



# Association Analysis of Peripheral and CSF Biomarkers in Late Mild Cognitive Impairment

Tao Zhang<sup>1†</sup>, Wei Kong<sup>1\*†</sup>, Shuaiqun Wang<sup>1</sup> and Xiaoyang Mou<sup>2</sup>

<sup>1</sup> College of Information Engineering, Shanghai Maritime University, Shanghai, China, <sup>2</sup> Department of Biochemistry, Rowan University and Guava Medicine, Glassboro, NJ, United States

Research shows that late mild cognitive impairment (LMCI) has a high risk of turning into Alzheimer's disease (AD). Due to the invasion of detection methods and physical damage to the patients, it is not a convenient way to diagnose and detect early AD and LMCI by cerebrospinal fluid (CSF) data. So there is an urgent need to find the correlation between peripheral biological data and CSF data in the brain, and to find new diagnostic methods through changes in the peripheral biological data. Studies have shown that during the pathogenesis of LMCI and AD, peripheral immune cells specifically infiltrate into the brain through the blood–brain barrier, causing an imbalance in the brain's immune response and dysregulating the clearance of A $\beta$  in CSF. Therefore, in this paper, canonical correlation analysis (CCA) algorithm is presented to derive the correlation between peripheral and CSF biomarkers based on LMCI peripheral gene expression data and plasma marker information. Firstly, to explore the influence of the infiltration of peripheral blood immune cells on the brain, the abundance of 28 immune cells were calculated by using the gene set enrichment analysis algorithm of LMCI samples. Then, to identify the correlation between biomarkers inside and outside of the brain, we performed CCA to calculate the relationship between CSF and peripheral biomarkers. Results of CCA showed significant correlations between the variable sets of 8 peripheral biomarkers and the variable sets of CSF biomarkers (at 0.794). Finally, according to Kyoto Encyclopedia of Genes and Genomes and Gene Ontology analysis, it was found that the obtained peripheral biomarkers are involved in many immune-related pathways and functions which can be activated in peripheral blood of LMCI patients. Most related genes enriched in immune-related pathways and functions were up-regulated. Through receiver operating characteristic curve (ROC) analysis, it was also found that FP40/FP42 and type 1 T helper can accurately predict the pathological changes of LMCI (at 0.747).

**Keywords:** peripheral and CSF biomarkers, blood–brain barrier, late mild cognitive impairment (LMCI), gene set enrichment analysis (GSEA), canonical correlation analysis (CCA), receiver operating characteristic curve (ROC)

## OPEN ACCESS

### Edited by:

Xiaoyong Pan,  
Shanghai Jiao Tong University, China

### Reviewed by:

Zhonglong Zheng,  
Zhejiang Normal University, China  
Hui Liu,  
Dalian University of Technology, China

### \*Correspondence:

Wei Kong  
weikong@shmtu.edu.cn

<sup>†</sup>These authors share first authorship

### Specialty section:

This article was submitted to  
Computational Genomics,  
a section of the journal  
Frontiers in Genetics

**Received:** 22 May 2020

**Accepted:** 10 July 2020

**Published:** 12 August 2020

### Citation:

Zhang T, Kong W, Wang S and  
Mou X (2020) Association Analysis  
of Peripheral and CSF Biomarkers  
in Late Mild Cognitive Impairment.  
*Front. Genet.* 11:834.  
doi: 10.3389/fgene.2020.00834

## INTRODUCTION

Mild cognitive impairment (MCI) is a transitional stage in which normal aging develops into dementia (Stephan et al., 2007), but it is also an unstable state. In the follow-up study of MCI patients, the abundance of MCI developing into Alzheimer's disease (AD) increased year by year. It can be seen that MCI is an early warning signal for the onset of AD (Hansson et al., 2006;

Jessen et al., 2014), and the risk of conversion of late mild cognitive impairment (LMCI) to AD is the highest.

Cerebrospinal fluid (CSF) marker analysis is an effective method for diagnosing MCI could be a differentiating marker for the detection of prodromal AD from clinically diagnosed MCI patients (Park et al., 2019). However, due to the invasion of detection methods and physical damage to patients, it is not a convenient way to diagnose and detect early AD and LMCI by CSF. And the study of MCI in Magaki et al. (2007) found that the production of cytokines IL-6, IL-8, and IL-10 increased in peripheral blood, indicating that immune activation is an early phenomenon before AD. Changes in cytokines produced by PBMCs can be detected in MCI and can occur before clinical AD occurs. In the plasma marker study, plasma  $\beta$  amyloid has a certain relationship with  $\beta$  amyloid in the brain, and plasma A $\beta$  measurement can be used as an effective marker to measure A $\beta$  and tau (Risacher et al., 2019). Therefore, in this paper, the easy-to-measure peripheral blood gene data and plasma biomarkers are used to establish an association with CSF markers in the brain of LMCI patients.

To extract the correlation between biological features inside and outside the brain, feature information of biological signals inside and outside the brain needs to be extracted separately. For the calculation of immune cell abundance in peripheral blood, many new calculation methods can greatly enhance our investigation of immune cell subtypes. Among them, the CIBERSORT (Newman et al., 2015) and TIMER (Li et al., 2016) algorithms are based on deconvolution methods, inferring immune cells through gene expression profiling reference matrices and machine learning-based methods. However, these results are obtained from a mixture of simulated samples of different tissues (brain, heart, liver, lung, and tumor tissue), while in blood samples, it is difficult to distinguish (Novershtern et al., 2011; Shoemaker et al., 2012). Bolen et al. (2011) proposed a computational method (from enrichment-related subset prediction, SPEC). The SPEC algorithm is based on gene set enrichment analysis (GSEA; Quesenberry and Colvin, 2001; Subramanian et al., 2005), but it cannot distinguish subpopulations of immune cells in the tagged gene set used to measure immune cells. We obtained a new set of labeled genes through the article (Charoentong et al., 2017) and improved the SPEC algorithm. The advantage of this new algorithm is its robustness, which is insensitive to sample impurities and noise during preparation compared to deconvolution. And the subpopulation of immune cells can also be measured with high resolution.

Biological data sets in CSF can be used to diagnose diseases, but changes in a single indicator in the data set cannot measure the status of the disease. For peripheral biological indicators, a single biological indicator cannot accurately reflect physiological processes. Therefore, to find indicators that can reflect changes in CSF biomarker data in peripheral blood and more comprehensively reflect the physiological process of disease development, the canonical correlation analysis (CCA; Hotelling, 1992) algorithm is used in this paper to calculate the correlation between CSF data sets and peripheral biological data sets. Canonical correlation analysis is a multivariate statistical

model that maximizes the correlation between the two composite variables (Kabir et al., 2014). There are more features in the peripheral data, including 28 types of peripheral blood immune cell abundance data and four types of plasma A $\beta$  data. Canonical correlation analysis can more reliably measure the correlation between brain and peripheral biomarkers.

By analyzing the Pearson correlation between CSF data and peripheral data, we obtained that there were significant correlations between peripheral information (including six types of immune cells and two types of A $\beta$  data) and CSF data. Receiver operating characteristic curve (ROC) analysis found that FP40/FP42 (area under the ROC curve, 0.709) and type 1 T helper cell (area under the ROC curve, 0.703) have clinical significance in the diagnosis of the disease. Receiver operating characteristic curve verification found that the combined diagnosis of two biomarkers further improved the accuracy (area under the ROC curve, 0.747). By constructing a protein-protein interaction (PPI) network (Szklarczyk et al., 2015) to find key genes, and performing Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis, the biological processes, and mechanism for the immune-related correlations between peripheral and CSF biomarkers were discovered.

## MATERIALS AND METHODS

In this section, we describe the measurement of the abundance of peripheral blood immune cells in patients with LMCI and introduce three sources of data. The correlation was calculated and the biological reasons were analyzed.

### Data Sources and Preprocessing

Data used in this study were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI; Weiner and Veitch, 2015) database<sup>1</sup>. The ADNI database was launched in 2003 as a public-private partnership, led by the principal investigator Michael W. Weiner, MD. The ADNI participants have been recruited from more than 50 sites across the United States and Canada. The primary objective of the ADNI has been to test whether serial MRI, PET, other biological markers, and clinical or neuropsychological assessment can be combined to measure the progression of MCI and early AD. Alzheimer's Disease Neuroimaging Initiative database consists of three parts, including the ADNI 1, the ADNI Grand Opportunities, and the ADNI 2. To date, these three protocols have recruited more than 1500 adults (age range, 55–90 years) to participate in the research, including CN older individuals, persons with early or late MCI, and patients with early AD. The follow-up duration for each study group was specified in the protocols for the ADNI 1, ADNI 2, and ADNI Grand Opportunities. Regional ethics committees of all institutions approved of the study. Written informed consent was obtained from all study participants.

Gene expression profiling from blood samples of ADNI participants was contributed by Bristol-Myers Squibb (BMS) and performed at the BMS laboratories for 811 ADNI participants

<sup>1</sup><http://adni.loni.usc.edu>

from the ADNI WGS cohort. The Affymetrix Human Genome U219 Array (Affymetrix<sup>2</sup>, Santa Clara, CA) was used for expression profiling. Peripheral blood samples were collected using PAXgene tubes for RNA analysis. Blood RNA samples from 64 participants did not pass QC and were excluded from further processing. And we identified three questionable subjects from the additional QC steps and removed them. The data we finally downloaded was peripheral blood gene expression data containing 744 samples. The plasma amyloid-beta (A $\beta$ ) biomarkers and CSF biomarkers were also obtained from the ADNI database. The plasma A $\beta$  biomarkers contained 305 samples and the CSF biomarkers contained 1250 samples. By screening out samples that existed in all three data and collected three types of biological data in the same year (the samples with missing data were deleted), 36 samples containing three biological data were obtained. The 36 samples were labeled with disease status by using the file “AD Challenge Training Data: Clinical (Updated)” downloaded from the ADNI database.

Finally, the peripheral blood gene expression profile, peripheral blood A $\beta$  biomarkers, and CSF markers contained 36 samples from 20 patients with LMCI and 16 control. The gene expression profile contained 49386 RNAs. The plasma A $\beta$  contained A $\beta$ 40 and A $\beta$ 42 free in plasma (FP40, FP42), A $\beta$ 40, and A $\beta$ 42 total in plasma (TP40, TP42). We analyzed two ratios, free plasma A $\beta$ 42 to free A $\beta$ 40 (FP40: FP42) and total plasma A $\beta$ 42 to total A $\beta$ 40 (TP40: TP42), as they had been previously shown to correlate with amyloid positivity (Perez-Grijalba et al., 2013, 2019; Fandos et al., 2017; de Rojas et al., 2018). The CSF biomarkers contained A $\beta$ , tau protein (TAU), and phosphorylated tau protein (PTAU).

## Immune Cell Abundance Calculation

In this study, GSEA was applied to calculate the relative abundance of immune cells in peripheral blood. As described in Subramanian et al. (2005) GSEA considers experiments with genomewide expression profiles from samples belonging to two classes, labeled 1 or 2. Genes are ranked based on the correlation between their expression and the class distinction by using any suitable metric. Given an *a priori* defined set of genes *S* (e.g., genes encoding products in a metabolic pathway, located in the same cytogenetic band, or sharing the same GO category), the goal of GSEA is to determine whether the members of *S* are randomly distributed throughout *L* or primarily found at the top or bottom. We calculate an enrichment score (ES) that reflects the degree to which a set *S* is overrepresented at the extremes (top or bottom) of the entire ranked list *L*. The score is calculated by walking down the list *L*, increasing a running-sum statistic when we encounter a gene in *S* and decreasing it when we encounter genes, not in *S*. The magnitude of the increment depends on the correlation of the gene with the phenotype. The ES is the maximum deviation from zero encountered in the random walk. Tag gene sets of 28 immune cells were obtained from Charoentong et al. (2017), as *a priori* defined set of genes *S*. The code for calculating the ES of the custom prior defined set of genes *S* was obtained from Bolen et al. (2011) (SPEC). We have improved SPEC by replacing the

tag gene sets of eight immune cells in the SPEC calculation code with the tag gene sets of 28 immune cells. Download normal and patient peripheral blood gene expression data from the ADNI database, and then GSEA was used to calculate the ESs of 28 immune cell signature genes in the normal and patient sample gene expression data. The ES is the relative abundance of the immune cells in the sample.

## Canonical Correlation Analysis

Canonical correlation analysis is a suitable technique that can establish interrelation between two sets of variables as well as quantify the percentage of variance common to the two sets (Ventura et al., 2011; Kim et al., 2017). Canonical correlation analysis indicates a correlation between two linear combinations of sets of dependent and independent variables as linear combinations of variables useful for predictive or comparative purposes (Akbas and Takma, 2005; Cankaya and Kayaalp, 2007; Sahin et al., 2011). Therefore, the goal of CCA is to find the best linear combination between two multivariate datasets that can maximize the correlation coefficient between them (Malacarne, 2014). Linear combinations of original variables can be defined by canonical variates (*U<sub>i</sub>* and *V<sub>i</sub>*) as follows:

$$U_i = a_{i1}X_1 + a_{i2}X_2 + \dots + a_{ip}X_p \quad (1)$$

$$V_i = b_{i1}Y_1 + b_{i2}Y_2 + \dots + b_{iq}Y_q \quad (2)$$

The correlation between *U<sub>i</sub>* and *V<sub>i</sub>* can be defined as canonical correlation. Canonical correlation analysis is repeatedly looking coefficients *a* and *b* to maximize the correlation between *U<sub>i</sub>* and *V<sub>i</sub>*. The maximum number of canonical functions that can be extracted equals to the number of variables in the smallest canonical variate (Dattalo, 2014) which is 3 in this study. The first canonical function is derived to maximize the correlation between *U<sub>i</sub>* and *V<sub>i</sub>* variables (Laessig and Duckett, 1979).

Standardized canonical coefficients and loadings were used to evaluate the relative importance of variables in the model (Dattalo, 2014). Standardized coefficients are interpreted similarly to standardized regression coefficients in multiple regressions. Therefore, CCA is used to estimate canonical coefficients (*a<sub>i1</sub>*, *a<sub>i2</sub>*, ..., *a<sub>ip</sub>* and *b<sub>i1</sub>*, *b<sub>i2</sub>*, ..., *b<sub>iq</sub>*) when the canonical correlation is at the maximum (Akbas and Takma, 2005). Canonical loading reflects the variance that the observed variable can be shared with canonical variate and interpreted like a factor loading in assessing the relative contribution of each variable to each canonical function (Safari et al., 2013). The result of canonical loading shows the contribution degree of the variable to the variable set. Redundancy index (RI) is proposed to calculate each canonical correlation to determine how much of the variance in one set of variables is accounted by the other set of variables (Sharma, 1996; Safari et al., 2013; Kim et al., 2017). The CCA can use a small number of features to analyze the correlation between the two sets of variables (CSF and peripheral biomarkers). The software SPSS 25.0 for Windows was used for statistical analysis of the data.

<sup>2</sup>www.affymetrix.com

## RESULTS

### Immune Cells Abundance Measurement

The computational method GSEA was applied to estimate the abundance of 28 kinds of peripheral blood immune cells including 15 kinds of adaptive immune cells and 13 kinds of innate immune cells: activated B cell, immature B cell, memory B cell, activated CD4 T cell, activated CD8 T cell, central memory CD4 T cell, central memory CD8 T cell, effector memory CD4 T cell, effector memory CD8 T cell, gamma delta T cell, regulatory T cell, T follicular helper cell, type 1 T helper cell, type 17 T helper cell, type 2 T helper cell, plasmacytoid, activated dendritic cell, immature dendritic cell, natural killer cell, natural killer T cell, CD56 dim natural killer cell, CD56 bright natural killer cell, eosinophil, macrophage, mast cell, MDSC, natural, and neutrophil. These 28 kinds of immune cells include most types of lymphocytes which produce important cytokines. The abundance of the 28 peripheral blood immune cells for LMCI and normal samples inferred by GSEA were shown in **Figure 1**.

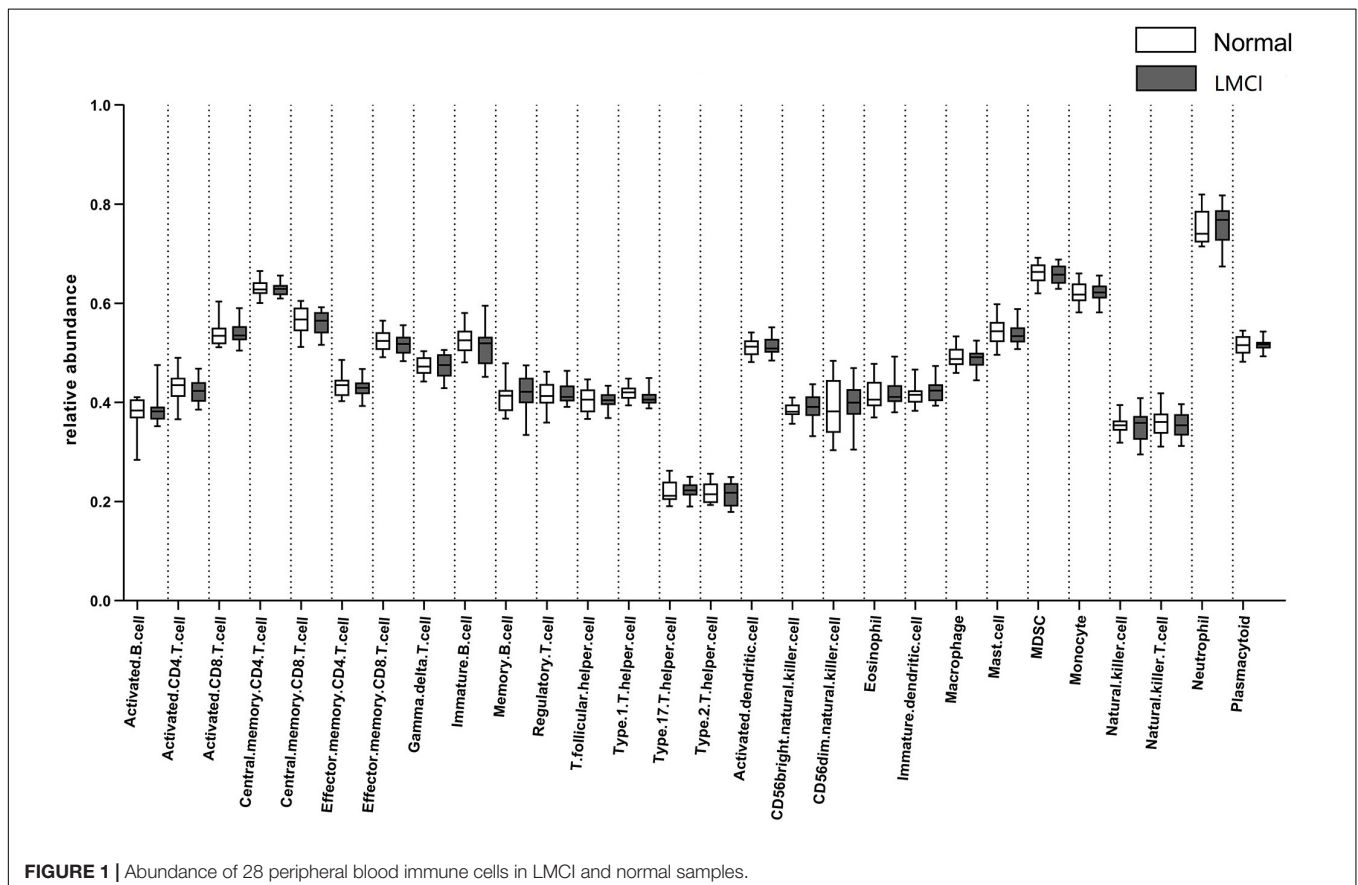
From **Figure 1** we can see, compared with normal samples, the activated CD4 T cell, activated B cell, central memory CD8 T cell, effector memory CD4 T cell, effector memory CD8 T cell, immature B cell, regulatory T cell, T follicular helper cell, type 1 T helper cell, activated dendritic cell, mast cell, MDSC, natural killer T cell appear to be in lower abundance in LMCI samples. However, gamma delta T cell, memory B cell, type 17

T helper cell, type 2 T helper cell, CD56bright natural killer cell, CD56dim natural killer cell, eosinophil, immature dendritic cell, macrophage, monocyte, natural killer cell, neutrophil were increased in LMCI samples.

### Results of Correlation Analysis

At first, to explore the relationship between LMCI and AD, we select proper biomarkers by comparing the changes of CSF biomarkers in LMCI with AD relative to their normal samples respectively (see **Supplementary Figure 1**). From the **Supplementary Figure 1A** to **Supplementary Figure 1C**, we can see that compared with the normal samples, the levels of A $\beta$  are lower, and the levels of TAU and PTAU are higher in the CSF of LMCI. From the **Supplementary Figure 1D** to **Supplementary Figure 1F**, the changes in the levels of A $\beta$ , TAU, and PTAU in AD are in the same situation. These results are consistent with the literature in Magaki et al. (2007), McKhann et al. (2011), Dubois et al. (2014) which found that A $\beta$  has a lower level, and TAU and PTAU have a higher level in AD. We can suggest that the LMCI sample has a very high risk of developing AD. Therefore, we selected A $\beta$ , TAU, and PTAU data at baseline as the significant biomarkers of the CSF of LMCI.

Not all peripheral biomarkers are related to CSF biomarkers. To evaluate the correlation between CSF and peripheral biomarkers, the Pearson correlation coefficient (Pearson, 1920) was calculated between CSF and peripheral biomarkers,





including three variables in the CSF biomarkers and 34 variables in the peripheral biomarkers. The Pearson correlation coefficients of CSF and peripheral marker variables are shown in **Table 1**.

The level of statistical significance was set at  $\alpha = 0.05$ . From the **Table 1** we can see 8 peripheral biomarkers were significantly associated with CSF biomarkers under significant conditions. The Pearson correlation coefficient between T follicular helper cell and A $\beta$  was  $-0.408$  ( $\alpha = 0.014 < 0.05$ ); the Pearson correlation coefficient between gamma delta T cell and PTAU was  $-0.332$  ( $\alpha = 0.048 < 0.05$ ); the Pearson correlation coefficient between immature B cell and TAU was  $-0.350$  ( $\alpha = 0.036 < 0.05$ ); the Pearson correlation coefficient between regulatory T cell and TAU was  $-0.373$  ( $\alpha = 0.025 < 0.05$ ); the Pearson correlation coefficient between regulatory T cell and PTAU was  $-0.372$

( $\alpha = 0.025 < 0.05$ ); the Pearson correlation coefficient between type 1 T helper cell and PTAU was  $-0.401$  ( $\alpha = 0.015 < 0.05$ ); the Pearson correlation coefficient between TP40 and TAU was  $-0.376$  ( $\alpha = 0.024 < 0.05$ ); Pearson correlation coefficient of FP40/FP42 and PTAU was  $0.346$  ( $\alpha = 0.038 < 0.05$ ). The levels of 8 peripheral biomarkers in LMCI and normal samples are shown in **Figure 2**.

Compared with the normal samples, we can see from **Figure 2** that the abundances of gamma delta T cell increased slightly; the abundance of immature B cell, regulatory T cell, T follicular cell, and TP40 decreased slightly; the abundance of CD56 bright natural killer cell and FP40/FP42 increased significantly; type 1 T helper cell abundance reduced significantly.

Bivariate correlations among variables of CSF biomarkers and variables of peripheral biomarkers are shown in **Table 1**. Results showed correlations between variable sets of gamma delta T cell, immature B cell, regulatory T cell, T follicular cell, type 1 T helper cell, CD56bright natural killer cell, TP40, or FP40/FP42(8 of the peripheral biomarkers). As we know that MCI was a complicated brain disease and the occurrence of MCI was accompanied by changes in the levels of multiple CSF biomarkers. One CSF biomarker cannot accurately determine the occurrence of disease. Since the Pearson correlation only could explain the correlation between two variables, it is impossible to simultaneously discover the relationship between the variable set of CSF and peripheral biomarkers. Base on that, in this study, CCA was introduced to explain the relationship between CSF biomarkers and peripheral biomarkers. **Table 2** shows direct results for the correlation between the two variable sets, which presented the canonical correlation coefficients and the significance of the research.

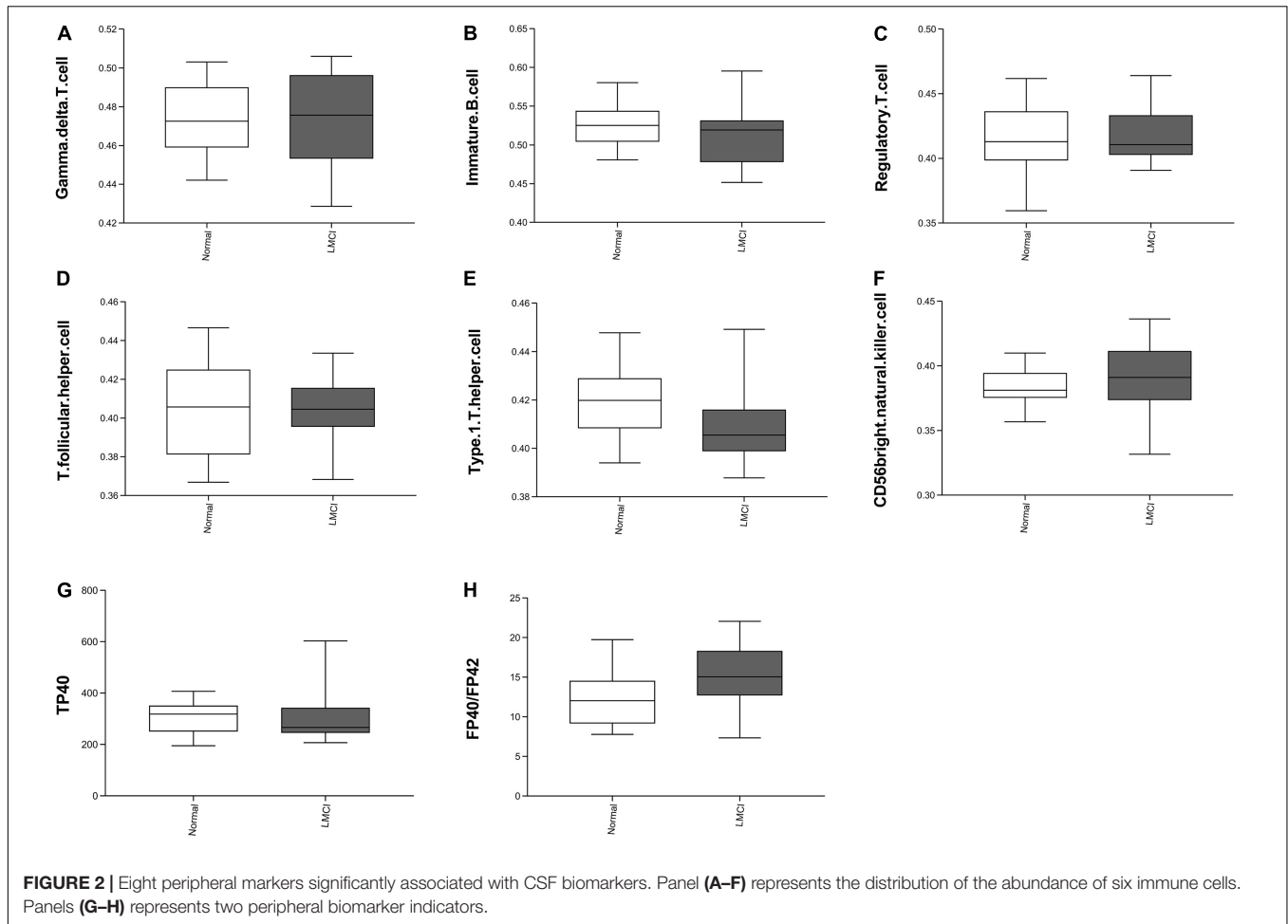
**Table 2** showed that two of the 3 confirmed canonical correlations were statistically significant. The first canonical correlation was 0.794. It represented the highest possible correlation between any linear combinations for three CSF biomarkers (U1) and eight peripheral biomarkers (V1) ( $p < 0.01$ ). The second canonical correlation was 0.709. It indicated that the relationship between canonical variates was significant ( $p = 0.008 < 0.01$ ). However, the canonical correlation of correlation 3 was not statistically significant.

Standardized canonical coefficients for pairs of all canonical variables [ $U_i$  and  $V_i$  in formula (1) and (2)] were shown in **Table 3**. Canonical variates representing optimal linear combinations of dependent and independent variables were defined by C1. Standardized canonical coefficients of A $\beta$  (0.435 in C1 and  $-0.493$  in C2), TAU (0.867 in C1 and  $-0.886$  in C2), and PTAU (0.339 in C1 and 1.062 in C2) as variables of CSF biomarkers suggested that they contributed importantly to the first and second canonical variate ( $U_1, U_2$ ). On the other hand, the first standardized canonical coefficients of immature B cell, regulatory T cell, CD56 bright natural killer cell, gamma delta T cell, TP40 as variables of peripheral biomarkers were  $-0.392, -0.585, 0.514, -0.423,$  and  $-0.456,$  respectively, indicating that they contributed importantly to the first canonical variate ( $V_1$ ). The second standardized canonical coefficients of FP40/FP42(0.61), T follicular helper cell (0.641), type 1 T helper

**TABLE 1** | Pearson correlation between CSF and peripheral biomarkers.

Immune Cells	CSF data		
	ABETA	TAU	PTAU
Activated B cell	-0.132	-0.111	0.171
Activated CD4 T cell	0.078	0.048	0.062
Activated CD8 T cell	-0.007	-0.135	-0.083
Central memory CD4 T cell	0.290	-0.187	-0.216
Central memory CD8 T cell	-0.003	0.135	0.121
Effector memory CD4 T cell	0.165	-0.181	-0.060
Effector memory CD8 T cell	0.018	-0.221	-0.180
Gamma delta T cell	0.135	-0.042	<b>-0.332*</b>
Immature B cell	0.162	<b>-0.350*</b>	-0.169
Memory B cell	-0.069	0.202	-0.023
Regulatory T cell	-0.051	-0.373*	<b>-0.372*</b>
T follicular helper cell	<b>-0.408*</b>	0.123	0.117
Type 1 T helper cell	0.290	-0.327	<b>-0.401*</b>
Type 17 T helper cell	-0.013	-0.276	-0.107
Type 2 T helper cell	-0.247	-0.011	-0.040
Activated dendritic cell	0.036	0.004	0.133
CD56bright natural killer cell	0.080	<b>0.413*</b>	0.163
CD56dim natural killer cell	-0.230	-0.074	-0.289
Eosinophil	0.231	0.031	-0.178
Immature dendritic cell	-0.046	-0.060	0.209
Macrophage	-0.063	-0.086	0.003
Mast cell	-0.128	-0.105	0.013
MDSC	0.212	-0.078	-0.200
Monocyte	-0.147	-0.116	-0.036
Natural killer cell	0.185	0.166	0.003
Natural killer T cell	-0.151	-0.187	-0.200
Neutrophil	0.116	0.099	0.014
Plasmacytoid	-0.029	0.118	0.182
FP40	-0.175	-0.238	-0.207
TP40	0.000	<b>-0.376*</b>	-0.266
FP42	-0.033	-0.243	-0.328
TP42	0.115	-0.222	-0.282
FP40/FP42	0.074	0.093	<b>0.346*</b>
TP40/TP42	0.093	-0.133	0.031

\*At the 0.05 level, the correlation is significant. The meaning of bolded values are significant results obtained in the article.



**FIGURE 2 |** Eight peripheral markers significantly associated with CSF biomarkers. Panel (A–F) represents the distribution of the abundance of six immune cells. Panels (G–H) represents two peripheral biomarker indicators.

cell (−0.390), gamma delta T cell (−0.580), and TP40 (0.340) contributed importantly to the second canonical variate (V2).

To find out the key factors in each group of variables, we presented the loadings for the canonical function in **Table 4**. Canonical loading presents a product-moment correlation between the original variable and its corresponding canonical variate. These values reflect the degree of a variable to be represented by a canonical variate. Canonical loadings for variables of CSF biomarkers suggested that TAU (0.905) and PTAU (0.723) had more effect than Aβ (0.07) to form the first fair for variables of CSF biomarkers (U1). In the second fair for variables of CSF biomarkers (U2), Aβ (−0.561) and PTAU (0.691) were more important factors. On the other hand, canonical loadings for regulatory T cell (−0.595), CD56bright natural

killer cell (0.565), TP40 (−0.525), type 1 T helper cell (−0.370), immature B cell (−0.366), and FP40/FP42 (0.290) had stronger effects compared to other factors to form the first fair for variables of peripheral biomarkers (V1). And the FP40/FP42 (0.351), T follicular helper cell (0.305), type 1 T helper cell, CD56bright

**TABLE 2 |** Summary of results from canonical correlation analysis.

	Correlation	Eigenvalue	Wilks Statistic	F	Sig.
1	<b>0.794</b>	1.702	0.131	3.094	0.000
2	<b>0.709</b>	1.013	0.354	2.530	0.008
3	0.536	0.404	0.712	1.818	0.133

The meaning of bolded values are significant results obtained in the article.

**TABLE 3 |** Standardized canonical correlation coefficients.

Variable	C1	C2	C3
Aβ	<b>0.435</b>	<b>−0.493</b>	0.914
TAU	<b>0.867</b>	<b>−0.886</b>	−0.530
PTAU	<b>0.339</b>	<b>1.062</b>	0.752
FP40/FP42	−0.019	<b>0.613</b>	0.657
Immature B cell	<b>−0.392</b>	0.171	0.240
Regulatory T cell	<b>−0.585</b>	−0.040	−0.247
T follicular helper cell	0.123	<b>0.641</b>	−0.573
Type 1 T helper cell	0.158	<b>−0.390</b>	0.422
CD56bright natural killer cell	<b>0.514</b>	−0.259	0.071
Gamma delta T cell	<b>−0.423</b>	<b>−0.580</b>	−0.146
TP40	<b>−0.456</b>	<b>0.340</b>	0.328

C1, C2, C3: 1st, 2nd, 3rd standardized canonical coefficients. The meaning of bolded values are significant results obtained in the article.

**TABLE 4** | Canonical loadings of original variables with their canonical variables.

	CSF group set			Peripheral group set							
	ABETA	TAU	PTAU	FP40/FP42	Immature B cell	Regulatory T cell	T follicular helper cell	Type 1 T helper cell	CD56bright natural killer cell	Gamma delta T cell	TP40
<b>U1</b>	-0.07	<b>0.905</b>	<b>0.723</b>	<b>0.29</b>	<b>-0.366</b>	<b>-0.595</b>	-0.039	<b>-0.37</b>	<b>0.565</b>	-0.114	<b>-0.525</b>
<b>U2</b>	<b>-0.561</b>	0.012	<b>0.691</b>	<b>0.351</b>	0.073	-0.056	<b>0.305</b>	<b>-0.393</b>	<b>-0.326</b>	<b>-0.538</b>	0.072
<b>U3</b>	0.825	-0.425	0.029	0.52	0.385	-0.241	-0.653	0.254	-0.042	-0.194	0

*U1, U2, and U3, 1st, 2nd, and 3rd CSF variables; V1, V2, and V3, 1st, 2nd, and 3rd peripheral variables; CSF group set: CSF biomarkers; Peripheral group set: Peripheral biomarkers. The meaning of bolded values are significant results obtained in the article.*

natural killer cell (-0.326), and gamma delta T (-0.538) cell were more important factors to form the second pair for variables of peripheral biomarkers (V2).

Numbers of dimensions explaining the relationships between variable sets were reduced from 11 to 2 by CCA. These results indicated a high correlation between the set of CSF biomarkers (U1) and the set of peripheral biomarkers (V1) (at 0.794). Cross loadings represent correlations between original variables and opposite canonical variables. According to cross-loading results, the first pair of canonical variables TAU (0.718) and PTAU (0.574) provided a relatively strong contribution to canonical variate V1 whereas regulatory T cell (-0.472), CD56bright natural killer cell (0.448), TP40 (-0.417), type 1 T helper cell (-0.294), immature B cell (-0.291), and FP40/FP42 (0.230) highly contributed to U1 (Table 5). The second pair of canonical variables Aβ (-0.398) and PTAU (0.49) provided a greater contribution to canonical variate V2 whereas FP40/FP42 (0.249), T follicular helper cell (0.216), type 1 T helper cell (-0.279), CD56bright natural killer cell (-0.231), and gamma delta T cell made an important contribution to U2 (Table 5).

### ROC Analysis

To verify the practical significance of the correlation between CSF and peripheral biomarkers, we conducted the ROC analysis. Through CCA analysis, variables with a significant correlation between CSF and peripheral biomarkers were found. Receiver operating characteristic curve analysis by using the peripheral biomarkers (type 1 T helper cell, CD56bright natural killer cell, and FP40F/P42 highly contributed to U1 and U2) which highly contributed to V1 and V2 was performed to find peripheral biological indicators with clinical significance for disease classification. Then multiple peripheral biomarkers were used for logistic regression, and ROC analysis of multi-index combined classification was performed. We used SPSS software to draw the ROC curve. The results of the ROC analysis were shown in Figure 3.

Through the ROC analysis, we found that FP40/FP42 (area under the ROC curve, 0.709) and type 1 T helper cell (area under the ROC curve, 0.703) had clinical significance for the diagnosis of the disease (Figures 3A,B). Logistic regression was performed by using two indicators (set the coefficient significance to  $p < 0.1$ ). The coefficients of FP40F/FP42 ( $p = 0.046 < 0.1$ ) and type 1 T helper cell ( $p = 0.093 < 0.1$ ) satisfy the significance. The ROC analysis of the combined diagnosis of the two indicators showed that the combined diagnosis further improved accuracy (the area under the ROC curve, 0.747) (Figure 3C).

### Reconstruction of the PPI Network

The immune-related PPI network was constructed to explore the biological significance of the correlation between CSF and peripheral biomarkers. For the differential analysis of the microarray data used to calculate the abundance of immune cells, differential RNA extraction was performed by using the Limma algorithm. We chose  $p$ -value  $< 0.05$  as the threshold for screening differentially expressed genes. This process is implemented using the R package 'Limma' (Ritchie et al., 2015). A total of 1324 differentially expressed genes (DEGs) were identified, of which

TABLE 5 | Cross loadings of original variables with opposite canonical variables.

	CSF group set			Peripheral group set							
	ABETA	TAU	PTAU	FP40/FP42	Immature B cell	Regulatory T cell	T follicular helper cell	Type 1 T helper cell	CD56bright natural killer cell	Gamma delta T cell	TP40
<b>U1</b>	-0.055	<b>0.718</b>	<b>0.574</b>	<b>V1</b>	<b>-0.291</b>	<b>-0.472</b>	-0.031	<b>-0.294</b>	<b>0.448</b>	-0.09	<b>-0.417</b>
<b>U2</b>	<b>-0.398</b>	0.008	<b>0.49</b>	<b>V2</b>	0.052	-0.04	<b>0.216</b>	<b>-0.279</b>	<b>-0.231</b>	<b>-0.382</b>	0.051
<b>U3</b>	0.442	-0.228	0.015	<b>V3</b>	0.207	-0.129	-0.35	0.136	-0.023	-0.104	0

**U1, U2 and U3, 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> CSF variables; V1, V2, and V3, 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> peripheral variables; CSF group set: CSF biomarkers; Peripheral group set: Peripheral biomarkers. The meaning of bolded values are significant results obtained in the article.**

801 were up-regulated and 523 were down-regulated in LMCI. For all the differentially expressed genes, we constructed the PPI network by using the STRING database<sup>3</sup>. The final DEGs PPI network was constructed with 1050 gene nodes and 4547 edges through the STRING v10 data database.

Protein-protein interaction network related to the abundances of immune cells was constructed by retaining genes related to immunity and node genes directly linked to them. In the correlation analysis, six immune cells were significantly correlated with CSF biomarkers. 1324 differentially expressed genes were matched with the tag genes of 6 immune cells, and 26 immune-related differentially expressed genes were obtained. We used 26 immune-related differentially expressed genes to search in the constructed PPI network and found that 1 gene does not exist in the PPI network and 1 gene is an isolated node. Finally, 2 genes were deleted and 24 immune-related differentially expressed genes were retained. In the PPI network, 24 immune-related differentially expressed genes and gene nodes directly connected to them were selected, the immune-related PPI network was constructed by 223 gene nodes and 1123 edges finally. The visualization of the immune-related PPI network was built using Cytoscape v3.6.0 software (Shannon et al., 2003) (Figure 4).

Then the 223 significant genes were analyzed by DAVID for GO and KEGG, and we investigated the role of these genes in biological functions and processes (see Figure 5). The GO and KEGG analysis used the online database DAVID v6.8 (Sherman and Lempicki, 2009), and visual display through R software (R Core Team, 2013).

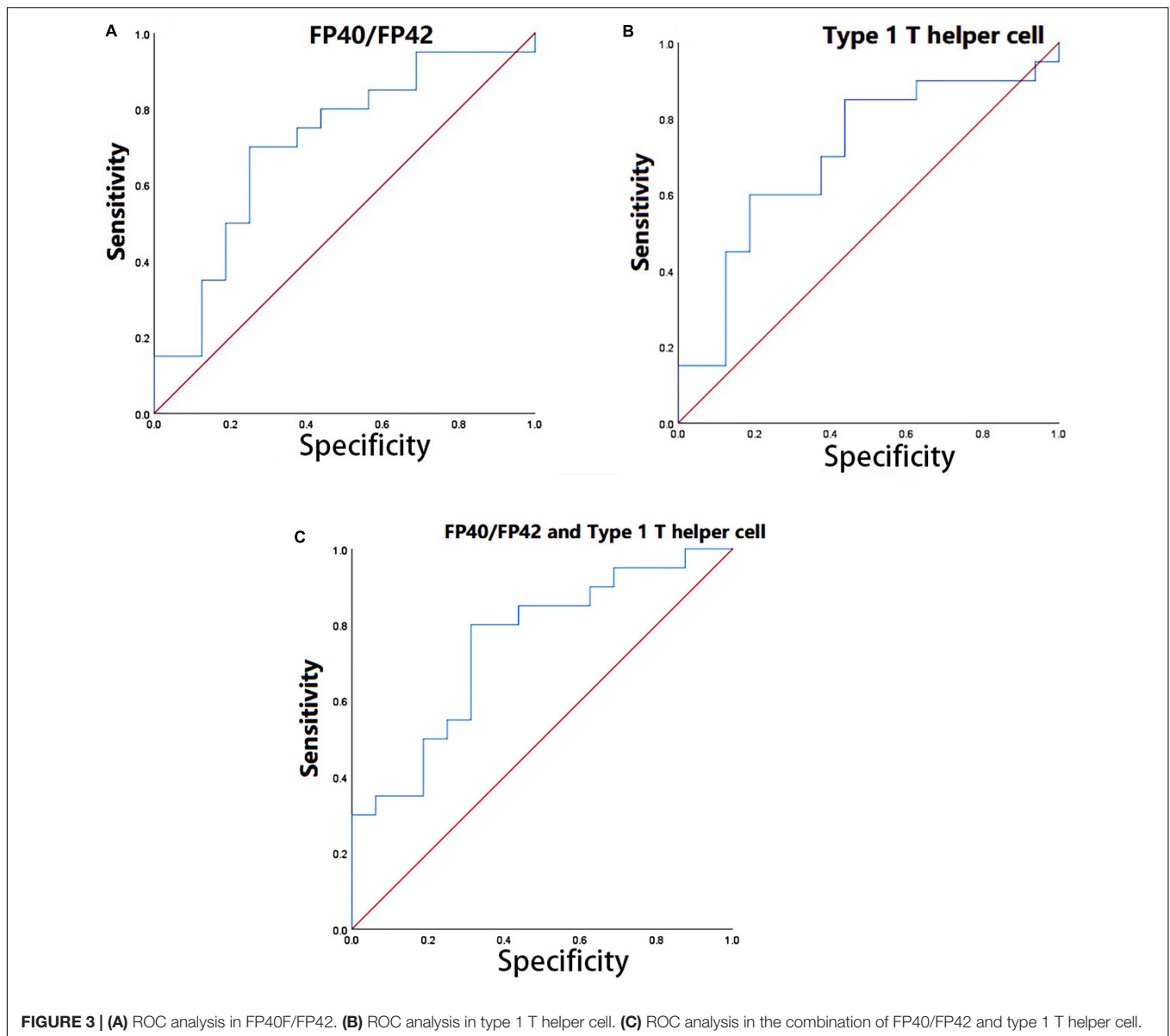
The results of the enrichment analysis showed that these genes were related to the inflammatory response and had a certain relationship with the expression of some cytokines. In the reconstructed PPI network, most genes were up-regulated (a total of 223 genes, 157 genes were significantly up-regulated and 68 genes were significantly down-regulated). Among the 24 immune-related genes, 19 genes were significantly up-regulated and 5 genes were down-regulated. It shows that LMCI is closely related to the systemic inflammatory response. LMCI may show similar physiological processes as AD, such as the recruitment of peripheral immune cells to the brain, and the permeability of blood-brain barrier (BBB). It may be the biological reason for the significant correlation between the LMCI peripheral data and the CSF data.

## DISCUSSION

MCI is a chronic degenerative disease of the nervous system, which refers to a state of cognitive impairment between normal aging and dementia. It appears that memory loss is not commensurate with age, but it has not yet reached the standard of AD. However, patients diagnosed with MCI are at high risk of developing AD. The AD conversion rate is 6–25% per year. In the AD research, many clinical cases have shown that the treatment effect is not good in the middle and advanced stages,

<sup>3</sup><https://string-db.org/>

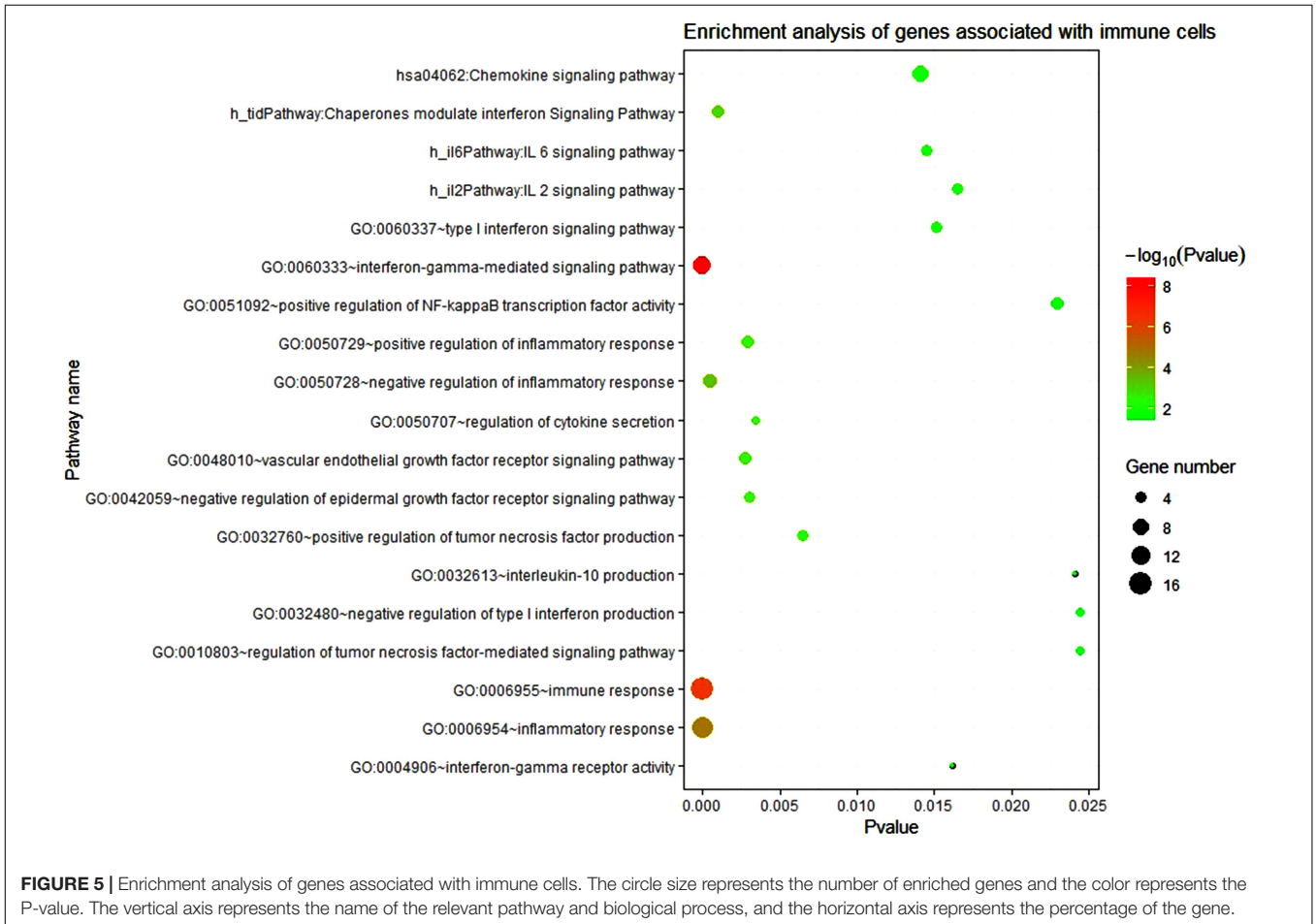
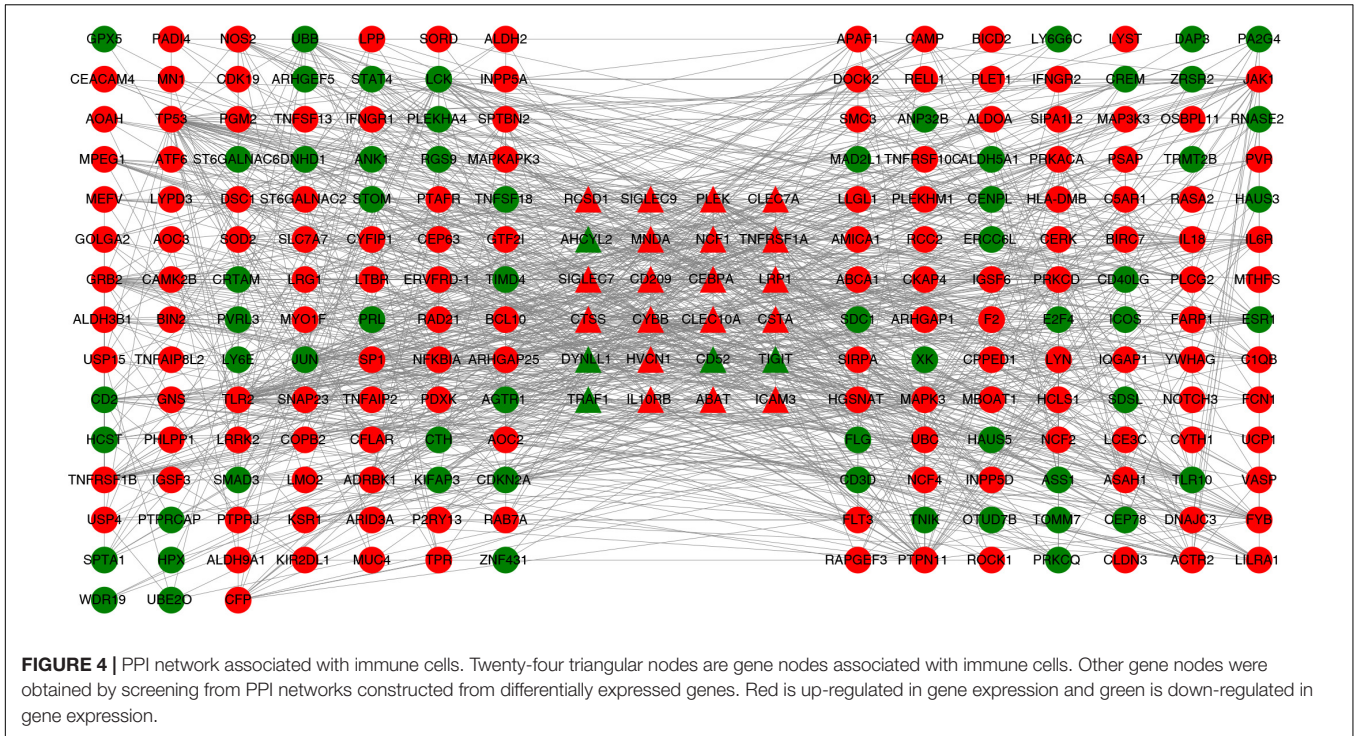




so researchers have turned their attention to its early diagnosis and preventive intervention in recent years. There is evidence that peripheral immune cells in AD recognize A $\beta$  and treat it, present it to T cells, and trigger adaptive immunity (Jóźwik et al., 2012; Begum et al., 2014). This indicates that peripheral biomarkers may be potentially associated with AD. The CSF biomarker changes in LMCI patients are very similar to those in AD, so the study of LMCI patients is of great significance for early detection and early intervention of AD. CCA was introduced to explain the relationship between CSF biomarkers and peripheral biomarkers. The results indicated a high correlation between variable sets of CSF biomarkers and variable sets of peripheral biomarkers. It may be due to A $\beta$  transported to the periphery and activated adaptive immunity. Receiver operating characteristic curve analysis found that the diagnostic accuracy of two of

the peripheral biomarkers (FP40/FP42 and type 1 T helper cell) for the disease was 0.747. It was found that the systemic immune response plays an important role in the correlation between peripheral biomarkers and CSF biomarkers through KEGG and GO analysis.

About 223 differentially expressed genes related to immunity were used for KEGG and GO analysis. It was found that these genes are closely related to the IL-6 signaling pathway, TNF signaling pathway, IFN- $\gamma$  signaling pathway,  $\gamma$ -interferon-mediated signaling pathways, tumor necrosis factor-mediated signaling pathway, chemokine signal pathway, inflammatory response function, immune response function, and so on. From the biological analysis we can see that these pathways and functions in KEGG and GO results are related to immunity, inflammation, and cytokines. We also found other biological



processes related to the immune response, such as positive and negative regulation of inflammatory response function, type I interferon signaling pathway, tumor necrosis factor-activated receptor activity function, positive regulation of tumor necrosis factor production function, regulation of  $\gamma$ -interferon-mediated signaling pathway. Most genes enriched in these pathways and functions are up-regulated.

The further molecular biological analysis shows that the expression of A $\beta$  in CSF and peripheral is closely related. It has been found in AD studies that A $\beta$  enters peripheral blood to trigger an inflammatory response. PBMCs from AD patients are induced to release pro-inflammatory cytokines such as IL-6, TNF, and IFN- $\gamma$  (Pellicano et al., 2010). It further promoted the production of amyloid precursor protein and enzymes that cleave it, leading to increased A $\beta$  production (Sutinen et al., 2012). A $\beta$  can also stimulate the pro-inflammatory NF- $\kappa$ B dependent signaling pathway (Kumar et al., 2014). Through regulated transport, peripheral inflammatory markers can cross the BBB and perform neuromodulation (Banks, 2005). INF- $\gamma$  stimulated the release of CXCL-10, and INF also increased. In AD, CXCL-10 has been found to bind to the chemokine receptor CXC chemokine receptor 3, which is involved in T cell initiation and maintenance of natural killer cells in the body, thereby inducing extracellular signal-regulated kinases Pathways eventually lead to neuronal dysfunction and apoptosis (Nelson and Gruol, 2004; Sui et al., 2006; Cho et al., 2009). Compared with normal samples, the concentrations of peripheral blood cytokines IL-2, IL-6, and epidermal growth factor (EGF) in AD patients were significantly increased (Lai et al., 2017). Physiological changes including inflammatory responses and changes in cytokines were found in AD. Kyoto Encyclopedia of Genes and Genomes and GO analysis of differentially expressed genes in this study also found similar physiological changes in pathways and functions in LMCI patients. Kyoto Encyclopedia of Genes and Genomes and GO analysis of differentially expressed genes in peripheral blood of LMCI patients revealed that most of the genes enriched in cytokine-related signaling pathways were up-regulated. A $\beta$  clearance in the brain mainly depends on peripheral clearance to achieve local lymph node degradation. In LMCI, A $\beta$  in the brain is transported to the peripheral blood and releases pro-inflammatory cytokines, activating immune-related pathways and functions. This indicates that patients with LMCI have shown a systemic inflammatory response similar to AD patients. These results show that peripheral biomarkers can reflect the pathological changes of LMCI in the brain. Peripheral biomarkers may develop new diagnostic methods for LMCI. The research on LMCI will also help to predict AD at an early stage, and also provide an important basis for AD immunotherapy.

## REFERENCES

- Akbas, Y., and Takma, C. H. (2005). Canonical correlation analysis for studying the relationship between egg production traits and body weight, egg weight and age at sexual maturity in layers. *Czech J. Anim. Sci.* 50, 163–168. doi: 10.17221/4010-CJAS
- Banks, W. A. (2005). Blood-brain barrier transport of cytokines: a mechanism for neuropathology. *Curr. Pharm. Des* 11, 973–984. doi: 10.2174/1381612053381684

## DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at ADNI datasets: <https://ida.loni.usc.edu/pages/access/studyData.jsp>; <https://ida.loni.usc.edu/pages/access/geneticData.jsp>.

## ETHICS STATEMENT

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## AUTHOR CONTRIBUTIONS

WK and TZ conceived the study, participated in its design and coordination, and performed all the molecular biological analyses of the data. TZ carried out the GSEA studies on the LMCI's gene expression data and drafted the manuscript. SW and TZ performed the pre-processing, GSEA, CCA algorithms, and some statistical analysis of LMCI expression data and plasma marker data. WK helped with data interpretation and manuscript drafting. XM participated in the final data analysis and interpretation. All authors read and approved the final manuscript.

## FUNDING

Natural Science Foundation of Shanghai (No. 18ZR1417200) and National Natural Science Foundation of China (No. 61803257).

## ACKNOWLEDGMENTS

The authors thank the Alzheimer's Disease Neuroimaging Initiative (ADNI) for contributing data.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2020.00834/full#supplementary-material>

**FIGURE S1** | Change of CSF biomarkers in LMCI (A–C) and AD (D–F).

- Begum, A. N., Cunha, C., Sidhu, H., Alkam, T., Scolnick, J., Rosario, E. R., et al. (2014). Women with the Alzheimer's risk marker ApoE4 lose A $\beta$ -specific CD4+ T cells 10–20 years before men. *Trans. Psychiatry* 4:e414. doi: 10.1038/tp.2014.51
- Bolen, C. R., Uduman, M., and Kleinstein, S. H. (2011). Cell subset prediction for blood genomic studies. *BMC Bioinformatics* 12:258. doi: 10.1186/1471-2105-12-258
- Cankaya, S., and Kayaalp, G. T. (2007). Estimation of relationship between live weights and some body measurements in German farm x hair



- crossbred by canonical correlation analysis. *Hayvansal Üretim* 48, 27–32.
- Charoentong, P., Finotello, F., Angelova, M., Mayer, C., Efreanova, M., Rieder, D., et al. (2017). Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep.* 18, 248–262. doi: 10.1016/j.celrep.2016.12.019
- Cho, J., Nelson, T. E., Bajova, H., and Gruol, D. L. (2009). Chronic CXCL10 alters neuronal properties in rat hippocampal culture. *J. Neuroimmunol.* 207, 92–100. doi: 10.1016/j.jneuroim.2008.12.007
- Dattalo, P. V. (2014). *A Demonstration of Canonical Correlation Analysis with Orthogonal Rotation to Facilitate Interpretation*. Richmond: Virginia Commonwealth University, 3–5.
- de Rojas, I., Romero, J., Rodriguez-Gomez, O., Pesini, P., Sanabria, A., Perez-Cordon, A., et al. (2018). Correlations between plasma and PET beta-amyloid levels in individuals with subjective cognitive decline: the Fundacio ACE healthy brain initiative (FACEHBI). *Alzheimer's Res. Ther.* 10:119. doi: 10.1186/s13195-018-0444-1
- Dubois, B., Feldman, H. H., Jacova, C., Hampel, H., Molinuevo, J. L., Blennow, K., et al. (2014). Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *Lancet Neurol.* 13, 614–629. doi: 10.1016/S1474-4422(14)70090-0
- Fandos, N., Pérez-Grijalba, V., Pesini, P., Olmos, S., Bossa, M., Villemagne, V. L., et al. (2017). Plasma amyloid  $\beta$  42/40 ratios as biomarkers for amyloid  $\beta$  cerebral deposition in cognitively normal individuals. *Alzheimer's Dement.* 8, 179–187. doi: 10.1016/j.dadm.2017.07.004
- Hansson, O., Zetterberg, H., Buchhave, P., Londos, E., Blennow, K., and Minthon, L. (2006). Association between CSF biomarkers and incipient Alzheimer's disease in patients with mild cognitive impairment: a follow-up study. *Lancet Neurol.* 5, 228–234. doi: 10.1016/S1474-4422(06)70355-6
- Hotelling, H. (1992). "Relations between two sets of variates," in *Breakthroughs in Statistics*, eds S. Kotz, and N. L. Johnson (New York, NY: Springer), 162–190. doi: 10.2307/2333955
- Jessen, F., Wolfsgruber, S., Wiese, B., Bickel, H., Mösch, E., Kaduszkiewicz, H., et al. (2014). AD dementia risk in late MCI, in early MCI, and in subjective memory impairment. *Alzheimer's Dement.* 10, 76–83. doi: 10.1016/j.jalz.2012.09.017
- Jóźwik, A., Landowski, J., Bidzan, L., Fülöp, T., Bryl, E., and Witkowski, J. M. (2012). Beta-amyloid peptides enhance the proliferative response of activated CD4+ CD28+ lymphocytes from Alzheimer disease patients and from healthy elderly. *PLoS One* 7:e33276. doi: 10.1371/journal.pone.0033276
- Kabir, A., Merrill, R. D., Shamim, A. A., Klemm, R. D., Labrique, A. B., Christian, P., et al. (2014). Canonical correlation analysis of infant's size at birth and maternal factors: a study in rural Northwest Bangladesh. *PLoS One* 9:e94243. doi: 10.1371/journal.pone.0094243
- Kim, T. W., Kim, I. S., Kwon, S. G., Hwang, J. H., Park, D. H., Kang, D. G., et al. (2017). Identification of relationship between pork colour and physicochemical traits in American Berkshire by canonical correlation analyses. *Anim. Product. Sci.* 57, 1179–1185. doi: 10.1071/an14975
- Kumar, A., Pate, K. M., Moss, M. A., Dean, D. N., and Rangachari, V. (2014). Self-propagative replication of A $\beta$  oligomers suggests potential transmissibility in Alzheimer disease. *PLoS One* 9:e111492. doi: 10.1371/journal.pone.0111492
- Laessig, R. E., and Duckett, E. J. (1979). Canonical correlation analysis: potential for environmental health planning. *Am. J. Public Health* 69, 353–359. doi: 10.2105/AJPH.69.4.353
- Lai, K. S. P., Liu, C. S., Rau, A., Lanctôt, K. L., Köhler, C. A., Pakosh, M., et al. (2017). Peripheral inflammatory markers in Alzheimer's disease: a systematic review and meta-analysis of 175 studies. *J. Neurol. Neurosurg. Psychiatry* 88, 876–882. doi: 10.1136/jnnp-2017-316201
- Li, B., Severson, E., Pignon, J. C., Zhao, H., Li, T., Novak, J., et al. (2016). Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. *Genome Biol.* 17, 1–16. doi: 10.1186/s13059-016-1028-7
- Magaki, S., Mueller, C., Dickson, C., and Kirsch, W. (2007). Increased production of inflammatory cytokines in mild cognitive impairment. *Exp. Gerontol.* 42, 233–240. doi: 10.1016/j.exger.2006.09.015
- Malacarne, R. L. (2014). Canonical correlation analysis. *Math. J.* 16, 1–22.
- McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R. Jr., Kawas, C. H., et al. (2011). The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dement.* 7, 263–269. doi: 10.1016/j.jalz.2011.03.005
- Nelson, T. E., and Gruol, D. L. (2004). The chemokine CXCL10 modulates excitatory activity and intracellular calcium signaling in cultured hippocampal neurons. *J. Neuroimmunol.* 156, 74–87. doi: 10.1016/j.jneuroim.2004.07.009
- Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12, 453–457. doi: 10.1038/nmeth.3337
- Novershtern, N., Subramanian, A., Lawton, L. N., Mak, R. H., Haining, W. N., McConkey, M. E., et al. (2011). Densely interconnected transcriptional circuits control cell states in human hematopoiesis. *Cell* 144, 296–309. doi: 10.1016/j.cell.2011.01.004
- Park, J. E., Choi, K. Y., Kim, B. C., Choi, S. M., Song, M. K., Lee, J. J., et al. (2019). Cerebrospinal fluid biomarkers for the diagnosis of prodromal Alzheimer's disease in amnesic mild cognitive impairment. *Dement. Geriatr. Cogn. Disord. Extra* 9, 100–113. doi: 10.1159/000496920
- Pearson, K. (1920). Notes on the history of correlation. *Biometrika* 13, 25–45. doi: 10.2307/2331722
- Pellicano, M., Bulati, M., Buffa, S., Barbagallo, M., Di Prima, A., Misiano, G., et al. (2010). Systemic immune responses in Alzheimer's disease: in vitro mononuclear cell activation and cytokine production. *J. Alzheimer's Dis.* 21, 181–192. doi: 10.3233/jad-2010-091714
- Perez-Grijalba, V., Pesini, P., Monleon, I., Boada, M., Tárraga, L., Ruiz-Laza, A., et al. (2013). Several direct and calculated biomarkers from the amyloid- $\beta$  pool in blood are associated with an increased likelihood of suffering from mild cognitive impairment. *J. Alzheimer's Dis.* 36, 211–219. doi: 10.3233/JAD-121744
- Perez-Grijalba, V., Romero, J., Pesini, P., Sarasa, L., Monleon, I., San-Jose, I., et al. (2019). Plasma A $\beta$ 42/40 ratio detects early stages of Alzheimer's disease and correlates with CSF and neuroimaging biomarkers in the AB255 study. *J. Prevent. Alzheimer's Dis.* 6, 34–41. doi: 10.14283/jpad.2018.41
- Quesenberry, P. J., and Colvin, G. A. (2001). Hematopoietic stem cells, progenitor cells, and cytokines. *Williams Hematol.* 338, 153–174.
- Risacher, S. L., Fandos, N., Romero, J., Sherriff, I., Pesini, P., Saykin, A. J., et al. (2019). Plasma amyloid beta levels are associated with cerebral amyloid and tau deposition. *Alzheimer's Dement.* 11, 510–519. doi: 10.1016/j.dadm.2019.05.007
- Ritchie, M. E., Phipson, B., Wu, D. I., Hu, Y., Law, C. W., Shi, W., et al. (2015). limma powers differential expression analyses for RNA-seq and microarray studies. *Nucleic Acids Res.* 43:e47. doi: 10.1093/nar/gkv007
- Safari, P., Danyali, S. F., Honarnejad, R., and Esfahani, M. (2013). Study of relationship between oil quality traits with agro-morphological traits in peanut genotypes by canonical correlation analysis. *Int. J. Biosci.* 3, 1–10. doi: 10.12692/ijb/3.8.1-10
- Sahin, M., Cankaya, S., and Ceyhan, A. (2011). Canonical correlation analysis for estimation of relationships between some traits measured at weaning time and six-month age in merino lambs. *Bulgarian J. Agricult. Sci.* 17, 680–686.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303
- Sharma, S. (1996). *Applied Multivariate Techniques*. Canada: John Wiley and Sons Inc, 391–404.
- Sherman, B. T., and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4:44. doi: 10.1038/nprot.2008.211
- Shoemaker, J. E., Lopes, T. J., Ghosh, S., Matsuoka, Y., Kawaoka, Y., and Kitano, H. (2012). CTen: a web-based platform for identifying enriched cell types from heterogeneous microarray data. *BMC Genomics* 13:460. doi: 10.1186/1471-2164-13-460
- Stephan, B. C., Matthews, F. E., McKeith, I. G., Bond, J., Brayne, C., Medical Research Council Cognitive Function and Aging Study (2007). Early cognitive change in the general population: how do different definitions work? *J. Am. Geriatr. Soc.* 55, 1534–1540. doi: 10.1111/j.1532-5415.2007.01386.x
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* 102, 15545–15550. doi: 10.1073/pnas.0506580102



- Sui, Y., Stehno-Bittel, L., Li, S., Loganathan, R., Dhillon, N. K., Pinson, D., et al. (2006). CXCL10-induced cell death in neurons: role of calcium dysregulation. *Eur. J. Neurosci.* 23, 957–964. doi: 10.1111/j.1460-9568.2006.04631.x
- Sutinen, E. M., Pirttilä, T., Anderson, G., Salminen, A., and Ojala, J. O. (2012). Pro-inflammatory interleukin-18 increases Alzheimer's disease-associated amyloid- $\beta$  production in human neuron-like cells. *J. Neuroinflamm.* 9, 1–14. doi: 10.1186/1742-2094-9-199
- Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., et al. (2015). STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 43, D447–D452. doi: 10.1093/nar/gku1003
- R Core Team (2013). R: a language and environment for statistical computing. *Computing* 14:12.
- Ventura, H. T., Lopes, P. S., Peloso, J. V., Guimarães, S. E., Carneiro, A. P. S., and Carneiro, P. L. (2011). A canonical correlation analysis of the association between carcass and ham traits in pigs used to produce dry-cured ham. *Genet. Mol. Biol.* 34, 415–455. doi: 10.1590/S1415-47572011000300014
- Weiner, M. W., and Veitch, D. P. (2015). Introduction to special issue: overview of Alzheimer's Disease Neuroimaging Initiative. *Alzheimer's Dement.* 11, 730–733. doi: 10.1016/j.jalz.2015.05.007

**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 © 2020 Zhang, Kong, Wang and Mou. 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.