Schizophrenia Bulletin vol. 43 no. 2 pp. 459–471, 2017
doi:10.1093/schbul/sbw102
Advance Access publication July 22, 2016

SZDB: A Database for Schizophrenia Genetic Research

Yong Wu[1,2], Yong-Gang Yao[1–4], and Xiong-Jian Luo[*,1,2,4]

[1]Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences and Yunnan Province, Kunming Institute of Zoology, Kunming, China; [2]Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, China; [3]CAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai, China

[4]YGY and XJL are co-corresponding authors who jointly directed this work.

*To whom correspondence should be addressed; Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China; tel: +86-871-68125413, fax: +86-871-68125413, e-mail: luoxiongjian@mail.kiz.ac.cn

Schizophrenia (SZ) is a debilitating brain disorder with a complex genetic architecture. Genetic studies, especially recent genome-wide association studies (GWAS), have identified multiple variants (loci) conferring risk to SZ. However, how to efficiently extract meaningful biological information from bulk genetic findings of SZ remains a major challenge. There is a pressing need to integrate multiple layers of data from various sources, eg, genetic findings from GWAS, copy number variations (CNVs), association and linkage studies, gene expression, protein–protein interaction (PPI), co-expression, expression quantitative trait loci (eQTL), and Encyclopedia of DNA Elements (ENCODE) data, to provide a comprehensive resource to facilitate the translation of genetic findings into SZ molecular diagnosis and mechanism study. Here we developed the SZDB database (http://www.szdb.org/), a comprehensive resource for SZ research. SZ genetic data, gene expression data, network-based data, brain eQTL data, and SNP function annotation information were systematically extracted, curated and deposited in SZDB. In-depth analyses and systematic integration were performed to identify top prioritized SZ genes and enriched pathways. Multiple types of data from various layers of SZ research were systematically integrated and deposited in SZDB. In-depth data analyses and integration identified top prioritized SZ genes and enriched pathways. We further showed that genes implicated in SZ are highly co-expressed in human brain and proteins encoded by the prioritized SZ risk genes are significantly interacted. The user-friendly SZDB provides high-confidence candidate variants and genes for further functional characterization. More important, SZDB provides convenient online tools for data search and browse, data integration, and customized data analyses.

Key words: schizophrenia/genetic study/database/gene expression/integrative analysis

Introduction

Schizophrenia (SZ) is a severe mental disorder characterized by abnormal perceptions, incoherent or illogical thoughts, and disorganized speech and behavior. It affects approximately 0.5%–1% of the world populations[1] and is one of the leading causes of disability worldwide.[2–4] With the substantial morbidity and high mortality, SZ imposes profound impact on patients' quality of life, and significant economic burden on families and society.[4] It is estimated that the overall cost of SZ in the United States in 2002 was $62.7 billion.[5] In Canada, the total cost of SZ was estimated to be 6.85 billion (Canadian dollars) in 2004.[6] The global economic burden of SZ is continuously increasing, which makes SZ an urgent global health issue.

Despite the fact that millions of people are suffering from SZ and the substantial personal and societal costs, current available antipsychotic medications can only alleviate the symptoms of SZ. Moreover, the efficacy of antipsychotic drugs varies greatly among patients (ie, many patients are treatment-resistant).[7,8] So far, there is no drug or treatment that can completely cure SZ or ensure that there will be no further psychotic episodes. A key reason for this therapeutic conundrum is due to the unknown pathophysiological mechanism of SZ. During the past decades, great efforts have been made to identify the causes of SZ and to develop new drugs and treatments. Though significant progress has been achieved, the etiology of SZ remains largely unknown.

A growing body of evidence strongly suggest that both genetic and environmental factors are involved in the pathophysiology of SZ.[9] Previous studies have shown that environmental factors such as obstetric complications, exposure to influenza viruses during gestation, and prenatal exposure to maternal stress can increase risk of SZ.[10–15] Though environmental factors have a role in the pathogenesis of SZ, high heritability clearly shows the

major role of inherited genetic variants in the etiology of SZ.[16] Family, twin and adoption studies have consistently demonstrated the key role of genetic components in the pathogenesis of SZ. The heritability estimate of SZ is about 0.80,[16] which is higher than most of other complex diseases.[17]

To elucidate the genetic susceptibility to SZ, large numbers of linkage and association studies, as well as meta-analyses have been performed, and numerous risk variants, genes, and chromosomal regions have been reported to be associated with SZ.[18–20] Despite the fact that many promising candidate genes have been identified, the sample sizes are relatively small and the coverage of genetic markers are relatively low in these studies, which limited the identification of credible susceptibility variants and genes. The advent of genome-wide association studies (GWAS) provides critical opportunities to explore the genetic landscape of SZ. In 2008, *ZNF804A* had been identified as the first gene that reached a genome-wide significant level for SZ.[21] Since then, multiple GWAS have been conducted in world populations and numerous SZ-associated variants and loci were identified.[22–35] A landmark of SZ genetics is the establishment of Psychiatric Genomics Consortium (PGC), which performed the largest GWAS meta-analysis of SZ so far and identified numerous high-confidence genetic risk variants and loci.[36]

With the continuous increase of sample size in PGC, novel risk variants and genes can now be identified at unprecedented rates. Despite the great success of GWAS in identifying SZ risk variants, understanding the pathological role of these risk variants in SZ remains a daunting task. First, most of the SZ-associated variants identified by GWAS reside in noncoding regions (with limited annotation and no obvious functional consequence), how these risk variants contribute to SZ risk is largely unknown. There is a growing gap between the identification of SZ risk variants and the extraction of meaningful biological information. Second, besides the common genetic variants (single nucleotide polymorphisms, SNPs) identified by GWAS, other types of genetic variations including rare variants, copy number variation (CNV), and structural genomic variations play a pivotal role in SZ.[37–41] Therefore, there is a pressing need to integrate data from various genetic studies of SZ to provide a global view of current genetic findings on SZ. Third, systematic integration and in-depth analysis of multi-type data from various sources (eg, genetic and gene expression studies) will help to provide useful guide for further experimental validation and to establish the connection among different types of data. For example, selecting high-confidence candidate gene for further functional validation is challenging due to a large number of genes identified by genetic association studies. Gene prioritization analysis (eg, convergent functional genomics)[42] provides a valuable solution for selecting the candidate genes for experimental verification. Fourth, accumulating evidence imply that common polygenic variation contributes to SZ risk.[22] That is, SZ is likely caused by multiple genes that disrupt one or more molecular pathways.[43,44] Thus, pathway and network-based analysis is an effective method to identify the molecular pathways dysregulated in SZ and helps with mechanism investigation and drug development. To better understand the genetic architecture of SZ and to facilitate the translation of genetic findings into molecular risk mechanisms for SZ, we developed SZDB, a comprehensive database containing multiple integrative data from various layers of SZ genetic researches. SZDB not only provides a user-friendly web interface for data search, browse and visualize, but also provides in-depth data integration and analysis functions such as SNP functional annotation, spatio-temporal gene expression pattern analysis in human brain, expression quantitative trait loci (eQTL) and network-based (PPI and co-expression) analyses.

Methods

SNP Associations From GWAS of SZ

Several GWAS of SZ have been performed during the past decade. To further increase the statistic power, PGC2 performed the largest meta-analysis of SZ GWAS so far.[36] In brief, genome-wide genotype data of 49 ancestry-matched, nonoverlapping case-control samples and 3 family-based samples of European ancestry were obtained. The total sample size included in the GWAS meta-analysis was 82 315, including 35 476 SZ cases and 46 839 controls. After quality control and SNP imputation, association test was performed separately in each sample and the results were combined and meta-analyzed. More detailed information about sample description, genotyping and statistical analyses can be found in the original paper.[36] We downloaded the summary statistics (SNP associations) of PGC2 from the PGC website (http://www.med.unc.edu). In total, summary statistics of 9 444 230 SNPs were downloaded.

Plotting of Regional Association Signal

To visualize the results of genome-wide association data of PGC2,[36] we utilized the LocusZoom[45] to plot regional association results. The *P* values of SNPs were downloaded from the PGC website (http://www.med.unc.edu/pgc/downloads).[36] We downloaded the Hg19 genome assembly and 1000 Genomes population data from http://locuszoom.sph.umich.edu/locuszoom/.

SNP Function Annotation

Most of the SZ-associated SNPs identified by GWAS lie in noncoding region and are predominately located within putative regulatory elements,[36] suggesting that these variants may confer risk to SZ through affecting

gene expression. We took advantage of the recent study by Boyle et al[46] that systematically annotated SNPs in human genome using experimental data sets from ENCODE[47] and other sources.[46,48] Briefly, high-throughput computational and experimental data sets, including DNase footprinting, binding motif, ChIP-seq, chromatin state, eQTL, as well as other sources were used to predict if a specific SNP has a regulatory potential. To make the results easier to use and understand, a score system was introduced and a corresponding database (RegulomeDB, http://RegulomeDB.org/) was developed by Boyle and coworkers.[46] Lower scores in the RegulomeDB suggest increasing evidence for a variant to be located in a functional region. More detailed information about variant annotation and RegulomeDB database can be found in the original paper.[46] We downloaded the RegulomeDB score data for all SNPs (including genotyped and imputed) analyzed in PGC2[36] and integrated these scores into SZDB. We also annotated the potential functional consequences of the non-synonymous (missense) variants using SIFT[49,50] and PolyPhen-2[51], 2 extensively used prediction tools to conduct in silico functional prediction of non-synonymous (missense) variants.

Extraction and Compilation of SZ Candidate Genes

Previous studies (including genetic linkage and association studies,[18,19] GWAS,[36] integrative studies such as convergent functional genomics (CFG)[42] and *Sherlock* integrative analysis,[52] and whole exome sequencing studies[53,54]) have identified numerous SZ candidate genes. We systematically extracted and compiled a comprehensive list of SZ candidate genes from these studies. Based on the study type, SZ candidate genes were classified into the following categories:

(1) Genes identified by genetic linkage and association studies. To systematically combine the results of previous genetic linkage studies, Lewis et al[18] conducted a rank-based genome scan meta-analysis and identified multiple chromosomal regions that may harbor SZ susceptibility variants and genes. In total, 160 genes located in these chromosomal regions were extracted (supplementary table S1). Recently, Ng et al[19] also performed a systematic meta-analysis using 32 genome-wide linkage studies of SZ. We extracted all 173 genes identified by Ng et al[19] (supplementary table S2).

In addition to linkage studies, numerous SZ candidate genes have been identified by association studies and we included these genes in our list. Briefly, we extracted 42 top genes from the SZGene database (http://www.szgene.org/) prioritized by Allen et al,[20] who performed a systematic meta-analysis and identified multiple SZ risk genes (supplementary table S3). We also extracted SZ candidate genes from the SZGR database (formerly available at http://bioinfo.mc.vanderbilt.edu/SZGR/, but now could not be accessible) developed by Jia et al.[55] Through using combined odds ratio (COR) method, Sun et al[56] ranked and prioritized 38 core (supplementary table S4) and 75 SZ candidate genes (supplementary table S5) and these genes were included in the SZGR database.[55] We downloaded and integrated all these SZ candidate genes into SZDB, and a total of 226 nonoverlapping SZ candidate genes (supplementary table S6) were included.

(2) Genes identified by GWAS of SZ. Multiple GWAS of SZ had reported numerous risk genes.[22–35] We manually extracted SZ risk genes that reached the genome-wide significance level. Totally, 375 nonoverlapping genes (supplementary table S7) were extracted and integrated into SZDB.

(3) Genes identified by convergent functional genomics (CFG). By using a convergent functional genomics approach, Ayalew et al[42] prioritized 42 promising SZ candidate genes (supplementary table S8) and we included these genes into SZDB.

(4) Genes affected by CNVs. Accumulating evidence suggests that CNVs play an important role in SZ. We systematically prioritized the genes affected by CNVs in SZ and identified pivotal genes that are frequently disrupted by CNVs in SZ cases.[57] We extracted these genes (supplementary table S9) and integrated them into SZDB.

(5) Genes identified by *Sherlock* integrative analysis.[58] Multiple lines of evidence support the important role of gene expression dysregulation in SZ. Recently, He et al[58] developed a statistical framework (named *Sherlock*) to identify disease-associated genes through matching eQTL and SNP associations from GWAS. We performed integrative analysis using *Sherlock* and identified 12 promising SZ candidate genes.[52] We integrated these 12 genes into SZDB.

(6) Genes differentially expressed in SZ cases. Dysregulation of gene expression has a key role in SZ pathogenesis. To detect genes that are differentially expressed between SZ patients and healthy controls, many large-scale gene expression studies have been performed. We downloaded 5 microarray datasets (which used brain tissues from SZ patients and controls) from the gene expression omnibus (GEO; [http://www.ncbi.nlm.nih.gov/geo/]), including GSE53987 (contains 3 brain tissues: prefrontal cortex, striatum and hippocampus; a total of 114 samples),[59] GSE12649 (prefrontal cortex; a total of 69 postmortem brains samples),[60] GSE21138 (prefrontal cortex; a total of 59 postmortem brains samples),[61] GSE35978 (cerebellum and parietal cortex brain; a total of 195 samples),[62] and GSE62191 (frontal cortex; a total of 59 samples).[63] We re-analyzed these microarray datasets, and top 1% differentially expressed genes in SZ cases and controls were integrated into SZDB. The differentially expressed genes identified by RNA sequencing (RNA-seq)[64] were also integrated into SZDB.

(7) Genes identified by *Pascal* gene-based test. The gene-based analysis is a strategy to improve statistical power and gain biological insight.[65] Through this test, one can get an overview from the gene and pathway level rather than the SNP level. Here we used the gene-based test Pathway scoring algorithm (*Pascal*)[65] to integrate PGC2 GWAS data. A total of 343 genes with *P*-values lower than 10^{-6} were included in the SZDB (supplementary table S10).

(8) Genes identified by whole exome sequencing. Accumulating evidence suggest that *de novo* mutations and rare disruptive mutations play an important role in SZ pathogenesis. Recently, Fromer et al[54] performed the largest exome sequencing study of *de novo* mutations in SZ and they found that *de novo* mutations in SZ are overrepresented in synaptic networks. We integrated genes (supplementary table S11) that carrying *de novo* mutations (including coding and canonical splice site variants) in SZ probands from the study of Fromer et al[54] into SZDB. In addition, Purcell et al[53] identified numerous genes that contain rare disruptive mutations in SZ cases through exome sequencing. We also integrated the genes (supplementary table S12) identified by Purcell et al[53] into SZDB.

(9) Genes differentially methylated in SZ. In addition to genetic components, environmental factors may also play an important role in SZ. DNA methylation is a common epigenetic modification that is influenced by environmental factors, and recent studies[66,67] have identified multiple differentially methylated genomic regions (or genes) among SZ cases and healthy controls. We integrated the differentially methylated genes identified by Montano et al[67] and Aberg et al[66] into SZDB. In addition, we also integrated the methylation quantitative trait loci from study of Hannon et al[68] into SZDB.

In-depth Data Analysis and Systematic Integration

Based on the above comprehensive data/ gene lists from various sources, we carried out the following in-depth data analyses and systematic integration:

(1) eQTL analysis. Genetic variants affecting gene expression in brain tissues may confer risk of SZ. To investigate if genome-wide significant SNPs identified by PGC2 are associated with gene expression in human brain, we downloaded genome-wide expression and genotype data from the study of Myers et al,[69] which included 193 neuropathologically normal human brain samples. The association between each SNP and transcript was assessed by using PLINK,[70] as described in the study of Myers et al.[69] In total, we assessed correlations among 366 140 SNPs and the expression of 14 078 transcripts. Briefly, before the analysis, rank-invariant normalized expression data were $\log_{10}$ transformed, and missing data were encoded as missing, not as a zero level of expression. We excluded transcripts that were expressed in less than 5% of the series from the analysis. The following minimum SNP cut-off values were used: per sample call rate at least 90%, per SNP call rate at least 90%, per SNP minor allele frequency of at least 1%, and lack of significance ($P > .05$) for the Hardy-Weinberg equilibrium tests. Only those pairs with transcript-specific empirical P values ≤ .05 were retained. SNP position and gene position were based on the Hg19 genome assembly.

(2) Spatio-temporal expression pattern analysis. A pivotal step to elucidate the role of risk genes in SZ pathogenesis is to understand the spatio-temporal expression pattern of risk genes in brain. To investigate if SZ risk genes are preferentially expressed in specific brain tissues (eg, prefrontal cortex and hippocampus) and at specific developmental stages (eg, fetal and postnatal), we performed spatio-temporal gene expression pattern analysis using the RNA-seq based gene expression data from the BrainSpan: Atlas of the Developing Human Brain (http://www.brainspan.org).[71] These expression data were from 16 brain regions (a total of 42 postmortem brains) that spanned 15 consecutive periods of neurodevelopment and adulthood from 8 post-conceptual weeks (PCW) to 40 years. The data were partitioned into different developmental stages and subsets of brain regions as described by Willsey et al.[72] Data processing, quality control and statistical analyses were performed as described previously.[73] In addition to expression data from the BrainSpan,[71] we also downloaded expression data from the BrainCloud (http://braincloud.jhmi.edu/),[74] which contains gene expression data from prefrontal cortex. The developmental stages were also partitioned into 15 periods. More detailed information about the BrainCloud can be found in the original study[74] and the SZDB website (http://www.szdb.org/index.html).

(3) Co-expression analysis. Recent study suggested that SZ risk genes are significantly co-expressed in patterns specific to developmental stage and neuroanatomical structure.[73] Therefore, co-expression analysis may help to prioritize SZ candidate genes from large-scale genetic studies (eg, GWAS or transcriptome). To explore if SZ risk genes were co-expressed in specific brain regions at specific developmental stages, we performed gene co-expression analyses using the RNA-seq based expression data from the BrainSpan: Atlas of the Developing Human Brain (http://www.brainspan.org/).[71] Briefly, we divided the tissues from the BrainSpan atlas into 4 anatomic regions and 15 developmental stages as described by Willsey et al[72] and Gilman et al.[75] The 4 anatomic regions are: (1) V1C-STC cluster (V1C, ITC, IPC, A1C, and STC); (2) Prefrontal and primary motor-somatosensory

cortex or M1C-S1C cluster (M1C, S1C, VFC, MFC, DFC, and OFC); (3) STR-AMY cluster (STR, HIP, AMY); (4) MD-CBC cluster (MD and CBC). The 15 developmental stages are: (1) Embryonic, from 4 to 8 post-conception weeks (PCW); (2) Early fetal, from 8 to 10 PCW; (3) Early fetal, from 10 to 13 PCW; (4) Early mid-fetal, from 13 to 16 PCW; (5) Early mid-fetal, from 16 to 19 PCW; (6) Late mid-fetal, from 19 to 24 PCW; (7) Late fetal, from 24 to 38 PCW; (8) Neonatal & early infancy, from 0 to 6 months (M); (9) Late infancy from 6 to 12 M; (10) Early childhood, from 1 to 6 years (Y); (11) Middle and late childhood, from 6 to 12 Y; (12) Adolescence, from 12 to 20 Y; (13) Young adulthood, from 20 to 40 Y; (14) Middle adulthood, from 40 to 60 Y; (15) Late adulthood, more than 60 years old. Pearson correlation coefficients were calculated in each brain region across developmental stages as described previously.[72,73] Only gene pair with a Pearson correlation coefficient equal or bigger than 0.8 was retained.

(4) Protein–protein interaction (PPI) analysis. Accumulating evidence suggest that proteins involved in the same disease are more likely to interact with each other (ie, guilt-by-association).[76,77] We had showed that SZ risk genes encode an highly interconnected network,[78,79] implying PPI information may help to prioritize SZ candidate genes. To explore if proteins encoded by SZ candidate genes are physically interacted with proteins encoded by other genes (including SZ risk genes and others), we downloaded PPI data from InWeb, a well-characterized PPI database developed by Lage et al.[80,81] Genes directly interact with SZ genes were regarded as the first-degree interaction genes.

(5) Analysis of differentially expressed genes. Genes differentially expressed in SZ cases and healthy controls may play a pivotal role in the etiology of SZ. Though multiple gene expression studies have been performed to identify differentially expressed genes in SZ,[59–63] we noticed that only limited overlapping genes were identified among different studies. To compare the differentially expressed from different studies, we preformed differentially expression analyses using microarray datasets from GEO. The main inclusion criteria are as follows: First, GEO datasets with a relatively large sample size (more than 20 SZ cases) were selected. Second, GEO datasets using brain tissues (including prefrontal cortex, hippocampus, and striatum) were given higher priority. Third, GEO datasets that contain genome-wide gene expression data were selected. It should be noted that only limited expression datasets are public available and can be downloaded from GEO. Based on these criteria, we selected 5 GEO datasets[59–63] for differential expression analysis. GEO2R (http://www.ncbi.nlm.nih.gov/geo/geo2r/), which uses limma (Linear Models for Microarray Analysis) in R package (an Empirical Bayes method), was used to analyze differentially expressed genes. GEO2R includes the following steps: (1) Load series and platform data from GEO; (2) Group names for all samples; (3) Transform expression data by log2 (if the data were not transformed, auto-detect by GEO2R); (4) Set up the data and proceed with analysis; (5) Load NCBI platform annotation; (6) Replace original platform annotation. To visualize these results, we used the jQuery plugin plotly (https://plot.ly/javascript/) to draw a boxplot for every gene in SZ cases and controls. We also downloaded the differentially expressed genes (a total of 798 genes) identified by the RNA-seq[64] and integrated these results into SZDB.

(6) Co-expression module analysis. To explore if SZ risk genes show similar expression pattern during cortical development, we performed co-expression network analysis using the RNA-seq based expression data from the BrainSpan.[71] Signed weighted gene co-expression network analysis was performed as previously described.[82]

(7) Gene-based analysis. Currently, most of the GWAS tested the association between a trait and individual genetic variant (usually SNP). Gene-based method is a useful complement to GWAS as it considers association between a trait and all variants (eg, SNPs) within a gene rather than each marker individually. To identify the genes that are associated with SZ at the gene level, we performed gene-based analysis using the *Pascal* software (summary statistics, ie, SNP *P* values from PGC2 were used as input).[65] More detailed information can be found in the supplementary material.

Prioritization of SZ Risk Genes

We used an arbitrary cumulative scoring method to prioritize SZ risk genes. The rationale of our method assumed that SZ-associated gene may be detected in several independent studies (or datasets). For example, if a specific gene was identified to be associated with SZ in genetic studies and gene expression study also showed that this gene is dysregulated in SZ cases, we would score this gene as a promising SZ candidate gene. More detailed information about this polyevidence scoring approach can be found in the paper of Ayalew et al[42] in which they identified promising SZ candidate genes through integrating multiple lines of evidence from different SZ researches. We took advantage of the following lines of evidence to prioritize SZ candidate genes in SZDB:

(1) Genes identified by GWAS. If a gene was reported to be associated with SZ in GWAS, the total score of this gene rises by 1 point.

(2) Genes identified by genetic linkage and association studies. If a gene was reported to be associated with SZ in previous genetic linkage or association studies, the total score of this gene rises by 1 point.

(3) Genes identified by convergent functional genomics (CFG). If a gene was in the top gene list (supplementary table S8) prioritized by CFG,[42] the total score of this gene rises by 1 point.
(4) Genes frequently disrupted by CNVs. If a gene was frequently disturbed by CNVs in SZ (supplementary table S9),[57] the total score of this gene rises by 1 point.
(5) *Sherlock* integrative analysis. If a gene was identified to be associated with SZ by *Sherlock* integrative analysis,[52] the total score of this gene rises by 1 point.
(6) Genes differentially expressed in SZ cases and controls. If the expression level of a specific gene was reported to be significantly changed in SZ cases compared with controls, the total score of this gene rises by 1 point.
(7) Gene-based evidence. If a gene showed a significant association with SZ in the gene-based test, the total score of this gene rises by 1 point.

We summed up all scores for each gene and those genes with a score ≥2 were regarded as the prioritized genes (supplementary table S13).

Drug Target Identification

We explored whether the top 29 prioritized genes in SZDB are drug targets through using the Drug-Gene interaction database (http://dgidb.genome.wustl.edu/)[83] and the list of 1030 druggable genes compiled by Rask-Andersen et al.[83] The druggable gene list by Rask-Andersen et al[83] was developed by systematically integration of public available druggable targets resources, including the DrugBank database (http://www.drugbank.ca),[84] the Drugs@FDA database (http://www.accessdata.fda.gov/scripts/cder/drugsatfda/), and the Drugs in Clinical Trials Database (http://www.centerwatch.com).

Results

Database Overview

On the basis of comprehensive data collection and curation, we performed in-depth data analysis and systematic integration and included all these results in the SZDB. Therefore, SZDB not only contains genetic risk variants and susceptibility genes for SZ, but also provides a demonstration of the in-depth analytical results, such as SNP function annotation, eQTL, spatio-temporal expression and co-expression, and PPI data. Specifically, SZDB integrates powerful online analysis tools and users can easily perform custom analyses such as eQTL, PPI, co-expression and spatio-temporal expression pattern analyses (figure 1). All of the preprocessed datasets can be freely downloaded at SZDB (http://www.szdb.org/download.html). The user-friendly web interface of the SZDB contains 3 core modules: Search module, Analysis and Tools module, and Genes module.

(1) Search module. This module provides a powerful search engine for SNP and gene query. Users can input interested SNPs and genes in the search bar. Multiple search items (eg, a set of interested SNPs or genes) are allowed for single query. For SNP query, SZDB will generate a detailed report (supplementary figure S1), including chromosomal location based on the Hg19 genome assembly and corresponding genomic coordinates of the query SNP, SNP alleles (polymorphism), SNP location in gene (flanking or nearby genes of the query SNP), SNP type (eg, synonymous or non-synonymous substitution), association significance (OR and *P* value) with SZ (from PGC2[36]), eQTL (genes whose expression are associated with the query SNP),[69] regulatory potential (RegulomeDB score[46,48]). If the query SNP is a non-synonymous SNP, SZDB will also predict the potential impact of amino acid change on protein structure and function using SIFT[49,50] and PolyPhen-2.[51] To visualize the association significance between the query SNP and SZ, a LocusZoom link for the query SNP is provided (supplementary figure S1). The LocusZoom generates and plots the regional association results between the query SNP (including SNPs that are located 200 kb up and downstream of the query SNP) and SZ using GWAS data from PGC2.[36] The query SNP is highlighted in the LocusZoom results, so the users can evaluate if the query SNP has the smallest association significance (*P* value) with SZ. For gene query, SZDB will return a "Results" page which includes the Entrez gene id, genomic coordinate, cytogenetic location, and brief description of the query gene (supplementary figure S2). Of note, a LocusZoom link for the query gene is also provided, and the users can download and browse the association significance between SNPs located in this gene and SZ (based on PGC2 data[36]; supplementary figure S3). Through using the LocusZoom, the users can get an overview about the association between the query gene and SZ (based on PGC2 data[36]). The most significant SNP is highlighted in the LocusZoom results (supplementary figure S3). Moreover, SZDB provides the differentially expressed genes in SZ cases and healthy controls.
(2) Analysis and Tools module. This is the most important and useful module. Results from our in-depth analyses were integrated by this module. In this module, users can perform customized analyses, including co-expression analysis, spatio-temporal expression pattern analysis, protein-protein interaction analysis, association analysis, and eQTL analysis. Detailed descriptions for each analysis are as follows: (a) Co-expression analysis. The co-expression analysis can provide exploratory data for candidate gene prioritization and suggestion for more work (eg, functional experiments) to validate if the candidate gene

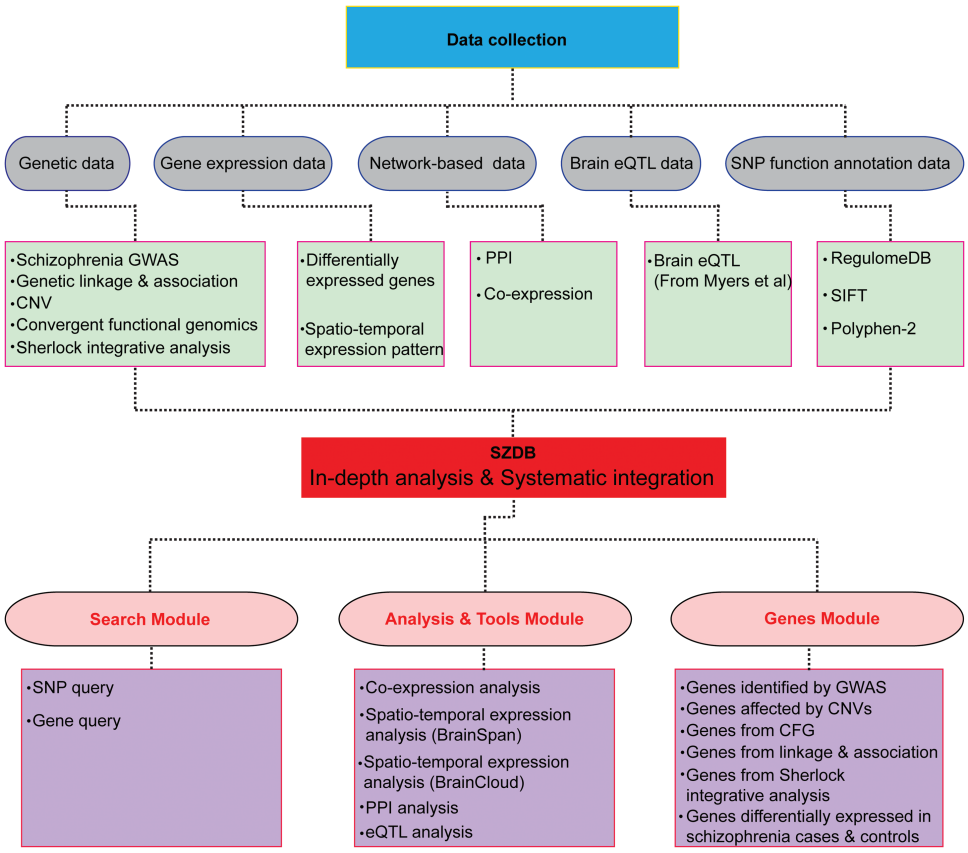

Fig. 1. Overview of SZDB database. Multi-type data from various sources were collected and curated. On the basis of comprehensive data collection and curation, in depth-analyses and systematic integration were performed. SZDB contains 3 core modules, which can provide comprehensive description information and perform customized analysis.

contributes to SZ pathogenesis. Users can explore if a specific gene (or multiple genes) is significantly co-expressed with other genes in a specific brain region at a specific developmental stage (supplementary figure S4). For each input gene, SZDB will return the genes that have the most significant co-expression correlation coefficient (default Pearson correlation coefficient ≥ 0.8) with the input gene. (b) Spatio-temporal expression pattern analysis. Two well-characterized expression datasets[74,85] were used to construct the spatio-temporal expression pattern of SZ risk genes. Multiple genes are allowed as input for a single query. For each query, SZDB generates the spatio-temporal plot of the query genes. The average expression level of all genes is also calculated and displayed. As the expression data from the BrainCloud[74] contains only one brain tissue (prefrontal cortex), the default mode will calculate the temporal expression patterns of the SZ risk genes in prefrontal cortex. For the Brainspan,[71] users can select different brain regions. Users can conduct customized analysis to examine the sparo-temporal expression pattern of the query genes (supplementary figure S5). (c) PPI analysis. For a specific input protein (or proteins), SZDB extracts all of the proteins that interacted with the query protein (supplementary figure S6). Multiple proteins are allowed as input for a single query. For visualization, SZDB draws the PPI plot based on the interactions among the input proteins. (d) Regional association plot and linkage disequilibrium analysis. To visualize the association results, SZDB generates regional association plot (supplementary figure S3). Users can extract the associations (*P* values) between genetic variants located in a specific chromosomal region and SZ using the PGC2 GWAS data.[36] Three types of query, including SNP, gene, and region are allowed. Users first input the interested gene (SNP or region), then select the flanking region (from several to hundreds kb). SZDB generates the regional association plot and highlights the most significant SNP in the queried region. (e) eQTL analysis. We integrated brain eQTL data from Myers et al[69] into SZDB and users can perform SNP or gene query. For SNP query, SZDB will carry out the eQTL analysis to identify the gene (or genes) whose expression is associated with the query SNP. For gene query, SZDB will extract the SNP (or SNPs) that is associated with the expression of the query gene (supplementary figure S7).

(3) Genes module. Multiple SZ candidate genes have been identified by different methods, including genetic

linkage and association studies, differential gene expression studies, convergent functional genomics, *Sherlock* integrative analysis and so on. Currently, SZ candidate genes from 9 resources were collated and integrated into SZDB: (a) Genes frequently affected by CNVs in SZ cases; (b) Genes identified by SZ GWAS; (c) Genes identified by genetic linkage and association studies; (d) Genes identified by convergent functional genomics (CFG); (e) Genes identified by Sherlock integrative analysis; and (f) Genes differentially expressed in SZ cases and controls; (g) Genes identified by the gene-based test (*Pascal*)[65]; (h) Genes identified by whole exome sequencing; and (i) Genes differentially methylated in SZ cases compared with controls. Users can perform customized queries to retrieve these genes.

Prioritized SZ Genes, Enriched Pathways and Drug Targets

In addition to providing a user-friendly online resources and powerful search and analysis tools for users, we also ranked SZ risk genes through using a cumulative scoring strategy (see methods). Polyevidence scoring identified 29 top prioritized SZ candidate genes that have a final score of 3 and above (figure 2 and supplementary table S13). Among these top prioritized genes, 3 (*DRD2*, *PGBD1* and *DOC2A*) have the highest final score (4 points, ie, 4 lines of evidence support these genes may be associated with SZ), suggesting that these 3 genes may represent promising SZ risk genes. We performed the pathway analysis[86,87] using genes that have a polyevidence score of 2 and above and identified significant enriched pathways, including synaptic transmission, dopamine metabolic process and ion channel activity (table 1).

To explore if the top prioritized genes in SZDB are drug targets, we examined the Drug-Gene interaction database[83] and the druggable gene list of Rask-Andersen et al.[83] We found that 7 of the top 29 genes prioritized in SZDB database are drug targets, including *DRD2*, *BDNF*, *COMT*, *CNR1*, *GRIA1*, *SLC1A2*, and *SRR* (supplementary table S14). Of note, *DRD2* is the most intensively studied gene and is the target of most of drugs.

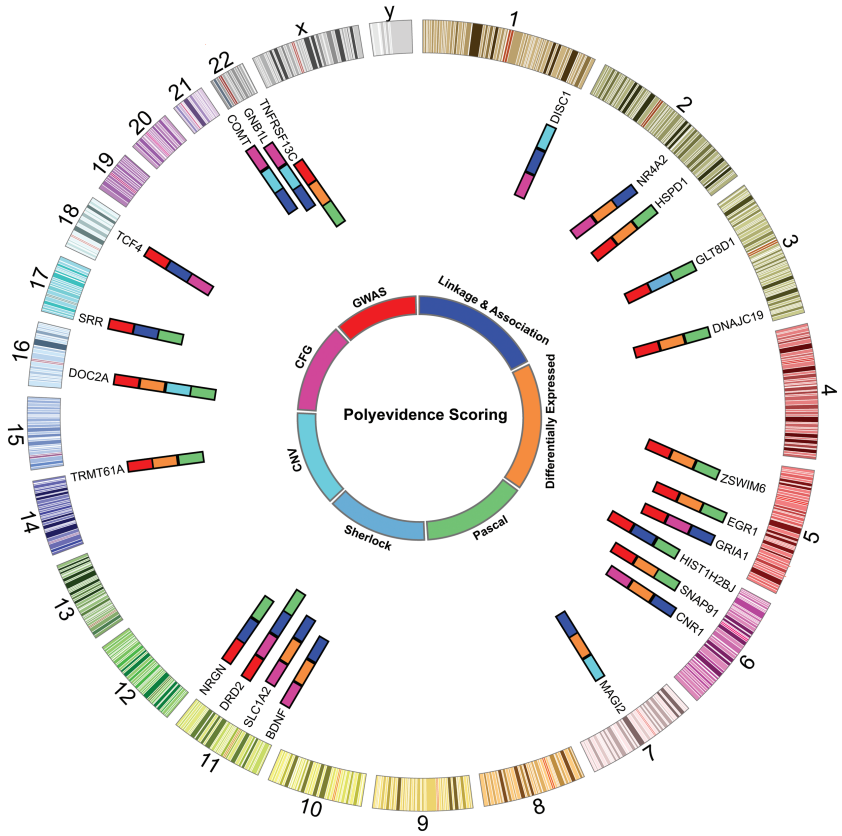

Fig. 2. Polyevidence scoring strategy and distribution of top prioritized schizophrenia risk genes on human genome. Schizophrenia risk genes were collected from 7 different sources, including genetic study (genome-wide association studies [GWAS], linkage and association, copy number variations [CNVs]), gene-based analysis, gene expression study, convergent functional genomics, *sherlock* integrative analysis, and differentially expressed genes in schizophrenia cases and controls. Evidence from these sources was integrated and genes were scored based on their presence in these sources.

Genes Implicated in SZ are Highly Co-expressed in Human Brain

On the basis of whole-genome transcriptomic data from the BrainSpan (http://www.brainspan.org),[71] we constructed gene co-expression networks by applying signed weighted gene co-expression network analysis (supplementary methods). In total, we identified 23 co-expression modules (figure 3). To further explore if SZ candidate genes are enriched in specific co-expression modules, we used 2 datasets of SZ candidate genes. The first dataset included the genes identified by recent GWAS of SZ (supplementary table S7). The second dataset included genes from our prioritized SZ candidate genes (ie, genes with a ployevidence score of 2 and above; supplementary table S13). We found that SZ candidate genes from GWAS are significantly enriched in co-expression module 2 (M2) (P = .0083) and M5 (P = .0079; figure 3). Intriguingly, genes from our polyevidence scoring were also significantly enriched in M2 (P = .018) and M5 (P = .019). These results suggested that SZ candidate genes are highly co-expressed in human brain.

Proteins Encoded by Prioritized SZ Risk Genes are Significantly Interacted

We performed PPI analysis (detailed analysis procedure can be found in our previous study[78]) using the prioritized SZ genes from SZDB (the genes that have a final score of 2 and above; supplementary table S13). We found that proteins encoded by prioritized SZ risk genes are significantly interacted (P < .05) and formed a densely

Table 1. Gene Ontology (GO) Analysis of Genes That Have a Polyevidence Score of 2 and Above

Category	Term	Genes	P_{adj}
GOTERM_BP_FAT	Synaptic transmission	*HTR2A, BDNF, CHRNA3, CHRNB4, DRD1, DRD2, DRD4, EGR1, GRIA4, GRIN2B, GRM3, NISCH, SLC1A3, ERBB4*	5.60×10^{-5}
GOTERM_BP_FAT	Dopamine metabolic process	*COMT, DRD1, DRD2, DRD4, MAOA, NR4A2, SLC6A3*	1.20×10^{-4}
GOTERM_BP_FAT	Ion channel activity	*CHRNA7, CACNA1I, CACNB2,CHRNA3, CHRNA5, CHRNB4, GABRB3, GRIA4, GRIA1, GRIN2B, HCN1, CACNA1C, KCTD13, KCNV1, KCNN3, KCNJ13, KCNB1*	1.90×10^{-2}

Note: The table shows GO terms identified by DAVID.[86,87] that are enriched among 302 genes that have a polyevidence score of 2 and above. P_{adj} values represent P-values corrected by the Benjamini-Hochberg procedure in DAVID.

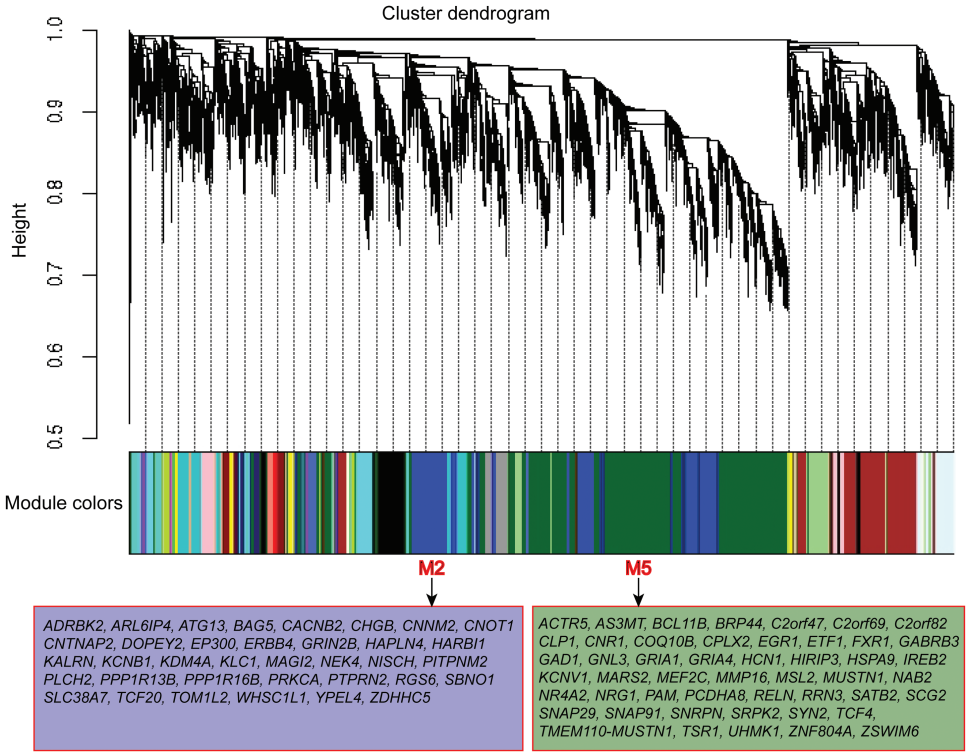

Fig. 3. Co-expression module analysis and enrichment of schizophrenia risk genes in co-expression module 2 and module 5. Schizophrenia candidate genes in module 2 (M2) and module 5 (M5) are shown in box.

interconnected PPI network (figure 4). This result further supported our previous findings[78] and suggested that perturbations of common underlying molecular processes or pathways modulate risk to SZ.

Discussion

In recent years, significant progresses have been made to dissect the genetic architecture of SZ. Multiple highly-significant risk variants have been successfully identified by GWAS.[21–36] Nevertheless, pinpointing the causal risk variant(s) remains a major challenge. In addition to risk genes identified by GWAS, other promising SZ candidate genes have also been revealed by other approaches, such as genetic linkage and association,[18] gene expression study, convergent functional genomics,[42] *Sherlock* integrative analysis.[52] With the increase of sample size and improvement of research approaches, novel SZ risk variants and genes are continuously being identified. Though a large number of risk variants and loci have been identified, the pathophysiology of SZ remains largely unknown. Evidently, it is the time to systematically integrate multi-type data from various studies and to harvest meaningful biological information from genetic findings of SZ. Therefore, we developed the SZDB, a comprehensive database that integrates multi-type resources from different layers of SZ researches, to fill the lacuna.

Compared with other SZ databases such as the SZGene[20] and SZGR,[55] the SZDB has several advantages. First, the SZ risk genes included in the SZGene and SZGR were mainly based on genetic linkage and association studies,

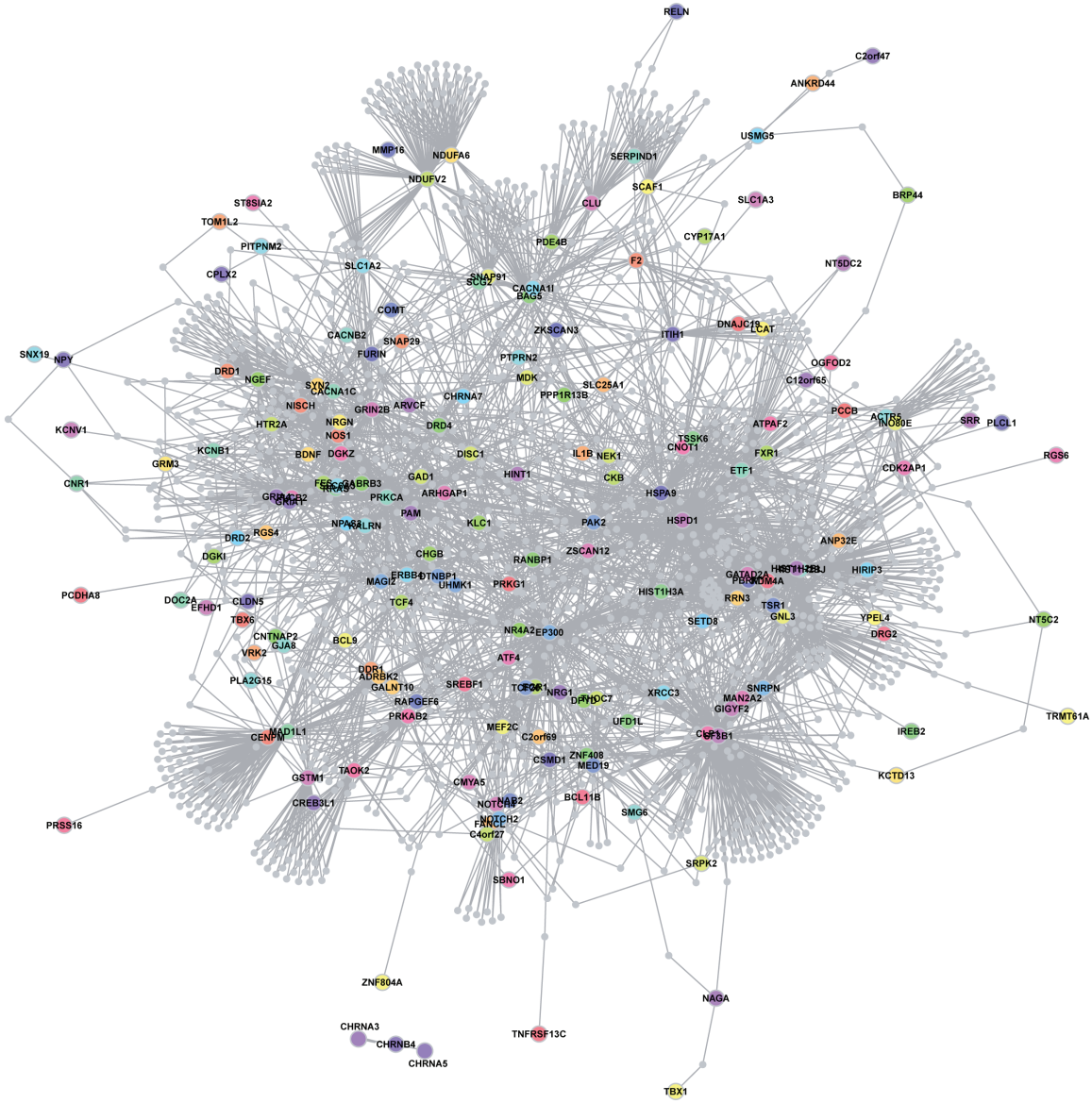

Fig. 4. Proteins encoded by prioritized schizophrenia risk genes form a densely interconnected protein–protein interaction (PPI) network. By using the high-confidence PPI data from InWeb, proteins encoded by prioritized schizophrenia risk genes (that have a final score of 2 and above) were used to construct the PPI network. Proteins encoded by prioritized schizophrenia risk genes are shown.

and these databases were not updated since the establishment in 2008 and 2010, respectively. Moreover, we found that the SZGR could not be even accessible during the reviewing of this manuscript. We noticed that the SZGR was recently moved to a new website (https://bioinfo.uth.edu/SZGR/). We did not check the content of the new website of the SZGR, but genes identified by GWAS and other integrative analyses such as convergent functional genomics[42] and *Sherlock*[52] integrative analysis were not included in the SZGene and the old SZGR databases. Second, the SZDB has a powerful analysis module and users can perform customized analysis, including eQTL, PPI, co-expression, spatio-temporal expression pattern analyses, to name a few. Third, on the basis of multi-type data from various resources, in-depth data integration was conducted and top SZ risk genes were prioritized in the SZDB. Interestingly, we noted that *DRD2* was among the top prioritized genes and had the highest final scores. Dysregulation of the dopamine system function has been well characterized in SZ and almost all available antipsychotic drugs exert their main therapeutic effects though the blockade of DRD2.[88–91] The top prioritized genes thus represent promising SZ risk genes which might be worthy of further functional characterization. Fourth, the SZDB will be updated periodically to keep up with the research progress. With the use of new technologies and analysis methods, novel SZ risk variants and genes are continuously being identified. These latest findings will be integrated into SZDB periodically.

This study has several limitations. First, we used a simple and arbitrary scoring algorithm to prioritize the promising SZ candidate genes from diverse sources (eg, genetic and gene expression studies). Though this scoring system is simple and operational, it ignores the potential overlapping between the sources used for scoring. For example, genes identified by convergent functional genomics (CFG)[42] also integrated information from genetic and gene expression studies. Second, we treated all evidence from different sources equally. Nevertheless, genes identified by GWAS represent the high-confidence SZ candidate genes, therefore, these genes should be given higher weight than genes identified by linkage or association study. These limitations can be addressed through developing an appropriate scoring algorithm in the future.

In conclusion, the SZDB aims to provide a comprehensive resource for SZ genetic research. To the best of our knowledge, SZDB is the most comprehensive SZ genetic database that integrates multiple layers of data from various sources so far. With more genes being identified and more data resources being available, SZDB will be updated every 6 months to maintain the up-to-date resources. The SZDB will integrate more types of data in future, including large-scale data from whole-genome sequencing, tissue-specific PPI interactions, gene expression studies from RNA sequencing. Evidence from epigenetics and animal models will also be integrated into SZDB in future. The ultimate goal of SZDB is to provide a one-stop service for SZ researchers, and they can find most of the information they wanted in the SZDB. With the increase of new genes and integration of more data type, the SZDB will become more powerful. We hope the SZDB will facilitate the translation of genetic findings into molecular risk mechanisms for SZ, and ultimately to the improvement of disease diagnosis and treatment.

Supplementary Material

Supplementary material is available at http://schizophreniabulletin.oxfordjournals.org.

Funding

Strategic Priority Research Program (B) of the Chinese Academy of Sciences (XDB02020003); 1000 Young Talents Program.

Acknowledgments

We thank Dr Long Fan for technical assistance and the 4 reviewers for their critical comments on the early version of the manuscript. The authors declared no potential conflicts of interest.

References

1. Saha S, Chant D, Welham J, McGrath J. A systematic review of the prevalence of schizophrenia. *PLoS Med*. 2005;2:e141.
2. Tandon R, Keshavan MS, Nasrallah HA. Schizophrenia, "Just the Facts": what we know in 2008 part 1: overview. *Schizophr Res*. 2008;100:4–19.
3. World Health Organization. *The Global Burden of Disease: 2004 Update*. Geneva, Switzerland: WHO Press; 2008.
4. Knapp M, Mangalore R, Simon J. The global costs of schizophrenia. *Schizophr Bull*. 2004;30:279–293.
5. Wu EQ, Birnbaum HG, Shi L, et al. The economic burden of schizophrenia in the United States in 2002. *J Clin Psychiatry*. 2005;66:1122–1129.
6. Goeree R, Farahati F, Burke N, et al. The economic burden of schizophrenia in Canada in 2004. *Curr Med Res Opin*. 2005;21:2017–2028.
7. Lieberman JA, Stroup TS, McEvoy JP, et al. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *N Engl J Med*. 2005;353:1209–1223.
8. Lieberman JA. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia: efficacy, safety and cost outcomes of CATIE and other trials. *J Clin Psychiatry*. 2007;68:e04.
9. Walker E, Kestler L, Bollini A, Hochman KM. Schizophrenia: etiology and course. *Annu Rev Psychol*. 2004;55:401–430.
10. Buka SL, Tsuang MT, Lipsitt LP. Pregnancy/delivery complications and psychiatric diagnosis. A prospective study. *Arch Gen Psychiatry*. 1993;50:151–156.
11. Dalman C, Allebeck P, Cullberg J, Grunewald C, Köster M. Obstetric complications and the risk of schizophrenia: a longitudinal study of a national birth cohort. *Arch Gen Psychiatry*. 1999;56:234–240.
12. Barr CE, Mednick SA, Munk-Jorgensen P. Exposure to influenza epidemics during gestation and adult schizophrenia. A 40-year study. *Arch Gen Psychiatry*. 1990;47:869–874.

13. Brown AS, Cohen P, Harkavy-Friedman J, et al. A.E. Bennett Research Award. Prenatal rubella, premorbid abnormalities, and adult schizophrenia. *Biol Psychiatry*. 2001;49:473–486.
14. van Os J, Selten JP. Prenatal exposure to maternal stress and subsequent schizophrenia. The May 1940 invasion of The Netherlands. *Br J Psychiatry*. 1998;172:324–326.
15. Quenstedt M, Parshall A. Prenatal exposure to maternal stress and subsequent schizophrenia. *Br J Psychiatry*. 1998;173:183.
16. Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Arch Gen Psychiatry*. 2003;60:1187–1192.
17. Sullivan PF, Daly MJ, O'Donovan M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat Rev Genet*. 2012;13:537–551.
18. Lewis CM, Levinson DF, Wise LH, et al. Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: Schizophrenia. *Am J Hum Genet*. 2003;73:34–48.
19. Ng MY, Levinson DF, Faraone SV, et al. Meta-analysis of 32 genome-wide linkage studies of schizophrenia. *Mol Psychiatry*. 2009;14:774–785.
20. Allen NC, Bagade S, McQueen MB, et al. Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database. *Nat Genet*. 2008;40:827–834.
21. O'Donovan MC, Craddock N, Norton N, et al. Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nat Genet*. 2008;40:1053–1055.
22. Purcell SM, Wray NR, Stone JL, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*. 2009;460:748–752.
23. Ripke S, O'Dushlaine C, Chambert K, et al. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat Genet*. 2013;45:1150–1159.
24. Ikeda M, Aleksic B, Kinoshita Y, et al. Genome-wide association study of schizophrenia in a Japanese population. *Biol Psychiatry*. 2011;69:472–478.
25. Hamshere ML, Walters JT, Smith R, et al. Genome-wide significant associations in schizophrenia to ITIH3/4, CACNA1C and SDCCAG8, and extensive replication of associations reported by the Schizophrenia PGC. *Mol Psychiatry*. 2013;18:708–712.
26. Rietschel M, Mattheisen M, Degenhardt F, et al. Association between genetic variation in a region on chromosome 11 and schizophrenia in large samples from Europe. *Mol Psychiatry*. 2012;17:906–917.
27. Irish Schizophrenia Genomics Consortium & Wellcome Trust Case Control Consortium. Genome-wide association study implicates HLA-C*01:02 as a risk factor at the major histocompatibility complex locus in schizophrenia. *Biol Psychiatry*. 2012;72:620–628.
28. Shi J, Levinson DF, Duan J, et al. Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature*. 2009;460:753–757.
29. Shi Y, Li Z, Xu Q, et al. Common variants on 8p12 and 1q24.2 confer risk of schizophrenia. *Nat Genet*. 2011;43:1224–1227.
30. Stefansson H, Ophoff RA, Steinberg S, et al. Common variants conferring risk of schizophrenia. *Nature*. 2009;460:744–747.
31. Li J, Zhou G, Ji W, et al. Common variants in the BCL9 gene conferring risk of schizophrenia. *Arch Gen Psychiatry*. 2011;68:232–240.
32. Steinberg S, de Jong S, Andreassen OA, et al. Common variants at VRK2 and TCF4 conferring risk of schizophrenia. *Hum Mol Genet*. 2011;20:4076–4081.
33. Wong EH, So HC, Li M, et al. Common variants on Xq28 conferring risk of schizophrenia in Han Chinese. *Schizophr Bull*. 2014;40:777–786.
34. Yue WH, Wang HF, Sun LD, et al. Genome-wide association study identifies a susceptibility locus for schizophrenia in Han Chinese at 11p11.2. *Nat Genet*. 2011;43:1228–1231.
35. Lencz T, Guha S, Liu C, et al. Genome-wide association study implicates NDST3 in schizophrenia and bipolar disorder. *Nat Commun*. 2013;4:2739.
36. Schizophrenia Working Group of the Psychiatric Genomics Consortium*. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 2014;511:421–427.
37. Stefansson H, Rujescu D, Cichon S, et al.; GROUP. Large recurrent microdeletions associated with schizophrenia. *Nature*. 2008;455:232–236.
38. The International Schizophrenia Consortium. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature*. 2008;455:237–241.
39. Xu B, Roos JL, Levy S, et al. Strong association of de novo copy number mutations with sporadic schizophrenia. *Nat Genet*. 2008;40:880–885.
40. Walsh T, McClellan JM, McCarthy SE, et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science*. 2008;320:539–543.
41. Malhotra D, McCarthy S, Michaelson JJ, et al. High frequencies of de novo CNVs in bipolar disorder and schizophrenia. *Neuron*. 2011;72:951–963.
42. Ayalew M, Le-Niculescu H, Levey DF, et al. Convergent functional genomics of schizophrenia: from comprehensive understanding to genetic risk prediction. *Mol Psychiatry*. 2012;17:887–905.
43. Horváth S, Mirnics K. Schizophrenia as a disorder of molecular pathways. *Biol Psychiatry*. 2015;77:22–28.
44. Hall J, Trent S, Thomas KL, O'Donovan MC, Owen MJ. Genetic risk for schizophrenia: convergence on synaptic pathways involved in plasticity. *Biol Psychiatry*. 2015;77:52–58.
45. Pruim RJ, Welch RP, Sanna S, et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics*. 2010;26:2336–2337.
46. Boyle AP, Hong EL, Hariharan M, et al. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res*. 2012;22:1790–1797.
47. The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012;489:57–74.
48. Xie D, Boyle AP, Wu L, et al. Dynamic trans-acting factor colocalization in human cells. *Cell*. 2013;155:713–724.
49. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res*. 2003;31:3812–3814.
50. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc*. 2009;4:1073–1081.
51. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7:248–249.
52. Luo XJ, Mattheisen M, Li M, et al. Systematic Integration of Brain eQTL and GWAS Identifies ZNF323 as a Novel Schizophrenia Risk Gene and Suggests Recent Positive Selection Based on Compensatory Advantage on Pulmonary Function. *Schizophr Bull*. 2015;41:1294–1308.
53. Purcell SM, Moran JL, Fromer M, et al. A polygenic burden of rare disruptive mutations in schizophrenia. *Nature*. 2014;506:185–190.

54. Fromer M, Pocklington AJ, Kavanagh DH, et al. De novo mutations in schizophrenia implicate synaptic networks. *Nature*. 2014;506:179–184.
55. Jia P, Sun J, Guo AY, Zhao Z. SZGR: a comprehensive schizophrenia gene resource. *Mol Psychiatry*. 2010;15:453–462.
56. Sun J, Kuo PH, Riley BP, Kendler KS, Zhao Z. Candidate genes for schizophrenia: a survey of association studies and gene ranking. *Am J Med Genet B Neuropsychiatr Genet*. 2008;147B:1173–1181.
57. Luo X, Huang L, Han L, et al. Systematic prioritization and integrative analysis of copy number variations in schizophrenia reveal key schizophrenia susceptibility genes. *Schizophr Bull*. 2014;40:1285–1299.
58. He X, Fuller CK, Song Y, et al. Sherlock: detecting gene-disease associations by matching patterns of expression QTL and GWAS. *Am J Hum Genet*. 2013;92:667–680.
59. Lanz TA, Joshi JJ, Reinhart V, et al. STEP levels are unchanged in pre-frontal cortex and associative striatum in post-mortem human brain samples from subjects with schizophrenia, bipolar disorder and major depressive disorder. *PLoS One*. 2015;10:e0121744.
60. Iwamoto K, Bundo M, Kato T. Altered expression of mitochondria-related genes in postmortem brains of patients with bipolar disorder or schizophrenia, as revealed by large-scale DNA microarray analysis. *Hum Mol Genet*. 2005;14:241–253.
61. Narayan S, Tang B, Head SR, et al. Molecular profiles of schizophrenia in the CNS at different stages of illness. *Brain Res*. 2008;1239:235–248.
62. Chen C, Cheng L, Grennan K, et al. Two gene co-expression modules differentiate psychotics and controls. *Mol Psychiatry*. 2013;18:1308–1314.
63. de Baumont A, Maschietto M, Lima L, et al. Innate immune response is differentially dysregulated between bipolar disease and schizophrenia. *Schizophr Res*. 2015;161:215–221.
64. Fillman SG, Cloonan N, Catts VS, et al. Increased inflammatory markers identified in the dorsolateral prefrontal cortex of individuals with schizophrenia. *Mol Psychiatry*. 2013;18:206–214.
65. Lamparter D, Marbach D, Rueedi R, Kutalik Z, Bergmann S. Fast and rigorous computation of gene and pathway scores from snp-based summary statistics. *PLoS Comput Biol*. 2016;12:e1004714.
66. Aberg KA, McClay JL, Nerella S, et al. Methylome-wide association study of schizophrenia: identifying blood biomarker signatures of environmental insults. *JAMA Psychiatry*. 2014;71:255–264.
67. Montano C, Taub MA, Jaffe A, et al. Association of DNA methylation differences with schizophrenia in an epigenome-wide association study. *JAMA Psychiatry*. 2016;73:506–514.
68. Hannon E, Spiers H, Viana J, et al. Methylation QTLs in the developing brain and their enrichment in schizophrenia risk loci. *Nat Neurosci*. 2016;19:48–54.
69. Myers AJ, Gibbs JR, Webster JA, et al. A survey of genetic human cortical gene expression. *Nat Genet*. 2007;39:1494–1499.
70. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559–575.
71. BrainSpan. BrainSpan: Atlas of the Developing Human Brain [Internet]. 2013. http://www.brainspan.org. Accessed April 28, 2015.
72. Willsey AJ, Sanders SJ, Li M, et al. Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell*. 2013;155:997–1007.
73. Gulsuner S, Walsh T, Watts AC, et al. Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. *Cell*. 2013;154:518–529.
74. Colantuoni C, Lipska BK, Ye T, et al. Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature*. 2011;478:519–523.
75. Gilman SR, Chang J, Xu B, et al. Diverse types of genetic variation converge on functional gene networks involved in schizophrenia. *Nat Neurosci*. 2012;15:1723–1728.
76. Oti M, Brunner HG. The modular nature of genetic diseases. *Clin Genet*. 2007;71:1–11.
77. Jia P, Zhao Z. Network.assisted analysis to prioritize GWAS results: principles, methods and perspectives. *Hum Genet*. 2014;133:125–138.
78. Luo X, Huang L, Jia P, et al. Protein-protein interaction and pathway analyses of top schizophrenia genes reveal schizophrenia susceptibility genes converge on common molecular networks and enrichment of nucleosome (chromatin) assembly genes in schizophrenia susceptibility loci. *Schizophr Bull*. 2014;40:39–49.
79. Luo XJ, Huang L, Li M, Gan L. Protein-protein interaction analysis reveals common molecular processes/pathways that contribute to risk of schizophrenia. *Schizophr Res*. 2013;143:390–392.
80. Lage K, Karlberg EO, Størling ZM, et al. A human phenome-interactome network of protein complexes implicated in genetic disorders. *Nat Biotechnol*. 2007;25:309–316.
81. Lage K, Hansen NT, Karlberg EO, et al. A large-scale analysis of tissue-specific pathology and gene expression of human disease genes and complexes. *Proc Natl Acad Sci U S A*. 2008;105:20870–20875.
82. Parikshak NN, Luo R, Zhang A, et al. Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell*. 2013;155:1008–1021.
83. Rask-Andersen M, Masuram S, Schiöth HB. The druggable genome: Evaluation of drug targets in clinical trials suggests major shifts in molecular class and indication. *Annu Rev Pharmacol Toxicol*. 2014;54:9–26.
84. Law V, Knox C, Djoumbou Y, et al. DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Res*. 2014;42:D1091–D1097.
85. Kang HJ, Kawasawa YI, Cheng F, et al. Spatio-temporal transcriptome of the human brain. *Nature*. 2011;478:483–489.
86. Huang DW, Sherman BT, Tan Q, et al. DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res*. 2007;35:W169–W175.
87. Huang DW, Sherman BT, Tan Q, et al. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol*. 2007;8:R183.
88. Masri B, Salahpour A, Didriksen M, et al. Antagonism of dopamine D2 receptor/beta-arrestin 2 interaction is a common property of clinically effective antipsychotics. *Proc Natl Acad Sci U S A*. 2008;105:13656–13661.
89. Frégeau MO, Carrier M, Guillemette G. Mechanism of dopamine D2 receptor-induced Ca(2+) release in PC-12 cells. *Cell Signal*. 2013;25:2871–2877.
90. Miyamoto S, Miyake N, Jarskog LF, Fleischhacker WW, Lieberman JA. Pharmacological treatment of schizophrenia: a critical review of the pharmacology and clinical effects of current and future therapeutic agents. *Mol Psychiatry*. 2012;17:1206–1227.
91. Seeman P. Targeting the dopamine D2 receptor in schizophrenia. *Expert Opin Ther Targets*. 2006;10:515–531.

Online Supplementary Materials

SZDB: A Database for Schizophrenia Genetic Research

Yong Wu[1,2], Yong-Gang Yao[1,2,3,4,*], and Xiong-Jian Luo[1,2,4,*]

[1]Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences and Yunnan Province, Kunming Institute of Zoology, Kunming, Yunnan, China

[2]Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, Yunnan 650204, China

[3]CAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai 200031, China

[4]YGY and XJL are co-corresponding authors who jointly directed this work

*To whom correspondence should be addressed: Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China. E-mail: luoxiongjian@mail.kiz.ac.cn (X-J L) or yaoyg@mail.kiz.ac.cn (Y-G Y).

Development of SZDB

The SZDB database, which runs on a dual-processor server with Ubuntu/linux operating system, is implemented under LAMP (Linux-Apache-MySQL-PHP) software stack. Data are stored in MySQL and administrated with the help of phpMyAdmin. Web interfaces are shaped by Javascript and Cascading Style Sheet. AJAX and some jQuery plugins (like Echarts and plotly) were used for the interface development. The data analysis programs were written by PHP. GBrowse uses MySQL as backend and was built following the configuration files provided by its developer (http://gmod.org/wiki/GBrowse_Configuration_HOWTO). The genomic data in SZDB are loaded into GBrowse after being converted into genome feature format (GFF). We got the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway annotation information of these genes from DAVID (https://david.ncifcrf.gov).[1, 2]

Co-expression module analysis

The BrainSpan developmental RNA-seq data[3] summarized to Gencode 10 gene-level reads per kilobase million mapped reads (RPKM) values were used to identify co-expressed genes. First, genes with no gene id and duplicated gene symbols were removed and this led to 12,897 unique genes. Second, as described in the study of Willsey et al.,[4] we used 16 brain tissues- Primary motor cortex (M1C), Primary somatosensory cortex (S1C), Ventrolateral prefrontal cortex (VFC), Anterior cingulate cortex (MFC), Dorsolateral prefrontal cortex (DFC), Orbital frontal cortex (OFC), Primary visual cortex (V1C), Inferolateral temporal cortex (ITC), Posteroventral parietal cortex (IPC), Primary auditory cortex (A1C), Posterior superior temporal cortex (STC), Striatum (STR), Hippocampus (HIP), Amygdaloid complex (AMY), Mediodorsal nucleus of thalamus (MD), and Cerebellar cortex (CBC). Totally, 357 samples were included to do subsequent analysis. Then, R package, weighted correlation network analysis (WGCNA),[5] was used to identify co-expressed modules. Expression values were log-transformed (log2[RPKM+1]). Modules were defined as branches of a hierarchical cluster tree using the top-down dynamic tree cut method.[6] A weighted signed network was computed based on a fit to scale-free topology, and a thresholding power of 7 was chosen (as it was the smallest threshold that resulted in a scale-free R^2 fit of 0.8). In total, 23 co-expression modules (labeled numerically from M1 to M23, also labeled by color like black, dark green and so on) which contain genes from 24 to 4011 were identified. To explore if schizophrenia risk genes are enriched in specific modules, we conducted enrichment analyses using the GWAS hit genes and SZDB prioritized genes (with a final score of 2 and above), respectively.

Table S1. 160 schizophrenia susceptibility genes from the SZGR database (by Lewis et al.[7])

Gene ID	Symbol	Official Full Name	Location
260425	MAGI3	membrane associated guanylate kinase, WW and PDZ domain containing 3	1p12-p11.2
2944	GSTM1	glutathione S-transferase mu 1	1p13.3
22854	NTNG1	netrin G1	1p13.3
2730	GCLM	glutamate-cysteine ligase, modifier subunit	1p22.1
2899	GRIK3	glutamate receptor, ionotropic, kainate 3	1p34-p33
4524	MTHFR	5,10-methylenetetrahydrofolate reductase (NADPH)	1p36.3
2703	GJA8	gap junction protein, alpha 8, 50kDa	1q21.1
23208	SYT11	synaptotagmin XI	1q21.2
3782	KCNN3	potassium intermediate/small conductance calcium-activated channel, subfamily N, member 3	1q21.3
5999	RGS4	regulator of G-protein signaling 4	1q23.3
5321	PLA2G4A	phospholipase A2, group IVA (cytosolic, calcium-dependent)	1q25
3586	IL10	interleukin 10	1q31-q32
1116	CHI3L1	chitinase 3-like 1 (cartilage glycoprotein-39)	1q32.1
5362	PLXNA2	plexin A2	1q32.2
7257	TSNAX	translin-associated factor X	1q42.1
27185	DISC1	disrupted in schizophrenia 1	1q42.1
5305	PIP4K2A	phosphatidylinositol-5-phosphate 4-kinase, type II, alpha	10p12.2
1103	CHAT	choline acetyltransferase	10q11.2
2894	GRID1	glutamate receptor, ionotropic, delta 1	10q22
355	FAS	Fas (TNF receptor superfamily, member 6)	10q24.1
627	BDNF	brain-derived neurotrophic factor	11p13
6506	SLC1A2	solute carrier family 1 (glial high affinity glutamate transporter), member 2	11p13-p12
7166	TPH1	tryptophan hydroxylase 1	11p15.3-p14
1815	DRD4	dopamine receptor D4	11p15.5
7054	TH	tyrosine hydroxylase	11p15.5
2893	GRIA4	glutamate receptor, ionotrophic, AMPA	11q22

Gene ID	Symbol	Official Full Name	Location
		4	
3606	IL18	interleukin 18 (interferon-gamma-inducing factor)	11q22.2-q22.3
2900	GRIK4	glutamate receptor, ionotropic, kainate 4	11q22.3
486	FXYD2	FXYD domain containing ion transport regulator 2	11q23
1813	DRD2	dopamine receptor D2	11q23
3359	HTR3A	5-hydroxytryptamine (serotonin) receptor 3A	11q23.1
53826	FXYD6	FXYD domain containing ion transport regulator 6	11q23.3
4900	NRGN	neurogranin (protein kinase C substrate, RC3)	11q24
2904	GRIN2B	glutamate receptor, ionotropic, N-methyl D-aspartate 2B	12p12
4908	NTF3	neurotrophin 3	12p13
2065	ERBB3	v-erb-b2 erythroblastic leukemia viral oncogene homolog 3 (avian)	12q13
1610	DAO	D-amino-acid oxidase	12q24
4842	NOS1	nitric oxide synthase 1 (neuronal)	12q24.2-q24.31
3356	HTR2A	5-hydroxytryptamine (serotonin) receptor 2A	13q14-q21
3843	IPO5	importin 5	13q32.2
267012	DAOA	D-amino acid oxidase activator	13q33.2|13q34
64067	NPAS3	neuronal PAS domain protein 3	14q12-q13
1113	CHGA	chromogranin A (parathyroid secretory protein 1)	14q32
207	AKT1	v-akt murine thymoma viral oncogene homolog 1	14q32.32|14q32.32
89832	CHRFAM7A	CHRNA7 (cholinergic receptor, nicotinic, alpha 7, exons 5-10) and FAM7A (family with sequence similarity 7A, exons A-E) fusion	15q13.1
1139	CHRNA7	cholinergic receptor, nicotinic, alpha 7	15q14
1544	CYP1A2	cytochrome P450, family 1, subfamily A, polypeptide 2	15q24.1
8128	ST8SIA2	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 2	15q26

Gene ID	Symbol	Official Full Name	Location
3240	HP	haptoglobin	16q22.1
63826	SRR	serine racemase	17p13
7157	TP53	tumor protein p53	17p13.1
6532	SLC6A4	solute carrier family 6 (neurotransmitter transporter, serotonin), member 4	17q11.1-q12
753	C18orf1	chromosome 18 open reading frame 1	18p11.2
5289	PIK3C3	phosphoinositide-3-kinase, class 3	18q12.3
718	C3	complement component 3	19p13.3-p13.2
9253	NUMBL	numb homolog (Drosophila)-like	19q13.13-q13.2
348	APOE	apolipoprotein E	19q13.2
57142	RTN4	reticulon 4	2p16.3
8807	IL18RAP	interleukin 18 receptor accessory protein	2q12
8809	IL18R1	interleukin 18 receptor 1	2q12
8685	MARCO	macrophage receptor with collagenous structure	2q12-q13
4867	NPHP1	nephronophthisis 1 (juvenile)	2q13
3552	IL1A	interleukin 1, alpha	2q14
3553	IL1B	interleukin 1, beta	2q14
3557	IL1RN	interleukin 1 receptor antagonist	2q14.2
4929	NR4A2	nuclear receptor subfamily 4, group A, member 2	2q22-q23
2571	GAD1	glutamate decarboxylase 1 (brain, 67kDa)	2q31
2066	ERBB4	v-erb-a erythroblastic leukemia viral oncogene homolog 4 (avian)	2q33.3-q34
1114	CHGB	chromogranin B (secretogranin 1)	20pter-p12
2937	GSS	glutathione synthetase	20q11.2
10215	OLIG2	oligodendrocyte lineage transcription factor 2	21q22.11
27037	TRMT2A	TRM2 tRNA methyltransferase 2 homolog A (S. cerevisiae)	22q11.1-22q13|22q11.21
54584	GNB1L	guanine nucleotide binding protein (G protein), beta polypeptide 1-like	22q11.2
421	ARVCF	armadillo repeat gene deletes in velocardiofacial syndrome	22q11.21
5625	PRODH	proline dehydrogenase (oxidase) 1	22q11.21
5902	RANBP1	RAN binding protein 1	22q11.21

Gene ID	Symbol	Official Full Name	Location
6899	TBX1	T-box 1	22q11.21
7122	CLDN5	claudin 5	22q11.21
7353	UFD1L	ubiquitin fusion degradation 1 like (yeast)	22q11.21
9342	SNAP29	synaptosomal-associated protein, 29kDa	22q11.21
9993	DGCR2	DiGeorge syndrome critical region gene 2	22q11.21
29801	ZDHHC8	zinc finger, DHHC-type containing 8	22q11.21
65078	RTN4R	reticulon 4 receptor	22q11.21
1312	COMT	catechol-O-methyltransferase	22q11.21-q11.23|22q11.21
8214	DGCR6	DiGeorge syndrome critical region gene 6	22q11.21|22q11
7494	XBP1	X-box binding protein 1	22q12.1|22q12
7533	YWHAH	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, eta polypeptide	22q12.3
1439	CSF2RB	colony stimulating factor 2 receptor, beta, low-affinity (granulocyte-macrophage)	22q13.1
6663	SOX10	SRY (sex determining region Y)-box 10	22q13.1
8398	PLA2G6	phospholipase A2, group VI (cytosolic, calcium-independent)	22q13.1
9145	SYNGR1	synaptogyrin 1	22q13.1
9463	PICK1	protein interacting with PRKCA 1	22q13.1
2847	MCHR1	melanin-concentrating hormone receptor 1	22q13.2
23209	MLC1	megalencephalic leukoencephalopathy with subcortical cysts 1	22q13.33
9223	MAGI1	membrane associated guanylate kinase, WW and PDZ domain containing 1	3p14.1
3269	HRH1	histamine receptor H1	3p25
6854	SYN2	synapsin II	3p25
10752	CHL1	cell adhesion molecule with homology to L1CAM (close homolog of L1)	3p26.1
2917	GRM7	glutamate receptor, metabotropic 7	3p26.1-p25.1
1814	DRD3	dopamine receptor D3	3q13.3
2932	GSK3B	glycogen synthase kinase 3 beta	3q13.3
886	CCKAR	cholecystokinin A receptor	4p15.1-p15.2

Gene ID	Symbol	Official Full Name	Location
1816	DRD5	dopamine receptor D5	4p16.1
2638	GC	group-specific component (vitamin D binding protein)	4q12-q13
10611	PDLIM5	PDZ and LIM domain 5	4q22
1950	EGF	epidermal growth factor (beta-urogastrone)	4q25
6531	SLC6A3	solute carrier family 6 (neurotransmitter transporter, dopamine), member 3	5p15.3
3350	HTR1A	5-hydroxytryptamine (serotonin) receptor 1A	5q11.2-q13
4762	NEUROG1	neurogenin 1	5q23-q31
9542	NRG2	neuregulin 2	5q23-q33
2246	FGF1	fibroblast growth factor 1 (acidic)	5q31
23305	ACSL6	acyl-CoA synthetase long-chain family member 6	5q31
3360	HTR4	5-hydroxytryptamine (serotonin) receptor 4	5q31-q33
3562	IL3	interleukin 3 (colony-stimulating factor, multiple)	5q31.1
3565	IL4	interleukin 4	5q31.1
51735	RAPGEF6	Rap guanine nucleotide exchange factor (GEF) 6	5q31.1
56990	CDC42SE2	CDC42 small effector 2	5q31.1
2566	GABRG2	gamma-aminobutyric acid (GABA) A receptor, gamma 2	5q31.1-q33.1
3593	IL12B	interleukin 12B (natural killer cell stimulatory factor 2, cytotoxic lymphocyte maturation factor 2, p40)	5q31.1-q33.1
3094	HINT1	histidine triad nucleotide binding protein 1	5q31.2
84466	MEGF10	multiple EGF-like-domains 10	5q33
2568	GABRP	gamma-aminobutyric acid (GABA) A receptor, pi	5q33-q34
9685	CLINT1	clathrin interactor 1	5q33.3
2890	GRIA1	glutamate receptor, ionotropic, AMPA 1	5q33|5q31.1
2559	GABRA6	gamma-aminobutyric acid (GABA) A receptor, alpha 6	5q34
2561	GABRB2	gamma-aminobutyric acid (GABA) A	5q34

Gene ID	Symbol	Official Full Name	Location
		receptor, beta 2	
2554	GABRA1	gamma-aminobutyric acid (GABA) A receptor, alpha 1	5q34-q35
629	CFB	complement factor B	6p21.3
721	C4B	complement component 4B (Childo blood group)	6p21.3
780	DDR1	discoidin domain receptor tyrosine kinase 1	6p21.3
4049	LTA	lymphotoxin alpha (TNF superfamily, member 1)	6p21.3
4277	MICB	MHC class I polypeptide-related sequence B	6p21.3
4855	NOTCH4	Notch homolog 4 (Drosophila)	6p21.3
7124	TNF	tumor necrosis factor (TNF superfamily, member 2)	6p21.3
2550	GABBR1	gamma-aminobutyric acid (GABA) B receptor, 1	6p21.31
84062	DTNBP1	dystrobrevin binding protein 1	6p22.3
3720	JARID2	jumonji, AT rich interactive domain 2	6p24-p23
4835	NQO2	NAD(P)H dehydrogenase, quinone 2	6pter-q12
1268	CNR1	cannabinoid receptor 1 (brain)	6q14-q15
319100	TAAR6	trace amine associated receptor 6	6q23.2
54806	AHI1	Abelson helper integration site 1	6q23.3
4852	NPY	neuropeptide Y	7p15.1
6804	STX1A	syntaxin 1A (brain)	7q11.23
9863	MAGI2	membrane associated guanylate kinase, WW and PDZ domain containing 2	7q21
2913	GRM3	glutamate receptor, metabotropic 3	7q21.1-q21.2
5649	RELN	reelin	7q22
93986	FOXP2	forkhead box P2	7q31
5803	PTPRZ1	protein tyrosine phosphatase, receptor-type, Z polypeptide 1	7q31.3
3361	HTR5A	5-hydroxytryptamine (serotonin) receptor 5A	7q36.1
27121	DKK4	dickkopf homolog 4 (Xenopus laevis)	8p11.2-p11.1
3084	NRG1	neuregulin 1	8p12
7976	FZD3	frizzled homolog 3 (Drosophila)	8p21
5533	PPP3CC	protein phosphatase 3 (formerly 2B),	8p21.3

Gene ID	Symbol	Official Full Name	Location
		catalytic subunit, gamma isoform	
6570	SLC18A1	solute carrier family 18 (vesicular monoamine), member 1	8p21.3
1808	DPYSL2	dihydropyrimidinase-like 2	8p22-p21
5108	PCM1	pericentriolar material 1	8p22-p21.3
1960	EGR3	early growth response 3	8p23-p21
2902	GRIN1	glutamate receptor, ionotropic, N-methyl D-aspartate 1	9q34.3
9968	MED12	mediator complex subunit 12	Xq13
1438	CSF2RA	colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage)	Xp22.32 and Yp11.3

These genes are prioritized by using the meta-analysis results in Lewis et al.[7] Multi-dimensional evidence-based gene prioritization approach was utilized to generate the candidate genes. Exhaustive search in a weight-matrix space was used to find the optimal weight matrix in order to combine four categories of data (association, linkage, expression, and literature). Please refer the SZGR database (http://bioinfo.mc.vanderbilt.edu/SZGR/) for details.[8]

Table S2. 173 schizophrenia susceptibility genes from the SZGR database (by Ng et al.[9])

Gene ID	Gene name	Official Full Name	Location
4853	NOTCH2	Notch homolog 2 (Drosophila)	1p13-p11
2944	GSTM1	glutathione S-transferase mu 1	1p13.3
22854	NTNG1	netrin G1	1p13.3
5142	PDE4B	phosphodiesterase 4B, cAMP-specific (phosphodiesterase E4 dunce homolog, Drosophila)	1p31
2899	GRIK3	glutamate receptor, ionotropic, kainate 3	1p34-p33
4524	MTHFR	5,10-methylenetetrahydrofolate reductase (NADPH)	1p36.3
405	ARNT	aryl hydrocarbon receptor nuclear translocator	1q21
6271	S100A1	S100 calcium binding protein A1	1q21
6281	S100A10	S100 calcium binding protein A10	1q21
9826	ARHGEF11	Rho guanine nucleotide exchange factor (GEF) 11	1q21
10712	C1orf2	chromosome 1 open reading frame 2	1q21
2532	DARC	Duffy blood group, chemokine receptor	1q21-q22
8991	SELENBP1	selenium binding protein 1	1q21-q22
1401	CRP	C-reactive protein, pentraxin-related	1q21-q23
2703	GJA8	gap junction protein, alpha 8, 50kDa	1q21.1
10628	TXNIP	thioredoxin interacting protein	1q21.1
2029	ENSA	endosulfine alpha	1q21.2
23208	SYT11	synaptotagmin XI	1q21.2
1141	CHRNB2	cholinergic receptor, nicotinic, beta 2 (neuronal)	1q21.3
3782	KCNN3	potassium intermediate/small conductance calcium-activated channel, subfamily N, member 3	1q21.3
126669	SHE	Src homology 2 domain containing E	1q21.3
4332	MNDA	myeloid cell nuclear differentiation antigen	1q22
10763	NES	Nestin	1q23.1
5999	RGS4	regulator of G-protein signaling 4	1q23.3
9722	NOS1AP	nitric oxide synthase 1 (neuronal) adaptor protein	1q23.3
127933	UHMK1	U2AF homology motif (UHM) kinase 1	1q23.3
5321	PLA2G4A	phospholipase A2, group IVA (cytosolic, calcium-dependent)	1q25
63827	BCAN	Brevican	1q31

Gene ID	Gene name	Official Full Name	Location
3586	IL10	interleukin 10	1q31-q32
1116	CHI3L1	chitinase 3-like 1 (cartilage glycoprotein-39)	1q32.1
5362	PLXNA2	plexin A2	1q32.2
8564	KMO	kynurenine 3-monooxygenase (kynurenine 3-hydroxylase)	1q42-q44
7257	TSNAX	translin-associated factor X	1q42.1
27185	DISC1	disrupted in schizophrenia 1	1q42.1
5305	PIP4K2A	phosphatidylinositol-5-phosphate 4-kinase, type II, alpha	10p12.2
1103	CHAT	choline acetyltransferase	10q11.2
2894	GRID1	glutamate receptor, ionotropic, delta 1	10q22
355	FAS	Fas (TNF receptor superfamily, member 6)	10q24.1
6571	SLC18A2	solute carrier family 18 (vesicular monoamine), member 2	10q25
627	BDNF	brain-derived neurotrophic factor	11p13
6506	SLC1A2	solute carrier family 1 (glial high affinity glutamate transporter), member 2	11p13-p12
7166	TPH1	tryptophan hydroxylase 1	11p15.3-p14
1815	DRD4	dopamine receptor D4	11p15.5
7054	TH	tyrosine hydroxylase	11p15.5
1813	DRD2	dopamine receptor D2	11q23
3359	HTR3A	5-hydroxytryptamine (serotonin) receptor 3A	11q23.1
4900	NRGN	neurogranin (protein kinase C substrate, RC3)	11q24
2904	GRIN2B	glutamate receptor, ionotropic, N-methyl D-aspartate 2B	12p12
4908	NTF3	neurotrophin 3	12p13
2065	ERBB3	v-erb-b2 erythroblastic leukemia viral oncogene homolog 3 (avian)	12q13
1610	DAO	D-amino-acid oxidase	12q24
4842	NOS1	nitric oxide synthase 1 (neuronal)	12q24.2-q24.31
3356	HTR2A	5-hydroxytryptamine (serotonin) receptor 2A	13q14-q21
3843	IPO5	importin 5	13q32.2
267012	DAOA	D-amino acid oxidase activator	13q33.2|13q34
64067	NPAS3	neuronal PAS domain protein 3	14q12-q13
1113	CHGA	chromogranin A (parathyroid secretory protein 1)	14q32
207	AKT1	v-akt murine thymoma viral oncogene	14q32.32|14q32.32

Gene ID	Gene name	Official Full Name	Location
		homolog 1	
89832	CHRFAM7A	CHRNA7 (cholinergic receptor, nicotinic, alpha 7, exons 5-10) and FAM7A (family with sequence similarity 7A, exons A-E) fusion	15q13.1
1139	CHRNA7	cholinergic receptor, nicotinic, alpha 7	15q14
1544	CYP1A2	cytochrome P450, family 1, subfamily A, polypeptide 2	15q24.1
3240	HP	Haptoglobin	16q22.1
7157	TP53	tumor protein p53	17p13.1
6532	SLC6A4	solute carrier family 6 (neurotransmitter transporter, serotonin), member 4	17q11.1-q12
718	C3	complement component 3	19p13.3-p13.2
9253	NUMBL	numb homolog (Drosophila)-like	19q13.13-q13.2
348	APOE	apolipoprotein E	19q13.2
8685	MARCO	macrophage receptor with collagenous structure	2q12-q13
1622	DBI	diazepam binding inhibitor (GABA receptor modulator, acyl-Coenzyme A binding protein)	2q12-q21
4867	NPHP1	nephronophthisis 1 (juvenile)	2q13
7851	MALL	mal, T-cell differentiation protein-like	2q13
150465	TTL	tubulin tyrosine ligase	2q13
5624	PROC	protein C (inactivator of coagulation factors Va and VIIIa)	2q13-q14
3553	IL1B	interleukin 1, beta	2q14
51141	INSIG2	insulin induced gene 2	2q14.1
3557	IL1RN	interleukin 1 receptor antagonist	2q14.2
4929	NR4A2	nuclear receptor subfamily 4, group A, member 2	2q22-q23
2571	GAD1	glutamate decarboxylase 1 (brain, 67kDa)	2q31
3623	INHA	inhibin, alpha	2q33-q36
130749	CPO	carboxypeptidase O	2q33.3
2066	ERBB4	v-erb-a erythroblastic leukemia viral oncogene homolog 4 (avian)	2q33.3-q34
2335	FN1	fibronectin 1	2q34
4133	MAP2	microtubule-associated protein 2	2q34-q35
1674	DES	Desmin	2q35
3577	IL8RA	interleukin 8 receptor, alpha	2q35

Gene ID	Gene name	Official Full Name	Location
79137	FAM134A	family with sequence similarity 134, member A	2q35
7857	SCG2	secretogranin II (chromogranin C)	2q35-q36
1114	CHGB	chromogranin B (secretogranin 1)	20pter-p12
2937	GSS	glutathione synthetase	20q11.2
10215	OLIG2	oligodendrocyte lineage transcription factor 2	21q22.11
54584	GNB1L	guanine nucleotide binding protein (G protein), beta polypeptide 1-like	22q11.2
5625	PRODH	proline dehydrogenase (oxidase) 1	22q11.21
6899	TBX1	T-box 1	22q11.21
7122	CLDN5	claudin 5	22q11.21
9993	DGCR2	DiGeorge syndrome critical region gene 2	22q11.21
29801	ZDHHC8	zinc finger, DHHC-type containing 8	22q11.21
1312	COMT	catechol-O-methyltransferase	22q11.21-q11.23|22q11.21
7494	XBP1	X-box binding protein 1	22q12.1|22q12
7533	YWHAH	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, eta polypeptide	22q12.3
1439	CSF2RB	colony stimulating factor 2 receptor, beta, low-affinity (granulocyte-macrophage)	22q13.1
6663	SOX10	SRY (sex determining region Y)-box 10	22q13.1
8398	PLA2G6	phospholipase A2, group VI (cytosolic, calcium-independent)	22q13.1
9463	PICK1	protein interacting with PRKCA 1	22q13.1
23209	MLC1	megalencephalic leukoencephalopathy with subcortical cysts 1	22q13.33
9223	MAGI1	membrane associated guanylate kinase, WW and PDZ domain containing 1	3p14.1
6854	SYN2	synapsin II	3p25
10752	CHL1	cell adhesion molecule with homology to L1CAM (close homolog of L1)	3p26.1
1814	DRD3	dopamine receptor D3	3q13.3
339855	KY	kyphoscoliosis peptidase	3q22.2
886	CCKAR	cholecystokinin A receptor	4p15.1-p15.2
1816	DRD5	dopamine receptor D5	4p16.1
2638	GC	group-specific component (vitamin D binding protein)	4q12-q13
10611	PDLIM5	PDZ and LIM domain 5	4q22

Gene ID	Gene name	Official Full Name	Location
1950	EGF	epidermal growth factor (beta-urogastrone)	4q25
6531	SLC6A3	solute carrier family 6 (neurotransmitter transporter, dopamine), member 3	5p15.3
3350	HTR1A	5-hydroxytryptamine (serotonin) receptor 1A	5q11.2-q13
5066	PAM	peptidylglycine alpha-amidating monooxygenase	5q14-q21
2246	FGF1	fibroblast growth factor 1 (acidic)	5q31
3702	ITK	IL2-inducible T-cell kinase	5q31-q32
5521	PPP2R2B	protein phosphatase 2 (formerly 2A), regulatory subunit B, beta isoform	5q31-q32
3360	HTR4	5-hydroxytryptamine (serotonin) receptor 4	5q31-q33
3562	IL3	interleukin 3 (colony-stimulating factor, multiple)	5q31.1
51735	RAPGEF6	Rap guanine nucleotide exchange factor (GEF) 6	5q31.1
56990	CDC42SE2	CDC42 small effector 2	5q31.1
2566	GABRG2	gamma-aminobutyric acid (GABA) A receptor, gamma 2	5q31.1-q33.1
3593	IL12B	interleukin 12B (natural killer cell stimulatory factor 2, cytotoxic lymphocyte maturation factor 2, p40)	5q31.1-q33.1
51520	LARS	leucyl-tRNA synthetase	5q32
2568	GABRP	gamma-aminobutyric acid (GABA) A receptor, pi	5q33-q34
2161	F12	coagulation factor XII (Hageman factor)	5q33-qter
134265	AFAP1L1	actin filament associated protein 1-like 1	5q33.1
9443	MED7	mediator complex subunit 7	5q33.3
9685	CLINT1	clathrin interactor 1	5q33.3
2890	GRIA1	glutamate receptor, ionotropic, AMPA 1	5q33|5q31.1
2559	GABRA6	gamma-aminobutyric acid (GABA) A receptor, alpha 6	5q34
2561	GABRB2	gamma-aminobutyric acid (GABA) A receptor, beta 2	5q34
2554	GABRA1	gamma-aminobutyric acid (GABA) A receptor, alpha 1	5q34-q35
6586	SLIT3	slit homolog 3 (Drosophila)	5q35
1812	DRD1	dopamine receptor D1	5q35.1
10814	CPLX2	complexin 2	5q35.2

Gene ID	Gene name	Official Full Name	Location
629	CFB	complement factor B	6p21.3
721	C4B	complement component 4B (Childo blood group)	6p21.3
780	DDR1	discoidin domain receptor tyrosine kinase 1	6p21.3
4855	NOTCH4	Notch homolog 4 (Drosophila)	6p21.3
7124	TNF	tumor necrosis factor (TNF superfamily, member 2)	6p21.3
2550	GABBR1	gamma-aminobutyric acid (GABA) B receptor, 1	6p21.31
84062	DTNBP1	dystrobrevin binding protein 1	6p22.3
1268	CNR1	cannabinoid receptor 1 (brain)	6q14-q15
54806	AHI1	Abelson helper integration site 1	6q23.3
4852	NPY	neuropeptide Y	7p15.1
6804	STX1A	syntaxin 1A (brain)	7q11.23
9863	MAGI2	membrane associated guanylate kinase, WW and PDZ domain containing 2	7q21
2913	GRM3	glutamate receptor, metabotropic 3	7q21.1-q21.2
5649	RELN	Reelin	7q22
93986	FOXP2	forkhead box P2	7q31
5803	PTPRZ1	protein tyrosine phosphatase, receptor-type, Z polypeptide 1	7q31.3
27121	DKK4	dickkopf homolog 4 (Xenopus laevis)	8p11.2-p11.1
3084	NRG1	neuregulin 1	8p12
7976	FZD3	frizzled homolog 3 (Drosophila)	8p21
148	ADRA1A	adrenergic, alpha-1A-, receptor	8p21-p11.2
2936	GSR	glutathione reductase	8p21.1
55806	HR	hairless homolog (mouse)	8p21.2
5533	PPP3CC	protein phosphatase 3 (formerly 2B), catalytic subunit, gamma isoform	8p21.3
6570	SLC18A1	solute carrier family 18 (vesicular monoamine), member 1	8p21.3
4023	LPL	lipoprotein lipase	8p22
11178	LZTS1	leucine zipper, putative tumor suppressor 1	8p22
1808	DPYSL2	dihydropyrimidinase-like 2	8p22-p21
5108	PCM1	pericentriolar material 1	8p22-p21.3
1960	EGR3	early growth response 3	8p23-p21
2902	GRIN1	glutamate receptor, ionotropic, N-methyl	9q34.3

Gene ID	Gene name	Official Full Name	Location
		D-aspartate 1	
4128	MAOA	monoamine oxidase A	Xp11.3
367	AR	androgen receptor	Xq11.2-q12
9968	MED12	mediator complex subunit 12	Xq13
1438	CSF2RA	colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage)	

Note: These genes are prioritized by using the meta-analysis results of Ng et al.[9]

Table S3. Top schizophrenia susceptibility genes from the SZGene database[10]

#[a]	Gene	Polymorphism	Ethnicity	N minor (Grade)[b]	I^2 (Grade)[c]	Overall Grade[d]
1	*PRSS16*	rs6932590	All	13934 (A)	n.a. (A)	A
2	*PGBD1*	rs13211507	All	5075 (A)	0 (A)	A
3	*NRGN*	rs12807809	All	12620 (A)	0 (A)	A
4	*NOTCH4*	rs3131296	All	7829 (A)	19 (A)	A
5	*HIST1H2BJ*	rs6913660	All	10335 (A)	0 (A)	C
6	*PDE4B*	rs910694	All	2393 (A)	2 (A)	A
7	*TCF4*	rs9960767	All	4143 (A)	20 (A)	C
8	*DRD4*	rs4646984	All	1513 (A)	0 (A)	A
9	*DRD2*	rs6277	Caucasian	6721 (A)	79 (C)	C
10	*DAOA*	rs778293	Asian	3609 (A)	18 (A)	C
11	*ZNF804A*	rs1344706	Caucasian	13839 (A)	55 (C)	C
12	*AHI1*	rs2064430	All	15114 (A)	0 (A)	C
13	*TPH1*	rs1800532	All	5411 (A)	20 (A)	A
14	*HTR2A*	rs6311	Caucasian	4665 (A)	22 (A)	C
15	*RPP21*	rs3130375	All	1392 (A)	63 (C)	C
16	*CCKAR*	rs1800857	Caucasian	704 (B)	0 (A)	C
17	*GABRB2*	rs1816072	Caucasian	1621 (A)	0 (A)	C
18	*DTNBP1*	rs3213207	Caucasian	4066 (A)	0 (A)	C
19	*C6orf217*	rs1475069	All	10686 (A)	19 (A)	C
20	*RELN*	rs7341475	Caucasian	4149 (A)	0 (A)	C
21	*MDGA1*	rs11759115	All	1347 (A)	15 (A)	A
22	*CMYA5*	rs10043986	All	4300 (A)	3 (A)	C
23	*DISC1*	rs3737597	Caucasian	102 (B)	0 (A)	B
24	*NRG1*	rs10503929	All	2654 (A)	0 (A)	C
25	*MTHFR*	rs1801133	All	6832 (A)	59 (C)	C
26	*AKT1*	rs3803300	Caucasian	574 (B)	18 (A)	B
27	*RGS4*	rs2661319	All	17088 (A)	17 (A)	C
28	*PPP3CC*	rs10108011	All	6263 (A)	0 (A)	C
29	*RPGRIP1L*	rs9922369	All	1292 (A)	64 (C)	C
30	*GRIK3*	rs6691840	All	773 (B)	4 (A)	C
31	*HP*	Hp_1/2	Caucasian	2602 (A)	0 (A)	C
32	*PLXNA2*	rs841865	Caucasian	506 (B)	24 (A)	B
33	*COMT*	rs737865	All	11219 (A)	4 (A)	C
34	*OPCML*	rs3016384	Caucasian	15673 (A)	62 (C)	C
35	*GSTM1*	GSTM1*0	All	1165 (A)	63 (C)	C
36	*GRIN2B*	rs1019385	All	1269 (A)	58 (C)	C
37	*DAO*	rs4623951	All	1871 (A)	0 (A)	C
38	*IL1B*	rs16944	All	2972 (A)	0 (A)	C
39	*IL10*	rs1800896	Caucasian	457 (B)	41 (B)	C
40	*SLC18A1*	rs2270641	All	979 (B)	85 (C)	C
41	*APOE*	APOE_e2/3/4	Caucasian	1136 (A)	0 (A)	A

42	*SRR*	rs408067	Asian	1362 (A)	49 (B)	C

This table lists all genetic loci with at least one nominally significant meta-analysis result in the SZGene database (Updated December 23, 2011), representing top 42 candidate genes for schizophrenia. Loci are ranked by statistical significance (*P*-value). For loci containing more than one polymorphism showing nominally significant association in meta-analysis, only the polymorphism with the smallest P-value is displayed. Please refer to the SZGene database (http://www.szgene.org/) for details.[10]

[a]Ranking based on HuGENet (Human Genome Epidemiology Network) interim guidelines for the assessment of genetic association studies.

[b,c,d]HuGENet interim criteria was used to grade each nominally significant meta-analysis in the SZGene according to cumulative evidence of genetic associations, including sample size, measured as "N minor"; heterogeneity across studies, measured as "I^2".

Table S4. 38 core genes from the SZGR database[11]

Gene	Official Full Name	Location
MTHFR	5,10-methylenetetrahydrofolate reductase (NADPH)	1p36.3
RGS4	regulator of G-protein signaling 4	1q23.3
PLXNA2	plexin A2	1q32.2
DISC1	disrupted in schizophrenia 1	1q42.1
TPH1	tryptophan hydroxylase 1	11p15.3-p14
DRD4	dopamine receptor D4	11p15.5
GRIK4	glutamate receptor, ionotropic, kainate 4	11q22.3
DRD2	dopamine receptor D2	11q23
FEZ1	fasciculation and elongation protein zeta 1 (zygin I)	11q24.2
OPCML	opioid binding protein/cell adhesion molecule-like	11q25
GRIN2B	glutamate receptor, ionotropic, N-methyl D-aspartate 2B	12p12
DAO	D-amino-acid oxidase	12q24
HTR2A	5-hydroxytryptamine (serotonin) receptor 2A	13q14-q21
DAOA	D-amino acid oxidase activator	13q33.2|13q34
NPAS3	neuronal PAS domain protein 3	14q12-q13
AKT1	v-akt murine thymoma viral oncogene homolog 1	14q32.32|14q32.32
CHRNA7	cholinergic receptor, nicotinic, alpha 7	15q14
RPGRIP1L	RPGRIP1-like	16q12.2
HP	Haptoglobin	16q22.1
TP53	tumor protein p53	17p13.1
SLC6A4	solute carrier family 6 (neurotransmitter transporter, serotonin), member 4	17q11.1-q12
APOE	apolipoprotein E	19q13.2
IL1B	interleukin 1, beta	2q14
GAD1	glutamate decarboxylase 1 (brain, 67kDa)	2q31
ZNF804A	zinc finger protein 804A	2q32.1
ERBB4	v-erb-a erythroblastic leukemia viral oncogene homolog 4 (avian)	2q33.3-q34
PRODH	proline dehydrogenase (oxidase) 1	22q11.21
COMT	catechol-O-methyltransferase	22q11.21-q11.23|22q11.21
GABRB2	gamma-aminobutyric acid (GABA) A receptor, beta 2	5q34
DRD1	dopamine receptor D1	5q35.1
DTNBP1	dystrobrevin binding protein 1	6p22.3
OFCC1	orofacial cleft 1 candidate 1	6p24.3
MUTED	muted homolog (mouse)	6p25.1-p24.3

Gene	Official Full Name	Location
GRM3	glutamate receptor, metabotropic 3	7q21.1-q21.2
RELN	Reelin	7q22
NRG1	neuregulin 1	8p12
PPP3CC	protein phosphatase 3 (formerly 2B), catalytic subunit, gamma isoform	8p21.3
SLC18A1	solute carrier family 18 (vesicular monoamine), member 1	8p21.3

Note: These genes were manually curated based on Ross et al. and Allen et al.'s studies.[10, 12]

Table S5. 75 candidate genes for schizophrenia from SZGR database[11]

Gene ID	Symbol	Official Full Name	Location
260425	*MAGI3*	membrane associated guanylate kinase, WW and PDZ domain containing 3	1p12-p11.2
2944	*GSTM1*	glutathione S-transferase mu 1	1p13.3
2899	*GRIK3*	glutamate receptor, ionotropic, kainate 3	1p34-p33
4524	*MTHFR*	5,10-methylenetetrahydrofolate reductase (NADPH)	1p36.3
5999	*RGS4*	regulator of G-protein signaling 4	1q23.3
3586	*IL10*	interleukin 10	1q31-q32
5362	*PLXNA2*	plexin A2	1q32.2
27185	*DISC1*	disrupted in schizophrenia 1	1q42.1
5305	*PIP4K2A*	phosphatidylinositol-5-phosphate 4-kinase, type II, alpha	10p12.2
1103	*CHAT*	choline acetyltransferase	10q11.2
627	*BDNF*	brain-derived neurotrophic factor	11p13
6506	*SLC1A2*	solute carrier family 1 (glial high affinity glutamate transporter), member 2	11p13-p12
7166	*TPH1*	tryptophan hydroxylase 1	11p15.3-p14
1815	*DRD4*	dopamine receptor D4	11p15.5
486	*FXYD2*	FXYD domain containing ion transport regulator 2	11q23
1813	*DRD2*	dopamine receptor D2	11q23
3359	*HTR3A*	5-hydroxytryptamine (serotonin) receptor 3A	11q23.1
2904	*GRIN2B*	glutamate receptor, ionotropic, N-methyl D-aspartate 2B	12p12
2065	*ERBB3*	v-erb-b2 erythroblastic leukemia viral oncogene homolog 3 (avian)	12q13
1610	*DAO*	D-amino-acid oxidase	12q24
4842	*NOS1*	nitric oxide synthase 1 (neuronal)	12q24.2-q24.31
3356	*HTR2A*	5-hydroxytryptamine (serotonin) receptor 2A	13q14-q21
267012	*DAOA*	D-amino acid oxidase activator	13q33.2|13q34
1113	*CHGA*	chromogranin A (parathyroid secretory protein 1)	14q32
207	*AKT1*	v-akt murine thymoma viral oncogene homolog 1	14q32.32|14q32.32
89832	*CHRFAM7A*	CHRNA7 (cholinergic receptor, nicotinic,	15q13.1

Gene ID	Symbol	Official Full Name	Location
		alpha 7, exons 5-10) and FAM7A (family with sequence similarity 7A, exons A-E) fusion	
1544	*CYP1A2*	cytochrome P450, family 1, subfamily A, polypeptide 2	15q24.1
8128	*ST8SIA2*	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 2	15q26
63826	*SRR*	serine racemase	17p13
7157	*TP53*	tumor protein p53	17p13.1
6532	*SLC6A4*	solute carrier family 6 (neurotransmitter transporter, serotonin), member 4	17q11.1-q12
9253	*NUMBL*	numb homolog (Drosophila)-like	19q13.13-q13.2
348	*APOE*	apolipoprotein E	19q13.2
3553	*IL1B*	interleukin 1, beta	2q14
4929	*NR4A2*	nuclear receptor subfamily 4, group A, member 2	2q22-q23
2571	*GAD1*	glutamate decarboxylase 1 (brain, 67kDa)	2q31
2066	*ERBB4*	v-erb-a erythroblastic leukemia viral oncogene homolog 4 (avian)	2q33.3-q34
1114	*CHGB*	chromogranin B (secretogranin 1)	20pter-p12
5625	*PRODH*	proline dehydrogenase (oxidase) 1	22q11.21
9342	*SNAP29*	synaptosomal-associated protein, 29kDa	22q11.21
9993	*DGCR2*	DiGeorge syndrome critical region gene 2	22q11.21
1312	*COMT*	catechol-O-methyltransferase	22q11.21-q11.23|22q11.21
6663	*SOX10*	SRY (sex determining region Y)-box 10	22q13.1
9145	*SYNGR1*	synaptogyrin 1	22q13.1
9223	*MAGI1*	membrane associated guanylate kinase, WW and PDZ domain containing 1	3p14.1
3269	*HRH1*	histamine receptor H1	3p25
6854	*SYN2*	synapsin II	3p25
10752	*CHL1*	cell adhesion molecule with homology to L1CAM (close homolog of L1)	3p26.1
1814	*DRD3*	dopamine receptor D3	3q13.3
2932	*GSK3B*	glycogen synthase kinase 3 beta	3q13.3
886	*CCKAR*	cholecystokinin A receptor	4p15.1-p15.2
10611	*PDLIM5*	PDZ and LIM domain 5	4q22
3350	*HTR1A*	5-hydroxytryptamine (serotonin) receptor 1A	5q11.2-q13

Gene ID	Symbol	Official Full Name	Location
3565	*IL4*	interleukin 4	5q31.1
2566	*GABRG2*	gamma-aminobutyric acid (GABA) A receptor, gamma 2	5q31.1-q33.1
2568	*GABRP*	gamma-aminobutyric acid (GABA) A receptor, pi	5q33-q34
9685	*CLINT1*	clathrin interactor 1	5q33.3
2559	*GABRA6*	gamma-aminobutyric acid (GABA) A receptor, alpha 6	5q34
2561	*GABRB2*	gamma-aminobutyric acid (GABA) A receptor, beta 2	5q34
2554	*GABRA1*	gamma-aminobutyric acid (GABA) A receptor, alpha 1	5q34-q35
4277	*MICB*	MHC class I polypeptide-related sequence B	6p21.3
4855	*NOTCH4*	Notch homolog 4 (Drosophila)	6p21.3
7124	*TNF*	tumor necrosis factor (TNF superfamily, member 2)	6p21.3
2550	*GABBR1*	gamma-aminobutyric acid (GABA) B receptor, 1	6p21.31
84062	*DTNBP1*	dystrobrevin binding protein 1	6p22.3
319100	*TAAR6*	trace amine associated receptor 6	6q23.2
9863	*MAGI2*	membrane associated guanylate kinase, WW and PDZ domain containing 2	7q21
2913	*GRM3*	glutamate receptor, metabotropic 3	7q21.1-q21.2
93986	*FOXP2*	forkhead box P2	7q31
5803	*PTPRZ1*	protein tyrosine phosphatase, receptor-type, Z polypeptide 1	7q31.3
3084	*NRG1*	neuregulin 1	8p12
5533	*PPP3CC*	protein phosphatase 3 (formerly 2B), catalytic subunit, gamma isoform	8p21.3
6570	*SLC18A1*	solute carrier family 18 (vesicular monoamine), member 1	8p21.3
2902	*GRIN1*	glutamate receptor, ionotropic, N-methyl D-aspartate 1	9q34.3
1438	*CSF2RA*	colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage)	Xp22.32 and Yp11.3

These candidate genes are prioritized by combining the odds ratio method and ranking more than 500 genes in more than 2000 association studies. Please refer the Schizophrenia Gene Resource (SZGR) database for detail.[13]

Table S6. A comprehensive and non-overlapping list of schizophrenia candidate genes from genetic linkage and association studies

Gene name	Gene name	Gene name	Gene name	Gene name	Gene name	Gene name
SNAP29	*PRODH*	*CPO*	*CHRNB2*	*RAPGEF6*	*MAOA*	*CSF2RA*
ARNT	*GRM3*	*SHE*	*NOS1*	*NPHP1*	*TRMT2A*	*TCF4*
SLC18A1	*LTA*	*NR4A2*	*GSK3B*	*CDC42SE2*	*PROC*	*MED12*
BDNF	*CCKAR*	*NUMBL*	*HR*	*ST8SIA2*	*DGCR6*	*CPLX2*
MLC1	*ARVCF*	*JARID2*	*IL10*	*KMO*	*TAAR6*	*SCG2*
ERBB4	*KY*	*SLC1A2*	*CHAT*	*GABRA1*	*APOE*	*INHA*
PRSS16	*TH*	*PPP2R2B*	*EGR3*	*ADRA1A*	*NRG2*	*IL18RAP*
GRIN1	*CMYA5*	*MALL*	*CHI3L1*	*COMT*	*NOTCH4*	*SLIT3*
FXYD2	*FOXP2*	*NRG1*	*HRH1*	*CSF2RB*	*IL12B*	*PTPRZ1*
GABBR1	*ZNF804A*	*PICK1*	*HIST1H2BJ*	*RPGRIP1L*	*CLINT1*	*CFB*
DISC1	*TNF*	*HTR3A*	*CHGA*	*RTN4R*	*TSNAX*	*AKT1*
DBI	*CHL1*	*AR*	*DRD4*	*GRIK3*	*FAM134A*	*NTF3*
IL8RA	*INSIG2*	*DES*	*CRP*	*PDLIM5*	*SYN2*	*GRIN2B*
C4B	*DTNBP1*	*GABRG2*	*RPP21*	*AHI1*	*MEGF10*	*FGF1*
PGBD1	*HTR1A*	*DGCR2*	*SELENBP1*	*MNDA*	*SRR*	*FXYD6*
DDR1	*MICB*	*CLDN5*	*NPAS3*	*SLC6A3*	*DRD1*	*DPYSL2*
GSR	*PIK3C3*	*CHRFAM7A*	*TXNIP*	*NTNG1*	*IPO5*	*PCM1*
IL4	*FEZ1*	*GABRA6*	*DARC*	*XBP1*	*ERBB3*	*PLA2G4A*
GRIA4	*IL3*	*HINT1*	*SYT11*	*IL1B*	*C3*	*ZDHHC8*
CHGB	*SLC6A4*	*SOX10*	*C18orf1*	*GRID1*	*FAS*	*DRD2*
RTN4	*MUTED*	*AFAP1L1*	*NQO2*	*PIP4K2A*	*F12*	*LPL*
SLC18A2	*GABRP*	*GJA8*	*TPH1*	*DRD5*	*HP*	*PLA2G6*
HTR5A	*DAO*	*UFD1L*	*GRIA1*	*MED7*	*KCNN3*	*PAM*
NEUROG1	*LARS*	*CYP1A2*	*GRIK4*	*FN1*	*OFCC1*	*DRD3*
C1orf2	*MAGI3*	*LZTS1*	*FZD3*	*MCHR1*	*PLXNA2*	*TTL*
GRM7	*RGS4*	*BCAN*	*MTHFR*	*CNR1*	*GSTM1*	*GC*
ENSA	*RELN*	*STX1A*	*GSS*	*MARCO*	*MAGI2*	*TP53*
HTR2A	*S100A1*	*GAD1*	*GNB1L*	*TBX1*	*IL1RN*	*YWHAH*
GCLM	*EGF*	*ACSL6*	*C6orf217*	*PDE4B*	*S100A10*	*CHRNA7*
MDGA1	*IL18R1*	*DAOA*	*MAP2*	*NPY*	*NOTCH2*	*DKK4*
ARHGEF11	*IL1A*	*OLIG2*	*SYNGR1*	*IL18*	*PPP3CC*	*RANBP1*
ITK	*NRGN*	*MAGI1*	*NES*	*UHMK1*	*NOS1AP*	*GABRB2*
OPCML	*HTR4*					

Table S7. Genes identified by schizophrenia GWAS

Gene name	Gene name	Gene name	Gene name	Gene name	Gene name	Gene name
RCN3	*FES*	*YPEL4*	*TAC3*	*ATXN7*	*MARS2*	*ATG13*
CREB3L1	*SMIM4*	*IMMP2L*	*RAI1*	*NEK4*	*GLT8D1*	*C4orf27*
CYP17A1	*NUTF2*	*HARS*	*SGSM2*	*ACD*	*FUT9*	*FAM109B*
CYP26B1	*SATB2*	*DDX28*	*TRIM26*	*DPP4*	*C11orf87*	*TSNAXIP1*
OTUD7B	*TMEM194A*	*SDCCAG8*	*ZMAT2*	*CDK2AP1*	*PLA2G15*	*FANCL*
KDM4A	*DPEP2*	*C3orf49*	*GID4*	*PLEKHO1*	*TSSK6*	*CNNM2*
IFT74	*MSL2*	*TMEM110-MUSTN1*	*TMX2*	*ZSCAN2*	*GDPD3*	*PCGF6*
HAPLN4	*AL049840.1*	*SCAF1*	*TCF20*	*HSPA9*	*BRP44*	*EFHD1*
VPS14C	*VRK2*	*NCK1*	*LRRIQ3*	*ARHGAP1*	*C2orf47*	*ZNF804A*
PRSS16	*GRM3*	*TLE1*	*TLE3*	*ESRP2*	*GIGYF2*	*CSMD1*
NKAPL	*EP300*	*DRD2*	*TSR1*	*FAM53C*	*STAG1*	*MPP6*
PSMA4	*PCDHA10*	*MPHOSPH9*	*FAM5B*	*PCDHA4*	*CILP2*	*ZDHHC5*
RILPL2	*RENBP*	*PCDHA5*	*F2*	*ANKRD63*	*ITIH3*	*SREBF2*
TSPAN18	*ARL6IP4*	*PCDHA1*	*GRIA1*	*HARS2*	*PAK6*	*STAT6*
STT3A	*PDCD11*	*ANP32E*	*MAD1L1*	*MYO15A*	*MKL1*	*MDK*
CHRM4	*GRAMD1B*	*L3MBTL2*	*LRP1*	*HIRIP3*	*APH1A*	*SBNO1*
SLC35G2	*MIR548AJ2*	*CTRL*	*MAN2A2*	*TBC1D5*	*LUZP2*	*PITPNM2*
GNL3	*SMG6*	*IGSF9B*	*NEK1*	*CNTN4*	*MAN2A1*	*SHISA8*
PTN	*C12orf65*	*KCTD13*	*INA*	*SLC7A6OS*	*ZSWIM6*	*BCL9*
NXPH4	*KLC1*	*PCDHA9*	*HARBI1*	*BOLL*	*MED19*	*CNOT1*
ITIH1	*NLGN4X*	*CTNND1*	*DGKI*	*WHSC1L1*	*SREBF1*	*USMG5*
PGBD1	*RGS6*	*CHADL*	*DPEP3*	*LCAT*	*PUS7*	*C2orf69*
INO80E	*PLCB2*	*CR1L*	*THOC7*	*PCGEM1*	*CUL3*	*GRIN2A*
CA8	*NT5C2*	*CHRNA3*	*TAF5*	*WDR55*	*PODXL*	*TBX6*
GFRA3	*FURIN*	*XRCC3*	*EPC2*	*RRAS*	*PRRG2*	*PRKD1*
NDUFA4L2	*BAG5*	*SFXN2*	*MUSTN1*	*AC027228.1*	*CKB*	*NRGN*
CACNA1C	*ZFYVE21*	*ERCC4*	*APOPT1*	*IK*	*SERPING1*	*QPCT*
TMCO6	*HCN1*	*NAB2*	*CHRNA5*	*ABCB9*	*NFATC3*	*SF3B1*
PBX4	*NDUFA6*	*ESAM*	*TAOK2*	*STAC3*	*NISCH*	*SLC45A1*
DOC2A	*PPP1R16B*	*WBP2NL*	*SEZ6L2*	*SRPK2*	*MMP16*	*TM6SF2*
GFOD2	*PPP1R13B*	*NGEF*	*C1orf132*	*SUGP1*	*ZNF408*	*TMTC1*
WBP1L	*R3HDM2*	*SLC12A4*	*CCDC39*	*PLCL1*	*CD46*	*FONG*
C12orf42	*KCNV1*	*LSM1*	*VSIG2*	*TYW5*	*ASPHD1*	*OGFOD2*
DUS2L	*AMBRA1*	*KCNB1*	*CKAP5*	*C11orf31*	*CLU*	*TSNARE1*
C2orf82	*SMDT1*	*CTNNA1*	*HLA-DRB9*	*AGPHD1*	*SNX19*	*AC005609.1*
SLC7A6	*C16orf92*	*CYP2D6*	*PRR12*	*HSPD1*	*ZKSCAN4*	*BCL11B*
NRN1L	*AC073043.2*	*OSBPL3*	*NDST3*	*FAM57B*	*PSMD6*	*ATP2A2*
NDUFA13	*DGKZ*	*PCDHA8*	*PTPRF*	*NCAN*	*ZNF536*	*MAPK3*
RANBP10	*TMEM219*	*CA14*	*THAP11*	*PARD6A*	*RIMS1*	*CLP1*
SLC39A8	*NOSIP*	*MYO1A*	*DND1*	*STAB1*	*TRANK1*	*NDUFA2*
ADAMTSL3	*MIR137*	*ENKD1*	*PCDHA6*	*RANGAP1*	*PCDHA2*	*SLC38A7*
DPYD	*PLAA*	*TRIM8*	*ZEB2*	*ADRBK2*	*FXR1*	*MAU2*

ANKRD44	*ATPAF2*	*PJA1*	*PSKH1*	*RERE*	*PBRM1*	*EDC4*
GATAD2A	*LRRC48*	*CENPM*	*AS3MT*	*COQ10B*	*ARL3*	*ALDOA*
PCDHA7	*SNAP91*	*EPHX2*	*ETF1*	*CCDC68*	*CNKSR2*	*HSPE1*
PPP2R3A	*AKT3*	*SHMT2*	*ACTR5*	*MEF2C*	*GPM6A*	*PLCH2*
CDC25C	*HIST1H2BJ*	*YPEL3*	*IREB2*	*C1orf54*	*ITIH4*	*PPP4C*
PSMB10	*NT5DC2*	*VPS45*	*PTGIS*	*NOTCH4*	*C16orf86*	*CACNB2*
HIST1H2BL	*REEP2*	*TRMT61A*	*C12orf79*	*SLC4A10*	*BTBD18*	*KCNJ13*
CENPT	*TOM1L2*	*CACNA1I*	*CLCN3*	*DNAJC19*	*DFNA5*	*CHRNB4*
SPCS1	*NAGA*	*C10orf32*	*PCDHA3*	*SETD8*	*PCCB*	*C1orf51*
MYO18B	*TCF4*	*EGR1*	*DRG2*	*SRR*	*TMEM110*	*KDM3B*
RFTN2	*TNFRSF13C*	*GALNT10*	*SLCO6A1*	*MLL5*	*GOLGA6L4*	*SLC32A1*
AC005477.1	*CD14*	*MSANTD2*				

Table S8. Top candidate genes for schizophrenia from convergent functional genomics (CFG) analysis of schizophrenia[14]

Gene name	CFG score
DISC1	5.0
HSPA1B	5.0
MBP	5.0
TCF4	5.0
MOBP	4.5
NCAM1	4.5
NDUFV2	4.5
NRCAM	4.5
RAB18	4.5
ADCYAP1	4.0
ALDH1A1	4.0
ANK3	4.0
BDNF	4.0
CD9	4.0
CNR1	4.0
COMT	4.0
CPLX2	4.0
DRD2	4.0
DTNBP1	4.0
FABP7	4.0
GABRB3	4.0
GAD1	4.0
GNB1L	4.0
GRIA1	4.0
GRIA4	4.0
GRIN2B	4.0
GRM5	4.0
GSN	4.0
HINT1	4.0
HTR2A	4.0
KALRN	4.0
KIF2A	4.0
NR4A2	4.0
NRG1	4.0
PDE4B	4.0
PRKCA	4.0
RELN	4.0
RGS4	4.0
SLC1A2	4.0
SNAP25	4.0
SYN2	4.0

TNIK	4.0

These top candidate genes for schizophrenia were generated by Convergent functional genomics (CFG) analysis integrating GWAS data with other genetic and gene expression studies in humans and animal models. For details, please refer to Ayalew et al.[15]

Table S9. Genes frequently affected by CNVs[16]

Gene name	Gene name	Gene name	Gene name	Gene name	Gene name	Gene name
NRXN1	*ASTN2*	*ANKRD35*	*FAM8A1*	*ATXN2L*	*TRIP13*	*HTF9C*
CHRNA7	*OTUD7A*	*PCYT1A*	*PTPRN2*	*OCA2*	*SERPINA4*	*RIT2*
CYFIP1	*RANBP1*	*PIAS3*	*MYOM2*	*SSTR5*	*LIX1L*	*PIGX*
GJA8	*TBX1*	*WDR1*	*HS3ST3B1*	*ZCWPW2*	*PEX11B*	*CTNND2*
NTAN1	*SCARF2*	*QPRT*	*GSTM1*	*PAK2*	*ITGA10*	*ZNF364*
NDE1	*SERPIND1*	*UHRF2*	*PRKG1*	*TTC26*	*PCQAP*	*PIGZ*
SNAP29	*CHD1L*	*NUP153*	*MAGI2*	*NDN*	*BRD9*	*CD160*
GJA5	*CDC45L*	*GRIK2*	*NKD2*	*C3orf67*	*CDRT4*	*RAPGEF6*
COMT	*ZNF74*	*NDNL2*	*WDR53*	*SENP5*	*SLC12A7*	*LRRC33*
BCL9	*DGCR8*	*UBXN7*	*NCBP2*	*USP7*	*P2RXL1*	*AIFM3*
PRKAB2	*DOC2A*	*TEKT3*	*NRG3*	*UFD1L*	*CIT*	*DLGAP2*
MYH11	*DLG2*	*DISC1*	*SNRPN*	*HIRA*	*PTPRM*	*KIAA0430*
PARK2	*ZDHHC8*	*APBA2*	*C16orf72*	*MPV17L*	*RBM8A*	*GSCL*
NIPA2	*CLDN5*	*P2RX6*	*CALN1*	*GNB1L*	*GABRA5*	*UBE3A*
VIPR2	*GP1BB*	*C3orf34*	*CABIN1*	*C22orf25*	*ZNHIT3*	*PMP22*
A2BP1	*C22orf29*	*CSMD3*	*GLDC*	*CLTCL1*	*FBXO45*	*TBX6*
PRODH	*MYT1L*	*FAM18B2*	*MTMR15*	*TXNRD2*	*CDRT15*	*TM4SF19*
KLHL22	*THAP7*	*SUMF1*	*KIF13A*	*RNF168*	*DGCR6L*	*PRRT2*
FMO5	*TRPM1*	*SLC7A4*	*MSRA*	*MTMR10*	*MAGEL2*	*DGCR2*
NIPA1	*CNTNAP2*	*ZDHHC19*	*GABRB3*	*CSMD1*	*TSSK2*	*LZTR1*
ACP6	*MRPL40*	*GABRG3*	*MKRN3*	*POLR3C*	*KLF13*	*ARVCF*
ABCC1	*NBPF11*	*TCTEX1D2*	*SLC1A3*	*NOMO3*	*MACROD2*	*RTN4R*
TUBGCP5	*DGCR14*	*WBSCR17*	*CAP2*	*SNURF*	*CRKL*	*SLC25A1*
RRN3	*PDXDC1*	*GSTT2*	*ABCC6*	*DGCR6*	*ERBB4*	*PI4KA*

Table S10. Genes identified by *Pascal*[17] with *P* value less than 10^{-6}

Gene	*P*-value	Gene	*P*-value	Gene	*P*-value
PLCH2	9.85E-07	*SEZ6L2*	1.44E-07	*FOXP1*	5.04E-09
SLC45A1	8.01E-09	*ASPHD1*	2.16E-09	*MSL2*	3.18E-07
KDM4A	4.95E-07	*KCTD13*	1.26E-09	*PCCB*	3.19E-09
DPYD	1.00E-12	*TMEM219*	3.33E-10	*FXR1*	1.07E-10
DPYD-AS1	4.02E-07	*TAOK2*	8.33E-10	*DNAJC19*	4.41E-10
MIR137HG	1.00E-12	*HIRIP3*	2.52E-09	*SOX2-OT*	4.12E-12
MIR2682	1.00E-12	*INO80E*	3.58E-09	*HCN1*	5.11E-07
MIR137	1.00E-12	*DOC2A*	7.68E-09	*ZSWIM6*	5.96E-10
CACNB2	1.42E-09	*C16orf92*	3.46E-08	*MAN2A1*	2.16E-07
WBP1L	1.54E-12	*FAM57B*	5.29E-08	*EGR1*	4.41E-07
CYP17A1	1.00E-12	*CNOT1*	1.96E-07	*ETF1*	2.41E-07
C10orf32-AS3MT	1.00E-12	*SLC38A7*	1.49E-08	*HSPA9*	1.98E-07
AS3MT	1.00E-12	*PLA2G15*	6.61E-07	*SNORD63*	3.92E-07
CNNM2	1.00E-12	*CPNE7*	6.17E-07	*GALNT10*	3.97E-08
NT5C2	1.00E-12	*SMG6*	1.77E-08	*SLC17A4*	2.20E-07
LOC729020	7.34E-07	*SRR*	4.71E-08	*SLC17A1*	5.38E-11
CREB3L1	1.33E-07	*TSR1*	1.07E-07	*SLC17A3*	1.00E-12
DGKZ	3.48E-09	*SNORD91B*	2.04E-07	*SLC17A2*	1.83E-10
MIR4688	2.52E-10	*SNORD91A*	2.07E-07	*TRIM38*	6.39E-10
MDK	2.77E-10	*SHISA6*	4.76E-07	*HIST1H1A*	1.00E-12
CHRM4	3.48E-10	*SREBF1*	2.67E-07	*HIST1H3A*	1.00E-12
AMBRA1	5.79E-10	*MIR33B*	3.58E-07	*HIST1H4A*	1.00E-12
MIR3160-1	1.06E-09	*TOM1L2*	8.55E-08	*HIST1H4B*	1.68E-12
MIR3160-2	1.23E-09	*LRRC48*	3.20E-08	*HIST1H3B*	1.56E-11
HARBI1	1.33E-09	*ATPAF2*	2.63E-08	*HIST1H2AB*	2.53E-11
ATG13	3.17E-09	*GID4*	2.84E-08	*HIST1H2BB*	1.05E-10
ARHGAP1	8.09E-09	*DRG2*	6.12E-08	*HIST1H3C*	1.01E-10
ZNF408	2.12E-08	*HAPLN4*	6.64E-07	*HIST1H1C*	3.00E-11
F2	8.28E-08	*TM6SF2*	6.04E-07	*HFE*	1.00E-12
MIR5582	9.74E-07	*SUGP1*	1.32E-07	*HIST1H4C*	1.00E-12
YPEL4	6.50E-07	*MAU2*	1.18E-07	*HIST1H1T*	1.00E-12
CLP1	4.52E-07	*GATAD2A*	3.35E-08	*HIST1H2BC*	1.00E-12
ZDHHC5	2.32E-07	*TSSK6*	1.16E-07	*HIST1H2AC*	1.00E-12
MED19	3.99E-07	*YJEFN3*	8.90E-08	*HIST1H1E*	1.89E-09
C11orf31	3.24E-07	*CILP2*	1.39E-07	*HIST1H2BD*	2.56E-07
BTBD18	3.19E-07	*PBX4*	2.03E-07	*HIST1H4H*	2.14E-07
DRD2	1.60E-07	*LPAR2*	4.74E-07	*BTN3A2*	1.00E-12
MIR4301	1.43E-07	*GMIP*	6.95E-07	*BTN2A2*	1.00E-12
NRGN	1.52E-10	*PRR12*	3.10E-07	*BTN3A1*	1.00E-12
VSIG2	4.06E-10	*RRAS*	3.44E-07	*BTN2A3P*	1.00E-12

ESAM	5.19E-09	*SCAF1*	4.71E-07	*BTN3A3*	4.71E-11
MSANTD2	8.64E-08	*MRPL33*	6.71E-07	*BTN2A1*	5.14E-12
SNX19	4.33E-11	*VRK2*	7.24E-12	*LOC285819*	1.00E-12
MIR4697HG	1.63E-07	*FANCL*	5.75E-09	*BTN1A1*	1.00E-12
IGSF9B	1.00E-12	*EMX1*	5.53E-07	*HCG11*	4.97E-11
LOC100128239	2.23E-08	*ANKRD44*	4.02E-07	*HMGN4*	7.33E-10
CACNA1C	1.00E-12	*SF3B1*	1.47E-10	*ABT1*	7.44E-11
CACNA1C-AS4	1.00E-12	*COQ10B*	1.06E-10	*ZNF322*	2.12E-11
CACNA1C-IT3	1.00E-12	*HSPD1*	7.58E-11	*GUSBP2*	3.54E-12
ABCB9	3.09E-10	*HSPE1-MOB4*	1.50E-10	*LINC00240*	1.00E-12
OGFOD2	1.27E-10	*MOB4*	9.06E-11	*LOC100270746*	1.07E-11
ARL6IP4	1.08E-10	*RFTN2*	2.79E-09	*HIST1H2BJ*	1.00E-12
PITPNM2	4.23E-10	*MARS2*	3.55E-07	*HIST1H2AG*	1.00E-12
MIR4304	9.05E-11	*PLCL1*	1.12E-07	*HIST1H2BK*	1.00E-12
LOC100507091	1.24E-10	*SATB2*	2.33E-07	*HIST1H4I*	1.00E-12
MPHOSPH9	3.69E-10	*FTCDNL1*	1.76E-07	*HIST1H2AH*	1.76E-12
C12orf65	1.25E-10	*C2orf69*	3.33E-12	*MIR3143*	1.76E-12
CDK2AP1	2.01E-10	*TYW5*	2.19E-11	*PRSS16*	6.07E-11
SBNO1	1.02E-09	*C2orf47*	3.02E-11	*POM121L2*	1.00E-12
SETD8	3.28E-09	*SPATS2L*	3.92E-07	*VN1R10P*	1.00E-12
RILPL2	1.80E-07	*EFHD1*	6.48E-07	*ZNF204P*	5.04E-09
MIR548AI	6.09E-07	*GIGYF2*	5.27E-09	*ZNF391*	7.10E-08
PCNX	7.73E-07	*KCNJ13*	3.70E-08	*ZNF184*	9.91E-08
BCL11B	3.62E-08	*C2orf82*	2.73E-11	*LOC100507173*	1.00E-12
EIF5	3.74E-07	*NGEF*	7.29E-12	*LOC100131289*	1.00E-12
SNORA28	7.13E-07	*SLC32A1*	1.72E-09	*HIST1H2BL*	1.00E-12
MARK3	8.49E-07	*ACTR5*	2.11E-09	*HIST1H3H*	1.00E-12
CKB	2.75E-09	*PPP1R16B*	9.61E-10	*HIST1H2AJ*	1.00E-12
TRMT61A	1.09E-10	*KCNB1*	4.27E-08	*HIST1H2BM*	1.00E-12
BAG5	6.88E-12	*PTGIS*	9.06E-08	*HIST1H4J*	1.00E-12
KLC1	3.16E-12	*PPDPF*	5.07E-07	*HIST1H4K*	1.00E-12
XRCC3	5.58E-12	*PTK6*	1.65E-07	*HIST1H2AK*	1.00E-12
ZFYVE21	5.16E-12	*SRMS*	3.40E-08	*HIST1H2BN*	1.00E-12
PPP1R13B	2.33E-11	*C20orf195*	5.80E-08	*HIST1H2AL*	1.00E-12
LINC00637	9.75E-11	*DOPEY2*	1.08E-07	*HIST1H1B*	1.00E-12
ANKRD63	1.76E-07	*SMCR7L*	2.31E-07	*HIST1H3I*	1.00E-12
PLCB2	5.01E-07	*ATF4*	3.06E-07	*HIST1H4L*	1.00E-12
IREB2	1.75E-07	*RPS19BP1*	1.95E-08	*HIST1H3J*	1.00E-12
AGPHD1	1.70E-07	*CACNA1I*	2.62E-12	*HIST1H2AM*	1.00E-12
CHRNA5	1.29E-07	*MIR1281*	7.33E-07	*HIST1H2BO*	1.00E-12
CHRNA3	4.99E-09	*TNFRSF13C*	8.17E-07	*OR2B2*	1.00E-12
CHRNB4	1.20E-09	*CENPM*	1.30E-07	*OR2B6*	5.56E-14
LOC80154	1.92E-07	*LINC00634*	4.15E-09	*ZNF165*	2.35E-12
RPS17L	1.02E-07	*SEPT3*	7.17E-09	*LOC100129195*	1.00E-12

LOC727751	2.98E-07	*WBP2NL*	2.26E-08	*ZSCAN12P1*	3.76E-12
CPEB1	8.69E-08	*NAGA*	5.39E-08	*ZSCAN16*	1.00E-12
LOC283692	5.48E-08	*FAM109B*	6.39E-08	*ZKSCAN8*	1.00E-12
AP3B2	4.99E-08	*SMDT1*	6.89E-08	*ZNF192P1*	1.00E-12
LOC338963	6.38E-08	*NDUFA6*	6.24E-08	*TOB2P1*	1.00E-12
LOC283693	5.21E-07	*NDUFA6-AS1*	4.11E-08	*ZSCAN9*	1.00E-12
EFTUD1P1	6.31E-10	*CYP2D6*	4.38E-08	*ZKSCAN4*	1.00E-12
LOC100505679	3.90E-10	*CYP2D7P1*	5.16E-08	*NKAPL*	1.00E-12
LOC440300	1.36E-09	*TCF20*	5.06E-07	*ZSCAN26*	1.00E-12
LOC388152	4.58E-09	*LOC388906*	5.30E-08	*PGBD1*	1.00E-12
LOC642423	4.66E-09	*RBMS3*	3.95E-07	*ZSCAN31*	1.00E-12
GOLGA6L4	7.32E-09	*TRANK1*	3.59E-09	*ZKSCAN3*	1.00E-12
DNM1P41	2.54E-08	*NT5DC2*	2.49E-07	*ZSCAN12*	1.00E-12
GOLGA6L5	4.55E-09	*SMIM4*	1.77E-07	*ZSCAN23*	1.00E-12
UBE2Q2P1	7.07E-10	*PBRM1*	1.40E-07	*ITPR3*	6.02E-07
LINC00933	4.87E-10	*GNL3*	1.65E-07	*LEMD2*	1.79E-07
ZSCAN2	2.51E-10	*SNORD19*	1.53E-07	*MLN*	7.27E-07
SCAND2P	2.64E-09	*SNORD19B*	1.53E-07	*SNAP91*	7.24E-09
WDR73	5.80E-09	*SNORD69*	1.52E-07	*MAD1L1*	5.05E-11
NMB	9.33E-09	*GLT8D1*	1.63E-07	*MIR4655*	4.10E-08
SEC11A	1.75E-08	*SPCS1*	1.85E-07	*LOC100216546*	1.17E-07
ZNF592	4.83E-08	*NEK4*	3.77E-08	*SRPK2*	1.91E-07
ALPK3	5.71E-07	*ITIH1*	1.48E-09	*IMMP2L*	3.57E-07
FURIN	3.67E-10	*ITIH3*	3.18E-10	*DGKI*	3.54E-07
FES	1.87E-08	*ITIH4*	1.52E-10	*FAM86B3P*	2.73E-07
MAN2A2	4.06E-07	*MUSTN1*	2.10E-10	*ZDHHC2*	9.48E-07
LOC440354	1.42E-08	*TMEM110-MUSTN1*	2.32E-10	*LETM2*	8.44E-07
LOC606724	1.24E-08	*TMEM110*	2.57E-10	*MMP16*	5.13E-07
SLX1A	1.22E-08	*SFMBT1*	7.59E-10	*LINC00051*	2.63E-11
SLX1B-SULT1A4	1.22E-08	*THOC7*	5.06E-07	*TSNARE1*	7.58E-08
LOC613038	1.22E-08				

PGC2 GWAS data[18] were used to do gene-based test by *Pascal*.[17]

Table S11. Genes affected by *de novo* mutations in schizophrenia probands[19]

Gene	*Gene*	*Gene*	*Gene*	*Gene*	*Gene*	*Gene*
EML6	*OBSCN*	*TTK*	*FADS2*	*SRCIN1*	*TAF13*	*FAM83H*
PCNXL3	*CAPN9*	*AKD1*	*SYVN1*	*STAC2*	*ALMS1*	*LINGO2*
C17orf97	*TSNAX*	*LAMA4*	*SPDYC*	*ERBB2*	*NEB*	*DCTN3*
BAIAP2	*ZP4*	*KIAA1244*	*NUMA1*	*KRT25*	*PHF7*	*CTSL1*
EPB41	*WDR64*	*KIAA1244*	*ARAP1*	*KRT20*	*TRH*	*ANGPTL2*
NUPL2	*CMPK2*	*UTRN*	*GDPD5*	*KRT15*	*PAQR9*	*NUP188*
ARRDC1	*GTF3C2*	*SYNE1*	*CAPN5*	*EFTUD2*	*NIPBL*	*ABL1*
SLC4A8	*NRXN1*	*ARID1B*	*CLNS1A*	*SPAG9*	*PCDHAC2*	*INPP5E*
ITM2B	*CTNNA2*	*AMZ1*	*MAML2*	*CA10*	*BTNL2*	*NOTCH1*
SCN2A	*ASTL*	*FBXL18*	*CEP164*	*BZRAP1*	*UFL1*	*DIP2C*
CNTN3	*CNNM3*	*CDCA7L*	*PHLDB1*	*MRC2*	*TNRC18*	*KIAA1462*
DLG2	*SLC5A7*	*SCRN1*	*HSPA8*	*TANC2*	*DAGLB*	*PAPSS2*
MARK4	*TMEM87B*	*NME8*	*HSPA8*	*KCNH6*	*ZEB1*	*KIF20B*
MYH7B	*ZC3H6*	*YAE1D1*	*FOXRED1*	*CCDC47*	*CUZD1*	*TMEM180*
KLHL17	*IL36B*	*GLI3*	*IQSEC3*	*SLC9A3R1*	*MKI67*	*PEX16*
TMEM201	*GLI2*	*POM121C*	*CD163L1*	*LLGL2*	*ZDHHC5*	*DGKZ*
EPHA2	*RIF1*	*RUNDC3B*	*CD163*	*TNRC6C*	*EPYC*	*DDB2*
TAF13	*TANC1*	*AKAP9*	*SLCO1B7*	*CBX8*	*PSPC1*	*INCENP*
MOV10	*XIRP2*	*PEX1*	*STK38L*	*BAIAP2*	*TP53I13*	*NRXN2*
POGZ	*SLC25A12*	*GNB2*	*ARID2*	*MYADML2*	*ZNF607*	*TECTA*
DARC	*CDCA7*	*MUC17*	*KRT82*	*FN3KRP*	*MGME1*	*CD9*
PIK3C2B	*TTN*	*RELN*	*ZC3H10*	*SMAD7*	*ITSN1*	*GXYLT1*
RASSF5	*AOX1*	*C7orf60*	*MDM1*	*ZNF532*	*TTC28*	*FAM186B*
TLK1	*FN1*	*C7orf60*	*CEP290*	*AP3D1*	*MASP2*	*PIP4K2C*
NCKIPSD	*SPEG*	*SLC13A1*	*IKBIP*	*ZNF77*	*CTRC*	*AGAP2*
YTHDC1	*CUL3*	*SND1*	*NUP37*	*TBXA2R*	*AHDC1*	*C12orf74*
ATG12	*ITM2C*	*CCDC136*	*ALDH1L2*	*SAFB2*	*COL16A1*	*DEPDC4*
PRRC2A	*KIF1A*	*TAS2R3*	*TCP11L2*	*EMR1*	*KIAA0319L*	*SART3*
AKD1	*ITPR1*	*SSPO*	*TCP11L2*	*EVI5L*	*HIVEP3*	*TBX5*
CLCN1	*RPUSD3*	*NOS3*	*ANAPC5*	*ADAMTS10*	*ZZZ3*	*ANAPC5*
KIAA1429	*NUP210*	*MLL3*	*PSPC1*	*ZNF426*	*FUBP1*	*NCOR2*
ZMYND11	*LSM3*	*TNKS*	*STARD13*	*FARSA*	*BCAR3*	*TBC1D4*
HSPA8	*STT3B*	*XKR6*	*NBEA*	*IER2*	*CAPZA1*	*CUL4A*
BCAT1	*ZNF445*	*XPO7*	*HTR2A*	*ILVBL*	*TRIM33*	*TMCO3*
SMARCC2	*LIMD1*	*FAM160B2*	*INTS6*	*WIZ*	*GDAP2*	*RPGRIP1*
XPOT	*CELSR3*	*PTK2B*	*NALCN*	*HAUS8*	*TDRKH*	*PSMB11*
KLHDC1	*RAD54L2*	*ESCO2*	*TINF2*	*UNC13A*	*GBA*	*KCNH5*
HSP90AA1	*ALAS1*	*KIF13B*	*NYNRIN*	*KXD1*	*HAPLN2*	*ZNF410*
DUOXA2	*DNAH1*	*TEX15*	*C14orf28*	*ZNF536*	*ARHGEF11*	*TMCO5A*
VPS13C	*PHF7*	*WRN*	*ESR2*	*MED29*	*TDRD5*	*OIP5*
NR2E3	*KCTD6*	*ANK1*	*PLEKHH1*	*SHKBP1*	*NID1*	*DNAJA4*
DNAH9	*PTPRG*	*TTPA*	*ACTN1*	*AXL*	*RYR2*	*BAIAP3*

ZNF407	*GXYLT2*	*XKR9*	*DICER1*	*ERF*	*KIF26B*	*GNPTG*
GTPBP3	*CRYBG3*	*PI15*	*WARS*	*CIC*	*VIT*	*USP7*
SHANK1	*TFG*	*FABP4*	*AKT1*	*ALDH16A1*	*HEATR5B*	*ACSM2B*
MIF	*ALCAM*	*YWHAZ*	*TRPM1*	*TSKS*	*DNAH6*	*DNAH3*
NAP1L2	*KIAA2018*	*TRHR*	*RYR3*	*LENG1*	*INPP4A*	*EARS2*
DVL1	*PLXNA1*	*SQLE*	*RYR3*	*BRSK1*	*CHRNG*	*ATXN2L*
ESPN	*HPS3*	*ASAP1*	*DLL4*	*ZNF551*	*BRPF1*	*ADCY7*
EMC1	*SHOX2*	*PUF60*	*INO80*	*ZBTB45*	*ATP2B2*	*JPH3*
LUZP1	*PSMD2*	*EPPK1*	*SPTBN5*	*CHMP2A*	*GLB1*	*SLC7A5*
NIPAL3	*DLG1*	*TONSL*	*DUOXA1*	*ZNF133*	*DCLK3*	*ANKRD11*
EXTL1	*ZNF721*	*RANBP6*	*GNB5*	*ABHD12*	*CCBP2*	*SPATA2L*
NR0B2	*PSAPL1*	*HAUS6*	*CLPX*	*DUSP15*	*DNAH12*	*SMYD4*
PTAFR	*ZNF518B*	*ACO1*	*UACA*	*ASXL1*	*DZIP3*	*MYO15A*
YARS	*SULT1B1*	*PAX5*	*CHSY1*	*FAM83C*	*CD96*	*WNK4*
PHC2	*ADAMTS3*	*PRUNE2*	*RHBDF1*	*EPB41L1*	*ALDH1L1*	*SP2*
PHC2	*PARM1*	*VPS13A*	*MPG*	*TGM2*	*PLXND1*	*MRPL27*
CSMD2	*SEC31A*	*COL15A1*	*SOLH*	*SULF2*	*IGSF10*	*FAM20A*
CLSPN	*GRID2*	*FSD1L*	*DNASE1L2*	*DPM1*	*EIF4G1*	*UBE2O*
THRAP3	*METTL14*	*ACTL7B*	*CDIP1*	*TSHZ2*	*EPHB3*	*LGALS3BP*
CCDC30	*MYOZ2*	*PTPN3*	*GRIN2A*	*SLMO2*	*HTT*	*KDSR*
PTPRF	*ANKRD50*	*COL27A1*	*CLEC16A*	*C20orf197*	*USP38*	*SERPINB8*
PIK3R3	*FAT4*	*PHF19*	*MYH11*	*GTPBP5*	*LPCAT1*	*LONP1*
TTC4	*PCDH10*	*LHX6*	*EEF2K*	*ADAMTS1*	*MAP3K1*	*SLC25A23*
DNAJC6	*DCHS2*	*C9orf114*	*COG7*	*URB1*	*MAST4*	*OLFM2*
LPHN2	*CBR4*	*POMT1*	*TMEM219*	*IFNAR2*	*VCAN*	*MAST1*
LPAR3	*SORBS2*	*EGFL7*	*TBC1D10B*	*SON*	*CHSY3*	*USF2*
ZNF326	*TLR3*	*MAMDC4*	*FBXL19*	*TSPEAR*	*KIF3A*	*ARHGAP33*
RPL5	*PDZD2*	*EXD3*	*PYCARD*	*COL6A2*	*VDAC1*	*PPP1R13L*
MYBPHL	*ZFR*	*EPC1*	*AKTIP*	*PCNT*	*PCDHGA11*	*FPR2*
SLC6A17	*ADAMTS6*	*MAPK8*	*NLRC5*	*CCT8L2*	*ARSI*	*SNTA1*
PDE4DIP	*TAF9*	*LIPK*	*NLRC5*	*CABIN1*	*GEMIN5*	*CBFA2T2*
DCST1	*SHROOM1*	*ANKRD22*	*CDH5*	*MYO18B*	*SH3PXD2B*	*KIAA1755*
ASH1L	*CD14*	*ACTA2*	*CES4A*	*EIF4ENIF1*	*DBN1*	*MAFB*
INSRR	*CD14*	*MYOF*	*DPEP2*	*APOL5*	*EXOC2*	*UBE2C*
ARHGEF11	*SH3TC2*	*TM9SF3*	*SLC7A6*	*MYH9*	*HIST1H3G*	*GNAS*
FCGR3B	*STK10*	*POLL*	*PMFBP1*	*EP300*	*OR11A1*	*DSCAM*
PBX1	*NSD1*	*PDCD11*	*CLEC3A*	*CERK*	*ANKS1A*	*TSPEAR*
PRRC2C	*GFOD1*	*ACSL5*	*TAF1C*	*ARSE*	*PIM1*	*KRTAP10-3*
KLHL20	*JARID2*	*FAM160B1*	*COX4I1*	*TLR7*	*ANKRD66*	*LZTR1*
SOAT1	*ALDH5A1*	*MUC6*	*FBXO31*	*DMD*	*TAAR8*	*MYH9*
TPR	*HIST1H2BE*	*MUC6*	*SGSM2*	*TSPYL2*	*TNRC18*	*SMDT1*
KIF14	*HIST1H4I*	*MUC5B*	*ALOX15*	*HUWE1*	*ABCA13*	*CELSR1*
CACNA1S	*FKBPL*	*PDE3B*	*NLRP1*	*HUWE1*	*ACN9*	*DENND6B*
PLEKHA6	*RGL2*	*IGSF22*	*FBXO39*	*MAGT1*	*CTTNBP2*	*MUM1L1*

PIK3C2B	*C6orf222*	*KIF18A*	*ARHGEF15*	*NXF3*	*SLC13A4*	*ATP5J2-PTCD1*
NFASC	*KIF6*	*KIF18A*	*NOS2*	*ACSL4*	*SSPO*	*PEG3*
CR1L	*TFAP2D*	*ACCSL*	*TNFAIP1*	*HTATSF1*	*GIMAP8*	*PSAP*
TLR5	*PAQR8*	*LRP4*	*MYO18A*	*NIPAL3*	*RP1*	*PTCD1*
NVL	*RIMS1*	*PTPRJ*	*GIT1*	*HIVEP3*	*CHMP4C*	*ZIM2*
PRSS38	*KCNQ5*	*SMTNL1*	*SSH2*	*LPHN2*		

Table S12. Genes affected by rare disruptive mutations in schizophrenia cases[20]

Gene	*Gene*	*Gene*	*Gene*	*Gene*	*Gene*	*Gene*
AAK1	*CACNG1*	*DUSP8*	*IGSF22*	*NCOA6*	*PYCRL*	*SYNJ1*
AARS	*CACNG2*	*DVL1*	*IGSF8*	*NCOR1*	*PYGB*	*SYNPO*
AATK	*CACNG3*	*DYNC1H1*	*IGSF9B*	*NCOR2*	*R3HDM1*	*SYP*
ABCA2	*CACNG4*	*DYNC1I1*	*IKBIP*	*NCS1*	*R3HDM2*	*SYT1*
ABCA3	*CACNG5*	*DYNC1LI2*	*IL36B*	*NDRG2*	*RAB13*	*SYT5*
ABCD3	*CACNG6*	*DYNLL1*	*ILVBL*	*NDRG4*	*RAB15*	*SYT7*
ABCG1	*CACNG7*	*DYNLL2*	*IMMT*	*NDST1*	*RAB1A*	*SYVN1*
ABCG4	*CACNG8*	*DYNLRB1*	*INA*	*NDUFA13*	*RAB1B*	*SZT2*
ABHD12	*CACYBP*	*DYNLRB2*	*INO80*	*NDUFA2*	*RAB35*	*TAF13*
ABI1	*CADPS*	*EDEM2*	*INPP4A*	*NDUFA4*	*RAB3A*	*TAF1C*
ABI2	*CALM1*	*EEF1A1*	*INPP5A*	*NDUFA7*	*RAB3GAP2*	*TAF7L*
ABLIM1	*CALM3*	*EEF1A2*	*INSRR*	*NDUFB10*	*RAB5B*	*TAF9*
ABR	*CAMK2A*	*EEF1D*	*INTS1*	*NDUFB6*	*RAB5C*	*TAGLN3*
ACBD5	*CAMK2B*	*EEF1G*	*INTS10*	*NDUFB7*	*RAB6A*	*TANC1*
ACCSL	*CAMK2D*	*EEF2*	*INTS6*	*NDUFS1*	*RAB6B*	*TANC2*
ACLY	*CAMK2G*	*EEF2K*	*IPO13*	*NDUFS2*	*RAB7A*	*TAOK1*
ACO1	*CAMK2N1*	*EFHD2*	*IPO4*	*NDUFS3*	*RAC1*	*TAOK2*
ACO2	*CAMK4*	*EFTUD2*	*IPO5*	*NEB*	*RAD54L2*	*TAS2R3*
ACOT6	*CAMKK1*	*EGFL7*	*IQSEC1*	*NEDD4*	*RALA*	*TBC1D10B*
ACP1	*CAMKK2*	*EGR1*	*IQSEC2*	*NEFH*	*RALGAPA1*	*TBC1D14*
ACSL3	*CAMKV*	*EHD1*	*IQSEC3*	*NEFM*	*RALGAPB*	*TBC1D24*
ACSL4	*CAMSAP1*	*EHD3*	*IRGQ*	*NEGR1*	*RALGDS*	*TBC1D9*
ACSL5	*CAMTA1*	*EHMT1*	*IRS1*	*NELF*	*RANBP6*	*TBC1D9B*
ACSL6	*CAMTA2*	*EHMT2*	*IRS2*	*NEURL4*	*RAP1A*	*TBCB*
ACTA2	*CAND1*	*EIF2C1*	*ITGA6*	*NF1*	*RAP1GAP*	*TBXA2R*
ACTB	*CANX*	*EIF2C2*	*ITM2B*	*NFASC*	*RAP1GAP2*	*TCF20*
ACTL7B	*CAP2*	*EIF4A2*	*ITM2C*	*NFIC*	*RAP1GDS1*	*TCF25*
ACTN1	*CAPG*	*EIF4ENIF1*	*ITPR1*	*NFIX*	*RAPGEF1*	*TCF4*
ACTN2	*CAPN5*	*EIF4G1*	*ITSN1*	*NGEF*	*RAPGEF2*	*TCP1*
ACTN4	*CAPN9*	*EIF4G2*	*JAK1*	*NHSL1*	*RAPGEF4*	*TCP11L2*
ACTR1A	*CAPZA1*	*EIF4G3*	*JARID2*	*NIPAL3*	*RAPGEFL1*	*TEF*
ACTR1B	*CAPZA2*	*EIF5*	*JPH3*	*NIPBL*	*RARG*	*TEKT5*
ACTR2	*CAPZB*	*ELFN2*	*JPH4*	*NISCH*	*RASGRF1*	*TEP1*
ACTR3	*CARD6*	*ELMO2*	*KALRN*	*NLGN2*	*RASGRP1*	*TEX15*
ADAMTS1	*CASKIN1*	*EMC1*	*KBTBD11*	*NLGN3*	*RASSF5*	*TFAM*
ADAMTS10	*CASP4*	*EML2*	*KCNA2*	*NLRC5*	*RB1CC1*	*TFAP2D*
ADAMTS3	*CBR1*	*EML6*	*KCNB1*	*NLRP1*	*RBX1*	*TFG*
ADAMTS6	*CBR3*	*EMR1*	*KCNC3*	*NME1*	*RC3H1*	*TGM2*
ADAP1	*CBR4*	*EMR3*	*KCND2*	*NME8*	*RC3H2*	*TGM6*
ADARB1	*CBX6*	*ENC1*	*KCNH1*	*NNT*	*RECK*	*THBS1*
ADCY1	*CBX8*	*ENO1*	*KCNH3*	*NOMO1*	*RELN*	*THRA*
ADCY5	*CC2D1A*	*ENO2*	*KCNH6*	*NOS2*	*RERE*	*THRAP3*

ADCY7	*CCDC108*	*ENO3*	*KCNH7*	*NOS3*	*REV3L*	*TIAM1*
ADD1	*CCDC136*	*ENPP6*	*KCNMA1*	*NPAS2*	*RFX3*	*TINF2*
ADD2	*CCDC137*	*EP300*	*KCNQ2*	*NPEPPS*	*RGL2*	*TJP1*
ADD3	*CCDC30*	*EP400*	*KCNQ3*	*NPRL2*	*RGS12*	*TJP2*
ADNP	*CCDC39*	*EPB41*	*KCNQ5*	*NPTN*	*RGS7*	*TLE3*
ADRBK1	*CCDC47*	*EPB41L1*	*KCNT1*	*NPTXR*	*RGS7BP*	*TLK1*
AFF3	*CCDC84*	*EPB41L3*	*KCTD12*	*NR0B2*	*RHBDF1*	*TLN2*
AFF4	*CCT2*	*EPC1*	*KCTD6*	*NR2E3*	*RHOB*	*TLR3*
AGAP1	*CCT3*	*EPHA2*	*KCTD8*	*NR2F1*	*RHOBTB2*	*TLR5*
AGAP2	*CCT4*	*EPHA4*	*KDM2B*	*NRCAM*	*RHOG*	*TLR7*
AGAP3	*CCT5*	*EPN1*	*KDM4B*	*NRGN*	*RIF1*	*TM9SF3*
AGPAT3	*CCT6A*	*EPPK1*	*KDM5C*	*NRIP1*	*RIMBP2*	*TMEM132A*
AGRN	*CCT7*	*EPRS*	*KDM6B*	*NRXN1*	*RIMS1*	*TMEM151A*
AGTPBP1	*CCT8*	*EPS15L1*	*KIAA0284*	*NRXN2*	*ROCK2*	*TMEM151B*
AHCYL1	*CCT8L2*	*EPYC*	*KIAA0528*	*NRXN3*	*RPH3A*	*TMEM201*
AHDC1	*CD14*	*ERBB2*	*KIAA1045*	*NSD1*	*RPL12*	*TMEM219*
AHNAK	*CD163*	*ERC1*	*KIAA1109*	*NSF*	*RPL14*	*TMEM63B*
AHNAK2	*CD163L1*	*ERC2*	*KIAA1217*	*NSFL1C*	*RPL38*	*TMEM87B*
AK1	*CDC42BPA*	*ERF*	*KIAA1244*	*NTRK2*	*RPL5*	*TMEM8B*
AK5	*CDC42BPB*	*ESAM*	*KIAA1429*	*NTRK3*	*RPL7*	*TMOD1*
AKAP6	*CDCA7*	*ESCO2*	*KIAA2018*	*NUMA1*	*RPL8*	*TMOD2*
AKAP9	*CDCA7L*	*ESPN*	*KIF13B*	*NUP210*	*RPLP0*	*TNC*
AKD1	*CDH13*	*ESR2*	*KIF14*	*NUP37*	*RPN1*	*TNFAIP1*
AKR1A1	*CDH2*	*EVC2*	*KIF18A*	*NUP54*	*RPRD2*	*TNIK*
AKR7A2	*CDH5*	*EVI5L*	*KIF1A*	*NUP98*	*RPS13*	*TNK2*
AKT1	*CDK16*	*EVL*	*KIF1B*	*NUPL2*	*RPS14*	*TNKS*
AKT3	*CDK17*	*EXD3*	*KIF21A*	*NVL*	*RPS18*	*TNPO2*
AKTIP	*CDK5*	*EXOC1*	*KIF21B*	*NWD1*	*RPS19*	*TNR*
ALAS1	*CDK5R1*	*EXOC2*	*KIF2A*	*NXF3*	*RPS25*	*TNRC18*
ALCAM	*CDK5R2*	*EXOC3*	*KIF3C*	*NYNRIN*	*RPS27*	*TNRC6B*
ALDH16A1	*CDKL5*	*EXOC4*	*KIF5A*	*OBSCN*	*RPS3*	*TNRC6C*
ALDH1L2	*CEACAM18*	*EXOC8*	*KIF5B*	*OGDH*	*RPTOR*	*TNS3*
ALDH2	*CELF2*	*EXTL1*	*KIF5C*	*OGT*	*RPUSD3*	*TOLLIP*
ALDH5A1	*CELF5*	*EXTL3*	*KIF6*	*OLFM1*	*RRP1B*	*TOM1L2*
ALDH6A1	*CELSR2*	*EZR*	*KIFC2*	*OMG*	*RTN1*	*TOMM20*
ALDH7A1	*CELSR3*	*FABP4*	*KLC1*	*OPA1*	*RTN3*	*TOMM70A*
ALDOA	*CEND1*	*FADS2*	*KLC2*	*OPCML*	*RTN4*	*TONSL*
ALDOC	*CEP164*	*FAM115A*	*KLF12*	*OR4C46*	*RTN4R*	*TOP3B*
ALMS1	*CEP170*	*FAM120A*	*KLHDC1*	*OXCT1*	*RUFY3*	*TP53I13*
ALOX15	*CEP290*	*FAM13C*	*KLHL17*	*OXR1*	*RUNDC3B*	*TPI1*
ALS2	*CERK*	*FAM160A2*	*KLHL20*	*P2RY2*	*RUSC1*	*TPM1*
ALS2CL	*CES4A*	*FAM160B1*	*KLHL22*	*PABPC1*	*RUSC2*	*TPM3*
AMPH	*CFL1*	*FAM160B2*	*KNDC1*	*PACS1*	*RYR2*	*TPM4*
AMZ1	*CHCHD3*	*FAM171B*	*KPNA1*	*PACS2*	*RYR3*	*TPPP*

ANAPC1	*CHCHD6*	*FAM179B*	*KPNB1*	*PACSIN1*	*SAFB2*	*TPR*
ANAPC5	*CHD3*	*FAM3D*	*KRAS*	*PAG1*	*SALL2*	*TRAK1*
ANK1	*CHD4*	*FAM49A*	*KRT15*	*PAICS*	*SAMD4B*	*TRAK2*
ANK2	*CHD5*	*FAM5B*	*KRT20*	*PAK1*	*SAP130*	*TRAP1*
ANK3	*CHD6*	*FAM65A*	*KRT25*	*PAK6*	*SAP30BP*	*TRAPPC10*
ANKFY1	*CHD8*	*FAM81A*	*KRT82*	*PALM*	*SASH1*	*TRAPPC3*
ANKRD11	*CHMP2A*	*FAM83C*	*KXD1*	*PALM2*	*SBF1*	*TREM2*
ANKRD17	*CHMP4B*	*FARP1*	*L1CAM*	*PAPPA2*	*SBNO1*	*TRH*
ANKRD22	*CHN1*	*FARSA*	*LAMA1*	*PAQR8*	*SCAF1*	*TRHR*
ANKRD50	*CHN2*	*FARSB*	*LAMA2*	*PAQR9*	*SCAP*	*TRIM2*
ANKRD52	*CHST2*	*FASN*	*LAMA4*	*PARM1*	*SCCPDH*	*TRIM3*
ANKS1B	*CHSY1*	*FASTKD5*	*LANCL2*	*PAX5*	*SCD*	*TRIM32*
ANO9	*CIC*	*FAT1*	*LARGE*	*PBX1*	*SCFD1*	*TRIM37*
ANXA2	*CISD1*	*FAT2*	*LARP7*	*PBXIP1*	*SCN2A*	*TRIM9*
ANXA5	*CIT*	*FAT3*	*LARS2*	*PCBP2*	*SCN8A*	*TRIO*
ANXA6	*CKAP4*	*FAT4*	*LCT*	*PCDH1*	*SCRN1*	*TRIP12*
AOX1	*CKAP5*	*FBXL16*	*LDHA*	*PCDH10*	*SDF4*	*TRO*
AP1B1	*CKB*	*FBXL18*	*LDHB*	*PCDH7*	*SEC14L1*	*TRPC4AP*
AP1S1	*CKMT1B*	*FBXL19*	*LENG1*	*PCDH9*	*SEC16A*	*TRPM1*
AP2A1	*CLASP1*	*FBXO31*	*LHFPL4*	*PCDHA4*	*SEC23A*	*TRPM3*
AP2A2	*CLASP2*	*FBXO39*	*LHX6*	*PCDHAC2*	*SEC31A*	*TRRAP*
AP2B1	*CLCN1*	*FBXO41*	*LIMD1*	*PCDHGA12*	*SECISBP2L*	*TSC2*
AP2M1	*CLCN3*	*FBXO7*	*LIN7B*	*PCDHGC3*	*SEPT10*	*TSC22D1*
AP2S1	*CLEC16A*	*FCGBP*	*LIN7C*	*PCLO*	*SEPT11*	*TSHZ1*
AP3D1	*CLEC3A*	*FCGR3B*	*LINGO1*	*PCMT1*	*SEPT2*	*TSHZ2*
APBA1	*CLIC5*	*FCHO1*	*LIPK*	*PCNT*	*SEPT3*	*TSKS*
APBB1	*CLIP3*	*FKBP8*	*LLGL1*	*PCNX*	*SEPT4*	*TSNAX*
APC	*CLNS1A*	*FKBPL*	*LLGL2*	*PCNXL2*	*SEPT5*	*TSPAN7*
APC2	*CLPX*	*FLNA*	*LMO7*	*PCNXL3*	*SEPT6*	*TSPEAR*
APLP1	*CLSPN*	*FLOT1*	*LMTK2*	*PDCD11*	*SEPT7*	*TSPYL2*
APOD	*CLSTN1*	*FLOT2*	*LMTK3*	*PDCD6IP*	*SEPT8*	*TSPYL4*
APOE	*CLTC*	*FMN2*	*LPAR3*	*PDE2A*	*SEPT9*	*TTBK1*
APOL2	*CLU*	*FMNL2*	*LPHN1*	*PDE3B*	*SETD5*	*TTBK2*
APOL5	*CMPK2*	*FMNL3*	*LPHN2*	*PDE4B*	*SETX*	*TTC28*
APP	*CNNM3*	*FN1*	*LPHN3*	*PDE4DIP*	*SEZ6L2*	*TTC3*
AQP4	*CNP*	*FN3KRP*	*LPIN2*	*PDE8B*	*SFN*	*TTC4*
ARAP1	*CNTN1*	*FOXK2*	*LRP1*	*PDHA1*	*SFXN1*	*TTC7B*
ARAP2	*CNTN2*	*FOXO3*	*LRP3*	*PDHB*	*SFXN5*	*TTK*
ARF3	*CNTN3*	*FOXRED1*	*LRP4*	*PDS5B*	*SGIP1*	*TTLL7*
ARF5	*CNTNAP1*	*FRMPD4*	*LRP8*	*PDZD2*	*SGSM2*	*TTN*
ARFGEF1	*COBL*	*FRY*	*LRPPRC*	*PDZD8*	*SH3BP4*	*TTPA*
ARHGAP20	*COG7*	*FSCN1*	*LRRC41*	*PEA15*	*SH3GL1*	*TTYH1*
ARHGAP21	*COL15A1*	*FSD1*	*LRRC47*	*PEBP1*	*SH3GL2*	*TTYH3*
ARHGAP23	*COL27A1*	*FSD1L*	*LRRC4B*	*PEG3*	*SH3GLB2*	*TUB*

ARHGAP32	*COL3A1*	*FTH1*	*LRRC7*	*PER1*	*SH3TC2*	*TUBA1A*
ARHGAP33	*COL6A2*	*FYN*	*LRRC8B*	*PEX1*	*SHANK1*	*TUBA1B*
ARHGAP35	*COPG1*	*GABBR1*	*LRRN2*	*PFKL*	*SHANK2*	*TUBA4A*
ARHGDIA	*CORO1A*	*GABBR2*	*LSAMP*	*PFKM*	*SHANK3*	*TUBB2A*
ARHGEF10	*CORO1C*	*GABRA1*	*LSM3*	*PFKP*	*SHKBP1*	*TUBB2B*
ARHGEF11	*CORO2B*	*GAP43*	*LUZP1*	*PFN2*	*SHOX2*	*TUBB3*
ARHGEF12	*COX4I1*	*GAPDH*	*LYN*	*PGAM1*	*SHROOM1*	*TUFM*
ARHGEF15	*COX5B*	*GARNL3*	*LYNX1*	*PGD*	*SIPA1L1*	*TULP4*
ARHGEF17	*COX6B1*	*GAS7*	*MACF1*	*PGK1*	*SIPA1L2*	*TXNL1*
ARHGEF2	*COX6C*	*GBAS*	*MADD*	*PGM2L1*	*SIPA1L3*	*UACA*
ARHGEF38	*COX7A2L*	*GBF1*	*MAGEC1*	*PHACTR1*	*SIRPA*	*UBA1*
ARHGEF4	*CPE*	*GCN1L1*	*MAGED1*	*PHB*	*SKI*	*UBAP2L*
ARHGEF7	*CPLX1*	*GDI1*	*MAGI2*	*PHC2*	*SLC12A5*	*UBC*
ARID1A	*CPLX2*	*GDI2*	*MAGT1*	*PHF12*	*SLC12A6*	*UBE2N*
ARID1B	*CPNE5*	*GDPD5*	*MAMDC4*	*PHF19*	*SLC13A1*	*UBE2O*
ARID2	*CPT1C*	*GFAP*	*MAML2*	*PHF20*	*SLC17A1*	*UBE2V2*
ARNT2	*CR1L*	*GFOD1*	*MAN2A2*	*PHF23*	*SLC17A7*	*UBE3B*
ARPC2	*CREBBP*	*GIF*	*MAOA*	*PHF7*	*SLC19A2*	*UBE3C*
ARPC3	*CRIP2*	*GIT1*	*MAOB*	*PHGDH*	*SLC1A2*	*UBQLN1*
ARPP21	*CRMP1*	*GJA1*	*MAP1A*	*PHLDB1*	*SLC22A17*	*UBQLN2*
ARRB1	*CRTAC1*	*GLI2*	*MAP1B*	*PHYHIP*	*SLC24A2*	*UBR3*
ARRDC1	*CRTC1*	*GLI3*	*MAP1LC3A*	*PI15*	*SLC25A11*	*UBR5*
ARSE	*CRYAB*	*GLUD1*	*MAP2*	*PI4KA*	*SLC25A12*	*UCHL1*
ARVCF	*CRYBG3*	*GLUL*	*MAP2K1*	*PIGQ*	*SLC25A13*	*UFL1*
ASAP1	*CRYM*	*GNA13*	*MAP3K12*	*PIK3C2B*	*SLC25A18*	*UGT1A3*
ASH1L	*CS*	*GNAI2*	*MAP4*	*PIK3CB*	*SLC25A22*	*UHRF1BP1L*
ASTL	*CSE1L*	*GNAL*	*MAP4K4*	*PIK3R3*	*SLC25A23*	*ULK1*
ASXL1	*CSMD2*	*GNAO1*	*MAP6*	*PIKFYVE*	*SLC25A3*	*ULK2*
ATF7IP	*CSPG4*	*GNAS*	*MAP6D1*	*PIN1*	*SLC25A4*	*UNC13A*
ATG12	*CST3*	*GNAZ*	*MAP7D1*	*PINK1*	*SLC25A5*	*UNC13C*
ATG2A	*CTBP1*	*GNB1*	*MAPK1*	*PIP4K2A*	*SLC25A6*	*UNC5A*
ATG2B	*CTNNA1*	*GNB2*	*MAPK3*	*PIP4K2B*	*SLC26A8*	*UQCRC1*
ATG9A	*CTNNA2*	*GNB4*	*MAPK4*	*PIP5K1C*	*SLC27A4*	*UQCRC2*
ATIC	*CTNNB1*	*GNB5*	*MAPK8*	*PITPNM1*	*SLC4A3*	*URB1*
ATMIN	*CTNND1*	*GNL1*	*MAPK8IP1*	*PITPNM2*	*SLC4A4*	*URB2*
ATN1	*CTNND2*	*GNPAT*	*MAPK8IP3*	*PJA2*	*SLC4A8*	*USP22*
ATP13A2	*CUL3*	*GOT1*	*MAPKBP1*	*PKD1*	*SLC5A7*	*USP32*
ATP1A1	*CUL9*	*GPAM*	*MAPRE2*	*PKP4*	*SLC6A1*	*USP34*
ATP1A2	*CUX1*	*GPHN*	*MAPRE3*	*PLA2G12B*	*SLC6A17*	*USP5*
ATP1A3	*CUX2*	*GPI*	*MAPT*	*PLCB1*	*SLC7A6*	*USP9X*
ATP1B1	*CUZD1*	*GPM6A*	*MARK2*	*PLCH2*	*SLC8A1*	*UTRN*
ATP1B2	*CYFIP1*	*GPR115*	*MARK4*	*PLCL2*	*SLC8A2*	*VAMP2*
ATP2A2	*CYFIP2*	*GPR153*	*MAST1*	*PLD3*	*SLC9A3R1*	*VAPA*
ATP2B1	*CYTH1*	*GPR158*	*MAST2*	*PLEC*	*SLCO1B7*	*VAPB*

ATP2B2	*DAB2IP*	*GPR162*	*MAST4*	*PLEKHA6*	*SLITRK5*	*VCAN*
ATP2B3	*DAGLA*	*GPRC5B*	*MAZ*	*PLEKHH1*	*SLMO2*	*VCL*
ATP2B4	*DAGLB*	*GPRIN1*	*MBD5*	*PLP1*	*SMAD7*	*VCP*
ATP5A1	*DAPK1*	*GPRIN3*	*MBP*	*PLXNA1*	*SMAP2*	*VDAC1*
ATP5B	*DARC*	*GRAMD1B*	*MBTPS1*	*PLXNA2*	*SMARCA2*	*VDAC2*
ATP5D	*DARS*	*GRIA3*	*MDH1*	*PLXNA4*	*SMARCA4*	*VDAC3*
ATP5I	*DBC1*	*GRID2*	*MDH2*	*PLXNB1*	*SMARCC2*	*VIM*
ATP5O	*DBNL*	*GRIK3*	*MDM1*	*PLXND1*	*SMG1*	*VN1R4*
ATP6V0A1	*DCAF6*	*GRIK5*	*MED13*	*PMFBP1*	*SMPD3*	*VPS13A*
ATP6V0D1	*DCHS2*	*GRIN1*	*MED13L*	*PML*	*SMTNL1*	*VPS13C*
ATP6V1A	*DCLK1*	*GRIN2A*	*MED14*	*POGZ*	*SNAP25*	*VPS13D*
ATP6V1B2	*DCST1*	*GRIN2B*	*MED16*	*POLL*	*SNAP91*	*VPS35*
ATP6V1C1	*DCTN1*	*GRIN2D*	*MED29*	*POLR2A*	*SND1*	*VPS39*
ATP6V1D	*DCTN2*	*GRM4*	*MEF2D*	*POM121C*	*SNPH*	*VPS41*
ATP6V1E1	*DDAH1*	*GRM5*	*METTL14*	*POMT1*	*SNTA1*	*VPS52*
ATP6V1G2	*DDHD2*	*GSK3B*	*MFHAS1*	*PPARGC1A*	*SNTB2*	*VSNL1*
ATP6V1H	*DDN*	*GSN*	*MGAT5B*	*PPFIA1*	*SNX27*	*WARS*
ATP8A1	*DDX1*	*GSTM3*	*MIB1*	*PPFIA2*	*SNX4*	*WASF1*
ATP9A	*DDX10*	*GTF3C1*	*MICAL2*	*PPFIA3*	*SOAT1*	*WASF3*
ATXN1	*DDX24*	*GTF3C2*	*MIF*	*PPFIA4*	*SOBP*	*WASL*
AUTS2	*DDX3X*	*GTPBP3*	*MINK1*	*PPIA*	*SOLH*	*WDFY3*
AXL	*DENND5A*	*GTPBP5*	*MKI67*	*PPM1E*	*SON*	*WDR11*
B3GAT1	*DGCR2*	*GXYLT2*	*MKL2*	*PPP1CA*	*SORBS2*	*WDR13*
BAG3	*DGKZ*	*H2AFV*	*MLL2*	*PPP1CB*	*SORL1*	*WDR6*
BAI1	*DHX30*	*HADHA*	*MLL3*	*PPP1R12A*	*SORT1*	*WDR64*
BAI2	*DHX8*	*HADHB*	*MLL5*	*PPP1R14D*	*SPAG9*	*WDR7*
BAIAP2	*DICER1*	*HAPLN2*	*MLLT4*	*PPP1R7*	*SPARCL1*	*WIZ*
BAP1	*DIDO1*	*HAUS6*	*MMP24*	*PPP2R1A*	*SPATA22*	*WNK1*
BASP1	*DIP2A*	*HAUS8*	*MOG*	*PPP2R2C*	*SPATA5*	*WNK2*
BAZ2A	*DIP2B*	*HCFC1*	*MON2*	*PPP2R5B*	*SPDYC*	*WRN*
BCAN	*DIP2C*	*HCK*	*MOV10*	*PPP3CA*	*SPECC1*	*WWC1*
BCAT1	*DIRAS2*	*HCN2*	*MPG*	*PPP3CB*	*SPEG*	*XIRP2*
BCL9L	*DISP2*	*HDAC4*	*MPP2*	*PRDX1*	*SPEN*	*XKR6*
BCORL1	*DLAT*	*HDAC5*	*MPRIP*	*PRDX2*	*SPHKAP*	*XKR9*
BCR	*DLC1*	*HDLBP*	*MRC2*	*PRDX6*	*SPIB*	*XPO6*
BIN1	*DLD*	*HEATR5B*	*MSN*	*PREX1*	*SPIRE1*	*XPO7*
BIRC6	*DLG1*	*HECTD1*	*MTCH2*	*PREX2*	*SPRED1*	*XPOT*
BLVRB	*DLG2*	*HERC1*	*MTMR4*	*PRICKLE2*	*SPRN*	*XPR1*
BMPR2	*DLG3*	*HERC2*	*MTOR*	*PRKACB*	*SPTAN1*	*YAE1D1*
BPTF	*DLG4*	*HIPK1*	*MTSS1L*	*PRKAR1A*	*SPTB*	*YARS*
BRD4	*DLG5*	*HIPK2*	*MUC17*	*PRKAR2A*	*SPTBN1*	*YES1*
BRPF1	*DLGAP1*	*HIPK3*	*MUC5B*	*PRKAR2B*	*SPTBN2*	*YTHDC1*
BRSK1	*DLGAP2*	*HIST1H1E*	*MUC6*	*PRKCB*	*SPTBN4*	*YWHAB*
BRSK2	*DLGAP3*	*HIST1H2AG*	*MYADML2*	*PRKCE*	*SPTBN5*	*YWHAE*

BSN	*DLGAP4*	*HIST1H2BE*	*MYBPHL*	*PRKCG*	*SQLE*	*YWHAG*
BTNL2	*DLL4*	*HIST1H4I*	*MYCBP2*	*PRPF8*	*SRC*	*YWHAH*
BZRAP1	*DLST*	*HIVEP1*	*MYH10*	*PRPH*	*SRCIN1*	*YWHAQ*
C10orf35	*DMD*	*HIVEP2*	*MYH11*	*PRR12*	*SREBF2*	*YWHAZ*
C14orf28	*DMWD*	*HIVEP3*	*MYH14*	*PRRC2A*	*SRGAP3*	*ZBTB40*
C16orf62	*DMXL2*	*HK1*	*MYH7B*	*PRRC2C*	*SRI*	*ZBTB45*
C17orf49	*DNAH1*	*HLA-C*	*MYH9*	*PRSS38*	*SRRM2*	*ZC3H10*
C17orf97	*DNAH9*	*HMGCR*	*MYL6*	*PRUNE2*	*SSBP3*	*ZC3H4*
C1orf185	*DNAJA1*	*HMGXB3*	*MYL6B*	*PSAP*	*SSH2*	*ZC3H6*
C1QC	*DNAJB4*	*HNRNPUL1*	*MYO10*	*PSAPL1*	*SSPO*	*ZC3H7B*
C20orf197	*DNAJB6*	*HOMER1*	*MYO16*	*PSD*	*STAC2*	*ZCCHC14*
C2CD2L	*DNAJC13*	*HPS3*	*MYO18A*	*PSD3*	*STAG1*	*ZDHHC5*
C6orf222	*DNAJC6*	*HSD17B4*	*MYO18B*	*PSG2*	*STAP2*	*ZEB1*
C7orf60	*DNASE1L2*	*HSP90AA1*	*MYO1D*	*PSMD2*	*STARD13*	*ZEB2*
C9orf114	*DNM2*	*HSP90AB1*	*MYO5A*	*PSPC1*	*STIP1*	*ZER1*
C9orf172	*DNM3*	*HSPA12A*	*MYO6*	*PTAFR*	*STK10*	*ZFHX2*
CA10	*DOCK10*	*HSPA2*	*MYOF*	*PTCH1*	*STK25*	*ZFP106*
CA2	*DOCK3*	*HSPA4*	*MYOZ2*	*PTEN*	*STK38L*	*ZFR*
CA4	*DOCK4*	*HSPA4L*	*MYT1L*	*PTK2*	*STOM*	*ZFYVE1*
CABIN1	*DOCK9*	*HSPA5*	*NACAD*	*PTK2B*	*STOX2*	*ZHX3*
CACNA1A	*DOPEY1*	*HSPA6*	*NALCN*	*PTPN11*	*STRN4*	*ZMIZ1*
CACNA1B	*DOPEY2*	*HSPA8*	*NAP1L2*	*PTPN23*	*STT3B*	*ZMIZ2*
CACNA1C	*DOT1L*	*HSPA9*	*NAPA*	*PTPN3*	*STX1A*	*ZMYND11*
CACNA1D	*DPEP2*	*HSPB1*	*NAPB*	*PTPN5*	*STX1B*	*ZNF133*
CACNA1E	*DPM1*	*HSPD1*	*NAPG*	*PTPRD*	*STXBP1*	*ZNF229*
CACNA1F	*DPP8*	*HSPH1*	*NAT8L*	*PTPRF*	*STXBP3*	*ZNF326*
CACNA1G	*DPYD*	*HTATSF1*	*NAV1*	*PTPRG*	*STXBP5*	*ZNF407*
CACNA1H	*DPYSL2*	*HTR2A*	*NAV2*	*PTPRJ*	*SUCLA2*	*ZNF426*
CACNA1I	*DPYSL3*	*HTR7*	*NAV3*	*PTPRM*	*SULF2*	*ZNF445*
CACNA1S	*DSCAM*	*HTT*	*NBEA*	*PTPRN2*	*SULT1B1*	*ZNF480*
CACNA2D1	*DSCAML1*	*HUWE1*	*NCAM1*	*PTPRS*	*SV2A*	*ZNF518B*
CACNA2D2	*DST*	*IDH2*	*NCAM2*	*PTPRT*	*SV2B*	*ZNF530*
CACNA2D3	*DSTN*	*IDS*	*NCAN*	*PTPRZ1*	*SYMPK*	*ZNF532*
CACNA2D4	*DTNA*	*IER2*	*NCDN*	*PUF60*	*SYN1*	*ZNF536*
CACNB1	*DTX1*	*IFNAR2*	*NCKAP1*	*PUM1*	*SYNE1*	*ZNF551*
CACNB2	*DUOXA1*	*IFT140*	*NCKIPSD*	*PUM2*	*SYNGAP1*	*ZNF565*
CACNB3	*DUOXA2*	*IFT81*	*NCOA1*	*PURA*	*SYNGR1*	*ZNF607*
CACNB4	*DUSP15*	*IGFL2*	*NCOA2*	*PYCARD*	*SYNGR3*	*ZNF721*
ZNF77	*ZNFX1*	*ZP4*	*ZYG11B*			

Table S13. Prioritized genes that have a score of 2 and above

Gene	Entrez Id	CFG[a]	CNV[b]	DE[c]	GWAS[d]	Linkage[e]	Sherlock[f]	PASCAL[g]	Score[h]
DRD2	**1813**	**1**	**0**	**0**	**1**	**1**	**0**	**1**	**4**
DOC2A	**8448**	**0**	**1**	**1**	**1**	**0**	**0**	**1**	**4**
PGBD1	**84547**	**0**	**0**	**1**	**1**	**1**	**0**	**1**	**4**
SNAP91	**9892**	**0**	**0**	**1**	**1**	**0**	**0**	**1**	**3**
MAGI2	**9863**	**0**	**1**	**1**	**0**	**1**	**0**	**0**	**3**
HIST1H2BJ	**8970**	**0**	**0**	**0**	**1**	**1**	**0**	**1**	**3**
TCF4	**6925**	**1**	**0**	**0**	**1**	**1**	**0**	**0**	**3**
RTN4R	**65078**	**0**	**1**	**1**	**0**	**1**	**0**	**0**	**3**
SLC1A2	**6506**	**1**	**0**	**1**	**0**	**1**	**0**	**0**	**3**
ZSCAN31	**64288**	**0**	**0**	**0**	**1**	**0**	**1**	**1**	**3**
SRR	**63826**	**0**	**0**	**0**	**1**	**1**	**0**	**1**	**3**
BDNF	**627**	**1**	**0**	**1**	**0**	**1**	**0**	**0**	**3**
ZSWIM6	**57688**	**0**	**0**	**1**	**1**	**0**	**0**	**1**	**3**
GLT8D1	**55830**	**0**	**0**	**0**	**1**	**0**	**1**	**1**	**3**
GNB1L	**54584**	**1**	**1**	**0**	**0**	**1**	**0**	**0**	**3**
NR4A2	**4929**	**1**	**0**	**1**	**0**	**1**	**0**	**0**	**3**
NRGN	**4900**	**0**	**0**	**0**	**1**	**1**	**0**	**1**	**3**
GOLGA6L5	**374650**	**0**	**0**	**1**	**1**	**0**	**0**	**1**	**3**
ITIH4	**3700**	**0**	**0**	**1**	**1**	**0**	**0**	**1**	**3**
HSPD1	**3329**	**0**	**0**	**1**	**1**	**0**	**0**	**1**	**3**
GRIA1	**2890**	**1**	**0**	**0**	**1**	**1**	**0**	**0**	**3**
DISC1	**27185**	**1**	**1**	**0**	**0**	**1**	**0**	**0**	**3**
EGR1	**1958**	**0**	**0**	**1**	**1**	**0**	**0**	**1**	**3**
COMT	**1312**	**1**	**1**	**0**	**0**	**1**	**0**	**0**	**3**
DNAJC19	**131118**	**0**	**0**	**1**	**1**	**0**	**0**	**1**	**3**
CNR1	**1268**	**1**	**0**	**1**	**0**	**1**	**0**	**0**	**3**
TRMT61A	**115708**	**0**	**0**	**1**	**1**	**0**	**0**	**1**	**3**
TNFRSF13C	**115650**	**0**	**0**	**1**	**1**	**0**	**0**	**1**	**3**
PRSS16	**10279**	**0**	**0**	**0**	**1**	**1**	**0**	**1**	**3**
DGCR2	9993	0	1	0	0	1	0	0	2
DOPEY2	9980	0	0	1	0	0	0	1	2
HS3ST3B1	9953	0	1	1	0	0	0	0	2
TRANK1	9881	0	0	0	1	0	0	1	2
ATG13	9776	0	0	0	1	0	0	1	2
ZSCAN12	9753	0	0	0	1	0	0	1	2
KDM4A	9682	0	0	0	1	0	0	1	2
PLCH2	9651	0	0	0	1	0	0	1	2
RGS6	9628	0	0	1	1	0	0	0	2
BAG5	9529	0	0	0	1	0	0	1	2
TAOK2	9344	0	0	0	1	0	0	1	2
SNAP29	9342	0	1	0	0	1	0	0	2
MARS2	92935	0	0	0	1	0	0	1	2

ZNF804A	91752	0	0	0	1	1	0	0	2
SMDT1	91689	0	0	0	1	0	0	1	2
ATPAF2	91647	0	0	0	1	0	0	1	2
DGKI	9162	0	0	0	1	0	0	1	2
C12orf65	91574	0	0	0	1	0	0	1	2
ANKRD44	91526	0	0	0	1	0	0	1	2
CREB3L1	90993	0	0	0	1	0	0	1	2
ESAM	90952	0	0	0	1	0	0	1	2
KALRN	8997	1	0	1	0	0	0	0	2
CACNA1I	8911	0	0	0	1	0	0	1	2
DGKZ	8525	0	0	0	1	0	0	1	2
USMG5	84833	0	0	1	1	0	0	0	2
PLCD4	84812	0	0	1	0	0	1	0	2
HIRIP3	8479	0	0	0	1	0	0	1	2
DTNBP1	84062	1	0	0	0	1	0	0	2
TSSK6	83983	0	0	0	1	0	0	1	2
IMMP2L	83943	0	0	0	1	0	0	1	2
MAD1L1	8379	0	0	0	1	0	0	1	2
FAM57B	83723	0	0	0	1	0	0	1	2
HIST1H3A	8350	0	0	1	0	0	0	1	2
HIST1H2BL	8340	0	0	0	1	0	0	1	2
CDC45	8318	0	1	1	0	0	0	0	2
DGCR6	8214	0	1	0	0	1	0	0	2
ANP32E	81611	0	0	1	1	0	0	0	2
ST8SIA2	8128	0	0	1	0	1	0	0	2
CDK2AP1	8099	0	0	0	1	0	0	1	2
FXR1	8087	0	0	0	1	0	0	1	2
PBX4	80714	0	0	0	1	0	0	1	2
ZKSCAN3	80317	0	0	0	1	0	0	1	2
EFHD1	80303	0	0	0	1	0	0	1	2
COQ10B	80219	0	0	0	1	0	0	1	2
THOC7	80145	0	0	0	1	0	0	1	2
ACTR5	79913	0	0	0	1	0	0	1	2
ZNF408	79797	0	0	0	1	0	0	1	2
MSANTD2	79684	0	0	0	1	0	0	1	2
OGFOD2	79676	0	0	0	1	0	0	1	2
C2orf47	79568	0	0	0	1	0	0	1	2
ZFYVE21	79038	0	0	0	1	0	0	1	2
CENPM	79019	0	0	0	1	0	0	1	2
GID4	79018	0	0	0	1	0	0	1	2
SCG2	7857	0	0	1	0	1	0	0	2
CACNB2	783	0	0	0	1	0	0	1	2
DDR1	780	0	0	1	0	1	0	0	2
CACNA1C	775	0	0	0	1	0	0	1	2

ZKSCAN8	7745	0	0	1	0	0	0	1	2
XRCC3	7517	0	0	0	1	0	0	1	2
VRK2	7444	0	0	0	1	0	0	1	2
UFD1L	7353	0	1	0	0	1	0	0	2
CLDN5	7122	0	1	0	0	1	0	0	2
TCF20	6942	0	0	0	1	0	0	1	2
TBX6	6911	0	1	0	1	0	0	0	2
TBX1	6899	0	1	0	0	1	0	0	2
SYN2	6854	1	0	0	0	1	0	0	2
NEK4	6787	0	0	0	1	0	0	1	2
SRPK2	6733	0	0	0	1	0	0	1	2
SREBF1	6720	0	0	0	1	0	0	1	2
SNRPN	6638	0	1	1	0	0	0	0	2
SLC25A1	6576	0	1	1	0	0	0	0	2
SLC6A3	6531	0	0	1	0	1	0	0	2
SLC1A3	6507	0	1	1	0	0	0	0	2
NT5DC2	64943	0	0	0	1	0	0	1	2
BCL11B	64919	0	0	0	1	0	0	1	2
EFTUD1P1	648809	0	0	0	1	0	0	1	2
CSMD1	64478	0	1	0	1	0	0	0	2
GOLGA6L4	643707	0	0	0	1	0	0	1	2
BTBD18	643376	0	0	0	1	0	0	1	2
SLC39A8	64116	0	0	1	1	0	0	0	2
NPAS3	64067	0	0	1	0	1	0	0	2
RRAS	6237	0	0	0	1	0	0	1	2
BCL9	607	0	1	0	1	0	0	0	2
RGS4	5999	1	0	0	0	1	0	0	2
RANBP1	5902	0	1	0	0	1	0	0	2
SCAF1	58506	0	0	0	1	0	0	1	2
PTPRN2	5799	0	1	1	0	0	0	0	2
SUGP1	57794	0	0	0	1	0	0	1	2
PITPNM2	57605	0	0	0	1	0	0	1	2
PRR12	57479	0	0	0	1	0	0	1	2
AS3MT	57412	0	0	0	1	0	0	1	2
PTGIS	5740	0	0	0	1	0	0	1	2
RELN	5649	1	0	0	0	1	0	0	2
PRODH	5625	0	1	0	0	1	0	0	2
PCDHA8	56140	0	0	1	1	0	0	0	2
SEPT3	55964	0	0	0	1	0	0	1	2
PRKG1	5592	0	1	1	0	0	0	0	2
PRKCA	5578	1	0	1	0	0	0	0	2
TSR1	55720	0	0	0	1	0	0	1	2
PRKAB2	5565	0	1	1	0	0	0	0	2
AMBRA1	55626	0	0	0	1	0	0	1	2

GALNT10	55568	0	0	0	1	0	0	1	2
SLC38A7	55238	0	0	0	1	0	0	1	2
SBNO1	55206	0	0	0	1	0	0	1	2
PBRM1	55193	0	0	0	1	0	0	1	2
MSL2	55167	0	0	0	1	0	0	1	2
FANCL	55120	0	0	0	1	0	0	1	2
ZSCAN2	54993	0	0	0	1	0	0	1	2
C4orf27	54969	0	0	1	1	0	0	0	2
WHSC1L1	54904	0	0	1	1	0	0	0	2
WBP1L	54838	0	0	0	1	0	0	1	2
GATAD2A	54815	0	0	0	1	0	0	1	2
AHI1	54806	0	0	1	0	1	0	0	2
CNNM2	54805	0	0	0	1	0	0	1	2
BTN2A3P	54718	0	0	1	0	0	0	1	2
RRN3	54700	0	1	1	0	0	0	0	2
TM6SF2	53345	0	0	0	1	0	0	1	2
PLCL1	5334	0	0	0	1	0	0	1	2
PLCB2	5330	0	0	0	1	0	0	1	2
RAPGEF6	51735	0	1	0	0	1	0	0	2
PDE4B	5142	1	0	0	0	1	0	0	2
ARL6IP4	51329	0	0	0	1	0	0	1	2
PCCB	5096	0	0	0	1	0	0	1	2
PAM	5066	0	0	1	0	1	0	0	2
SLC45A1	50651	0	0	0	1	0	0	1	2
PAK2	5062	0	1	1	0	0	0	0	2
FURIN	5045	0	0	0	1	0	0	1	2
NOTCH4	4855	0	0	0	1	1	0	0	2
NOTCH2	4853	0	0	1	0	1	0	0	2
NPY	4852	0	0	1	0	1	0	0	2
NOS1	4842	0	0	1	0	1	0	0	2
NEK1	4750	0	0	1	1	0	0	0	2
NDUFV2	4729	1	0	1	0	0	0	0	2
NDUFA6	4700	0	0	0	1	0	0	1	2
ATF4	468	0	0	1	0	0	0	1	2
NAGA	4668	0	0	0	1	0	0	1	2
NAB2	4665	0	0	1	1	0	0	0	2
SMIM4	440957	0	0	0	1	0	0	1	2
LOC440354	440354	0	0	1	0	0	0	1	2
MMP16	4325	0	0	0	1	0	0	1	2
ARVCF	421	0	1	0	0	1	0	0	2
MEF2C	4208	0	0	1	1	0	0	0	2
MDK	4192	0	0	0	1	0	0	1	2
MAOA	4128	0	0	1	0	1	0	0	2
MAN2A1	4124	0	0	0	1	0	0	1	2

MAN2A2	4122	0	0	0	1	0	0	1	2
MIR137	406928	0	0	0	1	0	0	1	2
HAPLN4	404037	0	0	0	1	0	0	1	2
MIR137HG	400765	0	0	0	1	0	0	1	2
SNX19	399979	0	0	0	1	0	0	1	2
LCAT	3931	0	0	1	1	0	0	0	2
ARHGAP1	392	0	0	0	1	0	0	1	2
MUSTN1	389125	0	0	0	1	0	0	1	2
C2orf82	389084	0	0	0	1	0	0	1	2
UBE2Q2P1	388165	0	0	0	1	0	0	1	2
SETD8	387893	0	0	0	1	0	0	1	2
ZKSCAN4	387032	0	0	0	1	0	0	1	2
KLC1	3831	0	0	0	1	0	0	1	2
KCNN3	3782	0	0	1	0	1	0	0	2
KCNJ13	3769	0	0	0	1	0	0	1	2
TMEM110	375346	0	0	0	1	0	0	1	2
YJEFN3	374887	0	0	0	1	0	0	1	2
KCNB1	3745	0	0	0	1	0	0	1	2
ITIH3	3699	0	0	0	1	0	0	1	2
ITIH1	3697	0	0	0	1	0	0	1	2
IREB2	3658	0	0	0	1	0	0	1	2
IL1B	3553	0	0	1	0	1	0	0	2
HCN1	348980	0	0	0	1	0	0	1	2
FTCDNL1	348751	0	0	0	1	0	0	1	2
SOX2-OT	347689	0	0	0	1	0	0	1	2
LINC00634	339674	0	0	0	1	0	0	1	2
HTR2A	3356	1	0	0	0	1	0	0	2
HSPA9	3313	0	0	0	1	0	0	1	2
HINT1	3094	1	0	0	0	1	0	0	2
NRG1	3084	1	0	0	0	1	0	0	2
SERPIND1	3053	0	1	1	0	0	0	0	2
ZDHHC8	29801	0	1	0	0	1	0	0	2
GSTM1	2944	0	1	0	0	1	0	0	2
GRM3	2913	0	0	0	1	1	0	0	2
GRIN2B	2904	1	0	0	0	1	0	0	2
SPCS1	28972	0	0	0	1	0	0	1	2
GRIA4	2893	1	0	0	0	1	0	0	2
FAM86B3P	286042	0	0	1	0	0	0	1	2
INO80E	283899	0	0	0	1	0	0	1	2
HARBI1	283254	0	0	0	1	0	0	1	2
C11orf31	280636	0	0	0	1	0	0	1	2
TRMT2A	27037	0	1	0	0	1	0	0	2
GJA8	2703	0	1	0	0	1	0	0	2
KCNV1	27012	0	0	1	1	0	0	0	2

MDGA1	266727	0	0	1	0	1	0	0	2
SEZ6L2	26470	0	0	0	1	0	0	1	2
GNL3	26354	0	0	0	1	0	0	1	2
GIGYF2	26058	0	0	0	1	0	0	1	2
PPP1R16B	26051	0	0	0	1	0	0	1	2
CNTNAP2	26047	0	1	1	0	0	0	0	2
SPATS2L	26010	0	0	1	0	0	0	1	2
ZDHHC5	25921	0	0	0	1	0	0	1	2
BRP44	25874	0	0	1	1	0	0	0	2
NGEF	25791	0	0	0	1	0	0	1	2
GAD1	2571	1	0	0	0	1	0	0	2
GABRB3	2562	1	1	0	0	0	0	0	2
ASPHD1	253982	0	0	0	1	0	0	1	2
KCTD13	253980	0	0	0	1	0	0	1	2
PLA2G15	23659	0	0	0	1	0	0	1	2
VSIG2	23584	0	0	0	1	0	0	1	2
ABCB9	23457	0	0	0	1	0	0	1	2
SF3B1	23451	0	0	0	1	0	0	1	2
MAU2	23383	0	0	0	1	0	0	1	2
PPP1R13B	23368	0	0	0	1	0	0	1	2
SATB2	23314	0	0	0	1	0	0	1	2
ACSL6	23305	0	0	1	0	1	0	0	2
SMG6	23293	0	0	0	1	0	0	1	2
CNOT1	23019	0	0	0	1	0	0	1	2
IGSF9B	22997	0	0	0	1	0	0	1	2
NT5C2	22978	0	0	0	1	0	0	1	2
FES	2242	0	0	0	1	0	0	1	2
NKAPL	222698	0	0	0	1	0	0	1	2
ZSCAN23	222696	0	0	0	1	0	0	1	2
MED19	219541	0	0	0	1	0	0	1	2
YPEL4	219539	0	0	0	1	0	0	1	2
F2	2147	0	0	0	1	0	0	1	2
ETF1	2107	0	0	0	1	0	0	1	2
ERBB4	2066	0	1	0	0	1	0	0	2
C2orf69	205327	0	0	0	1	0	0	1	2
EP300	2033	0	0	1	1	0	0	0	2
TSNARE1	203062	0	0	0	1	0	0	1	2
CMYA5	202333	0	0	1	0	1	0	0	2
RILPL2	196383	0	0	0	1	0	0	1	2
DRG2	1819	0	0	0	1	0	0	1	2
DRD4	1815	0	0	1	0	1	0	0	2
DRD1	1812	0	0	1	0	1	0	0	2
DPYD	1806	0	0	0	1	0	0	1	2
WBP2NL	164684	0	0	0	1	0	0	1	2

CYP17A1	1586	0	0	0	1	0	0	1	2
ADRBK2	157	0	0	1	1	0	0	0	2
CYP2D6	1565	0	0	0	1	0	0	1	2
FAM109B	150368	0	0	0	1	0	0	1	2
CILP2	148113	0	0	0	1	0	0	1	2
TOM1L2	146691	0	0	0	1	0	0	1	2
C16orf92	146378	0	0	0	1	0	0	1	2
SLC32A1	140679	0	0	0	1	0	0	1	2
CR1L	1379	0	0	1	1	0	0	0	2
RFTN2	130132	0	0	0	1	0	0	1	2
TYW5	129450	0	0	0	1	0	0	1	2
UHMK1	127933	0	0	1	0	1	0	0	2
TMEM219	124446	0	0	0	1	0	0	1	2
AGPHD1	123688	0	0	0	1	0	0	1	2
CLU	1191	0	0	1	1	0	0	0	2
SFXN2	118980	0	0	1	1	0	0	0	2
CKB	1152	0	0	0	1	0	0	1	2
CHRNB4	1143	0	0	0	1	0	0	1	2
CHRNA7	1139	0	1	0	0	1	0	0	2
CHRNA5	1138	0	0	0	1	0	0	1	2
CHRNA3	1136	0	0	0	1	0	0	1	2
CHRM4	1132	0	0	0	1	0	0	1	2
NXPH4	11247	0	0	1	1	0	0	0	2
NISCH	11188	0	0	1	1	0	0	0	2
CHI3L1	1116	0	0	1	0	1	0	0	2
CHGB	1114	0	0	1	0	1	0	0	2
CLP1	10978	0	0	0	1	0	0	1	2
CPLX2	10814	1	0	0	0	1	0	0	2
MPHOSPH9	10198	0	0	0	1	0	0	1	2
DPYD-AS1	100873932	0	0	0	1	0	0	1	2
HSPE1-MOB4	100529241	0	0	0	1	0	0	1	2
TMEM110-MUS									
TN1	100526772	0	0	0	1	0	0	1	2
LINC00933	100506874	0	0	0	1	0	0	1	2
MIR4304	100422931	0	0	0	1	0	0	1	2
MIR3160-1	100422827	0	0	1	0	0	0	1	2
MIR1281	100302237	0	0	0	1	0	0	1	2
LOC100216546	100216546	0	0	0	1	0	0	1	2
NDUFA6-AS1	100132273	0	0	1	0	0	0	1	2
ANKRD63	100131244	0	0	0	1	0	0	1	2
LOC378135	-	0	0	0	1	0	1	0	2

[a] CFG: gene identified by the Convergent functional genomics analysis; a total of 42 genes was adopted from the study of Ayalew et al.[14]

[b] CNV: genes affected by copy number variant; CNV was set as "1" if this gene was reported in at

least two of these studies.[21-52]

[c] DE: genes expressed differentially in schizophrenia; DE was set as "1" if this gene was differentially expressed in one of GSE53987, GSE12649, GSE21138, GSE35978, GSE62191, GSE62191 these datasets at ranking of top 1% most differentially expressed or in the report of.[53]

[d] GWAS: genes identified by genome wide association study; GWAS was set as “1” if this gene was identified in one of these studies.[16, 54-68]

[e] Linkage: genes identified by linkage and association study; Linkage was set as “1” if this gene was identified in one of these studies.[7, 9, 10]

[f] Sherlock: genes identified by Sherlock integrative analysis; a total of 12 genes was got from the study of Luo et al.[69]

[g] PASCAL: genes identified by Pathway scoring algorithm, we use gene-based test PASCAL[17] to calculate PGC2 data and if the *P*-value of a gene is less than 10-6, then we set its PASCAL to “1”.

[h] Score: number of “1” about this gene.

Table S14. Seven of the 29 top prioritized genes in SZDB database are drug targets

Gene	Drug (s)	Source (s)
BDNF	DESIPRAMINE	PharmGKB
CNR1	ANANDAMIDE,CANNABINOL,CBD CANNABIS DERIVATIVE,CP55940,DRONABINOL,HU-210,NABILONE, RIMONABANT,SURINABANT,TARANABANT	TdgClinicalTrial, TTD, TEND, GuideToPharmacologyInteractions, TdgClinicalTrial
COMT	2-METHOXYESTRADIOL,BUPROPION,ENTACAPONE,LEVODOPA,METHYLDOPA,NIALAMIDE,NICOTINE,RISPERIDONE,S-ADENOSYLMETHIONINE,TOLCAPONE	TEND, ChEMBL, TdgClinicalTrial, PharmGKB, TTD, DrugBank
DRD2	7-OH-DPAT,ABAPERIDONE HYDROCHLORIDE,AC1L2SUW,ACEPROMAZINE,ACETOPHENAZINE,ALIZAPRIDE,AMANTADINE,AMISULPRIDE,AMOXAPINE,AMPHETAMINE,ANIRACETAM,APLINDORE FUMARATE,APOMORPHINE,ARIPIPRAZOLE,ASENAPINE, ASENAPINE MALEATE,BENZQUINAMIDE,BICIFADINE,BIFEPRUNOX, BL-1020,BLONANSERIN,BRASOFENSINE,BREXPIPRAZOLE,BROMOCRIPTINE,BROMOPRIDE,BROMPERIDOL,BUPROPION,BUSPIRONE,CABERGOLINE,CAFFEINE,CARFENAZINE,CARIPRAZINE,CHLORPROMAZINE,CHLORPROTHIXENE,CINNARIZINE,CLOZAPINE,DESIPRAMINE,DEXPRAMIPEXOLE,DOMPERIDONE,DOPAMINE,DOXEPIN,DROPERIDOL,ERGOLOID MESYLATE,ETICLOPRIDE,ETILEVODOPA,FLUPENTIXOL, FLUPHENAZINE,FLUPHENAZINE DECANOATE,FLUPHENAZINE ENANTHATE,FLUSPIRILENE,HALOPERIDOL,ILOPERIDONE,IMIPRAMINE,ITOPRIDE,KETAMINE,LAMECTACIN,LEVODOPA,LISURIDE,LOXAPINE,LURASIDONE,MAPROTILINE,MELEVODOPA,MEMANTINE,MESORIDAZINE,MESORIDAZINE BESYLATE,METOCLOPRAMIDE,MIANSERIN,MINAPRINE,MIRTAZAPINE,MOLINDONE,MOLINDONE HYDROCHLORIDE,NAFADOTRIDE,NEMONAPRIDE,NICOTINE,NOMIFENSINE,NORTRIPTYLINE,OCAPERIDONE,OLANZAPINE,PALIPERIDONE,PALIPERIDONE PALMITATE,PARDOPRUNOX,PERGOLIDE,PEROSPIRONE, PERPHENAZINE,PIMOZIDE,PIPOTIAZINE,PIRIBEDIL,PRAMIPEXOLE,PROCHLORPERAZINE,PROCHLORPERAZINE MALEATE,PROMAZINE,PROMETHAZINE,PROPIOMAZIN	TEND, ChEMBL, GuideToPharmacologyInteractions, TdgClinicalTrial, PharmGKB, TTD, DrugBank

	E,QUETIAPINE,QUINAGOLIDE HYDROCHLORIDE,QUINELORANE,QUINPIROLE,RACLOPRIDE,REMOXIPRIDE,RISPERIDONE,ROPINIROLE,ROTIGOTINE,ROXINDOLE,SERTINDOLE,SIBENADET,SULPIRIDE,SUMANIROLE,TERGURIDE,TETRABENAZINE,THIETHYLPERAZINE,THIETHYLPERAZINE MALATE,THIOPROPERAZINE,THIORIDAZINE,THIOTHIXENE,THIOTHIXENE HYDROCHLORIDE,TRIFLUOPERAZINE,TRIFLUPROMAZINE,TRIFLUPROMAZINE HYDROCHLORIDE,TRIMIPRAMINE,TRIMIPRAMINE MALEATE,VILAZODONE,YOHIMBINE,ZIPRASIDONE,ZOTEPINE,ZUCLOPENTHIXOL	
GRIA1	AMPA,ANIRACETAM,ATPO,CYCLOTHIAZIDE,DESFLURANE,ENFLURANE,ETHANOL,FARAMPATOR,ISOFLURANE,L-GLUTAMIC ACID,METHOXYFLURANE,PERAMPANEL,PHENOBARBITAL,PIRACETAM,SEVOFLURANE,TALAMPANEL,TEZAMPANEL	TEND, GuideToPharmacologyInteractions, TdgClinicalTrial, TTD, DrugBank
SLC1A2	DL-TBOA,L-GLUTAMIC ACID,WAY-213613	GuideToPharmacologyInteractions, DrugBank
SRR	L-SERINE,PYRIDOXAL PHOSPHATE	DrugBank

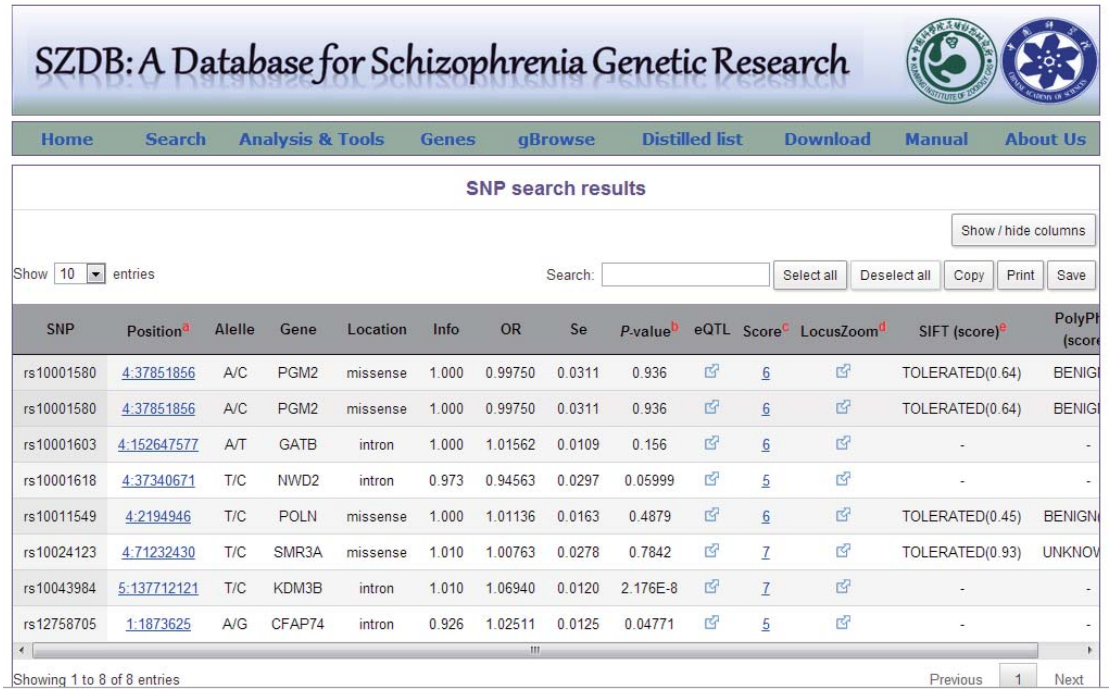

Figure S1. SNP query report. For SNP query (batch query is allowed), SZDB will generate a detailed report, including SNP id, chromosomal coordinates, alleles, genes that the query SNP located, SNP type (i.e., if the SNP is located in intron or coding region), association significance (P value) with schizophrenia (data from the PGC2), eQTL analysis and SNP functional annotation (from RegulomeDB). If the SNP is located in coding region, SIFT and PolyPhen-2 annotations are also generated.

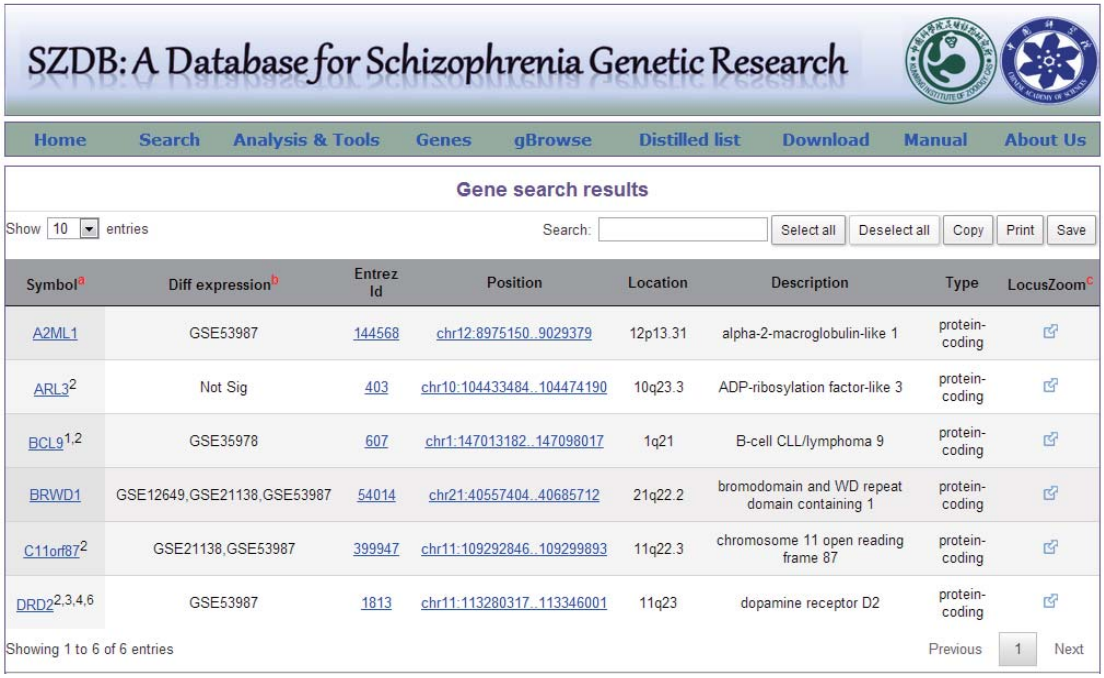

Figure S2. Gene query report. For gene query (batch query is allowed), SZDB will generate a detailed report, which contain following information: (1) If the query genes is differentially expressed in schizophrenia cases and controls; (2) The Entrez id of the query gene; (3) The genomic coordinate of the query gene; (4) The association significance (P value) between SNPs in the query gene and schizophrenia (from the PGC2).

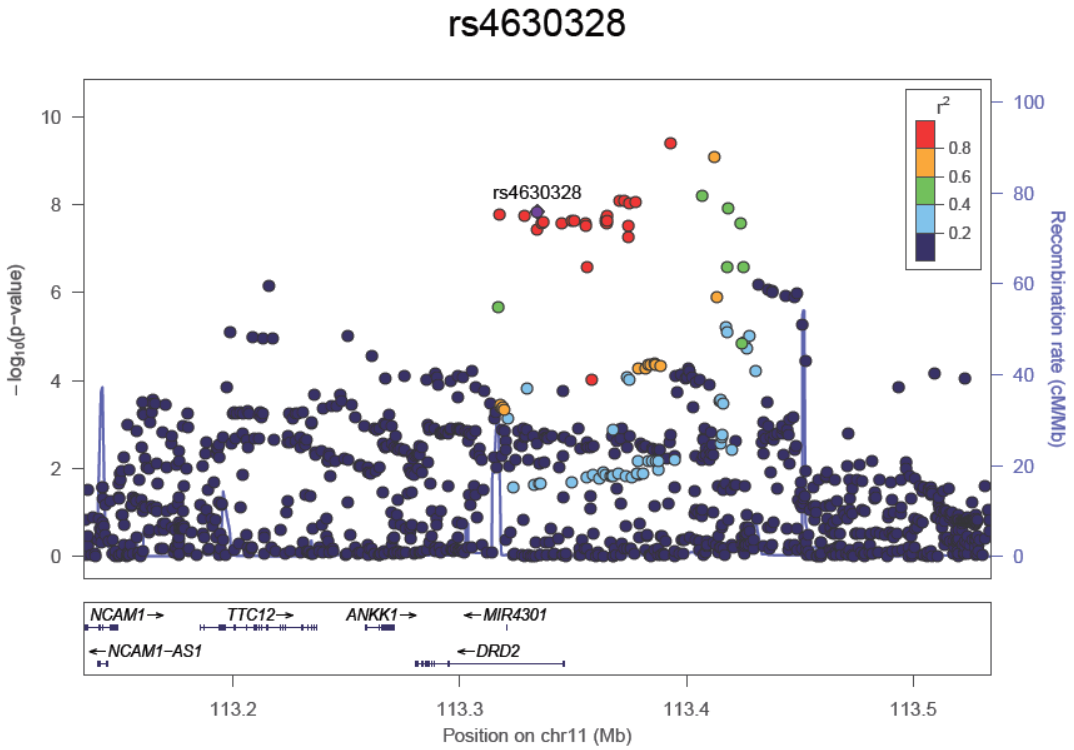

Figure S3. The Locuszoom view shows the association significance between SNPs in query gene and schizophrenia (summary statistics, i.e., *P* values are taken from the PGC2).

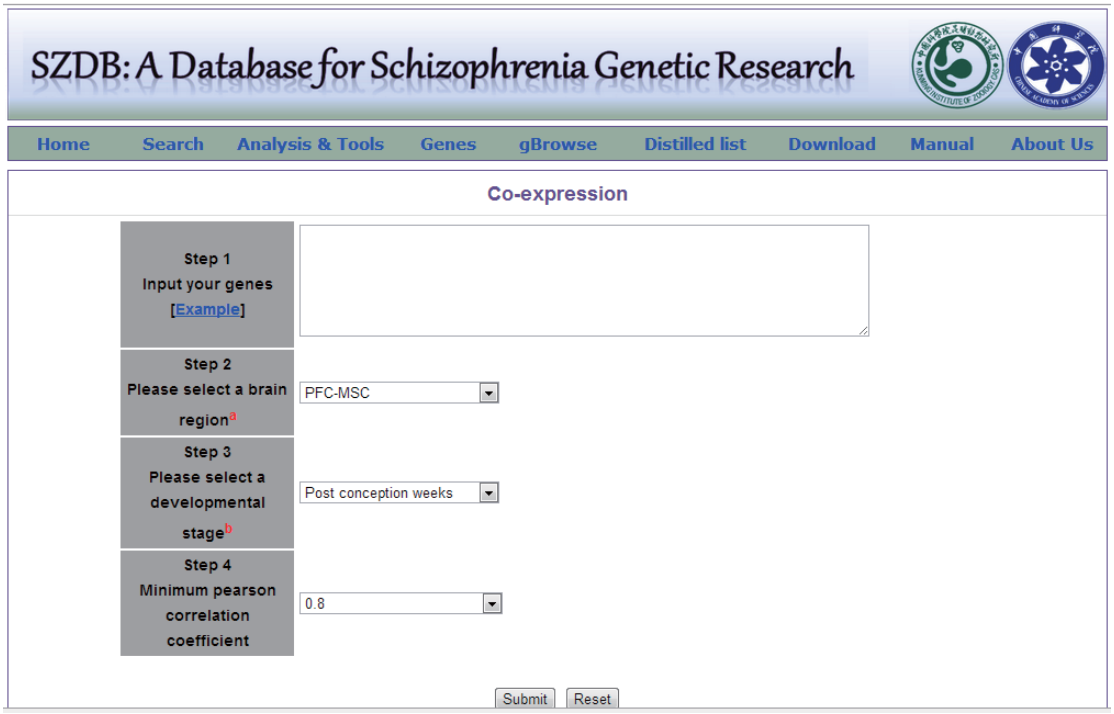

Figure S4. SZDB can perform co-expression analysis using brain expression data. If the users identified a novel schizophrenia risk gene and they want to know the other genes that are significantly co-expressed with this novel gene, SZDB can finish this job easily so the users do not need to download related data and perform additional analyses.

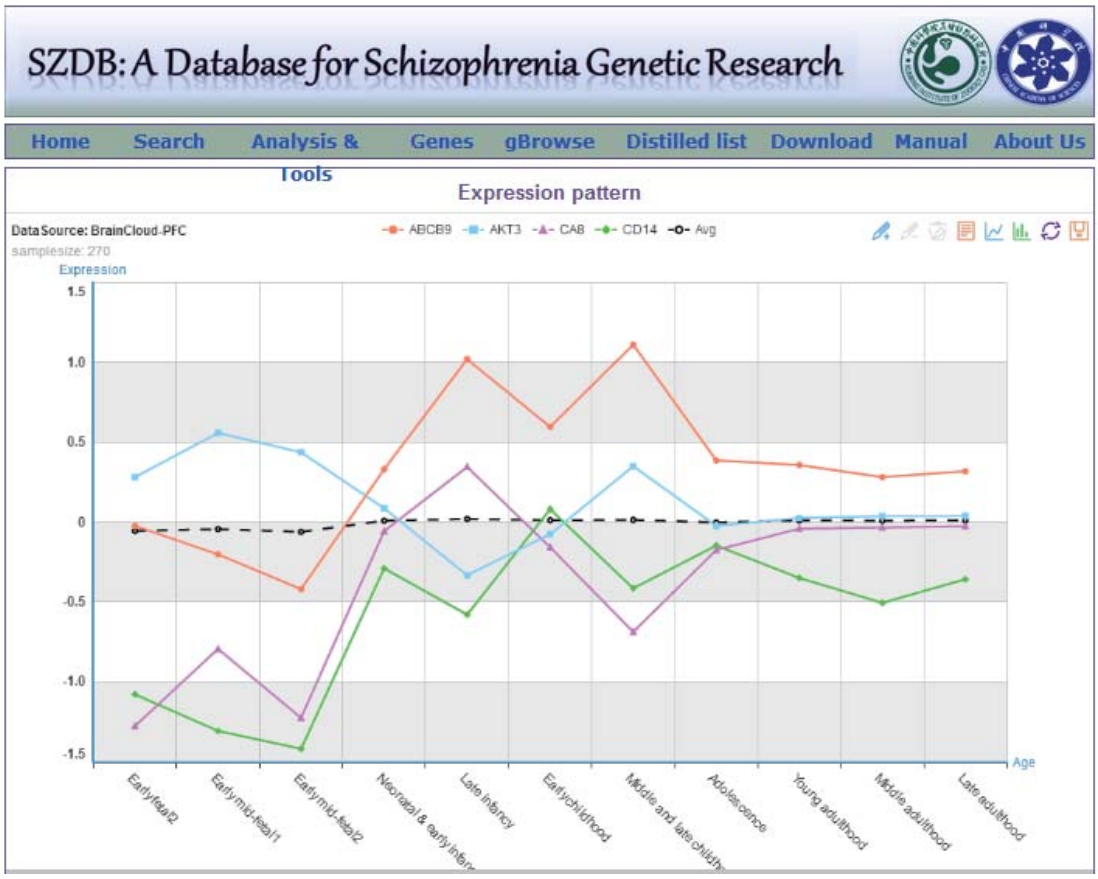

Figure S5. Spatio-temporal expression pattern analysis can be performed at SZDB. If the users identified novel schizophrenia risk genes and they want to know if these genes are expressed in human brain, SZDB can finish this job so the users do not need to download related data and perform additional analyses.

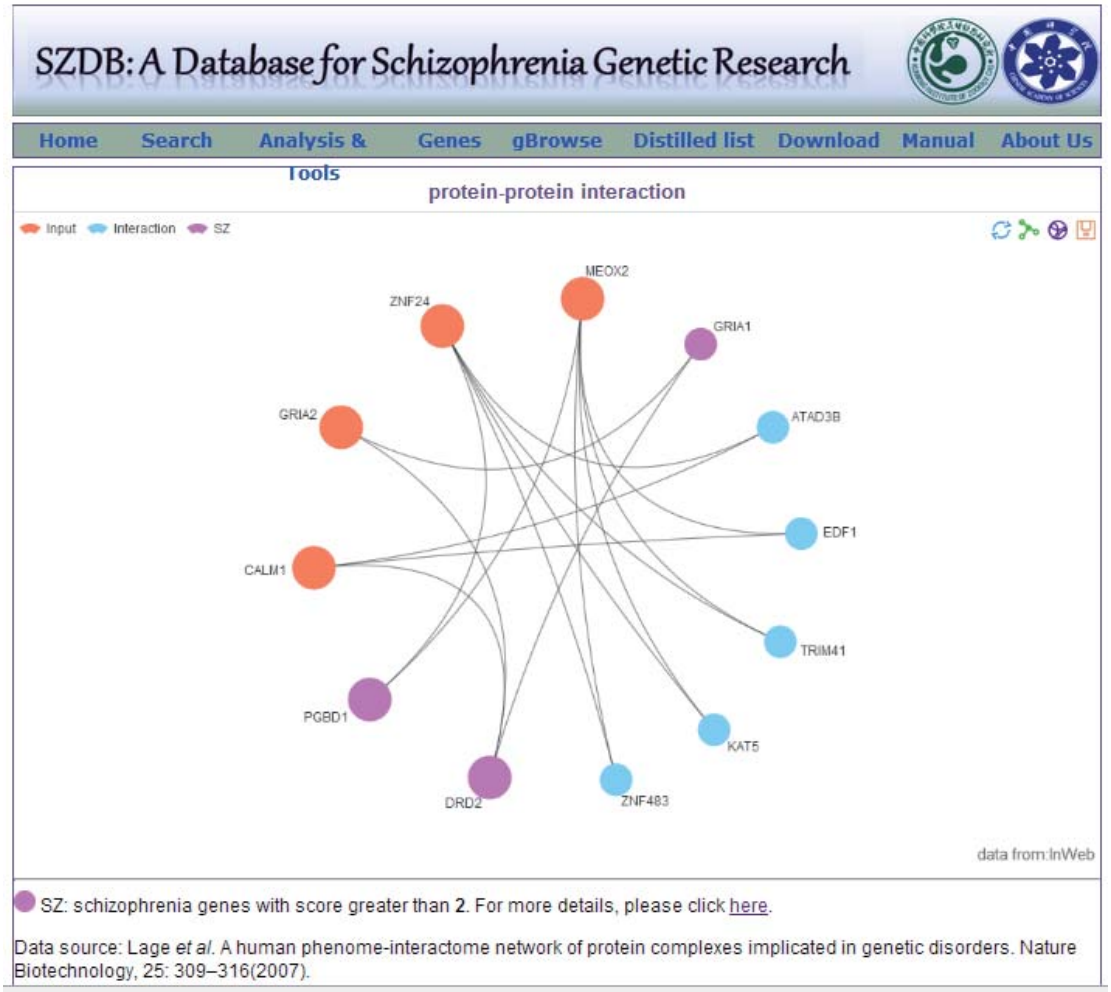

Figure S6. Users can perform protein-protein interaction analysis at the SZDB.

Figure S7. Users can perform eQTL analysis at the SZDB. Brain expression and genetic data are taken from Myers et al.[70]

References

1 Huang DW, Sherman BT, Tan Q, Collins JR, Alvord WG, Roayaei J *et al.* The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol* 2007; **8**: R183.

2 Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D *et al.* DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res* 2007; **35**: W169-175.

3 BrainSpan. BrainSpan: Atlas of the Developing Human Brain [Internet]. *http://wwwbrainspanorg* 2013.

4 Willsey AJ, Sanders SJ, Li M, Dong S, Tebbenkamp AT, Muhle RA *et al.* Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell* 2013; **155**: 997-1007.

5 Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC bioinformatics* 2008; **9**: 559.

6 Langfelder P, Zhang B, Horvath S. Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics* 2008; **24**: 719-720.

7 Lewis CM, Levinson DF, Wise LH, DeLisi LE, Straub RE, Hovatta I *et al.* Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: Schizophrenia. *Am J Hum Genet* 2003; **73**: 34-48.

8 Sun J, Jia P, Fanous AH, Webb BT, van den Oord EJ, Chen X *et al.* A multi-dimensional evidence-based candidate gene prioritization approach for complex diseases-schizophrenia as a case. *Bioinformatics* 2009; **25**: 2595-6602.

9 Ng MY, Levinson DF, Faraone SV, Suarez BK, DeLisi LE, Arinami T *et al.* Meta-analysis of 32 genome-wide linkage studies of schizophrenia. *Mol Psychiatry* 2009; **14**: 774-785.

10 Allen NC, Bagade S, McQueen MB, Ioannidis JP, Kavvoura FK, Khoury MJ *et al.* Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database. *Nat Genet* 2008; **40**: 827-834.

11 Jia P, Sun J, Guo AY, Zhao Z. SZGR: a comprehensive schizophrenia gene resource. *Mol Psychiatry* 2010; **15**: 453-462.

12 Ross CA, Margolis RL, Reading SA, Pletnikov M, Coyle JT. Neurobiology of schizophrenia. *Neuron* 2006; **52**: 139-153.

13 Sun J, Kuo PH, Riley BP, Kendler KS, Zhao Z. Candidate genes for schizophrenia: a survey of association studies and gene ranking. *Am J Med Genet B Neuropsychiatr Genet* 2008; **147B**: 1173-1181.

14 Ayalew M, Le-Niculescu H, Levey DF, Jain N, Changala B, Patel SD *et al.* Convergent functional genomics of schizophrenia: from comprehensive understanding to genetic risk prediction. *Mol Psychiatry* 2012; **17**: 887-905.

15 Ayalew M, Le-Niculescu H,Levey DF, Jain N, Changala B, Patel SD, Winiger E, Breier A, Shekhar A, Amdur R, Koller D, Nurnberger JI, Corvin A, Geyer M, Tsuang MT, Salomon D, Schork NJ, Fanous AH, O'Donovan MC, Niculescu AB. Convergent functional genomics of schizophrenia: from comprehensive understanding to genetic risk prediction. *Molecular Psychiatry advance online publication, May 15, 2012; doi:101038/mp201237*.

16 Luo X, Huang L, Han L, Luo Z, Hu F, Tieu R *et al.* Systematic prioritization and integrative

analysis of copy number variations in schizophrenia reveal key schizophrenia susceptibility genes. *Schizophr Bull* 2014; **40**: 1285-1299.

17 Lamparter D, Marbach D, Rueedi R, Kutalik Z, Bergmann S. Fast and Rigorous Computation of Gene and Pathway Scores from SNP-Based Summary Statistics. *PLoS Comput Biol* 2016; **12**: e1004714.

18 Schizophrenia Working Group of the Psychiatric Genomics Consortium*. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014; **511**: 421-427.

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

20 Purcell SM, Moran JL, Fromer M, Ruderfer D, Solovieff N, Roussos P *et al.* A polygenic burden of rare disruptive mutations in schizophrenia. *Nature* 2014; **506**: 185-190.

21 Stefansson H, Rujescu D, Cichon S, Pietilainen OP, Ingason A, Steinberg S *et al.* Large recurrent microdeletions associated with schizophrenia. *Nature* 2008; **455**: 232-236.

22 International Schizophrenia C. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* 2008; **455**: 237-241.

23 McCarthy SE, Makarov V, Kirov G, Addington AM, McClellan J, Yoon S *et al.* Microduplications of 16p11.2 are associated with schizophrenia. *Nature genetics* 2009; **41**: 1223-1227.

24 Xu B, Roos JL, Levy S, van Rensburg EJ, Gogos JA, Karayiorgou M. Strong association of de novo copy number mutations with sporadic schizophrenia. *Nature genetics* 2008; **40**: 880-885.

25 Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM *et al.* Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* 2008; **320**: 539-543.

26 Rodriguez-Santiago B, Brunet A, Sobrino B, Serra-Juhe C, Flores R, Armengol L *et al.* Association of common copy number variants at the glutathione S-transferase genes and rare novel genomic changes with schizophrenia. *Molecular psychiatry* 2010; **15**: 1023-1033.

27 Ingason A, Rujescu D, Cichon S, Sigurdsson E, Sigmundsson T, Pietilainen OP *et al.* Copy number variations of chromosome 16p13.1 region associated with schizophrenia. *Molecular psychiatry* 2011; **16**: 17-25.

28 Ikeda M, Aleksic B, Kirov G, Kinoshita Y, Yamanouchi Y, Kitajima T *et al.* Copy number variation in schizophrenia in the Japanese population. *Biological psychiatry* 2010; **67**: 283-286.

29 Tarabeux J, Champagne N, Brustein E, Hamdan FF, Gauthier J, Lapointe M *et al.* De novo truncating mutation in Kinesin 17 associated with schizophrenia. *Biological psychiatry* 2010; **68**: 649-656.

30 Levinson DF, Duan J, Oh S, Wang K, Sanders AR, Shi J *et al.* Copy number variants in schizophrenia: confirmation of five previous findings and new evidence for 3q29 microdeletions and VIPR2 duplications. *The American journal of psychiatry* 2011; **168**: 302-316.

31 Vacic V, McCarthy S, Malhotra D, Murray F, Chou HH, Peoples A *et al.* Duplications of the neuropeptide receptor gene VIPR2 confer significant risk for schizophrenia. *Nature* 2011; **471**: 499-503.

32 Kirov G, Gumus D, Chen W, Norton N, Georgieva L, Sari M *et al.* Comparative genome hybridization suggests a role for NRXN1 and APBA2 in schizophrenia. *Human molecular*

genetics 2008; **17**: 458-465.

33 Wilson GM, Flibotte S, Chopra V, Melnyk BL, Honer WG, Holt RA. DNA copy-number analysis in bipolar disorder and schizophrenia reveals aberrations in genes involved in glutamate signaling. *Human molecular genetics* 2006; **15**: 743-749.

34 Vrijenhoek T, Buizer-Voskamp JE, van der Stelt I, Strengman E, Genetic R, Outcome in Psychosis C *et al.* Recurrent CNVs disrupt three candidate genes in schizophrenia patients. *American journal of human genetics* 2008; **83**: 504-510.

35 Need AC, Ge D, Weale ME, Maia J, Feng S, Heinzen EL *et al.* A genome-wide investigation of SNPs and CNVs in schizophrenia. *PLoS genetics* 2009; **5**: e1000373.

36 Xu B, Woodroffe A, Rodriguez-Murillo L, Roos JL, van Rensburg EJ, Abecasis GR *et al.* Elucidating the genetic architecture of familial schizophrenia using rare copy number variant and linkage scans. *Proceedings of the National Academy of Sciences of the United States of America* 2009; **106**: 16746-16751.

37 Glessner JT, Reilly MP, Kim CE, Takahashi N, Albano A, Hou C *et al.* Strong synaptic transmission impact by copy number variations in schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America* 2010; **107**: 10584-10589.

38 Amar S, Ovadia O, Maier W, Ebstein R, Belmaker RH, Mishmar D *et al.* Copy number variation of the SELENBP1 gene in schizophrenia. *Behavioral and brain functions : BBF* 2010; **6**: 40.

39 Mulle JG, Dodd AF, McGrath JA, Wolyniec PS, Mitchell AA, Shetty AC *et al.* Microdeletions of 3q29 confer high risk for schizophrenia. *American journal of human genetics* 2010; **87**: 229-236.

40 Moreno-De-Luca D, Consortium S, Mulle JG, Simons Simplex Collection Genetics C, Kaminsky EB, Sanders SJ *et al.* Deletion 17q12 is a recurrent copy number variant that confers high risk of autism and schizophrenia. *American journal of human genetics* 2010; **87**: 618-630.

41 Kirov G, Pocklington AJ, Holmans P, Ivanov D, Ikeda M, Ruderfer D *et al.* De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. *Molecular psychiatry* 2012; **17**: 142-153.

42 Malhotra D, McCarthy S, Michaelson JJ, Vacic V, Burdick KE, Yoon S *et al.* High frequencies of de novo CNVs in bipolar disorder and schizophrenia. *Neuron* 2011; **72**: 951-963.

43 Buizer-Voskamp JE, Muntjewerff JW, Genetic R, Outcome in Psychosis Consortium M, Strengman E, Sabatti C *et al.* Genome-wide analysis shows increased frequency of copy number variation deletions in Dutch schizophrenia patients. *Biological psychiatry* 2011; **70**: 655-662.

44 Kirov G, Grozeva D, Norton N, Ivanov D, Mantripragada KK, Holmans P *et al.* Support for the involvement of large copy number variants in the pathogenesis of schizophrenia. *Human molecular genetics* 2009; **18**: 1497-1503.

45 Guha S, Rees E, Darvasi A, Ivanov D, Ikeda M, Bergen SE *et al.* Implication of a rare deletion at distal 16p11.2 in schizophrenia. *JAMA psychiatry* 2013; **70**: 253-260.

46 Liao HM, Chao YL, Huang AL, Cheng MC, Chen YJ, Lee KF *et al.* Identification and characterization of three inherited genomic copy number variations associated with familial schizophrenia. *Schizophrenia research* 2012; **139**: 229-236.

47 Karlsson R, Graae L, Lekman M, Wang D, Favis R, Axelsson T *et al.* MAGI1 copy number

variation in bipolar affective disorder and schizophrenia. *Biological psychiatry* 2012; **71**: 922-930.

48 Stewart LR, Hall AL, Kang SH, Shaw CA, Beaudet AL. High frequency of known copy number abnormalities and maternal duplication 15q11-q13 in patients with combined schizophrenia and epilepsy. *BMC medical genetics* 2011; **12**: 154.

49 Melhem N, Middleton F, McFadden K, Klei L, Faraone SV, Vinogradov S *et al.* Copy number variants for schizophrenia and related psychotic disorders in Oceanic Palau: risk and transmission in extended pedigrees. *Biological psychiatry* 2011; **70**: 1115-1121.

50 Van Den Bossche MJ, Strazisar M, Cammaerts S, Liekens AM, Vandeweyer G, Depreeuw V *et al.* Identification of rare copy number variants in high burden schizophrenia families. *American journal of medical genetics Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics* 2013; **162B**: 273-282.

51 Gauthier J, Siddiqui TJ, Huashan P, Yokomaku D, Hamdan FF, Champagne N *et al.* Truncating mutations in NRXN2 and NRXN1 in autism spectrum disorders and schizophrenia. *Human genetics* 2011; **130**: 563-573.

52 Rujescu D, Ingason A, Cichon S, Pietilainen OP, Barnes MR, Toulopoulou T *et al.* Disruption of the neurexin 1 gene is associated with schizophrenia. *Human molecular genetics* 2009; **18**: 988-996.

53 Fillman SG, Cloonan N, Catts VS, Miller LC, Wong J, McCrossin T *et al.* Increased inflammatory markers identified in the dorsolateral prefrontal cortex of individuals with schizophrenia. *Molecular psychiatry* 2013; **18**: 206-214.

54 Schizophrenia Psychiatric Genome-Wide Association Study C. Genome-wide association study identifies five new schizophrenia loci. *Nature genetics* 2011; **43**: 969-976.

55 Rietschel M, Mattheisen M, Degenhardt F, Genetic R, Outcome in P, Muhleisen TW *et al.* Association between genetic variation in a region on chromosome 11 and schizophrenia in large samples from Europe. *Molecular psychiatry* 2012; **17**: 906-917.

56 Steinberg S, de Jong S, Irish Schizophrenia Genomics C, Andreassen OA, Werge T, Borglum AD *et al.* Common variants at VRK2 and TCF4 conferring risk of schizophrenia. *Human molecular genetics* 2011; **20**: 4076-4081.

57 Hamshere ML, Walters JT, Smith R, Richards AL, Green E, Grozeva D *et al.* Genome-wide significant associations in schizophrenia to ITIH3/4, CACNA1C and SDCCAG8, and extensive replication of associations reported by the Schizophrenia PGC. *Molecular psychiatry* 2013; **18**: 708-712.

58 Ikeda M, Aleksic B, Yamada K, Iwayama-Shigeno Y, Matsuo K, Numata S *et al.* Genetic evidence for association between NOTCH4 and schizophrenia supported by a GWAS follow-up study in a Japanese population. *Molecular psychiatry* 2013; **18**: 636-638.

59 Shi Y, Li Z, Xu Q, Wang T, Li T, Shen J *et al.* Common variants on 8p12 and 1q24.2 confer risk of schizophrenia. *Nature genetics* 2011; **43**: 1224-1227.

60 Yue WH, Wang HF, Sun LD, Tang FL, Liu ZH, Zhang HX *et al.* Genome-wide association study identifies a susceptibility locus for schizophrenia in Han Chinese at 11p11.2. *Nature genetics* 2011; **43**: 1228-1231.

61 Li J, Zhou G, Ji W, Feng G, Zhao Q, Liu J *et al.* Common variants in the BCL9 gene conferring risk of schizophrenia. *Archives of general psychiatry* 2011; **68**: 232-240.

62 International Schizophrenia C, Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC

et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748-752.

63 Stefansson H, Ophoff RA, Steinberg S, Andreassen OA, Cichon S, Rujescu D *et al.* Common variants conferring risk of schizophrenia. *Nature* 2009; **460**: 744-747.

64 O'Donovan MC, Craddock N, Norton N, Williams H, Peirce T, Moskvina V *et al.* Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nature genetics* 2008; **40**: 1053-1055.

65 Ripke S, O'Dushlaine C, Chambert K, Moran JL, Kahler AK, Akterin S *et al.* Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nature genetics* 2013; **45**: 1150-1159.

66 Wong EH, So HC, Li M, Wang Q, Butler AW, Paul B *et al.* Common variants on Xq28 conferring risk of schizophrenia in Han Chinese. *Schizophrenia bulletin* 2014; **40**: 777-786.

67 Lencz T, Guha S, Liu C, Rosenfeld J, Mukherjee S, DeRosse P *et al.* Genome-wide association study implicates NDST3 in schizophrenia and bipolar disorder. *Nature communications* 2013; **4**: 2739.

68 Consortium SWGotPG. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014; **511**: 421-427.

69 Luo XJ, Mattheisen M, Li M, Huang L, Rietschel M, Borglum AD *et al.* Systematic Integration of Brain eQTL and GWAS Identifies ZNF323 as a Novel Schizophrenia Risk Gene and Suggests Recent Positive Selection Based on Compensatory Advantage on Pulmonary Function. *Schizophrenia bulletin* 2015; **41**: 1294-1308.

70 Myers AJ, Gibbs JR, Webster JA, Rohrer K, Zhao A, Marlowe L *et al.* A survey of genetic human cortical gene expression. *Nat Genet* 2007; **39**: 1494-1499.