

# A genome-wide multiphenotypic association analysis identified candidate genes and gene ontology shared by four common risky behaviors

Jing Ye<sup>1,\*</sup>, Li Liu<sup>1,\*</sup>, Xiaoqiao Xu<sup>2,\*</sup>, Yan Wen<sup>1</sup>, Ping Li<sup>1</sup>, Bolun Cheng<sup>1</sup>, Shiqiang Cheng<sup>1</sup>, Lu Zhang<sup>1</sup>, Mei Ma<sup>1</sup>, Xin Qi<sup>1</sup>, Chujun Liang<sup>1</sup>, Om Prakash Kafle<sup>1</sup>, Cuiyan Wu<sup>1</sup>, Sen Wang<sup>1</sup>, Xi Wang<sup>1</sup>, Yujie Ning<sup>1</sup>, Xiaomeng Chu<sup>1</sup>, Lin Niu<sup>2</sup>, Feng Zhang<sup>1</sup>

<sup>1</sup>Key Laboratory of Trace Elements and Endemic Diseases of National Health and Family Planning Commission, School of Public Health, Health Science Center, Xi'an Jiaotong University, Xi'an, China

<sup>2</sup>Clinical Research Center of Shaanxi Province for Dental and Maxillofacial Diseases, College of Stomatology, Xi'an Jiaotong University, Xi'an, China

\*Equal contribution

**Correspondence to:** Feng Zhang; email: [fzhxjtu@mail.xjtu.edu.cn](mailto:fzhxjtu@mail.xjtu.edu.cn)

**Keywords:** risky behaviors, regulatory SNP (rSNP), methylation quantitative trait loci (MeQTL), multivariate adaptive shrinkage (MASH)

**Received:** August 23, 2019

**Accepted:** January 25, 2020

**Published:** February 22, 2020

**Copyright:** Ye et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## ABSTRACT

**Background:** Risky behaviors can lead to huge economic and health losses. However, limited efforts are paid to explore the genetic mechanisms of risky behaviors.

**Result:** MASH analysis identified a group of target genes for risky behaviors, such as APBB2, MAPT and DCC. For GO enrichment analysis, FUMA detected multiple risky behaviors related GO terms and brain related diseases, such as regulation of neuron differentiation (adjusted  $P$  value =  $2.84 \times 10^{-5}$ ), autism spectrum disorder (adjusted  $P$  value =  $1.81 \times 10^{-27}$ ) and intelligence (adjusted  $P$  value =  $5.89 \times 10^{-15}$ ).

**Conclusion:** We reported multiple candidate genes and GO terms shared by the four risky behaviors, providing novel clues for understanding the genetic mechanism of risky behaviors.

**Methods:** Multivariate Adaptive Shrinkage (MASH) analysis was first applied to the GWAS data of four specific risky behaviors (automobile speeding, drinks per week, ever-smoker, number of sexual partners) to detect the common genetic variants shared by the four risky behaviors. Utilizing genomic functional annotation data of SNPs, the SNPs detected by MASH were then mapped to target genes. Finally, gene set enrichment analysis of the identified candidate genes were conducted by the FUMA platform to obtain risky behaviors related gene ontology (GO) terms as well as diseases and traits, respectively.

## INTRODUCTION

Risky behavior or risk-taking behavior has been defined as either a socially unacceptable volitional behavior with a potentially negative outcome in which precautions are not taken (e.g. speeding, drinking and driving), or a socially accepted behavior in which the danger is recognized (e.g. competitive sports and sky-

diving) [1]. It is well reported that risky behaviors were associated with high prevalence, low productivity and more generally with a decline of individual and collective well-being in the short, medium and long run [2], as well as high premature death and increased health care spending [3], leading to huge economic and health losses of society, especially automobile speeding, drinking, smoking, and multiple sexual partners. For

example, a large US representative cohort over a period of 15 years follow-up shown a significant linear relationship for females and males under 60 years of age at baseline relationship between alcohol consumption and all-cause mortality [4].

Risky behaviors are a multifactorial disease with severe health and social consequences. Interestingly, past studies showed that risky behaviors were clinically recognized as a feature of several psychiatric disorders, including Attention deficit and hyperactivity disorder (ADHD) [5], Novelty seeking (NS) [6]. Moreover, the relationship between brain disorders and risky behaviors can also be explained from the genetic domain. Multiple studies pointed out gene expression in the brain region, especially in the prefrontal cortex, basal ganglia and midbrain, was associated with risky behaviors [7]. Therefore, clarifying the genetic mechanism of risky behaviors has the potential to inform the biology of complex psychiatric disorders.

Apart from the environmental effects, there are indeed genetic factors affecting the risky behaviors. It is worth emphasizing that a latest research found that genetic correlations for risky behaviors were significantly higher than other phenotypes, and that many lead SNPs were shared among their phenotypes [8]. Overlapping genetic factors usually imply a sharing of possible etiological mechanisms, so integrating multiple risk traits could offer new insights for us. Cortes et al. clustered the genetic risk profiles of 3,025 genome-wide independent loci of 19,155 disease classification codes from 320,644 participants in the UK Biobank, and identified 339 distinct disease association profiles, and used multiple methods to association the clusters to potential biological pathways [9]. In addition, limited efforts were paid to explore the common genetic factors shared by various risky behaviors. Therefore, joint analysis of multiple risky behaviors has the potential to provide insight into the biological mechanism of risky behaviors.

In recent years, GWAS has identified many genetic variants associated with complex diseases and traits [10, 11]. A large part of significant loci identified by GWAS were in non-coding chromosomal regions [12]. Interestingly, previous studies demonstrated the important roles of regulatory genetic variants in the pathogenesis of complex diseases and traits, such as the expression quantitative trait locus (eQTLs) and methylation quantitative trait locus (meQTL) [13–17]. For instance, Huan et al. integrated whole blood meQTL and GWAS with disease-related variants, and illuminated the pathways for cardiovascular disease [17]. Through the comprehensive functional annotation of schizophrenia susceptibility SNPs identified by

GWAS, including eQTL and meQTL analysis, Niu et al. identified 447 target protein-coding genes [18]. Additionally, significant overlap between the SNPs identified by GWAS and regulatory genetic regions [19], functional SNPs located in protein-coding genes and non-coding genes, known as regulatory single nucleotide polymorphisms (rSNPs), playing a major and indirect role in regulating gene function [20]. Integrating GWAS data with rSNP and MeQTL had the potential to discover novel susceptibility genetic variants for human complex diseases [21, 22].

In this study, our goal was to explore the potential genetic factors shared by the four risky behaviors, especially those related to mental or psychological disorders. Accordingly, we first used MASH analysis to detect the common shared variants for the four common risky behaviors. Then the identified SNPs were mapped to target genes according to the genomic functional annotation data of eQTLs, MeQTLs and the SNPs near to known genes. Moreover, gene enrichment analysis was conducted to identify the significant pathways for risky behaviors.

## RESULTS

### Single gene analysis

For rSNP, we identified 4378 SNPs shared by the four risky behaviors, corresponding to 792 target regulatory genes, such as PLEKHM1, MAPT, APBB2, TFEC and KANSL1-AS1 (Supplementary Table 1). For MeQTL, 793 significant SNPs were detected, corresponding to 162 target genes, such as ADH1B, ADH1C, DCC, ARHGAP10 and ARHGAP27 (Supplementary Table 2). For the SNPs near to known genes, we detected 414 SNPs, corresponding to 89 target genes, such as PHC2, HYL1, PTPRF, ELAVL4, and LRP8 (Supplementary Table 3).

### Functional gene sets enrichment analysis

For rSNP, we detected 56 candidate GO terms, such as Gene silencing (adjusted P value=5.73×10<sup>-12</sup>), Negative regulation of gene expression (adjusted P=5.73×10<sup>-12</sup>) and Chromatin silencing (adjusted P=1.13×10<sup>-11</sup>). For MeQTL, we identified 79 candidate GO terms, such as Neurogenesis (adjusted P=9.84×10<sup>-3</sup>), Subpallium development (adjusted P=9.84×10<sup>-3</sup>) and Striatum development (adjusted P=2.17×10<sup>-2</sup>). For the SNPs mapping to known genes, 215 candidate GO terms were identified, such as Regulation of collateral sprouting (adjusted P=2.84×10<sup>-5</sup>), Regulation of extent of cell growth (adjusted P=2.84×10<sup>-5</sup>) and Regulation of neuron differentiation (adjusted P=2.84×10<sup>-5</sup>). Moreover, FUMA analysis also detected multiple

**Table 1. Gene set enrichment analysis results of rSNP related genes associated with the four risky behaviors.**

Gene Set	adjusted <i>P</i>
GO_GENE_SILENCING	5.73×10 <sup>-12</sup>
GO_NEGATIVE_REGULATION_OF_GENE_EXPRESSION_EPIGENETIC	5.73×10 <sup>-12</sup>
GO_CHROMATIN_SILENCING	1.13×10 <sup>-11</sup>
GO_CHROMATIN_ORGANIZATION	1.69×10 <sup>-08</sup>
GO_REGULATION_OF_GENE_EXPRESSION_EPIGENETIC	1.75×10 <sup>-08</sup>
GO_CHROMOSOME_ORGANIZATION	5.07×10 <sup>-08</sup>
GO_CHROMATIN_ASSEMBLY_OR_DISASSEMBLY	7.10×10 <sup>-07</sup>
GO_DNA_PACKAGING	7.10×10 <sup>-07</sup>
GO_PROTEIN_DNA_COMPLEX_SUBUNIT_ORGANIZATION	3.81×10 <sup>-06</sup>
GO_CHROMATIN_SILENCING_AT_RDNA	1.13×10 <sup>-04</sup>
Autism spectrum disorder or schizophrenia	1.81×10 <sup>-27</sup>
Alcohol use disorder (total score)	1.44×10 <sup>-12</sup>
Neuroticism	2.47×10 <sup>-12</sup>
General cognitive ability	2.47×10 <sup>-12</sup>
Intelligence (MTAG)	8.93×10 <sup>-09</sup>
Tacrolimus trough concentration in kidney transplant patients	1.22×10 <sup>-08</sup>
Facial emotion recognition (sad faces)	4.90×10 <sup>-07</sup>
Squamous cell lung carcinoma	3.10×10 <sup>-06</sup>
Cannabis use	3.28×10 <sup>-06</sup>
Lung cancer in ever smokers	4.03×10 <sup>-06</sup>

**Table 2. Gene set enrichment analysis results of MeQTL related genes associated with the four risky behaviors.**

Gene Set	adjusted <i>P</i>
GO_NEUROGENESIS	9.84×10 <sup>-03</sup>
GO_SUBPALLIUM_DEVELOPMENT	9.84×10 <sup>-03</sup>
GO_HEAD_DEVELOPMENT	9.84×10 <sup>-03</sup>
GO_CENTRAL_NERVOUS_SYSTEM_NEURON_DIFFERENTIATION	9.84×10 <sup>-03</sup>
GO_NEURON_DIFFERENTIATION	9.84×10 <sup>-03</sup>
GO_REGULATION_OF_NEURON_DIFFERENTIATION	9.84×10 <sup>-03</sup>
GO_ENDOCRINE_SYSTEM_DEVELOPMENT	9.84×10 <sup>-03</sup>
GO_REGULATION_OF_VOLTAGE_GATED_CALCIUM_CHANNEL_ACTIVITY	1.06×10 <sup>-02</sup>
GO_REGULATION_OF_NERVOUS_SYSTEM_DEVELOPMENT	1.13×10 <sup>-02</sup>
GO_DEVELOPMENTAL_INDUCTION	1.13×10 <sup>-02</sup>
General cognitive ability	2.23×10 <sup>-18</sup>
Intelligence (MTAG)	5.89×10 <sup>-15</sup>
Self-reported risk-taking behaviour	8.24×10 <sup>-13</sup>
Alcohol use disorder (total score)	5.11×10 <sup>-11</sup>
Hand grip strength	2.58×10 <sup>-10</sup>
Educational attainment	1.01×10 <sup>-08</sup>
Alcohol consumption in current drinkers	4.72×10 <sup>-08</sup>
Body mass index	1.24×10 <sup>-07</sup>
Cognitive ability (MTAG)	1.50×10 <sup>-07</sup>
Alcohol use disorder (consumption score)	1.96×10 <sup>-07</sup>

brain related diseases or traits enriched in the identified target genes of the four risky behaviors, such as Autism spectrum disorder (adjusted *P* value=1.81×10<sup>-27</sup>), Neuroticism (adjusted *P*=2.47×10<sup>-12</sup>) and Intelligence (adjusted *P*=5.89×10<sup>-15</sup>). More additional details were provided in Tables 1–3 and Supplementary Tables 4–6.

## DISCUSSION

Keller and Cannon’s watershed analogy of the genotype–phenotype relationship that “the more upstream the phenotype is, the closer its relationship to the genetic variants that affect it [23, 24]. Thereby, we conducted a large-scale genome-wide multiphenotypic

**Table 3. Gene set enrichment analysis results of genes with SNPs showing association with the four risky behaviors.**

Gene Set	adjusted <i>P</i>
GO_REGULATION_OF_COLLATERAL_SPROUTING	2.84×10 <sup>-05</sup>
GO_REGULATION_OF_EXTENT_OF_CELL_GROWTH	2.84×10 <sup>-05</sup>
GO_REGULATION_OF_NEURON_DIFFERENTIATION	2.84×10 <sup>-05</sup>
GO_NEGATIVE_REGULATION_OF_NITROGEN_COMPOUND_METABOLIC_PROCESS	2.84×10 <sup>-05</sup>
GO_REGULATION_OF_NERVOUS_SYSTEM_DEVELOPMENT	2.84×10 <sup>-05</sup>
GO_REGULATION_OF_CELL_DEVELOPMENT	1.03×10 <sup>-04</sup>
GO_NEUROGENESIS	1.03×10 <sup>-04</sup>
GO_NEGATIVE_REGULATION_OF_PROTEIN_OLIGOMERIZATION	1.03×10 <sup>-04</sup>
GO_REGULATION_OF_NEURON_PROJECTION_DEVELOPMENT	1.19×10 <sup>-04</sup>
GO_ETHANOL_METABOLIC_PROCESS	1.19×10 <sup>-04</sup>
Alcohol use disorder (total score)	2.77×10 <sup>-08</sup>
Cognitive ability (MTAG)	2.77×10 <sup>-08</sup>
General cognitive ability	2.77×10 <sup>-08</sup>
Alcohol use disorder (consumption score)	7.78×10 <sup>-08</sup>
Intelligence (MTAG)	3.27×10 <sup>-07</sup>
Alzheimer's disease in APOE e4- carriers	3.11×10 <sup>-06</sup>
Cognitive ability	3.17×10 <sup>-06</sup>
Hand grip strength	4.46×10 <sup>-06</sup>
Parental longevity (father's attained age)	5.35×10 <sup>-06</sup>
Autism spectrum disorder or schizophrenia	1.35×10 <sup>-05</sup>

integrative analysis for four common risky behaviors. We reported multiple candidate genes and GO terms shared by the four risky behaviors. Our study had the potential to explore the genetic mechanism underlying risky behaviors and further elucidate the biological mechanisms of the development of mental disorders.

Multiphenotypic analysis identified several common genes associated with four risky behaviors, such as APBB2, MAPT and DCC. APBB2, the protein encoded by this gene interacts with the cytoplasmic domains of amyloid beta (A4) precursor protein and amyloid beta (A4) precursor-like protein 2. This protein contains two phosphotyrosine binding (PTB) domains that are thought to play a role in signal transduction [25]. Polymorphisms in this gene have been associated with Alzheimer's disease [26]. Another interesting new finding was that they used PCR-RFLP to analyze the substitution of hCV1558625 (rs13133980) and rs13133980, and discovered that hcv1558625-rs13133980 AG haplotype increased the relative risk of severe cognitive impairment in centenarians [27]. In addition, the APBB2 rs13133980 G allele was more highly expressed in centenarians with severe cognitive impairment than in the individuals without cognitive impairment [27].

MAPT, a microtubule-related protein, whose transcript undergoes complex, regulated selective splicing,

producing a variety of mRNA species. MAPT transcription products are expressed differently in the nervous system, depending on the stage of neuron maturation and the type of neuron [25]. Mutations in the MAPT gene are associated with a variety of neurodegenerative diseases, such as Alzheimer's disease [26], frontotemporal dementia, and cortical basal cell degeneration [25]. By analyzing MAPT region expression, splicing and regulation in 2011 brain samples from 439 individuals, a survey found that regional differences in MAPT mRNA expression and splicing in human brain were highly correlated with the total expression level of tau protein [28]. Furthermore, they hypothesized that genetic risk factors for neurodegenerative disease at the MAPT site may play a role by altering mRNA splicing in different areas of the brain, rather than by the overall expression of the MAPT gene [28].

DCC encodes the netrin 1 receptor. Transmembrane proteins are members of the immunoglobulin superfamily of cell adhesion molecules that mediate neuronal growth cones and are the source of axon-directed netrin 1 ligands [25]. Variations in DCC may determine differential predisposition to mPFC disorders in humans, clarified by Manitt et al. [29] Their results found that DCC expression was elevated in the brains of antidepressant-free subjects who committed suicide.

Horn K et al. [30] found that the dendritic spines of the pyramidal neurons of wild-type mice were rich in DCC, and then demonstrated that selective deletion of DCC in the brain neurons in the adult forebrain resulted in long-term potential loss (LTP), complete long-term depression, short dendritic spines, and impaired spatial and recognition memory, through the DCC knockout mice experiment.

GO enrichment analysis identified multiple significant GO terms, which were involved in the development of the brain and mental, such as neurogenesis, regulation of neuron differentiation and striatum development. Neurogenesis, or called neural cell differentiation, a process of producing functional neurons from adult neural precursors throughout life in a limited number of mammalian brain regions [31]. Deng et al. suggested that an important role for adult hippocampal neurogenesis is learning and memory [32]. According to Kang et al.'s research, dysregulation in adult neurogenesis are implicated in psychiatric diseases in humans, such as affective disorders, schizophrenia, and drug addiction [33].

Regulation of neuron differentiation, which is defined as any process, modulates the frequency, rate or extent of neuron differentiation. A previous study combined the genome-wide disease risk profile of GWAS with the longitudinal *in vitro* gene expression profile of human neuronal differentiation using an analytical framework they developed to demonstrate that the cumulative impact of risk loci for specific psychiatric disorders is significantly correlated with genes that are differentially expressed during neuronal differentiation [34].

Striatum development refers to the process from initial formation to maturation of the striatum. The striatum is a region of the forebrain that consists of the caudate nucleus, putamen nucleus, and striatum base. Similarly, studies linked striatum development to mental illness, such as delayed development of the ventral striatum during adolescence reflects emotional neglect and predicts depressive symptoms [35].

Additionally, we detected multiple brain related diseases or traits enriched in the identified target genes of the four risky behaviors, such as SCZ and intelligence. SCZ is an idiopathic mental disorder with high heritability. Previous studies showed that schizophrenia is associated with neural calcium channels, which are one of the pathogenesis of bipolar disorder and autism [36]. Sullivan et al.'s study described the strong effects of eight rare copy number variants on schizophrenia, and these associations might also be associated with autism, mental retardation or

epilepsy, which are usually not disease-specific [37]. In addition, analysis of the genetic characteristics found a significant relationship between intelligence and changes in the expression of the brain and pituitary gland. It also suggested that neurogenesis was the process by which new neurons are created, a process previously associated with human intelligence using GWAS data [38]. Interestingly, combined with our findings and previous research, it seemed that many psychiatric disorders also have codependent biological mechanisms that rely on similar gene regulation or expression, or are partly due to common genetic influence. Our study confirmed this finding and provided some assistance in further elucidating the biological mechanism of risky behavior.

One strengths of this study is the combination of multiple risky behaviors and integrating it with genomic functional annotation of regulatory genetic variants. Complex disorders or traits are usually regulated by gene expression, methylation, microRNAs, and epigenomics as a whole. Due to some limitation of GWAS method [12] and the important role of regulatory genetic variants [13–17, 19], integrating GWAS and regulatory genetic variants data can discover novel candidate genes for mental disorders in this study. Meanwhile, GWAS of common diseases have revealed a wide range of pleiotropic, resulting in significant genetic correlations among different traits [9]. For instance, Cross-disorder Group of the Psychiatric Genomics Consortium observed the genetic correlation among five mental disorders based on genome-wide SNPs data [39]. Overlapping genetic risks imply a sharing of possible etiological mechanisms. Ellinghaus et al.'s analysis of five chronic inflammatory diseases identified 27 new associations and highlighted disease-specific patterns in shared loci [40]. Furthermore, previous studies provided evidence that different risky behaviors occurred simultaneously with the same mental disease [5–7]. Therefore, joint analysis of multiple risky behaviors has the potential to provide insight into the biological mechanism of risky behaviors. Our study identified multiple risk behaviors associated genes, which have been suggested to be involved in the development of mental disorders, such as APBB2 [26] and DCC [30]. Further GO analysis identified multiple risk behavior related GO terms, functionally involved in brain development and mental disorders, such as neurogenesis, regulation of neuron differentiation and striatum development.

In our study, we used MASH analysis to explore the common genetic factors shared by the four common risky behaviors. MASH was developed based on recent methods [41, 42], and it combines the advantages of existing methods while overcomes the major limitations

of them. MASH gained more power compared to a tissue-by-tissue analysis and ANOVA or simple linear regression. The most important feature of MASH is that it facilitates more estimation and assessment of effect-size heterogeneity than simple “shared/condition-specific” assessments [41]. In particular, MASH is generic and adaptive. It is generic in that it can take as input any matrix of Z scores (or, better, a matrix of effect estimates and their corresponding standard errors) testing many effects in many conditions. And MASH is adaptive in that it learns patterns of sharing of multivariate effects from the data, allowing it to maximize power and precision for each setting. Urbut et al. [41] conducted a detailed analysis of locally-acting (“cis”) eQTLs in 44 human tissues through MASH, and compared the performance of MASH with that of ASH (a univariate shrinkage procedure) [43] and BMATILE (a multivariate model) [42]. It turns out that MASH outperformed other methods, particularly in the shared and structured effects scenario. However, it should be noted that MASH does not distinguish the causal associations and those caused by linkage disequilibrium (LD). When jointly analyzing GWAS and eQTLs data, a SNP identified by GWAS may be a significant eQTL simply because it is in LD with another causal SNP.

Additionally, there are some limitations of this study. First, due to the inclusion criteria of risky behaviors from our GAWs data sources [8], we analyzed four common risk behaviors in this study. The four risky behaviors have been proved to be related to psychological or mental disorders [5–7, 44–46]. Furthermore, they have been demonstrated to be a powerful predictor for injury, risk tolerance and other problems [1, 3]. However, it is worthy to jointly analyze more risky behaviors in future studies. Second, our analysis only included the genetically regulated portion of gene expression. Therefore, it could not capture or interpret the variance of expression caused by environment factors, which may also contribute to development of psychiatric disorders. Third, we analyzed the genetic data from European cohorts in this study. Therefore, it should be carefully to apply our results to other populations.

In conclusion, multiple genes and GO terms shared by the four risky behaviors were reported in our study. And the results supported the functional relevance of brain development with risky behaviors from the genetic domain, which can offer some help to further elucidate the biological mechanisms of the relationship. Due to many mental disorders also had mutual biological mechanism reliant on the similar genetic regulation or expression, our results may do a lot help to construct better multi-gene scores to measure environmental, demographic, and genetic factors interacting or building neurological scores.

## MATERIALS AND METHODS

### GWAS data of risky behaviors

In this study, we analyzed four common risky behaviors, including automobile speeding, alcohol drinking, smoking, and multiple sexual partners. Automobile speeding, alcohol drinking, smoking, and multiple sexual partners are major common risky behaviors, which have been demonstrated to be a powerful predictor for injury, risk tolerance and other problems [2, 3]. The GAWs data of the four risky behaviors were obtained from the published study [8]. It consisted of over 1 million European-ancestry participants for the four risky behaviors, including 404,291 for automobile speeding propensity, 414,343 for drinks per week, 370,711 for number of sexual partners and 518,633 for smoker, respectively. Genotyping was performed by using a range of commercially available genotyping arrays, such as the UK BiLEVE array [47] and the UK Biobank Axiom array [48]. Extensive quality-control procedures were applied to the cohort-level summary statistics, including the EasyQC software developed by the GIANT consortium [49], and only SNPs with minor allele frequency (MAF) greater than 0.001 were analyzed. IMPUTE4 was applied for genotype imputation [48]. The top ten (or more) principal components of the genetic relatedness matrix, sex and birth year were controlled during the GWAS. At last, each behavior consisted of approximately 11,515,000 SNPs in this study. More detailed information about cohorts, inclusion criteria for risky behaviors, genotyping and imputation can be found in the published study [8].

### Multiple traits integrative analysis

The MASH analysis was applied to the GWAS datasets of the four risky behaviors to detect the SNPs associated with all of the four risky behaviors. MASH (<https://github.com/stephenslab/mashr>) can estimate and test multiple effects under multiple conditions. The approach improves on existing methods to allow arbitrary correlation of effect size between conditions and improves effect size assessment, which is helpful for more quantitative assessment of effect size heterogeneity [41]. The SNPs associated with the four risky behaviors were detected by MASH and then mapped to target genes according to the genomic annotation data of rSNP-target genes, MeQTL-target genes and the SNPs near to known genes, respectively. We used “canonical” type to set up covariance matrix and then to fit the model. The sharing important signals between each pair of conditions were selected. And the default definition

for sharing from MASH software is "the same sign and within a factor 0.5 of each other".

### Annotation data of regulatory SNPs, MeQTL and the SNPs mapping to near genes

The genomic annotation data of rSNPs was obtained from the rSNPBase 3.1 database (<http://rsnp3.psych.ac.cn/>). rSNPBase provides genomic annotation of SNP-related regulatory element-target gene pairs [50]. Currently, rSNPBase database contains nearly 119,630,196 rSNP annotation entries on SNP regulatory information. Genomic similarities or widely used reference databases were used to analyze functional associations between regulatory elements and target genes [51]. The MeQTLs–target gene annotation data were collected from published studies [52]. In short, about 4.5 million loci were measured in 697 subjects. SNPs were genotyped using Affymetrix genome-wide SNP Arrays 5.0 or 6.0 or Illumina OmniExpress. The Minimac method was used to input 1000 genomic reference versions of v3 with the lower MAF>0.05 and  $r^2 > 0.5$  as thresholds. 4,761,800 SNPs were identified for MeQTL correlation after quality control. For mapping SNPs to near genes, a physical short of 500 kb was to link a SNP and a gene, because most enhancers and repressors are < 500 kb away from genes, and most linkage disequilibrium blocks are < 500 kb.

### Functional gene sets enrichment analysis

The identified candidate genes shared by the four risky behaviors were subjected to gene set enrichment analysis, implemented by the GENE2FUNC of the FUMA tool [20]. FUMA [20] is a platform for annotating, sorting, visualizing, and interpreting GWAS results. For every input gene, GENE2FUNC provides information about tissue specificity, the enrichment of publicly available gene sets, and the expression of different tissue types. The genes were tested for representation in different functional gene sets, including GO terms and various diseases or traits related gene sets. The Benjamini-Hochberg false discovery rate (FDR) was recommended by the FUMA software [20], and used for controlling the potential impact of multiple testing problem in this study. The adjusted  $P$  value cut-off was 0.05 and minimum overlapping genes with gene-sets was assigned 2 during the FUMA analysis. Additionally, the ensemble version v92 and GTEx v8 were chosen for FUMA analysis.

### CONFLICTS OF INTEREST

There are no conflicts of interest to declare.

### FUNDING

This work was supported by the National Natural Scientific Foundation of China [81673112]; the Key projects of international cooperation among governments in scientific and technological innovation [2016YFE0119100]; the Natural Science Basic Research Plan in Shaanxi Province of China [2017JZ024]; and the Fundamental Research Funds for the Central Universities.

### REFERENCES

1. Turner C, McClure R, Pirozzo S. Injury and risk-taking behavior—a systematic review. *Accid Anal Prev*. 2004; 36:93–101.  
[https://doi.org/10.1016/S0001-4575\(02\)00131-8](https://doi.org/10.1016/S0001-4575(02)00131-8)  
PMID:[14572831](https://pubmed.ncbi.nlm.nih.gov/14572831/)
2. Favara M, Sanchez A. Psychosocial competencies and risky behaviours in Peru. *IZA Journal of Labor and Development*. 2017; 6.  
<https://doi.org/10.1186/s40175-016-0069-3>
3. Galizzi MM. Label, nudge or tax? A review of health policies for risky behaviours. *J Public Health Res*. 2012; 1:14–21.  
<https://doi.org/10.4081/jphr.2012.e5> PMID:[25170442](https://pubmed.ncbi.nlm.nih.gov/25170442/)
4. Rehm J, Sempos CT. Alcohol consumption and all-cause mortality. *Addiction*. 1995; 90:471–80.  
<https://doi.org/10.1111/j.1360-0443.1995.tb02177.x>  
PMID:[7773106](https://pubmed.ncbi.nlm.nih.gov/7773106/)
5. Shoham R, Sonuga-Barke EJ, Aloni H, Yaniv I, Pollak Y. ADHD-associated risk taking is linked to exaggerated views of the benefits of positive outcomes. *Sci Rep*. 2016; 6:34833.  
<https://doi.org/10.1038/srep34833> PMID:[27725684](https://pubmed.ncbi.nlm.nih.gov/27725684/)
6. Wang Y, Liu Y, Yang L, Gu F, Li X, Zha R, Wei Z, Pei Y, Zhang P, Zhou Y, Zhang X. Novelty seeking is related to individual risk preference and brain activation associated with risk prediction during decision making. *Sci Rep*. 2015; 5:10534.  
<https://doi.org/10.1038/srep10534> PMID:[26065910](https://pubmed.ncbi.nlm.nih.gov/26065910/)
7. Clifton EA, Perry JR, Imamura F, Lotta LA, Brage S, Forouhi NG, Griffin SJ, Wareham NJ, Ong KK, Day FR. Genome-wide association study for risk taking propensity indicates shared pathways with body mass index. *Commun Biol*. 2018; 1:36.  
<https://doi.org/10.1038/s42003-018-0042-6>  
PMID:[30271922](https://pubmed.ncbi.nlm.nih.gov/30271922/)
8. Karlsson Linnér R, Biroli P, Kong E, Meddens SF, Wedow R, Fontana MA, Lebreton M, Tino SP, Abdellaoui A, Hammerschlag AR, Nivard MG, Okbay A, Rietveld CA, et al, and 23andMe Research Team, and eQTLgen Consortium, and International Cannabis

- Consortium, and Social Science Genetic Association Consortium. Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. *Nat Genet.* 2019; 51:245–57.  
<https://doi.org/10.1038/s41588-018-0309-3>  
PMID:30643258
9. Cortes A, Albers PK, Dendrou CA, Fugger L, McVean G. Identifying cross-disease components of genetic risk across hospital data in the UK Biobank. *Nat Genet.* 2020; 52:126–134.  
<https://doi.org/10.1038/s41588-019-0550-4>  
PMID:31873298
10. Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetsky V, Mattheisen M, Wang Y, Coleman JR, Gaspar HA, de Leeuw CA, Steinberg S, Pavlides JM, et al, and eQTLGen Consortium, and BIOS Consortium, and Bipolar Disorder Working Group of the Psychiatric Genomics Consortium. Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet.* 2019; 51:793–803.  
<https://doi.org/10.1038/s41588-019-0397-8>  
PMID:31043756
11. Tin A, Marten J, Halperin Kuhns VL, Li Y, Wuttke M, Kirsten H, Sieber KB, Qiu C, Gorski M, Yu Z, Giri A, Sveinbjornsson G, Li M, et al, and German Chronic Kidney Disease Study, and Lifelines Cohort Study, and V. A. Million Veteran Program. Target genes, variants, tissues and transcriptional pathways influencing human serum urate levels. *Nat Genet.* 2019; 51:1459–74.  
<https://doi.org/10.1038/s41588-019-0504-x>  
PMID:31578528
12. Hill WD, Davies NM, Ritchie SJ, Skene NG, Bryois J, Bell S, Di Angelantonio E, Roberts DJ, Xueyi S, Davies G, Liewald DC, Porteous DJ, Hayward C, et al. Genome-wide analysis identifies molecular systems and 149 genetic loci associated with income. *Nat Commun.* 2019; 10:5741–5741.  
<https://doi.org/10.1038/s41467-019-13585-5>  
PMID:31844048
13. Grubert F, Zaugg JB, Kasowski M, Ursu O, Spacek DV, Martin AR, Greenside P, Srivas R, Phanstiel DH, Pekowska A, Heidari N, Euskirchen G, Huber W, et al. Genetic Control of Chromatin States in Humans Involves Local and Distal Chromosomal Interactions. *Cell.* 2015; 162:1051–65.  
<https://doi.org/10.1016/j.cell.2015.07.048>  
PMID:26300125
14. Kou Y, Koag MC, Lee S. N7 methylation alters hydrogen-bonding patterns of guanine in duplex DNA. *J Am Chem Soc.* 2015; 137:14067–70.  
<https://doi.org/10.1021/jacs.5b10172> PMID:26517568
15. Kou Y, Koag MC, Lee S. Structural and Kinetic Studies of the Effect of Guanine N7 Alkylation and Metal Cofactors on DNA Replication. *Biochemistry.* 2018; 57:5105–16.  
<https://doi.org/10.1021/acs.biochem.8b00331>  
PMID:29957995
16. Strunz T, Grassmann F, Gayán J, Nahkuri S, Souza-Costa D, Maugeais C, Fauser S, Nogoceke E, Weber BH. A mega-analysis of expression quantitative trait loci (eQTL) provides insight into the regulatory architecture of gene expression variation in liver. *Sci Rep.* 2018; 8:5865.  
<https://doi.org/10.1038/s41598-018-24219-z>  
PMID:29650998
17. Huan T, Joehanes R, Song C, Peng F, Guo Y, Mendelson M, Yao C, Liu C, Ma J, Richard M, Agha G, Guan W, Almlí LM, et al. Genome-wide identification of DNA methylation QTLs in whole blood highlights pathways for cardiovascular disease. *Nat Commun.* 2019; 10:4267.  
<https://doi.org/10.1038/s41467-019-12228-z>  
PMID:31537805
18. Niu HM, Yang P, Chen HH, Hao RH, Dong SS, Yao S, Chen XF, Yan H, Zhang YJ, Chen YX, Jiang F, Yang TL, Guo Y. Comprehensive functional annotation of susceptibility SNPs prioritized 10 genes for schizophrenia. *Transl Psychiatry.* 2019; 9:56.  
<https://doi.org/10.1038/s41398-019-0398-5>  
PMID:30705251
19. Schaub MA, Boyle AP, Kundaje A, Batzoglou S, Snyder M. Linking disease associations with regulatory information in the human genome. *Genome Res.* 2012; 22:1748–59.  
<https://doi.org/10.1101/gr.136127.111>  
PMID:22955986
20. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun.* 2017; 8:1826.  
<https://doi.org/10.1038/s41467-017-01261-5>  
PMID:29184056
21. Macintyre G, Bailey J, Haviv I, Kowalczyk A. is-rSNP: a novel technique for in silico regulatory SNP detection. *Bioinformatics.* 2010; 26:i524–30.  
<https://doi.org/10.1093/bioinformatics/btq378>  
PMID:20823317
22. Gu R, Xu J, Lin Y, Sheng W, Ma D, Ma X, Huang G. The role of histone modification and a regulatory single-nucleotide polymorphism (rs2071166) in the Cx43 promoter in patients with TOF. *Sci Rep.* 2017; 7:10435.  
<https://doi.org/10.1038/s41598-017-10756-6>  
PMID:28874875
23. Keller MC, Miller G. Resolving the paradox of common, harmful, heritable mental disorders: which

- evolutionary genetic models work best? *Behav Brain Sci.* 2006; 29:385–404; discussion 405–52.  
<https://doi.org/10.1017/S0140525X06009095>  
PMID:[17094843](https://pubmed.ncbi.nlm.nih.gov/17094843/)
24. Ormel J, Hartman CA, Snieder H. The genetics of depression: successful genome-wide association studies introduce new challenges. *Transl Psychiatry.* 2019; 9:114.  
<https://doi.org/10.1038/s41398-019-0450-5>  
PMID:[30877272](https://pubmed.ncbi.nlm.nih.gov/30877272/)
25. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, Franz M, Grouios C, Kazi F, Lopes CT, Maitland A, Mostafavi S, Montojo J, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* 2010 (suppl\_2); 38:W214–20.  
<https://doi.org/10.1093/nar/gkq537> PMID:[20576703](https://pubmed.ncbi.nlm.nih.gov/20576703/)
26. Kakati T, Kashyap H, Bhattacharyya DK. THD-Module Extractor: An Application for CEN Module Extraction and Interesting Gene Identification for Alzheimer's Disease. *Sci Rep.* 2016; 6:38046.  
<https://doi.org/10.1038/srep38046> PMID:[27901073](https://pubmed.ncbi.nlm.nih.gov/27901073/)
27. Golanska E, Sieruta M, Gresner SM, Pfeffer A, Chodakowska-Zebrowska M, Sobow TM, Klich I, Mossakowska M, Szybinska A, Barcikowska M, Liberski PP. APBB2 genetic polymorphisms are associated with severe cognitive impairment in centenarians. *Exp Gerontol.* 2013; 48:391–94.  
<https://doi.org/10.1016/j.exger.2013.01.013>  
PMID:[23384821](https://pubmed.ncbi.nlm.nih.gov/23384821/)
28. Trabzuni D, Wray S, Vandrovcova J, Ramasamy A, Walker R, Smith C, Luk C, Gibbs JR, Dillman A, Hernandez DG, Arepalli S, Singleton AB, Cookson MR, et al. MAPT expression and splicing is differentially regulated by brain region: relation to genotype and implication for tauopathies. *Hum Mol Genet.* 2012; 21:4094–103.  
<https://doi.org/10.1093/hmg/dds238>  
PMID:[22723018](https://pubmed.ncbi.nlm.nih.gov/22723018/)
29. Manitt C, Eng C, Pokinko M, Ryan RT, Torres-Berrío A, Lopez JP, Yogendran SV, Daubaras MJ, Grant A, Schmidt ER, Tronche F, Krimpenfort P, Cooper HM, et al. dcc orchestrates the development of the prefrontal cortex during adolescence and is altered in psychiatric patients. *Transl Psychiatry.* 2013; 3:e338.  
<https://doi.org/10.1038/tp.2013.105> PMID:[24346136](https://pubmed.ncbi.nlm.nih.gov/24346136/)
30. Horn KE, Glasgow SD, Gobert D, Bull SJ, Luk T, Girgis J, Tremblay ME, McEachern D, Bouchard JF, Haber M, Hamel E, Krimpenfort P, Murai KK, et al. DCC expression by neurons regulates synaptic plasticity in the adult brain. *Cell Rep.* 2013; 3:173–85.  
<https://doi.org/10.1016/j.celrep.2012.12.005>  
PMID:[23291093](https://pubmed.ncbi.nlm.nih.gov/23291093/)
31. Ming GL, Song H. Adult neurogenesis in the mammalian brain: significant answers and significant questions. *Neuron.* 2011; 70:687–702.  
<https://doi.org/10.1016/j.neuron.2011.05.001>  
PMID:[21609825](https://pubmed.ncbi.nlm.nih.gov/21609825/)
32. Deng W, Aimone JB, Gage FH. New neurons and new memories: how does adult hippocampal neurogenesis affect learning and memory? *Nat Rev Neurosci.* 2010; 11:339–50.  
<https://doi.org/10.1038/nrn2822> PMID:[20354534](https://pubmed.ncbi.nlm.nih.gov/20354534/)
33. Kang E, Wen Z, Song H, Christian KM, Ming GL. Adult Neurogenesis and Psychiatric Disorders. *Cold Spring Harb Perspect Biol.* 2016; 8:a019026.  
<https://doi.org/10.1101/cshperspect.a019026>  
PMID:[26801682](https://pubmed.ncbi.nlm.nih.gov/26801682/)
34. Ori APS, Bot MHM, Molenhuis RT, Olde Loohuis LM, Ophoff RA. A Longitudinal Model of Human Neuronal Differentiation for Functional Investigation of Schizophrenia Polygenic Risk. *Biol Psychiatry.* 2019; 85:544–53.  
<https://doi.org/10.1016/j.biopsych.2018.08.019>  
PMID:[30340753](https://pubmed.ncbi.nlm.nih.gov/30340753/)
35. Hanson JL, Hariri AR, Williamson DE. Blunted Ventral Striatum Development in Adolescence Reflects Emotional Neglect and Predicts Depressive Symptoms. *Biol Psychiatry.* 2015; 78:598–605.  
<https://doi.org/10.1016/j.biopsych.2015.05.010>  
PMID:[26092778](https://pubmed.ncbi.nlm.nih.gov/26092778/)
36. 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](https://pubmed.ncbi.nlm.nih.gov/23974872/)
37. Sullivan PF, Daly MJ, O'Donovan M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat Rev Genet.* 2012; 13:537–51.  
<https://doi.org/10.1038/nrg3240>  
PMID:[22777127](https://pubmed.ncbi.nlm.nih.gov/22777127/)
38. Hill WD, Marioni RE, Maghziyan O, Ritchie SJ, Hagenaars SP, McIntosh AM, Gale CR, Davies G, Deary IJ. A combined analysis of genetically correlated traits identifies 187 loci and a role for neurogenesis and myelination in intelligence. *Mol Psychiatry.* 2019; 24:169–81.

- <https://doi.org/10.1038/s41380-017-0001-5>  
PMID:[29326435](https://pubmed.ncbi.nlm.nih.gov/29326435/)
39. Lee SH, Ripke S, Neale BM, Faraone SV, Purcell SM, Perlis RH, Mowry BJ, Thapar A, Goddard ME, Witte JS, Absher D, Agartz I, Akil H, et al, and International Inflammatory Bowel Disease Genetics Consortium (IIBDGC). Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet.* 2013; 45:984–94.  
<https://doi.org/10.1038/ng.2711> PMID:[23933821](https://pubmed.ncbi.nlm.nih.gov/23933821/)
40. Ellinghaus D, Jostins L, Spain SL, Cortes A, Bethune J, Han B, Park YR, Raychaudhuri S, Pouget JG, Hübenthal M, Folseraas T, Wang Y, Esko T, et al, and International IBD Genetics Consortium (IIBDGC), and International Genetics of Ankylosing Spondylitis Consortium (IGAS), and International PSC Study Group (IPSCSG), and Genetic Analysis of Psoriasis Consortium (GAPC), and Psoriasis Association Genetics Extension (PAGE). Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. *Nat Genet.* 2016; 48:510–18.  
<https://doi.org/10.1038/ng.3528> PMID:[26974007](https://pubmed.ncbi.nlm.nih.gov/26974007/)
41. Urbut SM, Wang G, Carbonetto P, Stephens M. Flexible statistical methods for estimating and testing effects in genomic studies with multiple conditions. *Nat Genet.* 2019; 51:187–95.  
<https://doi.org/10.1038/s41588-018-0268-8>  
PMID:[30478440](https://pubmed.ncbi.nlm.nih.gov/30478440/)
42. Flutre T, Wen X, Pritchard J, Stephens M. A statistical framework for joint eQTL analysis in multiple tissues. *PLoS Genet.* 2013; 9:e1003486.  
<https://doi.org/10.1371/journal.pgen.1003486>  
PMID:[23671422](https://pubmed.ncbi.nlm.nih.gov/23671422/)
43. Stephens M. False discovery rates: a new deal. *Biostatistics.* 2017; 18:275–94.  
<https://doi.org/10.1093/biostatistics/kxw041>  
PMID:[27756721](https://pubmed.ncbi.nlm.nih.gov/27756721/)
44. Benowitz NL. Pharmacology of nicotine: addiction, smoking-induced disease, and therapeutics. *Annu Rev Pharmacol Toxicol.* 2009; 49:57–71.  
<https://doi.org/10.1146/annurev.pharmtox.48.113006.094742> PMID:[18834313](https://pubmed.ncbi.nlm.nih.gov/18834313/)
45. Room R, Babor T, Rehm J. Alcohol and public health. *Lancet.* 2005; 365:519–30.  
[https://doi.org/10.1016/S0140-6736\(05\)17870-2](https://doi.org/10.1016/S0140-6736(05)17870-2)  
PMID:[15705462](https://pubmed.ncbi.nlm.nih.gov/15705462/)
46. Tsutsumi A, Izutsu T, Matsumoto T. Risky sexual behaviors, mental health, and history of childhood abuse among adolescents. *Asian J Psychiatr.* 2012; 5:48–52.  
<https://doi.org/10.1016/j.ajp.2011.12.004>  
PMID:[26878947](https://pubmed.ncbi.nlm.nih.gov/26878947/)
47. Wain LV, Shrine N, Miller S, Jackson VE, Ntalla I, Soler Artigas M, Billington CK, Kheirallah AK, Allen R, Cook JP, Probert K, Obeidat M, Bossé Y, et al, and OxGSK Consortium. Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study in UK Biobank. *Lancet Respir Med.* 2015; 3:769–81.  
[https://doi.org/10.1016/S2213-2600\(15\)00283-0](https://doi.org/10.1016/S2213-2600(15)00283-0)  
PMID:[26423011](https://pubmed.ncbi.nlm.nih.gov/26423011/)
48. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O’Connell J, Cortes A, Welsh S, McVean G, et al. Genome-wide genetic data on ~500,000 UK Biobank participants. *bioRxiv.* 2017.  
<https://doi.org/10.1101/166298>
49. Winkler TW, Day FR, Croteau-Chonka DC, Wood AR, Locke AE, Mägi R, Ferreira T, Fall T, Graff M, Justice AE, Luan J, Gustafsson S, Randall JC, et al, and Genetic Investigation of Anthropometric Traits (GIANT) Consortium. Quality control and conduct of genome-wide association meta-analyses. *Nat Protoc.* 2014; 9:1192–212.  
<https://doi.org/10.1038/nprot.2014.071>  
PMID:[24762786](https://pubmed.ncbi.nlm.nih.gov/24762786/)
50. Guo L, Wang J. rSNPBase 3.0: an updated database of SNP-related regulatory elements, element-gene pairs and SNP-based gene regulatory networks. *Nucleic Acids Res.* 2018; 46:D1111–16.  
<https://doi.org/10.1093/nar/gkx1101> PMID:[29140525](https://pubmed.ncbi.nlm.nih.gov/29140525/)
51. Aken BL, Ayling S, Barrell D, Clarke L, Curwen V, Fairley S, Fernandez Banet J, Billis K, García Girón C, Hourlier T, Howe K, Kähäri A, Kokocinski F, et al. The Ensembl gene annotation system. *Database (Oxford).* 2016; 2016:baw093.  
<https://doi.org/10.1093/database/baw093>  
PMID:[27337980](https://pubmed.ncbi.nlm.nih.gov/27337980/)
52. McClay JL, Shabalin AA, Dozmorov MG, Adkins DE, Kumar G, Nerella S, Clark SL, Bergen SE, Hultman CM, Magnusson PK, Sullivan PF, Aberg KA, van den Oord EJ, and Swedish Schizophrenia Consortium. High density methylation QTL analysis in human blood via next-generation sequencing of the methylated genomic DNA fraction. *Genome Biol.* 2015; 16:291.  
<https://doi.org/10.1186/s13059-015-0842-7>  
PMID:[26699738](https://pubmed.ncbi.nlm.nih.gov/26699738/)

## **SUPPLEMENTARY MATERIALS**

Please browse the Full text version to see the data of Supplementary Tables 1-6.

**Supplementary Table 1. List of the four risky behaviors associated rSNPs and corresponding genes.**

**Supplementary Table 2. List of the four risky behaviors associated MeQTL SNPs and corresponding genes.**

**Supplementary Table 3. List of the four risky behaviors associated SNPs and adjacent genes.**

**Supplementary Table 4. Gene set enrichment analysis results of rSNP related genes associated with the four risky behaviors.**

**Supplementary Table 5. Gene set enrichment analysis results of MeQTL related genes associated with the four risky behaviors.**

**Supplementary Table 6. Gene set enrichment analysis results of genes with SNPs showing association with the four risky behaviors.**