Research Paper

Identification of 34 genes conferring genetic and pharmacological risk for the comorbidity of schizophrenia and smoking behaviors

Yunlong Ma[1], Jingjing Li[1], Yi Xu[1], Yan Wang[1], Yinghao Yao[1], Qiang Liu[1], Maiqiu Wang[1], Xinyi Zhao[1], Rongli Fan[1], Jiali Chen[1], Bin Zhang[1], Zhen Cai[1], Haijun Han[1], Zhongli Yang[1], Wenji Yuan[1], Yigang Zhong[2], Xiangning Chen[3], Jennie Z. Ma[4], Thomas J. Payne[5], Yizhou Xu[2], Yuping Ning[6], Wenyan Cui[1], Ming D. Li[1,7]

[1]State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, National Clinical Research Center for Infectious Diseases, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China
[2]Department of Cardiology, Affiliated Hangzhou First People's Hospital, Zhejiang University School of Medicine, Hangzhou, China
[3]Institute of Personalized Medicine, University of Nevada at Las Vegas, Las Vegas, NV 89154, USA
[4]Department of Public Health Sciences, University of Virginia, Charlottesville, VA 22904, USA
[5]Department of Otolaryngology and Communicative Sciences, University of Mississippi Medical Center, Jackson, MS 39216, USA
[6]The Affiliated Brain Hospital of Guangzhou Medical University, Guangzhou, China
[7]Research Center for Air Pollution and Health, Zhejiang University, Hangzhou, China

Correspondence to: Ming D. Li, Wenyan Cui; **email:** ml2km@zju.edu.cn, cuiwy@zju.edu.cn
Keywords: schizophrenia, tobacco smoking, multi-omics, susceptibility genes, pathways, GWAS
Received: November 11, 2019 **Accepted:** January 2, 2020 **Published:** February 3, 2020

ABSTRACT

The prevalence of smoking is significantly higher in persons with schizophrenia (SCZ) than in the general population. However, the biological mechanisms of the comorbidity of smoking and SCZ are largely unknown. This study aimed to reveal shared biological pathways for the two diseases by analyzing data from two genome-wide association studies with a total sample size of 153,898. With pathway-based analysis, we first discovered 18 significantly enriched pathways shared by SCZ and smoking, which were classified into five groups: postsynaptic density, cadherin binding, dendritic spine, long-term depression, and axon guidance. Then, by using an integrative analysis of genetic, epigenetic, and expression data, we found not only 34 critical genes (e.g., *PRKCZ*, *ARHGEF3*, and *CDKN1A*) but also various risk-associated SNPs in these genes, which convey susceptibility to the comorbidity of the two disorders. Finally, using both *in vivo* and *in vitro* data, we demonstrated that the expression profiles of the 34 genes were significantly altered by multiple psychotropic drugs. Together, this multi-omics study not only reveals target genes for new drugs to treat SCZ but also reveals new insights into the shared genetic vulnerabilities of SCZ and smoking behaviors.

INTRODUCTION

Cigarette smoking is a common brain disorder that is extremely harmful to the individual and society. Smoking prevalence is much higher among individuals with mental disorders than in the general population [1]. In developed countries, the smoking rate among the general population has decreased over recent decades, whereas there has been no decrease among mental health patients. In particular, smoking shows a significant

association with schizophrenia (SCZ) [2], with around 80% of schizophrenic patients being smokers [3, 4]. There are three primary hypotheses intended to elucidate the comorbidity of smoking and SCZ [5–8]. One dominant hypothesis contends that smoking is able to remedy, at least partly, the symptoms of SCZ [6]. There are two main lines of evidence for this hypothesis. One is that nicotine intake enhances the metabolizing of antipsychotic drugs [7]; the other is that nicotine promotes the release of several neurotransmitters (e.g., dopamine, glutamate, and serotonin) and improves patient performance in memory and attention [4, 5]. The second hypothesis is that given that both diseases are highly influenced by genetics [9, 10], there might exist shared genetic components predisposing to both SCZ and smoking behaviors. Recently, many genetics-based studies have lent support to this hypothesis [11–14]. The third hypothesis is that smoking leads to the onset of SCZ in view of the fact that smoking initiation typically predates the appearance of SCZ [8]. A meta-analysis of cross-sectional and prospective studies reported that daily cigarette smoking was associated with an earlier age of onset of a psychosis [15]. Consistently, a large prospective study of Swedish registry data demonstrated that both light and heavy smoking were highly associated with a greater risk for SCZ [16]. To some extent, these hypotheses are mutually non-exclusive and may collectively contribute to the correlation of SCZ with smoking.

Increasing neuroimaging evidence supports the presence of an association between SCZ and smoking [17]. An fMRI study [18] showed that nicotine can restore deficient sensorimotor gating, which is associated with activation of the limbic regions and striatum in both SCZ patients and healthy controls. Compared with non-smoking SCZ patients and healthy subjects, SCZ patients with concurrent nicotine addiction have reduced grey matter volumes [19]. Furthermore, it is well documented that nicotine increases the release of dopamine, acetylcholine, glutamate, norepinephrine, and serotonin, which are all implicated in the etiology of SCZ [4].

Accumulating evidence has revealed shared genetic components for SCZ and smoking [11–14]. One of the most promising findings is that variants in the *CHRNA5/A3/B4* cluster on chromosome 15q24 are associated not only with nicotine dependence (ND) [12] but also with SCZ [11]. By using experimental evidence from animal models of addiction and SCZ, Koukouli et al. [20] found that in rodents, nicotine addiction-associated polymorphisms in *CHRNA5* provoke a decrease in neuronal activity, which mirrors the hypofrontiality detected in SCZ or addictive patients. Furthermore, there were multiple susceptibility genes reported to be associated with both SCZ and smoking risk, including *DRD2* [11, 21–23], *BDNF* [24], and *COMT* [25, 26], to name a few. As we know, psychiatric disorders including SCZ are highly comorbid diseases [27–30]; thus there exists a great comorbidity on the relations between SCZ and substance addictions [31, 32]. For example, a good number of reports [33–37] have concentrated on the genetic effect of the Val158Met polymorphism in *COMT* on the comorbidity of SCZ and substance addictions. In addition, Chen et al. [13] reported that ND is positively correlated with a polygenic risk score for SCZ whereas SCZ is positively correlated with the polygenic risk score for cotinine concentration. A recent study [38] revealed a statistically significant genetic correlation between SCZ and several smoking-related phenotypes.

To make further progress in the prevention and treatment of SCZ and smoking, it is essential to identify the etiologic biological pathways and susceptibility genes underlying the comorbidity of both disorders. Earlier genetics association studies or pathway-based studies focused primarily on either SCZ or smoking [39–42]. To date, only two reports [13, 14] investigated the genetic relations between SCZ and smoking based on pathway analysis results for a limited number of candidates or significant genes. To the best our knowledge, there has been no study providing an integrative genomics analysis based on multi-omics data and biological pathways for both SCZ and smoking. Therefore, the primary objective of this study was to identify susceptibility SNPs, genes, and pathways for the comorbidity of SCZ and smoking with the use of multi-omics data from various sources (Supplementary Figure 1).

RESULTS

GWAS-based enrichment analysis for SCZ and smoking behaviors

Our pathway analysis of GWAS summary statistics on SCZ and smoking phenotypes revealed 175, 172, 233, 225, and 158 significantly enriched pathways (q value < 0.1) for SCZ, CPD, ever smoking, former smoking, and age at smoking initiation, respectively. There were 84 pathways with a q value < 0.1 for SCZ that were in common with that for at least one smoking phenotype (Figure 1A and Supplementary Table 2). Of them, 18 showed significant enrichment in SCZ and all four smoking phenotypes (Table 1 and Figure 1A). For SCZ, the most significant pathway was postsynaptic density ($P < 3.04 \times 10^{-14}$), which is consistent with a previous report [43]. This postsynaptic density pathway also showed highly significant enrichment in all smoking-related behaviors, with P values from 3.09×10^{-6} for age at smoking initiation to 1.43×10^{-13} for CPD (Table 1),

confirming an important role for postsynaptic density in these two disorders. In contrast, we found that the enriched pathways in smoking behaviors had a much higher overlap with SCZ than that in the traits of null and height (Figure 1B, Supplementary Figures 2 and 3). Next, by using multidimensional scaling (MDS) for shared genes, we found that the 18 common pathways were categorized into five clusters (Figure 1C and Table 1): postsynaptic density (Cluster #1), cadherin binding (Cluster #2), dendritic spine (Cluster #3), long-term depression (Cluster #4), and axon guidance (Cluster #5).

Following the pathway analysis of GWAS summary data on SCZ and smoking, we performed gene-based analysis on the same dataset. For SCZ, we found 590 significantly associated genes after correction for multiple testing. Of them, 236 were located in 108 previously identified loci [11]. Even though numerous genes were significantly associated with CPD and former smoking, no one gene reached significance for ever-smoking or age at smoking initiation. We found 208 genes significantly associated with SCZ that also were associated with at least one of the four smoking phenotypes (Supplementary Table 3). Of these 208 shared genes, 70 were located in 108 previously reported loci for SCZ (Supplementary Figure 4). In addition, by performing computer permutation analysis, we found that the genes identified in our gene-based analysis were significantly overrepresented in the identified common pathways ($P < 0.003$; Figure 1D and Supplementary Figures 5 and 6).

Brain gene expression of common pathways

By analyzing brain expression data, we found 1,443 genes from the 84 common pathways to be coexpressed

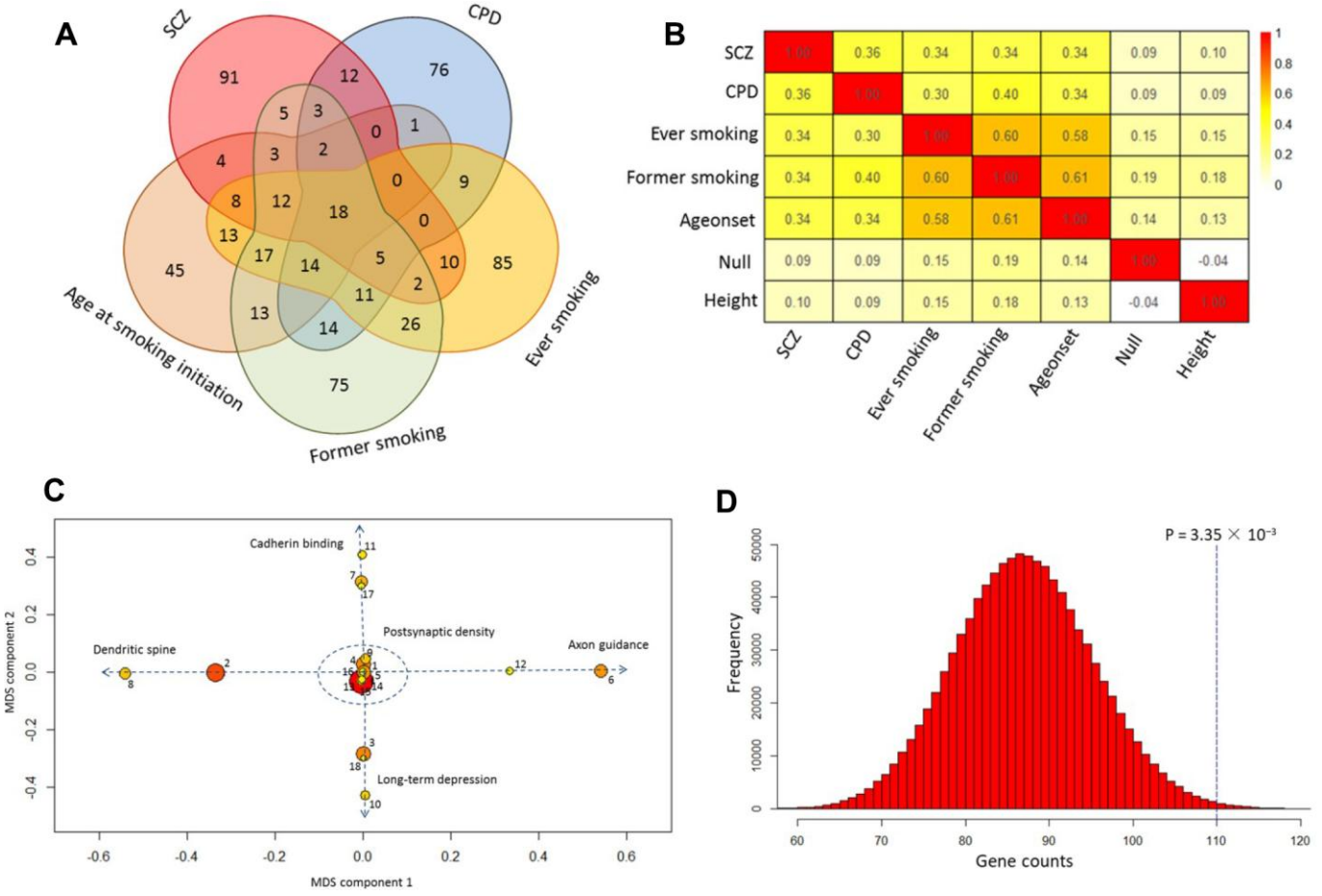


Figure 1. Shared pathways of SCZ and smoking phenotypes. (**A**) Venn diagram of significantly enriched pathways (q value < 0.1) for SCZ, CPD, ever smoking, former smoking, and age at smoking initiation. (**B**) Heatmap of the correlation among SCZ, CPD, ever smoking, former smoking, age at smoking initiation, null, and height based on the Z score of pathway enrichment. (**C**) Multidimensional scaling plot of 18 shared pathways for SCZ and smoking behaviors. Circular ring sizes reflect number of genes in the pathway (range 18–284). Color indicates the significance of the pathway (red marks the significant pathways with lowest P values). Arabic numerals are the pathway numbers as shown in Table 1. (**D**) Computer permutation analysis of 590 genes associated with SCZ in 84 shared pathways.

Table 1. Top 18 pathways shared between schizophrenia and smoking behaviors.

Pathway Number	Pathway ID	Description	Category	SCZ P-value	CPD P-value	Ever smoking P-value	Former smoking P-value	Age at smoking initiation P-value
1	GO:0014069	Postsynaptic density	Cluster #1	3.04×10^{-14}	1.43×10^{-13}	1.21×10^{-9}	1.50×10^{-6}	3.09×10^{-6}
2	GO:0045211	Postsynaptic membrane	Cluster #3	7.44×10^{-13}	4.97×10^{-11}	1.10×10^{-10}	5.92×10^{-9}	5.52×10^{-6}
3	GO:0045202	Synapse	Cluster #4	3.01×10^{-9}	2.76×10^{-13}	6.88×10^{-15}	1.91×10^{-9}	4.88×10^{-8}
4	GO:0030425	Dendrite	Cluster #1	9.46×10^{-9}	1.34×10^{-7}	0.0043	0.0011	0.0001
5	GO:0007268	Synaptic transmission	Cluster #1	3.22×10^{-8}	8.03×10^{-10}	1.85×10^{-11}	3.22×10^{-11}	9.84×10^{-6}
6	GO:0007411	Axon guidance	Cluster #5	5.89×10^{-8}	8.28×10^{-6}	6.04×10^{-12}	8.10×10^{-12}	3.64×10^{-9}
7	GO:0005001	Transmembrane receptor protein tyrosine phosphatase activity	Cluster #2	7.21×10^{-7}	0.0010	5.50×10^{-8}	0.0086	7.84×10^{-5}
8	GO:0043197	Dendritic spine	Cluster #3	4.34×10^{-6}	0.0009	0.0007	0.0011	0.0048
9	hsa05412	Arrhythmogenic right ventricular cardiomyopathy (ARVC)	Cluster #1	1.06×10^{-5}	0.0025	4.17×10^{-6}	1.36×10^{-5}	0.0008
10	hsa04912	GnRH signaling pathway	Cluster #4	7.76×10^{-5}	0.0056	8.06×10^{-5}	0.0042	0.0001
11	GO:0045296	Cadherin binding	Cluster #2	0.00014	0.0028	0.0008	0.0004	5.83×10^{-5}
12	GO:0007156	Homophilic cell adhesion	Cluster #5	0.00049	0.0011	1.48×10^{-5}	2.00×10^{-7}	1.50×10^{-7}
13	GO:0005216	Ion channel activity	Cluster #1	0.00051	2.95×10^{-13}	0.0001	1.29×10^{-5}	0.0001
14	GO:0051056	Regulation of small GTPase mediated signal transduction	Cluster #1	0.00053	0.0002	1.14×10^{-6}	4.88×10^{-5}	0.0002
15	hsa04020	Calcium signaling pathway	Cluster #1	0.0006	0.0056	2.74×10^{-10}	2.38×10^{-6}	2.99×10^{-6}
16	GO:0006813	Potassium ion transport	Cluster #1	0.00074	0.0038	2.73×10^{-6}	0.0017	9.14×10^{-5}
17	GO:0007626	Locomotory behavior	Cluster #2	0.0012	9.36×10^{-14}	0.0032	3.10×10^{-5}	0.0030
18	hsa04730	Long-term depression	Cluster #4	0.0042	7.10×10^{-5}	0.0012	2.85×10^{-6}	0.0009

in six modules (Supplementary Figures 7–9), and the top 10 "hub" genes in each module were demonstrated to have highly intramodular and intermodular connections (Supplementary Figures 10 and 11). Interestingly, we observed two distinct and dynamic expression patterns in different brain regions and at different developmental stages (Figure 2A). For example, one module (marked in yellow) has a greater than twofold difference between brain regions, and two modules (brown and red) showed twofold temporal changes (Supplementary Table 4).

Common methylated genes in both smoking and schizophrenia

By comparison with the previously reported findings [44], we found 149 module genes that overlapped with smoking-associated DNA-methylated genes; and 38 of these genes had at least two independent pieces of supporting evidence (Supplementary Table 5). Next, we performed a differential methylation analysis of 8,236 CpG loci in 149 genes related to SCZ and found that 822 CpG loci, mapped to 124 genes (124/149; binomial test: $P < 2.2 \times 10^{-16}$), showed significant association with SCZ ($P < 6.07 \times 10^{-6}$; Supplementary Table 5). For the 38 genes, 413 significantly methylated CpG loci were annotated in 34 of the genes, and there existed a high consistency of shared methylated genes between SCZ and smoking (34/38; binomial test: $P = 6.04 \times 10^{-7}$). Of note, 16 of these 34 candidate genes have been reported extensively to be associated with SCZ (Table 2 and Supplementary Table 6).

Through using these 34 genes to construct a subnetwork, we found 23 of 34 genes significantly enriched in that subnetwork (PPI enrichment: $P = 1.3 \times 10^{-4}$; Figure 2E). Many of these genes had several significantly SCZ-associated CpG loci (Figure 2B–2D and Supplementary Figures 12 and 13). For example, SCZ-associated cg10334053 ($P = 8.66 \times 10^{-21}$), cg02481000 ($P = 1.7 \times 10^{-19}$), and cg18188739 ($P = 3.56 \times 10^{-14}$) were all located in *PRKCZ* (Figure 2B). Of note, cg18188739 appeared to be associated with smoking at a genome-wide significance level ($P = 2.86 \times 10^{-8}$) [45].

Differential expression profiles of 34 candidate genes

We next examined the difference in the influence of smoking on gene expression between SCZ patients and healthy controls and found that 9 of 34 genes showed significantly different expression in SCZ subjects and

controls stratified by smoking status (Table 2; Supplementary Table 7; Figure 3; Supplementary Figures 14–16). Furthermore, several genes showed distinct expression profiles in schizophrenic hiPSC-derived neurons (Supplementary Figure 17) compared with the control neurons. Consistent with previous studies [11, 39], most of these candidate genes were highly expressed in human brain regions (Supplementary Figures 18–26).

The pharmacological effects of 34 candidate genes

Based on the DGIdb database, we found that 13 of the 34 genes (38.24%) were targeted by at least one drug. For example, *HRH1*, *CDKN1A*, and *TNF* are targeted by various psychotropic drugs. Also, 25 (73.53%) belong to one or more potentially "druggable" gene categories (Supplementary Figure 27). Consistently, the expression of several genes in the mouse brain was significantly modulated by quetiapine (Supplementary Figure 28) and nicotine (Supplementary Figure 29) in a dose-dependent pattern. Further, the count of significantly quetiapine- and nicotine-induced genes among these 34 genes were prominently higher than that of genes in 18 (Figure 3C, 3G) and 84 common pathways (Figure 3D, 3H) or all background genes (Figure 3E, 3I). For example, the expression of *Arhgef3* was downregulated by both quetiapine (Figure 3F: ANOVA P = 0.00018) and nicotine (Figure 3J: ANOVA P = 0.0084). In addition, we found that many genes showed different expression

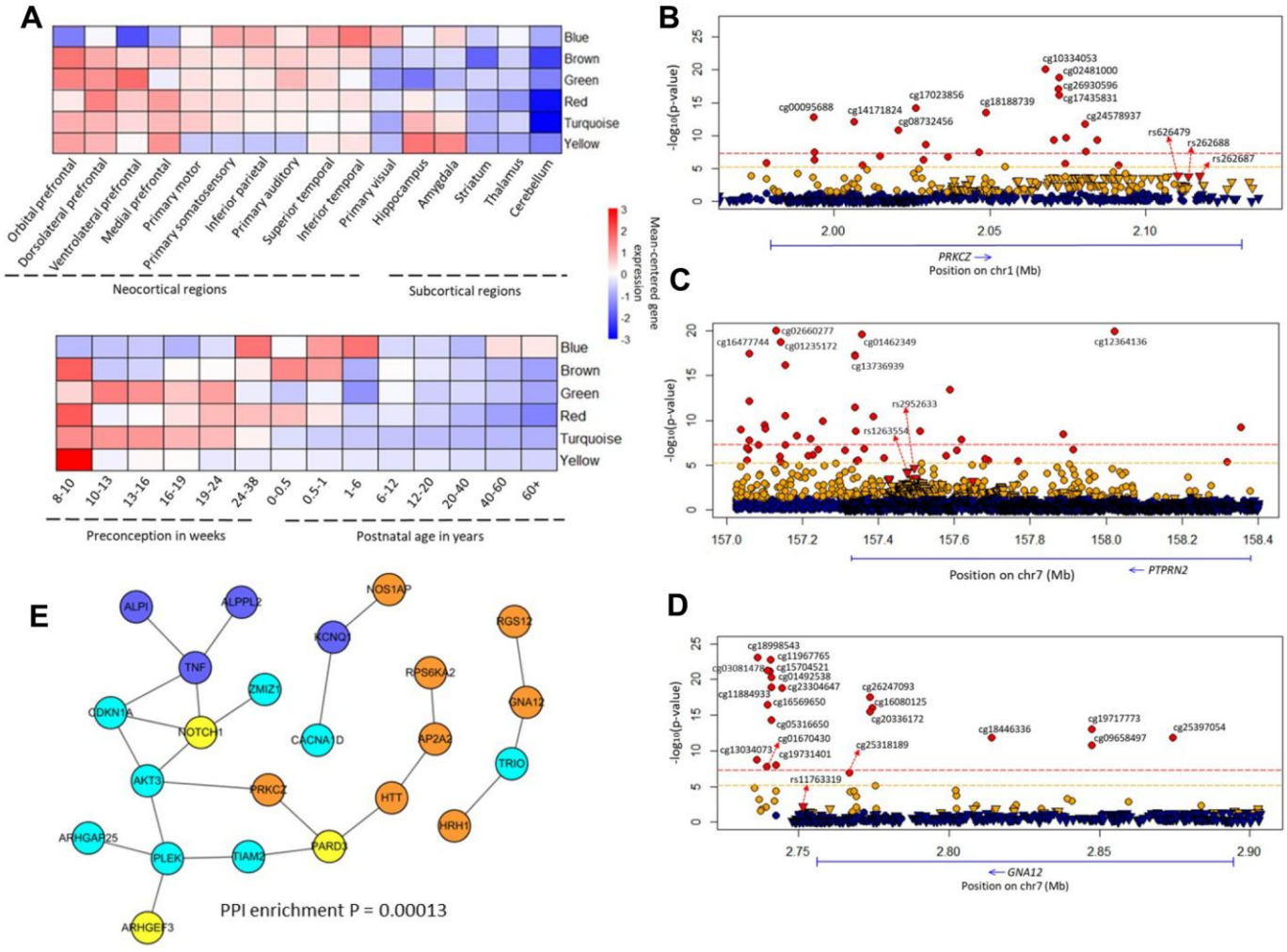


Figure 2. Brain co-expression modules and common methylated genes. (**A**) Regional and temporal patterns of gene expression mean-centered by the extent of gene expression within each module. (**B–D**) Regional plot of association between genetic and epigenetic variants of the *PRKCZ*, *PTPRN2*, *GNA12* loci, and SCZ, respectively. Circular symbols indicate the association of CpG loci with SCZ (red represents loci significantly associated with SCZ with $P \leq 6.07 \times 10^{-6}$; orange indicates loci with $6.07 \times 10^{-6} < P \leq 0.05$; blue marks loci with $P > 0.05$). Triangular symbols indicate association of SNPs with SCZ (red represents top-ranked SNPs associated with SCZ; orange indicates SNPs associated with SCZ with $P \leq 0.05$; blue marks SNPs with $P > 0.05$). (**E**) Gene subnetwork constituted from the 34 common genes. The protein–protein interactions are according to the database of STRING (v. 10.5). We used Cytoscape software to visualize the subnetwork. The color of a node indicates the co-expression module of the genes.

Table 2. Multiple lines of evidence support involvement of 34 genes in the comorbidity of both SCZ and smoking.

Gene names	Chromosome	Brain coexpression modules	Smoking-associated methylation genes	SCZ-associated methylation genes	cis-meQTLs in brain samples	Genes enriched in the PPI subnetwork	Differential expressed genes in SCZ patients stratified by smoking: Anova P values	Differential expression genes treated by quetiapine or nicotine: Anova P values	Genes risk to SCZ in previous studies
TRIO	Chr5	Turquoise	Yes	Yes	Yes	Yes	Not significant	0.031 (Quetiapine)	Yes (PMID: 21422296)
GNA12	Chr7	Brown	Yes	Yes	Yes	Yes	Not significant	0.049 (Quetiapine)	Yes (PMID: 22792057)
KCNQ1	Chr11	Blue	Yes	Yes	Yes	Yes	Not significant	0.0073 (Nicotine); 0.3 (Quetiapine)	Yes (PMID: 26971948; 28188958)
AKT3	Chr1	Turquoise	Yes	Yes	Yes	Yes	0.0395	0.018 (Nicotine)	Yes (PMID: 25056061; 28467426; 29173281; 23974872; 25599223)
ARHGEF3	Chr3	Yellow	Yes	Yes	Yes	Yes	0.00723	0.00018 (Quetiapine); 0.0084 (Nicotine)	No data
CACNA1D	Chr3	Turquoise	Yes	Yes	Yes	Yes	0.052	Not significant	Yes (PMID: 29214423; 26255836; 24996399)
RGS12	Chr4	Brown	Yes	Yes	Yes	Yes	0.0058	0.053 (Nicotine)	Yes (PMID: 25420024)
RPS6KA2	Chr6	Brown	Yes	Yes	Yes	Yes	Not significant	0.05 (Nicotine)	No data
NOTCH1	Chr9	Yellow	Yes	Yes	Yes	Yes	Not significant	0.0166 (Quetiapine)	Yes (PMID: 26232790)
PARD3	Chr10	Yellow	Yes	Yes	Yes	Yes	0.054	0.018 (Nicotine)	Yes (PMID: 22969987)
AP2A2	Chr11	Brown	Yes	Yes	No	Yes	Not significant	0.0032 (Nicotine)	Yes (PMID: 23811784)
NOS1AP	Chr1	Brown	Yes	Yes	Yes	Yes	Not significant	Not significant	Yes (PMID:26861996; 25542305; 20605702; 16146415; 15065015; 12116186; 19077434)
HRH1	Chr3	Brown	Yes	Yes	No	Yes	Not significant	Not significant	Yes (PMID:28400155; 27855565)
HTT	Chr4	Brown	Yes	Yes	No	Yes	Not significant	0.0085 (Quetiapine); 0.075 (Nicotine)	No data
CDKN1A	Chr6	Turquoise	Yes	Yes	No	Yes	Not significant	0.00085 (Quetiapine); 0.23 (Nicotine)	Yes (PMID: 23549417)
TIAM2	Chr6	Turquoise	Yes	Yes	Yes	Yes	0.062	Not significant	No data
TNF	Chr6	Blue	Yes	Yes	No	Yes	Not significant	0.06 (Nicotine)	Yes (PMID: 29499967; 29706448)
CNTNAP2	Chr7	Brown	Yes	Yes	No	No	0.087	0.0085 (Quetiapine)	Yes (PMID: 29610457;

									25852443; 23123147)
MAD1L1	Chr7	Brown	Yes	Yes	Yes	No	Not significant	0.024 (Nicotine)	Yes (PMID: 25056061; 26193471; 26528791)
PTPRN2	Chr7	Red	Yes	Yes	Yes	No	0.00142	0.0028 (Nicotine)	No data
CLCN6	Chr1	Green	Yes	Yes	Yes	No	Not significant	Not significant	No data
PRKCZ	Chr1	Brown	Yes	Yes	Yes	Yes	Not significant	Not significant	No data
ALPI	Chr2	Blue	Yes	Yes	Yes	Yes	Not significant	Not significant	No data
ALPPL2	Chr2	Blue	Yes	Yes	Yes	Yes	Not significant	Not significant	No data
ARHGAP25	Chr2	Turquoise	Yes	Yes	No	Yes	0.00026	Not significant	No data
PLEK	Chr2	Turquoise	Yes	Yes	No	Yes	0.043	0.018 (Nicotine)	No data
ZMIZ1	Chr10	Turquoise	Yes	Yes	Yes	Yes	0.041	0.0128 (Quetiapine)	No data
TBC1D14	Chr4	Turquoise	Yes	Yes	Yes	No	0.03	Not significant	No data
SORBS1	Chr10	Turquoise	Yes	Yes	Yes	No	Not significant	Not significant	No data
TGFBR3	Chr1	Turquoise	Yes	Yes	No	No	Not significant	Not significant	No data
CPOX	Chr3	Brown	Yes	Yes	No	No	Not significant	Not significant	No data
TIGIT	Chr3	Blue	Yes	Yes	No	No	Not significant	Not significant	No data
GPSM3	Chr6	Blue	Yes	Yes	No	No	0.0014	Not significant	No data
DUSP4	Chr8	Turquoise	Yes	Yes	No	No	Not significant	Not significant	No data

profiles in the mouse striatum at four time points of treatment with 18 major psychotropic drugs (Supplementary Figures 30–32).

Cis-regulatory effects of SNPs in 34 candidate genes

Based on two large GWAS datasets, there were 3,483 risk-suggested SNPs shared by SCZ and at least one smoking phenotype (Figure 4A and Supplementary Figure 33), and 228 of these were common to SCZ and all four smoking phenotypes (Figure 4B). A great number of SNPs among these identified 3,483 risk SNPs were located within different types of regulatory elements in brain tissues and neuroblastoma cell lines (Supplementary Figures 34 and 35). We performed a *cis*-meQTL analysis in human brain samples, which showed 7,558 significant SNP–CpG methylation pairs with 1,145 SNPs in 21 genes (FDR < 0.01; Figure 4A and Supplementary Figure 36). A number of the 392 variants in the 21 genes had *cis*-regulatory roles in both DNA methylation and gene expression (Figure 4A, 4C).

DISCUSSION

Cigarette smoking is highly concurrent with SCZ [2, 18, 46]. Individuals with mental disorders such as SCZ are at higher risk for developing smoking-related diseases, which include cardiovascular and respiratory diseases and various cancers [5]. Based on a national cohort study including about 6 million Swedish adults, Crump et al. [47] found a significant portion of the morbidity and premature death in persons with SCZ was ascribable to ischemic heart disease and cancers. Considering that the co-occurrence of SCZ and smoking has greatly impacted public health, it is important to understand the pathogenesis of the comorbidity of the two diseases. Previous studies concentrated largely on the investigation of the genetic mechanisms of either SCZ [11] or smoking [12]. Similarly, studies employing pathway-based enrichment analysis so far have focused mainly on either SCZ or smoking [39–42].

Recently, multiple large-scale GWAS studies [27–30, 38] consistently revealed that there exists a considerable genetic correlation between SCZ and smoking-related traits. However, susceptibility variants, genes, and biological pathways for the comorbidity remain largely unknown. In the present study, by conducting an integrative genomics analysis of large-scale GWAS data with multi-omics data, we intended to identify the risk SNPs, genes, and pathways implicated in the etiology of these two comorbid diseases.

We first conducted pathway-based enrichment analysis of GWAS summary data. For this hypothesis-free genome-wide approach, we utilized four commonly used resources (i.e., KEGG, GO, BioCarta, and Reactome) to determine the number of genes included in the GWAS pathway enrichment analyses [40], which could overcome the bias of previous studies for arbitrary gene

selection [13, 14]. This analysis revealed 18 significantly enriched pathways that were shared by SCZ and all four smoking phenotypes. These pathways were clustered into postsynaptic density, cadherin binding, dendritic spine, long-term depression, and axon guidance; all of them have been implicated in various psychiatric disorders [11, 13, 39, 43, 48–51]. Compared with the negative controls, the degrees of correlation between pathways across SCZ and smoking phenotypes were much higher, indicating that overlap between SCZ and smoking was nonrandom at any pathway level. These findings strongly indicate that the common pathways identified for both SCZ and smoking behaviors were attributable to the shared genetic vulnerability.

Given that leveraging multi-omics datasets is a better way to understand the molecular mechanism of complex diseases, we here provide robust evidence to explain the

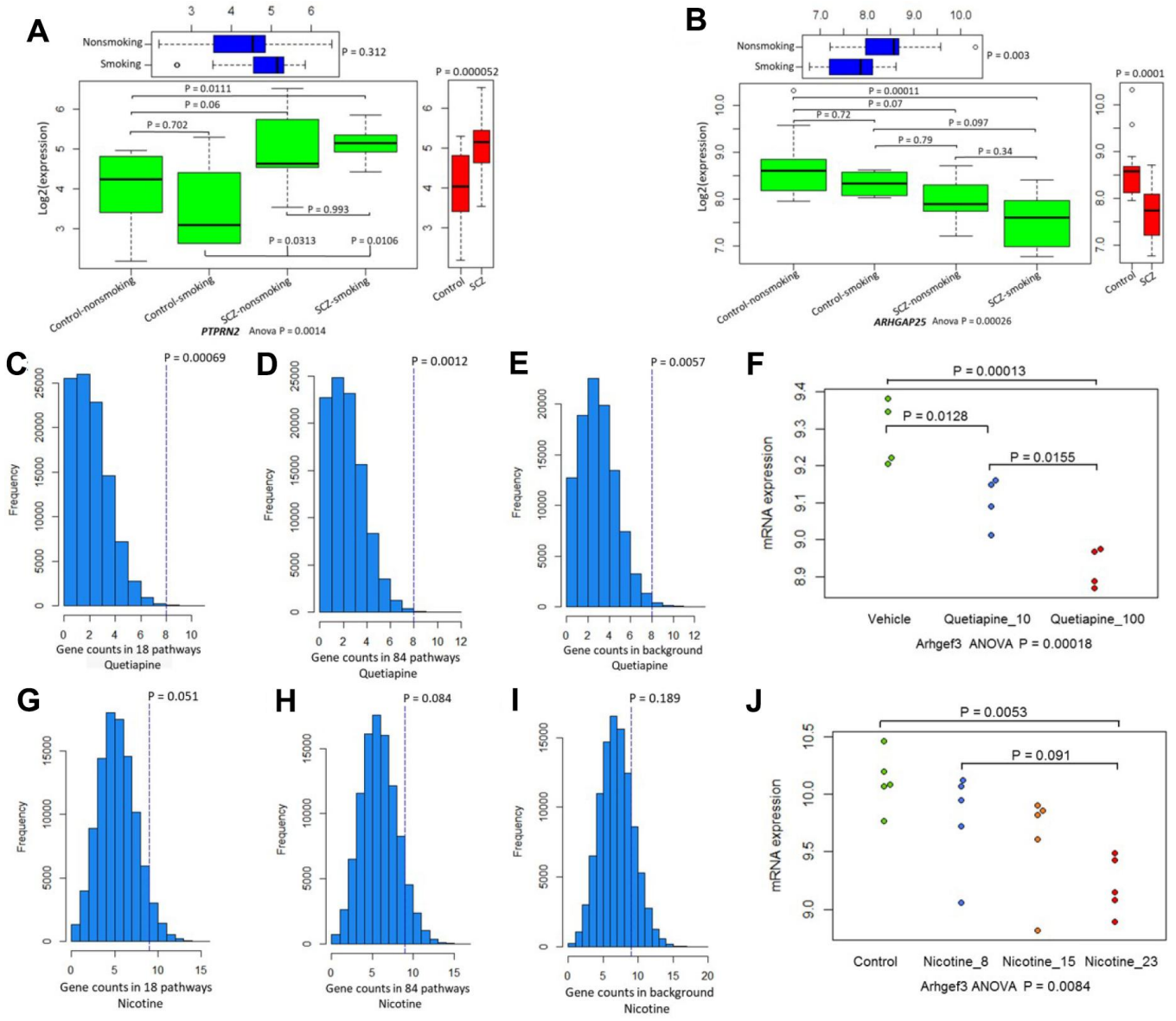


Figure 3. The differential expression patterns of 34 candidate genes. (**A**) Pattern of *PTPRN2* in SCZ patients and controls divided by smoking status. (**B**) Pattern of *ARHGAP25* in SCZ patients and controls divided by smoking status. (**C–E**) Permutation analysis of 34 candidate genes in 18 common pathways (N = 1,588 genes), 84 common pathways (N =3,334 genes), and background genes (N = 45,037) for quetiapine treatment. (**F**) Plot summarizing *Arhgef3* expression changes in quetiapine (at doses of 10 or 100 mg/kg)-treated mice. (**G–I**) Permutation analysis of 34 candidate genes in 18 common pathways (N = 1,588), 84 common pathways (N = 3,334), and background genes (N = 31,047) for nicotine treatment. (**J**) Plot summarizing *Arhgef3* expression alterations in nicotine (at doses of 8 μg, 15 μg, and 23 μg/L)-treated mice.

underlying mechanism of the comorbidity between SCZ and smoking from genetic, epigenetic, and expression points of view. Based on the gene co-expression data, we found a considerable number of genes in common pathways that showed differential co-expression patterns in different developmental time points and brain regions, providing supportive evidence that shared genetic risk for SCZ and smoking phenotypes has a vital influence on neurodevelopmentally regulated pathways. Similar to a previous study [43], by integrating gene co-expression data with pathway-based analyses on large-scale GWAS, we could identify even greater specificity of candidate genes conferring risk for the comorbidity.

Previous studies [11, 44, 52–54] have shown that aberrant methylated DNAs (DNAm) were involved in the etiology of SCZ and smoking. For example, our previous study, based on a sophisticated data mining of published papers [44], showed that 320 genes exhibited robust methylation evidence of involvement in the etiology of smoking. In the current study, we highlighted 34 genes with significant methylation evidence that

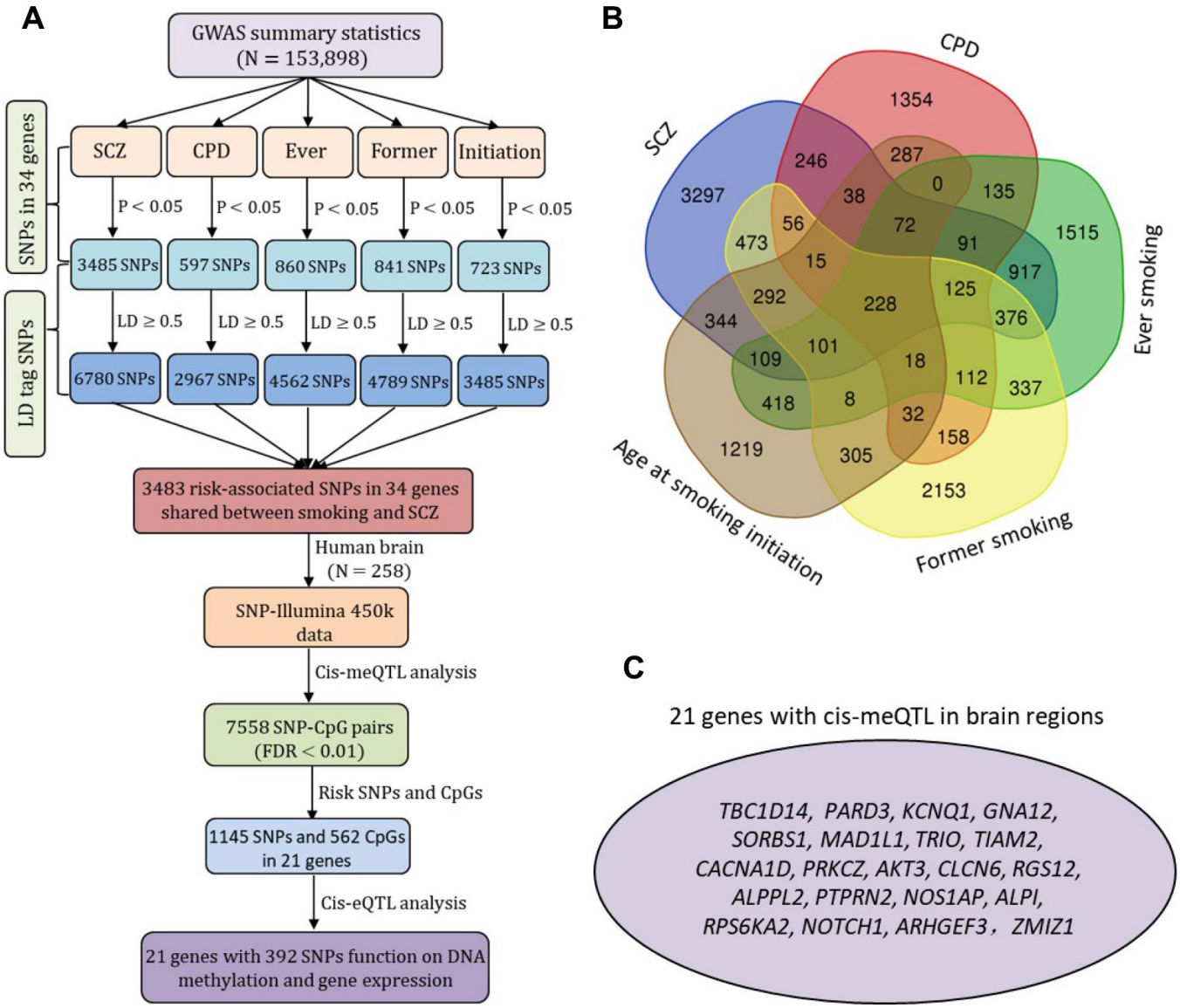


Figure 4. The *cis*-acting regulatory effects of risk-associated SNPs in 34 common genes on DNA methylation and gene expression. (**A**) Schematic of risk-associated SNPs (P < 0.05) compiled from two large-scale meta-GWASs on SCZ (N = 79,845) and smoking behaviors (N = 74,053) and 3,483 risk-associated SNPs in 34 genes shared by SCZ and smoking for *cis*-meQTL and *cis*-eQTL analysis. The SNPs in strong LD with risk-associated SNPs (LD cutoff $r^2 \geq 0.5$) were generated according to the 1000 Genome European Phase 3 panel as reference. (**B**) Venn diagram of risk-associated SNPs (P < 0.05) for SCZ, CPD, ever smoking, former smoking, and age at smoking initiation. (**C**) Plot shows the 21 promising genes with SNPs had *cis*-regulatory roles in both DNA methylation and gene expression.

were associated with both SCZ and smoking. More interestingly, we found that 23 of these 34 genes were located in a molecular subnetwork, which is consistent with the previous notion that function-related genes may collectively contribute the risk of the pathology of complex diseases [44]. Notably, 16 of the 34 methylated genes have often been reported to be associated with SCZ and smoking. For example, AKT3 is a critical molecule underlying psychiatric-related behaviors, including SCZ [11, 55], cognitive function [56], and smoking behaviors [44]. Multiple lines of evidence from genetic association studies [11, 57, 58] have indicated that *MAD1L1* is significantly associated with SCZ. Based on network-assisted investigation of combined causal signals from GWAS studies [59], *CNA12* showed a significant association with SCZ. *PRKCZ* and *PTPRN2* were reported to be SCZ-associated differential methylation regions [54]. The *KCNQ1* gene belongs to the potassium channel gene family and plays a vital role in signal transduction within the central nervous system. It may contribute to the shared risk of diminished processing speed, diminished white matter integrity, and increased risk of SCZ [60, 61]. In addition, we found that a large proportion of the 34 genes were targeted by various psychotropic drugs, and the expression patterns of these drug-responsive genes were significantly regulated in mouse brain by treatment with 18 major psychotropic drugs in a dose- or time-dependent pattern. Thus, these genes likely represent targets for pharmacotherapeutic intervention in SCZ and smoking behaviors.

Genetic variants influencing DNAm and mRNA expression can modulate gene transcript levels and thereby exert risk effects on diseases, as evidenced by meQTL and eQTL analyses of top GWAS risk-associated SNPs for various human phenotypes, including smoking [54, 62, 63] and SCZ [64, 65]. Thus, we first explored the vital influence of regulatory genomic elements (i.e., *cis*-meQTLs and *cis*-eQTLs) in neurodevelopmental processes of the comorbidity of smoking and SCZ. In human brain samples, we identified 21 of 34 candidate genes with a great number of SNP–CpG pairs. Interestingly, there was a large proportion, 39.4%, of these SNPs that showed allele-specific gene expression. By downloading SNPs from the GWAS Catalog (on 03 May 2019), some of the identified SNPs were found to be associated with psychiatric disorders in GWAS studies. For example, rs1107592 ($P = 2.0 \times 10^{-6}$) and rs802568 ($P = 2.0 \times 10^{-7}$) were suggested to be associated with bipolar disorder and SCZ [66]. Rs1107592 ($P = 5.0 \times 10^{-6}$), rs4721295 ($P = 6.0 \times 10^{-10}$), and rs12666575 ($P = 2.0 \times 10^{-9}$) were significantly associated with SCZ and other psychiatric disorders, such as autism spectrum, bipolar, and major depressive disorders [67–69]. Rs1403174 ($P = 3.0 \times 10^{-10}$) showed significant association with age at smoking initiation [70]. These findings were consistent with earlier reports where it was discovered that the majority of non-coding risk-associated SNPs for brain disorders influence gene expression via DNAm [71, 72].

Some limitations of the current study warrant comment. First, those SCZ- and smoking-related genes identified were prioritized from GWAS. In view of the inherent defects of the GWAS approach, some of these genes might not be truly associated with both of the disorders. Second, a great number of genes have not been characterized or mapped to computationally predicted or manually curated pathways. Thus, the contributions of these genes could not be delineated in pathway enrichment analysis. Third, although we have provided very strong evidence from genetic association, gene expression, and DNA methylation to support the 18 common pathways as being linked to both smoking and SCZ, we could not determine to what extent each specific pathway contributes to smoking or SCZ, nor could we quantify the relative influences of each pathway for the two disorders.

In sum, by employing a comprehensive bioinformatics analysis of genomic and pharmacogenomics data at the DNA, methylation, or expression levels, we first identified 18 biological pathways that are significantly associated with both SCZ and smoking phenotypes. Subsequently, we discovered 34 novel and promising susceptibility genes and variants within these genes with robust genetic, epigenetic, expression, and pharmacological evidence for both SCZ and smoking. Our findings lend considerable weight to the hypothesis that shared genetic vulnerabilities create a propensity for the comorbidity of SCZ and smoking. By pinpointing which and how SNPs in these candidate genes affect these disorders, this study provides novel insights into the biological mechanism and a solid foundation for understanding the shared disease biology of SCZ and smoking behaviors.

MATERIALS AND METHODS

Multi-omics datasets used in current study

In current investigation, we performed a series of multi-omics data analyses, which include large-scale meta-GWAS, BrainSpan exon array data on brain development and aging, methylation data, expression data, and pharmacogenomics data. The following is a brief summary of datasets used in our analysis. For details on these datasets, please infer to Supplementary Table 1.

1. GWAS data on SCZ: This dataset was obtained from a published GWAS [11] on SCZ from the

Psychiatric Genomics Consortium (PGC), of which the data from case-control samples (N = 79,845) were used for both gene and pathway analysis.

2. GWAS data on smoking: This dataset was from a published meta-GWAS on smoking behaviors (including smoking status, quantity smoked in ever-regular smokers, smoking cessation, and age at initiation) by the Tobacco and Genetics Consortium (TAG), which contained the genotype data for a total sample size of 74,053 [12].

3. GWAS data on height and null: To demonstrate that the identified pathways were attributable to shared biology between SCZ and smoking, we performed the GWAS enrichment analysis on two other independent data sets as negative controls. One of them was a GWAS of human height with a sample size of 183,727 [73], and the other was a null GWAS based on randomly distributed phenotypes that we constructed from a real GWAS with a sample size of 3,960 [74].

4. Expression data on brain development and aging: The BrainSpan exon array data related to brain development and aging were downloaded from the NCBI's Gene Expression Omnibus (GEO; Accession No. GSE25219), with a sample size of 1,340.

5. Methylation data on smoking and SCZ: Three datasets were used for this part of the analysis, with the first one from our previous study on smoking and consisting of 18,677 samples [44], the second one on 847 SCZ blood samples from NCBI GEO (Accession No. GSE84727) [75], and the third one on brain samples (N=258) from NCBI GEO (Accession No. GSE74193) [76].

6. Expression data on smoking and SCZ: Two expression datasets were used here, with the first one based on olfactory epithelium tissues (N = 31) and the second one on induced pluripotent human stem cells (hiPSCs; N = 8). Both were downloaded from NCBI GEO (Accession Nos. GSE73129 and GSE25673).

7. Pharmacogenomics data: To reveal potential druggable genes with therapeutic effects, we down-loaded the psychotropic drug-treated gene expression data from NCBI's GEO with Accession Nos. of GSE45229, GES50254, GSE48954, GSE15774, and GSE48951. We first explored the dosage influence of quetiapine and nicotine on gene expression changes in mouse striatum, then concentrated on the time-course (1, 2, 4, and 8 hours) of gene expression alterations in mouse striatum that were induced by 18 major psychotropic drugs.

Pathway- and gene-based analysis

We combined gene-set data from four sources: KEGG [77], GO [78], BioCarta [79], and Reactome [80], which were downloaded from their respective sources on or before May 12, 2017. Because of concerns that the gene set with > 300 genes or < 10 genes is considered to be either less specific and computationally inefficient or over-dispersed [40], we confined our analysis to those pathways with 10–300 genes.

We used the GSA-SNP program [81] to perform GWAS-based enrichment analysis. The SNPs of interest were assigned to genes if they lie within 20 kb upstream or downstream of the gene, and each SNP was assigned to only one gene. When multiple SNPs were mapped to the same gene, the GSA-SNP chose the most significantly associated SNP. Additionally, we used the MAGMA (https://ctg.cncr.nl/software/magma) for gene-based analysis of GWAS summary data. Because LD could result in a number of genes rather than one being counted as significant when genes are physically close in one region, we used the LD-pruned method to calculate the LD for the published GWAS data with the 1000 Genome European Phase 3 panel as the reference.

Computer permutation analysis

There were 3,334 genes (named Gene set 1) from 84 common pathways with a q value < 0.1 from our pathway enrichment analysis. To determine whether these genes were indeed significantly overrepresented with identified genes from our gene-based analysis, we conducted permutation analysis by randomly selecting 3,334 genes in the 84 common pathways from the total genes (N = 17,385) in all 2,532 pathways for 10^6 times. We then calculated how many times the counts of genes overlapped with the genes from gene-based analysis that were larger than the observed number among 10^6 trials. The probability of the observation was treated as the P-value, with a P value < 0.05 being considered significant.

Co-expression network analysis

We then conducted further analysis for 3,334 identified genes in Gene Set 1 to determine how these pathways were related to brain development and aging with the BrainSpan exon array data. RNA expression profiles for 1,340 samples were analyzed by using weighted gene co-expression network analysis (WGCNA), an R package used for clustering genes into modules according to co-expression data. The 10 most highly

connected genes within each module were displayed in the network plot by using Cytoscape (v. 3.5.1) (http://www.cytoscape.org/).

Smoking- and SCZ-associated differentially methylated loci and regions

To determine whether the identified genes from co-expression modules were involved in the smoking- and SCZ-associated methylation process, we first selected a list of 1,429 smoking-associated methylated genes from our previous study [44], and then examined whether these smoking-associated methylation module genes were significantly methylated in SCZ patients [75]. Bonferroni correction was used to determine significant association. An exact binomial analysis was applied to examine whether there existed a significant excess of consistence of observed methylated genes between smoking and SCZ more often than was expected by chance.

Candidate gene expression profiles in SCZ and smoking

We performed an ANOVA analysis to explore the differential expression profiles of the 34 genes in olfactory epithelium tissues among SCZ patients and controls grouped by smoking status. Turkey HSD test was used for multiple comparisons. Considering that human induced pluripotent stem cells (hiPSCs) provide a novel strategy for defining characteristics of schizophrenic neurons, we further used the first cell-based human model of SCZ by directly reprogramming fibroblasts from schizophrenic patients into hiPSCs and subsequently differentiating these disorder-specific hiPSCs into neurons *in vitro* to explore the different expression profiles of 34 genes between controls and schizophrenic patients.

Candidate genes in response to various psychotropic drugs

To identify potential druggable targets, we searched for these identified genes in the Drug-Gene Interaction database (DGIdb) (v. 3.0; http://www.dgidb.org/). Firstly, we searched the 20 databases for drug–gene interactions with FDA-approved pharmaceutical compounds according to 51 known interaction types with 34 common genes for both SCZ and smoking phenotypes. Secondly, we searched 10 databases for the potential drug ability for gene targets to reveal genes that might form targets for novel therapies in addition to existing medicines. To explore whether these genes had therapeutic effects, we applied the dosage treatment of quetiapine and nicotine, and the time-course (1, 2, 4, and 8 hours) treatment of 18 major psychotropic drugs on gene expression changes in mouse striatum. Furthermore, using the computer permutation analysis of 10^5 times, we determined whether these 34 candidate genes were more prone to drug-induced action than other 1,588 genes in 18 common pathways, 3,334 genes in 84 common pathways, as well as more than 30,000 background genes.

Cis-meQTLs/eQTLs of SNPs within 34 genes

To explore the relations between genotype and methylation status, we conducted cis-meQTL analysis of SNPs within 34 candidate genes. Based on the two large GWAS used in our pathway-based analysis, we collected suggested SNPs with a P value of < 0.05 in 34 candidate genes for SCZ and smoking behaviors. Considering that multiple lines of evidence have suggested that identified tag SNPs were more likely to be in LD with casual variants [82], we generated a list of SNPs that were in strong LD (LD cutoff $r^2 \geqslant 0.5$) with each tag SNP using the 1000 Genome European Phase 3 panel for reference genotyping.

By employing these identified SNPs, we downloaded *cis*-meQTL data from human brain samples (N = 258) [76] and used the Matrix eQTL (v. 2.1.1) R package [83] to examine the associations between SNPs and methylation loci with linear regression under an additive model. We restricted methylation loci to 20 kb upstream and downstream of each SNP. The intervals for nearby SNPs were combined if they overlapped. There were 7,426,085 common genotyped and imputed SNPs from the 1,000 Genomes reference panel and 477,636 qualified CpGs used for *cis*-meQTL analysis with a maximum distance of 20 kb between each SNP and CpG analyzed, resulting in 47,675,913 tests. A total of 4,107,214 significant SNP-CpG methylation associations at FDR < 0.01 were identified from this dataset. To further explore the *cis*-regulatory effects of SNPs on expression of the 34 candidate genes, we performed a *cis*-acting eQTL analysis in human tissues by using a web-based tool of GTEX PROTAL (https://gtexportal.org/home/).

Abbreviations

SCZ: schizophrenia; ND: nicotine dependence; GWAS: genome-wide association study; CPD: cigarette smoking per day; MDS: multidimensional scaling; PPI: protein-protein interaction; eQTL: expression quantitative trait loci; meQTL: methylation quantitative trait loci; PGC: the psychiatric genomics consortium; TAG: the tobacco and genetics consortium; GEO: Gene Expression Omnibus; hiPSCs: human induced pluripotent stem cells; KEGG: Kyoto Encyclopedia of Genes and Genomes; GO: gene ontology; WGCNA:

weighted gene co-expression network analysis; DGIdb: the Drug-Gene Interaction database; LD: linkage disequilibrium.

AUTHOR CONTRIBUTIONS

YM participated sample collection, data analysis and wrote the manuscript. JL, YX, and YY performed literature and public database searches. MW, QL, and XZ conducted the cis-meQTL analysis. RF, JC, and JL performed co-expression network analysis and other expression analysis. RF, JC, HH, BZ, YW, ZC, ZY, WC, WY, XC, YN, JZM, and TJP participated sample collection and/or reviewed the manuscript. MDL conceived the study and was involved in every step of the research. All authors approved the final manuscript.

ACKNOWLEDGMENTS

We thank Dr. Patrick F. Sullivan from the Departments of Genetics and Psychiatry in University of North Carolina for his critical comments and suggestions on the paper. We also thank Dr. David L. Bronson for excellent editing of this manuscript.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

FUNDING

This study was supported in part by the China Precision Medicine Initiative (2016YFC0906300), China Postdoctoral Science Foundation (2018M630667), Research Center for Air Pollution and Health of Zhejiang University, and the State Key Laboratory for Diagnosis and Treatment of Infectious Diseases of the First Affiliated Hospital of Zhejiang University. All authors declare no competing interest in relation to the work reported in this communication.

REFERENCES

1. Lasser K, Boyd JW, Woolhandler S, Himmelstein DU, McCormick D, Bor DH. Smoking and mental illness: A population-based prevalence study. JAMA. 2000; 284:2606–10. https://doi.org/10.1001/jama.284.20.2606 PMID:11086367

2. de Leon J, Diaz FJ. A meta-analysis of worldwide studies demonstrates an association between schizophrenia and tobacco smoking behaviors. Schizophr Res. 2005; 76:135–57. https://doi.org/10.1016/j.schres.2005.02.010 PMID:15949648

3. Huang W, Shen F, Zhang J, Xing B. Effect of Repetitive Transcranial Magnetic Stimulation on Cigarette Smoking in Patients with Schizophrenia. Shanghai Arch Psychiatry. 2016; 28:309–17. https://doi.org/10.11919/j.issn.1002-0829.216044 PMID:28638206

4. Dalack GW, Healy DJ, Meador-Woodruff JH. Nicotine dependence in schizophrenia: clinical phenomena and laboratory findings. Am J Psychiatry. 1998; 155:1490–501. https://doi.org/10.1176/ajp.155.11.1490 PMID:9812108

5. Morisano D, Bacher I, Audrain-McGovern J, George TP. Mechanisms underlying the comorbidity of tobacco use in mental health and addictive disorders. Can J Psychiatry. 2009; 54:356–67. https://doi.org/10.1177/070674370905400603 PMID:19527556

6. Smith RC, Singh A, Infante M, Khandat A, Kloos A. Effects of cigarette smoking and nicotine nasal spray on psychiatric symptoms and cognition in schizophrenia. Neuropsychopharmacology. 2002; 27:479–97. https://doi.org/10.1016/S0893-133X(02)00324-X PMID:12225705

7. Desai HD, Seabolt J, Jann MW. Smoking in patients receiving psychotropic medications: a pharmacokinetic perspective. CNS Drugs. 2001; 15:469–94. https://doi.org/10.2165/00023210-200115060-00005 PMID:11524025

8. Gage SH, Munafò MR. Smoking as a causal risk factor for schizophrenia. Lancet Psychiatry. 2015; 2:778–79. https://doi.org/10.1016/S2215-0366(15)00333-8 PMID:26236007

9. Li MD, Cheng R, Ma JZ, Swan GE. A meta-analysis of estimated genetic and environmental effects on smoking behavior in male and female adult twins. Addiction. 2003; 98:23–31. https://doi.org/10.1046/j.1360-0443.2003.00295.x PMID:12492752

10. Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. Arch Gen Psychiatry. 2003; 60:1187–92. https://doi.org/10.1001/archpsyc.60.12.1187 PMID:14662550

11. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014; 511:421–7. https://doi.org/10.1038/nature13595 PMID:25056061

12. Tobacco GC, and Tobacco and Genetics Consortium. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. Nat Genet. 2010;

42:441–47.
https://doi.org/10.1038/ng.571 PMID:20418890

13. Chen J, Bacanu SA, Yu H, Zhao Z, Jia P, Kendler KS, Kranzler HR, Gelernter J, Farrer L, Minica C, Pool R, Milaneschi Y, Boomsma DI, et al, and Cotinine meta-analysis group, and FTND meta-analysis group. Genetic Relationship between Schizophrenia and Nicotine Dependence. Sci Rep. 2016; 6:25671.
https://doi.org/10.1038/srep25671 PMID:27164557

14. Hu Y, Fang Z, Yang Y, Rohlsen-Neal D, Cheng F, Wang J. Analyzing the genes related to nicotine addiction or schizophrenia via a pathway and network based approach. Sci Rep. 2018; 8:2894.
https://doi.org/10.1038/s41598-018-21297-x
PMID:29440730

15. Gurillo P, Jauhar S, Murray RM, MacCabe JH. Does tobacco use cause psychosis? Systematic review and meta-analysis. Lancet Psychiatry. 2015; 2:718–25.
https://doi.org/10.1016/S2215-0366(15)00152-2
PMID:26249303

16. Kendler KS, Lönn SL, Sundquist J, Sundquist K. Smoking and schizophrenia in population cohorts of Swedish women and men: a prospective co-relative control study. Am J Psychiatry. 2015; 172:1092–100.
https://doi.org/10.1176/appi.ajp.2015.15010126
PMID:26046339

17. Ziedonis D, Hitsman B, Beckham JC, Zvolensky M, Adler LE, Audrain-McGovern J, Breslau N, Brown RA, George TP, Williams J, Calhoun PS, Riley WT. Tobacco use and cessation in psychiatric disorders: National Institute of Mental Health report. Nicotine Tob Res. 2008; 10:1691–715.
https://doi.org/10.1080/14622200802443569
PMID:19023823

18. Postma P, Gray JA, Sharma T, Geyer M, Mehrotra R, Das M, Zachariah E, Hines M, Williams SC, Kumari V. A behavioural and functional neuroimaging investigation into the effects of nicotine on sensorimotor gating in healthy subjects and persons with schizophrenia. Psychopharmacology (Berl). 2006; 184:589–99.
https://doi.org/10.1007/s00213-006-0307-5
PMID:16456657

19. Tregellas JR, Shatti S, Tanabe JL, Martin LF, Gibson L, Wylie K, Rojas DC. Gray matter volume differences and the effects of smoking on gray matter in schizophrenia. Schizophr Res. 2007; 97:242–49.
https://doi.org/10.1016/j.schres.2007.08.019
PMID:17890058

20. Koukouli F, Rooy M, Tziotis D, Sailor KA, O'Neill HC, Levenga J, Witte M, Nilges M, Changeux JP, Hoeffer CA, Stitzel JA, Gutkin BS, DiGregorio DA, Maskos U. Nicotine reverses hypofrontality in animal models of addiction and schizophrenia. Nat Med. 2017; 23:347–54.
https://doi.org/10.1038/nm.4274 PMID:28112735

21. Kaalund SS, Newburn EN, Ye T, Tao R, Li C, Deep-Soboslay A, Herman MM, Hyde TM, Weinberger DR, Lipska BK, Kleinman JE. Contrasting changes in DRD1 and DRD2 splice variant expression in schizophrenia and affective disorders, and associations with SNPs in postmortem brain. Mol Psychiatry. 2014; 19:1258–66.
https://doi.org/10.1038/mp.2013.165 PMID:24322206

22. Ma Y, Yuan W, Jiang X, Cui WY, Li MD. Updated findings of the association and functional studies of DRD2/ANKK1 variants with addictions. Mol Neurobiol. 2015; 51:281–99.
https://doi.org/10.1007/s12035-014-8826-2
PMID:25139281

23. Ma Y, Wang M, Yuan W, Su K, Li MD. The significant association of Taq1A genotypes in DRD2/ANKK1 with smoking cessation in a large-scale meta-analysis of Caucasian populations. Transl Psychiatry. 2015; 5:e686.
https://doi.org/10.1038/tp.2015.176
PMID:26624925

24. Novak G, LeBlanc M, Zai C, Shaikh S, Renou J, DeLuca V, Bulgin N, Kennedy JL, Le Foll B. Association of polymorphisms in the BDNF, DRD1 and DRD3 genes with tobacco smoking in schizophrenia. Ann Hum Genet. 2010; 74:291–98.
https://doi.org/10.1111/j.1469-1809.2010.00578.x
PMID:20456319

25. Beuten J, Payne TJ, Ma JZ, Li MD. Significant association of catechol-O-methyltransferase (COMT) haplotypes with nicotine dependence in male and female smokers of two ethnic populations. Neuropsychopharmacology. 2006; 31:675–84.
https://doi.org/10.1038/sj.npp.1300997
PMID:16395295

26. Twamley EW, Hua JP, Burton CZ, Vella L, Chinh K, Bilder RM, Kelsoe JR. Effects of COMT genotype on cognitive ability and functional capacity in individuals with schizophrenia. Schizophr Res. 2014; 159:114–17.
https://doi.org/10.1016/j.schres.2014.07.041
PMID:25139113

27. Anttila V, Bulik-Sullivan B, Finucane HK, Walters RK, Bras J, Duncan L, Escott-Price V, Falcone GJ, Gormley P, Malik R, Patsopoulos NA, Ripke S, Wei Z, et al, and Brainstorm Consortium. Analysis of shared heritability in common disorders of the brain. Science. 2018; 360:eaap8757.
https://doi.org/10.1126/science.aap8757
PMID:29930110

28. Bipolar Disorder and Schizophrenia Working Group of the Psychiatric Genomics Consortium. Electronic

address: douglas.ruderfer@vanderbilt.edu; Bipolar Disorder and Schizophrenia Working Group of the Psychiatric Genomics Consortium. Genomic Dissection of Bipolar Disorder and Schizophrenia, Including 28 Subphenotypes. Cell. 2018; 173:1705–1715.e16. https://doi.org/10.1016/j.cell.2018.05.046 PMID:29906448

29. Cross-Disorder Group of the Psychiatric Genomics Consortium. Electronic address: plee0@mgh.harvard.edu; Cross-Disorder Group of the Psychiatric Genomics Consortium. Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. Cell. 2019; 179:1469–1482.e11. https://doi.org/10.1016/j.cell.2019.11.020 PMID:31835028

30. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, Duncan L, Perry JR, Patterson N, Robinson EB, Daly MJ, Price AL, Neale BM, and ReproGen Consortium, and Psychiatric Genomics Consortium, and Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control Consortium 3. An atlas of genetic correlations across human diseases and traits. Nat Genet. 2015; 47:1236–41. https://doi.org/10.1038/ng.3406 PMID:26414676

31. de Leon J, Diaz FJ. Genetics of schizophrenia and smoking: an approach to studying their comorbidity based on epidemiological findings. Hum Genet. 2012; 131:877–901. https://doi.org/10.1007/s00439-011-1122-6 PMID:22190153

32. Carrà G, Johnson S, Crocamo C, Angermeyer MC, Brugha T, Azorin JM, Toumi M, Bebbington PE. Psychosocial functioning, quality of life and clinical correlates of comorbid alcohol and drug dependence syndromes in people with schizophrenia across Europe. Psychiatry Res. 2016; 239:301–07. https://doi.org/10.1016/j.psychres.2016.03.038 PMID:27046394

33. Henquet C, Rosa A, Krabbendam L, Papiol S, Fananás L, Drukker M, Ramaekers JG, van Os J. An experimental study of catechol-o-methyltransferase Val158Met moderation of delta-9-tetrahydrocannabinol-induced effects on psychosis and cognition. Neuropsychopharmacology. 2006; 31:2748–57. https://doi.org/10.1038/sj.npp.1301197 PMID:16936704

34. Carrà G, Nicolini G, Crocamo C, Lax A, Amidani F, Bartoli F, Castellano F, Chiorazzi A, Gamba G, Papagno C, Clerici M. Executive control in schizophrenia: a preliminary study on the moderating role of COMT Val158Met for comorbid alcohol and substance use disorders. Nord J Psychiatry. 2017; 71:332–39. https://doi.org/10.1080/08039488.2017.1286385 PMID:28635556

35. Carrà G, Nicolini G, Lax A, Bartoli F, Castellano F, Chiorazzi A, Gamba G, Bava M, Crocamo C, Papagno C. Facial emotion recognition in schizophrenia: an exploratory study on the role of comorbid alcohol and substance use disorders and COMT Val158Met. Hum Psychopharmacol. 2017; 32:e2630. https://doi.org/10.1002/hup.2630 PMID:28913946

36. Batalla A, Soriano-Mas C, López-Solà M, Torrens M, Crippa JA, Bhattacharyya S, Blanco-Hinojo L, Fagundo AB, Harrison BJ, Nogué S, de la Torre R, Farré M, Pujol J, Martín-Santos R. Modulation of brain structure by catechol-O-methyltransferase Val(158) Met polymorphism in chronic cannabis users. Addict Biol. 2014; 19:722–32. https://doi.org/10.1111/adb.12027 PMID:23311613

37. De Sousa KR, Tiwari AK, Giuffra DE, Mackenzie B, Zai CC, Kennedy JL. Age at onset of schizophrenia: cannabis, COMT gene, and their interactions. Schizophr Res. 2013; 151:289–90. https://doi.org/10.1016/j.schres.2013.10.037 PMID:24268936

38. Hartz SM, Horton AC, Hancock DB, Baker TB, Caporaso NE, Chen LS, Hokanson JE, Lutz SM, Marazita ML, McNeil DW, Pato CN, Pato MT, Johnson EO, Bierut LJ. Genetic correlation between smoking behaviors and schizophrenia. Schizophr Res. 2018; 194:86–90. https://doi.org/10.1016/j.schres.2017.02.022 PMID:28285025

39. Pers TH, Timshel P, Ripke S, Lent S, Sullivan PF, O'Donovan MC, Franke L, Hirschhorn JN, and Schizophrenia Working Group of the Psychiatric Genomics Consortium. Comprehensive analysis of schizophrenia-associated loci highlights ion channel pathways and biologically plausible candidate causal genes. Hum Mol Genet. 2016; 25:1247–54. https://doi.org/10.1093/hmg/ddw007 PMID:26755824

40. Minicã CC, Mbarek H, Pool R, Dolan CV, Boomsma DI, Vink JM. Pathways to smoking behaviours: biological insights from the Tobacco and Genetics Consortium meta-analysis. Mol Psychiatry. 2017; 22:82–88. https://doi.org/10.1038/mp.2016.20 PMID:27021816

41. Wang J, Li MD. Common and unique biological pathways associated with smoking initiation/progression, nicotine dependence, and smoking cessation. Neuropsychopharmacology. 2010; 35:702–19. https://doi.org/10.1038/npp.2009.178 PMID:19890259

42. Liu C, Bousman CA, Pantelis C, Skafidas E, Zhang D, Yue

W, Everall IP. Pathway-wide association study identifies five shared pathways associated with schizophrenia in three ancestral distinct populations. Transl Psychiatry. 2017; 7:e1037. https://doi.org/10.1038/tp.2017.8 PMID:28221366

43. Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. Nat Neurosci. 2015; 18:199–209. https://doi.org/10.1038/nn.3922 PMID:25599223

44. Ma Y, Li MD. Establishment of a Strong Link Between Smoking and Cancer Pathogenesis through DNA Methylation Analysis. Sci Rep. 2017; 7:1811. https://doi.org/10.1038/s41598-017-01856-4 PMID:28500316

45. Teschendorff AE, Yang Z, Wong A, Pipinikas CP, Jiao Y, Jones A, Anjum S, Hardy R, Salvesen HB, Thirlwell C, Janes SM, Kuh D, Widschwendter M. Correlation of Smoking-Associated DNA Methylation Changes in Buccal Cells With DNA Methylation Changes in Epithelial Cancer. JAMA Oncol. 2015; 1:476–85. https://doi.org/10.1001/jamaoncol.2015.1053 PMID:26181258

46. Kelly C, McCreadie R. Cigarette smoking and schizophrenia. Adv Psychiatr Treat. 2000; 6:327–31. https://doi.org/10.1192/apt.6.5.327

47. Crump C, Winkleby MA, Sundquist K, Sundquist J. Comorbidities and mortality in persons with schizophrenia: a Swedish national cohort study. Am J Psychiatry. 2013; 170:324–33. https://doi.org/10.1176/appi.ajp.2012.12050599 PMID:23318474

48. Konopaske GT, Lange N, Coyle JT, Benes FM. Prefrontal cortical dendritic spine pathology in schizophrenia and bipolar disorder. JAMA Psychiatry. 2014; 71:1323–31. https://doi.org/10.1001/jamapsychiatry.2014.1582 PMID:25271938

49. Børglum AD, Demontis D, Grove J, Pallesen J, Hollegaard MV, Pedersen CB, Hedemand A, Mattheisen M, Uitterlinden A, Nyegaard M, Ørntoft T, Wiuf C, Didriksen M, et al, and GROUP investigators10. Genome-wide study of association and interaction with maternal cytomegalovirus infection suggests new schizophrenia loci. Mol Psychiatry. 2014; 19:325–33. https://doi.org/10.1038/mp.2013.2 PMID:23358160

50. Telley L, Cadilhac C, Cioni JM, Saywell V, Jahannault-Talignani C, Huettl RE, Sarrailh-Faivre C, Dayer A, Huber AB, Ango F. Dual Function of NRP1 in Axon Guidance and Subcellular Target Recognition in Cerebellum. Neuron. 2016; 91:1276–91. https://doi.org/10.1016/j.neuron.2016.08.015 PMID:27618676

51. Van Battum EY, Brignani S, Pasterkamp RJ. Axon guidance proteins in neurological disorders. Lancet Neurol. 2015; 14:532–46. https://doi.org/10.1016/S1474-4422(14)70257-1 PMID:25769423

52. Grayson DR, Guidotti A. The dynamics of DNA methylation in schizophrenia and related psychiatric disorders. Neuropsychopharmacology. 2013; 38:138–66. https://doi.org/10.1038/npp.2012.125 PMID:22948975

53. Waterland RA, Michels KB. Epigenetic epidemiology of the developmental origins hypothesis. Annu Rev Nutr. 2007; 27:363–88. https://doi.org/10.1146/annurev.nutr.27.061406.093705 PMID:17465856

54. Hannon E, Dempster E, Viana J, Burrage J, Smith AR, Macdonald R, St Clair D, Mustard C, Breen G, Therman S, Kaprio J, Toulopoulou T, Hulshoff Pol HE, et al. An integrated genetic-epigenetic analysis of schizophrenia: evidence for co-localization of genetic associations and differential DNA methylation. Genome Biol. 2016; 17:176. https://doi.org/10.1186/s13059-016-1041-x PMID:27572077

55. Ripke S, O'Dushlaine C, Chambert K, Moran JL, Kähler AK, Akterin S, Bergen SE, Collins AL, Crowley JJ, Fromer M, Kim Y, Lee SH, Magnusson PK, et al, and Multicenter Genetic Studies of Schizophrenia Consortium, and Psychosis Endophenotypes International Consortium, and Wellcome Trust Case Control Consortium 2. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. Nat Genet. 2013; 45:1150–59. https://doi.org/10.1038/ng.2742 PMID:23974872

56. Howell KR, Floyd K, Law AJ. PKBγ/AKT3 loss-of-function causes learning and memory deficits and deregulation of AKT/mTORC2 signaling: relevance for schizophrenia. PLoS One. 2017; 12:e0175993. https://doi.org/10.1371/journal.pone.0175993 PMID:28467426

57. Chang S, Fang K, Zhang K, Wang J. Network-Based Analysis of Schizophrenia Genome-Wide Association Data to Detect the Joint Functional Association Signals. PLoS One. 2015; 10:e0133404. https://doi.org/10.1371/journal.pone.0133404 PMID:26193471

58. Su L, Shen T, Huang G, Long J, Fan J, Ling W, Jiang J. Genetic association of GWAS-supported MAD1L1 gene polymorphism rs12666575 with schizophrenia

susceptibility in a Chinese population. Neurosci Lett. 2016; 610:98–103. https://doi.org/10.1016/j.neulet.2015.10.061 PMID:26528791

59. Jia P, Wang L, Fanous AH, Pato CN, Edwards TL, Zhao Z, and International Schizophrenia Consortium. Network-assisted investigation of combined causal signals from genome-wide association studies in schizophrenia. PLOS Comput Biol. 2012; 8:e1002587. https://doi.org/10.1371/journal.pcbi.1002587 PMID:22792057

60. Rannals MD, Hamersky GR, Page SC, Campbell MN, Briley A, Gallo RA, Phan BN, Hyde TM, Kleinman JE, Shin JH, Jaffe AE, Weinberger DR, Maher BJ. Psychiatric Risk Gene Transcription Factor 4 Regulates Intrinsic Excitability of Prefrontal Neurons via Repression of SCN10a and KCNQ1. Neuron. 2016; 90:43–55. https://doi.org/10.1016/j.neuron.2016.02.021 PMID:26971948

61. Bruce HA, Kochunov P, Paciga SA, Hyde CL, Chen X, Xie Z, Zhang B, Xi HS, O'Donnell P, Whelan C, Schubert CR, Bellon A, Ament SA, et al. Potassium channel gene associations with joint processing speed and white matter impairments in schizophrenia. Genes Brain Behav. 2017; 16:515–21. https://doi.org/10.1111/gbb.12372 PMID:28188958

62. Takata A, Matsumoto N, Kato T. Genome-wide identification of splicing QTLs in the human brain and their enrichment among schizophrenia-associated loci. Nat Commun. 2017; 8:14519. https://doi.org/10.1038/ncomms14519 PMID:28240266

63. Hannon E, Spiers H, Viana J, Pidsley R, Burrage J, Murphy TM, Troakes C, Turecki G, O'Donovan MC, Schalkwyk LC, Bray NJ, Mill J. Methylation QTLs in the developing brain and their enrichment in schizophrenia risk loci. Nat Neurosci. 2016; 19:48–54. https://doi.org/10.1038/nn.4182 PMID:26619357

64. Hancock DB, Wang JC, Gaddis NC, Levy JL, Saccone NL, Stitzel JA, Goate A, Bierut LJ, Johnson EO. A multiancestry study identifies novel genetic associations with CHRNA5 methylation in human brain and risk of nicotine dependence. Hum Mol Genet. 2015; 24:5940–54. https://doi.org/10.1093/hmg/ddv303 PMID:26220977

65. Liu Q, Han H, Wang M, Yao Y, Wen L, Jiang K, Ma Y, Fan R, Chen J, Su K, Yang Z, Cui W, Yuan W, et al. Association and cis-mQTL analysis of variants in CHRNA3-A5, CHRNA7, CHRNB2, and CHRNB4 in relation to nicotine dependence in a Chinese Han population. Transl Psychiatry. 2018; 8:83. https://doi.org/10.1038/s41398-018-0130-x PMID:29666375

66. Wang KS, Liu XF, Aragam N. A genome-wide meta-analysis identifies novel loci associated with schizophrenia and bipolar disorder. Schizophr Res. 2010; 124:192–99. https://doi.org/10.1016/j.schres.2010.09.002 PMID:20889312

67. Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. Lancet. 2013; 381:1371–79. https://doi.org/10.1016/S0140-6736(12)62129-1 PMID:23453885

68. Sleiman P, Wang D, Glessner J, Hadley D, Gur RE, Cohen N, Li Q, Hakonarson H, Janssen CN, and Janssen-CHOP Neuropsychiatric Genomics Working Group. GWAS meta analysis identifies TSNARE1 as a novel Schizophrenia / Bipolar susceptibility locus. Sci Rep. 2013; 3:3075. https://doi.org/10.1038/srep03075 PMID:24166486

69. Bergen SE, O'Dushlaine CT, Ripke S, Lee PH, Ruderfer DM, Akterin S, Moran JL, Chambert KD, Handsaker RE, Backlund L, Ösby U, McCarroll S, Landen M, et al. Genome-wide association study in a Swedish population yields support for greater CNV and MHC involvement in schizophrenia compared with bipolar disorder. Mol Psychiatry. 2012; 17:880–86. https://doi.org/10.1038/mp.2012.73 PMID:22688191

70. Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, Chen F, Datta G, Davila-Velderrain J, McGuire D, Tian C, Zhan X, Choquet H, Docherty AR, et al, and 23andMe Research Team, and HUNT All-In Psychiatry. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. Nat Genet. 2019; 51:237–44. https://doi.org/10.1038/s41588-018-0307-5 PMID:30643251

71. Schulz H, Ruppert AK, Herms S, Wolf C, Mirza-Schreiber N, Stegle O, Czamara D, Forstner AJ, Sivalingam S, Schoch S, Moebus S, Pütz B, Hillmer A, et al. Genome-wide mapping of genetic determinants influencing DNA methylation and gene expression in human hippocampus. Nat Commun. 2017; 8:1511. https://doi.org/10.1038/s41467-017-01818-4 PMID:29142228

72. Barr CL, Misener VL. Decoding the non-coding genome: elucidating genetic risk outside the coding genome. Genes Brain Behav. 2016; 15:187–204. https://doi.org/10.1111/gbb.12269 PMID:26515765

73. Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, Willer CJ, Jackson AU, Vedantam S, Raychaudhuri S, Ferreira T, Wood AR, Weyant RJ, et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature.

2010; 467:832–38.
https://doi.org/10.1038/nature09410 PMID:20881960

74. Landi MT, Chatterjee N, Yu K, Goldin LR, Goldstein AM, Rotunno M, Mirabello L, Jacobs K, Wheeler W, Yeager M, Bergen AW, Li Q, Consonni D, et al. A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma. Am J Hum Genet. 2009; 85:679–91.
https://doi.org/10.1016/j.ajhg.2009.09.012
PMID:19836008

75. Diboun I, Wernisch L, Orengo CA, Koltzenburg M. Microarray analysis after RNA amplification can detect pronounced differences in gene expression using limma. BMC Genomics. 2006; 7:252.
https://doi.org/10.1186/1471-2164-7-252
PMID:17029630

76. Jaffe AE, Gao Y, Deep-Soboslay A, Tao R, Hyde TM, Weinberger DR, Kleinman JE. Mapping DNA methylation across development, genotype and schizophrenia in the human frontal cortex. Nat Neurosci. 2016; 19:40–47.
https://doi.org/10.1038/nn.4181
PMID:26619358

77. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000; 28:27–30.
https://doi.org/10.1093/nar/28.1.27 PMID:10592173

78. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, et al, and The Gene Ontology Consortium. Gene ontology: tool for the unification of biology. Nat Genet. 2000; 25:25–29.
https://doi.org/10.1038/75556 PMID:10802651

79. Nishimura D. BioCarta. Biotech Software & Internet Report: The Computer Software Journal for Scient. 2001; 2:117–20.
https://doi.org/10.1089/152791601750294344

80. Joshi-Tope G, Gillespie M, Vastrik I, D'Eustachio P, Schmidt E, de Bono B, Jassal B, Gopinath GR, Wu GR, Matthews L, Lewis S, Birney E, Stein L. Reactome: a knowledgebase of biological pathways. Nucleic Acids Res. 2005; 33:D428–32.
https://doi.org/10.1093/nar/gki072 PMID:15608231

81. Nam D, Kim J, Kim SY, Kim S. GSA-SNP: a general approach for gene set analysis of polymorphisms. Nucleic Acids Res. 2010 (suppl_2); 38:W749–54.
https://doi.org/10.1093/nar/gkq428 PMID:20501604

82. Cowper-Sal lari R, Zhang X, Wright JB, Bailey SD, Cole MD, Eeckhoute J, Moore JH, Lupien M. Breast cancer risk-associated SNPs modulate the affinity of chromatin for FOXA1 and alter gene expression. Nat Genet. 2012; 44:1191–8.
https://doi.org/10.1038/ng.2416 PMID:23001124

83. Shabalin AA. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. Bioinformatics. 2012; 28:1353–58.
https://doi.org/10.1093/bioinformatics/bts163
PMID:22492648

SUPPLEMENTARY MATERIALS

Supplementary References

1. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014; 511:421–27. https://doi.org/10.1038/nature13595 PMID:25056061

2. Tobacco and Genetics Consortium. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. Nat Genet. 2010; 42:441–47. https://doi.org/10.1038/ng.571 PMID:20418890

3. Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, Willer CJ, Jackson AU, Vedantam S, Raychaudhuri S, Ferreira T, Wood AR, Weyant RJ, et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature. 2010; 467:832–38. https://doi.org/10.1038/nature09410 PMID:20881960

4. Landi MT, Chatterjee N, Yu K, Goldin LR, Goldstein AM, Rotunno M, Mirabello L, Jacobs K, Wheeler W, Yeager M, Bergen AW, Li Q, Consonni D, et al. A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma. Am J Hum Genet. 2009; 85:679–91. https://doi.org/10.1016/j.ajhg.2009.09.012 PMID:19836008

5. Ma Y, Li MD. Establishment of a Strong Link Between Smoking and Cancer Pathogenesis through DNA Methylation Analysis. Sci Rep. 2017; 7:1811. https://doi.org/10.1038/s41598-017-01856-4 PMID:28500316

6. Hannon E, Dempster E, Viana J, Burrage J, Smith AR, Macdonald R, St Clair D, Mustard C, Breen G, Therman S, Kaprio J, Toulopoulou T, Hulshoff Pol HE, et al. An integrated genetic-epigenetic analysis of schizophrenia: evidence for co-localization of genetic associations and differential DNA methylation. Genome Biol. 2016; 17:176. https://doi.org/10.1186/s13059-016-1041-x PMID:27572077

7. Jaffe AE, Gao Y, Deep-Soboslay A, Tao R, Hyde TM, Weinberger DR, Kleinman JE. Mapping DNA methylation across development, genotype and schizophrenia in the human frontal cortex. Nat Neurosci. 2016; 19:40–47. https://doi.org/10.1038/nn.4181 PMID:26619358

8. Horiuchi Y, Kondo MA, Okada K, Takayanagi Y, Tanaka T, Ho T, Varvaris M, Tajinda K, Hiyama H, Ni K, Colantuoni C, Schretlen D, Cascella NG, et al. Molecular signatures associated with cognitive deficits in schizophrenia: a study of biopsied olfactory neural epithelium. Transl Psychiatry. 2016; 6:e915. https://doi.org/10.1038/tp.2016.154 PMID:27727244

9. Horiuchi Y, Kano S, Ishizuka K, Cascella NG, Ishii S, Talbot CC Jr, Jaffe AE, Okano H, Pevsner J, Colantuoni C, Sawa A. Olfactory cells via nasal biopsy reflect the developing brain in gene expression profiles: utility and limitation of the surrogate tissues in research for brain disorders. Neurosci Res. 2013; 77:247–50. https://doi.org/10.1016/j.neures.2013.09.010 PMID:24120685

10. Cascella NG, Takaki M, Lin S, Sawa A. Neurodevelopmental involvement in schizophrenia: the olfactory epithelium as an alternative model for research. J Neurochem. 2007; 102:587–94. https://doi.org/10.1111/j.1471-4159.2007.04628.x PMID:17488280

11. Brennand KJ, Simone A, Jou J, Gelboin-Burkhart C, Tran N, Sangar S, Li Y, Mu Y, Chen G, Yu D, McCarthy S, Sebat J, Gage FH. Modelling schizophrenia using human induced pluripotent stem cells. Nature. 2011; 473:221–25. https://doi.org/10.1038/nature09915 PMID:21490598

12. Kondo MA, Tajinda K, Colantuoni C, Hiyama H, Seshadri S, Huang B, Pou S, Furukori K, Hookway C, Jaaro-Peled H, Kano SI, Matsuoka N, Harada K, et al. Unique pharmacological actions of atypical neuroleptic quetiapine: possible role in cell cycle/fate control. Transl Psychiatry. 2013; 3:e243. https://doi.org/10.1038/tp.2013.19 PMID:23549417

13. Kogel U, Schlage WK, Martin F, Xiang Y, Ansari S, Leroy P, Vanscheeuwijck P, Gebel S, Buettner A, Wyss C, Esposito M, Hoeng J, Peitsch MC. A 28-day rat inhalation study with an integrated molecular toxicology endpoint demonstrates reduced exposure effects for a prototypic modified risk tobacco product compared with conventional cigarettes. Food Chem Toxicol. 2014; 68:204–17. https://doi.org/10.1016/j.fct.2014.02.034 PMID:24632068

14. Howell KR, Floyd K, Law AJ. PKBγ/AKT3 loss-of-function causes learning and memory deficits and deregulation of AKT/mTORC2 signaling: relevance for schizophrenia. PLoS One. 2017; 12:e0175993. https://doi.org/10.1371/journal.pone.0175993 PMID:28467426

15. Levenga J, Wong H, Milstead RA, Keller BN, LaPlante LE, Hoeffer CA. AKT isoforms have distinct

hippocampal expression and roles in synaptic plasticity. eLife. 2017; 6:6. https://doi.org/10.7554/eLife.30640 PMID:29173281

16. Ripke S, O'Dushlaine C, Chambert K, Moran JL, Kähler AK, Akterin S, Bergen SE, Collins AL, Crowley JJ, Fromer M, Kim Y, Lee SH, Magnusson PK, et al, and Multicenter Genetic Studies of Schizophrenia Consortium, and Psychosis Endophenotypes International Consortium, and Wellcome Trust Case Control Consortium 2. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. Nat Genet. 2013; 45:1150–59. https://doi.org/10.1038/ng.2742 PMID:23974872

17. Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. Nat Neurosci. 2015; 18:199–209. https://doi.org/10.1038/nn.3922 PMID:25599223

18. Sharp SI, Hu Y, Weymer JF, Rizig M, McQuillin A, Hunt SP, Gurling HM. The effect of clozapine on mRNA expression for genes encoding G protein-coupled receptors and the protein components of clathrin-mediated endocytosis. Psychiatr Genet. 2013; 23:153–62. https://doi.org/10.1097/YPG.0b013e32835fe51d PMID:23811784

19. Genis-Mendoza A, Gallegos-Silva I, Tovilla-Zarate CA, López-Narvaez L, González-Castro TB, Hernández-Díaz Y, López-Casamichana M, Nicolini H, Morales-Mulia S. Comparative Analysis of Gene Expression Profiles Involved in Calcium Signaling Pathways Using the NLVH Animal Model of Schizophrenia. J Mol Neurosci. 2018; 64:111–16. https://doi.org/10.1007/s12031-017-1013-y PMID:29214423

20. Guan F, Li L, Qiao C, Chen G, Yan T, Li T, Zhang T, Liu X. Evaluation of genetic susceptibility of common variants in CACNA1D with schizophrenia in Han Chinese. Sci Rep. 2015; 5:12935. https://doi.org/10.1038/srep12935 PMID:26255836

21. Berger SM, Bartsch D. The role of L-type voltage-gated calcium channels Cav1.2 and Cav1.3 in normal and pathological brain function. Cell Tissue Res. 2014; 357:463–76. https://doi.org/10.1007/s00441-014-1936-3 PMID:24996399

22. Gao R, Piguel NH, Melendez-Zaidi AE, Martin-de-Saavedra MD, Yoon S, Forrest MP, Myczek K, Zhang G, Russell TA, Csernansky JG, Surmeier DJ, Penzes P. CNTNAP2 stabilizes interneuron dendritic arbors through CASK. Mol Psychiatry. 2018; 23:1832–50. https://doi.org/10.1038/s41380-018-0027-3 PMID:29610457

23. Poot M. Connecting the CNTNAP2 Networks with Neurodevelopmental Disorders. Mol Syndromol. 2015; 6:7–22. https://doi.org/10.1159/000371594 PMID:25852443

24. Ji W, Li T, Pan Y, Tao H, Ju K, Wen Z, Fu Y, An Z, Zhao Q, Wang T, He L, Feng G, Yi Q, Shi Y. CNTNAP2 is significantly associated with schizophrenia and major depression in the Han Chinese population. Psychiatry Res. 2013; 207:225–28. https://doi.org/10.1016/j.psychres.2012.09.024 PMID:23123147

25. Jia P, Wang L, Fanous AH, Pato CN, Edwards TL, Zhao Z, and International Schizophrenia Consortium. Network-assisted investigation of combined causal signals from genome-wide association studies in schizophrenia. PLOS Comput Biol. 2012; 8:e1002587. https://doi.org/10.1371/journal.pcbi.1002587 PMID:22792057

26. Solismaa A, Kampman O, Lyytikäinen LP, Seppälä N, Viikki M, Mononen N, Lehtimäki T, Leinonen E. Histaminergic gene polymorphisms associated with sedation in clozapine-treated patients. Eur Neuropsychopharmacol. 2017; 27:442–49. https://doi.org/10.1016/j.euroneuro.2017.03.009 PMID:28400155

27. Tiwari AK, Zhang D, Pouget JG, Zai CC, Chowdhury NI, Brandl EJ, Qin L, Freeman N, Lieberman JA, Meltzer HY, Kennedy JL, Müller DJ. Impact of histamine receptors H1 and H3 polymorphisms on antipsychotic-induced weight gain. World J Biol Psychiatry. 2018; 19:S97–S105. https://doi.org/10.1080/15622975.2016.1262061 PMID:27855565

28. Rannals MD, Hamersky GR, Page SC, Campbell MN, Briley A, Gallo RA, Phan BN, Hyde TM, Kleinman JE, Shin JH, Jaffe AE, Weinberger DR, Maher BJ. Psychiatric Risk Gene Transcription Factor 4 Regulates Intrinsic Excitability of Prefrontal Neurons via Repression of SCN10a and KCNQ1. Neuron. 2016; 90:43–55. https://doi.org/10.1016/j.neuron.2016.02.021 PMID:26971948

29. Bruce HA, Kochunov P, Paciga SA, Hyde CL, Chen X, Xie Z, Zhang B, Xi HS, O'Donnell P, Whelan C, Schubert CR, Bellon A, Ament SA, et al. Potassium channel gene associations with joint processing speed and white matter impairments in schizophrenia. Genes Brain Behav. 2017; 16:515–21. https://doi.org/10.1111/gbb.12372 PMID:28188958

30. Chang S, Fang K, Zhang K, Wang J, and Network-Based Analysis of Schizophrenia Genome-Wide Association Data to Detect the Joint Functional Association Signals. Network-Based Analysis of Schizophrenia Genome-Wide Association Data to Detect the Joint Functional

Association Signals. PLoS One. 2015; 10:e0133404. https://doi.org/10.1371/journal.pone.0133404 PMID:26193471

31. Su L, Shen T, Huang G, Long J, Fan J, Ling W, Jiang J. Genetic association of GWAS-supported MAD1L1 gene polymorphism rs12666575 with schizophrenia susceptibility in a Chinese population. Neurosci Lett. 2016; 610:98–103. https://doi.org/10.1016/j.neulet.2015.10.061 PMID:26528791

32. Candemir E, Kollert L, Weißflog L, Geis M, Müller A, Post AM, O'Leary A, Harro J, Reif A, Freudenberg F. Interaction of NOS1AP with the NOS-I PDZ domain: implications for schizophrenia-related alterations in dendritic morphology. Eur Neuropsychopharmacol. 2016; 26:741–55. https://doi.org/10.1016/j.euroneuro.2016.01.008 PMID:26861996

33. Carrel D, Hernandez K, Kwon M, Mau C, Trivedi MP, Brzustowicz LM, Firestein BL. Nitric oxide synthase 1 adaptor protein, a protein implicated in schizophrenia, controls radial migration of cortical neurons. Biol Psychiatry. 2015; 77:969–78. https://doi.org/10.1016/j.biopsych.2014.10.016 PMID:25542305

34. Hadzimichalis NM, Previtera ML, Moreau MP, Li B, Lee GH, Dulencin AM, Matteson PG, Buyske S, Millonig JH, Brzustowicz LM, Firestein BL. NOS1AP protein levels are altered in BA46 and cerebellum of patients with schizophrenia. Schizophr Res. 2010; 124:248–50. https://doi.org/10.1016/j.schres.2010.05.009 PMID:20605702

35. Xu B, Wratten N, Charych EI, Buyske S, Firestein BL, Brzustowicz LM. Increased expression in dorsolateral prefrontal cortex of CAPON in schizophrenia and bipolar disorder. PLoS Med. 2005; 2:e263. https://doi.org/10.1371/journal.pmed.0020263 PMID:16146415

36. Brzustowicz LM, Simone J, Mohseni P, Hayter JE, Hodgkinson KA, Chow EW, Bassett AS. Linkage disequilibrium mapping of schizophrenia susceptibility to the CAPON region of chromosome 1q22. Am J Hum Genet. 2004; 74:1057–63. https://doi.org/10.1086/420774 PMID:15065015

37. Rosa A, Fañanás L, Cuesta MJ, Peralta V, Sham P. 1q21-q22 locus is associated with susceptibility to the reality-distortion syndrome of schizophrenia spectrum disorders. Am J Med Genet. 2002; 114:516–18. https://doi.org/10.1002/ajmg.10526 PMID:12116186

38. Kremeyer B, García J, Kymäläinen H, Wratten N, Restrepo G, Palacio C, Miranda AL, López C, Restrepo M, Bedoya G, Brzustowicz LM, Ospina-Duque J, Arbeláez MP, Ruiz-Linares A. Evidence for a role of the NOS1AP (CAPON) gene in schizophrenia and its clinical dimensions: an association study in a South American population isolate. Hum Hered. 2009; 67:163–73. https://doi.org/10.1159/000181154 PMID:19077434

39. Wang HN, Liu GH, Zhang RG, Xue F, Wu D, Chen YC, Peng Y, Peng ZW, Tan QR. Quetiapine Ameliorates Schizophrenia-Like Behaviors and Protects Myelin Integrity in Cuprizone Intoxicated Mice: The Involvement of Notch Signaling Pathway. Int J Neuropsychopharmacol. 2015; 19:pyv088. https://doi.org/10.1093/ijnp/pyv088 PMID:26232790

40. Kim SK, Lee JY, Park HJ, Kim JW, Chung JH. Association study between polymorphisms of the PARD3 gene and schizophrenia. Exp Ther Med. 2012; 3:881–85. https://doi.org/10.3892/etm.2012.496 PMID:22969987

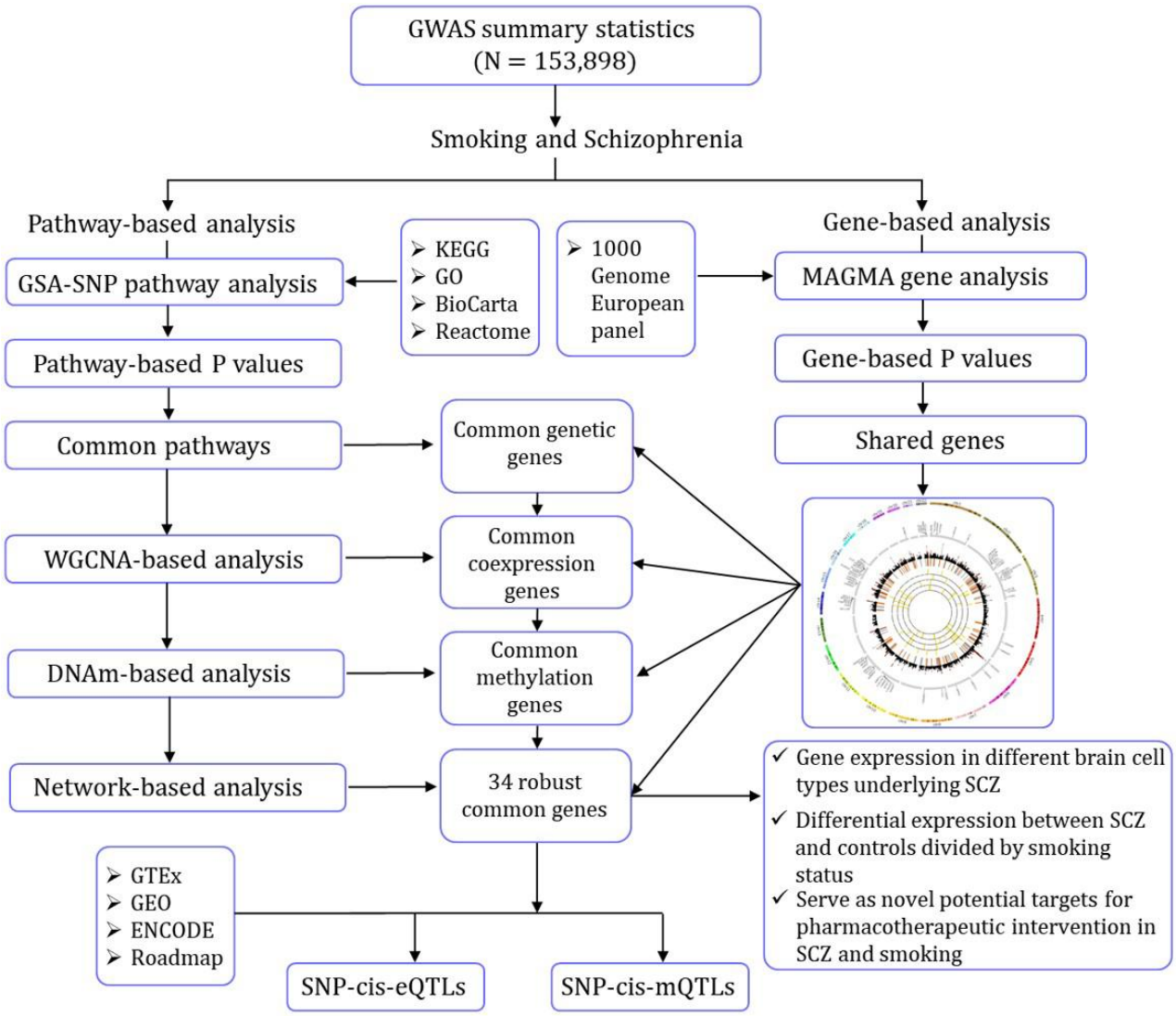
41. Guipponi M, Santoni FA, Setola V, Gehrig C, Rotharmel M, Cuenca M, Guillin O, Dikeos D, Georgantopoulos G, Papadimitriou G, Curtis L, Méary A, Schürhoff F, et al. Exome sequencing in 53 sporadic cases of schizophrenia identifies 18 putative candidate genes. PLoS One. 2014; 9:e112745. https://doi.org/10.1371/journal.pone.0112745 PMID:25420024

42. Goldsmith DR, Haroon E, Miller AH, Strauss GP, Buckley PF, Miller BJ. TNF-α and IL-6 are associated with the deficit syndrome and negative symptoms in patients with chronic schizophrenia. Schizophr Res. 2018; 199:281–84. https://doi.org/10.1016/j.schres.2018.02.048 PMID:29499967

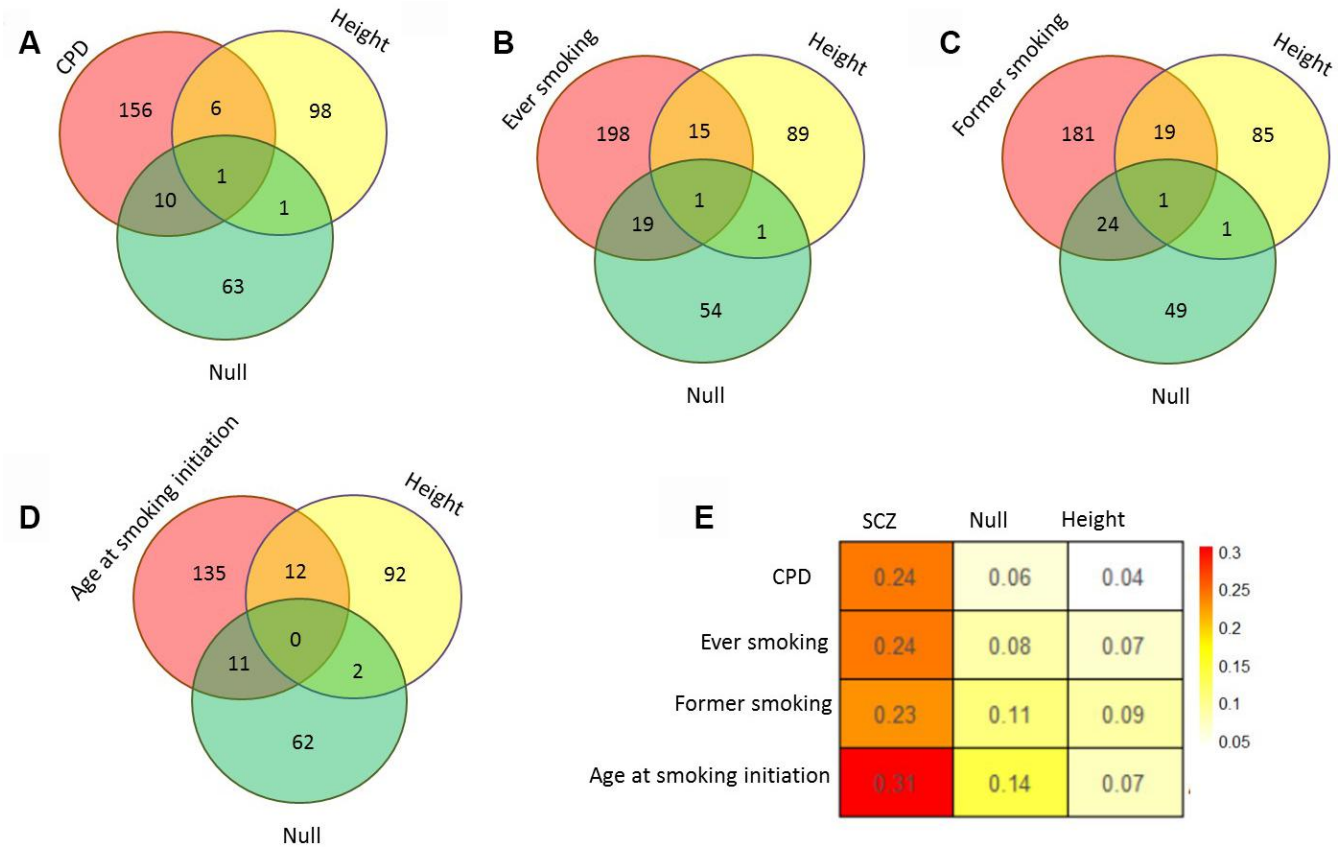
43. Mostaid MS, Pantelis C, Everall IP, Bousman CA. Decreased peripheral TNF alpha (TNF-α) mRNA expression in patients with treatment-resistant schizophrenia. Schizophr Res. 2018; 202:387–88. https://doi.org/10.1016/j.schres.2018.04.032 PMID:29706448

44. Chen SY, Huang PH, Cheng HJ. Disrupted-in-Schizophrenia 1-mediated axon guidance involves TRIO-RAC-PAK small GTPase pathway signaling. Proc Natl Acad Sci USA. 2011; 108:5861–66. https://doi.org/10.1073/pnas.1018128108 PMID:21422296

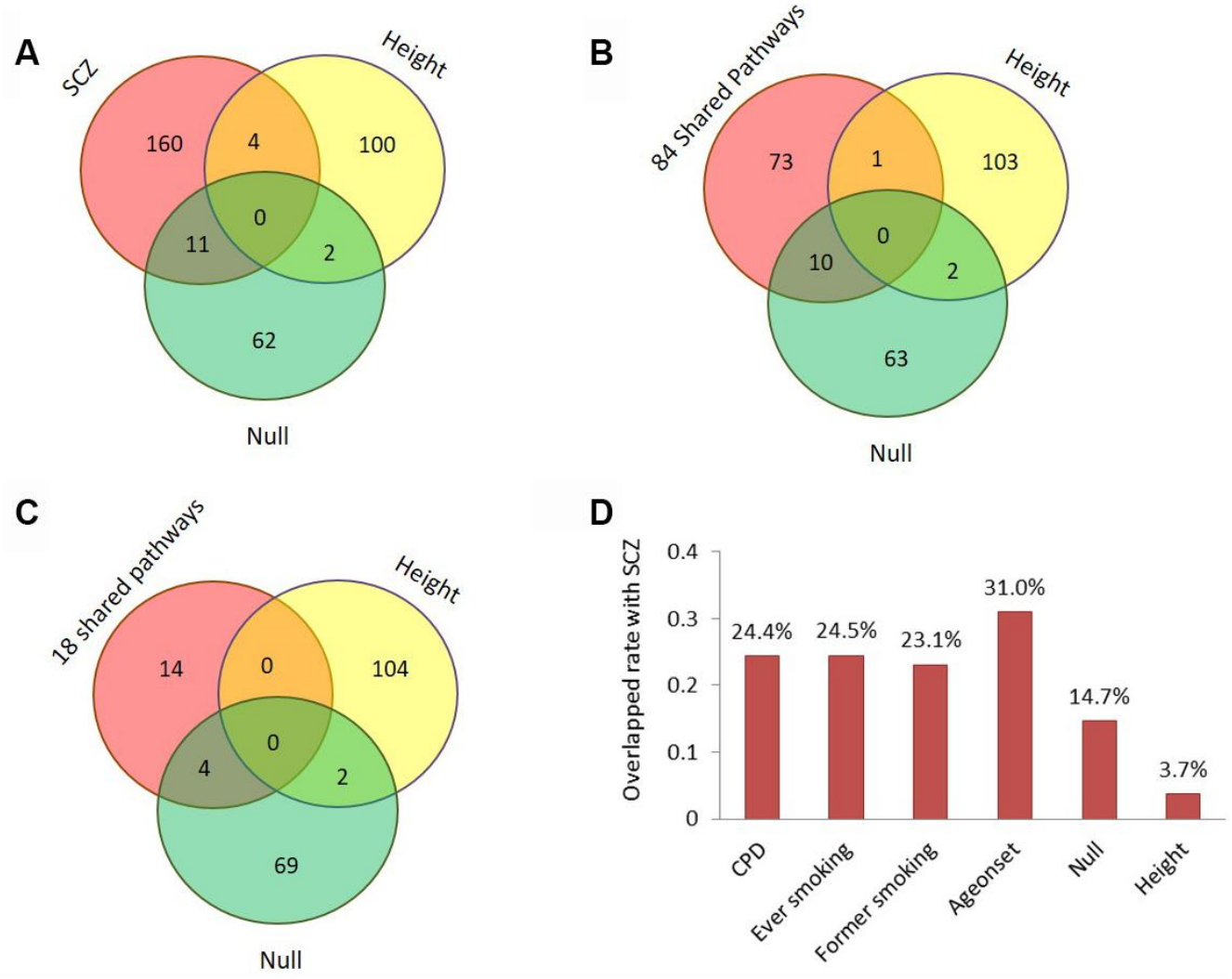
Supplementary Figures



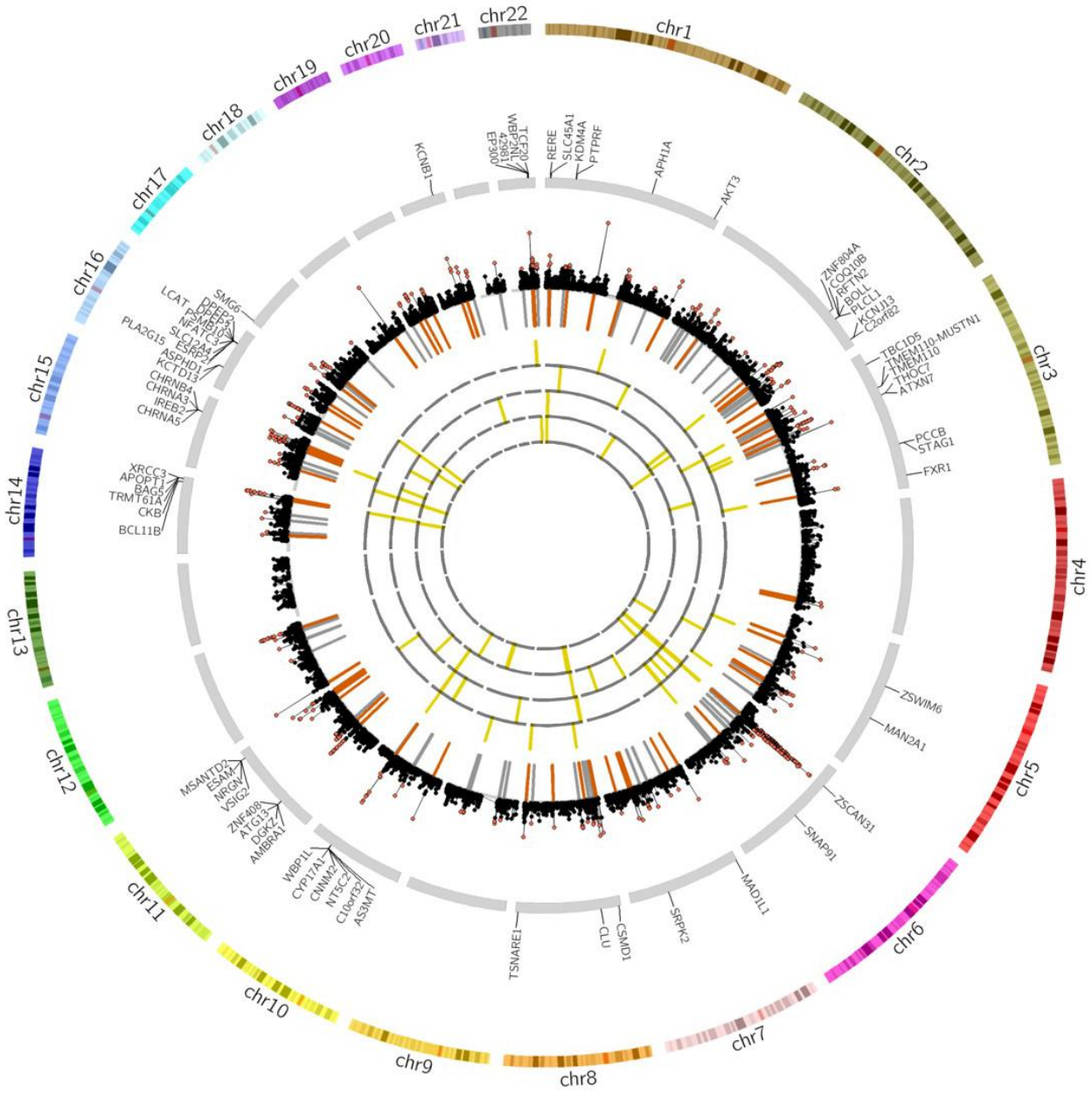
Supplementary Figure 1. Overview of statistical approaches for integrative pathway- and gene-based analysis of GWAS summary statistics. Using GSA-SNP and MAGMA software, pathway- and gene-based analyses were carried out. A number of novel genes were identified from MAGMA analysis. There was a total of 208 shared genes between SCZ and smoking behaviors. According to pathway-based enrichment analysis, 18 common pathways were identified. We performed a simulation analysis and found these 208 genes to be significantly enriched in common pathways identified from GSA-SNP pathway analysis. Further, to explore the underlying biological mechanism of the comorbidity of SCZ and smoking, we then integrated multi-dimensional omics datasets, including human brain and blood transcriptome data, on smoking and SCZ with well-established tools for interrogating such datasets. Finally, we identified 34 candidate genes conferring risk of the comorbidity of SCZ and smoking and revealed the regulatory mechanism of SNP-methylation expression of the comorbidity.



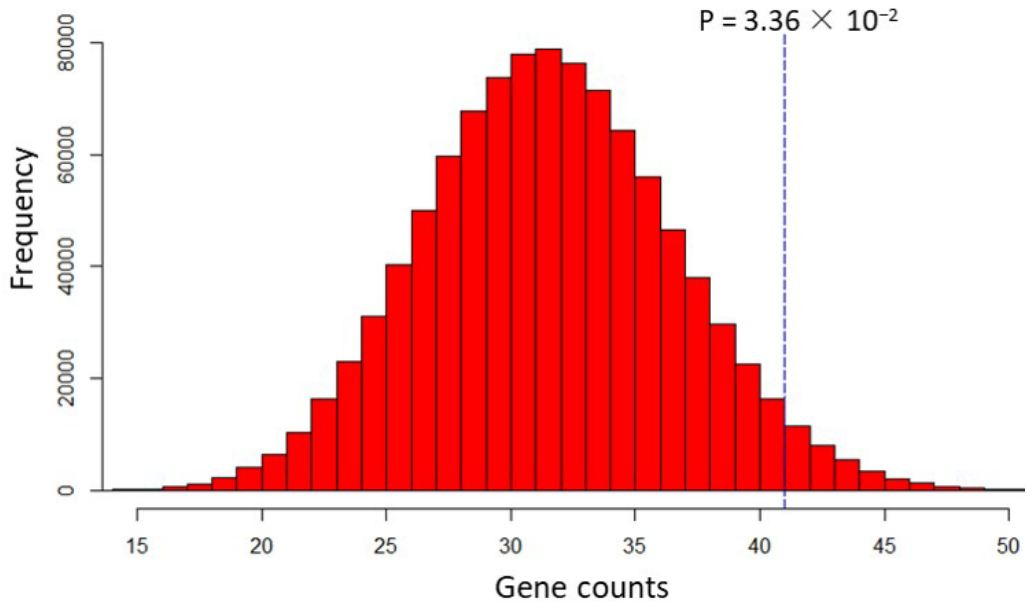
Supplementary Figure 2. Overlapped pathways among smoking behaviors, null, and height. (**A**) Overlapped pathways among CPD, null, and height; (**B**) Overlapped pathways among ever-smoking, null, and height; (**C**) Overlapped pathways among former smoking, null, and height; (**D**) Overlapped pathways among age at smoking initiation, null, and height; (**E**) Heatmap of the proportion of significantly enriched pathways in CPD, ever smoking, former smoking, and age at smoking initiation with SCZ, null, and height. Heatmap plots were generated using the *pheatmap* package in R.



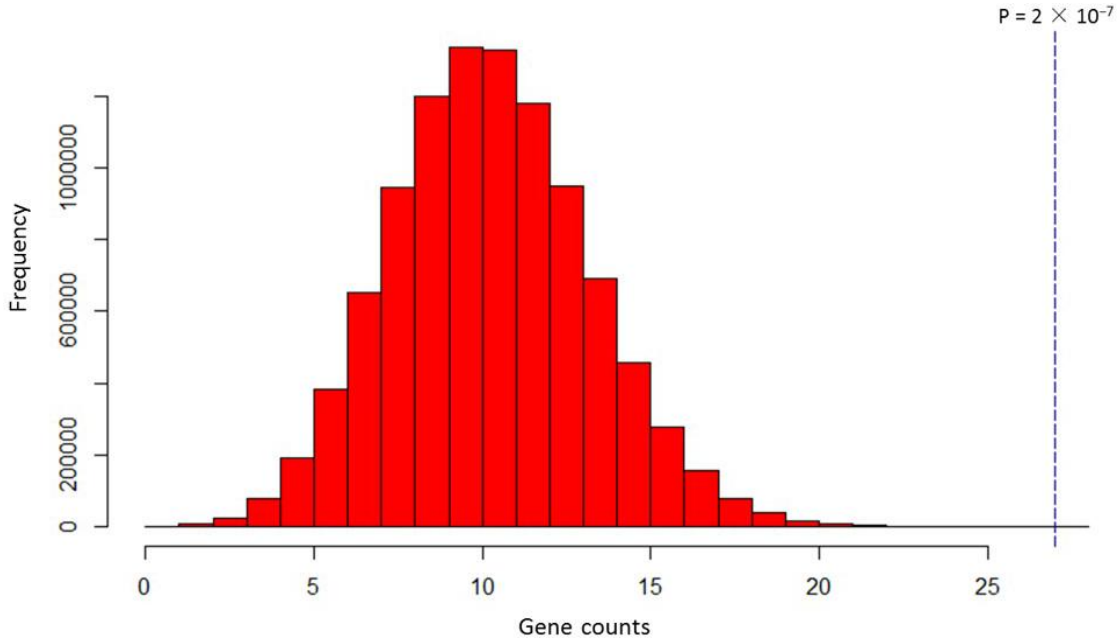
Supplementary Figure 3. Overlapped pathways among SCZ and smoking behaviors, null, and height. (**A**) Venn diagram of significantly enriched pathways for SCZ with the pathways for null and height. (**B**) Venn diagram of significantly enriched pathways among 84 shared pathways between SCZ and at least one of four smoking phenotypes with the pathways for null and height. (**C**) Venn diagram of significantly enriched pathways among 18 pathways shared by SCZ and all four smoking phenotypes with the pathways for null and height. (**D**) The proportion of significantly enriched pathways in CPD, ever smoking, former smoking, age at smoking initiation, null, and height with SCZ.



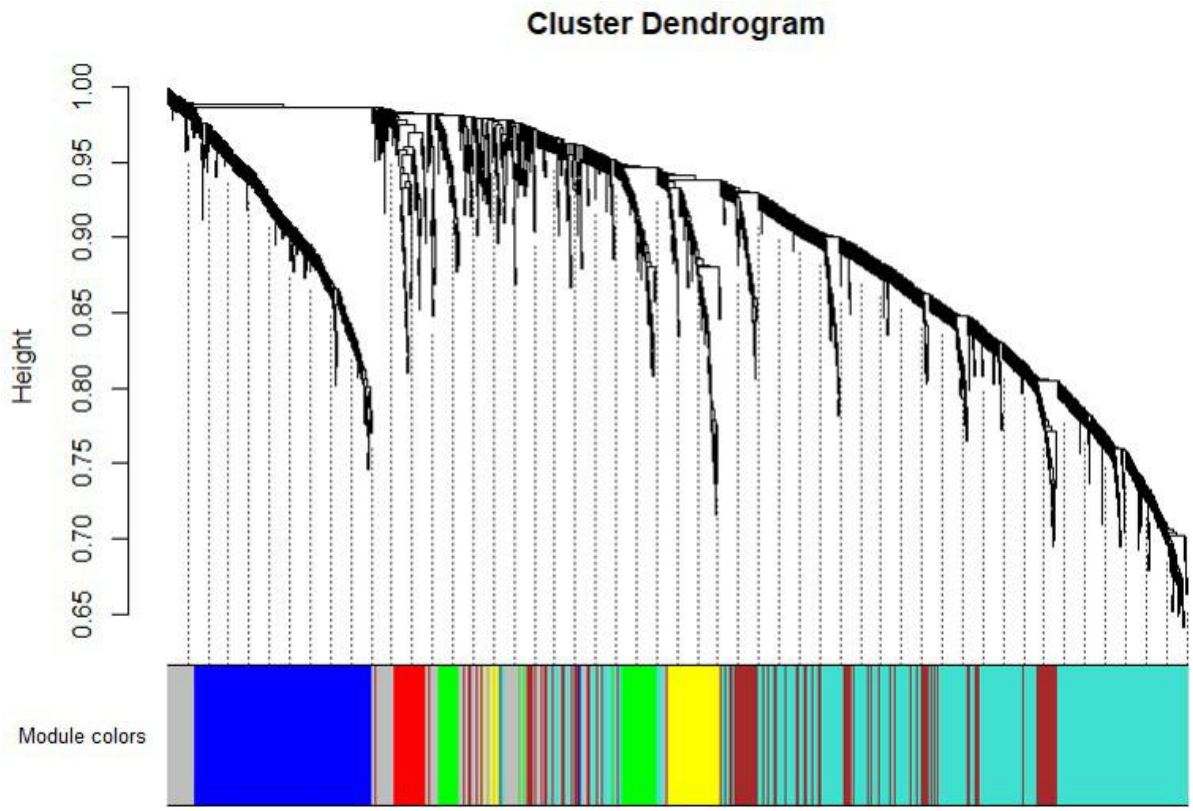
Supplementary Figure 4. Circus plot of results of gene-based analysis of SCZ and smoking phenotypes. The outer ring shows the 22 autosomal human chromosomes. Gene symbols marked in the second ring are the 70 of 208 common genes located in the previously reported 108 loci for SCZ. The third ring demonstrates the results of gene-based analysis of SCZ and 590 genes significantly associated with it ($P < 2.73 \times 10^{-6}$), which are marked with orange points. The orange bars represent 236 genes located in 108 previously identified SCZ-related loci. The grey bars mark novel genes associated with SCZ in the current study. The fourth through the seventh rings mark CPD, ever smoking, former smoking, and smoking initiation, respectively. The yellow bars indicate the distribution of the 208 genes shared by SCZ and smoking.



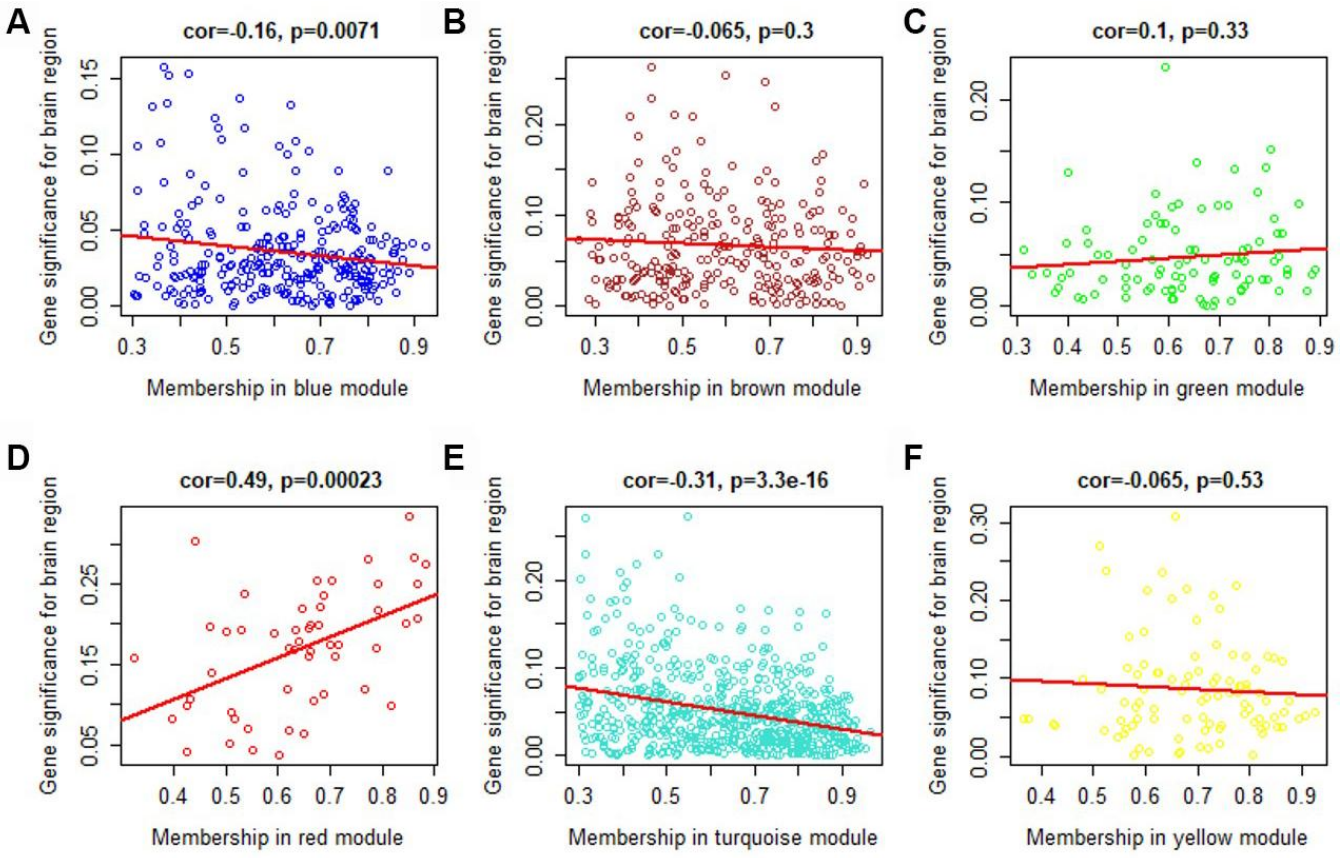
Supplementary Figure 5. Computer permutation analysis of 208 genes associated with SCZ and smoking behaviors for genes in 84 shared pathways.



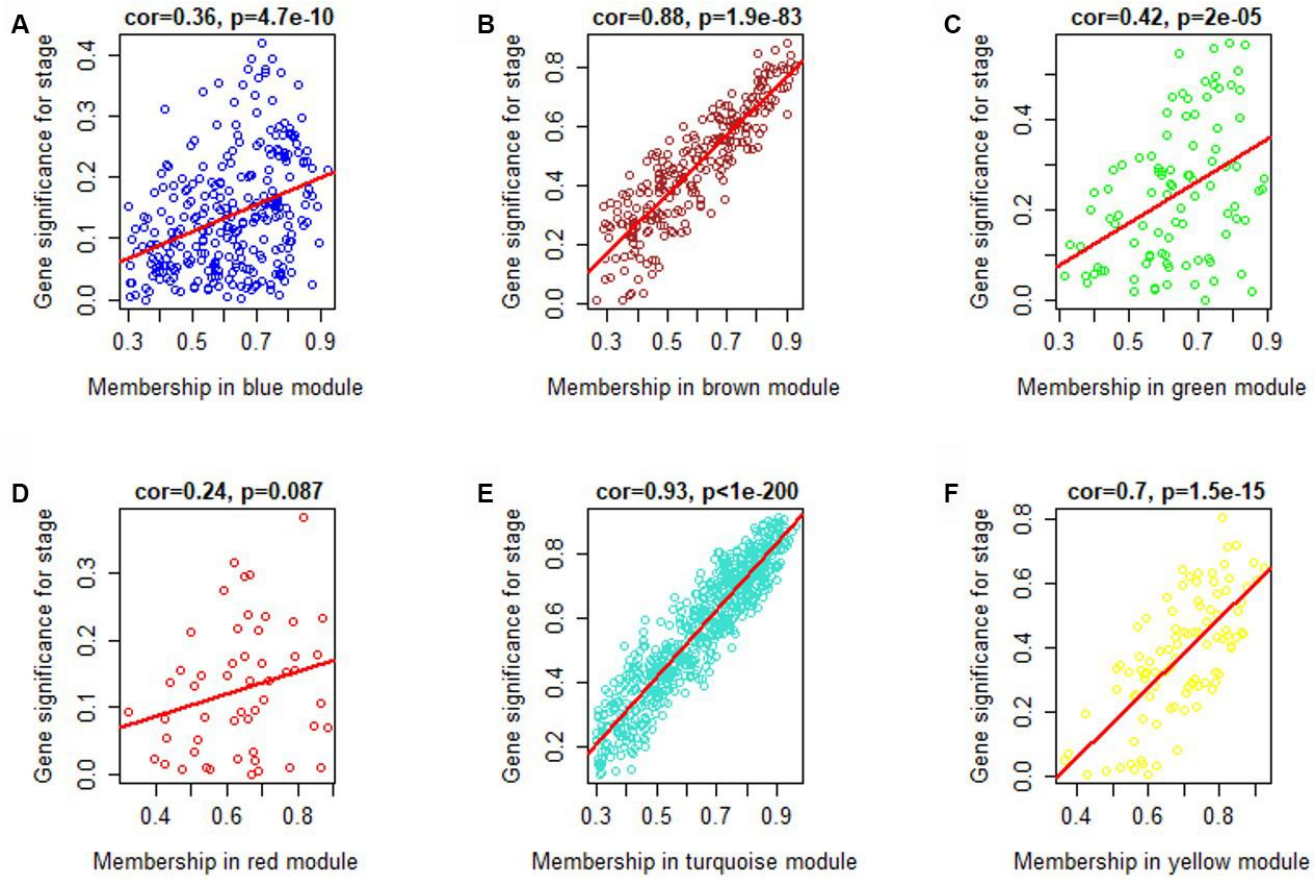
Supplementary Figure 6. Computer permutation analysis of 70 genes associated with SCZ and smoking behaviors for genes in 84 shared pathways.



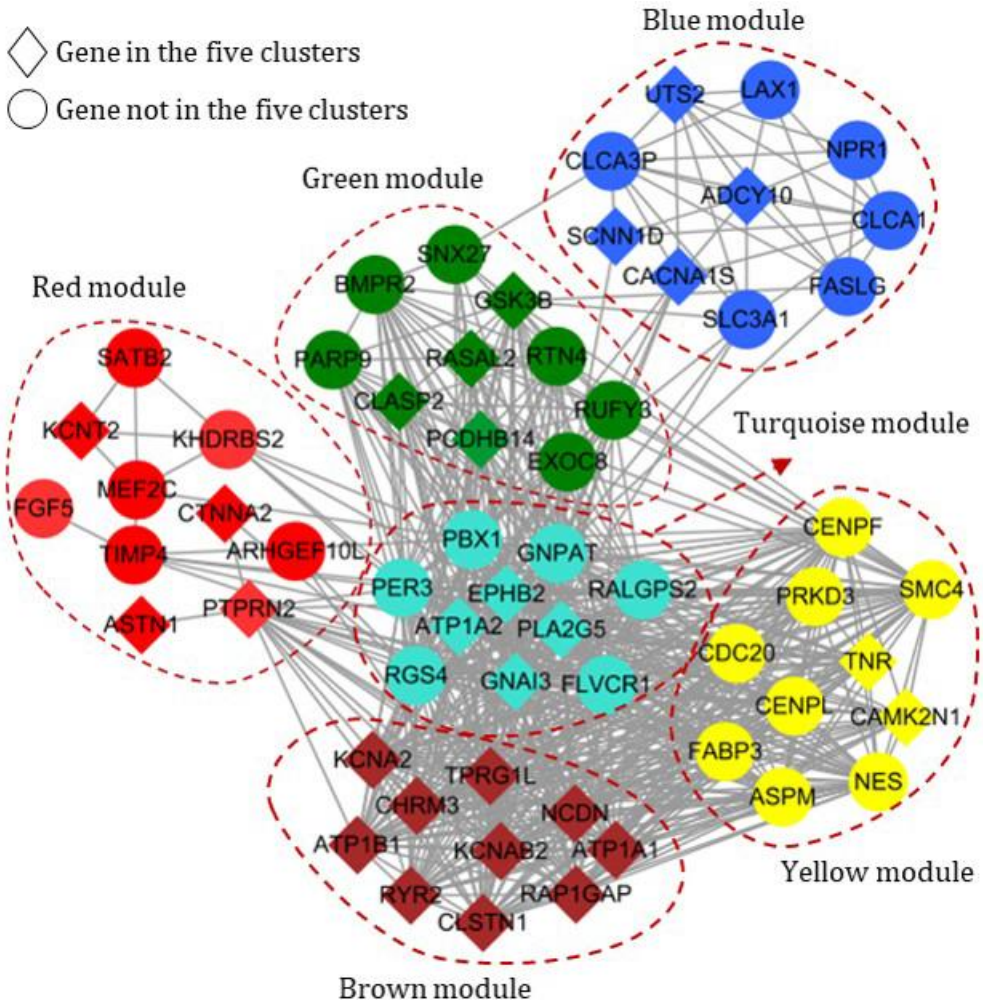
Supplementary Figure 7. WGCNA module-based analysis for genes in all shared pathways with *q* value < 0.1. The hierarchical clustering dendrogram is for all 1,644 expressed and highly connected genes. Each line represents an individual gene. Genes were clustered on the basis of a dissimilarity measure (1-TOM). The branches are related to modules of highly interconnected gene groups. Below the dendrogram of each gene group is a color to indicate the six module colors (i.e., blue, brown, green, red, turquoise, and yellow). Genes within grey boxes were not assigned to a module.



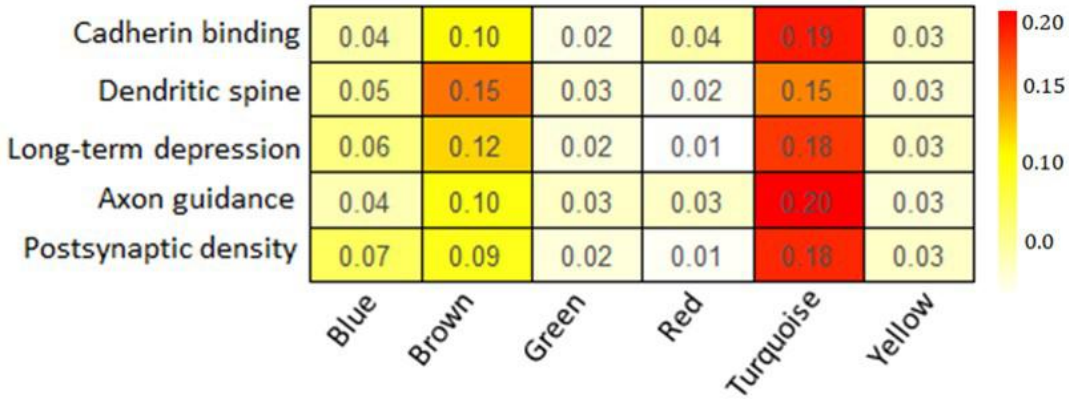
Supplementary Figure 8. Scatterplot of gene significance for brain regions (y axis) vs. module membership (x axis) in different modules. (**A**) For blue module; (**B**) for brown module; (**C**) for green module; (**D**) for red module; (**E**) for turquoise module; (**F**) for yellow module.



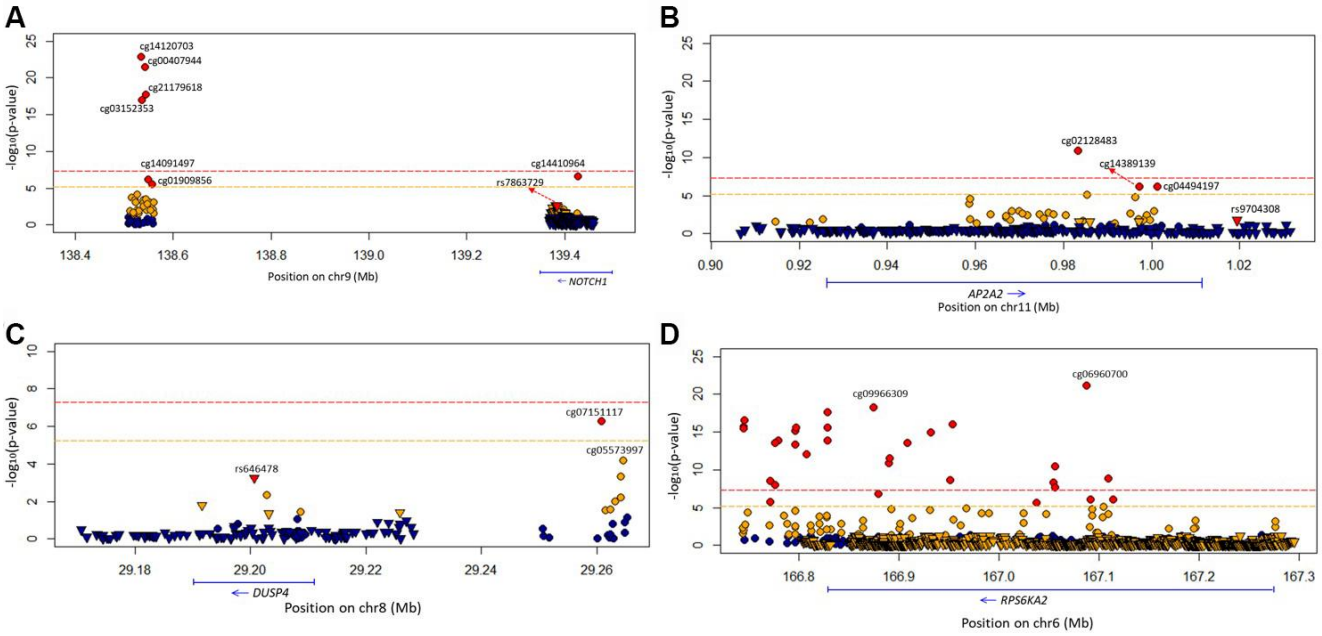
Supplementary Figure 9. Scatterplot of gene significance for developmental time points (stages; y axis) vs. module membership (x axis) in different modules. (**A**) For blue module; (**B**) for brown module; (**C**) for green module; (**D**) for red module; (**E**) for turquoise module; (**F**) for yellow module. For the most significant modules (i.e., blue, brown, green, turquoise, and yellow), genes with high module membership often also show high gene significance. Gene significance and module membership have a highly significant correlation, indicating that hub genes of these modules also predispose to high correlation with weight.



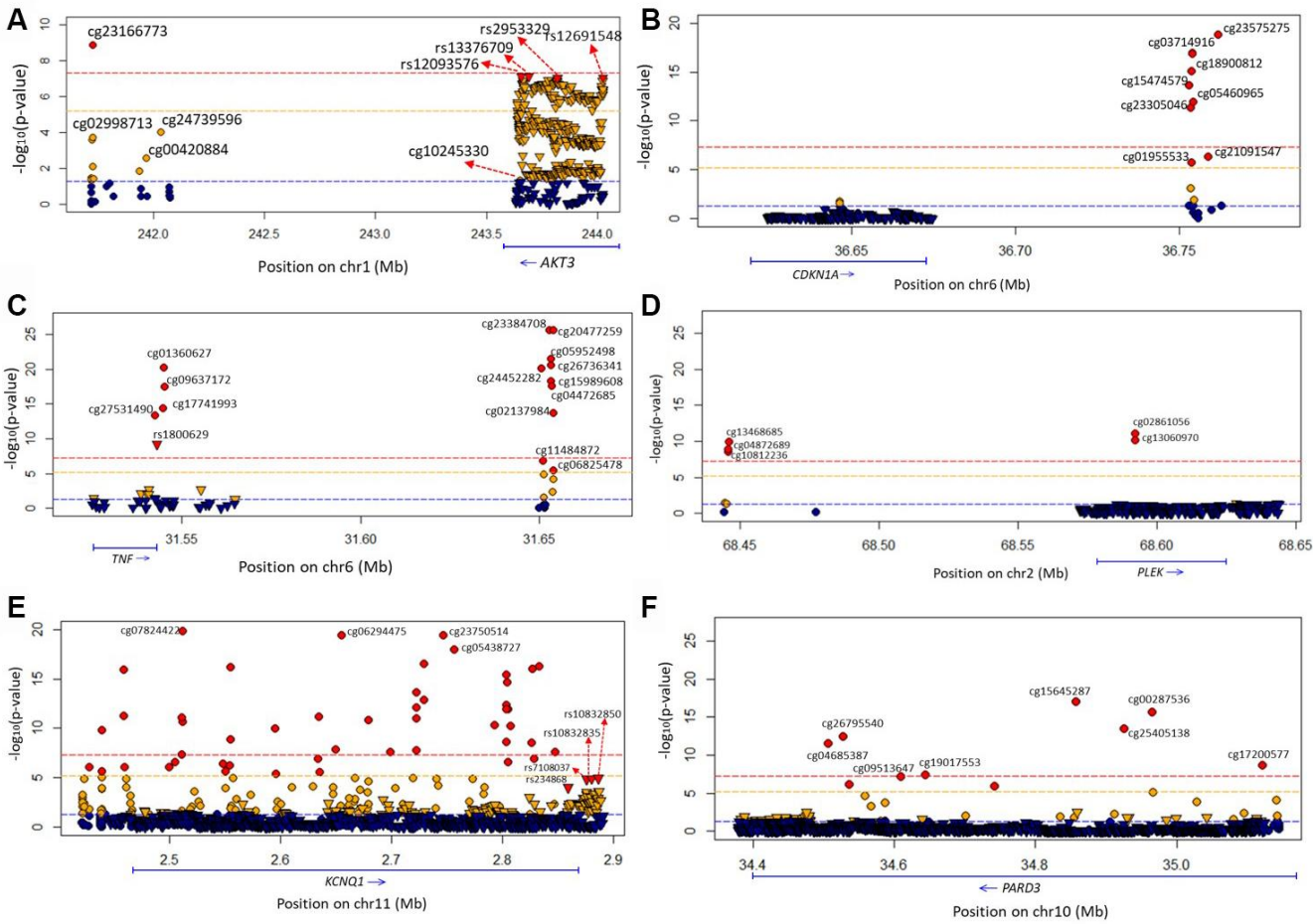
Supplementary Figure 10. Network plot of 10 hub genes from each module showing clustering across different brain regions and developmental time points. The nodes (genes) are colored by module, while the edges reflect positive correlations across brain regions and developmental epochs. The rectangular nodes represent genes clustered in the 18 shared pathways, and the circular nodes mark genes not so clustered.



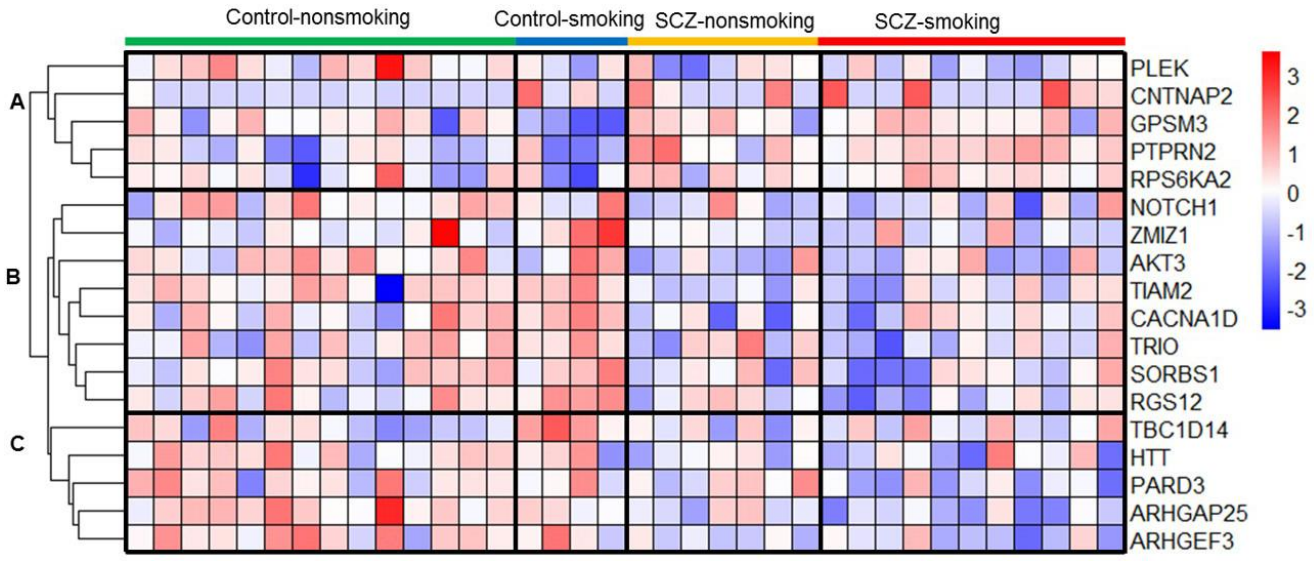
Supplementary Figure 11. Enrichment of the five clusters of 18 shared pathways in six co-expression modules (red = cluster with the highest proportion in a module). More than 15% of the genes in the five clusters among 18 shared pathways were in the turquoise module.



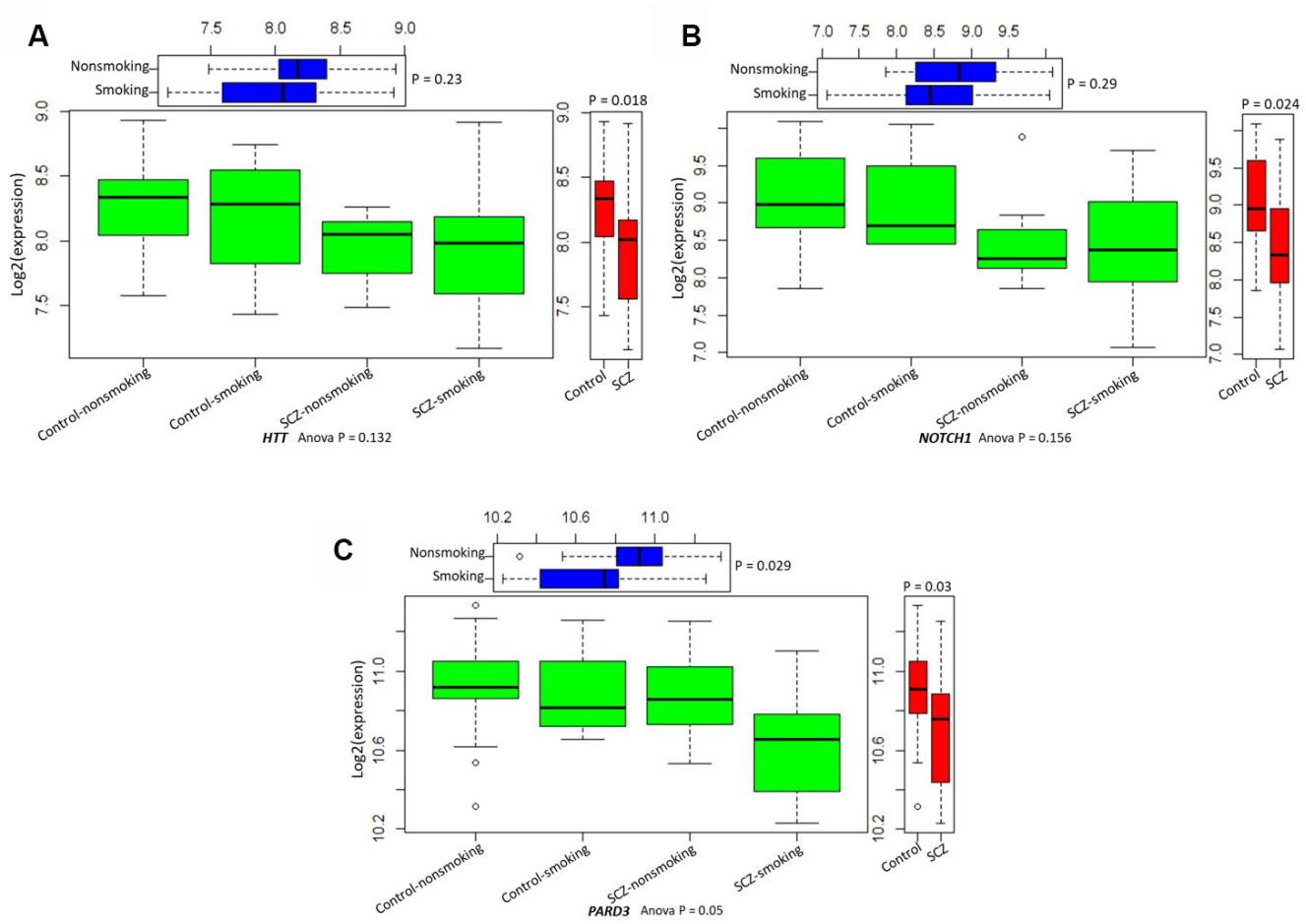
Supplementary Figure 12. Regional plot of association between genetic and epigenetic variation at the four genomic loci and SCZ. (**A**) For *NOTCH1*; (**B**) for *AP2A2*; (**C**) for *DUSP4*; (**D**) for *RPS6KA2*. Circular symbols indicate the association of CpG loci with SCZ (red represents these loci significantly associated with SCZ with $P \leq 6.07 \times 10^{-6}$; orange represents these loci with $6.07 \times 10^{-6} < P \leq 0.05$; blue represents these loci with $P > 0.05$). Triangular symbols indicate the association of SNPs with SCZ (red represents the top-ranked SNPs associated with SCZ; orange represents these SNPs associated with SCZ with $P \leq 0.05$; blue represents these SNPs with $P > 0.05$).



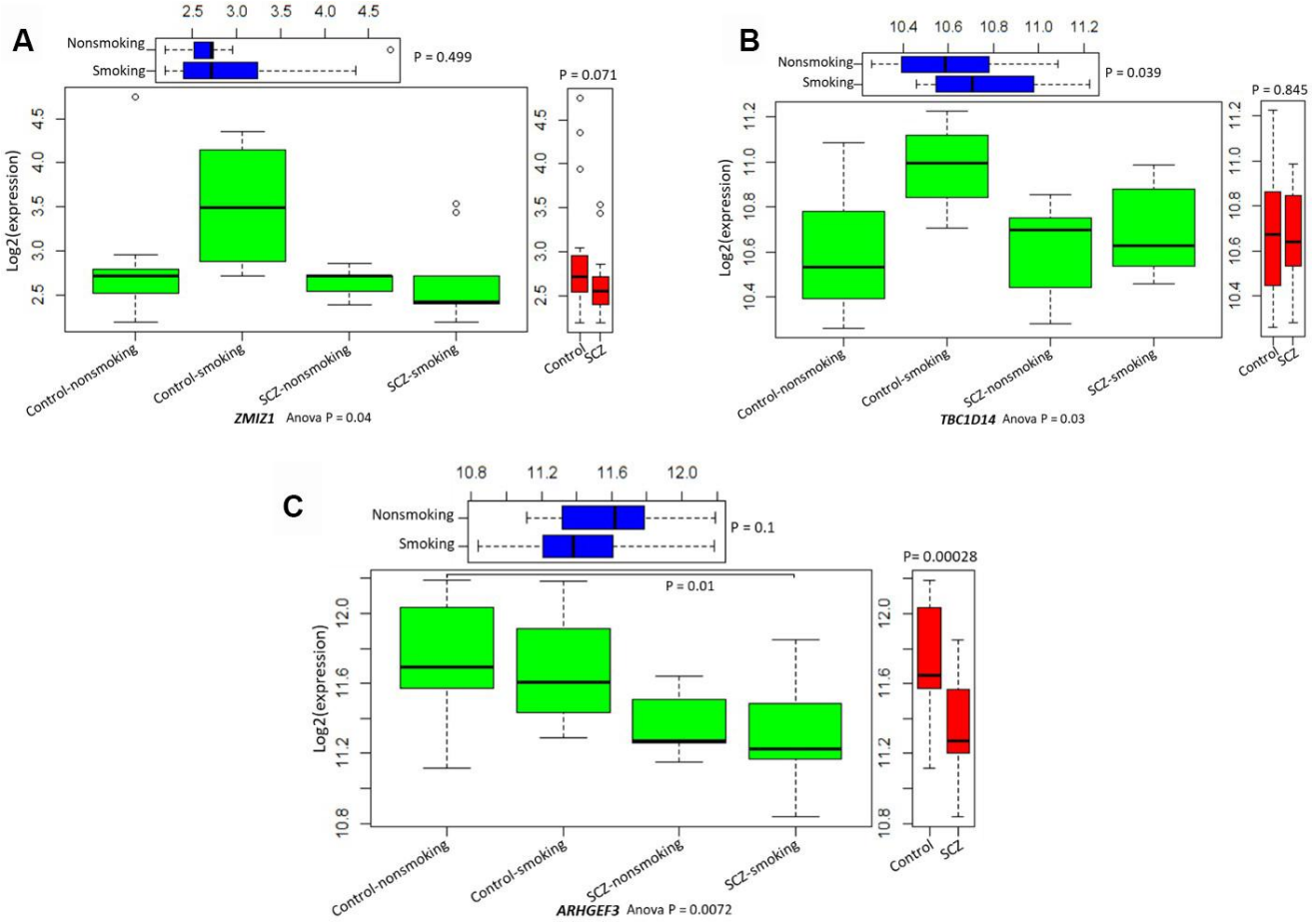
Supplementary Figure 13. Regional plot of association between genetic and epigenetic variation at the six genomic loci and SCZ. (**A**) For *AKT3*; (**B**) for *CDKN1A*; (**C**) for *TNF*; (**D**) for *PLEK*; (**E**) for *KCNQ1*; (**F**) for *PARD3*. Circular symbols indicate the association of CpG loci with SCZ (red represents these loci significantly associated with SCZ with $P \leq 6.07 \times 10^{-6}$; orange represents these loci with $6.07 \times 10^{-6} < P \leq 0.05$; blue represents these loci with $P > 0.05$). Triangular symbols indicate the association of SNPs with SCZ (red represents the top-ranked SNPs associated with SCZ; orange represents these SNPs associated with SCZ with $P \leq 0.05$; blue represents these SNPs with $P > 0.05$).



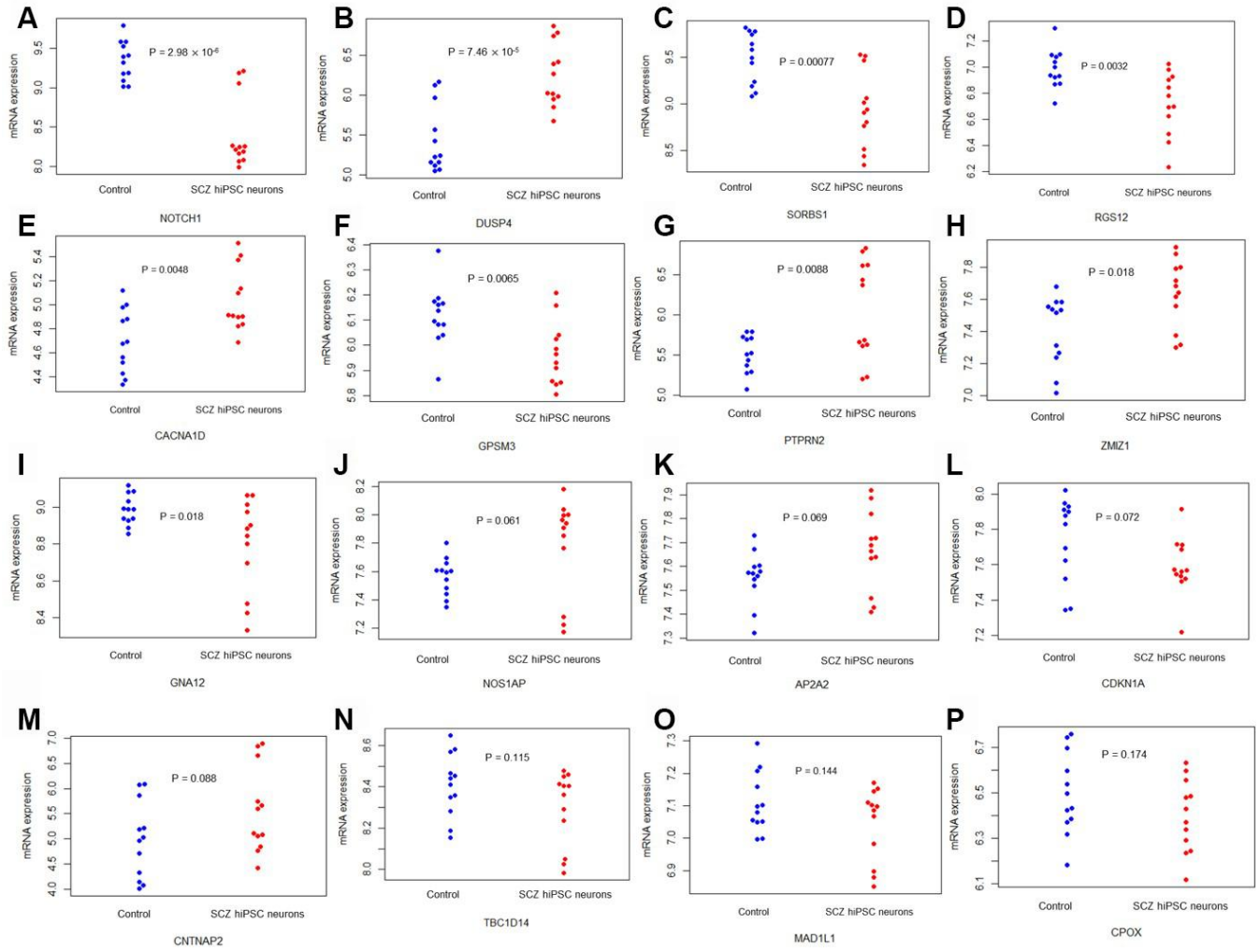
Supplementary Figure 14. Heatmap of significant genes from ANOVA by comparing the difference between SCZ patients and controls divided by smoking status. Colored rectangles represent transcript abundance indicated above the gene labeled on the right. The intensity of the color is proportional to the standardized values between -3 and 3 for each gene, as indicated on the bar on the right of the heat map plot. Clustering was performed using Euclidean distance according to the scale on the left. Major gene transcription patterns are arbitrarily described as A, B, and C. Heatmap plots were generated using the *pheatmap* package in R.



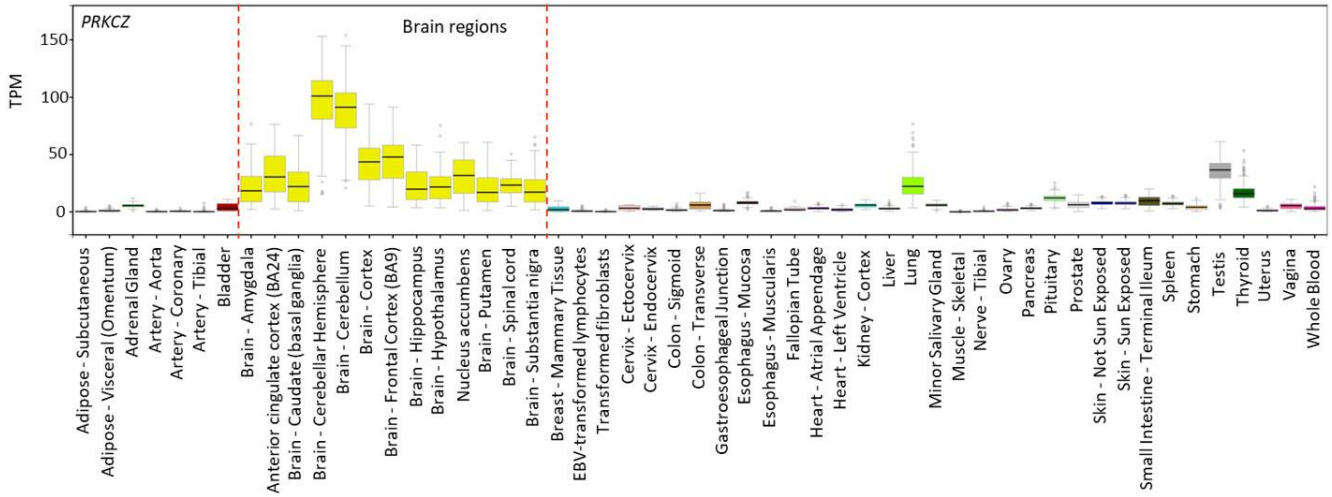
Supplementary Figure 15. Boxplots of 3 genes' expression patterns in SCZ patients and controls divided by smoking status. (**A**) For *HTT*; (**C**) for *NOTCH1*; (**D**) for *PARD3*.



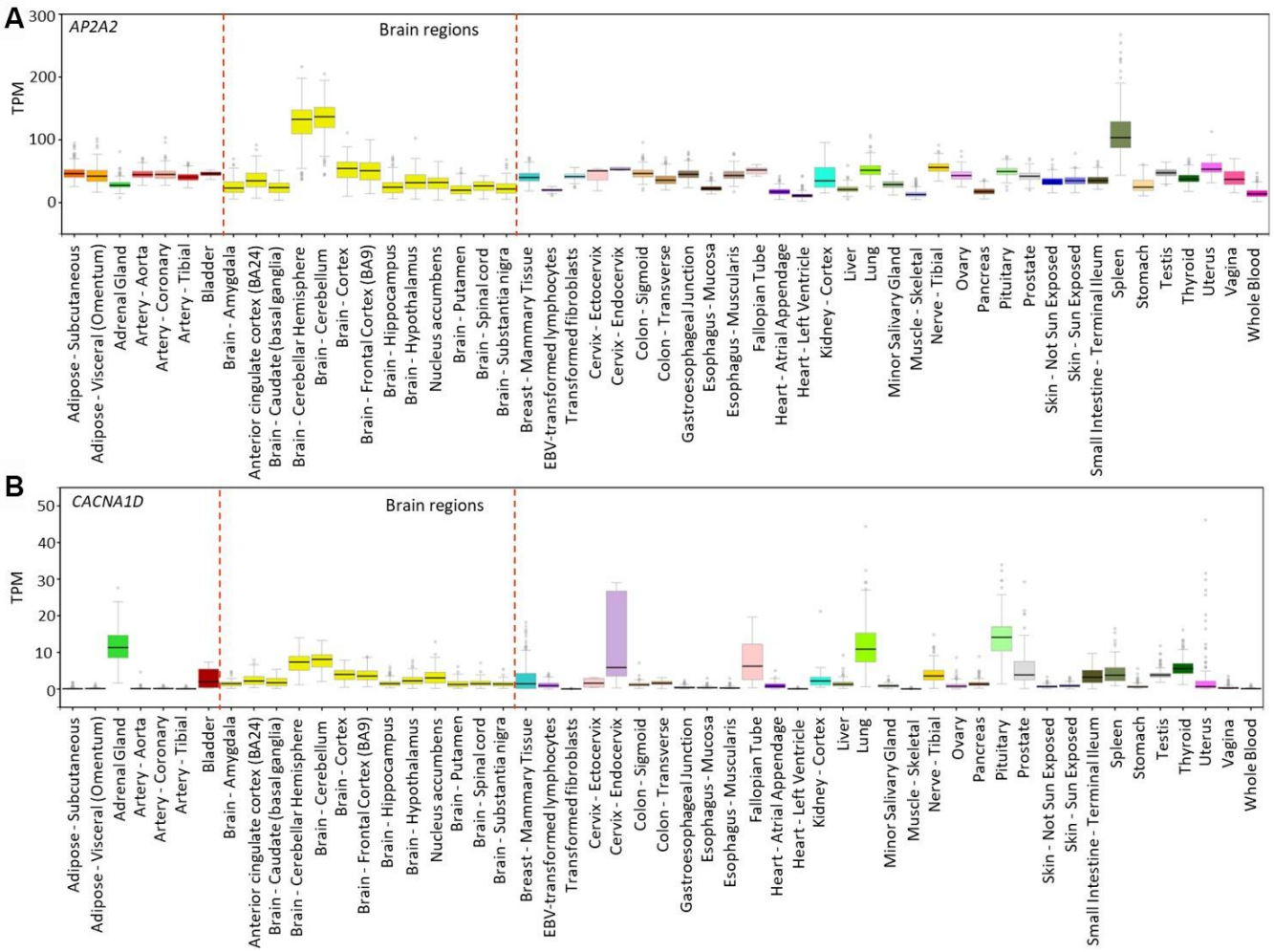
Supplementary Figure 16. Boxplots of 3 genes' expression patterns in SCZ patients and controls divided by smoking status. (**A**) For *ZMIZ1*; (**B**) for *TBC1D14*; (**C**) for *ARHGEF3*.



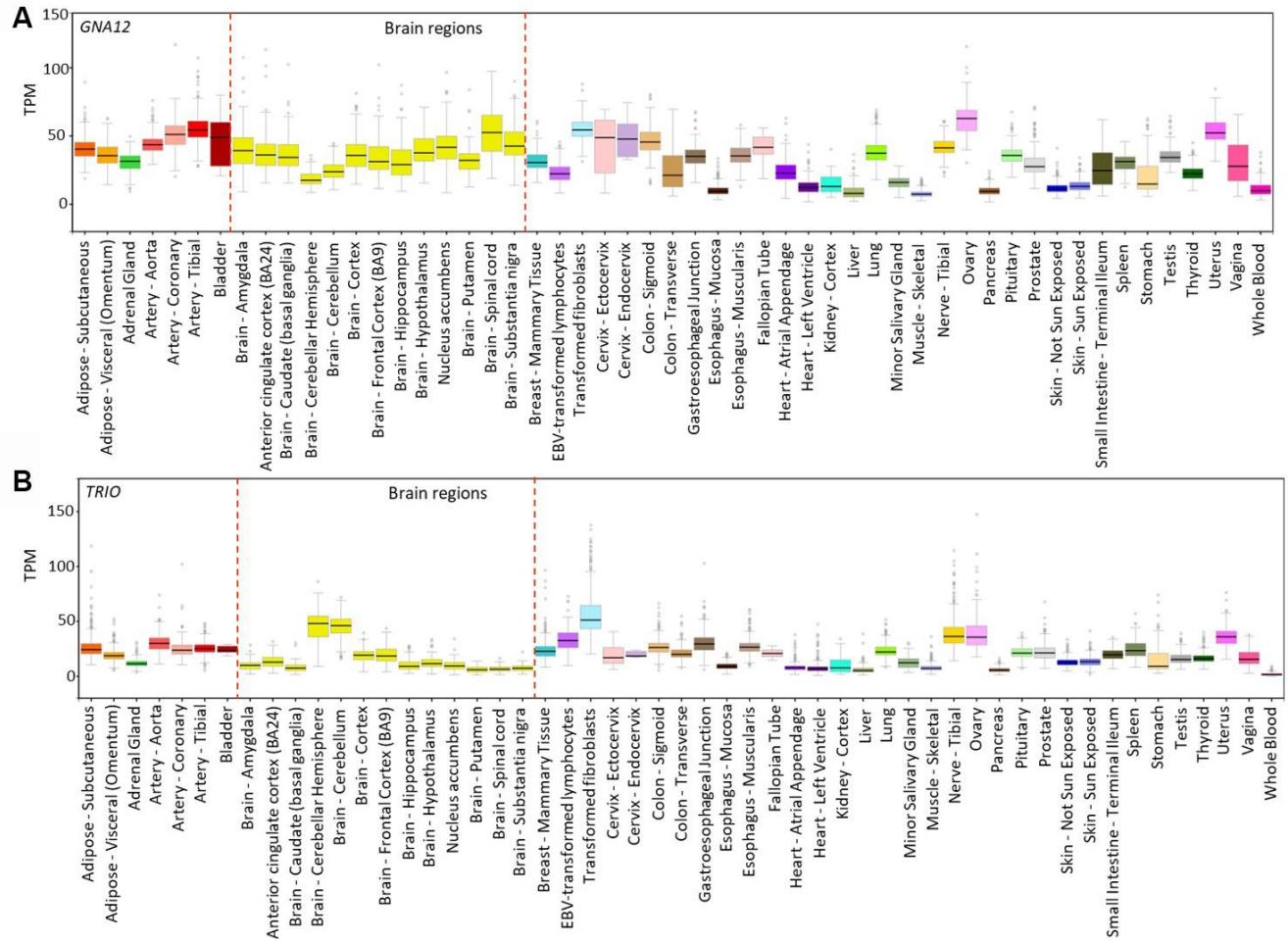
Supplementary Figure 17. Plots summarizing significant gene expression alterations in schizophrenic hiPSC-derived neurons. Student's *t*-test was used for comparing the difference between SCZ hiPSC neurons and controls. Gene expression data from the GEO dataset (Accession No. GSE25673). (**A**) For *NOTCH1*; (**B**) for *DUSP4*; (**C**) for *SORBS1*; (**D**) for *RGS12*; (**E**) for *CACNA1D*; (**F**) for *GPSM3*; (**G**) for *PTPRN2*; (**H**) for *ZMIZ1*; (**I**) for *GNA12*; (**J**) for *NOS1AP*; (**K**) for *AP2A2*; (**L**) for *CDKN1A*; (**M**) for *CNTNAP2*; (**N**) for *TBC1D14*; (**O**) for *MAD1L1*; (**P**) for *CPOX*. These plots were generated using the *beeswarm* package in R.



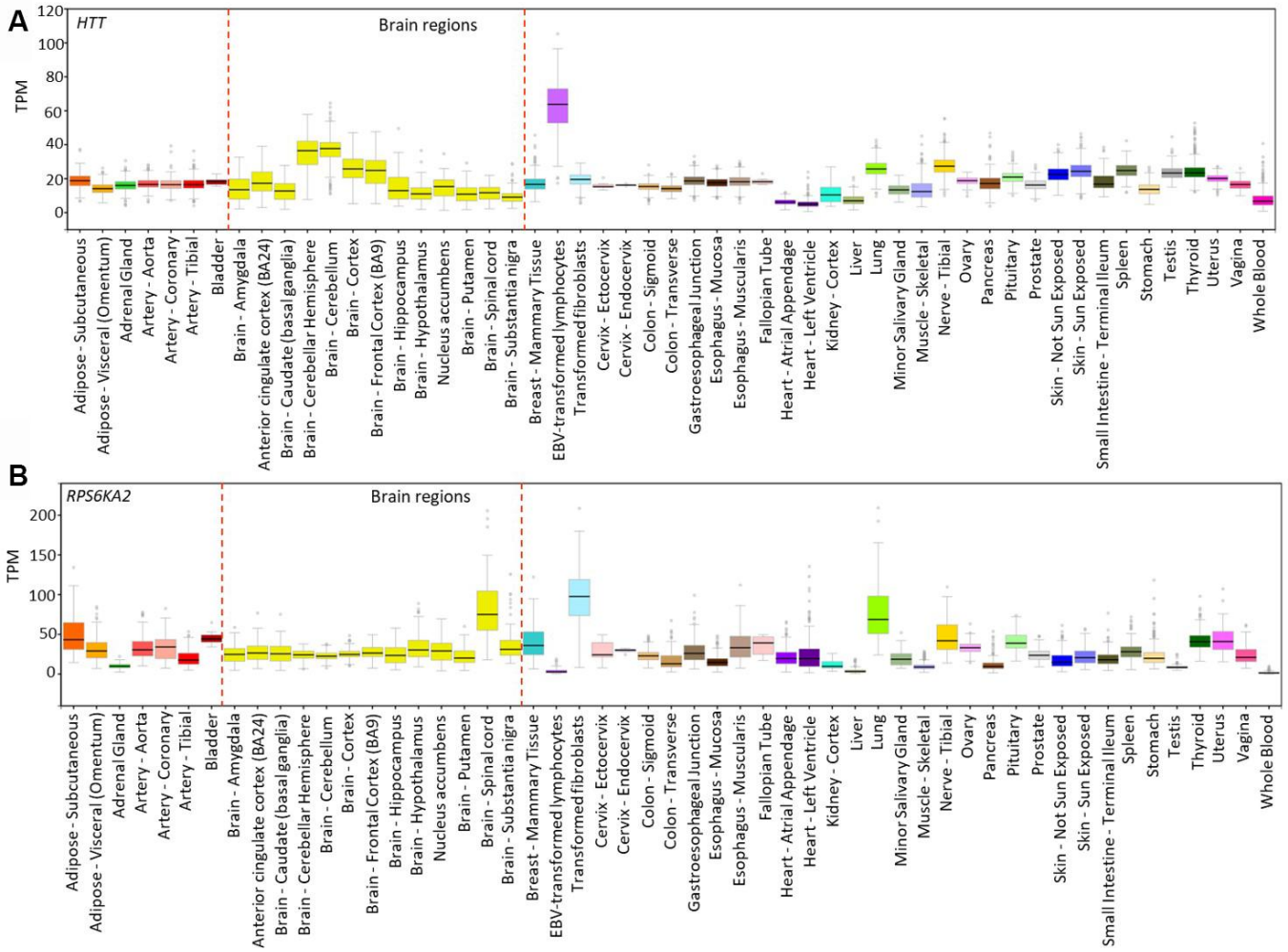
Supplementary Figure 18. An example of *PRKCZ* expression in 53 human tissues, which were obtained from GTEx analysis release V7 (dbGaP Accession No. phs000424.v7.p2). Expression values are shown in transcripts per million (TPM), calculated from a gene model with isoforms collapsed into a single gene. Box plots describe median and 25th and 75th percentiles; points are displayed as outliers if they are above or below 1.5 times the interquartile range.



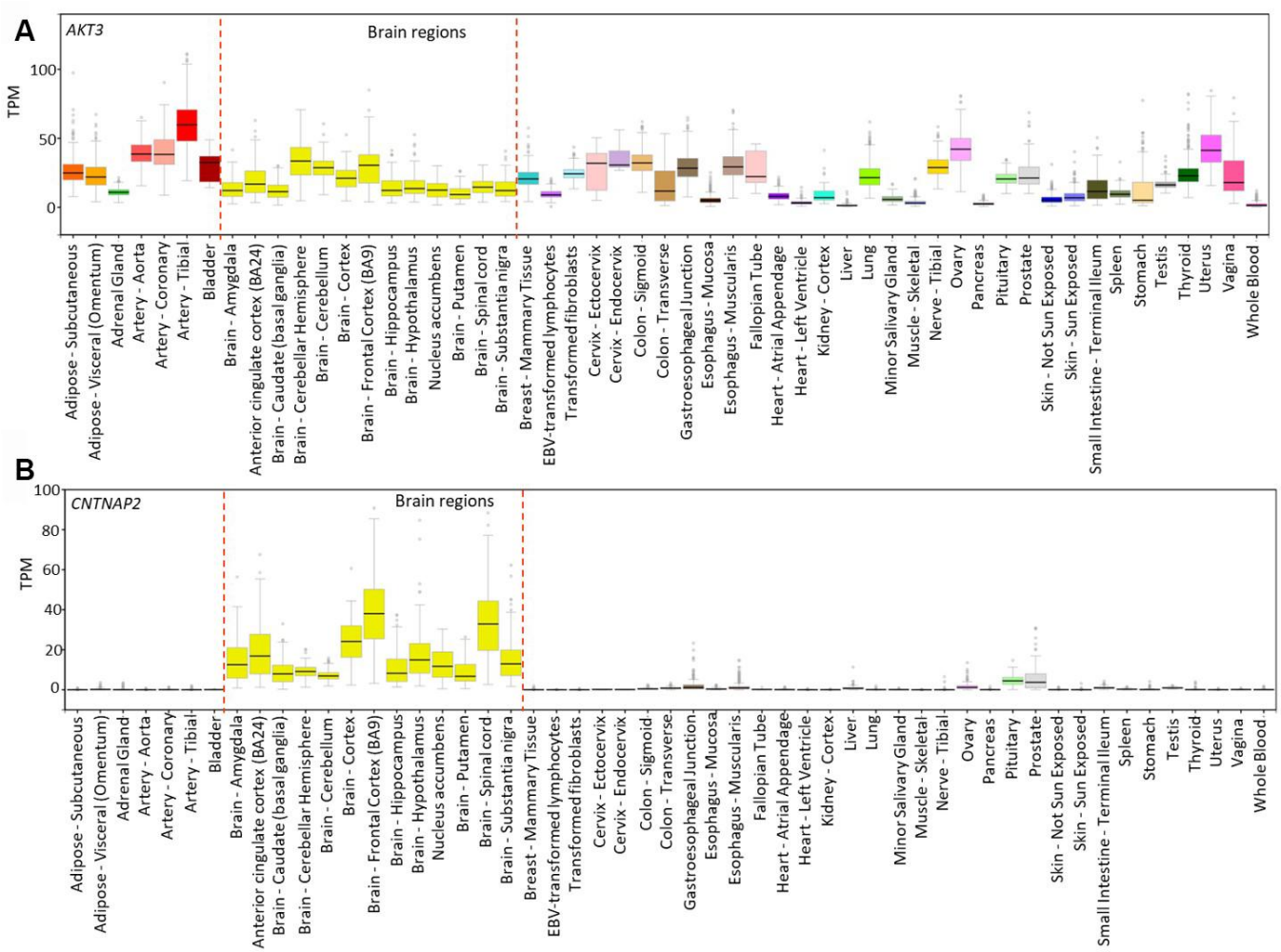
Supplementary Figure 19. Expression of (**A**) *AP2A2* and (**B**) *CACNA1D* in 53 human tissues. The expression data were obtained from GTEx analysis release V7 (dbGaP Accession No. phs000424.v7.p2). Expression values are shown in transcripts per million (TPM), calculated from a gene model with isoforms collapsed to a single gene. Box plots are shown as median and 25th and 75th percentiles; points are displayed as outliers if they are above or below 1.5 times the interquartile range.



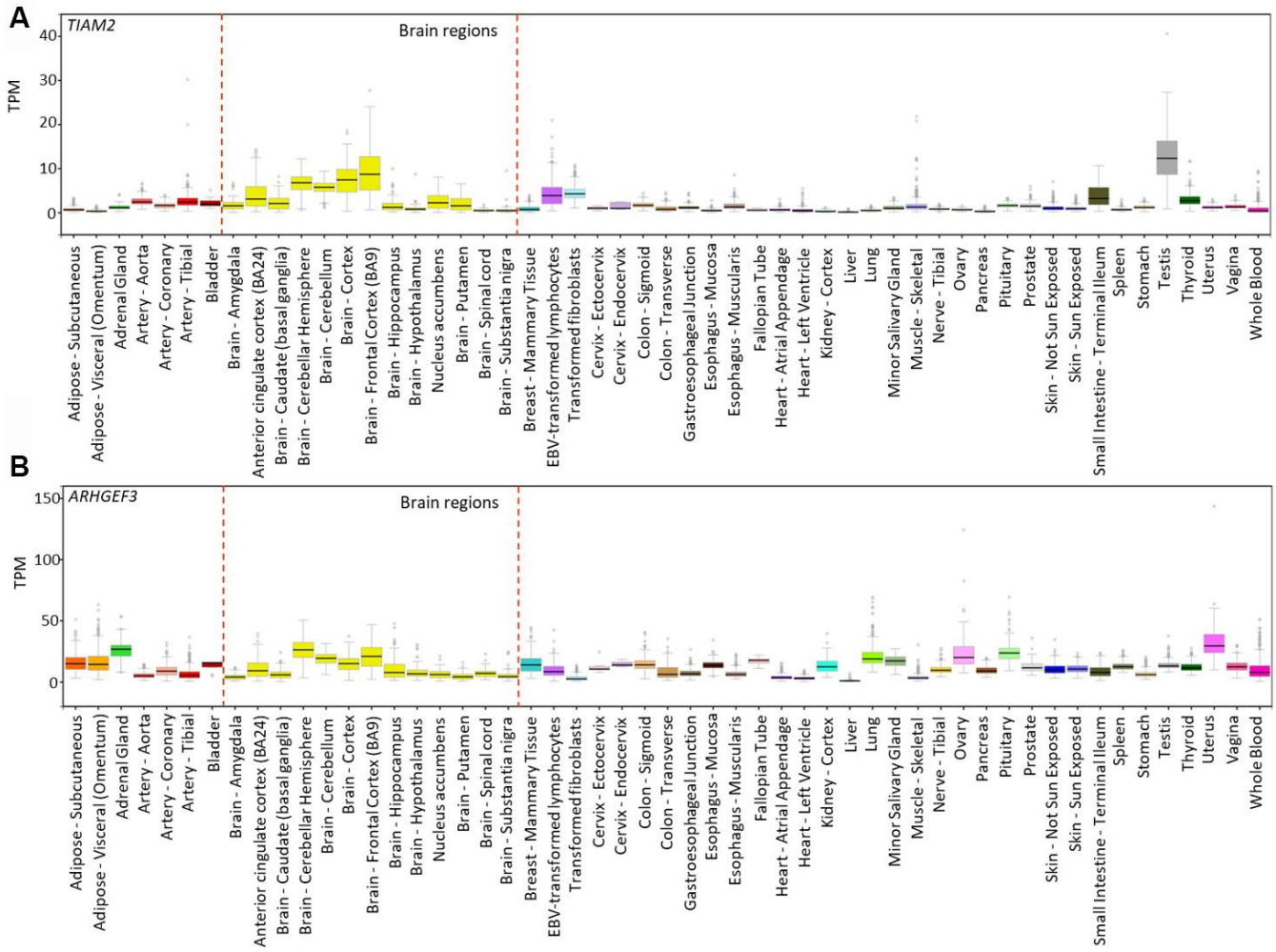
Supplementary Figure 20. Expression of (**A**) *GNA12* and (**B**) *TRIO* in 53 human tissues. The expression data were obtained from GTEx analysis release V7 (dbGaP Accession No. phs000424.v7.p2). Expression values are shown in transcripts per million (TPM), calculated from a gene model with isoforms collapsed to a single gene. Box plots are shown as median and 25th and 75th percentiles; points are displayed as outliers if they are above or below 1.5 times the interquartile range.



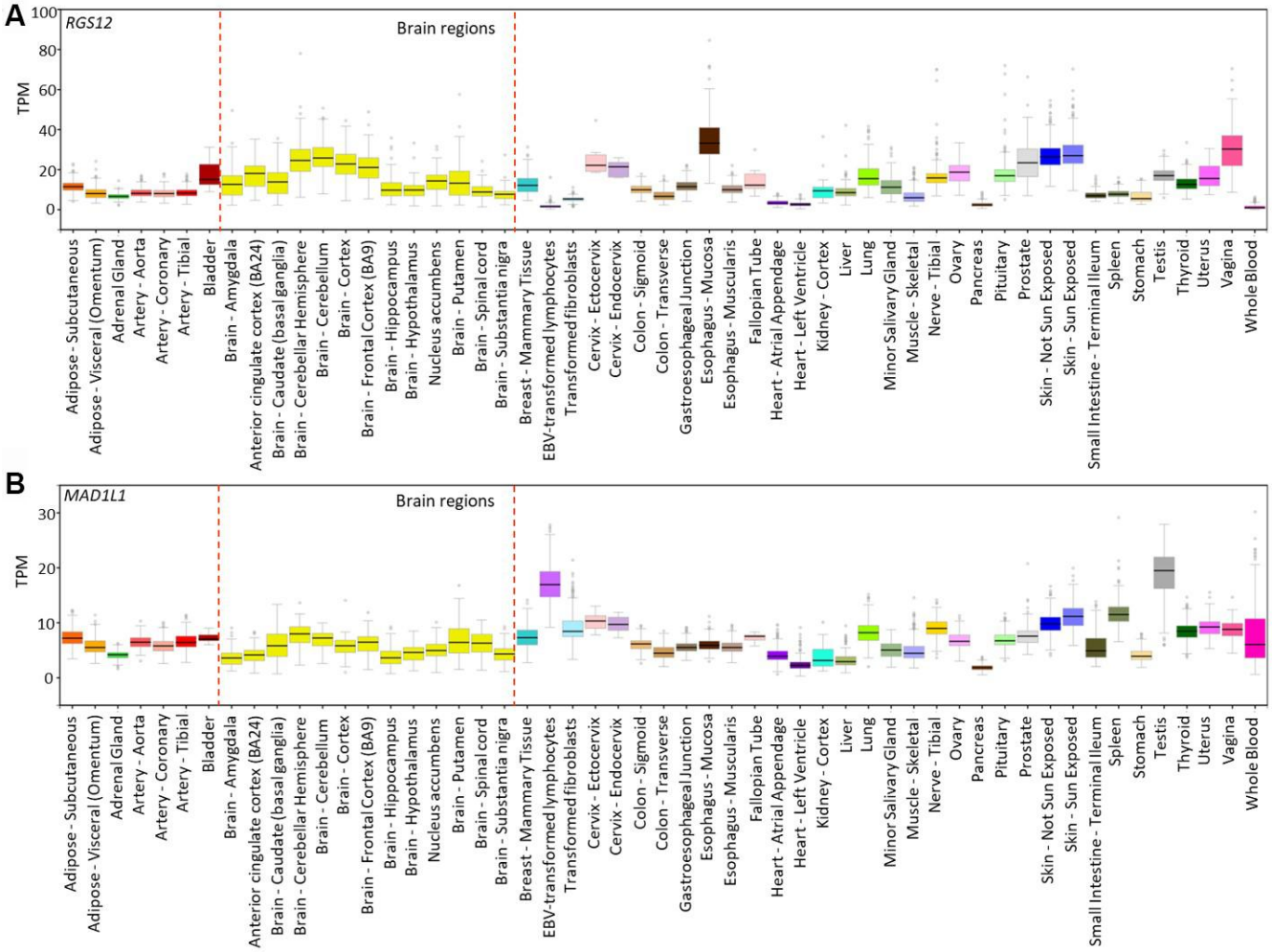
Supplementary Figure 21. Expression of (**A**) *HTT* and (**B**) *RPS6KA2* in 53 human tissues. The expression data were obtained from GTEx analysis release V7 (dbGaP Accession No. phs000424.v7.p2). Expression values are shown in transcripts per million (TPM), calculated from a gene model with isoforms collapsed to a single gene. Box plots are shown as median and 25th and 75th percentiles; points are displayed as outliers if they are above or below 1.5 times the interquartile range.



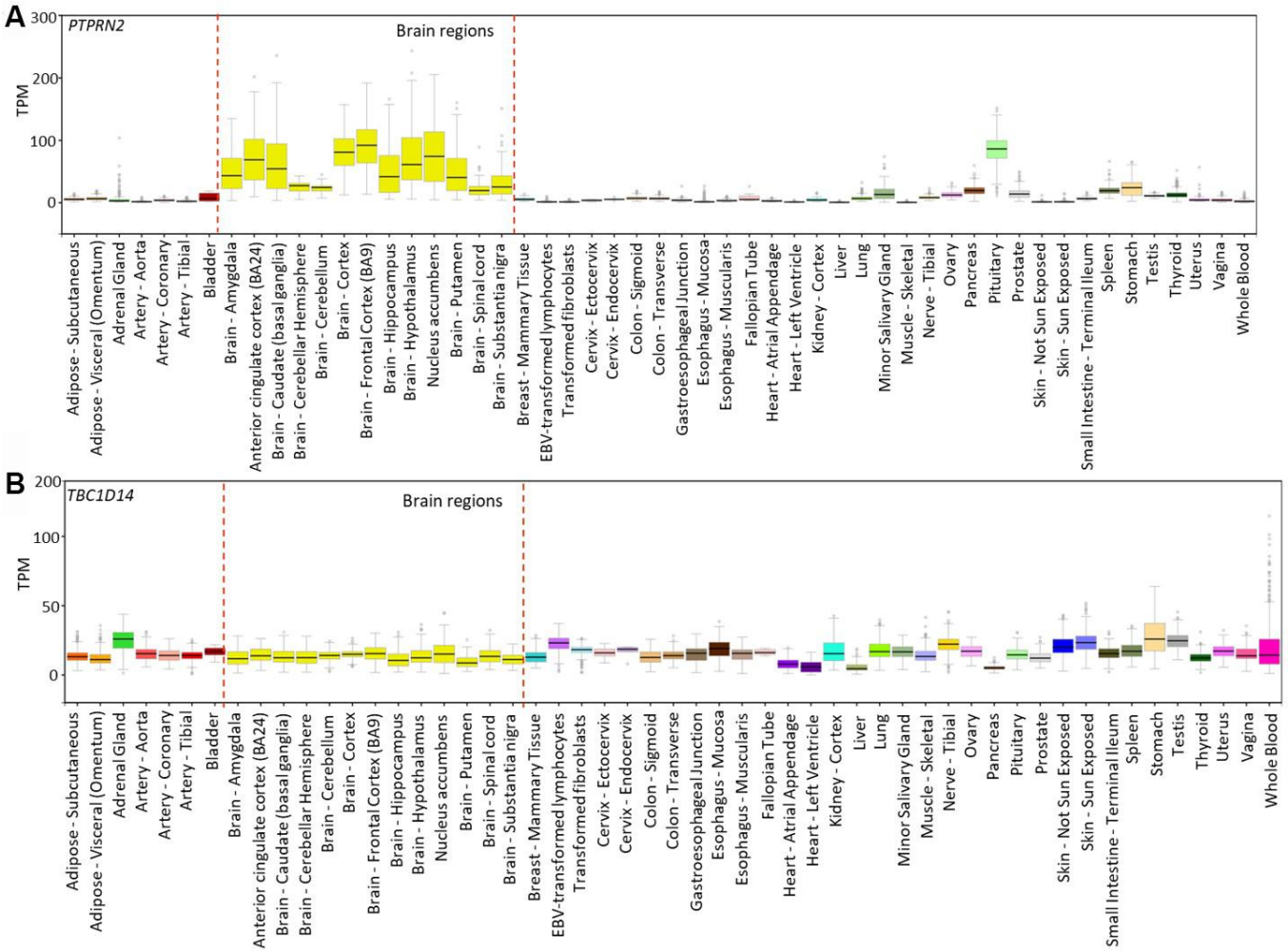
Supplementary Figure 22. Expression of (**A**) *AKT3* and (**B**) *CNTNAP2* in 53 human tissues. The expression data were obtained from GTEx analysis release V7 (dbGaP Accession No. phs000424.v7.p2). Expression values are shown in transcripts per million (TPM), calculated from a gene model with isoforms collapsed to a single gene. Box plots are shown as median and 25th and 75th percentiles; points are displayed as outliers if they are above or below 1.5 times the interquartile range.



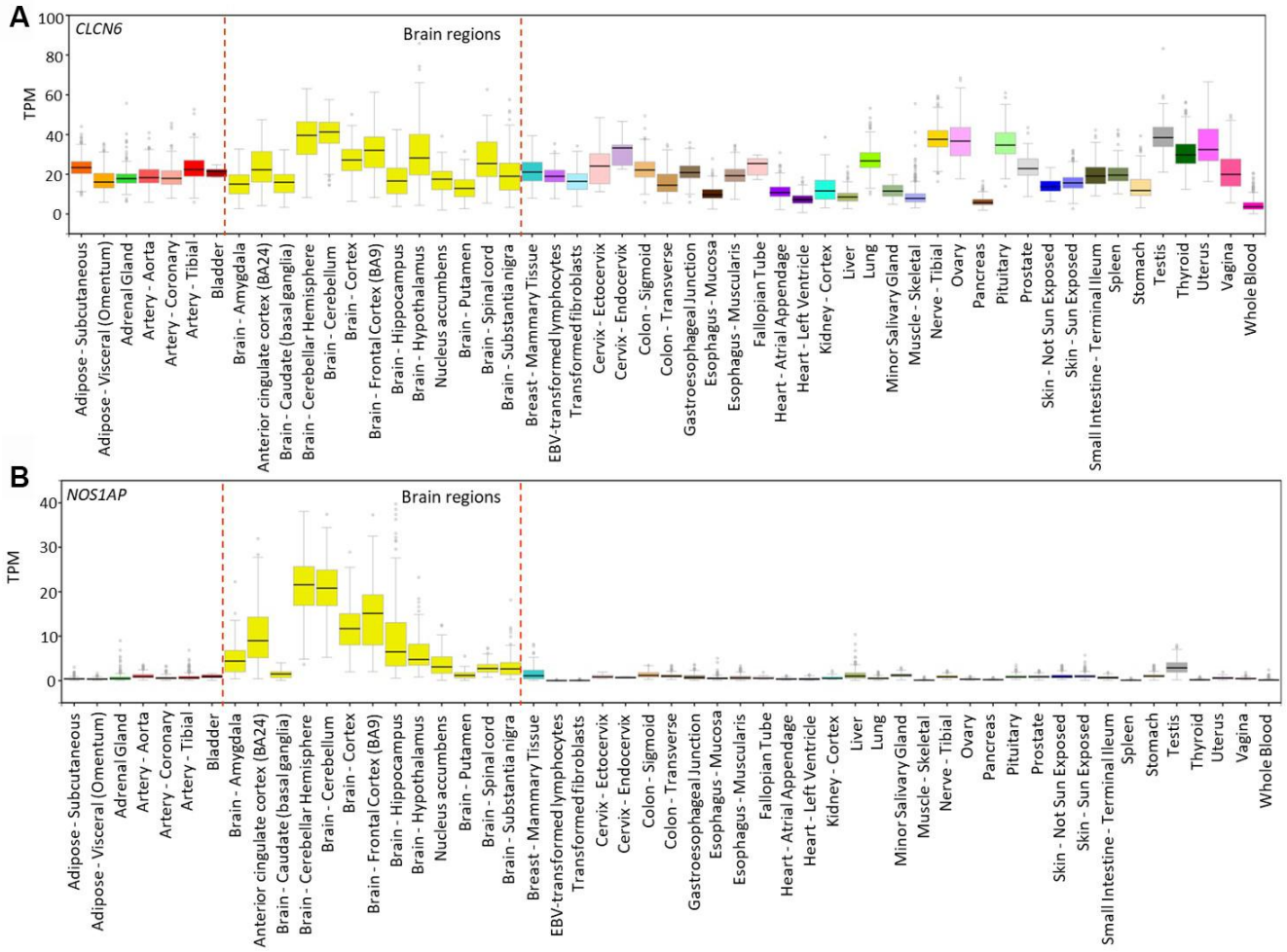
Supplementary Figure 23. Expression of (**A**) *TIAM2* and (**B**) *ARHGEF3* in 53 human tissues. The expression data were obtained from GTEx analysis release V7 (dbGaP Accession phs000424.v7.p2). Expression values are shown in transcripts per million (TPM), calculated from a gene model with isoforms collapsed to a single gene. Box plots are shown as median and 25th and 75th percentiles; points are displayed as outliers if they are above or below 1.5 times the interquartile range.



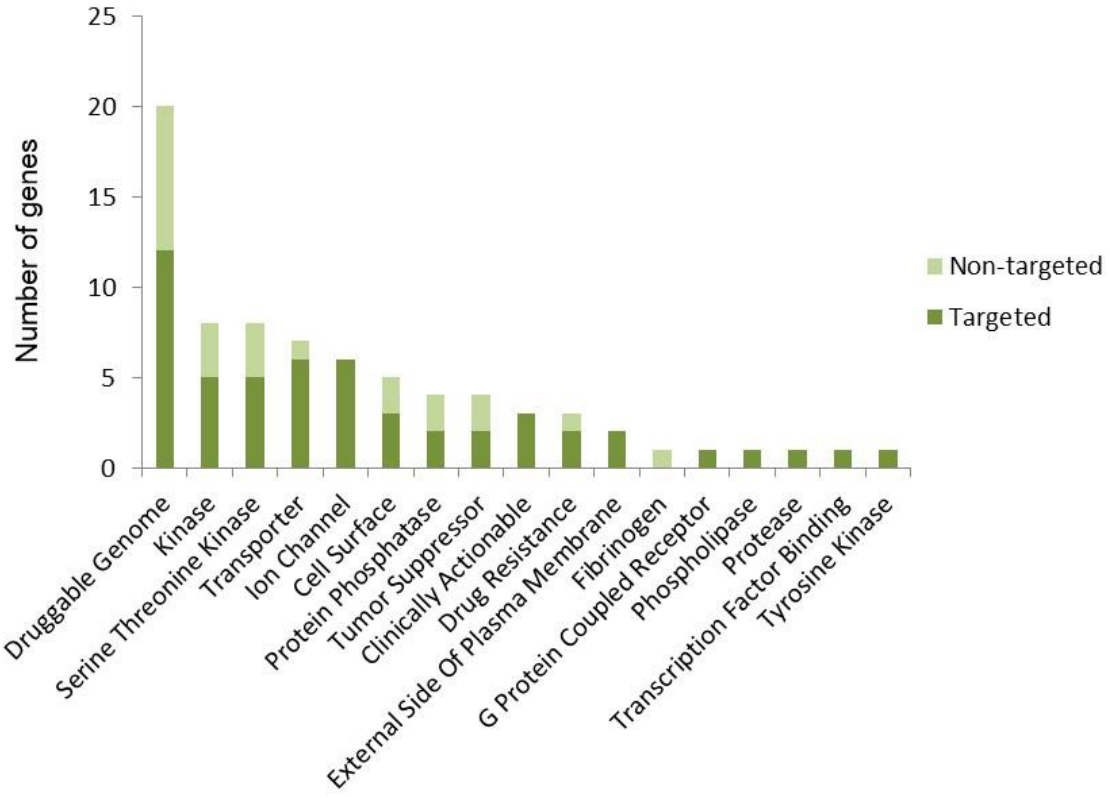
Supplementary Figure 24. Expression of (**A**) *RGS12* and (**B**) *MAD1L1* in 53 human tissues. The expression data were obtained from GTEx analysis release V7 (dbGaP Accession No. phs000424.v7.p2). Expression values are shown in transcripts per million (TPM), calculated from a gene model with isoforms collapsed to a single gene. Box plots are shown as median and 25th and 75th percentiles; points are displayed as outliers if they are above or below 1.5 times the interquartile range.



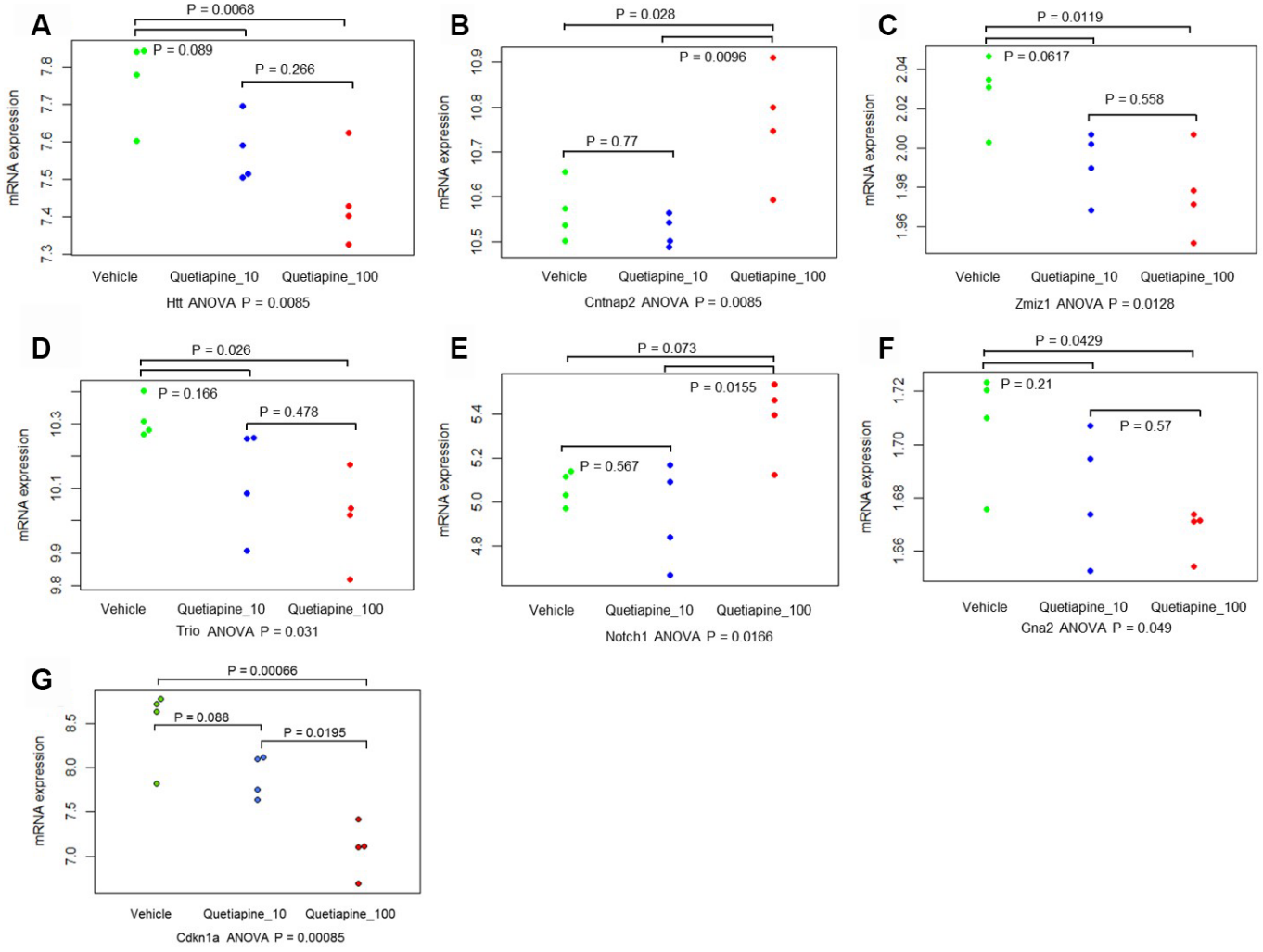
Supplementary Figure 25. Expression of (**A**) *PTPRN2* and (**B**) *TBC1D14* in 53 human tissues. The expression data were obtained from GTEx analysis release V7 (dbGaP Accession No. phs000424.v7.p2). Expression values are shown in transcripts per million (TPM), calculated from a gene model with isoforms collapsed to a single gene. Box plots are shown as median and 25th and 75th percentiles; points are displayed as outliers if they are above or below 1.5 times the interquartile range.



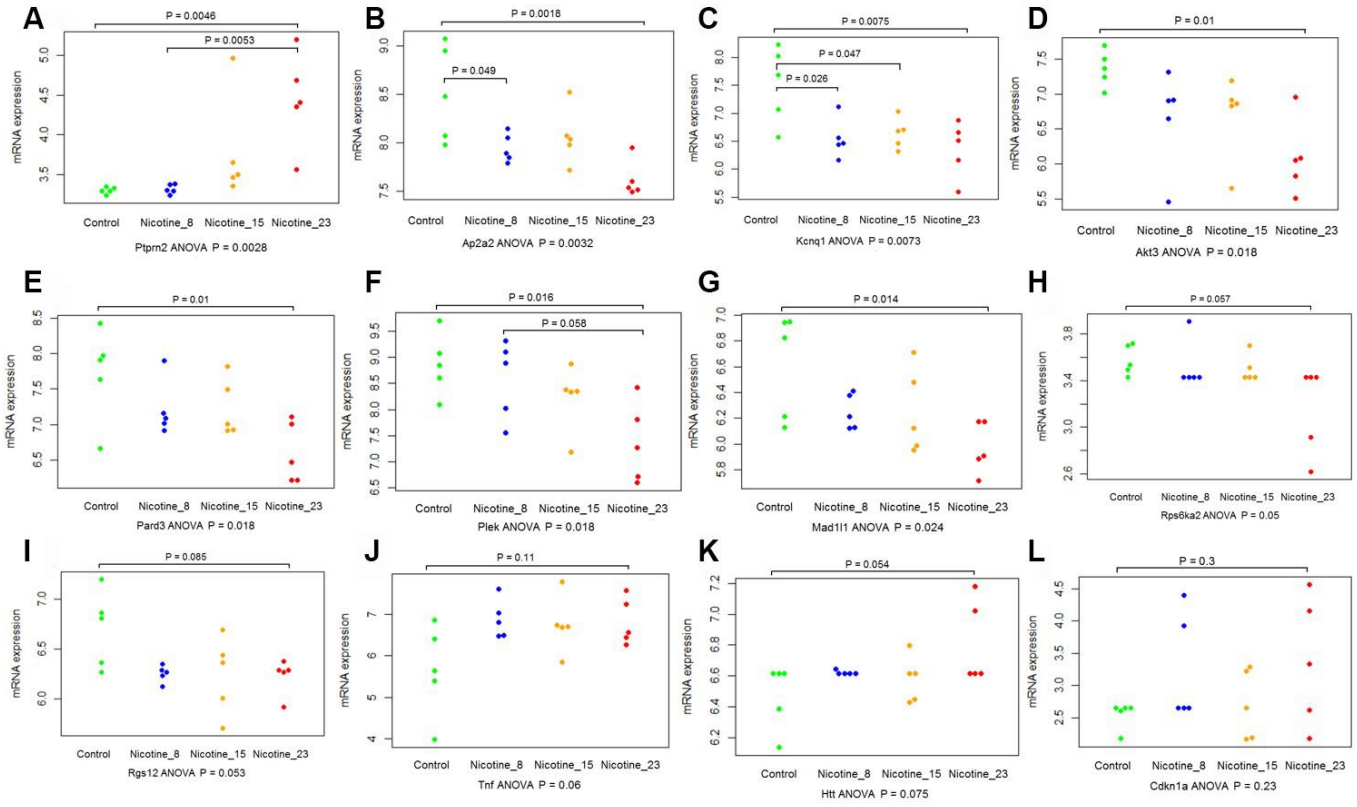
Supplementary Figure 26. Expression of (**A**) *CLCN6* and (**B**) *PRKCZ* in 53 human tissues. The expression data were obtained from GTEx analysis release V7 (dbGaP Accession phs000424.v7.p2). Expression values are shown in transcripts per million (TPM), calculated from a gene model with isoforms collapsed to a single gene. Box plots are shown as median and 25th and 75th percentiles; points are displayed as outliers if they are above or below 1.5 times the interquartile range.



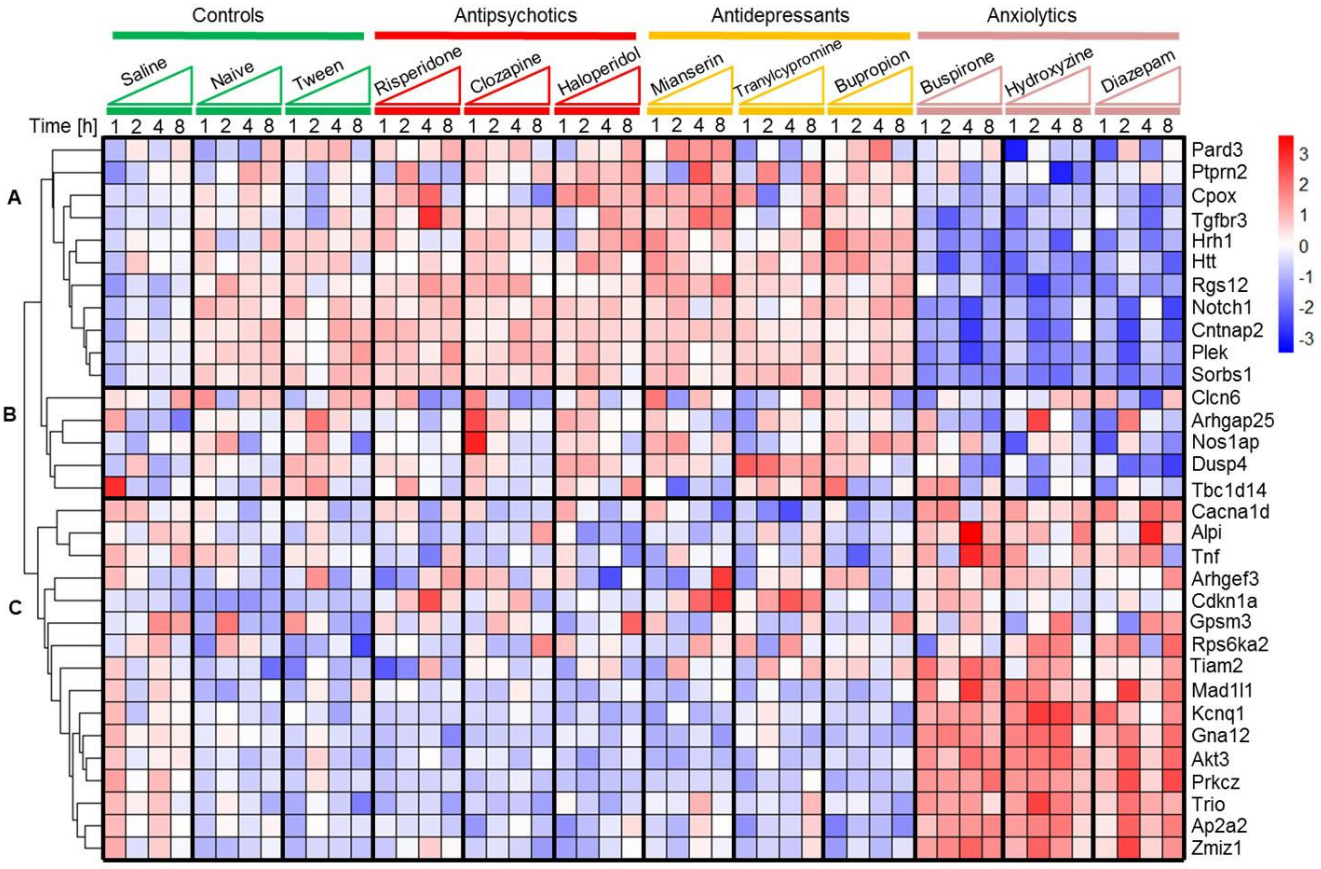
Supplementary Figure 27. Summary of candidate genes in potentially druggable categories. The numbers of these genes in potentially druggable categories, and the numbers of genes in these categories that are targeted by a known drug.



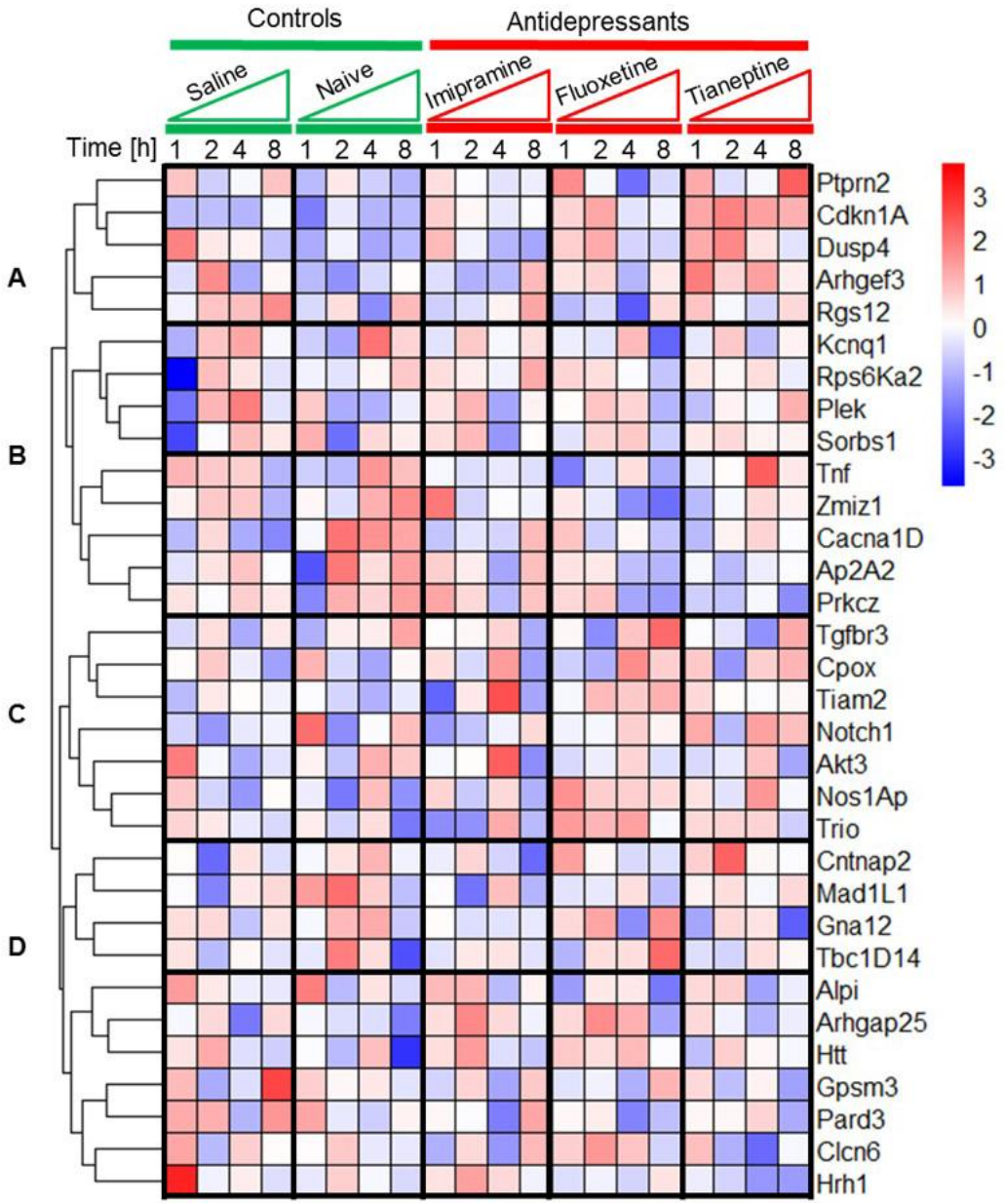
Supplementary Figure 28. Plots summarizing significant gene expression alterations in quetiapine (at doses of 10 or 100 mg/kg)-treated mice. One-way ANOVA was applied, and Tukey-Kramer HSD test was used for multiple comparisons. Gene expression data from the GEO dataset (Accession No. GSE45229). (**A**) For *Htt*; (**B**) for *Cntnap2*; (**C**) for *Zmiz1*; (**D**) for *Trio*; (**E**) for *Notch1*; (**F**) for *Gna2*. These plots were generated using the *beeswarm* package in R.



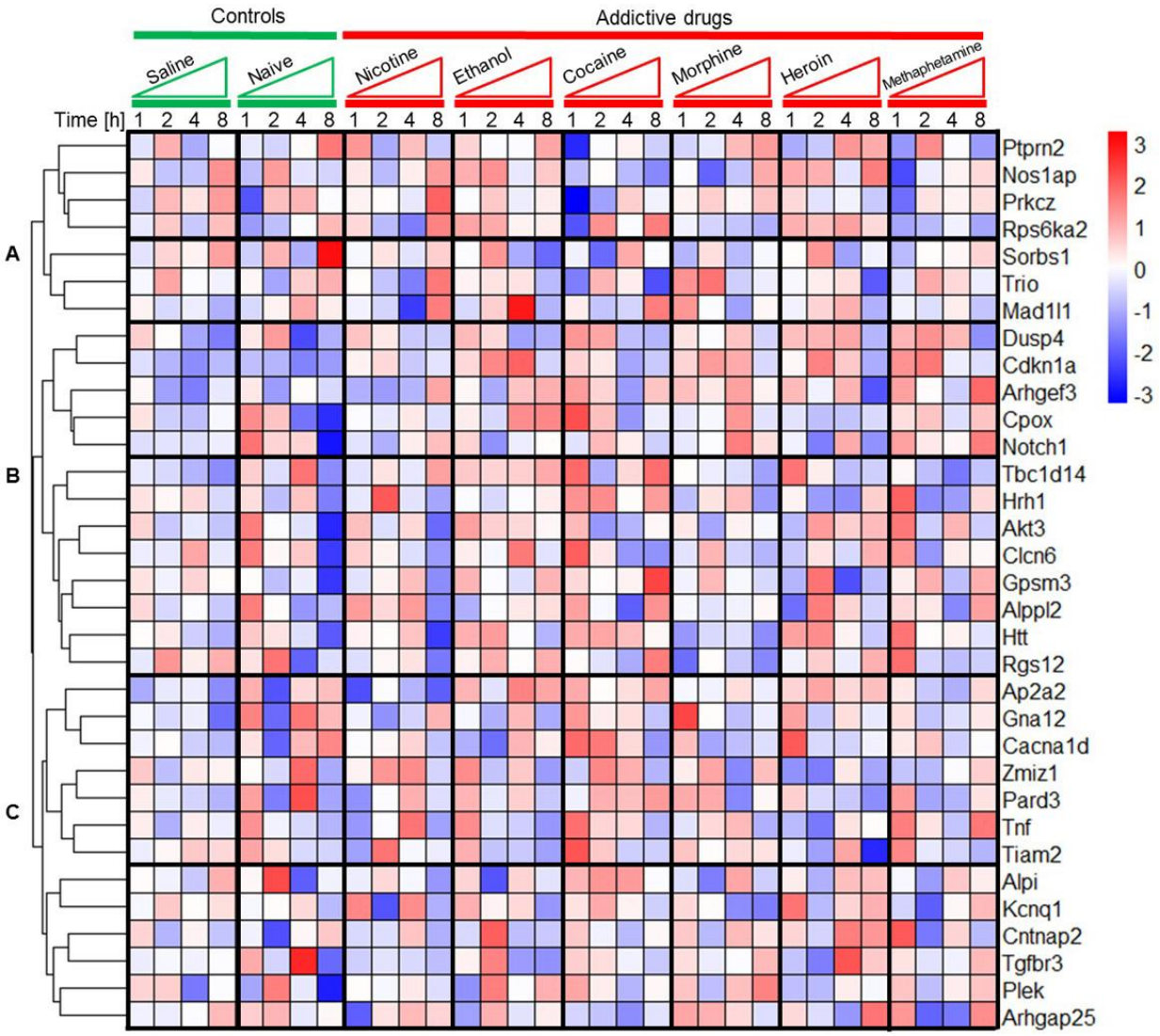
Supplementary Figure 29. Plots summarizing significant gene expression alterations in nicotine (at doses of 8 µg nicotine/l, 15 µg nicotine/l, and 23 µg nicotine/l)-treated mice. One-way ANOVA was applied, and Tukey-Kramer HSD test was used for multiple comparisons. Gene expression data from the GEO dataset (Accession No. GES50254). (**A**) For *Ptprn2*; (**B**) for *Ap2a2*; (**C**) for *Kcnq1*; (**D**) for *Akt3*; (**E**) for *Pard3*; (**F**) for *Plek*; (**G**) for *Mad1l1*; (**H**) for *Rps6ka2*; (**I**) for *Rgs12*; (**J**) for *Tnf*; (**K**) for *Htt*; (**L**) for *Cdkn1a*. These plots were generated using the *beeswarm* package in R.



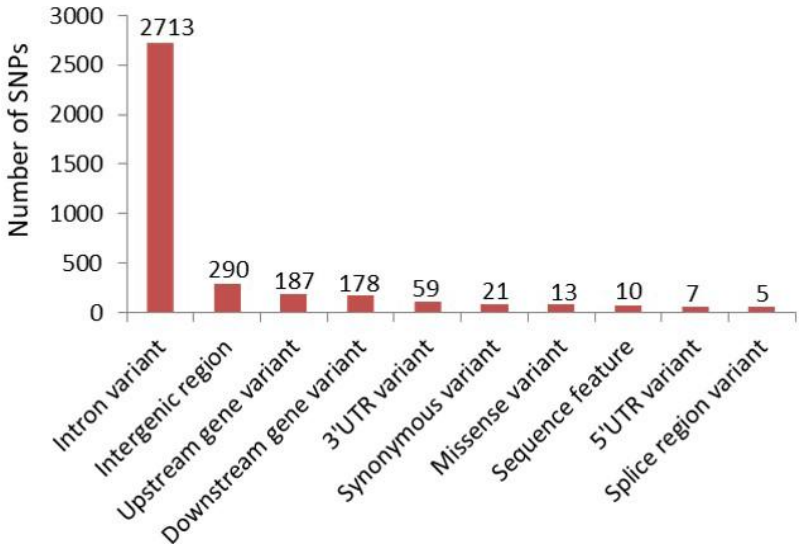
Supplementary Figure 30. Heatmap of the identified candidate genes treated by psychotropic drugs. Colored rectangles represent transcript abundance indicated above the gene labeled on the right. The intensity of the color is proportional to the standardized values between -3 and 3 for each gene, as indicated on the bar on the right of the heat map plot. Clustering was performed using Euclidean distance according to the scale on the left. Major gene transcription patterns are arbitrarily described as (**A**–**C**) Gene expression data from the GEO dataset (Accession No. GES50254). Three antidepressants (bupropion 20 mg/kg, tranylcypromine 20 mg/kg, mianserin 20 mg/kg, i.p.), three anxiolytics (diazepam 5 mg/kg, buspirone 10 mg/kg, hydroxyzine 10 mg/kg, i.p.), and three antipsychotics (clozapine 3 mg/kg, risperidone 0.5 mg/kg, haloperidol 1 mg/kg) were selected for the comparison. To analyze dynamics of early, intermediate, and relatively late changes of mRNA abundance, the experiment was performed at four time points (1, 2, 4, and 8h after drug administration). To exclude influence of drug injection and circadian rhythm on gene expression profile, control groups of saline or Tween (1% Tween 80)-treated and naïve animals were prepared for each time point. Heatmap plots were generated using the *pheatmap* package in R.



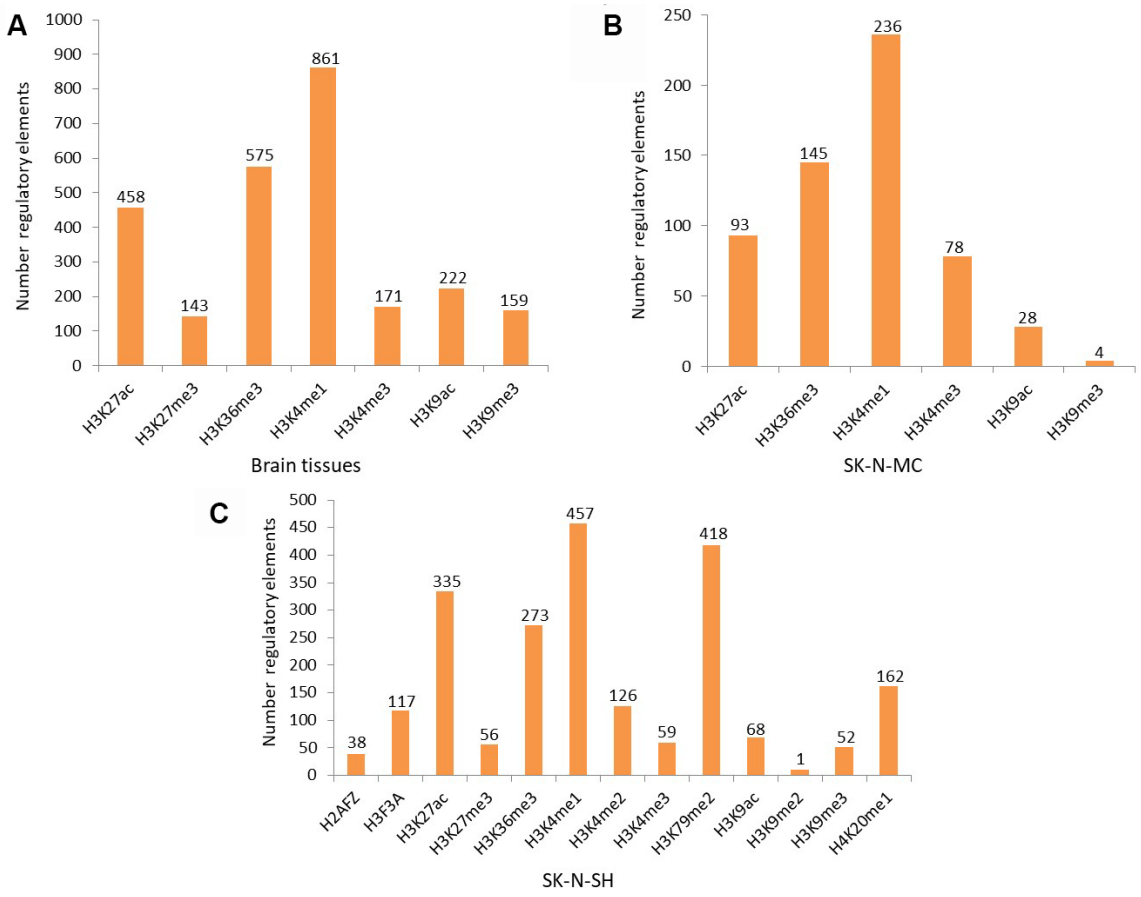
Supplementary Figure 31. Heatmap of the identified candidate genes treated by psychotropic drugs. Colored rectangles represent transcript abundance indicated above the gene labeled on the right. The intensity of the color is proportional to the standardized values between -3 and 3 for each gene, as indicated on the bar on the right of the heat map plot. Clustering was performed using Euclidean distance according to the scale on the left. Major gene transcription patterns are arbitrarily described as (**A**–**D**) Gene expression data from the GEO dataset (Accession No. GSE48951). Three antidepressants (imipramine 10 mg/kg, fluoxetine 20 mg/kg, and tianeptine 20 mg/kg, i.p.) were selected for the comparison. To analyze dynamics of early, intermediate, and relatively late changes of mRNA abundance, the experiment was performed at four time points (1, 2, 4, and 8h after drug administration). To exclude influence of drug injection and circadian rhythm on gene expression profile, control groups of saline or Tween (1% Tween 80)-treated and naïve animals were prepared for each time point. Heatmap plots were generated using the *pheatmap* package in R.



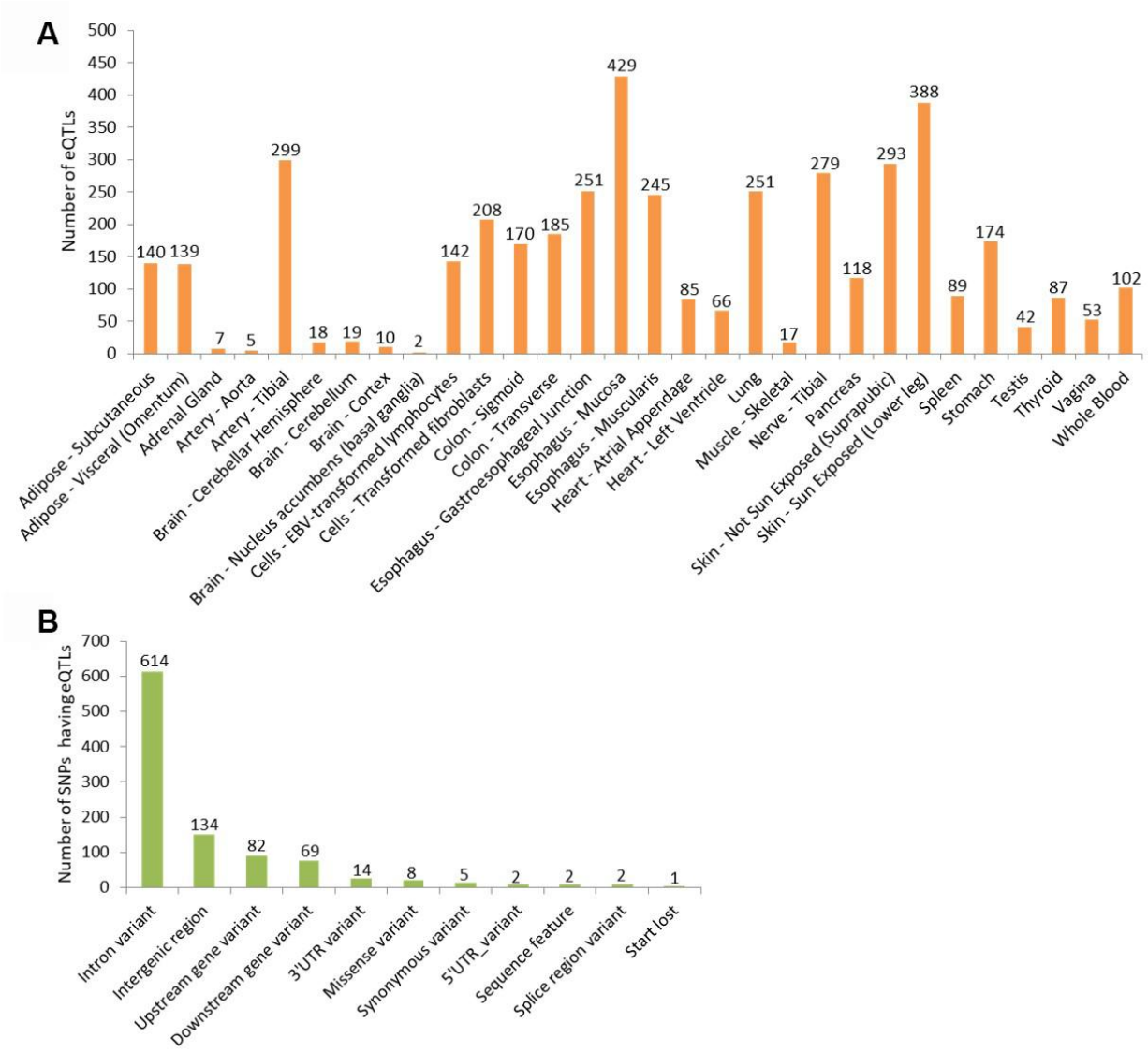
Supplementary Figure 32. Heatmap of the identified candidate genes treated by psychotropic drugs. Colored rectangles represent transcript abundance indicated above the gene labeled on the right. The intensity of the color is proportional to the standardized values between -3 and 3 for each gene, as indicated on the bar on the right of the heat map plot. Clustering was performed using Euclidean distance according to the scale on the left. Major gene transcription patterns are arbitrarily described as (**A**–**C**) Gene expression data from the GEO dataset (Accession No. GSE15774). Six of the most addictive and harmful drugs of abuse (morphine 20 mg/kg, heroin 10 mg/kg, ethanol 2 g/kg, nicotine 1 mg/kg, methamphetamine 2 mg/kg, or cocaine 25 mg/kg, i.p.) were selected for comparison. To analyze dynamics of early, intermediate, and relatively late changes of mRNA abundance, the experiment was performed at four time points (1, 2, 4, and 8h after drug administration). To exclude influence of drug injection and circadian rhythm on gene expression profile, control groups of saline-treated and naïve animals were prepared for each time point. Heatmap plots were generated using the *pheatmap* package in R.



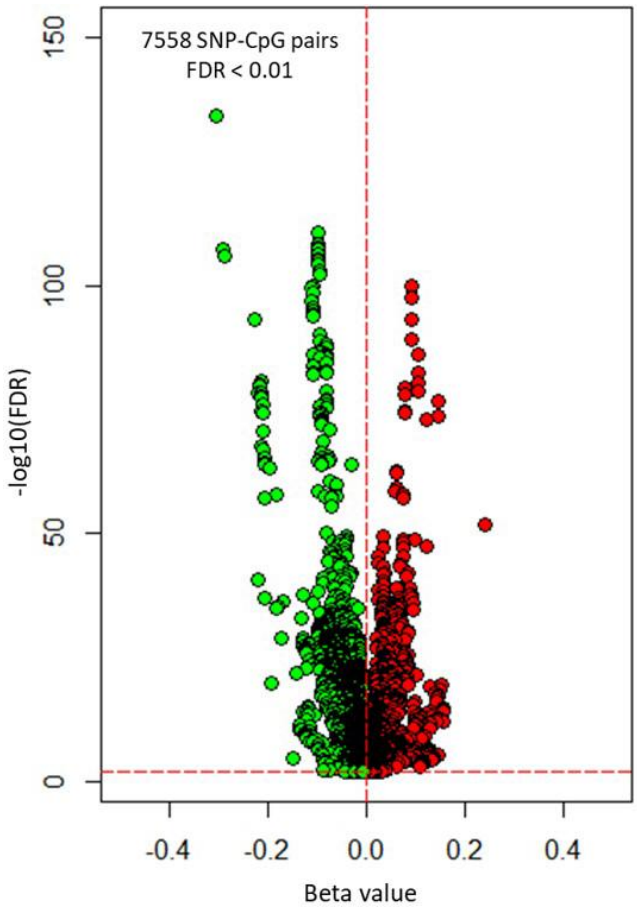
Supplementary Figure 33. Variants annotation of 3,483 SNPs shared by SCZ and at least one phenotype of smoking behavior. We used snpEff (http://snpeff.sourceforge.net/) to carry out the annotation analysis.



Supplementary Figure 34. 3843 SNPs shared by SCZ and smoking located within different types of regulatory elements in brain tissues and neuroblastoma cell lines. (**A**) For brain tissues. (**B**) for SK-N-MC cells. (**C**) for SK-N-SH. The data of different types of regulatory elements, including H2AFZ, H3F3A, H3K27ac, H3K27me3, H3K36me3, H3K4me1, H3K4me2, H3K4me3, H3K79me2, H3K9ac, H3K9me2, H3K9me3, and H4K20me1, were downloaded from ENCODE database (https://www.encodeproject.org/).



Supplementary Figure 35. 933 of 3843 SNPs shared by SCZ and smoking having 4313 eQTLs in 31 human tissues. (**A**) 4313 eQTLs distributed in 31 human tissues. (**B**) Variant annotation of 933 SNPs shared by SCZ and at least one phenotype of smoking behavior. We used the tool snpEff (http://snpeff.sourceforge.net/) to carry out the annotation analysis. Based on the tissue-specific eQTL association data from the GTEx database, 933 of 3,483 SNPs have *cis*-regulatory effects on gene expression with 4,313 SNP-eQTLs pairs in 31 human tissues.



Supplementary Figure 36. Volcano plot for significant SNP–CpG methylation associations in brain samples. Green represents these significant SNP-CpG pairs with negative beta value. Red represents these significant SNP-CpG pairs with positive beta value.

Supplementary Tables

Please browse Full Text version to see the data of Supplementary Tables 1, 2, 3, 5, 7.

Supplementary Table 1. Multiple omics datasets used in current study downloaded from public databases.

Supplementary Table 2. 84 significantly enriched pathways shared between schizophrenia and smoking behaviors.

Supplementary Table 3. Shared genes between schizophrenia and smoking behaviors.

Supplementary Table 4. Module-level statistics.

Statistics	Modules					
	Blue	Brown	Green	Red	Turquoise	Yellow
Orbital prefrontal (OFC)	4.93	8.42	8.96	7.57	7.89	8.27
Dorsolateral prefrontal (DFC)	4.95	8.39	8.95	7.64	7.90	8.23
Ventrolateral prefrontal (VFC)	4.93	8.40	8.97	7.62	7.90	8.10
Medial prefrontal (MFC)	4.94	8.36	8.92	7.61	7.88	8.25
Primary motor (M1C)	4.95	8.37	8.93	7.58	7.87	7.50
Primary somatosensory (S1C)	4.97	8.36	8.93	7.57	7.87	7.49
Inferior parietal (IPC)	4.97	8.37	8.93	7.56	7.86	7.49
Primary auditory (A1C)	4.96	8.37	8.94	7.55	7.86	7.49
Superior temporal (STC)	4.97	8.36	8.93	7.56	7.86	8.07
Inferior temporal (ITC)	4.97	8.37	8.91	7.57	7.85	8.07
Primary visual (V1C)	4.96	8.27	8.89	7.49	7.81	7.48
Hippocampus (HIP)	4.95	8.20	8.86	7.40	7.80	8.15
Amygdala (AMY)	4.96	8.23	8.88	7.45	7.80	8.17
Striatum (STR)	4.94	8.03	8.88	7.27	7.70	7.35
Thalamus (MD)	4.95	8.09	8.88	7.21	7.67	7.34
Cerebellum (CBC)	4.93	7.90	8.85	7.04	7.52	7.14
max(log2(expression))	4.97	8.42	8.97	7.64	7.90	8.27
min(log2(expression))	4.93	7.90	8.85	7.04	7.52	7.14
max(log2(FC))	0.05	0.52	0.12	0.60	0.38	1.12
8-10 weeks (stage2)	4.92	7.59	8.99	7.32	7.66	8.27
10-13 weeks (stage3)	5.01	8.12	9.16	8.18	7.78	8.08
13-16 weeks (stage4)	5.01	8.03	9.13	8.21	7.73	8.15
16-19 weeks (stage5)	4.97	8.22	9.03	8.27	7.77	8.06
19-24 weeks (stage6)	4.88	7.77	8.91	7.52	7.34	7.34
24-38 weeks (stage7)	5.13	8.41	9.17	8.56	7.98	8.05
0-0.5 years (stage8)	5.16	8.58	9.13	8.80	8.07	8.08
0.5-1 years (stage9)	5.25	8.74	8.97	8.72	7.85	7.78
1-6 years (stage10)	5.26	8.50	8.69	8.47	7.84	7.73
6-12 years (stage11)	5.04	8.72	8.98	8.52	7.87	7.70
12-20 years (stage12)	5.09	8.71	8.84	8.51	7.80	7.71
20-40 years (stage13)	5.08	8.73	8.94	8.43	7.83	7.67
40-60 years (stage14)	5.21	8.70	8.91	8.33	7.81	7.64
60+ years (stage15)	5.19	8.60	8.78	8.24	7.79	7.63
max(log_2(expression))	5.26	8.74	9.17	8.80	8.07	8.27
min(log_2(expression))	4.88	7.59	8.69	7.32	7.34	7.34
max(log_2(FC))	0.38	1.15	0.48	1.47	0.73	0.92

Note: For both regional expression patterns and temporal expression patterns, we used the method of mean log_2(expression) of all genes in each module.

Supplementary Table 5. 149 coexpression module genes involved in smoking- and SCZ-associated methylation.

Supplementary Table 6. 16 of 34 smoking- and SCZ-associated methylation genes have been reported to be associated with SCZ in previous studies.

Gene Name	Description	References
AKT3	Genetic variation in the *AKT3* locus (chr1:243503719–244002945) is a top GWAS signal in schizophrenia and pathway analysis identified 50 single nucleotide polymorphisms (SNPs) within the *AKT3* gene that contribute to four of the top pathways associated with risk for schizophrenia and bipolar disorder. *AKT3* shows prenatal enrichment during human neocortical development and recurrent copy number variations involving the 1q43-44 locus are associated with cortical malformations and intellectual disability, implicating an essential role in early brain development.	[1, 14–17]
AP2A2	Schizophrenia-based genetic association study shows the involvement of the clathrin-mediated endocytosis (CME)-related protein enthoprotein encoded by the clathrin interactor 1 (*CLINT1*) gene. The expression of genes encoding adaptor and clathrin assembly proteins, *AP2A2*, *AP2B1*, *AP180*, *CLINT1*, *HIP1*, *ITSN2*, and *PICALM*, increased relative to the control in SH-SY5Y cells incubated with 5–10 μmol/l clozapine for 24–72 h.	[18]
CACNA1D	The animal model of neonatal lesion in ventral hippocampus (NLVH) is a recognized animal model for schizophrenia. NLVH influenced change expressions in various genes involved in Ca^{2+} homeostasis, including Cacna1d, Atp2a2, Adcy2, Ppp3cb, and Ptk2b.	[19–21]
CDKN1A	Quetiapine is an atypical neuroleptic with a pharmacological profile distinct from classic neuroleptics that function primarily via blockade of dopamine D2 receptors. In the United States, quetiapine is currently approved for treating patients with schizophrenia, major depression and bipolar I disorder. Despite its widespread use, its cellular effects remain elusive. To address possible mechanisms, we chronically treated mice with quetiapine, haloperidol or vehicle and examined quetiapine-specific gene expression change in the frontal cortex. Through microarray analysis, we observed that several groups of genes were differentially expressed upon exposure to quetiapine compared with haloperidol or vehicle; among them, Cdkn1a, the gene encoding p21, exhibited the greatest fold change relative to haloperidol. The quetiapine-induced downregulation of p21/Cdkn1a was confirmed by real-time polymerase chain reaction and in situ hybridization. Consistent with single gene-level analyses, functional group analyses also indicated that gene sets associated with cell cycle/fate were differentially regulated in the quetiapine-treated group. In cortical cell cultures treated with quetiapine, p21/Cdkn1a was significantly downregulated in oligodendrocyte precursor cells and neurons, but not in astrocytes. We propose that cell cycle-associated intervention by quetiapine in the frontal cortex may underlie a unique efficacy of quetiapine compared with typical neuroleptics.	[12]
CNTNAP2	Contactin associated protein-like 2 (CNTNAP2) has emerged as a prominent susceptibility gene implicated in multiple complex neurodevelopmental disorders, including autism spectrum disorders (ASD), intellectual disability (ID), and schizophrenia (SCZ). Based on genomic rearrangements and copy number variations, the contactin-associated protein-like 2 gene (CNTNAP2) has been implicated in neurodevelopmental disorders such as Gilles de la Tourette syndrome, intellectual disability, obsessive compulsive disorder, cortical dysplasia-focal epilepsy syndrome, autism, schizophrenia, Pitt-Hopkins syndrome, and attention deficit hyperactivity disorder.	[22–24]
GNA12	*GNA12* is significantly associated with schizophrenia based on network-assisted investigation of combined causal signals from GWAS studies in schizophrenia.	[25]
GPSM3	*GPSM3* is significantly associated with schizophrenia based on a multi-stage schizophrenia GWAS.	[1]
HRH1	*HRH1* is significantly associated with schizophrenia based on genetic association study.	[26, 27]
KCNQ1	Patients with schizophrenia show decreased processing speed on neuropsychological testing and decreased white matter integrity as measured by diffusion tensor imaging, two traits shown to be both heritable and genetically associated indicating that there may be genes that influence both traits as well as schizophrenia disease risk. The potassium channel gene family is a reasonable candidate to harbor such a gene given the prominent role potassium channels play in the central nervous system in signal transduction, particularly in myelinated axons. *KCNQ1* may contribute to the shared risk for diminished processing speed, diminished white mater integrity and increased risk of	[28, 29]

	schizophrenia.	
MAD1L1	Multiple lines of evidence from genetic association studies indicate *MAD1L1* confer risk to schizophrenia.	[1, 30, 31]
NOS1AP	*NOS1AP* is a protein implicated in schizophrenia. Several independent studies reported linkage of schizophrenia to chromosome 1q21–22, containing *NOS1AP*. Many molecular functional-based and genetic-based studies have been identified *NOS1AP* gene as a schizophrenia susceptibility gene.	[32–38]
NOTCH1	The antipsychotic and myelin protective effects of quetiapine are mediated by Notch signaling in a mouse model of cuprizone-induced demyelination associated with schizophrenia-like behaviors. The Notch pathway might therefore be a novel target for the development of antipsychotic drugs.	[39]
PARD3	Based on a genetic-based association study, *PARD3* is associated with susceptibility to schizophrenia in a Korean population.	[40]
RGS12	*RGS12* is putative candidate genes for sporadic schizophrenia.	[41]
TNF	TNF-α is associated with the deficit syndrome and negative symptoms in patients with chronic schizophrenia.	[42, 43]
TRIO	Disrupted-in-Schizophrenia 1-mediated axon guidance involves TRIO-RAC-PAK small GTPase pathway signaling	[44]

Supplementary Table 7. Differential expressed genes in SCZ patients and controls divided by smoking status.

Gene Name	SCZ VS CTRL	Smoking VS Nonsmoking	Anova P value[a]
PTPRN2	0.0000518	0.316	0.00142
ARHGAP25	0.0001	0.0028	0.00026
ARHGEF3	0.000281	0.11	0.00723
AKT3	0.00111	0.381	0.0395
RGS12	0.00202	0.347	0.0058
SORBS1	0.0134	0.471	0.108
TIAM2	0.0161	0.954	0.062
CACNA1D	0.0162	0.403	0.052
HTT	0.0179	0.259	0.132
NOTCH1	0.0239	0.301	0.156
PLEK	0.0256	0.0383	0.043
RPS6KA2	0.0268	0.815	0.112
PARD3	0.0296	0.033	0.054
TRIO	0.0326	0.324	0.147
CNTNAP2	0.0344	0.062	0.087
GPSM3	0.0369	0.353	0.0014
ZMIZ1	0.0708	0.499	0.041
TBC1D14	0.845	0.0389	0.030

Note: *anova* statistical test was used for comparing the significant differences between SCZ patients and controls grouped by smoking status.