Joint Biotechnology Master Program

Palestine Polytechnic University

Deanship of Graduate Studies and Scientific Research

Bethlehem University

Faculty of Applied Sciences, Technology & Engineering

Shared Genetics Between Schizophrenia and Cardiometabolic Diseases: A Large-Scale Genome-Wide Cross-Trait Analysis

By

Tasnim Ahmad Mohammad Salameen

In Partial Fulfillment of the Requirements for the Degree

Master of Science

February 2026

The undersigned hereby certify that they have read and recommend to the Faculty of Scientific Research and Higher Studies at the Palestine Polytechnic University and the Faculty of Applied Sciences, Technology and Engineering at Bethlehem University for acceptance a thesis entitled:

Shared Genetics Between Schizophrenia and Cardiometabolic Diseases: A Large-Scale Genome-Wide Cross-Trait Analysis

by

Tasnim Ahmad Mohammad Salameen

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Biotechnology.

Graduate Advisory Committee:

Committee Member(Student's Supervisor)	Date
Dr. Yaqoub Ashhab – Palestine Polytechnic University	
Committee Member (Internal Examiner)	Date
Dr. Fouad Zahdeh	
Committee Member (External Examiner)	Date
Dr. Robin Abu Ghazaleh	

Approved for the Faculties

Dean of Graduate Studies and Scientific Research	Dean of Graduate Studies and Scientific Research
Palestine Polytechnic University	Bethlehem University
Date	Date

Abstract

Schizophrenia (SCZ) is a chronic mental disease, with a worldwide prevalence of about 1%. While epidemiological studies demonstrate the comorbidity of SCZ with cardiometabolic traits, the genetic basis, causality of association, and pathophysiology of this comorbidity are poorly understood.

Thus, this study intended to explore shared genetics between SCZ and cardiometabolic traits. Utilizing genome-wide association study (GWAS) summary data with sample sizes ranging from 50,364 to 898,130 of European ancestry, linkage-disequilibrium logistic regression (LDSC) was used to assess the heritability and the genetic correlation of SCZ with cardiometabolic traits, and a large-scale genome-wide cross-trait meta-analysis was conducted to investigate the genetic overlap between SCZ and cardiometabolic disorders. Multi-trait fine mapping and colocalization analyses were conducted to identify putative causative and shared causal variants, respectively. MAGMA gene-level analysis was integrated to identify significantly associated genes. Enrichment analyses were used to identify differential tissue expression and major shared pathways. Finally, the generalized summary statistic-based Mendelian randomization (GSMR) was applied to assess causation.

LDSC revealed a significant negative genetic correlation of SCZ with type 2 diabetes, cardiometabolic multimorbidity, obesity class 1, and obesity class 2. Cross-trait meta-analysis identified 124 independent SNPs, of which 7 SNPs are novel. Multi-trait fine mapping identified 90 putative causative SNPs, of which 15 SNPs are novel, and colocalization analysis revealed 9 shared causative SNPs between SCZ and T2D. MAGMA gene-level analysis revealed 94 significant shared genes. Enrichment analyses pinpointed a complex genetic interplay, with possible inflammatory, metabolic, and oxidative stress mechanisms. GSMR analysis suggested a possible weak causal link between SCZ and cardiometabolic traits in this study.

This large-scale genome-wide cross-trait analysis identified shared genetics, suggesting common biological mechanisms underlying their comorbidity.

ملخص الدراسة

الفصام هو مرض عقلي مزمن، ويبلغ معدل انتشاره حوالي 1% عالمياً. تُشير الدراسات الوبائية إلى الاعتلال المشترك بين الفصام والاضطرابات القلبية والأيضة، مع ذلك، فإن الأسس الجينيّة والعلاقة السببية والفسيولوجيا المَرَضيّة لهذا الاعتلال المشترك لا تزال غير مفهومة بشكل كامل .

لذا فإن هذه الدراسة تهدف إلى استقصاء العوامل الجينية المشتركة بين الفصام والسمات القلبية-الأيضية. باستخدام بيانات احصائية لأفراد من أصول أوروبية من دراسات الترابط على مستوى الجينوم (GWAS) وتتراوح أحجام العينات فيها من 50,364 و898,130، حيث استُخدم تحليل الانحدار اللوجستي لتحليل الارتباط غير المتكافئ (LDSC) لتقييم وراثة هذه الأمراض وكذلك التداخل الجيني بين الفصام والاضرابات المتعلقة بالقلب والأيض، كما أُجري تحليل التباين الشمولي عبر السمات (cross-trait meta-analysis) واسع النطاق على مستوى الجينوم لدراسة التداخل الجيني بين الفصام واضطرابات القلب والأيض. كما أُجري تحليل التحديد الدقيق مُتعدد السمات (multi-trait fine mapping) وتحليل التَمَوضُع المُشترك(Colocalization) لتحديد المتغيرات الجينيّة السببية المحتملة والمشتركة. كما تم استخدام MAGMA لتحديد الجينات المرتبطة وذات دلالة إحصائية. استُخدمت تحليلات الإثراء (enrichment analysis) لتحديد التعبير الجيني التفاضلي في الأنسجة وكذلك المسارات البيولوجيّة المشتركة. أخيراً، تم تطبيق أسلوب العشوائية المندليّة (GSMR)لتقييم السببيّة .

كشف تحليل LDSC عن وجود ارتباط جيني سلبي ذو دلالة إحصائية بين الفصام وداء السكري من النوع الثاني، وتعدد الأمراض القلبية الأيضية، والسمنة من الدرجة الأولى، والسمنة من الدرجة الثانية. كما نتج عن تحليل التباين الشمولي عبر السمات 124 متغيرًا جينيًا مستقلًا، منها 7 متغيرات مُحدَثة. كما كشف تحليل التحديد الدقيق مُتعدد السمات عن 90 متغيرًا جينيًا يُحتمل أن تكون مسببة، منها 15 متغيرًا مُحدَثاً وكشف تحليل التَمَوضُع المُشترك عن 11 متغيراً جينياً مسبباً للفصام ومرض السكري من النوع الأول. وأظهر تحليل MAGMA على مستوى الجينات 94 جينًا مشتركًا ذا دلالة إحصائية. وأشارت تحليلات الإثراء إلى تفاعل جيني معقد، مع آليات محتملة تشمل الالتهاب، الأيض، والإجهاد التأكسدي. وأشار تحليل العشوائية المدليّة إلى وجود مسؤولية جينية ضعيفة بين الفصام والأمراض القلبية الأيضية في هذه الدراسة.

كشف هذا التحليل الشامل للجينوم عن سمات جينية مشتركة، مما يُشير إلى وجود آليات بيولوجية مشتركة وراء التزامن المرضي.

DECLARATION

I declare that the master thesis entitled "**Shared Genetics Between Schizophrenia and Cardiometabolic Diseases: A Large-Scale Genome-Wide Cross-Trait Analysis**" is my own original work, and thereby certify that unless stated, all work contained within this thesis is my own independent research and has not been submitted for the award of any other degree at any institution, except where due acknowledgment is made in the text.

Name: Tasnim Ahmad Mohammad Salameen

Signature: ———————————

Date: ———————————

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for the joint master's degree in biotechnology at Palestine Polytechnic University and Bethlehem University, I agree that the joint library shall make it available to borrowers under the rules of the joint library. Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of the source is made. Permission for extensive quotation from, reproduction, or publication of this thesis may be granted by my supervisors, or in their absence, by the Dean of Higher Studies when, in the opinion of either, the proposed use of the material is for scholarly purposes. Any copying or use of the material in this thesis for financial gain shall not be allowed without my written permission.

Name: Tasnim Ahmad Mohammad Salameen

Signature: ———————————

Date: ———————————

DEDICATION

I dedicate my work to my parents, who have been sources of strength throughout my life and to whom I owe an enormous debt of appreciation. Your endless love, support, and sacrifices have been an umbrella for my academic journey. I hope that I made you proud.

ACKNOWLEDGMENTS

First and foremost, I want to convey profound appreciation to my dedicated supervisor, Dr. Yaqoub Ashhab. I have learned so much from you. Thank you for your counsel, guidance, and support throughout this journey. This thesis is undeniably a better piece of work owing to it.

I would like to express my sincere gratitude to the Palestine-Korea Biotechnology Center at the Palestine Polytechnic University (PPU) for providing me with the opportunity, resources, and academic environment that enriched this project.

My gratitude extends to Dr. Rami Arafeh, Dr. Robin Abu Ghazaleh, Asma Tamimi, Ayman Al Eadieh, Hassan Taradeh, and Zaid Taradeh for their support and guidance.

Besides, I would like to thank my thesis committee: Dr. Robin Abu Ghazaleh and Dr. Fouad Zahdeh.

I am forever grateful to my lovely pals. Haneen Ibraheem, thank you for always having my back (not my front). Yasmin Hisham, my lovely and gorgeous buddy, inspires me to be a better researcher. Naema Ali, who constantly motivates me to be a better person. Thank you, Ruba Abdoo, my sister Banan Ahmad, and your unwavering support!

At last, my heartfelt gratitude goes to my dear family for their ever-present love, patience, and sacrifices, which made everything possible.

LIST OF ABBREVIATIONS

SCZ	Schizophrenia
T2D	Type 2 diabetes mellitus
CVD	Cardiovascular disease
BMI	body mass index
OB	Obesity
CAD	Coronary artery disease
GWAS	Genome-wide association study
LDL	Low-density lipoprotein
HDL	High-density lipoprotein
TG	Triglycerides
TC	Total cholesterol
MR	Mendelian randomization
PGC	Psychiatric Genomics Consortium
GWAS ATLAS	Atlas of GWAS Summary Statistics
CMM	Cardiometabolic multimorbidity
OB-I	Obesity Class I
OB-II	Obesity Class II
OB-III	Obesity Class III
MAF	Minor allele frequency
MHC	Major histocompatibility complex
LD	Linkage disequilibrium
LDSC	LD score regression
SNP	Single-nucleotide polymorphisms
LD	Linkage disequilibrium
ASSET	A subset-based association analysis of heterogeneous traits and subtypes
PIP	Posterior inclusion probability
CSs	credible sets
eQTL	Expression quantitative trait loci
VEP	The Ensembl Variant Effect Predictor
MAGMA	Multi-marker Analysis of GenoMic Annotation
GSMR	Generalized Summary Data-Based Mendelian Randomization
FDR	False discovery rate
CADD	Combined Annotation Dependent Depletion
TF	Transcription factor

LIST OF FIGURES

LIST OF TABLES

LIST OF SUPPLEMENTAL TABLES

TABLE OF CONTENTS

CHAPTER ONE

Introduction

1.1. Overview of schizophrenia

Psychiatric disorders are a category of diseases that affect patients' thinking, feeling, mood, behavior, and ability to perform daily life functions (National Academies of Sciences, 2020). Among these, schizophrenia (SCZ) is a chronic, complex, relapsing, and debilitating mental disease. Schizophrenia complexity originates from its heterogeneity and the broad spectrum of symptoms and manifestations associated with it (Correll and Schooler, 2020; Correll et al., 2022). Moreover, SCZ is considered one of the most severe mental disorders due to the intensity of its symptoms and the fact that many individuals do not fully recover and experience stigma and social isolation (Jauhar, Johnstone and McKenna, 2022). Schizophrenia's global prevalence is about 1% (Millan et al., 2016; Velligan and Rao, 2023), with intra-population and geographical region variations (Patel *et al.*, 2014; Mandal *et al.*, 2022). Its prevalence is similar in both males and females, with the average age of onset ranging from 15 to 30 years (Fišar, 2023; Velligan and Rao, 2023).

There is a wide range of clinical manifestations in SCZ. Positive symptoms encompass hallucinations, delusions, disorganized thinking, and erratic behavior. Negative symptoms reflect a deficiency in functional capabilities, such as diminished motivation, emotional expression, and social engagement (Mandal et al., 2022; Velligan and Rao, 2023). Additionally, cognitive impairment is a fundamental aspect of the disorder, significantly impacting the functional deficits experienced by patients (Seidman and Mirsky, 2017; Fišar, 2023).

Schizophrenia etiology is still incompletely understood. It is believed that the etiopathogenesis of SCZ is based on abnormal neural circuit development and imbalanced synaptic communication, which results from dysregulated neurotransmitter systems and intracellular signaling pathways (Fišar, 2023). Core neurochemical hypotheses include

changes in dopaminergic, serotonergic, and glutamatergic neurotransmission (Jauhar, Johnstone and McKenna, 2022).

The conventional therapeutic management of SCZ includes antipsychotic pharmacotherapy to reduce acute psychotic outbursts and prevent relapse (Correll *et al.*, 2022). Weight gain and metabolic issues are notable side effects linked to atypical antipsychotics. On long-term use, there is an increased risk of T2D and CVD (Patel *et al.*, 2014).

Genetic studies of SCZ have made remarkable progress over the last ten years, from a few reproducible findings to many robust genetic associations (Sullivan and Geschwind, 2019). Because monozygotic twins had concordance rates between 41 to 65% compared to 0 to 28% in dizygotic ones, twin and family studies have demonstrated a significant genetic component (Henriksen, Nordgaard and Jansson, 2017). Additionally, the risk of SCZ increased about 10% among first-degree relatives and 3% among second-degree relatives, while the risk increased to about 40% when both parents are affected. All these findings emphasize the high level of heritability of SCZ (Reis Marques and Howes, 2020; Velligan and Rao, 2023).

The high heritability and extremely polygenic character of SCZ have been demonstrated by next-generation sequencing and molecular genetics. Copy number variations (CNVs) have been identified at over 15 loci, which, despite their rarity, account for roughly 2.5% of SCZ cases (Rees, O'Donovan and Owen, 2015; Sullivan, Yao and Hjerling-Leffler, 2024).

Large-scale genome-wide association studies (GWAS) continue expanding the number of associated loci. 108 loci were discovered in the initial GWAS in 2014 (Jauhar, Johnstone and McKenna, 2022; Sullivan, Yao and Hjerling-Leffler, 2024). Furthermore, over 300 relevant loci that account for roughly 24–33% of the genetic liability of SCZ in European populations have been found by extensive GWAS conducted by the Psychiatric Genomics Consortium (Legge *et al.*, 2021).

Overall, research indicates that SCZ is a highly polygenic and pleiotropic illness, with CNVs and both common and uncommon coding mutations raising the risk of this disease. (McCutcheon, Reis Marques and Howes, 2020).

Multimorbidity is defined as the condition in which two or more chronic diseases are present in an individual. The importance of multimorbidity is in its capacity to improve knowledge of the underlying mechanisms and, thus, facilitate the development of more successful health-care approaches (Arruda *et al.*, 2024).

Multimorbidity is more common in people with psychiatric problems than in the general population, including SCZ patients. Cardiovascular disease (CVD), respiratory conditions, Type 2 diabetes mellitus (T2D), obesity (OB), infections, and several types of cancer are among the comorbidities of SCZ, with prevalence rates between 40-70% (Dieset, Andreassen and Haukvik, 2016; Schneider *et al.*, 2019; Fiorillo and Sartorius, 2021). Evidence suggests that comorbidities associated with SCZ may arise from several factors, including lifestyle-related, health care system-related (Dieset, Andreassen and Haukvik, 2016; Schneider *et al.*, 2019), therapy-related, and genetic factors (Dieset, Andreassen and Haukvik, 2016).

1.2. Comorbidity of SCZ and T2D

Type 2 diabetes mellitus (T2D) is a chronic metabolic disease manifested by impaired insulin secretion and reduced insulin sensitivity, resulting in poor glycemic control and hyperglycemia (Huang *et al.*, 2018). T2D multimorbid diseases include metabolic syndrome, CVD, and psychiatric disorders (Guerrero Fernández de Alba *et al.*, 2020). The risk of T2D is two to five times higher in patients with SCZ, according to several studies (Lindekilde et al., 2021; Arruda et al., 2024).

Despite the well-known metabolic side effects of antipsychotic drugs, including the elevated risk of T2D, there is growing evidence that metabolic abnormalities present in SCZ may also play a significant role. For instance, in early psychosis, before antipsychotic treatment begins, there is often dysregulation of the appetite-hormonal axes, which are low leptin and increased insulin levels that raise the risk of T2D (Lis *et al.*, 2020). Moreover, this comorbidity is linked to allostatic load that accompanies SCZ and leads to biological changes, including low-grade inflammation, elevated oxidative stress, reduced neurotrophin levels, and dysregulation in HPA-axis activity (Mizuki *et al.*, 2021). Collectively, evidence points to the interaction of several factors, such as genetic predisposition, environmental

exposures, and disease-specific biochemical alterations, as the root of the comorbidity between SCZ and T2D (Gragnoli *et al.*, 2016).

Family-based linkage studies have identified several risk loci in different chromosomal regions that may contribute to this comorbidity, which include 1p13, 1p36, 1q25, 2q14, 2q33, 2q36, 3p22, 3q29, 9p24, 7q31, and 7q21, among others (Lin and Shuldiner, 2010). Many of them embrace gene-rich regions that potentially contain shared candidate genes contributing to both SCZ and T2D (Tziastoudi *et al.*, 2019).

Analyses using data from the Genetic Association Database, the GWAS Catalog, and the T2D Genetic Association Database identified 196 SCZ and 200 T2D susceptibility genes, of which 14 are common between the two diseases. Among these are genes involved in inflammation pathways (APOE, IL10, TNF) and oxidative-stress pathways (GSTM1, MTHFR, PON1, SOD2, UCP2). More recently, 402 SCZ and 890 T2D-associated genes were published in a GWAS catalog lists (Nagalski, Kozinski and Wisniewska, 2016), of which 26 genes are shared between SCZ and T2D (Wang *et al.*, 2013; Zhang *et al.*, 2013; Mizuki *et al.*, 2021). Additionally, there are also GWAS discoveries of pleiotropic shared loci, including TCF7L2, MPHOSPH9, and PROX1 (Arruda *et al.*, 2024).

Overall, the genetic interaction of the two diseases is complex, despite the shared variants between them; both SCZ and T2D represent widely heterogeneous phenotypes, and the genetic architecture of either disease negatively correlates with the other (Mizuki *et al.*, 2021; Arruda *et al.*, 2024). Moreover, the findings on common genetic influences of SCZ and T2D comorbidity are inconclusive, which underscores the importance of cross-disorder studies with large-scale GWAS data to shed light on the degree of genetic overlap between these two diseases and pathways that could mediate this comorbidity (Suvisaari *et al.*, 2016).

1.3. Comorbidity of SCZ and CVD

Cardiovascular disease (CVD) is a category of disorders affecting the heart and blood vessels (Lopez, Ballard and Jan, 2023). In SCZ, CVD is the main cause of premature death. Its relative risk is two to three times higher in SCZ patients than in the general population,

resulting in a 17–22% decrease in life expectancy (Ringen *et al.*, 2014; Azad *et al.*, 2016; Rødevand *et al.*, 2023; Polcwiartek *et al.*, 2025).

The precise mechanisms underlying this comorbidity are probably multifaceted and have not yet been fully identified (Rødevand *et al.*, 2023). Furthermore, the cardiovascular comorbidity in SCZ is significantly influenced by T2D, which is the primary risk factor for CVD, and is linked to a two- to four-fold increase in the incidence of SCZ. This increases the risk of coronary heart disease, stroke, and CVD mortality by approximately two times (Vancampfort *et al.*, 2016; Goldfarb *et al.*, 2022). Additionally, the metabolic syndrome seems to have a significant contribution to the augmented risk of T2D and CVD in patients with SCZ, driving hypertension, dyslipidemia, abdominal obesity, insulin resistance, inflammatory activity, endothelial dysfunction, and persistent stress (Ringen *et al.*, 2014; Azad *et al.*, 2016).

However, evidence indicates that SCZ and CVD have an underlying pathobiology that goes beyond lifestyle, healthcare, and treatment-related factors (Ringen *et al.*, 2014). Early evidence of glucose dysregulation in antipsychotics-naïve and first-episode patients, along with elevated CVD risk in their first-degree relatives, suggests a genetic involvement (Rødevand *et al.*, 2023; Polcwiartek *et al.*, 2025).

There is mounting evidence that the CVD risk factors and SCZ share a common genetic basis (Rødevand *et al.*, 2023). In a recent large-scale GWAS study in cohorts of European descent, 825 loci were found to have a common association with SCZ and CVD-related traits, including triglycerides, blood pressure, high-density lipoprotein (HDL), BMI, smoking initiation, waist-to-hip ratio, total cholesterol, low-density lipoprotein (LDL), T2D, and CAD (Rødevand *et al.*, 2023).

Several GWA studies exploring the shared genetics of SCZ and CVD identified candidate shared genes, including the cardiomyopathy-associated gene (CMYA5) (Chen *et al.*, 2011). The NRG1 gene, linked to sudden unexpected deaths from ventricular fibrillation, is also associated with SCZ (Huertas-Vazquez *et al.*, 2013), while the SLC22A23 gene is reported to be related to QTc prolongation in SCZ patients on antipsychotic treatment (Åberg *et al.*,

2012). However, the detailed mechanisms underlying SCZ remain poorly understood (Zhu *et al.*, 2018; Fry *et al.*, 2024).

1.4. Comorbidity of SCZ and OB

Adults who have a body mass index (BMI) of more than 30 kg/m^2 (body weight in kilograms divided by squared height in meters) are considered obese (Díaz-López and Gutiérrez-Aguilar, 2020; Yu et al., 2023). Patients with SCZ have a higher prevalence of obesity, more than twice as high as the overall population. Notably, a metabolically unfavorable body composition, characterized by increased fat percentage, abdominal obesity, and decreased muscular mass, is frequently observed in patients with SCZ (Suvisaari *et al.*, 2016; Yu *et al.*, 2023).

Multiple systematic reviews and meta-analyses indicate an elevated risk of OB and increased body fat among patients with SCZ. Several factors were suggested as contributors to this elevated risk, including a sedentary lifestyle, poor eating habits, metabolic effects of antipsychotics, genetic predisposition, hormonal changes, and inflammatory markers (Ringen *et al.*, 2014; Yu *et al.*, 2023). Furthermore, neurostructural abnormalities have been reported in obese schizophrenic patients, and variations in BMI are suggested to have an impact on brain volumes, pointing to the possibility that OB, SCZ, and brain structural changes might be linked to shared genetic pathways (Kharabian Masouleh et al., 2016; Yokum and Stice, 2017; Wolf et al., 2017).

Several biological mechanisms link psychiatric disorders and OB. Proposed mechanisms include alterations in inflammatory processes, dysregulated function of the hypothalamic-pituitary-adrenocortical axis, imbalanced hormones, and neurotransmitters (Chao, Wadden and Berkowitz, 2019). The role of LEP and LEPR in neurological diseases was addressed through their inhibition of the JAK/STAT3 pathway, thus protecting against dopaminergic neurodegeneration and neuroinflammation caused by α-synuclein (α-SYN) (Qin *et al.*, 2016). Furthermore, by lowering TAU phosphorylation, LEP promotes neuroprotection (Flores-Dorantes, Díaz-López and Gutiérrez-Aguilar, 2020).

Identified shared genes between OB and psychiatric disorders include FTO, POMC, ITIH4, TLR4, BDNF, and CREB1 (Amare *et al.*, 2017). Furthermore, several GWA studies identified 578 brain-expressed genes related to food intake and body weight, some of them linked to OB, neurotransmitter signaling, and synaptic function, including NEGR1, CADM2, NRXN3, ELAVL4, GRID1, and SCG3 (Locke *et al.*, 2015; Ignatieva *et al.*, 2016).

Nevertheless, shared genetics between SCZ and OB has only been the subject of a limited number of GWAS-based investigations, which explains why new studies might help to unveil potential therapeutic targets or risk factors (Yu *et al.*, 2023).

1.5. Statistical genetic epidemiology methods relevant to this thesis

1.5.1. Genetic correlation and heritability analysis using LDSC

Complex diseases have genetic and environmental predispositions. Furthermore, distinguishing the complicated interrelationships between diseases and features is an important goal of epidemiology. Therefore, it is important to interpret the genetic effects on phenotypic variations in a specific population, which can be quantified using heritability (Bulik-Sullivan, Finucane, *et al.*, 2015; Zhu and Zhou, 2020; Barry *et al.*, 2022). Recently, GWAS revolutionized heritability estimation methods into population-based approaches, in which the genetic similarity is estimated by utilizing single-nucleotide polymorphisms (SNPs). The percentage of phenotypic variation explained by all SNPs is referred to as SNP-based heritability estimates (Bulik-Sullivan, Finucane, *et al.*, 2015; Zhu and Zhou, 2020; Srivastava, Williams and Zhang, 2023).

LDSC only utilizes summary-level data, enabling heritability estimations of complex traits using publicly available GWAS (Barry et al., 2022). The LDSC approach takes into account that SNPs that have high LD scores are probably tagged causal SNPs, and thus have a greater association compared to those with low LD scores. To estimate heritability, LDSC regresses GWAS chi-square against LD Scores for all SNPs; the intercept measures bias, and the LDSC slope estimates heritability (Barry *et al.*, 2022).

The previously presented single-trait LD score regression approach is expanded to cross-trait LDSC to assess the genetic covariance (ρg) between LD and the product of effect sizes (z-

scores) for two distinct phenotypes. To determine genetic correlation (r_g), genetic covariance (ρ_g) normalized by SNP heritability (h^2):

$$r_g = \rho_g / \sqrt{h_1^2 h_2^2},$$

where h_1^2, h_2^2 are the heritabilities of phenotype 1 and phenotype 2, respectively.

The genetic correlation can be interpreted on a scale between −1 and 1 (Bulik-Sullivan, Finucane, *et al.*, 2015). Z-score matrices of the two phenotypes are inflated in case of overlapped samples. The amount of inflation is uniform throughout all markers and is unaffected by LD scores. Thus, overlapped samples only impact the intercept of the regression ($\rho N_s / \sqrt{N_1 N_2}$) not the slope, therefore, genetic correlation estimation will not be influenced by this overlap (Bulik-Sullivan, Finucane, *et al.*, 2015).

1.5.2. Cross-trait meta-analysis using ASSET

Traditionally, meta-analyses of GWAS are mostly focused on a single phenotype and the SNPs associated with it. However, there is mounting evidence of the critical role that pleiotropy and epistasis play in complex disorders. The influence of one SNP on another is known as epistasis (Deng *et al.*, 2020; Mortezaei and Tavallaei, 2021). Regarding pleiotropy, there are two types: vertical pleiotropy, which arises when a SNP impacts one phenotype, which then causes another, and horizontal pleiotropy, which appears when a SNP affects multiple phenotypes independently. Mendelian randomization (MR) is primarily used to detect the first type. Hence, by combining summary data of several different traits, cross-trait GWAS meta-analysis increases the statistical power for identifying shared loci among the traits (Li and Zhu, 2017; Jang *et al.*, 2020).

A subset-based association analysis of heterogeneous traits and subtypes (ASSET) employs a novel approach to conduct cross-trait fixed-effect meta-analysis by exploring different subsets of multiple traits to identify true association signals that could be present in one subset of the phenotypes, either in the same or opposite directions. In this method, SNP associations are tested across all possible subsets (S) of z-scores as per this equation:

$$(Z(S) = \sum_{k \in S} \sqrt{\pi_k(S)}\, Z_k),$$

Where $\pi(S) = n_k / \sum_{k \in S} n_k$ denotes the sample size for the kth study relative to the total sample size for the given subset *S*. The selected subset of traits is the one that offers the best meta-analyzed z-score ($Z_{\max-\text{meta}} = max_{S \in \boldsymbol{S}} | Z(S) |$).

ASSET cross-trait meta-analysis provides three options to perform cross-trait meta-analysis: (i) standard fixed-effect meta-analysis across all traits; (ii) one-sided fixed-effect meta-analysis that identifies the best subset of associated traits either in a positive or negative directions; and (iii) two-sided fixed-effect meta-analysis in which one-side subset search is separately applied for positively and negatively associated traits for a given SNP followed by combining the association signals from both directions into a single pooled chi-square type statistic. ASSET offers multiple methods for p-value computation, including Exact Importance Sampling (ID), Discrete Local Maximum (DLM), and Bonferroni (B). Meanwhile, the calculated p-value also accounts for the multiple testing penalty due to the subset search (Bhattacharjee *et al.*, 2012).

1.5.3. Fine mapping of multiple traits using mvSuSiE

GWAS meta-analysis outputs are the lead SNPs, which are associated with the phenotypes, and may be causal or not. Other SNPs in the locus, however, can be associated either because they are in LD with the causative SNP or because they are independently associated. Therefore, the putative causative variants from GWAS meta-analysis and causal SNPs linked to a particular trait are identified using fine-mapping approaches (Schaid, Chen and Larson, 2018; Broekema, Bakker and Jonkers, 2020). Methods to perform fine-mapping include the heuristic method and Bayesian fine-mapping (Schaid, Chen and Larson, 2018; Broekema, Bakker and Jonkers, 2020).

mvSuSiE is a multi-trait fine mapping method that jointly analyzes GWAS data of multiple traits to identify putative causal variants by integrating a multivariate linear regression model. The outputs of mvSuSiE include a list of credible sets (CSs) and a posterior inclusion probability (PIP) for each SNP (Zou *et al.*, 2024). The credible sets refer to the minimum set of SNPs that contains all causal SNPs with probability α. It can be estimated by cumulative

summing of PIP from the largest to the smallest, then selecting the SNPs with α more than 99% or 95% as putative causal SNPs (Adebiyi *et al.*, 2021).

1.5.4. Colocalization analysis using COLOC

Colocalization is an approach used to evaluate whether a causal SNP in a specific genomic locus is shared between traits (Wallace, 2021). The COLOC, which is the colocalization tool used in this analysis, computes the posterior probability of five hypotheses using a Bayesian statistical approach to determine the likelihood of shared and distinct causes of the associated SNPs, as follows: H0 denotes no association with either trait; H1 and H2 denote association with either of the two traits, respectively; H3 denotes link with both traits but with different casual SNPs for each trait; and H4 denotes association with both traits with shared causal SNPs. Strong evidence of shared causality between traits is seen in variants with PPH4 > 0.75 and/or PP4/PP3 ≥ 3 (Pan *et al.*, 2023), while 0.5 < PPH4 < 0.75 is considered as a moderate indication of colocalization (Shi *et al.*, 2024).

1.5.5. MAGMA analysis

Multi-marker Analysis of GenoMic Annotation (MAGMA) is a tool for analyzing GWAS data, including gene and gene-set analyses (de Leeuw *et al.*, 2015). MAGMA determines candidate significant genes of associations by aggregating SNP-level GWAS results, while taking into account LD between SNPs within each gene. Additionally, MAGMA conducts gene-set analysis by assessing the enrichment of significant genes in particular biological pathways, such as KEGG (Pilalis *et al.*, 2025).

1.5.6. Generalized Summary-Data-Based Mendelian Randomization (GSMR)

Mendelian randomization is a statistical method that excludes pleiotropy and employs SNPs as instrumental variables to assess the causal link between two diseases or phenotypes (exposure and outcome) (Adebiyi *et al.*, 2021). Three assumptions underlie MR. First, there must be a durable association between the exposure of concern and instrumental variables (IVs) at a genome-wide significant threshold of $P < 5 \times 10^{-8}$. Second, there must be no correlation between the IVs and any confounders that might bias the exposure-outcome link.

Third, the only way IVs can influence the outcome is through exposure (Nordestgaard, 2016).

Generalized Summary Data-Based Mendelian Randomization (GSMR) is a multi-SNP MR analysis integrating GWAS summary level data to test for causality between different traits. GSMR integrates the significant independent SNPs as IVs, which minimizes the bias originating from LD between IVs. In addition, the HEIDI-outlier method is implemented within GSMR to eliminate pleiotropic SNPs.

CHAPTER TWO

Aims and Objectives

Schizophrenia (SCZ) is often comorbid with multiple chronic diseases, including T2D, CVD, and OB. The genetic mechanisms underlying this multimorbidity remain largely unclear. Recent GWAS studies have started to explore shared genetic factors, but these are limited and often do not focus specifically on SCZ.

This study seeks to explore the shared genetics between SCZ and three of its major comorbid diseases (T2D, CVD, and OB) in individuals of European ancestry. This study will integrate multiple statistical genetic epidemiology methods using GWAS data of large sample sizes to shed more light on overlapping genetic loci and causal relationships.

The primary objective of this study is to explore overlapping genetic loci that contribute to susceptibility to SCZ, T2D, CVD, and OB using cross-trait meta-analysis.

Secondary objectives of this project include:

1. To investigate possible putative causative shared SNPs using multi-trait fine-mapping, and then to identify shared common causal SNPs through colocalization analysis.
2. To infer the causal relationship of SCZ with T2D, CVD, and obesity classes by utilizing Generalized summary data-based Mendelian randomization (GSMR).

CHAPTER THREE

Data and Methods

3.1. Study design

Briefly, the analysis was conducted using GWAS summary statistics of European ancestry cohorts. First, heritability estimations and genetic correlation analyses were conducted. Second, cross-trait meta-analysis using ASSET was utilized, followed by fine-mapping using (mvSuSiE) and pairwise colocalization analysis using (COLOC). Third, functional annotations of identified loci were conducted, followed by gene-level analysis. Then, pathway and tissue-enrichment analysis of the prioritized genes was conducted. Finally, Mendelian randomization analysis was performed using GSMR. **Figure 3-1** shows the design of this thesis project.

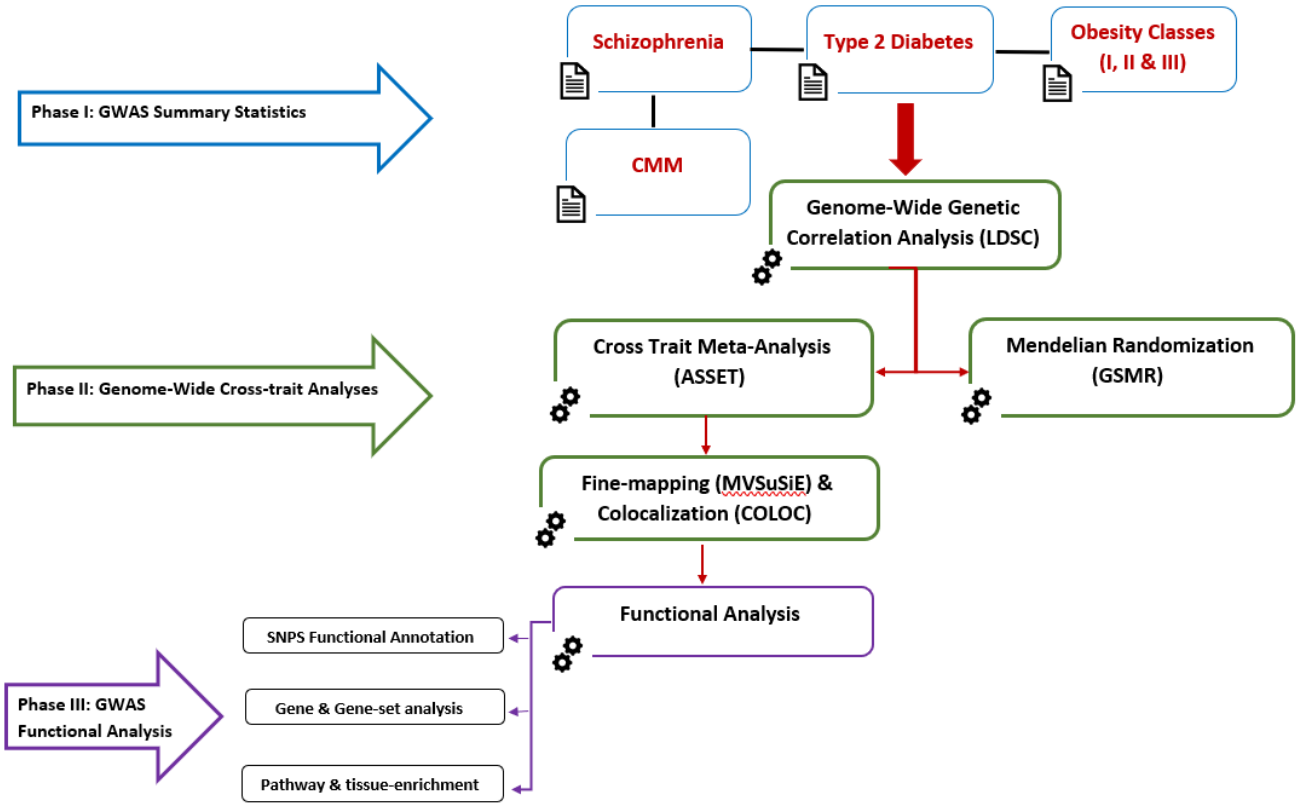


Figure 3-1. Study design

3.2. GWAS data

Summary statistics from publicly available databases. The SCZ (Lam *et al.*, 2019) summary statistics was obtained from the PGC (Psychiatric Genomics Consortium) database. The cardiometabolic multimorbidity (CMM) (Zhao *et al.*, 2024) summary data was sourced from the NHGRI-EBI GWAS Catalog. Summary statistics for T2D (Mahajan *et al.*, 2018), as well as OB Classes (Berndt *et al.*, 2013) were sourced from GWASATLAS. All GWAS summary data pertain to cohorts of European ancestry and are based on the hg19 human assembly. The sample sizes for each cohort are summarized in **Table 3-1**.

Table 3-1. Cohort sample sizes

Phenotype	Cohort	Total sample size	Number of cases	Number of controls
Schizophrenia	EUR	77,096	33,640	43,456
obesity class I	EUR	98,697	32,858	65,839
obesity class II	EUR	72,546	9,889	62,657
obesity class III	EUR	50,364	2,896	47,468
Type 2 diabetes	UKB2 (EUR meta)	898,130	74,124	824,006
Cardiometabolic multimorbidity	EUR	367,147	14,835	352,312

As a Quality control (QC) procedure, the munge_sumstats function in the LDSC (Bulik-Sullivan, Loh, *et al.*, 2015) Python package was applied to GWAS summaries. This function filtering and quality control include put not limited to removing SNPs with out-of-range p-values ($p \leq 0$ or > 1), SNPs with minor allele frequency (MAF) values less than 0.01, SNPs with imputation quality less than 0.9, and strand-ambiguous (A/T, C/G), dropping rows with missing required values, merging/aligning alleles, and removing duplicate SNPs. The extended major histocompatibility complex (MHC) region was excluded from this analysis owing to its complexity.

3.3. Statistical genetic analysis

LD score regression (LDSC) (Bulik-Sullivan, Loh, *et al.*, 2015) Python package (https://github.com/bulik/ldsc) was used to estimate the heritability of and the pairwise genetic correlation. Pre-calculated LD scores for European ancestry from the 1000 Genomes Project were used as a source of LD scores, and filtered GWAS summary results were used as input data.

Before running the analyses, GWAS summaries were merged with the HapMap3 SNPs panel to improve data quality, thus maintaining the accuracy of the LDSC estimations. Heritability was estimated for SCZ, T2D, CMM, and obesity classes, while pairwise genetic correlations were obtained for SCZ with T2D, CMM, and obesity classes.

The next step was to conduct cross-trait meta-analysis. By using the ASSET Bioconductor R package version 2.24.0 (https://doi.org/10.18129/B9.bioc.ASSET), a subset-based analysis across SCZ, T2D, CMM, and Obesity classes was conducted using the "h.traits" function, which is a suitable function to conduct meta-analysis of different traits when summary-level data are available. Other parameters were adjusted as follows: (i) the p-values of association were calculated by the Discrete Local Maximization (DLM) method, using "p.dlm" function, (ii) the "search" helper function retained "NULL", to perform meta, one-sided, and two-sided subset searches. Furthermore, a correlation matrix of Z-statistic was estimated empirically using null SNPs and incorporated in the ASSET framework to address possible sample overlap between GWAS data.

Following the ASSET analysis, SNPs with significant p-values at a threshold of less than 5×10^{-8} were extracted from ASSET outputs. Then, a function called' ld-clump' in PLINK (Purcell *et al.*, 2007) was used to extract lead SNPs (SNPs that are independent of one another at the LD threshold and have reached a minimal p-value threshold). In this step, the LD reference panel of the 1000 Genomes Project for European ancestry was used as a reference for LD scores. Finally, the clumping was performed for the significant SNPs with the following predefined parameters (-clump-p1 $1e^{-8}$ –clump-p2 $1e^{-2}$ –clump-r^2 0.05 –clump-500kb), which means, SNPs that reach a genome-wide significance threshold of 5×10^{-8}, and are independent of each other at LD $r^2 < 0.05$, and within a 500 kb window from each other were fused into the same locus.

Before performing downstream analysis, subset1sided and subset2sided lead SNPs were filtered so that any lead SNP was retained only if it was associated with schizophrenia and at least one of the other traits of this study; otherwise, the Lead SNP was excluded. Finally, three separate Lead SNP lists were created, including a list containing the shared lead SNPs and the other two lists containing lead SNPs only reported in either Subset1sided or Subset2sided analysis.

To investigate shared putative causal variants of SCZ with cardiometabolic traits, the mvSuSiE R package (https://github.com/stephenslab/mvsusieR) was used to fine-map the lead SNPs to GWAS summary data of SCZ, T2D, CMM, and obesity classes. The fine-mapping analysis window was defined with a region of size (±500 kb) around the lead SNPs, using the same LD reference panel as described above to build the LD matrix in this analysis. Finally, the resulting SNPs were filtered based on PIP; SNPs with PIP> 0.9 were considered as possible putative causal SNPs and called Top SNPs.

To determine which of the abovementioned top SNPs colocalize between pairs of traits, the "coloc.abf" function in the COLOC R package (https://github.com/chr1swallace/coloc) was implicated. COLOC computes the posterior probability under the assumption that each trait in the region has a single putative causal variant, using the same LD reference panel as described above. In this analysis, SNPs were considered to be colocalized if the posterior probability of colocalization for those SNPs between the selected pair of traits (PPH4) exceeded 50.0%.

Significant candidate genes were determined using MAGMA v1.10 (de Leeuw *et al.*, 2015) gene-level analysis, which was performed using lead SNPs and top SNPs as input data. A genome-wide significant threshold of $P < 2.75 \times 10^{-6}$ (FDR adjustment for testing 18,188 genes) was used to map SNPs to 18,188 genes.

The required files to conduct this analysis, such as the gene location file and gene-set files, were obtained from MAGMA (https://cncr.nl/research/magma/) and GSEA (https://www.gsea-msigdb.org/gsea/msigdb/human/collections.jsp) resources, respectively, using the same LD reference panel as described above. An annotation window of 35,10kb upstream and downstream of the genes was integrated to capture variants in the regulatory regions. Then, the resulting candidate genes were used as input for gene-set analysis. Significance threshold for either genes in gene-level analysis or for gene-sets was selected using FDR < 0.05.

Causal association between traits in this analysis was conducted using the GSMR2 R package (https://github.com/JianYang-Lab/gsmr2) under default parameters, which comprises using independent SNPs ($r2$<0.05) that are significantly associated with exposure

at p-value $<5\times10^{-8}$ as instrumental variables (IVs). In this analysis, bi-directional GSMR analyses were performed for SCZ with T2D, CMM, and obesity classes, while removing SNPs that display horizontal pleiotropy (HEIDI outlier $P< 0.05$). The required input data include GWAS summary statistics of all traits and the previously described LD reference panel.

3.4. Functional annotation of lead SNPs and top SNPs

3.4.1. Determination of SNPs' novelty

In this analysis, a SNP was considered a novel association if it was not previously reported in the GWAS catalog and was not within a distance of less than 1 Mb from a SNP previously reported with the selected traits. To accomplish this analysis, the "LDtrait" function in the LDlink (Myers, Chanock and Machiela, 2020) R package (https://github.com/CBIIT/LDlinkR) was used. Additionally, the GWAS Catalog trait associations selected to determine novelty included SCZ, schizoaffective disorder, BMI, OB, T2D, CAD, stroke, CVD, ischemic heart disease, overweight body mass index status, metabolic syndrome, and insulin resistance.

3.4.2. SNP-level functional annotation

Functional annotation of GWAS variants using multiple resources is a crucial step to elicit functional information regarding those variants (Zhou *et al.*, 2023). In this project, different tools were used to score and prioritize the lead and top SNPs, including Haploreg v4.2 web browser (Pilalis *et al.*, 2025), regulomeDB v2.2 web browser (Pilalis *et al.*, 2025), which were used to annotate regulatory SNPs in the human genome, and the Combined Annotation Dependent Depletion (CADD) v1.7 (https://cadd.bihealth.org/score), which was used to measure the deleteriousness of the SNPs (Kircher *et al.*, 2014), and the VEP Ensembl browser (https://grch37.ensembl.org/info/docs/tools/vep/index.html) was used for further annotations.

3.5. Tissue enrichment and pathway analysis

For the candidate genes obtained by MAGMA gene-level analysis, FUMA GENE2FUNC (https://fuma.ctglab.nl/gene2func), FLAME (https://pavlopoulos-lab-services.org/shiny/app/flame), Enricher-KG (https://maayanlab.cloud/enrichr-kg), and

STRING v12.0 (https://string-db.org/cgi/input?sessionId=bM0GCIzUda4s) were integrated in this analysis.

FUMA GENE2FUNC v1.8.1 (Watanabe *et al.*, 2017) was used in this analysis to conduct gene-set analysis, which gives insights about enrichment of the candidate genes in functional categories and biological pathways, as well as the differential expression of genes in a specific tissue in parallel with other tissues. Additionally, to identify pathways, diseases, biological functions, or phenotypes associated with the abovementioned candidate genes, Flame (v2.0) (Karatzas *et al.*, 2023) and Enrichr-KG (Evangelista *et al.*, 2023) functional enrichment was integrated (Karatzas *et al.*, 2023). Finally, STRING (Szklarczyk *et al.*, 2022) was used to investigate protein–protein interactions.

CHAPTER FOUR

Results

4.1. Genetic correlation analysis revealed a negative correlation of SCZ with cardiometabolic traits

To quantify the heritable component of each trait, SNP-based heritability was assessed for SCZ, T2D, CMM, and obesity classes using LDSC. The results (**Table 4-1**) revealed SNP-based heritability (h^2) estimations of 45.13% for SCZ, 4.84% for T2D, 2.6% for CMM, 16.37% for OB1, 14.87% for OB2, and 9.68% for OB3.

Table 4-1. SNP-based heritability

Phenotype	SNPs in analysis	h^2	SE
SCZ	1,176,813	0.4513	0.0163
T2D	1,173,994	0.0484	0.0022
CMM	1,142,683	0.026	0.0019
OB1	1,041,384	0.1637	0.0075
OB2	1,027,341	0.1487	0.0087
OB3	999,710	0.0968	0.0085

h^2: heritability estimate, SE: standard error of heritability, SE: standard error.

To comprehend the genetic correlation between SCZ, T2D, CMM, and obesity classes. The results revealed a statistically significant negative correlation between SCZ and multiple cardiometabolic phenotypes (**Table 4-2**). Specifically, SCZ showed significant negative correlations with T2D (rg =-0.0684, *P* =0.0009), CMM (rg =-.0696, *P* =0.0207), OB1 (rg =-0.0836, *P* =0.0013), and OB2 (rg =-0.0946, *P* =0.0022). Although SCZ and OB3 exhibited a negative correlation (rg =-0.0667), it did not reach statistical significance (*P* =0.1028).

Table 4-2. Genetic correlation between SCZ and cardiometabolic traits

Phenotype 1	Phenotype 2	rg	rg_se	z-score	p value
SCZ	T2D	-0.0684	0.0206	-3.3235	0.0009
	CMM	-0.0696	0.0301	-2.3129	0.0207
	OB1	-0.0836	0.0261	-3.2069	0.0013
	OB2	-0.0946	0.0309	-3.0583	0.0022
	OB3	-0.0667	0.0409	-1.6314	0.1028

rg: genetic correlation estimate, rg_se: standard error of genetic correlation

4.2. Cross-trait meta-analysis (ASSET) and multi-trait fine mapping identified novel shared loci between SCZ and cardiometabolic traits

The central theme of this analysis is to explore the shared genetics of SCZ with cardiometabolic traits. To accomplish this, a cross-trait meta-analysis employing an unconventional approach

(ASSET) was implemented. To concentrate on interpreting innovative signals from ASSET results, “meta” results were excluded from subsequent analyses while focusing on “subset1sided” and “subset2sided” results. The Manhattan plots (**Figure 4-1a-b**) display multiple regions of genome-wide significant signals ($p < 5 \times 10^{-8}$) across chromosomes 2, 3, 6, 10, 16, and 18.

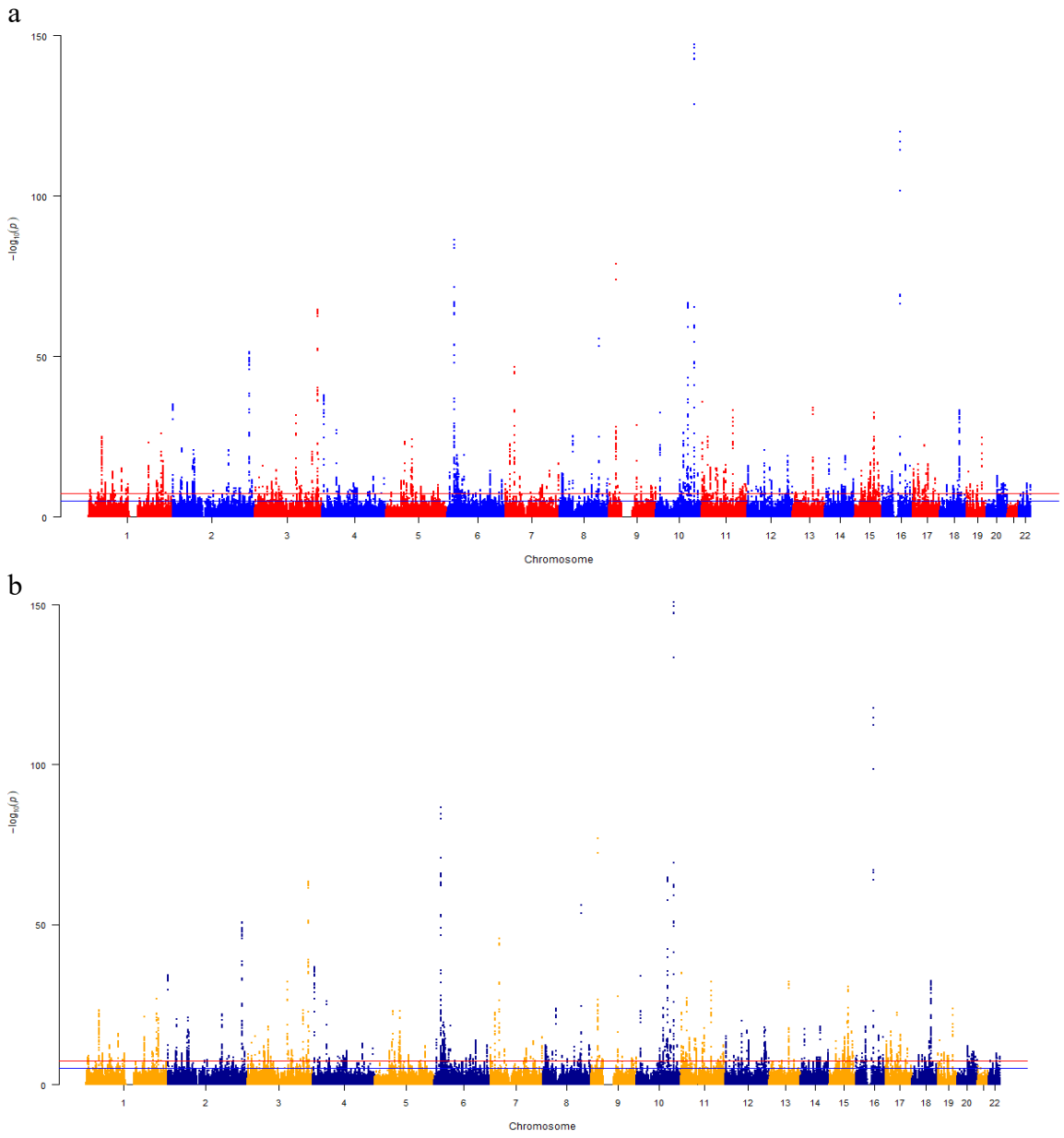
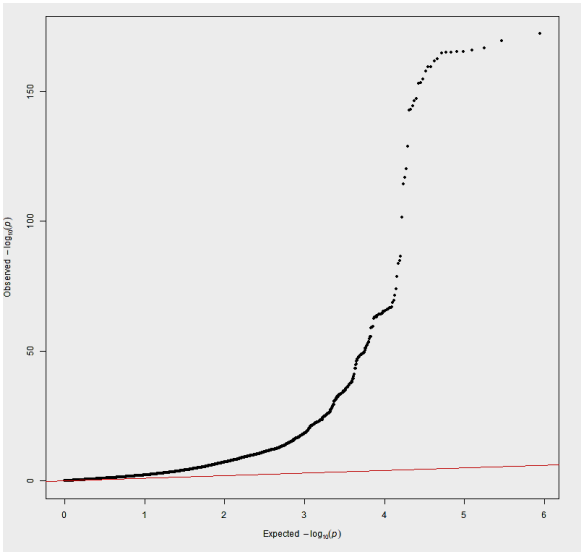


Figure 4-1. Manhattan plots of ASSET cross-trait meta-analysis. (**a**) Manhattan plot of ASSET “subset1sided” result set. (**b**) Manhattan plot of ASSET “subset2sided” result set. The Manhattan plots show significant genomic associations ($< 5 \times 10^{-8}$) across the genome. Every point on the plot represents an SNP; its height on the y axis indicates the significance of association (-log10 p-value), and its position on the x-axis reflects its genomic location

QQ plots (**Figure 4-2a-b**) show significant deviation from the null distribution in the tail, indicating a true association signal rather than inflation. Such trends are frequently observed in large cross-trait GWAS studies where multiple traits share polygenic influences.

a



b

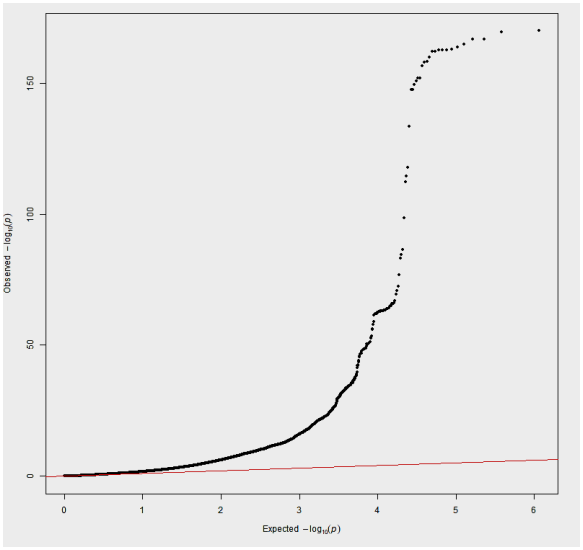


Figure 4-2. QQ plots of ASSET cross-trait meta-analysis. (**a**) QQ plot of ASSET "subset1sided" result set. (**b**) QQ plot of ASSET "subset2sided" result set.
The x-axis represents (-log10 p-value) of the theoretical distribution, y-axis represents (-log10 p-value) of the sample data.

This comprehensive analysis identified 124 lead SNPs distributed across 99 loci (see **Supplementary Table 4-1**). Among these, 90 SNPs were associated with SCZ, T2D, CMM, and at least one of the obesity classes; 10 SNPs were associated with SCZ, T2D, and CMM; 13 SNPs were associated with SCZ, T2D, and at least one of the obesity classes; 10 SNPs were associated with SCZ and at least one of the obesity classes; and 1 SNP was only associated with SCZ and T2D.

In a further advanced step to investigate putative causal shared SNPs, multi-trait fine-mapping using mvSuSiE identified 90 top SNPs across 76 loci (see **Supplementary Table 4-2**). Of these, 20 SNPs overlapped with lead SNPs (**Figure 4-3**), suggesting that these lead SNPs may represent putative causative signals.

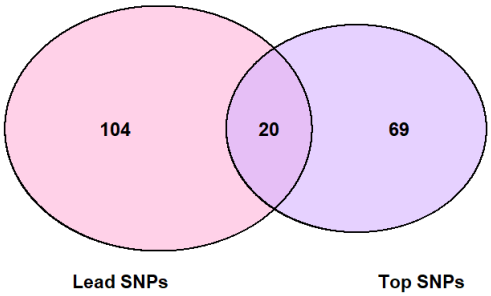


Figure 4-3. Overlap between Lead SNPs and Top SNPs.
A Venn diagram plot showing SNPs overlap between Lead SNPs and Top SNPs.

To get more insights about the previously reported associations of these shared signals, LDtrait in LDlink was used. Previous associations with a broad range of traits were uncovered, including psychiatric (e.g. cognitive performance, insomnia, anxiety, brain morphology, major depressive disorder (MDD), bi-polar disorder, neuroticism, and autism spectrum disorder), cardiometabolic (e.g. HDL levels, LDL levels, TG levels, TC levels, adiposity, dyslipidemia, leptin and leptin receptor levels, apolipoprotein A levels, waist-hip index, childhood body mass index, trunk fat mass, waist circumference, systolic and diastolic blood pressure, myocardial infarction, atrial fibrillation, dilated and hypertrophic cardiomyopathy, fasting insulin), and inflammatory markers (e.g. C-reactive protein levels) (see **Supplementary Table 4-1** and **Supplementary Table 4-2**).

In addition, this analysis revealed previous pleiotropic associations for 82 SNPs, including pleiotropy of SCZ or psychiatric disorders with either T2D or OB2 (see **Supplementary Table 4-1** and **Supplementary Table 4-2**). However, this analysis revealed an unreported pleiotropy pattern; for instance, rs2933203 (MDGA2) was reported between SCZ and OB2, whereas this analysis revealed pleiotropy between SCZ, T2D, CMM, and obesity classes.

Remarkably, novel SNPs were identified. Of the lead SNPs, 7 were novel, 5 of them were associated with SCZ and cardiometabolic traits, 1 with SCZ and obesity classes (rs1343424 mapped to AGBL4), and 1 with SCZ and metabolic traits (rs8009761 mapped to SYNDIG1L) (**Table 4-3**). Among the top SNPs, 15 were novel, with rs2904221 (FAM47E) shared with lead SNPs, and rs12958029 (NOL4) and rs830644 (FOXP1) being colocalized between SCZ and T2D (**Table 4-4**).

Table 4-3. Novel lead SNPs

Lead SNP	CHR	BP	ASSET associations	P value	Gene	Distance	Consequence
rs12196300	6	38963764	SCZ,T2D,OB2,OB3,CMM	9.95E-09	DNAH8	0	Intron
rs1343424	1	49376820	SCZ,OB1,OB2,OB3	7.10E-09	AGBL4	0	Intron
rs2904221	4	77182033	SCZT2D,OB1, OB2,OB3,CMM	1.67E-09	FAM47E	0	Intron
rs3799380	6	26467182	SCZ, OB2,OB3,CMM	4.43E-19	BTN2A1	0	Intron
rs3819720	6	32804570	SCZ,T2D,CMM	1.32E-09	TAP2	0	Intron
rs8009761	14	74888554	SCZ,T2D,OB1,OB2,OB3	3.20E-09	SYNDIG1L	0	Intron
rs17763551	17	3882309	SCZ,T2D,CMM,OB3	2.40E-14	ATP2A3	15kb	Intron, non-coding transcript

Table 4-4. Novel top SNPs

Top SNP	CHR	BP	PIP	Gene	Distance	Consequence
rs10515678	5	152323236	0.9962	NMUR2	538kb	Intergenic
rs6718758	2	60328802	1	MIR4432	286kb	Upstream gene
rs10906025	10	11869246	1	C10orf47	0	Intron
rs12958029*	18	31388680	1	NOL4	42kb	Intergenic
rs2588120	8	17465263	1	PDGFRL	0	Intron
rs2904221**	4	77182033	0.9999	FAM47E	0	Intron
rs2656865	19	1530060	1	PLK5P	0	Intron
rs6581977	12	71342952	1	PTPRR	28kb	Regulatory
rs4678874	3	36457372	1	STAC	0	Intron
rs8119937	20	32645551	0.9999	RALY	0	Intron
rs830644*	3	71665559	1	FOXP1	32kb	Intergenic
rs7217945	17	3497570	1	TRPV1	0	Intron
rs13425918	2	233714686	1	GIGYF2	0	Intron
rs7778318	7	122003326	1	CADPS2	0	Intron
rs4678398	3	137721875	1	CLDN18	0	Downstream gene

PIP: posterior inclusion probability.
* rs12958029, rs830644 Novel top SNPs colocalized between SCZ and T2D
** rs2904221 Novel lead and top SNP

To infer the novel associations, gene-level and enrichment analyses related to them were investigated. According to gene tissue expression by FUMA GENE2Function GTEx v7 heatmap, the majority of genes exhibited broad-range tissue expression, with remarkable expression across tissues of RALY, BTN2A1, and ATP2A3, whilst AGBL4, DNAH8, and NMUR2 showed minimal expression (**Figure 4-4a**).

Additionally, previous neurological and cardiometabolic associations were revealed for these genes by FUMA GENE2FUNC GWAS disease enrichment. FOXP1, GIGYF2, AGBL4, and NOL4 were associated with SCZ, with FOXP1 having associations with brain morphology and apolipoprotein A1 levels, GIGYF2 with cognitive ability, myocardial infarction, and CAD, and AGBL4 with cardio-metabolic traits, including waist circumference, BMI, and systolic blood pressure. RALY was associated with T2D and HDL, BTN2A1 and CADPS2

with SCZ-related and cardiometabolic-related traits, including brain morphology, waist circumference, but never reported with SCZ (**Figure 4-4b**).

Furthermore, enrichment patterns of these genes were unveiled by gene set analyses. Significant enrichment in gene sets of TF-binding sites and microRNA targets was disclosed for FOXP1, NOL4, GIGYF2, RALY, AGBL4, TAP2, DNAH8, BTN2A1, and CADPS2. Among these, TFs regulate neural development, stress response, lipid metabolism, metabolic processes, and immune and hormonal responses, such as STAT5A_04, OCT1_04, a TF, BRN2_01, YY1, SREBP1, CEPδ TF, P300, and ZNF596. Likewise, detected pathway enrichment was in the pathway of central nervous system neuron differentiation (AGBL4 and GIGYF2), in peptide and amide transport pathways (TAP2), and in secretion, signal release, and export from cell pathways (CADPS2) (**Figure 4-4c**).

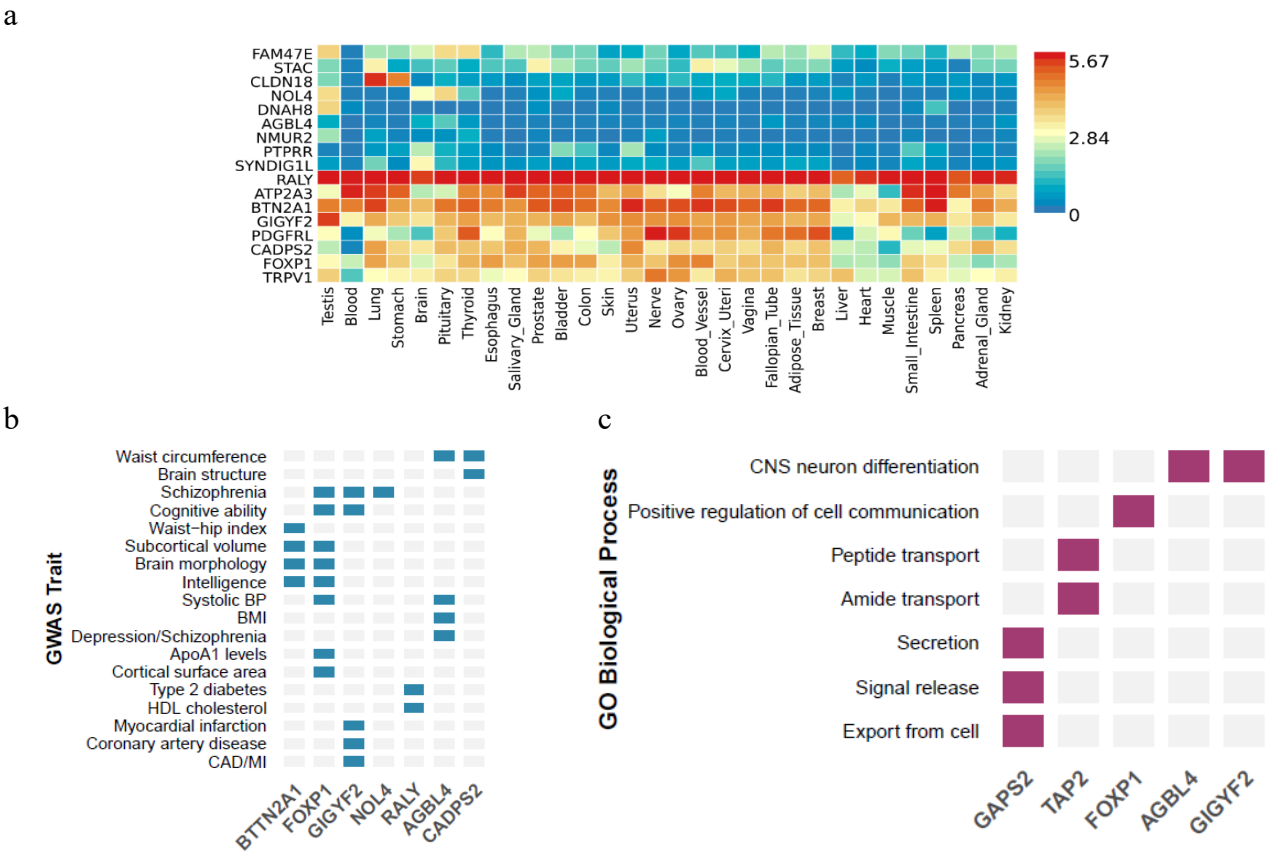


Figure 4-4. FUMA GENE2FUNC enrichment results of genes mapped to novel SNPs. (**a**) GTEx v7 heatmap revealed notable tissue enrichment of these genes, with minimal expression of AGBL4, DNAH8, and NMUR2. The x-axis represents tissues, and the y-axis represents genes. (**b**) GWAS disease enrichment among SCZ and cardiometabolic traits revealed enrichment of some mapped genes. The x-axis represents enriched genes, while the y-axis represents enriched diseases. (**c**) GO pathway analysis revealed enrichment in pathways of neuron differentiation, cellular transport, secretion, and communication. The x-axis represents genes while the y-axis represents GO:BP pathways.

4.3. Colocalization analysis uncovers shared causal SNPs between SCZ and T2D

To extend the multi-trait fine-mapping findings, shared causal associations were investigated by colocalization of the top SNPs. Shared causal signals were identified between SCZ and T2D in multiple loci (**Table 4-5**). Among these, rs6568686 (TRAF3IP2-AS1), rs2796441 (TLE1), and rs3903399 (CNTN2) exhibited high posterior probabilities of shared causality (PP.H4 > 0.75); rs12958029 (NOL4) and rs830644 (FOXP1) were novel; and SNPs with reported pleiotropy between T2D and SCZ or other psychiatric disorders (rs7660298, rs729599, rs1727302, and rs569255).

Table 4-5. Colocalized top SNPs

Top SNP	CHR	BP	PIP	PP.H4	Gene	Distance	Consequence	Previous pleiotropy
rs12958029*	18	31388680	1	53.69%	NOL4	42kb	Intergenic	---
rs6568686**	6	111872482	0.9999	92.48%	TRAF3IP2-AS1	0	Intron, non-coding transcript	---
rs7660298	4	103960768	1	56.42%	NHEDC2	0	Intron	T2D with ADHD/OCD
rs830644*	3	71665559	1	67.05%	FOXP1	32kb	Intergenic	---
rs729599	5	44842260	0.9997	74.38%	MRPS30	27kb	Intron	T2D with SCZ/OCD
rs1727302	12	123632930	1	86.48%	MPHOSPH9	8kb	Downstream gene	OB2 with ADHD/SCZ; T2D with BMI
rs2796441**	9	84308948	1	98.99%	TLE1	5.4kb	Upstream gene	---
rs3903399**	1	205041542	0.9991	77.25%	CNTN2	0	Intergenic	---
rs569255	3	124925934	1	52.89%	SLC12A8	0	Intron	T2D with SCZ/OCD

Posterior possibilities (PPH4) of colocalization were estimated with a region ± 500 Mb of lead SNPs. PIP: posterior inclusion probability, ADHD: attention-deficit hyperactivity disorder, OCD: obsessive-compulsive disorder.
* Novel colocalized SNP.
** Colocalized SNPs with no previous pleiotropy.

To deepen the understanding of these colocalized signals, related gene-level and enrichment analyses were scouted. Mapped genes exhibited notable expression across multiple tissues, including brain, pituitary, liver, adrenal gland, pancreas, blood vessels, and adipose tissues, as revealed by FUMA GENE2FUNC GTEx v7 heatmap (**Figure 4-5a**). Moreover, disease enrichment analyses highlighted previous associations of MPHOSPH9 and TLE1 with both SCZ and T2D, while other genes were previously reported with either of the traits (FOXP1, CNTN2, NOL4) (**Figure 4-5b**).

Similarly, these genes were enriched in shared gene sets of TF binding sites or microRNA targets. Furthermore, pathway analyses uncovered enrichment of FOXP1 and CNTN2 in the pathway of positive regulation of cell communication, with CNTN2 being enriched in pathways of regulation of neuron and nervous system neuron differentiation, protein

localization to the axon, and IL-18 signaling, whereas TLE1 was enriched in pathways of stress-induced DNA damage, the Wnt and Presenilin 1 signaling pathways (**Figure 4-5c-d**).

Figure 4-5. Enrichment results of genes mapped to colocalized SNPs between SCZ and T2D. (**a**) GTEx v7 heatmap revealed notable tissue enrichment of these genes. (**b**) GWAS disease enrichment among SCZ and cardiometabolic traits revealed enrichment of some mapped genes with either SCZ, T2D, or both. (**c**) GO pathway analysis revealed enrichment in pathways of neuron differentiation and cellular communication. (**d**) Enrichr-KG network revealed possible pathways of SCZ and T2D as indicated by TLE1 and FOXP1.

4.4. Functional analysis illuminates regulatory elements of the SNPs

Understanding the functionality of the identified SNPs is a vital step in bridging SNP-level findings with functional enrichment and pathway analyses, enabling interpretation of the latter in the context of the SNPs' functionality. Therefore, functional annotations for identified SNPs were carried out using different tools (see **Supplementary Table 4-3** and **Supplementary Table 4-4**). Around half of the variants were intronic, with the remainder distributed among intergenic, downstream gene variants, synonymous, 3'-UTR variants, upstream gene variants, non-coding transcript variants, downstream gene variants, missense variants, and regulatory region variants, as shown in **Figure 4-6(a-b)**.

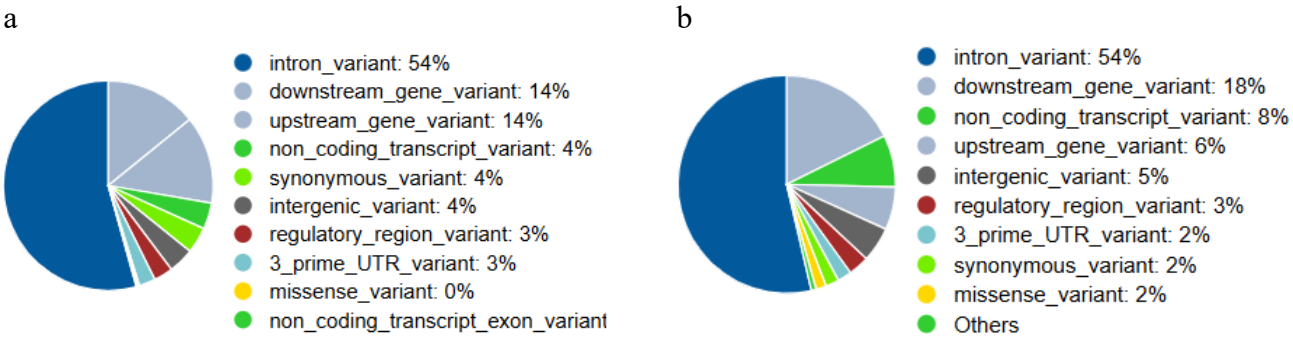


Figure 4-6. Biological functions of the lead and top SNPs. (**a**) Lead SNPs. (**b**) Top SNPs. About half of both lead and top SNPs are intronic variants.

Besides, RegulomeDB provides two measures: rank (a heuristic scoring based on experimental or computational evidence) and probability (machine learning-based predictions). According to the rank, approximately 85% of the SNPs were predicted to have a regulatory function (e.g., eQTL or transcription factor binding sites) (**Figure 4-7a-b**). by the same token, 101 SNPs revealed to have a RegulomeDP probability (ranging from 0 to 1, where 1 represents the highest regulatory) greater than 0.5.

Unexpectedly, some SNPs with rank 7 (no regulatory function) revealed to have a probability of more than 0.5, while the GRASP analysis revealed eQTL and meQTL related to some of these SNPs (**Table 4-6a-b**). For example, rs2275997 (CADM1) was reported to have an eQTL with CADM1 in the cerebellum in Alzheimer's disease, and rs4917985 (C10orf32) with meQTL chr10:104943598-104943648 in the temporal cortex.

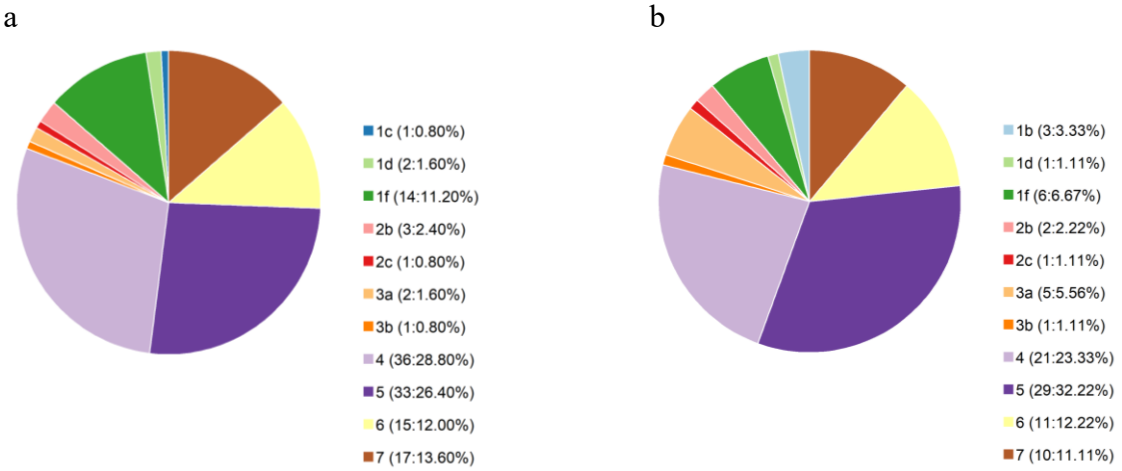


Figure 4-7. RegulomeDB ranking of the lead and top SNPs. Each color refers to RegulomeDB rank, accompanied by the number and percentage of SNPs for each rank. (**a**) RegulomeDB ranking of the lead SNPs. (**b**) RegulomeDB ranking of the top SNPs.

The RegulomeDB score is a heuristic scoring based on the degree of experimental or computational evidence.

RegulomeDB variant classification outline:

1a: eQTL + TF binding + matched TF motif + matched DNases footprint + DNase peak
1b: eQTL + TF binding + any motif + DNases footprint + DNase peak
1c: eQTL + TF binding + matched TF motif + DNase peak
1d: eQTL + TF binding + any motif + DNase peak
1e: eQTL + TF binding + matched TF motif
1f: eQTL + TF binding/DNase peak

Likely to affect binding
2a: TF binding + matched TF motif + matched DNases footprint + DNase peak
2b: TF binding + any motif + DNases footprint + DNase peak
2c: TF binding + matched TF motif + DNase peak

Less likely to affect binding:
3a: TF binding + any motif + DNase peak
3b: TF binding + matched TF motif

Minimal binding evidence:
4: TF binding + DNase peak
5: TF binding or DNase peak
6: Motif hit

Table 4-6a. Lead SNPs with a ranking of 7 and a probability of more than 0.5

Lead SNP	CHR	BP	Probability	Gene	GRASP eQTL	GRASP meQTL
rs2275997	chr11	115100185	0.51392	CADM1	CADM1 in cerebellum in Alzheimer's disease, IGSF4 in Caudal pons	---
rs3799380	chr6	26467181	0.51392	BTN2A1	BTN3A2 in cerebellum in Alzheimer's disease	chr6:26472772-26472822 in Caudal pons
rs10838158	chr11	43752521	0.51392	HSD17B12	HSD17B12 in Frontal cortex	---

Table 4-6b. Top SNPs with a ranking of 7 and a probability of more than 0.5

Top SNP	CHR	BP	Probability	Gene	GRASP eQTL	GRASP meQTL
rs12998587	chr2	161242294	0.51392	RBMS1	CADM1 in cerebellum in Alzheimer's disease, IGSF4 in Caudal pons	---
rs4917985	chr10	104624071	0.51392	C10orf32	BTN3A2 in cerebellum in Alzheimer's disease	chr6:26472772-26472822 in Caudal pons
					HSD17B12 in Frontal cortex	---

GRASB v2.0 (the Genome-Wide Repository of Associations between SNPs and Phenotypes database). Only QTLs related to traits in this analysis were investigated.

Lastly, the CADD score was used to investigate the deleteriousness of the SNPs, which corresponds to the increase in CADD score. **Table 4-7(a-b)** shows the SNPs with a CADD score greater than 10. Notably, rs5398 (SLC2A2) with a CADD score of 15.53, classified as benign, deleterious, and pathogenic by SIFT, Polyphen, and Alphamissense, respectively. This variant was previously reported with HDL levels and waist circumference, while in this analysis, it was associated with T2D, CMM, SCZ, OB1, OB2, and OB3.

Table 4-7a. Lead SNPs with a CADD score greater than 10

Lead SNP	CHR	BP	CADD	Gene	Distance	Consequence
rs6990912	8	9200471	10.92	LOC157273	7.9kb	Intergenic
rs12196300	**6**	**38963763**	**11.02**	**DNAH8**	0	**Intron**
rs7968682	12	66371879	14.2	HMGA2	12kb	Intergenic
rs17221259	12	14410484	16.33	ATF7IP	108kb	Downstream
rs5215	11	17408629	10.38	KCNJ11	0	Missense
rs8068351	17	34916878	10.78	GGNBP2	0	Intron
rs3115672	6	31727896	11.06	MSH5	0	Synonymous
rs340874	1	214159255	11.25	PROX1	2.6kb	Upstream gene
rs4702	15	91426559	12.59	FURIN	0	3' UTR
rs12615058	2	228989842	13.09	SPHKAP	0	Intron
rs12910334	15	84958317	13.7	LOC388152	59kb	Non-coding transcript exon
rs10189857	2	60713234	15	BCL11A	0	Intron
rs1573815	3	52870131	18.09	TMEM110-MUSTN1	0	Upstream gene
rs6126570	20	51094308	20.9	ZFP64	286kb	Intergenic
rs10838158	11	43752521	11.59	HSD17B12	0	Intron
rs10077814	5	44916788	12.1	MRPS30	101kb	Intergenic
rs3903399	1	205041541	12.94	CNTN2	0	Intron
rs7156625	14	79942646	13.79	NRXN3	0	Intron
rs2250377*	1	201860625	19.63	SHISA4	0	Missense
rs2796441	9	84308947	19.81	TLE1	5.4kb	Upstream gene

Table 4-9b. Top SNPs with a CADD score greater than 10

Top SNP	CHR	BP	CADD	Gene	Distance	
rs6990912	8	9200471	10.92	LOC157273	7.9kb	Intergenic
rs3737095	6	38957770	12.52	DNAH8		Splice polypyrimidine tract, intron
rs340874	1	214159255	11.25	PROX1	2.6kb	Upstream gene
rs4412207	6	98410755	11.4	MIR2113	62kb	Intron, non-coding transcript
rs13266634	8	118184782	15.11	SLC30A8	0	Missense
rs6097012	20	51146472	15.38	ZFP64	338kb	Regulatory region
rs5398	3	170715829	15.53	SLC2A2	0	Missense
rs2306590	17	34854279	16.88	MYO19	0	Missense
rs1016287	2	59305624	19.22	FLJ30838	15kb	Intergenic
rs2270576	17	47007962	11.14	SNF8	0	Synonymous
rs569255	3	124925933	11.52	SLC12A8	0	Intron
rs3903399	1	205041541	12.94	CNTN2	0	Intron
rs3784692	15	67988132	14.26	MAP2K5	0	Intron
rs2796441	9	84308947	19.81	TLE1	5.4kb	Upstream gene

rs12196300 is a Novel SNP. rs6990912, rs569255, rs3903399, and rs2796441 are colocalized SNPs between SCZ and T2D.

4.5. MAGMA gene-level analysis identifies 94 significant genes

The next step after assessing the functionality of the shared signals was to identify the shared significant candidate genes, which were obtained by aggregating SNP-level signals using MAGMA. This gene-level aggregation enhances the statistical power of detecting shared genes beyond individual SNP associations. MAGMA identified 94 significant genes mapped to 106 of the SNPs (see **Supplementary Table 4-5**). These candidate genes were used to conduct gene-set, enrichment, and pathway analyses.

4.6. Enrichment and pathway analyses reveal shared biological mechanisms

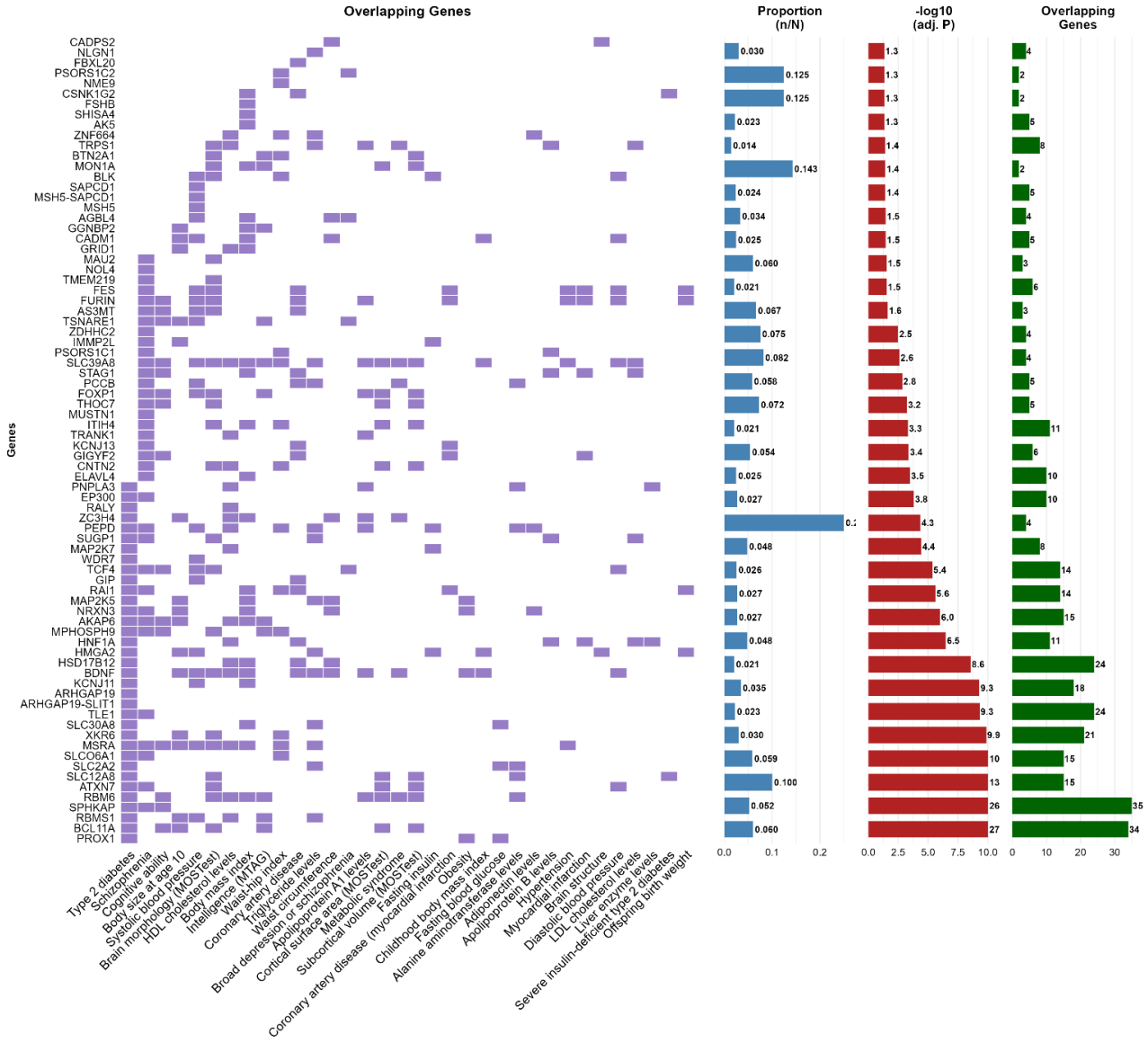
These analyses were intended to translate gene-level findings into a meaningful biological context and to understand possible mechanisms behind these associations. Gene-set analyses unveiled significant enrichment across curated gene sets (MsigDB c2), TF and microRNA binding sites gene sets (MsigDB c3), gene ontology (GO) and gene ontology biological processes (MsigDB c5), and immunologic signature gene sets (MsigDB c7). (see **Supplementary Table 4-6** and **Supplementary Table 4-7**).

GWAS disease enrichment analysis highlighted genes' previous associations; 35 genes were previously associated with SCZ, 35 with T2D, 24 with BMI, 22 with systolic blood pressure, 13 with CAD, 15 with the waist-to-hip index, 24 with brain morphology, 14 with TG levels, and 15 with cognitive ability (see **Supplementary Table 4-7**). Notably, MSRA was reported in SCZ, T2D, and BMI. FES, FURIN, AS3MT, TCF4, TSNARE1, PCCP, PEPD, GIGYF, FOXP1, SLC39A8, and KCNJ13 were associated with SCZ and cardiac traits, mainly CAD and systolic blood pressure, with PEPD being reported with T2D and SLC39A8 with MBI, RAI1, NRXN3, and AKAP6 with SCZ, T2D, and BMI (**Figure 4-8a**).

Genes with no GWAS disease enrichment include IL17REL, MDGA2, NLGN1, APOPTI (C14orf153), ARTNL, C12orf65, DNAH8, FAM212B, FAM47E, RBM6, SLC9B1, SLCO4A1, SPPL3, SYNDIGIL, DNAJC11, TMEM219, and TAP2. FLAME enrichment analysis revealed an association of NLGN1 (mapped to rs247975) with cell differentiation in the spinal cord and in several pathways of glutamatergic synaptic regulation and organization, and an association of FAM47E (mapped to rs2904221) in the regulation of

histone methylation (**Figure 4-8b**). Moreover, STRING network analysis revealed co-expression of NLGN1 with MDGA2 and with BDNF.

a



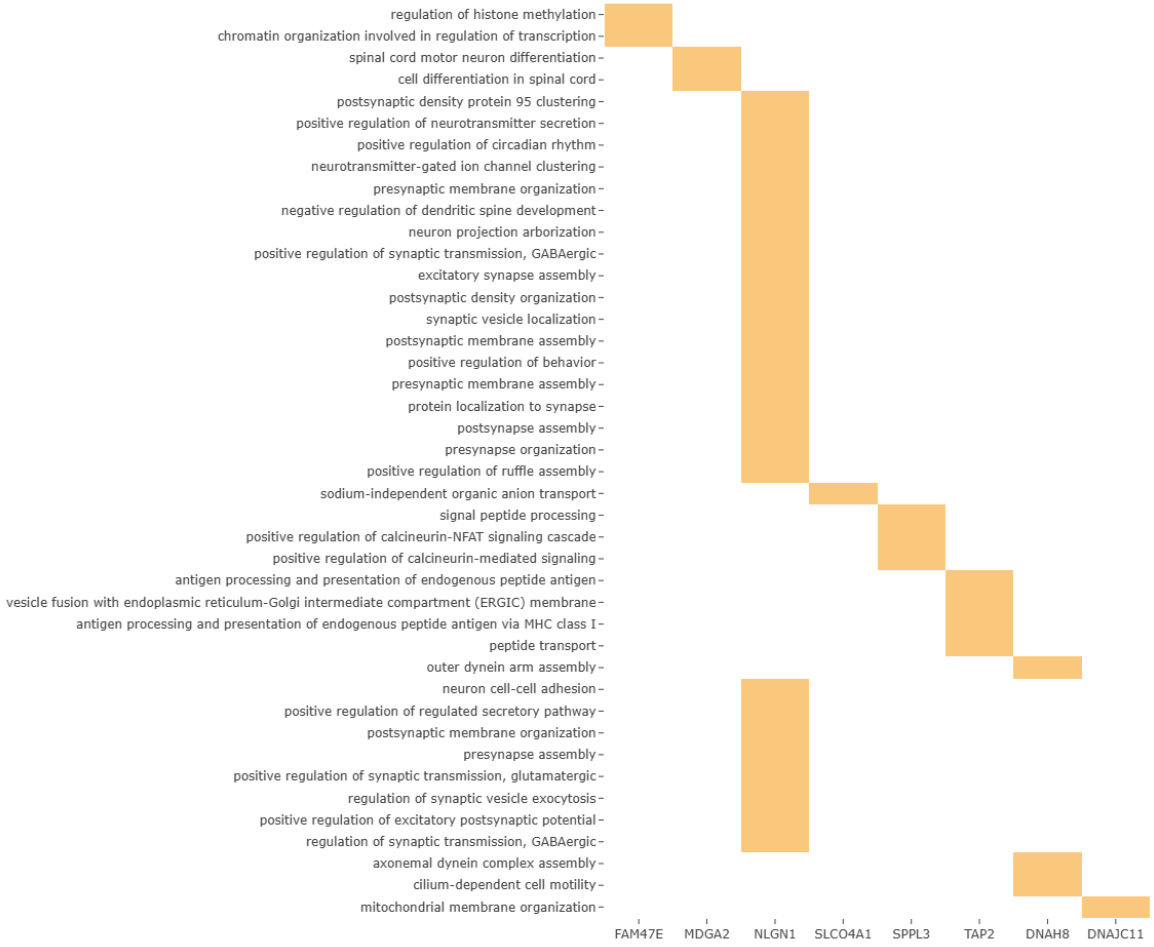


Figure 4-8. FUMA GENE2FUNC GWAS disease enrichment and FLAME GO:GP enrichment. (**a**) Heatmap of FUMA GENE2FUNC GWAS disease enrichment of MAGMA significant genes. The y-axis represents enriched genes, and the x-axis represents diseases enriched for these genes. The proportion side par represents the proportion of genes enriched for each disease, calculated by dividing the number of enriched genes per disease by the total genes annotated to that gene set. -log10 (adj. P) represents -log10 of the adjusted p-value of the enrichment. Overlapping genes represent the number of genes enriched in each set. (**b**) The FLAME enrichment heat map shows the GO:BP enrichment of genes with no GWAS disease enrichment.

Furthermore, the majority of genes were found to be enriched in shared gene-sets of TF and microRNA binding sites, suggesting a possible shared regulatory mechanism (**see Supplementary Table 4-6**). Among the microRNA gene sets, MIR495_3P enriched for (AGBL4, BCL11A, ATXN7, NLGN1, CADPS2, ARHGAP19-SLIT1, ARHGAP19, BDNF, CADM1, MPHOSPH9, AKAP6, NRXN3, TCF4), MIR548F_5P and MIR548AJ_5P_MIR548G_5P_MIR548X_5P enriched for (ELAVL4, PROX1, RBMS1, KCNJ13, SLC12A8, PCCB, AS3MT, HMGA2, TCF4, ZC3H4), and MIR129 enriched for (ELAVL4, RBMS1, AS3MT, CADM1, NRXN3, NOL4, TCF4, EP300), with ELAVL4, RBMS1, AS3MT, CADM1, NRXN3, NOL4, TCF4, EP300 being enriched among all of these gene sets (**Figure 4-9a**).

In terms of TF binding sites, genes showed high enrichment in these gene sets, with genes showing shared enrichment with microRNA gene sets as well. **Figure 4-9b** showed genes enriched in TF binding sites gene sets for the top 30 enriched gene sets, with BCL11A, BDNF, ELAVL4, KCNJI3, MAP2K5, NRXN3, RAI1, HNF1A, EP300, FOXP1, SLC12A8, HF1A, PROX1, RALY, CADM1, TRPS1, NOL4, and TLE1 being targets for more than one of these TFs. Moreover, STRING network analysis revealed enrichment of 68 of MAGMA significant genes in the alternative splicing mechanism, indicating protein products with at least two isoforms due to distinct pre-mRNA splicing events (**Table 4-8**).

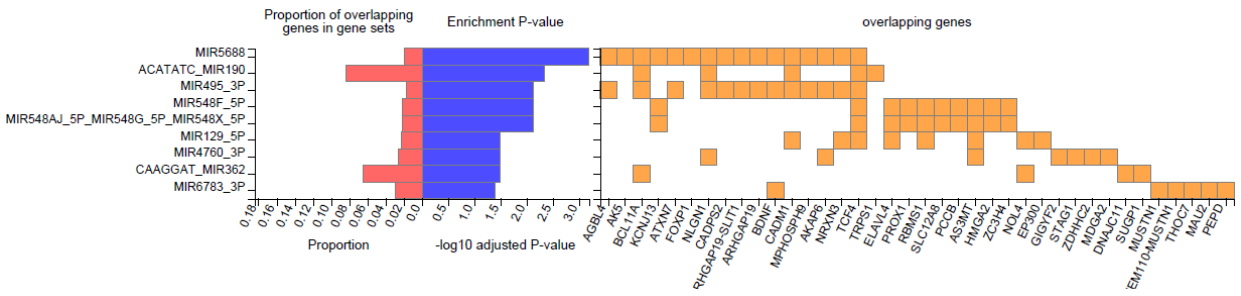
Table 4-8. STRING network analysis alternative splicing enrichment

Term description	Gene count	FDR	Genes
Alternative splicing	68	0.0416	MAP2K5,PNPLA3,SLCO4A1,KCNJ13,PEPD,RALY,SUGP1,WDR7,PSORS1C1,NOL4,MAU2,FBXL20,RBM6,ITIH4,HSD17B12,AKAP6,SPPL3,RBMS1,ATXN7,SLC9B1,MON1A,TSNARE1,BTN2A1,MSRA,RAI1,SLC2A2,CADM1,GRID1,FES,DNAH8,NME9,KCNJ11,AK5,ELAVL4,ARHGAP19,AS3MT,AGBL4,TAP2,MSH5,DNAJC11,STAG1,SPHKAP,SLC12A8,SLC39A8,TRPS1,MAP2K7,TCF4,ARNTL,IMMP2L,COA8,GIGYF2,C12orf65,NLGN1,CADPS2,MDGA2,FAM47E,FEZF1,SAPCD1,BDNF,SLC30A8,XKR6,PCCB,SLCO6A1,HNF1A,MPHOSPH9,GGNBP2,NRXN3,BCL11A

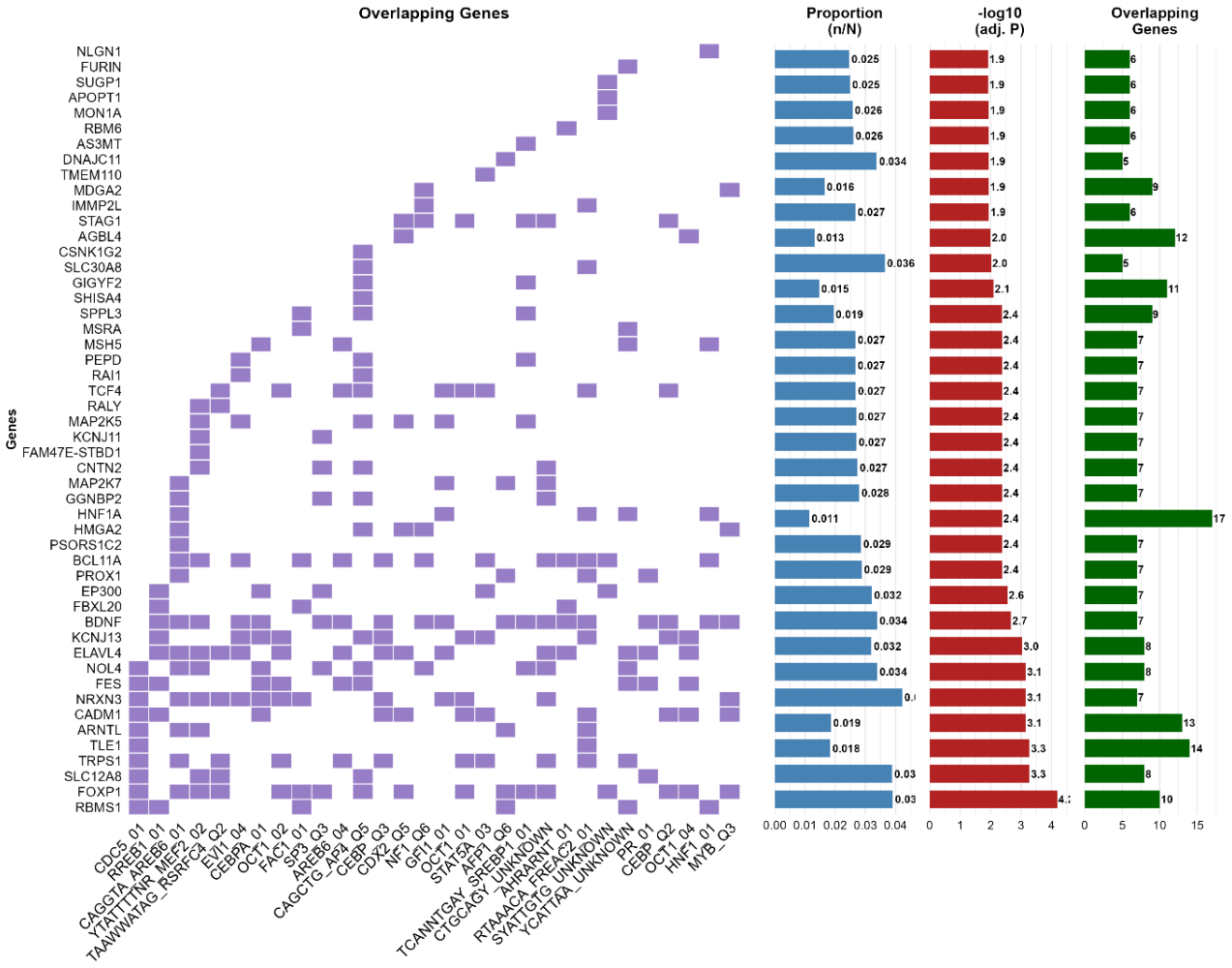
Gene count refers to the number of candidate genes that are enriched in alternative splicing. FDR represents the adjusted p-value of this enrichment.

Pathway enrichment analysis disclosed the enrichment of a cluster of genes (BLK, SLC30A8, KCNJ11, SLC2A2, ARNTL, GIP, HNF1A) in pathways, including insulin secretion, peptide transport, and hormone transport, with SLC2A2 and KCNJ11 being enriched in the type II diabetes pathway (KEGG) and the HNF (Hepatocyte Nuclear Factor) pathways. Another cluster of genes (BDNF, TCF4, BCL11A, FEZF1, MDGA2, GIGYF2, AGBL4, ELAVL4, CNTN2, PROX1) was enriched in pathways of regulation of neuron differentiation and central nervous system neuron development. TNF1 and TLE1 were found to be enriched in Presenilin action in Notch and Wnt signaling Pathways. TLE1**,** AKAP6 are enriched in the DNA damage by stress mechanisms pathway (**Figure 4-9c**).

a



b



c

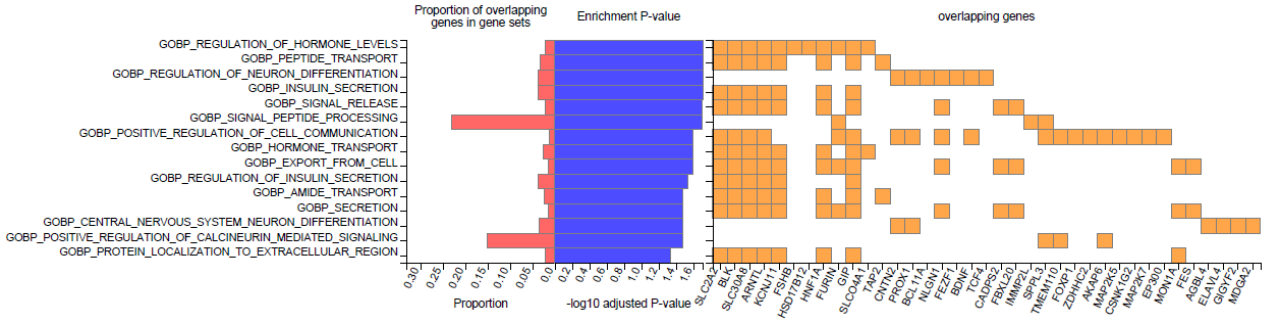


Figure 4-9. FUMA GENE2FUNC enrichment results. (**a**) Heatmap of FUMA GENE2FUNC microRNA gene set enrichment of MAGMA significant genes (The x-axis represents enriched genes, and the y-axis represents enriched microRNA gene sets). (**b**) Heatmap of FUMA GENE2FUNC shows the top 30 enriched TF binding site gene sets (The y-axis represents enriched genes, and the x-axis represents enriched TF binding site gene sets). (**b**) Heatmap of FUMA GENE2FUNC of GO:BP enrichment of MAGMA significant genes (The x-axis represents enriched genes, and the y-axis represents enriched GO:BP). In all figures, the proportion side par represents the proportion of genes enriched for each disease, which is calculated by dividing the number of enriched genes per disease by the total genes annotated to that gene set. -log10 (adj. P) represents -log10 of the adjusted p-value of the enrichment. Overlapping genes represent the number of genes enriched in each set.

Additionally, differentially expressed genes (DEG) were sought to look into gene differential expression across different tissues. **Figure 4-10** showed enrichment of the genes in different brain tissues, pancreas, liver, heart, and aortic artery. Some tissues of brain regions (hypothalamus and anterior cingulate cortex) showed both up-regulated and down-regulated expression; the brain frontal cortex showed up-regulated expression, other brain regions, liver, and pancreas associated with down-regulated expression, while both-sided expression was observed in heart and aortic artery tissues.

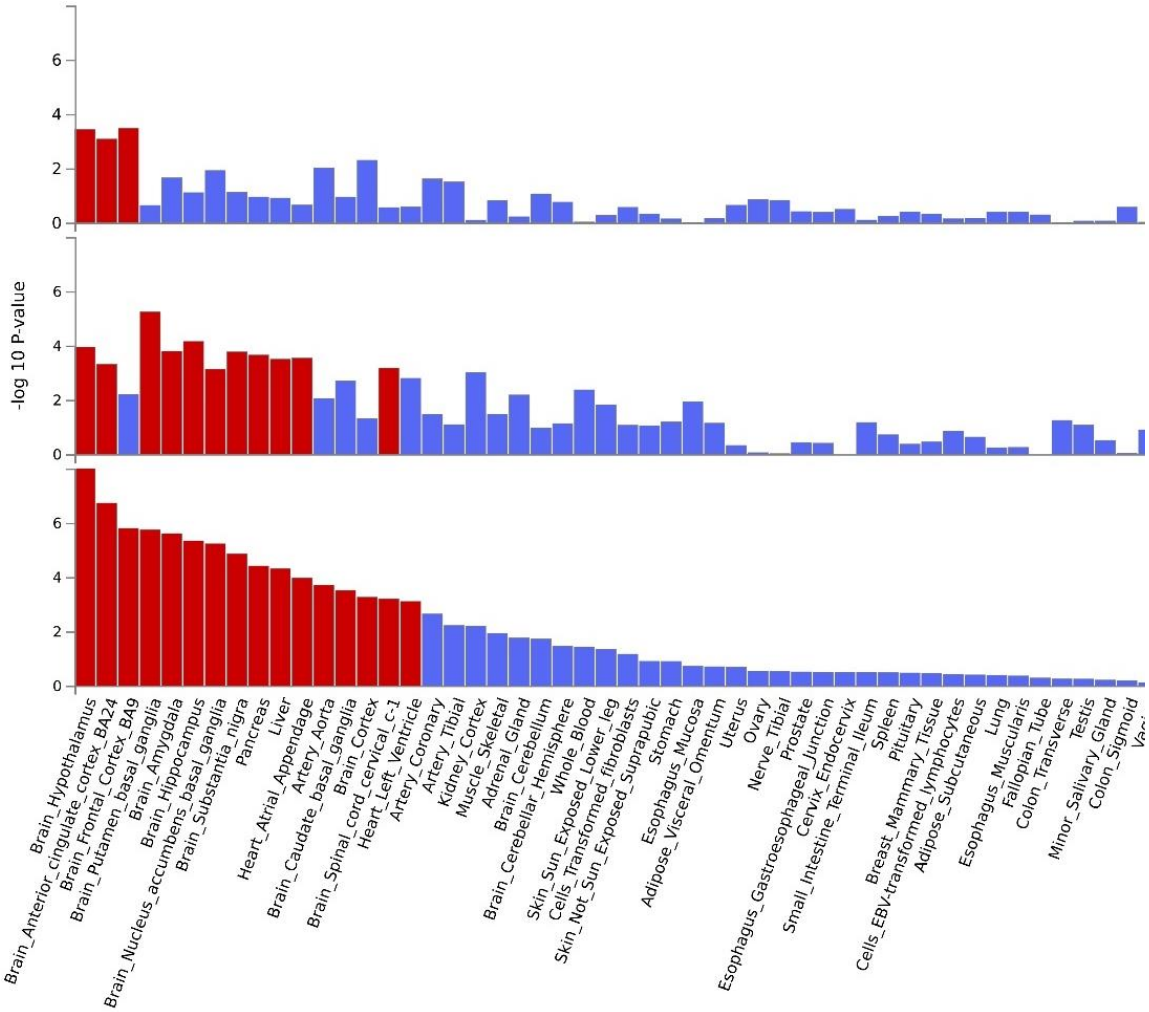


Figure 4-10. FUMA GENE2FUNC differential gene expression among tissues. The results revealed a noteworthy enrichment of mapped genes in different brain regions, including the hypothalamus, amygdala, putamen basal ganglia, accumbens basal ganglia, hippocampus, anterior cingulate cortex BA24, frontal cortex BA9, and substantia nigra, as well as in pancreas, liver, heart atrial appendage, heart left ventricle, and aortic artery.

4.7. Protein-protein interaction analysis highlights central hub genes

Accumulated Evidence of shared but complex genetic interplay was seen through the results of SNP-based, Gene-based, enrichment, and pathway analyses. In an attempt to further these findings, protein-protein interactions were explored. STRING network analysis revealed distinct clusters that map onto specific disease domains, with the histone coactivator EP300 possibly linking these clusters (**Figure 4-11**).

A cluster of metabolic regulators pointing to insulin secretion and glucose homeostasis, and was strongly tangled with T2D, including HNF1A, SLC2A2, KCNJ11, SLC30A8, PROX1, and BLK. Another cluster of neuronal and synaptic genes, corresponding to brain development and synaptic function, and was strongly linked to schizophrenia and other psychiatric disorders, includes TCF4, FOXP1, BDNF, CADPS2, ARNTL, FURIN, ELAVL4, CADM1, NRXN3, NLGN1, MDGA2, TRANK1, and GRID1. A third cluster involves hormone signaling and MAP kinase genes that are relevant to energy balance, metabolic syndrome, and obesity, including GIP, FSHB, MAP2K5, AKAP6, MUSTN1, and SLC39A8. A fourth cluster was enriched in immune and MHC-region genes, reflecting possible inflammatory and autoimmune pathways, including PSORS1C1, PSORS1C2, MSH5, and SAPCD1.

Remarkably, central hub genes connect the clusters, BLK links to immune activation, PROX1 to developmental and possibly metabolic regulation, and EP300 connects multiple clusters.

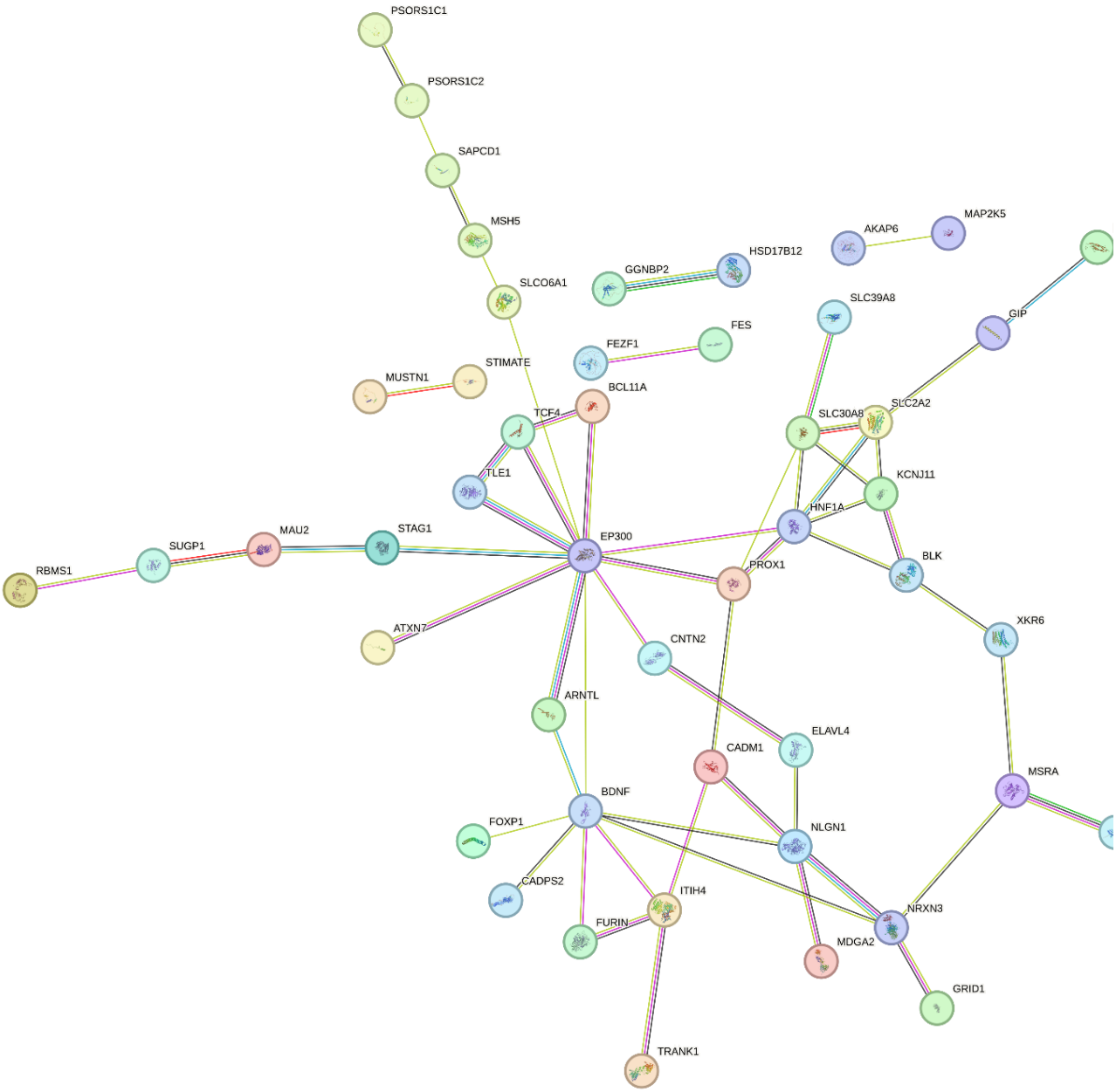


Figure 4-11. STRING network depicting association genes between SCZ and cardiometabolic diseases.
PPI enrichment p-value of this network is 7.91e-10. The expected number of edges is 23, while this network contains 58 edges. This indicates that there are more interactions between these proteins than would be predicted for a random collection of proteins from the genome with the same size and degree distribution. An enrichment of this kind suggests that the proteins are at least somewhat connected to one another. Particular interactions are represented by the edges that join the nodes. Purple and turquoise lines show known interactions. Predicted interactions are shown by green, dark blue, and red lines. Text-mining, co-expression, and protein homology relationships are shown in yellow, black, and light blue, respectively. For clarity, the disconnected nodes, which showed no physical or functional protein connections with genes in the core network, were buried.

4.8. GSMR analysis discloses weak causality between SCZ and cardiometabolic Diseases

Shared genetic architecture (pleiotropy) between SCZ and cardiometabolic traits was explored using multiple levels of evidence. GSMR was implemented to explore the causal relationship between the mentioned traits, adding the last piece to this mosaic. summarizes the results of GSMR analyses of SCZ with T2D, CMM, and obesity classes.

A small but significant negative causal effect was observed of SCZ on T2D (βSCZ→T2D= -0.01165, *P*=0.023), but not vice versa. CMM on SCZ showed a small but significant positive causal effect of CMM on SCZ (βCMM→SCZ=0.08958, *P*=0.040), but not vice versa. Additionally, a small but significant negative causal effects of SCZ on obesity classes were observed (βSCZ→OB1=-0.04656, $P=4.82\times10^{-3}$), (βSCZ→OB2=-0.07322, $P=6.14\times10^{-3}$), and (βSCZ→OB3=-0.17324, $P=7.91\times10^{-4}$), but not vice versa (**Figure 4-12**). While GSMR analysis from OB3 to SCZ was retained null due to insufficient number of instrumental variables (only 2 SNPs retained after filtration).

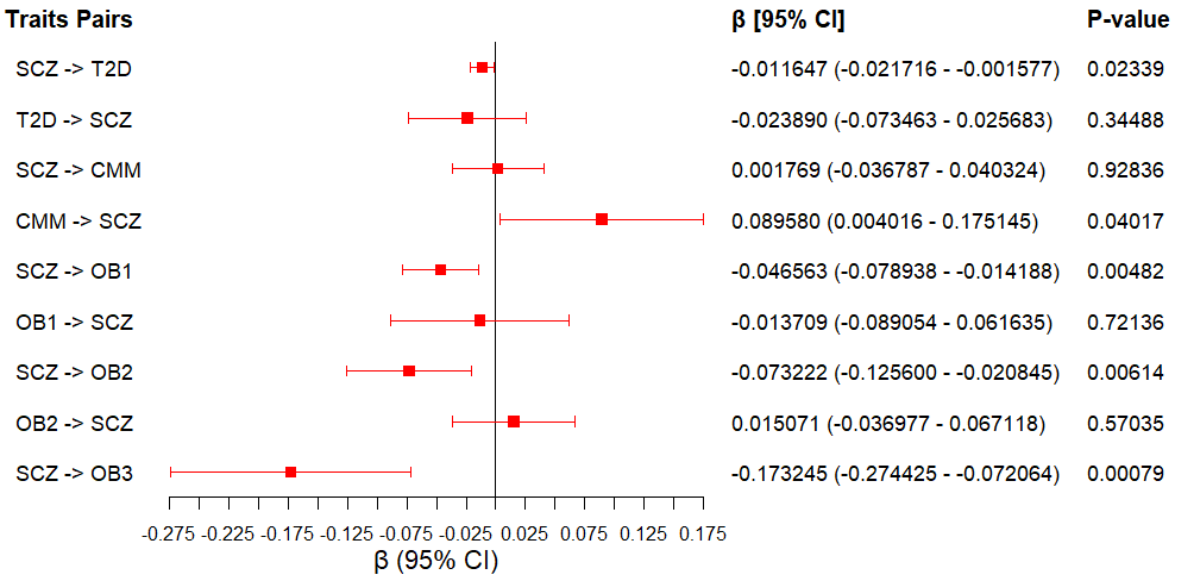


Figure 4-12. Forest plot of the GSMR results for the associations between SCZ and cardiometabolic disorders. This figure presents the pairwise results of GSMR using effect size (β) with 95% CI for pairs of traits. The squares represent the point estimates of the β, and the horizontal lines indicate the corresponding 95% CIs.

CHAPTER FIVE
Discussion

This study found genetic overlap between SCZ and other cardiometabolic traits, including T2D, CMM, and obesity classes. In line with previous research, this genetic overlap was confirmed through shared loci and common pathways, utilizing possible inflammatory, oxidative stress, and peripheral cardiometabolic alterations caused by dysregulated neuronal mechanisms. Nevertheless, utilizing an unconventional cross-trait meta-analysis approach (ASSET) along with comprehensive multi-trait fine-mapping and colocalization analyses gave more insight into these mechanisms as new shared loci and candidate genes were introduced.

SNP-level analyses unveiled several novel associations as well as colocalized signals between SCZ and T2D. Besides, new associations and pleiotropy were introduced for non-novel loci by ASSSET analysis. The regulatory predictions of the shared SNPs were comprehended through functional annotations. Remarkably, the majority of shared signals were in the regulatory regions with at least eQTL or a transcription factor binding site. The predominance of non-coding variants aligns with the current understanding that regulatory variants play a major role in complex traits.

However, careful assessment of SNPs' functionality is vital. As an example, rs5398 (3-170715830-G-A) is a missense variant previously reported to be associated with several metabolic traits, but has never been reported with SCZ. On one hand, the gene mapped to this SNP, SLC2A2, reduces GLUT2 expression and is linked to T2D and lower BDNF levels; on the other hand, greater BDNF levels lower T2D risk by exerting a protective effect on GLUT2 in pancreatic β-cells (Perry *et al.*, 2022). This example suggests a plausible shared mechanism of SLC2A2 in SCZ and T2D and provides a focal point regarding the importance of combining SNP-level and gene-level interpretations to understand shared pathways of complex traits.

All in all, SNP-based analysis findings were in line with the previous research findings. However, identifying novel common loci and the newly revealed pattern of pleiotropic associations suggests molecular connections between SCZ and metabolic disorders that were previously unidentified. This also applied to gene-level analysis findings, as a new pattern of association was reported. In addition, further investigation of genes with no GWAS enrichment for these traits is needed, including IL17REL, MDGA2, NLGN1, APOPTI (C14orf153), ARTNL, C12orf65, DNAH8, FAM212B, FAM47E, RBM6, SLC9B1, SLCO4A1, SPPL3, SYNDIG1L, DNAJC11, DNAH8, and TAP2.

Shedding light on some of these genes, IL17REL is possibly associated with T2D, through its role in the NF-κB pathway activation via LEPTIN, leading to the production of IL-1β, IL-6, and TNF-α, thereby contributing to insulin resistance (Abdel-Moneim, Bakery and Allam, 2018). However, no data is available regarding its possible role in either SCZ, CVD, or OB.

Complex, translational, and post-translational co-regulatory mechanisms were suggested based on the enrichment pattern found in gene sets of TF binding sites and microRNAs known to regulate neural development, stress response, lipid metabolism, metabolic processes, and immune and hormonal responses. Besides, the majority of candidate genes have alternative splicing mechanisms indicated more than one protein product for them and thus possible roles in different tissues or contexts.

Of the TFs that showed high enrichment in this analysis, STAT5 has been associated with regulating insulin production in pancreatic β-cells (Jackerott *et al.*, 2006). RREB1 has been linked to T2D; knockdown of RREB1 in adipocytes reduces fat cell formation and improves insulin sensitivity (Yu *et al.*, 2024). CEBPα has a crucial role in adipocyte biology; mice lacking adipose CEBPα develop abnormal fat distribution along with an increase in the hepatic triglycerides and plasma cholesterol, linking CEBPA to obesity-related dyslipidemia, a risk factor of CVD and atherosclerosis (Hu *et al.*, 2025). IRF1 (interferon-regulatory factor 1) expression is elevated in diabetic blood vessels, where it promotes endothelial inflammation and atherosclerotic changes (Zhou *et al.*, 2025).

Furthermore, a study of GWAS-reported cardiometabolic genes revealed that 80.6% of these genes are microRNA targets, and suggested a non-random convergence of genes to

microRNA, supporting the regulation of functionally related genes by the same microRNAs (Mustafa *et al.*, 2018). Notably, enriched microRNAs in this study were found to be linked to neurological disorders, T2D, OB, and adiposity. It has been suggested that miR-548 family members are associated with the risk of neurodevelopmental diseases, particularly miR-548 (Alsaqati *et al.*, 2022), miR-548ab shown to be significantly associated with fasting blood glucose and promotes obesity-related T2D (Pan *et al.*, 2025), and miR-362 shown to support brown fat differentiation and healthy adipose function, and thus indirectly protects against obesity and related metabolic disease (Mori *et al.*, 2014). Furthering findings related to TF binding sited and microRNAs enrichment obtained by this study in a meaningful tissue-specific context with advanced molecular techniques is the next required step.

This work contributed to a unifying interpretation: an inflammatory–metabolic–neurodevelopmental pathway may be shared by SCZ and cardiometabolic disorders. Pathway enrichment highlighted biological pathways, including pathways of insulin secretion, peptide transport, and hormone transport, the type II diabetes pathway (KEGG), HNF (Hepatocyte Nuclear Factor), regulation of neuron differentiation, central nervous system neuron development, Presenilin 1 action in Notch, Wnt signaling Pathways, and DNA damage by stress mechanisms. Moreover, CNTN2 (mapped to a colocalized SNP) was enriched in the pathway of protein colocalization to axon, highlighting its role in synaptic plasticity and neurotransmitter release, as well as in IL-18 signaling, thus its involvement in inflammatory responses via activation of PI3K/AKT signaling cascade.

Furthermore, A possible SCZ-cardiometabolic link via NLGN1-MDGA2-BDNF interaction was suggested. The rationale behind this is based on the pathway enrichment results that revealed enrichment of NLGN1 in pathways of glutamatergic synaptic regulation and organization, as well as the co-expression of NLGN1 with MDGA2 and with BDNF, which was revealed by STRING analysis. In addition, previously reported evidence linked MDAG2 to autism spectrum disorder (ASD), as well as to its role in suppressing excitatory synapse development by blocking the NLGN1 interaction with neurexins (Su *et al.*, 2018). Moreover, evidence from animal studies suggested that MDGA2 deficiency leads to activation of brain-derived neurotrophic factor/tyrosine kinase B (BDNF/TrkB) signaling, which is vital for

neuron survival, growth, differentiation, and synaptic plasticity (Zhao *et al.*, 2025). Peripherally, NLG1 is inhibited, and TNF-α, IL-1β, and IFN-γ are upregulated by overexpression of QKI-7 in blood, that mediated by excessive proliferation of endothelial cells and atherosclerosis mediated by T2D and high glucose (Xu *et al.*, 2022). All of these observations may give insights into the possible brain-cardiometabolic link of these genes.

ARTNTL1 of the circadian clock, also known as BMAL1, has shown suggestive evidence of association with psychiatric disorders (Prata *et al.*, 2019). Additionally, it is associated with the decreased expression of BDNF (Ali and von Gall, 2022), with insulin resistance through disruption of the circadian clock of beta cells of the pancreas, as well as with disturbed vascular endothelial regulation and adipogenesis (Ali and von Gall, 2022), implicating the association of ARNTL with T2D, OB, and CVD in humans (Škrlec *et al.*, 2019). Moreover, STRING analysis revealed co-occurrence and co-expression of ARNTL with EP300, which activates BMAL1/CLOCK in a redox-dependent manner (Healy, Morris and Liu, 2021). These findings suggest a potential association between ARNTL1 and these disorders.

STRING protein-protein interaction network further explained the functional relationships between genes associated with SCZ and cardiometabolic disorders. This analysis revealed a well-defined and highly interconnected network, implying that shared genes are not randomly dispersed but rather clustered into biologically relevant modules by unveiling distinct clusters including metabolic regulator genes, neuronal and synaptic genes, hormone signaling and MAP kinase genes, immune and MHC-region genes, with central hub genes that connect the clusters, which are BLK links to immune activation, PROX1 to developmental and possibly metabolic regulation; BDNF linked to epigenetic regulation, and EP300 connects multiple clusters.

An acetyltransferase involved in transcriptional control is encoded by EP300. Rises in body weight, total cholesterol, BMI, LDL, and TG concentrations linked with antipsychotic medication were all substantially correlated with EP300 expression levels. Simultaneously, EP300 is crucial for pancreatic islet development and insulin output (Martínez-Pinteño *et al.*, 2021). Furthermore, activation of the AKT signaling pathway via leptin stimulates histone acetyltransferase p300, leading to histone H3 acetylation and methylation at specific BDNF

promoters (Li *et al.*, 2021). On the other hand, FOXP1 plays a role in glucose homeostasis (Liu *et al.*, 2019). It also regulates human-specific aspects in early cortical development and transcriptional programs necessary for synaptic formation, and for regulating pathways including Wnt/β-catenin and MAPK signaling. Finally, BDNF/TrkB neurotrophic signaling has a crucial role in regulating food intake, while inadequate BDNF/TrkB activity promotes neurodegeneration by activating the JAK2/STAT3 pathway and elevating inflammatory cytokines in the human brain (Flores-Dorantes, Díaz-López and Gutiérrez-Aguilar, 2020).

Together, EP300, PROX1, and BLK comprise three complementary nodes that link important biological subsystems involved in SCZ and cardiometabolic diseases. EP300 functions as a chromatin co-activator and transcriptional integrator in energy metabolism, lipid homeostasis, and neural plasticity, bridging gene regulation in the brain and peripheral organs. PROX1, a transcription factor required for pancreatic and hepatic development, promotes glucose metabolism and vascular integrity, enhancing metabolic-neurodevelopmental crosstalk. In contrast, BLK acts as a signaling kinase at the interface of immunological and endocrine control, regulating both inflammatory responses and insulin production.

The convergence of these hub genes within the same network suggests a multi-layered molecular framework in which epigenetic regulation (EP300), developmental transcriptional control (PROX1), and immune-metabolic signaling (BLK) all influence susceptibility to psychiatric and metabolic pathology. This integrative model provides biological credibility to the observed genetic associations and emphasizes the need to take into account systemic mechanisms rather than isolated organ-specific effects when studying comorbid SCZ and cardiometabolic disease.

Overall, pathway analyses of the candidate genes implicated synaptic function, neurodevelopment, cellular processes, metabolic, immune, and inflammatory mechanisms. These results correspond to neuronal and immune system contributions to SCZ pathogenesis. Other mechanisms of central regulation of food intake and insulin resistance were also implicated (He *et al.*, 2022; Rødevand *et al.*, 2023).

These findings reinforce the hypothesis that diet, lifestyle, and antipsychotic treatment are the main contributors to the increased risk of cardiometabolic traits. However, the elevated risk of cardiometabolic risk in patients at their first psychotic episode or in antipsychotics-naïve patients could be explained by immune system activation, inflammation, and oxidative stress mechanisms during the prodromal phase before starting the treatment or even diagnosis. The prodromal phase lasts for 1 year before the first psychotic episode and manifests with changes in cognition, behavior, and mood due to GABAergic alterations, NDMA, and dopamine dysfunction. Increasingly, altered synaptogenesis, synaptic pruning, and glutathione deficit occur in childhood and early adolescence (McCutcheon, Reis Marques and Howes, 2020; Mandal *et al.*, 2022). These early neurological changes might contribute to cardiometabolic risk as a result of neural-cardiometabolic pathways introduced via pathway analysis. This possibility could assist in elucidating observations of cardiometabolic abnormalities in drug-naive patients with SCZ and patients at their first psychotic episode.

LDSC indicated a modest negative correlation, consistent with the previous evidence in the direction of correlation between SCZ and cardiometabolic traits (Zammarchi, Conversano and Pisanu, 2022; Ding *et al.*, 2023; Rødevand *et al.*, 2023). This negative correlation could be explained by diverging effect directions across loci associated with SCZ and cardiometabolic traits.

Furthermore, differentially expressed genes (DEG) showed expression patterns between brain and cardiometabolic tissues that could reflect complex genetic interactions, including heterogeneous regulation and parallel activation or suppression of different sets of genes. Demonstrating possible tissue-level specialization or central-peripheral bidirectional transcriptional mechanisms.

GSMR unveiled a small but significant negative causal effect of SCZ on T2D and obesity classes, but not vice versa, and a small but significant positive causal effect of CMM on SCZ but not vice versa. These findings suggest a weak possible causal link. However, vigilant interpretations of the directions of these correlations are required due to the complex genetic interplay shown in this analysis, warranting further investigation and experimental validation to verify these results.

Strengths and Limitations

The major strength of this master's thesis is represented by implicating multiple statistical approaches, which results in a comprehensive assessment of the topic area. Additionally, using large-scale GWAS data and large sample sizes contributes to enhanced discovery.

Furthermore, ASSET cross-trait meta-analysis fosters shared loci identification, as it employs a novel approach to conduct cross-trait fixed-effect meta-analysis by exploring different subsets of multiple traits to identify true association signals that could be present in one subset of the phenotypes, either in the same or opposite directions. Additionally, applying a multi-trait fine-mapping technique that jointly analyzes GWAS data of multiple traits enables the identification of the shared putative causative SNPs.

Employing enrichment analyses enables focusing on assessing the pooled effects of various SNPs and genes, rather than single-gene assessment, which contributes to enhanced clarity and a comprehensive vision in exploring the potential biological mechanisms, functions, or processes underlying complex disorders.

Limitations of this study include that the GWAS data in this project were mainly from European ancestry; hence, generalizing the findings to other ancestries might not be applicable. Second, the limited colocalization and modest genetic correlations observed do not negate shared biology, but rather reflect the complexity and heterogeneity of psychiatric metabolic genetic architectures, and reflect the importance of advanced colocalization techniques, integrating QTL data. Third, while FUMA provides valuable insights into the functional analyses of GWAS signals, further experimental validation is necessary to confirm these findings. Finally, GSMR analysis was not validated; multiple MR methods and colocalization analysis are needed.

Conclusion

Large-scale GWASs provide an enormous prospect to explore the shared genetics between diseases. This analysis applied different statistical approaches utilizing large-scale GWASs to investigate the shared genetic background between SCZ and cardiometabolic disorders.

This study provided multi-layered evidence of shared genetic background and introduced novel associations between SCZ and cardiometabolic disorders. Furthermore, functional enrichment analyses revealed some genetically controlled biological pathways underlying SCZ and cardiometabolic disorders, and revealed complex transcriptional and epigenetic regulation. Additionally, GSMR analysis unveils small protective effects of SCZ (exposure) on T2D and obesity classes (outcome), while CMM (exposure) was found to be a risk factor for SCZ (outcome). Hence, further analyses are needed to validate these results and to identify shared causal variants.

This study highlights the importance of systems-level approaches to psychiatric genomics, which bridge the gap between brain and peripheral physiology. Continued investigation of these overlapping processes could increase our understanding of comorbidity, refine predictive models, and inform targeted therapies for those at risk of both mental and metabolic disorders.

References

Abdel-Moneim, A., Bakery, H.H. and Allam, G. (2018) "The potential pathogenic role of IL-17/Th17 cells in both type 1 and type 2 diabetes mellitus," *Biomedicine and Pharmacotherapy*, 101, pp. 287–292. Available at: https://doi.org/10.1016/j.biopha.2018.02.103.

Åberg, K. *et al.* (2012) "Genome-wide association study of antipsychotic-induced QTc interval prolongation," *The pharmacogenomics journal*, 12(2), pp. 165–172. Available at: https://doi.org/10.1038/TPJ.2010.76.

Adebiyi, E. *et al.* (2021) "Performing post-genome-wide association study analysis: overview, challenges and recommendations," *F1000Research*, 10, p. 1002. Available at: https://doi.org/10.12688/F1000RESEARCH.53962.1.

Ali, A.A.H. and von Gall, C. (2022) "Adult Neurogenesis under Control of the Circadian System," *Cells*, 11(5). Available at: https://doi.org/10.3390/CELLS11050764.

Alsaqati, M. *et al.* (2022) "NRSF/REST lies at the intersection between epigenetic regulation, miRNA-mediated gene control and neurodevelopmental pathways associated with Intellectual disability (ID) and Schizophrenia," *Translational psychiatry*, 12(1). Available at: https://doi.org/10.1038/S41398-022-02199-Z.

Amare, A.T. *et al.* (2017) "The genetic overlap between mood disorders and cardiometabolic diseases: a systematic review of genome wide and candidate gene studies," *Translational psychiatry*, 7(1). Available at: https://doi.org/10.1038/TP.2016.261.

Arruda, A.L. *et al.* (2024) "Genomic insights into the comorbidity between type 2 diabetes and schizophrenia," *Schizophrenia 2024 10:1*, 10(1), pp. 22-. Available at: https://doi.org/10.1038/s41537-024-00445-5.

Azad, M.C. *et al.* (2016) "Cardiovascular diseases among patients with schizophrenia," *Asian Journal of Psychiatry*, 19, pp. 28–36. Available at: https://doi.org/10.1016/J.AJP.2015.11.012.

Barry, C.J.S. *et al.* (2022) "How to estimate heritability: a guide for genetic epidemiologists," *International Journal of Epidemiology*, 52(2), p. 624. Available at: https://doi.org/10.1093/IJE/DYAC224.

Berndt, S.I. *et al.* (2013) "Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture," *Nature genetics*, 45(5), pp. 501–512. Available at: https://doi.org/10.1038/NG.2606.

Bhattacharjee, S. *et al.* (2012) "A Subset-Based Approach Improves Power and Interpretation for the Combined Analysis of Genetic Association Studies of Heterogeneous Traits," *American Journal of Human Genetics*, 90(5), p. 821. Available at: https://doi.org/10.1016/J.AJHG.2012.03.015.

Biedermann, F. and Fleischhacker, W.W. (2016) "Psychotic disorders in DSM-5 and ICD-11," *CNS Spectrums*, 21(4), pp. 349–354. Available at: https://doi.org/10.1017/S1092852916000316.

Broekema, R. V., Bakker, O.B. and Jonkers, I.H. (2020) "A practical view of fine-mapping and gene prioritization in the post-genome-wide association era," *Open Biology*, 10(1). Available at: https://doi.org/10.1098/RSOB.190221;PAGE:STRING:ARTICLE/CHAPTER.

Bulik-Sullivan, B., Finucane, H.K., *et al.* (2015) "An Atlas of Genetic Correlations across Human Diseases and Traits," *Nature genetics*, 47(11), p. 1236. Available at: https://doi.org/10.1038/NG.3406.

Bulik-Sullivan, B., Loh, P.R., *et al.* (2015) "LD Score Regression Distinguishes Confounding from Polygenicity in Genome-Wide Association Studies," *Nature genetics*, 47(3), p. 291. Available at: https://doi.org/10.1038/NG.3211.

Chao, A.M., Wadden, T.A. and Berkowitz, R.I. (2019) "Obesity in Adolescents with Psychiatric Disorders," *Current Psychiatry Reports*, 21(1), pp. 3-. Available at: https://doi.org/10.1007/S11920-019-0990-7/METRICS.

Chen, X. *et al.* (2011) "GWA study data mining and independent replication identify cardiomyopathy-associated 5 (CMYA5) as a risk gene for schizophrenia," *Molecular psychiatry*, 16(11), pp. 1117–1129. Available at: https://doi.org/10.1038/MP.2010.96.

Correll, C.U. *et al.* (2022) "Systematic literature review of schizophrenia clinical practice guidelines on acute and maintenance management with antipsychotics," *Schizophrenia 2022 8:1*, 8(1), pp. 1–10. Available at: https://doi.org/10.1038/s41537-021-00192-x.

Correll, C.U. and Schooler, N.R. (2020) "Negative Symptoms in Schizophrenia: A Review and Clinical Guide for Recognition, Assessment, and Treatment," *Neuropsychiatric Disease and Treatment*, 16, p. 519. Available at: https://doi.org/10.2147/NDT.S225643.

Deng, Y. *et al.* (2020) "Genome-Wide Gene-Based Multi-Trait Analysis," *Frontiers in Genetics*, 11, p. 437. Available at: https://doi.org/10.3389/FGENE.2020.00437/FULL.

Dieset, I., Andreassen, O.A. and Haukvik, U.K. (2016) "Somatic Comorbidity in Schizophrenia: Some Possible Biological Mechanisms Across the Life Span," *Schizophrenia bulletin*, 42(6), pp. 1316–1319. Available at: https://doi.org/10.1093/SCHBUL/SBW028.

Ding, H. *et al.* (2023) "Shared genetics of psychiatric disorders and type 2 diabetes:a large-scale genome-wide cross-trait analysis," *Journal of Psychiatric Research*, 159, pp. 185–195. Available at: https://doi.org/10.1016/j.jpsychires.2023.01.037.

Evangelista, J.E. *et al.* (2023) "Enrichr-KG: bridging enrichment analysis across multiple libraries," *Nucleic Acids Research*, 51(W1), pp. W168–W179. Available at: https://doi.org/10.1093/NAR/GKAD393.

Fiorillo, A. and Sartorius, N. (2021) "Mortality gap and physical comorbidity of people with severe mental disorders: the public health scandal," *Annals of General Psychiatry 2021 20:1*, 20(1), pp. 52-. Available at: https://doi.org/10.1186/S12991-021-00374-Y.

Fišar, Z. (2023) "Biological hypotheses, risk factors, and biomarkers of schizophrenia," *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 120, p. 110626. Available at: https://doi.org/10.1016/J.PNPBP.2022.110626.

Flores-Dorantes, M.T., Díaz-López, Y.E. and Gutiérrez-Aguilar, R. (2020) "Environment and Gene Association With Obesity and Their Impact on Neurodegenerative and Neurodevelopmental Diseases," *Frontiers in Neuroscience*, 14, p. 565326. Available at: https://doi.org/10.3389/FNINS.2020.00863/FULL.

Fry, H. *et al.* (2024) "The Role of Furin and Its Therapeutic Potential in Cardiovascular Disease Risk," *International Journal of Molecular Sciences*, 25(17), p. 9237. Available at: https://doi.org/10.3390/IJMS25179237.

Goldfarb, M. *et al.* (2022) "Severe Mental Illness and Cardiovascular Disease: JACC State-of-the-Art Review," *Journal of the American College of Cardiology*, 80(9), pp. 918–933. Available at: https://doi.org/10.1016/J.JACC.2022.06.017;SUBPAGE:STRING:FULL.

Gragnoli, C. *et al.* (2016) "Dopamine–prolactin pathway potentially contributes to the schizophrenia and type 2 diabetes comorbidity," *Translational Psychiatry 2016 6:4*, 6(4), pp. e785–e785. Available at: https://doi.org/10.1038/tp.2016.50.

Guerrero Fernández de Alba, I. *et al.* (2020) "Association between mental health comorbidity and health outcomes in type 2 diabetes mellitus patients," *Scientific Reports 2020 10:1*, 10(1), pp. 19583-. Available at: https://doi.org/10.1038/s41598-020-76546-9.

He, Q. *et al.* (2022) "Exploring Lead loci shared between schizophrenia and Cardiometabolic traits," *BMC Genomics 2022 23:1*, 23(1), pp. 617-. Available at: https://doi.org/10.1186/S12864-022-08766-4.

Healy, K.L., Morris, A.R. and Liu, A.C. (2021) "Circadian Synchrony: Sleep, Nutrition, and Physical Activity," *Frontiers in network physiology*, 1. Available at: https://doi.org/10.3389/FNETP.2021.732243.

Henriksen, M.G., Nordgaard, J. and Jansson, L.B. (2017) "Genetics of schizophrenia: Overview of methods, findings and limitations," *Frontiers in Human Neuroscience*, 11, p. 250542. Available at: https://doi.org/10.3389/FNHUM.2017.00322/BIBTEX.

Hu, K.Y. *et al.* (2025) "Depot-Specific Roles for C/EBPα in White Adipose Tissue Development and Metabolism," *bioRxiv : the preprint server for biology* [Preprint]. Available at: https://doi.org/10.1101/2025.07.22.666007.

Huang, C.J. *et al.* (2018) "Schizophrenia in type 2 diabetes mellitus: Prevalence and clinical characteristics," *European Psychiatry*, 54, pp. 102–108. Available at: https://doi.org/10.1016/J.EURPSY.2018.08.002.

Huertas-Vazquez, A. *et al.* (2013) "A common missense variant in the neuregulin 1 gene is associated with both schizophrenia and sudden cardiac death," *Heart Rhythm*, 10(7), pp. 994–998. Available at: https://doi.org/10.1016/j.hrthm.2013.03.020.

Ignatieva, E. V. *et al.* (2016) "A compendium of human genes regulating feeding behavior and body weight, its functional characterization and identification of GWAS genes involved in brain-specific PPI network," *BMC Genetics*, 17(Suppl 3), p. 158. Available at: https://doi.org/10.1186/S12863-016-0466-2.

Jackerott, M. *et al.* (2006) "STAT5 activity in pancreatic beta-cells influences the severity of diabetes in animal models of type 1 and 2 diabetes," *Diabetes*, 55(10), pp. 2705–2712. Available at: https://doi.org/10.2337/DB06-0244.

Jang, S.K. *et al.* (2020) "Genetic correlation, pleiotropy, and causal associations between substance use and psychiatric disorder," *Psychological medicine*, 52(5), p. 968. Available at: https://doi.org/10.1017/S003329172000272X.

Jauhar, S., Johnstone, M. and McKenna, P.J. (2022) "Schizophrenia," *The Lancet*, 399(10323), pp. 473–486. Available at: https://doi.org/10.1016/S0140-6736(21)01730-X.

Karatzas, E. *et al.* (2023) "Flame (v2.0): advanced integration and interpretation of functional enrichment results from multiple sources," *Bioinformatics*, 39(8), p. btad490. Available at: https://doi.org/10.1093/BIOINFORMATICS/BTAD490.

Kharabian Masouleh, S. *et al.* (2016) "Higher body mass index in older adults is associated with lower gray matter volume: Implications for memory performance," *Neurobiology of Aging*, 40, pp. 1–10. Available at: https://doi.org/10.1016/j.neurobiolaging.2015.12.020.

Kircher, M. *et al.* (2014) "A general framework for estimating the relative pathogenicity of human genetic variants," *Nature genetics*, 46(3), p. 310. Available at: https://doi.org/10.1038/NG.2892.

Lam, M. *et al.* (2019) "Comparative genetic architectures of schizophrenia in East Asian and European populations," *Nature genetics*, 51(12), pp. 1670–1678. Available at: https://doi.org/10.1038/S41588-019-0512-X.

de Leeuw, C.A. *et al.* (2015) "MAGMA: Generalized Gene-Set Analysis of GWAS Data," *PLoS Computational Biology*, 11(4), p. e1004219. Available at: https://doi.org/10.1371/JOURNAL.PCBI.1004219.

Legge, S.E. *et al.* (2021) "Genetic architecture of schizophrenia: a review of major advancements," *Psychological Medicine*, 51(13), pp. 2168–2177. Available at: https://doi.org/10.1017/S0033291720005334.

Li, C. *et al.* (2021) "Leptin regulates exon-specific transcription of the Bdnf gene via epigenetic modifications mediated by an AKT/p300 HAT cascade," *Molecular psychiatry*, 26(8), pp. 3701–3722. Available at: https://doi.org/10.1038/S41380-020-00922-0.

Li, X. and Zhu, X. (2017) "Cross-phenotype association analysis using summary statistics from GWAS," *Methods in molecular biology (Clifton, N.J.)*, 1666, p. 455. Available at: https://doi.org/10.1007/978-1-4939-7274-6_22.

Lin, P.I. and Shuldiner, A.R. (2010) "Rethinking the genetic basis for comorbidity of schizophrenia and type 2 diabetes," *Schizophrenia Research*, 123(2–3), pp. 234–243. Available at: https://doi.org/10.1016/j.schres.2010.08.022.

Lindekilde, N. *et al.* (2021) "Prevalence of type 2 diabetes in psychiatric disorders: an umbrella review with meta-analysis of 245 observational studies from 32 systematic reviews," *Diabetologia 2021 65:3*, 65(3), pp. 440–456. Available at: https://doi.org/10.1007/S00125-021-05609-X.

Lis, M. *et al.* (2020) "Impaired hormonal regulation of appetite in schizophrenia: A narrative review dissecting intrinsic mechanisms and the effects of antipsychotics," *Psychoneuroendocrinology*, 119. Available at: https://doi.org/10.1016/j.psyneuen.2020.104744.

Liu, P. *et al.* (2019) "Foxp1 controls brown/beige adipocyte differentiation and thermogenesis through regulating β3-AR desensitization," *Nature communications*, 10(1). Available at: https://doi.org/10.1038/S41467-019-12988-8.

Locke, A.E. *et al.* (2015) "Genetic studies of body mass index yield new insights for obesity biology," *Nature*, 518(7538), pp. 197–206. Available at: https://doi.org/10.1038/NATURE14177.

Lopez, E.O., Ballard, B.D. and Jan, A. (2023) "Cardiovascular Disease," *StatPearls Publishing* [Preprint]. Available at: https://www.ncbi.nlm.nih.gov/books/NBK535419/ (Accessed: November 16, 2025).

Mahajan, A. *et al.* (2018) "Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps," *Nature genetics*, 50(11), pp. 1505–1513. Available at: https://doi.org/10.1038/S41588-018-0241-6.

Mandal, P.K. *et al.* (2022) "Schizophrenia, Bipolar and Major Depressive Disorders: Overview of Clinical Features, Neurotransmitter Alterations, Pharmacological Interventions, and Impact of Oxidative Stress in the Disease Process," *ACS chemical neuroscience*, 13(19), pp. 2784–2802. Available at: https://doi.org/10.1021/ACSCHEMNEURO.2C00420.

Martínez-Pinteño, A. *et al.* (2021) "Identification of EP300 as a Key Gene Involved in Antipsychotic-Induced Metabolic Dysregulation Based on Integrative Bioinformatics Analysis of Multi-Tissue Gene Expression Data," *Frontiers in Pharmacology*, 12. Available at: https://doi.org/10.3389/FPHAR.2021.729474/PDF.

McCutcheon, R.A., Reis Marques, T. and Howes, O.D. (2020) "Schizophrenia—An Overview," *JAMA Psychiatry*, 77(2), pp. 201–210. Available at: https://doi.org/10.1001/JAMAPSYCHIATRY.2019.3360.

Millan, M.J. *et al.* (2016) "Altering the course of schizophrenia: progress and perspectives," *Nature reviews. Drug discovery*, 15(7), pp. 485–515. Available at: https://doi.org/10.1038/NRD.2016.28.

Mizuki, Y. *et al.* (2021) "Mechanisms Underlying the Comorbidity of Schizophrenia and Type 2 Diabetes Mellitus," *International Journal of Neuropsychopharmacology*, 24(5), pp. 367–382. Available at: https://doi.org/10.1093/IJNP/PYAA097.

Mori, M.A. *et al.* (2014) "Altered miRNA processing disrupts brown/white adipocyte determination and associates with lipodystrophy," *The Journal of clinical investigation*, 124(8), pp. 3339–3351. Available at: https://doi.org/10.1172/JCI73468.

Mortezaei, Z. and Tavallaei, M. (2021) "Recent innovations and in-depth aspects of post-genome wide association study (Post-GWAS) to understand the genetic basis of complex phenotypes," *Heredity 2021 127:6*, 127(6), pp. 485–497. Available at: https://doi.org/10.1038/s41437-021-00479-w.

Mustafa, R. *et al.* (2018) "An Enrichment Analysis for Cardiometabolic Traits Suggests Non-Random Assignment of Genes to microRNAs," *International journal of molecular sciences*, 19(11). Available at: https://doi.org/10.3390/IJMS19113666.

Myers, T.A., Chanock, S.J. and Machiela, M.J. (2020) "LDlinkR: An R Package for Rapidly Calculating Linkage Disequilibrium Statistics in Diverse Populations," *Frontiers in Genetics*, 11, p. 513535. Available at: https://doi.org/10.3389/FGENE.2020.00157/BIBTEX.

Nagalski, A., Kozinski, K. and Wisniewska, M.B. (2016) "Metabolic pathways in the periphery and brain: Contribution to mental disorders?," *International Journal of Biochemistry and Cell Biology*, 80, pp. 19–30. Available at: https://doi.org/10.1016/j.biocel.2016.09.012.

National Academies of Sciences, E. and M.H. and M.D.B. on H.C.S.C. on I.D.M.C.L. to I. with T. (2020) "Mental Health Disorders," *Selected Health Conditions and Likelihood of Improvement with Treatment* [Preprint]. Available at: https://doi.org/10.17226/25662.

Nordestgaard, B.G. (2016) "Triglyceride-Rich Lipoproteins and Atherosclerotic Cardiovascular Disease: New Insights from Epidemiology, Genetics, and Biology," *Circulation Research*, 118(4), pp. 547–563. Available at: https://doi.org/10.1161/CIRCRESAHA.115.306249;WGROUP:STRING:PUBLICATION.

Pan, C. *et al.* (2025) "Integrated analysis reveals that miR-548ab promotes the development of obesity and T2DM," *Journal of Genetics and Genomics*, 52(2), pp. 231–244. Available at: https://doi.org/10.1016/j.jgg.2024.11.011.

Pan, S. *et al.* (2023) "COLOCdb: a comprehensive resource for multi-model colocalization of complex traits," *Nucleic Acids Research*, 52(D1), p. D871. Available at: https://doi.org/10.1093/NAR/GKAD939.

Patel, K.R. *et al.* (2014) "Schizophrenia: Overview and Treatment Options," *Pharmacy and Therapeutics*, 39(9), p. 638. Available at: https://pmc.ncbi.nlm.nih.gov/articles/PMC4159061/ (Accessed: November 11, 2025).

Perry, B.I. *et al.* (2022) "Evidence for Shared Genetic Aetiology Between Schizophrenia, Cardiometabolic, and Inflammation-Related Traits: Genetic Correlation and Colocalization Analyses," *Schizophrenia bulletin open*, 3(1). Available at: https://doi.org/10.1093/SCHIZBULLOPEN/SGAC001.

Pilalis, E. *et al.* (2025) "Genome-wide functional annotation of variants: a systematic review of state-of-the-art tools, techniques and resources," *Frontiers in Pharmacology*, 16, p. 1474026. Available at: https://doi.org/10.3389/FPHAR.2025.1474026/BIBTEX.

Polcwiartek, C. *et al.* (2025) "Prevalence of cardiovascular risk factors and disease in patients with schizophrenia: baseline results from a prospective cohort study with long-term clinical follow-up," *Schizophrenia 2025 11:1*, 11(1), pp. 95-. Available at: https://doi.org/10.1038/s41537-025-00642-w.

Prata, D.P. *et al.* (2019) "Unravelling the genetic basis of schizophrenia and bipolar disorder with GWAS: A systematic review," *Journal of Psychiatric Research*, 114, pp. 178–207. Available at: https://doi.org/10.1016/j.jpsychires.2019.04.007.

Purcell, S. *et al.* (2007) "PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses," *American Journal of Human Genetics*, 81(3), p. 559. Available at: https://doi.org/10.1086/519795.

Qin, H. *et al.* (2016) "Inhibition of the JAK/STAT Pathway Protects Against α-Synuclein-Induced Neuroinflammation and Dopaminergic Neurodegeneration," *The Journal of Neuroscience*, 36(18), p. 5144. Available at: https://doi.org/10.1523/JNEUROSCI.4658-15.2016.

Rees, E., O'Donovan, M.C. and Owen, M.J. (2015) "Genetics of schizophrenia," *Current Opinion in Behavioral Sciences*, 2, pp. 8–14. Available at: https://doi.org/10.1016/J.COBEHA.2014.07.001.

Ringen, P.A. *et al.* (2014) "Increased Mortality in Schizophrenia Due to Cardiovascular Disease – A Non-Systematic Review of Epidemiology, Possible Causes, and Interventions," *Frontiers in Psychiatry*, 5(SEP), p. 137. Available at: https://doi.org/10.3389/FPSYT.2014.00137.

Rødevand, L. *et al.* (2023) "Characterizing the Shared Genetic Underpinnings of Schizophrenia and Cardiovascular Disease Risk Factors," *American Journal of Psychiatry*, 180(11), pp. 815–826. Available at: https://doi.org/10.1176/APPI.AJP.20220660;REQUESTEDJOURNAL:JOURNAL:AJP;WGROUP:STRING:PUBLICATION.

Schaid, D.J., Chen, W. and Larson, N.B. (2018) "From genome-wide associations to candidate causal variants by statistical fine-mapping," *Nature reviews. Genetics*, 19(8), pp. 491–504. Available at: https://doi.org/10.1038/S41576-018-0016-Z.

Schneider, F. *et al.* (2019) "Mortality and Medical Comorbidity in the Severely Mentally Ill: A German Registry Study," *Deutsches Ärzteblatt International*, 116(23–24), p. 405. Available at: https://doi.org/10.3238/ARZTEBL.2019.0405.

Seidman, L.J. and Mirsky, A.F. (2017) "Evolving Notions of Schizophrenia as a Developmental Neurocognitive Disorder," *Journal of the International Neuropsychological Society*, 23(9–10), pp. 881–892. Available at: https://doi.org/10.1017/S1355617717001114.

Shi, J. *et al.* (2024) "Evidence based on Mendelian randomization and colocalization analysis strengthens causal relationships between structural changes in specific brain regions and risk of amyotrophic lateral sclerosis," *Frontiers in Neuroscience*, 18, p. 1333782. Available at: https://doi.org/10.3389/fnins.2024.1333782.

Škrlec, I. *et al.* (2019) "Circadian clock genes and myocardial infarction in patients with type 2 diabetes mellitus," *Gene*, 701, pp. 98–103. Available at: https://doi.org/10.1016/j.gene.2019.03.038.

Srivastava, A.K., Williams, S.M. and Zhang, G. (2023) "Heritability estimation approaches utilizing genome-wide data," *Current protocols*, 3(4), p. e734. Available at: https://doi.org/10.1002/CPZ1.734.

Stubbs, B. *et al.* (2015) "The prevalence and predictors of type two diabetes mellitus in people with schizophrenia: a systematic review and comparative meta-analysis," *Acta psychiatrica Scandinavica*, 132(2), pp. 144–157. Available at: https://doi.org/10.1111/ACPS.12439.

Su, S. *et al.* (2018) "Combined QTL and Genome Scan Analyses With the Help of 2b-RAD Identify Growth-Associated Genetic Markers in a New Fast-Growing Carp Strain," *Frontiers in genetics*, 9. Available at: https://doi.org/10.3389/FGENE.2018.00592.

Sullivan, P.F. and Geschwind, D.H. (2019) "Defining the Genetic, Genomic, Cellular, and Diagnostic Architectures of Psychiatric Disorders," *Cell*, 177(1), pp. 162–183. Available at: https://doi.org/10.1016/J.CELL.2019.01.015.

Sullivan, P.F., Yao, S. and Hjerling-Leffler, J. (2024) "Schizophrenia genomics: genetic complexity and functional insights," *Nature Reviews Neuroscience 2024 25:9*, 25(9), pp. 611–624. Available at: https://doi.org/10.1038/s41583-024-00837-7.

Suvisaari, J. *et al.* (2016) "Diabetes and Schizophrenia," *Current diabetes reports*, 16(2), pp. 1–10. Available at: https://doi.org/10.1007/S11892-015-0704-4.

Szklarczyk, D. *et al.* (2022) "The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest," *Nucleic Acids Research*, 51(D1), p. D638. Available at: https://doi.org/10.1093/NAR/GKAC1000.

Tziastoudi, M. *et al.* (2019) "Identification of Chromosomal Regions Linked to Diabetic Nephropathy: A Meta-Analysis of Genome-Wide Linkage Scans," *Genetic testing and molecular biomarkers*, 23(2), pp. 105–117. Available at: https://doi.org/10.1089/GTMB.2018.0209.

Vancampfort, D. *et al.* (2016) "Diabetes mellitus in people with schizophrenia, bipolar disorder and major depressive disorder: a systematic review and large scale meta-analysis," *World Psychiatry*, 15(2), pp. 166–174. Available at: https://doi.org/10.1002/WPS.20309.

Velligan, D.I. and Rao, S. (2023) "The Epidemiology and Global Burden of Schizophrenia," *Journal of Clinical Psychiatry*, 84(1). Available at: https://doi.org/10.4088/JCP.MS21078COM5.

Wallace, C. (2021) "A more accurate method for colocalisation analysis allowing for multiple causal variants," *PLoS Genetics*, 17(9), p. e1009440. Available at: https://doi.org/10.1371/JOURNAL.PGEN.1009440.

Wang, J. *et al.* (2013) "Association of rs12255372 in the TCF7L2 gene with type 2 diabetes mellitus: a meta-analysis," *Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e biologica*, 46(4), pp. 382–393. Available at: https://doi.org/10.1590/1414-431X20132677.

Watanabe, K. *et al.* (2017) "Functional mapping and annotation of genetic associations with FUMA," *Nature Communications 2017 8:1*, 8(1), pp. 1–11. Available at: https://doi.org/10.1038/s41467-017-01261-5.

Xu, J. *et al.* (2022) "Regulation of the Proliferation of Diabetic Vascular Endothelial Cells by Degrading Endothelial Cell Functional Genes with QKI-7," *Contrast media & molecular imaging*, 2022. Available at: https://doi.org/10.1155/2022/6177809.

Yu, G.Z. *et al.* (2024) "Loss of RREB1 reduces adipogenesis and improves insulin sensitivity in mouse and human adipocytes," *bioRxiv : the preprint server for biology* [Preprint]. Available at: https://doi.org/10.1101/2024.07.30.605923.

Yu, Yuefeng *et al.* (2023) "Investigating the shared genetic architecture between schizophrenia and body mass index," *Molecular Psychiatry 2023 28:6*, 28(6), pp. 2312–2319. Available at: https://doi.org/10.1038/s41380-023-02104-0.

Zammarchi, G., Conversano, C. and Pisanu, C. (2022) "Investigating Shared Genetic Bases between Psychiatric Disorders, Cardiometabolic and Sleep Traits Using K-Means Clustering and Local Genetic Correlation Analysis," *Algorithms*, 15(11). Available at: https://doi.org/10.3390/A15110409/S1.

Zhang, X. *et al.* (2013) "The type 2 diabetes mellitus susceptibility gene IGF2BP2 is associated with schizophrenia in a Han Chinese population," *The Journal of clinical psychiatry*, 74(4). Available at: https://doi.org/10.4088/JCP.12M07846.

Zhao, C. *et al.* (2024) "Genome-wide association study of cardiometabolic multimorbidity in the UK Biobank," *Clinical genetics*, 106(1), pp. 72–81. Available at: https://doi.org/10.1111/CGE.14513.

Zhao, D. *et al.* (2025) "Mdga2 deficiency leads to an aberrant activation of BDNF/TrkB signaling that underlies autism-relevant synaptic and behavioral changes in mice," *PLoS biology*, 23(4). Available at: https://doi.org/10.1371/JOURNAL.PBIO.3003047.

Zhou, H. *et al.* (2023) "FAVOR: functional annotation of variants online resource and annotator for variation across the human genome," *Nucleic Acids Research*, 51(D1), pp. D1300–D1311. Available at: https://doi.org/10.1093/NAR/GKAC966.

Zhou, Y. *et al.* (2025) "Interferon Regulatory Factor 1 Promotes Diabetic Vasculopathy by Mediating Endothelial Cell Pyroptosis," *Annals of Vascular Surgery*, 115, pp. 107–116. Available at: https://doi.org/10.1016/j.avsg.2025.02.018.

Zhu, B. *et al.* (2018) "Furin promotes dendritic morphogenesis and learning and memory in transgenic mice," *Cellular and molecular life sciences : CMLS*, 75(13), pp. 2473–2488. Available at: https://doi.org/10.1007/S00018-017-2742-3.

Zhu, H. and Zhou, X. (2020) "Statistical methods for SNP heritability estimation and partition: A review," *Computational and Structural Biotechnology Journal*, 18, p. 1557. Available at: https://doi.org/10.1016/J.CSBJ.2020.06.011.

Zou, Y. *et al.* (2024) “Fast and flexible joint fine-mapping of multiple traits via the Sum of Single Effects model,” *bioRxiv*, p. 2023.04.14.536893. Available at: https://doi.org/10.1101/2023.04.14.536893.

Appendix1

Supplementary Table 4-1a. ASSET 'subset1sided' lead SNPs.

SNP	CHR	BP	P	Pheno (ASSET)	Previous associations	Previous Multitrait associations	Previous associations (related)
rs12196300	**6**	**38963764**	**9.95E-09**	**SCZ,T2D,OB2,OB3,CMM**	**NA**	**NA**	**NA**
rs1343424	**1**	**49376820**	**7.10E-09**	**SCZ,OB1,OB2,OB3**	**NA**	**NA**	**Smoking initiation**
rs10515678	5	152323236	3.30E-09	SCZ,OB2	SCZ	SCZ_OBII	Educational attainment; Body size at age 10; Cognitive performance
rs11720523	3	71545170	3.55E-10	SCZ,T2D,CMM	SCZ, T2D	NA	Smoking initiation; Insomnia; Brain morphology; Brain region volumes; Cortical surface area; Cortical thickness; HDL levels; Intelligence; Dorsolateral prefrontal thickness; C-reactive protein levels
rs1359939	1	177820861	5.88E-09	SCZ,T2D,OB1,OB,CMM	BMI	NA	NA
rs1546924	1	112273485	3.29E-09	SCZ,T2D,OB1,OB2,CMM	BMI; T2D	T2D_ADHD/OCD	Metabolic syndrome; Waist-hip ratio; C-reactive protein levels; Basal ganglia structure; Educational attainment
rs1615350	12	123650335	9.59E-12	SCZ,OB3	T2D; BMI; SCZ; stroke	OBII_SCZ/ADHA; T2D_BMI	General cognitive ability; Insomnia; Intelligence; Whole body fat free mass; Cortical thickness; Diastolic blood pressure; Mean arterial pressure; Total cholesterol levels; Early-onset ischemic stroke; Brain morphology; Head circumference (infant); LDL levels; TG levels
rs17221259	12	14410485	1.00E-08	SCZ,T2D,OB1,OB3,CMM	SCZ; BMI	NA	Waist-hip ratio; Metabolic syndrome; Adiposity
rs1973675	2	233636450	4.06E-12	SCZ,OB2,OB3	SCZ	OBII_SCZ/ADHA	Educational attainment; Cognitive ability
rs2275997	11	115100186	3.29E-09	SCZ,T2D,OB1,OB2,OB3,CMM	BMI; OB	NA	Metabolic syndrome; TSH levels; Cigarettes smoked per day
rs2427363	20	61280662	4.69E-09	SCZ,T2D,OB1,OB3,CMM	T2D	T2D_SCZ	Hemoglobin A1c levels
rs247975	3	173107443	1.57E-09	SCZ,T2D,OB1,OB2,OB3,CMM	BMI; T2D	NA	Smoking initiation; Insomnia; Metabolic syndrome; C-reactive protein levels; HDL levels; Adiposity; MDD; TG levels; BMI and adiposity; Metabolic biomarkers; Body fat percentage
rs2933203	14	47308491	7.38E-09	SCZ,T2D,OB1,OB2,OB3,CMM	T2D	OBII_ADHD/Anorexia nervosa/MDD/SCZ	MDD; PTSD; Smoking initiation; BMI and adiposity; Metabolic syndrome
rs6059499	20	32447920	4.32E-09	SCZ,T2D,CMM	T2D	NA	Random glucose levels
rs6601414	8	9976748	2.92E-14	SCZ,T2D,OB1,OB2,CMM	BMI; T2D	T2D_ADHD/OCD	Hip circumference adjusted for BMI; Diastolic blood pressure; Waist circumference adjusted for BMI; Cortical surface area
rs6990912	8	9200472	4.92E-13	SCZ,T2D,OB1,CMM	T2D	T2D_ADHD/Anorexia nervosa/OCD	Waist-hip index
rs7432375	3	136288405	1.20E-09	SCZ,OB3	BMI; SCZ	NA	Neurociticism; Waist-hip ratio; Atrial fibrillation; HDL levels
rs7572970	2	161136656	8.12E-10	SCZ,T2D,OB1,CMM	BMI; T2D	T2D_ADHD	Metabolic syndrome; TG levels; Waist-hip ratio; Smoking initiation
rs769267	19	19446936	7.08E-09	SCZ,T2D,CMM	T2D; SCZ; BMI	OBII_ADHD/SCZ; T2D_ADHD/SCZ/OCD	Hemoglobin A1c; Childhood body mass index; Regional cortical thickness; Brain morphology; Subcortical volume
rs7968682	12	66371880	3.67E-16	SCZ,T2D,CMM	BMI; T2D	T2D_ADHD/Anorexia nervosa/OCD	BMI and adiposity; Cortical surface area; Systolic blood pressure; Whole body fat mass; Atrial fibrillation; Trunk fat mass; Hip circumference; TG levels; Fasting insulin; Head circumference (infant); Cortical volume; Brain region volumes; Cortical thickness; Educational attainment; Hypertension; Body size at age 10; Insomnia; Brain morphology
rs9449312	6	64169534	7.80E-09	SCZ,T2D,OB1,OB2,OB3,CMM	T2D; BMI	T2D_MDD/OCD	TSH levels; Metabolic syndrome

Supplementary Table 4-1b. ASSET 'subset2sided' lead SNPs.

SNP	CHR	BP	P	Pheno.1 (ASSET)	Pheno.2 (ASSET)	Previous associations	Previous Multitrait associations	Previous associations (related)
rs2904221	**4**	**77182033**	**1.67E-09**	**T2D,OB1,OB2,OB3,CMM**	**SCZ**	**NA**	**NA**	**NA**
rs3799380	**6**	**26467182**	**4.43E-19**	**SCZ**	**OB2,OB3,CMM**	**NA**	**NA**	Depression; Carotid intima-media thickness; Intelligence; General cognitive ability
rs3819720	**6**	**32804570**	**1.32E-09**	**T2D,CMM**	**SCZ**	**NA**	**NA**	**NA**
rs1009136	19	19440428	1.14E-10	SCZ,T2D	OB1,OB2,OB3	SCZ; T2D; BMI	OB2_ADHD/SCZ; T2D_ADHD/SCZ/OCD	Hemoglobin A1c; TG to HDL ratio; Childhood body mass index; Brain morphology; Subcortical volume; Cortical thickness
rs10189857	2	60713235	1.46E-10	SCZ,OB3	T2D,CMM	SCZ; T2D	T2D_Anorexia nervosa/SCZ/OCD	Cognitive ability; Educational attainment; Metabolic syndrome; Intelligence; Smoking behaviour
rs10282292	7	111092478	4.93E-12	SCZ	T2D,OB1,OB2,OB3,CMM	SCZ; BMI	NA	Lifetime smoking index
rs1042725	12	66358347	1.67E-17	SCZ,T2D,CMM	OB1,OB2,OB3	BMI; T2D	T2D_ADHD/Anorexia nervosa/OCD	BMI and adiposity; Cortical surface area; Systolic blood pressure; Atrial fibrillation; Hip circumference; Trunk fat mass; Birth weight; TG levels; Fasting insulin; Whole body fat free mass; Head circumference (infant); Cortical volume; Brain region volumes; Cortical thickness; Educational attainment; Hypertension; Body size at age 10; Insomnia; Brsin morphology
rs10515326	5	101813748	5.64E-10	T2D,OB2	SCZ	BMI; T2D	NA	Metabolic syndrome; Waist-to-hip ratio adjusted for BMI; Mean arterial pressure; Diastolic blood pressure
rs10872224	6	98435125	7.90E-09	T2D,OB1,OB2,OB3,CMM	SCZ	SCZ; BMI	NA	Educational attainment; Intelligence; Whole body fat mass; C-reactive protein levels; HDL levels; Metabolic biomarkers; Waist-hip ratio; Metabolic syndrome
rs11167136	8	143310815	6.89E-13	T2D,OB1,OB2,OB3,CMM	SCZ	SCZ	OB2_Anorexia nervosa/OCD/SCZ	Educational attainment; General cognitive ability; PTSD; Systolic blood pressure; Neuroticism; Diastolic blood pressure; Anxiety; Whole body fat free mass
rs11243150	6	7275941	1.06E-14	T2D,CMM	SCZ	T2D	T2D_Anorexia nervosa/SCZ/OCD	Glucose levels
rs11257655	10	12307894	1.23E-34	T2D,CMM	SCZ,OB1,OB2,OB3	T2D; CMM; BMI	T2D_OCD; T2D_BMI; T2D_BMI_CAD	Hemoglobin A1c levels; Blood glucose levels; Blood pressure; Childhood onset T2D; Circulating leptin levels or T2D
rs12437900	15	78134138	9.21E-09	T2D,OB1,OB2,CMM	SCZ	BMI	NA	Depression
rs12615058	2	228989843	1.52E-12	T2D,OB1,OB2,OB3,CMM	SCZ	SCZ; T2D; BMI	T2D_ADHD/Anorexia nervosa/OCD	C-reactive protein levels; Body fat percentage; Metabolic syndrome; Educational attainment; Predicted visceral adipose tissue; Trunk fat mass; Waist circumference; Frailty (Factor 4 - Metabolic Problems)
rs1265099	6	31105413	1.14E-13	SCZ	T2D,OB2,OB3,CMM	SCZ	OB2_Anorexia nervosa/MDD/SCZ	Broad depression or schizophrenia
rs12751064	1	78006124	6.53E-09	OB1,OB2,OB3	SCZ	BMI; OB	NA	Educational attainment; Bipolar I disorder; Diastolic blood pressure; Hip circumference; Personality traits or cognitive traits
rs12910334	15	84958318	1.82E-09	T2D,CMM	SCZ	SCZ; BMI	NA	Whole body fat mass; Hypertrophic cardiomyopathy; Dilated cardiomyopathy

rs1573815	3	52870132	7.03E-11	T2D,OB1,OB2,OB3,CMM	SCZ	SCZ; T2D; BMI	T2D_MDD/SCZ/OCD	Smoking initiation; Intelligence; General cognitive ability; Hypertrophic cardiomyopathy; Anxiety
rs1609520	12	123735937	8.18E-18	T2D,CMM	SCZ,OB3	SCZ; T2D; BMI; stroke	OB2_ADHD/SCZ; BMI_T2D	General cognitive ability; Waist-hip ratio; Intelligence; Cortical thickness; Diastolic blood pressure; Whole body fat free mass; Mean arterial pressure; HDL levels; Total cholesterol levels; Early-onset ischemic stroke; Brain morphology; Head circumference (infant); TG levels
rs17683730	18	54649293	3.01E-10	SCZ	T2D,OB1,OB2,OB3,CMM	T2D	NA	Metabolic syndrome
rs17747955	18	31582890	8.53E-10	SCZ	T2D,OB1,OB3,CMM	SCZ; T2D; BMI	NA	Insomnia
rs1782507	11	30243868	7.45E-09	SCZ	T2D,OB1,OB2,OB3,CMM	SCZ; T2D; BMI	T2D_ADHD/OCD	C-reactive protein levels
rs1861410	2	58933591	3.59E-12	SCZ	T2D,OB1,OB2,OB3,CMM	T2D; BMI	NA	Metabolic syndrome; TG; Hypertension; Frailty (Factor 4 - Metabolic Problems); Whole body fat mass; Insomnia; Hip circumference; Trunk fat mass; Waist circumference; HDL levels
rs1894951	16	49054994	5.74E-09	T2D,OB1,OB2,OB3,CMM	SCZ	T2D; BMI	NA	Metabolic syndrome; Adiposity
rs200440	1	6713282	1.40E-09	T2D,OB1,CMM	SCZ	SCZ; T2D; BMI; CMM	NA	Frailty (Factor 3 - Multimorbidity); Hemoglobin A1c; Adiposity; Metabolic syndrome; Childhood body mass index
rs2010390	8	9047178	4.22E-12	SCZ,T2D,OB1,OB2,CMM	NA	BMI	NA	Waist-hip index; General factor of neuroticism; Cortical surface area
rs215403	4	23425653	4.09E-09	SCZ	OB1,OB2,OB3	SCZ	NA	Cognitive ability, years of educational attainment or schizophrenia (pleiotropy)
rs2205829	6	27480227	9.19E-09	T2D,OB1,OB2	SCZ	SCZ	NA	Neuroticism general factor
rs2285268	8	17081010	2.46E-09	OB2,OB3,CMM	SCZ	SCZ	NA	Short sleep duration (<5 hours)
rs3115672	6	31727897	1.77E-18	T2D,OB3	SCZ	SCZ; BMI; CMM; stroke	NA	Total omega-3 fatty acid levels; Childhood body mass index; Waist-hip index; Ischemic stroke; Metabolic syndrome; Whole body fat free mass; Diastolic blood pressure
rs340874	1	214159256	2.32E-27	SCZ	T2D,CMM	T2D; BMI	T2D_OCD	Metabolic syndrome; Cognitive processing speed; Hemoglobin A1c; TG
rs3786897	19	33893008	7.51E-09	T2D	SCZ,B2,OB3	T2D; BMI	T2D_SCZ/OCD	Hip index; Waist-hip index; Gestational diabetes mellitus; Visceral adipose tissue volumes
rs3802177	8	118185025	8.56E-57	SCZ,OB1,OB2	T2D,CMM	T2D; BMI	T2D_BMI	Hemoglobin A1c levels; Fasting glucose; Gestational diabetes mellitus; Youth-onset type 2 diabetes; TG levels; Metabolic syndrome
rs3810291	19	47569003	1.49E-18	T2D,OB1,OB2,OB3,CMM	SCZ	T2D; BMI	NA	HDL levels; TG levels; Smoking initiation; Hip circumference; BMI and adiposity; Waist circumference; Metabolic syndrome; Metabolic biomarkers; Body fat percentage; Diastolic blood pressure; Childhood body mass index; TSH levels; Trunk fat mass; Hemoglobin A1c levels
rs3814614	10	88128281	1.93E-12	T2D,OB1,OB2,OB3	SCZ	T2D; BMI	T2D_SCZ/OCD	Childhood body mass index; Metabolic syndrome
rs3849850	8	116508778	1.92E-10	SCZ	T2D,OB1,OB2,OB3	SCZ; T2D; BMI	T2D_Anorexia nervosa/MDD/SCZ/OCD; OB2_Anorexia nervosa/SCZ	HDL levels; Whole body fat free; Blood pressure; Hip circumference; Metabolic syndrome; LDL levels; LEPR protein levels; Brain morphology; Apolipoprotein A levels; Hemoglobin A1c; Subcortical volume

rs4702	15	91426560	5.15E-11	OB3,CMM	SCZ	SCZ; CAD	NA	Systolic blood pressure; Mean arterial pressure; Hypertension; Smoking initiation; Apolipoprotein A1 levels; Brain morphology; Cortical thickness; Angina pectoris; Cortical surface area; Stable angina pectoris; Myocardial infarction
rs4804833	19	7970635	1.47E-14	T2D,OB2,OB3,CMM	SCZ	T2D; BMI	NA	TG levels; Metabolic syndrome; Fasting insulin; HDL levels; Whole body fat free mass
rs4807179	19	1956035	1.94E-10	T2D,OB1,OB2,OB3,CMM	SCZ	T2D; BMI; CAD	NA	Severe insulin-deficient T2D; C-reactive protein levels; Metabolic syndrome
rs483465	3	136047977	1.56E-11	SCZ	T2D,OB1,OB2,CMM	T2D; BMI; CAD; stroke	T2D_CAD	TG levels; HDL levels; Systolic blood pressure; Dyslipidemia; Ischemic stroke or fibrinogen levels; Myocardial infarction; Apolipoprotein A levels; C-reactive protein levels; Diastolic blood pressure; LEPR protein levels; Waist-hip ratio; Omega-3 fatty acid levels; Hypertension; LDL levels; Ischemic heart disease; Cigarettes smoked per day; Metabolic biomarkers; Body fat percentage; Waist circumference
rs5215	11	17408630	1.08E-27	SCZ,OB1,OB3	T2D,CMM	T2D; BMI; CMM	T2D_BMI	Systolic blood pressure; Hypertension; Diastolic blood pressure; Systolic blood pressure; Abdominal size; Waist-hip index; Frailty (Factor 3 - Multimorbidity); Adiposity; Hemoglobin A1c levels; Circulating leptin levels or T2D
rs6126570	20	51094309	8.36E-11	SCZ	T2D,OB1,OB2,OB3,CMM	BMI; OB		Educational attainment; Insomnia; Whole body fat mass; Metabolic biomarkers; Metabolic syndrome
rs6568686	6	111872482	8.18E-10	SCZ	T2D,CMM	SCZ; T2D		Smoking status; Metabolic syndrome; HDL levels; TG levels
rs6581987	12	71432168	1.32E-14	SCZ,OB2,OB3	T2D,CMM	T2D		
rs6590512	11	130717916	4.45E-13	T2D,OB1,OB2,OB3,CMM	SCZ	SCZ; BMI	OB2_Anorexia nervosa/SCZ	Anxiety; Metabolite levels; Frailty (Factor 4 - Metabolic Problems)
rs6704768	2	233592501	5.67E-11	T2D	SCZ,OB2,OB3	SCZ	OB2_ADHD/SCZ	Educational attainment; Cognitive ability
rs6779258	3	71549639	2.14E-09	SCZ,T2D,CMM	OB3	SCZ; T2D	NA	Smoking initiation; Insomnia; Brain morphology; Brain region volumes; Educational attainment; Cortical surface area; Cortical thickness; HDL levels; intelligence; General cognitive ability; C-reactive protein levels
rs6801189	3	180832739	5.66E-10	SCZ	T2D,OB1,OB2,OB3,CMM	SCZ; BMI	NA	Cognitive ability
rs6914422	6	33534880	5.41E-13	SCZ	T2D,OB1,CMM	T2D; BMI	NA	Inflammatory markers and poor diet; Metabolic syndrome
rs6983332	8	9977918	4.88E-13	SCZ,T2D,OB1,CMM	NA	T2D; BMI	T2D_ADHD/OCD	Body size or adipose distribution; Diastolic blood pressure; Cortical surface; Waist-hip index
rs7085104	10	104628873	8.98E-18	SCZ	OB1,OB2,OB3	SCZ; BMI	NA	Hip circumference; Waist circumference; Whole body fat mass; Age of smoking initiation; Blood pressure; Global brain arterial diameters; White matter lesion progression
rs7204797	16	29968015	7.37E-19	T2D,OB1,OB2,OB3	SCZ	SCZ; T2D; BMI; OB	OB2_OCD/SCZ; T2D_SCZ/OCD	C-reactive protein levels; Subcortical volume; BMI and adiposity; Waist-hip ratio; Systolic blood pressure; Adiposity; Insomnia; TG levels; HDL levels; Apolipoprotein A1 levels; Whole body fat mass; Brain morphology; Trunk fat mass; Metabolic syndrome
rs738408	22	44324730	5.45E-09	T2D,OB3	SCZ,OB1,CMM	T2D; BMI; OB; CAD	T2D_BMI; BMI_CAD	Metabolite levels; HDL levels; Total cholesterol levels; Apolipoprotein A1 levels; TG levels; Hemoglobin A1c; VLDL levels

rs7613875	3	49971514	9.87E-13	T2D,OB1,OB2,OB3,CMM	SCZ	SCZ, T2D; BMI	T2D_ADHD/Anorexia nervosa/MDD/OCD	C-reactive protein levels; Whole body fat mass; Intelligence; HDL levels; Metabolic biomarkers; Trunk fat mass; Diastolic blood pressure; Myocardial infarction; Educational attainment; General cognitive ability; Apolipoprotein A levels; Metabolic syndrome; Hip index; Insomnia
rs7676765	4	103937933	3.03E-11	T2D,OB1,OB2,OB3,CMM	SCZ	SCZ; T2D	T2D_ADHD/OCD	Fasting glucose; Hemoglobin A1c levels; TG levels; Atrial fibrillation; Hypertension
rs7975482	12	124481690	6.94E-11	T2D,CMM	SCZ,OB1,OB2	SCZ; T2D; BMI; CAD	T2D_Anorexia nervosa/MDD/SCZ/OCD; T2D/CAD; T2D_BMI; T2D_CAD	TG levels; HDL levels; Waist-hip ratio; Fasting insulin; Metabolic biomarkers; Insomnia; Whole body fat mass; Myocardial infarction; Apolipoprotein A levels; Trunk fat mass; Metabolic syndrome; Hemoglobin A1c; Adiponectin levels
rs8068351	17	34916879	2.84E-10	T2D,OB1,OB2,OB3,CMM	SCZ	SCZ; T2D; BMI; OB	NA	TG levels; Smoking initiation; Systolic blood pressure; Diastolic blood pressure; Educational attainment; Childhood body mass index; Metabolic biomarkers; Waist circumference; Intelligence; Body fat percentage; Trunk fat mass
rs815611	5	153518766	9.21E-13	SCZ	T2D,OB1,OB2,OB3	SCZ; T2D; BMI; OB	NA	Smoking initiation; Metabolic syndrome; Body size at age 10; Metabolic biomarkers; Body fat percentage
rs8192675	3	170724883	6.66E-24	T2D,CMM	SCZ,OB1,OB2,OB3	T2D; BMI	T2D_OCD	C-reactive protein levels; Hemoglobin A1c; Adiposity; Waist circumference; Whole body fat mass; Hip circumference; TG levels; Metabolic biomarkers; Educational attainment
rs832188	3	63841395	4.39E-12	SCZ,OB3	T2D,CMM	T2D	NA	Fasting plasma glucose
rs867743	8	60694647	1.40E-09	T2D,OB1,OB2,CMM	SCZ	SCZ; BMI	NA	Insomnia; Cortical surface area; Brain morphology; Educational attainment
rs9633835	11	13345593	7.93E-09	SCZ	T2D,OB1,OB2,OB3,CMM	SCZ; BMI	NA	Neuroticism; Sensitivity to environmental stress and adversity; Smoking initiation; Whole body fat mass; Adiposity
rs9636107	18	53200117	3.08E-14	T2D	SCZ	SCZ	NA	Cognitive ability
rs9811916	3	36878086	5.14E-11	T2D,CMM	SCZ	SCZ	NA	Cognitive ability
rs9941245	16	19916895	2.71E-10	T2D,OB1,OB2,OB3	SCZ,CMM	BMI; OB	OB2_OCD	Insomnia; Whole body fat mass; Total cholesterol levels; Childhood body mass index; Waist circumference; Metabolic syndrome; Body size at age 10; LDL levels; HDL levels; Aging

Supplementary Table 4-1c. ASSET 'shared' lead SNPs.

SNP	CHR	BP	P	Pheno.1 S2sided	Pheno.2 S2sided	Previous associations	Previous Multitrait associations	Previous associations (related)
rs1007090	20	32582871	1.29E-12	NA	SCZ,T2D,OB1,OB2,CMM	T2D; BMI	T2D_ADHD/OCD	Cortical surface area; Adult body size; Adiposity; Metabolic syndrome; Vertex-wise cortical thickness
rs10077814	5	44916789	3.92E-11	OB1,OB3,CMM	SCZ,T2D	SCZ; T2D	T2D_SCZ/OCD	PPY protein levels; Schizophrenia vs anorexia nervosa
rs10506133	12	39422793	3.81E-09	SCZ,T2D,OB1,OB2, OB3,CMM	NA	T2D	NA	Metabolic syndrome
rs10838158	11	43752522	5.18E-15	NA	SCZ,T2D,OB1,OB2,OB3, CMM	T2D; BMI; CAD	NA	Adiposity; Adult body size; Trunk fat mass; Waist circumference; Predicted visceral adipose tissue; Metabolic syndrome; Hip circumference; Whole body fat mass; BMI and adiposity
rs10888679	1	50566266	4.54E-10	NA	SCZ,OB1,OB2,OB3	BMI	NA	Whole body fat mass; Waist circumference; Cigarettes smoked per day
rs1169302	12	121432302	1.79E-18	OB3	SCZ,T2D,CMM	T2D	NA	Random glucose levels; LDL levels; C-reactive protein levels
rs12613687	2	200999256	2.96E-09	SCZ,OB3	OB1,OB2,CMM	SCZ; BMI	NA	Cognitive ability; Hip shape; QT interval
rs13746	12	121201167	3.10E-09	SCZ,T2D,OB2,OB3, CMM	NA	SCZ; T2D	T2D/CAD	Inflammatory markers and poor diet; TSH levels; Educational attainment; Whole body fat free mass
rs137845	22	50439430	2.66E-09	OB3	SCZ,T2D,CMM	T2D; BMI	NA	Gestational diabetes mellitus
rs151390	4	103183510	6.59E-09	SCZ,T2D,OB1,OB2, OB3,CMM	NA	BMI	NA	Volume of right medial nucleus; Volume of right cortical nucleus; Intelligence; General cognitive ability; Insomnia; Anterior amygdaloid area volume; Central nucleus volume; Adiposity; Volume of right central nucleus; Waist circumference adjusted for BMI; Medial nucleus volume
rs1569858	22	41541024	2.13E-10	NA	SCZ,T2D,OB1,OB2,CMM	T2D; BMI	NA	Intelligence; General cognitive ability
rs17522122	14	33302882	4.03E-18	SCZ,T2D,OB1,OB2, OB3,CMM	NA	SCZ; T2D; BMI; OB	OB2_ADHD/SCZ; BMI_ systolic blood pressure; T2D_ADHD	General cognitive ability; Intelligence; Cortical thickness; Frailty (Factor 4 - Metabolic Problems); Whole body fat mass; Adiposity; Hip circumference; Metabolic syndrome; C-reactive protein levels; Morbid obesity
rs1766377	1	49917084	1.59E-10	NA	SCZ,OB1,OB2,OB3	SCZ; BMI	OB2_SCZ	Diastolic blood pressure; Waist-hip ratio; Smoking status; Insomnia; Broad depression or schizophrenia
rs17763551	**17**	**3882309**	**2.40E-14**	**SCZ,T2D,CMM**	**OB3**	**NA**	**NA**	**NA**
rs2218378	7	121961096	2.19E-09	NA	SCZ,T2D,OB1,OB2,OB3	BMI	NA	Smoking initiation; Attention deficit hyperactivity disorder; BMI and adiposity; Weight; Whole body fat mass; Metabolic syndrome
rs2250377	1	201860626	4.81E-11	SCZ,T2D,OB1,OB2, OB3,CMM	NA	BMI; CAD	NA	Intelligence; General cognitive ability; Body fat percentage; TG levels; BMI and adiposity; HDL levels; Metabolic biomarkers; Waist circumference; Childhood body mass index; Trunk fat mass; Stable angina pectoris; Myocardial infarction; Metabolic syndrome; Educational attainment; C-reactive protein levels
rs2280141	10	124193181	5.39E-13	SCZ,T2D,OB3,CMM	OB1	T2D; BMI	T2D_ADHD/SCZ	Whole body fat free mass; Basal ganglia structure; Birth weight; Hypothyroidism; BMI and adiposity; Alzheimer's disease; Autoimmune thyroid disease; Intelligence; Cognitive performance
rs2618451	8	11376266	3.00E-11	NA	SCZ,T2D,OB1,OB2,CMM	BMI	NA	Childhood body mass index; Neuroticism; Pulse pressure; Systolic blood pressure; Diastolic blood pressure; Systolic blood pressure; Waist-hip index; Autoimmune thyroid disease; Adult body size; Hypertension

rs2796441	9	84308948	3.31E-28	OB3	SCZ,T2D,CMM	SCZ; T2D; BMI	NA	Hemoglobin A1c levels
rs2921077	8	8304502	1.33E-11	SCZ,T2D,OB1,OB2	NA	SCZ; T2D; BMI	NA	Waist-hip index; Neuroticism; Pulse pressure; Systolic blood pressure; Diastolic blood pressure; C-reactive protein levels; TG levels; Hypothyroidism; TSH levels; Brain morphology
rs3903399	1	205041542	2.45E-10	NA	SCZ,T2D,OB1,OB3,CMM	T2D	NA	Hemoglobin A1c; HDL cholesterol levels; Waist-hip ratio; Metabolic syndrome; HDL levels
rs4678408	3	138053187	1.23E-09	SCZ,T2D,OB2,CMM	NA	T2D; CMM; CAD	NA	Dyslexia; Hypertension; Metabolic syndrome; Hemoglobin A1c levels
rs4925109	17	17661802	2.04E-17	SCZ,T2D,CMM	OB1,OB2,OB3	SCZ; T2D; BMI	T2D/CAD; OB2_OCD/SCZ	Waist-hip index; Educational attainment; Brain morphology; General cognitive ability
rs531897	11	122013169	2.12E-09	NA	SCZ,T2D,OB1,OB2,OB3, CMM	SCZ; BMI; CMM	NA	Educational attainment; Cognitive ability; Metabolic syndrome; Waist-hip ratio; Frailty (Factor 3 - Multimorbidity)
rs6985109	8	10761585	2.13E-12	OB3	SCZ,T2D,OB1,OB2,CMM	T2D; BMI	NA	Neuroticism; Waist-hip index; LDL levels; Systolic blood pressure; Smoking initiation; Subcortical volume; Body fat distribution; Cognitive traits; Carotid intima media thickness; Inflammatory markers and poor diet
rs707939	6	31726688	1.37E-12	SCZ,CMM	OB1,OB2,OB3	SCZ; BMI	NA	Waist-hip index
rs7103411	11	27700125	3.14E-15	SCZ,OB1,OB2,OB3, CMM	NA	SCZ; T2D; BMI; OB; CAD	OB2_ADHD/SCZ	Smoking initiation; Brain morphology; Waist circumference; Educational attainment; Trunk fat mass; Waist-hip ratio; C-reactive protein levels; Inflammatory markers and poor diet; BMI and adiposity
rs7156625	14	79942647	1.07E-18	SCZ,T2D,OB1,OB2, OB3,CMM	NA	T2D; BMI; OB	T2D_ADHD	Hypertension; Atrial fibrillation; Hemoglobin A1c; Childhood body mass index; Waist-hip ratio; HDL levels; Metabolic syndrome; Whole body fat free mass
rs7776707	7	104605530	6.86E-09	SCZ,T2D,OB2,OB3, CMM	NA	SCZ; BMI	NA	Insomnia; Cognitive ability; Waist-hip ratio; Abdominal size
rs7848336	9	96454513	2.43E-09	NA	SCZ,T2D,OB1,OB2,OB3, CMM	BMI	NA	Educational attainment; Cigarettes smoked per day; Whole body fat mass; Metabolic syndrome
rs8009761	**14**	**74888554**	**3.20E-09**	**NA**	**SCZ,T2D,OB1,OB2,OB3**	**NA**	**NA**	Educational attainment; Cognitive performance; Metabolic syndrome
rs8017993	14	104047734	1.69E-11	OB3	SCZ,CMM	SCZ; BMI	NA	Cognitive ability; Insomnia; Cigarettes smoked per day; Blood pressure
rs8036171	15	95266810	1.32E-09	SCZ,T2D,OB1,OB2, OB3,CMM	NA	T2D; BMI	NA	Educational attainment; Metabolic syndrome; Adiposity
rs8069451	17	37504933	1.19E-09	NA	SCZ,T2D,OB3,CMM	T2D; CAD	T2D_SCZ/ADHD; T2D/CAD	Brain morphology; Smoking initiation; HDL cholesterol; Metabolic syndrome; Educational attainment
rs8078510	17	47045862	2.50E-16	NA	SCZ,T2D,OB2,CMM	SCZ; T2D; BMI; CMM	T2D_CAD; T2D_SCZ; BMI_T2D_CAD	Myocardial infarction; HDL levels; Insomnia; Brain morphology; Superior temporal area; Lifetime smoking index; Hypertension; Essential hypertension
rs938875	15	67952922	1.46E-15	CMM	SCZ,T2D,OB1,OB2,OB3	T2D; BMI; OB		Smoking initiation; Diastolic blood pressure; Whole body fat mass; HDL levels; Waist circumference; Metabolic syndrome; Morbid obesity
rs953097	10	98981914	8.28E-10	SCZ,T2D,OB1,OB2, OB3,CMM	NA	T2D; BMI	T2D_ADHD/MDD/OCD	Cognitive traits; Metabolic syndrome; Subcortical volume; Inflammatory markers and poor diet; Frailty (Factor 5 - Poorer Cognition)
rs9873519	3	124921457	4.33E-11	SCZ,T2D,OB3,CMM	OB1,OB2	T2D		Hemoglobin A1c; Fasting blood glucose

Supplementary Table 4-2a. 'subset1sided' top SNPs.

SNP	CHR	PIP	Previous associations	Previous Multitrait associations	Previous associations (related)
rs10515678	**5**	**0.996152**	**NA**	**NA**	**NA**
rs3737095	6	1	T2D; BMI; OB	NA	Smoking initiation; Morbid obesity; Hemoglobin A1c
rs6677300	1	0.999503	BMI	NA	BMI and adiposity; Cigarettes smoked per day; Body fat percentage; Trunk fat mass; Metabolic syndrome; Metabolic biomarkers; HDL levels
rs630372	1	1	T2D; BMI; OB	OB2_ADHD	C-reactive protein levels; HDL levels; Morbid obesity; BMI and adiposity; Childhood obesity; Metabolic biomarkers; Body fat percentage; Waist-hip ratio; Metabolic syndrome; Smoking initiation; Systolic blood pressure; Trunk fat mass; Inflammatory markers and poor diet
rs197419	1	1	T2D; BMI	NA	Metabolic syndrome; TG levels; HDL levels; C-reactive protein levels; Educational attainment
rs4938180	11	1	BMI; OB	NA	TSH levels; Metabolic syndrome; Educational attainment; Cigarettes smoked per day; Whole body fat mass
rs6062689	20	1	T2D	T2D_SCZ	Hemoglobin A1c levels; Fasting plasma glucose
rs247975	3	1	BMI; T2D	NA	Smoking initiation; Insomnia; Metabolic syndrome; C-reactive protein levels; HDL levels; Adiposity; MDD; TG levels; Metabolic biomarkers; Body fat percentage
rs3007105	14	0.999904	T2D; BMI	OB2_ADHD/Anorexia nervosa/MDD/SCZ	Smoking initiation; Trauma exposure; BMI and adiposity; Metabolic syndrome; Neuroticism
rs6990912	8	0.998063	T2D	T2D_ADHD/Anorexia nervosa/OCD	Waist-hip index
rs1145101	3	1	T2D; BMI; CAD; stroke	T2D_CAD	TG levels; HDL levels; Waist circumference; Systolic blood pressure; Myocardial infarction; Fasting blood glucose; C-reactive protein levels; LEPR protein levels; Waist-hip ratio; Apolipoprotein B levels; Omega-3 fatty acid levels; Coronary atherosclerosis; Cigarettes smoked per day; Metabolic biomarkers; Body fat percentage
rs12998587	2	0.99656	T2D; BMI	T2D_ADHD	Metabolic syndrome; TG levels; Waist-hip ratio; Smoking initiation
rs2758259	6	0.999999	T2D; BMI	T2D_MDD/SCZ/OCD	TSH levels; Metabolic syndrome

Supplementary Table 4-2b. 'subset2sided' top SNPs.

SNP	CHR	PIP	Previous associations	Previous Multitrait associations	Previous associations (related)
rs6718758	**2**	**1**	**NA**		**NA**
rs38744	7	1	SCZ	OB2_ADHD/OCD/SCZ	Smoking initiation
rs2600834	5	1	T2D; BMI	NA	Metabolic syndrome; Waist-to-hip ratio adjusted for BMI; Mean arterial pressure; Diastolic blood pressure
rs4412207	6	0.99683431	SCZ; BMI; CMM	NA	Intelligence; Systolic blood pressure; Cognitive ability; Brain region volumes; Educational attainment; C-reactive protein levels; HDL levels; Metabolic biomarkers; Body fat percentage; Waist-hip ratio; Metabolic syndrome
rs13207230	6	1	T2D; BMI	NA	Waist-hip index
rs10906025	**10**	**1**	**NA**	**NA**	**NA**
rs7178572	15	1	T2D; BMI	T2D_ADHD/SCZ	Fasting glucose; Hemoglobin A1c; Metabolic syndrome; Adiposity
rs3130573	6	1	BMI	NA	Waist-hip index
rs9729667	1	0.9987463	SCZ; BMI; OB	NA	Diastolic blood pressure; Waist circumference; Cognitive ability; Educational attainment
rs11635505	15	1	SCZ; BMI	NA	Whole body fat mass; Cognitive ability; Hypertrophic cardiomyopathy
rs2276825	3	1	SCZ; T2D; BMI	T2D_MDD/SCZ/OCD	Neuroticism; Smoking initiation; Intelligence; General cognitive ability; Hypertrophic cardiomyopathy; Hip circumference; Anxiety
rs12958029	**18**	**1**	**NA**	**NA**	Insomnia
rs642126	11	1	SCZ; T2D; BMI	T2D_ADHD/OCD	C-reactive protein levels
rs1016287	2	1	T2D; BMI; OB	NA	TG levels; HDL levels; Metabolic biomarkers; Body fat percentage; Systolic blood pressure; Atrial fibrillation; C-reactive protein levels; Waist-hip ratio; Metabolic syndrome; Trunk fat mass; Aging; BMI and adiposity; Pulse pressure; Smoking initiation
rs1894951	16	0.99995959	T2D; BMI	NA	Metabolic syndrome; Adiposity
rs3800316	6	0.99999999	schizophrenia	NA	General cognitive ability; Intelligence
rs2588120	**8**	**1**	**NA**	**NA**	**NA**
rs2904221	**4**	**0.99997056**	**NA**	**NA**	**NA**
rs340874	1	1	T2D; BMI	T2D_OCD	Metabolic syndrome; Cognitive processing speed; Hemoglobin A1c; TG levels
rs2546057	19	1	BMI	NA	NA
rs13266634	8	1	T2D; BMI	T2D_BMI	Fasting plasma glucose; Hemoglobin A1c levels; Gestational diabetes mellitus; TG levels; Youth-onset T2D; Metabolic syndrome
rs3810291	19	1	T2D; BMI	NA	HDL levels; TG levels; Smoking initiation; BMI and adiposity; Waist circumference; Metabolic syndrome; Metabolic biomarkers; Body fat percentage; Diastolic blood pressure; Childhood body mass index; TSH levels; Trunk fat mass; Hemoglobin A1c levels
rs4933391	10	1	T2D; BMI	T2D_SCZ/OCD	NA
rs4148870	6	1	T2D	T2D_ADHD/Anorexia nervosa/OCD	NA
rs10097784	8	1	SCZ; T2D; BMI	T2D_Anorexia nervosa/MDD/SCZ/OCD; OB2_Anorexia nervosa/SCZ	HDL levels; Whole body fat free mass; Blood pressure; Hip circumference; Metabolic syndrome; Hemoglobin A1c; LDL levels; LEPR protein levels; Brain morphology; Apolipoprotein A levels
rs12910825	15	0.99950405	T2D	SCZ_T2D	NA
rs2656865	**19**	**1**	**NA**	**NA**	**NA**

rs548288	3	0.9977314	T2D; BMI; CAD; stroke	T2D_CAD	TG levels, HDL levels; Waist circumference; Systolic blood pressure; C-reactive protein levels; Dyslipidemia; Myocardial infarction; Apolipoprotein A levels; LEPR protein levels; Hypertension; LDL levels; Disorders of lipid metabolism; Hyperlipidemia; Ischemic heart disease; Coronary atherosclerosis; Smoking cessation; Metabolic biomarkers; Body fat percentage
rs6097012	20	1	BMI; OB	NA	Insomnia; Body fat percentage; Metabolic biomarkers; Metabolic syndrome
rs6568686	6	0.99999999	SCZ; T2D	NA	Smoking status; Metabolic syndrome; HDL levels; TG levels
rs6581977	**12**	**1**	**NA**	**NA**	Apolipoprotein A1 levels
rs4678874	**3**	**1**	**NA**	**NA**	NA
rs10791097	11	0.87887879	SCZ; BMI	OB2_Anorexia nervosa/SCZ	Anxiety; Metabolite levels; Frailty (Factor 4 - Metabolic Problems)
rs4854912	3	0.94522593	SCZ; BMI	NA	NA
rs4917985	10	1	SCZ; BMI	NA	Waist circumference; Whole body fat mass; Smoking initiation
rs738408	22	0.99599829	T2D; BMI; OB; CAD	T2D_BMI; BMI_CAD	Metabolite levels; HDL levels; Total cholesterol levels; Apolipoprotein A1 levels; TG levels; Hemoglobin A1c; VLDL levels
rs7660298	4	1	SCZ; T2D	T2D_ADHD/OCD	Fasting glucose; Atrial fibrillation; Hemoglobin A1c; Hypertension
rs2306590	17	0.99999992	SCZ; T2D; BMI; OB	NA	TG levels; Smoking initiation; Systolic blood pressure; Diastolic blood pressure; Educational attainment; Cognitive performance; Childhood body mass index; Metabolic biomarkers; Body fat percentage; Adiposity; Intelligence; Non-HDL cholesterol levels; Trunk fat mass
rs815611	5	0.99988523	SCZ; T2D; BMI; OB	NA	Smoking initiation; Metabolic syndrome; Body size at age 10; Metabolic biomarkers; Body fat percentage
rs5398	3	1	T2D; BMI	T2D_OCD	C-reactive protein levels; Glucose levels; Hemoglobin A1c; Adiposity; Waist circumference; TG levels; Metabolic biomarkers; Body fat percentage; Educational attainment
rs832188	3	0.99556151	T2D	NA	Fasting plasma glucose
rs7018304	8	0.99203289	SCZ; BMI	NA	Insomnia; Cortical surface area; Brain morphology; Brain shape
rs9636107	18	1	SCZ	NA	Cognitive ability

Supplementary Table 4-2c. 'shared' top SNPs.

SNP	CHR	PIP	Previous associations	Previous Multitrait associations	Previous associations (related)
rs8119937	**20**	**0.999981**	**NA**	**NA**	**NA**
rs830644	**3**	**1**	**NA**	**NA**	NA
rs7217945	**17**	**1**	**NA**	**NA**	**NA**
rs13425918	**2**	**1**	**NA**	**NA**	NA
rs7778318	**7**	**1**	**NA**	**NA**	NA
rs4678398	**3**	**1**	**NA**	**NA**	NA
rs729599	5	0.999775	SCZ; T2D; BMI	T2D_SCZ/OCD	PPY protein levels
rs10506133	12	0.999967	T2D	NA	Metabolic syndrome
rs10838146	11	1	BMI	NA	Smoking initiation
rs4078484	1	0.999981	SCZ; BMI	NA	Smoking status; Cognitive traits
rs1169302	12	1	T2D	NA	Random glucose levels; LDL levels; C-reactive protein levels
rs12613687	2	0.998227	SCZ; BMI	NA	Cognitive ability; QT interval
rs6677300	1	0.999503	BMI	NA	BMI and adiposity; Body fat percentage; Trunk fat mass; Metabolic syndrome; Metabolic biomarkers; HDL levels
rs137862	22	0.999436	T2D; BMI	NA	Gestational diabetes mellitus
rs233816	4	1	BMI	NA	Intelligence; General cognitive ability; Insomnia; Central nucleus volume; Adiposity; Waist circumference
rs9611497	22	1	T2D; BMI	NA	Intelligence; Cognitive performance; Neuroticism
rs1727302	12	1	SCZ; T2D; BMI; stroke	OB2_ADHD/SCZ; T2D_BMI	Insomnia; Intelligence; Whole body fat free mass; Cortical thickness; Diastolic blood pressure; TC levels; Early-onset ischemic stroke; Waist-hip index; Cognitive performance; Educational attainment; HDL levels; Brain morphology; Head circumference (infant); LDL levels; TG levels
rs1766377	1	0.999783	SCZ; BMI	OB2_SCZ	Diastolic blood pressure; Waist-hip ratio; Smoking status; Insomnia; Broad depression or schizophrenia
rs2820312	1	0.680197	BMI; CAD	NA	Intelligence; Body fat percentage; TG levels; HDL levels; Metabolic biomarkers; Childhood body mass index; Trunk fat mass; Adiposity; MI; Metabolic syndrome; Educational attainment; CRP levels; Cognitive performance
rs2796441	9	1	SCZ; T2D; BMI	NA	Hemoglobin A1c levels
rs3903399	1	0.99908	T2D	NA	Hemoglobin A1c; HDL levels; Waist-hip ratio; Metabolic syndrome
rs513069	11	0.999998	SCZ; BMI; CMM	NA	Educational attainment; Cognitive ability; Metabolic syndrome; Waist-hip ratio
rs6601414	8	1	T2D; BMI	T2D_ADHD/OCD	Diastolic blood pressure; Waist-hip index
rs7832722	8	1	BMI	NA	Cortical surface area; Waist circumference
rs3134899	6	1	BMI	NA	Insomnia
rs10501087	11	1	SCZ; T2D; BMI; OB; CAD	OB2_ADHD/SCZ	Smoking initiation; Brain morphology; Educational attainment; Trunk fat mass; Waist-hip ratio; CRP levels; Inflammatory markers; BMI and adiposity
rs6909	19	0.999706	SCZ; T2D; BMI	OB2_ADHD/SCZ; T2D_ADHD/SCZ/OCD	Hemoglobin A1c; Childhood body mass index; Brain morphology; Subcortical volume; Cortical thickness
rs2398871	9	1	BMI	NA	Whole body fat mass; Metabolic syndrome
rs1125221	14	0.999997	T2D	NA	Educational attainment; Cognitive performance; Metabolic syndrome
rs11633626	15	0.999787	T2D; BMI	NA	Educational attainment; Metabolic syndrome; Adiposity
rs801426	17	0.999993	T2D; CAD	T2D_SCZ/OCD; T2D_CAD	Brain morphology; Smoking initiation; HDL levels; TC levels; Metabolic syndrome; Educational attainment
rs2270576	17	1	SCZ; T2D; BMI; CMM	Psycho_CMM; T2D_CAD; T2D_SCZ	MI; HDL levels; Insomnia; Brain morphology; Hypertension; Essential hypertension
rs3784692	15	1	T2D; BMI; OB	NA	Smoking initiation; Diastolic blood pressure; HDL levels; Body fat percentage; BMI and adiposity; Waist circumference; Metabolic syndrome; Morbid obesity
rs569255	3	1	T2D	T2D_SCZ/OCD	NA

Supplementary Table 4-3a. 'subset1sided' lead SNPs functional annotations.

Variants details		HaploReg									RegulomeDB		VEP	
chr	variant	Ref	Alt	EUR freq	Promoter histone marks	Enhancer histone marks	DNAse	Proteins bound	Motifs changed GWAS hits	RefSeq genes	probability	ranking	Consequence	CADD_ PHRED
1	**rs1343424**	**T**	**C,G**	**0.21**						**AGBL4**	**0.18412**	**7**	**intron**	**9.174**
1	rs1359939	G	A	0.29		FAT				77kb 3' of SEC16B	0	5	intergenic	0.04
1	rs1546924	T	C	0.51	HRT, MUS	12 tissues			CDP,Fox,Pax-4	C1orf183	0.13454	5	intron	0.108
2	rs1973675	A	G	0.59		ESDR, BLD, SKIN			4 altered motifs	KCNJ13	0.18412	7	intron	2.811
2	rs7572970	A	G	0.74		14 tissues	11 tissues	GATA2,YY1	Hdx	RBMS1	0.70497	4	intron	4.771
3	rs11720523	C	A	0.42	8 tissues	16 tissues	7 tissues	TCF12,YY1	RBP-Jkappa	FOXP1	0.60906	4	intron	1.459
3	rs247975	T	C	0.54					NRSF,Znf143	8.8kb 5' of NLGN1	0.13454	5	intergenic	4.204
3	rs7432375	G	A	0.44					4 altered motifs	STAG1	0	6	intron	1.227
5	rs10515678	C	T	0.28					RFX5	538kb 5' of NMUR2	0.41972	5	intergenic	0.447
6	**rs12196300**	**G**	**A**	**0.44**		**HRT**	**HRT**		**4 altered motifs**	**DNAH8**	**0.13454**	**5**	**intron**	**11.02**
6	rs9449312	A	C	0.57		5 tissues			Elf5,PU.1	112kb 5' of PTP4A1	0	5	intergenic	0.39
8	rs6601414	G	A	0.42					Arid5a,Cdx2	MSRA	0	5	intron	1.269
8	rs6990912	A	C	0.68		BRN			20 altered motifs	7.9kb 3' of LOC157273	0.63	6	intergenic	10.92
11	rs2275997	A	G	0.52		BRN, BLD				CADM1	0.51392	7	intron	2.614
12	rs1615350	C	T	0.73	LIV	5 tissues	GI,OVRY		PU.1,TAL1,Znf143	MPHOSPH9	0.60906	4	intron	7.477
12	rs17221259	T	C	0.22	13 tissues	18 tissues	42 tissues	POL2,TBP, YY1		108kb 5' of ATF7IP	0.60906	4	downstream_gene	16.33
12	rs7968682	G	T	0.54					5 altered motifs	12kb 3' of HMGA2	0.2714	5	intergenic	14.2
14	rs2933203	T	C	0.48					5 altered motifs	335bp 3' of MDGA2	0.027	6	downstream_gene	5.739
19	rs769267	G	A	0.69	BRST, SKIN	4 tissues	5 tissues			MAU2	0.22271	1f	synonymous	6.672
20	rs2427363	C	T	0.62		9 tissues	8 tissues		DEC,Myb,SP1	SLCO4A1	0.60906	4	intron	2.081
20	rs6059499	T	C	0.51		13 tissues	SKIN,LIV		Pax-5	5.7kb 3' of CHMP4B	0.55436	1f	intergenic	0.203

Supplementary Table 4-3b. 'subset2sided' lead SNPs functional annotations.

Variants details		HaploReg									RegulomeDB		VEP	
chr	variant	Ref	Alt	EUR freq	Promoter HM	Enhancer HM	DNAse	Proteins bound	Motifs changed GWAS hits	RefSeq genes	probability	ranking	Consequence	CADD_ PHRED
1	rs12751064	C	T	0.28	5 tissues	10 tissues	7 tissues		p300	AK5	0.60906	4	intron	4.983
1	rs200440	A	G	0.56		7 tissues	MUS,MUS	CMYC	4 altered motifs	DNAJC11	0.60906	4	intron	0.011
1	rs340874	T	C	0.54	7 tissues	14 tissues	6 tissues		Cdx2	2.6kb 5' of PROX1	0.60906	4	upstream_gene	11.25
2	rs10189857	A	G	0.44			BLD		4 altered motifs	BCL11A	0.58955	5	intron	15
2	rs12615058	G	A	0.33					Evi-1,NF kB,Pax-4	SPHKAP	0.5804	6	intron	13.09
2	rs1861410	C	T	0.56					HNF4,RXR::LXR,SETDB1	FLJ30838	0.18412	7	intron, noncoding_tran	1.07
2	rs6704768	G	A	0.53					9 altered motifs	GIGYF2	0.02167	6	intron	2.476
3	rs1573815	G	A	0.26	MUS	12 tissues	6 tissues		BDP1	TMEM110-MUSTN1	0.22271	1f	upstream_gene	18.09
3	rs483465	A	G	0.74		MUS			4 altered motifs	PCCB	0.58955	5	intron	3.107
3	rs6779258	T	C	0.42		5 tissues	BLD		SMC3	FOXP1	0.13454	5	intron	3.923
3	rs6801189	A	G	0.19					GR,HNF4	125kb 5' of DNAJC19	0.58955	5	intergenic	0.318
3	rs7613875	C	A	0.52		BLD	BLD		4 altered motifs	4.1kb 5' of MON1A	0.34847	6	upstream_gene	1.064
3	rs8192675	T	C	0.29		LIV, GI, BLD				SLC2A2	0.18412	7	intron	0.124
3	rs832188	T	C	0.8	ESC,IPSC, BLD	13 tissues	11 tissues	NANOG	Foxa	THOC7	0.60906	4	intron	0.736
3	rs9811916	A	G	0.36					Irf,STAT	TRANK1	0	6	intron	2.679
4	**rs2904221**	**T**	**C**	**0.74**		**LIV, GI**		**FOXA1 ,FOXA2, P300**	**5 altered motifs**	**FAM47E**	**1**	**3b**	**intron**	**1.632**
4	rs215403	C	T	0.33		5 tissues			Nanog,VDR,p300	MIR548AJ2	0.13454	5	intergenic	0.777
4	rs7676765	T	C	0.48					SREBP,Zbtb3	NHEDC1	0.80225	1f	intron	0.627
5	rs10515326	C	T	0.25					HMGIY,Mrg1::Hoxa9	SLCO6A1	0.18894	6	intron	1.619
5	rs815611	G	A	0.55		10 tissues			4 altered motifs	52kb 5' of GALNT10	0.951	1d	intergenic	3.481
6	**rs3799380**	**T**	**C**	**0.18**			**BLD**		**Pax-6**	**BTN2A1**	**0.51392**	**7**	**intron**	**1.463**
6	**rs3819720**	**G**	**A**	**0.37**		**8 tissues**			**Zbtb3**	**TAP2**	**0.22271**	**1f**	**intron**	**1.202**
6	rs10872224	G	T	0.6					GR	37kb 5' of MIR2113	0.18412	7	intron,non_C_trans	1.526
6	rs11243150	T	C	0.41		6 tissues	11 tissues	E2F6	E2F	5.3kb 3' of SSR1	0.60906	4	upstream_gene	6.307
6	rs1265099	A	C,G,T	0.42	15tissues	18 tissues	52 tissues	8 proteins		PSORS1C2	0.74401	4	3_prime_UTR	6.128
6	rs2205829	G	A	0.51	IPSC	ESC			Nrf1	39kb 5' of ZNF184	0.13454	5	downstream_gene	6.713
6	rs3115672	C	T	0.08		GI, THYM			Nr2f2	MSH5	0.86655	2b	synonymous	11.06
6	rs6568686	T	C	0.79		18 tissues	9 tissues		CEBPB,SRF	TRAF3IP2-AS1	0.60906	4	intron,non_C_trans	5.442
6	rs6914422	C	T	0.34		6 tissues	SKIN,SKIN		Smad3,Smad	5.4kb 3' of BAK1	0.55436	1f	intergenic	6.79
7	rs10282292	C	T	0.62			BRN,BRN		8 altered motifs	IMMP2L	0.18412	7	intron	9.786
8	rs11167136	G	A	0.44			BLD		ZEB1	TSNARE1	0.60906	4	intron	5.403
8	rs2010390	A	G	0.7	ESC, PLCNT, BLD	BLD, BRN	BLD		Hbp1	38kb 5' of PPP1R3B	0.995	1d	intron,non_C_trans	0.858

8	rs2285268	A	G	0.27		6 tissues	7 tissues		HDAC2,p300	768bp 3' of ZDHHC2	0.60906	4	3_prime_UTR	2.127
8	rs3802177	G	A	0.29		6 tissues			CEBPB,STAT,p300	SLC30A8	0.66702	5	3_prime_UTR	2.8
8	rs3849850	G	A	0.56					4 altered motifs	TRPS1	0.35857	6	intron	0.69
8	rs6983332	C	T	0.42					Hand1,Smad3	MSRA	0.22271	1f	intron	7.924
8	rs867743	G	A	0.6					5 altered motifs	407kb 3' of CA8	0.55195	6	intergenic	0.611
10	rs11257655	C	T	0.24	BLD, LNG	12 tissues	35 tissues	15 bound proteins	5 altered motifs	15kb 3' of CDC123	0.70497	4	downstream_gene	8.565
10	rs3814614	A	G	0.53						2kb 5' of GRID1	0.13454	5	upstream_gene	1.583
10	rs7085104	A	G	0.35	21 tissues	6 tissues	14 tissues		RXRA	C10orf32-AS3MT	0.60906	4	upstream_gene	5.851
11	rs1782507	G	T	0.68		ADRL			5 altered motifs	8.7kb 5' of FSHB	0.18412	7	intergenic	2.414
11	rs5215	C	T	0.66		7 tissues	4 tissues	CTCF	TCF4	KCNJ11	0.60906	4	missense	10.38
11	rs6590512	C	T	0.49	MUS, SKIN	9 tissues	7 tissues		CEBPG,HP1-site-factor,Sox	28kb 3' of SNX19	0.60906	4	intron,non_C_trans	1.539
11	rs9633835	G	A	0.36		5 tissues				ARNTL	0.18412	7	intron	0.129
12	rs1042725	C	T	0.51			6 tissues		8 altered motifs	HMGA2	0.60906	4	3_prime_UTR	6.735
12	rs1609520	G	A	0.73	6 tissues	18 tissues	6 tissues	AP2ALPHA,AP2GAMMA,BAF155		C12orf65	0.36649	3a	intron	0.058
12	rs6581987	G	T	0.67					STAT	87kb 3' of TSPAN8	0.58955	5	intron,non_C_trans	1.295
12	rs7975482	A	G	0.36		12 tissues	SKIN,GI		5 altered motifs	ZNF664	0.25211	5	intron	1.969
15	rs12437900	C	A	0.34	CRVX	10 tissues			Crx,GR,Lhx3	72kb 3' of LOC645752	0.18412	7	intergenic	1.632
15	rs12910334	G	A	0.29		GI				59kb 5' of LOC388152	0.60906	4	noncod_trans_exon	13.7
15	rs4702	G	A	0.57	BLD	17 tissues	4 tissues		CEBPB,Pax-4	FURIN	0.60906	4	3_prime_UTR	12.59
16	rs1894951	A	G	0.61	BRN	14 tissues	11 tissues		Myf	257kb 3' of CBLN1	0.60906	4	regulatory_region	0.846
16	rs7204797	C	T	0.44					NRSF,Sin3Ak-20	5.3kb 5' of TMEM219	0.22271	1f	intron	0.152
16	rs9941245	T	G	0.16					5 altered motifs	21kb 5' of GPRC5B	0.84	1c	intergenic	3.338
17	rs8068351	T	C	0.4					TATA	GGNBP2	0.17375	6	intron	10.78
18	rs17683730	G	A	0.24		ESDR, ESC, BRST			Cdc5,Pou5f1	WDR7	0.13454	5	intron	4.565
18	rs17747955	T	C	0.5		PANC				NOL4	0.58955	5	intron	7.562
18	rs9636107	A	G	0.5					Isl2	TCF4	0.18412	7	intron	6.492
19	rs1009136	A	G	0.69			BLD		4 altered motifs	MAU2	0.70497	4	intron	2.261
19	rs3786897	A	G	0.41		7 tissues	PANC		LF-A1,SP1,SZF1-1	PEPD	0.60906	4	intron	5.827
19	rs3810291	G	A	0.66		IPSC			4 altered motifs	ZC3H4	0.13454	5	intron	4.027
19	rs4804833	A	G	0.59	5 tissues	18 tissues	KID		AIRE	MAP2K7	0.60906	4	intron	3.07
19	rs4807179	G	A	0.61		11 tissues	ESC		Znf143	CSNK1G2	0.60906	4	3_prime_UTR	2.428
20	rs6126570	C	T	0.27					Irf	286kb 5' of ZFP64	0.13454	5	intergenic	20.9
22	rs738408	C	T	0.22		BLD, FAT			COMP1,EWSR1-FLI1	PNPLA3	0.60906	4	synonymous	0.113

Supplementary Table 4-3c. 'shared' Lead SNPs functional annotations.

Variants details		HaploReg									RegulomeDB		VEP	
chr	variant	Ref	Alt	EUR freq	Promoter HM	Enhancer HM	DNAse	Proteins bound	Motifs changed GWAS hits	RefSeq genes	probability	ranking	Consequence	CADD_ PHRED
1	rs10888679	T	C	0.3					Isl2	ELAVL4	0.16883	5	intron	8.185
1	rs1766377	C	T	0.81					4 altered motifs	AGBL4	0.504	6	intron	5.904
1	rs2250377	A	G	0.68		5 tissues			ATF4,CTCF,Rad21	SHISA4	0.13454	5	missense	19.63
1	rs3903399	T	C	0.21		BRN			7 altered motifs	CNTN2	0.13454	5	intron	12.94
2	rs12613687	C	T	0.18					Roaz,p53	170kb 3' of C2orf47	0.49385	6	intergenic	3.386
3	rs4678408	A	G	0.65						4.5kb 5' of TXNDC6	0.13454	5	upstream_gene	1.092
3	rs9873519	C	T	0.54					4 altered motifs	SLC12A8	0.58955	5	intron	0.258
4	rs151390	C	T	0.27			BRN,BRN,THYM		Bcl6b,STAT	SLC39A8	0.65852	5	3_prime_UTR	4.904
5	rs10077814	C	T	0.6					6 altered motifs	101kb 3' of MRPS30	0.51981	5	intergenic	12.1
6	rs707939	C	A	0.35			BLD		Pax-4	MSH5	0.60906	4	intron	3.765
7	rs2218378	A	G	0.37					5 altered motifs	CADPS2	0.16883	5	intron	1.54
7	rs7776707	C	A	0.34		7 tissues			Irf,STAT,ZEB1	17kb 3'of LOC100216546	0.60906	4	downstream_gene	8.454
8	rs2618451	T	C	0.39		6 tissues	MUS		12 altered motifs	BLK	0.57	1f	intron	0.176
8	rs2921077	G	A	0.46		LNG	IPSC		DMRT2,Hbp1	65kb 5' of SGK223	0.22271	1f	intergenic	1.606
8	rs6985109	G	A	0.53		ADRL	IPSC,ADRL		11 altered motifs	XKR6	0.172	1f	intron	3.486
9	rs2796441	G	A	0.42	PANC	8 tissues	13 tissues	GATA2	23 altered motifs	5.4kb 5' of TLE1	1	2b	upstream_gene	19.81
9	rs7848336	T	C	0.34					Sin3Ak-20	13kb 3' of PHF2	0.13454	5	intergenic	8.41
10	rs2280141	T	G	0.47		BLD, BRN	4 tissues		11 altered motifs	1.3kb 3' of PLEKHA1	0.49737	3a	downstream_gene	0.443
10	rs953097	T	C	0.64		ESC,ESDR, IPSC				ARHGAP19-SLIT1	0.18412	7	downstream_gene	6.144
11	rs10838158	G	A	0.69						HSD17B12	0.51392	7	intron	11.59
11	rs531897	C	T	0.68		8 tissues	OVRY		Cdx2	MIR100HG	0.18412	7	downstream_gene	1.756
11	rs7103411	C	T	0.76		7 tissues	SKIN,LNG		DMRT5	BDNF	0.60906	4	intron	3.616
12	rs10506133	T	C	0.57					10 altered motifs	123kb 5' of CPNE8	0.31869	6	intergenic	0.624
12	rs1169302	T	G	0.44					HNF4	HNF1A	0.60906	4	intron	1.238
12	rs13746	C	T	0.49		6 tissues	PLCNT,THYM		ERalpha-a	SPPL3	0.22271	1f	3_prime_UTR	1.039
14	**rs8009761**	**G**	**A**	**0.5**		**ESC,PANC**	**BLD**		**4 altered motifs**	**SYNDIG1L**	**0.60906**	**4**	**intron**	**0.232**
14	rs17522122	G	T	0.48					Pou5f1	613bp 3' of AKAP6	0.18412	7	downstream_gene	5.912
14	rs7156625	G	A	0.22		ESDR, MUS,LNG	ESDR,ADRL,MUS		7 altered motifs	NRXN3	0.60906	4	intron	13.79
14	rs8017993	A	G	0.28		BRST,SKIN,PLCNT	BRST,PLCNT,BRST		Nrf-2	C14orf153	0.22271	1f	intron	3.407
15	rs8036171	A	C	0.63		BRST			NF-I,PPAR	240kb 3' of MCTP2	0.2212	6	intergenic	0.769
15	rs938875	A	G	0.59		LIV			Pou5f1	MAP2K5	0.60906	4	intron	3.078
17	**rs17763551**	**G**	**A**	**0.15**		**4 tissues**	**ESC,PANC,BLD**		**SREBP,Zic**	**15kb 5' of ATP2A3**	**0.60906**	**4**	**intron,non_C_trans**	**4.043**
17	rs4925109	A	G	0.66	BLD, BRN,THYM	7 tissues	BLD,BLD,THYM		5 altered motifs	RAI1	0	5	intron	5.802
17	rs8069451	T	C	0.25					Brachyury,Hoxa5,Pax-5	FBXL20	0.18412	7	intron	0.236
17	rs8078510	G	A	0.26			BRN		GATA,Lmo2-complex	GIP	0.55436	1f	intron	1.167
20	rs1007090	T	C	0.7			5 tissues		7 altered motifs	RALY	0.81166	2b	intron	2.061
22	rs137845	A	G	0.53		4 tissues	ESDR		Pbx-1,SETDB1,SZF1-1	IL17REL	0.60906	4	intron	0.154
22	rs1569858	A	G	0.62			5 tissues			EP300	0.13454	5	intron	0.049

Supplementary Table 4-4a. 'subset1sided' top SNPs functional annotations.

Variants details		HaploReg									RegulomeDB		VEP	
chr	variant	Ref	Alt	EUR freq	Promoter histone marks	Enhancer histone marks	DNAse	Proteins bound	Motifs changed GWAS hits	RefSeq genes	probability	ranking	Consequence	CADD_ PHRED
1	rs6677300	G	A	0.3					THAP1,YY1	AGBL4	0.18412	7	intron	8.848
1	rs630372	G	A	0.23		SKIN			7 altered motifs	12kb 3' of SEC16B	0.462	5	intergenic	0.325
1	rs197419	A	C	0.42		LIV			13 altered motifs	3kb 3' of KCND3	0.90167	5	downstream_gene	5.876
2	rs12998587	T	C	0.71		14 tissues			CEBPB,SIX5	RBMS1	0.51392	7	intron	6.314
3	rs247975	T	C	0.54					NRSF,Znf143	8.8kb 5' of NLGN1	0.13454	5	intergenic	4.204
3	rs1145101	T	C	0.81	7 tissues	14 tissues	12 tissues	CEBPB		STAG1	0.60906	4	intron	2.391
5	**rs10515678**	**C**	**T**	**0.28**					**RFX5**	**538kb 5' of NMUR2**	**0.41972**	**5**	**intergenic**	**0.447**
6	rs3737095	T	C	0.44					14 altered motifs	DNAH8	0.78994	3a	splice_polypyrimidine_tract, intron	12.52
6	rs2758259	C	T	0.44					Irf	89kb 5' of PTP4A1	0.60906	4	downstream_gene	4.94
8	rs6990912	A	C	0.68		BRN			20 altered motifs	7.9kb 3' of LOC157273	0.63	6	intergenic	10.92
11	rs4938180	C	T	0.51		9 tissues	11 tissues	CTCF,RAD21	Brachyury,CTCF	3.3kb 3' of CADM1	0.55436	1f	3_prime_UTR	0.725
14	rs3007105	C	T	0.42					Foxa,LBP-1,Pax-5	MDGA2	0.15078	6	intron	3.18
20	rs6062689	G	A	0.33	MUS	15 tissues	MUS		ZNF263	SLCO4A1	0.8575	1b	intron	1.269

Supplementary Table 4-4b. 'subset2sided' top SNPs functional annotations.

Variants details		HaploReg									RegulomeDB		VEP	
chr	variant	Ref	Alt	EUR freq	Promoter histone marks	Enhancer histone marks	DNAse	Proteins bound	Motifs changed GWAS hits	RefSeq genes	probability	ranking	VEP	CADD_ PHRED
1	rs9729667	C	T	0.3					Hoxd10,Pax-4	5kb 5' of NEXN	0.13454	5	intron,non_coding_ transcript	0.076
1	rs340874	T	C	0.54	7 tissues	14 tissues	6 tissues		Cdx2	2.6kb 5' of PROX1	0.60906	4	upstream_gene	11.25
2	**rs6718758**	**C**	**A**	**0.31**	**BLD**	**BLD, THYM**			**Maf,Pbx3,SIX5**	**286kb 3' of MIR4432**	**0.20075**	**6**	**upstream_gene**	**9.3**
2	rs1016287	T	C	0.73					5 altered motifs	15kb 3' of FLJ30838	0.60906	4	intergenic	19.22
3	rs2276825	T	C	0.28		4 tissues	LNG		Ik-3,ZEB1	TMEM110-MUSTN1	0.55436	1f	intron	1.218
3	rs548288	T	C	0.75	23 tissues	SPLN	29 tissues	POL2		PCCB	0.60906	4	intron	0.736
3	**rs4678874**	**A**	**G**	**0.27**						**STAC**	**0.18412**	**7**	**intron**	**0.86**
3	rs4854912	C	T	0.18					Pou5f1,Sox	74kb 5' of DNAJC19	0.13454	5	intergenic	0.627
3	rs5398	G	A	0.29	STRM, VAS	15 tissues	10 tissues	11 bound proteins	11 altered motifs	SLC2A2	0.47552	3a	missense	15.53
3	rs832188	T	C	0.8	ESC, IPSC, BLD	13 tissues	11 tissues	NANOG	Foxa	THOC7	0.60906	4	intron	2.344
4	**rs2904221**	**T**	**C**	**0.74**		**LIV, GI**		**FOXA1,FOXA 2,P300**	**5 altered motifs**	**FAM47E**	**1**	**3b**	**intron**	**1.632**
4	rs7660298	T	C	0.49					4 altered motifs	NHEDC2	0.4855	6	intron	5.298
5	rs2600834	T	C	0.78					Pou2f2,Pou5f1	SLCO4C1	0.18412	7	intron	1.149
5	rs815611	G	A	0.55		10 tissues			4 altered motifs	52kb 5' of GALNT10	0.951	1d	intergenic	3.481
6	rs4412207	G	A	0.56		CRVX			Foxp3,HNF4,PPA R	62kb 5' of MIR2113	0.63123	5	intron,non_coding_ transcript	11.4
6	rs13207230	G	A	0.54	ESDR, BLD	14 tissues	ESDR,BLD	CEBPB,GATA 2,POL2	Smad	66kb 5' of RREB1	0.60906	4	upstream_gene	5.986
6	rs3130573	A	G	0.37		5 tissues	IPSC		p300	PSORS1C2	0.75833	1f	intron	2.568
6	rs3800316	A	C	0.25	11 tissues	7 tissues	8 tissues		15 altered motifs	21kb 3' of POM121L2	0.47552	3a	regulatory_region	0.953
6	rs4148870	C	T	0.45	24 tissues		45 tissues	13 bound proteins		TAP2	0.60906	4	intron	6.718
6	rs6568686	T	C	0.79		18 tissues	9 tissues		CEBPB,SRF	TRAF3IP2-AS1	0.60906	4	intron,non_coding_ transcript	5.442
7	rs38744	A	G	0.52		4 tissues	MUS		ATF3	IMMP2L	0.55436	1f	intron	5.959
8	**rs2588120**	**G**	**A**	**0.65**		**10 tissues**	**VAS**		**Hdx,Pou2f2**	**PDGFRL**	**0.13454**	**5**	**intron**	**2.863**
8	rs13266634	C	T	0.29	ESDR, BRN	12 tissues	17 tissues	5 bound proteins	8 altered motifs	SLC30A8	0.60906	4	missense	15.11
8	rs10097784	C	A	0.57						TRPS1	0.18412	7	intron	3.932
8	rs7018304	A	G	0.6						384kb 3' of CA8	0.18412	7	intergenic	0.709

10	**rs10906025**	**T**	**C**	**0.47**			**BLD,BLD,LNG**		**Pou2f2,RREB-1**	**C10orf47**	**0.13454**	**5**	**intron**	**0.58**
10	rs4933391	G	A	0.36		6 tissues	HRT,BLD		GATA,HMGN3,SREBP	GRID1	0.00875	5	intron	5.951
10	rs4917985	G	A	0.38					Brachyury	C10orf32	0.51392	7	3_prime_UTR	1.344
11	rs642126	G	A	0.67		10 tissues			4 altered motifs	9.5kb 3' of FSHB	0.18412	7	intergenic	5.618
11	rs10791097	T	G	0.52			SKIN		5 altered motifs	27kb 3' of SNX19	0.86056	3a	intron,non_coding_transcript	9.639
12	**rs6581977**	**C**	**T**	**0.44**		**FAT, VAS, MUS**	**11 tissues**	**CTCF,RAD21**	**8 altered motifs**	**28kb 5' of PTPRR**	**0.70497**	**4**	**regulatory_region**	**0.953**
15	rs7178572	A	G	0.68		9 tissues			THAP1	HMG20A	0.18412	7	intron	0.738
15	rs11635505	G	T	0.29			LNG			UBE2Q2P1	0.15348	6	intron,non_coding_transcript	0.726
15	rs12910825	A	G	0.39		BRST, SKIN	ESDR,IPSC,BLD		Hbp1	PRC1	0.55436	1f	intron	4.483
16	rs1894951	A	G	0.61	BRN	14 tissues	11 tissues		Myf	257kb 3' of CBLN1	0.60906	4	regulatory_region	0.846
17	rs2306590	G	A	0.39		4 tissues	THYM,BRST, SKIN		6 altered motifs	MYO19	0.51695	1b	missense	16.88
18	**rs12958029**	**T**	**C**	**0.38**		**PANC**			**DBP,Irf**	**42kb 3' of NOL4**	**0.86333**	**5**	**intergenic**	**4.639**
18	rs9636107	A	G	0.5					Isl2	TCF4	0.58955	5	intron	7.562
19	rs2546057	A	C	0.48					Pou5f1,Sox	14kb 3' of KCTD15	0.56066	5	intergenic	7.987
19	rs3810291	G	A	0.66		IPSC			4 altered motifs	ZC3H4	0.60906	4	3_prime_UTR	2.428
19	**rs2656865**	**C**	**T**	**0.3**		**SKIN**	**ESDR**			**PLK5P**	**0.60906**	**4**	**intron**	**0.596**
20	rs6097012	T	C	0.28		4 tissues	4 tissues		5 altered motifs	338kb 5' of ZFP64	0.60906	4	regulatory_region	15.38
22	rs738408	C	T	0.22		BLD, FAT			COMP1,EWSR1-FLI1	PNPLA3	0.60906	4	synonymous	0.113

Supplementary Table 4-4c. 'shared' top SNPs functional annotations.

Variants details		HaploReg									RegulomeDB		VEP	
chr	variant	Ref	Alt	EUR freq	Promoter histone marks	Enhancer histone marks	DNAse	Proteins bound	Motifs changed GWAS hits	RefSeq genes	probability	ranking	Consequence	CADD_ PHRED
1	rs4078484	C	T	0.18	BRN	BRN			GR,Pbx-1,Zic	ELAVL4	1	5	intron	3.246
1	rs1766377	C	T	0.81					4 altered motifs	AGBL4	0.504	6	intron	5.904
1	rs2820312	G	A	0.32		9 tissues		PU1	TAL1	LMOD1	0.41664	3a	missense	7.036
1	rs3903399	T	C	0.21		BRN			7 altered motifs	CNTN2	0.13454	5	intron	12.94
2	**rs13425918**	**A**	**G**	**0.25**		**BRN**			**Pax-5**	**GIGYF2**	**0.13454**	**5**	**intron**	**3.973**
2	rs12613687	C	T	0.18					Roaz,p53	170kb 3' of C2orf47	0.49385	6	intergenic	3.386
3	**rs830644**	**T**	**C**	**0.59**		**THYM, BLD**			**Nkx2**	**32kb 5' of FOXP1**	**0.05854**	**6**	**intergenic**	**2.728**
3	**rs4678398**	**A**	**G**	**0.27**	**GI**		**GI,GI**		**Cdx2,Hoxa13,Hoxa9**	**CLDN18**	**0.32932**	**5**	**downstream_gene**	**4.822**
3	rs569255	G	A	0.55	LIV	PANC, THYM			CTCF	SLC12A8	0.13454	5	intron	11.52
4	rs233816	G	T	0.21				MAFK		SLC39A8	0.58955	5	intron	2.182
5	rs729599	G	A	0.62		PANC			5 altered motifs	27kb 3' of MRPS30	0.58955	5	intergenic	0.425
6	rs3134899	C	T	0.76		IPSC			DEC,Ets,Nrf1	MICB	0.84	1f	intron	1.34
7	**rs7778318**	**C**	**T**	**0.25**		**ESDR, ESC, IPSC**	**4 tissues**		**GR,Pax-8**	**CADPS2**	**0.60906**	**4**	**intron**	**9.761**
8	rs6601414	G	A	0.42					Arid5a,Cdx2	MSRA	0	5	intron	1.269
8	rs7832722	G	A	0.41		ESDR			HEY1,Nrf1,Rad21	XKR6	0.13454	5	intron	3.565
9	rs2796441	G	A	0.42	PANC	8 tissues	13 tissues	GATA2	23 altered motifs	5.4kb 5' of TLE1	1	2b	upstream_gene	19.81
9	rs2398871	C	T	0.31					Myb	42kb 3' of PHF2	0.60906	4	intergenic	1.973
11	rs10838146	C	T	0.74					Elf5,NF-I,PU.1	13kb 5' of HSD17B12	0.13454	5	intergenic	2.081
11	rs513069	A	G	0.68		4 tissues	4 tissues		Rhox11	MIR100HG	0.14339	5	downstream_gene	0.72
11	rs10501087	T	C	0.22			LNG		Foxa,Pbx-1	BDNF-AS1	0.02167	6	intron,non_coding _transcript	0.469
12	rs10506133	T	C	0.57					10 altered motifs	123kb 5' of CPNE8	0.31869	6	intergenic	0.624
12	rs1169302	T	G	0.44					HNF4	HNF1A	0.60906	4	intron	1.238
12	rs1727302	G	A	0.73	BLD	6 tissues	4 tissues			8kb 3' of MPHOSPH9	0.60906	4	downstream_gene	6.962
14	rs1125221	G	A	0.53					7 altered motifs	20kb 5' of SYNDIG1L	0.45917	5	intergenic	3.831
15	rs11633626	C	A	0.63			PLCNT		Sox	244kb 3' of MCTP2	0.13454	5	intergenic	2.122
15	rs3784692	C	T	0.59					5 altered motifs	MAP2K5	0.20075	6	intron	14.26
17	**rs7217945**	**G**	**A**	**0.26**		**5 tissues**	**13 tissues**		**Foxo,STAT,Sp100**	**TRPV1**	**0.71614**	**5**	**intron**	**6.265**
17	rs801426	G	A	0.25		GI	GI,GI,LNG		7 altered motifs	FBXL20	0.01316	5	intron	0.729
17	rs2270576	C	T	0.26		8 tissues	8 tissues	POL2,POL24H8	8 altered motifs	SNF8	0.80136	1b	synonymous	11.14
19	rs6909	A	G	0.32					4 altered motifs	GATAD2A	0.60906	4	3_prime_UTR	3.189
20	**rs8119937**	**C**	**T**	**0.7**		**BLD**	**BLD**		**Mxi1,SREBP**	**RALY**	**0.18412**	**7**	**intron**	**2.993**
22	rs137862	C	A	0.53		4 tissues	GI,LNG		7 altered motifs	IL17REL	0.82852	2b	intron	0.021
22	rs9611497	A	G	0.62		MUS				EP300	0.01167	5	intron	0.406

Supplementary Table 4-5. MAGMA significant genes.

CHR	START	STOP	Gene	Gene's full name	SNP	Phenotype (ASSET)
1	48997527	50499626	AGBL4	AGBL Carboxypeptidase 4	rs1766377	SCZ,OB1,OB2,OB3
					rs1343424,rs6677300	SCZ,OB1,OB2,OB3
1	77737662	78026654	AK5	Adenylate Kinase 5	rs12751064	SCZ,OB1,OB2,OB3
14	32788479	33303268	AKAP6	A-kinase anchor protein 6	rs17522122	SCZ,T2D,OB1,OB2,OB3,CMM
14	104019299	104058236	APOPT1	Cytochrome c oxidase assembly factor 8	rs8017993	SCZ,CMM , OB3
10	98980930	99062430	ARHGAP19	Rho GTPase-activating protein 19	rs953097	SCZ,T2D,OB1,OB2,OB3,CMM
11	13289325	13409813	ARNTL	Aryl hydrocarbon receptor nuclear translocator-like protein 1	rs9633835	SCZ,T2D,OB1,OB2,OB3,CMM
10	104619183	104662656	AS3MT	Arsenite methyltransferase	rs7085104	SCZ, OB1,OB2,OB3
3	63839785	63990240	ATXN7	Ataxin-7	---	---
2	60677302	60790633	BCL11A	B-cell lymphoma/leukemia 11A	rs10189857	SCZ, T2D,CMM,OB3
11	27675440	27753605	BDNF	Brain-derived neurotrophic factor	rs10501087	SCZ,OB1,OB2,OB3,CMM
8	11341521	11423108	BLK	Tyrosine-protein kinase Blk	rs2618451	SCZ,T2D,OB1,OB2,CMM
6	26448132	26477849	BTN2A1	Butyrophilin subfamily 2 member A1	rs3799380	SCZ, OB2,OB3,CMM
12	123707844	123743651	C12orf65	Probable peptide chain release factor C12orf65	rs1609520	SCZ,OB3,T2D,CMM
11	115038933	115385241	CADM1	Cell adhesion molecule 1	rs4938180	SCZ,T2D,OB1,OB2,OB3,CMM
7	121957478	122536813	CADPS2	Calcium-dependent secretion activator 2	rs7778318	SCZ,T2D,OB1,OB2,OB3
1	205002340	205048173	CNTN2	Contactin-2	rs3903399	SCZ,T2D,OB1,OB3,CMM
19	1931148	1982337	CSNK1G2	Casein kinase I isoform gamma-2	rs4807179	SCZ,T2D,OB1,OB2,OB3,CMM
6	38673117	38999574	DNAH8	Dynein heavy chain 8	rs3737095	SCZ,T2D,OB2,OB3,CMM
1	6693228	6771966	DNAJC11	DnaJ homolog subfamily C member 11	rs4854912	SCZ,T2D,OB1,OB2,OB3,CMM
1	50503686	50670442	ELAVL4	ELAV-like protein 4	rs4078484	SCZ,OB1,OB2,OB3
22	41478614	41577081	EP300	Histone acetyltransferase p300	rs9611497	SCZ,T2D,OB1,OB2,CMM
1	112263686	112308419	FAM212B	Family With Sequence Similarity 212 Member B	rs1546924	SCZ,T2D,OB1,OB2,CMM
4	77125193	77205936	FAM47E	Family With Sequence Similarity 47 Member E	rs2904221	SCZ,T2D,OB1,OB2,OB3,CMM
17	37407897	37568530	FBXL20	F-Box And Leucine Rich Repeat Protein 20	rs801426	SCZ,T2D,OB3,CMM
15	91417665	91440006	FES	Tyrosine-protein kinase Fes/Fps	---	---
7	121940373	121961173	FEZF1	Fez family zinc finger protein 1	---	---
3	71002865	71643140	FOXP1	Forkhead box P1	rs830644, rs6779258	SCZ,T2D,CMM,OB3
					rs11720523	SCZ,T2D,CMM
11	30242563	30257824	FSHB	Follitropin subunit beta	rs642126	SCZ, T2D,OB1,OB2,OB3,CMM
15	91401885	91427687	FURIN	Furin, Paired Basic Amino Acid Cleaving Enzyme	rs4702	SCZ,OB3,CMM
17	34890737	34947278	GGNBP2	Gametogenetin-binding protein 2	rs8068351	SCZ,T2D,OB1,OB2,OB3,CMM
2	233552015	233726287	GIGYF2	GRB10-interacting GYF protein 2	rs13425918, rs6704768	SCZ,OB2,OB3,T2D
17	47034918	47055955	GIP	Gastric inhibitory polypeptide	rs8078510	SCZ,T2D,OB2,CMM
10	87358312	88136250	GRID1	Glutamate receptor ionotropic, delta-1	rs4933391	SCZ,T2D,OB1,OB2,OB3
12	66208240	66361071	HMGA2	High mobility group AT-hook 2	rs1042725	SCZ,T2D,CMM, OB1,OB2,OB3
					rs7968682	SCZ,T2D,CMM

12	121405861	121441315	HNF1A	Hepatocyte nuclear factor 1-alpha	rs1169302	SCZ,T2D,CMM,OB3
11	43692108	43879169	HSD17B12	Very-long-chain 3-oxoacyl-CoA reductase	rs10838146	SCZ,T2D,OB1,OB2,OB3,CMM
22	50431942	50461055	IL17REL	Putative interleukin-17 receptor E-like	rs137862	SCZ,T2D,CMM,OB3
7	110302106	111212588	IMMP2L	Mitochondrial inner membrane protease subunit 2	rs38744	SCZ,T2D,OB1,OB2,OB3,CMM
3	52846006	52874717	ITIH4	Inter-alpha-trypsin inhibitor heavy chain H4	rs1573815	SCZ,T2D,OB1,OB2,OB3,CMM
11	17405795	17420878	KCNJ11	ATP-sensitive inward rectifier potassium channel 11	rs5215	SCZ,OB1,OB3,T2D,CMM
2	233629512	233651275	KCNJ13	Inward rectifier potassium channel 13	rs1973675	SCZ,OB2,OB3
15	67825021	68100455	MAP2K5	Mitogen-activated protein kinase kinase 5	rs3784692	SCZ,T2D,OB1,OB2,OB3,CMM
19	7958665	7980363	MAP2K7	Mitogen-activated protein kinase kinase 7	rs3786897	SCZ,B2,OB3,T2D
19	19421496	19470563	MAU2	MAU2 chromatid cohesion factor homolog	rs769267	SCZ,T2D,CMM
					rs1009136	SCZ,T2D,OB1,OB2,OB3
14	47307826	48154157	MDGA2	MAM domain-containing glycosylphosphatidylinositol anchor protein 2	rs3007105	SCZ,T2D,OB1,OB2,OB3,CMM
3	49945302	49977445	MON1A	Vacuolar fusion protein MON1 homolog A	rs7613875	SCZ,T2D,OB1,OB2,OB3,CMM
12	123639943	123727785	MPHOSPH9	M-phase phosphoprotein 9	rs1727302	SCZ,OB3,T2D,CMM
6	31697725	31731455	MSH5	MutS protein homolog 5	rs3115672	SCZ,T2D,OB3
					rs707939	SCZ,CMM,OB1,OB2,OB3
8	9901830	10287401	MSRA	Mitochondrial peptide methionine sulfoxide reductase	rs6983332, rs6601414	SCZ,T2D,OB1,CMM
3	52866131	52879235	MUSTN1	Musculoskeletal embryonic nuclear protein 1	rs1573815	SCZ,T2D,OB1,OB2,OB3,CMM
3	173105570	174005434	NLGN1	Neuroligin-1	rs247975	SCZ,T2D,OB1,OB2,OB3,CMM
3	137979279	138059018	NME9	Thioredoxin domain-containing protein 6	rs4678408	SCZ,T2D,OB2,CMM
18	31430064	31813515	NOL4	Nucleolar protein 4	rs17747955,rs12958029	SCZ,T2D,OB1,OB3,CMM
14	78626716	80335633	NRXN3	Neurexin-3	rs7156625	SCZ,T2D,OB1,OB2,OB3,CMM
3	135959167	136057737	PCCB	Propionyl-CoA carboxylase beta chain	rs548288	SCZ,OB3,T2D,CMM
19	33876855	34022799	PEPD	Xaa-Pro dipeptidase	rs4804833	SCZ,T2D,OB2,OB3,CMM
22	44309619	44344451	PNPLA3	1-acylglycerol-3-phosphate O-acyltransferase PNPLA3	rs738408	SCZ,OB1,CMM,T2D,OB3
1	214151278	214215853	PROX1	Prospero homeobox protein 1	rs340874	SCZ,T2D,CMM
6	31072608	31108869	PSORS1C1	Psoriasis susceptibility 1 candidate 1	rs3130573	SCZ,T2D,OB2,OB3,CMM
6	31104311	31117127	PSORS1C2	Psoriasis susceptibility 1 candidate 2	rs3130573	SCZ,T2D,OB2,OB3,CMM
17	17574787	17715767	RAI1	Retinoic acid-induced protein 1	rs4925109	SCZ,T2D,CMM,OB1,OB2,OB3
20	32571458	32671991	RALY	RNA-binding protein Raly	rs8119937	SCZ,T2D,OB1,OB2,CMM
3	49967474	50115685	RBM6	RNA-binding protein 6	rs7613875	SCZ,T2D,OB1,OB2,OB3,CMM
2	161127662	161360366	RBMS1	RNA-binding motif	rs7572970	SCZ,T2D,OB1,CMM
6	31720773	31733627	SAPCD1	Suppressor APC domain containing 1	---	---
1	201847797	201862715	SHISA4	Shisa family member 4	rs2250377	SCZ,T2D,OB1,OB2,OB3,CMM
3	124800480	124941609	SLC12A8	Solute carrier family 12 member 8	rs569255	SCZ,T2D,OB3,CMM,OB1,OB2
3	170713137	170754768	SLC2A2	Solute carrier family 2	rs8192675, rs5398	SCZ,OB1,OB2,OB3,T2D,CMM
8	117952512	118189953	SLC30A8	Zinc transporter 8	rs3802177, rs13266634	SCZ,OB1,OB2,T2D,CMM
4	103171198	103276655	SLC39A8	Zinc transporter ZIP8	rs233816	SCZ,T2D,OB1,OB2,OB3,CMM
4	103805205	103957552	SLC9B1	Sodium/hydrogen exchanger 9B1	rs7676765	SCZ,T2D,OB1,OB2,OB3,CMM

20	61263795	61318137	SLCO4A1	Solute carrier organic anion transporter family member 4A1	rs2427363, rs6062689	SCZ,T2D,OB1,OB3,CMM
5	101706492	101844720	SLCO6A1	Solute carrier organic anion transporter family member 6A1	rs10515326	SCZ,T2D,OB2
2	228843670	229056836	SPHKAP	A-kinase anchor protein SPHKAP	rs12615058	SCZ,T2D,OB1,OB2,OB3,CMM
12	121199313	121352155	SPPL3	Signal peptide peptidase-like 3	rs13746	SCZ,T2D,OB2,OB3,CMM
3	136054077	136481245	STAG1	Cohesin subunit SA-1	rs1145101	SCZ,OB3
19	19386320	19441321	SUGP1	SURP and G-patch domain-containing protein 1	rs6909	SCZ,T2D,OB1,OB2,OB3
14	74871596	74902805	SYNDIG1L	Synapse differentiation-inducing gene protein 1-like	rs1125221	SCZ,T2D,OB1,OB2,OB3
6	32788610	32816547	TAP2	Antigen peptide transporter 2	rs4148870	SCZ,T2D,CMM
18	52888562	53313252	TCF4	Transcription factor 4	rs9636107	SCZ,T2D
3	63818546	63859597	THOC7	THO complex subunit 7 homolog	rs832188	SCZ,OB3,T2D,CMM
9	84197598	84313596	TLE1	Transducin-like enhancer protein 1	rs2796441	SCZ,T2D,CMM,OB3
3	52869772	52941597	TMEM110	STIM-Activating Enhancer Encoded By TMEM110	rs2276825	SCZ,T2D,OB1,OB2,OB3,CMM
16	29963351	29985373	TMEM219	Insulin-like growth factor-binding protein 3 receptor	rs7204797	SCZ,T2D,OB1,OB2,OB3
3	36867308	36996548	TRANK1	Tetratricopeptide repeat and ankyrin repeat containing 1	rs9811916	SCZ,T2D,CMM
8	116419724	116723299	TRPS1	Zinc finger transcription factor Trps1	rs10097784	SCZ,T2D,OB1,OB2,OB3
8	143292441	143494543	TSNARE1	t-SNARE domain containing 1	rs11167136	SCZ,T2D,OB1,OB2,OB3,CMM
18	54308590	54698036	WDR7	WD repeat domain 7	rs9636107	SCZ,T2D
8	10752654	11068875	XKR6	XK-related protein 6	rs7832722	SCZ,T2D,OB1,OB2,CMM,OB3
19	47566444	47627009	ZC3H4	Zinc finger CCCH-type containing 4	rs3810291	SCZ,T2D,OB1,OB2,OB3,CMM
8	17003836	17081241	ZDHHC2	Palmitoyltransferase ZDHHC2	rs2285268	SCZ,OB2,OB3,CMM
12	124447699	124500986	ZNF664	Zinc finger protein 664	rs7975482	SCZ,OB1,OB2,T2D,CMM

Supplementary Table 4-6. MAGMA gene-set analysis results.

microRNA targets (MsigDB c3)	
GeneSet	**genes**
MIR205_5P	PROX1, SLC30A8
MIR3202	SLC30A8, ZNF664
MIR4524A_5P	PROX1, SLC30A8, HMGA2
MIR4524B_5P	PROX1, SLC30A8, HMGA2
TF targets (MsigDB c3)	
GeneSet	**genes**
ZNF407_TARGET_GENES	GIGYF2, MSRA
LMTK3_TARGET_GENES	TMEM110, SLC30A8
BPTF_TARGET_GENES	AK5, PROX1, SLC30A8
SOX3_TARGET_GENES	TLE1, AKAP6
HNF4ALPHA_Q6	TLE1, HNF1A
HNF4_01_B	TLE1, HNF1A
GO Biological proc (MsigDB c5)	
GeneSet	**genes**
HP_ABNORMAL_WAIST_TO_HIP_RATIO	SLC2A2, SLC30A8
HP_DIABETES_MELLITUS	FOXP1, SLC2A2, SLC30A8, KCNJ11, HMGA2, TCF4
HP_INSULIN_RESISTANCE	SLC2A2, SLC30A8, KCNJ11
HP_LATE_ONSET	SLC2A2, SLC30A8
HP_TYPE_II_DIABETES_MELLITUS	SLC2A2, SLC30A8
GOBP_AMIDE_METABOLIC_PROCESS	SLC30A8, FURIN
GOBP_CARBOHYDRATE_HOMEOSTASIS	MUSTN1, SLC2A2, SLC30A8
GOBP_HORMONE_TRANSPORT	SLC2A2, SLC30A8, ARNTL, KCNJ11, HMGA2
GOBP_INORGANIC_ION_IMPORT_ACROSS_PLASMA_MEMBRANE	SLC30A8, KCNJ11
GOBP_INSULIN_SECRETION	SLC2A2, SLC30A8, ARNTL, KCNJ11
GOBP_PROTEIN_LOCALIZATION_TO_EXTRACELLULAR_REGION	MON1A, SLC2A2, SLC30A8, ARNTL, KCNJ11
GOBP_REGULATION_OF_HORMONE_SECRETION	SLC2A2, SLC30A8, ARNTL, KCNJ11, HMGA2
GOBP_REGULATION_OF_INSULIN_SECRETION	SLC2A2, SLC30A8, ARNTL, KCNJ11
GOBP_REGULATION_OF_PEPTIDE_TRANSPORT	SLC2A2, SLC30A8, ARNTL, KCNJ11, HMGA2
GOBP_REGULATION_OF_PROTEIN_SECRETION	SLC2A2, SLC30A8, ARNTL, KCNJ11
GOBP_REGULATION_OF_SECRETION	SLC2A2, SLC30A8, ARNTL, KCNJ11, HMGA2, FES
GOBP_SIGNALING_RECEPTOR_LIGAND_PRECURSOR_PROCESSING	SLC30A8, FURIN
GOBP_METAL_ION_TRANSPORT	TMEM110, SLC9B1, SLC30A8, KCNJ11
GOBP_MONOATOMIC_CATION_TRANSMEMBRANE_TRANSPORT	TMEM110, SLC9B1, SLC30A8, KCNJ11
GOBP_SIGNAL_RELEASE	SLC2A2, SLC30A8, ARNTL, KCNJ11, HMGA2

Supplementary Table 4-7. FUMA gene-set analysis results.

microRNA (MsigDB c3)					
GeneSet	**N**	**n**	**P-value**	**Adjust. P**	**genes**
MIR5688	768	15	2.50E-07	6.49E-04	AGBL4, AK5, BCL11A, KCNJ13, ATXN7, FOXP1, NLGN1, CADPS2, ARHGAP19-SLIT1, ARHGAP19, BDNF, CADM1, MPHOSPH9, AKAP6, NRXN3, TCF4
ACATATC_MIR190	60	5	3.59E-06	4.66E-03	BCL11A, NLGN1, TRPS1, CADM1, TCF4
MIR495_3P	688	12	1.38E-05	7.46E-03	AGBL4, BCL11A, ATXN7, NLGN1, CADPS2, ARHGAP19-SLIT1, ARHGAP19, BDNF, CADM1, MPHOSPH9, AKAP6, NRXN3, TCF4
MIR548F_5P	471	10	1.44E-05	7.46E-03	ELAVL4, PROX1, RBMS1, KCNJ13, SLC12A8, PCCB, AS3MT, HMGA2, TCF4, ZC3H4
MIR548AJ_5P_MIR548G_5P_MIR548X_5P	471	10	1.44E-05	7.46E-03	ELAVL4, PROX1, RBMS1, KCNJ13, SLC12A8, PCCB, AS3MT, HMGA2, TCF4, ZC3H4
MIR129_5P	360	8	7.89E-05	3.20E-02	ELAVL4, RBMS1, AS3MT, CADM1, NRXN3, NOL4, TCF4, EP300
MIR4760_3P	271	7	8.89E-05	3.20E-02	GIGYF2, STAG1, NLGN1, ZDHHC2, AS3MT, AKAP6, MDGA2, MDGA2
CAAGGAT_MIR362	62	4	9.86E-05	3.20E-02	DNAJC11, BCL11A, NOL4, SUGP1
MIR6783_3P	203	6	1.39E-04	4.03E-02	MUSTN1, TMEM110-MUSTN1, THOC7, BDNF, MAU2, PEPD
TF targets (MsigDB c3)					
GeneSet	**N**	**n**	**P-value**	**Adjust. P**	**genes**
CDC5_01	256	10	5.87E-08	6.55E-05	RBMS1, FOXP1, SLC12A8, TRPS1, TLE1, ARNTL, CADM1, NRXN3, FES, NOL4
RREB1_01	206	8	1.39E-06	5.25E-04	ELAVL4, RBMS1, KCNJ13, BDNF, CADM1, FES, FBXL20, EP300
CAGGTA_AREB6_01	765	14	1.41E-06	5.25E-04	ELAVL4, PROX1, BCL11A, FOXP1, PSORS1C2, TRPS1, ARNTL, BDNF, HMGA2, HNF1A, NRXN3, GGNBP2, NOL4, MAP2K7
YTATTTTNR_MEF2_02	698	13	2.88E-06	7.08E-04	ELAVL4, CNTN2, BCL11A, FOXP1, SLC12A8, FAM47E-STBD1, ARNTL, KCNJ11, BDNF, NRXN3, MAP2K5, NOL4, RALY
TAAWWATAG_RSRFC4_Q2	166	7	3.80E-06	7.08E-04	ELAVL4, FOXP1, SLC12A8, TRPS1, NRXN3, TCF4, RALY
EVI1_04	236	8	3.81E-06	7.08E-04	ELAVL4, BCL11A, KCNJ13, BDNF, NRXN3, MAP2K5, RAI1, PEPD
CEBPA_01	250	8	5.82E-06	9.27E-04	KCNJ13, MSH5, BDNF, CADM1, NRXN3, FES, NOL4, EP300
OCT1_02	206	7	1.56E-05	2.17E-03	ELAVL4, KCNJ13, FOXP1, TRPS1, NRXN3, FES, TCF4
FAC1_01	218	7	2.24E-05	2.78E-03	BCL11A, RBMS1, FOXP1, MSRA, SPPL3, NRXN3, FBXL20
SP3_Q3	243	7	4.48E-05	4.22E-03	CNTN2, FOXP1, KCNJ11, BDNF, GGNBP2, NOL4, EP300
AREB6_04	245	7	4.72E-05	4.22E-03	ELAVL4, BCL11A, MSH5, TRPS1, BDNF, FES, TCF4
CAGCTG_AP4_Q5	1479	17	4.73E-05	4.22E-03	SHISA4, CNTN2, GIGYF2, KCNJ13, FOXP1, SLC12A8, SLC30A8, HMGA2, SPPL3, MAP2K5, FES, RAI1, GGNBP2, NOL4, TCF4, CSNK1G2, PEPD
CEBP_Q3	250	7	5.36E-05	4.22E-03	ELAVL4, BCL11A, KCNJ13, TRPS1, BDNF, CADM1, NRXN3
CDX2_Q5	255	7	6.07E-05	4.22E-03	AGBL4, ELAVL4, FOXP1, STAG1, CADM1, HMGA2, MAP2K5
NF1_Q6	259	7	6.69E-05	4.22E-03	BCL11A, STAG1, IMMP2L, BDNF, HMGA2, MDGA2, MDGA2, NOL4
GFI1_01	260	7	6.86E-05	4.22E-03	ELAVL4, BDNF, HNF1A, NRXN3, MAP2K5, TCF4, MAP2K7
OCT1_01	261	7	7.03E-05	4.22E-03	KCNJ13, FOXP1, STAG1, TRPS1, CADM1, NRXN3, TCF4
STAT5A_03	262	7	7.20E-05	4.22E-03	BCL11A, KCNJ13, TMEM110, TRPS1, CADM1, TCF4, EP300
AFP1_Q6	262	7	7.20E-05	4.22E-03	DNAJC11, PROX1, RBMS1, FOXP1, ARNTL, BDNF, MAP2K7
TCANNTGAY_SREBP1_01	463	9	7.77E-05	4.33E-03	GIGYF2, FOXP1, STAG1, AS3MT, BDNF, SPPL3, MAP2K5, NOL4, PEPD
CTGCAGY_UNKNOWN	750	11	1.51E-04	8.03E-03	ELAVL4, CNTN2, BCL11A, FOXP1, STAG1, TRPS1, BDNF, NRXN3, GGNBP2, NOL4, MAP2K7
AHRARNT_01	137	5	1.94E-04	9.84E-03	ELAVL4, BCL11A, RBM6, BDNF, FBXL20
RTAAACA_FREAC2_01	915	12	2.11E-04	1.02E-02	PROX1, BCL11A, KCNJ13, IMMP2L, TRPS1, SLC30A8, TLE1, ARNTL, BDNF, CADM1, HNF1A, TCF4
SYATTGTG_UNKNOWN	225	6	2.43E-04	1.13E-02	BCL11A, MON1A, FOXP1, APOPT1, SUGP1, EP300
YCATTAA_UNKNOWN	548	9	2.74E-04	1.16E-02	ELAVL4, RBMS1, MSH5, MSRA, TRPS1, HNF1A, FURIN, FES, NOL4

PR_01	148	5	2.78E-04	1.16E-02	ELAVL4, PROX1, SLC12A8, BDNF, FES
CEBP_Q2	231	6	2.80E-04	1.16E-02	KCNJ13, FOXP1, STAG1, BDNF, CADM1, TCF4
OCT1_04	233	6	2.93E-04	1.17E-02	AGBL4, ELAVL4, KCNJ13, FOXP1, CADM1, FES
HNF1_01	241	6	3.51E-04	1.22E-02	BCL11A, RBMS1, NLGN1, MSH5, BDNF, HNF1A
MYB_Q3	243	6	3.67E-04	1.22E-02	FOXP1, BDNF, CADM1, HMGA2, MDGA2, MDGA2, NRXN3
E2F1_Q3_01	243	6	3.67E-04	1.22E-02	DNAJC11, ELAVL4, STAG1, TRPS1, ARNTL, CADM1
RSRFC4_01	243	6	3.67E-04	1.22E-02	ELAVL4, FOXP1, TRPS1, NRXN3, MAP2K5, RALY
ZNF596_TARGET_GENES	574	9	3.85E-04	1.22E-02	PROX1, BCL11A, BTN2A1, TRPS1, HMGA2, TMEM219, TCF4, ZC3H4, EP300
HNF4_01_B	246	6	3.92E-04	1.22E-02	PROX1, BCL11A, RBMS1, TLE1, HNF1A, PEPD
AREB6_02	246	6	3.92E-04	1.22E-02	PROX1, STAG1, FSHB, SPPL3, NRXN3, NOL4
HNF1_Q6	250	6	4.27E-04	1.22E-02	BCL11A, RBMS1, MSH5, HNF1A, NRXN3, MAP2K5
IPF1_Q4	250	6	4.27E-04	1.22E-02	ELAVL4, BCL11A, TLE1, BDNF, HNF1A, FES
IRF1_Q6	251	6	4.36E-04	1.22E-02	ELAVL4, PROX1, STAG1, MDGA2, MDGA2, NRXN3, FES
SRY_02	251	6	4.36E-04	1.22E-02	ELAVL4, SLC12A8, CADM1, HNF1A, ZNF664, TCF4
P300_01	251	6	4.36E-04	1.22E-02	ELAVL4, BCL11A, FOXP1, FSHB, FURIN, GGNBP2
ALPHACP1_01	253	6	4.55E-04	1.22E-02	ELAVL4, CNTN2, TLE1, HSD17B12, CADM1, GGNBP2
TST1_01	256	6	4.84E-04	1.22E-02	ELAVL4, STAG1, HMGA2, NRXN3, MAP2K5, FES
TGTTTGY_HNF3_Q6	726	10	5.02E-04	1.22E-02	ELAVL4, BCL11A, KCNJ13, FOXP1, TRPS1, CADM1, NRXN3, TCF4, MAP2K7, EP300
CATTGTYY_SOX9_B1	361	7	5.09E-04	1.22E-02	BCL11A, RBMS1, SLC39A8, TLE1, ZNF664, FBXL20, TCF4
NCX_01	169	5	5.11E-04	1.22E-02	ELAVL4, IMMP2L, BDNF, CADM1, HNF1A
OCT1_B	259	6	5.14E-04	1.22E-02	FOXP1, STAG1, BDNF, CADM1, NRXN3, FES
HNF4_Q6	259	6	5.14E-04	1.22E-02	PROX1, TRPS1, TLE1, HNF1A, GGNBP2, RALY
MSX1_01	171	5	5.39E-04	1.24E-02	FOXP1, ARNTL, BDNF, NOL4, MAP2K7
YNGTTNNNATT_UNKNOWN	365	7	5.44E-04	1.24E-02	MON1A, BDNF, HNF1A, NRXN3, MAP2K5, FES, MAP2K7
STAT6_01	264	6	5.69E-04	1.27E-02	BCL11A, TMEM110, TRPS1, CADM1, TCF4, EP300
CRX_Q4	266	6	5.92E-04	1.29E-02	BCL11A, NLGN1, MSH5, HNF1A, NRXN3, PNPLA3
LYF1_01	267	6	6.03E-04	1.29E-02	GIGYF2, SLC39A8, BDNF, HNF1A, FURIN, TCF4
CEBP_01	272	6	6.65E-04	1.40E-02	RBMS1, GIGYF2, FOXP1, CADM1, MDGA2, MDGA2, RALY
SMTTTTGT_UNKNOWN	396	7	8.78E-04	1.80E-02	ELAVL4, PROX1, BCL11A, FOXP1, SLC12A8, CADM1, TCF4
MYAATNNNNNNNGGC_UNKNOWN	110	4	8.86E-04	1.80E-02	TLE1, BDNF, NRXN3, FBXL20
GATAAGR_GATA_C	292	6	9.61E-04	1.91E-02	ELAVL4, KCNJ13, TRPS1, HNF1A, NRXN3, GIP
GR_01	196	5	9.96E-04	1.95E-02	ELAVL4, PROX1, SLC12A8, BDNF, FES
GCCATNTTG_YY1_Q6	407	7	1.03E-03	1.98E-02	ELAVL4, RBMS1, GIGYF2, FOXP1, MAP2K5, TCF4, EP300
RP58_01	201	5	1.11E-03	2.10E-02	DNAJC11, CNTN2, RBMS1, GIGYF2, FOXP1
ZNF22_TARGET_GENES	961	11	1.21E-03	2.24E-02	DNAJC11, BCL11A, FOXP1, SLC2A2, NLGN1, FEZF1, SPPL3, ZNF664, MAP2K5, TCF4, WDR7
TGACAGNY_MEIS1_01	821	10	1.28E-03	2.34E-02	ELAVL4, CNTN2, RBMS1, MON1A, PCCB, TLE1, MAP2K5, NOL4, SUGP1, EP300
GATTGGY_NFY_Q6_01	1125	12	1.31E-03	2.35E-02	ELAVL4, KCNJ13, TMEM110, STAG1, TLE1, ARNTL, BDNF, HSD17B12, CADM1, HMGA2, GGNBP2, NOL4
PAX3_TARGET_GENES	1449	14	1.33E-03	2.35E-02	RBMS1, ITIH4, MUSTN1, ATXN7, FOXP1, SLC12A8, NLGN1, ZDHHC2, ARHGAP19-SLIT1, ARHGAP19, MDGA2, MDGA2, MAP2K5, GIP, TCF4, WDR7
STAT5A_04	210	5	1.35E-03	2.36E-02	AGBL4, TRPS1, BDNF, NOL4, MAP2K7
RSRFC4_Q2	213	5	1.44E-03	2.45E-02	ELAVL4, TRPS1, NRXN3, MAP2K5, RALY
ZNF549_TARGET_GENES	1138	12	1.45E-03	2.45E-02	AK5, MON1A, ITIH4, NLGN1, TAP2, IMMP2L, HMGA2, NRXN3, APOPT1, MAP2K5, FBXL20, TCF4
ZNF513_TARGET_GENES	695	9	1.49E-03	2.47E-02	BCL11A, RBMS1, ATXN7, FEZF1, MAP2K5, RAI1, TCF4, ZC3H4, EP300
PBXIP1_TARGET_GENES	221	5	1.69E-03	2.78E-02	FEZF1, HMGA2, NRXN3, TMEM219, MAU2
TGCCAAR_NF1_Q6	710	9	1.72E-03	2.78E-02	PROX1, TRPS1, BDNF, HMGA2, HNF1A, MDGA2, MDGA2, GIP, TCF4, RALY

MZF1_02	229	5	1.98E-03	3.12E-02	ELAVL4, FOXP1, HNF1A, GGNBP2, NOL4
NKX3A_01	230	5	2.01E-03	3.12E-02	PROX1, BCL11A, GIGYF2, CADM1, NRXN3
NFKB_Q6_01	230	5	2.01E-03	3.12E-02	RBMS1, BDNF, CADM1, HNF1A, EP300
AACTTT_UNKNOWN	1873	16	2.12E-03	3.24E-02	ELAVL4, PROX1, BCL11A, RBMS1, GIGYF2, KCNJ13, STAG1, NLGN1, TRPS1, TLE1, BDNF, HSD17B12, CADM1, HNF1A, NOL4, TCF4
YY1_Q6	234	5	2.17E-03	3.25E-02	ELAVL4, BCL11A, GIGYF2, TCF4, EP300
CGTSACG_PAX3_B	141	4	2.21E-03	3.25E-02	BCL11A, GIGYF2, NOL4, PEPD
BRN2_01	235	5	2.21E-03	3.25E-02	BCL11A, DNAH8, CADM1, NRXN3, TCF4
FOXJ2_02	237	5	2.29E-03	3.32E-02	BCL11A, FOXP1, ARNTL, CADM1, TCF4
HMGIY_Q6	242	5	2.51E-03	3.56E-02	BCL11A, MON1A, FOXP1, CADM1, HNF1A
CTAWWWATA_RSRFC4_Q2	354	6	2.54E-03	3.56E-02	ELAVL4, FOXP1, SLC12A8, FAM47E-STBD1, MAP2K5, RALY
RORA1_01	243	5	2.56E-03	3.56E-02	ELAVL4, CNTN2, RBMS1, ARNTL, NOL4
GCCNNNWTAAR_UNKNOWN	148	4	2.63E-03	3.63E-02	ELAVL4, CADM1, HMGA2, FES
RTTTNNNYTGGM_UNKNOWN	149	4	2.70E-03	3.67E-02	BCL11A, FOXP1, TRPS1, HNF1A
NFY_Q6	247	5	2.74E-03	3.68E-02	ELAVL4, KCNJ13, HMGA2, GGNBP2, NOL4
FREAC4_01	151	4	2.83E-03	3.68E-02	KCNJ13, TRPS1, CADM1, HNF1A
NFKB_Q6	249	5	2.84E-03	3.68E-02	BDNF, CADM1, HNF1A, RAI1, EP300
ZIC3_01	249	5	2.84E-03	3.68E-02	BCL11A, RBMS1, FOXP1, FES, FBXL20
P53_02	251	5	2.94E-03	3.69E-02	BCL11A, AKAP6, NRXN3, MAP2K5, RAI1
NFY_Q6_01	251	5	2.94E-03	3.69E-02	ELAVL4, KCNJ13, TMEM110, GGNBP2, NOL4
MYOGENIN_Q6	252	5	2.99E-03	3.69E-02	CADM1, HMGA2, GGNBP2, FBXL20, NOL4
TCF11_01	254	5	3.09E-03	3.69E-02	AK5, CNTN2, BDNF, CADM1, MAP2K5
ZNF140_TARGET_GENES	632	8	3.14E-03	3.69E-02	TAP2, IMMP2L, MPHOSPH9, AKAP6, FES, GGNBP2, NOL4, WDR7
AAAYWAACM_HFH4_01	255	5	3.14E-03	3.69E-02	SLC39A8, BDNF, CADM1, NRXN3, TCF4
SOX10_TARGET_GENES	255	5	3.14E-03	3.69E-02	FOXP1, SLC12A8, TRPS1, ZNF664, SLCO4A1
OCT1_06	255	5	3.14E-03	3.69E-02	FOXP1, BDNF, NOL4, PEPD, EP300
AP2_Q6	255	5	3.14E-03	3.69E-02	ZDHHC2, CADM1, SPPL3, GGNBP2, EP300
TTAYRTAA_E4BP4_01	257	5	3.25E-03	3.78E-02	ELAVL4, BCL11A, KCNJ13, CADM1, EP300
MYB_Q5_01	258	5	3.30E-03	3.80E-02	ELAVL4, BDNF, CADM1, HMGA2, MDGA2, MDGA2
GATA_C	259	5	3.36E-03	3.82E-02	ELAVL4, BDNF, HNF1A, NRXN3, MAP2K5
TATAAA_TATA_01	1264	12	3.45E-03	3.87E-02	RBMS1, KCNJ13, MON1A, FOXP1, BLK, BDNF, FSHB, CADM1, HNF1A, MAP2K5, FURIN, TCF4
FREAC2_01	261	5	3.47E-03	3.87E-02	PROX1, BCL11A, BDNF, CADM1, HNF1A
HOXA4_Q2	262	5	3.53E-03	3.89E-02	MON1A, BDNF, CADM1, HMGA2, HNF1A
CEBP_Q2_01	265	5	3.70E-03	4.05E-02	KCNJ13, FOXP1, BDNF, CADM1, NRXN3
EVI1_05	168	4	4.15E-03	4.49E-02	ELAVL4, KCNJ13, MAP2K5, RAI1
ER_Q6_01	273	5	4.20E-03	4.50E-02	RBMS1, TRPS1, FSHB, CADM1, MAP2K7
HAND1E47_01	277	5	4.46E-03	4.74E-02	ELAVL4, BCL11A, SLC12A8, STAG1, EP300
MEF2_04	26	2	4.53E-03	4.75E-02	ELAVL4, RALY
CREBP1_01	173	4	4.60E-03	4.75E-02	ELAVL4, BCL11A, CADM1, EP300
GATA1_05	279	5	4.60E-03	4.75E-02	ELAVL4, TRPS1, BDNF, RAI1, MAP2K7
GO BP (MsigDB c5)					
GeneSet	**N**	**n**	**P-value**	**Adjust. P**	**genes**
GOBP_REGULATION_OF_HORMONE_LEVELS	522	11	5.59E-06	1.98E-02	SLC2A2, BLK, SLC30A8, ARNTL, KCNJ11, FSHB, HSD17B12, HNF1A, FURIN, GIP, SLCO4A1
GOBP_PEPTIDE_TRANSPORT	254	8	6.54E-06	1.98E-02	SLC2A2, TAP2, BLK, SLC30A8, ARNTL, KCNJ11, HNF1A, GIP

GOBP_REGULATION_OF_NEURON_DIFFERENTIATION	188	7	8.60E-06	1.98E-02	CNTN2, PROX1, BCL11A, NLGN1, FEZF1, BDNF, TCF4
GOBP_INSULIN_SECRETION	193	7	1.02E-05	1.98E-02	SLC2A2, BLK, SLC30A8, ARNTL, KCNJ11, HNF1A, GIP
GOBP_SIGNAL_RELEASE	470	10	1.41E-05	2.01E-02	SLC2A2, NLGN1, CADPS2, BLK, SLC30A8, ARNTL, KCNJ11, HNF1A, FBXL20, GIP
GOBP_SIGNAL_PEPTIDE_PROCESSING	13	3	1.56E-05	2.01E-02	IMMP2L, SPPL3, FURIN
GOBP_POSITIVE_REGULATION_OF_CELL_COMMUNICATION	1713	19	2.58E-05	2.58E-02	CNTN2, PROX1, TMEM110, FOXP1, SLC2A2, NLGN1, BLK, ZDHHC2, SLC30A8, ARNTL, BDNF, SPPL3, AKAP6, MAP2K5, FURIN, GIP, CSNK1G2, MAP2K7, EP300
GOBP_HORMONE_TRANSPORT	313	8	2.95E-05	2.58E-02	SLC2A2, BLK, SLC30A8, ARNTL, KCNJ11, HNF1A, GIP, SLCO4A1
GOBP_EXPORT_FROM_CELL	870	13	3.00E-05	2.58E-02	MON1A, SLC2A2, NLGN1, CADPS2, BLK, SLC30A8, ARNTL, KCNJ11, HNF1A, FURIN, FES, FBXL20, GIP
GOBP_REGULATION_OF_INSULIN_SECRETION	160	6	3.74E-05	2.90E-02	SLC2A2, BLK, SLC30A8, ARNTL, KCNJ11, GIP
GOBP_AMIDE_TRANSPORT	338	8	5.08E-05	3.37E-02	SLC2A2, TAP2, BLK, SLC30A8, ARNTL, KCNJ11, HNF1A, GIP
GOBP_SECRETION	927	13	5.75E-05	3.37E-02	MON1A, SLC2A2, NLGN1, CADPS2, BLK, SLC30A8, ARNTL, KCNJ11, HNF1A, FURIN, FES, FBXL20, GIP
GOBP_CENTRAL_NERVOUS_SYSTEM_NEURON_DIFFERENTIATION	173	6	5.79E-05	3.37E-02	AGBL4, ELAVL4, CNTN2, PROX1, GIGYF2, MDGA2, MDGA2
GOBP_POSITIVE_REGULATION_OF_CALCINEURIN_MEDIATED_SIGNALING	20	3	6.09E-05	3.37E-02	TMEM110, SPPL3, AKAP6
GOBP_PROTEIN_LOCALIZATION_TO_EXTRACELLULAR_REGION	367	8	9.01E-05	4.66E-02	MON1A, SLC2A2, BLK, SLC30A8, ARNTL, KCNJ11, HNF1A, GIP
GWAS enrichment					
GeneSet	**N**	**n**	**P-value**	**Adjust. P**	**genes**
Type 2 diabetes	563	34	1.15E-31	5.10E-28	PROX1, BCL11A, RBMS1, SPHKAP, RBM6, ATXN7, SLC12A8, SLC2A2, SLCO6A1, MSRA, XKR6, SLC30A8, TLE1, ARHGAP19-SLIT1, ARHGAP19, KCNJ11, BDNF, HSD17B12, HMGA2, HNF1A, MPHOSPH9, AKAP6, NRXN3, MAP2K5, RAI1, GIP, TCF4, WDR7, MAP2K7, SUGP1, PEPD, ZC3H4, RALY, EP300, PNPLA3
Schizophrenia	673	35	2.08E-30	4.61E-27	ELAVL4, CNTN2, SPHKAP, GIGYF2, KCNJ13, TRANK1, ITIH4, MUSTN1, THOC7, ATXN7, FOXP1, PCCB, STAG1, SLC39A8, SLCO6A1, PSORS1C1, IMMP2L, MSRA, ZDHHC2, TSNARE1, TLE1, AS3MT, MPHOSPH9, AKAP6, NRXN3, FURIN, FES, TMEM219, RAI1, NOL4, TCF4, SUGP1, MAU2, PEPD, EP300
Cognitive ability	150	15	2.15E-17	1.90E-14	BCL11A, SPHKAP, GIGYF2, RBM6, THOC7, FOXP1, STAG1, SLC39A8, MSRA, TSNARE1, AS3MT, MPHOSPH9, AKAP6, FURIN, TCF4
Body size at age 10	255	15	6.05E-14	3.82E-11	BCL11A, RBMS1, IMMP2L, MSRA, XKR6, TSNARE1, GRID1, BDNF, CADM1, HMGA2, AKAP6, NRXN3, MAP2K5, GGNBP2, ZC3H4
Systolic blood pressure	704	21	2.45E-13	1.35E-10	AGBL4, RBMS1, FOXP1, PCCB, SLC39A8, MSH5, MSH5-SAPCD1, SAPCD1, MSRA, BLK, TSNARE1, AS3MT, KCNJ11, BDNF, CADM1, HMGA2, FURIN, FES, GIP, TCF4, WDR7, PEPD
Brain morphology (MOSTest)	1049	24	1.10E-12	4.87E-10	CNTN2, BCL11A, MON1A, RBM6, ITIH4, THOC7, ATXN7, FOXP1, SLC12A8, SLC39A8, BTN2A1, MSRA, XKR6, BLK, TRPS1, AS3MT, BDNF, MPHOSPH9, FURIN, FES, TMEM219, TCF4, SUGP1, MAU2
HDL cholesterol levels	518	18	1.24E-12	4.99E-10	CNTN2, RBMS1, TRANK1, RBM6, SLC39A8, MSRA, TRPS1, GRID1, BDNF, HSD17B12, HNF1A, ZNF664, AKAP6, MAP2K7, PEPD, ZC3H4, RALY, PNPLA3
Body mass index	1153	24	8.08E-12	2.75E-09	AGBL4, ELAVL4, AK5, SHISA4, MON1A, RBM6, ITIH4, STAG1, SLC39A8, MSRA, SLC30A8, GRID1, KCNJ11, BDNF, FSHB, HSD17B12, CADM1, AKAP6, NRXN3, MAP2K5, RAI1, GGNBP2, CSNK1G2, ZC3H4
Intelligence (MTAG)	228	11	1.36E-09	3.34E-07	BCL11A, RBMS1, MON1A, RBM6, FOXP1, SLC39A8, BTN2A1, TSNARE1, MPHOSPH9, AKAP6, GGNBP2
Waist-hip index	565	15	4.47E-09	9.88E-07	CNTN2, ITIH4, NME9, SLC39A8, SLCO6A1, BTN2A1, PSORS1C1, PSORS1C2, MSRA, XKR6, BLK, MPHOSPH9, ZNF664, RAI1, PEPD
Coronary artery disease	522	14	1.35E-08	2.52E-06	GIGYF2, KCNJ13, PCCB, STAG1, AS3MT, BDNF, HSD17B12, HNF1A, FURIN, FES, RAI1, FBXL20, GIP, CSNK1G2
Triglyceride levels	548	14	2.48E-08	4.40E-06	RBMS1, PCCB, SLC2A2, NLGN1, SLC39A8, MSRA, TRPS1, SLC30A8, BDNF, HMGA2, ZNF664, MAP2K5, SUGP1, PEPD

Waist circumference	167	8	2.83E-07	3.80E-05	AGBL4, CADPS2, BDNF, HSD17B12, CADM1, NRXN3, MAP2K5, ZC3H4
Broad depression or SCZ	16	4	3.69E-07	4.53E-05	AGBL4, PSORS1C2, TSNARE1, TCF4
Apolipoprotein A1 levels	367	10	1.60E-06	1.54E-04	TRANK1, RBM6, FOXP1, SLC39A8, TRPS1, BDNF, FURIN, PEPD, ZC3H4, PNPLA3
Cortical surface area (MOSTest)	401	10	3.52E-06	3.18E-04	CNTN2, BCL11A, MON1A, RBM6, ITIH4, THOC7, ATXN7, FOXP1, SLC12A8, SLC39A8
Metabolic syndrome	111	6	4.65E-06	4.12E-04	RBM6, PCCB, SLC39A8, TRPS1, BDNF, ZC3H4
Subcortical volume (MOSTest)	522	11	5.59E-06	4.85E-04	CNTN2, BCL11A, MON1A, RBM6, ITIH4, THOC7, ATXN7, FOXP1, SLC12A8, SLC39A8, BTN2A1
Fasting insulin	69	5	7.19E-06	5.89E-04	IMMP2L, BLK, HMGA2, MAP2K7, PEPD
CAD/MI	86	5	2.11E-05	1.46E-03	GIGYF2, KCNJ13, FURIN, FES, RAI1
Obesity	49	4	3.89E-05	2.46E-03	PROX1, BDNF, NRXN3, MAP2K5
Childhood body mass index	53	4	5.32E-05	3.22E-03	SLC39A8, BDNF, CADM1, HMGA2
Fasting blood glucose	45	3	7.07E-04	2.35E-02	PROX1, SLC2A2, SLC30A8
Alanine aminotransferase levels	287	6	8.79E-04	2.82E-02	RBM6, SLC12A8, PCCB, SLC2A2, PEPD, PNPLA3
Adiponectin levels	50	3	9.63E-04	2.92E-02	ZNF664, NRXN3, PEPD
Apolipoprotein B levels	203	5	1.16E-03	3.48E-02	STAG1, PSORS1C1, TRPS1, HNF1A, SUGP1
Hypertension	119	4	1.19E-03	3.52E-02	SLC39A8, MSRA, FURIN, FES
Myocardial infarction	207	5	1.27E-03	3.72E-02	GIGYF2, STAG1, HNF1A, FURIN, FES
Brain structure	14	2	1.31E-03	3.75E-02	CADPS2, HMGA2
Diastolic blood pressure	555	8	1.40E-03	4.00E-02	ATXN7, SLC39A8, BLK, BDNF, CADM1, FURIN, FES, TCF4
LDL cholesterol levels	220	5	1.66E-03	4.60E-02	STAG1, SLC39A8, TRPS1, HNF1A, SUGP1
Liver enzyme levels	16	2	1.71E-03	4.60E-02	HNF1A, PNPLA3
Severe insulin-deficient T2D	16	2	1.71E-03	4.60E-02	SLC12A8, CSNK1G2
Offspring birth weight	132	4	1.74E-03	4.63E-02	HMGA2, FURIN, FES, RAI1