Comprehensive pathway analyses of schizophrenia risk loci point to dysfunctional postsynaptic signaling

Running title: Comprehensive pathway analyses in schizophrenia

Dick Schijven MSc [1,2,3,7], Daniel Kofink MSc [4], Vinicius Tragante PhD [4], Marloes Verkerke BSc [3], Sara L. Pulit PhD [2], René S. Kahn MD PhD [1,5], Jan H. Veldink MD PhD [2], Christiaan H. Vinkers MD PhD [1], Marco P. Boks MD PhD [1], Jurjen J. Luykx MD PhD [1,3,6]

1. Department of Psychiatry, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht University, the Netherlands
2. Department of Neurology and Neurosurgery, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht University, the Netherlands
3. Department of Translational Neuroscience, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht University, the Netherlands
4. Department of Cardiology, Division Heart and Lungs, University Medical Center Utrecht, Utrecht University the Netherlands
5. Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA
6. Department of Psychiatry, ZNA hospitals, Antwerp, Belgium

[7] Corresponding author:

Dick Schijven

Brain Center Rudolf Magnus, University Medical Center Utrecht, Universiteitsweg 100, 3584CG Utrecht, the Netherlands

Email: D.Schijven@umcutrecht.nl

Telephone: +31 88 75 676 81

Fax: +31 88 75 518 14

Abstract

Large-scale genome-wide association studies (GWAS) have implicated many low-penetrance loci in schizophrenia. However, its pathological mechanisms are poorly understood, which in turn hampers the development of novel pharmacological treatments. Pathway and gene set analyses carry the potential to generate hypotheses about disease mechanisms and have provided biological context to genome-wide data of schizophrenia. We aimed to examine which biological processes are likely candidates to underlie schizophrenia by integrating novel and powerful pathway analysis tools using data from the largest published schizophrenia GWAS (N = 79,845). By applying a primary unbiased analysis (Multi-marker Analysis of GenoMic Annotation; MAGMA) to weigh the role of biological processes from the Gene Ontology database, we identified enrichment of common variants in synaptic plasticity and neuron differentiation gene sets. We supported these findings using MAGMA, Meta-Analysis Gene-set Enrichment of variaNT Associations (MAGENTA) and Interval Enrichment Analysis (INRICH) on detailed synaptic signaling pathways from the Kyoto Encyclopedia of Genes and Genomes (KEGG) and found enrichment in mainly the dopaminergic synapse and long-term potentiation. Moreover, shared genes involved in these neurotransmitter systems had a large contribution to the observed enrichment. Finally, protein products of significant schizophrenia-associated genes showed more interactions than expected by chance and more indirect connections to other proteins. In conclusion, we provide strong and consistent genetics-informed evidence for the role of postsynaptic signaling processes in schizophrenia, opening avenues for future psychopharmacological studies.

Introduction

Antipsychotics target synaptic signaling by changing neurotransmission of the dopamine D2 receptor (DRD2) and the serotonin 5-HT_2 receptor (Celada *et al*, 2013; Kahn *et al*, 2015). These drugs are the mainstay treatment modality in schizophrenia and are fairly effective at reducing positive symptoms (Kane and Correll, 2010). However, improving the cognitive and negative symptoms, which substantially affect quality of life, has proven challenging (Kahn *et al*, 2015; Karow *et al*, 2014; Köster *et al*, 2014). Although post-mortem studies, imaging and human genetic studies have contributed to theories about pathophysiological mechanisms in schizophrenia, the underlying molecular processes have not been fully elucidated.

Genetic studies provide a valuable resource to investigate the mechanisms that are likely at play in schizophrenia. Schizophrenia is highly heritable (h^2 estimates ranging from 45-80%) and polygenic (Lichtenstein *et al*, 2006; Wang *et al*, 2017). Genome-wide association studies (GWAS) have identified 108 independent associated risk loci (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). More recently, 142 associated loci were reported in a meta-analysis (not yet peer-reviewed) combining this 2014 GWAS with a British study population (doi.org/10.1101/068593). Pathway and gene set enrichment analysis methods are widely used to provide biological context to the results of genetic association studies by testing whether biologically relevant pathways or sets of genes are enriched for genetic variants associated with a phenotype (de Leeuw *et al*, 2016). These analyses have been widely applied to schizophrenia, providing evidence for the involvement of synaptic and immune-related processes (Duncan *et al*, 2014; Lips *et al*, 2012). Such findings are supported by pathway analyses in combined psychiatric disorders (schizophrenia, depression and bipolar disorder) (Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium, 2015), revealing enrichment of genetic variants in neuronal, immune and histone pathways (Network and Pathway Analysis

Subgroup of Psychiatric Genomics Consortium, 2015). Involvement of calcium signaling and ion channels in schizophrenia has been reported in a gene set analysis paper combining GWAS data with post-mortem brain gene expression data (Hertzberg *et al*, 2015). Importantly, none of the abovementioned pathway analysis studies has used the full dataset reported in the latest (and largest) published schizophrenia GWAS (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). Moreover, several novel or widely used pathway analysis tools have not yet been applied to this schizophrenia GWAS. These tools constitute fast and powerful approaches to test gene set enrichment while correcting for confounding factors that can increase type 1 error rate (de Leeuw *et al*, 2016).

Aiming to comprehensively investigate the possible biological processes underlying schizophrenia, we set out to apply gene set and pathway enrichment analysis methods to the largest published GWAS in schizophrenia (Figure 1) (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). Our results elucidate the involvement of neuron differentiation and synaptic plasticity in schizophrenia, and reveal an accumulation of variants in post-synaptic signaling cascades. They moreover enable a more nuanced understanding of the several actionable classes of neurotransmitters implicated in the disease.

Materials and methods

Input data and analysis overview

We used summary-level results from the largest and most recent GWAS in schizophrenia, made publically available by the Psychiatric Genomics Consortium (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014) (www.med.unc.edu/pgc/results-and-downloads; downloaded on 10 May, 2017). This GWAS was performed in 34,341 schizophrenia cases and 45,604 healthy controls and association results for 9.4 million single-nucleotide polymorphisms (SNPs) were reported in the summary-level data. As detailed below (also see Figure 1), using Multi-marker Analysis of GenoMic Annotation (MAGMA) (de Leeuw *et al*, 2015) we successively (A) mapped SNPs to genes, (B) calculated gene p-values based on GWAS SNP p-values, (C) performed a primary gene set enrichment analysis using gene ontology (GO) terms, and (D) followed these findings up by a validation analysis on detailed molecular pathways derived from KEGG. Finally, we further investigated the results of the analysis on KEGG pathways using Meta-Analysis Gene-set Enrichment of variaNT Associations (MAGENTA) and Interval Enrichment Analysis (INRICH) (Lee *et al*, 2012; Segrè *et al*, 2010), and we applied protein-protein interaction analysis using Disease Association Protein-Protein Link Evaluator (DAPPLE) (Rossin *et al*, 2011).

Mapping SNPs to genes and assigning p-values to genes

Autosomal SNPs present in the European subset of the 1000 Genomes Phase 3 dataset were extracted from the GWAS summary-level results (The 1000 Genomes Project Consortium, 2015). Using MAGMA v1.06, we mapped SNPs to corresponding genes, extending gene footprints by an additional 20 kilobase (kb) up- and downstream, as a large proportion of regulatory elements involved in gene expression regulation is likely to be captured by including this region (Veyrieras *et al*, 2008). We then applied a gene analysis to obtain a p-

value for each gene to which at least one SNP was mapped. The p-value of a gene was based on the mean χ^2 statistic of SNPs contained in that gene, obtained using a known approximation of the sampling distribution. The European subset of 1000 Genomes Phase 3 was used as a reference dataset to estimate linkage disequilibrium (LD) between SNPs in the gene analysis, as the largest proportion of the schizophrenia GWAS meta-analysis was based on cohorts with northern and western European genetic ancestry (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014).

Gene Ontology (GO) term enrichment analysis

We downloaded 4,436 gene sets from the Molecular Signatures Database (MSigDB, release 6.0, April 2017), which are based on biological processes from the GO database (Gene Ontology Consortium, 2015; Liberzon *et al*, 2011). GO gene sets in MSigDB were limited to a minimum of 10 and a maximum of 2000 genes, and MSigDB merged original GO terms with high similarity (Jaccard's coefficient > 0.85) into one gene set. We used these gene sets in a competitive gene set analysis -which tests whether the genes in a gene set are more strongly associated with the phenotype of interest than random sets of genes- with the genes and assigned p-values from the previous step. The significance threshold was adjusted for multiple GO terms tested using a permutation procedure (10,000 times) implemented in MAGMA ($p = 3.9 \times 10^{-6}$). The competitive gene set analysis in MAGMA was corrected for confounding factors: gene size, linkage disequilibrium between genes that are close together (gene density) and minor allele count. We additionally tested the robustness of our analysis by implementing several sensitivity analyses: 1) excluding the MHC region on chromosome 6 (GRCh37/hg19 coordinates chr6:25500000-33500000); and 2) applying strict filtering for gene set size ($10 < n_{genes} < 200$) as larger gene sets are sometimes perceived as too broad or uninterpretable.

Follow-up using KEGG pathways

To further test the results of our primary GO term analysis, we used different gene sets representing synaptic signaling processes from the Kyoto Encyclopedia of Genes and Genomes (KEGG, www.kegg.jp, downloaded on 3 January, 2017) (Kanehisa and Goto, 2000) as input for MAGMA: hsa04724, Glutamatergic synapse (114 genes); hsa04725 Cholinergic synapse (111 genes); hsa04726 Serotonergic synapse (113 genes); hsa04727, GABAergic synapse (88 genes); hsa04728, Dopaminergic synapse (130 genes); hsa04720, Long-term potentiation (67 genes); hsa04730, Long-term depression (60 genes). We then ran a competitive gene set analysis on these pathways with correction for gene size, linkage disequilibrium (gene density) and minor allele count. The significance threshold was again adjusted for multiple testing using a permutation procedure (10,000 times) within MAGMA ($p = 0.009$). For each pathway enriched with schizophrenia-associated SNPs, we mapped significant schizophrenia-associated genes to pathway components using the *pathview* package in R version 3.3.3 (www.r-project.org) (Luo and Brouwer, 2013). We did not apply similar sensitivity tests as with the primary GO analysis, because the MHC region is not associated with synaptic signaling, and all gene sets contained more than 10 or less than 200 genes. Because many genes overlapped between two or more of the tested KEGG pathways (Figure S1), we applied a conditional competitive gene set analysis where we used gene sets different from the one tested as covariates to correct for overlapping genes present in the other gene sets in order to correct for the influence of shared genes on the enrichment signal. Additionally, we tested unique genes per KEGG gene set and one gene set containing overlapping genes in all tested KEGG sets for enrichment using MAGMA.

To support the results of the MAGMA analysis on KEGG pathways, we tested enrichment in the same KEGG pathways using MAGENTA and INRICH. MAGENTA maps SNPs onto

genes that are present in a gene set and assigns a score to each gene based on the lowest SNP p-value for that gene while correcting for confounding factors, such as gene size and gene overlap. It then assesses whether a gene set is enriched with low gene p-values at the 95th percentile cut-off (based on all gene p-values in the genome) compared to equally sized gene sets randomly sampled from the entire genome (Segrè *et al*, 2010). The MHC region was excluded in this analysis, we extended gene footprints 20 kb up- and downstream, and resulting enrichment p-values were corrected for multiple testing using false discovery rate (FDR < 0.05). INRICH tests whether sets of genes are enriched with association signals from independent genomic intervals (Lee *et al*, 2012). Intervals were calculated around genome-wide significant SNPs ($p < 5 \times 10^{-8}$) using the clump function in PLINK v1.90b3z (Chang *et al*, 2015). Only SNPs with a LD $R^2 > 0.5$ and an association p-value < 0.05 were included in an interval. Interval ranges were extended with 20 kb up- and downstream regions. Enrichment of 114 non-overlapping intervals in KEGG gene sets was tested in INRICH using a permutation procedure where the enrichment of a gene set was tested against the normal distribution of enrichment in background gene sets. The enrichment p-value for each gene set was corrected for multiple gene sets tested by a bootstrap method in INRICH.

Protein-protein interaction analysis

To increase insight into the biological impact of the pathways implicated in schizophrenia by our analyses, we performed protein-protein interaction analysis (PPI) using DAPPLE v0.17 (Rossin *et al*, 2011). Using permutations, DAPPLE tests whether (direct and indirect) networks of proteins that are coded by genes in these pathways showed more direct and indirect interactions with each other and with other proteins than expected by chance. We used 28 unique genes that were significantly enriched in the KEGG cholinergic synapse pathway (gene $p = 0.05/111 < 4.5 \times 10^{-4}$), KEGG dopaminergic synapse pathway (gene p =

$0.05/130 < 3.8 \times 10^{-4}$) or KEGG long-term potentiation pathway (gene $p = 0.05/67 < 7.5 \times 10^{-4}$).

Results

By performing a primary gene set enrichment analysis in MAGMA (in 4,436 biological processes described in the GO database) we identified two terms that were enriched for schizophrenia-associated SNPs (Figure 2A and Table S1): 'Regulation of synaptic plasticity' ($p = 1.13 \times 10^{-6}$) and 'Regulation of neuron differentiation' ($p = 1.18 \times 10^{-6}$). These significant GO terms are specialized terms of synaptic signaling and neuron differentiation, respectively (Figure S2). Excluding the extended MHC region did not affect the outcome this analysis (Figure S3). When we applied strict filtering for gene set size, we found that the 'Regulation of neuron differentiation' GO term was excluded from the analysis and only 'Regulation of synaptic plasticity' remained significant (Figure S3).

To gain a more nuanced understanding of molecular synaptic pathways enriched for schizophrenia-associated variants, we tested enrichment of SNPs in pathways that we derived from KEGG and that represented synaptic signaling resulting in synaptic plasticity, synaptic growth and survival (as indicated by the primary MAGMA enrichment analysis). We found a significant enrichment in pathways representing the dopaminergic synapse ($p = 1.5 \times 10^{-4}$), cholinergic synapse ($p = 2.0 \times 10^{-3}$) and long-term potentiation ($p = 3.8 \times 10^{-3}$) (Figure 2B, Table S2). For each significant pathway, we mapped genes enriched with schizophrenia SNPs on components within these KEGG pathways (schematic in Figure 3A, and detailed pathways in Figure S4). SNP enrichment was mostly restricted to trans-membrane and postsynaptic components in the cholinergic and dopaminergic synapses. The long-term potentiation pathway only included post-synaptic genes. We found high enrichment in signaling through extracellular signal-regulated kinase (ERK) and cAMP response element-binding protein (CREB), phospholipase C (PLC) and the inositol trisphosphate receptor (IP_3R), and signaling through protein kinase B (PKB/Akt). These cascades converge on mechanisms involved in synaptic growth regulation and synaptic plasticity. Furthermore,

voltage-gated calcium channels, glutamatergic NMDA and AMPA receptors, the dopamine D2 receptor (DRD2) and the muscarinic acetylcholine receptor M_4 were highly enriched (Figure 3A, Figure S4). Protein products of significant schizophrenia-associated genes (28 in total) from the three enriched pathways showed more interactions than expected by chance ($p = 1.0 \times 10^{-3}$) and on average more indirect connections to other proteins ($p = 2.0 \times 10^{-3}$), but not direct connections to other proteins ($p = 0.21$) (Figure 3).

When we conditioned the competitive gene set analysis of each KEGG pathway on the other pathways tested, thereby using overlapping genes in other pathways as covariates and adjusting for the effect of shared genes between pathways, the enrichment p-value of all pathways dropped (Figure S5): None of the pathways remained significant at the established significance threshold ($p = 0.009$), indicating a contribution of shared genes to the enrichment signal. Overlapping genes tested as a separate gene set were however not enriched after multiple testing correction in a combined analysis with gene sets containing unique genes ($p = 0.01$) (Figure S6).

Using MAGENTA and INRICH with the same data and gene sets, we also found significant enrichment of the dopaminergic synapse (MAGENTA corrected FDR = 5.0×10^{-4}, INRICH corrected p-value = 1.2×10^{-2}) and long-term potentiation (MAGENTA corrected FDR = 2.8×10^{-3}, INRICH corrected p-value = 3.0×10^{-3}). Enrichment for the cholinergic synapse was only confirmed in MAGENTA (corrected FDR = 4.8×10^{-3}), while enrichment of the glutamatergic synapse gene set was additionally found using INRICH (corrected p-value = 2.36×10^{-2}) (Figure 2C, D).

Discussion

By implementing complementary gene set enrichment analysis tools (MAGMA, MAGENTA and INRICH) and annotations from biological databases (GO and KEGG) into a comprehensive analysis, we aimed to elucidate biological processes underlying schizophrenia. Thus, we first detected enrichment of schizophrenia-associated SNPs in synaptic plasticity and neuronal differentiation processes, which is in line with biological hypotheses that previous studies have made (Hertzberg *et al*, 2015; Lips *et al*, 2012; Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium, 2015). We did however not replicate significant enrichment of immune and histone pathways (Figure S7), which has also been reported as highly associated to psychiatric disease (Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium, 2015). Possible reasons for this are the larger sample size and the exclusion of other phenotypes than schizophrenia in the current study. We followed this up in a targeted analysis on gene sets representing synaptic signaling in all major neurotransmitter systems, and demonstrated enrichment of schizophrenia SNPs in the dopaminergic system, long-term potentiation through the glutamatergic system and the cholinergic synapse, although the latter could not be confirmed using INRICH. Finally, we show that protein products of significant schizophrenia-associated genes in our enriched KEGG pathways show on average a higher rate of interactions with other proteins than expected by chance.

Dysfunctional synaptic transmission impacts synaptic plasticity and brain development, mediated through long-term potentiation (LTP) and long-term depression (LTD) (Bernardinelli *et al*, 2014). Although all five major neurotransmitter systems (dopamine, gamma-aminobutyric acid, glutamate, serotonin, and acetylcholine) have been implicated in schizophrenia, the extent to which each of them is involved had remained elusive (Kahn *et al*, 2015; Pocklington *et al*, 2014). We here studied SNP enrichment in these

processes, using gene sets of major neurotransmitter systems from the KEGG database, in which we also visualized this enrichment. Our results strongly support the involvement of the dopaminergic system, which has been extensively examined in schizophrenia. Previous studies have reported increased dopamine synthesis and release, and increased dopamine receptor expression in schizophrenia (Kaalund *et al*, 2014; Kahn *et al*, 2015). *DRD2* genetic variants are also implicated in schizophrenia and several of its intermediate phenotypes (Luykx *et al*, 2017; Vink *et al*, 2016). We here confirm an accumulation of *DRD2* genetic variants in schizophrenia. Moreover, we extend this evidence for involvement of the dopaminergic system by highlighting enrichment of variants in signaling cascades downstream of this receptor. Enrichment of the glutamate-induced LTP pathway was another finding that could be verified using MAGMA and MAGENTA. Mediation of LTP is, however, not limited to the glutamatergic system as post-synaptic signaling molecules such as the above mentioned CREB, IP_3R and PKB mediate synaptic plasticity in other neurotransmitter systems (e.g. the dopaminergic system). Multiple lines of evidence link LTP to cognitive deficits in schizophrenia (Salavati *et al*, 2015). Cholinergic transmission may also be relevant to symptomatology of schizophrenia, especially in light of the high rates of nicotine abuse and a range of cognitive symptoms (Carruthers *et al*, 2015; Parikh *et al*, 2016). The implication of acetylcholine in schizophrenia is further supported by a landmark study investigating chromatin interactions between enhancer regions containing schizophrenia-associated loci and promoter regions of target genes (Won *et al*, 2016). A recent, not yet peer-reviewed, GWAS meta-analysis on schizophrenia, that extended the data used in this study with a British cohort and also performed gene set enrichment analysis, also points towards synaptic transmission (doi.org/10.1101/068593). In contrast to our study, it also identified significant association at two genes involved in GABA signaling (*GABBR2* and *SCL6A11*). Although these genes had assigned gene p-values below 0.05 in our analysis, we

did not find and enrichment of the GABAergic synapse after multiple-testing correction. Possible explanations could be phenotypic differences between the 2014 schizophrenia GWAS and the recent meta-analysis (e.g. enrichment of relatively treatment-resistant patients in the latter), or the fact that the contribution of these genes to the enrichment of the GABAergic synapse pathway is diluted by non-associated genes present in this gene set. In a broader perspective, however, our findings are in line with rare variant studies in schizophrenia that reported mutations in genes coding for postsynaptic signaling components (CNV and Schizophrenia Working Groups of the Psychiatric Genomics Consortium, 2017; Fromer *et al*, 2014; Kirov *et al*, 2012; Purcell *et al*, 2014).

Our detailed analysis of downstream signaling cascades in all major neurotransmitter system gene sets revealed several of these cascades to be highly enriched for schizophrenia-associated variants: the phospholipase pathway, CREB signaling and the protein kinase B signaling cascade. All of these cascades may be linked to schizophrenia by numerous lines of neurobiological evidence, as outlined below. First, the phospholipase pathway (particularly PLC) controls neuronal activity and thereby maintains synaptic functioning and development. Activation of PLCβ and PLCγ results in cleavage of phosphatidylinositol 4,5-bisphosphate (PIP2) into the active form inositol 1,4,5-trisphosphate (IP_3) (Yang *et al*, 2016), in whose receptor (IP_3R) we report enrichment of schizophrenia-associated SNPs. Gene deletions in *PLC* are associated with schizophrenia and altered expression of *Plc* and schizophrenia-like behavior have been reported in *Plc* knock-out mice (Koh, 2013; Vasco *et al*, 2012). Second, signaling through the cellular transcription factor CREB modulates synaptic plasticity. A recent study focusing on the cyclic adenosine monophosphate (cAMP)/PKA/CREB pathway shows a significant association of a SNP in this system with schizophrenia (Forero *et al*, 2016). Additionally, ERK is part of the CREB signaling cascade and was also found to be enriched in our analyses. Impairment of signaling through ERK is hypothesized as a disease

mechanism in schizophrenia (Kyosseva, 2004; Yuan *et al*, 2010). Third, we found a significant enrichment of schizophrenia SNPs in postsynaptic protein kinase B (PKB or Akt). *AKT1* messenger RNA levels are higher in blood of schizophrenia patients compared to healthy controls and interactions between genetic variation in *AKT1* and cannabis use are associated with schizophrenia, possibly mediated through AKT signaling downstream of DRD2 (Liu *et al*, 2016; van Winkel, 2011). Interestingly, phosphorylation of glycogen synthase kinase 3 beta (Gsk3β) by the antipsychotic aripiprazole is mediated by Akt (Pan *et al*, 2015). Finally, we detected an accumulation of SNPs in protein phosphatase 1 (PP1) and protein phosphatase 2A (PP2A). PP2A is one of the mediators of sensorimotor gating, an intermediate phenotype for schizophrenia (Kapfhamer *et al*, 2010).

The finding of higher interaction rate of proteins coded by SNP-enriched genes in the dopaminergic, cholinergic and LTP pathways might highlight an important biological role of these genes in various biological processes, including synaptic signaling. This protein interaction finding is furthermore interesting in light of a novel hypothesis recently outlined in a hallmark paper about core disease genes that are connected to many other gene networks (Boyle *et al*, 2017).

Several limitations should be considered when interpreting our results. First, no significant SNP enrichment was found in other systems hypothesized to be dysregulated in schizophrenia, such as glutamatergic and GABAergic neurotransmission (Kahn *et al*, 2015; Schmidt and Mirnics, 2015; Yin *et al*, 2012). As our analyses are dependent on the power of GWAS, we cannot rule out the possibility that increased sample sizes in future studies may flag such systems, as illustrated by the identification of genes involved in GABA signaling in novel GWAS (doi.org/10.1101/068593). Second, we can only test for enrichment in gene sets and pathways that are annotated based on the knowledge currently available. Third, only protein-coding regions of the genome and up- and downstream regions in close proximity to

genes were considered in our analyses. Recently, it has become clear that non-coding stretches of the genome account for a major part of disease heritability and transcription regulation (Won *et al*, 2016). As the current analyses do not allow us to probe non-coding regions, we cannot take into account the effects that such genomic areas may have on neurotransmitter systems. Despite these limitations and slight technical differences between MAGMA, MAGENTA and INRICH (de Leeuw *et al*, 2015; 2016; Lee *et al*, 2012; Segrè *et al*, 2010), we found two pathways to be significantly enriched in results generated from each approach, and one enriched pathway in two out of three methods, suggesting our findings are fairly robust to analytic approach. Future integration of expression quantitative trait locus (eQTL) data and genomic interactions in pathway analysis tools has the potential to further deepen our understanding of the molecular mechanisms involved in the pathways underlying schizophrenia.

In conclusion, using complementary enrichment analysis approaches, we highlight downstream signaling cascades as the most likely part of the dopaminergic, cholinergic and glutamatergic (mediating LTP) systems to have a role in schizophrenia. Conditional analysis correcting mainly for shared genes showed a non-significant enrichment of all neurotransmitter gene sets, supporting the hypothesis that shared signaling mechanisms across the implicated pathways are likely to be a stronger underlying factor in schizophrenia than the independent neurotransmitter systems. Genes involved in these shared signaling mechanisms might be core genes which are more important in schizophrenia pathology, as supported by our protein interaction analyses. Our results open avenues for further research aimed at elucidating signaling pathways in schizophrenia, e.g. through tissue-specific manipulation of pathways in animal or cell models. Finally, our findings may aid the discovery of novel drug targets to hopefully reduce the burden imposed on quality of life in patients suffering from this disabling disorder.

Funding and disclosure

No funding was provided to carry out the current project.

Acknowledgments

The authors thank the Psychiatric Genomics Consortium for sharing summary-level genetic data on which this research is based.

Figures

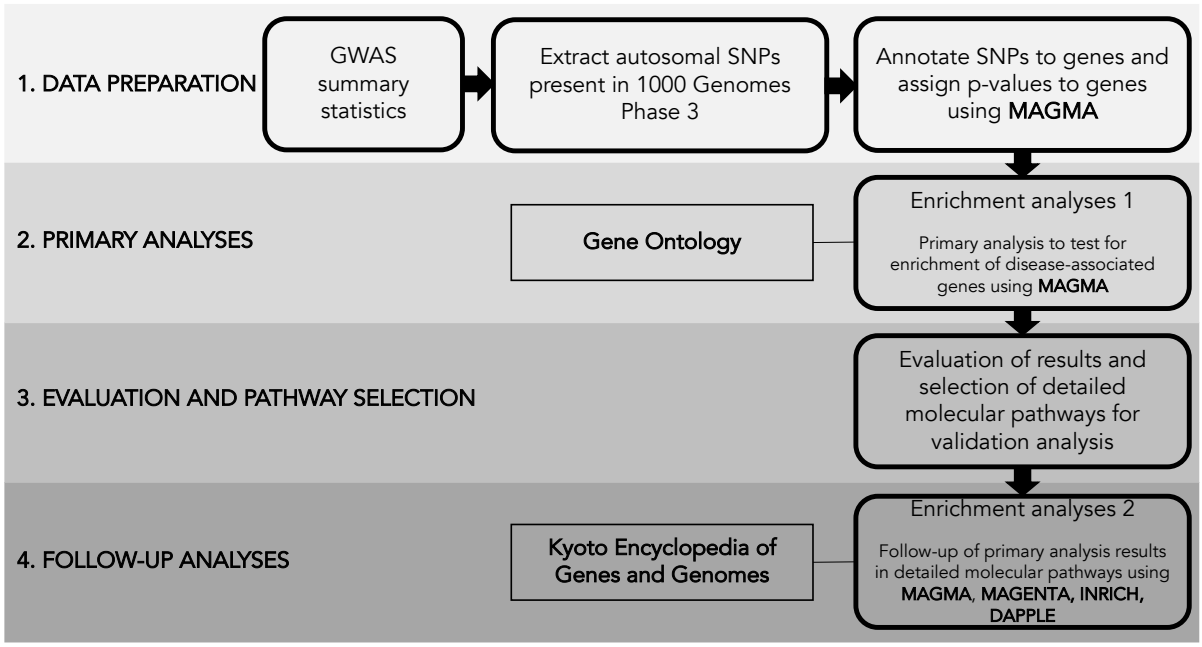

Figure 1 Overview of our pathway analysis pipeline. The analysis pipeline consists of four stages. In the first stage, data is prepared for analysis by assigning SNPs to genes and calculating gene p-values using 1000 Genomes Phase 3 as a reference for calculating linkage disequilibrium (LD). Second, a primary analysis is performed on biological processes included in the extensively annotated gene ontology database. Third, the results from the primary analysis are evaluated and biological pathways of interest are selected from detailed molecular databases such as the Kyoto Encyclopedia of Genes and Genomes (KEGG) are selected. Finally, pathway analysis on these detailed molecular pathways are performed to further investigate involvement of biological processes found in the primary analysis and to obtain a more nuanced understanding of SNP enrichment in these processes.

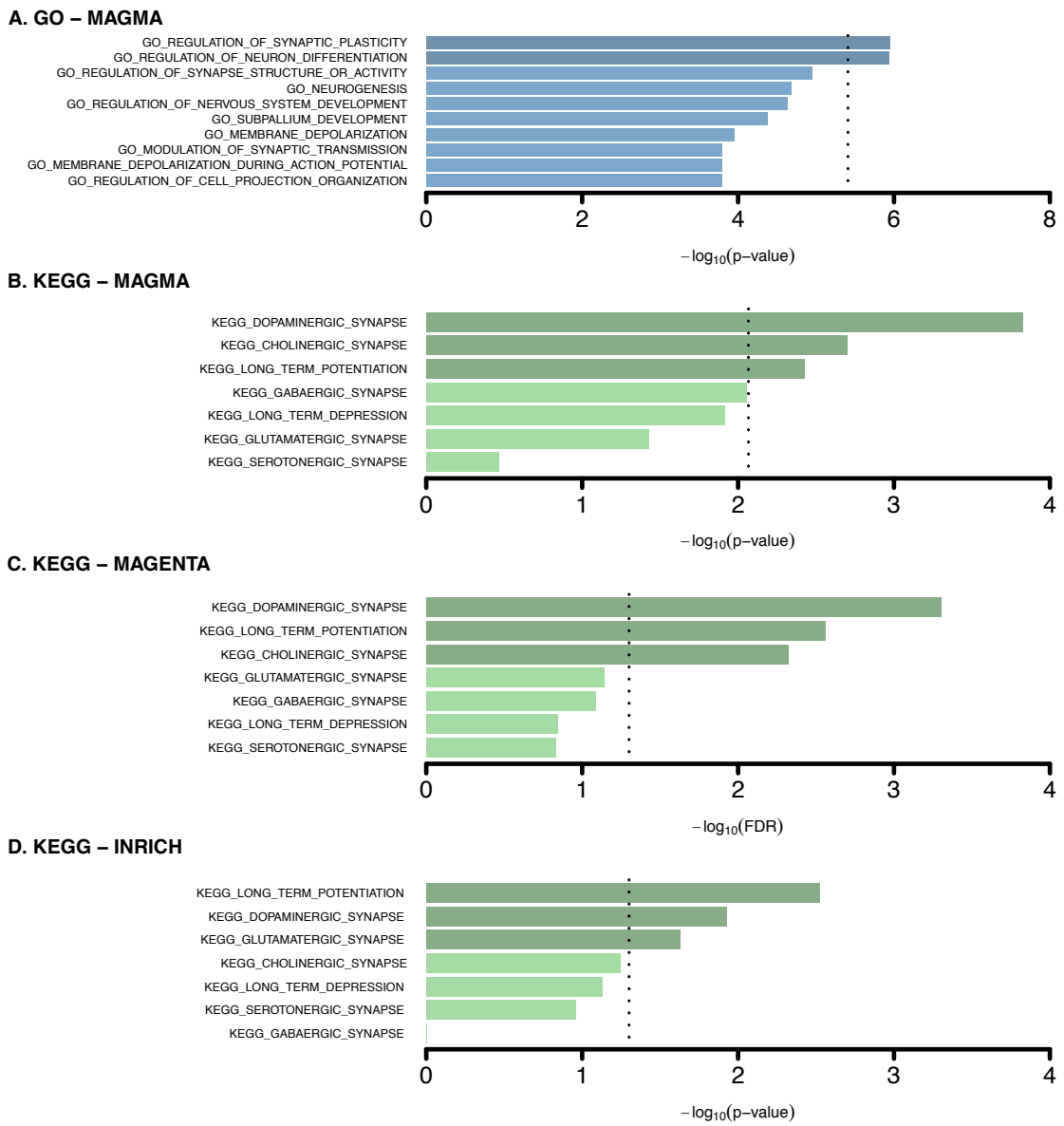

Figure 2 Results of gene set enrichment analysis on biological processes from Gene Ontology (GO) and detailed synaptic signaling pathways from KEGG. Reported p-values or FDR are $-\log_{10}$ converted. Significance thresholds adjusted for multiple gene sets tested are different per analysis depending on the used method and indicated with dotted lines. **(A)** Significantly enriched GO terms (shown in dark blue) are 'Regulation of synaptic plasticity' ($p = 1.13 \times 10^{-6}$) and 'Regulation of neuron differentiation' ($p = 1.18 \times 10^{-6}$) (significance threshold $p = 3.94 \times 10^{-6}$). Seven KEGG pathways representing synaptic signaling processes were tested for schizophrenia SNP enrichment using **(B)** MAGMA (significance threshold p

= 0.009), **(C)** MAGENTA (significance threshold FDR = 0.05) and **(D)** INRICH (significance threshold p = 0.05). Significantly enriched pathways (dark green bars) confirmed by three methods were dopaminergic synapse (MAGMA $p = 1.5 \times 10^{-4}$, MAGENTA FDR = 5.0×10^{-4} and INRICH $p = 1.2 \times 10^{-2}$), long-term potentiation (MAGMA $p = 3.8 \times 10^{-3}$, MAGENTA FDR = 2.8×10^{-3} and INRICH $p = 3.0 \times 10^{-3}$). Enrichment in cholinergic synapse (MAGMA $p = 2.0 \times 10^{-3}$ and MAGENTA FDR = 4.8×10^{-3}) and glutamatergic synapse (INRICH $p = 2.36 \times 10^{-2}$) was not confirmed by all three methods.

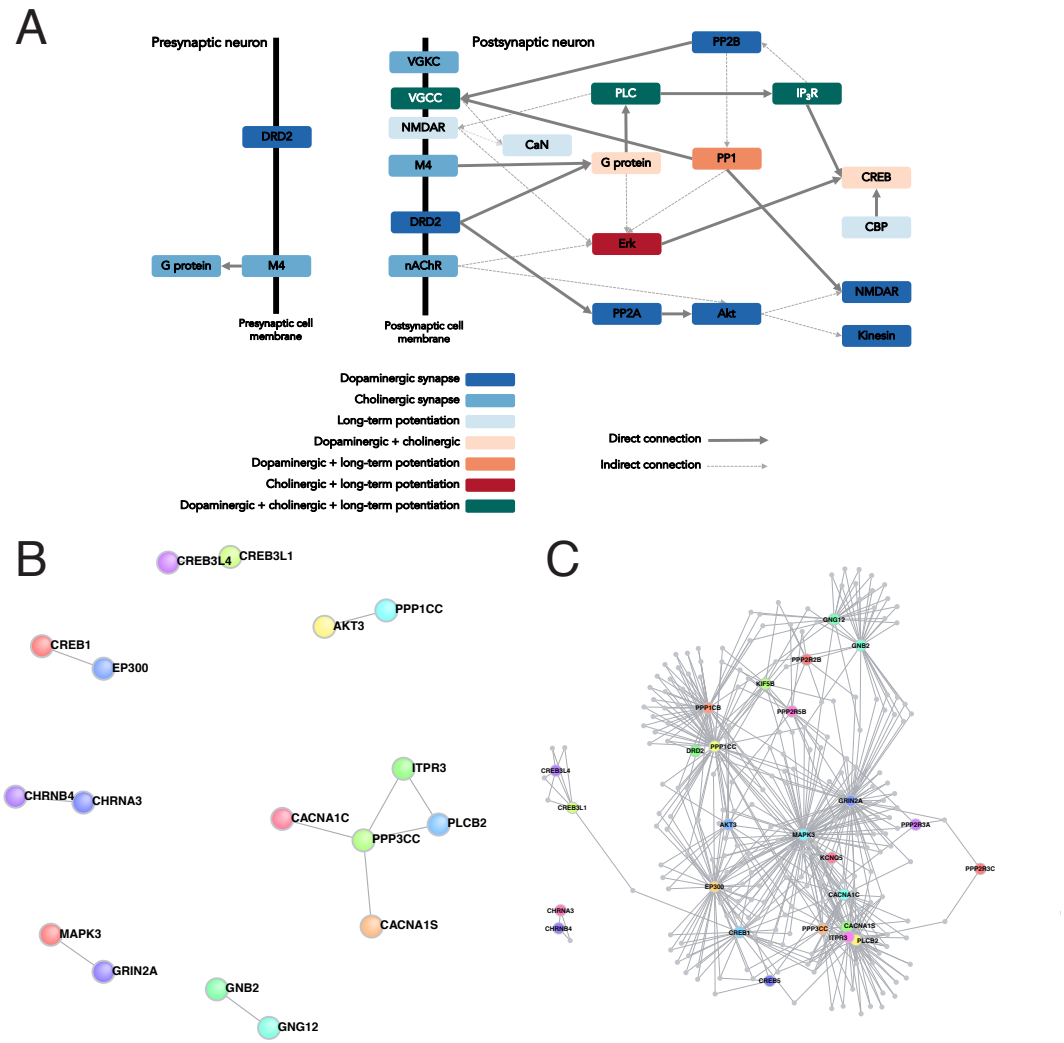

Figure 3 Enrichment and protein-protein interactions in synaptic signaling pathways. **(A)** Pathway components with gene p-values below 0.05/number of genes in the corresponding gene set were included in this overview (gene significance threshold $p = 0.05/130 = 3.8 \times 10^{-4}$ for dopaminergic synapse, $p = 0.05/111 = 4.5 \times 10^{-4}$ for cholinergic synapse, and $p = 0.05/67 = 7.5 \times 10^{-4}$ for long-term potentiation). Colors indicate in which gene sets the molecular components were present and enriched. Solid arrows indicate a direct connection between signaling components, dashed lines indicate an indirect connection (meaning the connection is separated by one or more non-enriched components).

Connections between components were based on KEGG (Kanehisa and Goto, 2000). **(B)** Direct (interaction $p = 1.0 \times 10^{-3}$) and **(C)** indirect (interaction $p = 2.0 \times 10^{-3}$) protein-protein interaction networks generated with DAPPLE (Rossin *et al*, 2011) for genes underlying pathway components from the dopaminergic synapse, long-term potentiation and cholinergic synapse gene sets. Dots indicate protein products of genes, lines indicate interactions with other proteins.

Abbreviations: DRD2, dopamine receptor D2; M4, muscarinic acetylcholine receptor M4; G protein, guanine nucleotide-binding protein; VGKC, voltage-gated potassium channel; VGCC, voltage-gated calcium channel; NMDAR, N-methyl-D-aspartate receptor; nAChR, nicotinic acetylcholine receptor; CaN, calcineurin; PP2A, protein phosphatase 2A; Akt, protein kinase B; Erk, Extracelluar signal-regulated kinase; CREB, cAMP response element binding protein; CBP, CREB-binding protein; PP1, protein-phosphatase 1; PLC, phospholipase C; IP_3R, inositol trisphosphate receptor; PP2B, protein-phosphatase 2B.

References

Bernardinelli Y, Nikonenko I, Muller D (2014). Structural plasticity: mechanisms and contribution to developmental psychiatric disorders. *Frontiers in Neuroanatomy* **8**:.

Boyle EA, Li YI, Pritchard JK (2017). An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell* **169**: 1177–1186.

Carruthers SP, Gurvich CT, Rossell SL (2015). The muscarinic system, cognition and schizophrenia. *Neurosci Biobehav Rev* **55**: 393–402.

Celada P, Bortolozzi A, Artigas F (2013). Serotonin 5-HT1A receptors as targets for agents to treat psychiatric disorders: rationale and current status of research. *CNS drugs* **27**: 703–716.

Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**: 7.

CNV and Schizophrenia Working Groups of the Psychiatric Genomics Consortium (2017). Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. *Nature genetics* **49**: 27–35.

de Leeuw CA, Mooij JM, Heskes T, Posthuma D (2015). MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLoS Comput Biol* **11**: e1004219.

de Leeuw CA, Neale BM, Heskes T, Posthuma D (2016). The statistical properties of gene-set analysis. *Nat Rev Genet* **17**: 353–364.

Duncan LE, Holmans PA, Lee PH, O'Dushlaine CT, Kirby AW, Smoller JW, *et al* (2014). Pathway analyses implicate glial cells in schizophrenia. *PloS one* **9**: e89441.

Forero DA, Herteleer L, De Zutter S, Norrback KF, Nilsson LG, Adolfsson R, *et al* (2016). A network of synaptic genes associated with schizophrenia and bipolar disorder. *Schizophrenia research* **172**: 68–74.

Fromer M, Pocklington AJ, Kavanagh DH, Williams HJ, Dwyer S, Gormley P, *et al* (2014). De novo mutations in schizophrenia implicate synaptic networks. *Nature* **506**: 179–184.

Gene Ontology Consortium (2015). Gene Ontology Consortium: going forward. *Nucleic acids research* **43**: D1049–56.

Hertzberg L, Katsel P, Roussos P, Haroutunian V, Domany E (2015). Integration of gene expression and GWAS results supports involvement of calcium signaling in Schizophrenia. *Schizophrenia research* **164**: 92–99.

Kaalund SS, Newburn EN, Ye T, Tao R, Li C, Deep-Soboslay A, *et al* (2014). Contrasting changes in DRD1 and DRD2 splice variant expression in schizophrenia and affective disorders, and associations with SNPs in postmortem brain. *Molecular psychiatry* **19**: 1258–1266.

Kahn RS, Sommer IE, Murray RM, Meyer-Lindenberg A, Weinberger DR, Cannon TD, *et al* (2015). Schizophrenia. *Nature Reviews Disease Primers* 15067doi:10.1038/nrdp.2015.67.

Kane JM, Correll CU (2010). Past and present progress in the pharmacologic treatment of schizophrenia. *The Journal of clinical psychiatry* **71**: 1115–1124.

Kanehisa M, Goto S (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research* **28**: 27–30.

Kapfhamer D, Berger KH, Hopf FW, Seif T, Kharazia V, Bonci A, *et al* (2010). Protein Phosphatase 2a and glycogen synthase kinase 3 signaling modulate prepulse inhibition of the acoustic startle response by altering cortical M-Type potassium channel activity. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **30**: 8830–8840.

Karow A, Wittmann L, Schöttle D, Schäfer I, Lambert M (2014). The assessment of quality of life in clinical practice in patients with schizophrenia. *Dialogues Clin Neurosci* **16**:

185–195.
Kirov G, Pocklington AJ, 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**: 142–153.
Koh HY (2013). Phospholipase C-beta1 and schizophrenia-related behaviors. *Advances in biological regulation* **53**: 242–248.
Köster L-S, Carbon M, Correll CU (2014). Emerging drugs for schizophrenia: an update. *Expert opinion on emerging drugs* **19**: 511–531.
Kyosseva SV (2004). The role of the extracellular signal-regulated kinase pathway in cerebellar abnormalities in schizophrenia. *Cerebellum* **3**: 94–99.
Lee PH, O'Dushlaine C, Thomas B, Purcell SM (2012). INRICH: interval-based enrichment analysis for genome-wide association studies. *Bioinformatics (Oxford, England)* **28**: 1797–1799.
Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P, Mesirov JP (2011). Molecular signatures database (MSigDB) 3.0. *Bioinformatics (Oxford, England)* **27**: 1739–1740.
Lichtenstein P, Björk C, Hultman CM, Scolnick E, Sklar P, Sullivan PF (2006). Recurrence risks for schizophrenia in a Swedish national cohort. *Psychological medicine* **36**: 1417–1425.
Lips ES, Cornelisse LN, Toonen RF, Min JL, Hultman CM, Holmans PA, *et al* (2012). Functional gene group analysis identifies synaptic gene groups as risk factor for schizophrenia. *Molecular psychiatry* **17**: 996–1006.
Liu L, Luo Y, Zhang G, Jin C, Zhou Z, Cheng Z, *et al* (2016). The mRNA expression of DRD2, PI3KCB, and AKT1 in the blood of acute schizophrenia patients. *Psychiatry research* **243**: 397–402.
Luo W, Brouwer C (2013). Pathview: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics* **29**: 1830–1831.
Luykx JJ, Broersen JL, de Leeuw M (2017). The DRD2 rs1076560 polymorphism and schizophrenia-related intermediate phenotypes: A systematic review and meta-analysis. *Neurosci Biobehav Rev* **74, Part A**: 214–224.
Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium (2015). Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nature neuroscience* **18**: 199–209.
Pan B, Chen J, Lian J, Huang X-F, Deng C (2015). Unique Effects of Acute Aripiprazole Treatment on the Dopamine D2 Receptor Downstream cAMP-PKA and Akt-GSK3β Signalling Pathways in Rats. *PLoS One* **10**: e0132722.
Parikh V, Kutlu MG, Gould TJ (2016). nAChR dysfunction as a common substrate for schizophrenia and comorbid nicotine addiction: Current trends and perspectives. *Schizophrenia research* **171**: 1–15.
Pocklington AJ, O'Donovan M, Owen MJ (2014). The synapse in schizophrenia. *Eur J Neurosci* **39**: 1059–1067.
Purcell SM, Moran JL, Fromer M, Ruderfer D, Solovieff N, Roussos P, *et al* (2014). A polygenic burden of rare disruptive mutations in schizophrenia. *Nature* **506**: 185–190.
Rossin EJ, Lage K, Raychaudhuri S, Xavier RJ, Tatar D, Benita Y, *et al* (2011). Proteins encoded in genomic regions associated with immune-mediated disease physically interact and suggest underlying biology. *PLoS Genet* **7**: e1001273.
Salavati B, Rajji TK, Price R, Sun Y, Graff-Guerrero A, Daskalakis ZJ (2015). Imaging-based neurochemistry in schizophrenia: a systematic review and implications for dysfunctional long-term potentiation. *Schizophrenia bulletin* **41**: 44–56.
Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014). Biological

insights from 108 schizophrenia-associated genetic loci. *Nature* **511**: 421–427.
Schmidt MJ, Mirnics K (2015). Neurodevelopment, GABA system dysfunction, and schizophrenia. *Neuropsychopharmacology* **40**: 190–206.
Segrè AV, Groop L, Mootha VK, Daly MJ, Altshuler D, DIAGRAM Consortium, *et al* (2010). Common Inherited Variation in Mitochondrial Genes Is Not Enriched for Associations with Type 2 Diabetes or Related Glycemic Traits. *PLoS Genet* **6**: e1001058.
The 1000 Genomes Project Consortium C (2015). A global reference for human genetic variation. *Nature* **526**: 68–74.
van Winkel R (2011). Family-based analysis of genetic variation underlying psychosis-inducing effects of cannabis: sibling analysis and proband follow-up. *Archives of general psychiatry* **68**: 148–157.
Vasco Lo VR, Cardinale G, Polonia P (2012). Deletion of PLCB1 gene in schizophrenia-affected patients. *Journal of cellular and molecular medicine* **16**: 844–851.
Veyrieras JB, Kudaravalli S, Kim SY, Dermitzakis ET, Gilad Y, Stephens M, *et al* (2008). High-resolution mapping of expression-QTLs yields insight into human gene regulation. *PLoS genetics* **4**: e1000214.
Vink M, de Leeuw M, Luykx JJ, van Eijk KR, van den Munkhof HE, van Buuren M, *et al* (2016). DRD2 Schizophrenia-Risk Allele Is Associated With Impaired Striatal Functioning in Unaffected Siblings of Schizophrenia Patients. *Schizophrenia bulletin* **42**: 843–850.
Wang K, Gaitsch H, Poon H, Cox NJ, Rzhetsky A (2017). Classification of common human diseases derived from shared genetic and environmental determinants. *Nat Genet* **6**: 111.
Won H, la Torre-Ubieta de L, Stein JL, Parikshak NN, Huang J, Opland CK, *et al* (2016). Chromosome conformation elucidates regulatory relationships in developing human brain. *Nature* **538**: 523–527.
Yang YR, Kang D-S, Lee C, Seok H, Follo MY, Cocco L, *et al* (2016). Primary phospholipase C and brain disorders. *Advances in biological regulation* **61**: 80–85.
Yin DM, Chen YJ, Sathyamurthy A, Xiong WC, Mei L (2012). Synaptic dysfunction in schizophrenia. *Advances in experimental medicine and biology* **970**: 493–516.
Yuan P, Zhou R, Wang Y, Li X, Li J, Chen G, *et al* (2010). Altered levels of extracellular signal-regulated kinase signaling proteins in postmortem frontal cortex of individuals with mood disorders and schizophrenia. *J Affect Disord* **124**: 164–169.