cognitive ability in both schizophrenia cases (Dickinson et al., 2008) and healthy individuals (Deary et al., 2010).

Neuropsychiatric CNVs affect cognition in controls, but that this has only been examined in small samples of those with clinical psychiatric disorders (Hanson et al., 2010; Citta et al., 2013; Niarchou et al., 2014). This study will be the first to specifically examine the effect of neuropsychiatric CNVs upon general cognitive ability in schizophrenia cases.

4.2.4 CNV burden and general cognitive ability in schizophrenia cases and healthy controls

Several studies have investigated rare CNV burden with respect to general cognitive ability in healthy populations. McRae and colleagues (2013) investigated CNVs greater than 20kb at a frequency of less than 5% with respect to IQ in 800 individuals. They found no associations between the number of CNVs, or their cumulative or average length, with IQ.

A substantially larger study in over 6000 individuals investigated the association between total CNV count and length (separately for deletions and duplications and both combined), and number of homozygous deletions with respect to full scale IQ (Kirkpatrick et al., 2014). Consistent with previous findings, they found

no robust evidence for any single CNV measure showing association with IQ. A trend was present showing larger CNVs were associated with lower IQ, which was largely driven by duplications rather than deletions.

The association of burden of large (>500kb) and rare (frequency < 1%) CNVs with fluid and crystallised intelligence was studied in 3133 and 3210 healthy controls respectively (MacLeod et al., 2012). They largely failed to identify associations between either fluid or crystallised intelligence and total CNV burden, length, or number of genes hit. Although a weak association between numbers of CNVs between 200-500kb and fluid intelligence was observed, this did not survive correction for multiple testing.

Other studies have investigated the association between rare CNVs and general cognitive ability in patients with a psychiatric diagnosis. Yeo and colleagues investigated whether the length and number of large rare CNVs in 74 patients with alcohol dependence were associated with performance on the Wechsler Abbreviated Scale of Intelligence (WASI) (Yeo et al., 2011). They found a negative correlation between full scale IQ and total length of CNV deletions (r=-0.3). Similar effects were also observed for the matrix reasoning task, but not for a verbal vocabulary task. In addition, they investigated whether carriers of known neuropsychiatric CNV deletions (from schizophrenia, autism and intellectual disability) performed significantly differently compared to all other carriers, but found no significant differences between them.

A recent study using 386 schizophrenia cases investigated the effect of rare CNVs hitting brain-expressed genes upon various measures of symptom domains and

cognitive tests (Merikangas et al., 2014). They found no significant associations between the number of brain-expressed genes hit, and performance, verbal or full scale IQ.

Overall, these studies provide evidence that the burden of rare CNVs (unselected for possible pathogenicity) is not associated with general cognitive ability in healthy individuals. However, rare CNV burden, particularly for deletions is associated with lower general cognitive ability in patients with schizophrenia and alcohol dependence. However, studies of rare CNV burden in clinical populations are substantially smaller compared to those in healthy individuals, and validation of these findings in large schizophrenia cohorts would be beneficial.

4.2.5 Aims & Hypotheses

The aim of this chapter was to identify associations between rare CNVs and general cognitive ability in a schizophrenia sample, and asks the following hypotheses:

1. Do schizophrenia carriers with well-supported neuropsychiatric CNVs have lower general cognitive ability compared to non-carriers?

2. Does total rare CNV burden of small (15-100kb) and large (>100kb) CNVs, their total length or number of genes hit show association with general cognitive ability? Furthermore, do different types of CNV (deletions or duplications) show the same, or differential associations?

3) Does the number of genes hit by CNVs within pathways relating to brain structure, function, behaviour and cognition show association with general cognitive ability in schizophrenia cases?

4.3 Methods

4.3.1 Schizophrenia Cognition Sample

Schizophrenia cases were recruited from the Cardiff Cognition in Schizophrenia sample (Cardiff COGS). These cases were recruited from South Wales and the wider UK, from community, in-patient and voluntary mental health settings. They underwent a comprehensive psychiatric interview including the Schedule for Clinical Assessment in Neuropsychiatry (Wing et al., 1990) as well as measures of current medication, psychotic and affective symptoms and demographic details. Clinical ratings and best estimate lifetime diagnosis (based on DSM-IV criteria (American Psychiatric Association, 1994)) were made based upon a review of interviews and case notes.

4.3.2 Cardiff COGS CNV calling / quality control

CNVs in Cardiff COGS were called as part of a larger batch including CLOZUK and controls by Professor George Kirov and Doctor Elliott Rees as described in (Kirov et al., 2014). A brief description of the genotyping and CNV calling performed by Kirov et al (2014) follows.

Genotyping of schizophrenia cases in CardiffCOGS was performed at the Broad Institute, Cambridge, Massachusetts on the HumanOmniExpressExome-8v1 (Combo array). These samples were analysed in tandem with other schizophrenia samples genotyped at the Broad Institute and controls obtained from the database of Genotypes and Phenotypes (dbGaP).

Case and control datasets were independently analysed to overcome possible batch effects. Raw intensity data from each case/control batch was processed using the Illumina Genome Studio software (v2011.1). This generated the log R ratios (LRR) and B-allele frequencies (BAF) required for CNV calling. CNVs were called using the PennCNV (http://www.openbioinformatics.org/penncnv/) algorithm following the standard protocol adjusting for GC content. As some of the cases and controls were genotyped on different Illumina arrays, CNVs were called using 520,766 probes common to all arrays used. Sample level quality control was performed using the following QC metrics generated by PennCNV: BAF drift, LRR standard deviation, wave factor and total number of CNVs called per individual. Samples were excluded from all subsequent analyses if for any one of these metrics they represented an outlier in their source dataset. Duplicate samples were checked using identity by descent analysis in PLINK (Purcell et al., 2007). Where duplicates were identified, the sample with the better QC was retained.

All CNVs underwent the following QC filtering. Firstly, raw CNVs in the same sample were joined when the distance between them was less than 50% of their combined length. CNVs were subsequently excluded if they had low probe

coverage (<10), had a length of 10kb or less, overlapped with low copy repeats by more than 50% of their length, or had a probe density (calculated by dividing the size of the CNV by the number of probes covering it) greater than 20K. CNV loci with a frequency greater than 1% were filtered out using PLINK. In total, rare CNVs were identified in 430 cases in Cardiff COGS.

The remaining rare CNVs underwent additional QC using an *in silico* median Z-score outlier method (Kirov et al., 2012). This method standardises SNP probe intensities per individual across all probes, then standardises the intensity of each probe across all individuals. These QC stages help to reduce noise caused by natural fluctuations in probe intensity. The median Z-score for standardized probe intensities within a potential CNV region is used to identify real deletions and duplications, which are represented as outliers in the median Z-score distribution. Z-score histograms of CNVs with extreme Z-Scores (between -6 and -4, 2 and 4) were inspected manually using Illumina GenomeStudio v2011.1 software.

4.3.3 Cognitive Phenotype

Patients completed the MATRICS cognitive battery (Kern et al., 2008; Nuechterlein et al., 2008), which measures ability on 10 separate tests encompassing 7 cognitive domains (attention/vigilance, reasoning/problem solving, speed of processing, social cognition, verbal learning/memory, visual learning/memory and working memory). Table 4.1 contains correlations between the individual MATRICS domains and the MATRICS composite score.

MATRICS Domain	Correlation with MATRICS Composite Score	P
Composite	1	N/A
Attention/Vigilance	0.77	<0.01
Reasoning/Problem Solving	0.72	<0.01
Social Cognition	0.55	<0.01
Speed of Processing	0.84	<0.01
Verbal Learning	0.80	<0.01
Visual Learning	0.80	<0.01
Working Memory	0.84	<0.01

Table 4-1 - Correlations between MATRICS composite score and individual domains

In total, 483 individuals had complete data for the MATRICS composite score and CNV call information.

4.3.4 Candidate pathways

The 155 pathways used in this analysis were previously outlined in Chapter 3, with definitions of all gene-sets available in Appendix B. These pathways are a priori more likely to be enriched for genes influencing general cognitive ability. For each pathway, the translational start and end of genes were identified using hg19 (NCBI Build 37) human reference gene coordinates. All pseudo-genes were removed. CNVs were classed as hitting a gene if at least one base pair overlapped

between them. For each proband, the total number of genes hit was calculated by summing the number of genes hit for each CNV.

4.3.5 Analyses

4.3.5.1 Neuropsychiatric CNVs and their association with general cognitive ability in schizophrenia cases

CNVs were denoted as neuropsychiatric if they showed prior association with schizophrenia, autism, intellectual disability or other congenital malformations (Kaminsky et al., 2011; Girirajan et al., 2012) (see Appendix F for list of neuropsychiatric CNVs considered pathogenic).

Analyses were performed using a linear regression model. The MATRICS composite score was regressed against a binary variable that denoted whether a proband had a known neuropsychiatric CNV (n=11) or not (n=472). Individuals were included in the non-neuropsychiatric CNV group if they either had no CNVs that passed QC, or had no neuropsychiatric CNVs called. Age was included as a covariate within the regression model. A permutation approach was used to generate an empirical p-value. The MATRICS composite score was permuted 10,000 times, and simulated regression p-values for neuropsychiatric CNV status were compared against the observed p-value.

4.3.5.2 Effect of CNV length and number of genes hit on general cognitive ability

To assess the contribution of small (between 15-100kb) and large (>100kb) CNVs upon the MATRICS composite score, a number of nested linear regression models were performed in R :

Model 1: MATRICS composite score ~ Age

Model 2: MATRICS composite score ~ Total Genes Hit (>100kb) + Age

Model 3: MATRICS composite score ~ Total CNV Length (>100kb) + Total Genes Hit (>100kb) + Age

Model 4: MATRICS composite score ~ Total Genes Hit (15-100kb) + Age

Model 5: MATRICS composite score ~ Total CNV Length (15-100kb) + Total Genes Hit (15-100kb) + Age

Model 6: MATRICS composite score ~ Total CNV Length (>100kb) + Total Genes Hit (>100kb) + Total CNV Length (between 15-100kb) + Total Genes Hit (between 15-100kb) + Age

Models 2 and 4 show whether the addition of number of genes hit by large and small CNVs respectively are associated with the MATRICS composite score. Models 3 and 5 show whether total CNV length of large and small CNVs is also associated with the MATRICS composite score in addition to number of genes hit. This shows whether the addition of total CNV length explains additional variance of the MATRICS composite score in addition to the other variables. Model 6 uses CNV burden terms from both large and small CNVs.

4.3.5.3 *Genes hit by CNVs in Candidate Pathways and their Association with the MATRICS composite score*

Two linear models were used to test the association between genes hit by CNVs in candidate pathways and general cognitive ability:

Model 7: Composite ~ Age + Total CNV Length + Number of genes hit (not in pathway)

Model 8: Composite ~ Total CNV Length + Age + Number of genes hit (not in pathway) + Number of genes hit (in pathway)

An ANOVA comparison of these two models informs whether the addition of number of pathway genes significantly improves the model fit.

A permutation approach was used to correct for the number of pathways tested. First, the MATRICS composite score was permuted across all individuals. ANOVA p-values comparing models 7 and 8 were obtained for each pathway, and the lowest p-value was retained. This was permuted 10,000 times. The corrected p-value represents the number of times the original ANOVA p-value for a pathway was less than the lowest permuted p-value across all pathways.

4.4 Results

4.4.1 *Known neuropsychiatric CNVs and their association with general cognitive ability*

12 schizophrenia cases in Cardiff COGS had at least one neuropsychiatric CNV across 10 loci (Table 4.2).

N Cases	Loci	Chromosome	Start	End	Type	CNV Type
1	1q21.1	1	146501348	147820342	Duplication	P
1	NRXN1	2	50793781	50924358	Deletion	P/SCZ*
1	WBS	7	72722981	74138603	Duplication	P/SCZ*
2	15q11.2	15	22750305	23226254	Deletion	P/SCZ*
1	15q11.2	15	22678398	23226254	Deletion	P/SCZ*
1	16p11.2 distal	16	28825605	29028905	Duplication	P
1	16p11.2	16	29672982	30192561	Duplication	P/SCZ*
1	16p12.1	16	21949122	22409463	Deletion	P
1	17q12	17	34815551	36223325	Duplication	P
2	22q11.21	22	18896464	21463730	Deletion	P/SCZ*

Table -2 - Rare CNVs in Cardiff COGS with prior association to schizophrenia/autism/intellectual disability. P=Pathogenic (Kamisky et al,2011; Girirajan et al, 2012), SCZ=Schizophrenia CNV (Rees et al, 2014)

One neuropsychiatric CNV carrier with a duplication at 16p11.2 was unable to complete the majority of the MATRICS tests, meaning a composite score could not be calculated in this individual. MATRICS composite scores were available for 11 probands with a neuropsychiatric CNV, and 472 patients with no known neuropsychiatric CNVs. A distribution of the composite z-scores in cases with and without a neuropsychiatric CNV can be found in Figure 4.1.

The z-score difference between the composite score of cases in neuropsychiatric and non-neuropsychiatric CNV groups was 0.839 standard deviations, with the neuropsychiatric CNV group displaying lower cognitive performance. The regression model showed neuropsychiatric CNV status was a significant term (P_{Emp}=0.0174), and explained 1% of the MATRICS composite score variance.

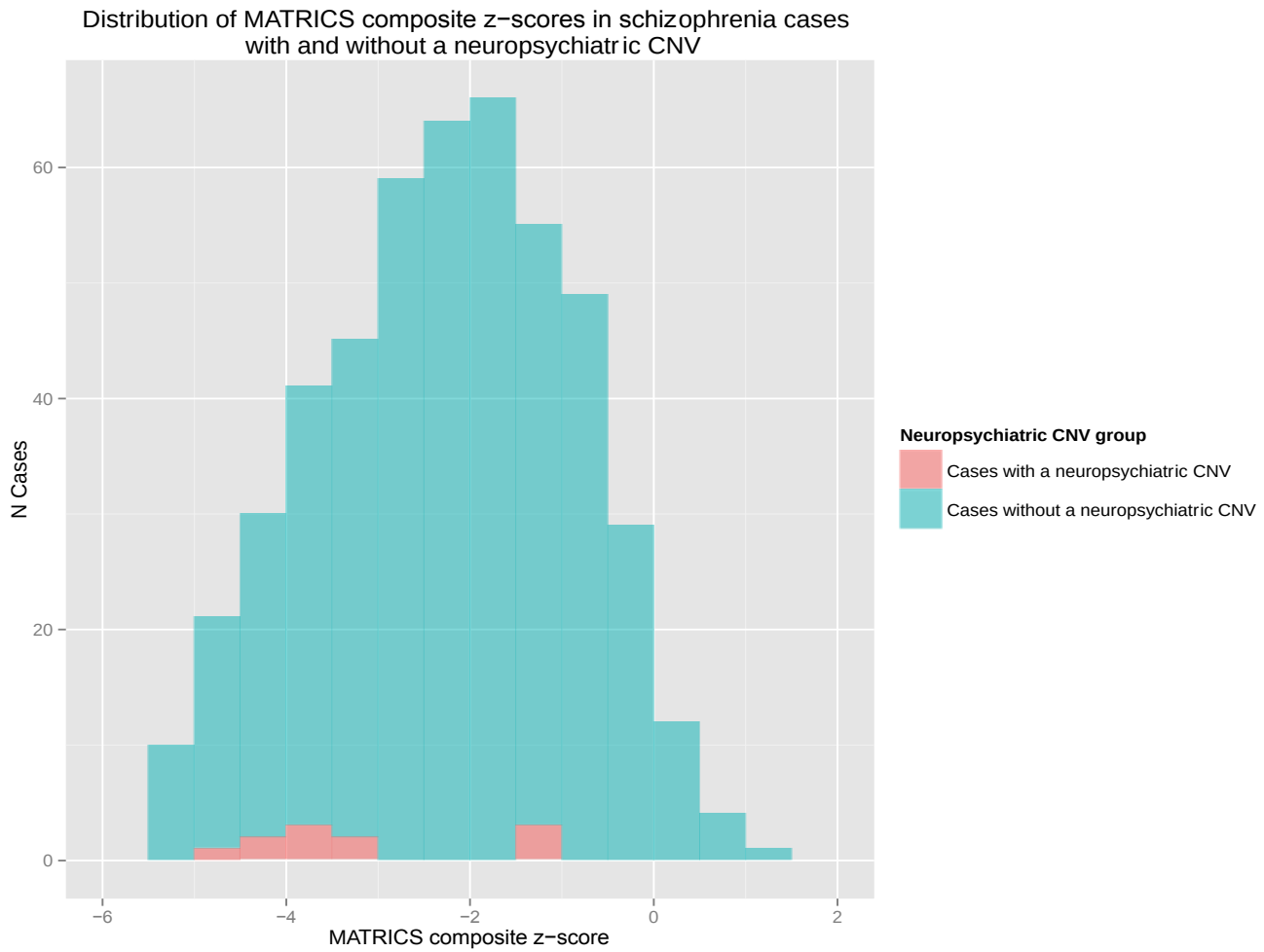

Figure 4.1 - Distribution of MATRICS composite z-scores in schizophrenia cases with and without a neuropsychiatric CNV

4.4.2 *Effect of CNV length and number of genes hit on general cognitive ability*

Summary information regarding the total length of rare CNVs can be found in table 4.2.

	15-100kb	*> 100kb*
N Individuals	165	281
Median Total Length (kb)	68.3	2998
Min/Max Total Length	15.6/ 281.7	102/10222.1

Table 4-3 - Summary information for individuals with rare CNV in Cardiff COGS

Several models were used to assess the contribution of total CNV length and number of genes hit in respect to general cognitive ability. The first model was used as a baseline, simply investigating the effect of age upon general cognitive ability. Age was significantly associated with the MATRICS composite score (p=3.78E-11) and explained 10.6% of the variance.

Models 2 and 3 investigated whether number of genes hit and total CNV length in large (>100kb) CNVs were associated with the MATRICS composite score separately. In model 2, an increased number of genes hit was associated with a lower composite score (p=0.012), explaining an additional 1.5% of the variance in addition to age. Model 3 added a term for total CNV length. The total number of genes hit remained significant (p=0.032), however total length of the CNV was not associated with composite performance (p=0.305).

Models 4 and 5 investigated whether number of genes hit and total CNV length for small (15-100kb) CNVs were associated with the MATRICS composite score separately. Model 4 showed the number of genes hit by small CNVs was not significantly predictive of the composite score (p=0.245). Model 5 used an additional term of CNV length. Neither number of genes hit (p=0.499) or total length of small CNVs (p=0.695) were associated with the MATRICS composite score.

Finally, model 6 used terms for total genes hit and total CNV length for both small and large CNVs together. Other than age, the total number of genes hit by CNVs greater than 100kb was the only significant term in the model (p=0.029).

Further analyses were performed separately for CNV deletions and duplications using the equivalent models described above.

4.4.2.1 Deletions

Summary information regarding the total length of rare CNV deletions can be found in table 4.3.

	15-100kb	*> 100kb*
N Individuals	92	104
Median Total Length (kb)	55.2	179.1
Min/Max Total Length (kb)	15.6/237.8	102.1/6022.2

Table 4-4 - Summary information for individuals with CNV deletions in Cardiff COGS

Models 2 and 3 investigated whether number of genes hit and total CNV length in large (>100kb) CNV deletions were associated with the MATRICS composite

score. In model 2, an increased number of genes hit by deletions were associated with a lower composite score (p=0.037), explaining an additional 2.5% of the variance in addition to age. Model 3 added the additional variable CNV length. The total number of genes hit in model 3 became less strongly associated with the addition of CNV length in the model (p=0.065). Total length of large CNV deletions was not associated with cognition (p=0.466).

Models 4 and 5 investigated whether number of genes hit by deletions and total length of small (15-100kb) CNV deletions were associated with the MATRICS composite score separately. Model 4 showed the number of genes hit by small CNVs was not significantly predictive of the composite score (p=0.453). Model 5 used an additional term of CNV length. Neither number of genes hit (p=0.352) nor total length of small CNV deletions (p=0.579) were associated with the MATRICS composite score.

Model 6 used terms for total genes hit and total CNV length for both small and large CNV deletions together. The total number of genes hit by CNV deletions greater than 100kb was associated at a trend level (p=0.059), but no other measure of CNV burden was associated with the MATRICS composite score.

4.4.2.2 Duplications

Summary information regarding the total length of rare CNV duplications can be found in table 4.4.

	15-100kb	*> 100kb*
N Individuals	73	177
Median Total Length (kb)	54	315.1

Min/Max Total Length (kb)	17/129	104.2/10222.1

Table 4-5 - Summary information for individuals with CNV duplications in Cardiff COGS

Models 2 and 3 investigated whether number of genes hit and total CNV length in large (>100kb) CNV duplications were associated with the MATRICS composite score separately. In model 2, an increased number of genes hit by duplications were associated at a trend level with a lower composite score (p=0.063), explaining an additional 1.6% of the variance in addition to age. Model 3 added total CNV length to the model. Neither total number of genes hit (p=0.128) nor total length of large CNV duplications were significantly associated with the MATRICS composite score (p=0.450).

Models 4 and 5 investigated whether number of genes hit and total CNV length in small (15-100kb) CNV duplications were associated with the MATRICS composite score. Model 4 showed the number of genes hit by small CNVs was not significantly predictive of the composite score (p=0.774). Model 5 used an additional term of CNV length. Neither number of genes hit (p=0.184) or total length of small CNVs (p=0.131) were associated with the MATRICS composite score.

4.4.3 CNV Pathway Results

Pathway analyses were restricted to genes hit by large CNVs greater than 100kb, as these previously showed the greatest association with the MATRICS composite score. Linear regression models were used to assess whether the addition of total number of genes hit in each pathway significantly improved

model fit for association with the MATRICS composite score, compared to a model with age, total CNV length and total number of genes hit (not in pathway).

85 pathways contained at least one hit by a rare CNV. An ANOVA was performed comparing models with and without the inclusion of number of pathway genes hit. No significant improvement over the baseline model was observed for any pathway, either before or after correction for multiple testing (complete results can be found in Appendix G).

4.5 Discussion

4.5.1 Known neuropsychiatric CNVs and their association with general cognitive ability in schizophrenia cases

Individuals carrying rare CNVs at a small number of well-supported loci have a substantially increased risk of developing schizophrenia, autism and intellectual disability (Malhotra & Sebat, 2012; Rees et al., 2014b). In addition, these risk CNVs also contribute to cognitive ability in apparently healthy individuals (Stefansson et al., 2014). However, no previous study has investigated whether schizophrenia cases carrying a neuropsychiatric CNV have greater deficits in general cognitive ability compared to other schizophrenia cases.

The first analysis in this chapter investigated the MATRICS composite score, a measure of general cognitive ability, with respect to whether a patient carried a neuropsychiatric CNV or not. We showed that schizophrenia cases with risk

CNVs have a considerable decreases in general cognitive ability, approximately one standard deviation below other schizophrenia cases. Furthermore, one proband with a CNV at 16p11.2 was not included in the analysis because they were too cognitively impaired to complete the MATRICS assessment. Although this is only one individual, it nonetheless corroborates the finding that schizophrenia cases with well-supported risk CNVs represent a subgroup with significant cognitive impairment. However, we recognise the number of individuals with neuropsychiatric CNVs is low (n=11), and replication in an independent sample is required to provide greater support to these findings.

There are several possible interpretations of these findings. First, biological pathways underlying general cognitive ability are "hit" twice in patients with schizophrenia, once through the effects of neuropsychiatric CNVs, and secondly via disease processes relating to schizophrenia. However, this is speculative as it is currently unknown whether genes in these CNVs contribute separately or in tandem regarding risk of psychiatric illness and general cognitive ability.

A second interpretation is that neuropsychiatric CNVs may increase the disease severity of schizophrenia, which in turn is associated with lower general cognitive ability in patients. There is limited evidence suggesting neuropsychiatric CNVs are associated with early onset schizophrenia (Ahn et al., 2014), and carriers of some neuropsychiatric CNVs such as 22q11.2 deletions report a higher frequency of psychotic experiences in childhood and adolescence (Baker & Skuse, 2005). Early onset of disease is predictive of greater disease severity and cognitive impairment (Haas & Sweeney, 1992; Frangou, 2010).

However, adding psychosis age of onset as a covariate within the model did not change the association between the MATRICS composite score and neuropsychiatric CNV status.

However, not all patients with risk CNVs performed poorly. One individual with a duplication at 17q.12 performed approximately 1 standard deviation below controls, showing less impairment relative to other schizophrenia cases within the sample. Duplications at 17q.12 are associated with a range of phenotypes including autism, developmental delay, seizures and other non-psychiatric conditions including diabetes and problems with renal functioning (Bierhals et al., 2013). Conversely, a small proportion of patients with 17q.12 duplications have no discernable neuropsychiatric abnormalities (Faguer et al., 2011). Taken together with our findings this suggests that duplications at this locus may not be fully penetrant for cognitive impairment, although the reasons for this are currently unclear.

Additional caution is warranted regarding the assumption that carriers of the same neuropsychiatric CNV have similar cognitive abilities. One example relates to the general cognitive ability of three probands with deletions at 15q11.2. One carrier was 1.3 standard deviations below the control mean, with two other carriers performing over 3.5 standard deviations lower. Interestingly, two cases shared the same CNV breakpoints at 15q11.2, with one showing substantial impairment, whilst the other was over two standard deviations higher. It is to be expected that environmental or other genetic factors contribute to differences in cognitive ability beyond the effects of neuropsychiatric CNVs and these factors

likely explain the observed variation in cognition between individuals with the same CNV.

The distribution of composite scores in patients with and without neuropsychiatric also CNVs overlapped. This suggests patients with neuropsychiatric CNVs do not represent a distinct subgroup based upon general cognitive ability alone.

Whilst understanding the neurobiology of CNVs and their functional relevance to disease is important, these findings may also be important for clinical settings. If patients with pathogenic CNVs can be identified through directed testing, as suggested by Rees *et al* (2014), these individuals may particularly benefit from intervention strategies or future pharmacological treatments that target cognitive symptoms.

4.5.2 Rare CNV burden, candidate pathways and their associations with general cognitive ability in schizophrenia cases

Measures of CNV burden were also investigated with respect to general cognitive ability in schizophrenia cases. CNVs were classified according to whether they were small (between 15-100kb) or large (>100kb). Using a series of nested linear regression models, the cumulative length and the number of genes hit were used as predictors of the MATRICS composite score. Lower general

cognitive ability was associated with a greater number of genes hit by large CNVs, even after covarying for total CNV length.

The correlation between the length and number of genes hit in CNVs greater than 100kb was high (r=0.855), showing these measures of CNV burden are strongly related. However, the number of genes hit is a more direct measure of potential biological impact, and CNV length may incorporate more noise by including regions with little or no biological consequence.

Similar, but weaker trends emerge when looking at deletions and duplications separately. An increase in the number of genes hit by large deletions was associated at a trend level after the addition of total CNV length, and was marginally stronger at predicting the MATRICS composite score compared to duplications. Conversely, neither length of large deletions or duplications was significant predictors of the composite score, mirroring the findings of the combined analysis. In addition, neither length nor number of genes hit by small deletions was associated with general cognitive ability.

When comparing these findings with similar studies a number of differences emerge. Firstly, Yeo and colleagues (2011) found the cumulative length of deleterious CNVs were predictive of IQ in 74 patients with alcohol dependence. In addition, Martin and colleagues (2014) reported the cumulative length of CNV deletions were associated with lower full scale IQ, which is more strongly driven verbal, rather than performance IQ in 78 schizophrenia cases. Finally, Yeo et al (2013) found total number of CNV deletions was a better predictor of general cognitive ability compared to CNV length in 79 schizophrenia cases.

Although these studies differed regarding their choice of measure of CNV burden, the direction of effect pertaining to their association with cognitive ability was consistent. However, these studies are in small numbers of schizophrenia cases and have not been universally replicated. One study found no association between number of CNVs, or number of genes hit in deletions, duplications or combined with IQ in 350 schizophrenia cases and 322 healthy controls (van Scheltinga et al., 2013). Furthermore, no interaction was observed between disease status and IQ for any CNV burden variable. There are several differences between this study and previous studies. First van Scheltinga and colleagues assigned to a CNV if they were within a 50kb border. In addition, they used CNVs called from their previous study (Buizer-Voskamp et al., 2011) using both common and rare CNVs. van Scheltinga *et al* (2013) did not provide details of a CNV population frequency cut off. If they included common CNVs that are not considered to contribute substantially to disease (Conrad et al., 2010) this may have masked associations with cognition from rare CNVs.

The functional impact of CNV burden is poorly defined, both regarding neural correlates, and biological mechanisms underlying cognitive ability. In part, this is because burden refers to the "load", rather than specific genes. This means it will be difficult to extract biological meaning because the functional effects of CNV burden may relate to completely independent biological processes in different individuals.

However, a small number of studies have investigated CNV burden with respect to association with structural brain changes. One study in 79 schizophrenia cases and 110 controls found increases in the number of rare (frequency < 3%) deletions were associated with larger ventricular volume in schizophrenia cases, however no such effect was observed in controls (Yeo et al., 2013).

A separate study has shown deletion burden is associated with numerous neuroanatomical changes (Martin et al., 2014). Increased grey matter in the striatum and superior temporal gyrus, and increased white matter in the corpus callosum were associated with larger cumulative length of rare CNVs (Martin et al., 2014). In addition, decreased functional connectivity between the prefrontal and dorsolateral prefrontal cortex (DLPFC), DLPFC and putamen, and DLPFC and the associative visual cortex was associated with increased cumulative deletion length. However, whilst this study did correct for multiple comparisons, it was nonetheless in relatively modest schizophrenia sample size (n=33). Nonetheless, disruption of prefrontal networks affects multiple cognitive phenotypes in schizophrenia including working memory (Wheeler et al., 2014), speed of processing (Woodward et al., 2013) and social cognition (Ursu et al., 2011). These findings may provide

However, other studies have reported no associations between whole brain volume, grey or white matter volume and CNV burden measured by the number of genes hit by deletions, duplications or combined in patients and controls (Terwisscha van Scheltinga et al., 2012), and the literature remains inconclusive.

There is limited evidence as to the effect of specific neuropsychiatric CNVs on brain structure. Steffanson and colleagues (2014) used structural magnetic resonance imaging to investigate differences between healthy carriers of 15q11.2 deletions (n=15) or duplications (n=55) and non-carrier controls (n=201). They found changes in the grey matter volume in the perigenual anterior cingulate cortex and the left insula. Interestingly, individuals who carry the deletion have lower grey matter volume compared to controls, whilst the duplication is associated with a reciprocal increase in grey matter volume of the same order of magnitude. These reciprocal differences extend to changes for white matter volume of temporal lobes, whereby carriers of deletions have lower densities. Finally, carriers of the duplication have decreased white matter volume in the corpus colosum, whereas carriers of deletions have increased white mater volumes. Although these results are from a single study, and for one CNV locus, the opposite effects on brain volume metrics observed are intriguing and raise confidence that they may point to biological insights that are applicable to other CNVs.

Unfortunately, the results from pathway analyses in the present, and previous studies have been unable to identify specific biological processes. We found no evidence of association between general cognitive ability and the number of genes hit in 85 candidate pathways. Furthermore, a previous study failed to identify an association between brain-enriched genes hit by CNV deletions and performance/verbal and full scale IQ in 386 schizophrenia cases (Merikangas et al., 2014). However, both our study and that of Merikangas and colleagues

(Merikangas et al., 2014) were conducted in relatively small samples, and likely to be underpowered for such pathway analyses.

4.5.3 Limitations

A first limitation is that of sample size. To date, studies investigating associations between CNV burden and cognition in schizophrenia cases have typically used relatively small samples. Whilst the present study was conducted in one of the largest schizophrenic case samples with cognitive data, it nonetheless remains a relatively small sample investigating rare events.

A second potential limitation is that CNVs were required to hit at least one base pair of the gene. Given that CNVs do not need to physically hit the gene to result in changes to gene expression, it is possible that CNVs hitting promoter regions that affect transcription, but do not hit the gene directly, may also contribute to gene expression and affect general cognitive ability.

4.5.4 Conclusion

We investigated the association between rare CNVs and general cognitive ability in a schizophrenia case sample. We found cases with a known pathogenic CNV performed approximately 1 standard deviation below other cases on the MATRICS composite score, a measure of generalised cognitive ability. In addition, we identified that the number of genes hit by large (>100kb), rare (<1%) CNVs were associated with lower general cognitive ability. We failed to

detect specific associations with CNV deletions identified in previous studies, although the number of genes hit by deletions approached nominal levels of significance. Finally, there was no evidence of association between the number of genes hit in candidate pathways and general cognitive ability. These findings are consistent with previous research investigating associations between CNVs hitting brain-expressed genes and general cognitive ability. However, given the small samples size across both studies, the lack of association with particular pathways likely reflects a lack of power.

Taken together, these findings suggest rare CNVs contribute to the generalised cognitive deficit in schizophrenia, although larger samples are required to identify the enrichment of gene-sets that could become targets for future pharmacological intervention to improve general cognitive ability in patients.

5 General Discussion

5.1 Overview

Patients with schizophrenia have substantial impairments in cognitive functioning (Heinrichs & Zakzanis, 1998), a substantial proportion of which is explained by a generalised cognitive factor (Dickinson et al., 2008). Importantly, current treatments do not significantly impact these cognitive impairments. Thus, the development of pharmacological targets that improve cognitive functioning in patients is seen as being a priority (Buchanan et al., 2005). Research examining the genetic relationships between schizophrenia and cognitive ability, and therefore provide insight about the underlying biology, offers potential to inform therapeutic development. The purpose of this thesis was to examine the association between both common and rare schizophrenia genetic variation and cognitive ability in schizophrenia cases and healthy controls. This chapter will present a summary of the rationale, main findings and interpretation of the three experimental chapters, as well as suggestions for further work and conclusions.

5.2 Polygenic risk of schizophrenia and its association with cognitive ability in healthy individuals

The studies presented in Chapter 2 were motivated by the hypothesis that there is common underlying biology to schizophrenia and cognition in the population, hence increased polygenic risk of schizophrenia would be associated with lower general cognitive ability in healthy individuals. During the course of the thesis

other groups have confirmed the polygenic overlap between schizophrenia and general cognitive ability (Lencz et al., 2014) but we sought to expand on these findings by using up to date and more powerful schizophrenia datasets, examining the question in relation to more specific domains of cognition and in more homogeneous populations. As well as pointing to shared biological underpinnings, within this context findings can provide evidence about whether specific traits, such as cognitive domains are suitable endophenotypic candidates for psychiatric disorders such as schizophrenia and bipolar disorder.

Chapter 2 consisted of three related studies. The first investigated whether polygenic risk of schizophrenia predicted performance on tests of cognition in a well-characterised and homogeneous German sample and separately in a developmental study cohort of children aged 8. The cognitive tests/domains examined were those affected in schizophrenia including attention, processing speed, social cognition, reasoning/problem solving, verbal learning and working memory, as well as performance, verbal and full scale IQ.

No study to date has investigated the association between polygenic risk of bipolar disorder and cognitive ability. The second study investigated whether polygenic risk of bipolar disorder predicted performance on the same cognitive domains and IQ scores described above. This approach has the potential to serve as a comparison for the schizophrenia/cognition polygenic analysis as well as further elucidate the nature of genetic overlap between bipolar disorder and cognition.

Finally, individuals with schizophrenia and bipolar disorder show differences in cognitive impairment. Differences in their common genetic architecture may therefore contribute to cognitive ability. Using GWAS summary statistics from a schizophrenia (case) versus bipolar disorder (control) analysis, the third study investigated whether polygenic risk scores (where higher scores represent increased polygenic risk of schizophrenia relative to bipolar disorder) were associated with the cognitive domains and IQ scores described above.

5.2.1 Schizophrenia polygenic risk and their associations with cognitive ability

Three schizophrenia discovery samples of increasing size were used; CLOZUK, PGC1 and PGC2 (Schizophrenia PGC, 2011; Schizophrenia Working Group of the Psychiatric Genetics Consortium, 2014). Two independent cognition samples were available. First, approximately 5500 children with cognitive data at age 8 were used from ALSPAC a large longitudinal cohort comprised of children born in the 1990s in Avon (Golding et al., 2001),. A second independent sample comprised of 936 healthy adults of German descent was also used (Rujescu et al., 2003).

In ALSPAC, increased polygenic risk of schizophrenia was significantly associated with lower performance IQ when training across all schizophrenia discovery datasets at multiple training thresholds. The best powered analysis used the schizophrenia PGC2 discovery sample and ALSPAC as the target cognition sample. Polygenic risk of schizophrenia explained approximately

0.34% of the variance for performance IQ, the strongest associated cognition measure. Using the schizophrenia PGC1 discovery set, increased schizophrenia polygenic risk was associated with lower performance IQ in an independent adult German sample, explaining 0.56% of performance IQ variance.

The relationship between schizophrenia genetic risk and performance IQ is supported from other analyses within our group (GCTA and polygenic analyses were completed by myself and Katherine Tansey and co-led by James Walters and Stan Zammit). We showed polygenic performance IQ scores are the strongest predictor of schizophrenia liability out of all the cognitive measures previously described (Hubbard et al, submitted). Furthermore, bivariate GCTA analyses show performance IQ has the strongest genetic correlation with schizophrenia (r_g=-0.38), more than verbal (r_g=-0.07) and full scale (r_g=-0.20) IQ. These findings are in keeping with phenotypic associations between performance IQ and schizophrenia. A meta-analysis has shown effect sizes deficits are larger for performance IQ (d=1.46) compared to verbal (d=0.98) and full scale IQ (d=1.24) (Heinrichs & Zakzanis, 1998) and since been replicated in a further meta-analysis (Dickinson et al., 2013b). Overall, these results indicate performance IQ may be a useful endophenotype for future studies examining the genetic aetiology of schizophrenia.

Differences in the amount of variance explained in performance IQ by schizophrenia polygenic risk in ALSPAC (0.34%) and German (0.56%) samples were observed. These differences could be attributable to differing sample characteristics. Cognition was measured at different ages within the two

samples; in ALSPAC cognition was measured at age 8, whilst individuals within the German samples were aged between 20-74, with a mean age of 52 years. Schizophrenia polygenic risk is associated with change across the lifespan in healthy controls (McIntosh et al., 2013), and heritability estimates show cognitive ability is subject to increasingly stronger genetic influence across the lifespan (Deary et al., 2012). The greater proportion of variance of performance IQ explained in the German sample may therefore be attributable to age related effects. However, an alternative explanation could be that a smaller schizophrenia discovery sample (PGC1) was used, but perhaps more importantly the target German sample was nearly six times smaller compared to that of ALSPAC. Whilst the size of the discovery sample is important for power, inadequately sized target samples are unlikely to affect the accuracy of the estimation of polygenic overlap. Thus the smaller German target sample will lead to a less accurate estimate of the polygenic signal and therefore the observed result could be a chance finding. In this respect it would be helpful to calculate confidence intervals for these estimates and there are emerging approaches to enable such calculations, which we will be pursuing in future work (Dudbridge, 2013).

Comparing the present findings with previous studies, some important differences are observed. We failed to fully replicate findings from previous studies investigating polygenic risk of schizophrenia and cognition (Lencz et al., 2014). The proportion of variance of total IQ explained by schizophrenia polygenic risk in our study ($r^2 \sim 0.1\%$) is lower than that reported by Lencz and colleagues ($\sim 0.5\%$ to 1%), although they measured general cognitive ability

through “g”, rather than full scale IQ. Whilst “g” and full scale IQ are both measures of general cognitive ability, they are derived using different methods, which could partially explain the differences between the two studies. Lencz and colleagues also used a heterogeneous cognition sample comprised of multiple nationalities. Furthermore, they used PGC1 as a discovery dataset, which contains substantially less individuals compared to PGC2. Better estimates of SNP effect sizes in PGC2 may have resulted in a more accurate estimation of polygenic overlap between schizophrenia and cognition, and may indicate a more modest overlap than previously stated (Lencz et al., 2014)

The negative findings from this analysis are also informative, with schizophrenia polygenic risk showing no association with problem solving, social cognition, verbal learning, working memory and verbal IQ even at trend levels. These findings are perhaps surprising given that tests measuring cognitive domains of attention, working memory and verbal memory are moderately heritable in families containing at least one schizophrenia proband (Greenwood et al., 2007). Other genetic factors (rare variants not tagged by polygene score or unidentified factors), gene-gene or gene-environmental interactions may therefore contribute to performance on specific cognitive domains.

5.2.2 Polygenic risk of bipolar disorder and their associations with cognitive ability

The second analysis in chapter 2 investigated associations between polygenic risk of bipolar disorder and cognitive ability within ALSPAC. No robust associations were identified with any cognitive domain or IQ score, although increased polygenic risk of bipolar disorder was associated with better social cognition at the 0.0001 training threshold ($p=0.039$, $\%r^2=0.08$), with trends emerging ($P<0.07$) in the same direction of effect for three other training thresholds. A weak association was observed between bipolar polygenic risk and processing speed at a training threshold of 0.1 ($p=0.012$, $\%r^2=0.11$), however this likely represents a false positive association given none of the other training thresholds were nominally significant.

This indicates polygenic risk of bipolar disorder is not associated with cognitive ability in healthy individuals. However, several points of consideration are warranted. Cognitive tests in ALSPAC were selected a-priori based upon their similarity with the MATRICS domains (Niarchou et al., 2013). These domains show the greatest impairment in schizophrenia cases (Kern et al., 2007) and may not accurately reflect cognitive deficits in bipolar disorder. Nevertheless, several studies have investigated the cognitive performance of bipolar cases using the MATRICS cognitive battery. Bipolar cases show impairment across the majority of the MATRICS domains, albeit with lower effect sizes when compared with schizophrenia cases (Burdick et al., 2011).

Other research has indicated bipolar risk is associated with both increased and decreased cognitive ability (MacCabe et al., 2010). They investigated the association between school performance at age 16 and risk of bipolar disorder in

a large population based study in Sweden. They showed superior school performance was associated with the great risk of developing bipolar disorder, although individuals with poorer performance were also at increased risk.

The findings from these studies are indicative that cognitive phenotypes in bipolar disorder are heterogeneous, and bipolar risk is associated with both increases and decreases in cognitive performance. These factors add a level of complexity regarding bipolar genetic risk and their associations with cognitive ability, with implications for the present study. Specifically, if increased polygenic risk of bipolar disorder is associated with both higher and lower cognitive ability for specific cognitive domains or IQ measures, using standard linear regressions may be inadequate for capturing non-linear associations and these effects would serve to nullify any associations present when conducting the current analyses.

5.2.3 Polygenic risk of schizophrenia versus bipolar disorder and their associations with cognitive ability

The third analysis used summary statistics from a schizophrenia (case) / bipolar (control) GWAS (Ruderfer et al., 2014) to predict cognition in ALSPAC. Increased polygenic risk of schizophrenia relative to bipolar disorder was associated with lower verbal and full scale IQ, explaining approximately 0.27% of the variance for both tests.

One interpretation of these results relates to differences in premorbid cognitive between schizophrenia and bipolar cases. Zammit *et al* (2004) showed low premorbid IQ was associated with an increased risk of developing schizophrenia, whereas no association was observed in individuals who developed bipolar disorder. These findings are also substantiated through meta-analyses of premorbid impairment in schizophrenia and bipolar disorder. Schizophrenia cases perform approximately half a standard deviation compared to controls relating to premorbid IQ (Woodberry et al., 2008), whereas no differences are found between bipolar cases and controls (Bora et al., 2009). Given that cognitive deficits in schizophrenia can be detected as early as age 4 (Cannon et al., 2000), these results are suggestive of a neurodevelopmental origin for cognitive defects in schizophrenia, but not bipolar disorder. Within the context of the present study, differences in common genetic variation between schizophrenia and bipolar disorder may be related to neurodevelopmental processes underlying general cognitive ability.

5.2.4 Methodological considerations of polygenic analyses

The studies in this chapter had several strengths. First, we used the largest schizophrenia training set available to date, providing greater power than previous analyses. Furthermore, ALSPAC represents the single largest population sample to date. All individuals completed the same cognitive tests at the same age within a fairly homogenous geographical area, potentially reducing effects of population stratification.

The bipolar case control, and schizophrenia (case) bipolar (control) discovery datasets were both modest in size in comparison with the most recent schizophrenia GWAS. In the bipolar analysis, increased polygenic risk of bipolar disorder was weakly associated with better social cognition, and to a lesser extent working memory. A larger bipolar training set with a comparable size to schizophrenia PGC2 may provide greater clarity regarding possible associations between polygenic risk of bipolar disorder and social cognition.

Secondly, cognitive variables within ALSPAC were measured in children aged 8. The availability of data restricted the current analyses to a single cross sectional time point, and the use of repeated measures of cognition would help minimize measurement error, although the fact we used cognitive data measured at equivalent ages can also be considered a strength.

Third, the individual cognitive tests selected were based upon their relatedness to the MATRICS domains. They represent cognitive domains most affected in schizophrenia cases, and thus are *a priori* more likely to be influenced by schizophrenia genetic risk. However, it is possible the lack of association with any particular cognitive sub-domain is because the selected tests do not adequately capture aspects of cognition that are most relevant to the aetiology of schizophrenia.

5.2.5 Polygenic summary

Schizophrenia, bipolar and schizophrenia versus bipolar polygenic scores were used to predict performance on tests of cognitive domains affected in schizophrenia and performance, verbal and full scale IQ.

Increased polygenic risk of schizophrenia was associated with decreased performance IQ using three schizophrenia discovery datasets, across multiple training thresholds and in two target samples. For the most significant associations, schizophrenia polygenic risk predicted 0.34% of performance IQ variance in ALSPAC, and 0.59% of the variance in the German sample.

Bipolar polygenic risk was not robustly associated with any single cognitive domain or IQ measure in ALSPAC.

Polygenic risk of schizophrenia versus bipolar disorder was predictive of verbal IQ and full scale IQ. For the most significant associations, increased polygenic risk predicted 0.27% of the variance of full scale IQ, and 0.27% of verbal IQ in ALSPAC.

These results provide the most detailed investigation of schizophrenia and bipolar polygenic risk with respect to cognition in healthy individuals to date. Whilst the proportion of cognitive variance explained by polygenic risk scores was modest, it nonetheless provided additional insight into the common genetic overlap between neuropsychiatric disorders and cognitive ability. In addition, selecting cognitive tests with the strongest genetic association with

schizophrenia and bipolar disorder are required to maximise the benefit from future endophenotypic studies.

5.3 Functional Pathways Underlying Cognitive Phenotypes in Schizophrenia and Control Populations

The studies presented in Chapter 3 used two methods to investigate the association of 155 gene-sets with cognitive ability in schizophrenia cases and healthy controls. The gene-sets were members of six overarching biological categories relating to behaviour/cognition, cellular physiology, cellular morphology, development, region tract morphology and subcellular neuronal. The first five categories contained gene-sets that were a subset of the full MGI pathways. Subcellular neuronal pathways were derived from proteomic studies in rodents and humans. Pathways within these categories are *a priori* more likely to be enriched for SNPs influencing general cognitive ability.

The first study applied a polygenic risk score approach to test whether schizophrenia polygenic pathway scores were associated with the MATRICS composite score in schizophrenia cases, and with performance IQ in healthy individuals.

To our knowledge no previous study has investigated whether SNPs associated with general cognitive ability in schizophrenia cases show enrichment in the biological pathways described above.

The second study in this chapter used Brown's method, a set-based test that uses information from all available SNPs within genes in the specified pathway. SNPs associated with the MATRICS composite score were tested for association with specific gene-sets. Rank tests were subsequently performed to test whether specific pathway categories showed greater association when compared against the remaining five categories.

5.3.1 Schizophrenia polygenic pathway scores and their association with general cognitive ability in schizophrenia cases and performance IQ in healthy controls

The results from Chapter 2 indicated increased polygenic risk of schizophrenia was associated with lower performance IQ in healthy controls. In Chapter 3, using a polygenic pathway approach, we tested whether polygenic risk scores in 155 candidate pathways were associated with performance IQ in healthy controls, and also whether they were associated with the MATRICS composite score in schizophrenia cases.

No polygenic pathway scores were associated with the MATRICS composite score in schizophrenia cases or performance IQ in healthy controls, after correction for multiple testing. In addition, there was no evidence suggesting gene-sets belonging to specific biological categories showed greater association with cognitive measures in cases or controls when compared to other biological gene-set categories. There are two possible explanations for these negative

findings. First, the size of our cognition samples was small (and thus there was insufficient variability in the polygenic risk scores) to detect associations with general cognitive ability. A second interpretation is that common genetic risk of schizophrenia does not strongly influence cognitive ability in specific pathways, instead exerting influence on cognitive ability via small effects across the genome.

Some studies have used schizophrenia polygenic pathway scores for ZNF804A and cell adhesion molecular pathways to test for association with cognition in schizophrenia cases (Hargreaves et al., 2013; Nicodemus et al., 2014). Importantly, significant associations between polygenic pathway scores and cognition in these studies were typically modest (p uncorrected ~ 0.01-0.05), and the size of their schizophrenia cognition sample is similar to Cardiff COGS. Given current schizophrenia cognition target samples are small, and the strength of association between schizophrenia polygenic risk and cognition is modest for single pathways, this indicates the testing of multiple pathways in unfeasible until samples are adequately sized and powered to detect robust associations that can withstand multiple correction burdens.

5.3.2 GWAS of MATRICS composite score and Brown's test

This study investigated whether SNPs associated with the MATRICS composite score in schizophrenia cases were enriched in individual gene-sets, or across broader biological categories.

No SNP reached genome wide-significance for association with the MATRICS composite score. Although rs10174400 in *SCN2A* recently reached genome-wide significance for association with general cognitive ability in a small schizophrenia cohort (Dickinson et al., 2014), this finding awaits independent replication. Associations between rs10174400 and the MATRICS composite score were not supported in the present study (p=0.749).

Using Brown's method, SNPs associated with general cognitive ability in schizophrenia cases were not enriched in any of the 155 pathways tested before correction for multiple testing. However, Brown's p-values for individual brain region/fibre tract morphology pathways were significantly higher than those in other pathway categories. The MGI definitions of region tract morphology gene-sets are based upon structural abnormalities of the mouse brain. There is perhaps a relatively sparse literature studying brain structure in relation to cognitive ability in schizophrenia. The largest study to date was performed by Andreasen et al (2011) using a longitudinal design in 202 schizophrenia cases and 125 healthy controls. They showed reductions in white matter in the frontal and temporal lobes were associated with lower attention, problem solving, fluency, verbal learning and working memory. However, these associations were weak, and subject to potential confounders including anti-psychotic medication, illicit drug use and smoking consumption. Within the context of the present study, evidence that brain region/fibre tract pathways are associated with general cognitive ability in schizophrenia cases requires replication in an independent sample. This would provide a basis for whether further work underlying their associations is warranted.

If the general cognitive deficit in schizophrenia is also highly polygenic, large schizophrenia cognition samples will be required to accurately estimate SNP effect sizes. Within the context of the present study, we were likely to be insufficiently powered to detect associations with general cognitive ability through GWAS, reducing the ability to detect associations from SNPs in specific gene sets. The lack of association between SNPs in the gene-sets tested is thus not reliable evidence against their involvement with general cognitive ability in schizophrenia cases. Larger studies with more accurately estimated effect sizes for SNPs associated with cognition would have greater power to detect whether specific gene-sets are implicated in the general cognitive deficit.

5.3.3 Methodological considerations for pathway analyses

To our knowledge, no previous studies have performed pathway analyses on a broad set of candidate pathways using SNPs associated with general cognitive ability in schizophrenia cases. Given that pathway analyses using SNP data primarily use results generated from GWAS, small samples are unlikely to be sufficiently powered (Duncan et al., 2014) to reliably detect SNPs that are causally related to general cognitive ability in schizophrenia cases.

Another potential limitation was that SNPs were assigned to genes only if they were within the transcriptional start and end sites of the gene boundary. There is inconsistency in the literature regarding the appropriate size of gene windows to

select for these types of analyses (Holmans, 2010). SNPs within wider gene windows may have regulatory effects on the gene up to 20kb away (Veyrieras et al., 2008), and others may be in LD with causal SNPs within the gene itself. However, other SNPs will have no functional role and thus no association with the trait under investigation (Holmans, 2010; Ramanan et al., 2012), possibly leading to a deflation in association with specific pathways.

A final limitation regards the selection of pathways tested. We selected a subset of pathways within broad biological categories that *a priori* were more likely to be enriched for SNPs influencing general cognitive ability. Whilst this has advantages with respect to minimising the burden of multiple testing, other biological processes may also be influenced by schizophrenia genetic risk that impact upon cognitive functioning, or contribute to general cognitive ability in schizophrenia cases.

5.3.4 Pathway analysis summary

We were unable to provide robust evidence that polygenic risk of schizophrenia was associated with performance IQ in healthy individuals, or the MATRICS composite score in schizophrenia cases using common polymorphisms within 155 candidate pathways.

Using Brown's method, SNPs associated with general cognitive ability in schizophrenia cases were not enriched in any of the candidate pathways tested. However, a weak association was observed showing pathways in the region tract

morphology pathway were ranked more highly than those in the other 5 pathway categories. Genes in region tract morphology pathways relate to abnormalities in brain structure in the mouse brain. Studies investigating brain structure and general cognitive ability in schizophrenia have not identified substantial associations and are also subject to multiple confounding factors.

Whilst analyses were limited to a relatively low number of pathways in comparison to all pathways in GO (Gene Ontology Consortium et al., 2013), or the MGI (Bult et al., 2013),, multiple correction burdens were still substantial. In addition, where studies are limited to small cognition sample sizes, they are unlikely to have sufficient power to detect associations. These two factors are important limitations of the analyses described above. Larger studies are required to identify whether or not common genetic risk of schizophrenia in biological pathways are associated with general cognitive ability in healthy individuals or schizophrenia cases, and whether common SNPs associated with general cognitive ability in schizophrenia cases are enriched for specific biological processes.

5.4 Rare CNVs and their association with general cognitive ability in schizophrenia cases

The studies presented in Chapter 4 investigated the association between rare CNVs and general cognitive ability in schizophrenia cases. There is robust evidence that a small number of rare CNVs and *de novo* CNVs are found at a higher rate in schizophrenia cases compared to controls (Rees et al., 2012; Rees

et al., 2014b). Furthermore, a number of rare CNVs are associated with neurodevelopmental disorders including autism, intellectual disability and attention deficit hyperactivity disorder (ADHD) (Malhotra & Sebat, 2012). These disorders share a common theme of cognitive impairment. During this thesis, work has been published showing apparently healthy carriers of neuropsychiatric CNVs have greater cognitive impairment when compared to non-carriers, although less impairment relative to schizophrenia cases (Stefansson et al., 2014). To date, a limited number of studies have investigated cognition in relatively small clinical samples of individuals carrying neuropsychiatric CNVs. However, no previous study has investigated whether schizophrenia cases with neuropsychiatric CNVs show greater cognitive impairment compared to cases with no known neuropsychiatric CNV. In addition little is known regarding the association between rare CNV burden and general cognitive ability in schizophrenia cases. Furthermore, it is not known whether genes within candidate pathways (previously described in Chapter 3) hit by rare CNVs are associated with general cognitive ability in schizophrenia cases.

Chapter 4 sought to address these gaps in the literature and expanded upon previous findings using three separate studies. The first study investigated whether schizophrenia cases with a known pathogenic CNV (n=11) had lower general cognitive ability (as measured by the MATRICS composite score) compared to other schizophrenia cases (n=472).

The second study investigated associations between the burden of rare (frequency <1%) CNVs and general cognitive ability in schizophrenia cases.

Several measures of CNV burden were investigated; total length of large (>100kb) and small (between 15-100kb) CNVs as well as total number of genes hit by large and small CNVs. Analyses were performed for deletions and duplications separately, and combined.

The third study investigated whether the total number of genes in candidate pathways (previously described in Chapter 3) hit by CNVs, was associated with general cognitive ability in schizophrenia cases.

5.4.1 Known pathogenic CNVs and their association with general cognitive ability

The first analysis in chapter 4 investigated differences in general cognitive ability between schizophrenia cases stratified by whether they carried a known pathogenic CNV.

We showed schizophrenia cases with a known pathogenic CNV performed almost one standard deviation lower compared to other schizophrenia cases on the MATRICS composite score (p=0.017), and explained approximately 1% of the composite score variance. The findings of a recent study suggested schizophrenia cases with pathogenic CNVs have an earlier onset of disease (Ahn et al., 2014), which is itself associated with greater cognitive impairment, although the authors did not examine cognitive outcomes in this study. After correction for age of onset, pathogenic CNV status remained significantly

associated with the MATRICS composite score and so this does not seem to explain the association between these CNVs and cognitive impairment.

This study extends the previous findings by Steffanson *et al* (2014) who showed apparently healthy carriers of neuropsychiatric CNVs had lower cognitive ability, between that of healthy non-carriers and schizophrenia cases. These results suggest that relatively subtle cognitive deficits are at the mildest end of the phenotypic spectrum of expression of neuropsychiatric CNVs , with additional genetic or environmental factors contributing towards a more severe neuropsychiatric phenotype. However, it is currently unknown whether cognitive deficits resulting from the effects of neuropsychiatric CNVs are independent from the cognitive symptoms observed in disorders for which the CNV confers risk. The findings of the present study may indicate that pathogenic CNVs negatively influence cognitive ability in addition to the deficits observed as part of schizophrenia psychopathology and these large, rare CNVs certainly do not account for the sum of the cognitive deficits seen in the disorder. The fact that smaller CNVs were not associated with impaired cognition, as described in the subsequent study in Chapter 4, may indicate that rare structural variants do not contribute majorly to the general cognitive impairment seen in schizophrenia. Studies in larger, though equally well phenotyped samples, will be required to answer this question definitively.

Not all schizophrenia cases carrying neuropsychiatric CNVs were cognitively impaired beyond other schizophrenia cases. The interpretation of this finding is unclear. Specifically, it is unknown whether these individuals have cognitive

impairments attributable to the effects of the neuropsychiatric CNV, with disease processes relating to schizophrenia having less impact upon cognitive functioning. A separate interpretation is that these individuals may have other genetic or environmental factors that have a protective effect on cognition.

These findings may have implications for future studies that identify single or multiple genes in pathogenic CNVs that contribute to genetic susceptibility of neuropsychiatric disorders and cognitive ability. A key challenge will be to identify which gene, or combination of genes is responsible for the phenotypic impact of the neuropsychiatric CNVs. If this is achieved then understanding the function of these genes individually, or as part of a pathway may provide insights into molecular targets that can form the basis for treatments designed to improve cognitive functioning in patients with schizophrenia.

5.4.2 Rare CNV burden and their association with general cognitive ability in schizophrenia cases

This study investigated measures of rare CNV burden with respect to association with general cognitive ability in schizophrenia cases. Associations were observed between the number of genes hit by large CNVs (>100kb) and a lower MATRICS composite score, both before ($p=0.012$) and after ($p=0.032$) the inclusion of total CNV length into the model. Neither total length nor numbers of genes hit by small (15-100kb) CNVs were associated with the MATRICS composite score.

Associations between CNV deletion and duplication burden and the MATRICS composite score showed similar although somewhat weaker findings . The total number of genes hit by large deletions (p=0.037) was a more significant predictor of the MATRICS composite score compared to duplications (p=0.063).

The findings from previous studies of CNV burden with relation to cognition in schizophrenia are not entirely consistent. IQ has been associated with the total length of rare (frequency < 1%) CNV deletions (Martin et al., 2014), and total number of rare (frequency < 3%) CNV deletions (Yeo et al., 2013), although this has not consistently replicated (van Scheltinga et al., 2013). The total length of CNV deletions have also been associated with IQ in cases with alcohol dependence (Yeo et al., 2011), possibly indicating that rare CNV burden contributes to cognition in multiple psychiatric phenotypes. Our findings may also suggest the number of genes hit by rare CNV duplications also contribute to general cognitive ability in schizophrenia patients, although larger and better powered studies will be required to provide more definitive evidence.

However, an increased number of genes hit by large and rare CNV deletions was associated with lower general cognitive ability, and thus consistent with the direction of effect of previous studies with respect to CNV deletion burden.

Interestingly, rare CNV burden is not associated with general cognitive ability in healthy controls (Kirkpatrick et al., 2014). The associations observed in schizophrenia cases may indicate the effects of rare CNVs upon cognition become

prominent against a backdrop of other genetic vulnerabilities that increase schizophrenia liability (Yeo et al., 2013).

5.4.3 Genes in pathways hit by large (>100kb) and rare (frequency < 1%) CNVs and their association with general cognitive ability in schizophrenia cases

The previous study showed an increased number of genes hit by large (>100kb) and rare (frequency <1%) CNVs were associated with lower general cognitive ability in schizophrenia cases. We then investigated whether the number genes hit by these CNVs in the pathways described previously were associated with the MATRICS composite score.

Of the 155 pathways tested, 85 pathways contained genes that were hit by at least one CNV. However, the number of genes hit within any of these 85 pathways was not significantly associated with the MATRICS composite score even prior to correction for multiple testing. A previous study failed to identify an association between brain-enriched genes hit by CNV deletions and performance/verbal and full scale IQ in 386 schizophrenia cases (Merikangas et al., 2014). However, both the present and Merikangas and colleagues study were performed in relatively small samples. Given that the CNVs tested have a low population frequency, and their effects on cognition in schizophrenia cases are modest, much larger samples are required to ascertain whether or not genes hit

by rare CNVs are enriched in specific gene-sets that influence general cognitive ability in schizophrenia cases.

5.4.4 Methodological considerations of CNV analyses

Only 11 schizophrenia cases in the present analysis carried known neuropsychiatric CNVs. Given these numbers are small, replication of the association between neuropsychiatric CNV status and general cognitive ability in an independent schizophrenia sample with cognitive data would provide greater support for our findings.

With regards to analyses of CNV burden, only rare CNVs with a population frequency of less than one percent were included in the analysis. Common CNV burden is not associated with cognition in healthy controls (Bagshaw et al., 2013), although their contribution to cognition in schizophrenia cases is unknown.

Finally, CNVs were required to hit at least one base pair of the gene. However, CNVs can also affect gene expression when hitting regulatory or functional regions beyond protein coding regions (Stranger et al., 2007). CNVs hitting promoter regions that affect transcription, but do not hit the gene directly, may also contribute to gene expression and affect general cognitive ability.

5.5 Future work

The findings from this thesis provide avenues for further research. First, when data becomes available from ALSPAC, we plan to investigate whether schizophrenia polygenic risk is associated with a negative change in cognitive ability in children between the ages of 8-15. Previous findings have indicated increased polygenic risk of schizophrenia is associated with a negative change across the lifespan (McIntosh et al., 2013). However schizophrenia genetic risk may exert stronger influence on change in cognitive ability during childhood and adolescence, a time period characterised by cognitive lag in individuals who later develop schizophrenia (Reichenberg et al., 2010).

Second, we intend to perform polygenic pathway analyses using the schizophrenia PGC2 discovery dataset to predict cognition within ALSPAC. This would be a substantially better powered analysis in comparison to the analyses presented, providing a greater opportunity to detect whether or not common genetic risk of schizophrenia influences cognition via specific pathways.

Third, we plan to investigate whether common CNV burden is associated with general cognitive ability in schizophrenia patients in Cardiff COGS. Although common CNV burden is not associated with cognition in healthy controls (Bagshaw et al., 2013), this has not been explicitly examined in schizophrenia cases.

Forth, we intend to model the effects of both common polygenic risk scores and rare CNV burden simultaneously upon general cognitive ability in schizophrenia patients.

Fifth, epigenetic functional genomic elements including DNA methylation, chromatin modification and other regulatory elements provide alternative mechanisms for gene expression (Encode Project Consortium, 2012), with relevance to the etiology of psychiatric disorders. Differences in DNA methylation have been identified between schizophrenia cases and controls in blood and frontal cortex (Wockner et al., 2014). DNA methylation has also been implicated in long term memory in mice (Oliveira et al., 2012), although its role on cognition in humans is currently unknown. Investigating the association between cognition in schizophrenia and epigenetic markers may provide new insights into schizophrenia pathophysiology.

5.6 Concluding Remarks

This thesis has investigated the impact of common and rare genetic risk factors for schizophrenia with respect to cognition in patients with schizophrenia and healthy controls. A number of novel findings have been identified. First, increased polygenic risk of schizophrenia is associated with lower performance IQ. This result was demonstrated to replicate in different training sets and in independent target samples. Second, no association was observed between bipolar polygenic risk and cognition. Third, common genetic differences between schizophrenia and bipolar disorder are associated with lower performance on

measures of verbal and full scale IQ. Fourth, schizophrenia cases carrying rare pathogenic CNVs have greater cognitive impairment compared to other schizophrenia cases. Fifth, increases in the number of genes hit by rare CNVs are associated with lower general cognitive ability in schizophrenia cases.

Additional understanding of the genetic mechanisms underlying cognitive impairment in schizophrenia is a step closer toward an understanding of the molecular nature of these deficits, and therefore toward the development of better treatments. This has the potential to improve the functional capacity of schizophrenia patients, leading to a better quality of life and integration within society.

6 References

Ahn, K., Gotay, N., Andersen, T. M., Anvari, A. A., Gochman, P., Lee, Y., et al. (2014). High rate of disease-related copy number variations in childhood onset schizophrenia. *Mol Psychiatry, 19*(5), 568-572.

Albanna, A., Choudhry, Z., Harvey, P. O., Fathalli, F., Cassidy, C., Sengupta, S. M., et al. (2014). TCF4 gene polymorphism and cognitive performance in patients with first episode psychosis. *Schizophr Res, 152*(1), 124-129.

Aleman, A., Hijman, R., de Haan, E. H., & Kahn, R. S. (1999). Memory impairment in schizophrenia: a meta-analysis. *Am J Psychiatry, 156*(9), 1358-1366.

Almeida, O. P., Schwab, S. G., Lautenschlager, N. T., Morar, B., Greenop, K. R., Flicker, L., et al. (2008). KIBRA genetic polymorphism influences episodic memory in later life, but does not increase the risk of mild cognitive impairment. *J Cell Mol Med, 12*(5A), 1672-1676.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (5th Ed.). Washington, DC.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed. ed.). Washington, DC.

An, Juan Ji, Gharami, Kusumika, Liao, Guey-Ying, Woo, Newton H., Lau, Anthony G., Vanevski, Filip, et al. (2008). Distinct Role of Long 3′ UTR BDNF mRNA in Spine Morphology and Synaptic Plasticity in Hippocampal Neurons. *Cell, 134*(1), 175-187.

Andersen, R., Fagerlund, B., Rasmussen, H., Ebdrup, B. H., Aggernaes, B., Gade, A., et al. (2013). The influence of impaired processing speed on cognition in first-episode antipsychotic-naive schizophrenic patients. *Eur Psychiatry, 28*(6), 332-339.

Andreasen, N. C., Nopoulos, P., Magnotta, V., Pierson, R., Ziebell, S., & Ho, B. C. (2011). Progressive brain change in schizophrenia: a prospective longitudinal study of first-episode schizophrenia. *Biol Psychiatry, 70*(7), 672-679.

Antila, M., Tuulio-Henriksson, A., Kieseppa, T., Soronen, P., Palo, O. M., Paunio, T., et al. (2007). Heritability of cognitive functions in families with bipolar disorder. *Am J Med Genet B Neuropsychiatr Genet, 144B*(6), 802-808.

Arancio, O., Kiebler, M., Lee, C. J., Lev-Ram, V., Tsien, R. Y., Kandel, E. R., et al. (1996). Nitric oxide acts directly in the presynaptic neuron to produce long-term potentiation in cultured hippocampal neurons. *Cell, 87*(6), 1025-1035.

Arndt, S., Alliger, R. J., & Andreasen, N. C. (1991). The distinction of positive and negative symptoms. The failure of a two-dimensional model. *Br J Psychiatry, 158*, 317-322.

Arts, B., Jabben, N., Krabbendam, L., & van Os, J. (2008). Meta-analyses of cognitive functioning in euthymic bipolar patients and their first-degree relatives. *Psychol Med, 38*(6), 771-785.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet, 25*(1), 25-29.

August, S. M., Kiwanuka, J. N., McMahon, R. P., & Gold, J. M. (2012). The MATRICS Consensus Cognitive Battery (MCCB): clinical and cognitive correlates. *Schizophr Res, 134*(1), 76-82.

Baddeley, A. (1981). The concept of working memory: a view of its current state and probable future development. *Cognition, 10*(1-3), 17-23.

Baethge, C., Baldessarini, R. J., Freudenthal, K., Streeruwitz, A., Bauer, M., & Bschor, T. (2005). Hallucinations in bipolar disorder: characteristics and comparison to unipolar depression and schizophrenia. *Bipolar Disorders, 7*(2), 136-145.

Bagshaw, A. T., Horwood, L. J., Liu, Y., Fergusson, D. M., Sullivan, P. F., & Kennedy, M. A. (2013). No effect of genome-wide copy number variation on measures of intelligence in a New Zealand birth cohort. *PLoS One, 8*(1), e55208.

Baker, K. D., & Skuse, D. H. (2005). Adolescents and young adults with 22q11 deletion syndrome: psychopathology in an at-risk group. *Br J Psychiatry, 186*, 115-120.

Balding, D. J. (2006). A tutorial on statistical methods for population association studies. *Nat Rev Genet, 7*(10), 781-791.

Ballif, B. C., Theisen, A., Coppinger, J., Gowans, G. C., Hersh, J. H., Madan-Khetarpal, S., et al. (2008). Expanding the clinical phenotype of the 3q29 microdeletion syndrome and characterization of the reciprocal microduplication. *Mol Cytogenet, 1*, 8.

Barch, D. M., Csernansky, J. G., Conturo, T., & Snyder, A. Z. (2002). Working and long-term memory deficits in schizophrenia: is there a common prefrontal mechanism? *J Abnorm Psychol, 111*(3), 478-494.

Barnett, J. H., McDougall, F., Xu, M. K., Croudace, T. J., Richards, M., & Jones, P. B. (2012). Childhood cognitive function and adult psychopathology: associations with psychotic and non-psychotic symptoms in the general population. *Br J Psychiatry, 201*, 124-130.

Barnett, J. H., Scoriels, L., & Munafo, M. R. (2008). Meta-analysis of the cognitive effects of the catechol-O-methyltransferase gene Val158/108Met polymorphism. *Biol Psychiatry, 64*(2), 137-144.

Barron, F., & Harrington, D. M. (1981). Creativity, Intelligence, and Personality. *Annual Review of Psychology, 32*, 439-476.

Bassett, A. S., & Chow, E. W. (2008). Schizophrenia and 22q11.2 deletion syndrome. *Curr Psychiatry Rep, 10*(2), 148-157.

Bassett, A. S., Chow, E. W., AbdelMalik, P., Gheorghiu, M., Husted, J., & Weksberg, R. (2003). The schizophrenia phenotype in 22q11 deletion syndrome. *Am J Psychiatry, 160*(9), 1580-1586.

Bayes, A., van de Lagemaat, L. N., Collins, M. O., Croning, M. D., Whittle, I. R., Choudhary, J. S., et al. (2011). Characterization of the proteome, diseases and evolution of the human postsynaptic density. *Nat Neurosci, 14*(1), 19-21.

Becamel, C., Alonso, G., Geleotti, N., Demey, E., Jouin, P., Ullmer, C., et al. (2002). Synaptic multiprotein complexes associated with 5-HT2C receptors: a proteomic approach. *Embo Journal, 21*(10), 2332-2342.

Benson, M. A., Newey, S. E., Martin-Rendon, E., Hawkes, R., & Blake, D. J. (2001). Dysbindin, a novel coiled-coil-containing protein that interacts with the dystrobrevins in muscle and brain. *J Biol Chem, 276*(26), 24232-24241.

Bergen, S. E., & Petryshen, T. L. (2012). Genome-wide association studies of schizophrenia: does bigger lead to better results? *Curr Opin Psychiatry, 25*(2), 76-82.

Bierhals, T., Maddukuri, S. B., Kutsche, K., & Girisha, K. M. (2013). Expanding the phenotype associated with 17q12 duplication: case report and review of the literature. *Am J Med Genet A, 161A*(2), 352-359.

Blair, C. (2006). How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. *Behav Brain Sci, 29*(2), 109-125; discussion 125-160.

Bora, E., Yucel, M., & Pantelis, C. (2009). Cognitive endophenotypes of bipolar disorder: a meta-analysis of neuropsychological deficits in euthymic patients and their first-degree relatives. *J Affect Disord, 113*(1-2), 1-20.

Bora, E., Yucel, M., Pantelis, C., & Berk, M. (2011). Meta-analytic review of neurocognition in bipolar II disorder. *Acta Psychiatr Scand, 123*(3), 165-174.

Boyd, A., Golding, J., Macleod, J., Lawlor, D. A., Fraser, A., Henderson, J., et al. (2013). Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol, 42*(1), 111-127.

Brandon, N. J., & Sawa, A. (2011). Linking neurodevelopmental and synaptic theories of mental illness through DISC1. *Nat Rev Neurosci, 12*(12), 707-722.

Brebion, G., Smith, M. J., Gorman, J. M., Malaspina, D., Sharif, Z., & Amador, X. (2000). Memory and schizophrenia: differential link of processing speed and selective attention with two levels of encoding. *J Psychiatr Res, 34*(2), 121-127.

Brown, A. S., & Susser, E. S. (2008). Prenatal nutritional deficiency and risk of adult schizophrenia. *Schizophr Bull, 34*(6), 1054-1063.

Brown, Morton B. (1975). A method for combining non-independent, one-sided tests of significance. *Biometrics*, 987-992.

Brunetti-Pierri, N., Berg, J. S., Scaglia, F., Belmont, J., Bacino, C. A., Sahoo, T., et al. (2008). Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities. *Nat Genet, 40*(12), 1466-1471.

Buchanan, R. W., Davis, M., Goff, D., Green, M. F., Keefe, R. S., Leon, A. C., et al. (2005). A summary of the FDA-NIMH-MATRICS workshop on clinical trial design for neurocognitive drugs for schizophrenia. *Schizophr Bull, 31*(1), 5-19.

Buchanan, R. W., Javitt, D. C., Marder, S. R., Schooler, N. R., Gold, J. M., McMahon, R. P., et al. (2007). The Cognitive and Negative Symptoms in Schizophrenia Trial (CONSIST): the efficacy of glutamatergic agents for negative symptoms and cognitive impairments. *Am J Psychiatry, 164*(10), 1593-1602.

Buizer-Voskamp, J. E., Muntjewerff, J. W., Genetic, Risk, Outcome in Psychosis Consortium, Members, Strengman, E., Sabatti, C., et al. (2011). Genome-wide analysis shows increased frequency of copy number variation deletions in Dutch schizophrenia patients. *Biol Psychiatry, 70*(7), 655-662.

Bult, C. J., Eppig, J. T., Blake, J. A., Kadin, J. A., Richardson, J. E., & Mouse Genome Database, Group. (2013). The mouse genome database: genotypes, phenotypes, and models of human disease. *Nucleic Acids Research, 41*(Database issue), D885-891.

Burdick, K. E., Goldberg, T. E., Cornblatt, B. A., Keefe, R. S., Gopin, C. B., Derosse, P., et al. (2011). The MATRICS consensus cognitive battery in patients with bipolar I disorder. *Neuropsychopharmacology, 36*(8), 1587-1592.

Burdick, K. E., Russo, M., Frangou, S., Mahon, K., Braga, R. J., Shanahan, M., et al. (2014). Empirical evidence for discrete neurocognitive subgroups in bipolar disorder: clinical implications. *Psychol Med, FirstView*, 1-14.

Calkins, M. E., Dobie, D. J., Cadenhead, K. S., Olincy, A., Freedman, R., Green, M. F., et al. (2007). The Consortium on the Genetics of Endophenotypes in Schizophrenia: model recruitment, assessment, and endophenotyping methods for a multisite collaboration. *Schizophr Bull, 33*(1), 33-48.

Canetta, S. E., & Brown, A. S. (2012). Prenatal Infection, Maternal Immune Activation, and Risk for Schizophrenia. *Transl Neurosci, 3*(4), 320-327.

Cannon, T. D., Bearden, C. E., Hollister, J. M., Rosso, I. M., Sanchez, L. E., & Hadley, T. (2000). Childhood cognitive functioning in schizophrenia patients and their unaffected siblings: a prospective cohort study. *Schizophr Bull, 26*(2), 379-393.

Cardno, A. G., Marshall, E. J., Coid, B., Macdonald, A. M., Ribchester, T. R., Davies, N. J., et al. (1999). Heritability estimates for psychotic disorders: the Maudsley twin psychosis series. *Arch Gen Psychiatry, 56*(2), 162-168.

Cattell, R. B. (1963). Theory of Fluid and Crystallized Intelligence - a Critical Experiment. *J Educ Psychol, 54*(1), 1-22.

Chadwick, M. J., Hassabis, D., Weiskopf, N., & Maguire, E. A. (2010). Decoding individual episodic memory traces in the human hippocampus. *Curr Biol, 20*(6), 544-547.

Chatterjee, N., Wheeler, B., Sampson, J., Hartge, P., Chanock, S. J., & Park, J. H. (2013). Projecting the performance of risk prediction based on polygenic analyses of genome-wide association studies. *Nat Genet, 45*(4), 400-405, 405e401-403.

Chen, M., Xu, Z., Zhai, J., Bao, X., Zhang, Q., Gu, H., et al. (2012). Evidence of IQ-modulated association between ZNF804A gene polymorphism and cognitive function in schizophrenia patients. *Neuropsychopharmacology, 37*(7), 1572-1578.

Christoforou, A., Espeseth, T., Davies, G., Fernandes, C. P. D., Giddaluru, S., Mattheisen, M., et al. (2014). GWAS-based pathway analysis differentiates between fluid and crystallized intelligence. *Genes, Brain and Behavior*, n/a-n/a.

Cirillo, M. A., & Seidman, L. J. (2003). Verbal declarative memory dysfunction in schizophrenia: From clinical assessment to genetics and brain mechanisms. *Neuropsychology Review, 13*(2), 43-77.

Cirulli, E. T., Kasperaviciute, D., Attix, D. K., Need, A. C., Ge, D., Gibson, G., et al. (2010). Common genetic variation and performance on standardized cognitive tests. *Eur J Hum Genet, 18*(7), 815-820.

Citta, S., Buono, S., Greco, D., Barone, C., Alfei, E., Bulgheroni, S., et al. (2013). 3q29 microdeletion syndrome: Cognitive and behavioral phenotype in four patients. *Am J Med Genet A, 161A*(12), 3018-3022.

Civelek, M., & Lusis, A. J. (2014). Systems genetics approaches to understand complex traits. *Nat Rev Genet, 15*(1), 34-48.
Collingridge, G. L., Peineau, S., Howland, J. G., & Wang, Y. T. (2010). Long-term depression in the CNS. *Nat Rev Neurosci, 11*(7), 459-473.
Conklin, H. M., Curtis, C. E., Calkins, M. E., & Iacono, W. G. (2005). Working memory functioning in schizophrenia patients and their first-degree relatives: cognitive functioning shedding light on etiology. *Neuropsychologia, 43*(6), 930-942.
Conrad, D. F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y., et al. (2010). Origins and functional impact of copy number variation in the human genome. *Nature, 464*(7289), 704-712.
Cooper, A., Grigoryan, G., Guy-David, L., Tsoory, M. M., Chen, A., & Reuveny, E. (2012). Trisomy of the G protein-coupled K+ channel gene, Kcnj6, affects reward mechanisms, cognitive functions, and synaptic plasticity in mice. *Proc Natl Acad Sci U S A, 109*(7), 2642-2647.
Corder, E. H., Saunders, A. M., Strittmatter, W. J., Schmechel, D. E., Gaskell, P. C., Small, G. W., et al. (1993). Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science, 261*(5123), 921-923.
Cousijn, Helena, Eissing, Marc, Fernández, Guillén, Fisher, Simon E., Franke, Barbara, Zwiers, Marcel, et al. (In Press). No effect of schizophrenia risk genes MIR137, TCF4, and ZNF804A on macroscopic brain structure. *Schizophr Res*(0).
Cousijn, J., Wiers, R. W., Ridderinkhof, K. R., van den Brink, W., Veltman, D. J., & Goudriaan, A. E. (2012). Grey matter alterations associated with cannabis use: results of a VBM study in heavy cannabis users and healthy controls. *Neuroimage, 59*(4), 3845-3851.
Coyle, T. R., Pillow, D. R., Snyder, A. C., & Kochunov, P. (2011). Processing speed mediates the development of general intelligence (g) in adolescence. *Psychol Sci, 22*(10), 1265-1269.
Craddock, N., O'Donovan, M. C., & Owen, M. J. (2009). Psychosis genetics: modeling the relationship between schizophrenia, bipolar disorder, and mixed (or "schizoaffective") psychoses. *Schizophr Bull, 35*(3), 482-490.
Craddock, N., & Owen, M. J. (2010). The Kraepelinian dichotomy - going, going... but still not gone. *Br J Psychiatry, 196*(2), 92-95.
Crow, T. J. (1981). Positive and negative schizophrenia symptoms and the role of dopamine. *Br J Psychiatry, 139*, 251-254.
Cuesta, M. J., & Peralta, V. (2001). Integrating psychopathological dimensions in functional psychoses: a hierarchical approach. *Schizophr Res, 52*(3), 215-229.
Darnell, J. C., Van Driesche, S. J., Zhang, C., Hung, K. Y., Mele, A., Fraser, C. E., et al. (2011). FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell, 146*(2), 247-261.
David, A. S., Malmberg, A., Brandt, L., Allebeck, P., & Lewis, G. (1997). IQ and risk for schizophrenia: a population-based cohort study. *Psychol Med, 27*(6), 1311-1323.
Davies, G., Harris, S. E., Reynolds, C. A., Payton, A., Knight, H. M., Liewald, D. C., et al. (2014). A genome-wide association study implicates the APOE locus in nonpathological cognitive ageing. *Mol Psychiatry, 19*(1), 76-87.

Davies, G., Tenesa, A., Payton, A., Yang, J., Harris, S. E., Liewald, D., et al. (2011). Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Mol Psychiatry, 16*(10), 996-1005.
Davis, O. S., Butcher, L. M., Docherty, S. J., Meaburn, E. L., Curtis, C. J., Simpson, M. A., et al. (2010). A three-stage genome-wide association study of general cognitive ability: hunting the small effects. *Behav Genet, 40*(6), 759-767.
De Jager, P. L., Shulman, J. M., Chibnik, L. B., Keenan, B. T., Raj, T., Wilson, R. S., et al. (2012). A genome-wide scan for common variants affecting the rate of age-related cognitive decline. *Neurobiol Aging, 33*(5), 1017 e1011-1015.
De Rubeis, S., Pasciuto, E., Li, K. W., Fernandez, E., Di Marino, D., Buzzi, A., et al. (2013). CYFIP1 coordinates mRNA translation and cytoskeleton remodeling to ensure proper dendritic spine formation. *Neuron, 79*(6), 1169-1182.
Deary, I. J. (1993). Inspection Time and Wais-R Iq Subtypes - a Confirmatory Factor-Analysis Study. *Intelligence, 17*(2), 223-236.
Deary, I. J., Penke, L., & Johnson, W. (2010). The neuroscience of human intelligence differences. *Nat Rev Neurosci, 11*(3), 201-211.
Deary, I. J., Yang, J., Davies, G., Harris, S. E., Tenesa, A., Liewald, D., et al. (2012). Genetic contributions to stability and change in intelligence from childhood to old age. *Nature, 482*(7384), 212-215.
DeLisi, L. E. (1996). Is there a viral or immune dysfunction etiology to schizophrenia? Re-evaluation a decade later. *Schizophr Res, 22*(1), 1-4.
Dick, D. M., Aliev, F., Bierut, L., Goate, A., Rice, J., Hinrichs, A., et al. (2006). Linkage analyses of IQ in the collaborative study on the genetics of alcoholism (COGA) sample. *Behav Genet, 36*(1), 77-86.
Dickinson, D. , Straub, R. E., Trampush, J. W., Gao, Y., Feng, N., Xie, B., et al. (2014). Differential Effects of Common Variants in SCN2A on General Cognitive Ability, Brain Physiology, and messenger RNA Expression in Schizophrenia Cases and Control Individuals. *JAMA Psychiatry, 71*(6), 647-656.
Dickinson, D., Goldberg, T. E., Gold, J. M., Elvevag, B., & Weinberger, D. R. (2011). Cognitive factor structure and invariance in people with schizophrenia, their unaffected siblings, and controls. *Schizophr Bull, 37*(6), 1157-1167.
Dickinson, D., & Harvey, P. D. (2009). Systemic hypotheses for generalized cognitive deficits in schizophrenia: a new take on an old problem. *Schizophr Bull, 35*(2), 403-414.
Dickinson, D., Ragland, J. D., Gold, J. M., & Gur, R. C. (2008). General and specific cognitive deficits in schizophrenia: Goliath defeats David? *Biol Psychiatry, 64*(9), 823-827.
Dickinson, D., Ramsey, M. E., & Gold, J. M. (2007). Overlooking the obvious: a meta-analytic comparison of digit symbol coding tasks and other cognitive measures in schizophrenia. *Arch Gen Psychiatry, 64*(5), 532-542.
Dickinson, Dwight, Schaefer, Jonathan, & Weinberger, Daniel R. (2013a). The multi-faceted," global" cognitive impairment profile in schizophrenia. *Cognitive Impairment in Schizophrenia: Characteristics, Assessment and Treatment*, 24.
Dickinson, Dwight, Schaefer, Jonathan, & Weinberger, Daniel R. (2013b). The multi-faceted," global" cognitive impairment profile in schizophrenia. In

P. D. Harvey (Ed.), *Cognitive Impairment in Schizophrenia: Characteristics, Assessment and Treatment* (pp. 24-49).
Dolmetsch, R. (2003). Excitation-transcription coupling: signaling by ion channels to the nucleus. *Sci STKE, 2003*(166), PE4.
Dominguez, M. G., Viechtbauer, W., Simons, C. J., van Os, J., & Krabbendam, L. (2009). Are psychotic psychopathology and neurocognition orthogonal? A systematic review of their associations. *Psychol Bull, 135*(1), 157-171.
Donohoe, G., Rose, E., Frodl, T., Morris, D., Spoletini, I., Adriano, F., et al. (2011a). ZNF804A risk allele is associated with relatively intact gray matter volume in patients with schizophrenia. *Neuroimage, 54*(3), 2132-2137.
Donohoe, G., Walters, J., Morris, D. W., Da Costa, A., Rose, E., Hargreaves, A., et al. (2011b). A neuropsychological investigation of the genome wide associated schizophrenia risk variant NRGN rs12807809. *Schizophr Res, 125*(2-3), 304-306.
Donohoe, G., Walters, J., Morris, D. W., Quinn, E. M., Judge, R., Norton, N., et al. (2009). Influence of NOS1 on verbal intelligence and working memory in both patients with schizophrenia and healthy control subjects. *Arch Gen Psychiatry, 66*(10), 1045-1054.
Dudbridge, F. (2013). Power and predictive accuracy of polygenic risk scores. *PLoS Genet, 9*(3), e1003348.
Duncan, L. E., Holmans, P. A., Lee, P. H., O'Dushlaine, C. T., Kirby, A. W., Smoller, J. W., et al. (2014). Pathway analyses implicate glial cells in schizophrenia. *PLoS One, 9*(2), e89441.
Elston, R. C. (1998). Methods of linkage analysis--and the assumptions underlying them [see comment]. *Am J Hum Genet, 63*(4), 931-934.
Elvevag, B., Weinberger, D. R., Suter, J. C., & Goldberg, T. E. (2000). Continuous performance test and schizophrenia: a test of stimulus-response compatibility, working memory, response readiness, or none of the above? *Am J Psychiatry, 157*(5), 772-780.
Encode Project Consortium. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature, 489*(7414), 57-74.
Erecinska, M., & Silver, I. A. (1990). Metabolism and role of glutamate in mammalian brain. *Prog Neurobiol, 35*(4), 245-296.
Esslinger, C., Kirsch, P., Haddad, L., Mier, D., Sauer, C., Erk, S., et al. (2011). Cognitive state and connectivity effects of the genome-wide significant psychosis variant in ZNF804A. *Neuroimage, 54*(3), 2514-2523.
Evangelou, M., Rendon, A., Ouwehand, W. H., Wernisch, L., & Dudbridge, F. (2012). Comparison of methods for competitive tests of pathway analysis. *PLoS One, 7*(7), e41018.
Faguer, S., Chassaing, N., Bandin, F., Prouheze, C., Arveiler, B., Rooryck, C., et al. (2011). A 17q12 chromosomal duplication associated with renal disease and esophageal atresia. *Eur J Med Genet, 54*(4), e437-440.
Farr, C. D., Gafken, P. R., Norbeck, A. D., Doneanu, C. E., Stapels, M. D., Barofsky, D. F., et al. (2004). Proteomic analysis of native metabotropic glutamate receptor 5 protein complexes reveals novel molecular constituents. *J Neurochem, 91*(2), 438-450.
Fernando, M. M., Stevens, C. R., Walsh, E. C., De Jager, P. L., Goyette, P., Plenge, R. M., et al. (2008). Defining the role of the MHC in autoimmunity: a review and pooled analysis. *PLoS Genet, 4*(4), e1000024.

Finkel, D., Pedersen, N. L., McGue, M., & McClearn, G. E. (1995). Heritability of cognitive abilities in adult twins: comparison of Minnesota and Swedish data. *Behav Genet, 25*(5), 421-431.
Fleming, K., Goldberg, T. E., Binks, S., Randolph, C., Gold, J. M., & Weinberger, D. R. (1997). Visuospatial working memory in patients with schizophrenia. *Biol Psychiatry, 41*(1), 43-49.
Forbes, N. F., Carrick, L. A., McIntosh, A. M., & Lawrie, S. M. (2009). Working memory in schizophrenia: a meta-analysis. *Psychol Med, 39*(6), 889-905.
Forrest, M. P., Waite, A. J., Martin-Rendon, E., & Blake, D. J. (2013). Knockdown of human TCF4 affects multiple signaling pathways involved in cell survival, epithelial to mesenchymal transition and neuronal differentiation. *PLoS One, 8*(8), e73169.
Fowler T, Zammit S. Owen M. J. Rasmussen F. (2012). A population-based study of shared genetic variation between premorbid iq and psychosis among male twin pairs and sibling pairs from sweden. *Arch Gen Psychiatry, 69*(5), 460-466.
Frangou, S. (2010). Cognitive function in early onset schizophrenia: a selective review. *Front Hum Neurosci, 3*, 79.
Freedman, R., Olincy, A., Buchanan, R. W., Harris, J. G., Gold, J. M., Johnson, L., et al. (2008). Initial phase 2 trial of a nicotinic agonist in schizophrenia. *Am J Psychiatry, 165*(8), 1040-1047.
Fromer, M., Pocklington, A. J., Kavanagh, D. H., Williams, H. J., Dwyer, S., Gormley, P., et al. (2014). De novo mutations in schizophrenia implicate synaptic networks. *Nature, 506*(7487), 179-184.
Fuller, R.; . (2002). Longitudinal Assessment of Premorbid Cognitive Functioning in Patients With Schizophrenia Through Examination of Standardized Scholastic Test Performance. *American Journal of Psychiatry, 159*(7), 1183-1189.
Fusar-Poli, P., Deste, G., Smieskova, R., Barlati, S., Yung, A. R., Howes, O., et al. (2012). Cognitive functioning in prodromal psychosis: a meta-analysis. *Arch Gen Psychiatry, 69*(6), 562-571.
Fusar-Poli, P., Smieskova, R., Kempton, M. J., Ho, B. C., Andreasen, N. C., & Borgwardt, S. (2013). Progressive brain changes in schizophrenia related to antipsychotic treatment? A meta-analysis of longitudinal MRI studies. *Neurosci Biobehav Rev, 37*(8), 1680-1691.
Galton, Francis. (1883). *Inquiries into human faculty and its development*: Macmillan.
Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The Children's Test of Nonword Repetition: a test of phonological working memory. *Memory, 2*(2), 103-127.
Ge, Y., Dong, Z., Bagot, R. C., Howland, J. G., Phillips, A. G., Wong, T. P., et al. (2010). Hippocampal long-term depression is required for the consolidation of spatial memory. *Proc Natl Acad Sci U S A, 107*(38), 16697-16702.
Gene Ontology Consortium, Blake, J. A., Dolan, M., Drabkin, H., Hill, D. P., Li, N., et al. (2013). Gene Ontology annotations and resources. *Nucleic Acids Research, 41*(Database issue), D530-535.
Girirajan, S., Rosenfeld, J. A., Coe, B. P., Parikh, S., Friedman, N., Goldstein, A., et al. (2012). Phenotypic heterogeneity of genomic disorders and rare copy-number variants. *N Engl J Med, 367*(14), 1321-1331.

Glahn, D. C., Almasy, L., Blangero, J., Burk, G. M., Estrada, J., Peralta, J. M., et al. (2007). Adjudicating neurocognitive endophenotypes for schizophrenia. *Am J Med Genet B Neuropsychiatr Genet, 144B*(2), 242-249.
Goddard, C. A., Butts, D. A., & Shatz, C. J. (2007). Regulation of CNS synapses by neuronal MHC class I. *Proc Natl Acad Sci U S A, 104*(16), 6828-6833.
Gogos, J. A., Morgan, M., Luine, V., Santha, M., Ogawa, S., Pfaff, D., et al. (1998). Catechol-O-methyltransferase-deficient mice exhibit sexually dimorphic changes in catecholamine levels and behavior. *Proc Natl Acad Sci U S A, 95*(17), 9991-9996.
Gold, J. M. (2004). Cognitive deficits as treatment targets in schizophrenia. *Schizophr Res, 72*(1), 21-28.
Gold, J. M., Hahn, B., Zhang, W. W., Robinson, B. M., Kappenman, E. S., Beck, V. M., et al. (2010). Reduced capacity but spared precision and maintenance of working memory representations in schizophrenia. *Arch Gen Psychiatry, 67*(6), 570-577.
Gold, J. M., Rehkemper, G., Binks, S. W., 3rd, Carpenter, C. J., Fleming, K., Goldberg, T. E., et al. (2000). Learning and forgetting in schizophrenia. *J Abnorm Psychol, 109*(3), 534-538.
Goldberg, T. E., Keefe, R. S., Goldman, R. S., Robinson, D. G., & Harvey, P. D. (2010). Circumstances under which practice does not make perfect: a review of the practice effect literature in schizophrenia and its relevance to clinical treatment studies. *Neuropsychopharmacology, 35*(5), 1053-1062.
Golding, J., Pembrey, M., Jones, R., & Team, Alspac Study. (2001). ALSPAC--the Avon Longitudinal Study of Parents and Children. I. Study methodology. *Paediatr Perinat Epidemiol, 15*(1), 74-87.
Gottesman, II, & Shields, J. (1967). A polygenic theory of schizophrenia. *Proc Natl Acad Sci U S A, 58*(1), 199-205.
Gottesman, II, & Shields, J. (1973). Genetic theorizing and schizophrenia. *Br J Psychiatry, 122*(566), 15-30.
Gottesman, I. I.; Gould, T.D. . (2003). The Endophenotype Concept in Psychiatry: Etymology and Strategic Intentions. *American Journal of Psychiatry, 160*(4), 636-645.
Gottesman, Irving I. (1991). *Schizophrenia genesis: The origins of madness*: WH Freeman/Times Books/Henry Holt & Co.
Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography (Reprinted from The Wall Street Journal, 1994). *Intelligence, 24*(1), 13-23.
Gould, T. D., & Gottesman, II. (2006). Psychiatric endophenotypes and the development of valid animal models. *Genes Brain Behav, 5*(2), 113-119.
Graf, E. R., Zhang, X., Jin, S. X., Linhoff, M. W., & Craig, A. M. (2004). Neurexins induce differentiation of GABA and glutamate postsynaptic specializations via neuroligins. *Cell, 119*(7), 1013-1026.
Gray, B. E., McMahon, R. P., & Gold, J. M. (2013). General intellectual ability does not explain the general deficit in schizophrenia. *Schizophr Res, 147*(2-3), 315-319.
Gray, J. A., & Roth, B. L. (2007). Molecular targets for treating cognitive dysfunction in schizophrenia. *Schizophr Bull, 33*(5), 1100-1119.
Green, M. F. (2006). Cognitive impairment and functional outcome in schizophrenia and bipolar disorder. *J Clin Psychiatry, 67*(10), e12.

Green, M. F., & Horan, W. P. (2010). Social Cognition in Schizophrenia. *Current Directions in Psychological Science, 19*(4), 243-248.
Green, M. F., Nuechterlein, K. H., Gold, J. M., Barch, D. M., Cohen, J., Essock, S., et al. (2004). Approaching a consensus cognitive battery for clinical trials in schizophrenia: the NIMH-MATRICS conference to select cognitive domains and test criteria. *Biol Psychiatry, 56*(5), 301-307.
Greenwood, T. A., Braff, D. L., Light, G. A., Cadenhead, K. S., Calkins, M. E., Dobie, D. J., et al. (2007). Initial heritability analyses of endophenotypic measures for schizophrenia: the consortium on the genetics of schizophrenia. *Arch Gen Psychiatry, 64*(11), 1242-1250.
Gronborg, M., Pavlos, N. J., Brunk, I., Chua, J. J., Munster-Wandowski, A., Riedel, D., et al. (2010). Quantitative comparison of glutamatergic and GABAergic synaptic vesicles unveils selectivity for few proteins including MAL2, a novel synaptic vesicle protein. *J Neurosci, 30*(1), 2-12.
Grozeva, D., Kirov, G., Ivanov, D., Jones, I. R., Jones, L., Green, E. K., et al. (2010). Rare copy number variants: a point of rarity in genetic risk for bipolar disorder and schizophrenia. *Arch Gen Psychiatry, 67*(4), 318-327.
Gulsuner, S., Walsh, T., Watts, A. C., Lee, M. K., Thornton, A. M., Casadei, S., et al. (2013). Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. *Cell, 154*(3), 518-529.
Gur, R. E., Calkins, M. E., Gur, R. C., Horan, W. P., Nuechterlein, K. H., Seidman, L. J., et al. (2007). The Consortium on the Genetics of Schizophrenia: neurocognitive endophenotypes. *Schizophr Bull, 33*(1), 49-68.
Haas, G. L., & Sweeney, J. A. (1992). Premorbid and onset features of first-episode schizophrenia. *Schizophr Bull, 18*(3), 373-386.
Hamshere, M. L., O'Donovan, M. C., Jones, I. R., Jones, L., Kirov, G., Green, E. K., et al. (2011). Polygenic dissection of the bipolar phenotype. *Br J Psychiatry, 198*(4), 284-288.
Hamshere, M. L., Walters, J. T., Smith, R., Richards, A. L., Green, E., Grozeva, D., et al. (2013). Genome-wide significant associations in schizophrenia to ITIH3/4, CACNA1C and SDCCAG8, and extensive replication of associations reported by the Schizophrenia PGC. *Mol Psychiatry, 18*(6), 708-712.
Hanscombe, K. B., Trzaskowski, M., Haworth, C. M., Davis, O. S., Dale, P. S., & Plomin, R. (2012). Socioeconomic status (SES) and children's intelligence (IQ): in a UK-representative sample SES moderates the environmental, not genetic, effect on IQ. *PLoS One, 7*(2), e30320.
Hanson, E., Nasir, R. H., Fong, A., Lian, A., Hundley, R., Shen, Y., et al. (2010). Cognitive and behavioral characterization of 16p11.2 deletion syndrome. *J Dev Behav Pediatr, 31*(8), 649-657.
Hargreaves, A., Anney, R., O'Dushlaine, C., Nicodemus, K. K., Schizophrenia Psychiatric Genome-Wide Association Study, Consortium, Wellcome Trust Case Control, Consortium, et al. (2013). The one and the many: effects of the cell adhesion molecule pathway on neuropsychological function in psychosis. *Psychol Med*, 1-11.
Harold, D., Abraham, R., Hollingworth, P., Sims, R., Gerrish, A., Hamshere, M. L., et al. (2009). Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat Genet, 41*(10), 1088-1093.

Harrison, P. J., & Weinberger, D. R. (2005). Schizophrenia genes, gene expression, and neuropathology: on the matter of their convergence. *Mol Psychiatry, 10*(1), 40-68; image 45.
Hartman, M., Steketee, M. C., Silva, S., Lanning, K., & McCann, H. (2003). Working memory and schizophrenia: evidence for slowed encoding. *Schizophr Res, 59*(2-3), 99-113.
Harvey, P. D., & Keefe, R. S. (2001). Studies of cognitive change in patients with schizophrenia following novel antipsychotic treatment. *Am J Psychiatry, 158*(2), 176-184.
Haworth, C. M., Wright, M. J., Luciano, M., Martin, N. G., de Geus, E. J., van Beijsterveldt, C. E., et al. (2010). The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Mol Psychiatry, 15*(11), 1112-1120.
Hayashi, N., Kazui, H., Kamino, K., Tokunaga, H., Takaya, M., Yokokoji, M., et al. (2010). KIBRA genetic polymorphism influences episodic memory in Alzheimer's disease, but does not show association with disease in a Japanese cohort. *Dement Geriatr Cogn Disord, 30*(4), 302-308.
Heaton, R. K.; Gladsjo, J.A.; Palmer, B.W.; Kuck, J.; Marcotte, T.D.; Jeste, D.V. (2001). Stability and Course of Neuropsychological Deficits in Schizophrenia. *Arch Gen Psychiatry, 58*(1), 24-32.
Heinrichs, R. Walter, & Zakzanis, Konstantine K. (1998). Neurocognitive deficit in schizophrenia: A quantitative review of the evidence. *Neuropsychology, 12*(3), 426-445.
Heller, E. A., Zhang, W., Selimi, F., Earnheart, J. C., Slimak, M. A., Santos-Torres, J., et al. (2012). The biochemical anatomy of cortical inhibitory synapses. *PLoS One, 7*(6), e39572.
Hill, M. J., & Bray, N. J. (2012). Evidence That Schizophrenia Risk Variation in the ZNF804A Gene Exerts Its Effects During Fetal Brain Development. *American Journal of Psychiatry, 169*(12), 1301-1308.
Hill, M. J., Jeffries, A. R., Dobson, R. J., Price, J., & Bray, N. J. (2012). Knockdown of the psychosis susceptibility gene ZNF804A alters expression of genes involved in cell adhesion. *Hum Mol Genet, 21*(5), 1018-1024.
Hill, W. D., Davies, G., van de Lagemaat, L. N., Christoforou, A., Marioni, R. E., Fernandes, C. P., et al. (2014a). Human cognitive ability is influenced by genetic variation in components of postsynaptic signalling complexes assembled by NMDA receptors and MAGUK proteins. *Transl Psychiatry, 4*, e341.
Hill, W. D., de Leeuw, C., Davies, G., Liewald, D. C., Payton, A., Craig, L. C., et al. (2014b). Functional gene group analysis indicates no role for heterotrimeric G proteins in cognitive ability. *PLoS One, 9*(3), e91690.
Hoff, A. L., Svetina, C., Shields, G., Stewart, J., & DeLisi, L. E. (2005). Ten year longitudinal study of neuropsychological functioning subsequent to a first episode of schizophrenia. *Schizophr Res, 78*(1), 27-34.
Holmans, P. (2010). Statistical methods for pathway analysis of genome-wide data for association with complex genetic traits. *Adv Genet, 72*, 141-179.
Hor, K., & Taylor, M. (2010). Suicide and schizophrenia: a systematic review of rates and risk factors. *J Psychopharmacol, 24*(4 Suppl), 81-90.

Hori, H., Yamamoto, N., Fujii, T., Teraishi, T., Sasayama, D., Matsuo, J., et al. (2012). Effects of the CACNA1C risk allele on neurocognition in patients with schizophrenia and healthy individuals. *Sci Rep, 2*, 634.

Huang, F. L., Huang, K. P., & Boucheron, C. (2007). Long-term enrichment enhances the cognitive behavior of the aging neurogranin null mice without affecting their hippocampal LTP. *Learn Mem, 14*(8), 512-519.

Hubbard, L., Tansley, T .E., Raj, D., Jones, P., Ripke, S., Chambert, K., et al. (submitted). Evidence of limited common genetic overlap between schizophrenia and cognition.

Hunt, D. L., & Castillo, P. E. (2012). Synaptic plasticity of NMDA receptors: mechanisms and functional implications. *Curr Opin Neurobiol, 22*(3), 496-508.

Husted, Janice A., Lim, Sooyeol, Chow, Eva W. C., Greenwood, Celia, & Bassett, Anne S. (2009). Heritability of neurocognitive traits in familial schizophrenia. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 150B*(6), 845-853.

Ingraham, L. J., & Kety, S. S. (2000). Adoption studies of schizophrenia. *Am J Med Genet, 97*(1), 18-22.

International Schizophrenia Consortium. (2008). Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature, 455*(7210), 237-241.

International Schizophrenia Consortium. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature, 460*(7256), 748-752.

Isohanni, M, Jones, PB, Moilanen, K, Rantakallio, P, Veijola, J, Oja, H, et al. (2001). Early developmental milestones in adult schizophrenia and other psychoses. A 31-year follow-up of the Northern Finland 1966 Birth Cohort. *Schizophr Res, 52*(1), 1-19.

Jaaskelainen, E., Juola, P., Hirvonen, N., McGrath, J. J., Saha, S., Isohanni, M., et al. (2013). A systematic review and meta-analysis of recovery in schizophrenia. *Schizophr Bull, 39*(6), 1296-1306.

Jablensky, A. (2000). Prevalence and incidence of schizophrenia spectrum disorders: implications for prevention. *Aust N Z J Psychiatry, 34 Suppl*, S26-34; discussion S35-28.

Jablensky, A. (2006). Subtyping schizophrenia: implications for genetic research. *Mol Psychiatry, 11*(9), 815-836.

Jones, Peter, Murray, R, Jones, P, Rodgers, B, & Marmot, M. (1994). Child developmental risk factors for adult schizophrenia in the British 1946 birth cohort. *The Lancet, 344*(8934), 1398-1402.

Joukamaa, M., Heliovaara, M., Knekt, P., Aromaa, A., Raitasalo, R., & Lehtinen, V. (2006). Schizophrenia, neuroleptic medication and mortality. *Br J Psychiatry, 188*, 122-127.

Kaminsky, E. B., Kaul, V., Paschall, J., Church, D. M., Bunke, B., Kunig, D., et al. (2011). An evidence-based approach to establish the functional and clinical significance of copy number variants in intellectual and developmental disabilities. *Genet Med, 13*(9), 777-784.

Kaprio, J. (2012). Twins and the mystery of missing heritability: the contribution of gene-environment interactions. *J Intern Med, 272*(5), 440-448.

Keefe, R. S., Bilder, R. M., Harvey, P. D., Davis, S. M., Palmer, B. W., Gold, J. M., et al. (2006). Baseline neurocognitive deficits in the CATIE schizophrenia trial. *Neuropsychopharmacology, 31*(9), 2033-2046.

Keefe, R. S., Buchanan, R. W., Marder, S. R., Schooler, N. R., Dugar, A., Zivkov, M., et al. (2013). Clinical trials of potential cognitive-enhancing drugs in schizophrenia: what have we learned so far? *Schizophr Bull, 39*(2), 417-435.

Keefe, R. S., Goldberg, T. E., Harvey, P. D., Gold, J. M., Poe, M. P., & Coughenour, L. (2004). The Brief Assessment of Cognition in Schizophrenia: reliability, sensitivity, and comparison with a standard neurocognitive battery. *Schizophr Res, 68*(2-3), 283-297.

Keefe, R. S., & Harvey, P. D. (2012). Cognitive impairment in schizophrenia. *Handb Exp Pharmacol*(213), 11-37.

Kern, R. S., Green, M. F., Nuechterlein, K. H., & Deng, B. H. (2004). NIMH-MATRICS survey on assessment of neurocognition in schizophrenia. *Schizophr Res, 72*(1), 11-19.

Kern, R. S., Nuechterlein, K. H., Green, M. F., Baade, L. E., Fenton, W. S., Gold, J. M., et al. (2008). The MATRICS Consensus Cognitive Battery, part 2: co-norming and standardization. *Am J Psychiatry, 165*(2), 214-220.

Kern, Robert S., Green, Michael F., & Marder, Stephen R. (2007). The NIMH MATRICS Initiative: Development of a Consensus Cognitive Battery. *Progress in Neurotherapeutics and Neuropsychopharmacology, 2*(01), 173-186.

Kirkpatrick, R. M., McGue, M., Iacono, W. G., Miller, M. B., Basu, S., & Pankratz, N. (2014). Low-Frequency Copy-Number Variants and General Cognitive Ability: No Evidence of Association. *Intelligence, 42*, 98-106.

Kirov, G., Grozeva, D., Norton, N., Ivanov, D., Mantripragada, K. K., Holmans, P., et al. (2009a). Support for the involvement of large copy number variants in the pathogenesis of schizophrenia. *Hum Mol Genet, 18*(8), 1497-1503.

Kirov, G., Pocklington, A. J., Holmans, P., Ivanov, D., Ikeda, M., Ruderfer, D., et al. (2012). De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. *Mol Psychiatry, 17*(2), 142-153.

Kirov, G., Rees, E., Walters, J. T., Escott-Price, V., Georgieva, L., Richards, A. L., et al. (2014). The penetrance of copy number variations for schizophrenia and developmental delay. *Biol Psychiatry, 75*(5), 378-385.

Kirov, G., Rujescu, D., Ingason, A., Collier, D. A., O'Donovan, M. C., & Owen, M. J. (2009b). Neurexin 1 (NRXN1) deletions in schizophrenia. *Schizophr Bull, 35*(5), 851-854.

Kline, Paul. (1991). *Intelligence: The Psychometric View*. London: Routledge

Klose, M. K., Dason, J. S., Atwood, H. L., Boulianne, G. L., & Mercier, A. J. (2010). Peptide-induced modulation of synaptic transmission and escape response in Drosophila requires two G-protein-coupled receptors. *J Neurosci, 30*(44), 14724-14734.

Knowles, E. E., David, A. S., & Reichenberg, A. (2010). Processing speed deficits in schizophrenia: reexamining the evidence. *Am J Psychiatry, 167*(7), 828-835.

Kobrynski, L. J., & Sullivan, K. E. (2007). Velocardiofacial syndrome, DiGeorge syndrome: the chromosome 22q11.2 deletion syndromes. *Lancet, 370*(9596), 1443-1452.
Kohler, C. G., Walker, J. B., Martin, E. A., Healey, K. M., & Moberg, P. J. (2010). Facial emotion perception in schizophrenia: a meta-analytic review. *Schizophr Bull, 36*(5), 1009-1019.
Kraft, P. (2008). Curses--winner's and otherwise--in genetic epidemiology. *Epidemiology, 19*(5), 649-651; discussion 657-648.
Kumari, V., & Postma, P. (2005). Nicotine use in schizophrenia: the self medication hypotheses. *Neurosci Biobehav Rev, 29*(6), 1021-1034.
Lachman, H. M., Papolos, D. F., Saito, T., Yu, Y. M., Szumlanski, C. L., & Weinshilboum, R. M. (1996). Human catechol-O-methyltransferase pharmacogenetics: description of a functional polymorphism and its potential application to neuropsychiatric disorders. *Pharmacogenetics, 6*(3), 243-250.
Lander, E. S. (1996). The new genomics: global views of biology. *Science, 274*(5287), 536-539.
Large, M., Sharma, S., Compton, M. T., Slade, T., & Nielssen, O. (2011). Cannabis use and earlier onset of psychosis: a systematic meta-analysis. *Arch Gen Psychiatry, 68*(6), 555-561.
Lasser, K., Boyd, J. W., Woolhandler, S., Himmelstein, D. U., McCormick, D., & Bor, D. H. (2000). Smoking and mental illness: A population-based prevalence study. *JAMA, 284*(20), 2606-2610.
Laursen, T. M. (2011). Life expectancy among persons with schizophrenia or bipolar affective disorder. *Schizophr Res, 131*(1-3), 101-104.
Lawrie, S. M., & Abukmeil, S. S. (1998). Brain abnormality in schizophrenia. A systematic and quantitative review of volumetric magnetic resonance imaging studies. *Br J Psychiatry, 172*, 110-120.
Lee, H., Brott, B. K., Kirkby, L. A., Adelson, J. D., Cheng, S., Feller, M. B., et al. (2014). Synapse elimination and learning rules co-regulated by MHC class I H2-Db. *Nature, 509*(7499), 195-200.
Lee, J., Folley, B. S., Gore, J., & Park, S. (2008). Origins of spatial working memory deficits in schizophrenia: an event-related FMRI and near-infrared spectroscopy study. *PLoS One, 3*(3), e1760.
Lee, J., & Park, S. (2005). Working memory impairments in schizophrenia: a meta-analysis. *J Abnorm Psychol, 114*(4), 599-611.
Lencz, T., Knowles, E., Davies, G., Guha, S., Liewald, D. C., Starr, J. M., et al. (2014). Molecular genetic evidence for overlap between general cognitive ability and risk for schizophrenia: a report from the Cognitive Genomics consorTium (COGENT). *Mol Psychiatry, 19*(2), 168-174.
Lencz, T., Morgan, T. V., Athanasiou, M., Dain, B., Reed, C. R., Kane, J. M., et al. (2007). Converging evidence for a pseudoautosomal cytokine receptor gene locus in schizophrenia. *Mol Psychiatry, 12*(6), 572-580.
Lennertz, L., Rujescu, D., Wagner, M., Frommann, I., Schulze-Rauschenbach, S., Schuhmacher, A., et al. (2011). Novel schizophrenia risk gene TCF4 influences verbal learning and memory functioning in schizophrenia patients. *Neuropsychobiology, 63*(3), 131-136.
Levinson, D. F., Duan, J. B., Oh, S., Wang, K., Sanders, A. R., Shi, J. X., et al. (2011a). Copy Number Variants in Schizophrenia: Confirmation of Five Previous

Findings and New Evidence for 3q29 Microdeletions and VIPR2 Duplications. *American Journal of Psychiatry, 168*(3), 302-316.
Levinson, D. F., Duan, J., Oh, S., Wang, K., Sanders, A. R., Shi, J., et al. (2011b). Copy number variants in schizophrenia: confirmation of five previous findings and new evidence for 3q29 microdeletions and VIPR2 duplications. *Am J Psychiatry, 168*(3), 302-316.
Li, Y., Willer, C. J., Ding, J., Scheet, P., & Abecasis, G. R. (2010). MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol, 34*(8), 816-834.
Lieberman, J. A., Stroup, T. S., McEvoy, J. P., Swartz, M. S., Rosenheck, R. A., Perkins, D. O., et al. (2005). Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *N Engl J Med, 353*(12), 1209-1223.
Linden, D. E., Lancaster, T. M., Wolf, C., Baird, A., Jackson, M. C., Johnston, S. J., et al. (2013). ZNF804A genotype modulates neural activity during working memory for faces. *Neuropsychobiology, 67*(2), 84-92.
Liu, S. K., Chiu, C. H., Chang, C. J., Hwang, T. J., Hwu, H. G., & Chen, W. J. (2002). Deficits in sustained attention in schizophrenia and affective disorders: stable versus state-dependent markers. *Am J Psychiatry, 159*(6), 975-982.
Luciano, M., Hansell, N. K., Lahti, J., Davies, G., Medland, S. E., Raikkonen, K., et al. (2011). Whole genome association scan for genetic polymorphisms influencing information processing speed. *Biol Psychol, 86*(3), 193-202.
Luciano, M., Posthuma, D., Wright, M. J., de Geus, E. J., Smith, G. A., Geffen, G. M., et al. (2005). Perceptual speed does not cause intelligence, and intelligence does not cause perceptual speed. *Biol Psychol, 70*(1), 1-8.
Luciano, M., Wright, M. J., Duffy, D. L., Wainwright, M. A., Zhu, G., Evans, D. M., et al. (2006). Genome-wide scan of IQ finds significant linkage to a quantitative trait locus on 2q. *Behav Genet, 36*(1), 45-55.
MacCabe, J. H., Brebion, G., Reichenberg, A., Ganguly, T., McKenna, P. J., Murray, R. M., et al. (2012). Superior intellectual ability in schizophrenia: neuropsychological characteristics. *Neuropsychology, 26*(2), 181-190.
MacCabe, J. H., Lambe, M. P., Cnattingius, S., Sham, P. C., David, A. S., Reichenberg, A., et al. (2010). Excellent school performance at age 16 and risk of adult bipolar disorder: national cohort study. *Br J Psychiatry, 196*(2), 109-115.
MacCabe, J. H., Wicks, S., Lofving, S., David, A. S., Berndtsson, A., Gustafsson, J. E., et al. (2013). Decline in cognitive performance between ages 13 and 18 years and the risk for psychosis in adulthood: a Swedish longitudinal cohort study in males. *JAMA Psychiatry, 70*(3), 261-270.
MacLeod, A. K., Davies, G., Payton, A., Tenesa, A., Harris, S. E., Liewald, D., et al. (2012). Genetic copy number variation and general cognitive ability. *PLoS One, 7*(12), e37385.
Makuch, L., Volk, L., Anggono, V., Johnson, R. C., Yu, Y., Duning, K., et al. (2011). Regulation of AMPA receptor function by the human memory-associated gene KIBRA. *Neuron, 71*(6), 1022-1029.
Malenka, R. C., & Nicoll, R. A. (1993). NMDA-receptor-dependent synaptic plasticity: multiple forms and mechanisms. *Trends Neurosci, 16*(12), 521-527.
Malhotra, D., McCarthy, S., Michaelson, J. J., Vacic, V., Burdick, K. E., Yoon, S., et al. (2011). High frequencies of de novo CNVs in bipolar disorder and schizophrenia. *Neuron, 72*(6), 951-963.

Malhotra, D., & Sebat, J. (2012). CNVs: harbingers of a rare variant revolution in psychiatric genetics. *Cell, 148*(6), 1223-1241.
Mandelman, S. D., & Grigorenko, E. L. (2012). BDNF Val66Met and cognition: all, none, or some? A meta-analysis of the genetic association. *Genes, Brain and Behavior, 11*(2), 127-136.
Mangalore, R., & Knapp, M. (2007). Cost of schizophrenia in England. *J Ment Health Policy Econ, 10*(1), 23-41.
Manrique-Garcia, E., Zammit, S., Dalman, C., Hemmingsson, T., Andreasson, S., & Allebeck, P. (2012). Cannabis, schizophrenia and other non-affective psychoses: 35 years of follow-up of a population-based cohort. *Psychol Med, 42*(6), 1321-1328.
Marchini, J., & Howie, B. (2010). Genotype imputation for genome-wide association studies. *Nat Rev Genet, 11*(7), 499-511.
Marder, S. R. (2006). The NIMH-MATRICS project for developing cognition-enhancing agents for schizophrenia. *Dialogues Clin Neurosci, 8*(1), 109-113.
Martin, A. K., Robinson, G., Reutens, D., & Mowry, B. (2014). Copy number deletion burden is associated with cognitive, structural, and resting-state network differences in patients with schizophrenia. *Behav Brain Res*.
Martin, N. W., Medland, S. E., Verweij, K. J., Lee, S. H., Nyholt, D. R., Madden, P. A., et al. (2011). Educational attainment: a genome wide association study in 9538 Australians. *PLoS One, 6*(6), e20128.
Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2001). Emotional intelligence as a standard intelligence. *Emotion, 1*(3), 232-242.
McCarthy, S. E., Makarov, V., Kirov, G., Addington, A. M., McClellan, J., Yoon, S., et al. (2009). Microduplications of 16p11.2 are associated with schizophrenia. *Nat Genet, 41*(11), 1223-1227.
McGrath, J. A., Nestadt, G., Liang, K. Y., Lasseter, V. K., Wolyniec, P. S., Fallin, M. D., et al. (2004a). Five latent factors underlying schizophrenia: analysis and relationship to illnesses in relatives. *Schizophr Bull, 30*(4), 855-873.
McGrath, J., Saha, S., Welham, J., El Saadi, O., MacCauley, C., & Chant, D. (2004b). A systematic review of the incidence of schizophrenia: the distribution of rates and the influence of sex, urbanicity, migrant status and methodology. *BMC Med, 2*, 13.
McIntosh, A. M., Gow, A., Luciano, M., Davies, G., Liewald, D. C., Harris, S. E., et al. (2013). Polygenic risk for schizophrenia is associated with cognitive change between childhood and old age. *Biol Psychiatry, 73*(10), 938-943.
McRae, Allan F, Wright, Margaret J, Hansell, Narelle K, Montgomery, Grant W, & Martin, Nicholas G. (2013). No Association Between General Cognitive Ability and Rare Copy Number Variation. *Behavior genetics*, 1-6.
Meier, M. H., Caspi, A., Reichenberg, A., Keefe, R. S., Fisher, H. L., Harrington, H., et al. (2014). Neuropsychological decline in schizophrenia from the premorbid to the postonset period: evidence from a population-representative longitudinal study. *Am J Psychiatry, 171*(1), 91-101.
Merikangas, A. K., Segurado, R., Cormican, P., Heron, E. A., Anney, R. J., Moore, S., et al. (2014). The phenotypic manifestations of rare CNVs in schizophrenia. *Schizophr Res*.
Merikangas, K. R., Akiskal, H. S., Angst, J., Greenberg, P. E., Hirschfeld, R. M., Petukhova, M., et al. (2007). Lifetime and 12-month prevalence of bipolar

spectrum disorder in the National Comorbidity Survey replication. *Arch Gen Psychiatry, 64*(5), 543-552.
Milnik, A., Heck, A., Vogler, C., Heinze, H. J., de Quervain, D. J., & Papassotiropoulos, A. (2012). Association of KIBRA with episodic and working memory: a meta-analysis. *Am J Med Genet B Neuropsychiatr Genet, 159B*(8), 958-969.
Mohn, Christine, Sundet, Kjetil, & Rund, Bjørn Rishovd. (2014). The relationship between IQ and performance on the MATRICS consensus cognitive battery. *Schizophrenia Research: Cognition, 1*(2), 96-100.
Morciano, M., Beckhaus, T., Karas, M., Zimmermann, H., & Volknandt, W. (2009). The proteome of the presynaptic active zone: from docked synaptic vesicles to adhesion molecules and maxi-channels. *J Neurochem, 108*(3), 662-675.
Moreno-De-Luca, D., Consortium, Sgene, Mulle, J. G., Simons Simplex Collection Genetics, Consortium, Kaminsky, E. B., Sanders, S. J., et al. (2010). Deletion 17q12 is a recurrent copy number variant that confers high risk of autism and schizophrenia. *Am J Hum Genet, 87*(5), 618-630.
Morice, R., & Delahunty, A. (1996). Frontal/executive impairments in schizophrenia. *Schizophr Bull, 22*(1), 125-137.
Morton, B. Brown. (1975). A Method for Combining Non-Independent, One-Sided Tests of Significance. *Biometrics, 31*(4), 987-992.
Moskvina, V., O'Dushlaine, C., Purcell, S., Craddock, N., Holmans, P., & O'Donovan, M. C. (2011). Evaluation of an approximation method for assessment of overall significance of multiple-dependent tests in a genomewide association study. *Genet Epidemiol, 35*(8), 861-866.
Muglia, P., Vicente, A. M., Verga, M., King, N., Macciardi, F., & Kennedy, J. L. (2003). Association between the BDNF gene and schizophrenia. *Mol Psychiatry, 8*(2), 146-147.
Muller, C. S., Haupt, A., Bildl, W., Schindler, J., Knaus, H. G., Meissner, M., et al. (2010). Quantitative proteomics of the Cav2 channel nano-environments in the mammalian brain. *Proc Natl Acad Sci U S A, 107*(34), 14950-14957.
Munafo, M. R., Bowes, L., Clark, T. G., & Flint, J. (2005). Lack of association of the COMT (Val158/108 Met) gene and schizophrenia: a meta-analysis of case-control studies. *Mol Psychiatry, 10*(8), 765-770.
Murphy, K. C., Jones, L. A., & Owen, M. J. (1999). High rates of schizophrenia in adults with velo-cardio-facial syndrome. *Arch Gen Psychiatry, 56*(10), 940-945.
Murray, R. M., Sham, P., Van Os, J., Zanelli, J., Cannon, M., & McDonald, C. (2004). A developmental model for similarities and dissimilarities between schizophrenia and bipolar disorder. *Schizophr Res, 71*(2-3), 405-416.
Murre, C. (2005). Helix-loop-helix proteins and lymphocyte development. *Nat Immunol, 6*(11), 1079-1086.
Need, A. C., Attix, D. K., McEvoy, J. M., Cirulli, E. T., Linney, K. N., Wagoner, A. P., et al. (2008). Failure to replicate effect of Kibra on human memory in two large cohorts of European origin. *Am J Med Genet B Neuropsychiatr Genet, 147B*(5), 667-668.
Need, A. C., McEvoy, J. P., Gennarelli, M., Heinzen, E. L., Ge, D., Maia, J. M., et al. (2012). Exome sequencing followed by large-scale genotyping suggests a

limited role for moderately rare risk factors of strong effect in schizophrenia. *Am J Hum Genet, 91*(2), 303-312.
Ng, M. Y., Levinson, D. F., Faraone, S. V., Suarez, B. K., DeLisi, L. E., Arinami, T., et al. (2009). Meta-analysis of 32 genome-wide linkage studies of schizophrenia. *Mol Psychiatry, 14*(8), 774-785.
Niarchou, M., Zammit, S., van Goozen, S. H., Thapar, A., Tierling, H. M., Owen, M. J., et al. (2014). Psychopathology and cognition in children with 22q11.2 deletion syndrome. *Br J Psychiatry, 204*(1), 46-54.
Niarchou, M., Zammit, S., Walters, J., Lewis, G., Owen, M. J., & van den Bree, M. B. (2013). Defective processing speed and nonclinical psychotic experiences in children: longitudinal analyses in a large birth cohort. *Am J Psychiatry, 170*(5), 550-557.
Nicodemus, K. K., Hargreaves, A., Morris, D., Anney, R., Gill, M., Corvin, A., et al. (2014). Variability in Working Memory Performance Explained by Epistasis vs Polygenic Scores in the ZNF804A Pathway. *JAMA Psychiatry, 71*(7), 778-785.
Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D. F., et al. (2012). Intelligence: new findings and theoretical developments. *Am Psychol, 67*(2), 130-159.
Nowicki, Stephen, Jr., & Duke, MarshallP. (1994). Individual differences in the nonverbal communication of affect: The diagnostic analysis of nonverbal accuracy scale. *Journal of Nonverbal Behavior, 18*(1), 9-35.
Nuechterlein, K. H., Barch, D. M., Gold, J. M., Goldberg, T. E., Green, M. F., & Heaton, R. K. (2004). Identification of separable cognitive factors in schizophrenia. *Schizophr Res, 72*(1), 29-39.
Nuechterlein, K. H., & Dawson, M. E. (1984). Information processing and attentional functioning in the developmental course of schizophrenic disorders. *Schizophr Bull, 10*(2), 160-203.
Nuechterlein, K. H., & Green, M. F. (2006). *MATRICS Consensus Cognitive Battery Manual*. Los Angeles: MATRICS Assessment Inc.
Nuechterlein, K. H., Green, M. F., Kern, R. S., Baade, L. E., Barch, D. M., Cohen, J. D., et al. (2008). The MATRICS Consensus Cognitive Battery, part 1: test selection, reliability, and validity. *Am J Psychiatry, 165*(2), 203-213.
Numakawa, T., Yagasaki, Y., Ishimoto, T., Okada, T., Suzuki, T., Iwata, N., et al. (2004). Evidence of novel neuronal functions of dysbindin, a susceptibility gene for schizophrenia. *Hum Mol Genet, 13*(21), 2699-2708.
O'Donovan, M. C., Craddock, N., Norton, N., Williams, H., Peirce, T., Moskvina, V., et al. (2008a). Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nat Genet, 40*(9), 1053-1055.
O'Donovan, M. C., Kirov, G., & Owen, M. J. (2008b). Phenotypic variations on the theme of CNVs. *Nat Genet, 40*(12), 1392-1393.
Olabi, B., Ellison-Wright, I., McIntosh, A. M., Wood, S. J., Bullmore, E., & Lawrie, S. M. (2011). Are there progressive brain changes in schizophrenia? A meta-analysis of structural magnetic resonance imaging studies. *Biol Psychiatry, 70*(1), 88-96.
Oliveira, A. L., Thams, S., Lidman, O., Piehl, F., Hokfelt, T., Karre, K., et al. (2004). A role for MHC class I molecules in synaptic plasticity and regeneration of neurons after axotomy. *Proc Natl Acad Sci U S A, 101*(51), 17843-17848.

Oliveira, A. M., Hemstedt, T. J., & Bading, H. (2012). Rescue of aging-associated decline in Dnmt3a2 expression restores cognitive abilities. *Nat Neurosci, 15*(8), 1111-1113.

Orozco, I. J., Koppensteiner, P., Ninan, I., & Arancio, O. (2014). The schizophrenia susceptibility gene DTNBP1 modulates AMPAR synaptic transmission and plasticity in the hippocampus of juvenile DBA/2J mice. *Mol Cell Neurosci, 58*, 76-84.

Owens, S. F., Picchioni, M. M., Rijsdijk, F. V., Stahl, D., Vassos, E., Rodger, A. K., et al. (2011). Genetic overlap between episodic memory deficits and schizophrenia: results from the Maudsley Twin Study. *Psychol Med, 41*(3), 521-532.

Pak, J. H., Huang, F. L., Li, J., Balschun, D., Reymann, K. G., Chiang, C., et al. (2000). Involvement of neurogranin in the modulation of calcium/calmodulin-dependent protein kinase II, synaptic plasticity, and spatial learning: a study with knockout mice. *Proc Natl Acad Sci U S A, 97*(21), 11232-11237.

Pantelis, C., Yucel, M., Wood, S. J., Velakoulis, D., Sun, D., Berger, G., et al. (2005). Structural brain imaging evidence for multiple pathological processes at different stages of brain development in schizophrenia. *Schizophr Bull, 31*(3), 672-696.

Papassotiropoulos, A., Stephan, D. A., Huentelman, M. J., Hoerndli, F. J., Craig, D. W., Pearson, J. V., et al. (2006). Common Kibra alleles are associated with human memory performance. *Science, 314*(5798), 475-478.

Park, S., Puschel, J., Sauter, B. H., Rentsch, M., & Hell, D. (1999). Spatial working memory deficits and clinical symptoms in schizophrenia: a 4-month follow-up study. *Biol Psychiatry, 46*(3), 392-400.

Paulo, J. A., Brucker, W. J., & Hawrot, E. (2009). Proteomic analysis of an alpha7 nicotinic acetylcholine receptor interactome. *J Proteome Res, 8*(4), 1849-1858.

Pedersen, N. L., Plomin, R., Nesselroade, J. R., & Mcclearn, G. E. (1992). A Quantitative Genetic-Analysis of Cognitive-Abilities during the 2nd-Half of the Life-Span. *Psychol Sci, 3*(6), 346-353.

PGC Bipolar Disorder Working Group. (2011). Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet, 43*(10), 977-983.

Phillips, M. L., & Kupfer, D. J. (2013). Bipolar disorder diagnosis: challenges and future directions. *Lancet, 381*(9878), 1663-1671.

Plomin, R., Haworth, C. M. A., Meaburn, E. L., Price, T. S., Davis, O. S. P., & Control, Wellcome Trust Case. (2013). Common DNA Markers Can Account for More Than Half of the Genetic Influence on Cognitive Abilities. *Psychol Sci, 24*(4), 562-568.

Plomin, R., Pedersen, N. L., Lichtenstein, P., & McClearn, G. E. (1994). Variability and stability in cognitive abilities are largely genetic later in life. *Behav Genet, 24*(3), 207-215.

Pocklington, A. J., Cumiskey, M., Armstrong, J. D., & Grant, S. G. (2006). The proteomes of neurotransmitter receptor complexes form modular networks with distributed functionality underlying plasticity and behaviour. *Mol Syst Biol, 2*, 2006 0023.

Porteous, D. J., Thomson, P. A., Millar, J. K., Evans, K. L., Hennah, W., Soares, D. C., et al. (2014). DISC1 as a genetic risk factor for schizophrenia and related major mental illness: response to Sullivan. *Mol Psychiatry, 19*(2), 141-143.
Posthuma, D., Baare, W. F. C., Pol, H. E. H., Kahn, R. S., Boomsma, D. I., & De Geus, E. J. C. (2003). Genetic correlations between brain volumes and the WAIS-III dimensions of verbal comprehension, working memory, perceptual organization, and processing speed. *Twin Research, 6*(2), 131-139.
Posthuma, D., de Geus, E. J., & Boomsma, D. I. (2001). Perceptual speed and IQ are associated through common genetic factors. *Behav Genet, 31*(6), 593-602.
Posthuma, D., Luciano, M., Geus, E. J., Wright, M. J., Slagboom, P. E., Montgomery, G. W., et al. (2005). A genomewide scan for intelligence identifies quantitative trait loci on 2q and 6p. *Am J Hum Genet, 77*(2), 318-326.
Poulain, F. E., & Sobel, A. (2007). The "SCG10-LIke Protein" SCLIP is a novel regulator of axonal branching in hippocampal neurons, unlike SCG10. *Mol Cell Neurosci, 34*(2), 137-146.
Poulton, R., Caspi, A., Moffitt, T. E., Cannon, M., Murray, R., & Harrington, H. (2000). Children's self-reported psychotic symptoms and adult schizophreniform disorder: a 15-year longitudinal study. *Arch Gen Psychiatry, 57*(11), 1053-1058.
Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet, 38*(8), 904-909.
Psychiatric GWAS Consortium Steering Committee. (2009). A framework for interpreting genome-wide association studies of psychiatric disorders. *Mol Psychiatry, 14*(1), 10-17.
Purcell, S. (2002). Variance components models for gene-environment interaction in twin analysis. *Twin Research, 5*(6), 554-571.
Purcell, S. M., Moran, J. L., Fromer, M., Ruderfer, D., Solovieff, N., Roussos, P., et al. (2014). A polygenic burden of rare disruptive mutations in schizophrenia. *Nature, 506*(7487), 185-190.
Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet, 81*(3), 559-575.
Quisenaerts, C., Morrens, M., Hulstijn, W., de Bruijn, E., Timmers, M., Streffer, J., et al. (2014). The nicotinergic receptor as a target for cognitive enhancement in schizophrenia: Barking up the wrong tree? *Psychopharmacology, 231*(3), 543-550.
Ramanan, V. K., Shen, L., Moore, J. H., & Saykin, A. J. (2012). Pathway analysis of genomic data: concepts, methods, and prospects for future development. *Trends Genet, 28*(7), 323-332.
Redon, R., Ishikawa, S., Fitch, K. R., Feuk, L., Perry, G. H., Andrews, T. D., et al. (2006). Global variation in copy number in the human genome. *Nature, 444*(7118), 444-454.
Rees, E., Kirov, G., O'Donovan, M. C., & Owen, M. J. (2012). De novo mutation in schizophrenia. *Schizophr Bull, 38*(3), 377-381.
Rees, E., Kirov, G., Sanders, A., Walters, J. T., Chambert, K. D., Shi, J., et al. (2014a). Evidence that duplications of 22q11.2 protect against schizophrenia. *Mol Psychiatry, 19*(1), 37-40.

Rees, E., Moskvina, V., Owen, M. J., O'Donovan, M. C., & Kirov, G. (2011). De novo rates and selection of schizophrenia-associated copy number variants. *Biol Psychiatry, 70*(12), 1109-1114.
Rees, E., Walters, J. T. R., Georgieva, L., Isles, A. R., Chambert, K. D., Richards, A. L., et al. (2014b). Analysis of copy number variations at 15 schizophrenia-associated loci. *British Journal of Psychiatry, 204*(2), 108-114.
Reichenberg, A., Caspi, A., Harrington, H., Houts, R., Keefe, R. S., Murray, R. M., et al. (2010). Static and dynamic cognitive deficits in childhood preceding adult schizophrenia: a 30-year study. *Am J Psychiatry, 167*(2), 160-169.
Reichenberg, Abraham. (2005). Cognitive impairment as a risk factor for psychosis. *Dialogues Clin Neurosci, 7*(1), 31-38.
Reske, M., Habel, U., Kellermann, T., Backes, V., Jon Shah, N., von Wilmsdorff, M., et al. (2009). Differential brain activation during facial emotion discrimination in first-episode schizophrenia. *J Psychiatr Res, 43*(6), 592-599.
Rhee, S. Y., Wood, V., Dolinski, K., & Draghici, S. (2008). Use and misuse of the gene ontology annotations. *Nature Reviews Genetics, 9*(7), 509-515.
Rietveld, C. A., Medland, S. E., Derringer, J., Yang, J., Esko, T., Martin, N. W., et al. (2013). GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science, 340*(6139), 1467-1471.
Rijsdijk, F. V., Vernon, P. A., & Boomsma, D. I. (2002). Application of hierarchical genetic models to Raven and WAIS subtests: a Dutch twin study. *Behav Genet, 32*(3), 199-210.
Riley, B., & Kendler, K. S. (2006). Molecular genetic studies of schizophrenia. *Eur J Hum Genet, 14*(6), 669-680.
Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1996). The structure of normal human attention: The Test of Everyday Attention. *J Int Neuropsychol Soc, 2*(6), 525-534.
Robinson, L. J., Thompson, J. M., Gallagher, P., Goswami, U., Young, A. H., Ferrier, I. N., et al. (2006). A meta-analysis of cognitive deficits in euthymic patients with bipolar disorder. *J Affect Disord, 93*(1-3), 105-115.
Rodriguez-Sanchez, J. M., Crespo-Facorro, B., Gonzalez-Blanch, C., Perez-Iglesias, R., Vazquez-Barquero, J. L., & Study, Pafip Group. (2007). Cognitive dysfunction in first-episode psychosis: the processing speed hypothesis. *Br J Psychiatry Suppl, 51*, s107-110.
Rosenheck, R., Leslie, D., Keefe, R., McEvoy, J., Swartz, M., Perkins, D., et al. (2006). Barriers to employment for people with schizophrenia. *Am J Psychiatry, 163*(3), 411-417.
Roussos, P., Giakoumaki, S. G., Georgakopoulos, A., Robakis, N. K., & Bitsios, P. (2011). The CACNA1C and ANK3 risk alleles impact on affective personality traits and startle reactivity but not on cognition or gating in healthy males. *Bipolar Disorders, 13*(3), 250-259.
Ruano, D., Abecasis, G. R., Glaser, B., Lips, E. S., Cornelisse, L. N., de Jong, A. P., et al. (2010). Functional gene group analysis reveals a role of synaptic heterotrimeric G proteins in cognitive ability. *Am J Hum Genet, 86*(2), 113-125.
Ruderfer, D. M., Fanous, A. H., Ripke, S., McQuillin, A., Amdur, R. L., Schizophrenia Working Group of Psychiatric Genomics, Consortium, et al. (2014).

Polygenic dissection of diagnosis and clinical dimensions of bipolar disorder and schizophrenia. *Mol Psychiatry, 19*(9), 1017-1024.
Rujescu, D., Hartmann, A. M., Gonnermann, C., Moller, H. J., & Giegling, I. (2003). M129V variation in the prion protein may influence cognitive performance. *Mol Psychiatry, 8*(11), 937-941.
Rujescu, D., Ingason, A., Cichon, S., Pietilainen, O. P., Barnes, M. R., Toulopoulou, T., et al. (2009). Disruption of the neurexin 1 gene is associated with schizophrenia. *Hum Mol Genet, 18*(5), 988-996.
Saha, S., Chant, D., & McGrath, J. (2007). A systematic review of mortality in schizophrenia: is the differential mortality gap worsening over time? *Arch Gen Psychiatry, 64*(10), 1123-1131.
Sanders, A. R. (2008). No significant association of 14 candidate genes with schizophrenia in a large European ancestry sample: Implications for psychiatric genetics (vol 165, pg 497, 2008). *American Journal of Psychiatry, 165*(10), 1359-1359.
Sanderson, T. M., Cotel, M. C., O'Neill, M. J., Tricklebank, M. D., Collingridge, G. L., & Sher, E. (2012). Alterations in hippocampal excitability, synaptic transmission and synaptic plasticity in a neurodevelopmental model of schizophrenia. *Neuropharmacology, 62*(3), 1349-1358.
Schadt, E. E. (2009). Molecular networks as sensors and drivers of common human diseases. *Nature, 461*(7261), 218-223.
Schaper, K., Kolsch, H., Popp, J., Wagner, M., & Jessen, F. (2008). KIBRA gene variants are associated with episodic memory in healthy elderly. *Neurobiol Aging, 29*(7), 1123-1125.
Schenkel, L. S., West, A. E., Jacobs, R., Sweeney, J. A., & Pavuluri, M. N. (2012). Cognitive dysfunction is worse among pediatric patients with bipolar disorder Type I than Type II. *J Child Psychol Psychiatry, 53*(7), 775-781.
Schiffman, J., Walker, E., Ekstrom, M., Schulsinger, F., Sorensen, H., & Mednick, S. (2004). Childhood videotaped social and neuromotor precursors of schizophrenia: a prospective investigation. *Am J Psychiatry, 161*(11), 2021-2027.
Schizophrenia PGC. (2011). Genome-wide association study identifies five new schizophrenia loci. *Nat Genet, 43*(10), 969-976.
Schizophrenia Working Group of the Psychiatric Genetics Consortium. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature, 511*(7510), 421-427.
Schmidt, S. J., Mueller, D. R., & Roder, V. (2011). Social cognition as a mediator variable between neurocognition and functional outcome in schizophrenia: empirical review and new results by structural equation modeling. *Schizophr Bull, 37 Suppl 2*, S41-54.
Sebat, J., Lakshmi, B., Troge, J., Alexander, J., Young, J., Lundin, P., et al. (2004). Large-scale copy number polymorphism in the human genome. *Science, 305*(5683), 525-528.
Sebat, J., Levy, D. L., & McCarthy, S. E. (2009). Rare structural variants in schizophrenia: one disorder, multiple mutations; one mutation, multiple disorders. *Trends Genet, 25*(12), 528-535.
Seidman, L. J., Lanca, M., Kremen, W. S., Faraone, S. V., & Tsuang, M. T. (2003). Organizational and visual memory deficits in schizophrenia and bipolar

psychoses using the Rey-Osterrieth complex figure: effects of duration of illness. *J Clin Exp Neuropsychol, 25*(7), 949-964.
Sheppard, L. D., & Vernon, P. A. (2008). Intelligence and speed of information-processing: A review of 50 years of research. *Personality and Individual Differences, 44*(3), 535-551.
Sheward, W. J., Lutz, E. M., & Harmar, A. J. (1995). The distribution of vasoactive intestinal peptide2 receptor messenger RNA in the rat brain and pituitary gland as assessed by in situ hybridization. *Neuroscience, 67*(2), 409-418.
Simonsen, C., Sundet, K., Vaskinn, A., Birkenaes, A. B., Engh, J. A., Faerden, A., et al. (2011). Neurocognitive dysfunction in bipolar and schizophrenia spectrum disorders depends on history of psychosis rather than diagnostic group. *Schizophr Bull, 37*(1), 73-83.
Smoller, J. W., Craddock, N., Kendler, K., Lee, P. H., Neale, B. M., Nurnberger, J. I., et al. (2013). Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet, 381*(9875), 1371-1379.
Snitz, B. E., Macdonald, A. W., 3rd, & Carter, C. S. (2006). Cognitive deficits in unaffected first-degree relatives of schizophrenia patients: a meta-analytic review of putative endophenotypes. *Schizophr Bull, 32*(1), 179-194.
Sorensen, H. J., Saebye, D., Urfer-Parnas, A., Mortensen, E. L., & Parnas, J. (2012). Premorbid intelligence and educational level in bipolar and unipolar disorders: a Danish draft board study. *J Affect Disord, 136*(3), 1188-1191.
Spearman, C. (1904). "General Intelligence," Objectively Determined and Measured. *The American Journal of Psychology, 15*(2), 201-292.
Splawski, I., Timothy, K. W., Sharpe, L. M., Decher, N., Kumar, P., Bloise, R., et al. (2004). Ca(V)1.2 calcium channel dysfunction causes a multisystem disorder including arrhythmia and autism. *Cell, 119*(1), 19-31.
Stathopoulou, A., Beratis, I. N., & Beratis, S. (2013). Prenatal tobacco smoke exposure, risk of schizophrenia, and severity of positive/negative symptoms. *Schizophr Res, 148*(1-3), 105-110.
Stclair, D., Blackwood, D., Muir, W., Carothers, A., Walker, M., Spowart, G., et al. (1990). Association within a Family of a Balanced Autosomal Translocation with Major Mental-Illness. *Lancet, 336*(8706), 13-16.
Stefansson, H., Meyer-Lindenberg, A., Steinberg, S., Magnusdottir, B., Morgen, K., Arnarsdottir, S., et al. (2014). CNVs conferring risk of autism or schizophrenia affect cognition in controls. *Nature, 505*(7483), 361-366.
Stefansson, H., Ophoff, R. A., Steinberg, S., Andreassen, O. A., Cichon, S., Rujescu, D., et al. (2009). Common variants conferring risk of schizophrenia. *Nature, 460*(7256), 744-747.
Stranger, B. E., Forrest, M. S., Dunning, M., Ingle, C. E., Beazley, C., Thorne, N., et al. (2007). Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science, 315*(5813), 848-853.
Straub, R. E., MacLean, C. J., O'Neill, F. A., Burke, J., Murphy, B., Duke, F., et al. (1995). A potential vulnerability locus for schizophrenia on chromosome 6p24-22: evidence for genetic heterogeneity. *Nat Genet, 11*(3), 287-293.
Strous, R. D., & Shoenfeld, Y. (2006). Schizophrenia, autoimmunity and immune system dysregulation: a comprehensive model updated and revisited. *J Autoimmun, 27*(2), 71-80.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A, 102*(43), 15545-15550.

Sullivan, P. F. (2010). The psychiatric GWAS consortium: big science comes to psychiatry. *Neuron, 68*(2), 182-186.

Sullivan, P. F. (2013). Questions about DISC1 as a genetic risk factor for schizophrenia. *Mol Psychiatry, 18*(10), 1050-1052.

Sullivan, P. F., Kendler, K. S., & Neale, M. C. (2003). Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Arch Gen Psychiatry, 60*(12), 1187-1192.

Sullivan, P. F., Lin, D., Tzeng, J. Y., van den Oord, E., Perkins, D., Stroup, T. S., et al. (2008). Genomewide association for schizophrenia in the CATIE study: results of stage 1. *Mol Psychiatry, 13*(6), 570-584.

Takamori, S., Holt, M., Stenius, K., Lemke, E. A., Gronborg, M., Riedel, D., et al. (2006). Molecular anatomy of a trafficking organelle. *Cell, 127*(4), 831-846.

Terwisscha van Scheltinga, A., Bakker, S., van Haren, N., Buizer-Voskamp, J., Boos, H., Vorstman, J., et al. (2012). Association study of copy number variants with brain volume in schizophrenia patients and healthy controls. *Psychiatry Res, 200*(2-3), 1011-1013.

Tewes, Uwe. (1991). Hamburg-Wechsler-Intelligenztest für Erwachsene-Revision 1991 (HAWIE-R). *Huber, Bern*.

Thibaut, F., Coron, B., Hannequin, D., Segard, L., Martin, C., Dollfus, S., et al. (1998). No association of apolipoprotein epsilon 4 allele with schizophrenia even in cognitively impaired patients. *Schizophr Res, 30*(2), 149-153.

Tiihonen, J., Haukka, J., Henriksson, M., Cannon, M., Kieseppa, T., Laaksonen, I., et al. (2005). Premorbid intellectual functioning in bipolar disorder and schizophrenia: results from a cohort study of male conscripts. *Am J Psychiatry, 162*(10), 1904-1910.

Torkamani, A., Topol, E. J., & Schork, N. J. (2008). Pathway analysis of seven common diseases assessed by genome-wide association. *Genomics, 92*(5), 265-272.

Toulopoulou, T., Goldberg, T. E., Mesa, I. R., Picchioni, M., Rijsdijk, F., Stahl, D., et al. (2010). Impaired intellect and memory: a missing link between genetic risk and schizophrenia? *Arch Gen Psychiatry, 67*(9), 905-913.

Toulopoulou, T., Picchioni, M., Rijsdijk, F., Hua-Hall, M., Ettinger, U., Sham, P., et al. (2007). Substantial genetic overlap between neurocognition and schizophrenia: genetic modeling in twin samples. *Arch Gen Psychiatry, 64*(12), 1348-1355.

Trzaskowski, M., Harlaar, N., Arden, R., Krapohl, E., Rimfeld, K., McMillan, A., et al. (2014). Genetic influence on family socioeconomic status and children's intelligence. *Intelligence, 42*(100), 83-88.

Trzaskowski, M., Yang, J., Visscher, P. M., & Plomin, R. (2013). DNA evidence for strong genetic stability and increasing heritability of intelligence from age 7 to 12. *Mol Psychiatry*.

Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., & Gottesman, II. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychol Sci, 14*(6), 623-628.
Ursu, S., Kring, A. M., Gard, M. G., Minzenberg, M. J., Yoon, J. H., Ragland, J. D., et al. (2011). Prefrontal cortical deficits and impaired cognition-emotion interactions in schizophrenia. *Am J Psychiatry, 168*(3), 276-285.
Vacic, V., McCarthy, S., Malhotra, D., Murray, F., Chou, H. H., Peoples, A., et al. (2011). Duplications of the neuropeptide receptor gene VIPR2 confer significant risk for schizophrenia. *Nature, 471*(7339), 499-503.
Valenstein, M., Blow, F. C., Copeland, L. A., McCarthy, J. F., Zeber, J. E., Gillon, L., et al. (2004). Poor antipsychotic adherence among patients with schizophrenia: medication and patient factors. *Schizophr Bull, 30*(2), 255-264.
van Os, J., & Kapur, S. (2009). Schizophrenia. *Lancet, 374*(9690), 635-645.
van Os, J., Kenis, G., & Rutten, B. P. (2010). The environment and schizophrenia. *Nature, 468*(7321), 203-212.
Van Rheenen, T. E., & Rossell, S. L. (2014). An empirical evaluation of the MATRICS Consensus Cognitive Battery in bipolar disorder. *Bipolar Disorders, 16*(3), 318-325.
van Scheltinga, A. F., Bakker, S. C., van Haren, N. E., Derks, E. M., Buizer-Voskamp, J. E., Cahn, W., et al. (2013). Schizophrenia genetic variants are not associated with intelligence. *Psychol Med*, 1-8.
Vassos, E., Bramon, E., Picchioni, M., Walshe, M., Filbey, F. M., Kravariti, E., et al. (2010). Evidence of association of KIBRA genotype with episodic memory in families of psychotic patients and controls. *J Psychiatr Res, 44*(12), 795-798.
Vassos, E., Pedersen, C. B., Murray, R. M., Collier, D. A., & Lewis, C. M. (2012). Meta-analysis of the association of urbanicity with schizophrenia. *Schizophr Bull, 38*(6), 1118-1123.
Veyrieras, J. B., Kudaravalli, S., Kim, S. Y., Dermitzakis, E. T., Gilad, Y., Stephens, M., et al. (2008). High-resolution mapping of expression-QTLs yields insight into human gene regulation. *PLoS Genet, 4*(10), e1000214.
Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era--concepts and misconceptions. *Nat Rev Genet, 9*(4), 255-266.
Walker, E. F., Grimes, K. E., Davis, D. M., & Smith, A. J. (1993). Childhood precursors of schizophrenia: facial expressions of emotion. *Am J Psychiatry, 150*(11), 1654-1660.
Walters, J. T., Corvin, A., Owen, M. J., Williams, H., Dragovic, M., Quinn, E. M., et al. (2010). Psychosis susceptibility gene ZNF804A and cognitive performance in schizophrenia. *Arch Gen Psychiatry, 67*(7), 692-700.
Walters, J. T., & Owen, M. J. (2007). Endophenotypes in psychiatric genetics. *Mol Psychiatry, 12*(10), 886-890.
Walters, J. T., Rujescu, D., Franke, B., Giegling, I., Vasquez, A. A., Hargreaves, A., et al. (2013). The role of the major histocompatibility complex region in cognition and brain structure: a schizophrenia GWAS follow-up. *Am J Psychiatry, 170*(8), 877-885.
Wang, K., Li, M., & Hakonarson, H. (2010). Analysing biological pathways in genome-wide association studies. *Nat Rev Genet, 11*(12), 843-854.

Wechsler, David, Golombok, J, & Rust, John. (1992). WISC-III UK Wechsler Intelligence Scale for Children: UK Manual. *Sidcup, UK: The Psychological Corporation*.

Weiser, M., Heresco-Levy, U., Davidson, M., Javitt, D. C., Werbeloff, N., Gershon, A. A., et al. (2012). A multicenter, add-on randomized controlled trial of low-dose d-serine for negative and cognitive symptoms of schizophrenia. *J Clin Psychiatry, 73*(6), e728-734.

Weiss, L. A., Shen, Y., Korn, J. M., Arking, D. E., Miller, D. T., Fossdal, R., et al. (2008). Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med, 358*(7), 667-675.

Weitzdoerfer, R., Hoeger, H., Engidawork, E., Engelmann, M., Singewald, N., Lubec, G., et al. (2004). Neuronal nitric oxide synthase knock-out mice show impaired cognitive performance. *Nitric Oxide, 10*(3), 130-140.

Wheeler, A. L., Chakravarty, M. M., Lerch, J. P., Pipitone, J., Daskalakis, Z. J., Rajji, T. K., et al. (2014). Disrupted prefrontal interhemispheric structural coupling in schizophrenia related to working memory performance. *Schizophr Bull, 40*(4), 914-924.

Williams, H. J., Glaser, B., Williams, N. M., Norton, N., Zammit, S., MacGregor, S., et al. (2005). No association between schizophrenia and polymorphisms in COMT in two large samples. *Am J Psychiatry, 162*(9), 1736-1738.

Williams, H. J., Norton, N., Dwyer, S., Moskvina, V., Nikolov, I., Carroll, L., et al. (2011a). Fine mapping of ZNF804A and genome-wide significant evidence for its involvement in schizophrenia and bipolar disorder. *Mol Psychiatry, 16*(4), 429-441.

Williams, H. J., Owen, M. J., & O'Donovan, M. C. (2007). Is COMT a susceptibility gene for schizophrenia? *Schizophr Bull, 33*(3), 635-641.

Williams, Hywel J., Moskvina, Valentina, Smith, Rhodri L., Dwyer, Sarah, Russo, Giancarlo, Owen, Michael J., et al. (2011b). Association between TCF4 and schizophrenia does not exert its effect by common nonsynonymous variation or by influencing cis-acting regulation of mRNA expression in adult human brain. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 156*(7), 781-784.

Wing, J. K., Babor, T., Brugha, T., Burke, J., Cooper, J. E., Giel, R., et al. (1990). SCAN. Schedules for Clinical Assessment in Neuropsychiatry. *Arch Gen Psychiatry, 47*(6), 589-593.

Wisdom, N. M., Callahan, J. L., & Hawkins, K. A. (2011). The effects of apolipoprotein E on non-impaired cognitive functioning: a meta-analysis. *Neurobiol Aging, 32*(1), 63-74.

Wisner, K. M., Elvevag, B., Gold, J. M., Weinberger, D. R., & Dickinson, D. (2011). A closer look at siblings of patients with schizophrenia: the association of depression history and sex with cognitive phenotypes. *Schizophr Res, 126*(1-3), 164-173.

Wockner, L. F., Noble, E. P., Lawford, B. R., Young, R. M., Morris, C. P., Whitehall, V. L., et al. (2014). Genome-wide DNA methylation analysis of human brain tissue from schizophrenia patients. *Transl Psychiatry, 4*, e339.

Wood, S. J., Brewer, W. J., Koutsouradis, P., Phillips, L. J., Francey, S. M., Proffitt, T. M., et al. (2007). Cognitive decline following psychosis onset: data from the PACE clinic. *Br J Psychiatry Suppl, 51*, s52-57.

Woodberry, K. A., Giuliano, A. J., & Seidman, L. J. (2008). Premorbid IQ in schizophrenia: a meta-analytic review. *Am J Psychiatry, 165*(5), 579-587.
Woodward, N. D., Duffy, B., & Karbasforoushan, H. (2013). Prefrontal cortex activity during response selection predicts processing speed impairment in schizophrenia. *J Int Neuropsychol Soc, 19*(7), 782-791.
Woodward, N. D., Purdon, S. E., Meltzer, H. Y., & Zald, D. H. (2005). A meta-analysis of neuropsychological change to clozapine, olanzapine, quetiapine, and risperidone in schizophrenia. *Int J Neuropsychopharmacol, 8*(3), 457-472.
Wright, M., De Geus, E., Ando, J., Luciano, M., Posthuma, D., Ono, Y., et al. (2001). Genetics of cognition: outline of a collaborative twin study. *Twin Research, 4*(1), 48-56.
Wu, M. C., Kraft, P., Epstein, M. P., Taylor, D. M., Chanock, S. J., Hunter, D. J., et al. (2010). Powerful SNP-set analysis for case-control genome-wide association studies. *Am J Hum Genet, 86*(6), 929-942.
Wykes, T., Huddy, V., Cellard, C., McGurk, S. R., & Czobor, P. (2011). A meta-analysis of cognitive remediation for schizophrenia: methodology and effect sizes. *Am J Psychiatry, 168*(5), 472-485.
Xu, B., Ionita-Laza, I., Roos, J. L., Boone, B., Woodrick, S., Sun, Y., et al. (2012). De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. *Nat Genet, 44*(12), 1365-1369.
Xu, B., Roos, J. L., Levy, S., van Rensburg, E. J., Gogos, J. A., & Karayiorgou, M. (2008). Strong association of de novo copy number mutations with sporadic schizophrenia. *Nat Genet, 40*(7), 880-885.
Xu, M. Q., St Clair, D., & He, L. (2006). Meta-analysis of association between ApoE epsilon4 allele and schizophrenia. *Schizophr Res, 84*(2-3), 228-235.
Yan, J., Noltner, K., Feng, J., Li, W., Schroer, R., Skinner, C., et al. (2008). Neurexin 1alpha structural variants associated with autism. *Neurosci Lett, 438*(3), 368-370.
Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., et al. (2010a). Common SNPs explain a large proportion of the heritability for human height. *Nat Genet, 42*(7), 565-569.
Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet, 88*(1), 76-82.
Yang, K., Lei, G., Jackson, M. F., & Macdonald, J. F. (2010b). The involvement of PACAP/VIP system in the synaptic transmission in the hippocampus. *J Mol Neurosci, 42*(3), 319-326.
Yang, Y., Paspalas, C. D., Jin, L. E., Picciotto, M. R., Arnsten, A. F., & Wang, M. (2013). Nicotinic alpha7 receptors enhance NMDA cognitive circuits in dorsolateral prefrontal cortex. *Proc Natl Acad Sci U S A, 110*(29), 12078-12083.
Yeo, R. A., Gangestad, S. W., Liu, J., Calhoun, V. D., & Hutchison, K. E. (2011). Rare copy number deletions predict individual variation in intelligence. *PLoS One, 6*(1), e16339.
Yeo, R. A., Gangestad, S. W., Liu, J., Ehrlich, S., Thoma, R. J., Pommy, J., et al. (2013). The impact of copy number deletions on general cognitive ability and ventricle size in patients with schizophrenia and healthy control subjects. *Biol Psychiatry, 73*(6), 540-545.

Yu, R., Zhao, L., & Lu, L. (2011). Regional grey and white matter changes in heavy male smokers. *PLoS One, 6*(11), e27440.
Zammit, S., Allebeck, P., David, A. S., Dalman, C., Hemmingsson, T., Lundberg, I., et al. (2004). A longitudinal study of premorbid IQ Score and risk of developing schizophrenia, bipolar disorder, severe depression, and other nonaffective psychoses. *Arch Gen Psychiatry, 61*(4), 354-360.
Zeng, L., Zhang, P., Shi, L., Yamamoto, V., Lu, W., & Wang, K. (2013). Functional impacts of NRXN1 knockdown on neurodevelopment in stem cell models. *PLoS One, 8*(3), e59685.
Zhabotinsky, A. M., Camp, R. N., Epstein, I. R., & Lisman, J. E. (2006). Role of the neurogranin concentrated in spines in the induction of long-term potentiation. *J Neurosci, 26*(28), 7337-7347.
Zuk, O., Hechter, E., Sunyaev, S. R., & Lander, E. S. (2012). The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc Natl Acad Sci U S A, 109*(4), 1193-1198.

7 Appendices

7.1 Appendix A – MATRICS scoring/imputation procedure

TMT: log transform the TMT score and then reverse the sign as lower means better performance

Z-scores: Calculate the means and standard for the controls for each test
Minus the control mean from each raw score and divide by the control standard deviation for each test. For domains with more than one test, sum the z-scores from the tests that make up the domain using just the control data again. Get a new mean and standard deviation from these summed z-scores. Minus the mean from each of these new summed z-scores and divide by the new standard deviation to get the domain z-score. For the composite score repeat as above, but sum all the new domain z-scores.

IMPUTATION

FORMULA:

Ydi = (Yd+) + (Y+i) - (Y++)

The imputed z-score for a missing task/domain for a individual "i" (Ydi) = (the mean z-score of that test for all cases) plus (the mean z-score of all completed tests for individual "i") minus (the mean z-score of all tests for all individuals)

For specific domains:

HVLT/BVMT: These only have summed scores in the control data. Not possible to impute one of the three subscores. Use imputed sum if more than two raw scores are missing. If first two scores present, get sum of these, and compute z-score from this, using control sd and mean of the first two subscores.

CPT - if all scores are missing, impute the mean score. However, if one or two subsets are missing, impute these separately and then derive the mean.

LNS/WMS - if one task is present and the other missing, impute the missing task and create a WM domain score with this. If both are missing, impute the domain score itself.

Speed of processing - individual missing test scores can be imputed for TMT/BACS and Fluency. If 2/3 tests are present, impute the missing test: this can be used to calculate domain score can be used towards composite.

If one or no tests completed impute the SOP domain score, but should not contribute towards the composite score.

Imputation scores should only be used when an individual has completed 5 or more domains. However, if hypothetically, an individual completes MSCEIT/BVMT/HVLT/Mazes but has a missing value for one of the SOP tasks, this individual task can be imputed, meaning the SOP domain can be calculated and thus the composite.

7.2 Appendix B – Candidate Pathways

7.2.1 MGI behavioral pathways

abnormal associative learning	anomaly in the ability to change the frequency or form of a behavior as a result of the influence of the environment
abnormal avoidance learning behavior	anomaly in the ability to associate a previously neutral stimulus with an unpleasant or punishing stimuli so that the animal learns to avoid the previously neutral stimulus
abnormal behavior	any anomaly in the actions, reactions, or performance of an organism in response to external or internal stimuli compared to controls
abnormal behavioral response to xenobiotic	any anomaly in the behavioral response induced by a foreign compound, such as consumption preference, induced hyperactivity or stereotypic behavior
abnormal chemical nociception	abnormal capability to sense pain elicited by chemical stimulation
abnormal circadian rhythm	deviation from the normal 24 hour biological activity cycle

abnormal conditioned place preference behavior	anomaly in the ability of an animal to learn and remember an association between a putative rewarding internal state produced by a xenobiotic or drug with a neutral, unchanging environment
abnormal contextual conditioning behavior	anomaly in the ability of an animal to learn and remember an association between an aversive experience (the unconditioned stimulus (US), usually a shock) and the neutral, unchanging environment (the conditioned stimulus (CS), or the environmental context in this case)
abnormal cued conditioning behavior	anomaly in the ability of an animal to learn and remember an association between an aversive experience (the unconditioned stimulus (US), usually a shock) and a neutral stimulus (the conditioned stimulus (CS), usually an auditory cue or light flash)
abnormal discrimination learning	anomaly in the ability to exhibit a differential response to different stimuli that is achieved by the reinforcement of the desired response for each particular stimulus

abnormal grooming behavior	defects in the standard of behavior of cleaning and/or keeping outward appearance tidy (self, mate or offspring)
abnormal involuntary movement	anomaly in movements that occur independent of planning (e.g. reflexive behavior)
abnormal mechanical nociception	abnormal capability to sense pain elicited by mechanical stimulation
abnormal motor learning	anomaly in the ability to repeat a motor task requiring well coordinated movements and balance; measures cerebellar dependent learning
abnormal nociception after inflammation	changes in pain thresholds after inflammation
abnormal object recognition memory	anomaly in the ability to recognize objects that the animal has previously encountered; recognition is measured by relative amount of time exploring objects, which should decrease upon subsequent or multiple presentations of the same object when presented with novel objects at the same time
abnormal pain threshold	increased or decreased average level of perception of pain
abnormal parental behavior	altered behavior of animals that affects the ability of offspring to survive

abnormal reflex	anomaly in an involuntary response to a peripheral stimulus
abnormal response to new environment	altered behavioral reaction associated with placing an animal in a new location
abnormal response to novelty	alteration in amount of exploration/investigation of a novel object, situation or environment
abnormal response to tactile stimuli	anomaly in the reflex action normally induced by touch or pain
abnormal seizure response to inducing agent	anomaly in the seizure activity response to an agent that normally can induce uncontrolled electrical activity in the brain, producing a physical convulsion and/or minor change in physical behavior
abnormal sexual interaction	altered initiation, failure of initiation or incomplete mating behavior
abnormal sleep behavior	any anomaly in the actions, reactions, or performance of an organism during a periodic, readily reversible state of reduced awareness and metabolic activity
abnormal social investigation	altered behavior of animals to approach and examine other animals
abnormal spatial learning	anomaly in the ability to ascertain or acquire spatial location information in

	order to improve navigation or other behavior using such location cues
abnormal spatial reference memory	anomaly in the ability to recall spatial location information from previous encounters or training sessions in order to navigate or perform other behavior using such location cues
abnormal spatial working memory	anomaly in the ability to spontaneously process spatial location information in order to navigate or perform other behavior using such location cues, without previous encounters or training at that location
abnormal temporal memory	anomaly in the ability to recall temporal events and stimuli
abnormal thermal nociception	abnormal capability to sense pain elicited by thermal stimulation
abnormal vocalization	an inability, decreased ability or altered ability to produce sound from the vocal organs; or a general increase or decrease in the production of vocal sound
abnormal voluntary movement	anomaly in coordinated movements executed with a purpose and can be improved by learning and/or experience
analgesia	inability to sense pain

convulsive seizures	seizures characterized by uncontrolled motor activity
nonconvulsive seizures	seizures without uncontrolled motor activity, but with impairment of consciousness
seizures	sudden and often acute manifestation of epileptic attack, sometimes convulsive
sporadic seizures	occasional seizures occurring at irregular intervals
stereotypic behavior	repetitive, invariant, perseverative motor patterns that do not appear to be purposeful

7.2.2 MGI cellular morphology pathways

Cellular Morphology Pathways	**MGI Description**
Abnormal CNS glial cell morphology	Any structural anomaly of non-neuronal cells of the central nervous system that form the myelin insulation of nervous pathways, guide neuronal migration during development, and exchange metabolites with neurons
Abnormal GABAergic neuron morphology	Any structural anomaly of the neurons that utilize gamma-aminobutyric acid as a neurotransmitter
Abnormal Muller cell morphology	Any structural anomaly of the elongated neuroglial cells that traverse all the layers of the retina and that act as supporting elements
Abnormal astrocyte morphology	Any structural anomaly of the large neuroglial (macroglial) cells in the central nervous system - the largest and most numerous neuroglial cells in the brain and spinal cord; astrocytes are irregularly shaped with mAny long processes, including those with 'end feet' which form the glial (limiting)

	membrane and directly and indirectly contribute to the blood-brain barrier; astrocytes regulate the extracellular ionic and chemical environment, and 'reactive astrocytes' (along with microglia) respond to injury
Abnormal axon morphology	Any structural anomaly of the single process of a nerve cell that normally conducts impulses away from the cell body
Abnormal brain interneuron morphology	Any structural anomaly of neurons that exclusively interact with other neurons in the brain; this includes most brain neuronal cell types
Abnormal cerebral cortex pyramidal cell morphology	Any structural anomaly of the projection neurons in the pyramidal cell layer of the cerebral cortex
Abnormal dendrite morphology	Any structural anomaly of the highly branched tree-like process of a neuron that serves as a receptive field and conducts impulses toward the cell body
Abnormal	Any structural anomaly of the neurons that utilize dopamine as a neurotransmitter

dopaminergic neuron morphology	
Abnormal glial cell morphology	Any structural anomaly of non-neuronal cells of the nervous system that form the myelin insulation of nervous pathways, guide neuronal migration during development, and exchange metabolites with neurons
Abnormal microglial cell morphology	Any structural anomaly of the small, migratory, phagocytic, interstitial cells derived from myeloid progenitor cells and found in the parenchyma of the central nervous system; microglia are scavengers, engulfing dead cells and other debris, and in Alzheimer's disease, microglia are found associated with dying nerve cells and amyloid plaques
Abnormal neurite morphology	Any structural anomaly of a neuronal process, either a dendrite or an axon in vivo, or a filamentous projection from a neuron such as is seen in tissue culture
Abnormal neuroendocrine cell morphology	Any structural anomaly of a neuron that has the specialized function to produce and secrete hormones, contains neurosecretory granules, and that constitutes, in whole or in part, an endocrine organ or system

Abnormal neuron morphology	Any structural anomaly of the cells of the nervous system that receive, conduct, and transmit impulses
Abnormal oligodendrocyte morphology	Any structural anomaly of the neuroglia of the central nervous system that form the insulating myelin sheath of axons in the CNS
Abnormal radial glial cell morphology	Any structural anomaly of the supporting cells of the developing central nervous system that guide neuronal migration during development and exchange metabolites with developing and migrating neurons; these cells differentiate into astrocytes and some neuronal types in the adult
Abnormal synapse morphology	Any structural anomaly of the membrane junction site of a nerve cell to a target cell, such as another nerve cell, an effector cell, or a sensory receptor cell; transmission of nerve impulses may be mediated by chemical or by electrical means

7.2.3 MGI cellular physiology pathways

Cellular Physiology Pathways	**MGI Description**
abnormal CNS synaptic transmission	defect in the communication from a neuron to a target across a synapse in the central nervous system
abnormal PNS synaptic transmission	defect in the communication from a neuron to a target across a synapse in the peripheral nervous system
abnormal excitatory postsynaptic currents	defect in the size or duration of currents detected in postsynaptic cells when an excitatory impulse arrives at the synapse causing depolarization
abnormal excitatory postsynaptic potential	defect in the potential detected in postsynaptic cells when an excitatory impulse arrives at the synapse causing depolarization
abnormal inhibitory postsynaptic currents	defect in the size or duration of currents detected in postsynaptic cells when an inhibitory impulse arrives at the synapse causing hyperpolarization

abnormal long term potentiation	alterations in a persistent robust synaptic response induced by synchronous stimulation of pre- and postsynaptic cells
abnormal miniature excitatory postsynaptic currents	defect in the size or duration of spontaneous currents detected in postsynaptic cells that occur in the absence of an excitatory impulse
abnormal miniature inhibitory postsynaptic currents	defect in the size or duration of spontaneous currents detected in postsynaptic cells that occur in the absence of an inhibitory impulse
abnormal neurotransmitter level	anomaly in the amount of endogenous signaling molecules into a synaptic cleft; neurotransmitters are released on excitation from the axon terminal of a presynaptic neuron of the central or peripheral nervous system and travel across the synaptic cleft to either excite or inhibit the target cell
abnormal prepulse inhibition	anomaly in the ability of a relatively mild stimulus to suppress the response to a strong, startle-eliciting stimulus

abnormal synaptic depression	changes in the duration of the reduction of effectiveness of synaptic connections between neurons and target after repetitive stimulation
abnormal synaptic plasticity	anomaly in the ability of a synapse to change its strength as a result of successive activations
abnormal synaptic transmission	defect in the communication from a neuron to a target across a synapse

7.2.4 MGI Development Pathways

Cellular Physiology Pathways	**MGI Description**
abnormal CNS synapse formation	any anomaly in the process of generating the initial connections between an axon and effector tissue or neuron
abnormal brain development	aberrant or incomplete differentiation of the brain
abnormal cerebellum development	aberrant or incomplete differentiation of the part of the metencephalon that lies dorsal to the pons and medulla behind the brain stem and controls balance for walking and standing, modulates the force and range of movement and is involved in the learning of motor skills
abnormal forebrain development	anomaly in the formation or patterning of the anterior of the three primary divisions of the developing chordate brain or the corresponding part of the adult brain (in vertebrates, includes especially the cerebral hemispheres, the thalamus, and the hypothalamus and especially in higher

	vertebrates is the main control center for sensory and associative information processing, visceral functions, and voluntary motor functions)
abnormal hindbrain development	anomaly in the formation or pattering of the caudal region of the brain
abnormal hippocampus development	improper differentiation of the hippocampus
abnormal midbrain development	anomaly in the formation of or the patterning of the part of the brainstem developing from the middle of the three primary cerebral vesicles of the embryo
abnormal nervous system development	impaired or altered growth of the components of the nervous system
abnormal neuron differentiation	abnormal growth or development of the cells of the nervous system that receive, conduct, and transmit impulses
abnormal telencephalon	anomaly in the progression of the enlarged anteriolateral part of the brain; consists of the paired cerebral hemispheres and olfactory bulbs, the basal ganglia and the connecting structures, and is

development	considered to be the seat of conscious mental processes; it develops from the anterior-most embryological division of the brain that develops from the prosencephalon

7.2.5 Region Tract Morphology Pathways

Region Tract Morphology Pathways	***MGI Description***
Abnormal autonomic nervous system morphology	Any structural anomaly of the function of the sensory and motor neurons that run between the central nervous system (especially the hypothalamus and medulla oblongata) and various internal organs (heart, lungs, endocrine and exocrine glands), responsible for controlling involuntary bodily functions
Abnormal basal ganglion morphology	Any structural anomaly of Any of a group of nuclei associated with the ability to carry out willed movements, including the caudate, putamen, nucleus accumbens, globus pallidus, substantia nigra, and subthalamic nucleus
Abnormal brain morphology	Any structural anomaly of the brain, one of the two components of the central nervous system and the center of thought and emotion; controls coordination, bodily activities and the interpretation of information from the senses (sight, hearing, smell, etc.)

Abnormal brain size	Deviation from the average range of brain size for an organism
Abnormal brain vasculature morphology	Any structural anomaly of the blood vessel network of the brain
Abnormal brain ventricle morphology	Any structural anomaly of the system of four communicating cavities within the brain that are continuous with the central canal of the spinal cord
Abnormal brain white matter morphology	Any structural anomaly of the regions of the brain that are largely or entirely composed of myelinated nerve cell axons and contain few or no neural cell bodies or dendrites
Abnormal brainstem morphology	Any structural anomaly of the stalk-like part of the brain that comprises the midbrain (aka mesencephalon), the pons (aka pons Varolii), and the medulla oblongata, and connects the cerebral hemispheres with the cervical spinal cord
Abnormal cerebral cortex morphology	Any structural anomaly of thin layer of grey matter on the surface of the cerebral hemisphere that develops from the telencephalon and folds into gyri; it is responsible for intellectual faculties and higher mental functions

Abnormal cerebrum morphology	Any structural anomaly of the largest part of the brain, derived from the telencephalon, and is composed of a right and a left hemisphere each which contains an outer cerebral cortex and a subcortical basal ganglia; cerebral functions include sensorimotor, emotional, and intellectual activities
Abnormal corpus callosum morphology	Any structural anomaly of a thick bundle of nerve fibers comprising a commissural plate connecting the two cerebral hemispheres; it consists of contralateral axon projections that provides communications between the right and left cerebral hemispheres
Abnormal diencephalon morphology	Any structural anomaly of the paired caudal parts of the prosencephalon from which the thalamus, hypothalamus, epithalamus and subthalamus are derived; these regions regulate autonomic, visceral and endocrine function, and process information directed to the cerebral cortex
Abnormal ependyma morphology	Any structural anomaly of the cellular membrane that lines the brain ventricles and the central canal of the spinal cord
Abnormal forebrain morphology	Any structural anomaly of the anterior of the three primary divisions of the developing chordate brain or the corresponding part of the adult brain (in vertebrates, includes especially the cerebral

	hemispheres, the thalamus, and the hypothalamus and especially in higher vertebrates is the main control center for sensory and associative information processing, visceral functions, and voluntary motor functions)
Abnormal fourth ventricle morphology	Any structural anomaly of the irregularly shaped cavity in the rhombencephalon, between the medulla oblongata, the pons, and the isthmus in front, and the cerebellum behind; it is continuous with the central canal of the cord below and with the cerebral aqueduct above, and through its lateral and median apertures it communicates with the subarachnoid space
Abnormal hindbrain morphology	Any structural anomaly of the part of the brain developed from the posterior of the three primary brain vesicles of the embryonic neural tube from which the metencephalon and myelencephalon are derived; the metencephalon (anterior part of the embryonic hindbrain), gives rise to the cerebellum and pons while the myelencephalon (posterior portion of the embryonic hindbrain) gives rise to the medulla oblongata
Abnormal hippocampus morphology	Any structural anomaly of the deep lying structure of the cerebrum involved with memory storage and spatial navigation

Abnormal hypothalamus morphology	Any structural anomaly of the ventral part of the diencephalon extending from the region of the optic chiasm to the caudal border of the mammillary bodies and forming the inferior and lateral walls of the third ventricle; this region regulates the autonomic nervous system via hormone production and release
Abnormal innervation	the malformation, misprojection, Abnormal fasciculation or Abnormal refinement of the connection, of nerve fibers to a target
Abnormal lateral ventricle morphology	Any structural anomaly of the cavity in each of the cerebral hemispheres derived from the cavity of the embryonic neural tube; they are separated from each other by the septum pellucidum, and each communicates with the third ventricle by the foramen of Monro, through which also the choroid plexuses of the lateral ventricles become continuous with that of the third ventricle
Abnormal limbic system morphology	Any structural anomaly of Any of a collection of structures in the brain involved in emotion, motivation and emotional aspects of memory; these structures act together to control the endocrine system and the autonomic nervous system
Abnormal midbrain	Any structural anomaly associated with the brain region derived from the middle of the three

morphology	cerebral vesicles of the embryo; this region controls sensory and motor functions in the adult, including eye movement and coordination of auditory and visual reflexes
Abnormal neocortex morphology	Any structural anomaly of the larger part of the mammalian cerebral cortex, distinguished from the allocortex by being composed of a larger number of nerve cells arranged in six layers
Abnormal nervous system morphology	Any structural anomaly of the extensive, intricate network of electochemical structures in the body that is comprised of the brain, spinal cord, nerves, ganglia and parts of the receptor organs that receive and interpret stimuli and transmit impulses to effector organs to control body functions
Abnormal nervous system tract morphology	Any structural anomaly in the structure of Any bundle of myelinated nerve fibers following a defined path through the brain and/or spinal cord
Abnormal parietal lobe morphology	Any structural anomaly of the upper central part of the cerebral hemisphere
Abnormal pituitary gland morphology	Any structural anomaly of the compound gland suspended from the base of the hypothalamus, which secretes somatotropins, prolactin, TSH (thyroid-stimulating hormone), gonadotropins,

	adrenal corticotropin from the anterior lobe; melanocyte stimulating hormone from the intermediate lobe and hormones involved in blood pressure regulation from the posterior lobe
Abnormal postnatal subventricular zone morphology	Any structural anomaly of the mitotically active layer of cells surrounding the brain ventricles in the adult that consists of migrating neuroblasts, astrocytes and transitory amplifying progenitor cells
Abnormal somatic nervous system morphology	Any structural anomaly of the part of the peripheral nervous system that is responsible for conveying voluntary motor and external sensory information, including all nerves controlling the skeletal muscular system and external sensory receptors (including balance, smell, sight, taste, touch and hearing sensory inputs)
Abnormal somatosensory cortex morphology	Any structural anomaly of the area of the parietal lobe that lies posterior to the central sulcus and is concerned with receiving and processing general sensations from the body surface
Abnormal spinal cord morphology	Any structural anomaly of the cylindrical tissue of the vertebral canal that extends from the medulla oblongata to the conus medullaris

Abnormal stratification in cerebral cortex	Abnormal formation or pattern of the layers of the cerebral cortex
Abnormal telencephalon morphology	Any structural anomaly of the enlarged anteriolateral part of the brain; consists of the paired cerebral hemispheres and olfactory bulbs, the basal ganglia and the connecting structures, and is considered to be the seat of conscious mental processes; it develops from the anterior-most embryological division of the brain that develops from the prosencephalon
Abnormal temporal lobe morphology	Any structural anomaly of the lower lateral part of the cerebral hemisphere
Abnormal thalamus morphology	Any structural anomaly of the large ovoid mass of paired bodies containing mostly grey matter and forming part of the lateral wall of the third ventricle of the brain
Abnormal third ventricle morphology	Any structural anomaly of the narrow cleft inferior to the corpus callosum, within the diencephalon, between the paired thalami; its floor is formed by the hypothalamus, its anterior wall by the lamina terminalis, and its roof by ependyma; it communicates with the fourth ventricle by the cerebral aqueduct, and with the lateral ventricles by the interventricular foramina

Decreased brain size	Smaller appearance of the brain
Dilated lateral ventricles	The luminal space of the lateral ventricles is increased in volume or area, usually with an increase in contained fluid, in the cavity in each of the cerebral hemispheres derived from the cavity of the embryonic neural tube; they are separated from each other by the septum pellucidum, and each communicates with the third ventricle by the foramen of Monro, through which also the choroid plexuses of the lateral ventricles become continuous with that of the third ventricle
Dilated third ventricle	The luminal space of the third ventricle is increased in volume or area, usually with an increase in contained fluid, in the narrow cleft inferior to the corpus callosum, within the diencephalon, between the paired thalami; its floor is formed by the hypothalamus, its anterior wall by the lamina terminalis, and its roof by ependyma; it communicates with the fourth ventricle by the cerebral aqueduct, and with the lateral ventricles by the interventricular foramina
Enlarged lateral ventricles	Increased size of the cavity in each of the cerebral hemispheres derived from the cavity of the embryonic neural tube; they are separated from each other by the septum pellucidum, and each communicates with the third ventricle by the foramen of Monro, through which also the choroid

	plexuses of the lateral ventricles become continuous with that of the third ventricle
Enlarged third ventricle	Increased size of the narrow cleft inferior to the corpus callosum, within the diencephalon, between the paired thalami; its floor is formed by the hypothalamus, its anterior wall by the lamina terminalis, and its roof by ependyma; it communicates with the fourth ventricle by the cerebral aqueduct, and with the lateral ventricles by the interventricular foramina
Increased brain size	Larger than the normal physical proportions of the brain
Thin cerebral cortex	Decreased depth of the mantle covering the surface of the cerebral hemispheres

7.2.6 Subcellular neuronal pathways

5HT 2C	(Becamel et al., 2002)
ARC	(Kirov et al., 2012)
CYFIP1 all	(De Rubeis et al., 2013)
Cav2 channels	(Muller et al., 2010)

Chrna7	(Paulo et al., 2009)
FMRP targets	(Darnell et al., 2011)
GABA PSD	(Heller et al., 2012)
NMDAR network	(Pocklington et al., 2006)
PSD human core	(Bayes et al., 2011)
Presynaptic active zone	(Morciano et al., 2009)
Presynapse	Pre-synaptic active zone + Synaptic vesicle
Synaptic vesicle	(Takamori et al., 2006) + Synaptic vesicle GABA enriched + Synaptic vesicle glutamate enriched
Synaptic vesicle GABA enriched	(Gronborg et al., 2010)
Synaptic vesicle glutamate enriched	(Gronborg et al., 2010)
mGluR5	(Farr et al., 2004)

7.3 Appendix C - Full PGC schizophrenia polygenic pathway linear regression results for German cognition sample

Pathway	Pathway Category	Training Threshold	P	P (Corrected)	R2	DOE
abnormal hypothalamus morphology	region tract morphology	0.05	**0.006**	**0.049**	0.0081	-
abnormal neuron morphology	cellular morphology	0.5	0.007	0.063	0.008	+
abnormal basal ganglion morphology	region tract morphology	0.5	0.01	0.088	0.0073	+
abnormal neurite morphology	cellular morphology	0.5	0.011	0.099	0.0066	-
abnormal spinal cord morphology	region tract morphology	0.05	0.013	0.118	0.0066	+
abnormal social conspecific interaction	behaviour	0.05	0.014	0.124	0.0065	+
abnormal social investigation	behaviour	0.5	0.015	0.131	0.0063	+
abnormal neuron morphology	cellular morphology	0.05	0.015	0.131	0.0066	+
abnormal nervous system development	development	0.05	0.015	0.134	0.0065	+
abnormal long term potentiation	cellular physiology	0.05	0.018	0.157	0.0061	-
abnormal telencephalon development	development	0.5	0.02	0.178	0.0058	+
abnormal learning memory	behaviour	0.05	0.021	0.186	0.0057	+
abnormal nervous system morphology	region tract morphology	0.05	0.021	0.19	0.0059	+
abnormal emotion affect behavior	behaviour	0.5	0.022	0.194	0.0056	-
abnormal midbrain morphology	region tract morphology	0.5	0.023	0.207	0.0057	-
abnormal brain morphology	region tract morphology	0.05	0.026	0.226	0.0055	+

abnormal nervous system morphology	region tract morphology	0.5	0.026	0.229	0.0056	+
abnormal associative learning	behaviour	0.5	0.027	0.243	0.0052	+
abnormal glial cell morphology	cellular morphology	0.5	0.03	0.262	0.0053	+
abnormal synaptic depression	cellular physiology	0.05	0.031	0.273	0.0049	+
abnormal forebrain morphology	region tract morphology	0.5	0.031	0.274	0.0053	+
abnormal spinal cord morphology	region tract morphology	0.5	0.036	0.317	0.0049	+
abnormal parental behavior	behaviour	0.05	0.037	0.324	0.0047	+
abnormal eating drinking behavior	behaviour	0.05	0.038	0.34	0.0047	+
abnormal paired pulse facilitation	cellular physiology	0.5	0.04	0.354	0.0046	+
abnormal postnatal subventricular zone morphology	region tract morphology	0.05	0.041	0.364	0.0047	-
abnormal behavioral response to xenobiotic	behaviour	0.5	0.042	0.368	0.0045	+
abnormal inhibitory postsynaptic currents	cellular physiology	0.5	0.042	0.368	0.0043	+
abnormal behavioral response to xenobiotic	behaviour	0.05	0.043	0.376	0.0044	+
abnormal social conspecific interaction	behaviour	0.5	0.045	0.395	0.0044	+
abnormal forebrain morphology	region tract morphology	0.05	0.045	0.397	0.0046	+
abnormal brain morphology	region tract morphology	0.5	0.047	0.418	0.0045	+
abnormal learning memory	behaviour	0.5	0.048	0.429	0.0042	+
abnormal basal ganglion morphology	region tract morphology	0.05	0.051	0.448	0.0042	+
abnormal eating drinking behavior	behaviour	0.5	0.055	0.4708	0.0041	+

abnormal object recognition memory	behaviour	0.5	0.056	0.4794	0.0041	-
abnormal GABAergic neuron morphology	cellular morphology	0.05	0.058	0.4965	0.0039	+
abnormal glial cell morphology	cellular morphology	0.05	0.058	0.4965	0.0041	+
abnormal dendrite morphology	cellular morphology	0.05	0.061	0.5222	0.0039	-
abnormal limbic system morphology	region tract morphology	0.5	0.064	0.5478	0.0037	+
CYFIP1 all	subcellular neuronal	0.5	0.068	0.5821	0.0036	+
dilated lateral ventricles	region tract morphology	0.5	0.07	0.5992	0.0036	+
abnormal avoidance learning behavior	behaviour	0.5	0.071	0.6078	0.0036	+
abnormal paired pulse facilitation	cellular physiology	0.05	0.073	0.6249	0.0034	+
abnormal telencephalon morphology	region tract morphology	0.5	0.073	0.6249	0.0036	+
abnormal hippocampus morphology	region tract morphology	0.5	0.074	0.6334	0.0035	+
abnormal autonomic nervous system morphology	region tract morphology	0.05	0.077	0.6591	0.0034	+
abnormal thalamus morphology	region tract morphology	0.05	0.078	0.6677	0.0032	+
nonconvulsive seizures	behaviour	0.05	0.084	0.7190	0.0033	-
abnormal axon morphology	cellular morphology	0.05	0.084	0.7190	0.0033	-
abnormal microglial cell morphology	cellular morphology	0.05	0.084	0.7190	0.0033	-
analgesia	behaviour	0.5	0.086	0.7362	0.0033	+
abnormal parental behavior	behaviour	0.5	0.087	0.7447	0.0032	+
sporadic seizures	behaviour	0.5	0.088	0.7533	0.0032	-
abnormal nervous system development	development	0.5	0.088	0.7533	0.0033	+

abnormal neurite morphology	cellular morphology	0.05	0.09	0.7704	0.0031	-
abnormal CNS glial cell morphology	cellular morphology	0.5	0.091	0.7790	0.0033	+
NMDAR network	subcellular neuronal	0.05	0.091	0.7790	0.0032	-
abnormal somatic nervous system morphology	region tract morphology	0.5	0.094	0.8046	0.0032	+
abnormal dendrite morphology	cellular morphology	0.5	0.099	0.8474	0.0029	-
abnormal thalamus morphology	region tract morphology	0.5	0.099	0.8474	0.0028	-
abnormal brain size	region tract morphology	0.5	0.101	0.8646	0.003	+
abnormal pain threshold	behaviour	0.5	0.102	0.8731	0.0029	+
abnormal temporal lobe morphology	region tract morphology	0.5	0.103	0.8817	0.0029	+
abnormal diencephalon morphology	region tract morphology	0.05	0.105	0.8988	0.0028	+
increased brain size	region tract morphology	0.5	0.106	0.9074	0.0028	-
Synaptic vesicle	subcellular neuronal	0.5	0.107	0.9159	0.003	+
abnormal cerebrum morphology	region tract morphology	0.5	0.111	0.9502	0.0029	+
abnormal brain ventricle morphology	region tract morphology	0.5	0.115	0.9844	0.0028	-
abnormal brainstem morphology	region tract morphology	0.05	0.125	1	0.0026	-
Presynapse	subcellular neuronal	0.5	0.127	1	0.0027	+
abnormal excitatory postsynaptic currents	cellular physiology	0.5	0.129	1	0.0024	-

Pre synaptic active zone	subcellular neuronal	0.5	0.133	1	0.0021	-
abnormal synaptic depression	cellular physiology	0.5	0.134	1	0.0023	+
abnormal cerebrum morphology	region tract morphology	0.05	0.134	1	0.0025	+
abnormal CNS synaptic transmission	cellular physiology	0.05	0.136	1	0.0023	+
abnormal CNS glial cell morphology	cellular morphology	0.05	0.14	1	0.0025	+
abnormal fourth ventricle morphology	region tract morphology	0.5	0.145	1	0.0023	+
Chrna7	subcellular neuronal	0.05	0.145	1	0.0022	+
abnormal glutamate mediated receptor currents	cellular physiology	0.5	0.148	1	0.0024	+
abnormal radial glial cell morphology	cellular morphology	0.05	0.156	1	0.0023	+
abnormal seizure response to inducing agent	behaviour	0.5	0.16	1	0.0022	+
abnormal long term potentiation	cellular physiology	0.5	0.164	1	0.0022	+
Cav2 channels	subcellular neuronal	0.5	0.165	1	0.0021	+
Pre synaptic active zone	subcellular neuronal	0.05	0.166	1	0.0018	-
abnormal involuntary movement	behaviour	0.05	0.17	1	0.002	-
abnormal synaptic transmission	cellular physiology	0.05	0.171	1	0.002	+
abnormal somatic nervous system morphology	region tract morphology	0.05	0.172	1	0.0022	+
abnormal sleep behavior	behaviour	0.5	0.176	1	0.002	+
abnormal miniature excitatory postsynaptic currents	cellular physiology	0.5	0.179	1	0.002	+
seizures	behaviour	0.5	0.182	1	0.002	+
abnormal diencephalon morphology	region tract morphology	0.5	0.182	1	0.0017	-
abnormal hindbrain morphology	region tract morphology	0.05	0.184	1	0.0019	+

abnormal involuntary movement	behaviour	0.5	0.188	1	0.0018	-
5HT 2C	subcellular neuronal	0.05	0.189	1	0.0017	+
abnormal nervous system tract morphology	region tract morphology	0.05	0.191	1	0.002	+
abnormal GABAergic neuron morphology	cellular morphology	0.5	0.207	1	0.0017	+
thin cerebral cortex	region tract morphology	0.5	0.208	1	0.0017	+
PSD 95 core SN	subcellular neuronal	0.5	0.208	1	0.0017	+
abnormal aggression related behavior	behaviour	0.5	0.215	1	0.0018	+
abnormal contextual conditioning behavior	behaviour	0.05	0.218	1	0.0017	+
GABA PSD	subcellular neuronal	0.05	0.219	1	0.0016	-
abnormal response to tactile stimuli	behaviour	0.05	0.223	1	0.0017	-
abnormal motor capabilities coordination movement	behaviour	0.5	0.23	1	0.0017	+
abnormal forebrain development	development	0.5	0.236	1	0.0015	+
abnormal contextual conditioning behavior	behaviour	0.5	0.24	1	0.0015	+
CYFIP1 all	subcellular neuronal	0.05	0.24	1	0.0016	+
abnormal associative learning	behaviour	0.05	0.241	1	0.0015	+
abnormal behavior	behaviour	0.5	0.241	1	0.0017	+
abnormal synaptic transmission	cellular physiology	0.5	0.242	1	0.0015	+
abnormal telencephalon development	development	0.05	0.249	1	0.0015	+
abnormal PNS synaptic transmission	cellular physiology	0.05	0.25	1	0.0015	-
abnormal brain ventricle choroid plexus morphology	region tract morphology	0.5	0.25	1	0.0014	+
abnormal stratification in cerebral cortex	region tract morphology	0.05	0.251	1	0.0013	-

abnormal third ventricle morphology	region tract morphology	0.5	0.251	1	0.0015	-
abnormal CNS synaptic transmission	cellular physiology	0.5	0.254	1	0.0014	+
enlarged third ventricle	region tract morphology	0.05	0.256	1	0.0014	-
abnormal midbrain hindbrain boundary development	development	0.5	0.263	1	0.0014	+
abnormal prepulse inhibition	cellular physiology	0.05	0.264	1	0.0013	-
abnormal excitatory postsynaptic potential	cellular physiology	0.05	0.265	1	0.0013	+
abnormal synaptic plasticity	cellular physiology	0.5	0.28	1	0.0012	+
abnormal conditioned place preference behavior	behaviour	0.5	0.292	1	0.0012	+
abnormal voluntary movement	behaviour	0.5	0.294	1	0.0013	+
convulsive seizures	behaviour	0.05	0.294	1	0.0013	-
abnormal post tetanic potentiation	cellular physiology	0.05	0.295	1	0.0012	+
abnormal brain ventricle morphology	region tract morphology	0.05	0.295	1	0.0012	-
mGluR5	subcellular neuronal	0.5	0.302	1	0.0012	-
abnormal temporal memory	behaviour	0.5	0.303	1	0.0012	+
abnormal cerebellum development	development	0.05	0.305	1	0.0012	+
abnormal brain size	region tract morphology	0.05	0.308	1	0.0011	+
abnormal stratification in cerebral cortex	region tract morphology	0.5	0.308	1	0.0011	+
dilated lateral ventricles	region tract morphology	0.05	0.309	1	0.0011	+
abnormal midbrain development	development	0.5	0.312	1	0.0011	-

abnormal brain white matter morphology	region tract morphology	0.05	0.312	1	0.0012	+
abnormal axon morphology	cellular morphology	0.5	0.315	1	0.0012	-
abnormal neuron differentiation	development	0.05	0.322	1	0.0011	+
abnormal sensory capabilities reflexes nociception	behaviour	0.05	0.325	1	0.0011	+
abnormal excitatory postsynaptic potential	cellular physiology	0.5	0.337	1	0.0009	+
abnormal miniature inhibitory postsynaptic currents	cellular physiology	0.5	0.347	1	0.0009	-
abnormal postnatal subventricular zone morphology	region tract morphology	0.5	0.355	1	0.0008	+
abnormal brain white matter morphology	region tract morphology	0.5	0.356	1	0.001	+
abnormal lateral ventricle morphology	region tract morphology	0.05	0.356	1	0.001	+
abnormal motor coordination balance	behaviour	0.5	0.357	1	0.001	+
abnormal Muller cell morphology	cellular morphology	0.05	0.357	1	0.0009	-
abnormal hypothalamus morphology	region tract morphology	0.5	0.358	1	0.0008	-
Presynapse	subcellular neuronal	0.05	0.365	1	0.0009	+
abnormal thermal nociception	behaviour	0.5	0.369	1	0.0008	+
abnormal depression related behavior	behaviour	0.5	0.37	1	0.0008	+
abnormal hippocampus morphology	region tract morphology	0.05	0.37	1	0.0009	+
abnormal midbrain morphology	region tract morphology	0.05	0.378	1	0.0008	-
5HT 2C	subcellular neuronal	0.5	0.378	1	0.0008	+
abnormal cerebral cortex pyramidal cell morphology	cellular morphology	0.05	0.379	1	0.0008	-

convulsive seizures	behaviour	0.5	0.384	1	0.0008	+
abnormal synaptic plasticity	cellular physiology	0.05	0.391	1	0.0008	+
abnormal mechanical nociception	behaviour	0.05	0.393	1	0.0008	+
abnormal grooming behavior	behaviour	0.5	0.396	1	0.0008	+
abnormal cued conditioning behavior	behaviour	0.05	0.397	1	0.0008	+
abnormal circadian rhythm	behaviour	0.5	0.398	1	0.0007	+
abnormal conditioned place preference behavior	behaviour	0.05	0.4	1	0.0008	-
abnormal inhibitory postsynaptic currents	cellular physiology	0.05	0.401	1	0.0006	+
abnormal brain interneuron morphology	cellular morphology	0.05	0.415	1	0.0008	+
abnormal synapse morphology	cellular morphology	0.05	0.415	1	0.0007	+
abnormal touch nociception	behaviour	0.5	0.431	1	0.0006	+
abnormal cerebral cortex morphology	region tract morphology	0.5	0.439	1	0.0007	+
abnormal hippocampus development	development	0.5	0.44	1	0.0006	+
abnormal innervation	region tract morphology	0.5	0.44	1	0.0006	-
abnormal ependyma morphology	region tract morphology	0.5	0.444	1	0.0007	-
abnormal hippocampus development	development	0.05	0.456	1	0.0006	+
abnormal forebrain development	development	0.05	0.463	1	0.0006	+
mGluR5	subcellular neuronal	0.05	0.469	1	0.0005	+
abnormal chemical nociception	behaviour	0.5	0.47	1	0.0006	+
abnormal Muller cell morphology	cellular morphology	0.5	0.473	1	0.0005	-
abnormal discrimination learning	behaviour	0.05	0.478	1	0.0005	-
abnormal brain development	development	0.5	0.479	1	0.0006	-

abnormal motor capabilities coordination movement	behaviour	0.05	0.492	1	0.0006	+
abnormal response to novelty	behaviour	0.05	0.5	1	0.0005	+
nonconvulsive seizures	behaviour	0.5	0.502	1	0.0005	-
enlarged lateral ventricles	region tract morphology	0.05	0.509	1	0.0004	+
abnormal nociception after inflammation	behaviour	0.05	0.515	1	0.0004	+
abnormal voluntary movement	behaviour	0.05	0.516	1	0.0005	+
abnormal seizure response to inducing agent	behaviour	0.05	0.519	1	0.0005	-
abnormal behavior	behaviour	0.05	0.52	1	0.0005	+
thin cerebral cortex	region tract morphology	0.05	0.523	1	0.0005	+
stereotypic behavior	behaviour	0.5	0.528	1	0.0004	-
abnormal brain vasculature morphology	region tract morphology	0.05	0.528	1	0.0004	+
analgesia	behaviour	0.05	0.536	1	0.0005	+
abnormal autonomic nervous system morphology	region tract morphology	0.5	0.541	1	0.0005	-
abnormal midbrain development	development	0.05	0.545	1	0.0004	-
abnormal vocalization	behaviour	0.5	0.55	1	0.0005	+
abnormal motor learning	behaviour	0.5	0.551	1	0.0004	-
abnormal sleep behavior	behaviour	0.05	0.554	1	0.0004	-
Synaptic vesicle GABA enriched	subcellular neuronal	0.05	0.56	1	0.0003	+
abnormal brainstem morphology	region tract morphology	0.5	0.567	1	0.0004	-
PSD human core SN	subcellular neuronal	0.05	0.567	1	0.0004	-

abnormal motor learning	behaviour	0.05	0.569	1	0.0003	+
abnormal limbic system morphology	region tract morphology	0.05	0.572	1	0.0004	+
ARC pathway SN	subcellular neuronal	0.5	0.576	1	0.0004	+
abnormal hindbrain morphology	region tract morphology	0.5	0.577	1	0.0004	-
Synaptic vesicle Glu enriched	subcellular neuronal	0.05	0.578	1	0.0003	+
abnormal glutamate mediated receptor currents	cellular physiology	0.05	0.583	1	0.0004	+
abnormal depression related behavior	behaviour	0.05	0.586	1	0.0004	-
abnormal telencephalon morphology	region tract morphology	0.05	0.588	1	0.0003	+
abnormal circadian rhythm	behaviour	0.05	0.59	1	0.0004	-
abnormal temporal memory	behaviour	0.05	0.596	1	0.0003	+
abnormal temporal lobe morphology	region tract morphology	0.05	0.596	1	0.0003	+
abnormal innervation	region tract morphology	0.05	0.609	1	0.0003	-
abnormal fourth ventricle morphology	region tract morphology	0.05	0.614	1	0.0003	-
enlarged lateral ventricles	region tract morphology	0.5	0.619	1	0.0003	+
abnormal aggression related behavior	behaviour	0.05	0.624	1	0.0003	+
abnormal pain threshold	behaviour	0.05	0.628	1	0.0002	+
abnormal spatial learning	behaviour	0.5	0.628	1	0.0002	-
abnormal hindbrain development	development	0.05	0.635	1	0.0003	+
abnormal post tetanic potentiation	cellular physiology	0.5	0.638	1	0.0002	+

abnormal somatosensory cortex morphology	region tract morphology	0.05	0.638	1	0.0003	-
abnormal nervous system tract morphology	region tract morphology	0.5	0.641	1	0.0003	+
abnormal discrimination learning	behaviour	0.5	0.642	1	0.0002	-
Synaptic vesicle Glu enriched	subcellular neuronal	0.5	0.656	1	0.0002	+
Synaptic vesicle	subcellular neuronal	0.05	0.661	1	0.0002	+
dilated third ventricle	region tract morphology	0.5	0.664	1	0.0002	+
stereotypic behavior	behaviour	0.05	0.665	1	0.0002	-
abnormal pituitary gland morphology	region tract morphology	0.05	0.668	1	0.0002	+
abnormal vocalization	behaviour	0.05	0.669	1	0.0002	-
abnormal microglial cell morphology	cellular morphology	0.5	0.67	1	0.0002	+
abnormal miniature excitatory postsynaptic currents	cellular physiology	0.05	0.67	1	0.0002	+
abnormal spatial reference memory	behaviour	0.05	0.671	1	0.0002	-
abnormal neocortex morphology	region tract morphology	0.5	0.675	1	0.0002	+
abnormal brain ventricle choroid plexus morphology	region tract morphology	0.05	0.676	1	0.0002	-
abnormal cerebral cortex morphology	region tract morphology	0.05	0.677	1	0.0002	-
ARC pathway SN	subcellular neuronal	0.05	0.686	1	0.0002	+
abnormal spatial learning	behaviour	0.05	0.692	1	0.0001	-
Chrna7	subcellular neuronal	0.5	0.695	1	0.0002	+
abnormal motor coordination balance	behaviour	0.05	0.698	1	0.0002	+

decreased brain size	region tract morphology	0.5	0.699	1	0.0002	+
abnormal neurotransmitter level	cellular physiology	0.5	0.709	1	0.0001	+
abnormal neuroendocrine cell morphology	cellular morphology	0.5	0.712	1	0.0002	+
abnormal astrocyte morphology	cellular morphology	0.05	0.715	1	0.0001	-
abnormal corpus callosum morphology	region tract morphology	0.05	0.716	1	0.0002	+
abnormal cerebral cortex pyramidal cell morphology	cellular morphology	0.5	0.717	1	0.0001	+
abnormal dopaminergic neuron morphology	cellular morphology	0.05	0.717	1	0.0001	+
Cav2 channels	subcellular neuronal	0.05	0.722	1	0.0001	+
abnormal spatial reference memory	behaviour	0.5	0.729	1	0.0001	+
abnormal hindbrain development	development	0.5	0.736	1	0.0001	-
abnormal sexual interaction	behaviour	0.5	0.754	1	0.0001	+
abnormal somatosensory cortex morphology	region tract morphology	0.5	0.768	1	0.0001	+
dilated third ventricle	region tract morphology	0.05	0.776	1	0.0001	+
abnormal fear anxiety related behavior	behaviour	0.05	0.778	1	0.0001	+
abnormal parietal lobe morphology	region tract morphology	0.5	0.781	1	0.0001	+
abnormal neurotransmitter level	cellular physiology	0.05	0.784	1	0.0001	+
abnormal corpus callosum morphology	region tract morphology	0.5	0.792	1	0.0001	+
abnormal third ventricle morphology	region tract morphology	0.05	0.797	1	0.0001	-
abnormal cued conditioning behavior	behaviour	0.5	0.802	1	0.0001	+

abnormal cerebellum development	development	0.5	0.806	1	0.0001	+
abnormal miniature inhibitory postsynaptic currents	cellular physiology	0.05	0.809	1	0.0001	+
abnormal reflex	behaviour	0.5	0.814	1	0.0001	-
abnormal neocortex morphology	region tract morphology	0.05	0.814	1	0.0001	+
NMDAR network	subcellular neuronal	0.5	0.814	1	0.0001	-
abnormal sensory capabilities reflexes nociception	behaviour	0.5	0.816	1	0.0001	+
abnormal response to new environment	behaviour	0.05	0.822	1	0.0001	+
abnormal touch nociception	behaviour	0.05	0.825	1	<0.0001	+
abnormal CNS synapse formation	development	0.05	0.828	1	0.0001	+
abnormal brain interneuron morphology	cellular morphology	0.5	0.837	1	0.0001	-
abnormal excitatory postsynaptic currents	cellular physiology	0.05	0.838	1	0.0001	+
PSD human core SN	subcellular neuronal	0.5	0.84	1	<0.0001	+
GABA PSD	subcellular neuronal	0.5	0.841	1	<0.0001	-
abnormal lateral ventricle morphology	region tract morphology	0.5	0.847	1	<0.0001	-
abnormal avoidance learning behavior	behaviour	0.05	0.848	1	<0.0001	+
abnormal social investigation	behaviour	0.05	0.852	1	0.0001	+
abnormal spatial working memory	behaviour	0.5	0.864	1	<0.0001	+
abnormal emotion affect behavior	behaviour	0.05	0.865	1	<0.0001	-
abnormal response to new environment	behaviour	0.5	0.865	1	<0.0001	-
abnormal brain vasculature morphology	region tract morphology	0.5	0.866	1	<0.0001	+
abnormal CNS synapse formation	development	0.5	0.87	1	<0.0001	-
decreased brain size	region tract	0.05	0.87	1	<0.0001	+

	morphology					
abnormal PNS synaptic transmission	cellular physiology	0.5	0.873	1	<0.0001	+
Synaptic vesicle GABA enriched	subcellular neuronal	0.5	0.878	1	<0.0001	+
abnormal sexual interaction	behaviour	0.05	0.889	1	<0.0001	+
sporadic seizures	behaviour	0.05	0.891	1	<0.0001	+
abnormal oligodendrocyte morphology	cellular morphology	0.05	0.891	1	<0.0001	-
abnormal ependyma morphology	region tract morphology	0.05	0.897	1	<0.0001	+
FMRP targets	subcellular neuronal	0.5	0.898	1	<0.0001	+
abnormal pituitary gland morphology	region tract morphology	0.5	0.9	1	<0.0001	-
abnormal GABA mediated receptor currents	cellular physiology	0.05	0.912	1	<0.0001	-
abnormal reflex	behaviour	0.05	0.915	1	<0.0001	+
abnormal parietal lobe morphology	region tract morphology	0.05	0.918	1	<0.0001	+
abnormal thermal nociception	behaviour	0.05	0.925	1	<0.0001	+
abnormal radial glial cell morphology	cellular morphology	0.5	0.93	1	<0.0001	+
abnormal mechanical nociception	behaviour	0.5	0.943	1	<0.0001	+
abnormal prepulse inhibition	cellular physiology	0.5	0.945	1	<0.0001	-
enlarged third ventricle	region tract morphology	0.5	0.945	1	<0.0001	+
abnormal grooming behavior	behaviour	0.05	0.948	1	<0.0001	-
FMRP targets	subcellular neuronal	0.05	0.953	1	<0.0001	-
abnormal spatial working memory	behaviour	0.05	0.955	1	<0.0001	-
abnormal midbrain hindbrain boundary development	development	0.05	0.956	1	<0.0001	+

abnormal nociception after inflammation	behaviour	0.5	0.957	1	<0.0001	-
PSD 95 core SN	subcellular neuronal	0.05	0.964	1	<0.0001	-
abnormal response to tactile stimuli	behaviour	0.5	0.966	1	<0.0001	-
seizures	behaviour	0.05	0.967	1	<0.0001	-
abnormal synapse morphology	cellular morphology	0.5	0.967	1	<0.0001	-
abnormal fear anxiety related behavior	behaviour	0.5	0.969	1	<0.0001	-
abnormal chemical nociception	behaviour	0.05	0.971	1	<0.0001	-
abnormal object recognition memory	behaviour	0.05	0.971	1	<0.0001	-
abnormal dopaminergic neuron morphology	cellular morphology	0.5	0.972	1	<0.0001	-
abnormal brain development	development	0.05	0.972	1	<0.0001	-
increased brain size	region tract morphology	0.05	0.972	1	<0.0001	-
abnormal oligodendrocyte morphology	cellular morphology	0.5	0.975	1	<0.0001	-
abnormal astrocyte morphology	cellular morphology	0.5	0.976	1	<0.0001	-
abnormal response to novelty	behaviour	0.5	0.981	1	<0.0001	-
abnormal GABA mediated receptor currents	cellular physiology	0.5	0.981	1	<0.0001	-
abnormal neuron differentiation	development	0.5	0.988	1	<0.0001	-
abnormal neuroendocrine cell morphology	cellular morphology	0.05	0.994	1	<0.0001	+

Direction of effect refers to the sign of the regression coefficient. “+” means increased schizophrenia polygenic risk of that pathway was associated with a better performance IQ score. “–“ means increased polygenic risk of that pathway was associated with a worse performance IQ score.

7.4 Appendix D – Full PGC2 Minus CLOZUK schizophrenia polygenic pathway linear regression results for Cardiff COGS

Pathway	Category	Thres	P	P (Corrected)	%R2	DOE
abnormal cerebellum development	development	0.5	0.0105	0.0854	1.3312	+
abnormal depression related behavior	behaviour	0.05	0.0119	0.0967	1.2864	-
abnormal hippocampus development	development	0.05	0.0194	0.1576	1.1126	+
abnormal oligodendrocyte morphology	cellular morphology	0.05	0.0217	0.1762	1.0731	+
abnormal depression related behavior	behaviour	0.5	0.0246	0.1996	1.0293	-
dilated third ventricle	region tract morphology	0.5	0.0252	0.2044	1.0100	-
abnormal temporal memory	behaviour	0.5	0.0277	0.2254	0.9866	-
abnormal hippocampus development	development	0.5	0.0307	0.2496	0.9511	+
abnormal glial cell morphology	cellular morphology	0.05	0.0366	0.2976	0.8900	+
abnormal fear anxiety related behavior	behaviour	0.5	0.0395	0.3207	0.8643	-
abnormal seizure response to inducing agent	behaviour	0.5	0.0451	0.3660	0.8189	-
abnormal contextual conditioning behavior	behaviour	0.5	0.0456	0.3703	0.8149	-
abnormal temporal memory	behaviour	0.05	0.0501	0.4068	0.7726	-
abnormal fear anxiety related behavior	behaviour	0.05	0.0516	0.4192	0.7668	-
abnormal behavioral response to xenobiotic	behaviour	0.5	0.0525	0.4263	0.7560	-
5HT 2C	subcellular neuronal	0.05	0.0542	0.4401	0.7433	-
abnormal CNS glial cell morphology	cellular morphology	0.05	0.0563	0.4569	0.7300	+
abnormal brain vasculature morphology	region tract morphology	0.5	0.0588	0.4778	0.7235	+

5HT 2C	subcellular neuronal	0.5	0.0596	0.4845	0.7163	-
abnormal grooming behavior	behaviour	0.5	0.0609	0.4950	0.7159	+
abnormal glutamate mediated receptor currents	cellular physiology	0.05	0.0610	0.4956	0.6997	-
abnormal chemical nociception	behaviour	0.05	0.0640	0.5200	0.6943	-
abnormal contextual conditioning behavior	behaviour	0.05	0.0651	0.5285	0.6885	-
abnormal seizure response to inducing agent	behaviour	0.05	0.0662	0.5378	0.6878	-
abnormal glutamate mediated receptor currents	cellular physiology	0.5	0.0663	0.5389	0.6636	-
abnormal conditioned place preference behavior	behaviour	0.5	0.0713	0.5793	0.6500	-
abnormal autonomic nervous system morphology	region tract morphology	0.5	0.0757	0.6151	0.6413	+
abnormal synaptic transmission	cellular physiology	0.05	0.0763	0.6194	0.6176	-
abnormal CNS glial cell morphology	cellular morphology	0.5	0.0819	0.6655	0.5903	+
abnormal hindbrain development	development	0.5	0.0836	0.6793	0.5703	+
abnormal brain development	development	0.5	0.0890	0.7231	0.5582	+
abnormal associative learning	behaviour	0.5	0.0946	0.7687	0.5544	-
abnormal response to novelty	behaviour	0.05	0.0950	0.7717	0.5543	+
seizures	behaviour	0.5	0.0982	0.7978	0.5500	-
abnormal grooming behavior	behaviour	0.05	0.0994	0.8071	0.5422	+
abnormal associative learning	behaviour	0.05	0.0994	0.8073	0.5400	-
abnormal spinal cord morphology	region tract morphology	0.5	0.1002	0.8143	0.5291	+
abnormal miniature excitatory postsynaptic currents	cellular physiology	0.5	0.1032	0.8381	0.5173	-

abnormal nervous system tract morphology	region tract morphology	0.5	0.1054	0.8561	0.5168	+
abnormal chemical nociception	behaviour	0.5	0.1074	0.8727	0.4962	-
Presynapse	subcellular neuronal	0.05	0.1114	0.9052	0.4859	-
abnormal mechanical nociception	behaviour	0.05	0.1116	0.9067	0.4764	+
abnormal axon morphology	cellular morphology	0.5	0.1131	0.9187	0.4742	+
abnormal cerebellum development	development	0.05	0.1190	0.9669	0.4573	+
abnormal eating drinking behavior	behaviour	0.5	0.1229	0.9987	0.4549	+
abnormal CNS synaptic transmission	cellular physiology	0.05	0.1267	1	0.4356	-
nonconvulsive seizures	behaviour	0.05	0.1276	1	0.4050	-
abnormal miniature excitatory postsynaptic currents	cellular physiology	0.05	0.1346	1	0.3946	-
abnormal synaptic transmission	cellular physiology	0.5	0.1356	1	0.3890	-
abnormal nervous system development	development	0.5	0.1442	1	0.3838	+
abnormal glial cell morphology	cellular morphology	0.5	0.1591	1	0.3810	+
abnormal pain threshold	behaviour	0.05	0.1607	1	0.3726	-
abnormal response to novelty	behaviour	0.5	0.1646	1	0.3600	+
Cav2 channels	subcellular neuronal	0.5	0.1676	1	0.3646	-
abnormal social conspecific interaction	behaviour	0.5	0.1705	1	0.3600	+
seizures	behaviour	0.05	0.1721	1	0.3587	-
abnormal oligodendrocyte morphology	cellular morphology	0.5	0.1769	1	0.3600	+
abnormal hindbrain morphology	region tract morphology	0.5	0.1805	1	0.3500	+
Presynapse	subcellular neuronal	0.5	0.1816	1	0.3487	-
abnormal axon morphology	cellular morphology	0.05	0.1844	1	0.3487	+
Cav2 channels	subcellular neuronal	0.05	0.1852	1	0.3422	-
Synaptic vesicle	subcellular neuronal	0.05	0.1865	1	0.3422	-
abnormal hippocampus morphology	region tract morphology	0.05	0.1894	1	0.3422	+

abnormal somatic nervous system morphology	region tract morphology	0.5	0.1900	1	0.3300	+
abnormal corpus callosum morphology	region tract morphology	0.5	0.1904	1	0.3300	+
Pre synaptic active zone	subcellular neuronal	0.5	0.1914	1	0.3239	-
abnormal hippocampus morphology	region tract morphology	0.5	0.1927	1	0.3065	+
abnormal limbic system morphology	region tract morphology	0.05	0.1950	1	0.3065	-
abnormal motor coordination balance	behaviour	0.5	0.1956	1	0.2900	+
abnormal brain size	region tract morphology	0.05	0.2018	1	0.2846	-
abnormal hindbrain morphology	region tract morphology	0.05	0.2042	1	0.2600	-
abnormal autonomic nervous system morphology	region tract morphology	0.05	0.2052	1	0.2604	-
abnormal CNS synaptic transmission	cellular physiology	0.5	0.2081	1	0.2560	-
abnormal discrimination learning	behaviour	0.05	0.2207	1	0.2556	+
abnormal cerebral cortex morphology	region tract morphology	0.5	0.2211	1	0.2500	-
abnormal parietal lobe morphology	region tract morphology	0.5	0.2283	1	0.2493	-
convulsive seizures	behaviour	0.5	0.2381	1	0.2400	-
abnormal midbrain morphology	region tract morphology	0.5	0.2518	1	0.2321	+
convulsive seizures	behaviour	0.05	0.2591	1	0.2321	-
abnormal long term potentiation	cellular physiology	0.5	0.2632	1	0.2321	-
abnormal behavioral response to xenobiotic	behaviour	0.05	0.2635	1	0.2264	-
decreased brain size	region tract morphology	0.5	0.2660	1	0.2243	+
abnormal reflex	behaviour	0.05	0.2672	1	0.2239	-
Pre synaptic active zone	subcellular neuronal	0.05	0.2695	1	0.2234	-
enlarged lateral ventricles	region tract morphology	0.5	0.2772	1	0.2157	+
abnormal neocortex morphology	region tract morphology	0.05	0.2781	1	0.2080	+
abnormal voluntary movement	behaviour	0.05	0.2866	1	0.2075	-

abnormal midbrain morphology	region tract morphology	0.05	0.2872	1	0.2052	+
abnormal brainstem morphology	region tract morphology	0.05	0.2895	1	0.2008	+
abnormal neurite morphology	cellular morphology	0.5	0.2927	1	0.2008	+
abnormal inhibitory postsynaptic currents	cellular physiology	0.5	0.2950	1	0.1957	+
abnormal avoidance learning behavior	behaviour	0.5	0.2953	1	0.2008	-
abnormal inhibitory postsynaptic currents	cellular physiology	0.05	0.2953	1	0.1938	+
abnormal parental behavior	behaviour	0.5	0.2959	1	0.1915	+
sporadic seizures	behaviour	0.05	0.3044	1	0.1884	-
abnormal spinal cord morphology	region tract morphology	0.05	0.3056	1	0.1862	-
abnormal parental behavior	behaviour	0.05	0.3132	1	0.1845	+
abnormal emotion affect behavior	behaviour	0.05	0.3138	1	0.1833	-
abnormal radial glial cell morphology	cellular morphology	0.05	0.3165	1	0.1799	+
sporadic seizures	behaviour	0.5	0.3218	1	0.1780	-
abnormal hypothalamus morphology	region tract morphology	0.05	0.3262	1	0.1780	-
abnormal excitatory postsynaptic currents	cellular physiology	0.05	0.3280	1	0.1776	-
abnormal somatic nervous system morphology	region tract morphology	0.05	0.3293	1	0.1735	-
abnormal cued conditioning behavior	behaviour	0.05	0.3304	1	0.1735	-
abnormal reflex	behaviour	0.5	0.3333	1	0.1654	-
abnormal limbic system morphology	region tract morphology	0.5	0.3362	1	0.1654	-
abnormal post tetanic potentiation	cellular physiology	0.5	0.3371	1	0.1649	+
Synaptic vesicle GABA enriched	subcellular neuronal	0.05	0.3400	1	0.1647	+
abnormal response to new environment	behaviour	0.5	0.3422	1	0.1647	+
stereotypic behavior	behaviour	0.05	0.3439	1	0.1621	-
abnormal brain development	development	0.05	0.3483	1	0.1613	+
abnormal brain vasculature morphology	region tract morphology	0.05	0.3504	1	0.1463	-

Synaptic vesicle	subcellular neuronal	0.5	0.3508	1	0.1505	-
abnormal paired pulse facilitation	cellular physiology	0.05	0.3515	1	0.1463	+
thin cerebral cortex	region tract morphology	0.5	0.3559	1	0.1462	-
abnormal long term potentiation	cellular physiology	0.05	0.3571	1	0.1460	-
abnormal GABAergic neuron morphology	cellular morphology	0.5	0.3686	1	0.1389	-
abnormal spatial reference memory	behaviour	0.05	0.3686	1	0.1372	+
abnormal neuroendocrine cell morphology	cellular morphology	0.05	0.3692	1	0.1353	+
mGluR5	subcellular neuronal	0.05	0.3696	1	0.1341	-
abnormal brain morphology	region tract morphology	0.05	0.3701	1	0.1281	-
abnormal pain threshold	behaviour	0.5	0.3734	1	0.1281	-
abnormal telencephalon development	development	0.5	0.3745	1	0.1281	+
abnormal cerebral cortex morphology	region tract morphology	0.05	0.3805	1	0.1253	-
Chrna7	subcellular neuronal	0.5	0.3911	1	0.1238	+
abnormal fourth ventricle morphology	region tract morphology	0.05	0.3955	1	0.1234	+
abnormal emotion affect behavior	behaviour	0.5	0.3979	1	0.1115	-
abnormal telencephalon development	development	0.05	0.3982	1	0.1079	+
abnormal GABAergic neuron morphology	cellular morphology	0.05	0.4099	1	0.1042	-
abnormal cued conditioning behavior	behaviour	0.5	0.4128	1	0.1007	-
abnormal social conspecific interaction	behaviour	0.05	0.4160	1	0.1005	+
abnormal sensory capabilities reflexes nociception	behaviour	0.05	0.4181	1	0.1005	-
abnormal basal ganglion morphology	region tract morphology	0.05	0.4191	1	0.0969	+
abnormal prepulse inhibition	cellular physiology	0.05	0.4288	1	0.0951	-
abnormal parietal lobe morphology	region tract morphology	0.05	0.4310	1	0.0951	+
abnormal response to new environment	behaviour	0.05	0.4339	1	0.0947	+
abnormal excitatory postsynaptic potential	cellular physiology	0.05	0.4367	1	0.0919	+

abnormal sensory capabilities reflexes nociception	behaviour	0.5	0.4373	1	0.0914	-
abnormal thalamus morphology	region tract morphology	0.5	0.4558	1	0.0907	+
abnormal eating drinking behavior	behaviour	0.05	0.4603	1	0.0947	+
abnormal CNS synapse formation	development	0.05	0.4678	1	0.0834	-
abnormal motor learning	behaviour	0.5	0.4755	1	0.0834	+
abnormal motor learning	behaviour	0.05	0.4829	1	0.0816	+
abnormal brain ventricle choroid plexus morphology	region tract morphology	0.5	0.4840	1	0.0810	+
abnormal telencephalon morphology	region tract morphology	0.5	0.4866	1	0.0808	+
abnormal astrocyte morphology	cellular morphology	0.5	0.4914	1	0.0791	+
abnormal excitatory postsynaptic currents	cellular physiology	0.5	0.4953	1	0.0781	-
abnormal brain interneuron morphology	cellular morphology	0.5	0.4954	1	0.0834	+
abnormal behavior	behaviour	0.05	0.4963	1	0.0756	-
abnormal ependyma morphology	region tract morphology	0.5	0.5028	1	0.0753	-
abnormal mechanical nociception	behaviour	0.5	0.5029	1	0.0750	+
CYFIP1 all	subcellular neuronal	0.5	0.5040	1	0.0741	+
abnormal somatosensory cortex morphology	region tract morphology	0.5	0.5051	1	0.0739	-
abnormal touch nociception	behaviour	0.05	0.5056	1	0.0731	-
abnormal forebrain morphology	region tract morphology	0.5	0.5070	1	0.0728	-
dilated third ventricle	region tract morphology	0.05	0.5090	1	0.0731	-
abnormal pituitary gland morphology	region tract morphology	0.5	0.5165	1	0.0690	-
abnormal brain size	region tract morphology	0.5	0.5179	1	0.0629	-
abnormal innervation	region tract morphology	0.5	0.5207	1	0.0626	-
abnormal midbrain hindbrain boundary	development	0.05	0.5232	1	0.0624	-

development						
abnormal touch nociception	behaviour	0.5	0.5276	1	0.0607	-
abnormal synaptic depression	cellular physiology	0.05	0.5285	1	0.0602	-
abnormal nociception after inflammation	behaviour	0.5	0.5291	1	0.0594	-
abnormal miniature inhibitory postsynaptic currents	cellular physiology	0.5	0.5299	1	0.0578	+
abnormal response to tactile stimuli	behaviour	0.05	0.5342	1	0.0554	+
abnormal microglial cell morphology	cellular morphology	0.5	0.5366	1	0.0531	-
abnormal third ventricle morphology	region tract morphology	0.05	0.5430	1	0.0531	-
abnormal paired pulse facilitation	cellular physiology	0.5	0.5432	1	0.0523	+
abnormal circadian rhythm	behaviour	0.05	0.5442	1	0.0521	+
abnormal involuntary movement	behaviour	0.05	0.5449	1	0.0521	-
abnormal miniature inhibitory postsynaptic currents	cellular physiology	0.05	0.5473	1	0.0518	+
abnormal cerebral cortex pyramidal cell morphology	cellular morphology	0.05	0.5478	1	0.0531	-
abnormal astrocyte morphology	cellular morphology	0.05	0.5500	1	0.0531	+
abnormal dopaminergic neuron morphology	cellular morphology	0.05	0.5509	1	0.0484	+
increased brain size	region tract morphology	0.05	0.5579	1	0.0531	-
abnormal social investigation	behaviour	0.5	0.5615	1	0.0531	+
abnormal forebrain development	development	0.5	0.5792	1	0.0455	+
abnormal learning memory	behaviour	0.05	0.5802	1	0.0440	-
abnormal motor coordination balance	behaviour	0.05	0.5809	1	0.0440	+
abnormal vocalization	behaviour	0.5	0.5862	1	0.0376	+
FMRP targets	subcellular neuronal	0.5	0.5867	1	0.0440	-
abnormal voluntary movement	behaviour	0.5	0.5876	1	0.0359	-

abnormal spatial working memory	behaviour	0.5	0.5900	1	0.0440	-
analgesia	behaviour	0.05	0.5950	1	0.0326	+
abnormal nociception after inflammation	behaviour	0.05	0.6007	1	0.0326	-
abnormal neuron differentiation	development	0.5	0.6030	1	0.0326	+
abnormal dendrite morphology	cellular morphology	0.05	0.6104	1	0.0303	-
enlarged third ventricle	region tract morphology	0.5	0.6120	1	0.0296	-
abnormal neurotransmitter level	cellular physiology	0.5	0.6133	1	0.0282	+
abnormal post tetanic potentiation	cellular physiology	0.05	0.6139	1	0.0271	+
abnormal hindbrain development	development	0.05	0.6140	1	0.0270	+
abnormal sleep behavior	behaviour	0.5	0.6150	1	0.0268	-
abnormal pituitary gland morphology	region tract morphology	0.05	0.6157	1	0.0265	-
increased brain size	region tract morphology	0.5	0.6235	1	0.0236	-
abnormal synapse morphology	cellular morphology	0.5	0.6266	1	0.0234	+
abnormal third ventricle morphology	region tract morphology	0.5	0.6282	1	0.0233	-
abnormal ependyma morphology	region tract morphology	0.05	0.6330	1	0.0228	-
stereotypic behavior	behaviour	0.5	0.6338	1	0.0215	-
abnormal behavior	behaviour	0.5	0.6373	1	0.0201	-
enlarged third ventricle	region tract morphology	0.05	0.6421	1	0.0197	-
CYFIP1 all	subcellular neuronal	0.05	0.6428	1	0.0175	+
abnormal somatosensory cortex morphology	region tract morphology	0.05	0.6604	1	0.0175	-
PSD 95 core SN	subcellular neuronal	0.5	0.6642	1	0.0167	-
FMRP targets	subcellular neuronal	0.05	0.6682	1	0.0175	-
abnormal brainstem morphology	region tract morphology	0.5	0.6713	1	0.0175	-
abnormal forebrain development	development	0.05	0.6752	1	0.0175	+
abnormal brain white matter morphology	region tract morphology	0.05	0.6772	1	0.0175	+

abnormal postnatal subventricular zone morphology	region tract morphology	0.05	0.6813	1	0.0151	+
abnormal midbrain development	development	0.5	0.6899	1	0.0138	-
abnormal motor capabilities coordination movement	behaviour	0.05	0.6928	1	0.0137	-
abnormal cerebrum morphology	region tract morphology	0.5	0.6968	1	0.0117	+
abnormal stratification in cerebral cortex	region tract morphology	0.05	0.6985	1	0.0117	+
nonconvulsive seizures	behaviour	0.5	0.7005	1	0.0117	-
Synaptic vesicle Glu enriched	subcellular neuronal	0.5	0.7037	1	0.0112	+
abnormal spatial reference memory	behaviour	0.5	0.7107	1	0.0107	+
abnormal involuntary movement	behaviour	0.5	0.7161	1	0.0117	-
abnormal object recognition memory	behaviour	0.5	0.7166	1	0.0117	-
abnormal synaptic plasticity	cellular physiology	0.5	0.7172	1	0.0095	+
abnormal excitatory postsynaptic potential	cellular physiology	0.5	0.7190	1	0.0117	+
GABA PSD	subcellular neuronal	0.5	0.7211	1	0.0087	-
abnormal discrimination learning	behaviour	0.5	0.7344	1	0.0117	+
abnormal brain interneuron morphology	cellular morphology	0.05	0.7355	1	0.0083	+
abnormal midbrain development	development	0.05	0.7356	1	0.0079	-
abnormal GABA mediated receptor currents	cellular physiology	0.05	0.7384	1	0.0076	-
abnormal neuron differentiation	development	0.05	0.7460	1	0.0117	+
GABA PSD	subcellular neuronal	0.05	0.7539	1	0.0065	-
abnormal circadian rhythm	behaviour	0.5	0.7562	1	0.0064	+
PSD human core SN	subcellular neuronal	0.5	0.7697	1	0.0060	-
abnormal lateral ventricle morphology	region tract morphology	0.5	0.7722	1	0.0059	+
abnormal neuron morphology	cellular morphology	0.05	0.7752	1	0.0117	-
abnormal basal ganglion morphology	region tract morphology	0.5	0.7788	1	0.0052	+

abnormal temporal lobe morphology	region tract morphology	0.05	0.7793	1	0.0052	+
abnormal CNS synapse formation	development	0.5	0.7807	1	0.0117	-
abnormal lateral ventricle morphology	region tract morphology	0.05	0.7823	1	< 0.01	+
abnormal postnatal subventricular zone morphology	region tract morphology	0.5	0.7826	1	< 0.01	+
abnormal cerebrum morphology	region tract morphology	0.05	0.7857	1	< 0.01	-
abnormal radial glial cell morphology	cellular morphology	0.5	0.7859	1	< 0.01	+
abnormal nervous system development	development	0.05	0.7950	1	< 0.01	+
abnormal prepulse inhibition	cellular physiology	0.5	0.7960	1	< 0.01	+
abnormal thalamus morphology	region tract morphology	0.05	0.8015	1	< 0.01	-
abnormal stratification in cerebral cortex	region tract morphology	0.5	0.8100	1	< 0.01	-
abnormal learning memory	behaviour	0.5	0.8110	1	< 0.01	-
abnormal neuron morphology	cellular morphology	0.5	0.8149	1	< 0.01	+
abnormal Muller cell morphology	cellular morphology	0.5	0.8192	1	< 0.01	+
abnormal microglial cell morphology	cellular morphology	0.05	0.8194	1	< 0.01	-
abnormal diencephalon morphology	region tract morphology	0.5	0.8195	1	< 0.01	-
dilated lateral ventricles	region tract morphology	0.5	0.8266	1	< 0.01	-
abnormal neurite morphology	cellular morphology	0.05	0.8293	1	< 0.01	+
abnormal forebrain morphology	region tract morphology	0.05	0.8306	1	< 0.01	-
NMDAR network	subcellular neuronal	0.5	0.8366	1	< 0.01	-
thin cerebral cortex	region tract morphology	0.05	0.8389	1	< 0.01	-
abnormal avoidance learning behavior	behaviour	0.05	0.8400	1	< 0.01	+
abnormal sexual interaction	behaviour	0.5	0.8447	1	< 0.01	+
abnormal conditioned place preference behavior	behaviour	0.05	0.8473	1	< 0.01	-
abnormal brain white matter morphology	region tract morphology	0.5	0.8537	1	< 0.01	-

ARC pathway SN	subcellular neuronal	0.05	0.8584	1	< 0.01	+
abnormal response to tactile stimuli	behaviour	0.5	0.8593	1	< 0.01	+
abnormal motor capabilities coordination movement	behaviour	0.5	0.8635	1	< 0.01	-
abnormal synaptic depression	cellular physiology	0.5	0.8642	1	< 0.01	-
PSD human core SN	subcellular neuronal	0.05	0.8657	1	< 0.01	-
abnormal neocortex morphology	region tract morphology	0.5	0.8674	1	< 0.01	-
abnormal PNS synaptic transmission	cellular physiology	0.05	0.8734	1	< 0.01	-
abnormal neuroendocrine cell morphology	cellular morphology	0.5	0.8735	1	< 0.01	+
abnormal fourth ventricle morphology	region tract morphology	0.5	0.8770	1	< 0.01	-
abnormal nervous system tract morphology	region tract morphology	0.05	0.8795	1	< 0.01	+
abnormal thermal nociception	behaviour	0.5	0.8801	1	< 0.01	+
abnormal diencephalon morphology	region tract morphology	0.05	0.8835	1	< 0.01	+
abnormal synapse morphology	cellular morphology	0.05	0.8855	1	< 0.01	+
mGluR5	subcellular neuronal	0.5	0.8878	1	< 0.01	-
abnormal midbrain hindbrain boundary development	development	0.5	0.8883	1	< 0.01	-
abnormal hypothalamus morphology	region tract morphology	0.5	0.8884	1	< 0.01	+
abnormal PNS synaptic transmission	cellular physiology	0.5	0.8935	1	< 0.01	-
abnormal aggression related behavior	behaviour	0.05	0.8966	1	< 0.01	-
abnormal Muller cell morphology	cellular morphology	0.05	0.8971	1	< 0.01	-
abnormal nervous system morphology	region tract morphology	0.5	0.8974	1	< 0.01	+
abnormal brain ventricle choroid plexus morphology	region tract morphology	0.05	0.9002	1	< 0.01	+
ARC pathway SN	subcellular neuronal	0.5	0.9037	1	< 0.01	-
NMDAR network	subcellular neuronal	0.05	0.9109	1	< 0.01	-

abnormal dopaminergic neuron morphology	cellular morphology	0.5	0.9152	1	< 0.01	+
abnormal social investigation	behaviour	0.05	0.9189	1	< 0.01	+
abnormal dendrite morphology	cellular morphology	0.5	0.9215	1	< 0.01	+
PSD 95 core SN	subcellular neuronal	0.05	0.9219	1	< 0.01	+
enlarged lateral ventricles	region tract morphology	0.05	0.9225	1	< 0.01	+
decreased brain size	region tract morphology	0.05	0.9226	1	< 0.01	+
abnormal vocalization	behaviour	0.05	0.9228	1	< 0.01	+
abnormal corpus callosum morphology	region tract morphology	0.05	0.9232	1	< 0.01	+
Synaptic vesicle Glu enriched	subcellular neuronal	0.05	0.9291	1	< 0.01	+
abnormal neurotransmitter level	cellular physiology	0.05	0.9295	1	< 0.01	-
abnormal cerebral cortex pyramidal cell morphology	cellular morphology	0.5	0.9324	1	< 0.01	-
abnormal object recognition memory	behaviour	0.05	0.9324	1	< 0.01	+
abnormal temporal lobe morphology	region tract morphology	0.5	0.9376	1	< 0.01	+
abnormal spatial learning	behaviour	0.05	0.9377	1	< 0.01	-
analgesia	behaviour	0.5	0.9393	1	< 0.01	+
abnormal sexual interaction	behaviour	0.05	0.9407	1	< 0.01	-
abnormal GABA mediated receptor currents	cellular physiology	0.5	0.9407	1	< 0.01	+
abnormal brain morphology	region tract morphology	0.5	0.9418	1	< 0.01	-
abnormal brain ventricle morphology	region tract morphology	0.05	0.9433	1	< 0.01	-
abnormal sleep behavior	behaviour	0.05	0.9447	1	< 0.01	+
abnormal spatial working memory	behaviour	0.05	0.9514	1	< 0.01	+
abnormal aggression related behavior	behaviour	0.5	0.9518	1	< 0.01	+
abnormal spatial learning	behaviour	0.5	0.9597	1	< 0.01	-
abnormal nervous system morphology	region tract morphology	0.05	0.9608	1	< 0.01	-
abnormal telencephalon morphology	region tract morphology	0.05	0.9639	1	< 0.01	-

abnormal brain ventricle morphology	region tract morphology	0.5	0.9667	1	< 0.01	-
abnormal synaptic plasticity	cellular physiology	0.05	0.9678	1	< 0.01	+
Synaptic vesicle GABA enriched	subcellular neuronal	0.5	0.9718	1	< 0.01	+
abnormal thermal nociception	behaviour	0.05	0.9803	1	< 0.01	-
abnormal innervation	region tract morphology	0.05	0.9883	1	< 0.01	-
dilated lateral ventricles	region tract morphology	0.05	0.9935	1	< 0.01	-
Chrna7	subcellular neuronal	0.05	0.9938	1	< 0.01	-

Direction of effect refers to the sign of the regression coefficient. “+” means increased schizophrenia polygenic risk of that pathway was associated with a better MATRICS composite score. “–“ means increased polygenic risk of that pathway was associated with a worse MATRICS composite score.

7.5 Appendix E – Brown's p-values for associations between SNPs in the 155 candidate pathways and general cognitive ability in schizophrenia cases

Pathway Category	Pathway	NSNP	P1	P (Corrected)
Behaviour	abnormal motor learning	12793	0.056	0.095
Behaviour	abnormal discrimination learning	5607	0.073	0.124
Development	abnormal CNS synapse formation	5657	0.075	0.127
Region Tract Morphology	abnormal hindbrain morphology	78393	0.083	0.141
Cellular Physiology	abnormal GABA mediated receptor currents	5186	0.094	0.160
Cellular Physiology	abnormal miniature inhibitory postsynaptic currents	5658	0.119	0.202
Region Tract Morphology	abnormal cerebrum morphology	87344	0.121	0.206
Region Tract Morphology	abnormal lateral ventricle morphology	16676	0.121	0.206
Region Tract Morphology	abnormal telencephalon morphology	109098	0.134	0.228
Region Tract Morphology	abnormal forebrain morphology	132002	0.136	0.231
Region Tract Morphology	abnormal midbrain morphology	20284	0.147	0.277
Region Tract Morphology	abnormal autonomic nervous system morphology	9261	0.162	0.305
Region Tract	abnormal cerebral cortex morphology	37594	0.178	0.335

Morphology				
Region Tract Morphology	abnormal brainstem morphology	37040	0.184	0.345
Region Tract Morphology	enlarged lateral ventricles	10680	0.189	0.355
Region Tract Morphology	abnormal brain morphology	215431	0.196	0.368
Behaviour	stereotypic behavior	37697	0.204	0.383
Development	abnormal cerebellum development	32712	0.204	0.384
Behaviour	abnormal response to new environment	40290	0.211	0.397
Cellular Morphology	abnormal neuroendocrine cell morphology	1150	0.222	0.416
Development	abnormal hindbrain development	37561	0.228	0.429
Cellular Morphology	abnormal cerebral cortex pyramidal cell morphology	6764	0.232	0.436
Development	abnormal hippocampus development	2849	0.233	0.437
Cellular Physiology	abnormal prepulse inhibition	25231	0.235	0.442
Development	abnormal brain development	85442	0.238	0.446
Behaviour	abnormal motor coordination balance	142486	0.241	0.452
Behaviour	abnormal involuntary movement	115326	0.249	0.467
Behaviour	abnormal response to novelty	43621	0.249	0.468
Behaviour	abnormal reflex	95790	0.251	0.471
Cellular Morphology	abnormal radial glial cell morphology	7569	0.251	0.472
Cellular Physiology	abnormal post tetanic potentiation	3393	0.265	0.498
Region Tract Morphology	abnormal nervous system morphology	350376	0.270	0.506
Region Tract Morphology	abnormal ependyma morphology	671	0.282	0.529
Region Tract	abnormal pituitary gland morphology	16432	0.285	0.535

Morphology				
Subcellular Neuronal	CYFIP1 all	7275	0.290	0.544
Region Tract Morphology	abnormal innervation	38718	0.290	0.544
Cellular Morphology	abnormal Muller cell morphology	934	0.291	0.547
Region Tract Morphology	abnormal limbic system morphology	53129	0.301	0.566
Region Tract Morphology	abnormal hippocampus morphology	47947	0.301	0.566
Cellular Physiology	abnormal neurotransmitter level	14912	0.312	0.586
Cellular Morphology	abnormal synapse morphology	29419	0.313	0.587
Region Tract Morphology	abnormal temporal lobe morphology	49656	0.316	0.594
Region Tract Morphology	abnormal brain ventricle morphology	24966	0.321	0.602
Region Tract Morphology	increased brain size	9253	0.321	0.603
Region Tract Morphology	abnormal brain ventricle choroid plexus morphology	30872	0.328	0.616
Behaviour	abnormal social investigation	21872	0.332	0.622
Development	abnormal nervous system development	150280	0.337	0.632
Cellular Physiology	abnormal paired pulse facilitation	20400	0.341	0.641
Region Tract Morphology	dilated lateral ventricles	5598	0.342	0.642
Region Tract Morphology	abnormal spinal cord morphology	47602	0.343	0.643
Cellular Physiology	abnormal synaptic plasticity	9020	0.343	0.643

Behaviour	abnormal learning memory	122683	0.345	0.648
Behaviour	abnormal behavioral response to xenobiotic	52232	0.361	0.677
Region Tract Morphology	abnormal neocortex morphology	1715	0.373	0.700
Subcellular Neuronal	Chrna7	13547	0.375	0.703
Behaviour	abnormal grooming behavior	18604	0.383	0.720
Cellular Morphology	abnormal dopaminergic neuron morphology	8503	0.387	0.727
Region Tract Morphology	abnormal diencephalon morphology	35091	0.389	0.731
Development	abnormal midbrain development	3652	0.390	0.732
Cellular Morphology	abnormal oligodendrocyte morphology	5615	0.396	0.744
Behaviour	nonconvulsive seizures	5223	0.399	0.749
Cellular Physiology	abnormal PNS synaptic transmission	9158	0.403	0.757
Behaviour	abnormal associative learning	54548	0.411	0.771
Development	abnormal telencephalon development	15894	0.411	0.772
Development	abnormal forebrain development	24338	0.417	0.782
Behaviour	abnormal sensory capabilities reflexes nociception	142991	0.430	0.807
Subcellular Neuronal	5HT 2C	2371	0.430	0.808
Cellular Morphology	abnormal neuron morphology	205349	0.431	0.809
Cellular Morphology	abnormal CNS glial cell morphology	34584	0.435	0.816
Region Tract Morphology	abnormal brain white matter morphology	36545	0.441	0.828
Cellular Morphology	abnormal dendrite morphology	23020	0.442	0.830
Region Tract Morphology	abnormal somatic nervous system morphology	127652	0.443	0.832
Region Tract Morphology	dilated third ventricle	2547	0.443	0.832

Region Tract Morphology	abnormal nervous system tract morphology	35792	0.443	0.832
Region Tract Morphology	abnormal brain size	32085	0.449	0.843
Region Tract Morphology	abnormal corpus callosum morphology	27791	0.452	0.849
Region Tract Morphology	abnormal thalamus morphology	6191	0.465	0.873
Behaviour	abnormal mechanical nociception	1420	0.470	0.881
Behaviour	abnormal chemical nociception	9115	0.487	0.915
Cellular Morphology	abnormal neurite morphology	50230	0.491	0.923
Cellular Physiology	abnormal CNS synaptic transmission	129079	0.492	0.923
Region Tract Morphology	abnormal hypothalamus morphology	3532	0.493	0.926
Region Tract Morphology	abnormal stratification in cerebral cortex	7829	0.496	0.931
Behaviour	abnormal sexual interaction	15301	0.500	0.938
Region Tract Morphology	abnormal somatosensory cortex morphology	6905	0.504	0.946
Development	abnormal midbrain hindbrain boundary development	1061	0.504	0.947
Region Tract Morphology	abnormal third ventricle morphology	8614	0.507	0.952
Behaviour	abnormal conditioned place preference behavior	3789	0.514	0.966
Cellular Physiology	abnormal synaptic transmission	147864	0.515	0.968
Behaviour	abnormal spatial working memory	11029	0.518	0.973
Behaviour	abnormal object recognition memory	3660	0.519	0.975
Cellular Physiology	abnormal excitatory postsynaptic currents	36747	0.523	0.981

Behaviour	abnormal spatial learning	55636	0.524	0.984
Region Tract Morphology	abnormal parietal lobe morphology	7344	0.524	0.985
Cellular Physiology	abnormal synaptic depression	28224	0.526	0.987
Cellular Physiology	abnormal excitatory postsynaptic potential	19984	0.531	0.997
Cellular Morphology	abnormal glial cell morphology	46229	0.542	1
Behaviour	abnormal cued conditioning behavior	23001	0.546	1
Behaviour	abnormal seizure response to inducing agent	46789	0.552	1
Behaviour	analgesia	13655	0.554	1
Behaviour	abnormal voluntary movement	202622	0.555	1
Behaviour	abnormal avoidance learning behavior	14004	0.557	1
Region Tract Morphology	abnormal fourth ventricle morphology	5094	0.566	1
Region Tract Morphology	abnormal brain vasculature morphology	3490	0.566	1
Region Tract Morphology	thin cerebral cortex	7160	0.571	1
Region Tract Morphology	abnormal basal ganglion morphology	10079	0.579	1
Cellular Physiology	abnormal inhibitory postsynaptic currents	16184	0.581	1
Cellular Morphology	abnormal brain interneuron morphology	14836	0.584	1
Behaviour	abnormal pain threshold	40551	0.592	1
Behaviour	abnormal response to tactile stimuli	23471	0.593	1
Behaviour	abnormal social conspecific interaction	65380	0.594	1
Behaviour	abnormal touch nociception	55421	0.595	1
Region Tract Morphology	decreased brain size	20799	0.600	1

Subcellular Neuronal	Cav2 channels	74186	0.601	1
Behaviour	abnormal nociception after inflammation	8630	0.602	1
Behaviour	abnormal contextual conditioning behavior	28093	0.618	1
Development	abnormal neuron differentiation	49642	0.619	1
Subcellular Neuronal	NMDAR network	21183	0.622	1
Region Tract Morphology	abnormal postnatal subventricular zone morphology	7084	0.623	1
Behaviour	abnormal thermal nociception	26987	0.628	1
Behaviour	abnormal temporal memory	29640	0.632	1
Behaviour	seizures	68289	0.638	1
Region Tract Morphology	enlarged third ventricle	3802	0.658	1
Subcellular Neuronal	PSD 95 core SN	21212	0.664	1
Subcellular Neuronal	mGluR5	18032	0.666	1
Cellular Morphology	abnormal axon morphology	19012	0.674	1
Behaviour	abnormal vocalization	12894	0.683	1
Behaviour	convulsive seizures	31270	0.691	1
Behaviour	abnormal emotion affect behavior	79890	0.704	1
Cellular Morphology	abnormal astrocyte morphology	16670	0.704	1
Behaviour	abnormal parental behavior	29994	0.705	1
Behaviour	abnormal eating drinking behavior	79612	0.709	1
Behaviour	abnormal spatial reference memory	9398	0.724	1
Subcellular Neuronal	Synaptic vesicle GABA enriched	4661	0.730	1
Cellular Physiology	abnormal long term potentiation	49049	0.734	1
Behaviour	abnormal sleep behavior	7749	0.740	1
Behaviour	abnormal depression related behavior	10651	0.744	1

Subcellular Neuronal	Synaptic vesicle Glu enriched	4110	0.745	1
Subcellular Neuronal	GABA PSD	9818	0.752	1
Behaviour	abnormal motor capabilities coordination movement	264692	0.756	1
Subcellular Neuronal	FMRP targets	267391	0.766	1
Cellular Physiology	abnormal glutamate mediated receptor currents	12340	0.766	1
Behaviour	abnormal circadian rhythm	8549	0.771	1
Behaviour	abnormal fear anxiety related behavior	48120	0.775	1
Subcellular Neuronal	ARC pathway SN	11767	0.784	1
Cellular Morphology	abnormal GABAergic neuron morphology	1763	0.792	1
Behaviour	abnormal behavior	372178	0.797	1
Cellular Morphology	abnormal microglial cell morphology	2574	0.806	1
Cellular Physiology	abnormal miniature excitatory postsynaptic currents	14889	0.825	1
Subcellular Neuronal	Pre synaptic active zone	31001	0.829	1
Behaviour	sporadic seizures	2368	0.834	1
Subcellular Neuronal	PSD human core SN	158492	0.856	1
Behaviour	abnormal aggression related behavior	13856	0.863	1
Subcellular Neuronal	Synaptic vesicle	71108	0.863	1
Subcellular Neuronal	Presynapse	88965	0.911	1

7.6 Appendix F – List of neuropsychiatric CNVs from (Kaminsky et al., 2011; Girirajan et al., 2012)

10q23
15q11.2
15q13.3
16p11.2
16p12.1
16p13.11
17p11.2
17q12
1p36
1q21.1
22q11.2
22q13
3q29
5q35
7q11.23
8p23.1
9q34
Potocki–Lupski
Prader Willi Syndrome/Angelman Syndrome
Williams-Beuren

7.7 Appendix G – Linear regression results for 85 pathways hit by genes in large (>100kb) CNVs and their association with the MATRICS composite score

Category	Pathway	P	Corrected P	%R2 Increase	Direction of Effect
behaviour	abnormal reflex	0.081	0.0996	0.329	Negative
behaviour	abnormal circadian rhythm	0.107	0.1316	0.318	Positive
behaviour	abnormal involuntary movement	0.125	0.1538	0.312	Negative
subcellular neuronal	Presynapse	0.142	0.1747	0.307	Negative
behaviour	stereotypic behavior	0.178	0.2189	0.299	Negative
development	abnormal hindbrain development	0.213	0.2620	0.292	Positive
development	abnormal midbrain development	0.213	0.2620	0.292	Positive
cellular physiology	abnormal excitatory postsynaptic potential	0.247	0.3038	0.287	Negative
cellular physiology	abnormal synaptic plasticity	0.247	0.3038	0.287	Negative
region tract morphology	abnormal hippocampus morphology	0.247	0.3038	0.287	Negative

region tract morphology	abnormal temporal lobe morphology	0.247	0.3038	0.287	Negative
behaviour	abnormal social investigation	0.250	0.3075	0.286	Negative
behaviour	abnormal avoidance learning behavior	0.275	0.3383	0.283	Positive
cellular physiology	abnormal miniature excitatory postsynaptic currents	0.275	0.3383	0.283	Positive
region tract morphology	abnormal basal ganglion morphology	0.275	0.3383	0.283	Positive
region tract morphology	abnormal brain size	0.275	0.3383	0.283	Positive
region tract morphology	abnormal lateral ventricle morphology	0.275	0.3383	0.283	Positive
region tract morphology	abnormal third ventricle morphology	0.275	0.3383	0.283	Positive
region tract morphology	dilated lateral ventricles	0.275	0.3383	0.283	Positive
region tract morphology	dilated third ventricle	0.275	0.3383	0.283	Positive
region tract morphology	increased brain size	0.275	0.3383	0.283	Positive
region tract morphology	abnormal diencephalon morphology	0.281	0.3456	0.282	Positive

behaviour	abnormal parental behavior	0.323	0.3973	0.277	Negative
subcellular neuronal	Chrna7	0.339	0.4170	0.276	Negative
development	abnormal forebrain development	0.345	0.4244	0.275	Positive
behaviour	abnormal behavioral response to xenobiotic	0.402	0.4945	0.271	Negative
development	abnormal brain development	0.403	0.4957	0.270	Positive
region tract morphology	abnormal limbic system morphology	0.408	0.5018	0.270	Negative
behaviour	abnormal associative learning	0.451	0.5547	0.267	Negative
cellular physiology	abnormal long term potentiation	0.451	0.5547	0.267	Negative
development	abnormal nervous system development	0.478	0.5879	0.266	Positive
cellular morphology	abnormal neurite morphology	0.524	0.6445	0.263	Positive
development	abnormal neuron differentiation	0.536	0.6593	0.263	Positive
behaviour	abnormal cued conditioning behavior	0.544	0.6691	0.262	Negative
behaviour	abnormal spatial learning	0.544	0.6691	0.262	Positive

region tract morphology	abnormal forebrain morphology	0.576	0.7085	0.261	Positive
behaviour	abnormal spatial working memory	0.578	0.7109	0.261	Positive
cellular morphology	abnormal dendrite morphology	0.578	0.7109	0.261	Positive
subcellular neuronal	mGluR5	0.578	0.7109	0.261	Positive
cellular morphology	abnormal glial cell morphology	0.588	0.7232	0.260	Negative
behaviour	abnormal object recognition memory	0.609	0.7491	0.260	Positive
region tract morphology	abnormal brain morphology	0.613	0.7540	0.259	Positive
region tract morphology	abnormal pituitary gland morphology	0.618	0.7601	0.259	Positive
region tract morphology	abnormal cerebral cortex morphology	0.634	0.7798	0.259	Positive
region tract morphology	abnormal hypothalamus morphology	0.640	0.7872	0.258	Positive
region tract morphology	abnormal cerebrum morphology	0.645	0.7934	0.258	Negative
region tract morphology	abnormal parietal lobe morphology	0.668	0.8216	0.258	Positive
region tract morphology	abnormal somatosensory cortex morphology	0.668	0.8216	0.258	Positive

behaviour	abnormal behavior	0.670	0.8241	0.258	Negative
behaviour	abnormal chemical nociception	0.680	0.8364	0.257	Negative
behaviour	abnormal response to tactile stimuli	0.680	0.8364	0.257	Negative
cellular morphology	abnormal synapse morphology	0.685	0.8426	0.257	Negative
cellular physiology	abnormal PNS synaptic transmission	0.685	0.8426	0.257	Negative
cellular physiology	abnormal neurotransmitter level	0.685	0.8426	0.257	Negative
behaviour	abnormal sexual interaction	0.699	0.8598	0.257	Positive
region tract morphology	abnormal brain white matter morphology	0.704	0.8659	0.257	Negative
region tract morphology	abnormal corpus callosum morphology	0.704	0.8659	0.257	Negative
region tract morphology	abnormal nervous system tract morphology	0.704	0.8659	0.257	Negative
development	abnormal telencephalon development	0.708	0.8708	0.256	Positive
cellular physiology	abnormal prepulse inhibition	0.721	0.8868	0.256	Positive
region tract morphology	abnormal nervous system morphology	0.727	0.8942	0.256	Positive

behaviour	abnormal contextual conditioning behavior	0.739	0.9090	0.256	Negative
behaviour	abnormal temporal memory	0.739	0.9090	0.256	Negative
behaviour	abnormal voluntary movement	0.743	0.9139	0.256	Positive
cellular morphology	abnormal neuron morphology	0.758	0.9323	0.255	Positive
region tract morphology	abnormal somatic nervous system morphology	0.765	0.9410	0.255	Negative
behaviour	abnormal pain threshold	0.774	1	0.255	Negative
cellular morphology	abnormal CNS glial cell morphology	0.783	1	0.255	Negative
cellular morphology	abnormal astrocyte morphology	0.783	1	0.255	Negative
cellular morphology	abnormal axon morphology	0.796	1	0.255	Positive
region tract morphology	abnormal brain ventricle morphology	0.816	1	0.254	Positive
behaviour	abnormal mechanical nociception	0.819	1	0.254	Negative
cellular physiology	abnormal synaptic transmission	0.826	1	0.254	Positive
region tract morphology	abnormal innervation	0.828	1	0.254	Negative

behaviour	abnormal grooming behavior	0.829	1	0.254	Negative
behaviour	abnormal seizure response to inducing agent	0.829	1	0.254	Negative
behaviour	seizures	0.829	1	0.254	Negative
behaviour	abnormal response to new environment	0.831	1	0.254	Positive
region tract morphology	abnormal telencephalon morphology	0.831	1	0.254	Negative
region tract morphology	abnormal neocortex morphology	0.833	1	0.254	Positive
cellular physiology	abnormal CNS synaptic transmission	0.837	1	0.254	Positive
behaviour	abnormal thermal nociception	0.911	1	0.253	Negative
region tract morphology	abnormal spinal cord morphology	0.929	1	0.253	Positive
region tract morphology	abnormal hindbrain morphology	0.937	1	0.253	Negative
behaviour	abnormal response to novelty	0.987	1	0.253	Positive

Category represents the broader biological function of the individual pathway. P represents whether the addition of the 'number of pathway genes hit' term in model 2 significantly improved fit to the data compared to model 1 using an ANOVA comparison. %R2

represents additional variance explained of the MATRICS composite score for addition of total number of pathway genes. Direction of effect refers to higher (positive) or lower (negative) MATRICS scores for the number of pathway genes hit.