bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Stratified Linkage Disequilibrium Score Regression reveals enrichment of eQTL effects on complex traits is not tissue specific

Authors

Hill F. Ip, [1]* Rick Jansen,[2] Abdel Abdellaoui,[1] Meike Bartels,[1,3,4] Dorret I. Boomsma,[1,3,4] & Michel G. Nivard[1]**

Affiliations

[1] Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, North-Holland, 1081 BT, the Netherlands

[2] Department of Psychiatry, VU University Medical Center, Amsterdam, North-Holland, 1081 HL, the Netherlands

[3] Amsterdam Public Health Research Institute, Amsterdam, North-Holland, 1081 BT, the Netherlands

[4] Neuroscience Amsterdam, Amsterdam, North-Holland, 1081 HV, the Netherlands

Corresponding authors

* h.f.ip@vu.nl

** m.g.nivard@vu.nl

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Both gene expression levels and eQTLs (expression quantitative trait loci) are partially tissue specific, complicating the detection of eQTLs in tissues with limited sample availability, such as the brain. However, eQTL overlap between tissues might be non-trivial, allowing for inference of eQTL functioning in the brain via eQTLs measured in readily accessible tissues, e.g. whole blood. Using Stratified Linkage Disequilibrium Score Regression (SLDSR), we quantify the enrichment of blood and brain eQTLs in genome-wide association study (GWAS) for three immune-related traits (Crohn's disease, rheumatoid arthritis, and ulcerative colitis), three brain-related (BMI, educational attainment, and schizophrenia), and five traits not associated with either tissue. Our analyses establish strong enrichments of blood (32-113 fold enrichment, mean=59) and brain (21-63 fold enrichment, mean=40) eQTLs in their effects across all traits. We find no evidence for tissue-specific enrichment in GWAS signal for either eQTLs uniquely found in the brain or whole blood. To extend our finding, we test tissue-specific enrichment of eQTLs discovered in 44 tissues by the Genotype-Tissue Expression (GTEx) consortium, and, again, find no tissue-specific eQTL effects. Finally, we integrate the GTEx eQTLs with tissue-specific epigenetic data and find substantially enriched effects on schizophrenia, though again not tissue specific. We conclude that, while eQTLs are strongly enriched in GWAS, the enrichment is not specific to the tissue used in eQTL discovery. Therefore, using relatively accessible tissues, such as whole blood, as proxy for eQTL discovery is sensible; and restricting lookups for GWAS hits to a specific tissue might not be advisable.

Introduction

The main aim of Genome-wide association studies (GWASs) is to detect statistically significant associations between genetic variants, such as single nucleotide polymorphisms (SNPs), and a trait of interest.[1] GWASs have identified many genetic variants and thereby provided insights into the genetic architecture of complex traits.[1,2] However, as a large number of variants identified through

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

GWASs are located outside of coding regions and specific knowledge of regulatory elements is limited, uncovering a relationship between GWAS hits and biological function has proven to be complicated.[3] Expression quantitative trait loci (eQTLs) are SNPs that influence gene expression, and are not necessarily located in coding regions. eQTLs may aid functional annotation of SNPs that have been identified in a GWAS and are located outside of coding regions.[3,4] Previous work has found substantial enrichment of eQTLs among GWAS hits and an enrichment in their genome-wide effect on complex traits.[5–8] Therefore, eQTLs are viewed as an important tool in moving from genome-wide association to biological interpretation.

As a result of difference in gene expression between cells originating from different tissues, eQTLs are potentially tissue specific.[9,10] Tissue-specificity poses no problem if the tissue of interest is readily available for research, such as whole blood. However, discovery of eQTLs gets complicated when measurement of expression levels in a tissue is limited by ethical and practical considerations, for example in brain tissue. Several studies have shown that the overlap between eQTLs from different tissues might actually be larger than initially assumed.[11,12] The Genotype-Tissue Expression (GTEx) consortium identified eQTLs in a wide range of human tissues and showed that 54-90% of the eQTLs identified in one tissue are also designated as an eQTL in at least one other tissue.[10] In another study, Liu *et al* found a high average pairwise genetic correlation (r_g=0.738) of local gene expression between tissues.[13] Nevertheless, small differences in terms of eQTL effect may be of considerable importance in terms of the effects of eQTLs on complex traits related to specific tissues. It is worthwhile to investigate the specific utility of tissue-specific eQTLs in their effect on complex traits, as studied in GWAS. It is plausible that the discovery of eQTLs for tissues such as the brain can be advanced by eQTLs discovered in more accessible tissues, such as whole blood. The use of accessible tissues, though, depends on a substantial degree of similarity of eQTL effect across tissue, and to what extend eQTL differences between tissues are important in complex trait etiology.

Stratified Linkage Disequilibrium Score Regression (SLDSR) is a technique that estimates the SNP-heritability (h^2_{SNP}) of a trait based on GWAS summary statistics.[14,15] By simultaneously analyzing

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

multiple categories of SNPs (annotations), SLDSR can also partition h^2_{SNP} by annotation (h^2_{annot}) and thereby provides a way to jointly quantify the enrichment in GWAS signal of several annotations. Here, we extend SLDSR by including annotations containing *cis*-eQTLs from several tissues, i.e. eQTLs located closely to the gene with which they associate.[16,17] To this end, we perform analyses based on representative eQTL resources, and consider a variety of traits as outcomes.

First, we selected the strongest eQTLs per gene discovered in large samples of measures taken in whole blood and brain.[17–19] We quantify the contribution of these blood and brain eQTLs to the genetic variance in complex traits captured in GWAS. We then attempt to detect tissue-specific eQTL effects on complex traits by estimating the enrichments of eQTLs uniquely found in whole blood or uniquely found in brain, conditional on the enrichment of the complete blood eQTL annotation or complete brain eQTL annotation, respectively. We consider the effect of eQTLs on three brain-related phenotypes: schizophrenia (MIM: 181500), BMI (MIM: 606641), and educational attainment; three immune disorders: Crohn's disease (MIM: 266600), rheumatoid arthritis (MIM: 180300), and ulcerative colitis (MIM: 266600); and five assorted traits and disorders: age at menarche (MIM: 612882), coronary artery disease (MIM: 607339), height (MIM: 606255), LDL levels, and smoking behavior (MIM: 188890).

Second, we retrieve all eQTLs identified in any of the 44 tissues from the GTEx consortium.[10] We consider the enrichment in GWAS signal of the union of all GTEx eQTLs, and, additionally, the enrichment of tissue-specific eQTL effects on top of the union of all GTEx eQTLs. We expect to observe tissue-specific enrichment of eQTLs in their effects on complex traits related to the tissue in question, e.g. eQTLs discovered in immune-related tissues are expected to show higher enrichments in their effect on immune-related traits compared to eQTLs found in skin tissue. We consider tissue-specific enrichment of *cis*-eQTLs in their effect on schizophrenia (a disorder where there is strong prior evidence for the involvement of processes in the brain) and rheumatoid arthritis (a disease with strong prior evidence for the involvement of processes in immune tissue). We further consider

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the enrichment of the intersection of *cis*-eQTLs discovered in any tissue, and histone modification in a specific tissue (i.e. tissue-specific epigenetically changed chromatin states in regulatory regions).

Our analyses are designed to elucidate the relation between eQTLs and complex traits, and to quantify the extent to which this relation is dependent on the tissue used in eQTL discovery. Our analysis further considers the enrichment of genomic regions related to gene expression and epigenetically modified in specific tissues.

Material and Methods

SLDSR method

A measure of linkage disequilibrium (LD) for each SNP, called an "LD score", can be computed by taking the sum of correlations between that SNP and all neighboring SNPs.[14,15] Under a polygenic model, LD scores are expected to show a linear relationship with GWAS test statistics of corresponding SNPs, where the slope is proportional to h^2_{SNP}. For SLDSR, LD scores are based on only (functional) parts of the genome and used as predictors in a multiple linear regression.[15] In this manner, SLDSR is able to partition h^2_{SNP} into parts that are explained by these parts of the genome (i.e. h^2_{annot}), while accounting for influences of the remaining annotations in the model. The enrichment of an annotation is then obtained by taking the ratio of h^2_{annot} over the proportion of SNPs that fall within that annotation.

Target traits

As outcome for SLDSR, we used summary statistics of GWASs on Crohn's disease,[20] rheumatoid arthritis,[21] ulcerative colitis,[20] BMI,[22] educational attainment,[23] schizophrenia,[24] age at menarche,[25] coronary artery disease,[26] height,[27] LDL levels,[28] and smoking behavior.[29] The first three traits were chosen because they are related to the immune system and are therefore expected to reveal considerable enrichment of blood eQTL signal.[20,21] Similarly, brain eQTLs are expected to show substantial enriched effects due to previous reports on the involvement of the central nervous

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

system (CNS) in schizophrenia,[24] educational attainment,[23] and BMI.[30] Of course, these traits do not perfectly align with either tissue, e.g. the immune system has been implicated in the etiology of schizophrenia and BMI,[31,32] and might therefore also be enriched in their effects for the other eQTL set. However, this is expected to occur at lower rates. Enrichment of blood and brain eQTL effects on the remaining traits was calculated to contrast the results with traits for which we do not have a strong *a priori* expectation of the relationship between trait and tissue.

The discovery sample for detection of blood eQTLs included participants from the Netherlands Twin Register (NTR) and Netherlands Study of Depression and Anxiety (NESDA).[18,19,33–35] Subjects from these studies, not necessarily the same ones, also participated in the GWAS for some of the traits examined.[22,23,25,27–29] To ensure that the discovery sample did not affect estimates of enrichments of eQTL effects in the various GWAS signals, we looked at trait-specific enrichment of blood and brain eQTL signal in GWAS signal for educational attainment and smoking behavior. We compared the results from using publicly available datasets with using summary statistics based on the same sample without subjects from the NTR or NESDA. The results did not reveal appreciable differences between the respective datasets (S1 Figure), therefore the remaining analyses for all traits with participants from the NTR or NESDA (age at menarche, BMI, educational attainment, height, LDL level, and smoking behavior) were performed using publicly available summary statistics.

Blood and brain eQTL enrichment

A catalog of whole blood *cis*-eQTLs was obtained from Jansen *et al*,[18,19] where the eQTL most strongly associated with gene expression in whole blood for each probe set was selected for inclusion in our whole blood eQTL annotation. A list of brain eQTLs was obtained from Ramasamy *et al*,[17] whom studied eQTL effects in 12 distinct brain regions. We included all eQTLs which were found to significantly associate to average gene expression across all 12 brain tissues. SLDSR annotations were constructed as per the instructions of Bulik-Sullivan and Finucane.[36] To guard against upward bias in the eQTL enrichment signal, two extra annotations containing SNPs within a 500 base pair (bp)

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and 100bp window around any eQTL were constructed for each eQTL set.[15] Finally, to ensure that the enrichment of eQTL effects in GWAS signal was not in fact caused by their proximity to the genes they influence, six additional "gene-centric" annotations were constructed per eQTL set. These annotations included all SNPs that were located in a gene promoter, coding region, or intron and their corresponding 500bp windows that were also within 1Mbp window from any blood or brain eQTL.

Tissue-specific eQTL enrichment

The availability of eQTLs from two tissues allowed for the consideration of eQTLs unique to either blood or brain gene expression, and those that overlap between the two tissues. We modeled the LD between the SNPs of the blood and brain eQTL annotations as a mixture of normal densities. Subsequently, we determined the tissue specificity of each eQTL based on the posterior probability of belonging to a class of SNPS in weak LD with the eQTLs in the other tissue (see appendix A1 for a full model description and results). Based on the split sets of eQTLs, we constructed two additional LD score annotations. The first annotation contained all eQTLs uniquely found in whole blood, and the second contained all eQTLs unique to brain. Here, we assume LD to be a good proxy for the co-localization of eQTLs. Strong LD does not guaranty that the blood and brain eQTL in strong LD tag the same underlying causal SNP effect on gene expression. However, given that our research question focuses on genome-wide enrichment, this metric of co-localization between brain and blood eQTLs was the optimal approach.

In the presence of tissue-specific eQTL effects, we expect eQTLs *unique* to a tissue to not be differentially enriched when compared to all eQTLs found in a tissue when the trait is related to that particular tissue (e.g. blood eQTLs are enriched in their effects on rheumatoid arthritis, regardless of their LD with brain eQTLs). We further expect eQTLs *unique* to a tissue to be *depleted* in their effect on traits not related to that particular tissue, when compared to eQTLs also found in a trait-relevant

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

tissue (e.g. blood eQTLs are expected to have an effect on educational attainment, but only if they are in strong LD with brain eQTLs).

Enrichment of eQTLs from 44 tissues

There are two limitations to above mentioned analyses of tissue-specific enrichments of eQTL effects in GWAS signal. First, the eQTLs are obtained from two different projects, which vary in terms of sample size and their definition of an eQTL. Second, both projects identified lead eQTLs (i.e. the statistically strongest eQTL out of many significant eQTL SNPs), which means annotations based on these lists of hits contained a very small portion of the SNPs in the genome, possibly limiting power in SLDSR. To mitigate these limitations, we performed additional analyses using eQTLs obtained by a common pipeline from 44 tissues (see S2 Table) and based on a broader eQTL locus definition.[10] For each of the 44 tissues, we made an annotation for analysis in SLDSR following the previously described procedure. Analogous to the procedure of Finucane *et al* for cell-type-specific analysis using SLDSR,[15] we additionally made an annotation that contained all GTEx eQTLs, i.e. a SNP would be included in this annotation if it was designated as part of at least one of the 44 tissue-specific GTEx annotations, and added a 100bp and 500bp window. No windows were made for the tissue-specific GTEx annotations. Using GWAS summary statistics for schizophrenia, we then ran one SLDSR model containing only the baseline categories and the union of GTEx eQTLs, and 44 additional models with the two previous annotations and one of the tissue-specific GTEx annotations at a time. We repeated the procedure using summary statistics for rheumatoid arthritis.

GTEx has relative small sample sizes for the brain eQTL discovery (mean=89 sample size, range=72-103) compared to other tissues (mean=160 sample size, range=70-361).[10] To investigate the effect of differences in sample size on estimates of enrichments in GWAS signal, we collapsed the union of individual brain eQTL annotations into a shared brain eQTL annotation (i.e. an eQTL found in at least one of the GTEx brain annotations was included in the shared brain eQTL

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

annotation). This annotation was then analyzed as an additional GTEx eQTL annotation. We further tested the relationship between tissue sample size and tissue eQTL enrichment.

Enrichment of the intersection between eQTLs and histone marks

The availability of annotations based on tissue-specific histone marks made it possible to create an annotation that represents the intersection between eQTLs and epigenetic modification related to enhancers and promoters of actively transcribed genes. We obtained LD score annotations of SNPs in regions that bare histone marks in cells from the CNS or immune system from Finucane *et al*.[15] Out of the 220 available cell-type-specific histone mark, 101 were found in the CNS or immune tissues. For each of the 101 annotations of SNPs in cell-type-specific histone marks, we extracted its intersection with the union of GTEx eQTLs and made a new annotation of eQTLs which intersected with histone marks (i.e. SNPs found in both annotations). We then analyzed each of the intersection annotations individually in a model together with the baseline categories, the union of GTEx eQTLs, and the corresponding cell-type-specific histone marks. Enrichments in GWAS signal of the intersection should be interpreted as enrichment of genome-wide SNP effects on a complex trait beyond the additive effects which work on all SNPs that are a *cis*-eQTL and histone mark in question. In fact, we test whether the interaction between tissue-specific chromatin state and eQTLs are enriched in their genome-wide effect on complex traits.

Results

Blood and brain eQTL enrichment

We fitted an SLDSR model containing the baseline categories; the complete annotation for both brain and blood eQTL tissues, their 100 and 500bp windows, and gene-centric annotations to all 11 GWAS traits (Crohn's disease, rheumatoid arthritis, ulcerative colitis, BMI, educational attainment, schizophrenia, age at menarche, coronary artery disease, height, LDL levels, and smoking behavior). Our analyses revealed substantial enrichment of both blood (32-113 fold enrichment, mean=59) and

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

brain (21-63, mean=40) eQTL effects in GWAS signals for all traits (Figure 1, S3 Table), exceeding enrichments of any of the baseline categories considered by Finucane *et al.*[15] Additionally, the coding region gene-centric annotation for both blood and brain eQTLs showed considerable enrichment in GWAS signal. Similar to Finucane *et al*,[15] the average enrichment in the "Enhancer Andersson"-annotation (called "FANTOM5 enhancer" in Finucane *et al*) is largely determined by the relatively high enrichments of GWAS signal in the three immune-related traits (red dots in Figure 1). These traits behave similarly in blood eQTLs and its gene-centric annotations, but not in brain eQTLs. Further inspection showed additional differences between the brain and blood eQTL sets in the three immune traits: in these traits, blood eQTLs were roughly 73—113 times enriched in their effect on GWAS signal, but only showed a 21-38-fold enrichment in GWAS signal for brain eQTLs (see S4 Table). The distinction between the two eQTL sets was less prominent for brain-related traits: brain eQTLs were 57 times enriched in their effect on educational attainment, and only showed a 34-37-fold enrichment for schizophrenia and BMI. Blood eQTLs showed a 38-45-fold enriched effect. The brain traits seem to behave similarly to traits without prior expectation on tissue-specific eQTL enrichment for either blood or brain (32-65 and 30-63, respectively).

Where the significance level associated with the enrichments do not account for the presence of the other annotations in the model, the significance test associated with the coefficient is corrected for the influence of the additional annotations. The Z-score for the coefficient allows us to determine the enrichment of SNPs in a particular annotation in their effect on complex traits is significant, conditional on the other annotation in the model. Results are largely as expected: blood eQTLs contribute substantially to h^2_{SNP} in rheumatoid arthritis and ulcerative colitis (though not statistically significant); whereas brain eQTLs do the same for all three brain-related traits (see supplemental S5 Table). The blood eQTL annotation further reached statistical significance for schizophrenia and BMI, while the same was seen for brain eQTLs in Crohn's disease.

Enrichment of eQTLs unique to, and shared between blood and brain tissue

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We fitted an SLDSR model for all 11 GWAS traits, containing the baseline categories, the complete blood eQTL annotation, its corresponding 100 and 500bp windows, and gene-centric annotations, as well as unique blood eQTLs (blood eQTLs not identified in the brain). The same was done using the complete brain eQTL annotation, and corresponding windows and gene-centric annotations. By definition, all the SNPs within the unique eQTL annotations are also part of their respective complete eQTL annotation. Adding both annotations simultaneously to the model will result in an estimate of enrichment or depletion of *cis*-eQTLs *unique* to a tissue on top of the enrichment of all eQTLs in the tissue.

Results for models which include both the complete blood eQTL annotation and the unique blood eQTLs are in table 1. For all the immune-related traits, unique blood eQTLs were neither enriched nor depleted in their genome-wide effects beyond the complete blood eQTL annotation. For the brain-related traits, unique blood eQTLs showed depletion in GWAS signal for schizophrenia and educational attainment, but not for BMI.

Unique brain eQTLs were not significantly depleted compared to brain eQTLs in strong LD with blood eQTLs in their effect on all three immune-traits (Table 1 left panel). Unique brain eQTLs were significantly depleted in their effect on schizophrenia and educational attainment compared to the complete brain eQTL annotation, but not for BMI. Informed by the unexpected depletion of enrichment in GWAS signal for unique brain eQTLs in brain-related traits, but not immune traits, we tested whether brain eQTLs in strong LD with blood eQTLs influenced gene expression in more brain regions than unique brain eQTLs. On average, brain eQTLs in strong LD with blood eQTLs were found to influence gene expression in 2.81 regions, whereas unique brain eQTLs influenced gene expression in 2.43 brain regions ($t=4.20$, $df=1967$, $p<3*10^{-5}$). These results indicated that cross-tissue LD between blood and brain eQTLs was weakly informative on cross-tissue eQTL effects across different brain regions.

Enrichment of eQTLs from 44 tissues in GTEx

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We interrogate the enrichment of the union of GTEx eQTLs and 44 tissue-specific GTEx annotations in their effect on schizophrenia and rheumatoid arthritis. Figure 2 shows the coefficient Z-scores of the 45 GTEx annotations, sorted from largest to smallest. In both cases, the union of GTEx eQTLs has a substantial Z-score (Z=5.501 and Z=3.802 for schizophrenia and rheumatoid arthritis, respectively, both $p<0.001$, S6 Table), indicating that eQTLs in general are significantly enriched in their effects on complex traits. The tissue-specific annotations, however, perform notably worse and in some cases even suggest depletion of genome-wide effects of tissue-specific eQTLs on schizophrenia and rheumatoid arthritis. For rheumatoid arthritis, the coefficient Z-scores of the whole blood annotation reached nominal significance (Z=2.036, $p=0.021$), but failed correction for multiple testing. None of the other annotations reached nominal significance. The union of all GTEx brain annotations did not contribute significantly to explaining h^2_{SNP} (Z=0.147, $p=0.441$). Sample size in the eQTL discovery phase appears to be a strong determinant of tissue-specific enrichment in GWAS signal. The correlation coefficients between the coefficient Z-scores and sample sizes were 0.6453 ($p=2.253*10^{-6}$) and 0.4247 ($p=0.004$) for schizophrenia and rheumatoid arthritis, respectively.

Enrichment of the intersection between eQTLs and histone marks

We interrogate the intersection of eQTLs and histone marks found in specific CNS and immune cells, and estimate the enrichment of the intersection in its effect on rheumatoid arthritis and schizophrenia. We find significant enrichment in GWAS signal for eQTLs that intersect with histones that bare modification H3K4me1, a modification thought to be present in the enhancer of actively transcribed genes,[37,38] in CNS cells for schizophrenia (see Figure 3). The results are further idiosyncratic in that there is some evidence for significant enrichment of eQTLs that intersect with immune cells baring the H3K4me1 mark in their effect on schizophrenia, but not on rheumatoid arthritis. Specifically, none of the intersecting annotations show evidence of enrichment for rheumatoid arthritis. For the separate annotations, we find significant enrichment in GWAS signal across all histone marks found in CNS cells and three significant immune cell-types that bare the

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

H3K4me3 modification, a modification associated with transcriptional start sites and promoters of actively transcribed genes,[37,38] for schizophrenia (S7 Figure). The opposite picture was seen for rheumatoid arthritis: a wide variety of immune-cell specific histone marks showed significant enrichments in GWAS signal, while all marks found in CNS cells were below zero. The union of GTEx eQTLs reached statistical significance for all models (S8 Figure).

Discussion

Stratified Linkage Disequilibrium Score Regression provides a way to partition h^2_{SNP} into parts explained by (functional) parts of the genome.[15] A "full baseline model" containing 24 non-cell-type-specific annotations of SNPs, such as SNPs located in promoters or coding regions, was developed previously for analysis using SLDSR. Here, we go beyond the baseline categories by adding annotations containing eQTLs derived from whole blood and brain tissue into the model, and showed that eQTLs were substantially stronger enriched in their effect on complex traits compared to all categories considered by Finucane *et al*.[15] The complete brain eQTL annotation was significantly enriched in GWAS signal for BMI and schizophrenia, and was consistent with previous estimates of eQTL effect enrichment.[8] Considerable enrichment for eQTLs, even for traits not apparently linked to the brain or immune system (e.g. coronary artery disease), suggested that non-trivial eQTL overlap across tissues might be present.

Inclusion of both brain and blood eQTLs into the SLDSR model did not fully separate the signal into tissue-specific effects, as shown by the roughly equal enrichments of both annotations within brain-related traits. Using LD scores to inform a mixture-based separation of the eQTL sets yielded an incomplete separation of tissue-specific enrichment. In general, we are not able to clearly identify tissue-specific eQTL signals using these datasets and SLDSR. Our second analysis of eQTL enrichment based on 44 tissue-specific *cis*-eQTL sets, obtained from the GTEx consortium, confirms the lack of tissue-specific eQTL enrichment. While an annotation containing all eQTLs identified in GTEx is significantly enriched (Z=5.501 and Z=3.802 for schizophrenia and rheumatoid arthritis,

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

respectively, both $p<0.001$), none of the analyzed brain tissues are enriched beyond all eQTLs in their effect on schizophrenia. Similarly, whole blood eQTLs are not significantly enriched beyond all GTEx eQTLs taken together in their effect on rheumatoid arthritis. Again, these findings are not consistent with the hypothesis of abundant tissue-specific eQTLs with effects on complex traits related to the specific tissue in question. Especially, when contrasted with tissue-specific gene expression levels and tissue-specific histone modifications,[13,39] tissue-specific eQTLs are of limited value in relating complex traits to a tissue. We do, however, find evidence for possible enrichment for eQTLs that intersect with tissue-specific H3K4me1 histone marks in the brain, but also immune cells, in their effect on schizophrenia but not rheumatoid arthritis. This means that eQTLs in H3k4me1 marks are enriched in their effect on schizophrenia above the expected enrichment based on the fact that these SNPs are both eQTLs and located in H3K4me1 histone marks. What is of substantial interest is that the enrichment in GWAS signal appears specific to H3K4me1 marks, suggesting that these marks specifically can aid in prioritizing genomic regions in which tissue-specific eQTLs may reside. Though, again, the totality of evidence is inconclusive on the relevance of tissue-specific eQTLs to variation in complex traits.

Our results are consistent with, and complimentary to, a study investigating the genetic correlation between gene expression levels across 15 tissues.[13] This study revealed substantial correlations between *cis*-genetic effects on gene expression across 15 tissues.[13] Our analyses confirm the value of using whole blood as discovery tissue for detection of eQTLs and further demonstrate the usefulness of techniques that use eQTLs discovered in whole blood to study the etiology of complex traits related to different tissues.[40,41] The results presented here highlight the overlap of *cis*-eQTL effects across tissues on a genome-wide level. However, the effect of a *cis*-eQTL might vary substantially across tissues for individual genes.[42] Our conclusions are based on genome-wide enrichments and therefore should not be interpreted as limited evidence for tissue-specific eQTL effects for individual genes. Therefore, eQTL discovery in the tissue most relevant to a specific trait or disorder remains important to further our understanding of the genetic regulation of tissue-

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

specific gene expression. What is also clear is that to find those tissue-specific eQTLs that are of relevance to the interpretation of GWASs of complex traits, tissue-specific eQTL discovery needs to be refined. The practice of, as a post-hoc analysis to GWAS, performing eQTL lookup in a specific tissue linked to a trait, when larger dataset for other accessible tissues are available may be sub optimal. In fact one may prefer to perform a lookup in the overlap between epigenetic modifications in a relevant tissue and eQTLs regardless of tissue. One can further consider utilizing eQTLs to link GWAS findings to a gene, and subsequently consider the differential expression of a gene to identify the tissue in which the gene is most likely to act in effecting the trait. Tissue-specific differential gene expression vastly outperforms eQTLs in tagging regions of the genome enriched in their effect on complex traits.[39]

It is also evident that a limited dichotomous definition of eQTL/no-eQTL may be insufficient to identify tissue-specific eQTL effects. An evident improvement would be to compute the difference in eQTL effect between tissues, and perform inference based on this difference in effect. eQTLs are strongly enriched SNPs, with clear biological function and utility for the translation of GWAS findings, though tissue-specific eQTL mechanisms remain elusive. The discovery of tissue-specific eQTL effects, which can aid in linking complex trait to tissue, may require novel research strategies.

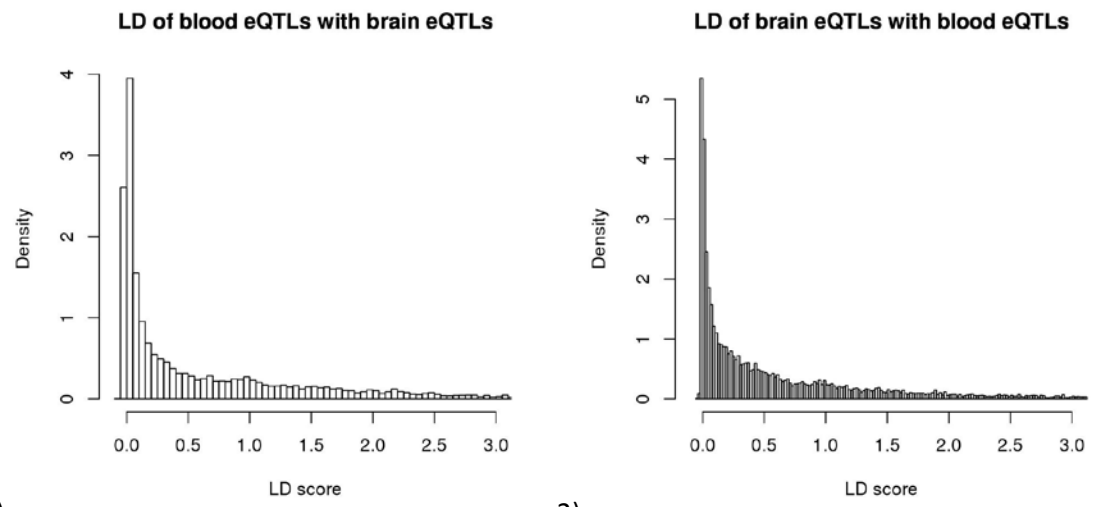
Appendices

A1. LD-based distinction between tissue-specific eQTLs and eQTLs in strong LD with eQTLs from the other tissue.

We attempted to separate the signal from all blood eQTLs collectively by looking at the degree of LD each blood eQTL has with all brain eQTLs. Fig A1 shows the distribution of LD scores for all blood eQTLs. A strong mass can be seen around 0 LD (eQTLs unique to whole blood) with a long tail to the right (blood eQTLs in [strong] LD with one or more brain eQTL[s]). To determine the cutoff value between unique blood eQTLs and those in (strong) LD with brain eQTLs, we analyzed the distribution of LD scores using the *normalmixEM*-function of the mixtools package in R.[43] This function

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

approximates the distribution of the LD scores using a mixture of a user-defined number of normal distributors. Based on their posterior probabilities, each eQTL is then assigned to one of the distributors. A five-class normal mixture distribution returned two components with a mean of roughly zero and three components that clearly deviated from zero (Table A). Fitting more than five mixture classes did not result in an appreciable different separation of blood eQTLs with near zero LD with brain eQTLs. Therefore, the results from fitting a five-class mixture were used to make the distinction, where the SNPs that had the highest posterior probability of belonging to one of the components with a mean LD near zero, i.e. the first two components, were considered unique blood eQTLs. Fig A2 shows the distribution of LD scores for brain eQTLs. Again, a strong mass around zero with a long tail to the right can be seen and fitting a mixture of five normal distributors resulted in optimal separation of unique brain eQTLs and brain eQTLs in (strong) LD with blood eQTLs (Table B). SNPs with the highest posterior probability of belonging to the first two components were categorized as unique brain eQTLs.



1) 2)

Figure A. Distribution of LD scores. a) LD scores of blood eQTLs with brain eQTLs. b) LD scores of brain eQTLs with blood eQTLs.

Table A. Results mixture modeling blood eQTLs.

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

	comp 1	comp 2	comp 3	comp 4	comp 5
Lambda	0,159244	0,220276	0,202397	0,346083	0,071999
Mu	0,000063	0,037581	0,239015	1,267693	4,871118
Sigma	0,004075	0,030574	0,137518	0,716810	3,328837
loglik at estimate: -4933,98					

	comp 1	comp 2	comp 3	comp 4	comp 5	comp 6
Lambda	0,145539	0,207641	0,186234	0,221515	0,196115	0,042957
Mu	-0,000245	0,029492	0,172924	0,738651	1,931109	6,179965
sigma	0,003462	0,025442	0,097901	0,353329	0,868389	3,674714
loglik at estimate: -4526,028						

	comp 1	comp 2	comp 3	comp 4	comp 5	comp 6	comp 7
lambda	0,088138	0,152810	0,164436	0,161404	0,177390	0,215443	0,040379
Mu	-0,000528	0,007292	0,058076	0,231495	2,039302	0,834543	6,354239
sigma	0,001545	0,011553	0,035202	0,114845	0,890447	0,374960	3,709027
loglik at estimate: -4253,795							

Lambda indicates the proportion of SNPs falling within the component. Mu signifies the mean LD score. Sigma represents the standard error around mu.

Table B. Results mixture modeling brain eQTLs.

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

	comp 1	comp 2	comp 3	comp 4	comp 5
lambda	0,139315	0,180625	0,276473	0,273498	0,130090
mu	0,000225	0,045226	0,279774	0,986414	2,825510
sigma	0,005331	0,033993	0,152244	0,466854	1,516570
loglik at estimate: -6336,214					

	comp 1	comp 2	comp 3	comp 4	comp 5	comp 6
lambda	0,123778	0,158111	0,156313	0,237922	0,205140	0,118735
mu	-0,000392	0,030385	0,143963	1,106905	0,407980	2,935880
sigma	0,004401	0,025348	0,071816	0,481418	0,171124	1,529066
loglik at estimate: -6068,594						

	comp 1	comp 2	comp 3	comp 4	comp 5	comp 6	comp 7
lambda	0,121341	0,150736	0,155579	0,199655	0,121956	0,204690	0,046044
mu	-0,000462	0,133097	0,028572	0,374076	2,045583	0,938595	3,956472
sigma	0,004272	0,066737	0,024453	0,153883	0,792295	0,366171	1,746882
loglik at estimate: -5907,881							

Lambda indicates the proportion of SNPs falling within the component. Mu signifies the mean LD score. Sigma represents the standard error around mu.

Supplemental Data

Supplemental Data includes 3 figures and 5 tables

Acknowledgements

MGN is supported by the Royal Netherlands Academy of Science Professor Award (PAH/6635) to DIB. HFI is supported by the “Aggression in Children: Unraveling gene-environment interplay to inform Treatment and InterventiON strategies” (ACTION) project. ACTION receives funding from the European Union Seventh Framework Program (FP7/2007-2013) under grant agreement no 602768. MB is supported by a University Research Chair of the Vrije Universiteit. The discovery of blood eQTL was funded by the US National Institute of Mental Health (RC2 MH089951, principal Investigator PFS) as part of the American Recovery and Reinvestment Act of 2009. We thank T. Lehner (National

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Institute of Mental Health) for his support. We acknowledge Hillary Finucane, Raymond Walters and Benjamin Neale for critical comments on our methods, design and manuscript.

Web Resources

Age at menarche summary statistics, www.reprogen.org/data_download.html

Blood eQTLs, https://eqtl.onderzoek.io/

Brain eQTLs:, http://www.braineac.org/

Coronary artery disease summary statistics, www.cardiogramplusc4d.org/data-downloads/

Crohn's disease and ulcerative colitis summary statistics, www.ibdgenetics.org/downloads.html

Educational attainment summary statistics, http://www.thessgac.org/data

Full baseline model LD scores, http://data.broadinstitute.org/alkesgroup/LDSCORE/

GTEx dataset, http://www.gtexportal.org/home/datasets

Height and BMI summary statistics,
www.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files

LDL levels summary statistics, www.broadinstitute.org/mpg/pubs/lipids2010/

OMIM, http://www.omim.org

Rheumatoid arthritis summary statistics, http://plaza.umin.ac.jp/yokada/datasource/software.htm

Schizophrenia and smoking behavior summary statistics, www.med.unc.edu/pgc/results-and-downloads

SLDSR software, https://github.com/bulik/ldsc/

References

1. Hirschhorn, J.N., and Daly, M.J. (2005). Genome-wide association studies for common diseases and complex traits. Nat. Rev. Genet. *6*, 95–108.

2. Visscher, P.M., Brown, M.A., McCarthy, M.I., and Yang, J. (2012). Five Years of GWAS Discovery. Am.

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

J. Hum. Genet. *90*, 7–24.

3. Lowe, W.L., and Reddy, T.E. (2015). Genomic approaches for understanding the genetics of complex disease. Genome Res. *25*, 1432–1441.

4. Morley, M., Molony, C.M., Weber, T.M., Devlin, J.L., Ewens, K.G., Spielman, R.S., and Cheung, V.G. (2004). Genetic analysis of genome-wide variation in human gene expression. Nature *430*, 743–747.

5. Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., et al. (2009). Finding the missing heritability of complex diseases. Nature *461*, 747–753.

6. Nicolae, D.L., Gamazon, E., Zhang, W., Duan, S., Dolan, M.E., and Cox, N.J. (2010). Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. PLoS Genet. *6*, e1000888.

7. Torres, J.M., Gamazon, E.R., Parra, E.J., Below, J.E., Valladares-Salgado, A., Wacher, N., Cruz, M., Hanis, C.L., and Cox, N.J. (2014). Cross-tissue and tissue-specific eQTLs: partitioning the heritability of a complex trait. Am. J. Hum. Genet. *95*, 521–534.

8. Davis, L.K., Yu, D., Keenan, C.L., Gamazon, E.R., Konkashbaev, A.I., Derks, E.M., Neale, B.M., Yang, J., Lee, S.H., Evans, P., et al. (2013). Partitioning the heritability of Tourette syndrome and obsessive compulsive disorder reveals differences in genetic architecture. PLoS Genet. *9*, e1003864.

9. Hernandez, D.G., Nalls, M.A., Moore, M., Chong, S., Dillman, A., Trabzuni, D., Gibbs, J.R., Ryten, M., Arepalli, S., Weale, M.E., et al. (2012). Integration of GWAS SNPs and tissue specific expression profiling reveal discrete eQTLs for human traits in blood and brain. Neurobiol. Dis. *47*, 20–28.

10. GTEx Consortium (2015). The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. Science (80-.). *348*, 648–660.

11. Ding, J., Gudjonsson, J.E., Liang, L., Stuart, P.E., Li, Y., Chen, W., Weichenthal, M., Ellinghaus, E., Franke, A., Cookson, W., et al. (2010). Gene expression in skin and lymphoblastoid cells: Refined statistical method reveals extensive overlap in cis-eQTL signals. Am. J. Hum. Genet. *87*, 779–789.

12. Nica, A.C., Parts, L., Glass, D., Nisbet, J., Barrett, A., Sekowska, M., Travers, M., Potter, S.,

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Grundberg, E., Small, K., et al. (2011). The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. PLoS Genet. *7*, e1002003.

13. Liu, X., Finucane, H.K., Gusev, A., Bhatia, G., Gazal, S., O'Connor, L., Bulik-Sullivan, B., Wright, F., Sullivan, P., Neale, B., et al. (2016). Functional partitioning of local and distal gene expression regulation in multiple human tissues. bioRxiv.

14. Bulik-Sullivan, B.K., Loh, P.-R., Finucane, H.K., Ripke, S., Yang, J., Patterson, N., Daly, M.J., Price, A.L., and Neale, B.M. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat. Genet. *47*, 291–295.

15. Finucane, H.K., Bulik-Sullivan, B., Gusev, A., Trynka, G., Reshef, Y., Loh, P.-R., Anttila, V., Xu, H., Zang, C., Farh, K., et al. (2015). Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat. Genet. *47*, 1228–1235.

16. Brem, R.B., Yvert, G., Clinton, R., and Kruglyak, L. (2002). Genetic dissection of transcriptional regulation in budding yeast. Science *296*, 752–755.

17. Ramasamy, A., Trabzuni, D., Guelfi, S., Varghese, V., Smith, C., Walker, R., De, T., Coin, L., de Silva, R., Cookson, M.R., et al. (2014). Genetic variability in the regulation of gene expression in ten regions of the human brain. Nat. Neurosci. *17*, 1418–1428.

18. Jansen, R., Hottenga, J.-J., Nivard, M.G., Abdellaoui, A., Laport, B., de Geus, E.J., Wright, F.A., Penninx, B.W., and Boomsma, D.I. (2017). Conditional eQTL Analysis Reveals Allelic Heterogeneity of Gene Expression. Hum. Mol. Genet.

19. Wright, F.A., Sullivan, P.F., Brooks, A.I., Zou, F., Sun, W., Xia, K., Madar, V., Jansen, R., Chung, W., Zhou, Y.-H., et al. (2014). Heritability and genomics of gene expression in peripheral blood. Nat. Genet. *46*, 430–437.

20. Jostins, L., Ripke, S., Weersma, R.K., Duerr, R.H., McGovern, D.P., Hui, K.Y., Lee, J.C., Schumm, L.P., Sharma, Y., and Anderson, C.A. (2012). Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature *491*, 119–124.

21. Okada, Y., Wu, D., Trynka, G., Raj, T., Terao, C., Ikari, K., Kochi, Y., Ohmura, K., Suzuki, A., and

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Yoshida, S. (2014). Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature *506*, 376–381.

22. Speliotes, E.K., Willer, C.J., Berndt, S.I., Monda, K.L., Thorleifsson, G., Jackson, A.U., Allen, H.L., Lindgren, C.M., Luan, J., and Mägi, R. (2010). Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat. Genet. *42*, 937–948.

23. Rietveld, C.A., Medland, S.E., Derringer, J., Yang, J., Esko, T., Martin, N.W., Westra, H.-J., Shakhbazov, K., Abdellaoui, A., and Agrawal, A. (2013). GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. Science (80-.). *340*, 1467–1471.

24. Ripke, S., Neale, B.M., Corvin, A., Walters, J.T.R., Farh, K.-H., Holmans, P.A., Lee, P., Bulik-Sullivan, B., Collier, D.A., Huang, H., et al. (2014). Biological insights from 108 schizophrenia-associated genetic loci. Nature *511*, 421–427.

25. Perry, J.R.B., Day, F., Elks, C.E., Sulem, P., Thompson, D.J., Ferreira, T., He, C., Chasman, D.I., Esko, T., and Thorleifsson, G. (2014). Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. Nature *514*, 92–97.

26. Schunkert, H., König, I.R., Kathiresan, S., Reilly, M.P., Assimes, T.L., Holm, H., Preuss, M., Stewart, A.F.R., Barbalic, M., and Gieger, C. (2011). Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat. Genet. *43*, 333–338.

27. Allen, H.L., Estrada, K., Lettre, G., Berndt, S.I., Weedon, M.N., Rivadeneira, F., Willer, C.J., Jackson, A.U., Vedantam, S., and Raychaudhuri, S. (2010). Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature *467*, 832–838.

28. Teslovich, T.M., Musunuru, K., Smith, A. V, Edmondson, A.C., Stylianou, I.M., Koseki, M., Pirruccello, J.P., Ripatti, S., Chasman, D.I., and Willer, C.J. (2010). Biological, clinical and population relevance of 95 loci for blood lipids. Nature *466*, 707–713.

29. Furberg, H., Kim, Y., Dackor, J., Boerwinkel, E., Franceshini, N., Ardissino, D., Bernardinelli, L., Mannucci, P.M., Maur, F., Merlini, P.A., et al. (2010). Genome-wide meta-analyses identify multiple loci associated with smoking behavior. Nat. Genet. *42*, 441–447.

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

30. Vimaleswaran, K.S., Tachmazidou, I., Zhao, J.H., Hirschhorn, J.N., Dudbridge, F., and Loos, R.J.F. (2012). Candidate genes for obesity-susceptibility show enriched association within a large genome-wide association study for BMI. Hum. Mol. Genet. *21*, 4537–4542.

31. Andreassen, O.A., Harbo, H.F., Wang, Y., Thompson, W.K., Schork, A.J., Mattingsdal, M., Zuber, V., Bettella, F., Ripke, S., Kelsoe, J.R., et al. (2015). Genetic pleiotropy between multiple sclerosis and schizophrenia but not bipolar disorder: differential involvement of immune-related gene loci. Mol. Psychiatry *20*, 207–214.

32. Karalis, K.P., Giannogonas, P., Kodela, E., Koutmani, Y., Zoumakis, M., and Teli, T. (2009). Mechanisms of obesity and related pathology: linking immune responses to metabolic stress. FEBS J. *276*, 5747–5754.

33. Willemsen, G., De Geus, E.J.C., Bartels, M., Van Beijsterveldt, C.E.M., Brooks, A.I., Estourgie-van Burk, G.F., Fugman, D.A., Hoekstra, C., Hottenga, J.-J., and Kluft, K. (2010). The Netherlands Twin Register biobank: a resource for genetic epidemiological studies. Twin Res. Hum. Genet. *13*, 231–245.

34. Willemsen, G., Vink, J.M., Abdellaoui, A., den Braber, A., van Beek, J.H.D.A., Draisma, H.H.M., van Dongen, J., van't Ent, D., Geels, L.M., and van Lien, R. (2013). The Adult Netherlands Twin Register: twenty-five years of survey and biological data collection. Twin Res. Hum. Genet. *16*, 271–281.

35. Penninx, B.W.J.H., Beekman, A.T.F., Smit, J.H., Zitman, F.G., Nolen, W.A., Spinhoven, P., Cuijpers, P., De Jong, P.J., Van Marwijk, H.W.J., and Assendelft, W.J.J. (2008). The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods. Int. J. Methods Psychiatr. Res. *17*, 121–140.

36. Bulik-Sullivan, B.K., and Finucane, H.K. (2015). LD Score Estimation Tutorial.

37. Allis, C.D., and Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. Nat Rev Genet *17*, 487–500.

38. Zhou, V.W., Goren, A., and Bernstein, B.E. (2011). Charting histone modifications and the functional organization of mammalian genomes. Nat Rev Genet *12*, 7–18.

39. Finucane, H., Reshef, Y., Anttila, V., Slowikowski, K., Gusev, A., Byrnes, A., Gazal, S., Loh, P.-R., Genovese, G., Saunders, A., et al. (2017). Heritability enrichment of specifically expressed genes

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

identifies disease-relevant tissues and cell types. bioRxiv.

40. Gamazon, E.R., Wheeler, H.E., Shah, K.P., Mozaffari, S. V, Aquino-Michaels, K., Carroll, R.J., Eyler, A.E., Denny, J.C., Nicolae, D.L., Cox, N.J., et al. (2015). A gene-based association method for mapping traits using reference transcriptome data. Nat. Genet. *47*, 1091–1098.

41. Gusev, A., Ko, A., Shi, H., Bhatia, G., Chung, W., Penninx, B.W.J.H., Jansen, R., de Geus, E.J.C., Boomsma, D.I., Wright, F.A., et al. (2016). Integrative approaches for large-scale transcriptome-wide association studies. Nat. Genet. *48*, 245–252.

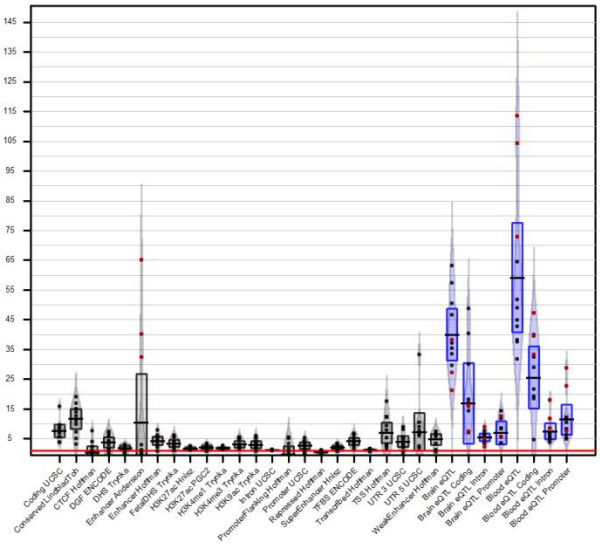
42. Grundberg, E., Small, K.S., Hedman, Å.K., Nica, A.C., Buil, A., Keildson, S., Bell, J.T., Yang, T.-P., Meduri, E., Barrett, A., et al. (2012). Mapping cis- and trans-regulatory effects across multiple tissues in twins. Nat. Genet. *44*, 1084–1089.

43. Benaglia, T., Chauveau, D., Hunter, D., and Young, D. (2009). mixtools: An R Package for Analyzing Finite Mixture Models. J. Stat. Softw. *32*, 1–29.

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure Titles and Legends

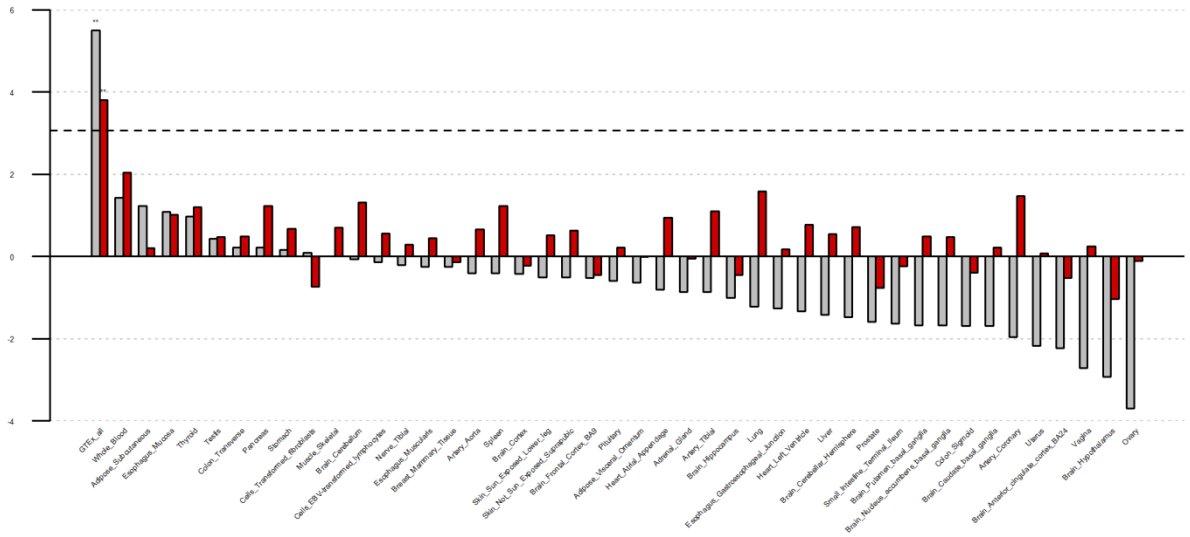
Figure 1. Average enrichment in GWAS signal of the 24 baseline annotations, 4 brain eQTL annotations and 4 blood eQTL annotations.



Violin plot of the average enrichment in GWAS signal across all traits for the 24 main baseline annotations and 8 main eQTL annotations. Grey beans represent the baseline categories. Blue beans represent eQTLs. Black bars indicate average enrichment. Boxes show upper- and lower-bounds of the 95% confidence interval of the mean. Red dots show enrichments for immune-related traits. Horizontal red line indicates enrichment of 1, i.e. no enrichment.

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

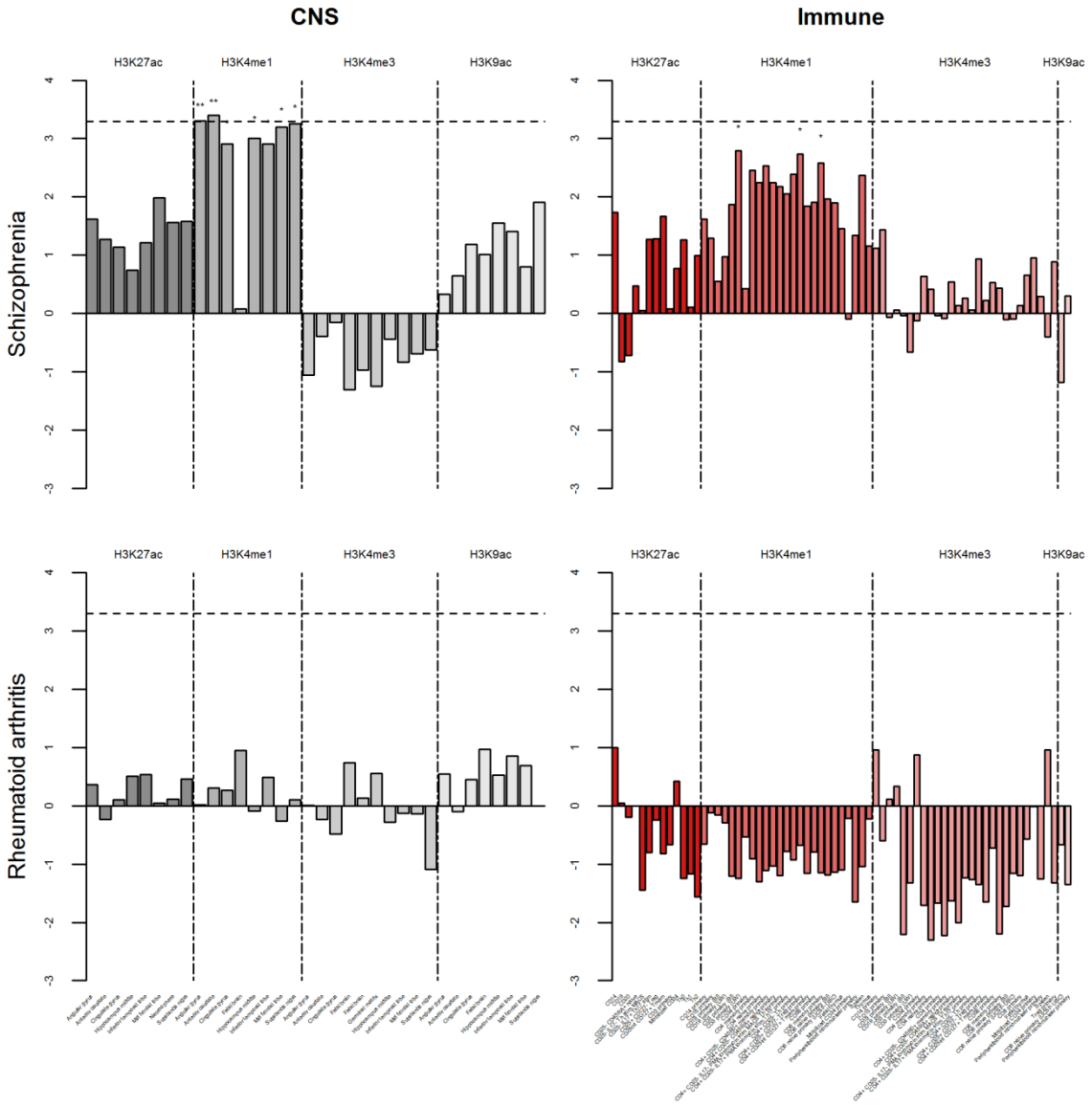
Figure 2. Coefficient Z-scores of the 45 GTEx annotations



Barplot of coefficient z-scores for all GTEx annotations for schizophrenia (grey) and rheumatoid arthritis (red). Bars are sorted from highest to lowest based on the results from schizophrenia. Horizontal dotted line indicates Bonferroni threshold for 45 tests. Two asterisks indicate bars passing Bonferroni correction for multiple testing.

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 3. Coefficient Z-score of intersection between union of GTEx eQTLs and cell-type-specific histone marks



Top two graphs show coefficient Z-scores for schizophrenia. Bottom two graphs show the same for rheumatoid arthritis. Grey bars indicate histone marks found in cells from the central nervous system. Red bars represent histone marks found in cells from the immune system. From dark to light, shades of the bars indicate histone marks H3K27ac, H3K4me1, H3K4me3, and H3K9ac. Vertical dotted lines indicate separation between histone marks. One asterisk above the bars indicate annotations passing FDR correction for multiple testing. Two asterisks indicate bars passing

bioRxiv preprint doi: https://doi.org/10.1101/107482; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Bonferroni correction for multiple testing. Horizontal dotted line indicates Bonferroni threshold for 101 tests.

Tables

Table 1. Coefficient Z-scores of simultaneously modelling complete and unique eQTLs

	Trait	Complete brain eQTL	Unique brain eQTL	Complete blood eQTL	Unique blood eQTL
Immune	Crohn's disease	3.257 (<0.001)**	0.049 (0.520)	1.593 (0.056)	0.311 (0.622)
	Rheumatoid arthritis	2.474 (0.007)*	-1.576 (0.057)	2.297 (0.011)*	-0.243 (0.404)
	Ulcerative colitis	0.600 (0.274)	-1.969 (0.024)	2.550 (0.005)*	0.415 (0.661)
Brain	BMI	4.260 (<0.001)**	0.109 (0.544)	3.088 (0.001)**	-0.954 (0.170)
	Educational attainment	3.454 (<0.001)**	-2.882 (0.002)**	1.961 (0.025)*	-2.710 (0.003)**
	Schizophrenia	5.041 (<0.001)**	-3.772 (<0.001)**	5.000 (<0.001)**	-2.619 (0.004)**
Other	Age at menarche	3.155 (<0.001)**	-1.964 (0.025)	2.916 (0.002)**	-1.176 (0.120)
	Coronary artery disease	2.184 (0.014)*	-1.309 (0.095)	1.509 (0.066)	-0.094 (0.463)
	Height	4.319 (<0.001)**	-0.337 (0.368)	3.782 (<0.001)**	-2.156 (0.016)
	LDL level	2.135 (0.016)*	-0.806 (0.210)	2.003 (0.023)*	0.391 (0.652)
	Smoking behavior	0.969 (0.166)	0.514 (0.696)	0.880 (0.189)	0.876 (0.810)

Z-scores associated with specific annotations in SLDSR models that distinguish enrichment in GWAS signal of unique from all eQTLs in a tissue. Values between brackets show the corresponding *p*-values. One asterisk indicates coefficients passing FDR correction for multiple testing. Two asterisks indicate coefficients passing Bonferroni correction for multiple testing. Negative z-score means depletion (i.e. unique eQTLs have less effect than eQTLs present in brain and blood) positive z-score means enrichment. Unique brain eQTLs were not significantly depleted compared to brain eQTLs in strong LD with blood eQTLs in their effect on all three immune-traits (Table 1 left panel). Unique brain eQTLs were significantly depleted in their effect on schizophrenia and educational attainment compared to the complete brain eQTL annotation, but not for BMI.