



Agonal Factors Distort Gene-Expression Patterns in Human Postmortem Brains

Jiacheng Dai^{1,2}, Yu Chen¹, Rujia Dai³, Yi Jiang¹, Jianghua Tian¹, Sihan Liu¹, Meng Xu¹, Miao Li¹, Jiaqi Zhou¹, Chunyu Liu⁴ and Chao Chen^{1,5*}

¹ Center for Medical Genetics, Department of Psychiatry, School of Life Sciences, National Clinical Research Center on Mental Disorders, The Second Xiangya Hospital, Central South University, Changsha, China, ² State Key Laboratory of Genetic Engineering, Human Phenome Institute, and School of Life Sciences, Fudan University, Shanghai, China, ³ Upstate Medical University, Syracuse, NY, United States, ⁴ Department of Psychiatry, Upstate Medical University, Syracuse, NY, United States, ⁵ National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, China

OPEN ACCESS

Edited by:

Ruth Luthi-Carter,
University of Leicester,
United Kingdom

Reviewed by:

Zhisong He,
ETH Zurich, Switzerland
Michael Oldham,
University of California,
San Francisco, United States

*Correspondence:

Chao Chen
chenchao@skimg.edu.cn

Specialty section:

This article was submitted to
Neurogenomics,
a section of the journal
Frontiers in Neuroscience

Received: 09 October 2020

Accepted: 16 February 2021

Published: 25 March 2021

Citation:

Dai J, Chen Y, Dai R, Jiang Y, Tian J, Liu S, Xu M, Li M, Zhou J, Liu C and Chen C (2021) Agonal Factors Distort Gene-Expression Patterns in Human Postmortem Brains. *Front. Neurosci.* 15:614142. doi: 10.3389/fnins.2021.614142

Agonal factors, the conditions that occur just prior to death, can impact the molecular quality of postmortem brains, influencing gene expression results. Our study used gene expression data of 262 samples from ROSMAP with the detailed terminal state recorded for each donor, such as fever, infection, and unconsciousness. Fever and infection were the primary contributors to brain gene expression changes, brain cell-type-specific gene expression, and cell proportion changes. Furthermore, we also found that previous studies of gene expression in postmortem brains were confounded by agonal factors. Therefore, correction for agonal factors is important in the step of data preprocessing. Our analyses revealed fever and infection contributing to gene expression changes in postmortem brains and emphasized the necessity of study designs that document and account for agonal factors.

Keywords: agonal factors, post-mortem brain studies, gene expression, gene co-expression, brain cell types

INTRODUCTION

Postmortem brain samples are widely used for human genetic studies, primarily for the study of neuropsychiatric disorders such as schizophrenia, bipolar disorders and Alzheimer's disease (Jaffe, 2016; Wang et al., 2019a). Genetic mechanisms studied in postmortem brains from individuals with neuropsychiatric disorders such as transcriptome patterns reflect the genetic impacts of a lifetime of severe mental illness as well as countless environmental factors. Correcting for the extraneous environmental factors that influence gene expression in postmortem tissue (McCullumsmith et al., 2014) is critical to accurate data analyses. One such category of environmental factors, agonal factors, describe the conditions that occur before death. While they represent an inevitable environmental component, agonal factors are historically neglected in postmortem brain studies.

Agonal factors refer to the manner of death and the terminal state before death. The manners of death include slow, intermediate, fast from natural causes, and violent fast (Hardy et al., 1985). Terminal states of consequence include coma, inadequate oxygen, fever, infection, and artificial respiration (Harrison et al., 1991; Johnston et al., 1997). Previous research has shown that messenger RNA (mRNA) is vulnerable to a greater or lesser degree to agonal factors

(Burke et al., 1991; Barton et al., 1993). Some studies have provided an agonal factor score to account for specific agonal conditions and terminal states per individual (Tomita et al., 2004; Hagenauer et al., 2018). Agonal factors have also been reported to negatively affect the gene expression profile (Tomita et al., 2004; Hagenauer et al., 2018). Similar influences are induced by low tissue pH (Li et al., 2004). Such factors cause irreversible decomposition before and up to the moment of death. While some studies have incorporated agonal factors, the basic design typically retains the following shortcomings: (1) Studies focused primarily on the correlation of agonal factors to aspects such as the RNA integrity number, tissue pH, and whole gene expression pattern, neglect to identify the gene expression pathways that underlie agonal related regulation; (2) Such studies (Li et al., 2004; Tomita et al., 2004; Atz et al., 2007; Hagenauer et al., 2018) often roughly combine various terminal states together, thereby overlooking the potential interactions or conflicts between them; and, (3) Analysis is often limited to Pearson correlation and differential gene expression when further methods are needed to reveal the expanse of gene co-expression relationships.

We hypothesize that individual agonal factors uniquely alter postmortem brain gene expression and that successful adjustment of agonal-related variants within gene expression data is necessary and achievable. To investigate this matter, our study included 262 samples of postmortem human brain tissue from the Religious Orders Study and Memory and Aging Project (ROSMAP) study. The data included complete agonal information for the following durations for all samples: surgery, fever, infection, unconsciousness, difficulty breathing, and on mechanical ventilation. We performed differential gene expression and identified the gene co-expression network. We performed linear regression analysis to correct for agonal-related surrogate variables and hidden batch effects.

MATERIALS AND METHODS

Samples

Discovery data was bulk brain RNA-seq gene expression matrix which was collected from the ROSMAP project¹, including 263 samples with detailed information about each donor's agonal conditions. Terminal state information included *fever*, *infection*, *surgery*, “unconc” for *unconsciousness*, “brthprb” for *difficulty breathing*, and “ventilat” for *artificial ventilation* within the hour (fever, infection, surgery, unconsciousness and difficulty breathing) or days (artificial ventilation) prior to death. The phenotypes *fever*, *infection* and *difficulty breathing* indicated that any of those experiences occurred within the 3 days prior to death; and the phenotype of *surgery* indicated a major surgery with anesthesia in the 2 weeks prior to death. The sample size for each terminal state is shown in **Supplementary Figure 1**.

We also try to adjust latent variables of agonal factors in microarray data which did not documented manner of death or terminal states. We performed meta-analysis of different source of data (**Supplementary Table 1**), including 5 microarray data of

Schizophrenia (Iwamoto et al., 2005; Narayan et al., 2008; Maycox et al., 2009; Chen et al., 2013; Reinhart et al., 2015), 3 microarray data of Autism Spectrum Disorder (Garbett et al., 2008; Voineagu et al., 2011; Chow et al., 2012), and 2 microarray data of Inflammatory Bowel Disease (IBD) (Noble et al., 2008; Granlund et al., 2013). The source of microarray data of neuropsychiatric disorders is brain tissue, while the source of IBD is bowel tissue.

Data Preprocess

We performed quality control and data normalization. First, we performed Principal Component Analysis (PCA) and hierarchical clustering to filter outliers. We found sample ID 23690880 was potential outliers, so we removed this sample. Next, we filtered genes which were low expressed. The threshold of filtering is that expression value is less than 0.1 in 20% samples. We included 9092 genes in ROSMAP datasets for following analysis. Then, we performed linear regression models corrected for the biological factors of the data. We corrected for factors including sex, racial group, Spanish ethnicity, years of education, age at death, postmortem interval, CERAD score (Consortium to Establish a Registry for Alzheimer's Disease), Braak stage, and the National Institute on Aging (NIA)-Reagan Institute diagnosis criteria for Alzheimer's disease. We calculated the residuals of each gene to the factors; then, we added the residuals to the mean values of genes of original data matrix, as the corrected data matrix; next, we used corrected data matrix for downstream analysis. Finally, we used quantile normalization to make the datasets in the same scale.

Statistical Analysis

We performed differential gene expression using DESeq2 (Love et al., 2014), which transformed the data type into a gene expression matrix of integer values. In the discovery data, we used *fever*, *infection*, *unconsciousness*, *difficulty breathing*, and *artificial ventilation*, respectively, as the design matrix in DESeq2. Resulting *P*-values were corrected using the Benjamin-Hochberg (BH) procedure to control for multiple comparisons. We compared the differential gene expression effect size of the 5 agonal conditions using the pSI package and Pearson's correlation. We performed all statistical analyses using R (v3.6.0).

Cell Deconvolution

We used cell deconvolution to estimate cell proportion of brain cell types in ROSMAP datasets. We used R package MuSiC (Wang et al., 2019b), which utilizes cell-type specific gene expression from single-cell RNA sequencing data to characterize cell type compositions from bulk tissue RNA-seq data. In our analysis, we used single-cell RNA sequencing data from ROSMAP (Mathys et al., 2019) as reference of cell deconvolution.

Module Construction and Preservation Testing

We identified the gene co-expression network using WGCNA (Langfelder and Horvath, 2008). Before the analysis, we log transformed the data matrix to ensure normal distribution. We calculated a correlation matrix for all genes and chose

¹<https://www.synapse.org/#!Synapse:syn3388564>

the soft-threshold power of 12 to construct an approximate scale-free topology network. Networks were constructed using the `blockwiseModules` function. We chose the signed network type. The network dendrogram was created using an average linkage hierarchical clustering of the topological overlap dissimilarity matrix (1-TOM). Modules were defined as branches of the dendrogram using the hybrid dynamic tree-cutting method. Modules were summarized by their first principal component (ME, module eigengene) and modules with eigengene correlations of >0.9 were merged together. Modules were defined using Pearson correlation. Module (eigengene)-disease associations were evaluated using Pearson correlation. Significance values were FDR-corrected to account for multiple comparisons.

Cell-Type Enrichment Analyses

Cell type enrichment was determined using the Zhang dataset, which uses cell type-specific expression datasets from human cortex brain samples from populations of neurons, astrocytes, oligodendrocytes, microglia and endothelial cells (Zhang et al., 2016). After normalizing and averaging replicate expression profiles for each cell type, a specificity index statistic (pSI) was calculated using the pSI package.

Gene Ontology Enrichment

We performed Gene Ontology (GO) enrichment for biological process, molecular function and cellular components using the `clusterProfiler` v3.12.0 package in R (Yu et al., 2012). The enrichment P -values were BH-corrected to control for multiple comparisons.

Unknown Factors Detection

To detect unknown factors within the ROSMAP dataset, we employed SVA (Leek et al., 2012) and PEER (Stegle et al.,

2012) algorithms in R. The SVA package contains functions for identifying and building surrogate variables for gene expression data that could be used in subsequent analyses to adjust for unknown, unmodeled, or latent sources of noise. In SVA, we created a full model matrix, including race as a variable of interest. We detected 16 surrogate variables in total using the SVA package. PEER was used first to unearth patterns of common variation across the whole dataset and to create up to 15 assumed global hidden factors. Next, the correlation between terminal states and each of the 16 surrogate variables or 15 hidden factors was tested in the ROSMAP dataset. After that, factors showing a Pearson's correlation test FDR-adjusted P -value smaller than 0.05 were included in linear regression analysis. The residual values from the regression analysis were used to correct the variables.

RESULTS

Agonal Factors Are Associated With Gene Expression in the Human Brain

Terminal state was recorded in the discovery data (ROSMAP) in our study. Details included *surgery*, *fever*, *infection*, *unconsciousness*, *difficulty breathing*, and *artificial ventilation*, which are recorded 3 days prior to death. In the following analysis, we focused primarily on these terminal states, performing differential gene expression analyses and gene co-expression analyses to uncover the potential effects that terminal states have on the gene regulatory network.

Differential gene expression (DEG) analysis showed that only fever, infection and unconsciousness were significantly associated with various gene expressions. We found 344 differentially expressed genes (DEGs) that were associated with *fever* (adjust $p < 0.05$, **Figure 1A**), 51 DEGs that were associated with *infection* (adjust $p < 0.05$, **Figure 1B**), and 5 DEGs that were

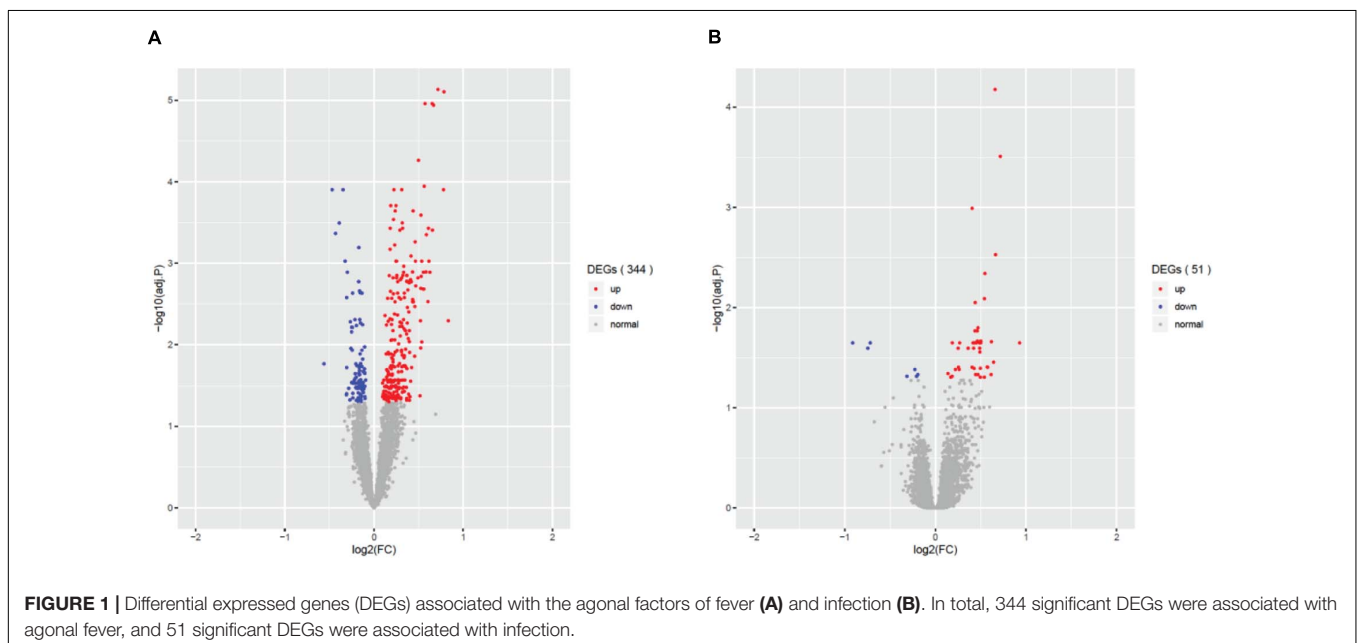


TABLE 1 | P value of cell proportion differential analysis.

	Ast	End	Ex	Mic	Oli	Per
Fever	0.926	0.936	0.36	0.022	0.936	0.06
Infection	0.36	0.936	0.36	0.36	0.98	0.753

associated with *unconsciousness* (adjust $p < 0.05$). However, no significant DEGs were induced by *surgery*, *difficulty breathing* and *artificial ventilation*. The list of significant DEGs is tabulated in **Supplementary Table 2**. These results showed that the agonal-related variants in data were mainly comprised of *fever* and *infection*, rather than *difficulty breathing* or *artificial ventilation*. We also performed PVCA to evaluate the contribution of sample variation of each terminal states (**Supplementary Figure 2**). We found that *fever* and *infection* showed higher contribution comparing to *difficulty breathing* and *unconsciousness*. We also found that terminal state *surgery* and *artificial ventilation* have the highest contribution in PVCA. The reason why terminal state *surgery* and *artificial ventilation* did not show significance contribution to differential gene expression analysis is possibly due to few individuals involving *surgery* and *artificial ventilation*.

The major variance components of bulk brain tissue may be driven by cell proportion, so we performed cell deconvolution to evaluate cell proportion in fever or infection. We used *T*-test to calculate the significance of cell proportion between terminal state and non-terminal state. We found microglia ($P = 0.022$) that showed significant differences in fever and non-fever sample (**Table 1**). The average cell proportion of microglia in fever sample is 7.6%, which is higher than the average cell proportion of microglia in non-fever sample (**Supplementary Table 3**). After we correct the cell proportion with linear regression model, most difference of DEGs are not significant. We only found 11 fever-related DEGs (CRELD1, AHS1, MKNK2, STIP1, CDK2AP2, NQO1, RBM3, PCBD1, CIRBP, AUP1, FKBP4) and 1 infection-related DEGs (SPARC), which is the subset of DEGs before correcting for cell proportion. These gene functioned as mRNA splicing and processing, protein or RNA binding. These results showed that fever may increase the cell proportion of microglia in pre-mortem brain.

DEG gene ontology of *fever* and *infection* showed significant enrichment in synapse- and immune-related pathways (**Figure 2**). In the top gene ontology enrichment of DEGs for *fever* (**Figure 2A**), we identified genes strongly enriched for immune-related pathways. In addition, we found gene enrichment for other pathways, including the response to unfolded protein pathway (adjust $p = 0.00087$), a protective response induced during periods of cellular stress that aims to restore protein homeostasis (Granlund et al., 2013). In the top gene ontology enrichment of DEGs for infection (**Figure 2B**), the strongest enrichment was that of the synapse organization pathway (adjust $p = 0.018831$), followed by the cell killing (adjust $p = 0.018831$), the regulation of synapse organization (adjust $p = 0.01983$) and the regulation of synapse structure or activity (adjust $p = 0.01983$) pathways, indicating that in addition to immune-related pathways, synapse-related gene

expression pathways also play an important role in the brain during *infection*.

Correlation Between Terminal States

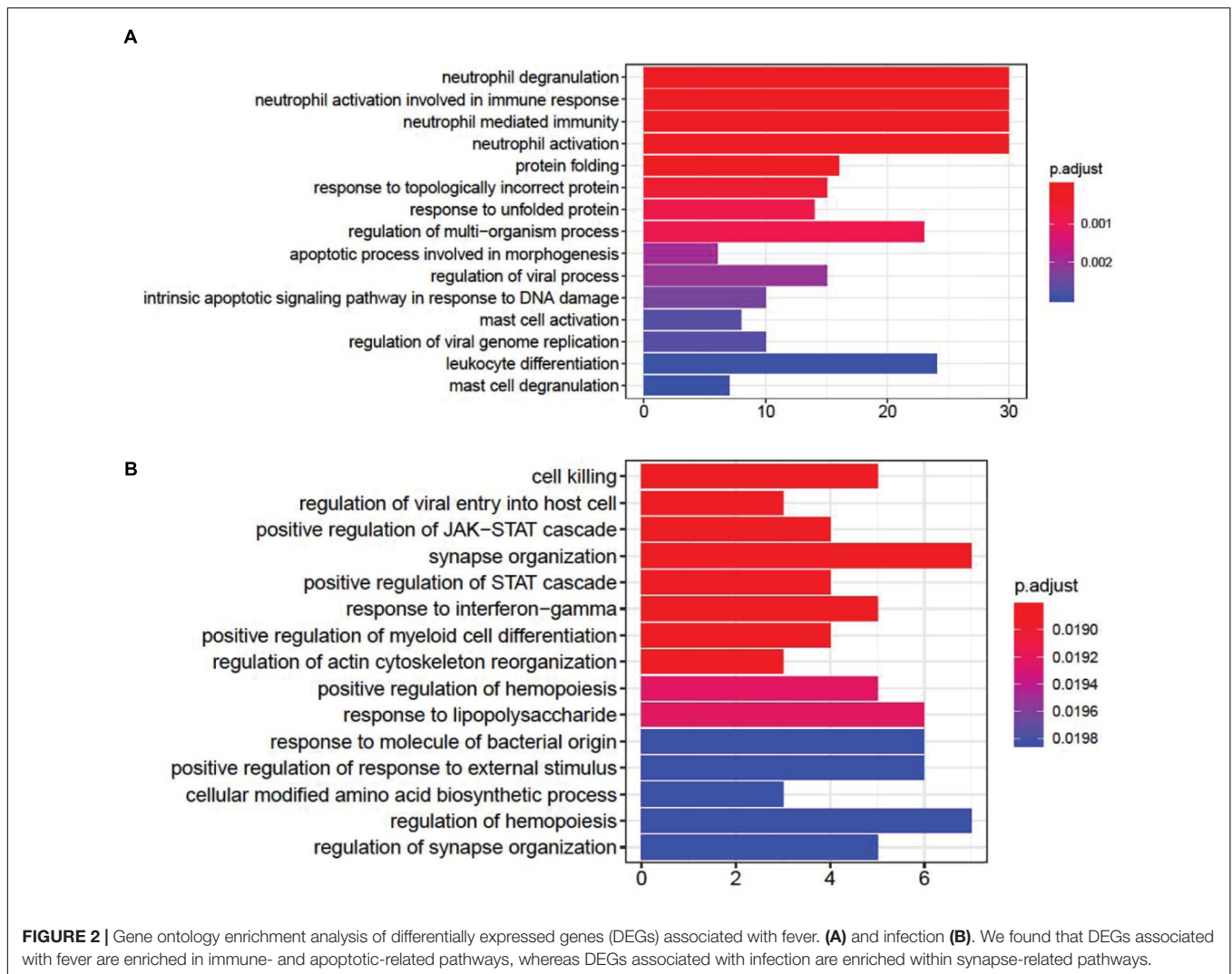
The correlation of log₂-FC effect sizes across terminal states also show contrasting directional effects on gene expression related to terminal state (**Figure 3**). We found that *fever* positively correlated with *infection* and *unconsciousness* ($\rho = 0.49, 0.42$); whereas *difficulty breathing* positively correlated with *artificial ventilation* ($\rho = 0.37$). We also observed that *fever* negatively correlated with *artificial ventilation* ($\rho = -0.48$), *infection* negatively correlated with *surgery* ($\rho = -0.44$), and *unconsciousness* negatively correlated with *difficulty breathing* and *artificial ventilation* ($\rho = -0.31, -0.38$). Based on the log₂-FC correlation, we concluded that the relationships among terminal states are complex and may have an opposite effect on gene expression.

Considering that the correlation of effect size was caused by the correlation of clinical phenotypes, we checked whether terminal states correlated clinically to participants. We found that *fever* and *infection* were significantly correlated ($P = 0.0002$, Fisher's exact test), and that *fever* and *difficulty breathing* were significantly correlated ($P = 0.01$, Fisher's exact test). Only *fever* and *infection* showed significant correlation both in gene expression level and clinical level, indicating a close relationship within this data. Other terminal states which have correlation in gene expression level showed insignificant correlation clinically. These results indicated that the correlation of gene expression in different terminal state was not decided by clinical correlation.

Agonal-Related Co-expression Networks Reveal Cell Type-Related Modules

Investigating a possible relationship between gene co-expression and agonal factors, we performed weighted correlation network analysis (WGCNA) on constructed gene co-expression networks (**Figure 4A**). We further analyzed whether gene co-expression modules (ME) using the following methods: cell-type enrichment analysis, psychiatric disorders' candidate gene enrichment analysis and gene ontology analysis. Of the 18 identified modules, two upregulated modules (ME11, ME12) were significantly associated with the combination of *fever* and *infection* (FDR $p < 0.05$); the ME8 was significantly associated with *fever*, while the ME17 was significantly associated with *infection*; whereas, only one downregulated module (ME10) was associated with *fever*. Yet, no modules were significantly associated with *surgery*, *unconsciousness*, *difficulty breathing* or *artificial ventilation*.

Correlating the gene co-expression module and terminal states shows us that terminal states introduce different effects on brain gene co-expression patterns (**Figure 4B**). Since *fever* and *infection* are frequently comorbid conditions in the clinical setting, we calculated AFS using *fever* and *infection* as a combined phenotype in the ROSMAP data for more significant results. We found five modules (ME17, ME8, ME11, ME12, ME10 and ME1) associated with the AFS calculated in this way. The ME1 subthreshold in the correlation analysis of *fever* and *infection* separated but reached significance in the correlation analysis of AFS.



An analysis of cell-type enrichment of the gene co-expression module revealed brain cell type-specific modules (**Figure 5**). The ME12 is enriched for microglia and endothelial cell types, the ME11 is enriched for microglia cell types, and the ME1 is enriched for neuron cell types. Different cell type played different role in brain function, so it is surprising that the upregulated modules are enriched for microglia and endothelial cell types, while the downregulated modules are enriched for neurons. The cell-type enrichment analysis indicates that agonal factors may have a cell type-specific impact on gene expression.

We selected five cell type-enriched agonal-related gene co-expression modules (ME1, ME8, ME11, ME12, and ME17) as well as the fever/infection-related DEG to perform an analysis of psychiatric-disorder candidate gene enrichment. We used candidate genes of autism spectrum disorder (ASD) (Basu et al., 2009; Darnell et al., 2011; Voineagu et al., 2011; Gupta et al., 2014; Li et al., 2016; Parikshak et al., 2016; Wang and Kaufman, 2016; Gandal et al., 2018a,b), major depression disorder (MDD) (Gandal et al., 2018a), and schizophrenia

(SCZ) (Lewis et al., 2003; Allen et al., 2008; The International Schizophrenia Consortium, 2008; Basu et al., 2009; Ng et al., 2009; Ayalew et al., 2012; Chen et al., 2013; He et al., 2013; Luo et al., 2014; Wang and Kaufman, 2016; Gandal et al., 2018a,b) from previous studies. The results showed the ME1, ME8, ME11 and ME12 significantly enriched in ASD and SCZ candidate genes (**Figure 6**). Also, we discovered that DEGs associated with fever were enriched in ASD and SCZ candidate genes which were discovered by co-expression network analysis.

The module shows a strong gene enrichment by cell-type, which can be associated with various biological processes. The ME1 was enriched for cell activity pathways, such as multicellular organismal homeostasis (adjust $p = 2.27E-05$), extracellular structure organization (adjust $p = 5.0E-06$), regulation of DNA-binding transcription factor activity (adjust $p = 4.23E-07$, **Figure 7A**). The gene expression level of ME1 was down-regulated, suggesting that related gene expression pathway was suppressed. The ME12 strongly enriched for pathways such as the mRNA catabolic process, RNA catabolic process, translational initiation, and protein localization to endoplasmic reticulum

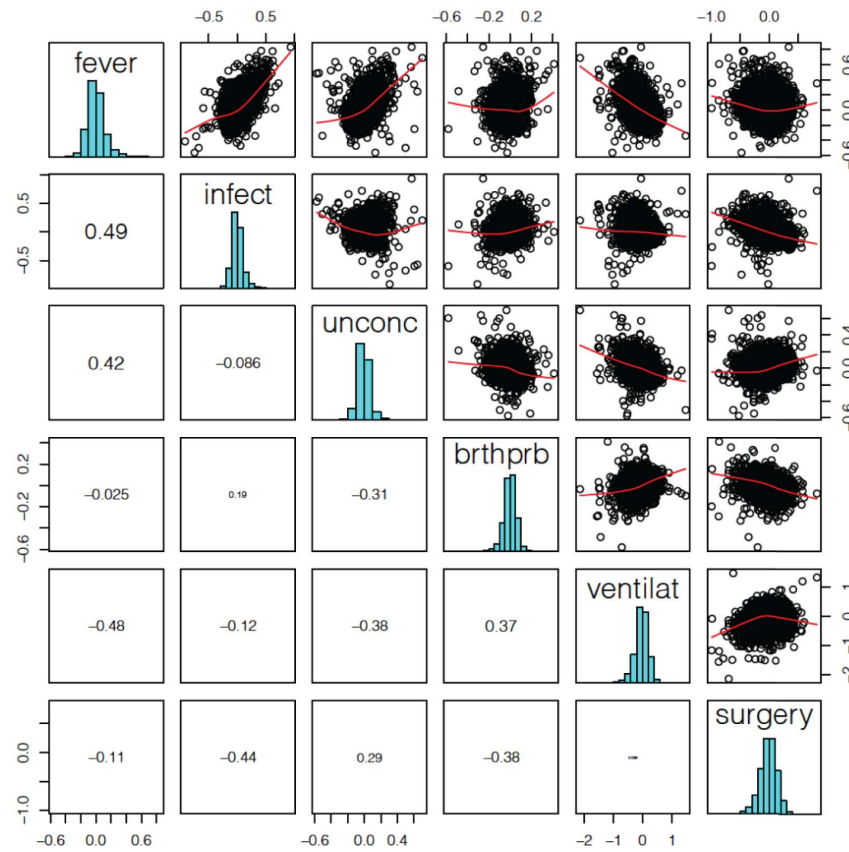


FIGURE 3 | Terminal state associated differentially expressed gene (DEG) log₂-FC effect size correlation. We found a significant positive correlation between fever and infection, difficulty breathing and artificial ventilation. We also found a negative correlation between fever and ventilation, infection and surgery, and unconsciousness and ventilation.

(**Figure 7B**). The gene expression level of the ME12 was up-regulated, which means that these cellular functions are activated to rescue the cell's basic function. The ME11, which was enriched in microglia and endothelial cells, represents the biological processes of glial cells (**Figure 7C**). The top gene ontology enrichment pathways of this module were synapse-related, including pre-synaptic, endomembrane system organization, modulation of chemical synaptic transmission and synapse organization. The up-regulating of synapse-related function indicated that microglia and endothelial cells are activated to protect neuron cells during the terminal state.

Utilizing Unknown Factors Analysis to Predict Agonal Factors

In modern, high-throughput biomolecular experiments, unmeasured or unmodeled factors can confound the primary variables and confuse the results. Researchers usually use a hidden factors estimation method to model large-scale noise dependence. This dependence can be caused by unmeasured or unmodeled factors. These models include surrogate variable analysis (SVA) for gene expression data and PEER, designed for transcriptomic data from eQTL analysis. In our study, we

used SVA and PEER to detect and correct for agonal factors. This also assisted in simulating agonal factors that were not otherwise recorded. For SVA, we detected 16 surrogate variables and correlated the 16 surrogate variables (sv) with terminal states (**Figure 8A**). Results showed that 10 of the 16 surrogate variables are significantly correlated with at least one agonal factor. We also observed that several surrogate variables correlated with more than 3 agonal factors. For example, sv2 is positively correlated with *surgery* but negatively correlated with *infection* and AFS; while, sv10 is negatively correlated with breathing difficulty, *fever*, *infection* and the AFS. These results also suggest a reverse effect for surgery and infection and a similar effect for *fever* and *infection*. In PEER analysis, we detected 15 hidden factors that explained the variants in the data (**Figure 8B**). For 15 hidden factors, we found 10 factors significantly correlated with terminal states. Similarly, to surrogate variable correlation, *surgery* and *infection* showed an opposite correlation to hidden factor 2; while *fever*, *infection*, and the AFS showed correlation in the same direction for hidden factors 4 and 13.

Linear regression analysis succeeded in correcting for agonal-related surrogate variables (SVA) and for hidden factors (PEER). We performed principal variance component analysis (PVCA) of the gene expression matrix before and after correcting for

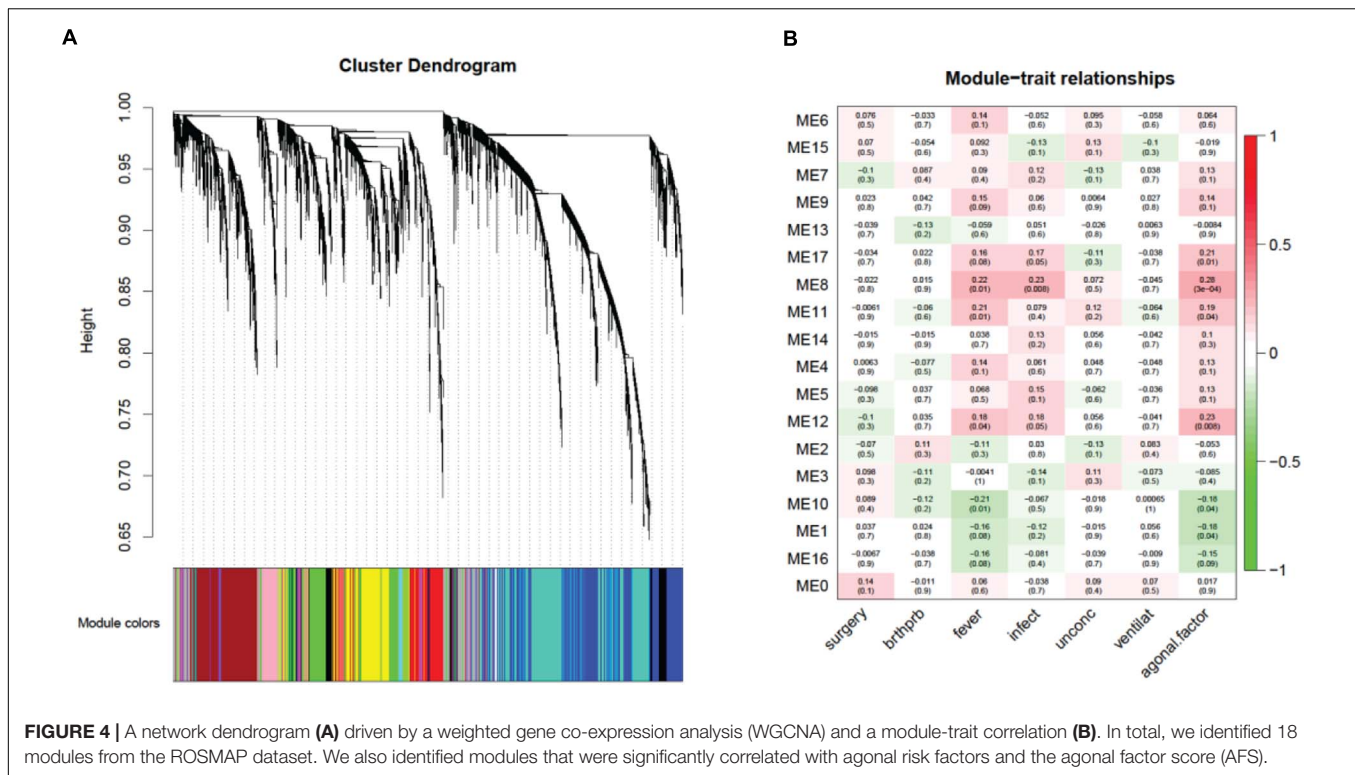


FIGURE 4 | A network dendrogram (A) driven by a weighted gene co-expression analysis (WGCNA) and a module-trait correlation (B). In total, we identified 18 modules from the ROSMAP dataset. We also identified modules that were significantly correlated with agonal risk factors and the agonal factor score (AFS).

10 agonal-related surrogate variables (sv 1, 2, 4, 5, 6, 7, 9, 10, 12, and 14). Results showed a decreased variance for most phenotypes (Figures 9A,B). After selecting race as a phenotype of interest, we found that the variance of race increased after correction in SVA. The variance of mechanical ventilation also increased, due to a lack of its correlation to surrogate variables. We also performed PVCA on the quantile normalized gene expression matrix before and after correcting for the 10 agonal-related hidden factors (factors 1, 2, 4, 5, 8, 9, 10, 12, 13, and 15). Terminal state variation decreased except for the factor of mechanical ventilation, due to the lack of its correlation also with hidden factors (Figures 9C,D). After performing differential gene expression analysis following correction for hidden factors in PEER and surrogate variables in SVA, we found no DEGs for *fever* nor for *infection*. In conclusion, if researchers have not collected agonal related phenotypes, correction for unknown factors can still occur using such methods as SVA or PEER.

We investigated previous studies' results and found them enriched in genes associated with the agonal-related module. We then applied SVA to these datasets and evaluated the variables from agonal factors. We used microarray data of SCZ, ASD and IBD, and we compared the DEGs before and after the SVA adjustment (Table 2). Performed a meta-analysis of 5 SCZ microarray data, we found 2044 DEGs of SCZ (FDR<0.05). After correcting 28 surrogate variables which had no significant correlation by disease group, we found 474 DEGs with 1628 genes filtered. Filtered genes were overlapped with fever-related DEGs (63 overlapped genes) and enriched in the ME8 ($p = 0.006$) and the ME12 ($p = 0.004$). Filtered genes were also enriched in hypoxia and oxygen level related pathways (Figure 10A).

We also found filtered genes in ASD as well as filtered genes enriched in oxygen levels-related pathways (Figure 10B). However, IBD filtered genes were not enriched in agonal-related pathways (Figure 10C).

DISCUSSION

We present a transcriptomic data analysis of human postmortem brain from a public database, which provides a framework for understanding how different terminal states contribute to gene expression changes in postmortem brain tissue. We performed data analysis to identify genes that are expressed differently in various abnormal agonal conditions compare to normal post-mortem controls. Many genes show altered expression after undergoing a terminal state characterized mainly by fever and infection, enriched for cellular stress-related pathways. Cell proportion of microglia and neuron also showed potential alteration in fever and infection samples. To determine whether terminal states related to other established disease candidate genes and to annotate its functional role in the human brain, we generated gene co-expression networks. We identified gene co-expression modules that significantly correlate with combination of fever and infection, drawing in positively co-expressed oligodendrocyte and microglia cell types as well as negatively co-expressed neuron cell types strongly enriched for SCZ and ASD candidate genes. Being a biological factor, terminal states can be adjusted by linear regression analysis or unknown factors adjustment methods, if it is not recorded. Surprisingly, we found that data from previous studies may

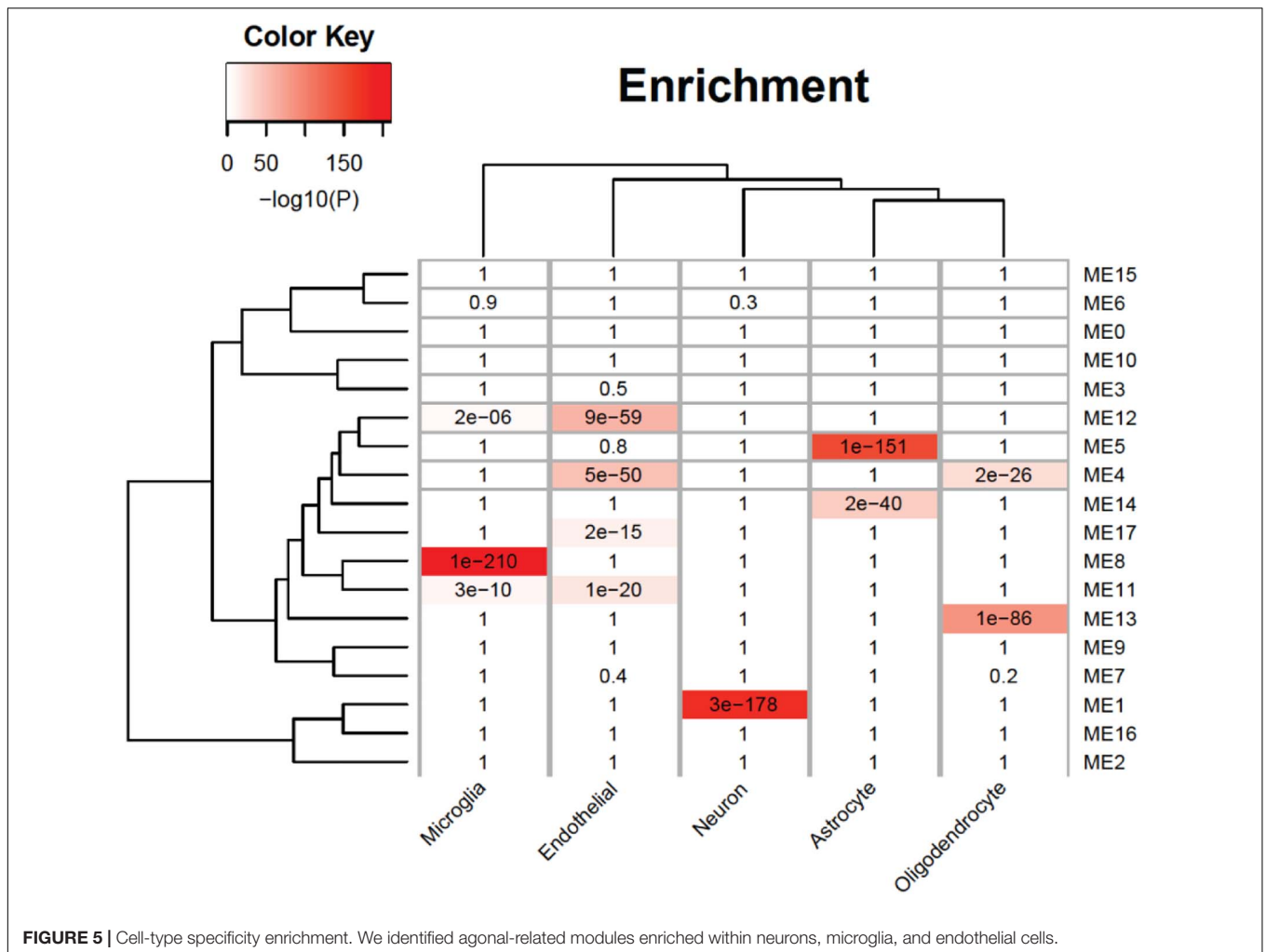
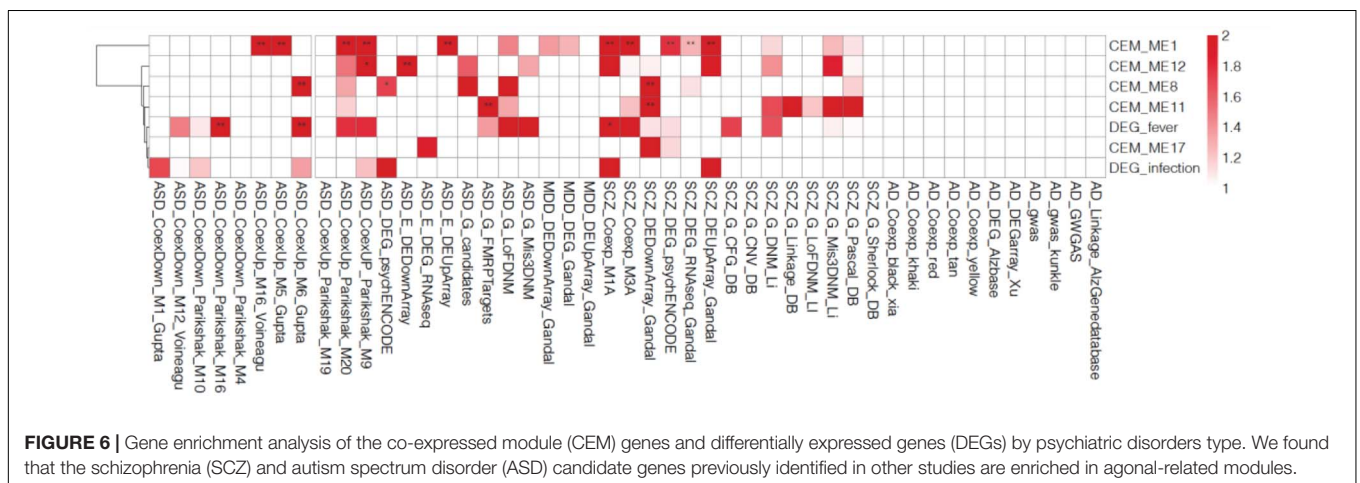


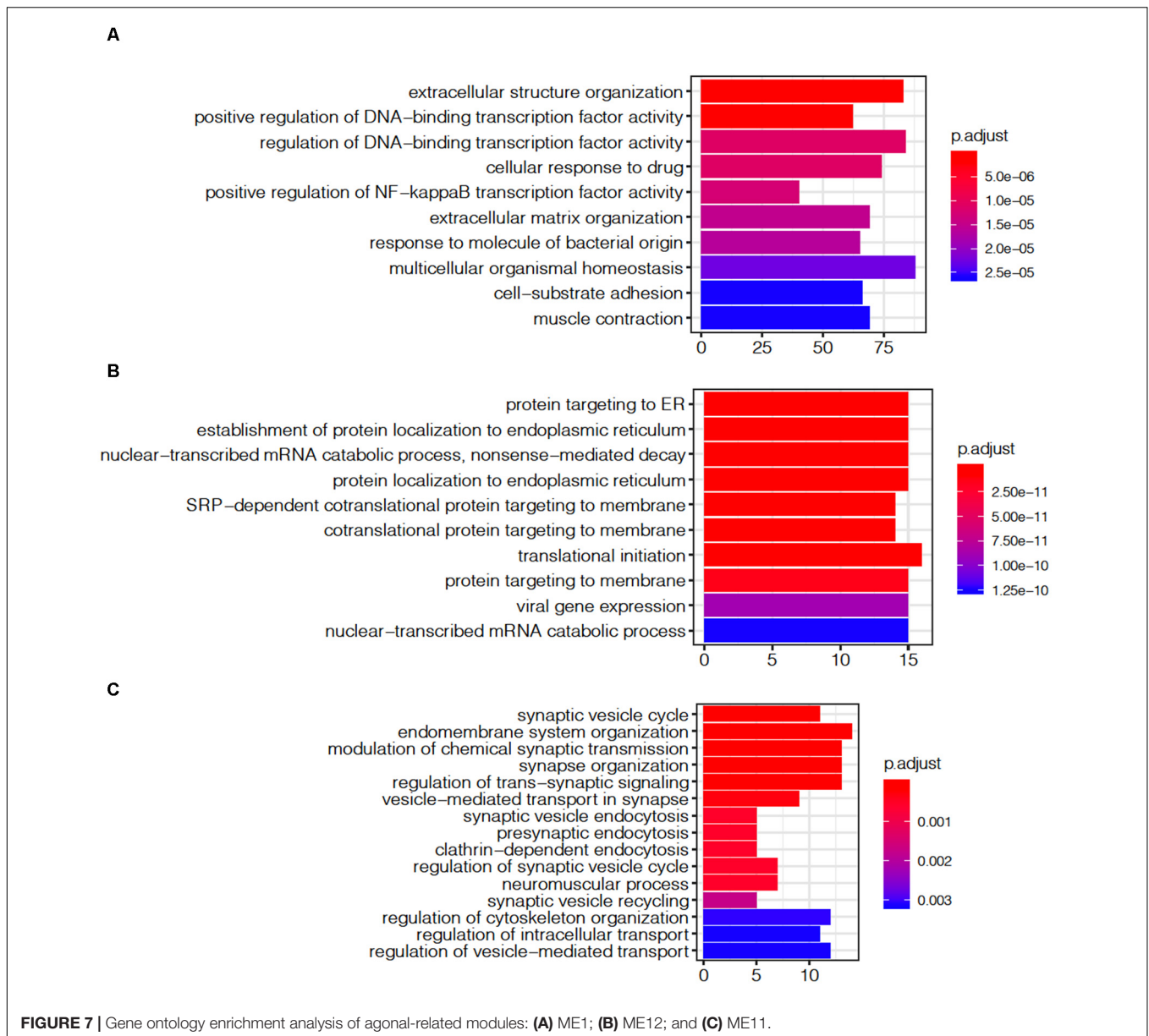
FIGURE 5 | Cell-type specificity enrichment. We identified agonal-related modules enriched within neurons, microglia, and endothelial cells.



be confounded by agonal factors that were not accounted for. We found gene sets from studies results were enriched for hypoxia- and oxygen level-related pathways, while the gene sets can be filtered by SVA. Our study emphasizes that agonal factors are important biological factors, and that agonal

factors should be documented and adjusted in postmortem brain tissue studies.

We performed data analysis aimed at *fever, infection, unconsciousness, breathing difficulty, artificial respiration and pre-mortem surgery*, only fever and infection related to altered



gene expression levels during the agonal period. We hypothesize the alteration possibly reflected by brain cell proportion changes (The International Schizophrenia Consortium, 2008). According to cell deconvolution, we found cell proportion of microglia rises in fever sample. Similar pattern was found in gene co-expression network. We found fever- and infection-related modules are enriched for microglia-specific cell markers, which showed up-regulation for gene expression. The coincident results of cell type changes indicate that specific brain cell types have different sensitivity to terminal states. In fever, parts of the brain become inflamed, which may cause microglia high expressed as active immune defense. Besides, gene co-expression modules are also enriched for neuron- and endothelial-specific cell markers. The results indicated that, in stressed brain environment of agonal, neuron may be more vulnerable to

the agonal state as compared with microglia and endothelial cells. One previous study reported that hypoxia was associated with increasing endothelial-specific expression and decreasing neuron-specific expression (Leek et al., 2012). The combination of fever and infection may also be associated with brain hypoxia, reflecting the vulnerability of neurons to low oxygen and/or severe infection (Banasiak and Haddad, 1998). Besides, we found fever may activate a cellular response to unfolded protein pathways after exposure to a stressed environment. It was reported that in cancer cells, hypoxia can activate components of this pathway (Fels and Koumenis, 2006). We also need more experiments to confirm the cell proportion changes in fever state.

Agonal factors other than fever and infection did not show any significant DEGs nor any associated modules within

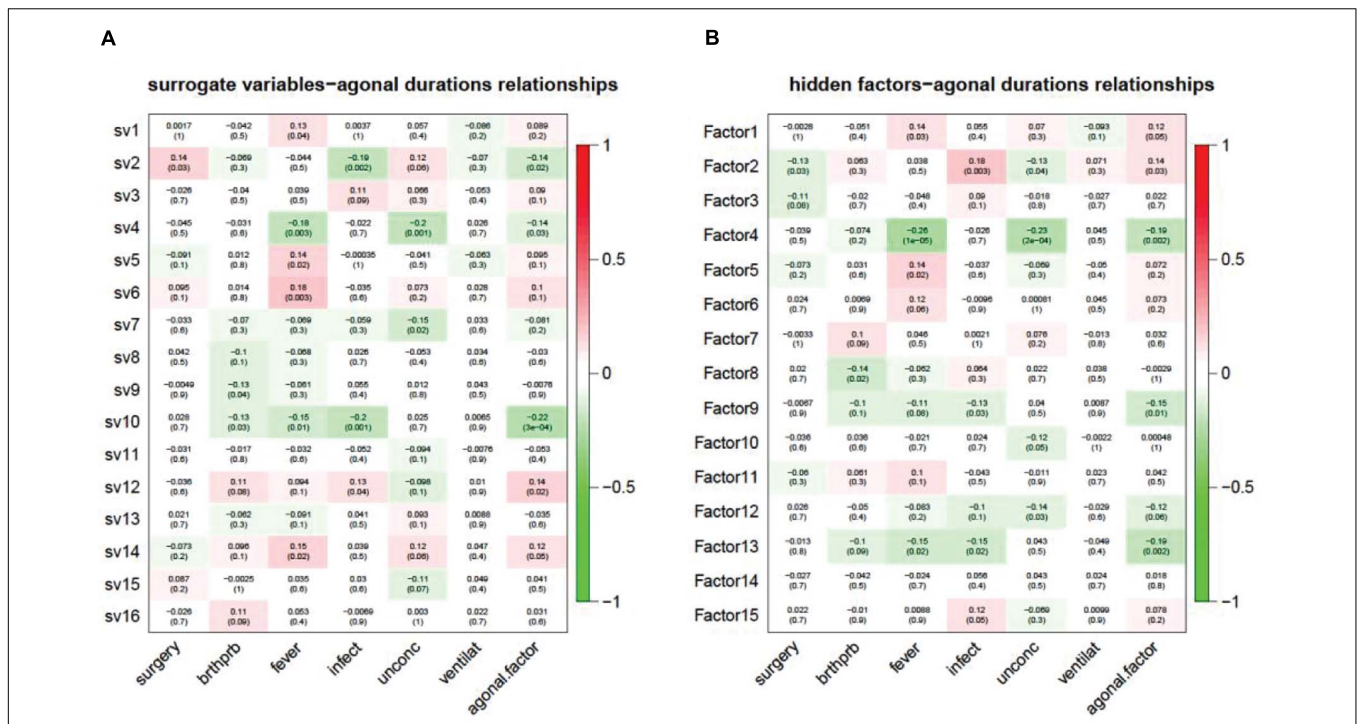


FIGURE 8 | Unknown factors discovered by SVA (A) and PEER (B) and terminal state correlation. (A) Surrogate variables 1, 2, 4, 5, 6, 7, 9, 10, 12, and 14 have a significant correlation with terminal states; (B) Hidden factors 1, 2, 4, 5, 8, 9, 10, 12, 13, and 15 have a significant correlation with terminal states.

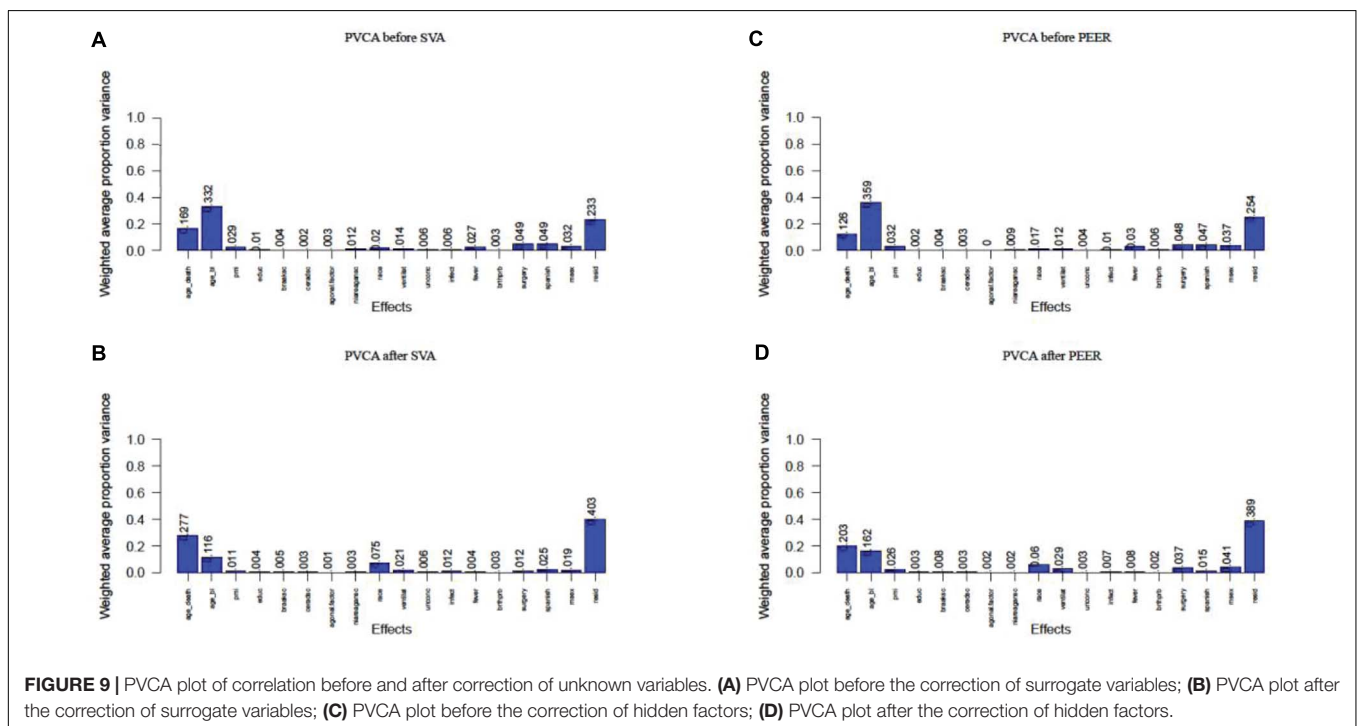


FIGURE 9 | PVCA plot of correlation before and after correction of unknown variables. (A) PVCA plot before the correction of surrogate variables; (B) PVCA plot after the correction of surrogate variables; (C) PVCA plot before the correction of hidden factors; (D) PVCA plot after the correction of hidden factors.

the dataset. We compared the log₂-FC of all agonal factors and found a negative correlation between artificial respiration and fever. Likewise, our module-trait correlation analysis yielded similar phenomena. When we combined all agonal

factors to calculate AFS (i.e., surgery, difficulty breathing, fever, infection, unconsciousness, and artificial respiration), no modules correlated significantly to AFS. However, when we calculated AFS based on fever and infection, a greater

TABLE 2 | Data replication for SVA correction.

	Raw DEGs (without SVA)	Number of variables	DEGs (with SVA)	Filtered gene from raw DEGs
SCZ	2044	28	474	1628
ASD	2470	2	1934	958
IBD	9165	1	7639	1539

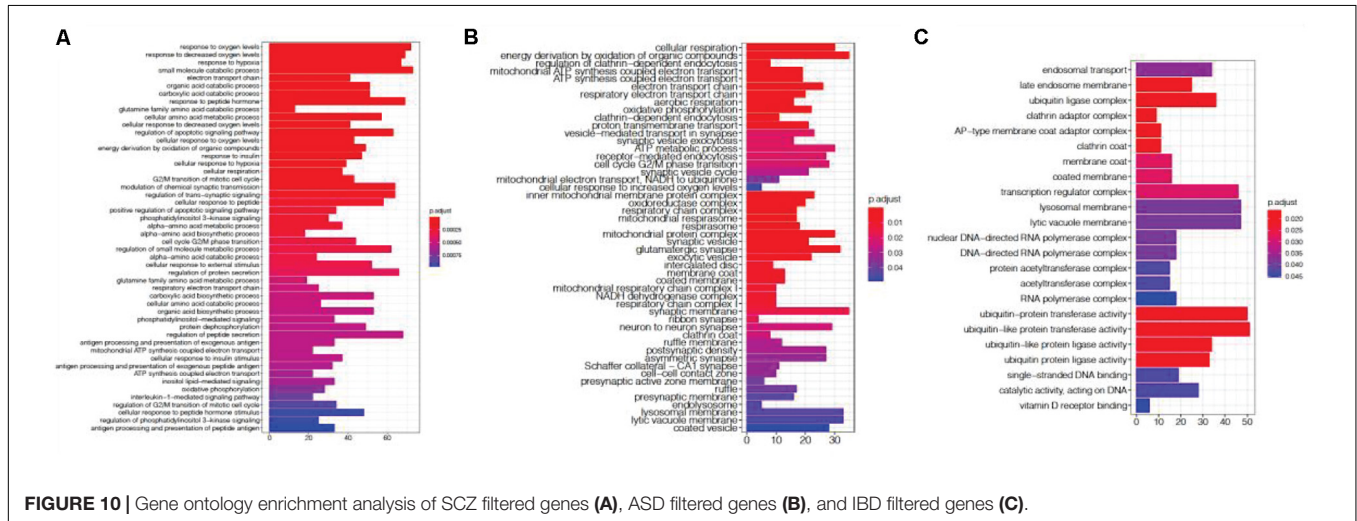


FIGURE 10 | Gene ontology enrichment analysis of SCZ filtered genes (A), ASD filtered genes (B), and IBD filtered genes (C).

number of gene co-expression modules showed significant correlation with AFS.

Previous studies simply added the manner of death with terminal states to define the severity of agonal conditions, but our findings suggest that that their method may confound the various agonal effects. We found that terminal states may contribute uniquely to gene expression. Therefore, we hypothesize that differences of gene expression come from the consequences brought on by the various agonal factors. While some agonal factors, such as medical interventions like artificial respiration try to return the brain’s extracellular environment to normal, other conditions, such as fever and infection, intensify stress upon the brain’s extracellular environment.

Some researchers have addressed the concern that agonal factors may represent considerable confounders. For example, the Netherland Brain Bank suggested that it is necessity of recording agonal factors; furthermore, they emphasize that researchers should ensure that patient and control groups match for as many known confounding factors as possible, including agonal states and stress of dying (Bao and Swaab, 2018). In another example, Ramaker et al. sought to avoid the variability of agonal factors connected to an extended dying process. They included only post-mortem brain samples from individuals who experienced violent fast deaths (Ramaker et al., 2017). Nevertheless, most postmortem brain studies still did not account for agonal confounders. We checked the expression data from the public database BrainEXP (Jiao et al., 2019) and found that most of the contributing datasets did not collect agonal information. This is also the reason that we were unable to find another comprehensive dataset to replicate results. Moreover, studies proved to have an inconsistent definition to the various agonal

states, making result replication problematic. Datasets with larger sample sizes are needed in addition to comprehensive recordings of agonal factors in order to accurately evaluate their effects.

Previous gene expression analysis of post-mortem brains may be confounded by terminal states, which may have resulted in data errors and false positive results. Our agonal-related modules revealed the enrichment of several candidate genes from previous neuropsychiatric studies. Those studies had not collected agonal factors nor had they corrected for unknown factors. This phenomenon especially exists in post-mortem brain samples. After we applied SVA to microarray data of samples from patients with SCZ, ASD and IBD, we found hypoxia-related pathways and oxygen level-related pathways. In the IBD data, which was not from brain tissue, we did not find any agonal-related pathways. This phenomenon suggests that the post-mortem brain is especially vulnerable to agonal factors. Moreover, we recommend a standardized data correction method to minimize the contribution of agonal factors. Normally, researchers can use linear regression analysis to correct for agonal factors similar to the measures used to correct for biological factors. These methods can detect variants induced by agonal factors. This is an important step for quality control in data preprocess. However, we cannot rule out the possible false negative or positive associations created by the use of PEER/SVA. Large replicate data is needed. We strongly suggest that researchers recheck previous results, since we identified several study results that were enriched in agonal-related modules and one study’s data confirmed to have been confounded by agonal factors. Our results provide a clear guidance for taking agonal factors into account and correcting for them in future research.

Our sample size was relatively small. We attribute the lack of significant DEGs or gene co-expression modules to these relatively small sample sizes, specifically, because of the limited number of *surgery* and *artificial respiration* phenotypes. A greater sample size will be necessary to validate the effects of unconsciousness, breathing difficulty, artificial respiration and pre-mortem surgery. We lacked a suitable data to replicate our findings. We did not find a replicate data that carry the terminal state information that match our discovery data and serve the purpose. For this reason, we also lack a comprehensive record of agonal factors of gene expression data and are unable to evaluate overall agonal factors systematically. Furthermore, to quantify and compare the agonal factors, we can compare the gene expression patterns of neurosurgical samples and postmortem samples in the future studies.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

JD is the first author. CC is the corresponding author. YC and CL helped with the study design and manuscript writing.

REFERENCES

- Allen, N. C., Bagade, S., McQueen, M. B., Ioannidis, J. P., Kavvoura, F. K., Khoury, M. J., et al. (2008). Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database. *Nat. Genet.* 40, 827–834. doi: 10.1038/ng.171
- Atz, M., Walsh, D., Cartagena, P., Li, J., Evans, S., Choudary, P., et al. (2007). Methodological considerations for gene expression profiling of human brain. *J. Neurosci. Methods* 163, 295–309. doi: 10.1016/j.jneumeth.2007.03.022
- Ayalew, M., Le-Niculescu, H., Levey, D. F., Jain, N., Changala, B., Patel, S. D., et al. (2012). Convergent functional genomics of schizophrenia: from comprehensive understanding to genetic risk prediction. *Mol. Psychiatry* 17, 887–905. doi: 10.1038/mp.2012.37
- Banasiak, K. J., and Haddad, G. G. (1998). Hypoxia-induced apoptosis: effect of hypoxic severity and role of p53 in neuronal cell death. *Brain Res.* 797, 295–304. doi: 10.1016/S0006-8993(98)00286-8
- Bao, A. M., and Swaab, D. F. (2018). The art of matching brain tissue from patients and controls for postmortem research. *Handb. Clin. Neurol.* 150, 197–217. doi: 10.1016/B978-0-444-63639-3.00015-3

RD and YJ helped with the study design. JT, SL, and MX helped with manuscript writing. ML and JZ helped with data collection. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (Grant Nos. 82022024, 31970572, 31571312, and 81401114), Innovation-driven Project of Central South University (Grant No. 2020CX003), the National Key R&D Project of China (Grant No. 2016YFC1306000) (to CC), the National Natural Science Foundation of China (Grant No. 31871276), the National Key R&D Project of China (Grant No. 2017YFC0908701) (to CL), and the Fundamental Research Funds for the Central Universities of Central South University (Grant No. 1053320184146) (to JD).

ACKNOWLEDGMENTS

We thank our patient donors and their families who have contributed to this research via postmortem brain tissue donations. We thank Rush University and the ROSMAP project for the RNA-sequencing data. We also thank Richard F. Kopp and Liz Kuney from SUNY Upstate Medical University, for their language editing contribution, which greatly improved the manuscript. This manuscript has been released as a pre-print at <https://www.biorxiv.org/content/10.1101/2020.07.11.198523v1> (Dai et al., 2020).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2021.614142/full#supplementary-material>

- Barton, A., Pearson, R., Najlerahim, A., and Harrison, P. (1993). Pre-and postmortem influences on brain RNA. *J. Neurochem.* 61, 1–11. doi: 10.1111/j.1471-4159.1993.tb03532.x
- Basu, S. N., Kollu, R., and Banerjee-Basu, S. (2009). AutDB: a gene reference resource for autism research. *Nucleic Acids Res.* 37, D832–D836. doi: 10.1093/nar/gkn835
- Burke, W. J., O'malley, K. L., Chung, H. D., Harmon, S. K., Miller, J. P., and Berg, L. (1991). Effect of pre-and postmortem variables on specific mRNA levels in human brain. *Mol. Brain Res.* 11, 37–41. doi: 10.1016/0169-328X(91)90018-S
- Chen, C., Cheng, L., Grennan, K., Pibiri, F., Zhang, C., Badner, J. A., et al. (2013). Two gene co-expression modules differentiate psychotics and controls. *Mol. Psychiatry* 18, 1308–1314. doi: 10.1038/mp.2012.146
- Chow, M. L., Pramparo, T., Winn, M. E., Barnes, C. C., Li, H. R., Weiss, L., et al. (2012). Age-dependent brain gene expression and copy number anomalies in autism suggest distinct pathological processes at young versus mature ages. *PLoS Genet.* 8:e1002592. doi: 10.1371/journal.pgen.1002592
- Dai, J., Chen, Y., Chen, C., and Liu, C. (2020). Agonal factors distort gene-expression patterns in human postmortem brains. *bioRxiv* [Preprint]. doi: 10.1101/2020.07.11.198523

- Darnell, J. C., Van Driesche, S. J., Zhang, C., Hung, K. Y. S., Mele, A., Fraser, C. E., et al. (2011). FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* 146, 247–261. doi: 10.1016/j.cell.2011.06.013
- Fels, D. R., and Koumenis, C. (2006). The PERK/eIF2 α /ATF4 module of the UPR in hypoxia resistance and tumor growth. *Cancer Biol. Ther.* 5, 723–728. doi: 10.4161/cbt.5.7.2967
- Gandal, M. J., Haney, J. R., Parikshak, N. N., Leppa, V., Ramaswami, G., Hartl, C., et al. (2018a). Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science* 359, 693–697. doi: 10.1126/science.aad6469
- Gandal, M. J., Zhang, P., Hadjimichael, E., Walker, R. L., Chen, C., Liu, S., et al. (2018b). Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science* 362:eaat8127.
- Garbett, K., Ebert, P. J., Mitchell, A., Lintas, C., Manzi, B., Mirnics, K., et al. (2008). Immune transcriptome alterations in the temporal cortex of subjects with autism. *Neurobiol. Dis.* 30, 303–311. doi: 10.1016/j.nbd.2008.01.012
- Granlund, A., Flatberg, A., Østvik, A. E., Drozdov, I., Gustafsson, B. I., Kidd, M., et al. (2013). Whole genome gene expression meta-analysis of inflammatory bowel disease colon mucosa demonstrates lack of major differences between Crohn's disease and ulcerative colitis. *PLoS One* 8:e56818. doi: 10.1371/journal.pone.0056818
- Gupta, S., Ellis, S. E., Ashar, F. N., Moes, A., Bader, J. S., Zhan, J., et al. (2014). Transcriptome analysis reveals dysregulation of innate immune response genes and neuronal activity-dependent genes in autism. *Nat. Commun.* 5:5748. doi: 10.1038/ncomms6748
- Hagenauer, M. H., Schulmann, A., Li, J. Z., Vawter, M. P., Walsh, D. M., Thompson, R. C., et al. (2018). Inference of cell type content from human brain transcriptomic datasets illuminates the effects of age, manner of death, dissection, and psychiatric diagnosis. *PLoS One* 13:e0200003. doi: 10.1371/journal.pone.0200003
- Hardy, J., Wester, P., Winblad, B., Gezelius, C., Bring, G., and Eriksson, A. (1985). The patients dying after long terminal phase have acidotic brains; implications for biochemical measurements on autopsy tissue. *J. Neural. Transm.* 61, 253–264. doi: 10.1007/BF01251916
- Harrison, P. J., Barton, A. J., Najlerahim, A., McDonald, B., and Pearson, R. C. A. (1991). Regional and neuronal reductions of polyadenylated messenger RNA in Alzheimer's disease. *Psychol. Med.* 21, 855–866. doi: 10.1017/S0033291700029858
- He, X., Fuller, C. K., Song, Y., Meng, Q., Zhang, B., Yang, X., et al. (2013). Sherlock: detecting gene-disease associations by matching patterns of expression QTL and GWAS. *Am. J. Hum. Genet.* 92, 667–680. doi: 10.1016/j.ajhg.2013.03.022
- Iwamoto, K., Bundo, M., and Kato, T. (2005). Altered expression of mitochondria-related genes in postmortem brains of patients with bipolar disorder or schizophrenia, as revealed by large-scale DNA microarray analysis. *Hum. Mol. Genet.* 14, 241–253. doi: 10.1093/hmg/ddi022
- Jaffe, A. E. (2016). Postmortem human brain genomics in neuropsychiatric disorders—how far can we go? *Curr. Opin. Neurobiol.* 36, 107–111. doi: 10.1016/j.conb.2015.11.002
- Jiao, C., Yan, P., Xia, C., Shen, Z., Tan, Z., Tan, Y., et al. (2019). BrainEXP: a database featuring with spatiotemporal expression variations and co-expression organizations in human brains. *Bioinformatics* 35, 172–174. doi: 10.1093/bioinformatics/bty576
- Johnston, N. L., Cerevna, J., Shore, A. D., Torrey, E. F., Yolken, R. H., and Consortium, S. N. (1997). Multivariate analysis of RNA levels from postmortem human brains as measured by three different methods of RT-PCR. *J. Neurosci. Methods* 77, 83–92. doi: 10.1016/S0165-0270(97)00115-5
- Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559. doi: 10.1186/1471-2105-9-559
- Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., and Storey, J. D. (2012). The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 28, 882–883. doi: 10.1093/bioinformatics/bts034
- Lewis, C. M., Levinson, D. F., Wise, L. H., Delisi, L. E., Straub, R. E., Hovatta, I., et al. (2003). Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: Schizophrenia. *Am. J. Hum. Genet.* 73, 34–48. doi: 10.1086/376549
- Li, J., Cai, T., Jiang, Y., Chen, H., He, X., Chen, C., et al. (2016). Genes with de novo mutations are shared by four neuropsychiatric disorders discovered from NPdenovo database. *Mol. Psychiatry* 21, 290–297. doi: 10.1038/mp.2015.40
- Li, J. Z., Vawter, M. P., Walsh, D. M., Tomita, H., Evans, S. J., Choudary, P. V., et al. (2004). Systematic changes in gene expression in postmortem human brains associated with tissue pH and terminal medical conditions. *Hum. Mol. Genet.* 13, 609–616. doi: 10.1093/hmg/ddh065
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8
- Luo, X., Huang, L., Han, L., Luo, Z., Hu, F., Tieu, R., et al. (2014). Systematic prioritization and integrative analysis of copy number variations in schizophrenia reveal key schizophrenia susceptibility genes. *Schizophr. Bull.* 40, 1285–1299. doi: 10.1093/schbul/sbu045
- Mathys, H., Davila-Velderrain, J., Peng, Z., Gao, F., Mohammadi, S., Young, J. Z., et al. (2019). Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* 570, 332–337. doi: 10.1038/s41586-019-1195-2
- Maycox, P. R., Kelly, F., Taylor, A., Bates, S., Reid, J., Logendra, R., et al. (2009). Analysis of gene expression in two large schizophrenia cohorts identifies multiple changes associated with nerve terminal function. *Mol. Psychiatry* 14, 1083–1094. doi: 10.1038/mp.2009.18
- McCullumsmith, R. E., Hammond, J. H., Shan, D., and Meador-Woodruff, J. H. (2014). Postmortem brain: an underutilized substrate for studying severe mental illness. *Neuropsychopharmacology* 39, 65–87. doi: 10.1038/npp.2013.239
- Narayan, S., Tang, B., Head, S. R., Gilmartin, T. J., Sutcliffe, J. G., Dean, B., et al. (2008). Molecular profiles of schizophrenia in the CNS at different stages of illness. *Brain Res.* 1239, 235–248. doi: 10.1016/j.brainres.2008.08.023
- Ng, M. Y., Levinson, D. F., Faraone, S. V., Suarez, B. K., Delisi, L. E., Arinami, T., et al. (2009). Meta-analysis of 32 genome-wide linkage studies of schizophrenia. *Mol. Psychiatry* 14, 774–785. doi: 10.1038/mp.2008.135
- Noble, C. L., Abbas, A. R., Cornelius, J., Lees, C. W., Ho, G. T., Toy, K., et al. (2008). Regional variation in gene expression in the healthy colon is dysregulated in ulcerative colitis. *Gut* 57, 1398–1405. doi: 10.1136/gut.2008.148395
- Parikshak, N. N., Swarup, V., Belgard, T. G., Irimia, M., Ramaswami, G., Gandal, M. J., et al. (2016). Genome-wide changes in lncRNA, splicing, and regional gene expression patterns in autism. *Nature* 540, 423–427. doi: 10.1038/nature20612
- Ramaker, R. C., Bowling, K. M., Lasseigne, B. N., Hagenauer, M. H., Hardigan, A. A., Davis, N. S., et al. (2017). Post-mortem molecular profiling of three psychiatric disorders. *Genome Med.* 9:72. doi: 10.1186/s13073-017-0458-5
- Reinhart, V., Bove, S. E., Volfson, D., Lewis, D. A., Kleiman, R. J., and Lanz, T. A. (2015). Evaluation of TrkB and BDNF transcripts in prefrontal cortex, hippocampus, and striatum from subjects with schizophrenia, bipolar disorder, and major depressive disorder. *Neurobiol. Dis.* 77, 220–227. doi: 10.1016/j.nbd.2015.03.011
- Stegle, O., Parts, L., Piipari, M., Winn, J., and Durbin, R. (2012). Using probabilistic estimation of expression residuals (PEER) to obtain increased power and interpretability of gene expression analyses. *Nat. Protoc.* 7, 500–507. doi: 10.1038/nprot.2011.457
- The International Schizophrenia Consortium (2008). Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* 455, 237–241. doi: 10.1038/nature07239
- Tomita, H., Vawter, M. P., Walsh, D. M., Evans, S. J., Choudary, P. V., Li, J., et al. (2004). Effect of agonal and postmortem factors on gene expression profile: quality control in microarray analyses of postmortem human brain. *Biol. Psychiatry* 55, 346–352. doi: 10.1016/j.biopsych.2003.10.013
- Voineagu, I., Wang, X., Johnston, P., Lowe, J. K., Tian, Y., Horvath, S., et al. (2011). Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature* 474, 380–384. doi: 10.1038/nature10110
- Wang, L., Xia, Y., Chen, Y., Dai, R., Qiu, W., Meng, Q., et al. (2019a). Brain banks spur new frontiers in neuropsychiatric research and strategies for analysis and validation. *Genomics Proteomics Bioinformatics* 17, 402–414. doi: 10.1016/j.gpb.2019.02.002
- Wang, M., and Kaufman, R. J. (2016). Protein misfolding in the endoplasmic reticulum as a conduit to human disease. *Nature* 529, 326–335. doi: 10.1038/nature17041

- Wang, X., Park, J., Susztak, K., Zhang, N. R., and Li, M. (2019b). Bulk tissue cell type deconvolution with multi-subject single-cell expression reference. *Nat. Commun.* 10, 1–9. doi: 10.1038/s41467-018-08023-x
- Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics* 16, 284–287. doi: 10.1089/omi.2011.0118
- Zhang, Y., Sloan, S. A., Clarke, L. E., Caneda, C., Plaza, C. A., Blumenthal, P. D., et al. (2016). Purification and characterization of progenitor and mature human astrocytes reveals transcriptional and functional differences with mouse. *Neuron* 89, 37–53. doi: 10.1016/j.neuron.2015.11.013

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Dai, Chen, Dai, Jiang, Tian, Liu, Xu, Li, Zhou, Liu and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.