Integrated Post-GWAS Analysis Sheds New Light on the Disease Mechanisms of Schizophrenia

Jhih-Rong Lin, Ying Cai, Quanwei Zhang, Wen Zhang, Rubén Nogales-Cadenas, and Zhengdong D. Zhang[1]
Department of Genetics, Albert Einstein College of Medicine, Bronx, New York 10461

ABSTRACT Schizophrenia is a severe mental disorder with a large genetic component. Recent genome-wide association studies (GWAS) have identified many schizophrenia-associated common variants. For most of the reported associations, however, the underlying biological mechanisms are not clear. The critical first step for their elucidation is to identify the most likely disease genes as the source of the association signals. Here, we describe a general computational framework of post-GWAS analysis for complex disease gene prioritization. We identify 132 putative schizophrenia risk genes in 76 risk regions spanning 120 schizophrenia-associated common variants, 78 of which have not been recognized as schizophrenia disease genes by previous GWAS. Even more significantly, 29 of them are outside the risk regions, likely under regulation of transcriptional regulatory elements contained therein. These putative schizophrenia risk genes are transcriptionally active in both brain and the immune system, and highly enriched among cellular pathways, consistent with leading pathophysiological hypotheses about the pathogenesis of schizophrenia. With their involvement in distinct biological processes, these putative schizophrenia risk genes, with different association strengths, show distinctive temporal expression patterns, and play specific biological roles during brain development.

KEYWORDS schizophrenia; GWAS; disease risk gene prioritization

SCHIZOPHRENIA is a debilitating brain disorder with a worldwide prevalence of ~1% that results in substantial morbidity and mortality. It is characterized by constellations of symptoms such as hallucinations, delusions, and cognitive impairments. Most cases of schizophrenia start during adolescence and early adulthood, and often have a lifelong course. Converging evidence indicates that schizophrenia results from a disruption in brain development (du Bois and Huang 2007) caused by genetic predisposition and environmental factors, the latter of which include prenatal infection, maternal nutrition, and stress. Schizophrenia is a highly heritable disease, with an estimated heritability between 64 and 81% (Sullivan *et al.* 2003; Lichtenstein *et al.* 2009), confirming the major role of genetic factors in contributing to disease risk. Therefore, further dissection of the genetic underpinnings of schizophrenia is crucial toward advancing our understanding of its pathogenesis.

The genetic basis of schizophrenia involves complex interactions among risk variants across an allelic frequency spectrum. While no Mendelian inheritance patterns have been observed for schizophrenia risk variants (Giusti-Rodriguez and Sullivan 2013), accumulating evidence indicates that the polygenic component of risk is substantial (International Schizophrenia Consortium *et al.* 2009). Rare copy number variants (CNVs) have shown relatively high penetrance for schizophrenia: the majority of 11 known risk CNVs with genome-wide significance for schizophrenia association have minor allele frequencies (MAFs) <0.1%, and odds ratios (ORs) between 2 and 60 (Rees 2015). In addition, significant progress has been made recently, using large-scale exome-sequencing and genome-wide association studies (GWAS), on the role of risk variants with subtle effects. Enrichment of disruptive rare (MAF <0.1%) single nucleotide variants (SNVs) of small effect sizes (OR = 1.12), as well as enrichment of nonsynonymous *de novo* SNVs, was found in several gene sets associated with synaptic function (Fromer *et al.* 2014; Purcell *et al.* 2014). Previous studies suggest that common

doi: 10.1534/genetics.116.187195
Manuscript received January 15, 2016; accepted for publication September 30, 2016; published Early Online October 17, 2016.
Supplemental material is available online at http://www.genetics.org/cgi/content/full/genetics.116.187195/DC1.
[1]Corresponding author: Department of Genetics, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461. E-mail: zhengdong.zhang@einstein.yu.edu

single nucleotide polymorphisms (SNPs) associated with schizophrenia generally have a small effect size (OR < 1.2), but, collectively, thousands of independent SNPs could account for up to 50% of variance in schizophrenia liability (Ripke *et al.* 2013). In particular, a recent large-scale meta-analysis based on past GWAS identified 108 schizophrenia risk regions with genome-wide significance (Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014), and thus further confirmed the important contribution that common variants make to the genetic risk of schizophrenia. To date, over 20 GWAS studies have been conducted in schizophrenia, providing valuable data for downstream analysis.

Identification of genes that confer risk for developing schizophrenia is crucial to providing insight into the underlying disease mechanisms, and for identifying new drug targets. One of the best-known schizophrenia genes encodes the dopamine receptor D2 (DRD2). The fact that it can be used as a drug target to treat schizophrenia supports a major etiological hypothesis that abnormal brain signaling involving dopamine is a substantial factor in the pathophysiology of schizophrenia (Di Forti *et al.* 2007). In addition, genes implicated in schizophrenia by previous studies of common or rare variants (Fromer *et al.* 2014; Purcell *et al.* 2014; Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014) include genes involved in glutamatergic neurotransmission (*GRM3*, *GRIN2A*, *SRR*, *GRIA1*, and *SLC38A7*), calcium channel signaling (*CACNA1C*, *CACNB2*, *CAMKK2*, *CACNA1I*, *NRGN*, and *RIMS1*), and synaptic plasticity such as *N*-methy-D-aspartate receptor (NMDAR) and activity-regulated cytoskeleton-associated scaffold protein (ARC). However, these findings are mostly limited to the level of gene set enrichment due to difficulty in pinpointing risk genes. In contrast to exome sequencing studies, in which risk genes are directly implicated by risk exonic variants, GWAS can identify only risk regions instead of risk genes. This intrinsic limitation of GWAS cannot be resolved by increasing the sample size. Thus, in order to investigate the biological effects of common variants, new methodologies are required to track down risk genes responsible for the GWAS signals found in schizophrenia (Need and Goldstein 2014).

The challenge of pinpointing risk genes in disease-associated risk regions lies in several aspects. Most risk regions cover and implicate multiple genes, which, without other information, makes it exceedingly difficult to determine the true risk gene(s) within them. Furthermore, risk genes may reside outside risk regions, and be affected through regulatory elements. In this study, we propose a framework to tackle this challenge. To cover risk genes that reside outside of risk regions, we incorporated gene regulatory information to include candidate genes outside risk regions. In addition, we developed a computational method to score schizophrenia candidate genes based on Gene Ontology (GO) annotations and functional network characteristics of a group of known (and well-accepted) schizophrenia genes. We prioritized 132 schizophrenia risk gene candidates as putative schizophrenia risk genes in risk regions that we constructed from previous GWAS studies. Subsequent multiple integrated functional analyses of these putative susceptibility genes provide us with novel and deeper biological insight into the genetic architecture, enriched pathways, gene expression profiles, and penetrance of schizophrenia.

Materials and Methods

The overall strategy of our approach is depicted in Figure 1.

Identification of genomic risk regions for schizophrenia

We collected SNPs/indels from the PGC study (Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014), and additional SNPs from the GWAS catalog (Hindorff *et al.* 2015) that were identified to be associated with schizophrenia ($P < 1 \times 10^{-5}$). The final set included 128 SNPs/indels from the PGC study, and 137 SNPs from the GWAS catalog. Using VCFtools (Danecek *et al.* 2011), and the 1KG reference panel (1000 Genomes Project Consortium *et al.* 2012), we calculated the linkage disequilibrium (LD) between each schizophrenia variant, and every 1KG variant in its 400-kb neighborhood. The neighboring SNPs with $r^2 > 0.5$ define the LD block indexed by the enclosed schizophrenia variant. Finally, we combined overlapping or close (within 250 kb) LD blocks to form genomic risk regions for schizophrenia.

Identification of schizophrenia risk gene candidates

After pinpointing the schizophrenia risk regions, we identified schizophrenia risk gene candidates that are linked to these risk regions. Based on the genomic distance, a schizophrenia risk gene candidate is either proximal or distal to the schizophrenia risk regions. Proximal candidate genes are candidate genes inside or closest to risk regions, while distal candidate genes are candidate genes outside, and not closest to, risk regions (if there are genes inside risk regions, they are closest to risk regions). The proximal candidates were identified with the same approach as used in the PGC meta-analysis (Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014): they are genes overlapping risk regions after extending them by 20 kb on both ends, or the closest genes to risk regions within 500 kb, when they contain no genes. In addition, in our analysis we also included possible distal risk genes by incorporating transcriptional regulatory interactions between expression quantitative trait loci (eQTL) or transcriptional regulatory elements (TREs) and their target genes. Both ENCODE and FANTOM5 provided enhancer-promoter connections, based on the correlation between their DNase hypersensitivity in different cell types, and between their expression activity, respectively. We used such enhancer–promoter connections to connect transcriptional regulatory elements to genes. Thus, the distal candidates are genes that are neither directly covered by, nor closest to, the risk regions (within 500 kb), but are likely regulated by eQTL or TREs within them. We collected eQTL, DHS, and

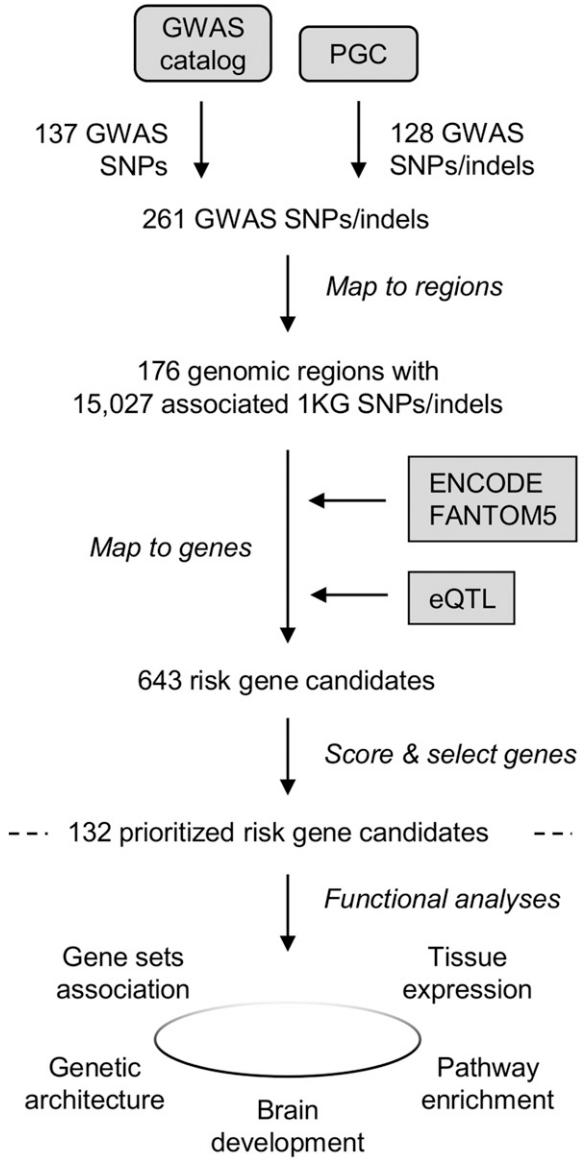


Figure 1 The flowchart of the integrated post-GWAS study of schizophrenia. The study consisted of two major parts: prioritization of schizophrenia risk gene candidates and subsequent functional analyses.

enhancers, each set with their target genes from GTEx (Analysis Pilot V3), ENCODE (Thurman *et al.* 2012), and FANTOM5 (Andersson *et al.* 2014), respectively. To minimize inclusion of irrelevant distal genes, we only considered eQTL, DHS, or enhancers that are in the risk regions, and also contain at least one SNP or indel in strong LD ($r^2 > 0.5$) with the schizophrenia GWAS SNPs or indels.

Scoring schizophrenia risk gene candidates

We have developed a statistical method to score the disease-relatedness of schizophrenia risk gene candidates, with predictive features extracted from gene networks, and annotation based on a set of training schizophrenia genes (Figure S1) (Supplemental Material, File S1). We used 56 training genes in our analysis, including (1) 38 manually curated schizophrenia genes with strong evidence (aka "core genes") (Jia *et al.* 2010), (2) eight schizophrenia susceptibility genes cataloged in the Online Mendelian Inheritance in Man (OMIM) database (McKusick 2007), (3) six well-accepted schizophrenia genes from recent genetics studies (Hall *et al.* 2015; Kotlar *et al.* 2015), and (4) four schizophrenia genes with solid support from other sources (Canetta *et al.* 2014; Nawa *et al.* 2014; Bossu *et al.* 2015; Lv *et al.* 2015) (Table S1). The gene network that we used is the functional linkage network (Linghu *et al.* 2009), in which the functional association (the edge) between a pair of genes was predicted based on 16 genomic features.

The predictive features are either network features or annotation features. Network features are the frequent combinations of the neighbors of schizophrenia genes in the functional linkage network (Linghu *et al.* 2009), while annotation features are the frequent combinations of GO terms associated with schizophrenia genes. We extracted those frequent combinations of the network neighbors or GO terms by using the frequent item set mining algorithm (Tan *et al.* 2006). Network features characterize schizophrenia genes, indirectly, by a combination of genes that schizophrenia genes are usually functionally associated with, while those functionally associated genes can be any genes, not restricted to schizophrenia genes. In other words, network features characterize the context of functionally associated genes in which schizophrenia genes are usually enriched. Annotation features characterize schizophrenia genes directly in terms of GO terms. The two features characterize schizophrenia genes from different angles, and are complementary according to our evaluation (shown in the section "*Evaluation of schizophrenia gene candidate scoring*"). The final score integrates two scores, based on GO terms and functional linkage network (Figure S2), respectively.

Our evaluation showed that a final-score cutoff set at 80 could achieve a high prediction precision (Figure S3A). Meanwhile, the majority of the training genes, and the majority of genes with strong literature support for connection with schizophrenia, both have scores higher than 80 (Figure S3B and Table S2). Moreover, because our analysis showed that it is unlikely to observe scores higher than 80 from the same set of schizophrenia candidate genes by training with random or irrelevant genes ($P = 0$) (Figure S4, A and B), we set the high-scoring cutoff at 80. The majority of the prioritized genes have scores higher than 160 (Figure S4C).

Evaluation of schizophrenia gene scoring

We used two complementary approaches—binary classification tests and Wilcoxon rank sum tests—to evaluate our scoring method in discerning schizophrenia genes. The former directly assessed how well our scoring method distinguished schizophrenia genes from nonschizophrenia genes, while the latter compared the scores between schizophrenia genes and nonschizophrenia genes. For binary classification tests, the 56 training genes were used as the only positive testing gene set, while 56 genes randomly selected from the "background"

gene set (Figure S5A) as the negative testing gene set in different binary classification tests. For Wilcoxon rank sum tests, we prepared an "enriched" gene set, which is composed of genes implicated in schizophrenia by rare mutations other than the 56 schizophrenia genes (Figure S5A), and compared both the schizophrenia gene set and the "enriched" gene set with the "background" gene set.

Gene sets association analysis

We compiled the following three gene sets, and compared our putative schizophrenia risk genes with each of them using Fisher's exact test of association.

1. We collected from the Mouse Genome Informatics (MGI) database (as on April 23, 2015 at http://www.informatics.jax.org,) a list of 3765 genes whose knock-outs in mouse models generated phenotypes of nervous systems and neurological behaviors.
2. Using text-mining techniques, we compiled a list of 54 genes with strong literature support for connection with schizophrenia (Table S2).
3. We assembled a list of 1401 genes that have been shown in previous studies to be differential expressed between schizophrenia patients and normal controls.

Pathway and GO term enrichment analysis

We used GeneCoDis3 (Tabas-Madrid *et al.* 2012) to identify KEGG and Panther pathways enriched among schizophrenia risk genes. Briefly, putative schizophrenia risk genes that we identified and Ensemble human genes were used as the input and the reference gene sets, respectively. Pathway annotations from KEGG and Panther were searched and compared in both the input and the reference gene sets to find pathways significantly enriched in the putative schizophrenia risk genes. To measure the significance of enrichment, the hypergeometric distribution was used to calculate *P*-values. Then, the false discovery rate was calculated for multiple test correction. Biological pathways with significant corrected *P*-values are candidates for involvement in the pathogenesis of schizophrenia. We used GO::TermFinder (Boyle *et al.* 2004) to analyze the enrichment of GO terms in the putative schizophrenia risk genes. To avoid potential confounding effects from the functional linkage network, we excluded associations between GO terms and genes based on Electronic Annotation (evidence code = IEA) from our enrichment analysis, and thus ensured that all associations between GO terms and genes were assigned manually by curators. *P*-values were adjusted for multiple tests using the Bonferroni method.

Tissue gene expression analysis

To examine the expression profiles of the putative schizophrenia risk genes in different tissues, we used the Gene Enrichment Profiler (http://xavierlab2.mgh.harvard.edu/EnrichmentProfiler/) (Benita *et al.* 2010), which catalogs normalized expression values of ~12,000 genes across 126 primary human tissues. To investigate gene-tissue expression specificity, we grouped the putative schizophrenia risk genes into different clusters according to their different expression patterns across tissues using the Euclidean distance, and the Ward's clustering method (Legendre 2014).

Data availability

The authors state that all the source of data necessary for reproducing the results are presented within the article. Strains are available upon request.

Results

Schizophrenia-associated common variants and genomic risk regions

We collected 261 schizophrenia-associated common variants (SNPs and indels) from 25 GWAS of the disease (see *Materials and Methods*). With a few exceptions, the associated variants reported by each study are within a range of effect size similar to one another (Figure S6). They represent at least 60 genomic loci harboring schizophrenia-associated variants that have been replicated in multiple independent studies (Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014; Hindorff *et al.* 2015). Interestingly, the vast majority of these loci act independently of known risk factors, promising the discovery of hitherto unknown mechanisms influencing risk. These variants are distributed throughout the human genome, with local clustering (Figure 2A). Using neighboring 1000 Genomes Project (1KG) variants that are in high ($r^2 > 0.5$) LD with schizophrenia GWAS variants, we identified 176 genomic schizophrenia-risk regions. After analyzing how schizophrenia GWAS variants, protein-coding genes (GENCODE v19) (Harrow *et al.* 2012), and TREs (we used ENCODE enhancers) (Encode Project Consortium 2012) are distributed together in human genome (Figure 2B), we found that many human genomic regions, such as the 1q43, 6p21, and 18q21 loci, are enriched with both schizophrenia GWAS SNPs and enhancers. These schizophrenia risk regions are either gene-rich or gene-poor.

Evaluation of schizophrenia gene candidate scoring

Risk gene candidates can be scored with predictive features extracted from either the functional linkage network (Linghu *et al.* 2009), or GO annotation, or the two sources combined. In comparison, they may also be scored by another two approaches: one is with their node degrees alone, a simple but informative network characteristic; the other is the number of connections to schizophrenia training genes (aka "risk degree"), which characterizes the degree of relationship with training genes. We evaluated the performance of these five scoring designs by two different but complementary approaches. First, for each design, we constructed the receiver operating characteristic (ROC) curve, and calculated the area under the curve (AUC) (Figure S5B). When we used both network and annotation predictive features, our method

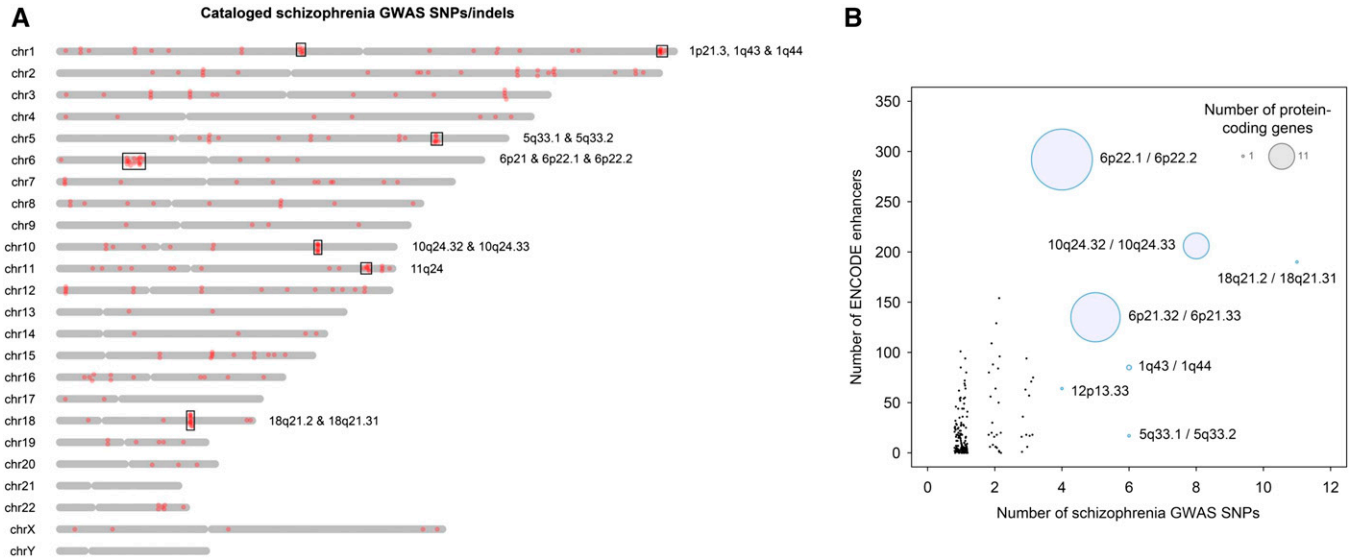


Figure 2 Currently cataloged schizophrenia GWAS SNPs. (A) Genomic distribution of schizophrenia GWAS SNPs. Each red dot represents a schizophrenia GWAS SNP. Several local clusters are highlighted. (B) The numbers of schizophrenia GWAS SNPs, ENCODE enhancers, and protein-coding genes in risk regions. The bubble size indicates the number of genes. The risk regions are labeled with the chromosome bands, and the ones with the number of schizophrenia SNPs <4 are shown only by dots.

achieved the best performance. Likewise, our method consistently outperformed the other two methods, and exhibited the most stable performance when confounding factors, such as the network degree and the gene size in evaluation gene sets, were controlled.

Next, we scored all 16,906 genes, and compared scores of 15,130 background genes with those of 56 known and 1718 "enriched" schizophrenia genes, respectively (Figure S5C). Wilcoxon rank sum tests showed that, with all five scoring designs, each of the schizophrenia gene sets scored significantly better than the background genes. For the test of evaluating score differences between 56 known/well-accepted schizophrenia genes and "background" genes, our method using both network and annotation predictive features exhibited similar performance to the method using risk degree as scores. However, for the test of evaluating score differences between the "enriched" gene set and the "background" gene set, our method using either network or annotation predictive features significantly outperformed the method of using risk degree as scores. This indicates that our network and annotation features are very effective in scoring unknown schizophrenia genes, not limited to scoring known schizophrenia genes that are highly functionally associated with other known schizophrenia genes. When both network and annotation predictive features were used, the differences in scores were the most significant.

Both evaluation approaches—the classification test and the Wilcoxon rank sum test—clearly showed that our scoring method can effectively prioritize schizophrenia genes. A large number of different combinations of functionally related genes as network features generated by training with seed genes can effectively capture the underlying genetic risk of schizophrenia. Network predictive features consider only genes functionally associated with query genes (but not query genes themselves), and may be insufficient to differentiate schizophrenia genes from genes with spurious functional linkage to the same neighbors in the network. On the other hand, scoring relying on GO annotation alone runs the risk of prediction being biased toward well-studied genes. Because the functional linkage gene network was built using high-throughput genomic data sets, by integrating both gene network and annotation, the risk of biased prediction was minimized.

High-scoring schizophrenia risk gene candidates

Using human genome annotation and transcriptional regulatory information, to 176 schizophrenia risk regions, we linked 643 schizophrenia risk gene candidates, of which 487 are proximal, covered by, or closest to, the schizophrenia risk regions, and the other 156 are distal, linked to the risk regions through long-range gene regulation. Our schizophrenia risk gene candidates show a size distribution very similar to that of all coding genes (Figure S7). By contrast, the set of genes closest to schizophrenia-associated SNPs, which were considered as risk genes by the current GWAS approach, show a striking bias toward large genes.

Due to the lack of GO annotation, or their absence from the functional gene linkage network, 58 candidates cannot be scored. Among the remaining 585 scored candidates, 132 genes from 76 schizophrenia risk regions achieve scores greater than the threshold (Table 1 and Table S3). Referred to as "schizophrenia risk genes" hereafter, these high-scoring candidates include 103 (78%) genes proximal to the schizophrenia risk genomic regions, and 29 (22%) distal genes that are likely regulated by TREs or eQTL in the risk regions (Table 1 and Table S4). For lack of a better approach, in most,

Table 1 The 132 high-scoring schizophrenia genes

Gene	Score	Gene	Score	Gene	Score	Gene	Score
FGFR1	**926.8**	*CD14*	1051.8	*GRIK3*	200.3	*PDC*	111.9
MAPK3	**915.7**	*DPP4*	933.8	*SOX2*	198.1	*NUCB2*	111.3
RAD51	**481.9**	*PLA2G15*	854.0	*MDK*	197.0	*HLA-DRB1*	107.4
PRKCD	**439.5**	*STAR*	822.7	*EPHX2*	192.7	*CTNNA1*	106.0
SREBF1	**388.5**	*LRP1*	692.5	*AIF1*	185.2	*AGER*	105.8
OPN1LW	**359.6**	*GNAL*	612.4	*CNTN4*	184.2	*CTNND1*	105.5
CD34	**346.3**	*CACNA1C*	597.1	*PTN*	174.6	*SREBF2*	105.2
FLNA	**246.1**	*RELN*	562.0	*MEF2C*	170.5	*SRR*	104.4
HBEGF	**189.8**	*NLGN4X*	554.6	*CHRNA5*	161.0	*L1CAM*	103.6
PAM	**188.1**	*TCF4*	522.1	*CHRM4*	160.6	*CACNB2*	101.7
TPR	**185.0**	*CYP17A1*	416.5	*PTGIS*	155.5	*BTG1*	100.8
MCL1	**182.2**	*GRM3*	378.2	*HSPD1*	154.5	*SLC12A4*	100.3
TAP1	**173.0**	*ARNTL*	359.6	*IMMP2L*	153.6	*PARD3*	100.2
NMUR2	**167.7**	*RIMS1*	357.2	*MSH6*	152.5	*RGS6*	99.8
TIE1	**167.5**	*CLU*	348.9	*NOTCH4*	151.4	*CETP*	99.7
CCL22	**159.4**	*GRIA1*	344.5	*PAK6*	149.7	*TENM3*	98.4
PJA2	**145.9**	*FURIN*	326.3	*DGKZ*	148.9	*TAOK2*	93.7
NISCH	**138.6**	*CHRNA3*	323.5	*LCAT*	148.3	*RANGAP1*	91.0
ADD1	**129.7**	*MECP2*	310.3	*ERCC4*	147.7	*NRGN*	90.1
NEURL	**127.1**	*MMP16*	303.7	*APOM*	147.1	*GNL3*	89.9
HTR3B	**125.1**	*IRAK1*	301.1	*HCN1*	146.9	*ARHGAP4*	89.0
PRMT1	**116.7**	*CYP21A2*	299.8	*STAT6*	144.1	*KCNV1*	88.8
SV2B	**107.8**	*FES*	286.1	*ITIH4*	140.8	*HLA-DRA*	86.5
MAPK11	**97.4**	*CYP2D6*	275.3	*ZEB2*	140.3	*MYO15A*	84.0
DOCK4	**90.2**	*CHRNB4*	269.3	*MYO1A*	138.6	*CNKSR2*	83.6
GCA	**88.6**	*PPARGC1A*	250.8	*SERPING1*	137.5	*CLIC1*	80.6
FLII	**87.9**	*DOC2A*	249.6	*TNXB*	137.0	*CD46*	80.6
FMR1	**85.7**	*CDH13*	231.7	*CHEK1*	136.9		
MLC1	**80.7**	*AVPR2*	229.9	*ANK3*	135.1		
DRD2	1420.9	*EI24*	220.3	*IK*	134.5		
GRIN2A	1385.0	*ATP2A2*	215.8	*CLCN3*	131.4		
PTGS2	1364.2	*PIK3C2A*	213.8	*RRAS*	127.3		
CXCL12	1114.3	*BMP7*	210.8	*HCFC1*	122.0		
PRKD1	1104.4	*EP300*	210.8	*HLA-DQB1*	114.6		
EGR1	1060.1	*PLCB2*	209.8	*TAC3*	113.4		

Note: 29 distal genes are highlighted in bold (see Table S3 and Table S4 for details).

if not all, GWAS in complex diseases, genes closest to the disease-associated SNPs were considered as the risk genes. This approach will certainly miss the distal risk genes, but, even among the proximal ones, only 189 out of 487 (39%) are genes closest to the schizophrenia GWAS signals.

We carefully examined the predicted schizophrenia risk genes to validate the effectiveness of our method. In one case (Figure 3A), a risk region on chromosome 8 indexed by schizophrenia-associated SNP rs16887244 is linked to nine protein-coding genes. Among them, rs16887244 is located in an intron of *LSM1*, and thus *LSM1* was reported as a putative schizophrenia risk gene (Hindorff *et al.* 2015). Our method, however, identified different risk genes for rs16887244. While *LSM1* scored low (4.8), we identified with high scores (822.7 and 926.8, respectively), two risk genes, *STAR* and *FGFR1*. *STAR* encodes a steroidogenic acute regulatory protein, which regulates the onset of steroidogenesis. Widely expressed throughout human brain, *STAR* may play a role in maintaining several brain functions, such as neurogenesis, neuroprotection, and synaptic plasticity (Sierra 2004). *FGFR1* encodes fibroblast growth factor receptor 1, and is involved in many important signaling pathways, whose impairment could lead to abnormal brain development, and confer risk of schizophrenia (Terwisscha van Scheltinga *et al.* 2010). In contrast to *STAR*, which resides in the risk region, *FGFR1* is a distal gene located outside the risk region. It was connected to the GWAS signal through two TREs in the risk region that may regulate its expression. This connection is strengthened by the strong LD between the schizophrenia SNP rs16887244 and the two SNPs, rs6999796 and rs16887343, each located in one of the TREs.

Another risk region on chromosome 16 indexed by schizophrenia-associated SNP rs12691307 is linked to 13 protein-coding genes (Figure 3B). The gene closest to the index SNP, *KCTD13*, was given a low score by our method (37.1). Instead, three other genes, *DOC2A*, *MAPK3*, and *TAOK2*, scored high (249.6, 915.7, and 93.7, respectively). In fact, the risk region is located at chromosome 16p11.2—a known risk locus for autism (Kumar *et al.* 2008). Autism is another neurodevelopmental disorder that shares a number of features with schizophrenia (Goldstein *et al.* 2002). Interestingly, all three genes have been implicated in autism in

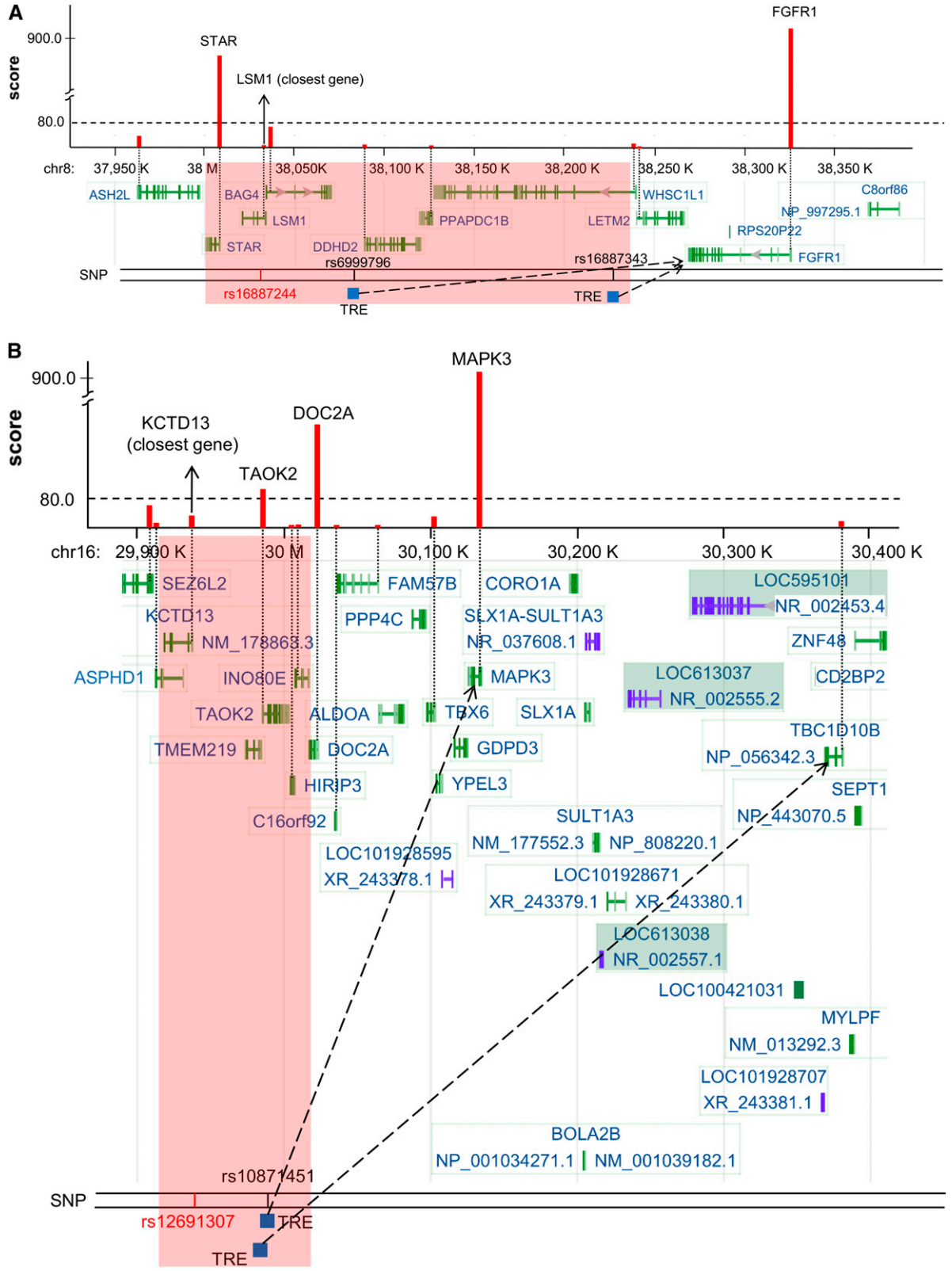


Figure 3 Prioritization of schizophrenia risk gene candidates. Schizophrenia-associated SNPs (shown with red "rs" IDs) were used to define schizophrenia risk genomic regions (red rectangles). Schizophrenia risk gene candidates are genes either overlapping schizophrenia risk regions or linked to them by TREs. The scores of candidates are indicated by the red bars. (A) A schizophrenia risk genomic region in 8p11.23. Among nine schizophrenia risk gene candidates, *STAR* and *FGFR1* achieved high scores. *FGFR1* is linked to this risk region through two TREs. (B) A schizophrenia risk genomic region in 16p11.2. 13 schizophrenia risk gene candidates are connected to this risk region (TBK6 is linked through eQTL). Three of them—*TAOK2*,

previous studies (Kumar *et al.* 2008; de Anda *et al.* 2012). *DOC2A* encodes calcium-signaling proteins responsible for neurotransmission (Glessner *et al.* 2010). *MAPK3* encodes a serine/threonine protein kinase that plays an important role in the regulation of synaptic plasticity (Thomas and Huganir 2004). *TAOK2* also encodes a serine/threonine protein kinase that affects basal dendrite formation (de Anda *et al.* 2012). Their biological roles are consistent with the current knowledge of schizophrenia etiology. In contrast to *DOC2A* and *TAOK2*, which reside in the risk region, *MAPK3* is a distal gene ~0.1 Mb downstream to the risk region. The connection between the GWAS signal and *MAPK3* was established through a TRE in the risk region. SNP rs10871451 in this TRE is in strong LD with the nearby schizophrenia-associated SNP, and thus implicated as the underlying risk variant.

In the aforementioned cases, our method predicted risk genes different from genes closest to schizophrenia-associated GWAS SNPs. The high scores assigned by our method to those predicted risk genes were calculated based on solid gene annotation and functional linkage. Our post-GWAS analysis generated a high-confidence set of schizophrenia risk genes, many of which are new. Although their ultimate validation and confirmation can be achieved only experimentally (and thus beyond the scope of this work), we carried out computational analyses and the results—described in the following sections—show that our predictions are well supported by other resources.

Association among schizophrenia-related gene sets

To validate and characterize our schizophrenia risk genes, among 585 scored candidates, we compiled three sets of schizophrenia-related genes based on phenotypes found in transgenic mice, schizophrenia research literature, and differential gene expression studies in schizophrenia. Fisher's exact tests of association among these four gene sets (Table S5) using 585 scored candidate genes as the background shows that our schizophrenia risk genes are highly associated with genes either rendering relevant phenotypes in transgenic mice ($P = 9.58 \times 10^{-19}$), or with schizophrenia literature support ($P = 2.92 \times 10^{-5}$). However, we did not detect association ($P = 0.135$) between our schizophrenia risk genes, and genes from differential expression studies of schizophrenia.

Tissue gene expression analysis

Although recognized as a brain disorder, accumulating evidence also shows that the etiology of schizophrenia is associated with immune dysfunction (Muller and Schwarz 2010). We examined the expression profiles of schizophrenia risk genes across different human tissues to investigate the tissue specificity of their transcriptional activities. Based on their expression patterns, we can cluster them into three groups (Figure S8). The first group of 39 genes is expressed almost exclusively in the central nervous system (CNS), especially the prefrontal cortex and the hippocampus. Many genes in this group, such as *CACNA1C*, *CACNB2*, and *RIMS1* have been implicated in the pathogenesis of schizophrenia (Table S2).

Thirty-five genes in the second group are highly expressed in immune cells: B-c and T-lymphocytes. Genes in this group include the major histocompatibility complex (MHC) genes, such as *HLA-DQB1*, *HLA-DRB1*, and *HLA-DRA*. MHC genes code for proteins that regulate immune functions (Janeway 2001), while the MHC region on chromosome 6p implicated in schizophrenia in replicated GWAS (Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014). Other immune-associated genes in this group, such as *PTGS2* and *FMR1*, have been linked to schizophrenia in previous studies (Wei and Hemmings 2004; Kelemen *et al.* 2013). Recent studies have found the anatomical connection between the immune system and the CNS (Aspelund *et al.* 2015; Louveau *et al.* 2015), which could explain the involvement of immune-associated genes in schizophrenia. The third group consists of 54 genes that are expressed across a wide range of different tissues including the CNS. In contrast to genes in the first group, genes in this group are not exclusive to the CNS, and are expressed more ubiquitously. Many genes in this group, such as *EGR1*, *FGFR1*, *CHRNA5*, *SREBF1*, *SREBF2*, and *PARD3*, are known to be involved in schizophrenia (Table S2). According to our results, an unexpectedly high percentage (~25%) of schizophrenia risk genes are not expressed in the CNS. How those genes expressed in the immune system play a role in the pathogenesis of schizophrenia requires further investigation.

Overlaps in schizophrenia genetic architecture

The common variant part of the genetic architecture of schizophrenia has been studied extensively in recent SNP array-based GWAS, which have identified a large number of associated SNPs, as noted above. A new frontier for schizophrenia genetics is to identify rare variants, and *de novo* mutations, associated with schizophrenia risk by whole exome (WES) or whole genome sequencing (WGS). Two recent studies—WES of 2536 schizophrenia individuals and 2543 healthy controls (Purcell *et al.* 2014), and WES of 623 schizophrenia trios (Fromer *et al.* 2014)—are the two largest sequencing-based studies to fill in the rare variant and *de novo* mutation part of the genetic architecture of schizophrenia. Although strong evidence from these large-scale genetics studies suggests that there is convergence of rare and common variants in genetic architecture of schizophrenia at broad gene functional levels (Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014), it remains unclear, however, how commonly at gene levels, rare variants underlie schizophrenia GWAS signals (form

DOC2A, and *MAPK3*—achieved high scores. Note that, in either case, the gene closest to the schizophrenia-associated SNPs has a score lower than the threshold (at 80, shown by the dashed line).

common variants), and how commonly schizophrenia risk genes may exert their pathogenic effects through both common and rare variants. We were able to shed some new light on these questions by comparing our GWAS-derived schizophrenia risk genes with genes containing rare variants or *de novo* mutations implicated in schizophrenia by the previous exome-sequencing studies (Girard *et al.* 2011; Xu *et al.* 2012; Fromer *et al.* 2014; Purcell *et al.* 2014).

Of our schizophrenia risk genes, 37, 7, and 7 contain rare variants, *de novo* mutations, or both, respectively (Figure S9A). We conducted two statistical tests to assess the significance of overlap between schizophrenia risk genes that we predicted, and schizophrenia risk genes implicated by rare mutations (Figure S9, B and C). There is a statistically significant association ($P = 6.7 \times 10^{-4}$) between high scoring genes linked to schizophrenia GWAS loci and schizophrenia genes implicated by rare variants (Figure S9B). After eliminating the confounding effect of "high scoring," the overlap between these two sets of schizophrenia risk genes remains significant ($P = 8.3 \times 10^{-4}$) (Figure S9C). Such overlaps indicate the possibility that some schizophrenia risk genes may contribute to the disease through both common and rare variants. Among the aforementioned 37 schizophrenia risk genes, we also found genes involved in glutamatergic neurotransmission (*GRIK3* and *GRIN2A*), and genes encode calcium channels (*CACNA1C* and *CACNB2*) and synaptic plasticity (NMDAR genes such as *FLNA* and *MAPK3*) (Kirov *et al.* 2012). All these three gene classes have been implicated in schizophrenia by both rare and common variants in a previous study (Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014).

Pathway enrichment

The Psychiatric Genomics Consortium (PGC) meta-analysis of schizophrenia (Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014) could not, with statistical significance after multi-test correction, identify any enriched pathways among genes within the 108 loci. By focusing only on high scoring risk genes, and expanding gene candidates to include distal genes and genes associated with weak GWAS signals, many biologically plausible pathways were overrepresented (Table S6 and Table S7). In addition, we also found pathways not enriched with training schizophrenia genes, including pathways involved in neural development (FGF signaling and Adherens junction), synaptic function and plasticity (Endothelin signaling pathway), and immune system (B cell activation and intestinal immune network for IgA production), all of which are consistent with the current knowledge of the etiology of schizophrenia. By integrating GWAS signals and regulatory information, our approach can identify disease risk genes to uncover novel disease-related pathways.

Schizophrenia risk genes with different association strengths

In GWAS, variants show different degrees of association with the disease. Variants with smaller *P*-values in the same study imply higher risks than variants with larger *P*-values. To identify the biological factors underlying different genetic risks, we divided the range of schizophrenia association strength of the risk regions into three classes based on the single large PGC study (Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014). According to the *P*-value distribution of GWAS SNPs (Figure S10), we divided 176 risk regions into three classes: 62 weak ($P > 5 \times 10^{-8}$), 70 moderate ($10^{-10} < P < 5 \times 10^{-8}$), and 39 strong ($P < 10^{-10}$) regions—with different disease-association strengths based on the lowest *P*-values of associated PGC GWAS signals in each region (Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014). The "weak" class consists of risk regions that contain no genome-wide significant GWAS signals from the PGC study. Five weak risk regions with either no, or contradictory, imputation signals in the PGC study were excluded from the analysis (Table S3). We then assigned the schizophrenia risk genes to these three association classes based on the GWAS variants to which they are linked (Table S8). GO term analysis reveals that genes in these three disease-association classes are enriched with GO terms of distinct biological processes (Figure 4): schizophrenia risk genes with weak association are enriched in biological processes related to cellular regulation and differentiation; ones with moderate association function mainly in response to stimulus and organismal processes; and strong association is connected with synaptic transmission and signaling. For example, weak associations involve many schizophrenia risk genes that play a role in cellular regulation of neural development, such as cell motion and axongenesis (*L1CAM*, *ANK3*, *BMP7*, *CXCL12*, and *RELN*). In contrast, strong associations involve many schizophrenia risk genes encoding calcium channels (*CACNA1C* and *CACNB2*) and neurotransmitter receptors (*DRD2*, *CHRNA3*, *CHRNA5*, *CHRM4*, and *HTR3B*) that are directly involved in synaptic transmission. To provide some biological context to the three disease-association classes, we compiled a set of 20 genes connected to schizophrenia from the OMIM database (http://www.omim.org, accessed November 2014) (McKusick 2007). These OMIM genes do not overlap with our predicted schizophrenia risk genes. As cataloged in the OMIM database, these genes have identifiable genetic factors that may have larger effect sizes on schizophrenia risk in general. Interestingly, like schizophrenia risk genes with strong association, these OMIM genes are also enriched in the biological process of synaptic transmission and signaling.

Consistent with the widely accepted hypothesis that schizophrenia symptoms are caused by the imbalance of neurotransmitter in brain, our result suggests that genes involved in synaptic transmission and signaling tend to have strong association with schizophrenia due to their direct influence on the balance of neurotransmitter in brain. The mutations of many genes involved in cellular regulation in brain may contribute to brain defects in the brain developmental process. However, this consequence may have implicit

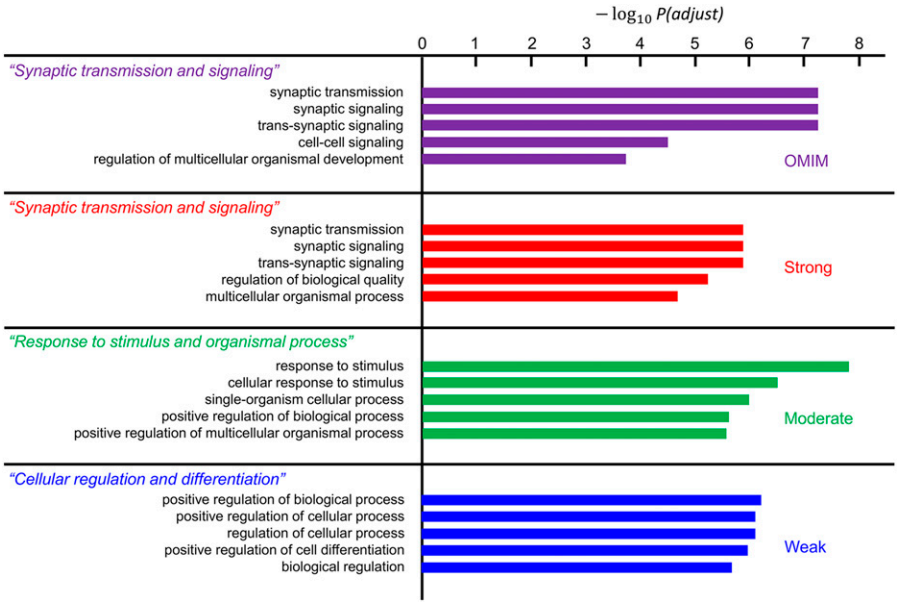


Figure 4 GO terms enriched among schizophrenia risk genes with different association strengths. The five most significantly enriched GO terms, and their *P*-values adjusted for multiple testings, are shown for each gene set. The label "OMIMS" in purple denotes 20 schizophrenia risk genes that we curated from the OMIM database. The labels "Strong," "Moderate," and Weak denote 36, 49, and 35 putative schizophrenia risk genes implicated by strong, moderate, and weak GWAS signals, respectively (see Table S8).

connection to the outcome of neurotransmitter imbalance, which is reflected by their weaker associations in general.

Expression of schizophrenia risk genes during brain development

Strong research findings indicate that schizophrenia is a complex neurodevelopmental disorder (Fatemi and Folsom 2009; Catts *et al.* 2013). Thus, we investigated how schizophrenia risk genes are expressed during brain development. Instead of studying them individually, or together as a whole, we examined the spatiotemporal expression profiles of the aforementioned three disease-association classes at eight brain locations, and 12 time points during brain development, using RNA-Seq data from BrainSpan (http://www.brainspan.org/, accessed March 2016) (Figure S11) (File S1). Expression analysis reveals that the timing of their transcriptional activity during brain development correlates well with the strength of their association with schizophrenia (Figure 5): schizophrenia risk genes with weak, moderate, and strong association tend to be more actively transcribed during the early, middle, and late time periods, respectively, during brain development. Again, like schizophrenia risk genes with strong associations, the OMIM schizophrenia genes tend to be transcribed more actively during the late time period. We generated new sets of prioritized genes using specially controlled training genes. Our comprehensive analysis of these genes showed essentially the same spatiotemporal expression patterns during brain development as before (Figure S12), and thus excluded the possibilities that the properties of training genes drive the patterns of transcriptional activities of schizophrenia risk genes with different association strengths. Although the binarization process used in the approach discards some transcriptional information, the advantage of our approach to identifying spatiotemporal expression patterns is the interpretability of its result, which shows the proportion of genes in the gene set that tend to be transcriptionally active, or suppressed at the corresponding time stage and brain region. To ensure that the observed spatiotemporal expression patterns are robust, we used a different transformation of the expression data, which gave results (Figure S13A) consistent with our previous observation. Moreover, we conducted statistical tests to assess the significance of transcriptional activities. The test results show consistent spatiotemporal expression patterns (Figure S13, B and C), indicating that the distinct patterns of transcriptional activities of our prioritized genes in different association classes are not due to the overall characteristics of genes linked to the genomic regions (Figure S13C). The three transcriptionally active time periods correspond to distinct brain developmental stages (Figure S14). The early time period is from 4 to 12 postconception weeks (PCW), when cell birth and migration occur in the embryonic and early prenatal brain. The middle time period includes 25–38 PCW (late prenatal), and 6–18 months after birth (late infancy), a major development stage for synaptogenesis. The late time period mainly consists of 8–19 years and 20–40 years, which include adolescence and early adulthood, when the onset of schizophrenia usually occurs.

The significantly enriched GO terms of biological processes among genes with weak association is consistent with the formation of brain "hardware" at the cellular level, for which early neurodevelopmental stages are critical times when these genes are most transcriptionally active. In addition to early stages of neurodevelopment, perinatal development is also potentially vulnerable to perturbations in schizophrenia susceptibility genes that may contribute to the future onset of the disorder (Catts *et al.* 2013). Considering that emerging evidence implicates postnatal development changes in schizophrenia (Catts *et al.* 2013), the observation that many schizophrenia risk genes with strong association are more active during this period is intriguing. The developmental trajectories of eight schizophrenia risk genes with strong

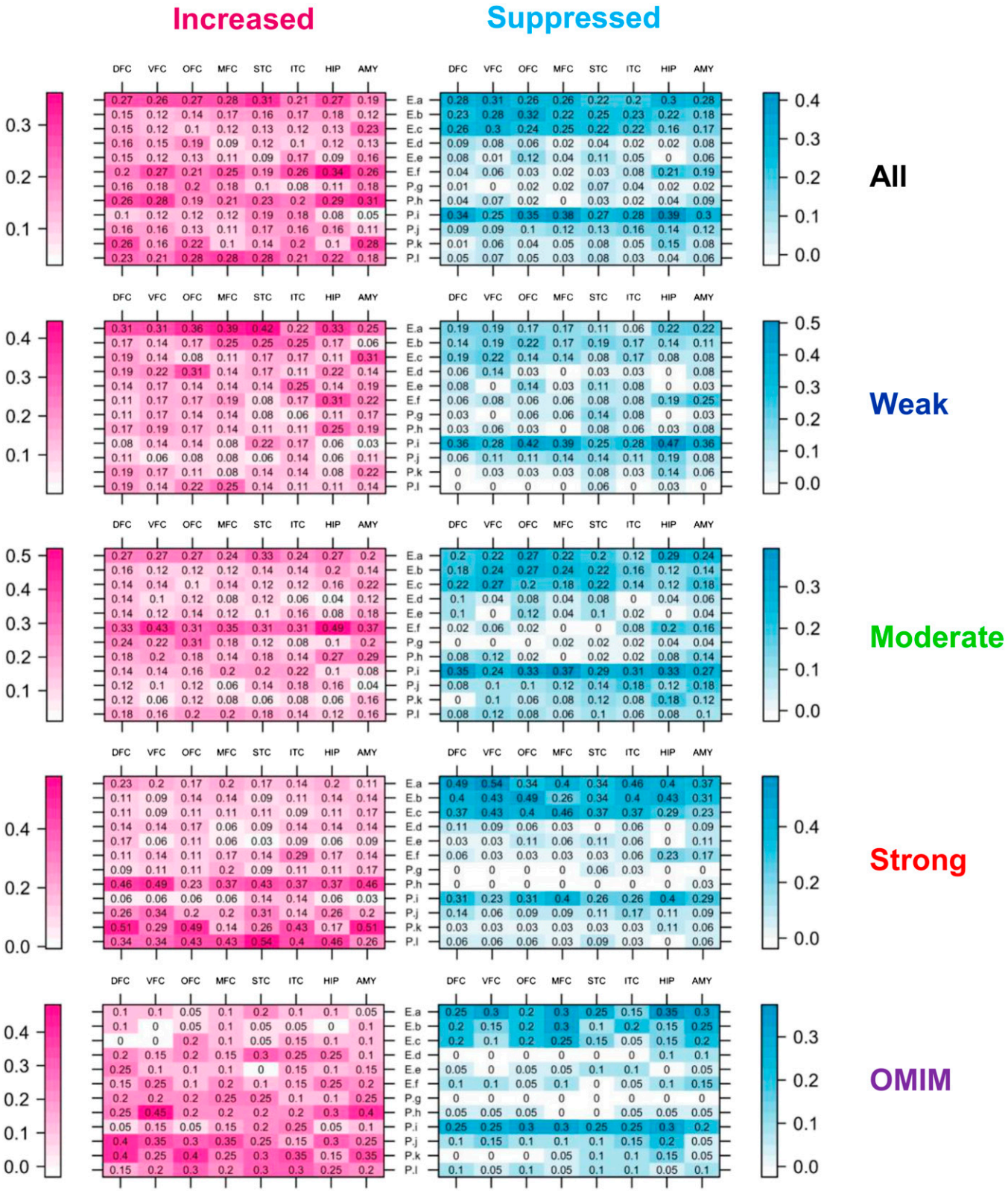


Figure 5 Spatiotemporal expression patterns of schizophrenia risk genes during brain development. The heat maps show both the active (red), and the suppressed (blue), expression, respectively, of different gene sets. The rows are 12 developmental stages in a chronological order, and the columns are eight brain regions. The shade of the color in a heat map is proportional to the ratio of genes that manifest active (or suppressed) activities, at the corresponding brain location and time stage, to the total number of genes in the specific gene set. E.a-f and P.g-l denote six embryonic, and six

associations (Figure S15) suggest that they are more active during the postnatal period, including adolescence.

Discussion

Schizophrenia is a complex genetic disease. As a severe lifelong mental disorder affecting ~1% of the United States population, it creates an enormous burden to patients, their families and the community. In the past several years, GWAS have been applied successfully to schizophrenia, and a large number of associated genetic loci have been identified, which could lead to the development of targeted therapies. Interpreting the GWAS results, however, remains difficult due to both the design of GWAS, and the nature of many identified risk loci. First, SNPs used in GWAS are tagging SNPs, each representing a large LD block, which may contain a large number of genes and regulatory elements (and thus possibly affecting genes elsewhere). Second, most variants found in GWAS to be associated with diseases including schizophrenia lie outside of protein-coding regions, and this observation remains true even after fine-mapping around the associated loci (Wellcome Trust Case Control Consortium *et al.* 2012).

For lack of a better approach, the genes closest to, or in the vicinity of, disease-associated SNPs found in GWAS are generally assumed to be the risk genes. However, this assumption may be overly simplistic, and identifying putative disease risk genes using new computational tools is critical in properly interpreting GWAS signals for diagnostic and therapeutic purposes. Responding to this need, we used an integrated post-GWAS analysis, and identified 132 putative schizophrenia risk genes, and determined their functional roles in schizophrenia. In our analysis framework, we used new computational methods based on rigorous statistical modeling to integrate a large number of heterogeneous genomic data sets from diverse sources, and, with a sensible score threshold, achieved high accuracy in our risk gene prediction. Two advantages of our method are immediately clear from our analysis results. First, our method can identify putative disease risk genes not only in the vicinity of GWAS signals, but also at a distance by regulatory elements in the risk region that affect gene expression. Disease genes distal to GWAS signals have never been identified before. Second, our method can also identify putative disease risk genes for GWAS variants that did not reach the genome-wide significance level ($P < 5 \times 10^{-8}$). Such weak GWAS signals are usually ignored. In this study of schizophrenia, we identified 29 putative distal risk genes, and 36 putative risk genes with weak association. Together, there are 55 novel schizophrenia risk genes that were missed by previous GWAS.

Our pathway analysis result indicates that, even though our gene scoring method is based on the functional properties of known risk genes, by integrating with GWAS signals and regulatory information, our approach has potential to uncover novo risk pathways in which common risk variants are involved. The underlying reason is that, although high-scoring genes must have certain functional similarities with seed genes, they are also likely involved in other risk factors not associated with seed genes. Therefore, benefitting from the fact that GWAS is non-hypothesis-driven, the analysis of high scoring genes implicated by GWAS signals may reveal novel risk factors associated with common risk variants.

The extended MHC region is a gene-dense region with long LD blocks, and often drives false-positive predictions. Six risk regions are located in this complex region (Table S3), and they involve 98 candidate genes, of which 11 are high scoring (Table S9 and Table S10). If the extended MHC region is excluded from our analysis, the results stay essentially the same. The set of high scoring genes remains highly associated with genes with relevant phenotypes of transgenic mice ($P = 3.11 \times 10^{-16}$), and genes with literature support ($P = 6.26 \times 10^{-5}$). The percentage of high scoring genes expressed in immune related tissues but not in the CNS remains high (~25%). The enrichment of GWAS risk genes among schizophrenia risk genes implicated by rare variants stays significant ($P = 8 \times 10^{-5}$). The extended MHC region is not involved in the analysis of schizophrenia risk genes with different association strengths, due to the uncertainty about the association strength of the risk regions within it (Table S3).

To explain the lack of association between 132 schizophrenia risk genes and genes from differential expression studies of schizophrenia, we investigated their topological arrangements in the functional linkage network. There are 932 differentially expressed genes among the neighbors of all 132 schizophrenia risk genes. On average, there are more differentially expressed genes among the neighbors of each of 132 schizophrenia risk genes, compared to 132 random genes (Figure S16). The result indicates that, although schizophrenia risk genes themselves may not be differentially expressed between schizophrenia patients and normal individuals, compared to nonrisk genes, they are more likely ($P = 0$, with 1000 replicates) to be functionally associated with differentially expressed genes.

In this study, we focused functional analyses on 132 prioritized genes out of 643 candidate genes. Despite the presence of potential false negatives [*e.g.*, *ZNF804A*], the overall characteristics of the remaining 511 candidate genes are very different from our prioritized genes. For example, genes with relevant phenotypes in transgenic mice, and genes with literature support for schizophrenia risk, are both overrepresented in our prioritized genes, but not in the remaining candidate genes (Figure S17). As expected, the patterns of transcriptional activities for prioritized genes with different

postnatal, developmental stages (see Figure S14 for details). DFC, dorsolateral prefrontal cortex; VFC, ventrolateral prefrontal cortex; OFC, orbitofrontal cortex; MFC, medial prefrontal cortex; STC, posterior superior temporal cortex; ITC, inferior temporal cortex; HIP, hippocampus; AMY, amygdala.

association strengths are not observed for the remaining candidate genes (Figure S13C and Figure S18).

Of the 176 schizophrenia risk regions derived from GWAS signals, 100 do not contain genes with high scores. Several reasons could account for this absence. First, for risk regions with weak associations, the possibility that the associated GWAS signals were false positives could not be excluded, especially for regions that do not contain genes with high scores. Second, some distal risk genes might not be included in the candidate gene list due to incomplete TRE/eQTL regulatory information. Third, our schizophrenia gene scoring method relied on previous knowledge of functional linkage network and GO annotations, and thus was limited by them. Fourth, our schizophrenia gene scoring method was trained by using the schizophrenia training gene set. Some schizophrenia risk genes exerting pathogenic effects through very different mechanisms from schizophrenia training genes would not score highly. Fifth, our method considered only coding schizophrenia genes, while noncoding RNAs, such as miRNAs, were not considered. It should be noted that emerging evidence showed that miRNAs could also be risk factors for schizophrenia (Mellios and Sur 2012).

We identified 132 putative schizophrenia risk genes using our method, of which the majority have not been recognized in previous schizophrenia GWAS. In particular, 36 putative risk genes associated with GWAS signals at genome wide significance level were identified. Those weak signals are usually ignored due to the lack of an approach to avoid false positive GWAS signals. However, identification of risk genes with weak association is important to investigate the disease mechanisms underlying association strength. Our analysis suggests that, despite the high diversity of risk factors involved in schizophrenia, genes involved in certain biological processes are more likely to have higher degrees of penetrance, which indicates that certain biological processes have a stronger linkage to developing the disorder. Our analysis also shows that schizophrenia risk genes that are transcriptionally active in certain brain developmental stages are more likely to have higher degrees of penetrance, implicating a stronger linkage between the biological events in those brain developmental stages, and developing the disorder.

Acknowledgments

The authors thank Herbert M. Lachman of the Department of Psychiatry and Behavioral Sciences at Albert Einstein College of Medicine, and Anne S. Bassett of the Department of Psychiatry at the University of Toronto, for comments and suggestions. This work was supported by the National Institutes of Health grant MH101720 from the National Institute of Mental Health to the International Consortium on Brain and Behavior in 22q11.2 Deletion Syndrome. The authors declare that they have no competing interests.

Literature Cited

1000 Genomes Project ConsortiumAbecasis, G. R., A. Auton, L. D. Brooks, M. A. DePristo, R. M. Durbin *et al.*, 2012 An integrated map of genetic variation from 1,092 human genomes. Nature 491: 56–65.

Andersson, R., C. Gebhard, I. Miguel-Escalada, I. Hoof, J. Bornholdt *et al.*, 2014 An atlas of active enhancers across human cell types and tissues. Nature 507: 455–461.

Aspelund, A., S. Antila, S. T. Proulx, T. V. Karlsen, S. Karaman *et al.*, 2015 A dural lymphatic vascular system that drains brain interstitial fluid and macromolecules. J. Exp. Med. 212: 991–999.

Benita, Y., Z. Cao, C. Giallourakis, C. Li, A. Gardet *et al.*, 2010 Gene enrichment profiles reveal T-cell development, differentiation, and lineage-specific transcription factors including ZBTB25 as a novel NF-AT repressor. Blood 115: 5376–5384.

Bossu, P., F. Piras, I. Palladino, M. Iorio, F. Salani *et al.*, 2015 Hippocampal volume and depressive symptoms are linked to serum IL-18 in schizophrenia. Neurol. Neuroimmunol. Neuroinflamm. 2: e111.

Boyle, E. I., S. Weng, J. Gollub, H. Jin, D. Botstein *et al.*, 2004 GO::TermFinder–open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. Bioinformatics 20: 3710–3715.

BrainSpan: Atlas of the Developing Human Brain [Internet]. Funded by ARRA Awards 1RC2MH089921–01, 1RC2MH090047–01, and 1RC2MH089929–01. 2011. Available at: http://developinghumanbrain.org. Accessed: March 28, 2016.

Canetta, S., A. Sourander, H. M. Surcel, S. Hinkka-Yli-Salomaki, J. Leiviska *et al.*, 2014 Elevated maternal C-reactive protein and increased risk of schizophrenia in a national birth cohort. Am. J. Psychiatry 171: 960–968.

Catts, V. S., S. J. Fung, L. E. Long, D. Joshi, A. Vercammen *et al.*, 2013 Rethinking schizophrenia in the context of normal neurodevelopment. Front. Cell. Neurosci. 7: 60.

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks *et al.*, 2011 The variant call format and VCFtools. Bioinformatics 27: 2156–2158.

de Anda, F. C., A. L. Rosario, O. Durak, T. Tran, J. Graff *et al.*, 2012 Autism spectrum disorder susceptibility gene TAOK2 affects basal dendrite formation in the neocortex. Nat. Neurosci. 15: 1022–1031.

Di Forti, M., J. M. Lappin, and R. M. Murray, 2007 Risk factors for schizophrenia–all roads lead to dopamine. Eur. Neuropsychopharmacol. 17(Suppl. 2): S101–S107.

du Bois, T. M., and X. F. Huang, 2007 Early brain development disruption from NMDA receptor hypofunction: relevance to schizophrenia. Brain Res. Brain Res. Rev. 53: 260–270.

Encode Project Consortium, , 2012 An integrated encyclopedia of DNA elements in the human genome. Nature 489: 57–74.

Fatemi, S. H., and T. D. Folsom, 2009 The neurodevelopmental hypothesis of schizophrenia, revisited. Schizophr. Bull. 35: 528–548.

Fromer, M., A. J. Pocklington, D. H. Kavanagh, H. J. Williams, S. Dwyer *et al.*, 2014 De novo mutations in schizophrenia implicate synaptic networks. Nature 506: 179–184.

Girard, S. L., J. Gauthier, A. Noreau, L. Xiong, S. Zhou *et al.*, 2011 Increased exonic de novo mutation rate in individuals with schizophrenia. Nat. Genet. 43: 860–863.

Giusti-Rodriguez, P., and P. F. Sullivan, 2013 The genomics of schizophrenia: update and implications. J. Clin. Invest. 123: 4557–4563.

Glessner, J. T., M. P. Reilly, C. E. Kim, N. Takahashi, A. Albano *et al.*, 2010 Strong synaptic transmission impact by copy number variations in schizophrenia. Proc. Natl. Acad. Sci. USA 107: 10584–10589.

Goldstein, G., N. J. Minshew, D. N. Allen, and B. E. Seaton, 2002 High-functioning autism and schizophrenia: a comparison of an early and late onset neurodevelopmental disorder. Arch. Clin. Neuropsychol. 17: 461–475.

Hall, J., S. Trent, K. L. Thomas, M. C. O'Donovan, and M. J. Owen, 2015 Genetic risk for schizophrenia: convergence on synaptic pathways involved in plasticity. Biol. Psychiatry 77: 52–58.

Harrow, J., A. Frankish, J. M. Gonzalez, E. Tapanari, M. Diekhans *et al.*, 2012 GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res. 22: 1760–1774.

Hindorff, L. A., J. MacArthur (European Bioinformatics Institute), J. Morales (European Bioinformatics Institute), H. A. Junkins, P. N. Hall, A. K. Klemm, and T. A. Manolio A Catalog of Published Genome-Wide Association Studies. Available at: http://www.genome.gov/gwastudies. Accessed: March 31, 2015.

International Schizophrenia ConsortiumPurcell, S. M., N. R. Wray, J. L. Stone, P. M. Visscher, M. C. O'Donovan *et al.*, 2009 Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature 460: 748–752.

Janeway, C. A., P. Travers, M. Walport, and M. J. Shlomchik 2001 *Immunobiology*. Garland Science, New York.

Jia, P., J. Sun, A. Y. Guo, and Z. Zhao, 2010 SZGR: a comprehensive schizophrenia gene resource. Mol. Psychiatry 15: 453–462.

Kelemen, O., T. Kovacs, and S. Keri, 2013 Contrast, motion, perceptual integration, and neurocognition in schizophrenia: the role of fragile-X related mechanisms. Prog. Neuropsychopharmacol. Biol. Psychiatry 46: 92–97.

Kirov, G., A. J. Pocklington, P. Holmans, D. Ivanov, M. Ikeda *et al.*, 2012 De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. Mol. Psychiatry 17: 142–153.

Kotlar, A. V., K. B. Mercer, M. E. Zwick, and J. G. Mulle, 2015 New discoveries in schizophrenia genetics reveal neurobiological pathways: a review of recent findings. Eur. J. Med. Genet. 58: 704–714.

Kumar, R. A., S. KaraMohamed, J. Sudi, D. F. Conrad, C. Brune *et al.*, 2008 Recurrent 16p11.2 microdeletions in autism. Hum. Mol. Genet. 17: 628–638.

Legendre, P., 2014 Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion? J. Classif. 31: 274–295.

Lichtenstein, P., B. H. Yip, C. Bjork, Y. Pawitan, T. D. Cannon *et al.*, 2009 Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. Lancet 373: 234–239.

Linghu, B., E. S. Snitkin, Z. Hu, Y. Xia, and C. Delisi, 2009 Genome-wide prioritization of disease genes and identification of disease-disease associations from an integrated human functional linkage network. Genome Biol. 10: R91.

Louveau, A., I. Smirnov, T. J. Keyes, J. D. Eccles, S. J. Rouhani *et al.*, 2015 Structural and functional features of central nervous system lymphatic vessels. Nature 523: 337–341.

Lv, M. H., Y. L. Tan, S. X. Yan, L. Tian, D. C. Chen *et al.*, 2015 Decreased serum TNF-alpha levels in chronic schizophrenia patients on long-term antipsychotics: correlation with psychopathology and cognition. Psychopharmacology (Berl.) 232: 165–172.

McKusick, V. A., 2007 Mendelian inheritance in man and its online version, OMIM. Am. J. Hum. Genet. 80: 588–604.

Mellios, N., and M. Sur, 2012 The emerging role of microRNAs in Schizophrenia and Autism spectrum disorders. Front. Psychiatry 3: 39.

Muller, N., and M. J. Schwarz, 2010 Immune system and Schizophrenia. Curr. Immunol. Rev. 6: 213–220.

Nawa, H., H. Sotoyama, Y. Iwakura, N. Takei, and H. Namba, 2014 Neuropathologic implication of peripheral neuregulin-1 and EGF signals in dopaminergic dysfunction and behavioral deficits relevant to schizophrenia: their target cells and time window. BioMed Res. Int. 2014: 697935.

Need, A. C., and D. B. Goldstein, 2014 Schizophrenia genetics comes of age. Neuron 83: 760–763.

Purcell, S. M., J. L. Moran, M. Fromer, D. Ruderfer, N. Solovieff *et al.*, 2014 A polygenic burden of rare disruptive mutations in schizophrenia. Nature 506: 185–190.

Rees, E., M. C. O'Donovan, and M. J. Owen, 2015 Genetics of schizophrenia. Current Opinion in Behavioral Sciences 2: 8–14.

Ripke, S., C. O'Dushlaine, K. Chambert, J. L. Moran, A. K. Kahler *et al.*, 2013 Genome-wide association analysis identifies 13 new risk loci for schizophrenia. Nat. Genet. 45: 1150–1159.

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

Sierra, A., 2004 Neurosteroids: the StAR protein in the brain. J. Neuroendocrinol. 16: 787–793.

Sullivan, P. F., K. S. Kendler, and M. C. Neale, 2003 Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. Arch. Gen. Psychiatry 60: 1187–1192.

Tabas-Madrid, D., R. Nogales-Cadenas, and A. Pascual-Montano, 2012 GeneCodis3: a non-redundant and modular enrichment analysis tool for functional genomics. Nucleic Acids Res. 40: W478–W483.

Tan, P.-N., M. Steinbach, and V. Kumar, 2006 Introduction to Data Mining. Pearson Addison-Wesley, Boston.

Terwisscha van Scheltinga, A. F., S. C. Bakker, and R. S. Kahn, 2010 Fibroblast growth factors in schizophrenia. Schizophr. Bull. 36: 1157–1166.

Thomas, G. M., and R. L. Huganir, 2004 MAPK cascade signalling and synaptic plasticity. Nat. Rev. Neurosci. 5: 173–183.

Thurman, R. E., E. Rynes, R. Humbert, J. Vierstra, M. T. Maurano *et al.*, 2012 The accessible chromatin landscape of the human genome. Nature 489: 75–82.

Wei, J., and G. P. Hemmings, 2004 A study of a genetic association between the PTGS2/PLA2G4A locus and schizophrenia. Prostaglandins Leukot. Essent. Fatty Acids 70: 413–415.

Wellcome Trust Case Control ConsortiumMaller, J. B., G. McVean, J. Byrnes, D. Vukcevic, K. Palin *et al.*, 2012 Bayesian refinement of association signals for 14 loci in 3 common diseases. Nat. Genet. 44: 1294–1301.

Xu, B., I. Ionita-Laza, J. L. Roos, B. Boone, S. Woodrick *et al.*, 2012 De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. Nat. Genet. 44: 1365–1369.

Communicating editor: C. Sabatti

SUPPLEMENTARY METHODS

Scoring the schizophrenia risk gene candidates

We have developed a statistical method to score the disease-relatedness of candidate genes with predictive features extracted from gene networks and annotation based on a set of training disease genes using frequent item set mining algorithm (Figure S1). For schizophrenia, we will first curate a set of genes, *D*, known to be associated with this disease from the SZGR database (JIA *et al.* 2010). Given *D* and the set of all known genes *G* (from GENCODE v19), we obtain the background genes $B = G - D$. First, from *D* we will extract the predictive features – i.e., the frequent combinations of either the direct neighbors of schizophrenia genes in the functional linkage network (LINGHU *et al.* 2009) (with the functional linkage weight cutoff = 1) or the gene ontology (GO) terms of schizophrenia genes – using the frequent item set mining algorithm (AGRAWAL *et al.* 1995) (with the support = 0.093) . GO terms of schizophrenia genes include not only annotated GO terms but also their ancestors GO terms along the paths of the "is a" relationship in the GO hierarchy structure. The considered predictive features are limit to frequent combinations with sizes no greater than 3 to avoid redundancy and intensive computation. Then, each predictive feature will be scored by the frequency with which it appears in *D* and *B*:

$$S_f = (F_D/N_D)/(F_B/N_B) \qquad (1),$$

in which F_D is the frequency with which the predictive feature, *f*, occurs in *D* and N_D the number of genes in *D*. F_B and N_B have similar meanings. Next, we will score the schizophrenia risk gene candidates. For each candidate, we identify all the predictive features that it contains and assign it the highest score of its predictive features.

Since the network and annotation features are treated separately, the final score is a combination of the two:

$$S_g = \alpha S_f^{(n)} + (1-\alpha) S_f^{(a)} \quad (0 < \alpha < 1) \qquad (2),$$

in which $S_f^{(n)}$ and $S_f^{(a)}$ are the highest network- and annotation-based scores, respectively, assigned to the candidate gene and α is a coefficient, controlling the amount of influence that these two scores have, relative to each other, on the final gene score. Setting α at 0.4 yields the best predictive power according to the result of our evaluation (Figure S2). Every candidate is scored using gene sets B and D excluding the candidate to avoid biased scoring.

Identification of spatiotemporal gene expression patterns

Our gene set expression data from BrainSpan for a gene set is a 3-dimensional data. To summarize this data, we first binarized each gene's activity at different brain locations and time stages (Figure S11). In the binary matrix for each gene, a cell value is set 1 if the expression value at the corresponding time stage and brain location is higher than the sum of the mean and standard variation of gene expression values among different time stages; otherwise, it is set 0. We then summarize the activities of all genes in the gene set by adding up all binary matrixes and obtain the spatiotemporal matrix. From this matrix, we can observe the spatiotemporal gene expression patterns for the gene set. Each cell value in the matrix represents the ratio of genes that are active to the total number of genes in the gene set at the corresponding brain location and time stage. This method can be used to detect suppressed activities as well.

SUPPLEMENTARY FIGURES

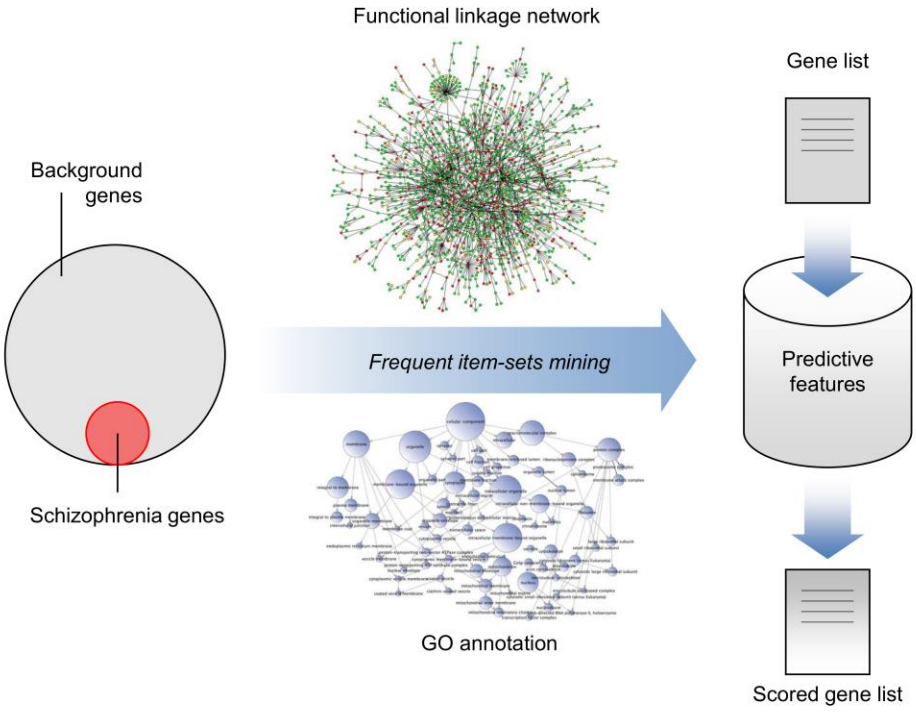


Figure S1. The schematic of scoring schizophrenia risk gene candidates.

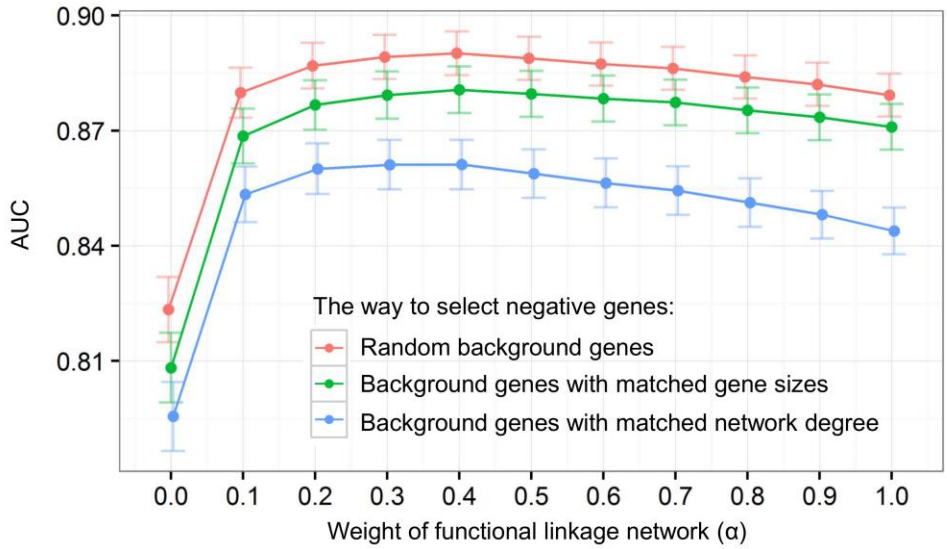


Figure S2. Evaluation different network weights in the mixture model. Using different negative gene sets in training, a network weight (α) set at 0.4 consistently yielded the best performance (see equation (2) in Supplemental Material), while the performance of scoring monotonically decreases when the network weight deviates from 0.4.

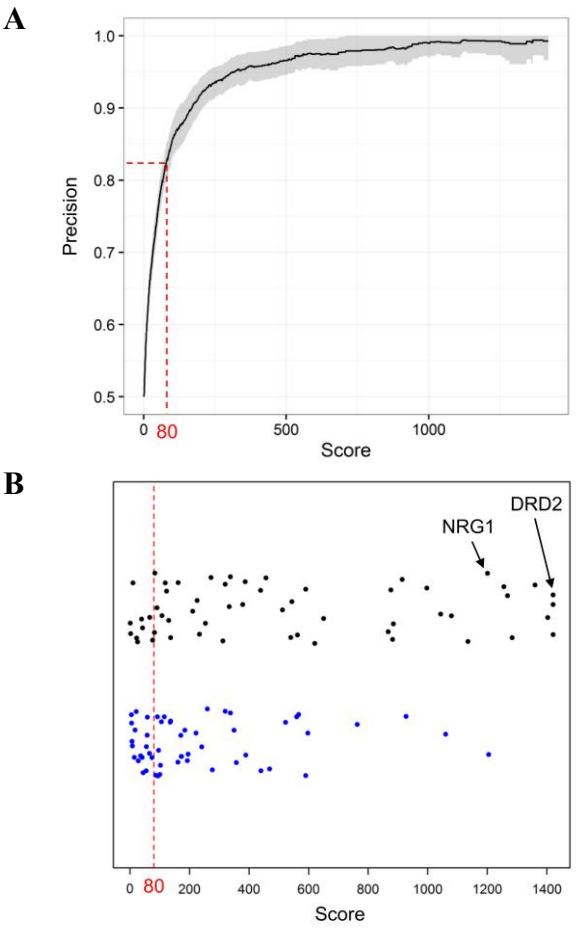


Figure S3. Evaluation of score cutoff for schizophrenia risk gene prediction. (**A**) Precision of prediction. The prediction precision is evaluated based on the classification test mentioned in the section "Evaluation of schizophrenia gene scoring". The dark gray area surrounding the black curve indicates the 95% confidence interval. A score cutoff at 80 can achieve a high prediction precision of 82.6%. (**B**) Sensitivity of prediction. Black and blue dots represent the schizophrenia genes from the training gene set and other schizophrenia genes with literature support (Table S2), respectively. A score cutoff at 80 can achieve a prediction sensitivity of 83.9% and 68.5% for these two schizophrenia gene sets respectively. DRD2 and NRG1, two widely recognized schizophrenia genes, achieved high scores.

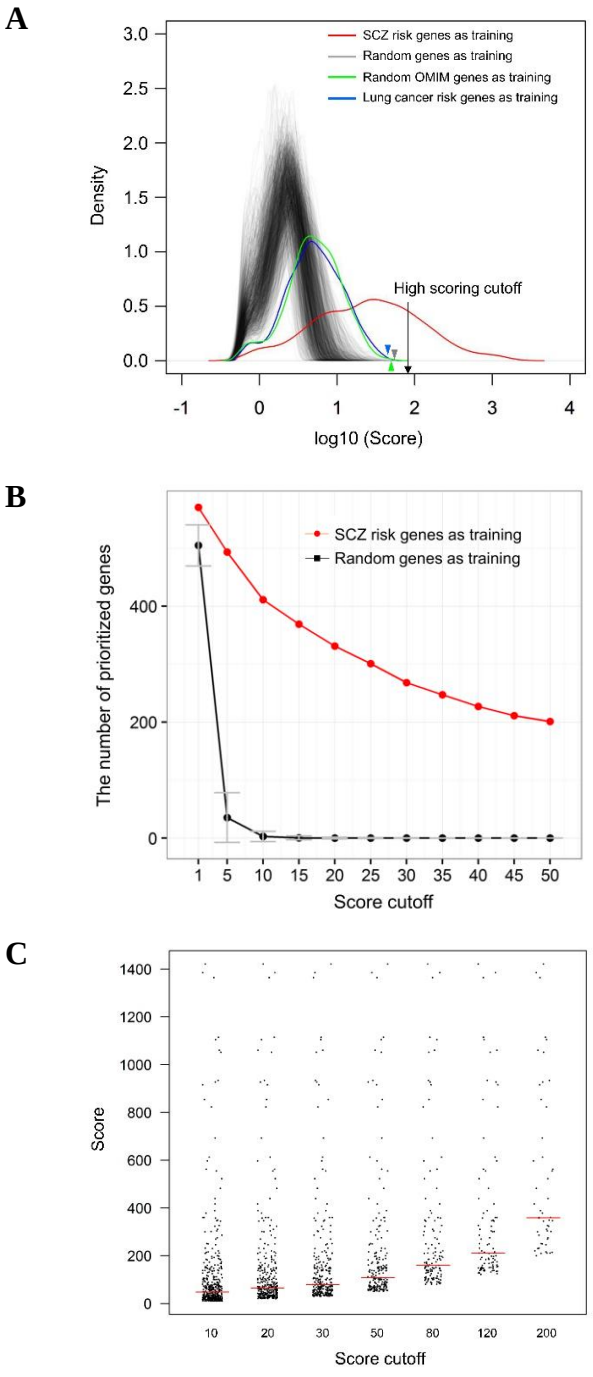


Figure S4. The statistical properties of the scoring method on candidate genes. Different from the evaluation of background genes, here we assess the statistical properties of our scoring method on 585 candidate genes. (**A**) Score distributions. In addition to the set of 56 known

schizophrenia genes, we also used three other size-matched gene sets – lung cancer risk genes collected from MalaCards (RAPPAPORT et al. 2014), OMIM genes, and random genes – for training to score 585 candidate genes. We compared the derived score distributions to assess the significance of the scores obtained by using the schizophrenia training gene set. Of random genes, we generated 1,000 different sets, each of which produced a separate score distribution. The highest scores of 585 candidate genes derived from training with lung cancer risk genes, OMIM genes, and random genes are 38.0, 44.7, and 47.8 (denoted by the blue, green, and grey triangles), respectively, all of which are far lower than the score cutoff (the black arrow). (**B**) The number of prioritized genes using different score cutoffs. When using random genes as training genes, the vast majority of candidate genes have scores lower than 5 and only a few have scores higher than 10. The error bar indicates the standard deviation. (**C**) The distribution of scores in a prioritized gene set above a score cutoff (using schizophrenia training genes). The red line indicates the median score, which is 160 in our prioritized gene set (cutoff at 80).

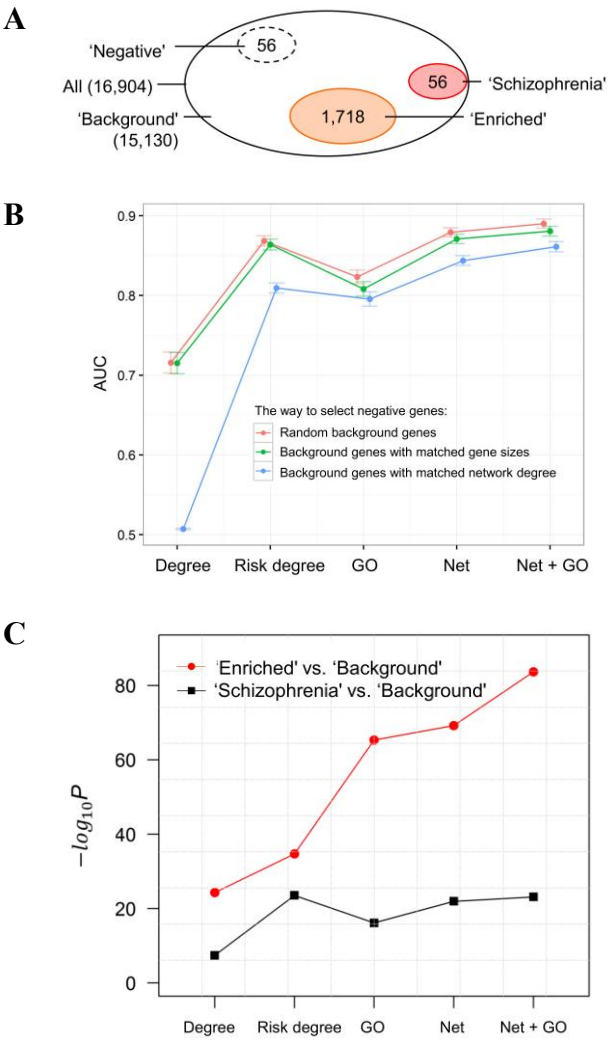


Figure S5. Evaluation of schizophrenia risk gene prediction. (**A**) Gene sets used for evaluation. The 'schizophrenia' gene set consists of the 56 known schizophrenia risk genes for training, while the 'enriched' gene set consists of other 1,718 genes implicated by rare mutations (PURCELL *et al.* 2014). (**B**) Classification test. Schizophrenia and each of 1,000 randomly generated negative gene sets were used to calculate the area under the receiver operating characteristic curve (AUC) with five different scoring schemes and three different ways to select negative genes. The error bar indicates the 95% confidence interval. For all methods, there is a decrease in AUC when matched genes were selected as negative genes because matched genes

have network features less distinguishable from schizophrenia seed genes. (**C**) Method comparison. For a scoring method, the Wilcoxon rank sum test was used to compare the scores of both the schizophrenia and the enriched gene sets with those of the background gene set.

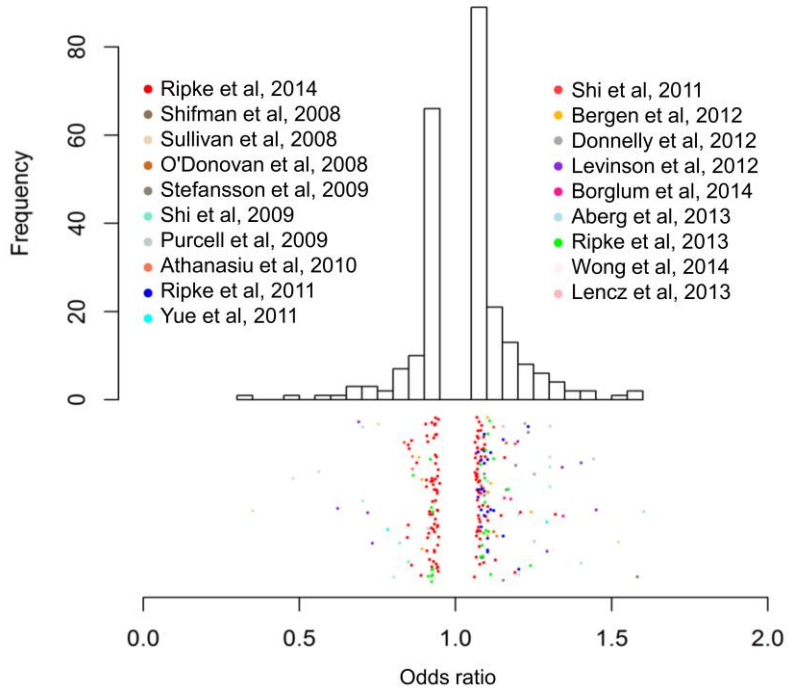


Figure S6. Odds-ratio distribution of schizophrenia GWAS variants. Each dot represents a schizophrenia GWAS variant annotated in a study. Most schizophrenia GWAS variants have odds ratio close to 1.

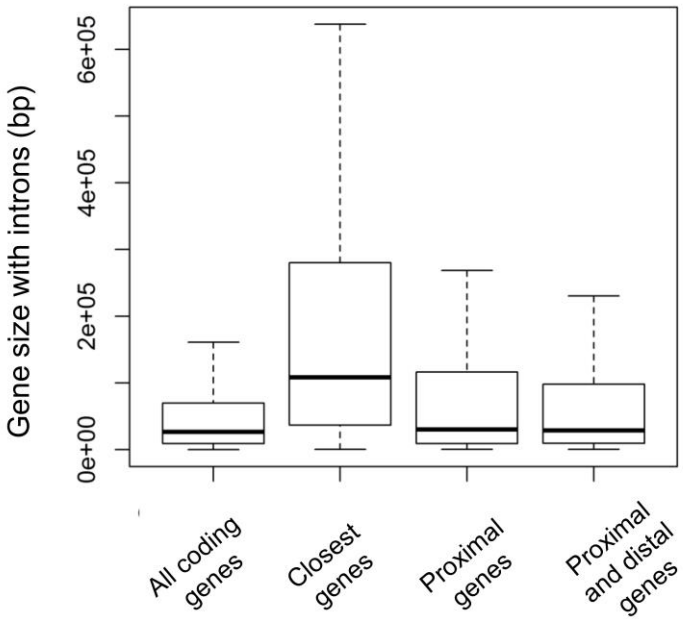


Figure S7. The size distributions of candidate genes identified by different strategies. 261 schizophrenia SNPs were used to identify candidate risk genes using different strategies. The strategy that considers only closest genes to the schizophrenia SNPs tends to include large genes compared to all protein-coding genes in the human genome. Our strategy that considers genes proximal and distal to risk regions includes candidate genes with a size distribution similar to that of all protein-coding genes.

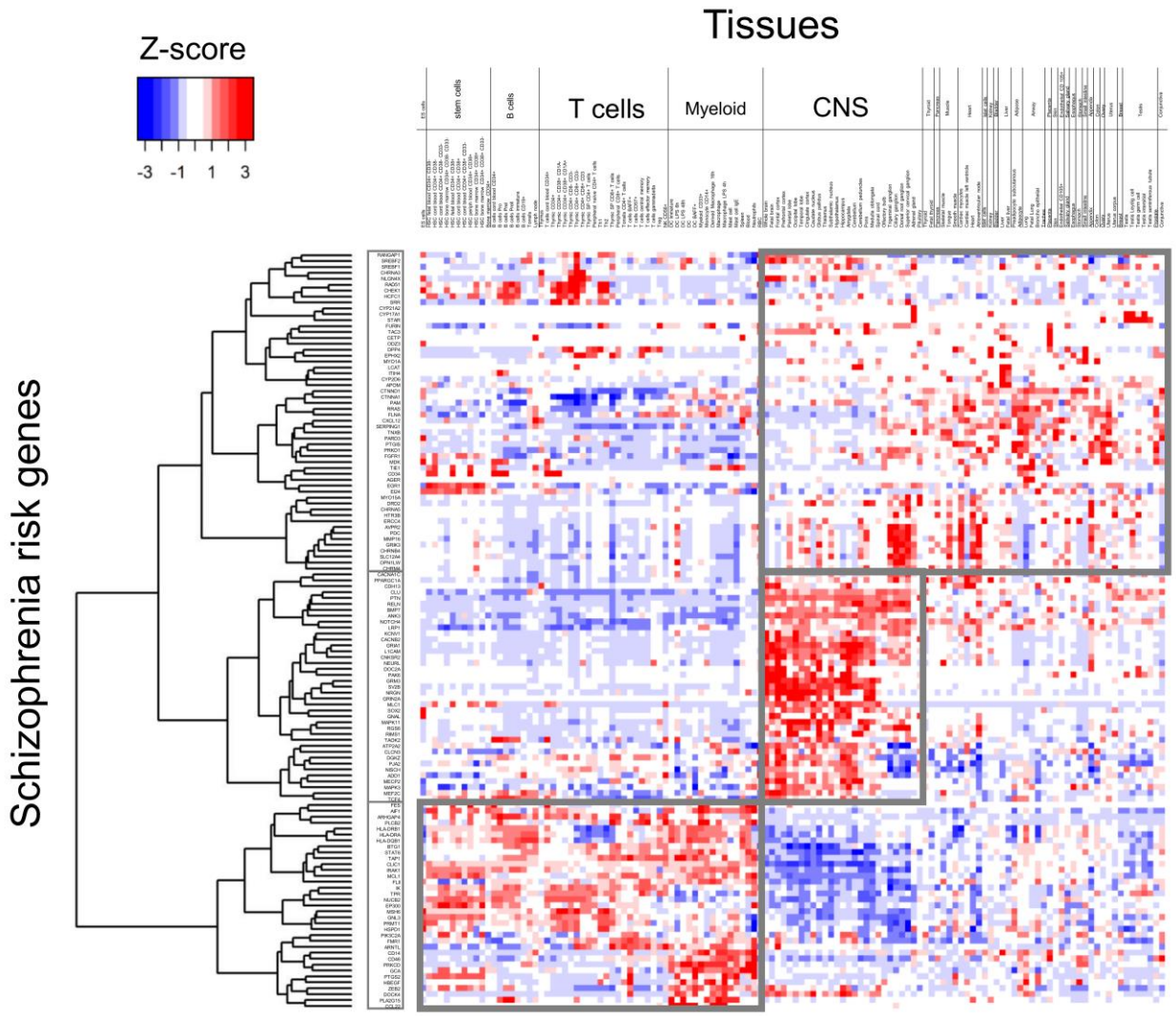
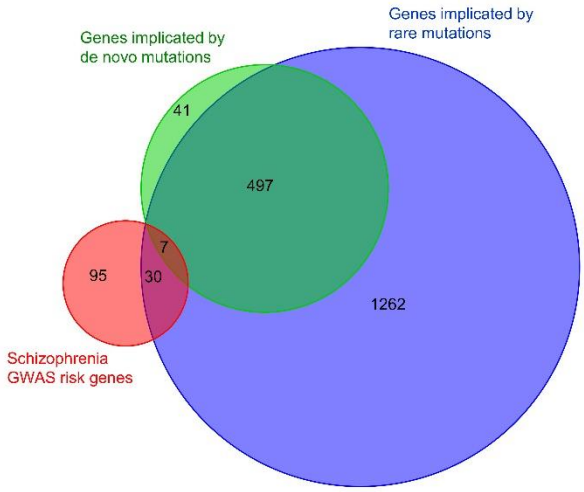
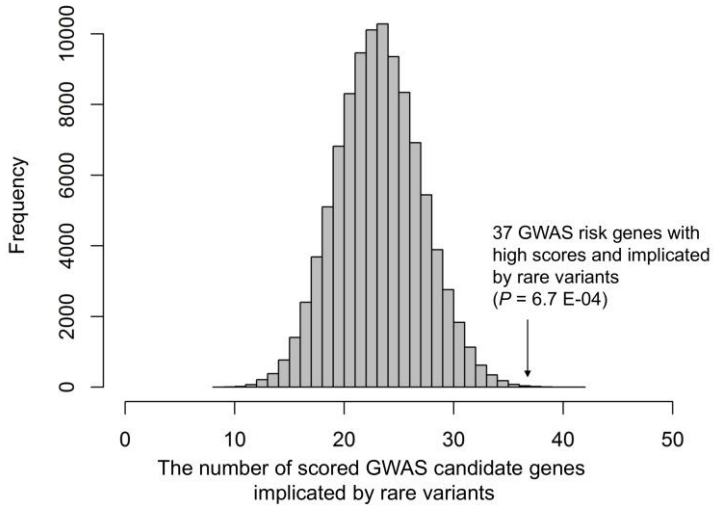


Figure S8. Expression of schizophrenia risk genes in different tissues. Genes were clustered according to their expressions among different tissues. Three grey boxes indicate gene clusters and corresponding tissues in which they are more transcriptionally active. 4 out of 132 schizophrenia risk genes were not included due to lack of gene expression data from the Gene Enrichment Profiler (http://xavierlab2.mgh.harvard.edu/EnrichmentProfiler/).

A


Genes implicated by
de novo mutations
Genes implicated by
rare mutations
41
497
7
95
30
1262
Schizophrenia
GWAS risk genes

B


Frequency
0
2000
4000
6000
8000
10000
37 GWAS risk genes with
high scores and implicated
by rare variants
(P = 6.7 E-04)
0
10
20
30
40
50
The number of scored GWAS candidate genes
implicated by rare variants

C

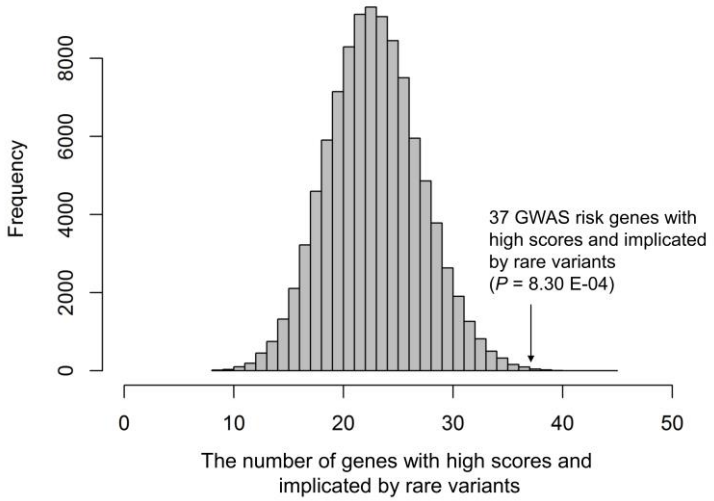

Frequency
0
2000
4000
6000
8000
37 GWAS risk genes with
high scores and implicated
by rare variants
(P = 8.30 E-04)
0
10
20
30
40
50
The number of genes with high scores and
implicated by rare variants

Figure S9. Schizophrenia genetic architecture. (A) Overlaps among genes implicated in schizophrenia by common variants (through GWAS), rare variants, and *de novo* mutations. Shown in the figure are the numbers of genes in different areas of the Vann diagram. **(B)** The evaluation of association between high scoring genes linked to schizophrenia GWAS loci and genes implicated by rare variants. We performed a permutation test with 100,000 iterations to construct the null distribution. In each iteration, a set of 132 candidate genes were randomly selected from 585 scored candidate genes linked to schizophrenia GWAS loci. The null hypothesis is that for genes linked to schizophrenia GWAS loci, the enrichment of schizophrenia genes implicated by rare variants in high scoring genes is not greater than the enrichment in random genes. **(C)** The permutation test of the significance of overlap between schizophrenia GWAS risk genes and schizophrenia risk genes implicated by exonic rare mutations. Genome-wide, there are 3,303 genes with high scores, including 132 ones connected to GWAS signals. To eliminate the confounding effect of "high scoring" (genes implicated by rare mutations are also enriched with high-scoring genes), we performed a permutation test with 100,000 iterations to construct the null distribution. In each iteration, a set of 132 genes were randomly selected from those 3,303 high-scoring genes. The null hypothesis is that for schizophrenia risk genes implicated by rare variants, the enrichment of schizophrenia GWAS risk genes is not greater than the enrichment of genes with high scores.

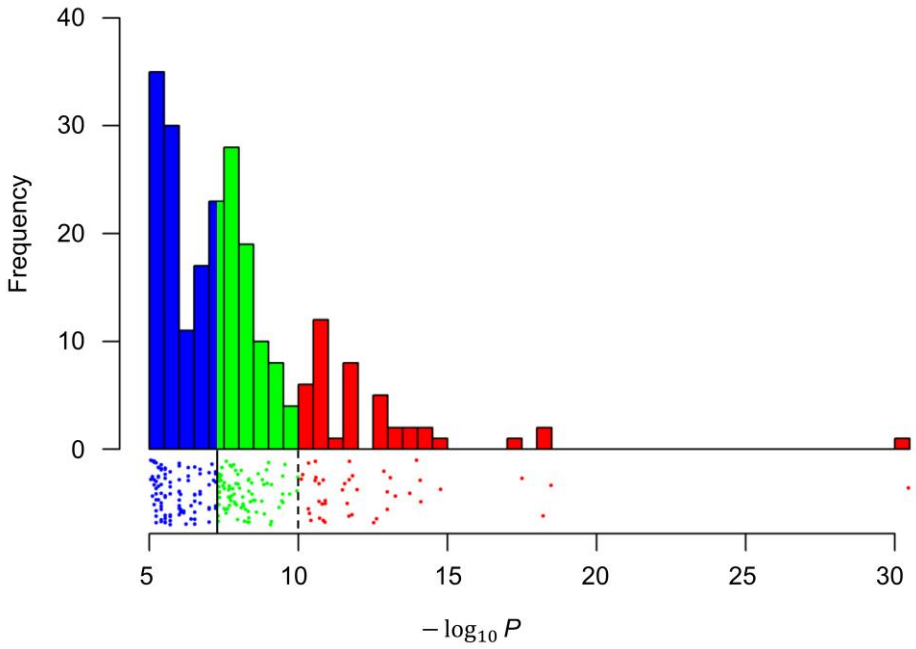


Figure S10. *P*-value distribution of schizophrenia GWAS variants. The solid and the dashed lines, representing the genome-wide significant level (*P*-value = 5×10^{-8}) and a much more highly significant association (*P*-value = 10^{-10}), divide schizophrenia GWAS variants into 3 classes with weak (blue), moderate (green), and strong (red) GWAS signals.

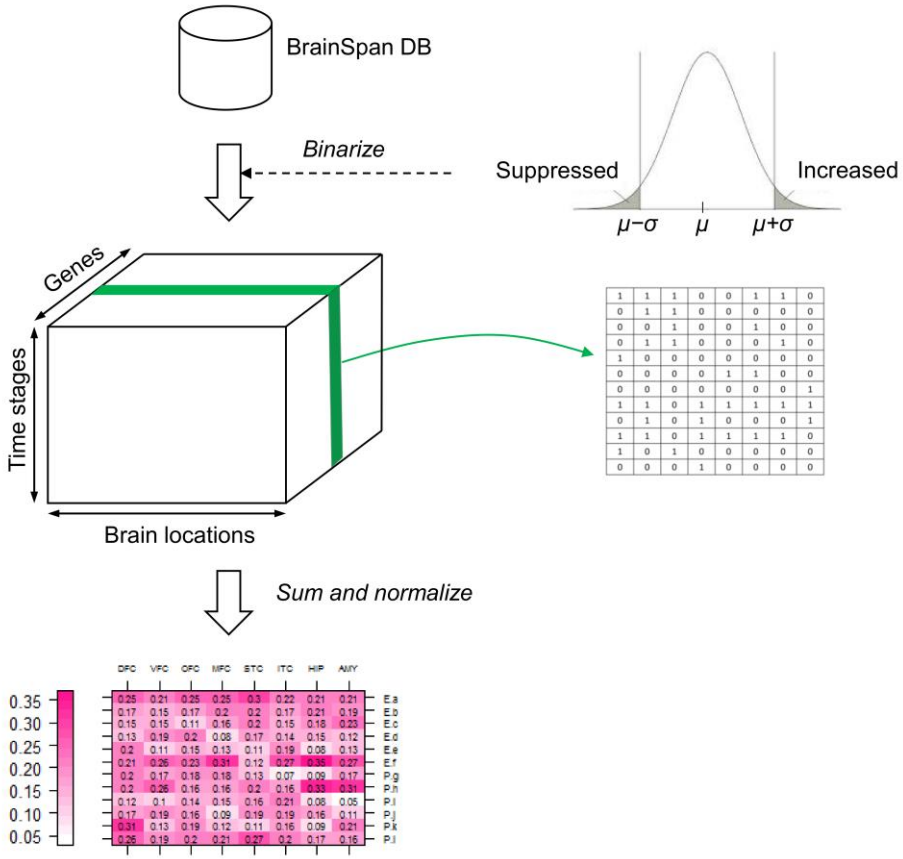


Figure S11. The schematic of identifying spatiotemporal gene expression patterns.

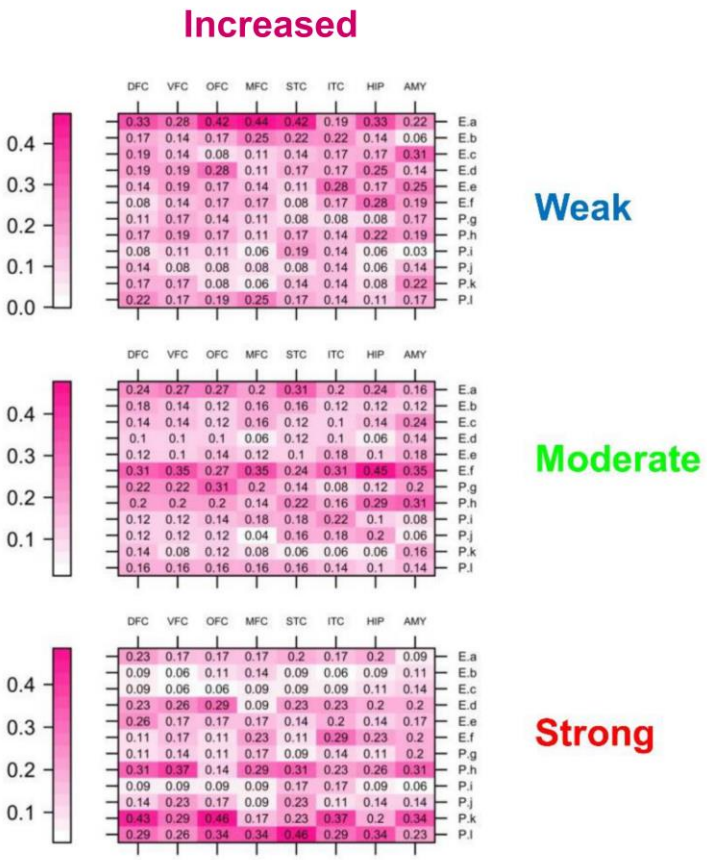


Figure S12. Spatiotemporal expression patterns of genes prioritized with controlled training genes during brain development. Similar to Figure 5, this figure shows the spatiotemporal expression patterns of prioritized genes in different classes of association strength based on a 'controlled' set of training genes that do not show the similar spatiotemporal expression patterns. This analysis was carried out in two steps: First, for each class of association strength (Weak, Moderate, or Strong), we re-compiled a controlled training gene set by selecting training genes not in high correlation in spatiotemporal expression with the prioritized genes in the corresponding class (Table S9). Second, we re-scored candidate genes in each class of association strength using the corresponding controlled training gene set, and then prioritized the same number of genes with the original prioritized genes in each association class. The

spatiotemporal expression patterns of those new prioritized genes by using controlled training genes exhibit the same characteristics with the previous result in Figure 5.

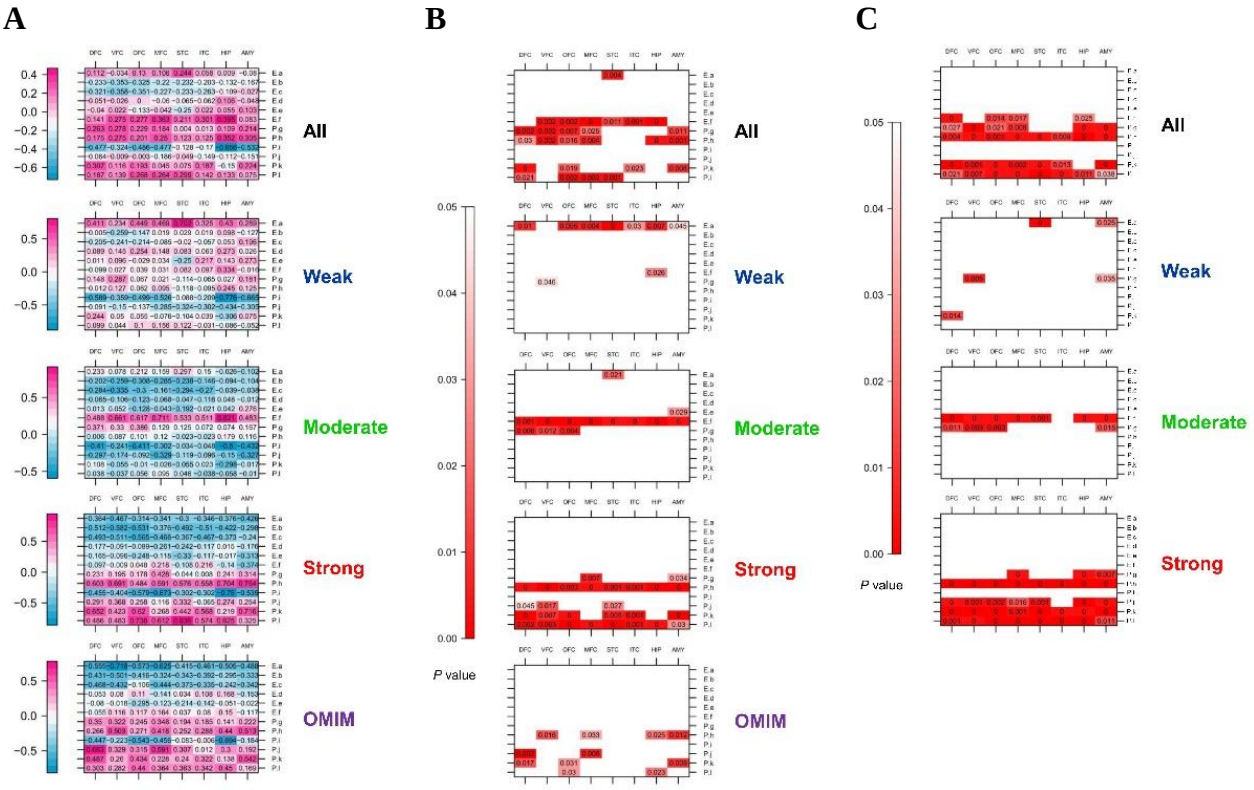


Figure S13. The z-score transformation and the significance tests of the transcriptional activities of schizophrenia risk genes during brain development. (A) Spatiotemporal expression based on the z-score transformation. For every risk gene, we standardized its gene expression value (i.e., z-scores) at each brain location and time stage by subtracting the mean of its expression values across stages from the original expression value and then dividing the result by the standard deviation across stages. The value in each cell of a heat map is the average z-score of genes in the corresponding gene set for the corresponding brain region and time stage. Positive average z-scores indicates that the transcriptional activities of schizophrenia risk genes tend to be increased, while negative average z-scores indicates that the transcriptional activities of schizophrenia risk genes tend to be suppressed. **(B)** The statistical significance of the increased transcriptional activities of schizophrenia risk genes assuming they do not tend to be more active at any stage. For each combination of time stage and brain region with a positive

average z-score, we assessed the significance of the active activity for a gene set by performing a permutation test with 100,000 iterations. In each iteration, we calculated the average z-scores of the same set of schizophrenia risk genes with their expression values randomly permuted by stages. The null hypothesis is that schizophrenia risk genes do not tend to be more transcriptionally active (i.e., have a higher average z-score) in the corresponding brain region and time stage compared with the same region but other stages. **(C)** The statistical significance of the increased transcriptional activities of schizophrenia risk genes compared with the rest un-prioritized candidate genes. For each combination of time stage and brain region with a positive average z-score, we assessed the significance of the active activity for a gene set by performing a permutation test 100,000 iterations. In each iteration, we calculated the average z-scores for a random size-matched set of un-prioritized genes in the same association class with the corresponding gene set. The null hypothesis is that schizophrenia risk genes do not tend to be more transcriptionally active than un-prioritized candidate genes in the corresponding time stage and brain region.

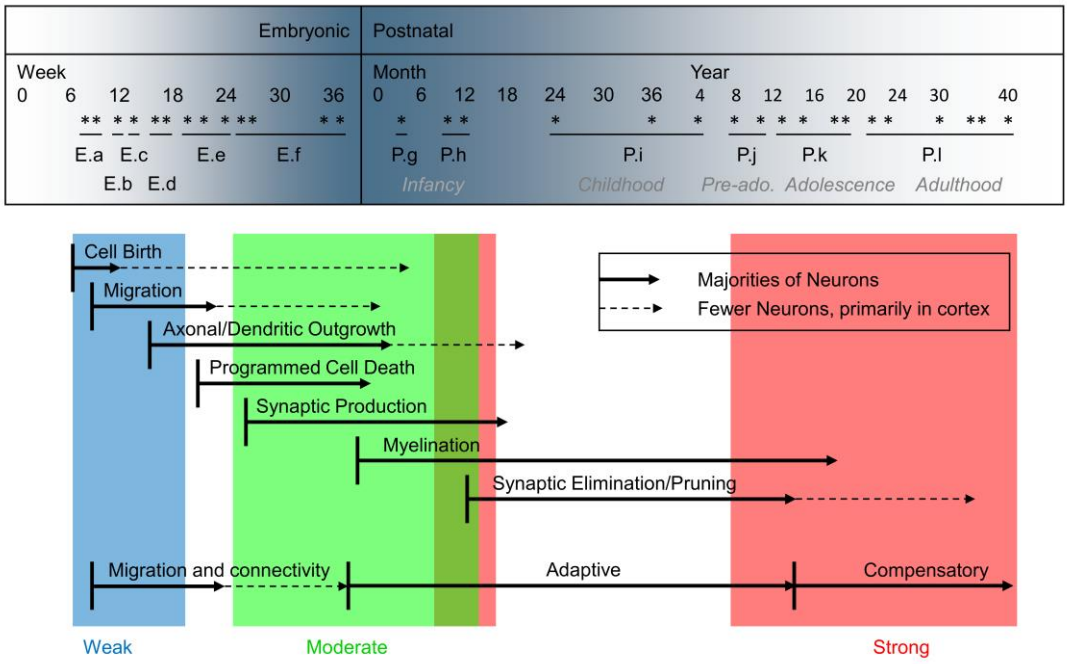


Figure S14. Expression of schizophrenia risk gens during brain development. The stages of brain development were defined by S. L. Anderson (ANDERSEN 2003).

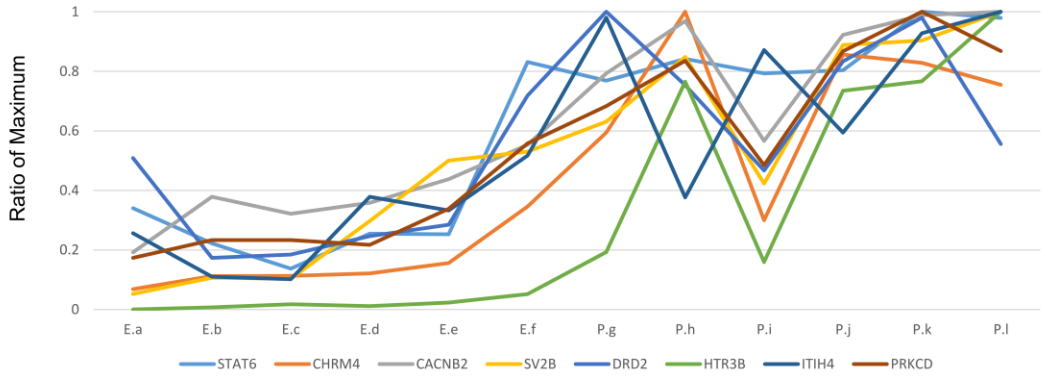


Figure S15. Transcriptional activities of 8 schizophrenia risk genes with strong association during brain development. Our post-GWAS analysis identified 8 genes in strong association with schizophrenia. Rebased relatively to the maximum expression level of each gene, their transcriptional activities in dorsolateral prefrontal cortex during brain development are plotted together. See Figure S14 for the meaning of the time stage symbols.

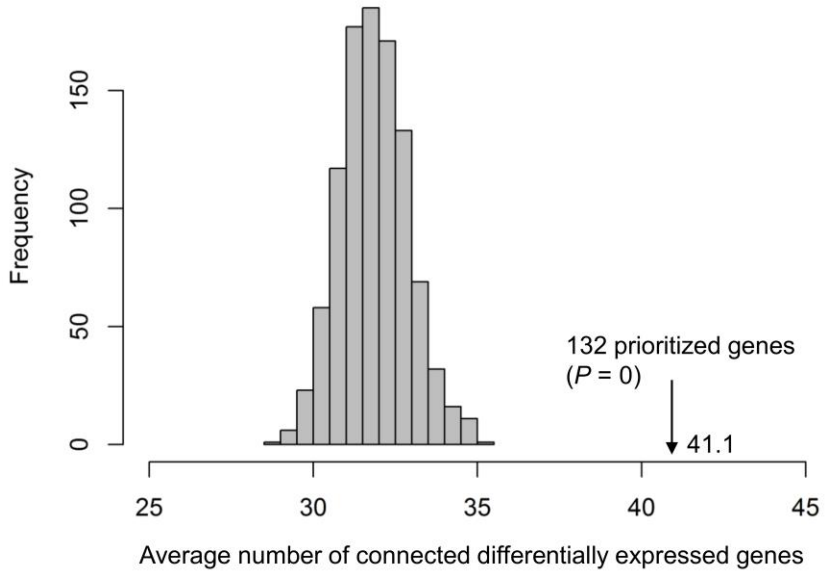


Figure S16. Statistical test for association in the gene network. In our permutation tests, the null distribution was constructed by 1,000 replicates of random sampling of 132 random genes in the functional linkage network with degree similar to 132 schizophrenia risk genes. The result indicates that schizophrenia risk genes are more likely to be functionally associated with differentially expressed genes. For example, *HBEGF* is a homologue of epidermal growth factor (*EGF*), which has been implicated in the etiology of schizophrenia (FUTAMURA *et al.* 2002). Compared to random genes with similar network degrees, *HBEGF* has significantly more (114 vs. 58 on average) differentially expressed genes among its direct neighbors in the network. In particular, pathway analysis showed that there is a significant enrichment of cytokine-cytokine receptor interaction pathway among its 114 differentially expressed neighbors ($P = 1.84\times10^{-21}$). Cytokines are crucial mediators of neurodevelopment processes, and its dysregulation has been connected to the pathogenesis of schizophrenia (WATANABE *et al.* 2010). However, *HBEGF* itself may not be differentially expressed in schizophrenia (FUTAMURA *et al.* 2002).

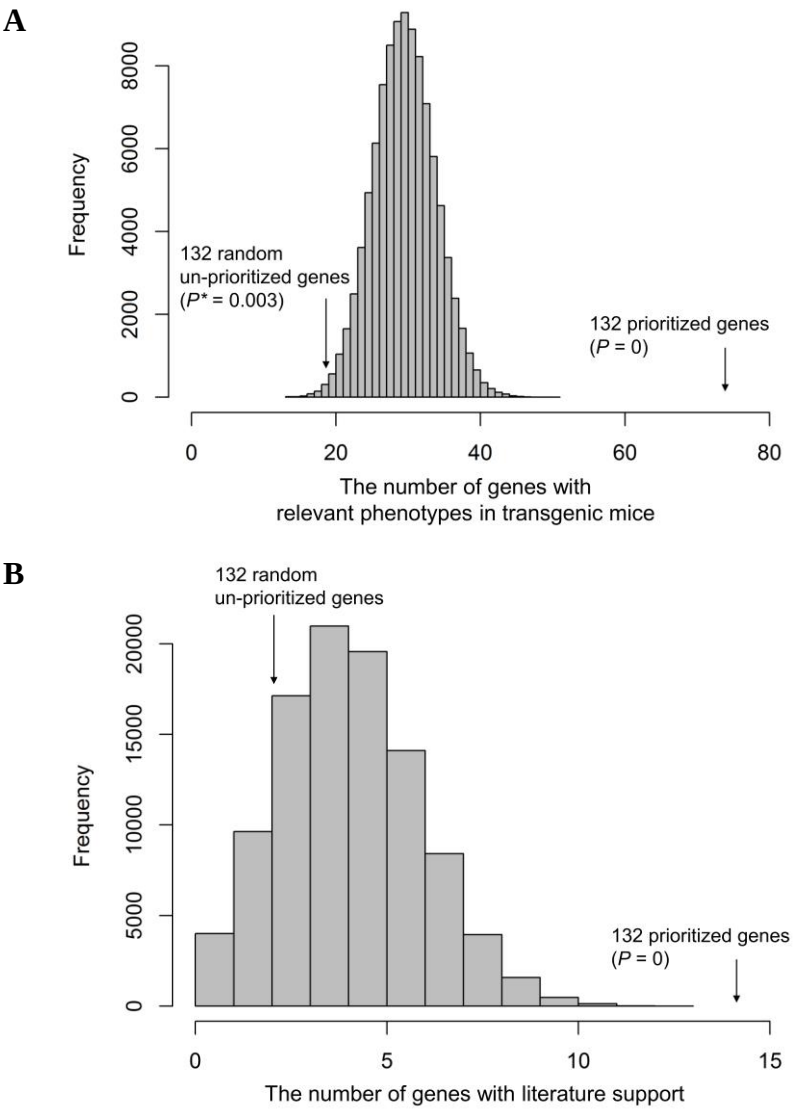


Figure S17. The comparison between 132 prioritized genes and the rest 511 candidate genes. We compared prioritized genes and the rest candidate genes in their association with schizophrenia-related genes. Since some un-prioritized candidates cannot be scored, the Fisher's exact test used in Table S5 with 585 scored candidate genes as the background is not applicable in this case. Here we assessed the association with a schizophrenia-related gene set using the permutation test with 100,000 iterations. In each iteration, a random set of 132 genes were selected from 643 candidate genes. **(A)** The number of genes with relevant phenotypes in transgenic mice. Consistent with the result in Table S5, genes with relevant phenotypes in

transgenic mice (mouse knock-out genes) are over-represented in our prioritized genes ($P = 0$). The average number of mouse knock-out genes in random 132 un-prioritized candidate genes is 18.8, indicating that mouse knock-out genes are even under-represented ($P^* = 0.003$, * denotes under-representation) in the rest 511 candidates, considering all 643 candidate genes as the background. **(B)** The number of genes with literature support. Consistent with the result in Table S5, genes with literature support are over-represented ($P = 0$) in our prioritized genes, which is not the case in the un-prioritized candidates.

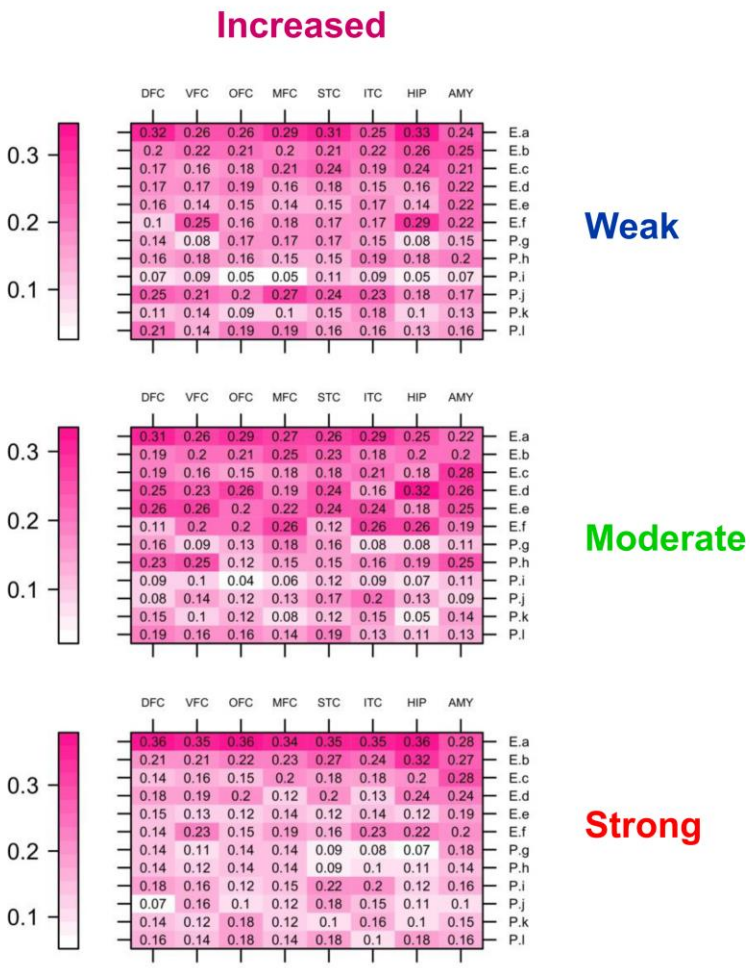


Figure S18. Spatiotemporal expression patterns of un-prioritized candidate genes during brain development. The interpretation of this figure is the same as Figure 5. The un-prioritized candidate genes do not exhibit the difference in transcriptional activities for different association strengths as observed for prioritized genes.

SUPPLEMENTARY TABLES

Table S1. 56 schizophrenia training genes.

Gene symbol	Entrez gene ID	Reference[1-4]
AKT1	207	(1)
APOE	348	(1)
APOL2	23780	(2)
APOL4	80832	(2)
CHI3L1	1116	(2)
CHRNA7	1139	(1)
COMT	1312	(1)
CRP	1401	(4)
CYFIP1	23191	(3)
DAO	1610	(1)
DAOA	267012	(1)
DISC1	27185	(1)
DLG1	1739	(3)
DLG2	1740	(3)
DRD1	1812	(1)
DRD2	1813	(1)
DRD3	1814	(2)
DRD4	1815	(1)
DTNBP1	84062	(1)
EGF	1950	(4)
EHMT1	79813	(3)
ERBB4	2066	(1)
FEZ1	9638	(1)
GABRB2	2561	(1)
GAD1	2571	(1)
GRIK4	2900	(1)
GRIN2B	2904	(1)
GRM3	2913	(1)
HP	3240	(1)
HTR2A	3356	(1)
IL18	3606	(4)
IL1B	3553	(1)
MTHFR	4524	(1)
MUTED	63915	(1)
NPAS3	64067	(1)
NRG1	3084	(1)
NRGN	4900	(3)
NRXN1	9378	(2)
OFCC1	266553	(1)
OPCML	4978	(1)

PLXNA2	5362	(1)
PPP3CC	5533	(1)
PRODH	5625	(1)
RELN	5649	(1)
RGS4	5999	(1)
RPGRIP1L	23322	(1)
RTN4R	65078	(2)
SHANK3	85358	(2)
SLC18A1	6570	(1)
SLC1A1	6505	(2)
SLC6A4	6532	(1)
TNF	7124	(4)
TP53	7157	(1)
TPH1	7166	(1)
VIPR2	7434	(3)
ZNF804A	91752	(1)

Notes:

(1) denotes the manually curated 'core genes' from the SZGR database (JIA *et al.* 2010).

(2) denotes schizophrenia susceptibility genes cataloged in the OMIM database (MCKUSICK 2007).

(3) denotes well-accepted schizophrenia genes from recent genetics studies (HALL *et al.* 2015; KOTLAR *et al.* 2015).

(4) denotes schizophrenia genes with solid support from other sources (CANETTA *et al.* 2014; NAWA *et al.* 2014; BOSSU *et al.* 2015; LV *et al.* 2015).

Table S2. Schizophrenia genes with literature support.

Gene symbol[1]	Entrez gene ID	Score	PMID[2]
DLG4	1742	1204.5	21151988
*EGR1**	1958	1060.1	22691714
*FGFR1**	2260	926.8	23231877
OPRM1	4988	762.8	23560613
*CACNA1C**	775	597.1	24262814
THBS1	7057	589.9	22311024
IFNG	3458	566.2	22623148
GRB2	2885	559.7	21195589
*TCF4**	6925	522.1	21932083
CIT	11113	468.6	20084519
FYN	2534	439.5	23250004
*SREBF1**	6720	388.5	18936756
*RIMS1**	22999	357.2	22682706
*CLU**	1191	348.9	20738160
MTNR1A	4543	337.1	21526376
CNTNAP2	26047	319.2	23123147
PLAT	5327	276	21898905
PDE4A	5141	259.5	21898905
SIRT1	23411	240.8	20977650
HOMER2	9455	221.5	19914345
NDEL1	81565	194.7	20084519
LIF	3976	192.2	19879916
HDAC3	8841	183.9	20471694
NRXN3	9369	173.1	23306218
NRG3	10718	171.5	20713722
*MEF2C**	4208	170.5	23380319
*CHRNA5**	1138	161	21418140
HDAC4	9759	135.9	20471694
*ANK3**	288	135.1	23109352
MDGA1	266727	115.2	21146959
*SREBF2**	6721	105.2	18936756
*CACNB2**	783	101.7	24901509
*PARD3**	56288	100.2	22969987
HSPA1A	3303	95.5	23893339
DLGAP2	9228	93.7	24416398
KPNA3	3839	91.5	22960338
*FMR1**	2332	85.7	23838275
*NCAN**	1463	73.3	23795679

STON2	85439	64.9	21407139
*ADAMTSL3**	57188	58.1	21239144
ASAH1	427	57.6	21375364
SIGMAR1	10280	54.9	21549171
GSTT1	2952	54.5	23107768
CMYA5	202333	43.8	23778016
SEMA3D	223117	40.9	20684831
LASP1	3927	35.6	23040864
*VRK2**	7444	27.8	23102693
*CNNM2**	54805	20.7	24160291
AMACR	23600	16.5	20875727
*TSNARE1**	203062	14.4	24166486
SETD1A	9739	7.8	24853937
*BRD1**	23774	6.3	19693800
*PKNOX2**	63876	4.9	22648509
*LSM1**	27257	4.8	24035562

Notes:

1. * denotes genes which are one of 643 candidate genes.
2. We only consider publications more recent than the functional linkage network (LINGHU *et al.* 2009) and exclude any of our collected 25 GWAS studies from the list.

Table S3. The summary data of 176 schizophrenia risk regions and their linked candidate genes.

Chromosome region (chr:start-end)	*P*-value[1]	GWAS SNPs	Genes[2]
6:28179560-28712247	3.48E-31	rs1635, rs115329265	*ZKSCAN3* (63.5), *GPX5* (58.3), *GPX6* (35.5), *ZSCAN12* (19.8), *PGBD1* (9.3), *ZSCAN23* (8.2), *ZBED9* (4.9), *ZKSCAN4* (2.9), *ZSCAN9* (2.9), *ZSCAN31* (2.9), *ZSCAN26*, *NKAPL*
1:98325796-98559093	3.36E-19	rs1625579, rs1702294, rs1198588	*DPYD* (72.2)
10:104423800-105059896	6.20E-19	rs7897654, rs7085104, rs55833108, rs11191419, rs7907645, rs7914558, rs11191580, chr10_104957618_I	*CYP17A1* (416.5), *NEURL* (127.1), *ARL3* (76.7), *INA* (64.7), *ACTR1A* (30.1), *TRIM8* (25.3), *CNNM2* (20.7), *FBXL15* (12.4), *SFXN2* (11.5), *NT5C2* (6.4), *PCGF6* (5.5), *AS3MT* (1.8), *C10orf95* (0.8), *C10orf32*, *CALHM2*, *WBP1L*
12:2292690-2523772	3.22E-18	rs1006737, rs4765905, rs2007044, rs2239063	*CACNA1C* (597.1), *ITFG2* (3.6)
8:143306126-143340566	1.74E-15	rs4129585	*TSNARE1* (14.4), *LY6K* (6.4)
4:103112470-103198082	7.98E-15	rs35518360	*SLC39A8*
7:1909865-2190100	8.20E-15	rs12666575, rs6461049, chr7_2025096_I	*LFNG* (78.9), *NUDT1* (33.1), *MAD1L1* (31.2), *SNX8* (26.7), *TMEM184A* (18.1), *EIF3B* (8.8), *FTSJ2* (7.4), *CHST12* (6.8), *IQCE* (5.3), *GRIFIN*, *BRAT1*
5:60484179-60843706	1.10E-14	rs4391122, rs7709645, rs171748	*ZSWIM6* (31.8)
12:123466111-123758235	1.86E-14	rs11532322, rs2851447	*EIF2B1* (33.8), *ABCB9* (26.8), *GTF2H3* (20.1), *PITPNM2* (19.1), *OGFOD2* (11.1), *MPHOSPH9* (8.5), *VPS37B* (8.0), *CDK2AP1* (6.3), *C12orf65* (5.6), *SBNO1* (4.7), *SNRNP35* (3.6), *ARL6IP4* (1.0), *DDX55* (0.9)
2:200696352-201022952	5.65E-14	rs2949006, chr2_200825237_I	*FTCDNL1* (11.6), *SGOL2* (8.6), *TYW5* (6.3), *C2orf69* (1.3), *C2orf47* (1.1)
15:91412848-91429042	8.30E-14	rs4702	*FURIN* (326.3), *FES* (286.1), *SV2B* (107.8), *MAN2A2* (6.6)
3:36843149-36945794	1.05E-13	rs4624519, rs75968099, rs6550435	*TRANK1* (1.9)
14:103991478-104121939	1.36E-13	rs12887734	*KLC1* (42.8), *CKB* (42.1), *PPP1R13B* (36.3), *BAG5* (18.2), *TRMT61A* (2.6), *APOPT1*
15:78785544-78930510	2.44E-13	rs8042374, rs190065944	*CHRNA3* (323.5), *CHRNB4* (269.3), *CHRNA5* (161.0), *CTSH* (70.9), *IREB2* (59.4), *WDR61* (35.9), *PSMA4* (27.4), *HYKK* (2.9)
7:110843795-111092478	3.03E-13	rs13240464	*IMMP2L* (153.6), *DOCK4* (90.2), *LRRN3* (31.3)
11:130714613-130894131	1.09E-12	rs10894294,	*SNX19* (24.2)

		rs10791097, rs7940866	
2:185545033-185926285	1.53E-12	rs4380187, rs11693094, rs1344706	*ZNF804A* (0.9)
X:21193426-21569920	1.61E-12	rs1378559	*CNKSR2* (83.6)
10:18680963-18782777	1.97E-12	rs17691888, rs7893279	*CACNB2* (101.7)
12:57428353-57682956	2.02E-12	rs324017, rs12826178	*LRP1* (692.5), *STAT6* (144.1), *MYO1A* (138.6), *TAC3* (113.4), *SHMT2* (56.4), *STAC3* (41.5), *NAB2* (29.6), *NXPH4* (5.3), *TMEM194A* (1.1), *R3HDM2* (0.8), *NDUFA4L2*
1:73729188-73991651	2.03E-12	rs10789369, rs12129573	
2:233559312-233806771	2.32E-12	rs6704768, rs778371	*GIGYF2* (79.1), *NGEF* (63.7), *INPP5D* (55.6), *EFHD1* (43.6), *ATG16L1* (22.6), *KCNJ13* (8.8), *C2orf82* (1.0)
11:124599063-124620147	2.80E-12	rs12807809, rs55661361	*NRGN* (90.1), *ESAM* (44.6), *VSIG2* (34.2), *SLC37A2* (5.3), *MSANTD2*, *TMEM218*
18:52747689-53804156	3.34E-12	rs17512836, rs1261117, rs4801131, rs78322266, rs9960767, rs72934570, rs715170, chr18_52749216_D, rs17594526, rs9636107, rs12966547	*TCF4* (522.1)
11:46339597-46548754	1.26E-11	chr11_46350213_D	*MDK* (197.0), *CHRM4* (160.6), *DGKZ* (148.9), *LRP4* (48.9), *CREB3L1* (24.0), *AMBRA1* (16.6), *CKAP5* (12.4), *ATG13*
3:180524764-180793432	1.30E-11	rs1879248, chr3_180594593_I, rs6782299	*FXR1* (31.1), *DNAJC19* (19.4), *CCDC39*
20:37422829-37485986	1.46E-11	rs6065094	*PPP1R16B* (26.0), *ACTR5* (4.9), *DHX35* (1.7)
2:57943567-58399905	1.47E-11	rs11682175, rs75575209, rs2312147	*FANCL* (28.5), *VRK2* (27.8)
15:84641125-84861420	1.62E-11	rs950169	*ADAMTSL3* (58.1)
2:198155170-198498316	2.06E-11	rs6434928	*HSPD1* (154.5), *SF3B1* (43.7), *HSPE1* (15.9), *MARS2* (15.4), *MOB4* (7.9), *ANKRD44* (3.5), *COQ10B* (1.4), *RFTN2*
22:41429084-41637119	2.07E-11	rs9607782	*EP300* (210.8), *RANGAP1* (91.0), *CHADL* (24.7), *PMM1* (23.8), *L3MBTL2* (4.8)
8:111460027-111630275	2.61E-11	rs36068923	*KCNV1** (88.8)
3:2530143-2576007	2.69E-11	rs17194490	*CNTN4* (184.2)
11:113317745-113424042	2.75E-11	rs2514218	*DRD2* (1420.9), *HTR3B* (125.1)
11:133808038-133853008	3.87E-11	rs75059851	*IGSF9B* (28.2), *B3GAT1* (12.0)
3:52638482-52960859	4.26E-11	rs2239547,	*PRKCD* (439.5), *ITIH4* (140.8), *NISCH* (138.6),

		rs2535627, rs4687552	*GNL3* (89.9), *SPCS1* (74.9), *ITIH1* (68.1), *ITIH3* (39.7), *TKT* (33.9), *NEK4* (31.9), *DNAH1* (27.1), *SFMBT1* (21.8), *PBRM1* (18.7), *GLT8D1* (3.0), *NT5DC2* (0.9), *TMEM110-MUSTN1*, *TMEM110*, *MUSTN1*
16:29923510-30018500	4.55E-11	rs12691307	*MAPK3* (915.7), *DOC2A* (249.6), *TAOK2* (93.7), *SEZ6L2* (69.7), *KCTD13* (37.1), *TBX6* (29.4), *TBC1D10B* (7.4), *ASPHD1* (6.6), *INO80E* (3.5), *FAM57B* (2.2), *C16orf92* (2.1), *HIRIP3* (1.7), *TMEM219*
22:39945791-40016767	4.73E-11	rs9611198, chr22_39987017_D	*GRAP2* (44.2), *SMCR7L* (40.1), *MGAT3* (32.7), *CACNA1I* (31.0), *RPS19BP1* (1.1), *FAM83F*
3:136097576-136473728	7.26E-11	rs7432375	*PCCB* (24.0), *STAG1* (7.5), *MSL2* (4.3)
5:151941138-152847217	1.06E-10	rs17504622, rs2973155, rs12522290, rs2910032, rs79212538, rs111294930	*GRIA1** (344.5), *NMUR2* (167.7)
X:68377126-68384580	1.98E-10	rs5937157	*PJA1* (2.7)
17:2096441-2220814	2.86E-10	rs4523957	*SRR* (104.4), *SMG6* (35.5), *HIC1* (35.5), *SGSM2* (6.7), *TSR1* (5.2)
7:86356183-86459347	3.33E-10	rs12704290	*GRM3* (378.2)
15:61831680-61909712	3.38E-10	rs12592967, rs4775413, rs12903146	*VPS13C**
1:44029353-44248230	3.39E-10	rs11210892	*TIE1* (167.5), *PTPRF* (71.3), *KDM4A* (58.6), *ARTN* (36.6), *IPO13* (18.5), *B4GALT2* (11.7), *ST3GAL3* (9.9), *MED8* (3.2), *SZT2*
19:19358672-19657632	3.63E-10	rs2905426, rs2905424	*NDUFA13* (77.3), *NCAN* (73.3), *GMIP* (57.1), *CILP2* (44.1), *PBX4* (37.6), *GATAD2A* (24.3), *ATP13A1* (16.5), *HAPLN4* (14.1), *TSSK6* (11.6), *TM6SF2* (10.1), *ZNF101* (2.9), *SUGP1* (2.6), *MAU2* (1.2), *YJEFN3*
1:149998923-150226321	4.49E-10	rs140505938	*MCL1* (182.2), *APH1A* (60.8), *ANP32E* (35.1), *VPS45* (34.4), *ADAMTSL4* (19.8), *OTUD7B* (15.6), *CA14* (9.8), *HIST2H3D* (9.7), *TARS2* (8.9), *PLEKHO1* (6.8), *PRPF3* (6.4), *C1orf54*
6:84279922-84409255	8.15E-10	chr6_84280274_D	*CYB5R4 (55.5)*, *SNAP91* (47.8)
1:2372321-2402499	8.70E-10	rs4648845	*PLCH2* (45.4)
16:13713926-13763942	1.01E-09	rs7405404	*ERCC4** (147.7)
7:104741842-105063372	1.13E-09	rs6466055	*SRPK2* (40.6), *KMT2E* (17.5), *RINT1* (11.0), *PUS7* (2.6), *ATXN7L1*
1:8392592-8605667	1.17E-09	chr1_8424984_D, rs2252865	*RERE* (32.3), *SLC45A1* (8.0)
12:110662327-110723245	1.40E-09	rs4766428	*ATP2A2* (215.8), *IFT81* (27.7), *TCTN1* (13.7)
4:170448464-170646003	1.47E-09	rs10520163	*CLCN3* (131.4), *NEK1* (34.4), *SH3RF1* (31.6), *C4orf27* (0.7)
6:96424182-96476028	1.64E-09	rs117074560	*FUT9* (5.6)
22:42315790-42689370	1.71E-09	rs6002655,	*CYP2D6* (275.3), *SREBF2* (105.2),

		rs1023500	*TNFRSF13C* (64.8), *NDUFA6* (38.0), *TCF20* (29.5), *WBP2NL* (23.2), *FAM109B* (18.8), *CENPM* (10.6), *MEI1* (9.9), *SEPT3* (8.5), *NAGA* (8.5), *SERHL2* (5.2), *SMDT1* (1.6), *SHISA8*
2:146416874-146441828	1.81E-09	chr2_146436222_I	
11:57369008-57681828	2.24E-09	rs9420	*SERPING1* (137.5), *CTNND1* (105.5), *TMX2* (43.2), *ZDHHC5* (29.0), *CLP1* (5.6), *C11orf31* (5.3), *MED19* (2.8), *BTBD18* (2.5), *YPEL4* (0.7), *SMTNL1*, *SLC43A3*
11:24367339-24412992	2.55E-09	rs11027857	*LUZP2** (1.3)
1:30412503-30437268	2.86E-09	rs1009080, rs1498232	
7:137039670-137085250	3.28E-09	rs3735025	*PTN* (174.6), *DGKI* (33.0)
9:84630452-84843011	3.61E-09	rs11139497	*SPATA31D1** (4.3)
1:243487861-244002773	3.73E-09	rs14403, rs77149735, rs6703335, rs1538774, rs10803138, chr1_243881945_I	*AKT3* (49.1), *SDCCAG8*
15:40566759-40602256	4.18E-09	rs1869901, rs56205728	*RAD51* (481.9), *PLCB2* (209.8), *PAK6* (149.7), *ANKRD63*
19:30981639-31038995	4.49E-09	rs2053079	*ZNF536* (38.7)
5:88637006-88825791	4.61E-09	rs16867576	*MEF2C** (170.5)
3:17668766-17888256	4.64E-09	rs4330281	*TBC1D5* (9.8)
5:137598340-137948140	4.67E-09	rs3849046, rs10043984	*EGR1* (1060.1), *CTNNA1* (106.0), *GFRA3* (77.1), *CDC25C* (54.9), *HSPA9* (42.8), *REEP2* (35.1), *ETF1* (21.0), *KDM3B* (3.8), *FAM53C*
14:99692254-99722290	4.80E-09	rs2693698	*BCL11B* (55.8)
14:72402431-72434002	4.86E-09	rs2332700	*RGS6* (99.8)
5:45291514-45393754	5.05E-09	rs1501357	*HCN1* (146.9)
8:60513088-60897721	5.97E-09	rs6984242	*CA8** (9.0), *TOX* (6.5)
2:72357336-72369876	7.39E-09	rs3768644	*CYP26B1* (33.7)
11:123394636-123395987	7.54E-09	rs77502336	*GRAMD1B* (1.0)
2:200154552-200314206	8.33E-09	rs6704641	*SATB2* (46.7)
2:193848340-194010884	8.41E-09	rs59979824, rs17662626	
4:176851045-176904037	9.47E-09	rs1106568	*GPM6A* (63.1)
8:4177231-4229379	1.06E-08	rs10503253, rs10503256	*CSMD1*
2:225334070-225467840	1.12E-08	rs11685299	*CUL3* (55.0), *DOCK10* (10.8)
8:89389761-89761163	1.22E-08	rs11995572, rs7819570, rs7004633	*MMP16** (303.7)
16:9875513-9971728	1.28E-08	rs9922678	*GRIN2A* (1385.0)
14:30189985-30192618	1.41E-08	rs2068012	*PRKD1* (1104.4)
3:63792668-64002247	1.43E-08	rs832187	*ATXN7* (76.9), *PSMD6* (42.5), *THOC7* (8.3), *C3orf49*
16:67989523-68305708	1.51E-08	rs8044995	*PLA2G15* (854.0), *LCAT* (148.3), *SLC12A4*

			(100.3), *NFATC3* (61.6), *DPEP3* (29.1), *PSMB10* (26.2), *SLC7A6* (21.6), *DUS2* (17.8), *PRMT7* (12.6), *DPEP2* (10.7), *ESRP2* (6.0), *DDX28* (2.8), *SLC7A6OS* (2.0)
2:149390792-149520186	1.59E-08	chr2_149429178_D	*EPC2* (7.4)
17:17760789-18036283	1.77E-08	rs8082590	*SREBF1* (388.5), *FLII* (87.9), *MYO15A* (84.0), *SMCR7* (33.8), *TOP3A* (19.5), *SHMT1* (19.2), *DRG2* (12.9), *LGALS9C* (8.5), *TOM1L2* (5.3), *LRRC48* (4.8), *ATPAF2* (2.4), *GID4*, *EVPLL*, *SMCR8*
15:70569536-70631100	1.79E-08	rs12148337	*TLE3** (20.4)
16:58669273-58682833	1.87E-08	rs12325245	*CNOT1* (69.7), *SLC38A7* (25.8)
8:27411100-27453579	2.10E-08	rs73229090	*CLU* (348.9), *EPHX2* (192.7)
X:5859733-6042430	2.21E-08	rs12845396	*NLGN4X* (554.6)
3:181061017-181205593	2.35E-08	rs9841616	*SOX2** (198.1)
1:97792597-97834525	2.64E-08	rs76869799	*DPYD* (72.2)
6:73132745-73171881	2.69E-08	rs1339227	*RIMS1* (357.2)
7:24619811-24844736	2.85E-08	chr7_24747494_D	*MPP6* (35.9), *DFNA5* (21.2), *OSBPL3* (10.6)
5:109030041-109209342	3.05E-08	rs4388249	*PJA2* (145.9), *MAN2A1* (49.7)
4:23348610-23443426	3.06E-08	rs215411	*PPARGC1A** (250.8)
5:153671061-153688682	3.15E-08	rs11740474	*GALNT10* (22.4), *LARP1* (14.9)
11:109317548-109573244	3.70E-08	rs12421382	*C11orf87* (1.0)
7:110029087-110106697	3.71E-08	rs211829	*IMMP2L** (153.6)
12:29864416-29940392	3.91E-08	rs16934812, rs679087	*TMTC1* (8.5)
7:131533816-131590785	4.42E-08	rs7801375	*PLXNA4** (78.2)
1:177237533-177323124	4.45E-08	rs6670165	*BRINP2* (11.0)
1:207893266-208024062	4.47E-08	rs7523273	*CD34* (346.3), *CD46* (80.6), *PLXNA2* (65.8), *CR1L* (15.6)
20:48097481-48131649	4.56E-08	rs7267348	*PTGIS* (155.5), *KCNB1* (72.1)
12:92243186-92258265	4.59E-08	rs4240748	*BTG1** (100.8)
2:162796517-162910223	4.62E-08	rs2909457	*DPP4* (933.8), *GCA* (88.6), *IFIH1* (40.3), *SLC4A10* (14.6)
19:50067508-50138023	4.69E-08	rs56873913	*RRAS* (127.3), *PRMT1* (116.7), *PRRG2* (70.1), *NOSIP* (63.9), *FUZ* (21.0), *SIGLEC11* (20.7), *RCN3* (19.5), *SCAF1* (1.0), *PRR12* (0.7)
12:103559817-103622466	4.84E-08	rs10860964	*C12orf42*
5:140024042-140222641	4.85E-08	chr5_140143664_I	*CD14* (1051.8), *HBEGF* (189.8), *IK* (134.5), *NDUFA2* (34.4), *PCDHB16* (29.4), *PCDHA2* (25.1), *PCDHA5* (25.1), *PCDHA1* (25.1), *PCDHA6* (25.1), *PCDHA4* (25.1), *PCDHA7* (17.9), *PCDHA10* (17.9), *PCDHA3* (17.9), *HARS2* (13.5), *PCDHA8* (11.9), *TMCO6* (8.9), *DND1* (8.1), *PCDHA9* (7.9), *HARS* (7.1), *ZMAT2* (3.4), *WDR55* (2.1)
6:30137209-30236038	N.S.#	rs2523722, rs2021722	*TRIM15* (6.8), *PPP1R18* (4.8), *TRIM10* (4.3), *TRIM31* (3.4), *TRIM26* (2.6), *C6orf136* (1.1)
6:31580539-31732547	N.S.#	rs1046089	*CYP21A2* (299.8), *AIF1* (185.2), *APOM* (147.1), *AGER* (105.8), *CLIC1* (80.6), *HLA-B* (74.9), *BAG6* (72.1), *CSNK2B* (69.7), *DDAH2* (60.5),

			MSH5 (50.2), *LY6G5B* (39.3), *ATF6B* (37.1), *NCR3* (27.6), *C6orf25* (27.6), *LY6G6F* (19.0), *VARS* (13.1), *ABHD16A* (11.2), *LY6G6D* (9.8), *PRRC2A* (8.8), *VWA7* (3.6), *GPANK1* (3.5), *LY6G6C* (1.7), *LY6G5C* (1.3), *C6orf47*, *SAPCD1*
X:153176959-153376436	N.S.	rs2269372	*OPN1LW* (359.6), *MECP2* (310.3), *IRAK1* (301.1), *FLNA* (246.1), *AVPR2* (229.9), *HCFC1* (122.0), *L1CAM* (103.6), *ARHGAP4* (89.0), *GDI1* (68.6), *ATP6AP1* (67.7), *RPL10* (61.0), *SLC10A3* (24.9), *RENBP* (23.4), *TAZ* (19.3), *TMEM187* (4.7), *NAA10* (3.5), *FAM50A* (3.5)
X:147287097-147480725	N.S.	rs2159767	*FMR1* (85.7), *AFF2** (66.6), *FMR1NB* (1.5)
9:78018249-78058588	N.S.	rs489332	*OSTF1** (59.8)
9:26756661-27066934	N.S.	rs7045881	*PLAA* (57.2), *IFT74* (25.2), *LRRC19* (4.8), *EQTN*, *CAAP1*
9:121326409-121364867	N.S.	rs1572299	
8:38014429-38231314	N.S.	rs16887244	*FGFR1* (926.8), *STAR* (822.7), *BAG4* (78.7), *ASH2L* (37.3), *WHSC1L1* (7.1), *DDHD2* (5.2), *LSM1* (4.8), *PPAPDC1B* (3.8), *LETM2* (1.6)
8:10007345-10028396	N.S.	rs7017212	*MSRA* (17.2)
7:71681396-71849677	N.S.	rs12699131	*CALN1* (13.5)
7:103401621-103465023	N.S.	rs7341475	*RELN* (562.0)
6:475489-475489	N.S.	rs12210050	*EXOC2* (23.1)
6:33844014-33862507	N.S.	rs16869652	*MLN** (50.4)
6:32014828-32609061	N.S.#	rs3132935, rs4530903, rs9272219, rs114002140, rs3131296	*CYP21A2* (299.8), *TAP1* (173.0), *NOTCH4* (151.4), *TNXB* (137.0), *HLA-DQB1* (114.6), *HLA-DRB1 (107.4)*, *AGER* (105.8), *HLA-DRA* (86.5), *C4A* (70.2), *HLA-C* (59.0), *HLA-DQA1* (58.3), *FKBPL* (57.0), *PBX2* (56.0), *BTNL2* (54.4), *HLA-DPA1* (52.8), *C4B* (50.6), *HLA-DQA2* (45.3), *HLA-DPB1* (44.4), *ATF6B* (37.1), *HLA-DQB2* (24.5), *RNF5* (22.4), *AGPAT1* (18.6), *EGFL8* (17.2), *PPT2* (11.3), *HLA-DRB5* (11.0), *GPSM3* (5.5), *HLA-DQB3*, *HLA-DRB6*, *C6orf10*, *PRRT1*
6:26978700-27905509	N.S.#	rs6932590, rs13194053, rs17693963, rs16897515	*HIST1H4L* (18.1), *HIST1H4K* (18.1), *HIST1H2BK* (13.7), *HIST1H2BJ* (13.7), *BTN3A2* (12.4), *OR2B2* (10.5), *OR2B6* (10.5), *HIST1H1B* (10.2), *HIST1H2BN* (9.1), *HIST1H2AL* (8.0), *PRSS16* (7.4), *HIST1H3I* (6.9), *HIST1H2AK* (6.8), *HIST1H2AG* (6.5), *HIST1H2BM* (6.5), *HIST1H2BL* (6.5), *HIST1H2BO* (6.5), *HIST1H2AJ* (3.7), *ZSCAN9* (2.9), *ZNF391* (2.9), *ZNF184* (2.9), *HIST1H2AI*, *HIST1H3J*, *HIST1H2AM*, *HIST1H3H*, *HIST1H4I*, *POM121L2*, *HIST1H4J*
5:64432141-64519335	N.S.	rs17206232	*ADAMTS6* (23.6)
5:56641049-56677611	N.S.	rs10052004	*GPBP1** (10.5)
5:101581848-101871853	N.S.	rs1502844,	*PAM* (188.1), *SLCO4C1* (22.1), *SLCO6A1* (4.4)

		rs6878284	
4:2359807-2404699	N.S.	rs959770	*ADD1* (129.7), *ZFYVE28* (16.4)
4:183135115-183151151	N.S.	rs2726807	*TENM3* (98.4)
4:118630975-118780055	N.S.	rs11098403	*NDST3** (9.8)
3:62058717-62081262	N.S.	rs11130874	*PTPRG* (68.0)
3:151676355-151781175	N.S.	rs1351267	*SUCNR1** (6.5)
3:119044502-119094288	N.S.	rs17203055	*ARHGAP31* (22.0), *C3orf30*
22:50162136-50321623	N.S.	rs138880	*MAPK11* (97.4), *MLC1* (80.7), *MAPK12* (73.7), *PLXNB2* (70.2), *CRELD2* (26.4), *TUBGCP6* (15.5), *BRD1* (6.3), *MOV10L1* (6.3), *ZBED4* (6.1), *ALG12* (3.1), *FAM116B* (1.1)
20:55673623-55686994	N.S.	rs11699237	*BMP7** (210.8)
2:47994837-48178441	N.S.	rs4381823	*MSH6* (152.5), *FBXO11* (22.1), *FOXN2* (7.4)
2:37586681-37592628	N.S.	rs2373000	*QPCT* (12.5)
2:236769459-236827263	N.S.	rs13025591	*AGAP1* (73.1)
2:145139727-145186749	N.S.	rs12991836	*ZEB2* (140.3)
2:124934949-125035016	N.S.	rs1170612	*CNTNAP5* (16.3)
19:42057604-42094734	N.S.	rs4803480	*BCKDHA* (35.2), *CEACAM3* (30.6), *CEACAM21* (4.4), *TMEM91*
19:40173197-40238407	N.S.	rs12611334	*CLC* (19.7), *CNTD2* (9.3), *LGALS14* (8.7)
18:77364313-77392379	N.S.	rs7233060	*CTDP1** (19.1), *PQLC1* (1.0)
18:75887984-75919598	N.S.	rs4798896	
18:41981849-42117685	N.S.	rs2048485	*SETBP1** (47.7)
18:11450290-11531630	N.S.	rs1455244	*GNAL** (612.4), *CHMP1B* (2.8)
16:82683758-82696317	N.S.	rs8057927	*CDH13* (231.7)
16:57021433-57040093	N.S.	rs17290922	*CCL22* (159.4), *CETP* (99.7), *NLRC5* (29.6), *MT1H* (20.5)
16:20565815-20880040	N.S.	rs433598, rs151222	*DCUN1D3* (19.4), *ACSM1* (17.0), *ACSM3* (14.4), *ACSM2B* (3.9), *ERI2* (2.3), *THUMPD1* (0.7)
16:13021759-13096582	N.S.	rs7192086	*SHISA9*
16:12070183-12086263	N.S.	rs12922317	*TNFRSF17* (48.1), *SNX29* (25.2)
15:86834376-87114576	N.S.	rs16977195	*AGBL1* (10.7)
15:61332682-61336442	N.S.	rs7172342	*RORA* (59.2)
13:61952782-61971074	N.S.	rs2323266	*PCDH20* (7.9)
13:27987643-28111185	N.S.	rs9512730	*GTF3A* (51.7), *LNX2* (18.1), *MTIF3* (4.7)
12:81706189-81862315	N.S.	rs12426725	*PPFIA2* (24.8)
12:119771465-119853676	N.S.	rs11064768	*CCDC60*
12:114697688-114705586	N.S.	rs1920592	*TBX5** (69.4)
11:44834764-44873282	N.S.	rs11038167	*TSPAN18* (4.4), *PRDM11* (1.1)
11:29115058-29242684	N.S.	rs1602565	
11:17056796-17305650	N.S.	rs4356203	*PIK3C2A* (213.8), *NUCB2* (111.3), *SOX6* (61.1), *USH1C* (52.2), *RPS13* (36.3), *PLEKHA7* (10.9)
11:13288698-13350131	N.S.	rs4757144	*ARNTL* (359.6)
11:125294082-125627917	N.S.	rs7930295, rs11220082, rs548181	*EI24* (220.3), *CHEK1* (136.9), *FEZ1* (42.1), *STT3A* (6.4), *ACRV1* (6.2), *PATE1* (5.1), *PKNOX2* (4.9), *PATE2*
10:62040118-62349324	N.S.	rs16915157,	*ANK3* (135.1), *RHOBTB1* (26.7)

		rs10761482	
10:44821814-44842980	N.S.	rs10900020	*CXCL12* (1114.3)
10:34068356-34100580	N.S.	rs1412115	*PARD3** (100.2)
10:21565732-21567565	N.S.	rs3847375	*NEBL** (28.7)
1:97078413-97284496	N.S.	rs7544736, rs12071951	*PTBP2* (11.4)
1:37093026-37194103	N.S.	rs589249	*GRIK3** (200.3)
1:245726054-245749914	N.S.	rs10924245	*KIF26B* (25.0), *SMYD3* (4.0)
1:244387175-244390564	N.S.	rs10429924	*C1orf100**
1:210519851-210577528	N.S.	rs7527939	*HHAT* (10.8)
1:186434518-186711910	N.S.	rs10911902	*PTGS2* (1364.2), *TPR* (185.0), *PDC* (111.9)
1:177653296-177747310	N.S.	rs12140439	*SEC16B** (28.3)
1:167844896-168096620	N.S.	rs10489202	*CD247* (64.5), *GPR161* (58.1), *MPZL1* (47.5), *ADCY10* (35.0), *RCSD1* (6.3), *DCAF6* (6.1), *MPC2* (5.0)
1:11788564-11788564	N.S.#	rs4846033	*AGTRAP* (27.6), *DRAXIN* (9.9)

Notes:

1. 'N.S.' denotes the *P*-value that is not genome-wide significant according to the analysis of the PGC study (SCHIZOPHRENIA WORKING GROUP OF THE PSYCHIATRIC GENOMICS 2014). Five genomic regions with either no or contradictory imputation signals in the PGC study are excluded from our analysis for risk genes with different association strengths (indicated by #).
2. Schizophrenia gene scores are in parentheses. Genes with mouse knock-outs phenotypes of nervous system and neurological behaviors from the MGI database (http://www.informatics.jax.org/) are underlined. Genes in black colors are proximal candidate genes covered by LD blocks or the closest genes within 500 kb (indicated by *) if no genes are covered by LD blocks; genes highlighted in blue are distal candidate genes linked through information of TREs and their interacted genes (ENCODE (THURMAN *et al.* 2012) and FANTOM5 (ANDERSSON *et al.* 2014)) or eQTL (GTEx (CONSORTIUM 2013)).
3. The result can be viewed in the UCSC Genome Browser (http://zdzlab.einstein.yu.edu/1/sz.html).

Table S4. High scoring genes with support of regulatory information.

Gene symbol	Score	Risk region (chr:start-end)	Linked TREs or eQTL (chr:start-end)	Implicated risk variant[1]	Reference of interaction[2]	Distal gene?
DRD2	1420.9	chr11:113317745-113424042	chr11:113344860-113345010	rs61902807	ENCODE	No
FGFR1	926.8	chr8:38014429-38231314	chr8:38081200-38081350	rs6999796	ENCODE	Yes
	926.8	chr8:38014429-38231314	chr8:38226140-38226290	rs16887343	ENCODE	Yes
MAPK3	915.7	chr16:29923510-30018500	chr16:29986080-29986230	rs10871451	ENCODE	Yes
STAR	822.7	chr8:38014429-38231314	chr8:38014320-38014470	rs75168396	ENCODE	No
CACNA1C	597.1	chr12:2292690-2523772	chr12:2368640-2368790	rs10774035	ENCODE	No
RAD51	481.9	chr15:40566759-40602256	chr15:40569820-40569970	rs28676999	ENCODE	Yes
PRKCD	439.5	chr3:52638482-52960859	chr3:52721240-52721390	rs11177	ENCODE	Yes
	439.5	chr3:52638482-52960859	chr3:52744220-52744370	chr3:52744224-52744224	ENCODE	Yes
	439.5	chr3:52638482-52960859	chr3:52746520-52746670	rs4687644	ENCODE	Yes
	439.5	chr3:52638482-52960859	chr3:52755560-52755710	rs11130319	ENCODE	Yes
	439.5	chr3:52638482-52960859	chr3:52766600-52766750	rs2268027	ENCODE	Yes
	439.5	chr3:52638482-52960859	chr3:52826500-52826650	rs746694	ENCODE	Yes
	439.5	chr3:52638482-52960859	chr3:52864020-52864170	rs4687554	ENCODE	Yes
	439.5	chr3:52638482-52960859	chr3:52870600-52870750	rs4687663	ENCODE	Yes
SREBF1	388.5	chr17:17760789-18036283	chr17:17820860-17821010	chr17:17820998-17820998	ENCODE	Yes
	388.5	chr17:17760789-18036283	chr17:17847860-17848010	rs6502624	ENCODE	Yes
OPN1LW	359.6	chrX:153176959-153376436	chrX:153241340-153241490	rs4898374	ENCODE	Yes
CD34	346.3	chr1:207893266-208024062	chr1:207919840-207919990	rs2796264	ENCODE	Yes
	346.3	chr1:207893266-208024062	chr1:207992400-207992550	rs12123251	ENCODE	Yes
	346.3	chr1:207893266-208024062	chr1:208017040-208017190	rs2724394	ENCODE	Yes
MECP2	310.3	chrX:153176959-153376436	chrX:153265580-153265730	rs11795678	ENCODE	No
CYP21A2	299.8	chr6:31580539-31732547	chr6:31688080-31688230	rs116316082	ENCODE	No
FES	286.1	chr15:91412848-91429042	chr15:91423500-91423650	rs6224	ENCODE	No
CHRNB4	269.3	chr15:78785544-	chr15:78907320-	rs7177514	ENCODE	No

		78930510	78907470			
PPARGC1A	250.8	chr4:23348610-23443426	chr4:23355480-23355630	rs6834404	ENCODE	No
FLNA	246.1	chrX:153176959-153376436	chrX:153211620-153211770	rs762656	ENCODE	Yes
GRIK3	200.3	chr1:37093026-37194103	chr1:37162320-37162470	rs589249	ENCODE	No
	200.3	chr1:37093026-37194103	chr1:37164780-37164930	rs631416	ENCODE	No
EPHX2	192.7	chr8:27411100-27453579	chr8:27442080-27442230	rs73229090	ENCODE	No
	192.7	chr8:27411100-27453579	chr8:27453520-27453670	rs35236974	ENCODE	No
HBEGF	189.8	chr5:140024042-140222641	chr5:140024400-140024550	rs702399	ENCODE	Yes
	189.8	chr5:140024042-140222641	chr5:140024400-140024550	rs60115373	ENCODE	Yes
PAM	188.1	chr5:101581848-101871853	chr5:101630860-101631010	rs841921	ENCODE	Yes
TPR	185.0	chr1:186434518-186711910	chr1:186451940-186452090	rs3131554	ENCODE	Yes
MCL1	182.2	chr1:149998923-150226321	chr1:150152820-150152970	rs1105209	ENCODE	Yes
	182.2	chr1:149998923-150226321	chr1:150194840-150194990	rs11576997	ENCODE	Yes
TAP1	173.0	chr6:32014828-32609061	chr6:32430480-32430630	rs114664081	ENCODE	Yes
	173.0	chr6:32014828-32609061	chr6:32430480-32430630	rs116580588	ENCODE	Yes
NMUR2	167.7	chr5:151941138-152847217	chr5:152022020-152022170	rs17454953	ENCODE	Yes
TIE1	167.5	chr1:44029353-44248230	chr1:44083800-44083950	rs639929	ENCODE	Yes
CCL22	159.4	chr16:57021433-57040093	chr16:57040020-57040170	rs34218679	ENCODE	Yes
NOTCH4	151.4	chr6:32014828-32609061	chr6:32154880-32155030	rs115219661	ENCODE	No
PJA2	145.9	chr5:109030041-109209342	chr5:109047260-109047410	rs13153918	ENCODE	Yes
	145.9	chr5:109030041-109209342	chr5:109185320-109185470	rs12656712	ENCODE	Yes
NISCH	138.6	chr3:52638482-52960859	chr3:52833100-52833250	rs2535629	ENCODE	Yes
ANK3	135.1	chr10:62040118-62349324	chr10:62094060-62094210	rs991405	ENCODE	No
	135.1	chr10:62040118-62349324	chr10:62103140-62103290	rs2061488	ENCODE	No
	135.1	chr10:62040118-62349324	chr10:62103140-62103290	rs2061489	ENCODE	No
	135.1	chr10:62040118-62349324	chr10:62103140-62103290	rs4456215	ENCODE	No
	135.1	chr10:62040118-62349324	chr10:62320320-62320470	rs2068043	ENCODE	No
ADD1	129.7	chr4:2359807-	chr4:2392660-	rs9994065	ENCODE	Yes

		2404699	2392810			
NEURL	127.1	chr10:104423800-105059896	chr10:104786200-104786350	rs4917992	ENCODE	Yes
	127.1	chr10:104423800-105059896	chr10:104837720-104837870	rs1046411	ENCODE	Yes
	127.1	chr10:104423800-105059896	chr10:104856080-104856230	rs12412038	ENCODE	Yes
	127.1	chr10:104423800-105059896	chr10:104877200-104877350	rs79082900	ENCODE	Yes
	127.1	chr10:104423800-105059896	chr10:104945200-104945350	rs4917997	ENCODE	Yes
	127.1	chr10:104423800-105059896	chr10:104947400-104947550	rs10883843	ENCODE	Yes
	127.1	chr10:104423800-105059896	chr10:104949660-104949810	rs11191600	ENCODE	Yes
	127.1	chr10:104423800-105059896	chr10:104949660-104949810	rs11191601	ENCODE	Yes
	127.1	chr10:104423800-105059896	chr10:104951440-104951590	chr10:104951466-104951466	ENCODE	Yes
HTR3B	125.1	chr11:113317745-113424042	chr11:113317700-113317850	rs17601612	ENCODE	Yes
	125.1	chr11:113317745-113424042	chr11:113318820-113318970	rs4936272	ENCODE	Yes
HCFC1	122.0	chrX:153176959-153376436	chrX:153214460-153214610	rs201119463	ENCODE	No
PRMT1	116.7	chr19:50067508-50138023	chr19:50091400-50091550	rs10406941	ENCODE	Yes
HLA-DQB1	114.6	chr6:32014828-32609061	chr6:32165300-32165450	rs115344853	ENCODE	No
	114.6	chr6:32014828-32609061	chr6:32603380-32603530	rs111529210	ENCODE	No
	114.6	chr6:32014828-32609061	chr6:32603380-32603530	rs113008958	ENCODE	No
SV2B	107.8	chr15:91412848-91429042	chr15:91419340-91419490	chr15:91419432-91419432	ENCODE	Yes
CTNNA1	106.0	chr5:137598340-137948140	chr5:137652180-137652330	rs11740078	ENCODE	No
AGER	105.8	chr6:31580539-31732547	chr6:31692160-31692310	rs116193838	ENCODE	No
MAPK11	97.4	chr22:50162136-50321623	chr22:50244560-50244710	rs3788730	ENCODE	Yes
	97.4	chr22:50162136-50321623	chr22:50319900-50320050	rs10854860	ENCODE	Yes
RANGAP1	91.0	chr22:41429084-41637119	chr22:41434800-41434950	rs9611474	ENCODE	No
DOCK4	90.2	chr7:110843795-111092478	chr7:110917580-110917730	rs13239254	ENCODE	Yes
GNL3	89.9	chr3:52638482-52960859	chr3:52642440-52642590	rs3774365	ENCODE	No
GCA	88.6	chr2:162796517-162910223	chr2:162866140-162866290	rs2909448	ENCODE	Yes
FLII	87.9	chr17:17760789-18036283	chr17:17823100-17823250	rs9890341	ENCODE	Yes
	87.9	chr17:17760789-	chr17:17853780-	rs8078105	ENCODE	Yes

		18036283	17853930			
FMR1	85.7	chrX:147287097-147480725	chrX:147351240-147351390	rs59460742	ENCODE	Yes
MLC1	80.7	chr22:50162136-50321623	chr22:50309360-50309510	rs4074304	ENCODE	Yes
CD46	80.6	chr1:207893266-208024062	chr1:207981320-207981470	rs761276	ENCODE	No
GRIN2A	1385.0	chr16:9875513-9971728	chr16:9912515-9912659	rs11645219	FANTOM5	No
SERPING1	137.5	chr11:57369008-57681828	chr11:57371918-57372022	rs28362950	FANTOM5	No
HLA-DQB1	114.6	chr6:32014828-32609061	chr6:32576157-32576293	rs34811813	FANTOM5	No
	114.6	chr6:32014828-32609061	chr6:32590770-32591146	rs115195925	FANTOM5	No
	114.6	chr6:32014828-32609061	chr6:32590770-32591146	rs115553940	FANTOM5	No
CD46	80.6	chr1:207893266-208024062	chr1:207999163-207999611	rs12132780	FANTOM5	No
SREBF1	388.5	chr17:17760789-18036283	chr17:17958402-17958402	rs8082590	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17760789-17760789	rs8079321	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17762457-17762457	rs9911281	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17764061-17764061	rs35451946	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17764502-17764502	rs4924823	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17765655-17765655	rs12941039	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17767165-17767165	rs1889014	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17767767-17767767	rs9895335	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17767819-17767819	rs9895750	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17770355-17770355	rs11657074	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17770965-17770965	rs9907246	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17774118-17774118	rs9907287	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17774422-17774422	rs9908017	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17774568-17774568	rs9908299	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17776389-17776389	rs8080061	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17777245-17777245	rs12936037	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17843378-17843378	rs11657845	GTEx	Yes
	388.5	chr17:17760789-	chr17:17891781-	rs7207821	GTEx	Yes

		18036283	17891781			
	388.5	chr17:17760789-18036283	chr17:17894750-17894750	rs62072048	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17896090-17896090	rs4368210	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17896205-17896205	rs4584886	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17896673-17896673	rs9912096	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17899839-17899839	rs4506969	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17902135-17902135	rs28537385	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17903505-17903505	chr17:17903505-17903505	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17906520-17906520	rs62072049	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17906564-17906564	rs62072050	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17913057-17913057	rs6502632	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17913504-17913504	rs7212167	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17915791-17915791	rs7223696	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17919749-17919749	rs9913277	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17924060-17924060	rs8079418	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17924868-17924868	rs9896837	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17926605-17926605	rs7215524	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17930253-17930253	rs2955378	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17932818-17932818	rs2955377	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17932931-17932931	rs7224047	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17939247-17939247	rs2955385	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17939573-17939573	rs2955384	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17940305-17940305	rs7503738	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17941037-17941037	rs6502633	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17942613-17942613	rs7406982	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17944349-17944349	rs2955380	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17945500-17945500	rs2955381	GTEx	Yes
	388.5	chr17:17760789-	chr17:17946401-	rs2955357	GTEx	Yes

		18036283	17946401			
	388.5	chr17:17760789-18036283	chr17:17946730-17946730	rs2955359	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17947710-17947710	rs2955382	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17948716-17948716	rs2955354	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17948979-17948979	rs2955353	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17949789-17949789	rs7207461	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17949802-17949802	rs12948749	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17950001-17950001	rs12940282	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17952439-17952439	rs11652894	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17952868-17952868	rs4925135	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17953548-17953548	rs2955350	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17954728-17954728	rs8080602	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17954764-17954764	rs8080334	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17955344-17955344	rs6502634	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17960613-17960613	rs11650021	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17961349-17961349	rs2955368	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17961407-17961407	rs2955369	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17964717-17964717	rs2955356	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17966524-17966524	chr17:17966524-17966524	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17966945-17966945	rs2955370	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17967397-17967397	rs2955371	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17970229-17970229	rs2955372	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17972973-17972973	rs4643387	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17980671-17980671	rs12943914	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17983817-17983817	rs6502636	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17987285-17987285	rs2955351	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17990474-17990474	rs4506967	GTEx	Yes
	388.5	chr17:17760789-	chr17:17992793-	rs4925138	GTEx	Yes

		18036283	17992793			
	388.5	chr17:17760789-18036283	chr17:17994332-17994332	rs12943202	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17995166-17995166	rs12950562	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17995619-17995619	rs2974998	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17997209-17997209	rs2230316	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:17997547-17997547	rs2974999	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:18003648-18003648	rs854814	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:18003845-18003845	rs854813	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:18005073-18005073	rs721669	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:18006421-18006421	rs854810	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:18006539-18006539	rs854809	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:18006634-18006634	rs854808	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:18008447-18008447	rs712265	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:18009028-18009028	rs2056842	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:18009102-18009102	rs854762	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:18010095-18010095	rs854763	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:18011140-18011140	rs6826	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:18011750-18011750	rs854764	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:18012730-18012730	rs854765	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:18012775-18012775	rs854766	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:18016148-18016148	rs1101727	GTEx	Yes
	388.5	chr17:17760789-18036283	chr17:18022039-18022039	rs854818	GTEx	Yes
ARNTL	359.6	chr11:13288698-13350131	chr11:13288851-13288851	rs34560638	GTEx	No
	359.6	chr11:13288698-13350131	chr11:13288885-13288885	rs61882109	GTEx	No
	359.6	chr11:13288698-13350131	chr11:13291931-13291931	rs4146385	GTEx	No
	359.6	chr11:13288698-13350131	chr11:13292864-13292864	rs2219998	GTEx	No
	359.6	chr11:13288698-13350131	chr11:13294268-13294268	rs900144	GTEx	No
	359.6	chr11:13288698-	chr11:13301875-	rs72867447	GTEx	No

		13350131	13301875			
CYP2D6	275.3	chr22:42315790-42689370	chr22:42315790-42315790	rs763263	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42343091-42343091	rs6002555	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42361336-42361336	rs5996096	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42364057-42364057	rs4822076	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42603814-42603814	rs6002655	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42534148-42534148	rs2743449	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42534682-42534682	rs2743451	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42534864-42534864	rs35711087	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42537597-42537597	rs1800754	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42538103-42538103	rs3021083	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42540551-42540551	rs2743461	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42541349-42541349	rs5758605	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42542870-42542870	rs2142694	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42543288-42543288	rs2743462	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42547739-42547739	rs2743465	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42557710-42557710	rs5758619	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42566314-42566314	rs5758623	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42571028-42571028	rs760648	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42574830-42574830	rs2413684	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42579309-42579309	rs67588321	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42579520-42579520	rs2899355	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42592239-42592239	rs5751239	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42598951-42598951	rs5758645	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42600589-42600589	rs5751241	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42613485-42613485	rs5758653	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42618340-42618340	rs2143138	GTEx	No
	275.3	chr22:42315790-	chr22:42622003-	rs5758659	GTEx	No

		42689370	42622003			
	275.3	chr22:42315790-42689370	chr22:42623718-42623718	rs5758660	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42624445-42624445	rs5758661	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42636687-42636687	rs5758670	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42639645-42639645	rs5751250	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42640606-42640606	rs5751251	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42643039-42643039	rs5758677	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42648408-42648408	rs5751255	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42650323-42650323	rs68037805	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42650663-42650663	rs134866	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42652074-42652074	rs134869	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42652317-42652317	rs134870	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42652716-42652716	rs134871	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42657566-42657566	rs134873	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42662371-42662371	rs134877	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42664201-42664201	rs134879	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42670965-42670965	rs134882	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42675960-42675960	rs66607825	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42680800-42680800	rs86669	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42683343-42683343	rs134900	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42683997-42683997	rs134902	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42689140-42689140	rs80506	GTEx	No
	275.3	chr22:42315790-42689370	chr22:42689370-42689370	rs134906	GTEx	No
HLA-DQB1	114.6	chr6:32014828-32609061	chr6:32605982-32605982	chr6:32605982-32605982	GTEx	No
	114.6	chr6:32014828-32609061	chr6:32586934-32586934	chr6:32586934-32586934	GTEx	No
HLA-DRB1	107.4	chr6:32014828-32609061	chr6:32605982-32605982	chr6:32605982-32605982	GTEx	No
	107.4	chr6:32014828-32609061	chr6:32586934-32586934	chr6:32586934-32586934	GTEx	No

Notes:

1. Implicated risk variants refer to variants residing in the linked TREs or eQTL that are in strong LD ($r^2 > 0.5$) with schizophrenia GWAS signals.
2. Reference of transcriptional regulatory interactions between TREs or eQTL and their target genes are collected from ENCODE (THURMAN et al. 2012), FANTOM5 (ANDERSSON et al. 2014), and GTEx (blood tissue) (CONSORTIUM 2013).

Table S5. Association among schizophrenia gene sets.

	High score	Mouse KO phenotypes	Literature support	Differential expression
High score		9.58E-19	2.92E-05	0.135
Mouse KO phenotypes			1.21E-05	0.067
Literature support				0.678
Differential expression				

Notes:

1. Gene set association was tested by Fisher's exact test.
2. Mouse knock-out phenotypes are for nervous systems and neurological behaviors.
3. The differential expression gene set are based on previous studies (GLATT *et al.* 2005; MAYCOX *et al.* 2009; PEREZ-SANTIAGO *et al.* 2012; FILLMAN *et al.* 2013; GARDINER *et al.* 2013; HWANG *et al.* 2013; SAINZ *et al.* 2013; SANDERS *et al.* 2013; SELLMANN *et al.* 2014).

Table S6. KEGG pathways overrepresented in 132 schizophrenia risk genes.

KEGG pathways[1,2]	Adjusted *P*-value
Neuroactive ligand-receptor interaction	7.93E-08
MAPK signaling pathway	5.03E-06
Tuberculosis	7.84E-06
Long-term potentiation	8.28E-06
Adherens junction*	8.58E-06
Leishmaniasis*	9.22E-06
Glutamatergic synapse	1.00E-05
GnRH signaling pathway*	3.62E-05
Alzheimer's disease	5.32E-04
Pathways in cancer	5.49E-04
Chagas disease (American trypanosomiasis)	5.56E-04
Chemokine signaling pathway*	9.22E-04
Toxoplasmosis*	1.14E-03
Cell adhesion molecules (CAMs)*	1.15E-03
Axon guidance	1.20E-03
Regulation of actin cytoskeleton*	1.31E-03
Arrhythmogenic right ventricular cardiomyopathy (ARVC)*	1.38E-03
Type I diabetes mellitus	2.14E-03
Intestinal immune network for IgA production*	3.94E-03
Calcium signaling pathway	3.99E-03
Toll-like receptor signaling pathway	4.05E-03
Type II diabetes mellitus*	4.29E-03
Vascular smooth muscle contraction*	4.93E-03
Leukocyte transendothelial migration*	4.93E-03
Focal adhesion	4.96E-03
Amyotrophic lateral sclerosis (ALS)	5.24E-03
Arachidonic acid metabolism*	5.29E-03
Neurotrophin signaling pathway	6.20E-03
Tight junction*	7.13E-03

Notes:

1. * denotes pathways not enriched (Adjusted $P < 0.05$) in training genes.
2. Only pathways with an adjusted *P*-value < 0.01 are shown.

Table S7. Panther pathways overrepresented in 132 schizophrenia risk genes.

Panther pathways[1,2]	**Adjusted *P*-value**
EGF receptor signaling pathway	3.15E-07
Endothelin signaling pathway*	3.96E-07
Alzheimer disease-amyloid secretase pathway*	4.26E-07
Nicotinic acetylcholine receptor signaling pathway*	9.47E-07
Toll receptor signaling pathway*	2.66E-04
TGF-beta signaling pathway*	2.87E-04
Ionotropic glutamate receptor pathway	3.11E-04
FGF signaling pathway*	4.10E-04
VEGF signaling pathway*	4.24E-04
Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway	4.36E-04
Oxytocin receptor mediated signaling pathway*	4.73E-04
5HT2 type receptor mediated signaling pathway	5.72E-04
Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway	1.41E-03
Angiogenesis*	1.43E-03
Beta2 adrenergic receptor signaling pathway*	2.20E-03
Beta1 adrenergic receptor signaling pathway*	2.20E-03
Inflammation mediated by chemokine and cytokine signaling pathway*	3.78E-03
Alzheimer disease-presenilin pathway*	3.92E-03
Thyrotropin-releasing hormone receptor signaling pathway*	4.53E-03
B cell activation*	4.76E-03
Metabotropic glutamate receptor group III pathway	5.72E-03
Integrin signalling pathway*	8.68E-03
Alpha adrenergic receptor signaling pathway*	9.00E-03

Notes:

1. * denotes pathways not enriched (Adjusted $P < 0.05$) in training genes.
2. Only pathways with an adjusted *P*-value < 0.01 are shown.

Table S8. Classes of loci with different association strengths.

	Total	**Weak**	**Moderate**	**Strong**
GWAS SNPs	248	68	93	87
Risk genomic regions	171	62	70	39
Risk gene candidates	554	146	220	188
Risk genes (High scoring genes)	120	36	49	35

Note: Two candidate genes (DPYD and IMMP2L) link to loci with different class of association strength were excluded from the table due to their ambiguous association strength.

Table S9. Correlation of spatiotemporal expression between training genes and prioritized genes.

Weak		Moderate		Strong	
Training gene	Correlation coefficient	Training gene	Correlation coefficient	Training gene	Correlation coefficient
TP53	0.606018	*CYFIP1*	0.560728	*HTR2A*	0.783905
MUTED	0.601462	*NPAS3*	0.54149	*GABRB2*	0.745154
EHMT1	0.497874	*IL1B*	0.540428	*DRD1*	0.668092
AKT1	0.480992	*COMT*	0.522435	*RTN4R*	0.6541
MTHFR	0.473531	*DRD4*	0.504125	*SHANK3*	0.652142
CYFIP1	0.466414	*TP53*	0.441944	*DLG1*	0.651256
SLC6A4	0.426591	*DISC1*	0.261806	*RGS4*	0.650333
RELN	0.405606	*APOE*	0.226015	*NRGN*	0.620919
RPGRIP1L	0.39084	*MUTED*	0.214408	*APOL2*	0.601066
EGF	0.307793	*EGF*	0.188993	*GAD1*	0.512738
NRG1	0.306834	*GRIK4*	0.153226	*CHRNA7*	0.506124
PLXNA2	0.29508	*AKT1*	0.147056	*DTNBP1*	0.480157
SLC18A1	0.202755	*NRG1*	0.111276	*CHI3L1*	0.466395
NPAS3	0.128998	*CRP*	0.104327	*DLG2*	0.43909
ZNF804A	0.114728	*PRODH*	0.091627	*TNF*	0.390873
APOL4	0.09571	*SLC18A1*	0.083151	*GRIK4*	0.358498
DRD4	0.066545	*IL18*	0.082069	*IL18*	0.332816
PPP3CC	0.036021	*VIPR2*	0.063812	*EGF*	0.314053
GRIN2B	0.023557	*APOL4*	0.063256	*VIPR2*	0.310065
DRD3	0.019007	*RTN4R*	0.058439	*RPGRIP1L*	0.309849
IL1B	0.003939	*SHANK3*	0.047446	*GRIN2B*	0.305772
TPH1	-0.006	*SLC6A4*	0.039909	*OPCML*	0.281525
OPCML	-0.0086	*EHMT1*	0.031161	*ERBB4*	0.160328
DRD2	-0.0567	*RELN*	0.008486	*DRD2*	0.14598
CHRNA7	-0.06326	*TPH1*	0.003565	*NRXN1*	0.145017
COMT	-0.08903	*GAD1*	-0.00551	*SLC1A1*	0.118932
DRD1	-0.09707	*HP*	-0.00838	*MTHFR*	0.118212
OFCC1	-0.1288	*RPGRIP1L*	-0.02015	*PRODH*	0.105492
RTN4R	-0.12902	*CHI3L1*	-0.02618	*OFCC1*	0.096665
DAOA	-0.1606	*PPP3CC*	-0.03326	*DAO*	0.078967
ERBB4	-0.16922	*APOL2*	-0.04071	*DAOA*	0.034813
HTR2A	-0.18204	*DRD3*	-0.04177	*IL1B*	0.022151
TNF	-0.18853	*DRD2*	-0.05543	*ZNF804A*	0.012573
DISC1	-0.18967	*DRD1*	-0.06693	*DRD3*	0.009339
APOL2	-0.20121	*MTHFR*	-0.09122	*FEZ1*	-0.00409
NRXN1	-0.20459	*DLG1*	-0.09304	*DISC1*	-0.04301
GABRB2	-0.22834	*GABRB2*	-0.10974	*SLC18A1*	-0.05426
CRP	-0.23874	*NRGN*	-0.12217	*CRP*	-0.05648
APOE	-0.24943	*OPCML*	-0.12587	*HP*	-0.07287
DAO	-0.27801	*NRXN1*	-0.13672	*NRG1*	-0.10715
CHI3L1	-0.28662	*GRM3*	-0.15224	*COMT*	-0.13578
DLG1	-0.28976	*HTR2A*	-0.15805	*APOE*	-0.14437
VIPR2	-0.29655	*DAOA*	-0.16778	*GRM3*	-0.17905

RGS4	-0.29852	*OFCC1*	-0.16787	*RELN*	-0.2322
IL18	-0.30341	*ERBB4*	-0.18177	*DRD4*	-0.23676
GAD1	-0.31423	*RGS4*	-0.18988	*SLC6A4*	-0.24208
SHANK3	-0.31754	*PLXNA2*	-0.20177	*TPH1*	-0.24774
SLC1A1	-0.32269	*DTNBP1*	-0.21695	*TP53*	-0.27373
GRM3	-0.33125	*GRIN2B*	-0.22215	*MUTED*	-0.29632
PRODH	-0.33168	*ZNF804A*	-0.22713	*NPAS3*	-0.29857
GRIK4	-0.33192	*CHRNA7*	-0.2499	*PLXNA2*	-0.31343
DLG2	-0.34973	*TNF*	-0.26524	*APOL4*	-0.32681
DTNBP1	-0.37558	*DAO*	-0.27631	*CYFIP1*	-0.36984
NRGN	-0.38989	*DLG2*	-0.27655	*PPP3CC*	-0.38205
FEZ1	-0.39014	*SLC1A1*	-0.33532	*EHMT1*	-0.519
HP	-0.41778	*FEZ1*	-0.42472	*AKT1*	-0.5222

Note: In the table, for each class of association strength (Weak/Moderate/Strong), training genes are sorted according to their Pearson correlation in spatiotemporal expression with the prioritized genes in the corresponding class. The correlation is calculated based on the spatiotemporal expression pattern of a gene set (shown in Figure 5) and the expression values of a training gene across the same combinations of brain regions and time stages. Training genes that have a correlation coefficient greater than 0.5 (in grey shading) are considered having high correlation in spatiotemporal expression with the corresponding prioritized gene set.

Table S10. The comparison between excluding and without excluding the extended MHC region.

Extended MHC region	Candidate genes	Scoring genes	High scoring genes
Not excluded	643	585	132
Excluded	545	502	121

Note: The 11 high scoring genes in the extended MHC region (chromosome 6 between 26Mb and 34 Mb in the hg19 assembly): *CYP21A2, AIF1, TAP1, NOTCH4, APOM, TNXB, HLA-DQB1, HLA-DRB1, AGER, HLA-DRA, and CLIC1*.

SUPPLEMETARY REFERENCES

Agrawal, R., H. Mannila, R. Srikant, H. Toivonen and A. I. Verkamo, 1995 Fast discovery of association rules in *Advances in Knowledge Discovery and Data Mining*. AAAI/MIT Press, Cambridge, MA.

Andersen, S. L., 2003 Trajectories of brain development: point of vulnerability or window of opportunity? Neurosci Biobehav Rev 27**:** 3-18.

Andersson, R., C. Gebhard, I. Miguel-Escalada, I. Hoof, J. Bornholdt *et al.*, 2014 An atlas of active enhancers across human cell types and tissues. Nature 507**:** 455-461.

Bossu, P., F. Piras, I. Palladino, M. Iorio, F. Salani *et al.*, 2015 Hippocampal volume and depressive symptoms are linked to serum IL-18 in schizophrenia. Neurol Neuroimmunol Neuroinflamm 2**:** e111.

Canetta, S., A. Sourander, H. M. Surcel, S. Hinkka-Yli-Salomaki, J. Leiviska *et al.*, 2014 Elevated maternal C-reactive protein and increased risk of schizophrenia in a national birth cohort. Am J Psychiatry 171**:** 960-968.

Consortium, G. T., 2013 The Genotype-Tissue Expression (GTEx) project. Nat Genet 45**:** 580-585.

Fillman, S. G., N. Cloonan, V. S. Catts, L. C. Miller, J. Wong *et al.*, 2013 Increased inflammatory markers identified in the dorsolateral prefrontal cortex of individuals with schizophrenia. Mol Psychiatry 18**:** 206-214.

Futamura, T., K. Toyooka, S. Iritani, K. Niizato, R. Nakamura *et al.*, 2002 Abnormal expression of epidermal growth factor and its receptor in the forebrain and serum of schizophrenic patients. Mol Psychiatry 7**:** 673-682.

Gardiner, E. J., M. J. Cairns, B. Liu, N. J. Beveridge, V. Carr *et al.*, 2013 Gene expression analysis reveals schizophrenia-associated dysregulation of immune pathways in peripheral blood mononuclear cells. J Psychiatr Res 47**:** 425-437.

Glatt, S. J., I. P. Everall, W. S. Kremen, J. Corbeil, R. Sasik *et al.*, 2005 Comparative gene expression analysis of blood and brain provides concurrent validation of SELENBP1 up-regulation in schizophrenia. Proc Natl Acad Sci U S A 102**:** 15533-15538.

Hall, J., S. Trent, K. L. Thomas, M. C. O'Donovan and M. J. Owen, 2015 Genetic risk for schizophrenia: convergence on synaptic pathways involved in plasticity. Biol Psychiatry 77**:** 52-58.

Hwang, Y., J. Kim, J. Y. Shin, J. I. Kim, J. S. Seo *et al.*, 2013 Gene expression profiling by mRNA sequencing reveals increased expression of immune/inflammation-related genes in the hippocampus of individuals with schizophrenia. Transl Psychiatry 3**:** e321.

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

Kotlar, A. V., K. B. Mercer, M. E. Zwick and J. G. Mulle, 2015 New discoveries in schizophrenia genetics reveal neurobiological pathways: A review of recent findings. Eur J Med Genet 58**:** 704-714.

Linghu, B., E. S. Snitkin, Z. Hu, Y. Xia and C. Delisi, 2009 Genome-wide prioritization of disease genes and identification of disease-disease associations from an integrated human functional linkage network. Genome Biol 10**:** R91.

Lv, M. H., Y. L. Tan, S. X. Yan, L. Tian, C. Chen da *et al.*, 2015 Decreased serum TNF-alpha levels in chronic schizophrenia patients on long-term antipsychotics: correlation with psychopathology and cognition. Psychopharmacology (Berl) 232**:** 165-172.

Maycox, P. R., F. Kelly, A. Taylor, S. Bates, J. Reid *et al.*, 2009 Analysis of gene expression in two large schizophrenia cohorts identifies multiple changes associated with nerve terminal function. Mol Psychiatry 14**:** 1083-1094.

McKusick, V. A., 2007 Mendelian Inheritance in Man and its online version, OMIM. Am J Hum Genet 80**:** 588-604.

Nawa, H., H. Sotoyama, Y. Iwakura, N. Takei and H. Namba, 2014 Neuropathologic implication of peripheral neuregulin-1 and EGF signals in dopaminergic dysfunction and behavioral deficits relevant to schizophrenia: their target cells and time window. Biomed Res Int 2014**:** 697935.

Perez-Santiago, J., R. Diez-Alarcia, L. F. Callado, J. X. Zhang, G. Chana *et al.*, 2012 A combined analysis of microarray gene expression studies of the human prefrontal cortex identifies genes implicated in schizophrenia. J Psychiatr Res 46**:** 1464-1474.

Purcell, S. M., J. L. Moran, M. Fromer, D. Ruderfer, N. Solovieff *et al.*, 2014 A polygenic burden of rare disruptive mutations in schizophrenia. Nature 506**:** 185-190.

Rappaport, N., M. Twik, N. Nativ, G. Stelzer, I. Bahir *et al.*, 2014 MalaCards: A Comprehensive Automatically-Mined Database of Human Diseases. Curr Protoc Bioinformatics 47**:** 1 24 21-19.

Sainz, J., I. Mata, J. Barrera, R. Perez-Iglesias, I. Varela *et al.*, 2013 Inflammatory and immune response genes have significantly altered expression in schizophrenia. Mol Psychiatry 18**:** 1056-1057.

Sanders, A. R., H. H. Goring, J. Duan, E. I. Drigalenko, W. Moy *et al.*, 2013 Transcriptome study of differential expression in schizophrenia. Hum Mol Genet 22**:** 5001-5014.

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

Sellmann, C., L. Villarin Pildain, A. Schmitt, F. Leonardi-Essmann, P. F. Durrenberger *et al.*, 2014 Gene expression in superior temporal cortex of schizophrenia patients. Eur Arch Psychiatry Clin Neurosci 264**:** 297-309.

Thurman, R. E., E. Rynes, R. Humbert, J. Vierstra, M. T. Maurano *et al.*, 2012 The accessible chromatin landscape of the human genome. Nature 489**:** 75-82.

Watanabe, Y., T. Someya and H. Nawa, 2010 Cytokine hypothesis of schizophrenia pathogenesis: evidence from human studies and animal models. Psychiatry Clin Neurosci 64**:** 217-230.