

Impaired type I interferon signaling activity implicated in the peripheral blood transcriptome of preclinical Alzheimer's disease

Liting Song

Fudan University

Jingqi Chen

Fudan University

Chun-Yi Lo

Fudan University

Qihao Guo

Shanghai Jiao Tong university

Jianfeng Feng

Fudan University

Xing-Ming Zhao (✉ xmzhao@fudan.edu.cn)

Fudan University

Article

Keywords: Subjective cognitive decline (SCD), Alzheimer's disease (AD), peripheral blood transcriptome, type I interferon signaling, progression biomarkers

Posted Date: September 23rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-895522/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Subjective cognitive decline (SCD) is considered the preclinical manifestation of Alzheimer's disease (AD), which is a potentially crucial window for preventing or delaying the progression of the disease. To explore the potential mechanism of disease progression and identify relevant biomarkers, we comprehensively assessed the peripheral blood transcriptomic alterations in SCD, covering lncRNA, mRNA, and miRNA. Dysregulated protein-coding mRNA at both gene and isoform levels implicated impairment in the type I interferon signaling pathway in SCD. Specifically, this pathway was regulated by the transcription factor *STAT1* and ncRNAs *NR1R* and *has-miR-146a-5*. The miRNA-mRNA-lncRNA coexpression network revealed hub genes for the interferon module. Individuals with lower interferon signaling activity and lower expression of hub genes *STAT1* and *TRIM22* exhibited a higher conversion rate to mild cognitive impairment (MCI). Our findings illustrated the down-regulation of interferon signaling activity would potentially increase the risk of disease progression and thus serve as a pre-disease biomarker.

Introduction

Alzheimer's disease (AD), the most common cause of dementia, is a progressive neurodegenerative disorder that causes deficits in memory, thinking, decision making, and social skills [1]. The pathophysiological events of AD begin years before the onset of clinical dementia [2]. The National Institute on Aging-Alzheimer's Association (NIA-AA) proposed three subsequent stages of AD, including the preclinical stage of AD, mild cognitive impairment (MCI), and AD dementia [3–5]. The primary hypothesis on the progression of AD pathologies stated that the initial deposition of amyloid- β ($A\beta$) leads to subsequent tau hyperphosphorylation, neuron and synaptic loss, and cognitive decline [6]. However, the continued failure of clinical trials aimed at decreasing $A\beta$ plaques implies that there may be alternative causes of disease progression or the neuronal injury of the late MCI and AD dementia may be irreversible.[7]. The long preclinical phase with mild neuronal damage provides a potentially crucial opportunity for exploring the potential pathogenic mechanisms of AD and preventing or delaying the progression of this disease [5, 8].

Subjective cognitive decline (SCD), characterized by a self-reported persistent decline in cognitive capacity compared to a previously normal status prior to the appearance of objective cognitive impairment, occurs at the preclinical stage of AD and may serve as a symptomatic indicator of preclinical AD [9]. The occurrence of biomarker abnormalities associated with AD in SCD, including decreased concentrations of $A\beta_{42}$ and increased concentrations of tau in CSF (cerebrospinal fluid) [10], provides supportive evidence for SCD as the earliest manifestation of AD. Moreover, SCD is associated with an increased risk for future cognitive decline [11–13]—longitudinal data analyses revealed that individuals with subjective cognitive concerns carried a four-fold and a six-fold increased risk of incident MCI and dementia, respectively, compared to those without memory complaints [13, 14]. The cumulative conversion rate of SCD progression to MCI and dementia is about 20.76% and 7.23% with high heterogeneity [15]. Therefore, exploring the early changes of subjective cognitive decline in preclinical Alzheimer's disease and predicting the conversion risk are urgently required.

The transcriptome represents a quantitative phenotype that reveals the biological processes disrupted in disorders. Despite the blood-brain barrier, blood expression data captured the majority of predictive molecular pathways for neurodegenerative progression as identified in brain expression data for AD [16]. Compared with

brain and CSF biomarkers, blood-based biomarkers are more obtainable and require less invasive collection methods. However, to our knowledge, no previous study has comprehensively evaluated the blood transcriptomic profiling for SCD in preclinical Alzheimer's disease.

Here, we performed RNA sequencing and microRNA sequencing on peripheral blood samples of SCD (preclinical AD) subjects and age-matched elderly control subjects with normal cognition (NC). The same procedures were also performed on MCI and AD blood samples as a comparison. We characterized the transcriptomic changes across multiple levels, including gene expression, transcript isoform expression, splicing, lncRNA and miRNA regulation networks, and the co-expression network (Fig. 1). We found that the type I interferon signaling pathway was remarkably downregulated in SCD while upregulated in later stages, possibly due to a comprehensive orchestration in multiple levels of transcriptional regulation. Outside data showed that normal samples with lower expression of hub genes in the co-expression network module for the type I interferon signaling pathway exhibited a higher disease conversion rate to MCI. The dysregulated molecular pathways and hub genes in SCD would provide insight into the etiology of disease conversion and serve as candidate biomarkers.

Results

Participant characteristics

To explore the potential mechanism of disease progression and identify candidate blood transcriptome biomarkers, we performed RNA sequencing and microRNA sequencing on peripheral blood samples from 44 patients with subjective cognitive decline (SCD) in preclinical AD and 82 individuals with normal cognition (NC), and assessed the peripheral blood transcriptomic dysregulation for SCD (Fig. 1). To observe the dynamics of these changes as the disease progresses, 51 MCI, and 25 AD samples were also sequenced following the same procedures as the SCD samples. To reduce the impact of comorbidities on the transcriptome, individuals with a history of hematological diseases, tumor and brain diseases (e.g., Parkinson's disease) were removed. Table 1 showed the demographic and cognitive characteristics of the enrolled participants from the four groups. The characteristics of the RNA sequencing data for the participants can be found in Supplementary Fig. 1. Compared with the NC group, the SCD and MCI groups did not show any significant difference with respect to age, gender, or education, whereas the AD patients were much older and had a higher proportion of APOE4 carriers than other groups. Both MCI and AD groups were significantly different from the NC group according to their cognition assessments, i.e. MMSE (Mini-Mental State Examination), MoCA_B (Montreal cognitive assessment-basic), and ACEIII scores (Addenbrooke's Cognitive Examination – III). The matched confounding factors between SCD and NC ensured that the transcriptomic changes identified in SCD mainly resulted from the disease state.

Table 1
The characteristics of participants of each group.

	NC (n = 82)	SCD (n = 44)		MCI (n = 51)		AD (n = 25)		P-value ^d
	Mean (SD) or %	Mean (SD) or %	P-value ^a	Mean (SD) or %	P-value ^b	Mean (SD) or %	P-value ^c	
Age	64.37 (8)	64.73 (7.18)	0.995	66.06 (7.72)	0.62	71.8 (8.6)	2.83e-4	0.00055
Gender (Male %)	32.93%	31.82%	1	23.53%	0.34	36%	0.97	0.62
Education	11.67 (4.31)	12.07 (3.49)	0.94	10.16 (2.68)	0.13	10.08 (4.76)	0.28	0.029
APOE4 + (%)	23.17%	11.36%	0.17	25.49%	0.92	40%	0.16	0.056
MMSE	27.76 (2.22)	27.59 (1.63)	0.99	26.51 (2.05)	0.032	17.24 (4.79)	4.11e-14	5.07e-44
MoCA_B	25.45 (2.95)	24.43 (2.94)	0.38	22 (3.21)	3.06e-07	12.32 (5.46)	4.12e-14	3.22e-39
ACEIII	80.51 (11.14)	78.98 (7.54)	0.88	70.86 (9.36)	9.04e-06	51.08 (16.71)	4.65e-14	1.89e-24

NC = normal cognition; SCD = subjective cognitive decline; MCI = mild cognitive impairment; AD = Alzheimer's disease; MMSE = Mini-Mental State Examination; MoCA_B = Montreal cognitive assessment-basic; ACEIII = Addenbrooke's Cognitive Examination - III; SD = standard deviation.

Results presented as mean \pm SD or frequencies with proportions. Quantitative and categorical characteristics differences were assessed with ANOVA and chi-square test, respectively.

^a P-value for comparison between SCD and NC.

^b P-value for comparison between MCI and NC.

^c P-value for comparison between AD and NC.

^d P-value for comparison among NC, SCD, MCI and AD.

The type I interferon signaling pathway is down-regulated in SCD

Transcript isoform diversity and dysregulation show higher disease specificity and are being increasingly implicated in neurological and neurodegenerative diseases [17, 18], we therefore quantified gene expression at both gene and isoform-level. With the transcriptome data, we identified 101 protein-coding genes and 360 transcript isoforms that were differentially expressed ($P < 0.05$ & $|\log_2(\text{Foldchange, FC})| > \log_2(1.3)$) in the SCD

group as compared with the NC group (Fig. 2A and Supplementary Table 1). Notably, although there was a substantial overlap ($P = 4.4 \times 10^{-155}$) between differentially expressed (DE) genes and isoforms, isoform-level alterations exhibited larger fold changes (Fig. 2A) than the gene-level and disease specificity (Supplementary Fig. 2B-C), highlighting the importance of splicing dysregulation in preclinical AD pathogenesis. In accordance with SCD as the earliest manifestation of AD, some previously reported AD-associated genes were identified here, such as *NR3C1* and *GSK3B* [19, 20]. Consistent with previous results [21], principal component analysis (PCA) based on the DE genes and isoforms (Fig. 2B, Supplementary Fig. 2D-E) revealed moderately separated clusters between disorders and normal controls. We also noted a scattered distribution of samples in each group. Both of which indicate a high heterogeneity of expression profile of peripheral blood.

As shown in the volcano plot (Fig. 2C), some interferon (IFN) stimulated genes and their isoforms were substantially downregulated in SCD, such as *IFI27*, *OAS2*, *IFI44L*, and *RSAD2*. Consistent with this finding, the interferon signaling pathway activity was significantly inhibited in SCD (Supplementary Fig. 3A), as confirmed by ingenuity pathway analysis (IPA) [22], especially for the type I interferon signaling pathway (Fig. 2D, Supplementary Table 2). Notably, *STAT1*, a key mediator of IFN signaling [23], was downregulated in SCD (Fig. 2E, Supplementary Fig. 3B). *STAT1* mediates cellular response to interferons, cytokines, and other growth factors and activates the transcription of IFN-stimulated genes that target almost any step in a virus life cycle [24]. Therefore, the downregulation of *STAT1* and IFN stimulated genes and type I IFN signaling in SCD would increase the replication of viral, like *Herpesviridae*. The bacterial and viral infections were reported to contribute to the pathophysiology of AD or to cognitive decline, most frequently implicating *Herpesviridae* [25].

Notably, the type I IFN signaling was activated in AD (Supplementary Fig. 3C-D, Supplementary Table 2), which may arise from the feedback regulation of antiviral response. Furthermore, gene set enrichment analysis (GSEA), takes all of the genes into consideration rather than only DE genes, which also gives validation to type I IFN signaling being down-regulated in SCD while up-regulated in MCI and AD compared with NC (Fig. 2F). The ISG (IFN-stimulated gene) score defining IFN signaling signature based on the mean of expression of six ISGs (*IFI44L*, *IFI27*, *RSAD2*, *SIGLEC1*, *IFIT1*, and *IS15*) in this pathway [26, 27], was significantly reduced in SCD compared with NC and gradually increased in MCI and AD (Fig. 2G and 2H). Together, these results demonstrated that in contrast to the MCI and AD stages, the SCD stage exhibited a down-regulation of type I interferon signaling pathways.

STAT1, a key transcription regulator of type I IFN signaling, is differentially spliced in SCD

Alternative splicing is an important post-transcriptional regulation mechanism, contributing to isoform diversity and protein complexity [28, 29]. Here, we identified 138 genes with differentially local splicing (DS, $P < 0.001$) in SCD compared with NC (Fig. 3A, Supplementary Table 3). DS genes in SCD overlapped significantly ($P = 0.0032$) with DE transcript isoforms, indicating that local splicing could partially explain isoform dysregulation. There were significant pairwise overlaps of DS genes between three stages of AD (Supplementary Fig. 4A). For example, differentially excised introns of *NRF1*, a transcription factor contributing to the pathogenesis of neurodegenerative diseases via perturbation of diverse mitochondrial and extra-mitochondrial functions [30], were found in SCD and AD (Supplementary Fig. 4D). However, DS genes in

AD exhibited few overlaps with those previously identified in brain transcriptome data (Supplementary Fig. 4B-C) [31, 32], highlighting the tissue-specificity of splicing events.

Differentially spliced type I IFN signaling genes in SCD significantly overlapped with those that were differentially expressed at the isoform level, including *STAT1*, *STAT2*, and *MX1* (Fig. 3B and 3C), which may partially resolve the dysregulation of the type I IFN signaling pathway. Notably, a DS intron cluster (chr2:191874730–191878744) in *STAT1* exhibited significantly increased exon skipping in SCD ($P=1.22e-05$) and MCI ($P=7.4e-4$) compared with NC (Fig. 3E). We identified splicing quantitative trait locus (sQTLs) driving or contributing to this DS using FastQTL [33], adjusting for known and inferred covariates (Supplementary Fig. 5 and Supplementary Table 4). The most significant SNP rs118149197 ($P=4.3e-20$, Fig. 3F-G), located in the UTR5' region of *STAT1*, with a higher mutation burden in SCD and MCI than NC, was predicted to affect RNA splicing by SPIDEX [34]. Individuals with this variant had a significantly higher PSI (percent spliced in) of this intron cluster (Fig. 3F). Consistent with a previous observation that DS events might predict the aberrant expression of isoforms [17], some transcripts of *STAT1* showed significantly decreased expression and transcript usage in SCD, like *STAT1-201* and *STAT1-223* (Fig. 3H and Supplementary Fig. 6). As described above, *STAT1* is one of the key transcription regulators of type I interferon signaling and activates the transcription of IFN-stimulated genes. Therefore, the dysregulated splicing and expression of *STAT1* in SCD patients may reduce their response to interferons and the defense to viruses (e.g, herpes simplex).

NRIR and has-miR-146a-5p, upstream regulators of type I IFN signaling genes, are dysregulated in SCD

Long noncoding RNAs (lncRNAs) and miRNAs regulate gene expression at epigenetic, transcriptional, and post-transcriptional levels [35, 36]. They play a key role in neurogenesis, neuronal maturation, neuronal function, and neuronal survival, and thus are involved in many neurological diseases like epilepsy and AD [37–39]. Here, we sought to explore the regulation of ncRNAs for the IFN signaling genes in SCD.

Based on GENCODE annotations, we identified 8 lncRNAs exhibiting differential expression in SCD compared with NC (Fig. 4A, Supplementary Table 5). Notably, as shown in the volcano plot (Fig. 4B), *NRIR* was remarkably down-regulated ($\log_2FC = -0.80$; $P=1.9e-4$) in SCD. The expression of this lncRNA was slightly increased in MCI ($\log_2FC = 0.2$), and significantly elevated in AD ($\log_2FC = 0.6$; $P=0.06$) (Fig. 4C). Consistent with *NRIR* driving interferon response in human monocytes [40], *NRIR* exhibited significantly positive correlation with the ISG score ($R=0.85$, $P<2.2e-16$, Fig. 4D) for the type I IFN signaling and *STAT1* ($R=0.85$, $P<2.2e-16$, Supplementary Fig. 7), indicating that decreased expression of *NRIR* might play a regulatory role in the impaired interferon activity of SCD.

To assess the potential role of miRNAs dysregulation in transcriptomic alteration, we performed genome-wide miRNA expression profiling in samples with mRNA sequencing. We identified differentially expressed miRNAs (Fig. 4E, Supplementary Table 6) and predicted their target genes. We noted *has-miR-125b-5p* and *hsa-miR-33a-5p*, two blood-based miRNA biomarkers for diagnosis of AD [41], were dysregulated in both SCD and AD. Several miRNAs significantly targeted genes in the type I IFN signaling pathway (FDR < 0.05, Supplementary Table 7). Notably, *has-miR-146a-5p* showing the most significant enrichment (FDR = $6.3e-15$), significantly targeted and negatively regulated down-expressed type I IFN genes in SCD, such as *STAT1* (Fig. 4F), *IFIT1*, and *ISG15*. Compared with NC, it was slightly up-regulated in SCD and returned to normal in MCI and AD (Fig. 4G).

The negative correlation ($R = -0.38$, $P = 3.0e-8$) between the PC1 of type I IFN signaling genes and the PC1 of miRNAs that targeted these genes (Fig. 4G) suggests that the accumulation of slightly upregulated miRNA negatively regulates the type I IFN signaling pathway.

Co-expression network module negatively associated with SCD is enriched for IFN signaling pathway

To further gain a systematic understanding of the relationship between expression changes and disease status and regulatory interactions among molecules, we performed integrated weighted gene co-expression network analysis (WGCNA) for protein-coding mRNAs and lncRNAs at gene and isoform levels, and miRNA to assign individual RNAs into network modules [42].

We identified 18 gene modules (Fig. 5A) summarized by eigengenes (i.e., PC1) in SCD, and assessed the association between them and disease status and covariates. The module M5 was significantly negatively correlated with disease status (Pearson's correlation, $R = -0.23$, $P = 0.01$) but not with any confounding factors such as age or gender, suggesting that this module was primarily driven by the SCD status. It showed a remarkably positive correlation with the ISG score (Pearson's correlation, $R = 0.93$, $P = 3e-55$) and was enriched for the interferon signaling pathway (Fig. 5B). WGCNA permits screening for the hub genes that may be promising biomarkers for diagnosis and prediction of outcomes of disease [43]. We evaluated the module membership (MM, correlation of gene expression with the module eigengene) and gene significance (GS, mediated P -value of each gene ($GS = -\log_{10}P$) in the linear regression between gene expression and the clinical traits) for each gene in the identified modules [43]. Several hub genes with both higher correlation with diagnosis ($GS > 0.2$) and higher module connectivity ($MM > 0.80$) were identified in module M5 (Fig. 5C, Supplementary Fig. 8), including the transcription factor *STAT1* and lncRNA regulator *NRIR* of IFN signaling identified above. Besides, *STAT2*, another key mediator of the JAK-STAT pathway, had the highest within-module connectivity ($MM = 0.94$) and a strong correlation with SCD ($GS = -0.22$), suggesting an association with SCD and a key role in the regulation of the interferon signaling. The isoform-level network of SCD (Supplementary Fig. 9) captured the generally equivalent interferon module (i.e., module M5), as well as a module M12 related to neutrophil degranulation, mRNA metabolism, and mRNA splicing, demonstrating the importance of splicing dysregulation. In contrast to what we found in SCD, modules of virus infection (module M11) and interferon signaling (module M19) in the AD co-expression network were significantly upregulated (Supplementary Fig. 10).

STAT1 and TRIM22 may serve as candidate biomarkers for conversion to MCI

We further sought to determine whether the hub genes could be candidate biomarkers for disease conversion in a longitudinal dataset (i.e., Alzheimer's Disease Neuroimaging Initiative, ADNI). A total of 743 participants (245 objectively NC, 382 MCI, and 116 AD) underwent Affymetrix array sequencing for blood transcriptome. Objectively NC participants were divided into two groups based on the median expression level of each hub gene. For 16 out of 23 hub genes (Supplementary Table 8), individuals of the lower-expression group carried a higher risk of conversion to MCI/AD compared with the higher-expression group, particularly for *TRIM22* ($P = 0.0022$, Fig. 5C) and *STAT1* ($P = 0.0031$, Fig. 5D). In the two groups with lower expression of *STAT1* or *TRIM22*, 26 of 122 samples converted to MCI and 1 converted to AD, which was more than twice the

conversion rate of the higher-expression group. STAT1 mediates the actions of IFNs and cytokines and upregulates genes causing pathogen response [44]. TRIM22 is an IFN-inducible TRIM family protein that restricts the replication of viruses via either regulating innate signaling pathways or serving as a viral restriction factor [45]. Thus, the down regulation of *STAT1*, *TRIM22*, and interferon signaling would weaken the cellular antiviral ability, increase virus replication (like replication of *Herpesviridae*) [46], and increase the risk of disease conversion. As expected, samples with low interferon activity had a high conversion rate when taking the first quartile of the ISG score as the threshold (Fig. 5E). Overall, this indicates that *STAT1* and *TRIM22*, hub genes of the interferon signaling module, could be candidate biomarkers for disease conversion.

Discussion

We present a comprehensive RNA-seq analysis for preclinical AD, including protein-coding and long noncoding genes at gene and isoform levels, local splicing, and miRNAs, as well as an integrative miRNA-mRNA-lncRNA co-expression network. We noted transcript isoforms exhibited larger effect sizes and disease specificity, highlighting the importance of splicing dysregulation in disease pathogenesis. We found that the activity of interferon signaling pathways was significantly down-regulated in SCD but up-regulated in MCI and AD. The impaired IFN activity in SCD may arise from differential splicing, the regulation of lncRNA (e.g., *NRIR*) and miRNA (e.g., *has-miR-146a-5*), and regulation of hub transcription factors (e.g., *STAT1* and *STAT2*). We also identified two positive regulatory hub genes (*STAT1* and *TRIM22*) of the interferon signaling module to be candidate biomarkers for disorder conversion. Objectively NC individuals with lower interferon signaling activity and expression of *STAT1* or *TRIM22* in ADNI exhibited a significantly higher conversion rate to MCI.

The accumulating evidence supports the long-standing infectious hypothesis for AD etiology, such as the presence of the herpes virus in brains and amyloid plaques of AD patients, and the increased risk of dementia in samples with herpes infection [47–49]. The inflammation for AD hallmarks has long been recognized; inflammation leads to the aggregation of A β plaques and tangles, which in turn can lead to more inflammation [50]. This cascade could be initiated by microbial pathogens. Alternatively, the presence of A β and tau in brain is the protective response against infection [51]. INFs are a group of cytokines released by host cells to protect the cells from viral infections [52]. Individuals with declined interferon signaling activity might have a higher risk of viral infections like *Herpesviridae*, and thus increased risk of progression to AD. Increased viral infection or replication would cause a feedback up-regulation of the antiviral response [53], which explains the up-regulation of defence response pathways in MCI and AD.

Antiviral treatment was associated with decreased risk of dementia [54]. A pilot trial of IFN β 1a for early mild-to-moderate Alzheimer's patients showed the treatment group revealed significant improvements in the instrumental activities of daily living and physical self-maintenance scale compared with the placebo group, indicating a potential protective role for antiviral treatment against dementia development [55]. Therefore, interferon supplementation for preclinical AD patients with impaired interferon activity might prevent or delay disease progression. Nevertheless, the epidemiological evidence indicated treatment of patients with non-steroidal anti-inflammatory drugs prior to the development of AD reduced the possibility of developing the disease. Therefore, it's also crucial to assess disease heterogeneity, the time point of initiation, and the duration of treatment.

There are some limitations in this study. First, although blood-based biomarkers are attractive for their easy availability, gene expression data using peripheral blood is easily affected by underlying factors. We noted samples exhibited higher intra-group heterogeneity and moderate inter-group differences. Second, Alzheimer's disease is a heterogeneous disease with diverse pathophysiologic mechanisms [53], the current sample size is not enough to capture molecular subtypes, which is critical for precision medicine. Third, the study was designed as a cross-sectional analysis, although several dysregulated pathways and candidate biomarker features were proposed, further longitudinal large-scale studies are needed to support the dynamics of signalings and the validity of high-risk features.

Collectively, our integrative analysis of different transcriptome biotypes at multiple levels revealed the blood transcriptional changes across progression stages of AD, particularly in SCD. These data consistently demonstrate that interferon signaling pathways are significantly down-regulated in SCD. We propose two candidate biomarkers, *STAT1* and *TRIM22*, for conversion to MCI/AD. Our findings provide an alternative hypothesis of disease progression mechanisms and guidance for disease prevention and intervention.

Methods

Sample collection

Whole blood samples and clinical data of 222 participants were obtained from Shanghai Sixth People's Hospital, including 49 SCD (preclinical AD) samples, 57 MCI samples, 28 AD samples, and samples from 88 elderly subjects with normal cognition (NC). They were diagnosed according to the criteria of National Institute on Aging-Alzheimer's Association workgroups and the Subjective Cognitive Decline Initiative, SCD-I and NIA-AA. All case metadata, including age, sex, education, cognitive function, disease history, and *APOE* genotype are provided in Supplementary Table 9. Written informed consent was obtained from all participants under protocols approved by the Institutional Review Board. After excluding individuals with comorbidities, a total of 202 participants, comprising 82 NC, 44 SCD, 51 MCI, and 25 AD samples, were finally enrolled in this study.

mRNA library preparation, sequencing, mapping, and quantification

2.5 ml human whole blood was collected in PAXgene Blood RNA Tubes. Total RNA was extracted using the PAXgene Blood RNA Kit as per manufacturer's instructions (Qiagen). Then the integrity of the total RNA was determined by a 2100 Bioanalyser (Agilent) and quantified using a NanoDrop (Thermo Scientific). About 1 μ g high-quality or media-quality RNA sample ($OD_{260/280} = 1.9 \sim 2.0$, RNA integrity number (RIN) ≥ 8) was used to construct the sequencing library. Then, rRNA was depleted (RiboZero Gold, Illumina) and libraries were prepared using the TruSeq v2 kit (Illumina) to construct stranded libraries. Libraries underwent 150-bp paired-end sequencing on an Illumina NovaSeq™ 6000.

All mRNA-seq samples were analyzed following the long-RNA-seq-pipeline of ENCODE Consortium. RNA-seq reads in FASTQ format were inspected using the FastQC [56]. Barcode and adapter contamination and low-quality bases (Phred score < 30 at ends) were trimmed using cutadapt (v1.18). Trimmed reads were then

aligned to hg19 reference genome using Gencode V33lift37 annotations via STAR (v2.4.2a) in 2-Pass mode[57]. Gene and isoform-level quantifications were evaluated using RSEM (v1.3.3) [58]. There was no significant difference in RNA integrity number (RIN), number of raw reads, and uniquely mapped reads among those four groups. On average, 46.0 million raw reads were sequenced, with 35.0 million reads uniquely mapped to the human reference genome (hg19) (Supplementary Fig. 1).

miRNA library preparation, sequencing, mapping, and quantification

miRNA libraries were prepared with the QIAseq miRNA Library Kit (QIAGEN) according to the manufacturer's instructions. Libraries were sequenced on the Illumina NovaSeq™ 6000 with 150bp paired-end sequencing. Only R1 reads were used for later analysis. RNA-seq reads in FASTQ format were inspected using FastQC. Adaptor sequences were removed from the raw R1 reads using Cutadapt. Reads < 18 nt and untrimmed reads were discarded.

For quantification of mature miRNAs, the genome (hg19) index file was first built by *bowtie-build* [59]. Then trimmed miRNA reads were mapped to the indexed hg19 reference genome using the *mapper.pl* script in the miRDeep2 package [60]. The *miRDeep2.pl* was run with default settings to quantify the number of reads mapped to mature miRNAs in miRBase (release 22.1) [61, 62].

Differential expression analysis

Expected counts were compiled from gene and isoform-level RSEM quantifications and imported into R for downstream analyses. To avoid the effect of comorbidity on blood transcriptome profile, samples with a disease history of hematological disease, tumor, brain disease, and psychiatric disease (e.g., Parkinson's disease) were excluded. Genes/isoforms were filtered to include those with TPM > 1 in at least 25% of samples. MiRNA samples were filtered to include those with count > 3 in at least 50% of samples. Removing low expressed genes/miRNAs and comorbidity samples resulted in 12,664 genes and 36,005 isoforms, and 811 miRNAs of 202 participants.

Differentially expressed genes (DEG), differential expressed transcript isoforms (DET), and differential expressed miRNA (P -value < 0.05, $|\log_2(\text{FC})| > \log_2(1.3)$) were identified using R package Deseq2 [63] with controlling for covariates (age, gender, and RNA integrity number (RIN)). Normalized expression matrices after correcting for library size were obtained for downstream analysis.

Differential Local Splicing (DS) analysis

We used LeafCutter [64] to obtain clusters of variably spliced introns. First, the above bam files produced by the STAR 2-pass were converted into junction files. Then, the introns found in the junction files were clustered together with default settings. Third, the LeafCutter *leafcutter_ds.R* script was used to perform differential splicing between each stage (SCD, MCI, and AD) and the normal cognition (NC) group. Only introns with at least one supporting read in > 5 samples, and > 3 samples in each group having at least 20 reads were retained. A significance file (P -value) and an effect-sizes (DeltaPSI, Percent-spliced-in) file were returned. Next, differential local splicings were further annotated using *prepare_results.R* script as annotated for cryptic (cryptic 5', cryptic 3', cryptic unanchored, or novel annotated pair) based on Gencode V33lift37 annotations.

Finally, DS clusters were schematically visualized using the LeafCutter shiny app. DS intron clusters and domains were mapped onto transcripts using gViz (v1.32) and ensemblldb (v2.12.1) R packages.

Splicing quantitative trait loci (sQTL) discovery

FastQTL [33] was used to identify cis-splicing QTLs (sQTLs) within regions of ± 100 kb of the intron clusters. SNPs genotypes were called from whole-genome sequencing data following the GATK4 pipeline [65], and filtered based on Hardy-Weinberg equilibrium ($-hwe\ 1e6$), minor allele frequency ($-maf\ 0.01$), individual missing genotype rate ($-mind\ 0.10$), and variant missing genotype rate ($-geno\ 0.05$) using PLINK v1.9 [66]. The intron excision ratios phenotype data and hidden covariates (top5 PCs) were prepared using *prepare_phenotype_table.py* in LeafCutter. sQTLs were identified adjusting for the inferred hidden covariates and known covariates (i.e, diagnosis, gender, age, education, and cognitive score) in both nominal and permutation pass modes.

Prediction of miRNA targets

miRNA targets were extracted from an experimentally supported miRNA-mRNA interactions database (miRTarBase v7.0; <http://mirtarbase.mbc.nctu.edu.tw/>) [67] and a predicted miRNA-targets interactions database (miRDB v6.0; <http://mirdb.org>) [68, 69] that predicted from high-throughput sequencing experiments with the miRNA sequences from miRBase version 22.

Enrichment analysis

To gain further insight into the biological functions of differentially expressed genes, we performed enrichment analysis of gene ontology (GO) biological processes, KEGG, and Reactome pathways using R package gProfileR (v0.7.0; <https://biit.cs.ut.ee/gprofiler/>) [70], where only pathways containing less than 500 genes were assessed. We also performed IPA [22] (Ingenuity Pathway Analysis; QIAGEN) to predict whether the pathway is activated or inhibited. To further determine whether the gene sets defined based on prior biological knowledge shows statistically concordant up or down-regulation in disease (SCD, MCI, and AD) compared with NC, the enrichment of GO and KEGG and Reactome gene set in the molecular signature database (MSigDB) was performed using normalized expression matrix of all genes as input by applying GSEA (V4.1.0) [71, 72]. To explore the biological processes enrichment of the WGCNA module (below), we performed enrichment analysis of gene ontology (GO) biological processes, KEGG, and Reactome pathways using R package gProfileR (v0.7.0; <https://biit.cs.ut.ee/gprofiler/>) [70], where only pathways containing less than 500 genes were assessed.

Weighted gene coexpression network analysis (WGCNA)

To explore the pattern of correlations between all possible pairs of genes (including protein-coding genes, lncRNA, and miRNA) and identify their modules, signed networks were constructed with the R package WGCNA [42] using a normalized expressed matrix at gene and isoform-levels separately. Top 5000 high median absolute deviation (MAD) genes/isoforms were used for analysis since low-expressed or non-varying genes usually represent noise. Soft thresholding power was picked as the minimum value satisfying $R^2 > 0.8$. Networks were constructed using the *blockwiseModules* function using parameters of *mergeCutHeight* = 0.15 and *deep split* = 2. Modules were summarized by module eigengene (ME) that was defined as the first principal component (PC1) of the module. Module (eigengene)-disease associations were evaluated using

Pearson's correlation. For each gene in the modules, the module membership (MM, correlation of its gene expression profile with the module eigengene of a given module) and gene significance (GS, mediated P -value of each gene ($GS = -\log_{10}P$) in the linear regression between gene expression and the clinical traits) were evaluated. Hub genes were identified based on the cut-off criteria ($|MM| > 0.8$ and $|GS| > 0.2$).

Differential expression analysis and Survival analysis in ADNI

An independent AD blood gene expression data, sequenced using the Affymetrix Human Genome U219 Array and normalized using the RMA (Robust Multi-chip Average), were obtained from the ADNI (<http://adni.loni.usc.edu>) [73]. Samples with normal objective cognition were divided into two groups based on their expression higher or lower than the median level. To compare the conversion rate of the two, Kaplan-Meier survival curves were fitted and plotted using survival and survminer R packages.

Statistical analyses

Statistical analysis was performed using R (version 4.0.2). Demographics were compared using ANOVA (analysis of variance) test for continuous variables and a chi-squared test for categorical variables. Comparison of gene expression and PSI between disorders and NC was performed using a two-tailed Student's t -test and Wilcoxon test. Correlations between module eigengenes and phenotype traits were assessed with Pearson's correlation. Overlaps of DE genes between two sets were assessed using a hypergeometric test. To correct for multiple comparisons, the Benjamini-Hochberg method was applied.

Data availability

All the RNA and miRNA sequencing data have been deposited in the Genome Sequence Archive for Human (GSA-Human) in National Genomics Data Center (NGDC). Please access them from the following link: <https://bigd.big.ac.cn/gsa-human/browse/HRA000942>. The ADNI data used in this study were obtained from the ADNI database (<https://adni.loni.usc.edu>).

Code availability

All custom code used in this work is available at the following GitHub repository: https://github.com/ZhaoXM-Lab/SCD_preAD.

Declarations

Acknowledgements

The authors thank the patients and their families for their selfless donations. Part of the data used here is from the ZIB (Zhangjiang International Brain Bank) Consortium. We thank Dr. Yang T. Yucheng for his suggestions on our project proposal.

Author contributions

Zhao X. and Chen J. conceived and designed the experiments. Song L. carried out data processing and analyses. Song L. wrote the manuscript. Zhao X. and Chen J. revised the manuscript. Guo Q. and Luo C.

collected samples. All authors read and proved the final manuscript.

Funding

This work was partly supported by National Key R&D Program of China (2020YFA0712403), National Natural Science Foundation of China (61932008, 61772368), Shanghai Municipal Science and Technology Major Project (2018SHZDZX01), the 111 Project (No.B18015) of China, ZJ Lab, and Shanghai Center for Brain Science and Brain-Inspired Technology.

Competing interests

The authors report no competing interests.

References

1. Ballard, C., et al., *Alzheimer's disease*. *Lancet*, 2011. **377**(9770): p. 1019–31.
2. Morris, J.C., *Early-stage and preclinical Alzheimer disease*. *Alzheimer Dis Assoc Disord*, 2005. **19**(3): p. 163–5.
3. McKhann, G.M., et al., *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*. *Alzheimers & Dementia*, 2011. **7**(3): p. 263–269.
4. Albert, M.S., et al., *The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease*. *Alzheimers & Dementia*, 2011. **7**(3): p. 270–279.
5. Sperling, R.A., et al., *Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease*. *Alzheimers & Dementia*, 2011. **7**(3): p. 280–292.
6. Hardy, J.A. and G.A. Higgins, *Alzheimer's disease: the amyloid cascade hypothesis*. *Science*, 1992. **256**(5054): p. 184–5.
7. Sperling, R.A., C.R. Jack, and P.S. Aisen, *Testing the Right Target and Right Drug at the Right Stage*. *Science Translational Medicine*, 2011. **3**(111).
8. Mantovani, E., et al., *Towards a Redefinition of Cognitive Frailty*. *Journal of Alzheimers Disease*, 2020. **76**(3): p. 831–843.
9. Jessen, F., et al., *A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease*. *Alzheimers & Dementia*, 2014. **10**(6): p. 844–852.
10. Visser, P.J., et al., *Prevalence and prognostic value of CSF markers of Alzheimer's disease pathology in patients with subjective cognitive impairment or mild cognitive impairment in the DESCRIPA study: a prospective cohort study*. *Lancet Neurol*, 2009. **8**(7): p. 619–27.
11. Dufouil, C., R. Fuhrer, and A. Alperovitch, *Subjective cognitive complaints and cognitive decline: Consequence or predictor? The epidemiology of vascular aging study*. *Journal of the American Geriatrics Society*, 2005. **53**(4): p. 616–621.

12. Glodzik-Sobanska, L., et al., *Subjective memory complaints: Presence, severity and future outcome in normal older subjects*. *Dementia and Geriatric Cognitive Disorders*, 2007. **24**(3): p. 177–184.
13. Donovan, N.J., et al., *Subjective cognitive concerns and neuropsychiatric predictors of progression to the early clinical stages of Alzheimer disease*. *Am J Geriatr Psychiatry*, 2014. **22**(12): p. 1642–51.
14. Abner, E.L., et al., *Baseline subjective memory complaints associate with increased risk of incident dementia: the PREADVISE trial*. *J Prev Alzheimers Dis*, 2015. **2**(1): p. 11–16.
15. Parfenov, V.A., et al., *Subjective cognitive decline as a predictor of future cognitive decline: a systematic review*. *Dementia & neuropsychologia*, 2020. **14**(3): p. 248–257.
16. Iturria-Medina, Y., et al., *Blood and brain gene expression trajectories mirror neuropathology and clinical deterioration in neurodegeneration*. *Brain*, 2020. **143**(2): p. 661–673.
17. Gandal, M.J., et al., *Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder*. *Science*, 2018. **362**(6420).
18. Vuong, C.K., D.L. Black, and S. Zheng, *The neurogenetics of alternative splicing*. *Nat Rev Neurosci*, 2016. **17**(5): p. 265–81.
19. de Quervain, D.J.F., et al., *Glucocorticoid-related genetic susceptibility for Alzheimer's disease*. *Human Molecular Genetics*, 2004. **13**(1): p. 47–52.
20. Llorens-Martin, M., et al., *GSK-3 beta, a pivotal kinase in Alzheimer disease*. *Frontiers in Molecular Neuroscience*, 2014. **7**.
21. Shigemizu, D., et al., *Identification of potential blood biomarkers for early diagnosis of Alzheimer's disease through RNA sequencing analysis*. *Alzheimers Res Ther*, 2020. **12**(1): p. 87.
22. Kramer, A., et al., *Causal analysis approaches in Ingenuity Pathway Analysis*. *Bioinformatics*, 2014. **30**(4): p. 523–30.
23. Au-Yeung, N., R. Mandhana, and C.M. Horvath, *Transcriptional regulation by STAT1 and STAT2 in the interferon JAK-STAT pathway*. *JAKSTAT*, 2013. **2**(3): p. e23931.
24. O'Shea, J.J., et al., *The JAK-STAT pathway: impact on human disease and therapeutic intervention*. *Annu Rev Med*, 2015. **66**: p. 311–28.
25. Lovheim, H., et al., *Herpes simplex infection and the risk of Alzheimer's disease: A nested case-control study*. *Alzheimers & Dementia*, 2015. **11**(6): p. 587–592.
26. Hadjadj, J., et al., *Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients*. *Science*, 2020. **369**(6504): p. 718–724.
27. Jeremiah, N., et al., *Inherited STING-activating mutation underlies a familial inflammatory syndrome with lupus-like manifestations*. *J Clin Invest*, 2014. **124**(12): p. 5516–20.
28. Kornblihtt, A.R., et al., *Alternative splicing: a pivotal step between eukaryotic transcription and translation*. *Nat Rev Mol Cell Biol*, 2013. **14**(3): p. 153–65.
29. Wang, E.T., et al., *Alternative isoform regulation in human tissue transcriptomes*. *Nature*, 2008. **456**(7221): p. 470–6.
30. Satoh, J., N. Kawana, and Y. Yamamoto, *Pathway Analysis of ChIP-Seq-Based NRF1 Target Genes Suggests a Logical Hypothesis of their Involvement in the Pathogenesis of Neurodegenerative Diseases*. *Gene Regul Syst Bio*, 2013. **7**: p. 139–52.

31. Li, H.D., et al., *Integrative functional genomic analysis of intron retention in human and mouse brain with Alzheimer's disease*. *Alzheimers Dement*, 2021.
32. Raj, T., et al., *Integrative transcriptome analyses of the aging brain implicate altered splicing in Alzheimer's disease susceptibility*. *Nature Genetics*, 2018. **50**(11): p. 1584+.
33. Ongen, H., et al., *Fast and efficient QTL mapper for thousands of molecular phenotypes*. *Bioinformatics*, 2016. **32**(10): p. 1479–1485.
34. Mendes de Almeida, R., et al., *Whole gene sequencing identifies deep-intronic variants with potential functional impact in patients with hypertrophic cardiomyopathy*. *PLoS One*, 2017. **12**(8): p. e0182946.
35. Cai, Y., et al., *A brief review on the mechanisms of miRNA regulation*. *Genomics Proteomics Bioinformatics*, 2009. **7**(4): p. 147–54.
36. Gil, N. and I. Ulitsky, *Regulation of gene expression by cis-acting long non-coding RNAs*. *Nature Reviews Genetics*, 2020. **21**(2): p. 102–117.
37. Batista, P.J. and H.Y. Chang, *Long noncoding RNAs: cellular address codes in development and disease*. *Cell*, 2013. **152**(6): p. 1298–307.
38. Kopp, F., *Molecular functions and biological roles of long non-coding RNAs in human physiology and disease*. *J Gene Med*, 2019. **21**(8): p. e3104.
39. Delay, C., W. Mandemakers, and S.S. Hebert, *MicroRNAs in Alzheimer's disease*. *Neurobiology of Disease*, 2012. **46**(2): p. 285–290.
40. Mariotti, B., et al., *The Long Non-coding RNA NRIR Drives IFN-Response in Monocytes: Implication for Systemic Sclerosis*. *Frontiers in Immunology*, 2019. **10**.
41. Nagaraj, S., et al., *microRNA diagnostic panel for Alzheimer's disease and epigenetic trade-off between neurodegeneration and cancer*. *Ageing Research Reviews*, 2019. **49**: p. 125–143.
42. Zhang, B. and S. Horvath, *A general framework for weighted gene co-expression network analysis*. *Statistical Applications in Genetics and Molecular Biology*, 2005. **4**.
43. Langfelder, P. and S. Horvath, *WGCNA: an R package for weighted correlation network analysis*. *BMC Bioinformatics*, 2008. **9**: p. 559.
44. Casanova, J.L., S.M. Holland, and L.D. Notarangelo, *Inborn errors of human JAKs and STATs*. *Immunity*, 2012. **36**(4): p. 515–28.
45. Ozato, K., et al., *TRIM family proteins and their emerging roles in innate immunity*. *Nature Reviews Immunology*, 2008. **8**(11): p. 849–860.
46. Cramer, M., et al., *MxB is an interferon-induced restriction factor of human herpesviruses*. *Nature Communications*, 2018. **9**.
47. Carbone, I., et al., *Herpes virus in Alzheimer's disease: relation to progression of the disease*. *Neurobiol Aging*, 2014. **35**(1): p. 122–9.
48. Sochocka, M., K. Zwolinska, and J. Leszek, *The Infectious Etiology of Alzheimer's Disease*. *Curr Neuropharmacol*, 2017. **15**(7): p. 996–1009.
49. Middleton, P.J., et al., *Herpes-Simplex Viral Genome and Senile and Presenile Dementias of Alzheimer and Pick*. *Lancet*, 1980. **1**(8176): p. 1038–1038.

50. Akiyama, H., et al., *Inflammation and Alzheimer's disease*. Neurobiol Aging, 2000. **21**(3): p. 383–421.
51. Bourgade, K., et al., *Protective Effect of Amyloid-beta Peptides Against Herpes Simplex Virus-1 Infection in a Neuronal Cell Culture Model*. J Alzheimers Dis, 2016. **50**(4): p. 1227–41.
52. De Andrea, M., et al., *The interferon system: an overview*. Eur J Paediatr Neurol, 2002. **6 Suppl A**: p. A41-6; discussion A55-8.
53. Neff, R.A., et al., *Molecular subtyping of Alzheimer's disease using RNA sequencing data reveals novel mechanisms and targets*. Science Advances, 2021. **7**(2).
54. Lopatko Lindman, K., et al., *Herpesvirus infections, antiviral treatment, and the risk of dementia-a registry-based cohort study in Sweden*. Alzheimers Dement (N Y), 2021. **7**(1): p. e12119.
55. Grimaldi, L.M.E., et al., *A pilot study on the use of interferon beta-1a in early Alzheimer's disease subjects*. Journal of Neuroinflammation, 2014. **11**.
56. Andrews, S., *FASTQC. A quality control tool for high throughput sequence data*. 2010.
57. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner*. Bioinformatics, 2013. **29**(1): p. 15–21.
58. Li, B. and C.N. Dewey, *RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome*. BMC Bioinformatics, 2011. **12**.
59. Langmead, B., et al., *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome*. Genome Biol, 2009. **10**(3): p. R25.
60. Friedlander, M.R., et al., *miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades*. Nucleic Acids Res, 2012. **40**(1): p. 37–52.
61. Kozomara, A. and S. Griffiths-Jones, *miRBase: annotating high confidence microRNAs using deep sequencing data*. Nucleic Acids Research, 2014. **42**(D1): p. D68-D73.
62. Kozomara, A., M. Birgaoanu, and S. Griffiths-Jones, *miRBase: from microRNA sequences to function*. Nucleic Acids Research, 2019. **47**(D1): p. D155-D162.
63. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2*. Genome Biology, 2014. **15**(12).
64. Li, Y.I., et al., *Annotation-free quantification of RNA splicing using LeafCutter*. Nature Genetics, 2018. **50**(1): p. 151+.
65. Van der Auwera, G.A., et al., *From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline*. Curr Protoc Bioinformatics, 2013. **43**: p. 11 10 1–11 10 33.
66. Chang, C.C., et al., *Second-generation PLINK: rising to the challenge of larger and richer datasets*. Gigascience, 2015. **4**: p. 7.
67. Huang, H.Y., et al., *miRTarBase 2020: updates to the experimentally validated microRNA-target interaction database*. Nucleic Acids Research, 2020. **48**(D1): p. D148-D154.
68. Liu, W.J. and X.W. Wang, *Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data*. Genome Biology, 2019. **20**.
69. Chen, Y.H. and X.W. Wang, *miRDB: an online database for prediction of functional microRNA targets*. Nucleic Acids Research, 2020. **48**(D1): p. D127-D131.

70. Reimand, J., et al., *g:Profiler-a web server for functional interpretation of gene lists (2016 update)*. *Nucleic Acids Research*, 2016. **44**(W1): p. W83-W89.
71. Subramanian, A., et al., *Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles*. *Proceedings of the National Academy of Sciences of the United States of America*, 2005. **102**(43): p. 15545–15550.
72. Mootha, V.K., et al., *PGC-1 alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes*. *Nature Genetics*, 2003. **34**(3): p. 267–273.
73. Mueller, S.G., et al., *Alzheimer's disease neuroimaging initiative*. *Advances in Alzheimer's and Parkinson's Disease: Insights, Progress, and Perspectives*, 2008. **57**: p. 183–189.
74. Bateman, A., et al., *The Pfam protein families database*. *Nucleic Acids Res*, 2000. **28**(1): p. 263–6.

Figures

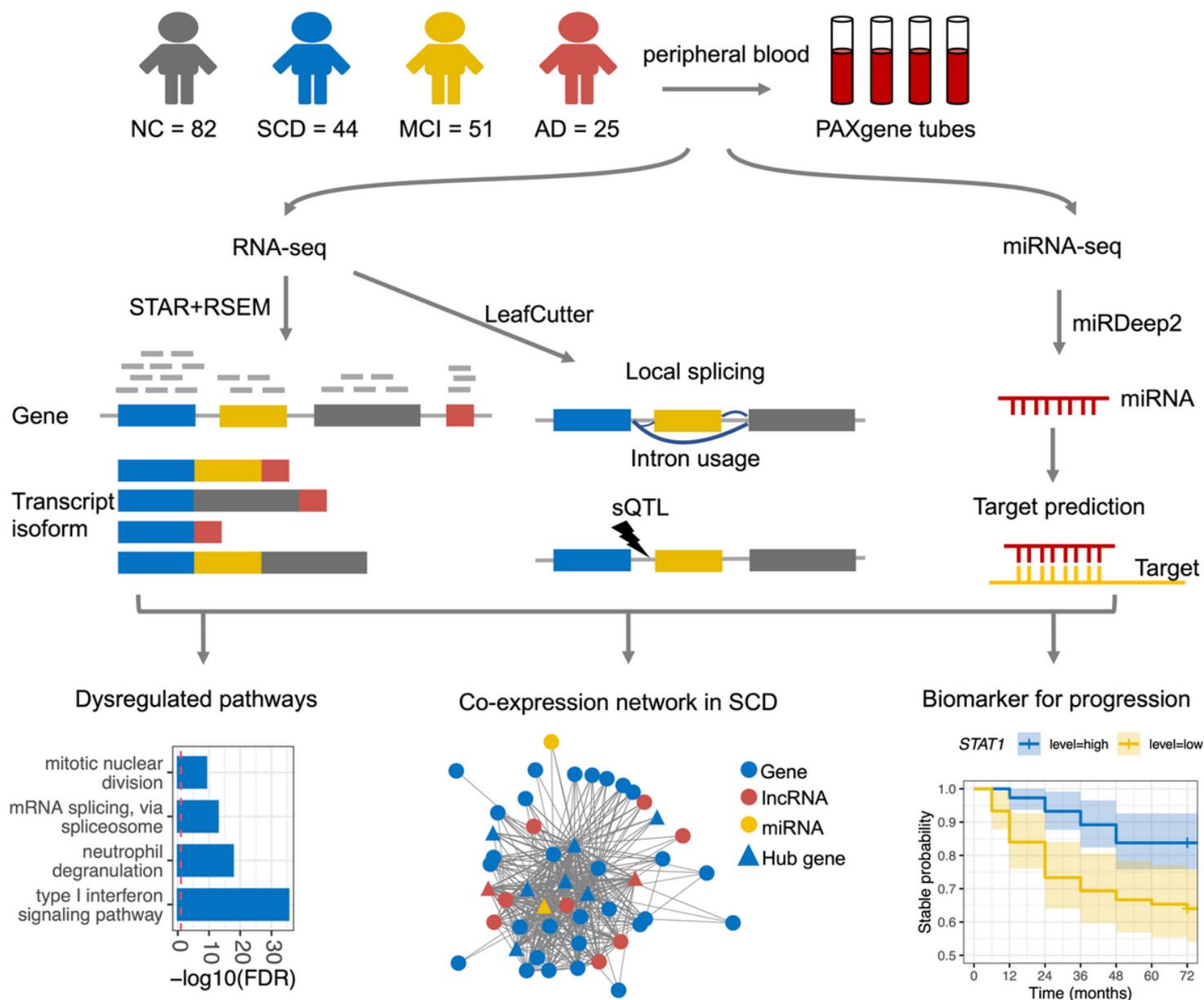


Figure 1

Study design and overview of the study.

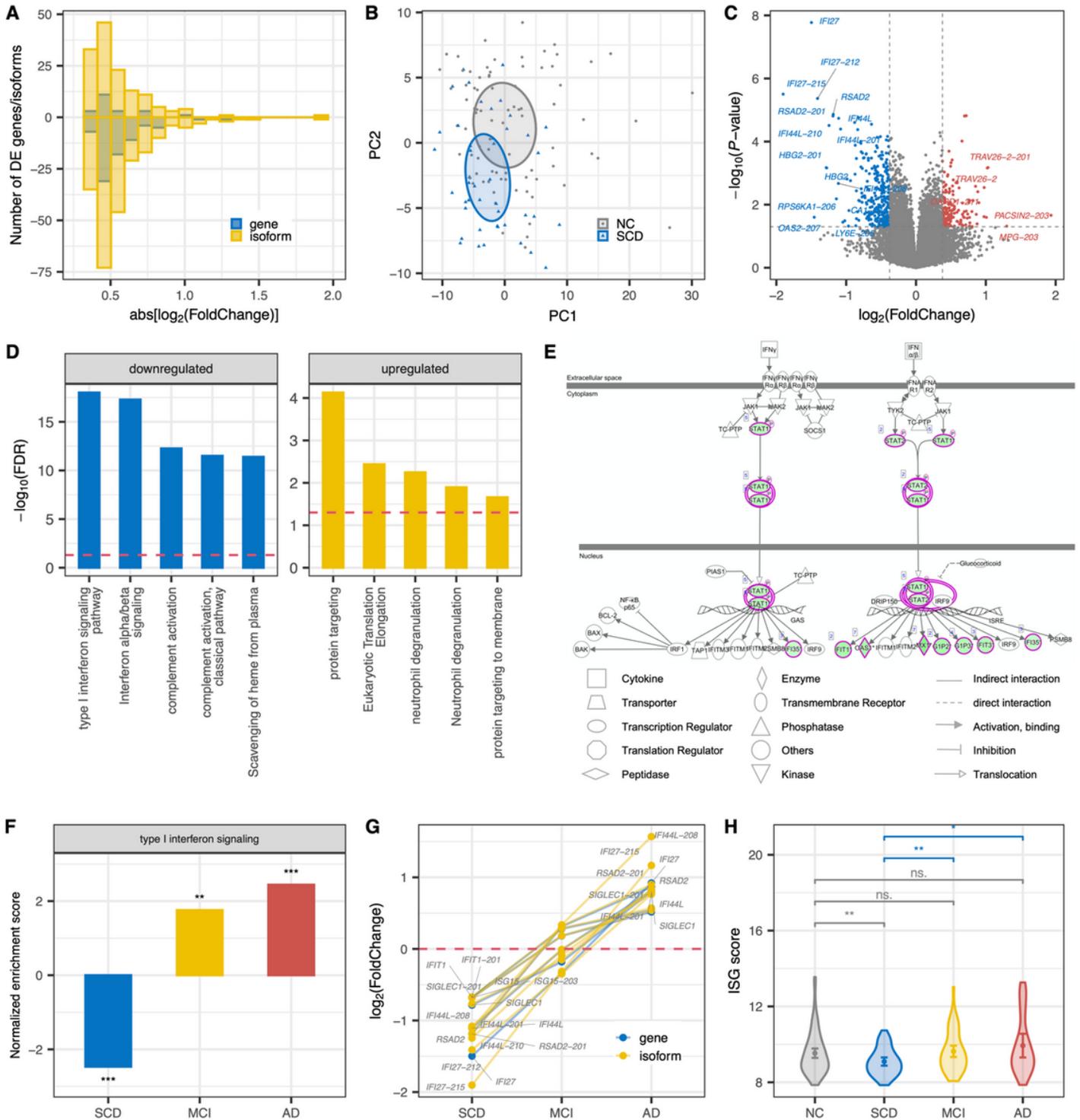


Figure 2

Downregulation of the interferon signaling pathway in SCD. (A) Fold change (FC) histograms for up (positive number) or down-regulated (negative number) protein-coding genes (blue) and isoforms (yellow) in SCD. Abs: absolute. (B) PCA plot on differentially expressed genes and isoforms of SCD. (C) Volcano plot displaying the statistical significance (P-value) versus magnitude of change (fold change) of protein-coding genes and

transcripts in SCD compared with NC. DE genes/isoforms with $|\log_2(\text{FC})| > 1$ are highlighted. (D) Pathway enrichment analysis for genes that were differentially expressed at gene or isoform level in SCD. The top 5 enriched pathways are shown. Blue and yellow represent down and up-regulated pathways, respectively. The red dashed line indicates the significance threshold (FDR = 0.05). (E) IPA canonical pathway for interferon signaling. Blue shading represents down-regulated genes. (F) GSEA in type I interferon signaling pathways across disorders. Barplot shows the normalized enrichment score (NES) and significance. A positive/negative NES value indicates up/down-regulation of this geneset in disease compared with NC. Three stars represent for P-value < 0.001, two stars for P-value < 0.01. (G) Fold changes of ISG genes (IFI44L, IFI27, RSAD2, SIGLEC1, IFIT1, and IS15) defining the IFN signature at gene and isoform level across disorders. (H) ISG score of each group based on the mean expression of these six signature genes. Error bar indicates 95% confidence interval of ISG score. P values were determined with the two-tailed t-test.

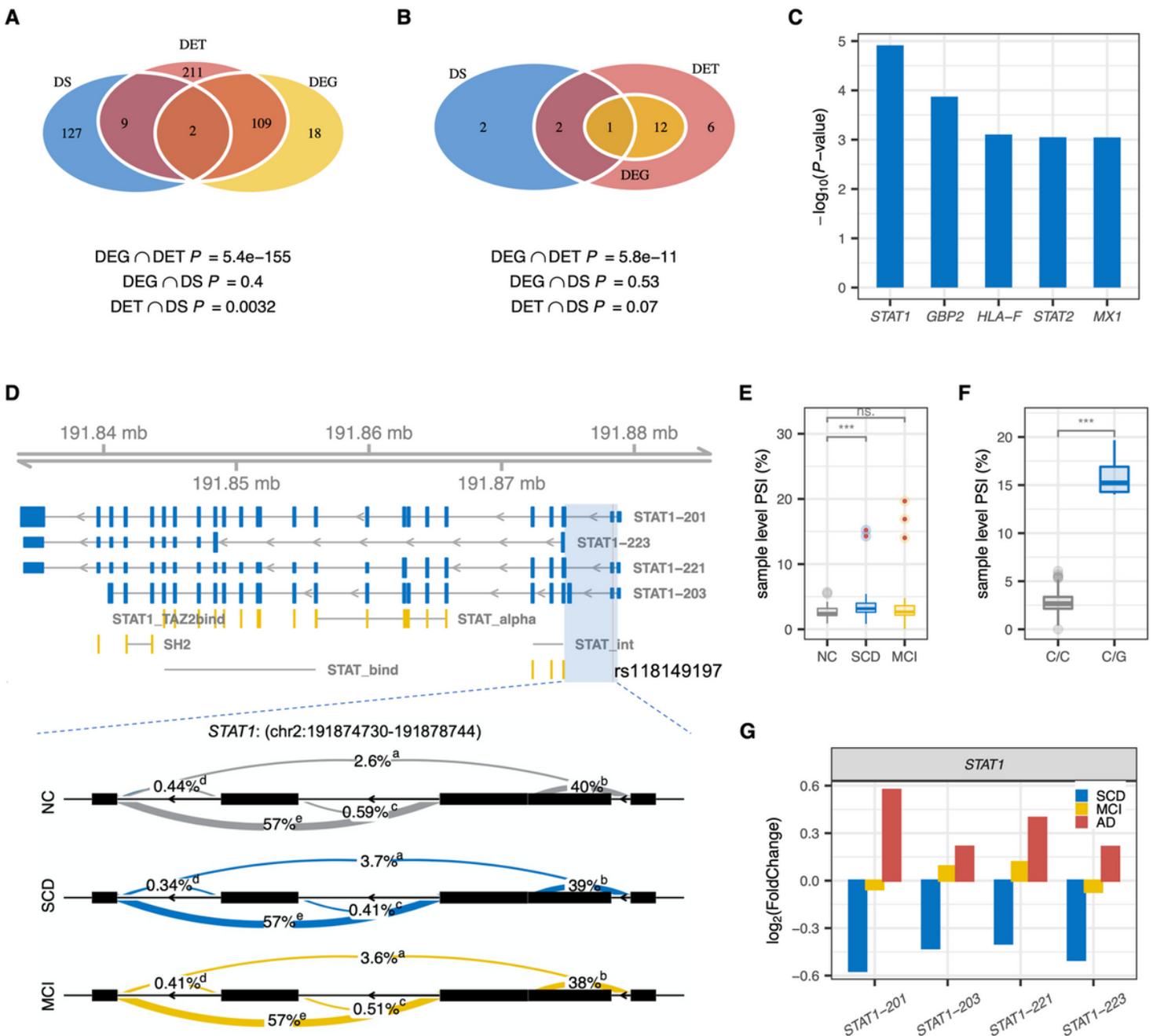


Figure 3

Differential local splicing in SCD. (A) Overlaps between significant DE genes (DEG), DE transcript isoforms (DET), and genes with significant differentially spliced (DS) intron clusters in SCD. P values for hypergeometric tests of pairwise overlaps between data types are shown at the bottom. Blue, yellow, and red indicate DS, DEG, and DET, respectively. (B) Overlaps between significant DE genes (DEG), DE transcript isoforms (DET), and DS of IFN signaling genes in SCD. (C) The significance of DS IFN signaling genes in SCD. (D) Whole-gene view of STAT1 highlighting (blue block, chr2:191874730-191878744) the intron cluster with significant DS in SCD ($P = 1.2e-5$) and MCI ($P = 7.4e-4$), as well as transcripts STAT1-201, 203, 221 and STAT1-223 that differentially expressed in SCD. Protein domain mappings from the Pfam database [74] are shown in yellow. STAT_int, STAT protein, protein interaction domain; STAT_alpha, STAT protein, all-alpha domain; STAT_bind, STAT protein, DNA binding domain; SH2, SH2 domain; STAT1_TAZ2bind, STAT1 TAZ2 binding domain. The red line denotes the sQTL rs118149197. (E) The distribution of individual PSI (percent spliced in) for the intron cluster in (D). Samples with rs118149197 were denoted with red points. P values were determined with the Wilcoxon test. (F) PSI distribution for samples with different rs118149197 genotypes. (G) Fold Changes of differentially expressed transcripts of SCD in each disorder.

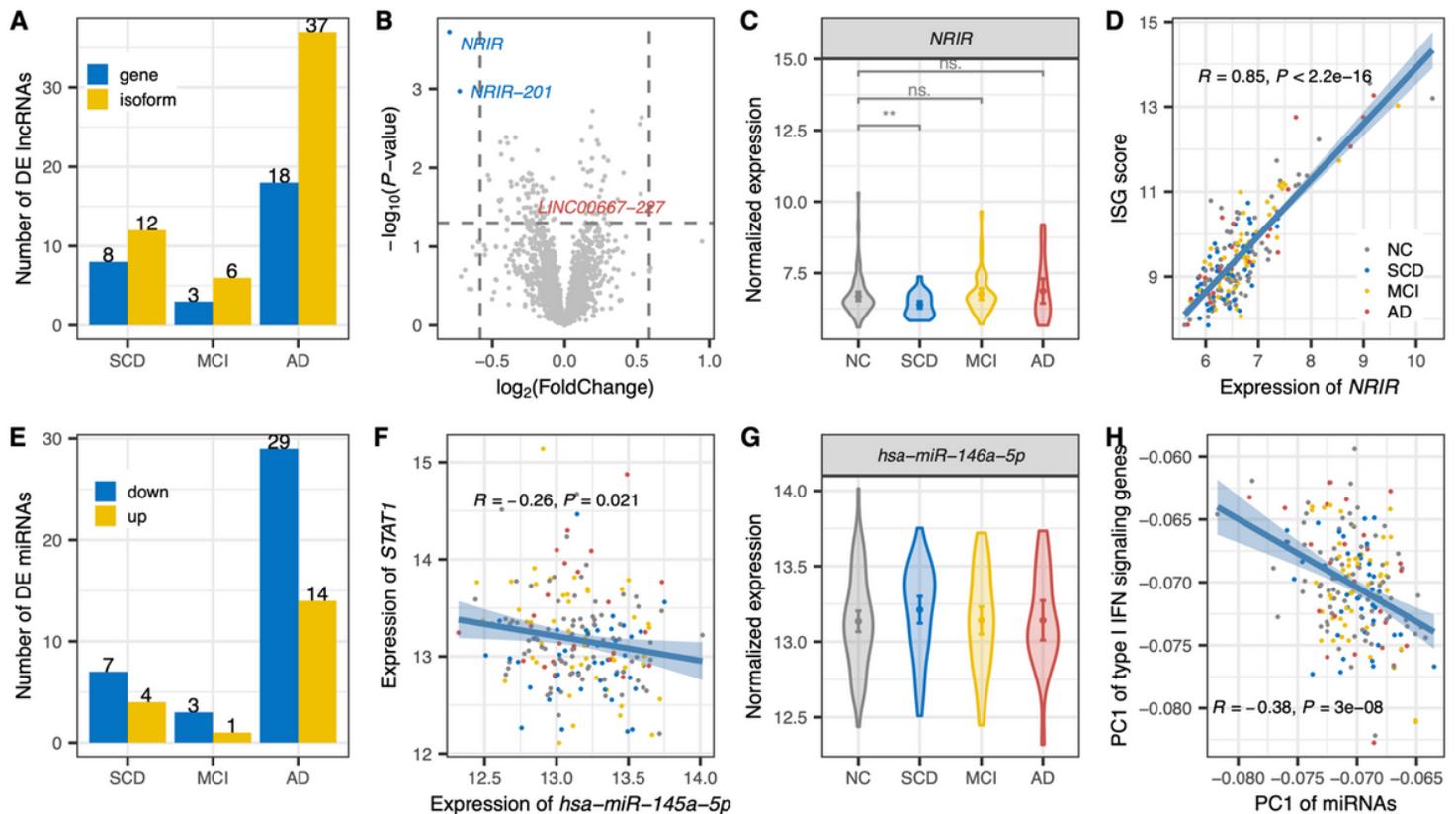


Figure 4

Dysregulation of ncRNAs regulated type I interferon signaling. (A) The number of differentially expressed long non-coding genes at gene (blue) and isoform (yellow) levels in SCD, MCI, and AD. (B) Volcano plot displaying the statistical significance (P-value) versus magnitude of change (FC) of long noncoding genes and transcripts in SCD compared with NC. (C) Distribution of normalized expression level of NRIR in each group. (D) Correlation between ISG scores and normalized expression of NRIR. (E) Up and down-regulated miRNAs among disorders. (F) Correlation between normalized expression of has-miR-145a-5p and STAT1. (G)

Distribution of normalized expression level of has-miR-145a-5p in each group. (H) Correlation between the PC1 of type I IFN signaling genes and the PC1 of miRNAs that target these genes.

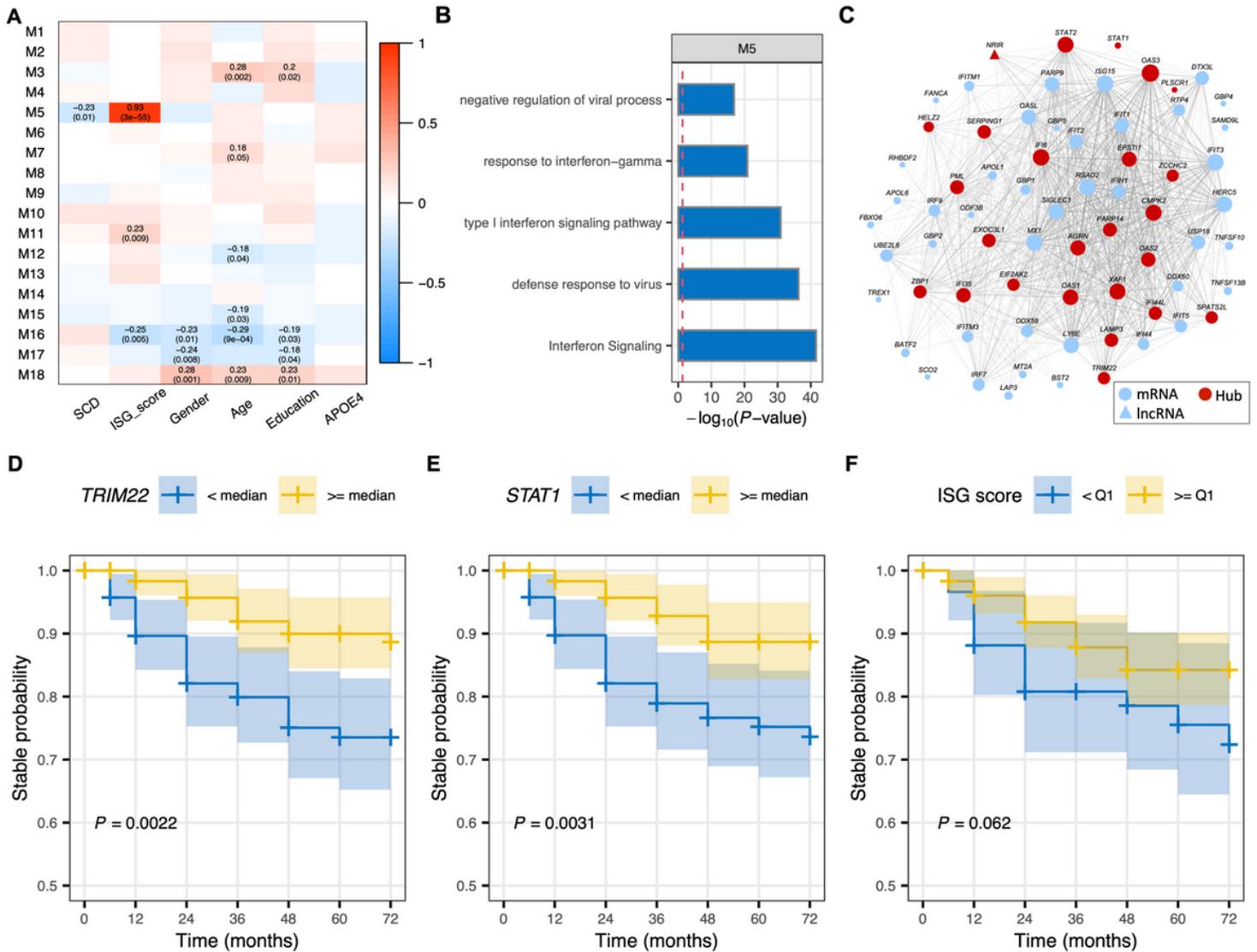


Figure 5

Hub genes of co-expression interferon signaling module serve as candidate biomarkers for conversion to MCI. (A) Pearson's correlation between module eigengenes and phenotype traits. Left module-blocks represent the miRNA-lncRNA-mRNA co-expression modules at the gene level defined in SCD. Correlation coefficients (top) and P-value (bottom) are shown in the grid, where red and blue colors indicate positive and negative correlations, separately. Only significant associations ($P < 0.05$) are displayed. (B) Top 5 enriched pathways of genes in the module M5. (C) The network of genes in the module M5, where only interaction edges with adjacency threshold > 0.15 are plotted. Edge thickness and node size correspond to pair-wise correlation and node connectivity, separately. Node shapes represent different biotypes. Hub genes are highlighted with the red nodes. (D-E) The Kaplan-Meier curves showing the 6-year disease conversion of cognitively normal participants in ADNI (Alzheimer's Disease Neuroimaging Initiative) grouped by the median expression of marker genes. (D) for TRIM22 and (E) for STAT1. Yellow and blue denote higher or lower than the median level, respectively. (F) The Kaplan-Meier curves showing the disease conversion of samples grouped by the

first quantile expression of the ISG score. Yellow and blue represent higher and lower than the first quartile (Q1) of the ISG score, respectively.

Supplementary Files

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

- [suppletables.xlsx](#)
- [supple.docx](#)