

# Multi-omics and Mendelian Randomization network analysis of the association between metabolic and cognitive functions in the UK Biobank Database

Paolo Alberto Lorenzini<sup>1,2\*</sup> and Sarah Raye Langley<sup>2\*</sup>

<sup>1</sup>Infectious Disease Labs (ID Labs), Agency for Science, Technology and Research (A\*STAR), 8a Biomedical Grove, Immunos, Singapore 138648, Republic of Singapore.

<sup>2</sup>Lee Kong Chian School of Medicine, Nanyang Technological University, 11 Mandalay Rd, Singapore 308232, Republic of Singapore

\* **Correspondence:**

Paolo Alberto Lorenzini [paoloal@IDLabs.a-star.edu.sg](mailto:paoloal@IDLabs.a-star.edu.sg)

Sarah Raye Langley [sarah.langley@ntu.edu.sg](mailto:sarah.langley@ntu.edu.sg)

## Abstract

In obese individuals, low levels of leptin, an hormone regulating energy expenditure, have been claimed to be associated with neurodegeneration during ageing. Since leptin signalling system is less potent in the elderly, its protective anti-inflammatory effect in the brain is inhibited and people overweight in middle-age are at higher risk of developing dementia later in life. However, the causal association between obesity and cognitive decline has been inconsistent so far. We incorporated several metabolic and cognitive phenotypes from primary-care health records from European individuals in the UK Biobank to study comprehensively the genetic architecture of metabolic function and cognition. By GWAS analysis, we identified genetic loci in various traits associated with metabolic function and cognitive performance and investigated their genetic association by Mendelian Randomization. Importantly, we identified putatively causality amongst metabolism related traits and that higher adiposity is causally associated with a worse cognitive performance. We conducted a transcriptome-wide association study to identify regulatory effects of the susceptibility variants to pinpoint genes associated with gene expression changes. We integrated the GWAS results with transcriptome and single cell public data to identify tissues and cell types associated with the pathogenesis of metabolic dysfunction and cognition. Overall, this *multi-omics* approach have proven to be invaluable for demonstrating the causal association between adiposity and cognition and for the discovery of causal variants, genes, tissues and cell types underlying these phenotypes.

**NOTE:** This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

## Introduction:

Increases in life expectancy coupled with decreases in fertility rates worldwide are leading towards an aging society (WHO, 2020). As a person grows older, the cognitive functions can deteriorate leading to difficulties with a person's processing, remembering and concentrating<sup>1</sup>. A certain baseline level of cognitive decline is normal during aging. However, when the cognitive impairment has become severe enough to compromise the ability to perform activities of daily life, dementia is diagnosed<sup>2,3</sup>. As the global population ages, more people will be affected by cognitive decline and more likely at risk to develop dementia<sup>4,5</sup>. According to recent reports, there are about 55 million people currently living with dementia and it is forecasted that cases will triple by 2050<sup>6</sup>. This projected trend underscores the need for public health planning efforts and public policies to promote healthy aging and reduce the burden that an aging population is posing on society<sup>5-7</sup>.

In addition to developing therapeutics, in order to develop pre-emptive strategies, it is very important to assess which are the modifiable risk factors that could potentially exacerbate, or reduce, one's risk of developing dementia in later life<sup>8</sup>. Type 2 diabetes (T2D), hypertension, and obesity (i.e. metabolic traits) as well as depression, smoking and low educational attainment have been identified as exacerbating factors of dementia<sup>9-11</sup>. In particular, epidemiological studies have demonstrated that T2D can increase the risk of cognitive impairment and that patients with T2D are more likely to be diagnosed with dementia<sup>12-14</sup>. Furthermore, obesity, T2D and dementia share several biological features such as chronic inflammation, severe oxidative stress and impairment in insulin and energy metabolism<sup>13,15-17</sup>.

Given this relationship among obesity, T2D, and dementia, understanding the mechanistic links among them is essential to developing effective strategies for prevention and treatment of these

diseases<sup>18,19</sup>. Despite this strong observational evidence that T2D is associated with dementia, causality cannot be reliably established since these observational studies are prone to confounding effects and reverse causation. Mendelian Randomization (MR) is an alternative approach to randomized clinical trials (RCTs) that can be utilised to determine causality by circumventing confounding effects and reverse causation<sup>20,21</sup>. MR is an epidemiological methodology which uses genetic variants, as instrumental variables (IVs), in order to determine causal relationship between a modifiable risk factor and an outcome<sup>20,22,23</sup>. Because of the random allocation of alleles at conception, genetic variants associated with a modifiable exposure are randomly distributed in relation to potential confounders. Furthermore, considering that the genotype inherited at birth stays constant throughout life, the possibility of reverse causation can be ruled out<sup>20,22,23</sup>. Importantly, in order to estimate direct causal effects, multivariable MR techniques can also be used to take into account of indirect effects due to “pleiotropy”, a widespread phenomenon in the genome by which a single SNP influences multiple traits<sup>24</sup>.

Previous MR studies have assessed the causality between T2D, insulin sensitivity and fasting glucose on the risk of neurological diseases with contradictory results. Several studies did not find significant associations between metabolic factors and dementias<sup>25-27</sup>, while others identified significant associations between glycaemic traits and increased risk of dementia and/or cognitive decline<sup>28-31</sup>. We conducted a GWAS analysis of traits related to metabolic function and cognitive performance in European individuals in the UK Biobank database and utilised the results to determine causality (Figure 1A-B).

While these genome-wide associations utilised in MR analyses provide a useful framework for identifying putative susceptibility loci, they rarely identify causal genes, predominantly due to the complicated Linkage Disequilibrium (LD) structure of the genome, as well as to the fact

that genetic variants can affect phenotype via distant regulation of gene expression<sup>32</sup>. In order to prioritise functionally relevant genes from Genome-Wide Association Study (GWAS) loci for follow-up studies, we integrated the GWAS results with expression Quantitative Trait Loci (eQTLs) from human skeletal muscle tissue (Figure 1A), which allowed us to discover genetic variants associated with transcriptional changes<sup>33</sup>. In addition, we utilised this multi-omics analysis to assess for pleiotropy. We performed functional mapping and annotation to further explore genetic variants and genomic loci in the pathogenesis of metabolic dysfunction and cognitive decline.

Furthermore, the underlying cell types mediating predisposition to metabolic dysfunction remain largely obscure, despite identifying such cell types is a very important step towards a better understanding of mechanisms causing the disease<sup>34</sup>.

For example, it has been recently reported that a dysregulation in specific adaptive immune cell types triggers the onset, development and progression of obesity and T2D<sup>35</sup>. As such, we integrated the GWAS results with public gene expression data in order to identify genes and cell types underlying susceptibility to obesity and cognitive health (Figure 1A). Overall, our multi-omics and MR analysis (Figure 1A) would provide important leads to better understand the pathogenesis of metabolic dysfunction and cognitive decline and to develop novel therapeutic targets for precision medicine.

## **Results:**

### **Establishing causal association between obesity and cognitive functions**

We used genetic and phenotypic data from the UK Biobank Database (European cohort)<sup>36</sup> to estimate the causal association between obesity and cognitive performance (Figure 1A-B). We

selected several metabolic and cognitive traits (Table S1) to cover as many as possible phenotypes. To have a comprehensive assessment of cognitive health status, we selected cognitive traits covering all the four major cognitive domains (episodic and working memory, executive function, processing speed and attention)<sup>37</sup> (Table S1). We obtained a final list of 22 traits (17 metabolic and 5 cognitive) (Table S1). For each trait, we performed a GWAS analysis, using both the entire population (sex-combined) or as individual sexes. Functional analysis of the sex-combined GWAS results confirmed that the significant hits ( $p$ -value  $< 5e-08$ ) are involved in lipid metabolism and homeostasis (Supp. Fig. S1A). With the sex-split summary statistics, we built a Mendelian Randomization network (Figure 1B) by utilising an algorithm which allows to estimate direct causal effects not mediated by other measured factors<sup>38</sup>. Several associations were in line with previous epidemiological observations, with a few exceptions. Most of the direct causal associations (36/39) excluding some involving cholesterol and hip circumference showed an expected epidemiological association (Table S2). For example, an increase in waist circumference was causally associated with an increase in body fat percentage, glycated haemoglobin, hip circumference, triglycerides and a decrease in HDL (Figure 1B). Interestingly, an increase in body fat percentage was causally associated with a lower performance in the FIQ test (Figure 1B), assessing verbal and numerical reasoning (executive function domain). An increase in HDL was causally associated with a decrease in LDL and a higher number of digits remembered correctly in the working memory test and an higher time to correctly identifying matches in the reaction test assessing processing speed, was causally associated with an increase in waist circumference and weight (Figure 1B).

Collectively, these results show a strong causal association amongst metabolic traits related to adiposity and that metabolic function, in particular adiposity, is causally associated with a worse performance in certain cognitive test (Figure 1B, Table S2).

## Identifying pleiotropically causal genes in obesity and cognitive functions

In order to identify and prioritize functionally relevant causal genes in obesity or cognitive functions, we integrated the genetic data with gene expression data (cis-eQTLs) and conducted a pleiotropic gene-trait association analysis by SMR-HEIDI<sup>39</sup>. In brief, by combining cis-eQTL data from a specific tissue with GWAS summary statistics of a single trait, the SMR test allows to assess whether the effect size of a SNP on that specific trait is mediated by gene expression<sup>39,40</sup>. Importantly, the test allows as well to discriminate if the observed association could be due to a single variant or alternatively two distinct genetic variants in high LD with each other (pleiotropy)<sup>33,39</sup>.

We integrated the sex-combined GWAS summary statistics of the metabolic and cognitive traits with cis-eQTLs data from skeletal human muscle tissue obtained from the GTEx database<sup>41</sup>. We focused on skeletal muscle tissue in particular because obesity-linked insulin resistance is mainly due to fatty acid overload in non-adipose tissues, particularly skeletal muscle and liver<sup>42</sup>. Since 10282 eQTLs probes (1.000.905 SNPs) were included, the genome-wide significance level ( $P_{\text{SMR}}$ ) is defined as  $0.05/10282 = 5.0 \times 10^{-6}$ .

We ran the SMR analysis with default settings (see methods). For the metabolic traits, we identified 140 significant probes tagging 82 unique genes which passed the  $P_{\text{SMR}}$  threshold, whereas for the cognitive traits only 4 genes/probes showed significant association (Tables S3-4).

Since significant SMR results could also reflect linkage (i.e. the top associated cis-eQTL being in LD with two distinct causal variants, one affecting gene expression and the other affecting trait variation), for the significant probes with  $P_{\text{SMR}} < 5.0 \times 10^{-6}$ , we conducted a HEIDI test. If a probe passes the HEIDI test ( $P_{\text{HEIDI}} \geq 0.05$ ), it signifies that there is a single causal variant

affecting trait and gene expression (pleiotropy or causality), which is of more important biological interest and should be prioritized for follow-up studies. After application of the HEIDI test ( $P_{\text{HEIDI}} \geq 0.05$ ), we retained 62 probes (60 metabolic, 2 cognitive) tagging to 39 unique genes (37 metabolic, 2 cognitive) (Table S5). Functional analysis of these genes showed enrichment in lipid metabolism (Supp. Fig. S1B) and a Protein-Protein Interaction (PPI) network of these 39 genes showed strong evidence of interaction between some of those proteins, including those involved in lipid metabolism (Supp. Fig. S2A-B).

Interestingly, several of the significant probes which passed the  $P_{\text{SMR}}$  threshold (33/144), tagging 13 unique genes were located on a chromosome 16 region well known be associated in the pathogenesis of human obesity and neural development<sup>43-46</sup> (Figure 2A-B, Table 1, Table S4).

Table 1

Most significant SMR probes in Europeans located on chromosome 16:

Trait	Probe ID	Gene	Chr	Top SNP	Allele Freq	$P_{\text{eQTL}}$	$P_{\text{GWAS}}$	$P_{\text{SMR}}$
BMI	ENSG00000140718	FTO	16	rs1477196	0.620664	2.95E-12	1.71E-39	7.03E-10
CHOL	ENSG00000040199	PHLPP2	16	rs1058750	0.832374	6.47E-44	1.74E-10	6.62E-09
HC	ENSG00000103510	KAT8	16	rs9925964	0.361484	3.03E-23	1.45E-11	2.35E-08
HC	ENSG00000260442	RP11-22P6.3	16	rs8049439	0.405005	4.53E-35	4.11E-10	2.45E-08
HC	ENSG00000197165	SULT1A2	16	rs4788084	0.419946	1.27E-61	9.81E-09	6.01E-08
BFM	ENSG00000158865	SLC5A11	16	rs2287783	0.195463	9.03E-76	1.46E-07	4.28E-07
HC	ENSG00000167397	VKORC1	16	rs2884737	0.258031	2.13E-16	9.17E-10	9.14E-07
Weight	ENSG00000149926	FAM57B	16	rs9928448	0.466783	2.51E-27	5.53E-08	1.20E-06
BMI	ENSG00000103550	KNOP1	16	rs62025019	0.171962	3.17E-94	8.49E-07	1.68E-06
Weight	ENSG00000176476	CCDC101	16	rs480400	0.543293	2.12E-25	8.00E-08	1.83E-06
HDL	ENSG00000213398	LCAT	16	rs1134760	0.163383	1.62E-11	2.09E-11	2.03E-06
HC	ENSG00000251417	RP11-1348G14.4	16	rs8049439	0.405005	3.16E-12	4.11E-10	3.27E-06
HC	ENSG00000260911	RP11-196G11.2	16	rs35468353	0.376089	4.28E-31	4.05E-07	3.44E-06

BMI Body Mass Index, CHOL Cholesterol, HC Hip Circumference, BFM, Body Fat Mass, HDL High Density Lipoproteins, Chr Chromosome number, Allele Freq Allele Frequency,  $P_{\text{eQTL}}$  p-value of the top associated cis-eQTL of the probe,  $P_{\text{GWAS}}$  GWAS p-value of the top cis-eQTL,  $P_{\text{SMR}}$  p-value for gene-trait association from the SMR test,  $P_{\text{HEIDI}}$  p-value from HEIDI test to indicate whether the gene-trait association is due to a single shared genetic variant (the smaller  $P_{\text{HEIDI}}$  the more likely that there are more than one genetic variant).

In addition, for those 144 probes, we detected high SMR signals as well on chromosomes 1,2,6,11,19,20, which suggests an involvement of these chromosomes in the pathogenesis of human metabolic dysfunctions and neurological diseases as well (Table 2, Table S4).

Table 2

Top most significant SMR probes in Europeans located on chromosomes 1,2,6,11,19,20:

Trait	Probe ID	Gene	Chr	Top SNP	Allele Freq	$P_{eQTL}$	$P_{GWAS}$	$P_{SMR}$
LDL	ENSG00000143126	CELSR2	1	rs7528419	0.22256	2.23E-76	1.12E-121	8.55E-48
TRIG	ENSG00000115234	SNX17	2	rs715325	0.392272	7.11E-87	6.55E-26	1.55E-20
TRIG	ENSG00000198522	GPN1	2	rs3749147	0.252445	6.86E-18	1.17E-53	5.37E-14
LDL	ENSG00000110080	ST3GAL4	11	rs11220462	0.128581	2.38E-47	1.15E-16	6.48E-13
LDL	ENSG00000134222	PSRC1	1	rs646776	0.777289	4.83E-14	1.23E-122	7.08E-13
TRIG	ENSG00000149485	FADS1	11	rs174549	0.308738	2.79E-19	5.26E-31	1.30E-12
HDL	ENSG00000172247	C1QTNF4	11	rs17788930	0.348354	1.36E-41	1.82E-16	2.05E-12
HDL	ENSG00000100979	PLTP	20	rs6065904	0.211698	4.45E-22	1.31E-24	2.11E-12
HDL	ENSG00000109919	MTCH2	11	rs4752856	0.35156	1.40E-33	5.77E-17	6.02E-12
LDL	ENSG00000175164	ABO	9	rs8176719	0.346567	1.70E-86	3.15E-13	8.16E-12
LDL	ENSG00000142252	GEMIN7	19	rs3178166	0.494128	3.74E-52	4.21E-14	1.34E-11
HB	ENSG00000179344	HLA-DQB1	6	rs3828790	0.609058	0	1.06E-11	2.05E-11
HC	ENSG00000065060	UHRF1BP1	6	rs9462026	0.327944	4.31E-79	3.64E-12	7.01E-11
HB	ENSG00000232629	HLA-DQB2	6	rs3828790	0.609058	3.19E-117	1.06E-11	7.06E-11
TRIG	ENSG00000134824	FADS2	11	rs174538	0.315157	7.00E-15	2.13E-29	1.53E-10
HDL	ENSG00000134575	ACP2	11	rs2167079	0.298227	1.00E-20	5.44E-17	4.52E-10
Weight	ENSG00000101019	UQCC1	20	rs4911179	0.626121	2.15E-36	9.61E-13	5.33E-10

LDL Low Density Lipoproteins, TRIG Triglycerides, HDL High Density Lipoproteins, HB glycated hemoglobin, HC Hip Circumference, Chr Chromosome number, Allele Freq Allele Frequency,  $P_{eQTL}$  p-value of the top associated cis-eQTL of the probe,  $P_{GWAS}$  GWAS p-value of the top cis-eQTL,  $P_{SMR}$  p-value for gene-trait association from the SMR test,  $P_{HEIDI}$  p-value from HEIDI test to indicate whether the gene-trait association is due to a single shared genetic variant (the smaller  $P_{HEIDI}$  the more likely that there are more than one genetic variant)

Conforming, in part, what was seen from the MR network, the only probes which passed the significant SMR threshold ( $P_{SMR} < 5.0 \times 10^{-6}$ ) for the cognitive traits, were detected for the Reaction Test (processing speed domain) (Table 3, Table S4).



Table 3

SMR probes for the cognitive traits in Europeans which passed the SMR threshold:

Trait	Probe ID	Gene	Chr	Top SNP	Allele Freq	$P_{eQTL}$	$P_{GWAS}$	$P_{SMR}$
RT	ENSG00000260075	NSFP1	17	rs35732828	0.160345	6.58E-62	5.58E-09	3.80E-08
RT	ENSG00000176681	LRRC37A	17	rs35732828	0.160345	2.60E-29	5.58E-09	2.29E-07
RT	ENSG00000073969	NSF	17	rs35732828	0.160345	2.28E-25	5.58E-09	3.67E-07
RT	ENSG00000228696	ARL17B	17	rs35732828	0.160345	1.96E-21	5.58E-09	6.72E-07

RT Reaction Test, Chr Chromosome number, Allele Freq Allele Frequency,  $P_{eQTL}$  p-value of the top associated cis-eQTL of the probe,  $P_{GWAS}$  GWAS p-value of the top cis-eQTL,  $P_{SMR}$  p-value for gene-trait association from the SMR test,  $P_{HEIDI}$  p-value from HEIDI test to indicate whether the gene-trait association is due to a single shared genetic variant (the smaller  $P_{HEIDI}$  the more likely that there are more than one genetic variant)

Functional analysis of the 86 genes shows that the most of the significant hits in GO enrichment analysis are involved in MHC protein complex assembly, peptide assembly, antigen binding, Endoplasmic Reticulum (ER) biology (Figure 3A). KEGG enrichment analysis showed significant association with allograft rejection, Type II diabetes, Graft-versus-host disease, autoimmune thyroid disease (Figure 3A). Reactome pathway analysis showed immunological synapse as top hit (Figure 3A). Wikipathways analysis showed that top hit was Ebola virus infection in host (Figure 3A). Overall, these functional results mainly point towards a prominent role of the immune system in the metabolic dysfunction and neurological diseases (Figure 3A).

## Identification of tissue and cell types underlying susceptibility to obesity and cognitive functions

Previous reports have detected an exclusive enrichment of BMI GWAS variants in brain tissue<sup>47</sup>. In order to replicate this and to extend it to other metabolic and cognitive traits, we performed tissue enrichment analysis of the GWAS summary statistics with a slightly relaxed

threshold ( $p$ -value  $< 5e-05$ ). We observed strong association of BMI in brain tissues, especially in the frontal cortex, in addition to Body Fat Mass and Percentage (Figure 3B).

Cell-type-specific enrichment analysis (CSEA)<sup>48</sup> of the BMI GWAS summary statistics in the Europeans, showed strong association with inhibitory and limbic system neurons (Figure 4A). In addition, we performed a cell type enrichment analysis<sup>48</sup> of the 86 genes which passed the SMR threshold in the Europeans. We identified macrophages, oligodendrocytes and mucus secreting cells as top 3 cell types based on combined  $p$ -values (Figure 4B). For the 62 probes which passed the HEIDI cut-off, we identified gastric chief cells, fasciculata cells and oligodendrocytes as top 3 hits based on combine  $p$ -values (Figure 4C).

## Discussion:

In this study, we carried out a multi-omics and genetic approach to investigate the causality between metabolic-related risk factors potentially linked to the development of dementia later in life (Figure 1). In particular, we used a network based Mendelian randomization analysis<sup>38</sup> to estimate the bidirectional effects of metabolic and cognitive health in Europeans (Figure 1B). In line with previous studies<sup>30,31</sup>, we did not find evidence of causal association between certain glycaemic traits (e.g. glucose, glycated haemoglobin) with cognition (Figure 1B). Instead, we detected direct causal associations with traits which are more related to adiposity and lipid metabolism such as HDL and fat mass, and body shape traits such weight/waist circumference (Figure 1B). This is in agreement with a recent MR study assessing causality between abdominal adiposity and cognition in different human populations<sup>49</sup>. These results point towards a direct causal role of adiposity on cognitive decline as previously suggested<sup>27,50</sup>. In addition, in our MR analysis, we detected metabolic traits negatively influencing cognitive health and cognitive traits negatively influencing obesity (Figure 1B, Table S2). Indeed, there is evidence that metabolic-related cognitive decline is likely not just a consequence of the

downstream effects of excess adipose tissue and insulin resistance, but a cause as well<sup>50</sup> . This results in a “vicious cycle” where lower cognitive abilities perpetuate metabolic dysfunction, which in turn perpetuate cognitive decline through the lifespan<sup>50,51</sup> .

As previously suggested by other reports, our tissue enrichment analysis of the top GWAS hits (Figure 3B) confirms that susceptibility to metabolic dysfunction is distributed across numerous brain areas that receive signals emanating from internal or external stimuli acting in concert to regulate feeding behaviour and energy metabolism<sup>47,52</sup> . In particular, these results confirm that susceptibility genes for common metabolic dysfunction may have an effect on eating addiction and reward behaviours through their high expression regions of the brains related to addition and reward such as the frontal cortex and the anterior cingulate cortex, in contrast to different pattern from monogenic obesity genes that act in the hypothalamus and cause hyperphagia<sup>52</sup> .

In addition, our cell type analysis of GWAS and SMR-HEIDI hits (Figure 4A-C) to investigate the identity of cell types that drive susceptibility to human metabolic dysfunction has shown strong association with different types of neurons as compared to other cell types. These results further suggest that this dysfunction is distributed across multiple, mainly neuronal, cell types across the brain, thus acting on a more broadly distributed set of neuronal circuits across the brain.

We integrated the GWAS summary statistics with eQTLs data of skeletal muscle tissue by SMR and identified many reported and novel functionally relevant genes in metabolic functions and neuronal development in Europeans (Tables S3-S5).

Regarding the top hits in our SMR analysis of the Europeans, the top most significant hit was detected on the CELSR2 gene (Supp. Fig. S3A, Table S4). This gene encodes a cadherin protein with a main function in neuronal development<sup>53</sup>, but recently found to have a role in

lipid homeostasis as well<sup>54</sup>. Furthermore, it has been found associated with cardiovascular diseases in the gene cluster CELSR2–PSRC1–SORT1<sup>55</sup>. Indeed, we detected significant signal on PSRC1 as well (Supp. Fig. S3A, Table S4). Instead, our second top hit was detected on the SNX17 gene (Supp. Fig. S3B, Table S4), recently found associated with cardiovascular diseases<sup>56</sup>. The gene belongs to the sorting nexin (SNX) family which consists of a diverse group of cytoplasmic and membrane-associated proteins orchestrating the process of cargo sorting<sup>57</sup>. Interestingly, it has been previously reported that SNX17 can regulate the trafficking of the ApoER2 receptor, member of the LDL-R family, which participates in neuronal migration during development<sup>58</sup>.

The majority of the significant SMR probes (33/144) (Table S4) were mapped to chromosome 16 (Figure 2A-B, Table S2), known to be associated with both obesity and neurodevelopmental disorders<sup>43,46,59</sup>. In particular, rearrangements on chromosome 16 including a proximal ~200-kb deletion and a distal ~600-kb deletion are seen in patients with the 16p11.2 deletion syndrome<sup>60,61</sup>. People with this disorder have developmental delay and intellectual disability and they are also at increased risk of obesity compared with the general population<sup>45,46,61</sup>. For example, the distal deletion carries the SH2B1 gene, a mediator of energy homeostasis, which is involved in leptin and insulin signalling<sup>43,59</sup>. We detected SMR signals with high significance and with suggested significance on several genes included in these deleted regions such as LAT, INO80E, FAM57B, TBX6, YPEL3, MAPK3, RP11-22P6.3, RP11-196G11.2 and BOLA2 (Figure 2A, Tables S3-S4). For example, confirming a connection between the immune system and metabolic function, the LAT (linker for activation of T cells) adaptor molecule participates in AKT activation and plays an important role in the regulation of lymphocyte maturation, activation and differentiation<sup>62</sup>. Instead, FAM57B is a gene which regulates adipogenesis by modulating ceramide synthesis and its haploinsufficiency

contributes to changes in neuronal activity in the brain through its crucial role in lipid metabolism<sup>63</sup>.

We also detected significant signals on reported and novel genes on the regions flanking those deletions. For example, a well-studied gene is the Lysine Acetyltransferase KAT8 (Figure 2A). KAT8 is an important gene reported to be involved in both metabolism, being a critical regulator of central carbon metabolism, and neuron development as well<sup>64,65</sup>. Some of the novel significant associations are the HSD3B7 gene, which encodes an enzyme involved in the initial stages of the synthesis of bile acids from cholesterol, the Warfarin target gene VKORC1, involved in vitamin K metabolism and the two sulfotransferase SULT1A1 and SULT1A2, involved in resveratrol metabolism in adipocytes (Fig 2A). We are also detected signals on 4 long non coding RNAs (lncRNA), RP11-1348G14.4, RP11-22P6.3, RP11-196G11.2, RP11-347C12.2 (Figure 2A-B) supporting recent evidence recognizing them as contributing intermediates in obesity and inflammation<sup>66</sup>.

In addition, the top most significant SMR signal on chromosome 16 was detected on the well-known Fat Mass and Obesity-Associated Protein (FTO) gene which has been linked to obesity and altered connectivity of the dopaminergic neurocircuitry, being highly expressed in the hypothalamus<sup>67</sup> (Figure 2B, Table S4). For this gene, we detected a specific association with BMI confirming what previously reported<sup>47</sup>. Previous studies reported that FTO variants are influencing gene expression levels of downstream genes in the brain such as IRX3<sup>68,69</sup>, a gene that plays a role in an early step of neural development, rather than FTO gene expression levels. In our analysis, we detected a very strong co-localization of GWAS signal with gene expression in the skeletal muscle (Figure 2B), thus pointing to a different mechanism of FTO gene regulation between skeletal muscle tissue and the brain. The second top most significant SMR signal on chromosome 16 was detected on the PHLPP2 gene, a phosphatase shown to be

increased during obesity leading to lipid accumulation and glucose dysfunction<sup>70</sup>. In line with this, we detected a specific gene-trait association with LDL and cholesterol (Table S4).

Similarly to chromosome 16, microdeletions on chromosome 11 have been associated with obesity and developmental delay as well<sup>71</sup>. Indeed, we detected many top hits on chromosome 11 as well (Table 2, Table S4). For example, in line with a previous study showing association of the ST3GAL4 gene with plasma concentration levels of triglycerides and LDL in Europeans<sup>72</sup>, we detected highly significant SMR association of this gene with LDL (Supp. Fig. S3D, Table S4). Whereas, confirming the evidence that altered activities in polyunsaturated fatty acids (PUFAs) metabolism are seen in obesity and a number of chronic diseases<sup>73</sup>, we detected highly significant SMR signals on the FADS1 and FADS2 genes (Supp. Fig. S3C, Table S4), encoding key enzymes in PUFA metabolism.

Regarding the cognitive traits, we detected significant gene-trait associations with the Reaction Test (Table S5) on a gene cluster on chromosome 17 which includes the N-ethylmaleimide-sensitive factor (NSF) gene, which is highly expressed in the amygdala and it has been suggested to be linked with molecular mechanisms related with memory formation<sup>74</sup> and two genes (ARL17B, LRRC37A) involved in the pathogenesis of the Koolen-De Vries Syndrome<sup>75</sup>, a developmental delay disorder in children characterised by intellectual disability. Interestingly, the Koolen-De Vries Syndrome is characterised by 17q21.31 microdeletions, the same region where we detected high SMR signal, and by truncating variants of the KAT8 regulatory NSL complex unit 1 (KANSL1) gene, which falls as well in the deleted region. KANSL1 is known to be part of the non-specific lethal (NSL) complex, which regulates global transcription by histone modification<sup>76</sup>. The complex includes several proteins including KAT8 on chromosome 16, on which we also detected high SMR signal as aforementioned (Figure 2A). This points to an important role of epigenetics integrity in both adipogenesis and neuronal development as recently reported<sup>77,78</sup>.

Overall, our functional analysis of the SMR results (Figure 3A) is in agreement with the fact that metabolic function is associated with increased MHC class II antigen presentation in adipocytes, which leads to increased proinflammatory T-cell activity in adipose tissue<sup>79</sup>. This generalized inflammation in adipocyte characterized by alteration in lipid metabolism would result in a dysregulation of glucose metabolism which ultimately may impact as well cognitive functions<sup>80</sup>.

Finally, our cell-type specific enrichment analysis (CSEA) of the SMR results, supports the existence of a “gut–brain axis” which includes interactions through the nervous system, and mutual crosstalk with the immune and the endocrine systems as previously reported<sup>81,82</sup>.

## **Conclusions**

In conclusion, this study provided novel insights into the genetic architecture of metabolic dysfunction and highlighted its putative causal relationship with cognitive decline. Results in this study can contribute to developing biomarkers, identifying drug targets and may have significant implications for global public health policies providing dietary recommendations for the management of obesity which may lead of lowering the risk of developing dementia during ageing.

## **Materials and Methods:**

### **Study Cohorts**

Established with the aim of discovering genetic and non-genetic contributors to human diseases, the UK Biobank is a very large prospective cohort initiated in 2006 with an ongoing follow-up, including more than 500.000 participants living in the UK<sup>83,84</sup>. The cohort comprises people aged between 40-69 years with a 0.84 male/female ratio<sup>83</sup>. The majority of participants

(94.6%) are of white British ethnicity, whereas the rest are South Asians and Black individuals<sup>83</sup>. Recruitment of the participants was based on proximity to assessment centres. Participants gave informed consent and provided detailed demographic, socioeconomic, lifestyle and health-related data via a touchscreen questionnaire, including medication history. As previously described, participants then underwent a large range of physical and biological assessment measures, including repeated blood pressure measurements, height and weight, from which BMI was derived, accelerometry and neuroimaging visits<sup>83,85</sup>; in addition to urine, saliva and blood collection for biomarker and genetic assessments as previous<sup>84,86</sup>.

Genotyping has been performed using the Affymetrix UK BiLEVE Axiom array on an initial 50,000 participants and the remaining 450,000 participants have been genotyped using the Affymetrix UK Biobank Axiom array<sup>83</sup>. Details of the quality control measures, imputation and reference genome utilised were described as previously<sup>83</sup>.

Furthermore, cognitive tests have been administered at several time points to UK Biobank participants. These tests were developed specifically for UK Biobank, to enable computerized administration at scale without staff involvement, and are thus non-standardized. As previously described<sup>85</sup>, a battery of several baseline and follow-up cognitive tests were administered in English language to participants; covering four major cognitive domains (episodic and working memory, executive function, processing speed and attention).

## Genetic Data Pre-processing

In order to run the MR analysis of the individuals with European ethnicity, we obtained genotype calls of more than 400,000 individuals, available following UK Biobank 2022 release, after centrally performing quality control (QC) procedures and imputation<sup>83,87</sup>. From these individuals, in order to avoid population stratification, we excluded participants with non-European ancestry. We removed from the analysis individuals for which a genetic kinship to



other participants of at least one relative was identified, individuals with a missingness > 10% and with an heterozygosity  $\pm 3$  SD from the mean. In addition, we removed individuals for which gender information was ambiguous or incorrect. After applying these filters, genotype calls of 261.089 individuals remained. The genotype calls consisted of 140.180 females and 120.909 males. Further stringent QC filters were applied using PLINK v.1.9<sup>88</sup>. Markers with a call rate < 95% (--geno 0.05), or those deviating from Hardy-Weinberg equilibrium (Bonferroni-corrected p-value threshold = 0.001) were removed. We filtered out rare variants (--maf 0.01) and we set a minor allele count (--mac) cut-off equal to 3.

## Genome-wide association study (GWAS) analysis

We estimated genome wide associations (GWAS) for several quantitative or binary metabolic and cognitive phenotypes obtained from the UK Biobank (Table S1), by fitting mixed-effects regression models accounting for population stratification and relatedness using REGENIE v2.0<sup>89</sup>. After QC procedures, the Genotype calls of the female dataset included 587.951 SNPs, the male dataset included 591.769 SNPs and the one altogether 577.742 SNPs. To run the GWAS analysis of the Genotype calls, we included as covariates: sex (when required), age, age<sup>2</sup> and 10 ancestry-informative principal components (PCs) derived from the analysis of LD-pruned (200 variant windows, 100 variant sliding window and  $r^2 < 0.2$ ) common variants (--maf 0.01) from Genotype calls, as well 10 principal components derived from the analysis of LD-pruned (200 variant windows, 100 variant sliding window and  $r^2 < 0.2$ ) rare variants from the Genotype calls with a frequency in between a minor allele count of 3 (--mac 3) and 1% (--max-maf 0.01), in order to further avoid population stratification. We obtained PCA values by PLINK v2.0<sup>88</sup>. To run REGENIE, we used the female, male or altogether as input in step 1, in which the whole genome regression model is fit to the traits, and a set of genomic predictions are produced as output. The predictions are then used in step 2 where the association model

was performed by including the aforementioned covariates. We run GWAS analysis on females, males or altogether.

## Mendelian Randomization (MR) Network analysis

We built the MR networks by the R package “bimmer” as previously described<sup>38</sup>. Briefly, to run “bimmer”, we used the male GWAS summary statistics for SNP selection and weight estimation, whereas the female GWAS summary statistics were used for model fitting. Initially, we performed pre-processing steps where we clumped the male GWAS summary statistics to  $p = 0.05$  with  $r^2 < 0.05$  and distance of 500 kilobases by PLINK 1.9<sup>88</sup>. The clumped SNPs were then utilized to select SNPs in the female GWAS summary statistics by using a “p\_thresh” value of 0.05. We used the “egger\_w” method to calculate a total causal effect (TCE) matrix on the female summary statistics by utilising the selected SNPs from the male dataset. Subsequently, we filtered the TCE matrix to remove problematic phenotypes with the parameters  $\text{max\_SE} = 0.9$  and  $\text{max\_nan\_perc} = 0.9$ . Then, we converted the TCE matrix to a direct causal effect (DCE) matrix. To do this, we constructed at first weights with a  $\text{max\_min\_ratio}$  equal to 1000. Finally, we used the `fit_inspire()` function to fit the DCE matrix and then we obtained the estimate of the DCEs at the selected index which was used to build the network by `ggnet2`<sup>38</sup>.

## Pleiotropy analysis

We conducted an SMR&HEIDI analysis with the SMR software as previously described. (<https://yanglab.westlake.edu.cn/software/smr/>)<sup>33</sup>. To run this analysis, we utilised GWAS summary statistics obtained previously for the European individuals as outcome. We performed SMR analysis on the sex-combined for participants of European ancestry, using

eQTL data as exposure. We downloaded the lite version of V8 release ( $n = 73\text{-}670$ ) eQTLs summary statistics (only SNPs with  $P < 1e\text{-}5$  were included) obtained from the GTEx Consortium<sup>41</sup> for European individuals. We adopted the default settings in SMR (e.g.,  $peqtl\text{-}smr = 5.0e\text{-}8$ ,  $peqtl\text{-}heidi = 1.57e\text{-}3$ ,  $diff\text{-}freq = 0.2$ ,  $diff\text{-}freq\text{-}prop = 0.05$ ,  $cis\text{-}window = 2000$  kb, minor allele frequency [MAF]  $> 0.01$ , removing SNPs in very strong linkage disequilibrium [LD,  $r^2 > 0.9$ ] with the top associated eQTL, and removing SNPs in low LD or not in LD [ $r^2$  less than 0.05] with the top associated eQTL).

We also conducted the heterogeneity in dependent instruments (HEIDI) test to evaluate the existence of linkage in the observed association<sup>33</sup>.

We performed functional enrichment analysis with the R packages `gprofiler2`<sup>90</sup> and `clusterProfiler`<sup>91</sup> with Bonferroni correction. We built a PPI network with the online database STRING v11.5 (<https://string-db.org/>)<sup>92</sup>.

## Integration with RNA-seq Public Data

We performed a tissue enrichment analysis of the sex-combined GWAS summary statistics of the metabolic and cognitive traits, with the `deTS` R package<sup>93</sup>. To run the enrichment analysis, phenotypes with less than 20 genes were excluded and we utilised bulk RNA-seq gene expression data from the GTEx database<sup>41</sup>.

We performed a cell-type-specific enrichment analysis (CSEA) of the sex-combined GWAS summary statistics of the metabolic and cognitive traits with the online web-based application WebCSEA<sup>48</sup>, which contains a total of 111 single cell RNA-seq (scRNA-seq) panels of human tissues across 11 human organ systems.

## Code and Data Availability Statement

The UKBB phenotypic data and genetic data analysed during the current study are available from UK Biobank. GWAS pre-processing and analysis was performed on the UK Biobank Research Analysis Platform (RAP) following UK Biobank coding guidelines (<https://dnanexus.gitbook.io>). Bimmer is implemented as an open-source R package available at <https://github.com/brielin/bimmer>. The deTS tissue enrichment tool is implemented as an open-source R package available at <https://github.com/bsml320/deTS>.

## Online resources

UK Biobank <https://www.ukbiobank.ac.uk/>

<https://dnanexus.gitbook.io>

SMR <https://yanglab.westlake.edu.cn/software/smr/>

g:Profiler <https://biit.cs.ut.ee/gprofiler/gost>

WebCSEA <https://bioinfo.uth.edu/webcsea/>

STRING <https://string-db.org/>

## Funding Statement

The research has been conducted using the UK Biobank Resource (project # 43769). UK Biobank was established by the Wellcome Trust medical charity, Medical Research Council, Department of Health, Scottish Government and Northwest Regional Development Agency. It also had funding from the Welsh Assembly Government, British Heart Foundation and Diabetes UK. The authors are grateful to the UK Biobank participants for making such research possible. This work was supported by the Singapore Ministry of Education under its Singapore Ministry of Education Academic Research Fund Tier 1.

## Author Contributions

P.A.L designed methodology, performed the analyses and summarized the results, writing code and data preparation, writing-original manuscript draft. S.R.L conceptualized the study, funding acquisition, designed methodology, resources acquisition, supervised the study and reviewed and edited the manuscript.

## Conflicts of Interest

All authors declare no conflict of interest.

## Acknowledgments

This research has been conducted using the UK Biobank Resource under Application Number # 43769 and uses data provided by patients and collected by the NHS as part of their care and support. We thank the UK Biobank (UKBB) for their assistance in preparing the data and providing access to computational resources. Part of the analyses were conducted on the Research Analysis Platform (<http://ukbiobank.dnanexus.com>). We are grateful to the National Supercomputing Center, Singapore (<https://www.nsc.sg>), where part of this work was performed. We thank Matthew Tham and the members of the Integrative Biology of Disease Group for their insightful comments and suggestions.

## References

1. Murman, D.L. (2015). The Impact of Age on Cognition. *Semin Hear* 36, 111-121. 10.1055/s-0035-1555115.
2. Arvanitakis, Z., Shah, R.C., and Bennett, D.A. (2019). Diagnosis and Management of Dementia: Review. *JAMA* 322, 1589-1599. 10.1001/jama.2019.4782.
3. Hugo, J., and Ganguli, M. (2014). Dementia and cognitive impairment: epidemiology, diagnosis, and treatment. *Clin Geriatr Med* 30, 421-442. 10.1016/j.cger.2014.04.001.
4. Larson, E.B., and Langa, K.M. (2012). Aging and incidence of dementia: A critical question. *Neurology* 78, 1452-1453. 10.1212/WNL.0b013e318256c00d.
5. Brayne, C., and Miller, B. (2017). Dementia and aging populations-A global priority for contextualized research and health policy. *PLoS Med* 14, e1002275. 10.1371/journal.pmed.1002275.

6. Prince, M.J., Wu, F., Guo, Y., Gutierrez Robledo, L.M., O'Donnell, M., Sullivan, R., and Yusuf, S. (2015). The burden of disease in older people and implications for health policy and practice. *Lancet* 385, 549-562. 10.1016/S0140-6736(14)61347-7.
7. Avan, A., and Hachinski, V. (2022). Global, regional, and national trends of dementia incidence and risk factors, 1990-2019: A Global Burden of Disease study. *Alzheimers Dement.* 10.1002/alz.12764.
8. Livingston, G., Huntley, J., Sommerlad, A., Ames, D., Ballard, C., Banerjee, S., Brayne, C., Burns, A., Cohen-Mansfield, J., Cooper, C., et al. (2020). Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet* 396, 413-446. 10.1016/S0140-6736(20)30367-6.
9. Coon, D.W., and Gomez-Morales, A. (2022). Modifiable Risk Factors for Brain Health and Dementia and Opportunities for Intervention: A Brief Review. *Clin Gerontol*, 1-12. 10.1080/07317115.2022.2114396.
10. Ranson, J.M., Rittman, T., Hayat, S., Brayne, C., Jessen, F., Blennow, K., van Duijn, C., Barkhof, F., Tang, E., Mummery, C.J., et al. (2021). Modifiable risk factors for dementia and dementia risk profiling. A user manual for Brain Health Services-part 2 of 6. *Alzheimers Res Ther* 13, 169. 10.1186/s13195-021-00895-4.
11. Villani, S., Ferraro, O.E., Poloni, T.E., and Guaita, A. (2022). Dementia and Risk Factors: Results from a Prospective, Population-Based Cohort Study. *Life (Basel)* 12. 10.3390/life12071055.
12. Biessels, G.J., and Despa, F. (2018). Cognitive decline and dementia in diabetes mellitus: mechanisms and clinical implications. *Nat Rev Endocrinol* 14, 591-604. 10.1038/s41574-018-0048-7.
13. Luchsinger, J.A. (2012). Type 2 diabetes and cognitive impairment: linking mechanisms. *J Alzheimers Dis* 30 Suppl 2, S185-198. 10.3233/JAD-2012-111433.
14. Verhagen, C., Janssen, J., Biessels, G.J., Johansen, O.E., and Exalto, L.G. (2022). Females with type 2 diabetes are at higher risk for accelerated cognitive decline than males: CAROLINA-COGNITION study. *Nutr Metab Cardiovasc Dis* 32, 355-364. 10.1016/j.numecd.2021.10.013.
15. Cholerton, B., Baker, L.D., Montine, T.J., and Craft, S. (2016). Type 2 Diabetes, Cognition, and Dementia in Older Adults: Toward a Precision Health Approach. *Diabetes Spectr* 29, 210-219. 10.2337/ds16-0041.
16. Muriach, M., Flores-Bellver, M., Romero, F.J., and Barcia, J.M. (2014). Diabetes and the brain: oxidative stress, inflammation, and autophagy. *Oxid Med Cell Longev* 2014, 102158. 10.1155/2014/102158.
17. Forny-Germano, L., De Felice, F.G., and Vieira, M. (2018). The Role of Leptin and Adiponectin in Obesity-Associated Cognitive Decline and Alzheimer's Disease. *Front Neurosci* 12, 1027. 10.3389/fnins.2018.01027.
18. Fink, A., Doerre, A., Demuth, I., and Doblhammer, G. (2022). Potential of prevention strategies for the modifiable risk factor type 2 diabetes with relation to the future number of dementia patients in Germany- a multi-state projection through 2040. *BMC Neurol* 22, 157. 10.1186/s12883-022-02682-6.
19. Irwin, K., Sexton, C., Daniel, T., Lawlor, B., and Naci, L. (2018). Healthy Aging and Dementia: Two Roads Diverging in Midlife? *Front Aging Neurosci* 10, 275. 10.3389/fnagi.2018.00275.
20. Burgess, S., Davey Smith, G., Davies, N.M., Dudbridge, F., Gill, D., Glymour, M.M., Hartwig, F.P., Holmes, M.V., Minelli, C., Relton, C.L., and Theodoratou, E. (2019).

- Guidelines for performing Mendelian randomization investigations. *Wellcome Open Res* 4, 186. [10.12688/wellcomeopenres.15555.2](https://doi.org/10.12688/wellcomeopenres.15555.2).
21. Lawlor, D.A., Harbord, R.M., Sterne, J.A., Timpson, N., and Davey Smith, G. (2008). Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med* 27, 1133-1163. [10.1002/sim.3034](https://doi.org/10.1002/sim.3034).
  22. Davies, N.M., Holmes, M.V., and Davey Smith, G. (2018). Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* 362, k601. [10.1136/bmj.k601](https://doi.org/10.1136/bmj.k601).
  23. Didelez, V., and Sheehan, N. (2007). Mendelian randomization as an instrumental variable approach to causal inference. *Stat Methods Med Res* 16, 309-330. [10.1177/0962280206077743](https://doi.org/10.1177/0962280206077743).
  24. Hemani, G., Bowden, J., and Davey Smith, G. (2018). Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum Mol Genet* 27, R195-R208. [10.1093/hmg/ddy163](https://doi.org/10.1093/hmg/ddy163).
  25. Ostergaard, S.D., Mukherjee, S., Sharp, S.J., Proitsi, P., Lotta, L.A., Day, F., Perry, J.R., Boehme, K.L., Walter, S., Kauwe, J.S., et al. (2015). Associations between Potentially Modifiable Risk Factors and Alzheimer Disease: A Mendelian Randomization Study. *PLoS Med* 12, e1001841; discussion e1001841. [10.1371/journal.pmed.1001841](https://doi.org/10.1371/journal.pmed.1001841).
  26. Thomassen, J.Q., Tolstrup, J.S., Benn, M., and Frikke-Schmidt, R. (2020). Type-2 diabetes and risk of dementia: observational and Mendelian randomisation studies in 1 million individuals. *Epidemiol Psychiatr Sci* 29, e118. [10.1017/S2045796020000347](https://doi.org/10.1017/S2045796020000347).
  27. Anand, S.S., Friedrich, M.G., Lee, D.S., Awadalla, P., Despres, J.P., Desai, D., de Souza, R.J., Dummer, T., Parraga, G., Larose, E., et al. (2022). Evaluation of Adiposity and Cognitive Function in Adults. *JAMA Netw Open* 5, e2146324. [10.1001/jamanetworkopen.2021.46324](https://doi.org/10.1001/jamanetworkopen.2021.46324).
  28. Benn, M., Nordestgaard, B.G., Tybjaerg-Hansen, A., and Frikke-Schmidt, R. (2020). Impact of glucose on risk of dementia: Mendelian randomisation studies in 115,875 individuals. *Diabetologia* 63, 1151-1161. [10.1007/s00125-020-05124-5](https://doi.org/10.1007/s00125-020-05124-5).
  29. Pan, Y., Chen, W., Yan, H., Wang, M., and Xiang, X. (2020). Glycemic traits and Alzheimer's disease: a Mendelian randomization study. *Aging (Albany NY)* 12, 22688-22699. [10.18632/aging.103887](https://doi.org/10.18632/aging.103887).
  30. Ware, E.B., Morataya, C., Fu, M., and Bakulski, K.M. (2021). Type 2 Diabetes and Cognitive Status in the Health and Retirement Study: A Mendelian Randomization Approach. *Front Genet* 12, 634767. [10.3389/fgene.2021.634767](https://doi.org/10.3389/fgene.2021.634767).
  31. Walter, S., Marden, J.R., Kubzansky, L.D., Mayeda, E.R., Crane, P.K., Chang, S.C., Cornelis, M., Rehkopf, D.H., Mukherjee, S., and Glymour, M.M. (2016). Diabetic Phenotypes and Late-Life Dementia Risk: A Mechanism-specific Mendelian Randomization Study. *Alzheimer Dis Assoc Disord* 30, 15-20. [10.1097/WAD.0000000000000128](https://doi.org/10.1097/WAD.0000000000000128).
  32. Tam, V., Patel, N., Turcotte, M., Bosse, Y., Pare, G., and Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nat Rev Genet* 20, 467-484. [10.1038/s41576-019-0127-1](https://doi.org/10.1038/s41576-019-0127-1).
  33. Wu, Y., Zeng, J., Zhang, F., Zhu, Z., Qi, T., Zheng, Z., Lloyd-Jones, L.R., Marioni, R.E., Martin, N.G., Montgomery, G.W., et al. (2018). Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. *Nat Commun* 9, 918. [10.1038/s41467-018-03371-0](https://doi.org/10.1038/s41467-018-03371-0).



34. Hekselman, I., and Yeger-Lotem, E. (2020). Mechanisms of tissue and cell-type specificity in heritable traits and diseases. *Nat Rev Genet* 21, 137-150. 10.1038/s41576-019-0200-9.
35. SantaCruz-Calvo, S., Bharath, L., Pugh, G., SantaCruz-Calvo, L., Lenin, R.R., Lutshumba, J., Liu, R., Bachstetter, A.D., Zhu, B., and Nikolajczyk, B.S. (2022). Adaptive immune cells shape obesity-associated type 2 diabetes mellitus and less prominent comorbidities. *Nat Rev Endocrinol* 18, 23-42. 10.1038/s41574-021-00575-1.
36. Mahase, E. (2023). UK Biobank: Major genome sequencing release could "unlock" new treatments and cures. *BMJ* 383, 2818. 10.1136/bmj.p2818.
37. Harvey, P.D. (2019). Domains of cognition and their assessment. *Dialogues Clin Neurosci* 21, 227-237. 10.31887/DCNS.2019.21.3/pharvey.
38. Brown, B.C., and Knowles, D.A. (2020). Phenome-scale causal network discovery with bidirectional mediated Mendelian randomization. *bioRxiv*, 2020.2006.2018.160176. 10.1101/2020.06.18.160176.
39. Zhu, Z., Zhang, F., Hu, H., Bakshi, A., Robinson, M.R., Powell, J.E., Montgomery, G.W., Goddard, M.E., Wray, N.R., Visscher, P.M., and Yang, J. (2016). Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet* 48, 481-487. 10.1038/ng.3538.
40. Adolphe, C., Xue, A., Fard, A.T., Genovesi, L.A., Yang, J., and Wainwright, B.J. (2021). Genetic and functional interaction network analysis reveals global enrichment of regulatory T cell genes influencing basal cell carcinoma susceptibility. *Genome Med* 13, 19. 10.1186/s13073-021-00827-9.
41. Consortium, G.T. (2020). The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* 369, 1318-1330. 10.1126/science.aaz1776.
42. Wu, H., and Ballantyne, C.M. (2017). Skeletal muscle inflammation and insulin resistance in obesity. *J Clin Invest* 127, 43-54. 10.1172/JCI88880.
43. Bachmann-Gagescu, R., Mefford, H.C., Cowan, C., Glew, G.M., Hing, A.V., Wallace, S., Bader, P.I., Hamati, A., Reitnauer, P.J., Smith, R., et al. (2010). Recurrent 200-kb deletions of 16p11.2 that include the SH2B1 gene are associated with developmental delay and obesity. *Genet Med* 12, 641-647. 10.1097/GIM.0b013e3181ef4286.
44. Perrone, L., Marzuillo, P., Grandone, A., and del Giudice, E.M. (2010). Chromosome 16p11.2 deletions: another piece in the genetic puzzle of childhood obesity. *Ital J Pediatr* 36, 43. 10.1186/1824-7288-36-43.
45. Walters, R.G., Jacquemont, S., Valsesia, A., de Smith, A.J., Martinet, D., Andersson, J., Falchi, M., Chen, F., Andrieux, J., Lobbens, S., et al. (2010). A new highly penetrant form of obesity due to deletions on chromosome 16p11.2. *Nature* 463, 671-675. 10.1038/nature08727.
46. Zufferey, F., Sherr, E.H., Beckmann, N.D., Hanson, E., Maillard, A.M., Hippolyte, L., Mace, A., Ferrari, C., Kutalik, Z., Andrieux, J., et al. (2012). A 600 kb deletion syndrome at 16p11.2 leads to energy imbalance and neuropsychiatric disorders. *J Med Genet* 49, 660-668. 10.1136/jmedgenet-2012-101203.
47. Locke, A.E., Kahali, B., Berndt, S.I., Justice, A.E., Pers, T.H., Day, F.R., Powell, C., Vedantam, S., Buchkovich, M.L., Yang, J., et al. (2015). Genetic studies of body mass index yield new insights for obesity biology. *Nature* 518, 197-206. 10.1038/nature14177.



48. Dai, Y., Hu, R., Liu, A., Cho, K.S., Manuel, A.M., Li, X., Dong, X., Jia, P., and Zhao, Z. (2022). WebCSEA: web-based cell-type-specific enrichment analysis of genes. *Nucleic Acids Res.* 10.1093/nar/gkac392.
49. Wang, S.H., Su, M.H., Chen, C.Y., Lin, Y.F., Feng, Y.A., Hsiao, P.C., Pan, Y.J., and Wu, C.S. (2022). Causality of abdominal obesity on cognition: a trans-ethnic Mendelian randomization study. *Int J Obes (Lond)* 46, 1487-1492. 10.1038/s41366-022-01138-8.
50. Farruggia, M.C., and Small, D.M. (2019). Effects of adiposity and metabolic dysfunction on cognition: A review. *Physiol Behav* 208, 112578. 10.1016/j.physbeh.2019.112578.
51. Kanoski, S.E., and Davidson, T.L. (2011). Western diet consumption and cognitive impairment: links to hippocampal dysfunction and obesity. *Physiol Behav* 103, 59-68. 10.1016/j.physbeh.2010.12.003.
52. Ndiaye, F.K., Huyvaert, M., Ortalli, A., Canouil, M., Lecoeur, C., Verbanck, M., Lobbens, S., Khamis, A., Marselli, L., Marchetti, P., et al. (2020). The expression of genes in top obesity-associated loci is enriched in insula and substantia nigra brain regions involved in addiction and reward. *Int J Obes (Lond)* 44, 539-543. 10.1038/s41366-019-0428-7.
53. Petrova, V., and Hakim, S. (2022). CELSR2, a new player in motor neuron axon growth and regeneration. *Brain* 145, 420-422. 10.1093/brain/awac021.
54. Tan, J., Che, Y., Liu, Y., Hu, J., Wang, W., Hu, L., Zhou, Q., Wang, H., and Li, J. (2021). CELSR2 deficiency suppresses lipid accumulation in hepatocyte by impairing the UPR and elevating ROS level. *FASEB J* 35, e21908. 10.1096/fj.202100786RR.
55. Arvind, P., Nair, J., Jambunathan, S., Kakkar, V.V., and Shanker, J. (2014). CELSR2-PSRC1-SORT1 gene expression and association with coronary artery disease and plasma lipid levels in an Asian Indian cohort. *J Cardiol* 64, 339-346. 10.1016/j.jjcc.2014.02.012.
56. Yang, J., Villar, V.A.M., Rozyyev, S., Jose, P.A., and Zeng, C. (2019). The emerging role of sorting nexins in cardiovascular diseases. *Clin Sci (Lond)* 133, 723-737. 10.1042/CS20190034.
57. Gallon, M., and Cullen, P.J. (2015). Retromer and sorting nexins in endosomal sorting. *Biochem Soc Trans* 43, 33-47. 10.1042/BST20140290.
58. Sotelo, P., Farfan, P., Benitez, M.L., Bu, G., and Marzolo, M.P. (2014). Sorting nexin 17 regulates ApoER2 recycling and reelin signaling. *PLoS One* 9, e93672. 10.1371/journal.pone.0093672.
59. Mefford, H.C., Muhle, H., Ostertag, P., von Spiczak, S., Buysse, K., Baker, C., Franke, A., Malafosse, A., Genton, P., Thomas, P., et al. (2010). Genome-wide copy number variation in epilepsy: novel susceptibility loci in idiopathic generalized and focal epilepsies. *PLoS Genet* 6, e1000962. 10.1371/journal.pgen.1000962.
60. Dell'Edera, D., Dilucca, C., Allegretti, A., Simone, F., Lupo, M.G., Liccese, C., and Davanzo, R. (2018). 16p11.2 microdeletion syndrome: a case report. *J Med Case Rep* 12, 90. 10.1186/s13256-018-1587-1.
61. Fetit, R., Price, D.J., Lawrie, S.M., and Johnstone, M. (2020). Understanding the clinical manifestations of 16p11.2 deletion syndrome: a series of developmental case reports in children. *Psychiatr Genet* 30, 136-140. 10.1097/YPG.0000000000000259.
62. Fuller, D.M., Zhu, M., Ou-Yang, C.W., Sullivan, S.A., and Zhang, W. (2011). A tale of two TRAPs: LAT and LAB in the regulation of lymphocyte development, activation, and autoimmunity. *Immunol Res* 49, 97-108. 10.1007/s12026-010-8197-3.

63. Tomasello, D.L., Kim, J.L., Khodour, Y., McCammon, J.M., Mitalipova, M., Jaenisch, R., Futerman, A.H., and Sive, H. (2022). 16pdel lipid changes in iPSC-derived neurons and function of FAM57B in lipid metabolism and synaptogenesis. *iScience* 25, 103551. [10.1016/j.isci.2021.103551](https://doi.org/10.1016/j.isci.2021.103551).
64. Li, L., Ghorbani, M., Weisz-Hubshman, M., Rousseau, J., Thiffault, I., Schnur, R.E., Breen, C., Oegema, R., Weiss, M.M., Waisfisz, Q., et al. (2020). Lysine acetyltransferase 8 is involved in cerebral development and syndromic intellectual disability. *J Clin Invest* 130, 1431-1445. [10.1172/JCI131145](https://doi.org/10.1172/JCI131145).
65. Pessoa Rodrigues, C., Chatterjee, A., Wiese, M., Stehle, T., Szymanski, W., Shvedunova, M., and Akhtar, A. (2021). Histone H4 lysine 16 acetylation controls central carbon metabolism and diet-induced obesity in mice. *Nat Commun* 12, 6212. [10.1038/s41467-021-26277-w](https://doi.org/10.1038/s41467-021-26277-w).
66. Wijesinghe, S.N., Nicholson, T., Tsintzas, K., and Jones, S.W. (2021). Involvements of long noncoding RNAs in obesity-associated inflammatory diseases. *Obes Rev* 22, e13156. [10.1111/obr.13156](https://doi.org/10.1111/obr.13156).
67. Melka, M.G., Gillis, J., Bernard, M., Abrahamowicz, M., Chakravarty, M.M., Leonard, G.T., Perron, M., Richer, L., Veillette, S., Banaschewski, T., et al. (2013). FTO, obesity and the adolescent brain. *Hum Mol Genet* 22, 1050-1058. [10.1093/hmg/dd504](https://doi.org/10.1093/hmg/dd504).
68. Osorio, J. (2014). Obesity. BMI-associated genetic variants in FTO mastermind IRX3 expression. *Nat Rev Endocrinol* 10, 313. [10.1038/nrendo.2014.36](https://doi.org/10.1038/nrendo.2014.36).
69. Smemo, S., Tena, J.J., Kim, K.H., Gamazon, E.R., Sakabe, N.J., Gomez-Marin, C., Aneas, I., Credidio, F.L., Sobreira, D.R., Wasserman, N.F., et al. (2014). Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature* 507, 371-375. [10.1038/nature13138](https://doi.org/10.1038/nature13138).
70. Cha, J.H., Jeong, Y., Oh, A.R., Lee, S.B., Hong, S.S., and Kim, K. (2021). Emerging roles of PHLPP phosphatases in metabolism. *BMB Rep* 54, 451-457. [10.5483/BMBRep.2021.54.9.095](https://doi.org/10.5483/BMBRep.2021.54.9.095).
71. Shinawi, M., Sahoo, T., Maranda, B., Skinner, S.A., Skinner, C., Chinault, C., Zascavage, R., Peters, S.U., Patel, A., Stevenson, R.E., and Beaudet, A.L. (2011). 11p14.1 microdeletions associated with ADHD, autism, developmental delay, and obesity. *Am J Med Genet A* 155A, 1272-1280. [10.1002/ajmg.a.33878](https://doi.org/10.1002/ajmg.a.33878).
72. Teslovich, T.M., Musunuru, K., Smith, A.V., Edmondson, A.C., Stylianou, I.M., Koseki, M., Pirruccello, J.P., Ripatti, S., Chasman, D.I., Willer, C.J., et al. (2010). Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 466, 707-713. [10.1038/nature09270](https://doi.org/10.1038/nature09270).
73. D'Angelo, S., Motti, M.L., and Meccariello, R. (2020). omega-3 and omega-6 Polyunsaturated Fatty Acids, Obesity and Cancer. *Nutrients* 12. [10.3390/nu12092751](https://doi.org/10.3390/nu12092751).
74. Gomez-Chacon, B., Gamiz, F., Foster, T.C., and Gallo, M. (2016). Increased N-Ethylmaleimide-Sensitive Factor Expression in Amygdala and Perirhinal Cortex during Habituation of Taste Neophobia. *Neural Plast* 2016, 2726745. [10.1155/2016/2726745](https://doi.org/10.1155/2016/2726745).
75. Koolen, D.A., Pfundt, R., Linda, K., Beunders, G., Veenstra-Knol, H.E., Conta, J.H., Fortuna, A.M., Gillissen-Kaesbach, G., Dugan, S., Halbach, S., et al. (2016). The Koolen-de Vries syndrome: a phenotypic comparison of patients with a 17q21.31 microdeletion versus a KANSL1 sequence variant. *Eur J Hum Genet* 24, 652-659. [10.1038/ejhg.2015.178](https://doi.org/10.1038/ejhg.2015.178).

76. Sheikh, B.N., Guhathakurta, S., and Akhtar, A. (2019). The non-specific lethal (NSL) complex at the crossroads of transcriptional control and cellular homeostasis. *EMBO Rep* 20, e47630. [10.15252/embr.201847630](https://doi.org/10.15252/embr.201847630).
77. Gao, W., Liu, J.L., Lu, X., and Yang, Q. (2021). Epigenetic regulation of energy metabolism in obesity. *J Mol Cell Biol* 13, 480-499. [10.1093/jmcb/mjab043](https://doi.org/10.1093/jmcb/mjab043).
78. Geng, H., Chen, H., Wang, H., and Wang, L. (2021). The Histone Modifications of Neuronal Plasticity. *Neural Plast* 2021, 6690523. [10.1155/2021/6690523](https://doi.org/10.1155/2021/6690523).
79. Majdoubi, A., Kishta, O.A., and Thibodeau, J. (2016). Role of antigen presentation in the production of pro-inflammatory cytokines in obese adipose tissue. *Cytokine* 82, 112-121. [10.1016/j.cyto.2016.01.023](https://doi.org/10.1016/j.cyto.2016.01.023).
80. de, A.B.A.P., de, O.C.P.H., BE, F.F., PS, E.S., de Moraes, L., and Migliolo, L. (2023). Adipose tissue, systematic inflammation, and neurodegenerative diseases. *Neural Regen Res* 18, 38-46. [10.4103/1673-5374.343891](https://doi.org/10.4103/1673-5374.343891).
81. Niccolai, E., Boem, F., Russo, E., and Amedei, A. (2019). The Gut(-)Brain Axis in the Neuropsychological Disease Model of Obesity: A Classical Movie Revised by the Emerging Director "Microbiome". *Nutrients* 11. [10.3390/nu11010156](https://doi.org/10.3390/nu11010156).
82. Agusti, A., Garcia-Pardo, M.P., Lopez-Almela, I., Campillo, I., Maes, M., Romani-Perez, M., and Sanz, Y. (2018). Interplay Between the Gut-Brain Axis, Obesity and Cognitive Function. *Front Neurosci* 12, 155. [10.3389/fnins.2018.00155](https://doi.org/10.3389/fnins.2018.00155).
83. Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L.T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J., et al. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature* 562, 203-209. [10.1038/s41586-018-0579-z](https://doi.org/10.1038/s41586-018-0579-z).
84. Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., et al. (2015). UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 12, e1001779. [10.1371/journal.pmed.1001779](https://doi.org/10.1371/journal.pmed.1001779).
85. Campbell, T., and Cullen, B. (2023). Estimating the effect of physical activity on cognitive function within the UK Biobank cohort. *Int J Epidemiol*. [10.1093/ije/dyad009](https://doi.org/10.1093/ije/dyad009).
86. Che, B., Zhong, C., Zhang, R., Pu, L., Zhao, T., Zhang, Y., and Han, L. (2023). Triglyceride-glucose index and triglyceride to high-density lipoprotein cholesterol ratio as potential cardiovascular disease risk factors: an analysis of UK biobank data. *Cardiovasc Diabetol* 22, 34. [10.1186/s12933-023-01762-2](https://doi.org/10.1186/s12933-023-01762-2).
87. Halldorsson, B.V., Eggertsson, H.P., Moore, K.H.S., Hauswedell, H., Eiriksson, O., Ulfarsson, M.O., Palsson, G., Hardarson, M.T., Oddsson, A., Jensson, B.O., et al. (2022). The sequences of 150,119 genomes in the UK Biobank. *Nature* 607, 732-740. [10.1038/s41586-022-04965-x](https://doi.org/10.1038/s41586-022-04965-x).
88. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4, 7. [10.1186/s13742-015-0047-8](https://doi.org/10.1186/s13742-015-0047-8).
89. Mbatchou, J., Barnard, L., Backman, J., Marcketta, A., Kosmicki, J.A., Ziyatdinov, A., Benner, C., O'Dushlaine, C., Barber, M., Boutkov, B., et al. (2021). Computationally efficient whole-genome regression for quantitative and binary traits. *Nat Genet* 53, 1097-1103. [10.1038/s41588-021-00870-7](https://doi.org/10.1038/s41588-021-00870-7).

90. Kolberg, L., Raudvere, U., Kuzmin, I., Vilo, J., and Peterson, H. (2020). gprofiler2 -- an R package for gene list functional enrichment analysis and namespace conversion toolset g:Profiler. *F1000Res* 9. 10.12688/f1000research.24956.2.
91. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)* 2, 100141. 10.1016/j.xinn.2021.100141.
92. Szklarczyk, D., Gable, A.L., Nastou, K.C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N.T., Legeay, M., Fang, T., Bork, P., et al. (2021). The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res* 49, D605-D612. 10.1093/nar/gkaa1074.
93. Pei, G., Dai, Y., Zhao, Z., and Jia, P. (2019). deTS: tissue-specific enrichment analysis to decode tissue specificity. *Bioinformatics* 35, 3842-3845. 10.1093/bioinformatics/btz138.

## Figure Captions.

**Figure 1. (A)** Summary of the workflow conducted in this study. (1) GWAS analysis was performed on genetic and phenotypic data of individuals with European ancestry obtained from UK Biobank database. (2) The GWAS results were then utilized to determine causality by Mendelian Randomization (MR). Subsequently, GWAS were integrated with eQTLs gene expression data in skeletal muscle of individual of European ancestry to perform pleiotropy analysis (3). Finally, GWAS were integrated with bulk RNA-seq public data to perform tissue enrichment analysis (4) and with single cell RNA-seq public data to perform cell-type specific analysis (5). **(B)** Mendelian Randomization Network (MRN) of metabolic (blue) and cognitive (red) traits in Europeans. Single-headed arrows indicate direct causal effects. Bi-directional double-headed arrows indicate that traits causally influence each other.

Systolic Blood Pressure (SBP), Diastolic Blood Pressure (DBP), Low Density Lipoproteins (LDL), High Density Lipoproteins (HDL), Waist Circumference (WC), Hip Circumference (HC), Cholesterol (CHOL), Triglycerides (TRIG), Body Mass Index (BMI), Glycated Hemoglobin (HB), Fat Mass (FM), Total Fat (TF), Total Fat Percentage (TFP), Visceral Adipose Tissue Volume (VATV), Visceral Adipose Tissue Mass (VATM), Abdominal Adipose Tissue Volume (AATV), Body Fat Mass (BFM), Body Fat Percentage (BFP), Working Memory (WM), Episodic Memory (EM), Fluid IQ (FIQ), Reaction Test (RT), Visual Attention Test (VAT). GWAS genome-wide association study, SMR summary data-based Mendelian randomization, eQTL expression quantitative trait loci.

**Figure 2. (A)** SMR plot of the K(lysine) acetyltransferase 8 (KAT8) locus for the Hip Circumference (HC) GWAS trait in Europeans. **(B)** SMR plot of the Fat mass and obesity-associated (FTO) locus for the Body Mass Index (BMI) GWAS trait in Europeans. On the y-axis at the top,  $-\log_{10}(P \text{ values})$  for SNPs reported in this GWAS meta-analysis are represented by grey dots and rhombuses represent the P-values for probes from the reverse SMR test. Solid rhombuses are indicating that the probes passed the HEIDI test and hollow rhombuses indicating that the probes did not pass the HEIDI test. The red dashed line indicates significant pSMR threshold in the SMR test. In the middle plot, magenta x-crosses represent the cis-eQTL  $-\log_{10}(P \text{ values})$  of SNPs associated with gene expression in skeletal muscle (GTex). Bottom plot, yellow arrows indicate genes over the genome positions in hg38 (x-axis). GWAS genome-wide association study, SMR summary data-based Mendelian randomization, HEIDI heterogeneity in dependent instruments, eQTL expression quantitative trait loci, GTEX Genotype-Tissue Expression.

**Figure 3. (A)** Graphical representation of the gProfiler Gene Ontology (GO) analysis of the 86 unique significant genes detected in the SMR analysis which passed the SMR threshold ( $P_{\text{SMR}} < 5.0 \times 10^{-6}$ ) for the European individuals. Top significant results based on adjusted p-values for Gene Ontology (GO) and pathways enrichment were selected and highlighted in black borders. GO molecular function (GO:MF); GO biological process (GO:BP); GO cellular component (GO:CC); Kyoto encyclopedia of genes and genomes (KEGG); Reactome (REAC); WikiPathways (WP); CORUM Protein Complexes (CORUM); Human Phenotype Ontology (HP). **(B)** Tissue Enrichment analysis of the GWAS summary statistics for the Europeans. Traits are on the x-axis and tissue types on the y-axis. Dark red indicate high enrichment. Numbers inside the red boxes indicate top 3 most associated tissues. Gene expression data in various tissues based on Genotype-Tissue Expression (GTEX) database. Systolic Blood Pressure (SBP), Low Density Lipoproteins (LDL), High Density Lipoproteins (HDL), Waist Circumference (WC), Hip Circumference (HC), Cholesterol (CHOL), Triglycerides (TRIG),



Body Mass Index (BMI), Glycated Hemoglobin (HB), Body Fat Mass (BFM), Body Fat Percentage (BFP), Working Memory (WM), Reaction Test (RT), Visual Attention Test (VAT).

**Figure 4.** (A) Jitter plot of the cell-type-specific enrichment analysis (CSEA) for the top significant genes detected in the GWAS analysis for the Body Mass Index (BMI) in Europeans. (B) Jitter plot of the cell-type-specific enrichment analysis (CSEA) for the 86 SMR significant genes in the Europeans. (C) Jitter plot of the cell-type-specific enrichment analysis (CSEA) for 39 HEIDI significant genes in Europeans. Red dashed line indicates Bonferroni-corrected significance threshold ( $P = 3.69 \times 10^{-5}$ ). The grey solid line indicates the nominal significance ( $p = 1 \times 10^{-3}$ ). Each dot represents one tissue-cell type in the group on the x-axis differentiated by colours. Results based on combined p-values. Top 5 general cell types are highlighted.

**Fig S1.** (A) Plot of the gProfiler ontology analysis of the significant GWAS hits for both the metabolic and cognitive traits which passed the threshold ( $p$ -value  $< 5e-08$ ) for the European individuals. (B) Plot of the gProfiler ontology analysis of the 39 unique significant genes that passed both the SMR and HEIDI test in the SMR analysis for the European individuals. GO molecular function (GO:MF); Kyoto encyclopedia of genes and genomes (KEGG); Reactome (REAC); WikiPathways (WP); CORUM Protein Complexes (CORUM).

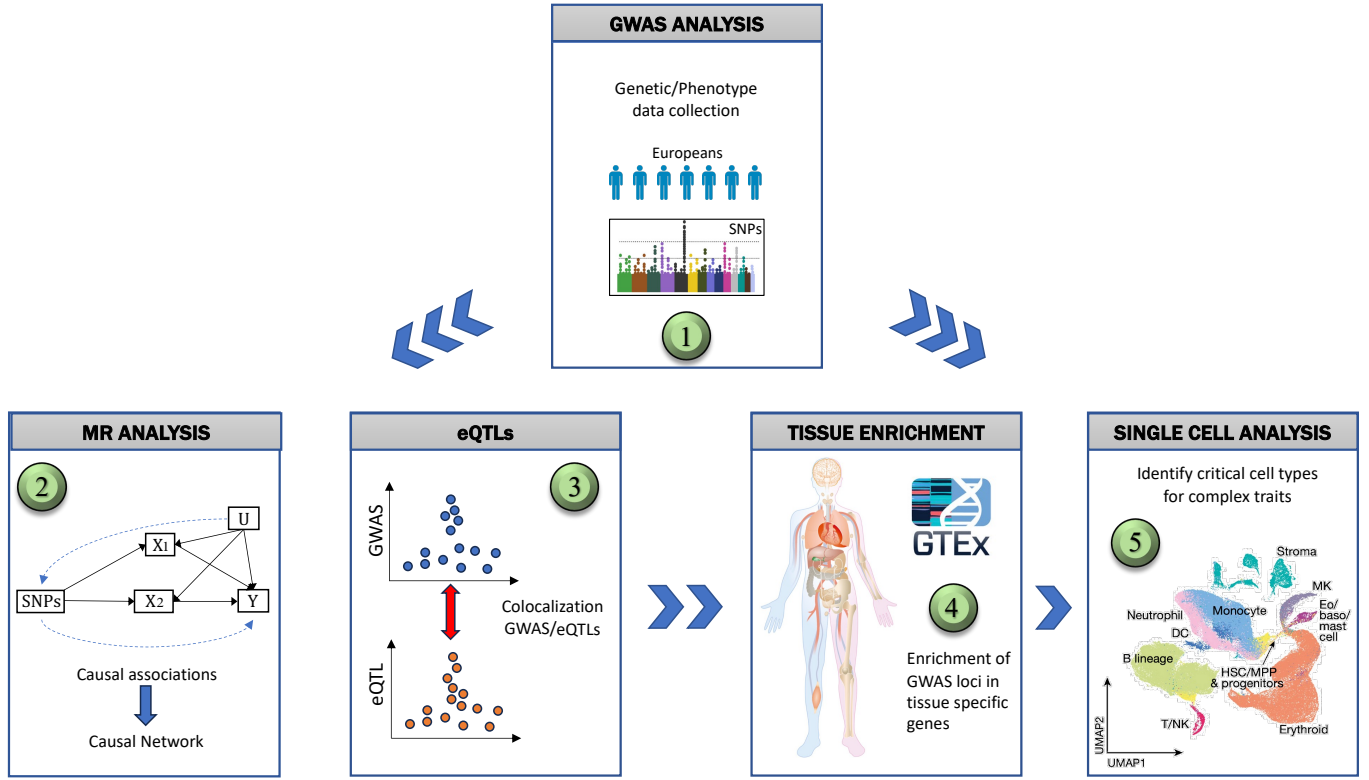
**Fig S2.** (A) Protein-Protein Interaction (PPI) network generated by STRING database of the 39 unique significant genes which passed both the SMR and HEIDI threshold in Europeans. (B) Jitter plot of the cell-type-specific enrichment analysis (CSEA) for the GWAS significant genes detected in the Reaction Test (RT) trait of the Europeans individuals. Red dashed line indicates Bonferroni-corrected significance threshold ( $P = 3.69 \times 10^{-5}$ ) by 1355 TCs. The grey solid line indicates the nominal significance ( $p = 1 \times 10^{-3}$ ). Each dot represents one tissue-cell type in the group on the x-axis differentiated by color. Results based on combined p-values.

**Fig S3.** (A) SMR locus plot of the CELSR2 locus for the Low-Density Lipoproteins (LDL) trait in Europeans. (B) SMR locus plot of the SNX17 locus for the Triglycerides (TRIG) trait in Europeans. (C) SMR locus plot of the FADS2 locus for the Triglycerides (TRIG) trait in Europeans. (D) SMR locus plot of the ST3GAL4 locus for the Low-Density Lipoproteins (LDL) trait in Europeans. On the y-axis, P-values for SNPs reported in this GWAS meta-analysis are represented by grey dots and diamonds represent the P-values for probes from the reverse SMR test. Magenta crosses in the middle plot represent the cis-eQTL P-values of SNPs associated with gene expression in skeletal muscle (GTEx). Bottom yellow arrows indicate genes over the genome positions in hg38 (x-axis).

Figure 1

A

Workflow



B

MR Network

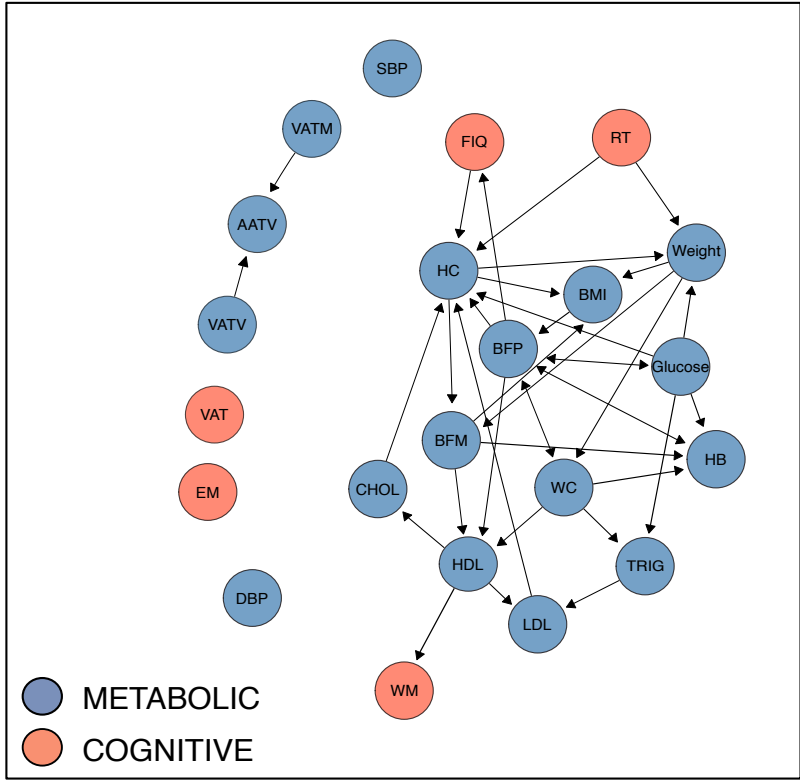


Figure 2

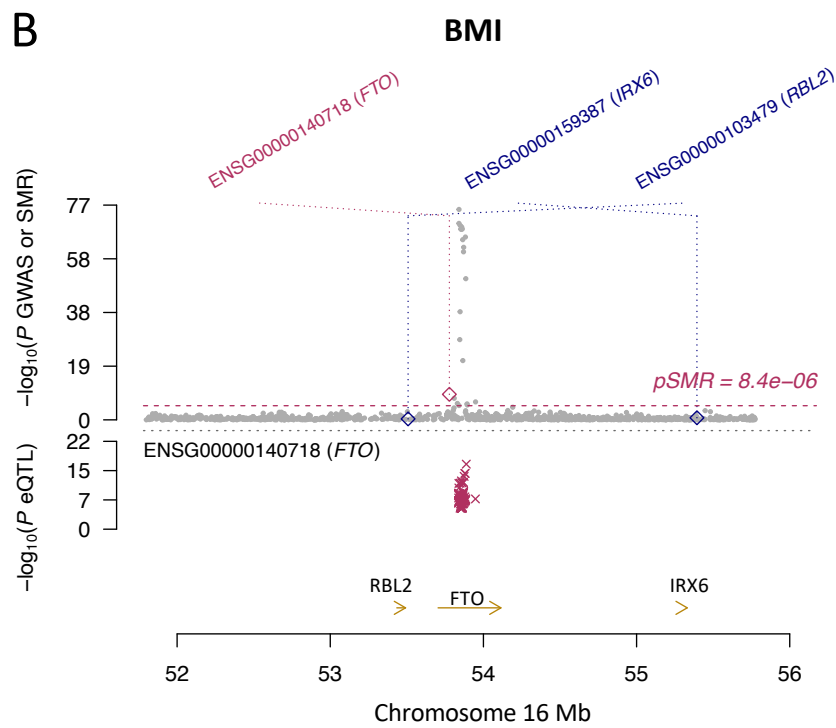
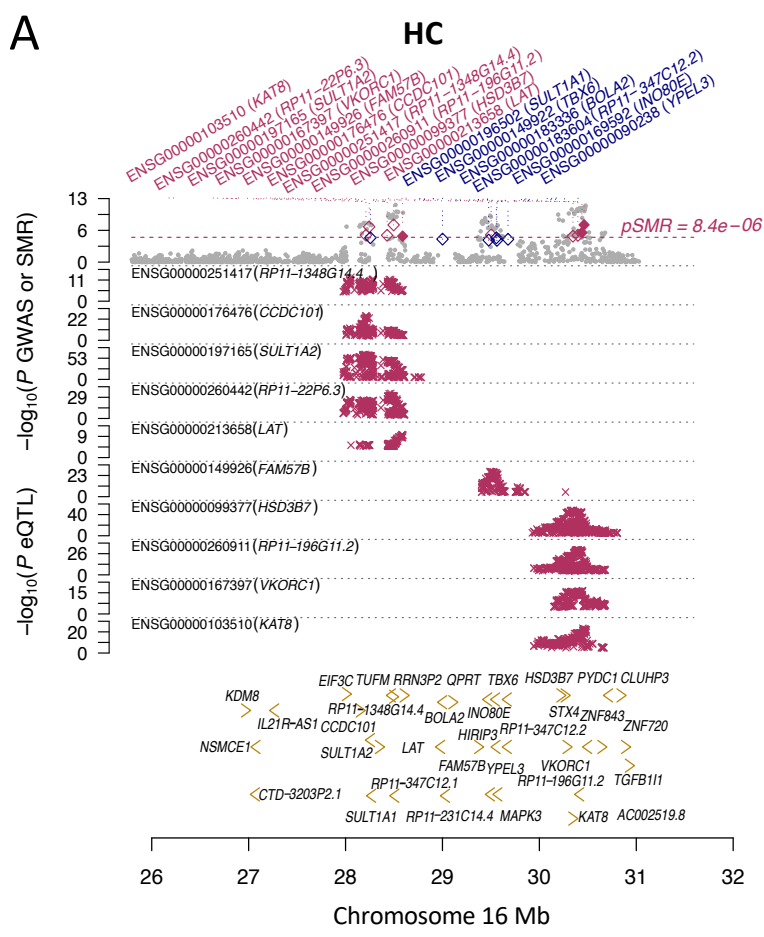






Figure 4

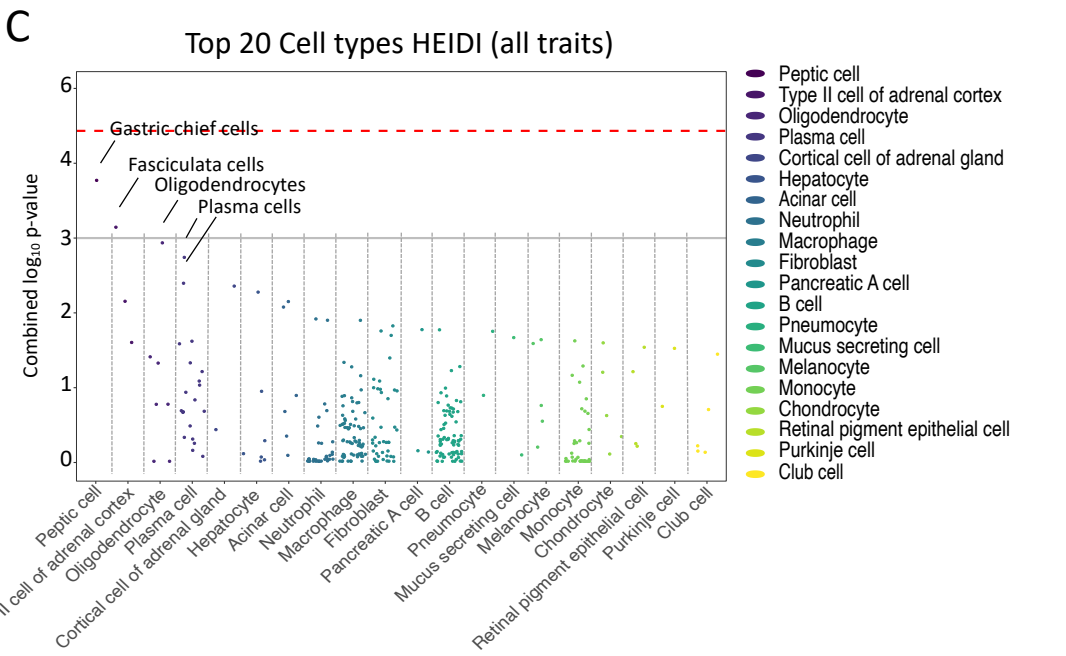
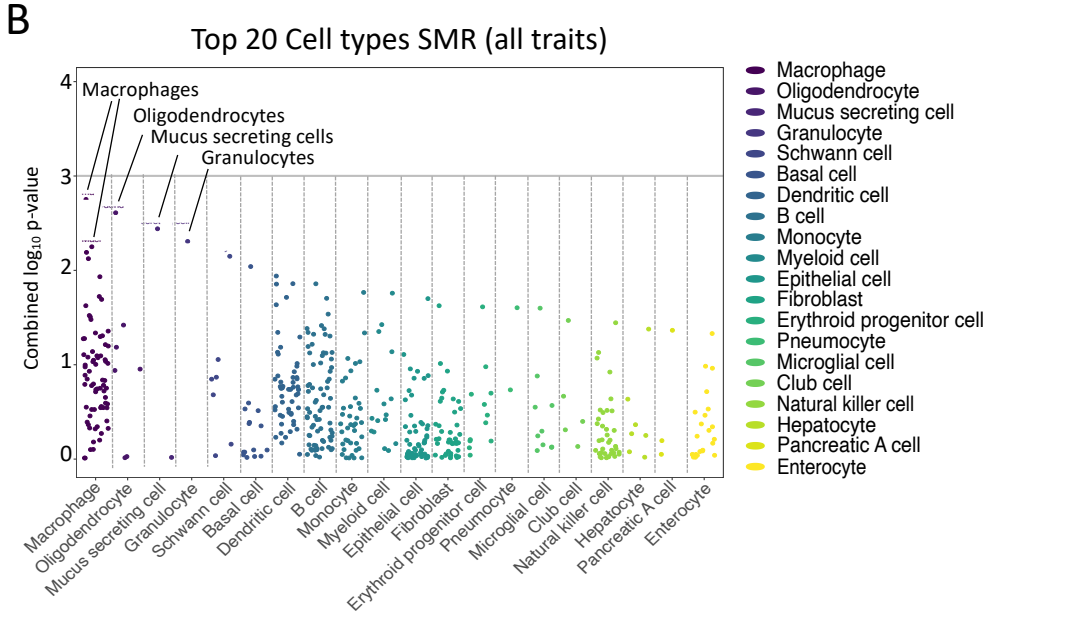
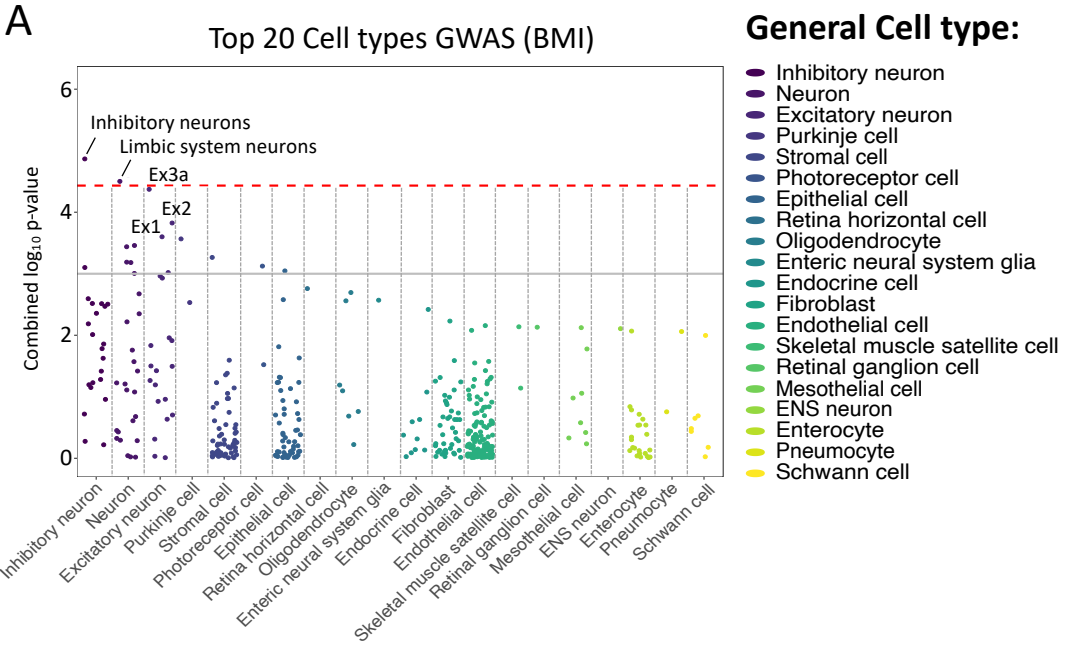
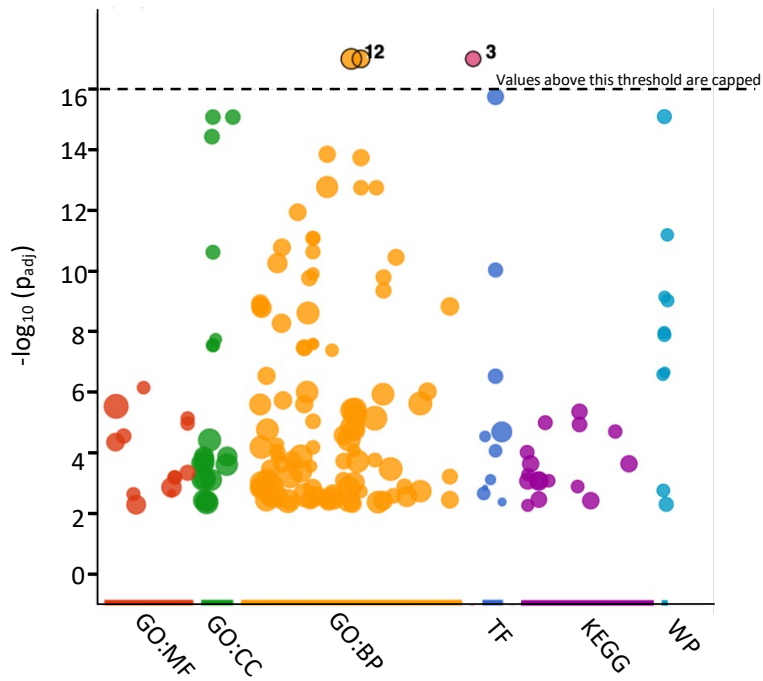


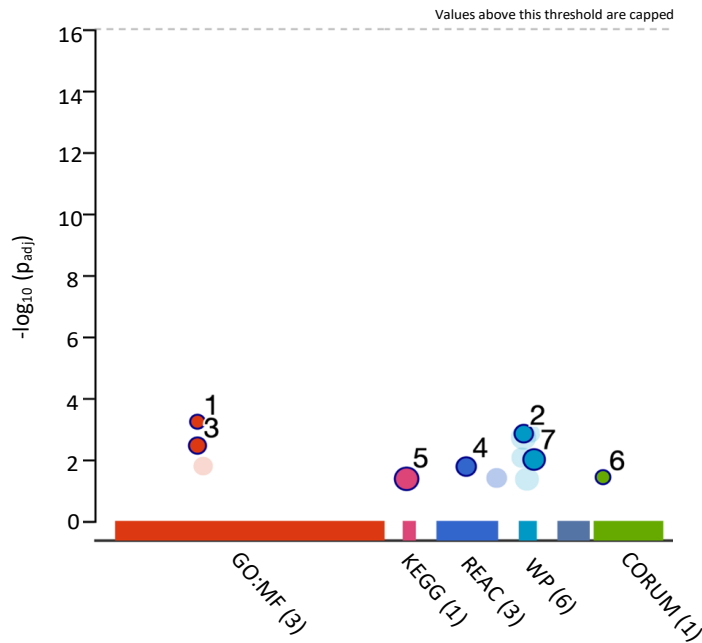
Figure S1

A



id	source	term_id	term_name	term_size	p_value
1	GO:BP	GO:0048878	chemical homeostasis	1014	1.7e-17
2	GO:BP	GO:0055088	lipid homeostasis	159	6.6e-22
3	KEGG	KEGG:04979	Cholesterol metabolism	51	1.0e-23

B



ID	Source	Term ID	Term Name	P <sub>adj</sub> (query_1)
1	GO:MF	GO:0016213	linoleoyl-CoA desaturase activity	5.928×10 <sup>-4</sup>
2	WP	WP:WP4586	Metabolism of alpha-linolenic acid	1.450×10 <sup>-3</sup>
3	GO:MF	GO:0016215	acyl-CoA desaturase activity	3.550×10 <sup>-3</sup>
7	WP	WP:WP4723	Omega-3 / omega-6 fatty acid synthesis	1.007×10 <sup>-2</sup>
4	REAC	REAC:R-HSA-2...	Linoleic acid (LA) metabolism	1.712×10 <sup>-2</sup>
6	CORUM	CORUM:724	MAF2 complex	3.780×10 <sup>-2</sup>
5	KEGG	KEGG:01040	Biosynthesis of unsaturated fatty acids	4.279×10 <sup>-2</sup>

Figure S2

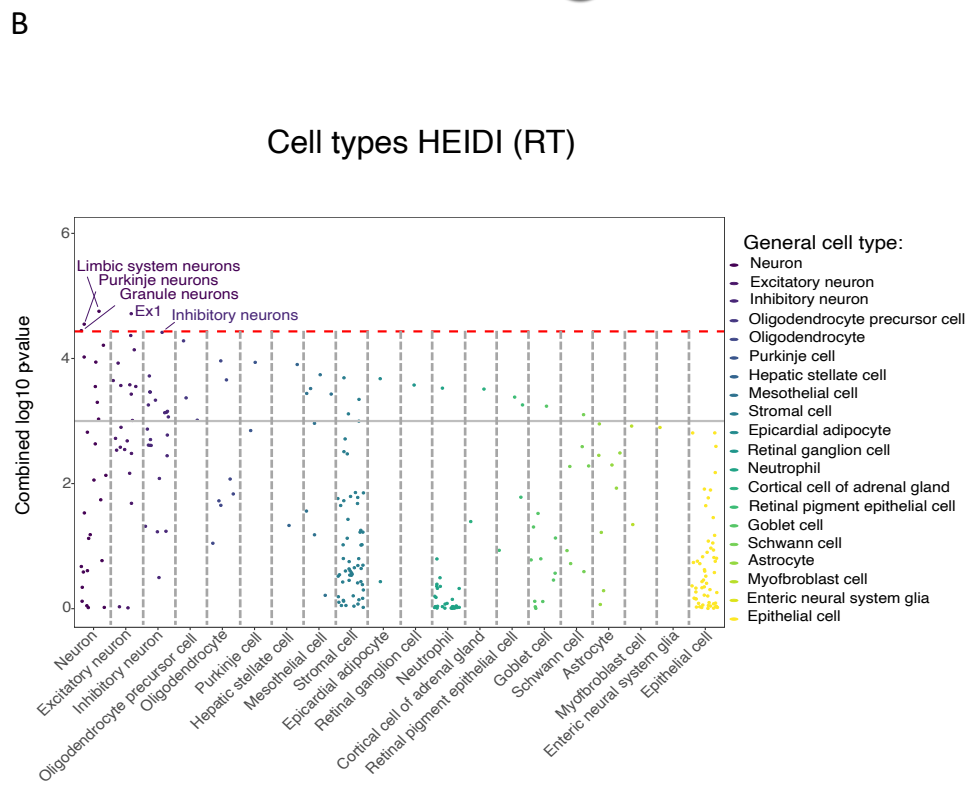
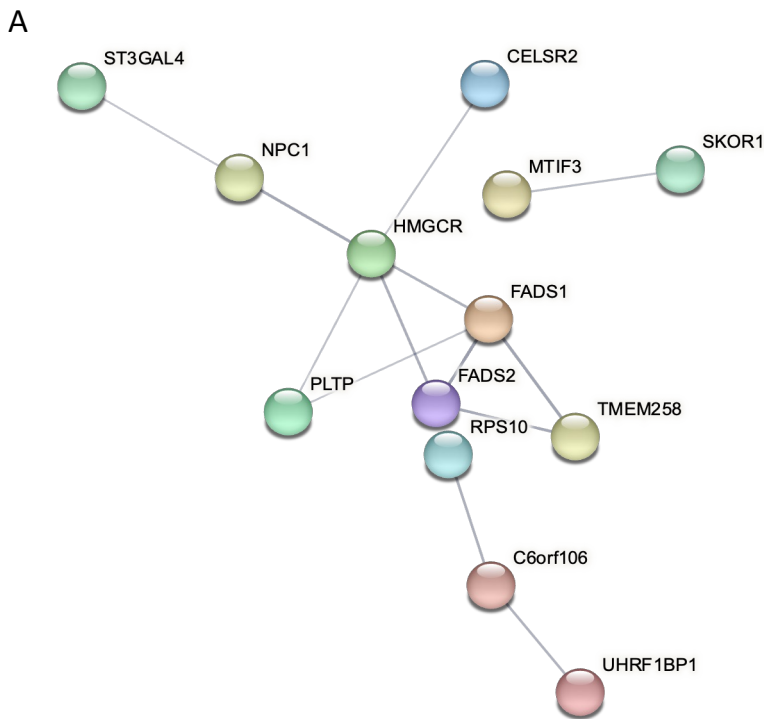


Figure S3

