Leveraging eQTLs to identify individual-level tissue of interest for a complex trait

Arunabha Majumdar[1*], Claudia Giambartolomei[1], Na Cai[2,3], Malika Kumar Freund[4], Tanushree Haldar[5], Tommer Schwarz[7], Jonathan Flint[6], and Bogdan Pasaniuc[1,4,7*]

[1]Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA

[2]Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton CB10 1SA, United Kingdom

[3]European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Hinxton CB10 1SD, United Kingdom

[4]Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA

[5]Institute for Human Genetics, University of California, San Francisco, CA, USA

[6]Department of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA

[7]Bioinformatics Interdepartmental Program, University of California, Los Angeles, Los Angeles, CA, USA

[*]Correspondence should be addressed to A.M. (arundar@ucla.edu) or B.P. (pasaniuc@ucla.edu)

Abstract

Genetic predisposition for complex traits is often manifested through multiple tissues of interest at different time points during their development. For example, the genetic predisposition for obesity could be manifested either through inherited variants that control metabolism through regulation of genes expressed in the brain, or through the control of fat storage by dysregulation of genes expressed in adipose tissue, or both. Here we describe a statistical approach that leverages tissue-specific expression quantitative trait loci (eQTLs) to prioritize the tissue of interest underlying the genetic predisposition of a given individual for a complex trait. Unlike existing approaches that prioritize tissues of interest for the trait in the population, our approach probabilistically quantifies the tissue-specific genetic contribution to the trait for a given individual. Through simulations using the UK Biobank genotype data, we show that our approach can predict the relevant tissue of interest accurately and can cluster individuals according to their tissue-specific genetic architecture. We analyze body mass index (BMI) and waist to hip ratio adjusted for BMI (WHRadjBMI) in the UK Biobank to identify individuals who have their genetic predisposition manifested through their brain versus adipose tissue, and adipose versus muscle tissue, respectively. Notably, we find that the individuals with a particular tissue of interest have specific phenotypic features beyond BMI and WHRadjBMI that distinguish them from random individuals in the data, demonstrating the role of tissue-specific genetic predisposition for these traits.

Introduction

Multiple clinical, pathologic, and molecular evidence suggest that many phenotypes and diseases show heterogeneity and can be viewed as a collection of multiple traits (i.e. subtypes) in the population [1–5]. Traditional subtype identification has relied on detecting biomarkers or subphenotypes that distinguish subsets of individuals in a biologically meaningful way. For example, individuals with Type 1 diabetes (T1D) can be sub-grouped by auto-antibody positivity [6]; breast cancer has two well-known subtypes, estrogen receptor positive and negative, [7–9]; and psychiatric disorder patients can have different severities [10]. With the advent of large scale genome-wide association studies (GWAS) that have robustly identified thousands of risk variants for complex traits, multiple approaches have investigated the use of genetic risk variants to define classes of individuals that show genetic heterogeneity across subtypes [11–16]. For example, autism can be subtyped by grouping together individuals with recurrent mutations in the same autism-associated gene [11,13]; Type 2 diabetes (T2D) can be subtyped using clusters of genetic variants previously associated with this disease [14]. Other examples include adiposity traits such as body mass index (BMI), waist-to-hip ratio (WHR), and WHR adjusted for BMI (WHRadjBMI), that can be subtyped based on genetic variants with distinct patterns of fat depots and metabolisms [15].

Genetic sub-typing offers an advantage over phenotypic sub-typing in that germline genetic characteristics are more stable than phenotypic characteristics of an individual [13,14], but comes with a drawback that genetic effect sizes are typically very small, thus requiring large sample sizes for prediction from genetics to be meaningful. A significant component of genetic susceptibility of complex traits is mediated through genetic control of gene expression in one or multiple tissues [17,18], with several studies highlighting the relevance of tissue-specific biological mechanisms underlying the pathogenesis of complex traits [19–22]. Such studies rely on integration of expression quantitative loci (eQTLs) with GWAS often in a tissue or cell-type specific manner to prioritize tissues and cell-types that are relevant for a given complex trait, and have often identified multiple tissues relevant to any given trait (e.g., brain and adipose for BMI [20,22]; muscle skeletal connective and adipose for WHRadjBMI [22]; liver, pancreas and thyroid tissues for total cholesterol [21], etc.). While these studies focused on prioritizing tissues or cell-types associated with a complex trait in the population, it remains an open question whether we can predict the tissue of interest for each individual in the population for a trait with multiple tissues of interest.

In this work, we present an approach that integrates tissue-specific eQTL with genetic association data for a complex trait to probabilistically assign a tissue of interest to the phenotype of each individual in the

study. We focus on traits where multiple tissues have been implicated (e.g., brain and adipose for BMI) and hypothesize that individuals have their genetic susceptibility mediated in a tissue specific manner (i.e. one group of individuals have their genetic predisposition through regulation in brain and another group of individuals through adipose). We propose eGST (eQTL-based Genetic Sub-Typer), an approach that estimates the posterior probability that an individual's phenotype can be assigned to a tissue based on individual-level genotype data of tissue-specific eQTLs and marginal phenotype data. eGST implements a Bayesian framework of mixture model by employing a computationally efficient maximum a posteriori (MAP) expectation-maximization (EM) algorithm to estimate the tissue-specific posterior probabilities per individual.

We perform extensive simulations using real genotypes from the UK Biobank and show that eGST accurately infers the simulated tissue of interest for each individual. We also show that a Bayesian framework of the mixture model performs better than the corresponding frequentist framework. By integrating expression data from the GTEx consortium [17,18], we apply eGST to two obesity related measures (BMI and WHRadjBMI) in the UK Biobank [23,24]. We consider brain and adipose tissues for BMI to identify 17,172 individuals with tissue-specific genetic predisposition, and muscle and adipose tissues for WHRadjBMI to identify 17,634 individuals with tissue-specific predisposition. Interestingly, the groups of individuals classified into each tissue show distinct genetic and phenotypic characteristics. 82 out of 106 phenotypes tested in the UK Biobank were differentially distributed between the BMI-adipose (BMI-brain) group of individuals and the remaining population, with 65 out of 82 remaining significant after adjusting for BMI. For example, diabetes proportion, various mental health phenotypes, alcohol intake frequency, and smoking status were differentially distributed between one or both of the tissue-specific subtype groups of BMI and the remaining population. Overall, our results suggest that tissue-specific eQTLs can be successfully utilized to prioritize the tissue of interest at an individual level in the study.

Results

Overview of methods

We start by depicting the main intuition underlying our hypothesis and model (Figure 1). For simplicity, consider two tissues of interest and assume that gene A is only expressed in tissue 1 whereas gene B is only expressed in tissue 2. The main hypothesis underlying our model is that genetic susceptibility of a complex trait for a given individual is mediated through regulation of either gene A or gene B, but not

through both. Having gene expression measurements in every individual at both genes in both tissues can be used to test this hypothesis. Unfortunately, gene expression measurements in large sample sizes such as the UK biobank are not typically available. To circumvent this, we use the top eQTL SNP for each gene as a proxy for the measured expression. In details, eGST takes as input the phenotype values and the genotype values at a set of variants known to be eQTLs for tissue-specific genes (Figure 1). Formally, the phenotype for individual i under the tissue of interest k is modeled as $y_i = \alpha_k + \boldsymbol{x}_{ki}^{'}\boldsymbol{\beta}_k + \epsilon_{ki}$, where α_k is the baseline tissue-specific trait mean, $\boldsymbol{x}_{ki}$ is the vector of normalized genotype values of individual i at the eQTL SNPs specific to tissue k, $\boldsymbol{\beta}_k$ are their effects on the trait, and ϵ_{ki} is a noise term, $i = 1, \ldots, n$ and $k = 1, \ldots, K$. For simplicity of exposition, we introduce indicator variables for each individual $C_i = k$ iff the individual i has its tissue of interest k. Thus, $\mathrm{P}(C_i = k) = w_k$ is the prior proportion of individuals for whom the phenotype has k^{th} tissue-specific genetic effect. We assume that the eQTL SNP sets across k tissues are non-overlapping and that each element in $\boldsymbol{\beta}_k$, the genetic effect of kth tissue-specific eQTLs on the trait, is drawn from $N(0, \sigma^2_{x_k})$. If $\sigma^2_{y_k}$ is the variance of the trait under $C_i = k$, then $\sigma^2_{x_k} = \frac{h^2_k \sigma^2_{y_k}}{m_k}$, where h^2_k is the heritability of the trait under $C_i = k$ due to k^{th} tissue-specific m_k eQTLs, and is termed as k^{th} tissue-specific subtype heritability. Under this mixture model, the likelihood of individual i takes the form: $f(y_i|\Theta) = \sum_{k=1}^{K} f(y_i|C_i = k, \Theta)\mathrm{P}(C_i = k) = \sum_{k=1}^{K} w_k \phi(y_i|\mu = \mu_{ik}, \sigma^2 = \sigma^2_{\epsilon_k})$, where $\phi(.|.)$ denotes the normal density, $\mu_{ik} = \alpha_k + x_{ki}^{'}\boldsymbol{\beta}_k$, $\Theta = (\theta_1, \ldots, \theta_K)$ with θ_k denoting k^{th} tissue-specific set of model parameters. We propose a Bayesian inference approach based on a maximum a posteriori (MAP) expectation maximization (EM) algorithm to estimate the posterior probability that the phenotype of individual i is mediated through the genetic effects of eQTLs specific to tissue k ($\mathrm{P}(C_i = k|X, Y)$).

Simulations

We performed simulations to assess the performance of eGST with respect to the accuracy of classifying the tissue of interest across individuals under various scenarios. We simulated phenotypes using the real genotype data from the UK Biobank, in which two tissue-specific eQTL effects generate the phenotype (see Methods). We evaluated the classification accuracy of eGST with respect to the variance explained in the trait specific to the two tissues by the two sets of tissue-specific SNPs' effects. As expected, the average area under the curve (AUC) increases with the tissue-specific subtype heritability, ranging from AUC of 50% when $h^2_1 = h^2_2 = 0\%$ to 95% when $h^2_1 = h^2_2 = 90\%$ (Figure 2). This is likely due to larger tissue-specific effect sizes (due to higher subtype heritability) inducing better differentiation between the tissue-specific genetic effects. Next, we assessed the performance of eGST compared to a variation of our approach that assumes

the parameters of the model are known (the performance obtained by this gold-standard strategy can be viewed as the maximum achievable under our proposed framework). We find that eGST loses $1.4\% - 3.9\%$ AUC on average compared to this strategy across all simulation scenarios considered (Figure 2, Figure S2, S3). We also considered a thresholding scheme on the tissue-specific posterior probabilities to balance total discoveries versus accuracy. As expected, the true discovery rate of classifying the tissue of interest increases with the posterior probability threshold but the proportion of discovery decreases (Figure S1).

We then explored the effect of other parameters on the classification accuracy. First, we found that increasing sample size n from $40,000$ to $100,000$ marginally increases the AUC by an average of 1% across different simulation scenarios (Table S2), which indicates that increasing sample size improves the overall classification accuracy. Second, we observed that as the number of causal SNPs explaining a fixed heritability of each subtype increases, the average AUC marginally decreases. For example, for a fixed subtype heritability explained, the average AUC for 2000 causal SNPs (1000 per tissue) is 1% higher than that for 3000 causal SNPs (1500 per tissue) across different choices of other simulation parameters (Table S3). Third, as the difference in the baseline tissue-specific mean of the trait across tissues increases, the classification accuracy also increases. For example, we find that the AUC increased from 60% (for no difference in tissue-specific phenotype means, $\alpha_1 = \alpha_2 = 0$) to 63% when $\alpha_1 = 0, \alpha_2 = 1$ (Table S4). We also explored the impact of the difference between the mean of tissue-specific genetic effect size distributions and observed that the classification accuracy improves compared to zero mean of both causal effects. For example, the AUC increases from 60% to 63% if we consider $E(\beta_{1j}) = -0.02$ and $E(\beta_{2j}) = 0.02, j = 1, \ldots, 1000$, instead of zero means of $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ (Table S5).

Finally, we explored the comparative performance of the MAP-EM algorithm under Bayesian framework and the EM algorithm under frequentist framework. Although both approaches yield similar AUC, MAP-EM performed better than EM with respect to the true discovery rate (TDR) at different posterior probability thresholds in nearly all of the simulation scenarios (Figure 3). MAP-EM offered an average of $0.05\% - 20\%$ higher TDR than EM across various posterior probability thresholds (Figure 3, S4, S5).

Inferring individual-level tissue of interest for BMI and WHRadjBMI

Having established in simulations that our approach is effective in correctly classifying the individual-level tissue of interest, we next analyzed BMI and WHRadjBMI, two phenotypes that are known to have multiple tissues of interest mediating their genetic susceptibility [20,22]. For BMI analysis, we used phenotype and genotype data for a selected group of $150,706$ individuals in the UK Biobank [23,24] at 1705 adipose specific

eQTLs and 1478 brain specific eQTLs (see Methods). BMI residuals were obtained after adjustment for age, sex, and 20 PCs of genetic ancestry. The group of $150,706$ individuals were selected from the full sample of $337,205$ unrelated white-British individuals (UK Biobank full release) by an iterative strategy using the Kolmogorov-Smirnov test of equality between two univariate distributions so that the BMI residual of the selected individuals follows a normal distribution (see Methods). Each SNP eQTL is the top cis-association of a tissue-specific expressed gene in the corresponding tissue [17, 18] (see Methods). At 65% threshold of tissue-specific subtype posterior probability, only 11% of the individuals where assigned a tissue; eGST classified the genetic susceptibility on the BMI of $12,543$ individuals through adipose eQTLs and for $4,629$ individuals through brain eQTLs (Table S1). Individuals classified to each of the tissues are distributed across different bins of BMI (Table S6).

For WHRadjBMI, we included 953 adipose subcutaneous (abbreviated and referred as AS in the following) tissue-specific eQTLs and 1052 muscle skeletal connective (abbreviated as MS) tissue-specific eQTLs; and WHRadjBMI residuals (adjusting WHRadjBMI for age, sex, and top 20 PCs) for a selected group of $271,818$ individuals chosen by the above-mentioned iterative strategy. Similarly to the BMI analysis, the tissue of interest for WHRadjBMI of a small percentage (6.5%) of individuals were classified (Table S1). Individuals assigned to the two tissues are both spread across different bins of WHRadjBMI (Table S7).

We permuted the phenotype data across individuals while keeping the eQTL assignment to tissues fixed as it is in the original data. For BMI, the average number of individuals classified as a tissue-specific subtype (based on 65% threshold of subtype posterior probability) across 500 random permutations of the phenotype was 12,321 compared to 17,172 individuals classified as real adipose or brain specific subtype of BMI in the original data. For WHRadjBMI, the average number of individuals classified as a tissue-specific subtype across 500 random permutations of the phenotype was 5,499 compared to 17,634 individuals classified as real AS and MS specific subtype groups of WHRadjBMI.

To confirm that eGST identified groups of individuals with different genetic basis, we contrasted SNP effects of adipose and brain-specific eQTLs on BMI in those individuals assigned to the adipose or the brain specific subtype. As expected, we find that in individuals classified as having the brain-specific subtype in their genetic contribution to BMI, the magnitude of the effect size of a brain eQTL SNP is larger than the corresponding effect size magnitude of an adipose eQTL SNP (Wilcoxon rank sum (WRS) right tail test p-value $< 2 \times 10^{-16}$ (Table 1)). The opposite is true for individuals classified as having the adipose-specific subtype of BMI. We also find that the magnitude of effect size of adipose eQTLs are larger in adipose-specific individuals than that in brain-specific individuals, and we find statistical evidence supporting the analogous

hypothesis about brain eQTLs, brain-specific, and adipose-specific individuals of BMI (Table 1). We observe the same pattern in our analogous analysis for WHRadjBMI (Table 1).

Phenotypic characteristics of individuals with a prioritized tissue

Next, we explored the phenotypic characteristics of the individuals assigned with a prioritized tissue. We considered 106 phenotypes in the UK Biobank and tested each one for being differentially distributed (heterogeneous) between individuals of each tissue-specific subtype and the remaining population (see Methods). In aggregate for BMI, 44 quantitative traits and 38 qualitative traits (total 82 among 106) were significantly heterogeneous between at least one of the BMI-adipose or BMI-brain specific groups versus the remaining population. None of these 106 traits was found to be differentially distributed between a random set of individuals from the population (with the same size as a tissue-specific subtype group) and the remaining population (see Methods). 23 quantitative and 11 categorical traits showed heterogeneity in both the adipose group versus the population, and the brain group versus population. We found 13 quantitative and 21 categorical traits heterogeneous for individuals in the adipose group but not the brain group, and found 8 quantitative and 6 categorical traits heterogeneous for individuals in the brain group but not the adipose group (Table S8, S10, Table 2).

For example, *hemoglobin concentration* and *snoring* were heterogeneous for both the adipose and brain groups, *platelet count* was heterogeneous only for the adipose group, and *age completed full time education* only for the brain group (Table S8, S10). We observe that *hemoglobin concentration* was lower in individuals from both groups as compared to the population (Figure 4), whereas *white blood cell leukocyte count* was relatively higher in individuals of the adipose but lower in those with the brain tissue when compared to the population (Figure 4). Among binary traits, *diabetes*, *snoring* and *fed up feeling* were more prevalent in those from the adipose subtype and less prevalent in brain group compared to the population (Figure 5), while many neuroticism related binary traits (e.g., *guilty feelings*) were more prevalent among individuals from both adipose and brain subtypes than in the population. Of note, when the tissue-specific relative change of some traits (see Methods) were in the same direction across tissues, they were of different magnitude for a majority of the traits (Figure 4, 5). For example, the relative change were 1.5% and 15.1% for *mean sphered cell volume*. Similar to BMI, we observed phenotypic heterogeneity across individuals with AS (MS) as the prioritized tissue for WHRadjBMI (Figure S6, S7, Table S9, S11, S12, Supplementary Note).

Since BMI itself was differentially distributed between the individuals of the adipose subtype as well as brain subtype compared to the remaining population, we investigated whether the heterogeneity of 81 non-

BMI traits (Table S8 and S10, Table 2) were induced due to BMI heterogeneity (see Methods). After BMI adjustment, 65 (out of 81) traits remained heterogeneous (39 quantitative traits (Figure 4, Table S13) and 26 qualitative traits (Figure 5, Table 2)) consistent with unique phenotypic characteristics of these individuals beyond the main phenotype effect. As previously mentioned, some phenotypes have the same direction in BMI-adjusted tissue-specific relative change (see Methods) in both adipose and brain subtypes compared to the population, while others have opposite effects (Figure 4, Table S13). Since we used linear regression while evaluating BMI-adjusted tissue-specific relative change of heterogeneous non-BMI quantitative traits, we also investigated a model-free BMI random matching strategy. We assessed the magnitude of relative change of a trait between individuals with the adipose (or brain) subtype and a group of BMI-matched random individuals drawn from the population (see Methods). For example, the magnitude of primary adipose-specific relative change (prior to BMI matching) for *hemoglobin concentration* (*standing height*) decreased from 9% (12%) to 6% (5%) after BMI matching (Table S15). Of note, it is very difficult to exactly match BMI between a tissue-specific subtype group and the corresponding random group of individuals, because bins of BMI in the tail of its distribution contain very few individuals (Table S6), the majority of whom were assigned to a tissue-specific subtype. We observed the same pattern in the results from analogous analyses for WHRadjBMI (Table S14, S16, Supplementary Note).

To better understand the phenotypic characteristics of the individuals classified to a specific tissue, we performed following two experiments. First, we shuffled the tissue-specific eQTL SNPs between tissues to create an artificial tissue-specific eQTL set and implemented eGST to identify groups of individuals having subtype specific to the artificial tissues (see Methods). We found that the mean of artificial tissue-specific means of a quantitative trait (found primarily heterogeneous between adipose and/or brain specific group versus the remaining population (Table S8)) across eQTL shuffles was close to the population mean of the trait but significantly further from the original corresponding tissue-specific trait mean (Table S17 for BMI and Table S18 for WHRadjBMI). For example, for *high light scatter reticulocyte percentage*, the mean of pseudo tissue-specific means over the random eQTL shuffles is 0.39 for adipose and 0.4 for brain, which are both very close to the population mean of 0.4 (Table S17), but significantly different from the original adipose-specific mean 0.44 ($P < 10^{-100}$) and brain-specific mean 0.34 ($P < 10^{-100}$). The same pattern was observed for the primary phenotypes BMI (Table S17) and WHRadjBMI (S18) themselves. Second, we permuted the phenotype data across individuals while keeping the eQTL assignment to tissues fixed as it is in the original data. As before, we also observed that the mean of tissue-specific means of a trait across random phenotype permutations was close to the population mean of the trait but significantly further from

the original corresponding tissue-specific trait mean (Table S19 for BMI and Table S20 for WHRadjBMI).

Computational efficiency

The MAP-EM algorithm underlying eGST is computationally efficient. 70 MAP-EM iterations in the BMI analysis ($150K$ individuals with 1705 adipose-specific eQTLs and 1478 brain-specific eQTLs) had a runtime of 30 minutes and yielded a log likelihood improvement of 5×10^{-8} in the final iteration. Though we ran eGST for a pair of tissues only considering the top eQTL per gene, it is computationally feasible to analyze larger datasets considering more eQTLs and multiple tissues simultaneously.

Discussion

We proposed a new method to prioritize tissue of interest for a complex trait for every individual in the study, integrating genotype and phenotype data and an external expression panel data. We applied our approach to infer individual-level tissue of interest for BMI and WHRadjBMI in the UK Biobank, integrating expression data in brain, adipose, and muscle tissues from the GTEx consortium, previously shown to be enriched in heritability for these phenotypes [20,22]. Interestingly, multiple metabolic traits, neuropsychiatric traits, and other traits attained significant differences between the tissue-specific groups of individuals and the remaining population, suggesting a biologically meaningful interpretation for these groups of individuals. Even after adjusting these traits for the primary phenotype, the majority of the traits that were initially heterogeneous remained differentially distributed between a tissue-specific group and the remaining population.

Although in this work we demonstrated the utility of eGST for BMI and WHRadjBMI, the MAP-EM algorithm underlying eGST is general and can be applied to any number of tissues. We note that our model can alternatively be viewed as an approach to assign individuals' phenotypes to a collection of tissues that are biologically important for the trait based on tissue-specific polygenic risk score [25].

We conclude with several caveats and limitations of our work and opportunities for future improvement. First, we investigated the utility of eGST using adipose and brain specific eQTLs for BMI [20,22], and using adipose and muscle specific eQTLs for WHRadjBMI [22]. However, the true tissues of interest could be different. Second, it is challenging to identify a set of SNPs that adequately represent a tissue-specific genetic architecture. Other types of tissue-specific QTLs (e.g., methylation QTLs, histone QTLs, splicing QTLs, etc. [26]) could be combined with eQTLs to create a set of SNPs that better represent a tissue-specific genetic architecture. In this case, the accuracy of inference will depend on how efficiently we can create the

set of tissue-specific expressed genes and corresponding set of tissue-specific QTLs. Third, we developed the model for continuous traits, meaning that to extend the method for case-control data, we would need to use a logistic regression likelihood. Another future methodological investigation is to extend the model under penalized regression framework; if the number of SNPs characterizing the genetic architecture of a tissue becomes large and the ratio between the number of individuals and number of SNPs in the data decreases, model fitting issues can arise. Finally, Fig. 1b motivates that if gene expression data across tissues are available, it is possible to use the expression data itself to identify expression subtypes of the trait. However, since expression data is not available in most GWAS cohorts, an alternative avenue will be to impute the genetically regulated component of gene expression, e.g., using PrediXcan [27], EpiXcan [28] and identify tissue of interest based on imputed gene expression. A possible advantage of such approach will be that all cis eQTLs can be unified to impute tissue-specific expression (instead of top few eQTLs only). Another limitation of our analysis was that we focused on the GTEx data which does not have a large sample size across tissues.

URLs

Tissue-specific expressed genes: `https://data.broadinstitute.org/alkesgroup/LDSCORE/LDSC_SEG_ldscores/`

GTEx portal: `https://gtexportal.org/home/`

UK Biobank: `https://www.ukbiobank.ac.uk`

PLINK: `https://www.cog-genomics.org/plink2`

Acknowledgments

This research was conducted using the UK Biobank Resource under applications 24129 and 33297. We thank the participants of UK Biobank for making this work possible. We thank Huwenbo Shi, Kathryn Burch, Nicholas Mancuso, Hilary Finucane, Eran Halperin, Michael Gandal for helpful discussions relating to this work. This work was funded in part by the National Institutes of Health (NIH) under awards R01HG009120, R01MH115676, R01HG006399, U01CA194393, T32NS048004, T32MH073526 and T32HG002536.

Methods

Model

For simplicity, we describe the model assuming $K = 2$ tissues of interest. Suppose for n unrelated individuals we have phenotype data $Y = (y_1, \ldots, y_n)$ and expression data for two sets of tissue-specific expressed genes E_1, E_2 characterizing two tissues. We define a subtype indicator variable C such that for an individual, $C = 1$ implies that the genetic susceptibility of the phenotype of the individual is mediated through first tissue, and $C = 2$ implies that the genetic susceptibility is mediated through the second tissue (Figure 1). The model for individual i based on tissue-specific expression is given by:

$$\begin{aligned} y_i &= a_1 + e_{1i}^{'}\boldsymbol{b}_1 + e_{2i}^{'}\boldsymbol{d}_1 + \varepsilon_{1i} \text{ if } C_i = 1 \\ &= a_2 + e_{1i}^{'}\boldsymbol{d}_2 + e_{2i}^{'}\boldsymbol{b}_2 + \varepsilon_{2i} \text{ if } C_i = 2 \end{aligned} \tag{1}$$

Here a_1 and a_2 represent the baseline tissue-specific trait means, and e_{1i} (or e_{2i}) denotes the vector of expression values of the first (or second) tissue-specific set of genes for individual i. Under $C_i = 1$, $\boldsymbol{b}_1$ and $\boldsymbol{d}_1$ denote the effects of expression of the two gene sets in first tissue on the trait. When $C_i = 1$, we assume that the expression of 2nd tissue-specific genes (low expressed in first tissue) have no effect ($\boldsymbol{d}_1 = \boldsymbol{0}$) on the phenotype. Similarly, under $C_i = 2$, we assume that the first tissue-specific expressed genes (with low expression in the second tissue) have no effect ($\boldsymbol{d}_2 = \boldsymbol{0}$) on the phenotype of the individual. Thus, we obtain the following model under these assumptions:

$$\begin{aligned} y_i &= a_1 + e_{1i}^{'}\boldsymbol{b}_1 + \varepsilon_{1i} \text{ if } C_i = 1 \\ &= a_2 + e_{2i}^{'}\boldsymbol{b}_2 + \varepsilon_{2i} \text{ if } C_i = 2 \end{aligned} \tag{2}$$

Since expression datasets in general have limited sample size and are not available in large GWAS cohorts (e.g., UK Biobank), we use genotypes of tissue-specific eQTLs as a proxy for the expressions of the corresponding tissue-specific genes. Suppose in a GWAS cohort, we have phenotype data and genotype data for the two sets of tissue-specific eQTL SNPs corresponding to the two sets of genes specifically expressed in two tissues, one comprising m_1 SNPs and the other comprising m_2 SNPs. Then,

$$
\begin{aligned}
y_i &= \alpha_1 + x_{1i}^{'}\boldsymbol{\beta}_1 + \epsilon_{1i} \text{ if } C_i = 1 \\
&= \alpha_2 + x_{2i}^{'}\boldsymbol{\beta}_2 + \epsilon_{2i} \text{ if } C_i = 2
\end{aligned} \tag{3}
$$

Here, α_1 and α_2 denote the baseline tissue-specific trait means, x_{1i} is the vector of normalized genotype values of the individual at the first tissue-specific eQTL SNPs. $\boldsymbol{\beta}_1$ denotes the effect of the first tissue-specific eQTL SNPs on the first tissue-specific subtype of the trait. Denote $\boldsymbol{\beta}_1 = (\beta_{11}, \beta_{12} \ldots, \beta_{1m_1})$, where β_{1j} is the effect of jth first tissue-specific eQTL SNP on the trait under $C_i = 1$. Similarly, we define $\boldsymbol{\beta}_2 = (\beta_{21}, \beta_{22}, \ldots, \beta_{2m_2})$. The random errors are distributed as $\epsilon_{1i} \sim \mathrm{N}(0, \sigma^2_{\epsilon_1})$ and $\epsilon_{2i} \sim \mathrm{N}(0, \sigma^2_{\epsilon_2})$. We note that in our context, the mixture model is identifiable because the mean term depends on the genotype vector of the set of tissue-specific eQTLs, which is distinct across tissues.

Prior distributions

For two tissues, the subtype indicator variable has the following distribution: $P(C_i = 1) = w_1$ and $P(C_i = 2) = w_2$, $w_1 + w_2 = 1$. The first tissue-specific genetic effects are distributed as $\beta_{1j} \sim N(0, \sigma^2_{x_1})$; $j = 1, \ldots, m_1$ (i.i.d. across j). We define h^2_1 as the first tissue-specific subtype heritability, $h^2_1 = \frac{m_1 \sigma^2_{x_1}}{\sigma^2_{y_1}}$, where $\sigma^2_{y_1}$ is the variance of the trait under $C_i = 1$. Thus, h^2_1 is the heritability of the trait under $C_i = 1$ due to the first tissue-specific m_1 eQTL SNPs. We also assume: $\alpha_1 \sim N(0, \sigma^2_\alpha)$ and $\alpha_2 \sim N(0, \sigma^2_\alpha)$, with fixed σ^2_α.

For two tissues, we assume that $(w_1, w_2) \sim \mathrm{Beta}(s_1, s_2)$ which will be a Dirichlet distribution for more than two tissues. We consider fixed values of $s_1 = s_2 = 1$. Next, we assume: $\sigma^2_{x_1}$ and $\sigma^2_{x_2} \sim$ Inverse-Gamma(a_x, b_x); $\sigma^2_{\epsilon_1}$ and $\sigma^2_{\epsilon_2} \sim$ Inverse-Gamma(a_ϵ, b_ϵ). We choose fixed values of $a_x, b_x, a_\epsilon, b_\epsilon$ such that in the prior expectation, 5% of total variance of each subtype (under $C_i = 1$ or 2) is explained by the corresponding set of tissue-specific eQTL SNPs and 95% of the variance remains unexplained.

Inference procedure

Our main objective is to learn $C_1, \ldots, C_n$ from the data. Under the Bayesian framework that we previously described, we implemented the maximum a posteriori (MAP) expectation-maximization (EM) algorithm (Algorithm 1) to estimate the posterior probability that the genetic susceptibility of an individual's phenotype is mediated through eQTLs specific to a tissue. We note that it is also possible to consider a frequentist framework of the mixture model, i.e., instead of having a distribution, (w_1, w_2), (α_1, α_2), $(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$, $(\sigma^2_{\epsilon_1}, \sigma^2_{\epsilon_2})$ can be assumed to have a fixed unknown true value. Since $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ are assumed to be fixed, no σ^2_x term

appears in the model. We implemented an EM algorithm to estimate the subtype posterior probability across individuals under the frequentist framework. Next, we outline the MAP-EM algorithm (Algorithm 1) that implements the Bayesian framework which eGST is based on and the EM algorithm (Algorithm 2) that implements the frequentist framework of the mixture model [29,30].

We describe the algorithms for a general K number of tissues, $K \geq 2$. For individual i and tissue k, $P(C_i = k) = w_k$; $i = 1, \ldots, n$; $k = 1, \ldots, K$; $\sum_k w_k = 1$. Denote k^{th} tissue-specific set of parameters by $\theta_k = (w_k, \alpha_k, \boldsymbol{\beta}_k, \sigma^2_{x_k}, \sigma^2_{\epsilon_k})$ and full set of parameters by $\Theta = (\theta_1, \ldots, \theta_K)$. In our context, for $i = 1, \ldots, n$,

$$f(y_i|\Theta) = \sum_{k=1}^{K} w_k \phi(y_i|\mu = \mu_{ik}, \sigma^2 = \sigma^2_{\epsilon_k}); \text{ with } \mu_{ik} = \alpha_k + x'_{ki}\boldsymbol{\beta}_k, \tag{4}$$

where $\phi(.|.)$ denotes the normal density. Thus the full data log-likelihood conditioned on Θ is given by: $\log f(Y|\Theta) = \sum_{i=1}^{n} \log f(y_i|\Theta)$. The prior log-likelihood of $(C_1, \ldots, C_n)$ is: $\sum_{i=1}^{n} \log f(C_i)$, where $f(C_i)$ is given by: $P(C_i = k) = w_k$. For $k = 1, \ldots, K$, the prior distribution of kth tissue-specific parameters θ_k has the following hierarchical structure: $\boldsymbol{\beta}_k|\sigma^2_{x_k} \sim \mathrm{N}_{m_k}(\mathbf{0}, \sigma^2_{x_k} I_{m_k})$, $\sigma^2_{x_k} \sim$ Inverse-Gamma(a_k, b_k), $\sigma^2_{\epsilon_k} \sim$ Inverse-Gamma(a_ϵ, b_ϵ), $\alpha_k \sim N(0, \sigma^2_\alpha)$, $(w_1, w_2) \sim$ Beta(s_1, s_2). For a given k, only the prior of $\boldsymbol{\beta}_k$ depends on $\sigma^2_{x_k}$, but the other parameters are independently distributed in the prior; and $\theta_1, \ldots, \theta_K$ are independently distributed.

Define the posterior probability that the phenotype of individual i be assigned tissue k as:

$$\gamma_{ik} = P(C_i = k|y_i, \Theta) = \frac{w_k \phi(y_i|\mu_{ik}, \sigma^2_{\epsilon_k})}{\sum_{t=1}^{K} w_t \phi(y_i|\mu_{it}, \sigma^2_{\epsilon_t})}; \quad \sum_{k=1}^{K} \gamma_{ik} = 1 \tag{5}$$

Thus, the choice of tissue of interest across individuals is quantified by $\Gamma = \{\gamma_{ik}; i = 1, \ldots, n; k = 1, \ldots, K\}$. Next, we define the total membership weight of kth tissue-specific subtype: $n_k = \sum_{i=1}^{n} \gamma_{ik}$, $\sum_k n_k = n$.

The main component in the expectation-maximization algorithm which we maximize is given by: $Q(\Theta|\Theta^{(r)}) = \sum_{i=1}^{n} Q_i(\Theta|\Theta^{(r)})$, where the conditional expectation $Q_i(\Theta|\Theta^{(r)}) = \mathrm{E}_{C_i|y_i,\Theta^{(r)}}\{\log f(y_i, C_i|\Theta)\} = \sum_{k=1}^{K} \gamma_{ik}[\log w_k + \log \phi(y_i|\mu_{ik}, \sigma^2_{\epsilon_k})]$. To obtain $\Theta^{(r+1)}$, we maximize $\{Q(\Theta|\Theta^{(r)}) + \log f(\Theta)\}$ in the MAP-EM algorithm implementing the Bayesian framework of the mixture model, and maximize only $Q(\Theta|\Theta^{(r)})$ in the EM algorithm implementing the frequentist framework of the mixture model [29,30].

Algorithm 1 Maximum a posteriori (MAP) expectation maximization (EM) algorithm under Bayesian framework of the mixture model

1: **Initialization**: For $k = 1, \ldots, K$, choose $\sigma^{2^{(0)}}_{\epsilon_k} = 0.95 \times var(Y)$, $\sigma^{2^{(0)}}_{x_k} = \frac{0.05 \times var(Y)}{m_k}$, $\alpha^{(0)}_k = 0$, $w^{(0)}_k = \frac{1}{K}$; and simulate $\boldsymbol{\beta}^{(0)}_k$ from $\mathrm{N}(0, \sigma^{2^{(0)}}_{x_k} I_{m_k})$. Compute the initial log-likelihood: $\mathrm{logL}^{(0)} = \frac{1}{n}\sum_{i=1}^{n} \log\Big(\sum_{k=1}^{K} w^{(0)}_k \phi(y_i | \mu = \mu^{(0)}_{ik}, \sigma^2 = \sigma^{2^{(0)}}_{\epsilon_k})\Big)$. Next, for iteration $r = 0, 1, \ldots$

2: **E-step**: Compute:

$$\gamma^{(r)}_{ik} = \frac{w^{(r)}_k \phi(y_i | \mu^{(r)}_{ik}, \sigma^{2^{(r)}}_{\epsilon_k})}{\sum_{t=1}^{K} w^{(r)}_t \phi(y_i | \mu^{(r)}_{it}, \sigma^{2^{(r)}}_{\epsilon_t})};\ \mu_{ik} = \alpha_k + x^{'}_{ki}\boldsymbol{\beta}_k;\ i = 1, \ldots, n;\ k = 1, \ldots, K$$

$$n^{(r)}_k = \sum_{i=1}^{n} \gamma^{(r)}_{ik},\ k = 1, \ldots, K$$

3: **M-step**: For $k = 1, \ldots, K$, update:

$$w^{(r+1)}_k = \frac{n^{(r)}_k}{n}$$

$$\alpha^{(r+1)}_k = \frac{1}{n^{(r)}_k + \frac{\sigma^{2^{(r)}}_{\epsilon_k}}{\sigma^2_\alpha}} \times \sum_{i=1}^{n} \gamma^{(r)}_{ik}(y_i - x^{'}_{ki}\boldsymbol{\beta}^{(r)}_k)$$

$$\boldsymbol{\beta}^{(r+1)}_k = \Big(\sum_{i=1}^{n} \gamma^{(r)}_{ik} x_{ki} x^{'}_{ki} + \frac{\sigma^{2^{(r)}}_{\epsilon_k}}{\sigma^{2^{(r)}}_{x_k}} I_{m_k}\Big)^{-1} \sum_{i=1}^{n} \gamma^{(r)}_{ik}(y_i - \alpha^{(r)}_k) x_{ki}$$

$$\sigma^{2^{(r+1)}}_{\epsilon_k} = \frac{1}{n^{(r)}_k + 2a_\epsilon + 1}\Big[2b_\epsilon + \sum_{i=1}^{n} \gamma^{(r)}_{ik}(y_i - \mu^{(r)}_{ik})^2\Big]$$

$$\sigma^{2^{(r+1)}}_{x_k} = \frac{\boldsymbol{\beta}^{'(r)}_k \boldsymbol{\beta}^{(r)}_k + 2b_x}{m_k + 2a_x + 1}$$

4: **Convergence check**: Compute the new log-likelihood:

$$\mathrm{logL}^{(r+1)} = \frac{1}{n}\sum_{i=1}^{n} \log\Big(\sum_{k=1}^{K} w^{(r+1)}_k \phi(y_i | \mu^{(r+1)}_{ik}, \sigma^{2^{(r+1)}}_{\epsilon_k})\Big)$$

Return to step 2, if $|\mathrm{logL}^{(r+1)} - \mathrm{logL}^{(r)}| > \delta$, for a pre-fixed threshold δ (e.g. 10^{-5}).

Algorithm 2 Expectation Maximization (EM) algorithm under frequentist framework of the mixture model

1: **Initialization**: For $k = 1, \ldots, K$, choose $\sigma^{2^{(0)}}_{\epsilon_k} = 0.95 \times var(Y)$, $\sigma^{2^{(0)}}_{x_k} = \frac{0.05 \times var(Y)}{m_k}$, $\alpha^{(0)}_k = 0$, $w^{(0)}_k = \frac{1}{K}$; and simulate $\boldsymbol{\beta}^{(0)}_k$ from $\mathrm{N}(0, \sigma^{2^{(0)}}_{x_k} I_{m_k})$. Compute the initial log-likelihood: $\mathrm{logL}^{(0)} = \frac{1}{n}\sum_{i=1}^{n} \log\Big(\sum_{k=1}^{K} w^{(0)}_k \phi(y_i|\mu = \mu^{(0)}_{ik}, \sigma^2 = \sigma^{2^{(0)}}_{\epsilon_k})\Big)$. Next, for iteration $r = 0, 1, \ldots$

2: **E-step**: Compute:

$$\gamma^{(r)}_{ik} = \frac{w^{(r)}_k \phi(y_i|\mu^{(r)}_{ik}, \sigma^{2^{(r)}}_{\epsilon_k})}{\sum_{t=1}^{K} w^{(r)}_t \phi(y_i|\mu^{(r)}_{it}, \sigma^{2^{(r)}}_{\epsilon_t})};\ \mu_{ik} = \alpha_k + x^{'}_{ki}\boldsymbol{\beta}_k;\ i = 1, \ldots, n;\ k = 1, \ldots, K$$

$$n^{(r)}_k = \sum_{i=1}^{n} \gamma^{(r)}_{ik},\ k = 1, \ldots, K$$

3: **M-step**: For $k = 1, \ldots, K$, update:

$$w^{(r+1)}_k = \frac{n^{(r)}_k}{n}$$

$$\alpha^{(r+1)}_k = \frac{1}{n^{(r)}_k} \times \sum_{i=1}^{n} \gamma^{(r)}_{ik}(y_i - x^{'}_{ki}\boldsymbol{\beta}^{(r)}_k)$$

$$\boldsymbol{\beta}^{(r+1)}_k = \Big(\sum_{i=1}^{n} \gamma^{(r)}_{ik} x_{ki} x^{'}_{ki}\Big)^{-1} \sum_{i=1}^{n} \gamma^{(r)}_{ik}(y_i - \alpha^{(r)}_k) x_{ki}$$

$$\sigma^{2^{(r+1)}}_{\epsilon_k} = \frac{1}{n^{(r)}_k} \sum_{i=1}^{n} \gamma^{(r)}_{ik}(y_i - \mu^{(r)}_{ik})^2$$

4: **Convergence check**: Compute the new log-likelihood:

$$\mathrm{logL}^{(r+1)} = \frac{1}{n}\sum_{i=1}^{n} \log\Big(\sum_{k=1}^{K} w^{(r+1)}_k \phi(y_i|\mu^{(r+1)}_{ik}, \sigma^{2^{(r+1)}}_{\epsilon_k})\Big)$$

Return to step 2, if $|\mathrm{logL}^{(r+1)} - \mathrm{logL}^{(r)}| > \delta$, for a pre-fixed threshold δ.

Simulation design and choice of parameters

Consider n individuals and two non-overlapping sets of m_1 SNPs and m_2 SNPs representing eQTL SNP sets specific to two tissues. We chose the SNPs on chromosome $8-17$ from the array SNPs in the UK Biobank (UKB). We pruned for LD between the SNPs such that two consecutive SNPs (on a chromosome) included in a SNP set had $r^2 < 0.25$ (based on UKB in-sample LD). Each SNP had MAF $> 1\%$ and satisfied Hardy Weinberg Equilibrium (HWE). We collected genotype data at both sets of SNPs for n individuals that were randomly selected from 337,205 white-British individuals in the UKB.

Let $w = (w_1, w_2)$ denote the proportions of individuals in the sample assigned to the two tissues where $(100 \times w_1)\%$ individuals are assigned as the first tissue-specific subtype and $(100 \times w_2)\%$ individuals are assigned as the second tissue-specific subtype. We assume that m_k SNPs explain $(100 \times h_k^2)\%$ of the total variance of kth tissue-specific subtype, $k = 1, 2$. So h_k^2 is the heritability of kth tissue-specific subtype of the trait due to m_k SNPs representing k^{th} tissue-specific eQTLs, $k = 1, 2$. Thus, if first subtype of Y has a total variance $\sigma_{y_1}^2$, we draw each element of $\boldsymbol{\beta}_1$ as: $\beta_{1j} \sim \mathrm{N}(0, \frac{h_1^2 \sigma_{y_1}^2}{m_1})$, $j = 1, \ldots, m_1$. Similarly we simulate $\boldsymbol{\beta}_2$, the genetic effect of second set of m_2 SNPs on the second subtype of Y. For simplicity, we assume $\sigma_{y_1}^2 = \sigma_{y_2}^2 = \sigma_y^2$, but the performance of eGST remains similar for other choices of this parameter. If the genetic susceptibility of an individual's phenotype was assigned to be mediated through first tissue, we simulated the phenotype as: $y = \alpha_1 + x_1^{'}\boldsymbol{\beta}_1 + \epsilon_1$, where x_1 is the normalized genotype values of the individual at the first set of SNPs. While simulating the phenotype, we normalized the genotypes at each of first tissue-specific m_1 SNPs only based on the individuals assigned to the first tissue-specific subtype. However, when applying eGST on a simulated dataset, we normalized the genotypes at each SNP based on all n individuals in the sample, because the tissue of interest across individuals are unknown. The random error components have the following distribution: $\epsilon_1 \sim \mathrm{N}(0, (1 - h_1^2)\sigma_{y_1}^2)$ and $\epsilon_2 \sim \mathrm{N}(0, (1 - h_2^2)\sigma_{y_2}^2)$.

We varied the choice of parameters to evaluate eGST in various simulation scenarios. We chose $\sigma_y^2 = 10$, and initially assumed $\alpha_1 = \alpha_2 = 0$ and simulate $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ from zero-mean normal distributions. We considered all possible combinations of (w_1, w_2) where $w_1 \in (\frac{1}{2}, \frac{1}{3})$ and $w_2 \in (\frac{1}{2}, \frac{1}{3})$, and all possible combinations of (h_1^2, h_2^2), where h_1^2 and $h_2^2 \in (10\%, 20\%, 30\%, 40\%, 50\%)$. We also considered two unrealistic scenarios of null and high subtype heritability: $h_1^2 = h_2^2 = 0\%$ and $h_1^2 = h_2^2 = 90\%$ to evaluate if eGST is performing as expected in these extreme scenarios. We chose (m_1, m_2) with $m_1 = 1000, 1500$ and $m_2 = 1000, 1500$. Initially we chose $n = 40,000$, and later $n = 100,000$ to explore the effects of an increased sample size. For each choice of the complete set of simulation parameters, we summarized the results of eGST for 50 simulated

datasets. We also performed simulations for $\alpha_1 \neq \alpha_2$ and different non-zero mean of $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ distributions.

BMI and WHRadjBMI analysis in the UK Biobank

We implemented eGST to infer the individual-level tissue of interest for two obesity related measures, BMI and WHRadjBMI, in the UK Biobank [23,24], integrating expression data from the Genotype-Tissue Expression (GTEx) project [17,18]. Sets of tissue-specific expressed genes were obtained from Finucane et al. [22]. They considered a gene to be specifically expressed in a tissue of interest if the gene's mean expression in the tissue is substantially higher than its mean expression in other tissues combined, and calculated a t-statistic to rank the genes with respect to higher expression in a specific tissue. Similar to their work [22], we considered the top 10% of all genes (2485 such genes) in a tissue, ranked according to descending value of the t-statistic, as the set of genes specifically expressed in the tissue.

We focused on the adipose and brain tissue for BMI, and the adipose and muscle tissue for WHRadjBMI. We took the union of the sets of genes specifically expressed in adipose subcutaneous and adipose visceral tissues, and considered it as the adipose-specific gene set. Similarly, we took the union of sets of genes specifically expressed in the brain cerebellum and brain cortex regions (these two had maximum sample size among different brain regions) to create a brain-specific set of genes. We excluded the genes overlapping between these two sets to consider non-overlapping sets of adipose and brain specific genes. For WHRadjBMI, we considered adipose subcutaneous and muscle skeletal connective tissue, and excluded the genes overlapping between the two sets of top 10% expressed genes within the tissues. We considered genes on the autosomal chromosomes 1-22. In BMI analysis, the main reason behind merging two type of adipose (or brain) tissues together to represent adipose (or brain) was to increase the number of tissue-specific expressed genes per tissue. For WHRadjBMI analysis, we considered the adipose subcutaneous and muscle skeletal tissues to find different possible patterns in the performance of eGST that might be missed in BMI analysis due to merging tissues.

The primary set of tissue-specific genes that were found to be eGenes in GTEx were included in subsequent analyses. For WHRadjBMI analysis, among the initially selected 2228 adipose subcutaneous tissue-specific genes, 1152 genes were found to be eGenes for which at least one bi-allelic SNP was reported to be an eQTL in the GTEx summary-level data (version v7). Similarly, we had 1272 eGenes for muscle skeletal tissue. And in BMI analysis, we had 1887 eGenes for adipose and 1653 eGenes for brain. For each gene in a tissue, we took the top bi-allelic eQTL SNP (smallest expression association p-value) with MAF $> 1\%$. In BMI analysis, while creating an adipose-specific set of eQTLs, if a gene was both adipose subcutaneous and

visceral tissue-specific gene, we included the top eQTL of the gene in both tissues, one in subcutaneous and one in visceral. We implemented the same strategy for brain tissue, as well.

Next, we obtained the subset of SNPs from each set of tissue-specific eQTL SNPs (obtained from GTEx), which were genotyped or imputed in UKB (imputation accuracy score > 0.9). The SNPs were also screened for HWE (p-value $> 10^{-6}$) in UKB. Using UKB in-sample LD, we LD-pruned each set of tissue-specific eQTL SNPs based on r^2 threshold 0.25. In a tissue-specific set, if two eQTL SNPs had $r^2 > 0.25$, we excluded the one for which the minimum of SNP-expression association p-value (in GTEx) across the genes (for which it was found to be the top eQTL) was larger. Finally, after LD pruning, we had 1705 eQTL SNPs specific to adipose and 1478 eQTL SNPs specific to brain in BMI analysis. We obtained 953 eQTL SNPs specific to adipose subcutaneous tissue and 1052 eQTL SNPs specific to muscle skeletal tissue for WHRadjBMI analysis. We used individual-level genotype data for the tissue-specific SNP sets in UKB to infer tissue of interest across individuals. Before running eGST, we normalized genotypes at each SNP based on the whole sample of individuals.

Phenotype data

We considered the BMI of 337,205 unrelated white-British individuals from the UK Biobank (full release) and excluded individuals for whom BMI or relevant covariates (age, sex, etc.) were missing. We then adjusted BMI for age, sex, and the top 20 principal components (PCs) of genetic ancestry by linear regression and obtained the BMI residuals. We initially developed eGST assuming that each tissue-specific subtype of the trait follows a normal distribution. The BMI residuals obtained after the adjustment of covariates deviated substantially from the normal distribution (p-value of Kolmogorov Smirnov (KS) test for deviation from normal distribution $< 10^{-16}$). One possible approach is to transform the residuals to a normally distributed random variable (e.g., using inverse rank normal transformation). However, for primary investigation of our hypothesis and performance of eGST, we preferred to implement eGST on the original scale of the phenotype. We employed an iterative strategy to extract a subset of individuals for whom the BMI residual follows a normal distribution. In each iteration, we removed a random subset of 6 individuals so that the magnitude of KS test statistic for testing normality of BMI residual in the remaining pool of individuals decreases compared to the remaining pool of individuals in the previous iteration. Thus, at each iteration, we randomly select a subset of individuals for whom the BMI residual gets increasingly distributed as a normal random variable. Finally, we obtained a set of 150,706 individuals for whom the BMI residual follows a normal distribution (KS test p-value = 0.99). We adjusted WHR for BMI to obtain WHRadjBMI. We

then adjusted WHRadjBMI for age, sex, and top 20 PCs of genetic ancestry. The WHRadjBMI residuals significantly deviated from the normal distribution. Implementing the same iterative strategy, we obtained a subset of 271,818 individuals from all individuals for whom the WHRadjBMI residual follows a normal distribution (KS test p-value = 0.99).

Genetic characteristics

We contrasted the genetic basis of the groups of individuals assigned to the adipose and brain-specific subtypes of BMI. Let $\boldsymbol{\beta}_1$ denote the joint SNP-effects of the adipose-specific eQTLs on the BMI of the individuals classified as the adipose-specific subtype for whom the adipose-specific posterior probability obtained by eGST was $> 50\%$. We chose a relaxed threshold of posterior probability because we used multiple linear regression (MLR) to estimate the joint SNP effects of a set of tissue-specific eQTLs on the BMI of a tissue-specific group of individuals, and MLR requires sufficiently large number of individuals (assigned to the corresponding tissue-specific subtype) in the sample for efficient estimation of the model parameters. Let $\boldsymbol{\gamma}_1$ be the joint SNP effects of the brain eQTLs on BMI of individuals assigned to the adipose subtype. Since the BMI of individuals assigned to the adipose subtype should have larger effects from adipose-specific eQTLs than from brain-specific eQTLs, we should expect that the magnitude of a general element in $\boldsymbol{\beta}_1$ would be larger than the magnitude of a general element in $\boldsymbol{\gamma}_1$. Based on the individuals assigned to the adipose subtype, we estimated $\boldsymbol{\beta}_1$ and $\boldsymbol{\gamma}_1$ using multiple linear regression of BMI residual on the genotypes of adipose eQTLs and brain eQTLs in UKB, respectively. Based on the estimated $\boldsymbol{\beta}_1$ and $\boldsymbol{\gamma}_1$ vectors, we performed the non-parametric Wilcoxon rank sum (WRS) test to evaluate $H_0 : |\beta_1| = |\gamma_1|$ versus $H_1 : |\beta_1| > |\gamma_1|$, where β_1 and γ_1 represent a general element in $\boldsymbol{\beta}_1$ and $\boldsymbol{\gamma}_1$ vectors, respectively. Similarly, we tested if the magnitude of the effect of a brain eQTL SNP on the BMI of individuals assigned to brain-specific subtype (β_2) was larger than the corresponding effect magnitude of an adipose eQTL SNP (γ_2). We also tested whether the adipose eQTLs had a larger SNP effect on the adipose subtype of BMI than on the brain subtype, and whether brain eQTLs had a larger effect on the brain subtype than on the adipose subtype. We performed the analogous experiments for the groups of individuals assigned to AS and MS tissue-specific subtype of WHRadjBMI.

Phenotypic characteristics

We explored if the group of individuals whose BMI were classified as a tissue-specific genetic subtype is phenotypically distinct from the rest of the population, with respect to various other phenotypes collected in

the UK Biobank. We considered 106 such phenotypes and individually tested each trait for being differentially distributed between individuals of each tissue-specific subtype and the remaining population (for BMI, 12543 individuals assigned to adipose subtype and 4629 individuals assigned to brain subtype classified based on 65% threshold of subtype posterior probability [Table S1]). We performed the Wilcoxon rank sum (WRS) test for a quantitative trait and χ^2 test based on the contingency table for a qualitative/categorical trait. We corrected the p-values for multiple testing across traits using the Bonferroni correction procedure. The same approach was adopted to find the traits differentially distributed between individuals classified as a tissue-specific subtype of WHRadjBMI and the remaining population (for WHRadjBMI, 11038 individuals with AS subtype and 6596 individuals of MS subtype [Table S1]). For a binary/case-control trait, we term the percentage of individuals (among those assigned to the tissue) who had the disorder as tissue-specific risk of the disease.

A random group of individuals is phenotypically homogeneous with the remaining population

From the pool of 150,706 individuals included in BMI analysis, we randomly selected two groups of individuals with the same size as the groups of tissue-specific BMI subtype (12,543 and 4,629) and evaluated phenotypic heterogeneity across 106 traits between each of the two random groups and the rest of the population using WRS test for a continuous trait and contingency table χ^2 test for a qualitative trait (as before). We repeated the random selection of individuals from the population to replicate the experiment. We did the same experiment for WHRadjBMI.

BMI (or WHRadjBMI) adjusted phenotypic heterogeneity

In the above analysis of BMI, BMI itself was found to be differentially distributed between the individuals with the adipose (as well as brain) specific subtype and the remaining population. Therefore, we further investigated whether the heterogeneity of non-BMI traits between a subtype group and the remaining population were induced due to BMI heterogeneity. For each quantitative trait initially found to be heterogeneous between individuals assigned to one of the subtype groups and the remaining population (Table S8), we first adjusted the trait for BMI in the whole population (150,706 individuals in BMI analysis) and obtained the trait residuals. We then tested for heterogeneity between the trait residual in the adipose (or the brain) subtype group and the remaining population using WRS test. Similarly for qualitative/categorical traits that were initially heterogeneous (Fig. 5, Table 2, Table S10), we performed a binomial or multinomial (depending on the number of categories of the trait) logistic regression adjusting for BMI in the population. We

adopted the same strategy for WHRadjBMI (which itself was found to be differentially distributed between AS as well as MS group and the remaining population) to find which among the non-WHR traits remain heterogeneous after WHRadjBMI adjustment.

Tissue-specific relative change

For each quantitative trait that was differentially distributed between the individuals of a tissue-specific subtype and the remaining population, we measured the relative change (or difference) of the trait between the tissue-specific subtype group and the remaining population as: $\frac{\text{tissue specific mean} - \text{remaining population mean}}{\text{population s.d.}} \times 100$, where the tissue-specific mean of the trait is calculated only in the individuals classified as the corresponding tissue-specific subtype of BMI (or WHRadjBMI). To quantify BMI-adjusted tissue-specific relative change of a primarily heterogeneous quantitative trait, we computed the same measure for BMI-adjusted trait residual (instead of the trait itself). To evaluate the tissue-specific relative change in the risk of a binary/case-control trait, we calculated $\frac{\text{tissue specific prevalence} - \text{population prevalence}}{\text{population s.d.}} \times 100$, where the tissue-specific prevalence is computed only in the individuals assigned to the corresponding tissue-specific subtype.

BMI (or WHRadjBMI) matched tissue-specific relative change

In order to further investigate the role of tissue-specific genetics (uncoupled from the role of BMI heterogeneity) underlying phenotypic characteristics of the individuals assigned to tissue-specific subtype of BMI, we performed the following experiment. We split the range of BMI of the individuals assigned to the adipose subtype (12543 individuals [Table S1]) into 30 consecutive non-overlapping bins. In each BMI bin, we counted the number of individuals assigned to the adipose subtype, and randomly sampled the same number of individuals from all of the individuals contained in the bin. In this way, we randomly selected a pool of individuals (with the same size as the adipose-specific group) from the population, who are matched with the BMI of the adipose subtype individuals. Next, for each non-BMI quantitative trait which was found to be heterogeneous between the adipose group and the remaining population after BMI adjustment (Table S13), we computed: $\left| \frac{\text{adipose specific mean} - \text{BMI matched random mean}}{\text{population s.d.}} \right| \times 100$, where the adipose specific mean of the trait is calculated only in the individuals of the adipose subtype and the BMI matched random mean is the trait mean calculated only in the BMI matched (with adipose group) random pool of individuals. This measure quantifies the relative change/difference of the trait between the individuals assigned to the adipose subtype and the corresponding BMI-matched individuals randomly selected from the population. This should provide insights into the phenotypic characteristics of the individuals with the adipose sub-

type, which is solely mediated through adipose-specific genetics (uncoupled from the corresponding effect of BMI heterogeneity between the adipose group and the remaining population). We repeated the random selection of BMI-matched individuals 500 times and computed the mean and s.d. of the above measure of BMI-matched tissue-specific relative change of a quantitative trait across random selections. We replicated the same experiment for individuals with brain subtype of BMI. For WHRadjBMI, we performed the same experiment to characterize the phenotypic characteristics of AS (or MS) subtype group induced due to AS (or MS) specific genetics only.

Tissue-specificity of phenotypic characteristics

To investigate tissue-specificity of the phenotypic characteristics of the individuals assigned to adipose and brain specific subtype of BMI, we randomly shuffled/exchanged 739 (half of the minimum of number of adipose and brain specific eQTLs = $\frac{1478}{2}$) eQTLs between the set of adipose and brain specific eQTLs to create artificial tissue-specific eQTL sets. We considered 500 such random shuffles. Keeping the phenotype data fixed, for the genotype data at each set of artificial tissue-specific eQTLs, we ran eGST to identify the groups of individuals with the BMI subtype specific to the artificial adipose and brain tissues (based on the posterior probability threshold of 65%). Next, for each quantitative trait that was found to be primarily heterogeneous between the individuals assigned to the original adipose (or brain) subtype of BMI and the remaining population (Table S8), we computed the artificial adipose and brain tissue-specific trait mean only in the individuals classified into the corresponding artificial tissue-specific subtype of BMI (based on the artificial tissue-specific eQTLs). Then for each trait, we computed central tendency measures of the artificial tissue-specific trait means across 500 sets of artificial tissue-specific eQTLs. For each trait, we also tested whether the overall mean of the artificial tissue-specific trait means is significantly different from the corresponding original (adipose or brain) tissue-specific trait mean. We performed the same experiment for WHRadjBMI.

Permuting phenotype data across individuals

Next, we performed a similar experiment for permuted phenotype (BMI or WHRadjBMI) data while keeping the eQTL assignment to tissues fixed as it is in the original data. We consider 500 random permutations of BMI across individuals. Keeping the genotype data fixed, we ran eGST for each permuted phenotype data and classified the tissue of interest across individuals based on 65% threshold of subtype posterior probability. As before, in each of these 500 pairs of subtype groups of individuals thus obtained, subtype-

specific means were computed for each quantitative trait that was found to be primarily heterogeneous between the individuals of the original adipose (or brain) subtype of BMI and the remaining population (Table S8). For each trait, we then computed central tendency measures of the tissue-specific means across 500 random BMI permutations. For each trait, we tested whether the overall mean of the tissue-specific trait means obtained across random permutations was significantly different from the corresponding original (adipose or brain) tissue-specific trait means. We conducted the same experiment for WHRadjBMI.

References

[1] Jaana M Leppala, Jarmo Virtamo, Rainer Fogelholm, Demetrius Albanes, and Olli P Heinonen. Different risk factors for different stroke subtypes: association of blood pressure, cholesterol, and antioxidants. *Stroke*, 30(12):2535–2540, 1999.

[2] BM Schamberger, HC Geiss, MM Ritter, P Schwandt, and KG Parhofer. Influence of ldl apheresis on ldl subtypes in patients with coronary heart disease and severe hyperlipoproteinemia. *Journal of lipid research*, 41(5):727–733, 2000.

[3] John M Ringman, Alison Goate, Colin L Masters, Nigel J Cairns, Adrian Danek, Neill Graff-Radford, Bernardino Ghetti, John C Morris, Dominantly Inherited Alzheimer Network, et al. Genetic heterogeneity in alzheimer disease and implications for treatment strategies. *Current neurology and neuroscience reports*, 14(11):499, 2014.

[4] Paul Gibson, Yiai Tong, Giles Robinson, Margaret C Thompson, D Spencer Currle, Christopher Eden, Tanya A Kranenburg, Twala Hogg, Helen Poppleton, Julie Martin, et al. Subtypes of medulloblastoma have distinct developmental origins. *Nature*, 468(7327):1095, 2010.

[5] Judy H Cho and Marc Feldman. Heterogeneity of autoimmune diseases: pathophysiologic insights from genetics and implications for new therapies. *Nature medicine*, 21(7):730, 2015.

[6] James Liley, John A Todd, and Chris Wallace. A method for identifying genetic heterogeneity within phenotypically defined disease subgroups. *Nature genetics*, 49(2):310, 2017.

[7] Javaid Iqbal, Ophira Ginsburg, Paula A Rochon, Ping Sun, and Steven A Narod. Differences in breast cancer stage at diagnosis and cancer-specific survival by race and ethnicity in the united states. *Jama*, 313(2):165–173, 2015.

[8] Roger L Milne, Karoline B Kuchenbaecker, Kyriaki Michailidou, Jonathan Beesley, Siddhartha Kar, Sara Lindström, Shirley Hui, Audrey Lemaçon, Penny Soucy, Joe Dennis, et al. Identification of ten variants associated with risk of estrogen-receptor-negative breast cancer. *Nature genetics*, 49(12):1767, 2017.

[9] Montserrat Garcia-Closas, Fergus J Couch, Sara Lindstrom, Kyriaki Michailidou, Marjanka K Schmidt, Mark N Brook, Nick Orr, Suhn Kyong Rhie, Elio Riboli, Heather S Feigelson, et al. Genome-wide association studies identify four er negative–specific breast cancer risk loci. *Nature genetics*, 45(4):392, 2013.

[10] Mark Zimmerman, Theresa A Morgan, and Kasey Stanton. The severity of psychiatric disorders. *World Psychiatry*, 17(3):258–275, 2018.

[11] Raphael Bernier, Christelle Golzio, Bo Xiong, Holly A Stessman, Bradley P Coe, Osnat Penn, Kali Witherspoon, Jennifer Gerdts, Carl Baker, Anneke T Vulto-van Silfhout, et al. Disruptive chd8 mutations define a subtype of autism early in development. *Cell*, 158(2):263–276, 2014.

[12] Shafali S Jeste and Daniel H Geschwind. Disentangling the heterogeneity of autism spectrum disorder through genetic findings. *Nature Reviews Neurology*, 10(2):74, 2014.

[13] Holly A Stessman, Raphael Bernier, and Evan E Eichler. A genotype-first approach to defining the subtypes of a complex disease. *Cell*, 156(5):872–877, 2014.

[14] Miriam S Udler, Jaegil Kim, Marcin von Grotthuss, Sílvia Bonàs-Guarch, Joanne B Cole, Joshua Chiou, Michael Boehnke, Markku Laakso, Gil Atzmon, Benjamin Glaser, et al. Type 2 diabetes genetic loci informed by multi-trait associations point to disease mechanisms and subtypes: A soft clustering analysis. *PLoS medicine*, 15(9):e1002654, 2018.

[15] Thomas W Winkler, Felix Günther, Simon Höllerer, Martina Zimmermann, Ruth JF Loos, Zoltán Kutalik, and Iris M Heid. A joint view on genetic variants for adiposity differentiates subtypes with distinct metabolic implications. *Nature communications*, 9(1):1946, 2018.

[16] Andy Dahl, Na Cai, Arthur Ko, Markku Laakso, Paivi Pajukanta, Jonathan Flint, and Noah Zaitlen. Reverse gwas: Using genetics to identify and model phenotypic subtypes. *bioRxiv*, page 446492, 2018.

[17] GTEx Consortium et al. The genotype-tissue expression (gtex) pilot analysis: multitissue gene regulation in humans. *Science*, 348(6235):648–660, 2015.

[18] GTEx Consortium et al. Genetic effects on gene expression across human tissues. *Nature*, 550(7675):204, 2017.

[19] Jason M Torres, Eric R Gamazon, Esteban J Parra, Jennifer E Below, Adan Valladares-Salgado, Niels Wacher, Miguel Cruz, Craig L Hanis, and Nancy J Cox. Cross-tissue and tissue-specific eqtls: partitioning the heritability of a complex trait. *The American Journal of Human Genetics*, 95(5):521–534, 2014.

[20] Ruo-Han Hao, Tie-Lin Yang, Yu Rong, Shi Yao, Shan-Shan Dong, Hao Chen, and Yan Guo. Gene expression profiles indicate tissue-specific obesity regulation changes and strong obesity relevant tissues. *International Journal of Obesity*, 2017.

[21] Halit Ongen, Andrew A Brown, Olivier Delaneau, Nikolaos I Panousis, Alexandra C Nica, Emmanouil T Dermitzakis, GTEx Consortium, et al. Estimating the causal tissues for complex traits and diseases. *Nature genetics*, 49(12):1676, 2017.

[22] Hilary K Finucane, Yakir A Reshef, Verneri Anttila, Kamil Slowikowski, Alexander Gusev, Andrea Byrnes, Steven Gazal, Po-Ru Loh, Caleb Lareau, Noam Shoresh, et al. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nature genetics*, 50(4):621, 2018.

[23] Cathie Sudlow, John Gallacher, Naomi Allen, Valerie Beral, Paul Burton, John Danesh, Paul Downey, Paul Elliott, Jane Green, Martin Landray, et al. Uk biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS medicine*, 12(3):e1001779, 2015.

[24] Clare Bycroft, Colin Freeman, Desislava Petkova, Gavin Band, Lloyd T Elliott, Kevin Sharp, Allan Motyer, Damjan Vukcevic, Olivier Delaneau, Jared OConnell, et al. The uk biobank resource with deep phenotyping and genomic data. *Nature*, 562(7726):203, 2018.

[25] Yiming Hu, Qiongshi Lu, Ryan Powles, Xinwei Yao, Can Yang, Fang Fang, Xinran Xu, and Hongyu Zhao. Leveraging functional annotations in genetic risk prediction for human complex diseases. *PLoS computational biology*, 13(6):e1005589, 2017.

[26] Farhad Hormozdiari, Steven Gazal, Bryce van de Geijn, Hilary K Finucane, Chelsea J-T Ju, Po-Ru Loh, Armin Schoech, Yakir Reshef, Xuanyao Liu, Luke OConnor, et al. Leveraging molecular quanti-

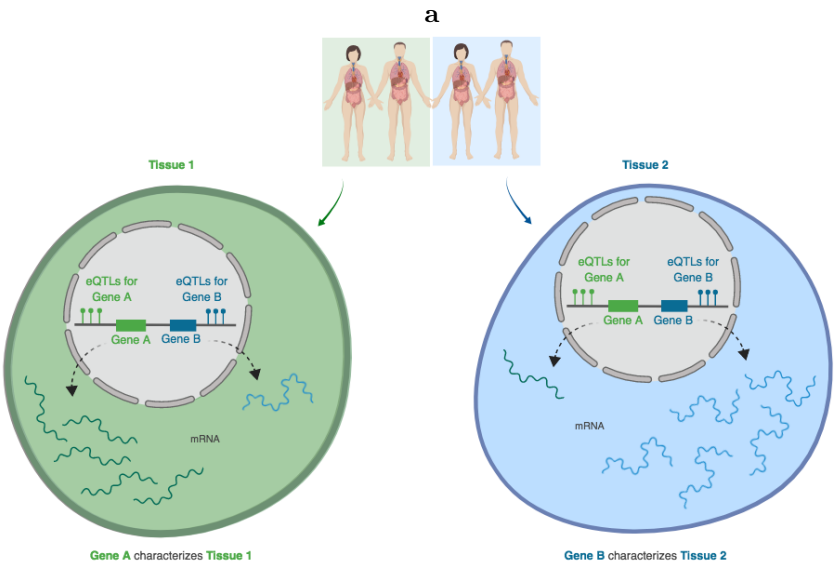
tative trait loci to understand the genetic architecture of diseases and complex traits. *Nature genetics*, 50(7):1041, 2018.

[27] Eric R Gamazon, Heather E Wheeler, Kaanan P Shah, Sahar V Mozaffari, Keston Aquino-Michaels, Robert J Carroll, Anne E Eyler, Joshua C Denny, Dan L Nicolae, Nancy J Cox, et al. A gene-based association method for mapping traits using reference transcriptome data. *Nature genetics*, 47(9):1091, 2015.

[28] Wen Zhang, Georgios Voloudakis, Veera M Rajagopal, Ben Reahead, Joel T Dudley, Eric E Schadt, Johan LM Bjorkegren, Yungil Kim, John F Fullard, Gabriel E Hoffman, et al. Integrative transcriptome imputation reveals tissue-specific and shared biological mechanisms mediating susceptibility to complex traits. *bioRxiv*, page 532929, 2019.

[29] Geoffrey McLachlan and Thriyambakam Krishnan. *The EM algorithm and extensions*, volume 382. John Wiley & Sons, 2007.

[30] Maya R Gupta, Yihua Chen, et al. Theory and use of the em algorithm. *Foundations and Trends® in Signal Processing*, 4(3):223–296, 2011.



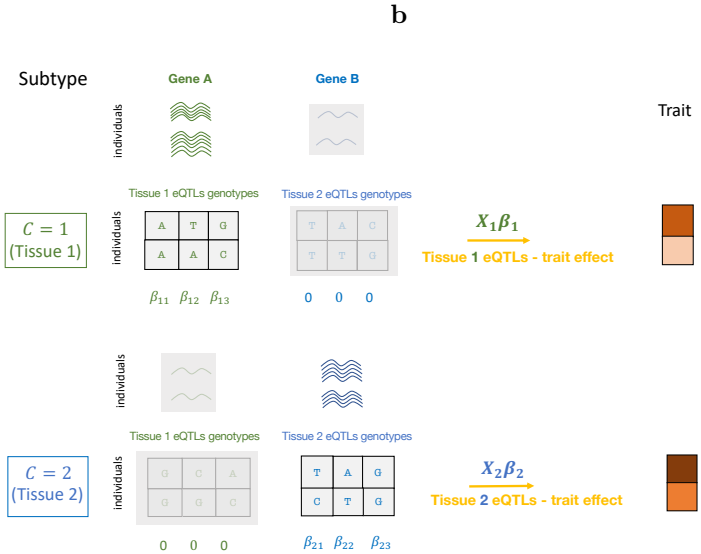


Figure 1: The top diagram (a) explains our main hypothesis and the bottom diagram (b) explains our model. Consider two tissues of interest for a phenotype, tissue 1 and tissue 2, where gene A has higher expression but gene B has lower expression in tissue 1, and gene B has higher expression but gene A has lower expression in tissue 2. The key hypothesis is that the susceptibility of the phenotype for the first two individuals is mediated through the effect of gene A in tissue 1, in which case we can assign tissue 1 as the tissue of interest for these individuals (similarly tissue 2 for last two individuals). We refer to the phenotype of the first two individuals as tissue 1 specific subtype. We use genotypes at the tissue-specific eQTLs as a proxy for the expressions of corresponding tissue-specific genes (b). We consider a finite mixture model with each of its components being a linear model regressing the trait on the genotypes at each set of tissue-specific eQTLs. Our method takes as input the individual-level measurements of the phenotype and genotypes at the sets of tissue-specific eQTLs and provides per-individual tissue-specific posterior probabilities as the main output.

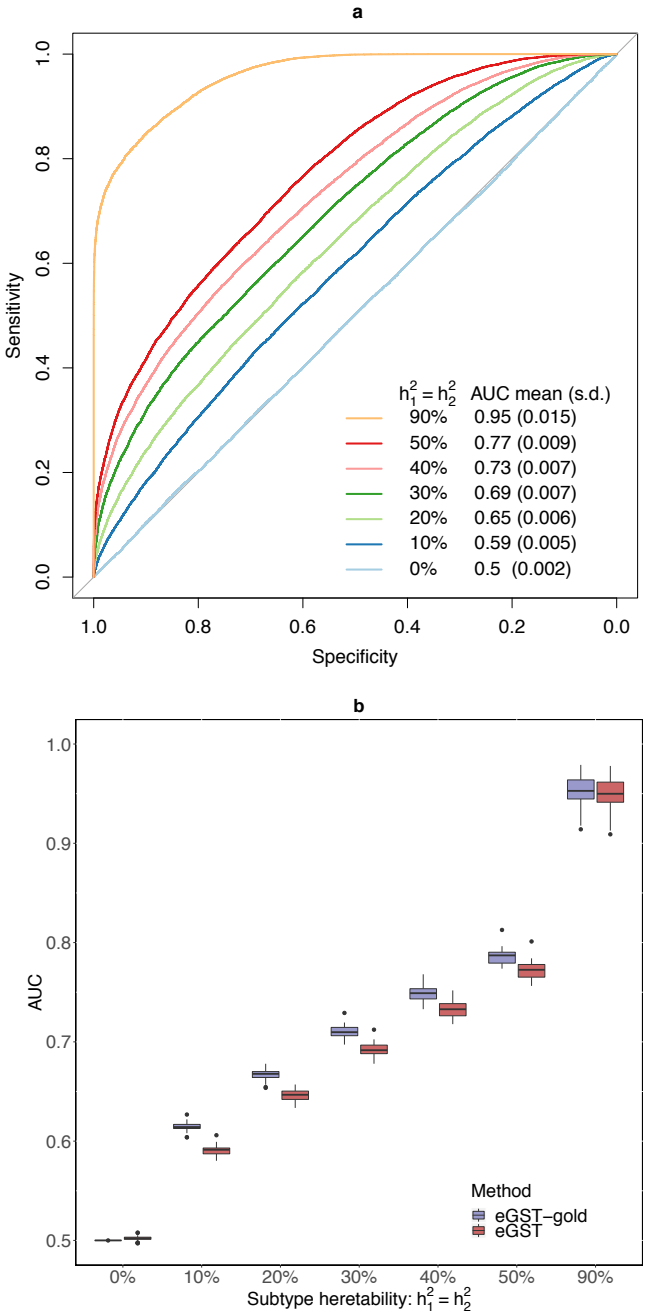


Figure 2: In the first diagram (a), we present the receiver operating characteristic (ROC) curve evaluating the classification accuracy of eGST for a single dataset simulated under the following scenarios: $h_1^2 = h_2^2 = 0\%, 10\%, 20\%, 30\%, 40\%, 50\%, 90\%$, $w_1 = w_2 = \frac{1}{2}$, $m_1 = m_2 = 1000$, $n = 40000$. The mean (over 50 simulated datasets) area under the curve (AUC) obtained by eGST under the same scenarios are also provided. Here h_1^2 and h_2^2 are the heritability of tissue-specific subtypes of the trait due to m_1 and m_2 SNPs representing two sets of tissue-specific eQTL SNPs, w_1 and w_2 are the proportions of individuals in the sample assigned to the two tissues, n is the total number of individuals. In the second diagram (b), box plots of AUCs obtained by eGST and the gold-standard strategy implementing our model in which true model parameters were assumed to be known while estimating the tissue-specific posterior probabilities are presented across the same simulation scenarios.

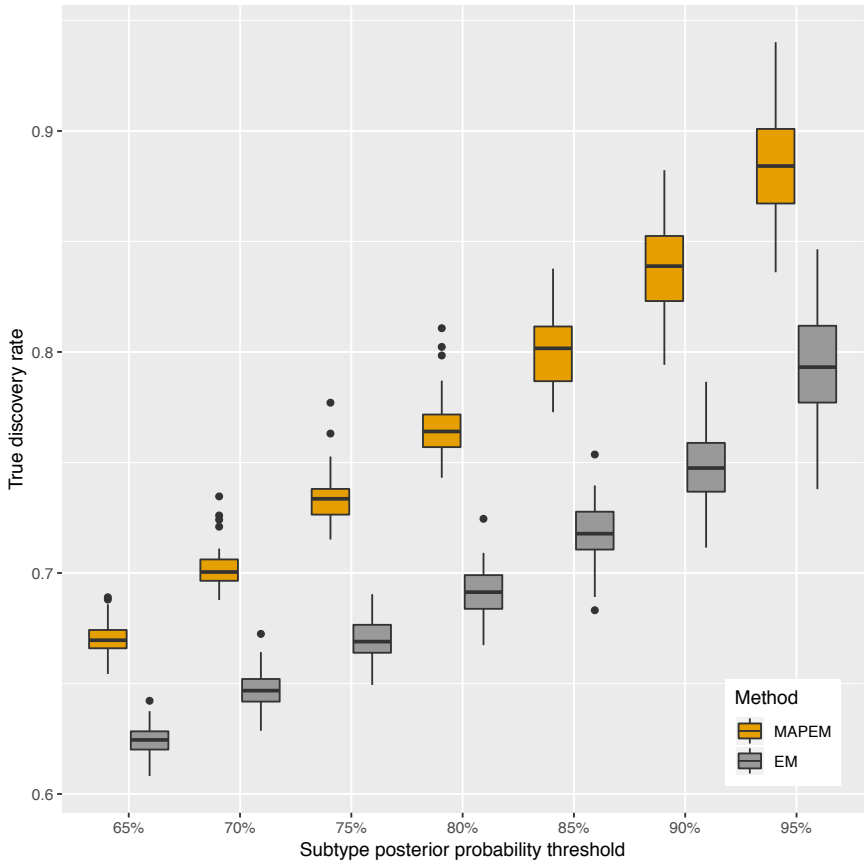


Figure 3: Comparison between the true discovery rate (TDR) of classifying tissue-specific subtypes by the MAP-EM algorithm (under the Bayesian framework of the mixture model which eGST employs) versus the EM algorithm (under the frequentist framework of the mixture model) based on the threshold of tissue-specific subtype posterior probability as $65\%, 70\%, 75\%, 80\%, 85\%, 90\%, 95\%$, respectively. Box plots of TDR across 50 datasets simulated under $h_1^2 = h_2^2 = 10\%$, $w_1 = w_2 = \frac{1}{2}$, $m_1 = m_2 = 1000$, $n = 40,000$ are presented. Here h_1^2 and h_2^2 are the heritability of tissue-specific subtypes of the trait due to m_1 and m_2 SNPs representing two sets of tissue-specific eQTL SNPs, w_1 and w_2 are the proportions of individuals in the sample assigned to the two tissues, n is the total number of individuals.

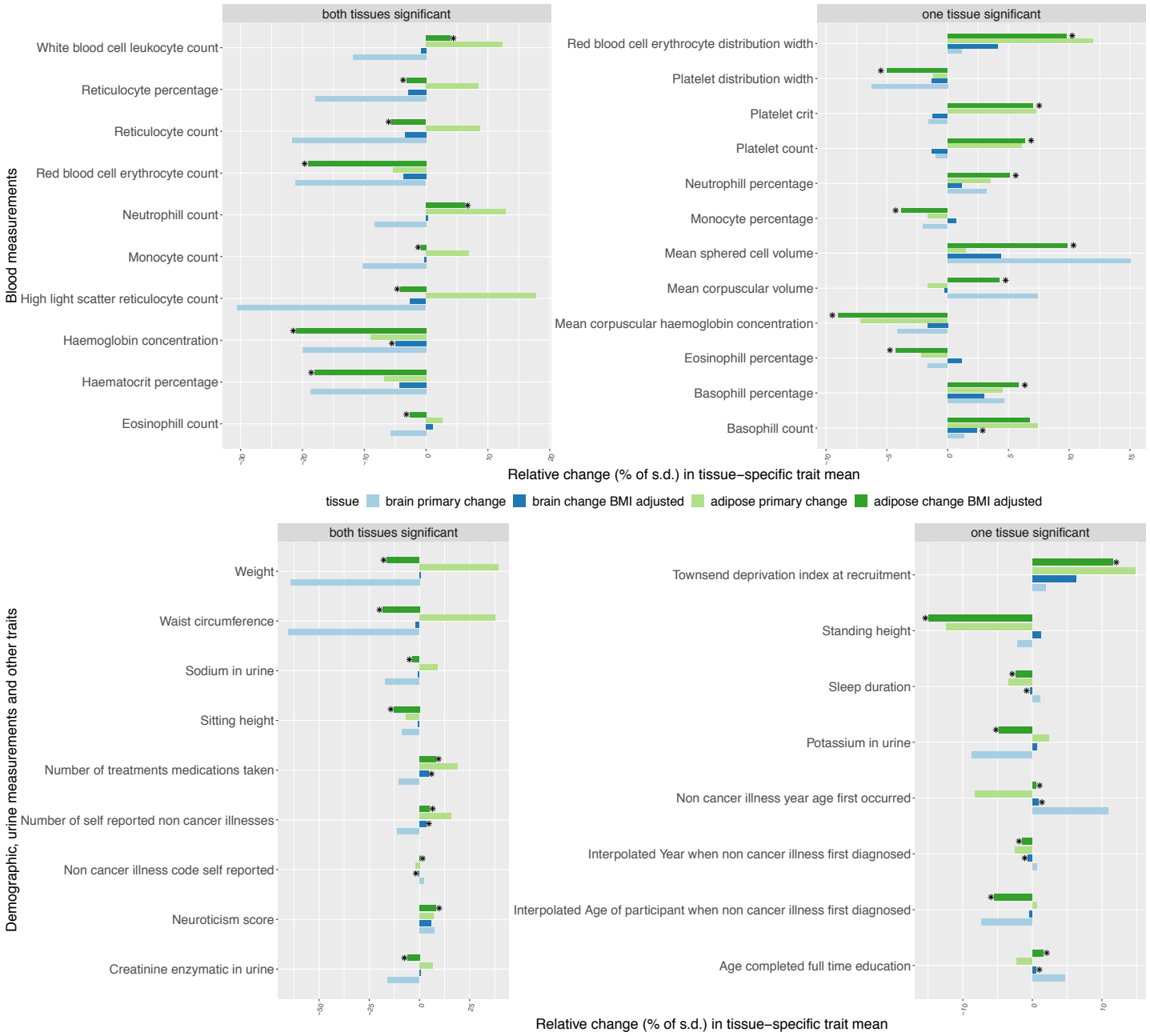


Figure 4: Percentage of tissue-specific relative change of the quantitative traits that were differentially distributed between the individuals assigned to a tissue-specific subtype of BMI and the remaining population. Traits in the left panels are primarily heterogeneous for both tissue-specific groups and traits in the right panels are heterogenous for one tissue-specific group. We measure the tissue-specific relative change of a trait by: $\frac{\text{tissue specific mean - remaining population mean}}{\text{population s.d.}} \times 100$, where the tissue-specific mean is computed only in the individuals with the corresponding tissue-specific subtype. The same measure is calculated for a trait residual obtained after adjusting for BMI to quantify the tissue-specific relative change of the trait after BMI adjustment. The faded green (or blue) bar presents primary adipose (or brain) tissue-specific relative change of a trait compared to the remaining population. The dark green (or blue) bar presents the BMI-adjusted adipose (or brain) specific relative change of a trait. Each trait listed here was found to be differentially distributed between at least one of the adipose or brain specific groups and the remaining population after BMI adjustment. For each trait, the asterisk mark attached to the bars indicates which tissue-specific group remains significantly heterogeneous after BMI adjustment.

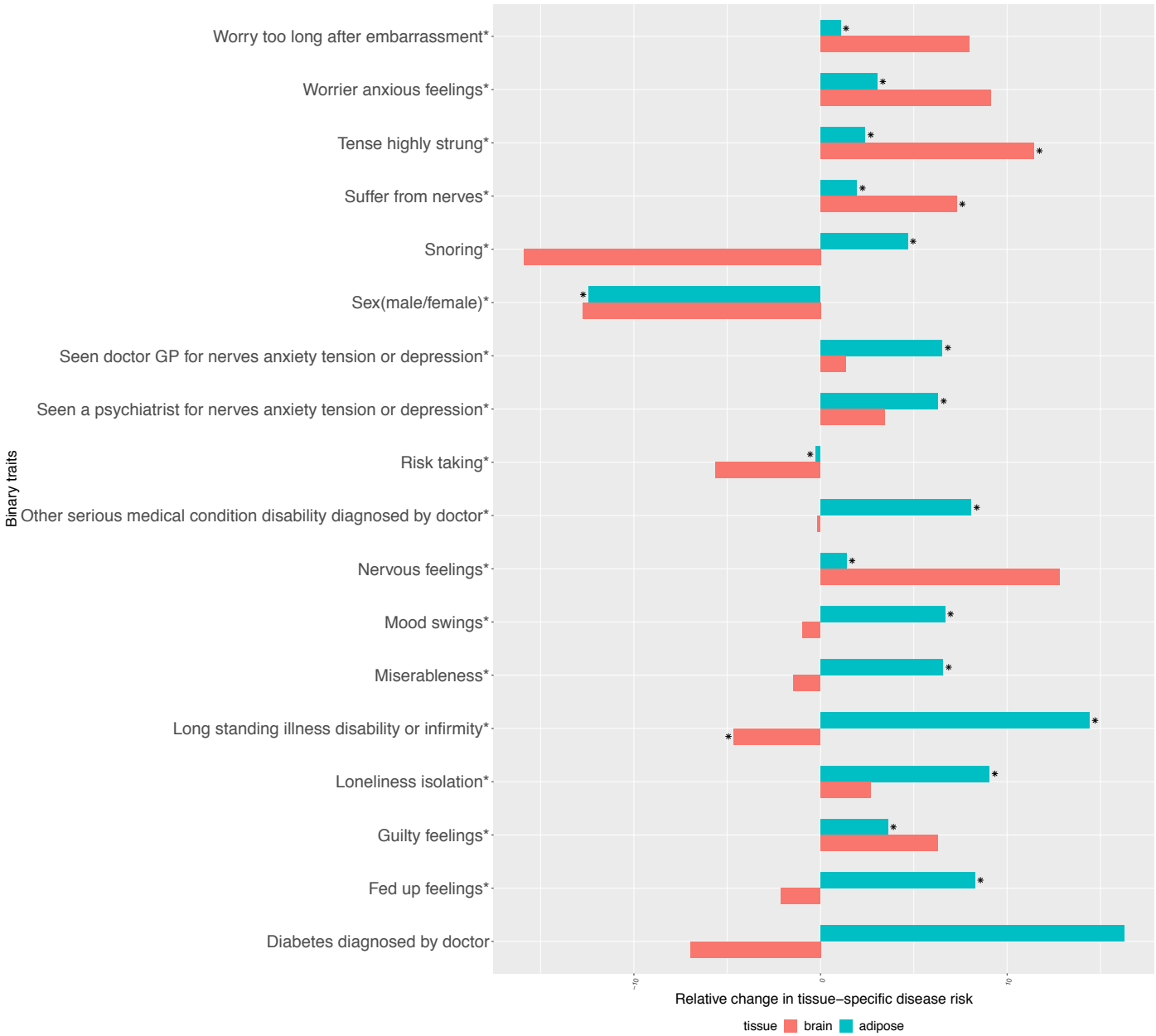


Figure 5: Percentage of tissue-specific relative change in the risk of case-control traits between the individuals assigned to a tissue-specific subtype of BMI and the population. The tissue-specific relative change of a disease risk is measured by: $\frac{\text{tissue specific prevalence - population prevalence}}{\text{population s.d.}} \times 100$. Tissue-specific prevalence of the disorder was computed only in the individuals classified as the corresponding tissue-specific subtype of BMI. The asterisk mark attached to the traits indicate which trait remains differentially distributed between at least one of the adipose and brain tissue-specific groups of individuals and the remaining population after BMI adjustment. For each trait the asterisk mark attached to the bars indicate which tissue-specific group of individuals remains significantly heterogeneous for the trait after BMI adjustment.

BMI analysis				
	Mean magnitude of eQTLs' effects			
BMI subtype	Adipose	Brain	WRS test p-values	
Adipose	0.045 ($\lvert\beta_1\rvert$)	0.031 ($\lvert\gamma_1\rvert$)	$<$ 2E-16 ($H_1 : \lvert\beta_1\rvert > \lvert\gamma_1\rvert$)	$<$ 2E-16 ($H_1 : \lvert\beta_2\rvert > \lvert\gamma_1\rvert$)
Brain	0.017 ($\lvert\gamma_2\rvert$)	0.034 ($\lvert\beta_2\rvert$)	0.002 ($H_1 : \lvert\beta_1\rvert > \lvert\gamma_2\rvert$)	$<$ 2E-16 ($H_1 : \lvert\beta_2\rvert > \lvert\gamma_2\rvert$)

WHRadjBMI analysis				
	Mean magnitude of eQTLs' effects			
WHRadjBMI subtype	AS	MS	WRS test p-values	
AS	0.00072 ($\lvert\beta_1\rvert$)	0.00044 ($\lvert\gamma_1\rvert$)	$<$ 2E-16 ($H_1 : \lvert\beta_1\rvert > \lvert\gamma_1\rvert$)	$<$ 2E-16 ($H_1 : \lvert\beta_2\rvert > \lvert\gamma_1\rvert$)
MS	0.00044 ($\lvert\gamma_2\rvert$)	0.00062 ($\lvert\beta_2\rvert$)	$<$ 2E-16 ($H_1 : \lvert\beta_1\rvert > \lvert\gamma_2\rvert$)	$<$ 2E-16 ($H_1 : \lvert\beta_2\rvert > \lvert\gamma_2\rvert$)

Table 1: Genetic heterogeneity between groups of individuals assigned to tissue-specific subtypes of BMI (or WHRadjBMI). In the BMI analysis, β_1 denotes the SNP-effect of an adipose eQTL on BMI in those individuals assigned to the adipose-specific subtype of BMI, β_2 is the SNP-effect of a brain eQTL on the brain-specific subtype of BMI, γ_1 is the SNP-effect of a brain eQTL on the adipose subtype and γ_2 is the effect of an adipose eQTL on the brain subtype. We provide the mean magnitude of the effect sizes of a tissue-specific eQTLs on the BMI of the corresponding tissue-specific group of individuals (e.g., joint SNP-effect of the adipose eQTLs on the BMI of individuals with adipose subtype), and the p-values obtained from the Wilcoxon rank sum (WRS) right tail tests of effect heterogeneity. For each test, the alternative hypotheses are listed in parentheses, while the null hypothesis is the equality between the corresponding pair of parameters. These parameters are defined in the same way for the adipose subcutaneous (AS) and muscle skeletal (MS) tissue-specific subtype of WHRadjBMI and the same analyses are performed.

Trait	P adipose		P brain		signif tissue		#categ
	primary	BMIadj	primary	BMIadj	primary	BMIadj	
*Overall health rating	1.52E-178	5.09E-14	5.16E-17	0.008	both	adipose	4
*Weight change compared with 1 year ago	3.13E-22	2.38E-08	7.86E-58	3.00E-05	both	both	3
Qualifications	8.44E-11	0.02	1.47E-15	0.63	both	none	7
Alcohol intake versus 10 years previously	6.23E-15	0.26	1.25E-06	0.27	both	none	3
Nap during day	1.74E-06	0.001	7.61E-09	0.86	both	none	3
*Past tobacco smoking	0.0004	7.76E-11	1.87E-06	0.61	both	adipose	4
*Smoking status	2.07E-09	2.81E-15	1.54E-05	0.57	both	adipose	3
*Alcohol intake frequency	1.99E-80	1.60E-06	0.0008	0.89	adipose	adipose	6
Frequency of tiredness lethargy in last weeks	1.04E-45	0.05	0.19	0.44	adipose	none	4
*Falls in the last year	5.45E-26	1.45E-06	0.17	0.66	adipose	adipose	3
Frequency of unenthusiasm disinterest in last 2 weeks	2.57E-21	0.002	0.36	0.88	adipose	none	4
Illness injury bereavement stress in last 2 years	6.18E-21	0.20	0.04	0.82	adipose	none	7
Blood clot DVT bronchitis emphysema asthma rhinitis eczema allergy diagnosed by doctor	5.11E-13	0.07	0.02	0.34	adipose	none	6
*Alcohol drinker status	1.03E-26	2.70E-22	0.06	0.002	adipose	adipose	3
*Frequency of depressed mood in last 2 weeks	2.75E-24	4.61E-08	0.20	0.06	adipose	adipose	4
Getting up in morning	2.74E-18	0.34	0.07	0.49	adipose	none	4
*Current tobacco smoking	8.16E-16	2.96E-15	0.02	0.57	adipose	adipose	3
Frequency of tenseness restlessness in last 2 weeks	1.09E-11	0.02	0.003	0.90	adipose	none	4
Sleeplessness insomnia	1.39E-09	0.003	0.87	0.55	adipose	none	3
Morning evening person chronotype	2.30E-06	0.65	0.32	1.00	adipose	none	4

Table 2: Qualitative/categorical traits with three or more categories that are differentially distributed between at least one of the adipose and brain-specific subtype groups of individuals for BMI and the remaining population. For each trait, we provide the p-values of testing heterogeneity between each tissue-specific subtype group of individuals and the remaining population before (primary) and after BMI adjustment (BMIadj). For each trait, tissue-specific groups which appear to be significantly heterogeneous (signif tissue) before (primary) and after BMI adjustment (BMIadj) are also provided. The asterisk mark attached to the traits indicate which trait remains differentially distributed between at least one of the tissue-specific groups and the remaining population after BMI adjustment. The number of categories for each trait (#categ) are also listed.