# Exploring the impact of complex diseases on non-diseased human tissue transcriptomes

**Chi-Lam Poon**
Weill Cornell Graduate School of Medical Sciences https://orcid.org/0000-0001-6298-7099
**Cho-Yi Chen** ( choyichen@nycu.edu.tw )
Weill Cornell Graduate School of Medical Sciences https://orcid.org/0000-0002-7915-7814

**Primary research**

**Keywords:** Complex disease, Cerebrovascular disease, Major depression, Inflammation, Transcriptome, GTEx

**Posted Date:** March 18th, 2021

**DOI:** https://doi.org/10.21203/rs.3.rs-317782/v1

## Abstract

# Background

The development of complex diseases is contributed by the combination of multiple factors and complicated interactions between them. Inflammation has recently been associated with many complex diseases and may cause long-term damage to the human body. In this study, we examined whether two types of complex disease systematically altered the transcriptomes of non-diseased human tissues and whether inflammation was linked to identifiable molecular signatures, using post-mortem samples from the Genotype-Tissue Expression (GTEx) project.

# Results

Following a series of differential expression (DE) analyses, dozens to hundreds of DE genes were identified in multiple tissues between subjects with and without a history of cerebrovascular disease (CVD) or major depression (MD). DE genes from these disease-associated tissues—the visceral adipose, tibial artery, caudate, and spinal cord for CVD; and the hypothalamus, putamen, and spinal cord for MD—were further analyzed for functional enrichment. Many pathways associated with immunological events were positively enriched in the DEGs of the CVD-associated tissues, as were the neurological and metabolic pathways in the MD-associated tissues. Eight gene–tissue pairs were found to overlap with those prioritized by our transcriptome-wide association studies (TWAS), indicating a potential genetic effect on gene expression for circulating cytokine phenotypes.

# Conclusions

Complex diseases like CVD and MD may cause observable changes in the gene expression of non-diseased tissues, suggesting that a long-term impact of diseases, lifestyles and environmental factors may together contribute to the appearance of transcriptomic “scars” on the human body. Furthermore, inflammation is probably one of the systemic and long-lasting effects of cerebrovascular events.

## Introduction

Complex diseases are caused by genetic, environmental and lifestyle factors and their interactions, most of which have not yet been identified. Recent studies have revealed that the immune system and inflammatory responses are involved in a wide range of complex diseases, such as cardiovascular disease [1], stroke [2], cancer [3], and psychiatric disorders [4]. Inflammation is generally defined as the immune system’s response that defends against injury or stress [5]. In a normal inflammatory response, the upregulation of inflammatory activity is strictly regulated. However, with psychological, environmental and biological factors [6–9], the regulated process can become uncontrolled in the resolution phase,

causing a systemic chronic inflammation that contributes to damage in all tissues and organs and increases the risk of diseases that remain global leading causes of disability and mortality [10, 11].

The Genotype-Tissue Expression (GTEx) project [12] has established a database of expression data, whole genome sequences, and whole-exome sequences of 54 non-diseased tissue sites across nearly 1,000 individuals (as of the current v8 release). GTEx has also collected subject phenotype data, including demographic information, general medical histories, histories at the time of death, the circumstances of death, and so on. The medical histories were provided by the hospital systems, which recorded the prior care of deceased donors.

Here we evaluate whether past chronic inflammatory diseases can leave biological alterations ('scars') in non-diseased tissues of the human body by comparing the expression profiles of subjects with and without a history of chronic inflammatory disease. We have focused on two types of complex disease—cerebrovascular disease (CVD) and major depression (MD)—since they are typical chronic inflammatory diseases with heritable components [13–16], and there are enough cases of these in the GTEx project. CVD comprises clinical conditions that impair blood flow to the brain, such as strokes, transient ischemic attacks, intracranial aneurysms, and other vessel diseases [17]. MD is one of the most common psychiatric illnesses, ranging from 3 to 16.9 percent worldwide [18, 19], and has a significant impact on society. It is characterized by a persistent feeling of sadness or a loss of interest or pleasure in outside stimuli. Previous large-scale genome-wide association studies (GWAS) and meta-analyses have identified a large number of genetic loci associated with stroke [20, 21] and depression [16, 22, 23] in multi-ancestry groups. However, genetic variability contributing to the susceptibility mechanism underlying CVD and depression and their interactions with inflammation remains not fully identified or characterized.

In this study, we aimed to answer the following two questions: 1) is there any significant transcriptomic difference in non-diseased tissues between subjects with and without CVD or MD? 2) if yes, is there any evidence to indicate that inflammation may play a role in shaping the transcriptomic landscape? We performed a differential expression (DE) analysis on each GTEx tissue by comparing the expression profiles between subjects with and without a history of CVD or MD. Significant differentially expressed genes (DEGs) identified in multiple tissues from the series of DE analyses were included in the downstream functional enrichment analysis. We also performed transcriptome-wide association studies (TWAS) on inflammation biomarkers to find any overlaps with the DEGs.

# Results

Cohorts and risk factors

Multi-tissue RNA-seq data were compiled from the GTEx project, as described in the Materials and Methods section. Subjects with an explicitly reported medical history of cerebrovascular disease or major depression were considered in this study. A total of 16,412 samples across 46 tissues, obtained from 928 subjects, were included in the CVD analysis (Fig. S1), and 16,221 samples across 45 tissues from 926 subjects in the MD analysis (Fig. S2).

Risk factors of complex diseases include clinical variables such as age [24], sex [25], and BMI [26]. The average age of the cohort with a history of CVD was significantly higher than that of the non-CVD cohort (CVD 52.09 ± 13.09, non-CVD 57.74 ± 10.79 yrs, $P = 3.87 \times 10^{-6}$, *t*-test), while BMI, sex and race showed no significant differences between the two groups (Fig. S3). The average age of the MD cohort was younger than that of the non-MD cohort (MD 49.69 ± 13.33, non-MD 53.07 ± 12.91 yrs, $P = 1.82 \times 10^{-2}$, *t*-test). Moreover, females had a higher incidence of developing depression ($P = 0.01$, Chi-squared test) (Fig. S4). This is consistent with the higher prevalence of major depressive disorder in women than in men [27].

Differentially expressed genes identified in multiple tissues

To investigate whether past CVD or MD left “scars” on any tissues or organs, we implemented the voom-limma [28, 29] pipeline to identify genes differentially expressed between the cohorts with and without a history of CVD or MD, using the linear model described in the Materials and Methods section. Seventeen out of 46 and 16 out of 45 tissues displayed significant differential expression (false discovery rate < 0.05 and absolute fold change > 1.5) in the CVD and MD analyses, respectively (Fig. 1a, b). The top four tissues with the highest number of significant DEGs were included in the functional enrichment analysis (Adipose - Visceral, Artery - Tibial, Brain - Caudate, and Brain - Spinal cord for CVD), as well as the top three MD tissues (Brain - Hypothalamus, Brain - Putamen, and Brain - Spinal cord).

There was no common DEG shared by the four CVD or the three MD tissues (Fig. 2a, b). A large number of significant DEGs identified by our CVD model (Table S1) were associated with inflammation. For instance, the most significant upregulated gene in the spinal cord, *CHI3L1*, is related to a variety of inflammatory disorders [30–32] and coronary artery disease [33]. Gene *FCGR3A*—upregulated in three of the CVD tissues (Brain - Spinal cord adj.P.val = 0.02; Brain - Caudate adj.P.val = 0.03; Artery - Tibial adj.P.val = 0.02)—encodes a receptor that binds the Fc portion of IgG, and it affects the pharmacokinetics in patients with Crohn’s Disease [34]. *LPAR5*, which was overexpressed in both the brain caudate and the spinal cord of the subjects with a history of CVD, has been reported to be activated during nerve injury [35], and it transmits pro-inflammatory signals [36]. *DLG2*, which was downregulated in hypothalamus tissues with MD (Table S2), has been reported to be associated with interferon production [37]. These results suggest that the systematic effects left by CVD and MD, on which inflammation may play a role, can still be identified in several post-mortem human tissues on the transcriptomic level.

Inflammatory events enriched in differentially expressed genes

A functional overview can be gained through gene set enrichment analysis. CVD DEGs from the Adipose - Visceral, Artery - Tibial, Brain - Caudate, and Brain - Spinal cord; and MD DEGs from the Brain - Hypothalamus, Brain - Putamen, and Brain - Spinal cord were further analyzed using the Gene Set Enrichment Analysis method [38]. A broad spectrum of Gene Ontology (GO) terms, with the top significantly enriched GO terms in the CVD spinal cord, are presented as an example in Fig. 3.

Strikingly, upregulated genes in all four CVD tissues were significantly enriched in immunological events, including antigen binding, T cell proliferation, the interferon-gamma-mediated signaling pathway, and so on (Fig. 3, Table 1), although the visceral adipose only had one significant upregulated gene. This is consistent with a large body of evidence that shows that inflammation plays a crucial role in cerebrovascular diseases. Inflammation can rupture the intracranial aneurysm wall [39], lead to secondary injury after an ischemic stroke [40], and impact the progression of symptomatic intracranial atherosclerosis [41]. Inflammation has also been linked to blood–brain barrier dysfunction [42] and tissue injury [43] in cerebrovascular diseases.

For the MD tissues, inflammatory events were only mainly enriched in downregulated genes in the spinal cord (Table 1) and a few in upregulated genes in the hypothalamus (Additional File 3). Likewise, depression has been associated with increased inflammatory activation in both the periphery and the central nervous system. Many antidepressant agents reduce inflammatory activation in immune cells and lower circulating inflammatory cytokine levels, supporting this association [44]. Furthermore, it is worth mentioning that mitochondrial events and cellular respiration were significantly upregulated in the putamen. Since mitochondrial energy metabolism in the putamen has been reported to be highly correlated with emotional and intellectual impairment in Schizophrenics [45], it might also have some hidden links with depression as well. Another interesting result is that hypothalamus upregulated genes were mapped to terms related to cilia (Additional File 3). There is still no obvious evidence connecting cilia with depression so far, but it is an underexplored area worth investigating [46].

We further explored our differential expression results using Disease Ontology (DO) [47] and Human Phenotype Ontology (HPO) [48] annotations. Top enriched diseases and human phenotypes are similar to the biological phenotypes found by our GO analysis (Additional File 3). Immune responses and cerebrovascular lesions were significantly enriched in the upregulated genes of the CVD tissues. For example, six DO terms—human immunodeficiency virus infectious diseases, temporal arteritis, alopecia areata, autoimmune thrombocytopenic purpura, intracranial aneurysms, and primary immunodeficiency diseases—were shared by all four CVD tissues' overexpressed genes (Fig. S5), and they are all diseases associated with inflammation and cerebrovascular accidents.

Our enrichment results reinforce the strong evidence linking inflammation to CVD, as well as other interesting biological phenomena that probably have associations with CVD and MD in analyzed tissues.

Table 1
Top three Gene Ontology (GO) terms with the highest normalized enrichment score in tissues with many immunological terms

| GO term | NES | FDR q-val | FWER p-val |
|---|---|---|---|
| **Adipose - Visceral (CVD)** | | | |
| Extracellular matrix structural constituent conferring tensile strength | 2.351 | < 0.01 | < 0.01 |
| fc receptor mediated stimulatory signaling pathway | 2.336 | < 0.01 | < 0.01 |
| Antigen binding | 2.258 | 0.005 | 0.012 |
| **Artery - Tibial (CVD)** | | | |
| Immunoglobulin complex | 3.722 | < 0.01 | < 0.01 |
| Antigen binding | 3.520 | < 0.01 | < 0.01 |
| Co-translational protein targeting to membrane | 3.386 | < 0.01 | < 0.01 |
| **Brain - Caudate (CVD)** | | | |
| Positive regulation of T cell proliferation | 2.980 | < 0.01 | < 0.01 |
| Positive regulation of leukocyte cell cell adhesion | 2.934 | < 0.01 | < 0.01 |
| T cell proliferation | 2.929 | < 0.01 | < 0.01 |
| **Brain - Spinal cord (CVD)** | | | |
| Response to interferon gamma | 2.981 | < 0.01 | < 0.01 |
| Leukocyte proliferation | 2.936 | < 0.01 | < 0.01 |
| Leukocyte cell cell adhesion | 2.926 | < 0.01 | < 0.01 |
| **Brain - Spinal cord (MD)** | | | |
| Antigen binding | -2.670 | < 0.01 | < 0.01 |
| T cell receptor complex | -2.610 | < 0.01 | < 0.01 |
| Immunoglobulin complex | -2.451 | < 0.01 | < 0.01 |

Transcriptome-Wide Association Analysis

We then enquired whether the transcriptomic ‘scars’ were associated with the subject’s genotype. Transcriptome-wide association analysis (TWAS) is a powerful approach to prioritize target genes by combining genetic variants identified in GWAS with transcriptome data, and can help shed light on possible associations between genetic loci and human complex diseases. Here, we carried out TWAS analysis with S-PrediXcan [49] on four public GWAS summary statistics datasets, which evaluated

human circulating levels of C-reactive protein (CRP) [50], monocyte chemotactic protein-1 (MCP1), interleukin-6 (IL-6), and interferon gamma [51] (Table 3). CRP is known as a systemic biomarker of inflammation and has been shown to be a CVD risk biomarker [52] and to increase in patients with MD [53]. Cytokines MCP1, IL-6 and interferon gamma have been reported to have a great probability of contributing to both CVD [54–56] and MD [57–59]. Our aim was to find whether any overlaps between gene–tissue pairs identified from TWAS and from our DE analyses amongst the tissues with expression variance (including Adipose - Visceral, Artery - Tibial, Brain - Caudate, Brain - Hypothalamus, Brain - Putamen, and Brain - Spinal cord), plus whole blood, since circulating cytokine levels were measured in serum or plasma samples. There were 1,575 and 25 significant gene–tissue pairs passing the corrected *P*-value threshold (*P*-value/number of genes) found in the CRP and MCP1 data respectively, regardless of tissue types (Table S3). There were eight overlapping gene–tissue pairs (Table 2), amongst which *PPP1R18, RP11-238F2.1, FRK* had the same direction of variations but others had the opposite direction. All seven protein-coding genes were more or less related to immune responses. Specifically, it was demonstrated that complement gene *C2* was expressed in human post-mortem brain-derived cerebrovascular smooth muscle cells and may amplify the pro-inflammatory effects in brain vessels [60]. The major histocompatibility complex class I chain-related gene A (MICA) is a highly polymorphic gene that encodes protein variants functioning in immune activation and surveillance; our results therefore indicate that there may be a link between MICA and depression.

Table 2
Overlap of TWAS associations and DE genes

| Gene | Tissue | TWAS z-score | TWAS P.val | DE *t* statistic | DE adj.P.val | Study |
|---|---|---|---|---|---|---|
| *C2* | Spinal cord | -6.43 | $1.27 \times 10^{-10}$ | 5.36 | $7.93 \times 10^{-4}$ | CVD |
| *PSD4* | Spinal cord | -4.34 | $1.45 \times 10^{-5}$ | 4.25 | $8.15 \times 10^{-3}$ | CVD |
| *PPP1R18* | Spinal cord | 4.78 | $1.75 \times 10^{-6}$ | 4.21 | $8.73 \times 10^{-3}$ | CVD |
| *PTPRJ* | Spinal cord | -5.02 | $5.10 \times 10^{-7}$ | 3.82 | 0.017 | CVD |
| *RP11-238F2.1* | Spinal cord | -5.07 | $4.08 \times 10^{-7}$ | -3.21 | 0.046 | CVD |
| *FRK* | Spinal cord | -4.59 | $4.44 \times 10^{-6}$ | -3.15 | 0.050 | CVD |
| *CD1E* | Whole blood | -8.74 | $2.25 \times 10^{-18}$ | 4.43 | 0.040 | CVD |
| *MICA* | Hypothalamus | -5.16 | $2.50 \times 10^{-7}$ | 3.90 | 0.046 | MD |

# Discussion

Most CVD and MD transcriptome analyses [61–67] are restricted to mouse models or a limited sample size of human expression data. In this study, we systematically analyzed expression data for over 16,000 healthy human samples across multiple tissues from GTEx, investigating global transcriptomic alterations on the human body in cases with a history of CVD or MD. We first built a linear mixed model and applied it to the expression data. Dozens to hundreds of differentially expressed genes were identified in the visceral adipose, tibial artery, caudate, and spinal cord for CVD, and in the hypothalamus, putamen, and spinal cord for MD. Furthermore, functional enrichment analysis showed that a large number of annotations pertaining to inflammatory responses were positively enriched in CVD DEGs from all four tissues, and that MD DEGs were mostly associated with neurological and metabolic events. Our results suggest that the long-term sequelae of cerebrovascular accidents and depressive symptoms can still be reflected in post-mortem samples, and that inflammation may be maintained for a long period of time after recovery from CVD.

A growing body of evidence indicates that inflammation not only contributes to the initiation and development of CVD [68, 69], it also persists globally in the brain for the long-term after CVD [70]. Neuroinflammation followed by cerebrovascular accidents may promote recovery and further injury, playing both beneficial and detrimental roles [71]. A large-scale GWAS discovered one genetic variant (rs1842681) in the gene *LOC105372028* associated with post-stroke outcomes [72]. Furthermore, proteomic studies of post-stroke depression (PSD) reveal that immune dysfunction in stroke survivors is associated with PSD [73, 74]. The connection between inflammation and depression is undeniable [53, 75]. However, unlike the CVD results, only a small portion of the MD DEGs were enriched in immune responses. This is understandable, since MD is highly heterogeneous and not all individuals exposed to inflammatory challenges develop depression. Still, inflammatory responses that occur before and after cerebrovascular accidents or depression are very complicated, and the underlying mechanism is yet to be elucidated.

Our study also provides evidence of the general and long-term effects left by cerebrovascular events and depression from the transcriptomic aspect. Non-brain tissues with significant CVD DEGs were related to vascular diseases and may pose risks to CVD. To be more specific, adipose tissue and its secreted inflammatory proteins contributed to obesity-associated vasculopathy and cardiovascular risk [76], and they may contribute to CVD as well. Peripheral arterial disease occurring in the tibial artery shared similar risk factors with CVD [77]. Moreover, DEGs such as *CHI3L1* and *LPAR5* may reveal possible mechanisms for post-CVD outcomes, but further experiments are necessary for validation. Interestingly, the hypothalamus had the highest number of MD DEGs, which is compatible with one of the most enduring and replicated findings in psychiatry — the activation of the hypothalamic-pituitary-adrenal axis in a subset of MD patients. The identified DEGs may play a role in the neuroendocrine function of the hypothalamus. The putamen, which also had many MD DEGs, indicates that ageing is accelerated in patients with major depressive disorders [78]. The pathways enriched in putamen positive DEGs were

mainly about mitochondrial functions and the electric transport chain, which replicates previous results and provides new insights into the long-term effects of depression.

Only a few DEGs identified by our linear model overlapped with genes prioritized by TWAS for selected cytokine phenotypes. This was expected and is probably due to the small fraction of genetic risk factors shared by complex diseases and these circulating cytokine levels. Additionally, only about 11 percent of heritability was explained by bulk tissue expression quantitative trait loci, according to this study [79]. Therefore, long-term transcriptomic alterations across tissues and organs are probably caused by external factors such as lifestyle and the social and physical environment. Nevertheless, we used bulk RNA-seq data for our analyses, and further utilizing techniques with higher resolution—like single-cell sequencing [80] or decomposition—could reveal more precise signals on specific cell types.

# Conclusions

This study reveals molecular signatures of chronic effects and damage on multiple tissues potentially contributed by complex diseases and associated factors. These signatures may be linked to inflammation and other disease-related pathways. Together, these results indicate that suffering from a complex disease can cause a tissue-wide impact on the transcriptomes, and they also suggest that treatment to attenuate inflammation may improve the body's health in patients recovering from CVD. Our study not only provides insights into these disease mechanisms but also offers a possible route to studying the long-lasting changes caused by chronic diseases on multiple tissues or organs.

# Materials And Methods

GTEx data

Multi-tissue RNA-seq data were collected from the GTEx project [12] v8 release (dbGaP: phs000424.v8.p2). The genes and samples were filtered and quantile-normalized in a tissue-aware manner, as described in the YARN pipeline paper [81].

Subjects with an explicitly reported medical history of cerebrovascular disease (MHCVD, phv00169142.v8.p2) or major depression (MHDPRSSN, phv00169145.v8.p2) were considered in this study. We removed subjects with missing values in their Hardy scale (DTHHRDY), ischemic time (SMTSISCH), or batch ID (SMNABTCH). All cell lines and tissues with less than 12 samples with a history of CVD or less than 10 samples with a history of MD were excluded from our analyses. Finally, we used a total of 16,412 human post-mortem samples (1,498 with and 14,914 without a history of CVD), covering 46 tissues from 928 subjects (99 with and 829 without a history of CVD) in the CVD analysis (Fig. S1); and a total of 16,221 samples (1,602 with and 14,619 without a history of depression) across 45 tissues, including at least 10 samples with a history of MD, from 926 subjects (91 with and 835 without a history of MD) (Fig. S2).

Differential Expression Analysis

Differential expression analysis between the samples with and without a history of CVD/MD was conducted using the voom-limma pipeline [28, 29]. Briefly, RNA-seq read counts were transformed to log counts per million (log-cpm) with associated precision weights to stabilize the variance in the data using the *voom* function, followed by linear model fitting and the empirical Bayes procedure. According to this paper [82], the multivariate linear regression model that adjusted for known confounders outperforms other methods correcting for hidden confounders, which may remove some of the desired biological signals. Hence, we adopted the linear regression model but replaced the experimental batch (SMGEBTCH) with another batch information (SMNABTCH). This model fits for gender (GENDER), the interval between the time of the donor's death and the sample collection (SMTSISCH), age (AGE), the type of nucleic acid isolation batch (SMNABTCH), and the type of death (DTHHRDY) for the gene expression data (Y):

$$Y \sim \beta_1\, GENDER + \beta_2\, SMTSISCH + \beta_3\, AGE + \beta_4\, SMNABTCH + \beta_5\, DTHHRDY + \beta_6\, MHCVD + \varepsilon$$

$$Y \sim \beta_1\, GENDER + \beta_2\, SMTSISCH + \beta_3\, AGE + \beta_4\, SMNABTCH + \beta_5\, DTHHRDY + \beta_6\, MHDPRSSN + \varepsilon$$

The *GENDER* term was removed from sex-specific tissues, and the *SMNABTCH* term was removed from tissues in only one batch. *P*-values from the regression model were adjusted for multiple testing using the Benjamini-Hochberg method [83].

Functional Enrichment Analysis

Pre-ranked Gene Set Enrichment Analysis (GSEA) [38] was conducted with gene lists ranked by the *t*-statistics from the results of our DE analyses, with default program parameters and a default background set on GSEA v4.0.1. The Gene Matrix Transposed (GMT) files of Gene Ontology were obtained from the Molecular Signatures Database v7.1. Disease Ontology [47] data were downloaded from the Alliance of Genome Resources (https://build.alliancegenome.org). Human Phenotype Ontology [48] annotations were acquired from the HPO website (https://hpo.jax.org/app/download/annotation).

Association Detection from GWAS summary statistics

GWAS summary statistics datasets were downloaded from the NHGRI-EBI GWAS Catalog [84] for study GCST009777 [50] and study GCST004421 [51] downloaded on 19/10/2020. These GWAS datasets examined biomarkers of inflammatory responses, and they were obtained from Caucasian subjects (Table. 3). Gene expression variation was inferred using S-PrediXcan [49] with GTEx v8 elastic-net prediction models (http://predictdb.org/) for the four tissues with expression variation between CVD and non-CVD cohorts: Adipose - Visceral, Artery - Tibial, Brain - Caudate, and Brain - Spinal cord; and the three tissues with expression variation between MD and non-MD cohorts: Brain - Hypothalamus, Brain - Putamen, and Brain - Spinal cord. We ran S-PrediXcan on these tissues one by one in each phenotype. Tissue–gene pairs with *P*-value < 0.05/(number of tested genes) were considered as significant.

**Table 3**. GWAS datasets used in this study

| Phenotype | Data source | Sample size |
|---|---|---|
| C-reactive protein [50] | UK BioBank | 418,642 |
| MCP1 [51] | The Cardiovascular Risk in Young Finns Study, FINRISK | 8,293 |
| IL-6 [51] | The Cardiovascular Risk in Young Finns Study, FINRISK | 8,293 |
| Interferon gamma [51] | The Cardiovascular Risk in Young Finns Study, FINRISK | 8,293 |

# Abbreviations

GTEx
The Genotype-Tissue Expression project
DE
differential expression
DEG
differentially expressed gene
CVD
cerebrovascular disease
MD
major depression
MHCVD
medical history of cerebrovascular disease
MHDPRSSN
medical history of major depression
TWAS
Transcriptome-Wide Association Analysis
GWAS
genome-wide association studies
GSEA
Gene Set Enrichment Analysis
GO
Gene Ontology
DO
Disease Ontology
HPO
Human Phenotype Ontology
CRP
C-reactive protein
MCP1
monocyte chemotactic protein-1

IL-6
interleukin-6
Log-cpm
log counts per million
DTHHRDY
Hardy scale variable name
SMTSISCH
ischemic time variable name
SMNABTCH
nucleic acid isolation batch id variable name
SMGEBTCH
experimental batch id variable name
GMT
Gene Matrix Transposed

## Declarations

# Ethics approval and consent to participate

This work was conducted under dbGaP-approved protocol 22839 (dbGaP: phs000424).

# Consent for publication

Not applicable.

# Availability of data and materials

This project is under the approval of access request #84958 for the dataset General Research Use in Genotype-Tissue Expression (dbGaP: phs000424). The GTEx data were downloaded from dbGaP.

# Competing interests

The authors declare that they have no competing interests.

# Funding

This work was financially supported by the Young Scholar Fellowship Program by the Ministry of Science and Technology (MOST) in Taiwan, under Grant MOST 110-2636-E-010-001, MOST109-2636-E-010-003, and MOST110-2811-B-010-507. This work was also supported by the Sustainable Translational Research

Orchestrated Knowledge Ecosystem (STROKE) and by the Brain Research Center, National Yang Ming Chiao Tung University, from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan.

# Authors' contributions

C.-L.P. and C.-Y.C. designed the study. C.-L.P. analyzed the data and drafted the manuscript. C.-L.P. and C.-Y.C. wrote and reviewed the manuscript. C.-Y.C. acquired the funding.

# Acknowledgments

We thank Yi-Sian Lin, Pin-Jou Wu, Yao-Chung Chen, and Hao-Jen Deng for their valuable discussions. We thank the National Center for High-performance Computing (NCHC) for providing computational and storage resources.

## References

1. Willerson JT, Ridker PM. Inflammation as a cardiovascular risk factor. Circulation. 2004.
2. Anrather J, Iadecola C. Inflammation and Stroke: An Overview. Neurotherapeutics. 2016.
3. Coussens LM, Werb Z. Inflammation and cancer. Nature. 2002.
4. Yuan N, Chen Y, Xia Y, Dai J, Liu C. Inflammation-related biomarkers in major psychiatric disorders: a cross-disorder assessment of reproducibility and specificity in 43 meta-analyses. Transl. Psychiatry. 2019.
5. Netea MG, Balkwill F, Chonchol M, Cominelli F, Donath MY, Giamarellos-Bourboulis EJ, et al. A guiding map for inflammation. Nat. Immunol. 2017.
6. Schneiderman N, Ironson G, Siegel SD. Stress and health: Psychological, behavioral, and biological determinants. Annu. Rev. Clin. Psychol. 2005.
7. Sears ME, Genuis SJ. Environmental determinants of chronic disease and medical approaches: Recognition, avoidance, supportive therapy, and detoxification. J. Environ. Public Health. 2012.
8. Zhu Y, Armstrong JL, Tchkonia T, Kirkland JL. Cellular senescence and the senescent secretory phenotype in age-related chronic diseases. Curr. Opin. Clin. Nutr. Metab. Care. 2014.
9. Furman D, Campisi J, Verdin E, Carrera-Bastos P, Targ S, Franceschi C, et al. Chronic inflammation in the etiology of disease across the life span. Nat Med. 2019;
10. Virani SS, Alonso A, Benjamin EJ, Bittencourt MS, Callaway CW, Carson AP, et al. Heart disease and stroke statistics—2020 update: A report from the American Heart Association. Circulation. 2020.
11. Murray CJL, Lopez AD. Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. Lancet. 1997;

12. Aguet F, Barbeira AN, Bonazzola R, Brown A, Castel SE, Jo B, et al. The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science (80- ). 2020;
13. Dichgans M, Pulit SL, Rosand J. Stroke genetics: discovery, biology, and clinical applications. Lancet Neurol. 2019.
14. Boehme AK, Esenwa C, Elkind MSV. Stroke Risk Factors, Genetics, and Prevention. Circ. Res. 2017.
15. Flint J, Kendler KS. The Genetics of Major Depression. Neuron. 2014.
16. Ormel J, Hartman CA, Snieder H. The genetics of depression: successful genome-wide association studies introduce new challenges. Transl Psychiatry. 2019;
17. Goldstein LB, Lichtman JH. Epidemiology of Cerebrovascular Disease. Vasc Med A Companion to Braunwald's Hear Dis Second Ed. 2013.
18. Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, et al. The Epidemiology of Major Depressive Disorder: Results from the National Comorbidity Survey Replication (NCS-R). J. Am. Med. Assoc. 2003.
19. Demyttenaere K, Bruffaerts R, Posada-Villa J, Gasquet I, Kovess V, Lepine JP, et al. Prevalence, severity, and unmet need for treatment of mental disorders in the World Health Organization World Mental Health Surveys. J. Am. Med. Assoc. 2004.
20. Keene KL, Hyacinth HI, Bis JC, Kittner SJ, Mitchell BD, Cheng YC, et al. Genome-Wide Association Study Meta-Analysis of Stroke in 22 000 Individuals of African Descent Identifies Novel Associations with Stroke. Stroke. 2020;
21. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. Nat Genet. 2018;
22. Howard DM, Adams MJ, Clarke TK, Hafferty JD, Gibson J, Shirali M, et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. Nat Neurosci. 2019;
23. Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A, et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. Nat Genet. 2018;
24. Niccoli T, Partridge L. Ageing as a risk factor for disease. Curr. Biol. 2012.
25. Ober C, Loisel DA, Gilad Y. Sex-specific genetic architecture of human disease. Nat. Rev. Genet. 2008.
26. Knight JA. Diseases and disorders associated with excess body weight. Ann Clin Lab Sci. 2011;
27. Kessler RC. Epidemiology of women and depression. J Affect Disord. 2003.
28. Law CW, Chen Y, Shi W, Smyth GK. Voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol. 2014;
29. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;

30. Johansen JS, Bojesen SE, Tybjærg-Hansen A, Mylin AK, Price PA, Nordestgaard BG. Plasma YKL-40 and total and disease-specific mortality in the general population. Clin Chem. 2010;
31. Kastrup J, Johansen JS, Winkel P, Hansen JF, Hildebrandt P, Jensen GB, et al. High serum YKL-40 concentration is associated with cardiovascular and all-cause mortality in patients with stable coronary artery disease. Eur Heart J. 2009;
32. Im JH, Yeo IJ, Park PH, Choi DY, Han SB, Yun J, et al. Deletion of Chitinase-3-like 1 accelerates stroke development through enhancement of Neuroinflammation by STAT6-dependent M2 microglial inactivation in Chitinase-3-like 1 knockout mice. Exp Neurol. 2020;
33. Rathcke CN, Vestergaard H. YKL-40 - an emerging biomarker in cardiovascular disease and diabetes. Cardiovasc. Diabetol. 2009.
34. Termant D, Berkane Z, Picon L, Gouilleux-Gruart V, Colombel JF, Allez M, et al. Assessment of the Influence of Inflammation and FCGR3A Genotype on Infliximab Pharmacokinetics and Time to Relapse in Patients with Crohn's Disease. Clin Pharmacokinet. 2015;
35. Santos-Nogueira E, López-Serrano C, Hernández J, Lago N, Astudillo AM, Balsinde J, et al. Activation of lysophosphatidic acid receptor type 1 contributes to pathophysiology of spinal cord injury. J Neurosci. 2015;
36. Plastira I, Bernhart E, Goeritzer M, Reicher H, Kumble VB, Kogelnik N, et al. 1-Oleyl-lysophosphatidic acid (LPA) promotes polarization of BV-2 and primary murine microglia towards an M1-like phenotype. J Neuroinflammation. 2016;
37. Ali S, Hoven A, Dress RJ, Schaal H, Alferink J, Scheu S. Identification of a novel Dlg2 isoform differentially expressed in IFNβ-producing plasmacytoid dendritic cells. BMC Genomics. 2018;
38. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;
39. Tulamo R, Frösen J, Hernesniemi J, Niemelä M. Inflammatory changes in the aneurysm wall: A review. J. Neurointerv. Surg. 2010.
40. Ahmad M, Graham SH. Inflammation After Stroke: Mechanisms and Therapeutic Approaches. Transl. Stroke Res. 2010.
41. Arenillas JF, Álvarez-Sabín J, Molina CA, Chacón P, Fernández-Cadenas I, Ribó M, et al. Progression of symptomatic intracranial large artery atherosclerosis is associated with a proinflammatory state and impaired fibrinolysis. Stroke. 2008;
42. de Vries HE, Kooij G, Frenkel D, Georgopoulos S, Monsonego A, Janigro D. Inflammatory events at blood-brain barrier in neuroinflammatory and neurodegenerative disorders: implications for clinical disease. Epilepsia. 2012.
43. Jin R, Yang G, Li G. Inflammatory mechanisms in ischemic stroke: role of inflammatory cells. J Leukoc Biol. 2010;
44. Lee CH, Giuliani F. The Role of Inflammation in Depression and Fatigue. Front. Immunol. 2019.

45. Prince JA, Harro J, Blennow K, Gottfries CG, Oreland L. Putamen mitochondrial energy metabolism is highly correlated to emotional and intellectual impairment in schizophrenics. Neuropsychopharmacology. 2000;
46. Pruski M, Lang B. Primary cilia–an underexplored topic in major mental illness. Front. Psychiatry. 2019.
47. Schriml LM, Mitraka E, Munro J, Tauber B, Schor M, Nickle L, et al. Human Disease Ontology 2018 update: Classification, content and workflow expansion. Nucleic Acids Res. 2019;
48. Köhler S, Carmody L, Vasilevsky N, Jacobsen JOB, Danis D, Gourdine JP, et al. Expansion of the Human Phenotype Ontology (HPO) knowledge base and resources. Nucleic Acids Res. 2019;
49. Barbeira AN, Dickinson SP, Bonazzola R, Zheng J, Wheeler HE, Torres JM, et al. Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. Nat Commun. 2018;
50. Han X, Ong JS, An J, Hewitt AW, Gharahkhani P, MacGregor S. Using Mendelian randomization to evaluate the causal relationship between serum C-reactive protein levels and age-related macular degeneration. Eur J Epidemiol. 2020;
51. Ahola-Olli A V., Würtz P, Havulinna AS, Aalto K, Pitkänen N, Lehtimäki T, et al. Genome-wide Association Study Identifies 27 Loci Influencing Concentrations of Circulating Cytokines and Growth Factors. Am J Hum Genet. 2017;
52. Di Napoli M, Elkind MS, Godoy DA, Singh P, Papa F, Popa-Wagner A. Role of C-reactive protein in cerebrovascular disease: A critical review. Expert Rev. Cardiovasc. Ther. 2011.
53. Miller AH, Raison CL. The role of inflammation in depression: From evolutionary imperative to modern treatment target. Nat. Rev. Immunol. 2016.
54. Georgakis MK, Malik R, Björkbacka H, Pana TA, Demissie S, Ayers C, et al. Circulating monocyte chemoattractant protein-1 and risk of stroke meta-analysis of population-based studies involving 17180 individuals. Circ Res. 2019;
55. Moreno VP, Subirá D, Meseguer E, Llamas P. IL-6 as a biomarker of ischemic cerebrovascular disease. Biomark. Med. 2008.
56. Seifert HA, Collier LA, Chapman CB, Benkovic SA, Willing AE, Pennypacker KR. Pro-Inflammatory Interferon Gamma Signaling is Directly Associated with Stroke Induced Neurodegeneration. J Neuroimmune Pharmacol. 2014;
57. Eyre HA, Air T, Pradhan A, Johnston J, Lavretsky H, Stuart MJ, et al. A meta-analysis of chemokines in major depression. Prog Neuro-Psychopharmacology Biol Psychiatry. 2016;
58. Hodes GE, Ménard C, Russo SJ. Integrating Interleukin-6 into depression diagnosis and treatment. Neurobiol. Stress. 2016.
59. Franscina Pinto E, Andrade C. Interferon-Related Depression: A Primer on Mechanisms, Treatment, and Prevention of a Common Clinical Problem. Curr Neuropharmacol. 2016;

60. Walker DG, Dalsing-Hernandez JE, Lue LF. Human postmortem brain-derived cerebrovascular smooth muscle cells express all genes of the classical complement pathway: A potential mechanism for vascular damage in cerebral amyloid angiopathy and Alzheimer's disease. Microvasc Res. 2008;
61. Cai Y, Zhang Y, Ke X, Guo Y, Yao C, Tang N, et al. Transcriptome Sequencing Unravels Potential Biomarkers at Different Stages of Cerebral Ischemic Stroke. Front Genet. 2019;
62. Androvic P, Kirdajova D, Tureckova J, Zucha D, Rohlova E, Abaffy P, et al. Decoding the Transcriptional Response to Ischemic Stroke in Young and Aged Mouse Brain. Cell Rep. 2020;
63. Kim J, Kang SW, Mallilankaraman K, Baik SH, Lim JC, Balaganapathy P, et al. Transcriptome analysis reveals intermittent fasting-induced genetic changes in ischemic stroke. Hum Mol Genet. 2018;
64. Grond-Ginsbach C, Hummel M, Wiest T, Horstmann S, Pfleger K, Hergenhahn M, et al. Gene expression in human peripheral blood mononuclear cells upon acute ischemic stroke. J Neurol. 2008;
65. Wittenberg GM, Greene J, Vértes PE, Drevets WC, Bullmore ET. Major Depressive Disorder Is Associated With Differential Expression of Innate Immune and Neutrophil-Related Gene Networks in Peripheral Blood: A Quantitative Review of Whole-Genome Transcriptional Data From Case-Control Studies. Biol Psychiatry. 2020;
66. Pantazatos SP, Huang YY, Rosoklija GB, Dwork AJ, Arango V, Mann JJ. Whole-transcriptome brain expression and exon-usage profiling in major depression and suicide: Evidence for altered glial, endothelial and ATPase activity. Mol Psychiatry. 2017;
67. Zhu Y, Strachan E, Fowler E, Bacus T, Roy-Byrne P, Zhao J. Genome-wide profiling of DNA methylome and transcriptome in peripheral blood monocytes for major depression: A Monozygotic Discordant Twin Study. Transl Psychiatry. 2019;
68. Lakhan SE, Kirchgessner A, Hofer M. Inflammatory mechanisms in ischemic stroke: Therapeutic approaches. J. Transl. Med. 2009.
69. Gu Y, Gutierrez J, Meier IB, Guzman VA, Manly JJ, Schupf N, et al. Circulating inflammatory biomarkers are related to cerebrovascular disease in older adults. Neurol Neuroimmunol NeuroInflammation. 2019;
70. Shi K, Tian DC, Li ZG, Ducruet AF, Lawton MT, Shi FD. Global brain inflammation in stroke. Lancet Neurol. 2019.
71. Jayaraj RL, Azimullah S, Beiram R, Jalal FY, Rosenberg GA. Neuroinflammation: Friend and foe for ischemic stroke. J. Neuroinflammation. 2019.
72. Söderholm M, Pedersen A, Lorentzen E, Stanne TM, Bevan S, Olsson M, et al. Genome-wide association meta-analysis of functional outcome after ischemic stroke. Neurology. 2019;
73. Zhan Y, Yang YT, You HM, Cao D, Liu CY, Zhou CJ, et al. Plasma-based proteomics reveals lipid metabolic and immunoregulatory dysregulation in post-stroke depression. Eur Psychiatry. 2014;
74. Nguyen VA, Carey LM, Giummarra L, Faou P, Cooke I, Howells DW, et al. A pathway proteomic profile of ischemic stroke survivors reveals innate immune dysfunction in association with mild symptoms of depression - a pilot study. Front Neurol. 2016;

75. Dantzer R, O'Connor JC, Freund GG, Johnson RW, Kelley KW. From inflammation to sickness and depression: When the immune system subjugates the brain. Nat. Rev. Neurosci. 2008.
76. Berg AH, Scherer PE. Adipose tissue, inflammation, and cardiovascular disease. Circ. Res. 2005.
77. Banerjee A, Fowkes FG, Rothwell PM. Associations between peripheral artery disease and ischemic stroke: Implications for primary and secondary prevention. Stroke. 2010.
78. Sacchet MD, Camacho MC, Livermore EE, Thomas EAC, Gotlib IH. Accelerated aging of the putamen in patients with major depressive disorder. J Psychiatry Neurosci. 2017;
79. Yao DW, O'Connor LJ, Price AL, Gusev A. Quantifying genetic effects on disease mediated by assayed gene expression levels. Nat Genet. 2020;
80. Nagy C, Maitra M, Tanti A, Suderman M, Théroux JF, Davoli MA, et al. Single-nucleus transcriptomics of the prefrontal cortex in major depressive disorder implicates oligodendrocyte precursor cells and excitatory neurons. Nat Neurosci. 2020;
81. Paulson JN, Chen CY, Lopes-Ramos CM, Kuijjer ML, Platig J, Sonawane AR, et al. Tissue-aware RNA-Seq processing and normalization for heterogeneous and sparse data. BMC Bioinformatics. 2017;
82. Somekh J, Shen-Orr SS, Kohane IS. Batch correction evaluation framework using a-priori gene-gene associations: Applied to the GTEx dataset. BMC Bioinformatics. 2019;
83. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B. 1995;
84. Buniello A, Macarthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res. 2019;

# Figures

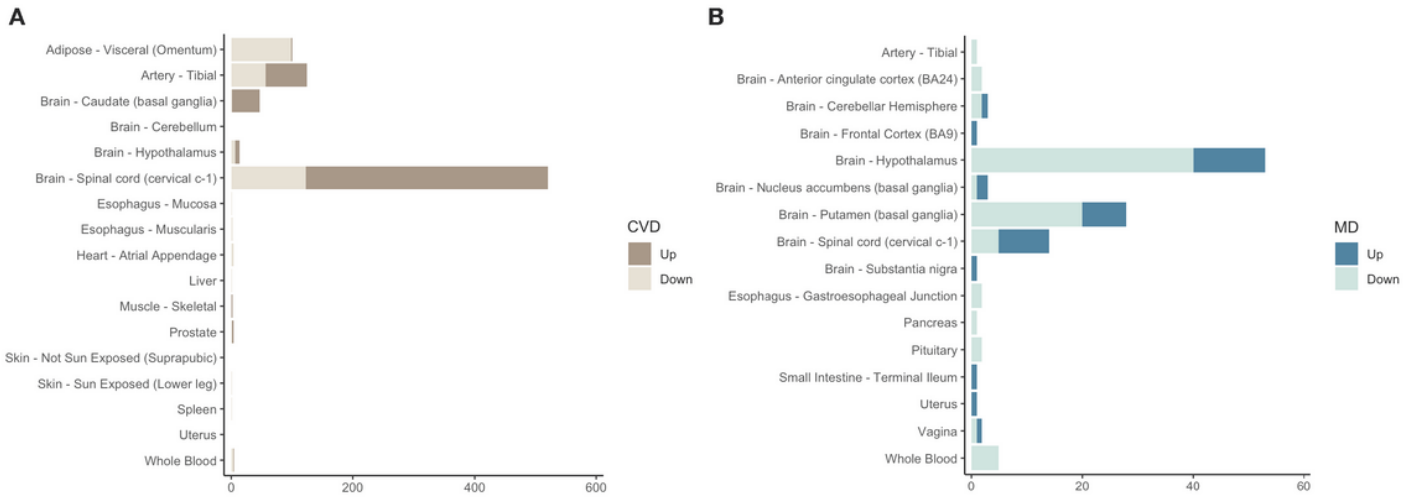
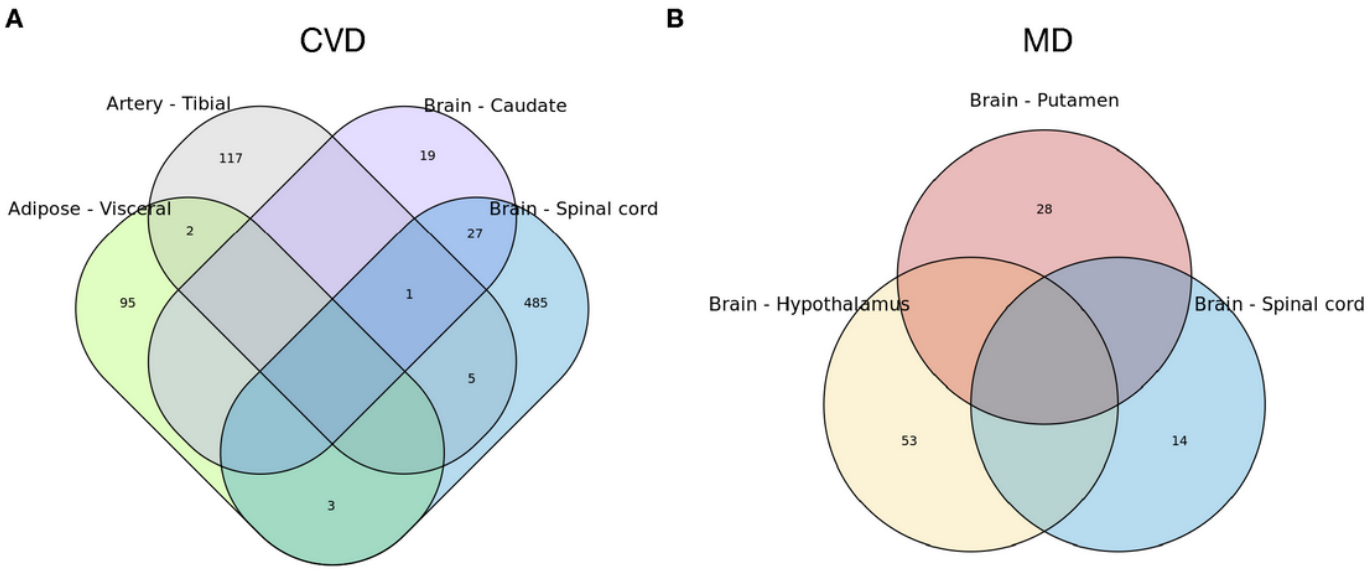


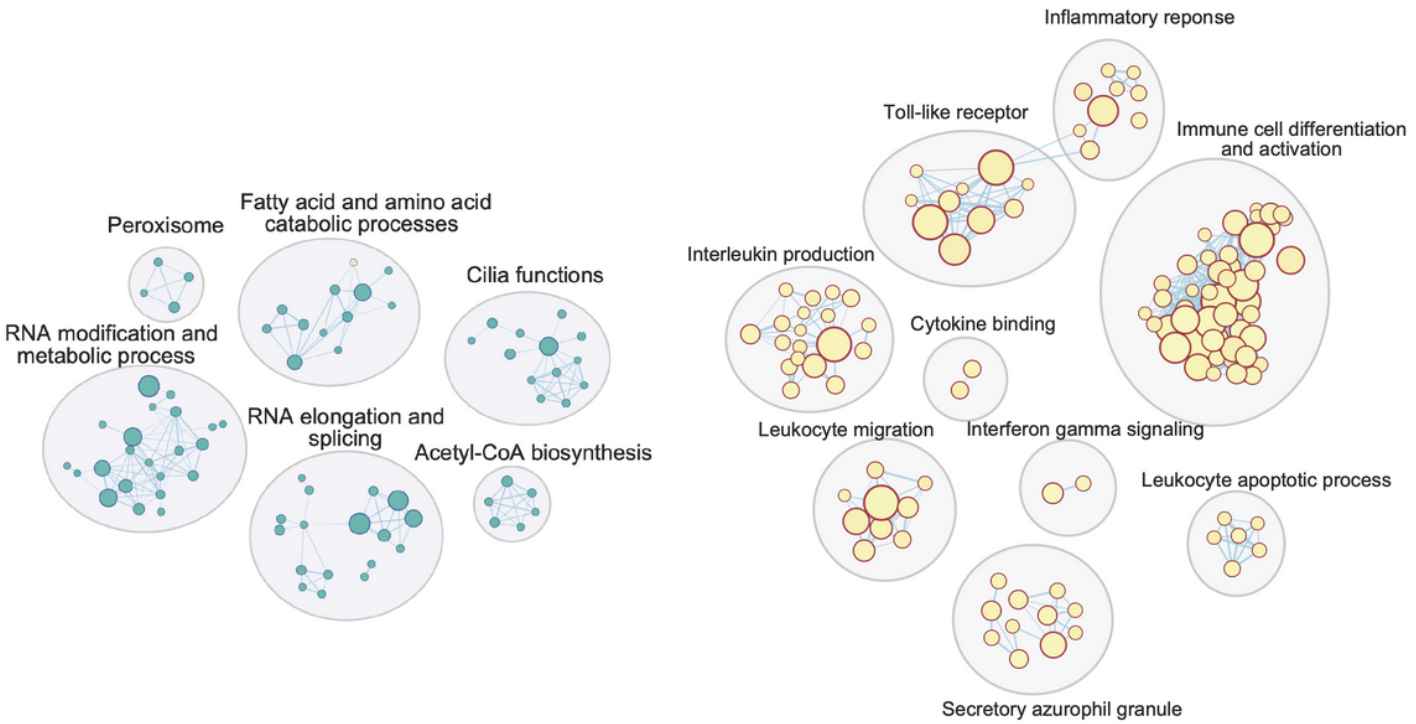

**Figure 1**

The number of significant upregulated and downregulated genes (FDR < 0.05 and absolute log2 fold change > 1.5) identified in GTEx tissues in (A) CVD and (B) MD analyses.




## Figure 2

Overlaps of significant differentially expressed genes. (A) CVD-associated tissues: Adipose - Visceral, Artery - Tibial, Brain - Caudate, and Brain - Spinal cord. (B) MD-associated tissues: Brain - Hypothalamus, Brain - Putamen, and Brain - Spinal cord.




## Figure 3

Cytoscape network image for GO terms significantly enriched in DEGs from the spinal cord. Circles in blue (left) are terms enriched for downregulated genes, those in red (right) are enriched for upregulated genes (cutoff: q-value < 0.1).

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryfigures.pdf
- significantcvddeg.xlsx
- significantdprssndeg.xlsx
- significanttwasresults.xlsx