Supplementary Material

Functional enrichment analysis of AD potential target genes

The “Clusterprofiler” (4.7.1) R software package was used to enrich functions of the potential AD target genes[1]. Firstly, the enrichGO function of Gene Ontology (GO) analysis identified the potential target biological processes (BPs), molecular functions (MFs), and cellular components (CCs). Subsequently, Kyoto Encyclopedia of Genomes (KEGG) analysis was performed to look for signaling pathways enriched in potential targets[2].

Supplementary Figures

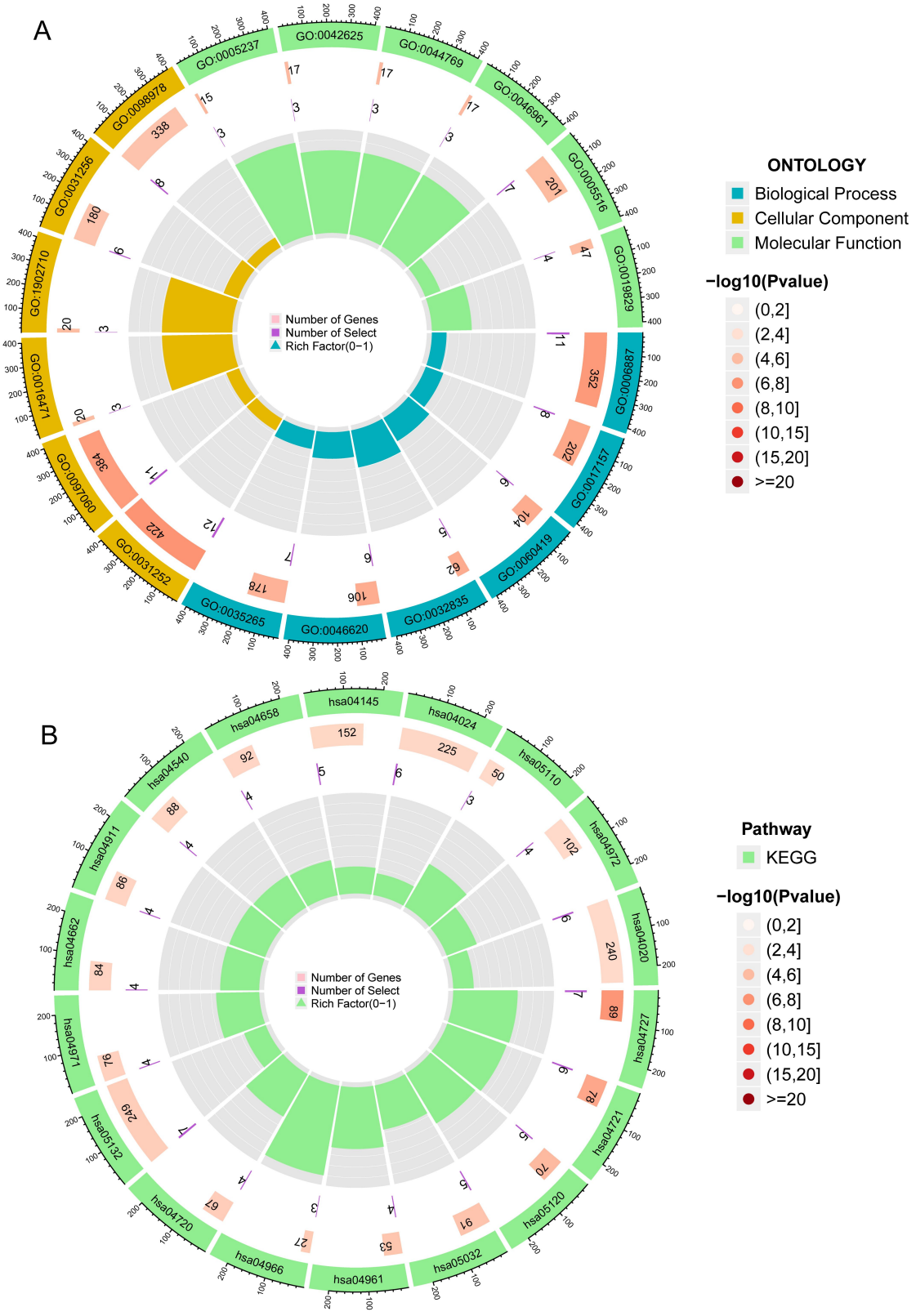

Fig. S1. Functional enrichment analysis of DEG. **(A)** The GO circle chart shows 18 GO terms. The height of the inner column represents the enriched proportion of genes to the total genes in the GO pathway. The higher the column, the greater the proportion of genes. The pink color represents the adjusted *p* value. The color was darker; the enrichment was more obvious. **(B)** The KEGG circle chart shows 18 KEGG terms. The height of the inner column represents the enriched proportion of genes to the total genes in the KEGG pathway. The higher the column, the greater the proportion of genes. The pink color represents the adjusted *p* value. The color was darker; the enrichment was more obvious.

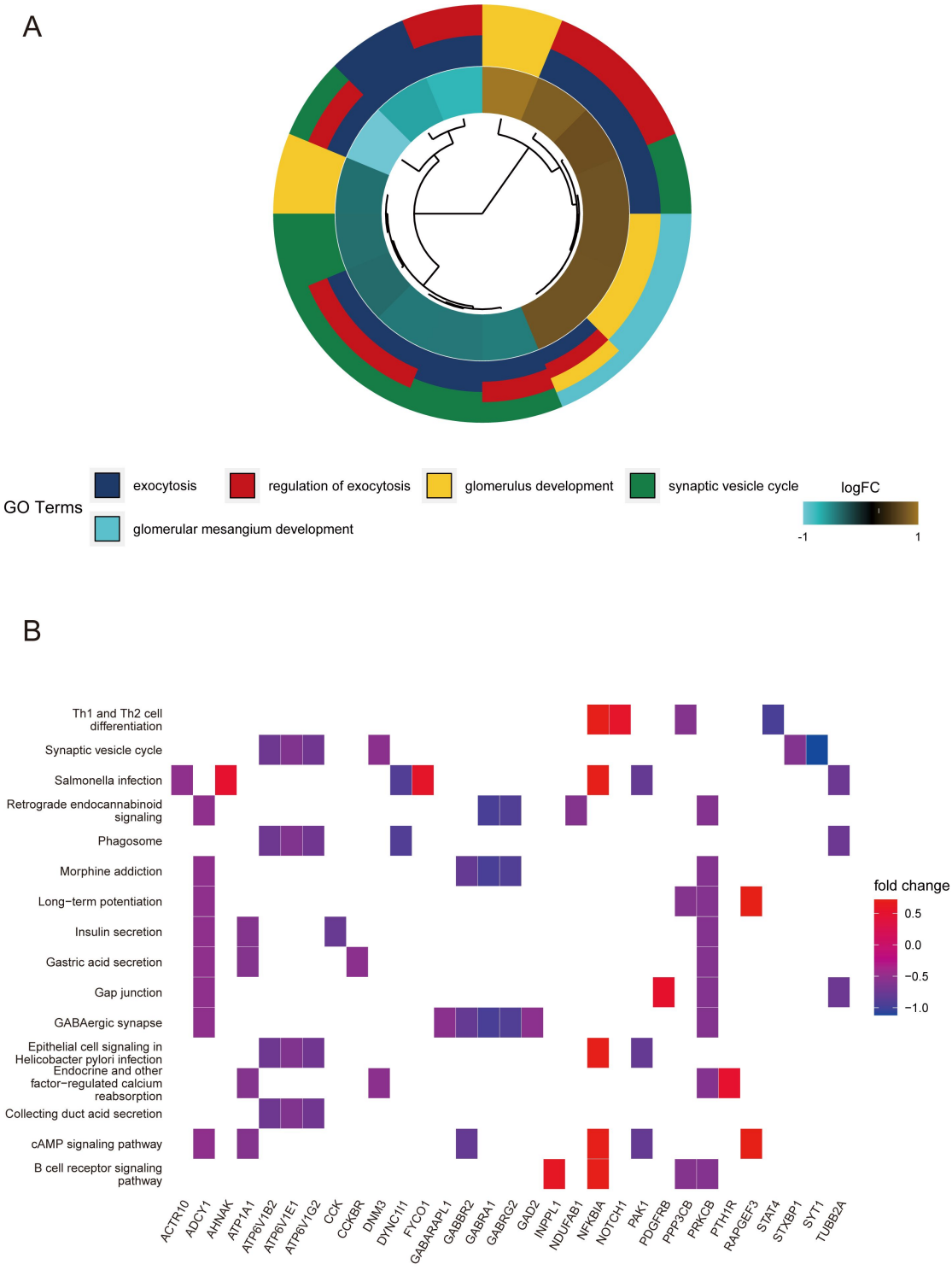

Fig. S2. Enrichment analysis of key gene functions. **(A)** 109 characteristic genes were performed with GO analysis. The inner circle with different color were indicated the multiple differences of the genes. The outer circle with different colors indicated the genes were enriched into different signal pathways. **(B)** KEGG functional analysis was also conducted on 109 characteristic genes,

with colors ranging from purple to red. The darker the color of the bar chart, the greater the multiplicity of differences in genes.

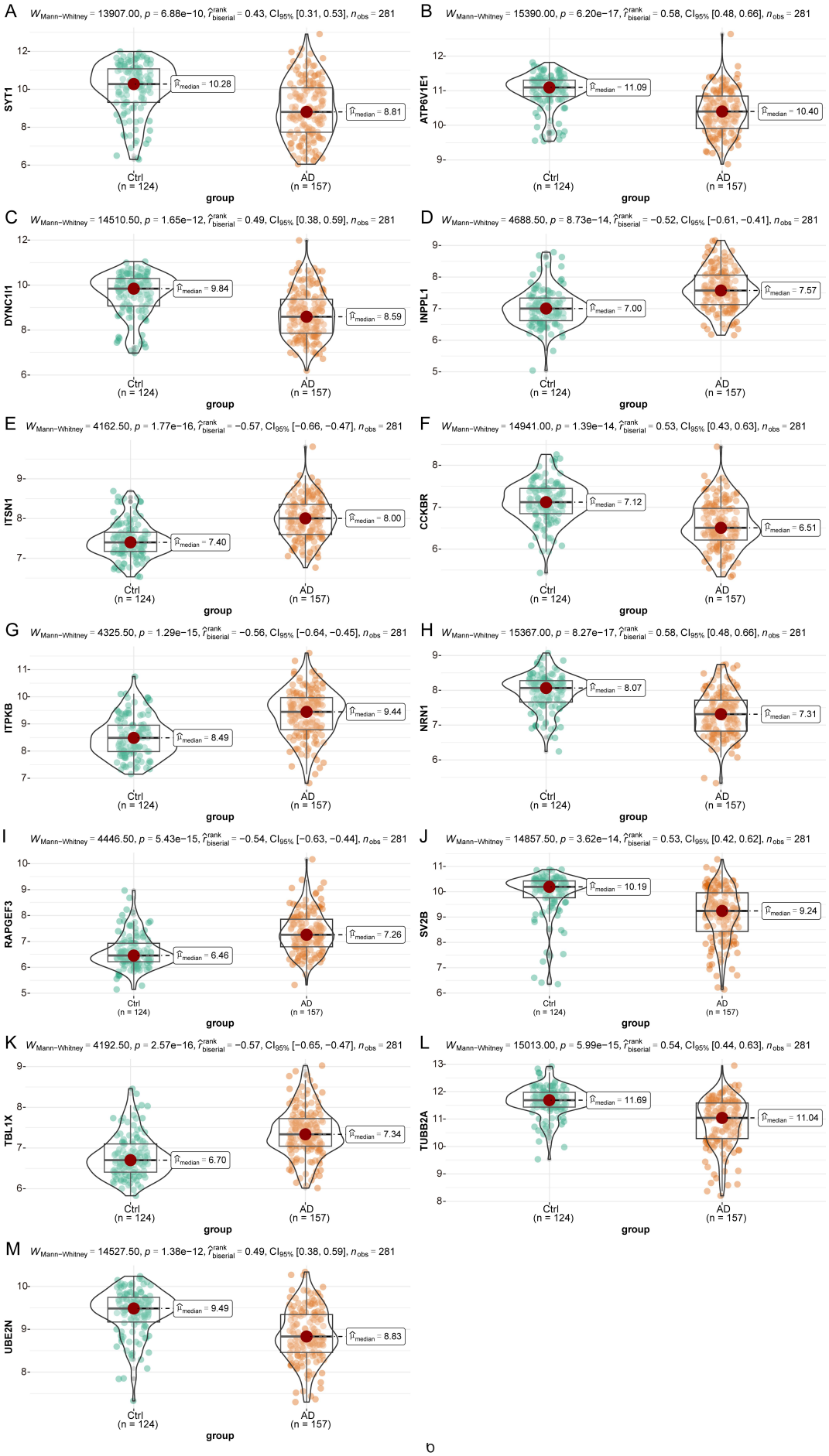
A
$W_{Mann-Whitney}$ = 13907.00, p = 6.88e−10, $\hat{r}^{rank}_{biserial}$ = 0.43, $CI_{95\%}$ [0.31, 0.53], n_{obs} = 281
SYT1
$\hat{\pi}_{median}$ = 10.28
$\hat{\pi}_{median}$ = 8.81
Ctrl (n = 124)
AD (n = 157)
group
B
$W_{Mann-Whitney}$ = 15390.00, p = 6.20e−17, $\hat{r}^{rank}_{biserial}$ = 0.58, $CI_{95\%}$ [0.48, 0.66], n_{obs} = 281
ATP6V1E1
$\hat{\pi}_{median}$ = 11.09
$\hat{\pi}_{median}$ = 10.40
Ctrl (n = 124)
AD (n = 157)
group
C
$W_{Mann-Whitney}$ = 14510.50, p = 1.65e−12, $\hat{r}^{rank}_{biserial}$ = 0.49, $CI_{95\%}$ [0.38, 0.59], n_{obs} = 281
DYNC1I1
$\hat{\pi}_{median}$ = 9.84
$\hat{\pi}_{median}$ = 8.59
Ctrl (n = 124)
AD (n = 157)
group
D
$W_{Mann-Whitney}$ = 4688.50, p = 8.73e−14, $\hat{r}^{rank}_{biserial}$ = −0.52, $CI_{95\%}$ [−0.61, −0.41], n_{obs} = 281
INPPL1
$\hat{\pi}_{median}$ = 7.00
$\hat{\pi}_{median}$ = 7.57
Ctrl (n = 124)
AD (n = 157)
group
E
$W_{Mann-Whitney}$ = 4162.50, p = 1.77e−16, $\hat{r}^{rank}_{biserial}$ = −0.57, $CI_{95\%}$ [−0.66, −0.47], n_{obs} = 281
ITSN1
$\hat{\pi}_{median}$ = 7.40
$\hat{\pi}_{median}$ = 8.00
Ctrl (n = 124)
AD (n = 157)
group
F
$W_{Mann-Whitney}$ = 14941.00, p = 1.39e−14, $\hat{r}^{rank}_{biserial}$ = 0.53, $CI_{95\%}$ [0.43, 0.63], n_{obs} = 281
CCKBR
$\hat{\pi}_{median}$ = 7.12
$\hat{\pi}_{median}$ = 6.51
Ctrl (n = 124)
AD (n = 157)
group
G
$W_{Mann-Whitney}$ = 4325.50, p = 1.29e−15, $\hat{r}^{rank}_{biserial}$ = −0.56, $CI_{95\%}$ [−0.64, −0.45], n_{obs} = 281
ITPKB
$\hat{\pi}_{median}$ = 8.49
$\hat{\pi}_{median}$ = 9.44
Ctrl (n = 124)
AD (n = 157)
group
H
$W_{Mann-Whitney}$ = 15367.00, p = 8.27e−17, $\hat{r}^{rank}_{biserial}$ = 0.58, $CI_{95\%}$ [0.48, 0.66], n_{obs} = 281
NRN1
$\hat{\pi}_{median}$ = 8.07
$\hat{\pi}_{median}$ = 7.31
Ctrl (n = 124)
AD (n = 157)
group
I
$W_{Mann-Whitney}$ = 4446.50, p = 5.43e−15, $\hat{r}^{rank}_{biserial}$ = −0.54, $CI_{95\%}$ [−0.63, −0.44], n_{obs} = 281
RAPGEF3
$\hat{\pi}_{median}$ = 6.46
$\hat{\pi}_{median}$ = 7.26
Ctrl (n = 124)
AD (n = 157)
group
J
$W_{Mann-Whitney}$ = 14857.50, p = 3.62e−14, $\hat{r}^{rank}_{biserial}$ = 0.53, $CI_{95\%}$ [0.42, 0.62], n_{obs} = 281
SV2B
$\hat{\pi}_{median}$ = 10.19
$\hat{\pi}_{median}$ = 9.24
Ctrl (n = 124)
AD (n = 157)
group
K
$W_{Mann-Whitney}$ = 4192.50, p = 2.57e−16, $\hat{r}^{rank}_{biserial}$ = −0.57, $CI_{95\%}$ [−0.65, −0.47], n_{obs} = 281
TBL1X
$\hat{\pi}_{median}$ = 6.70
$\hat{\pi}_{median}$ = 7.34
Ctrl (n = 124)
AD (n = 157)
group
L
$W_{Mann-Whitney}$ = 15013.00, p = 5.99e−15, $\hat{r}^{rank}_{biserial}$ = 0.54, $CI_{95\%}$ [0.44, 0.63], n_{obs} = 281
TUBB2A
$\hat{\pi}_{median}$ = 11.69
$\hat{\pi}_{median}$ = 11.04
Ctrl (n = 124)
AD (n = 157)
group
M
$W_{Mann-Whitney}$ = 14527.50, p = 1.38e−12, $\hat{r}^{rank}_{biserial}$ = 0.49, $CI_{95\%}$ [0.38, 0.59], n_{obs} = 281
UBE2N
$\hat{\pi}_{median}$ = 9.49
$\hat{\pi}_{median}$ = 8.83
Ctrl (n = 124)
AD (n = 157)
group

Fig. S3. Expression of characterized genes in the training set. **(A-M)** Feature gene expression in the GSE5281, GSE84422 and GSE132903 datasets. **(A)** SYT1 in AD was significantly lower compared to control, p = 6.88e-10. **(B)** ATP6V1E1 in AD was significantly lower compared to control, p = 6.20e-17. **(C)** DYNC1I1 in AD was significantly lower compared to control, p = 1.65e-12. **(D)** INPPL1 in AD was significantly elevated compared to controls, p = 8.73e-14. **(E)** ITSN1 in AD was significantly elevated compared to controls, p = 1.77e-16. **(F)** CCKBR in AD was significantly lower compared to the control group, p=1.39e-14. **(G)** ITPKB in AD was significantly higher compared to controls, p = 1.29e-15. **(H)** NRN1 was significantly lower compared to controls, p = 8.27e-17. **(I)** RAPGEF3 in AD was significantly elevated compared to controls, p = 5.43e-15. **(J)** SV2B was significantly lower compared to controls, p=3.62e-14. **(K)** TBL1X in AD was significantly higher compared to controls, p = 2.57e-16. **(L)** TUBB2A in AD was significantly lower compared to control, p = 5.99e-15. **(M)** UBE2N in AD was significantly lower compared to control, p = 1.38e-12.

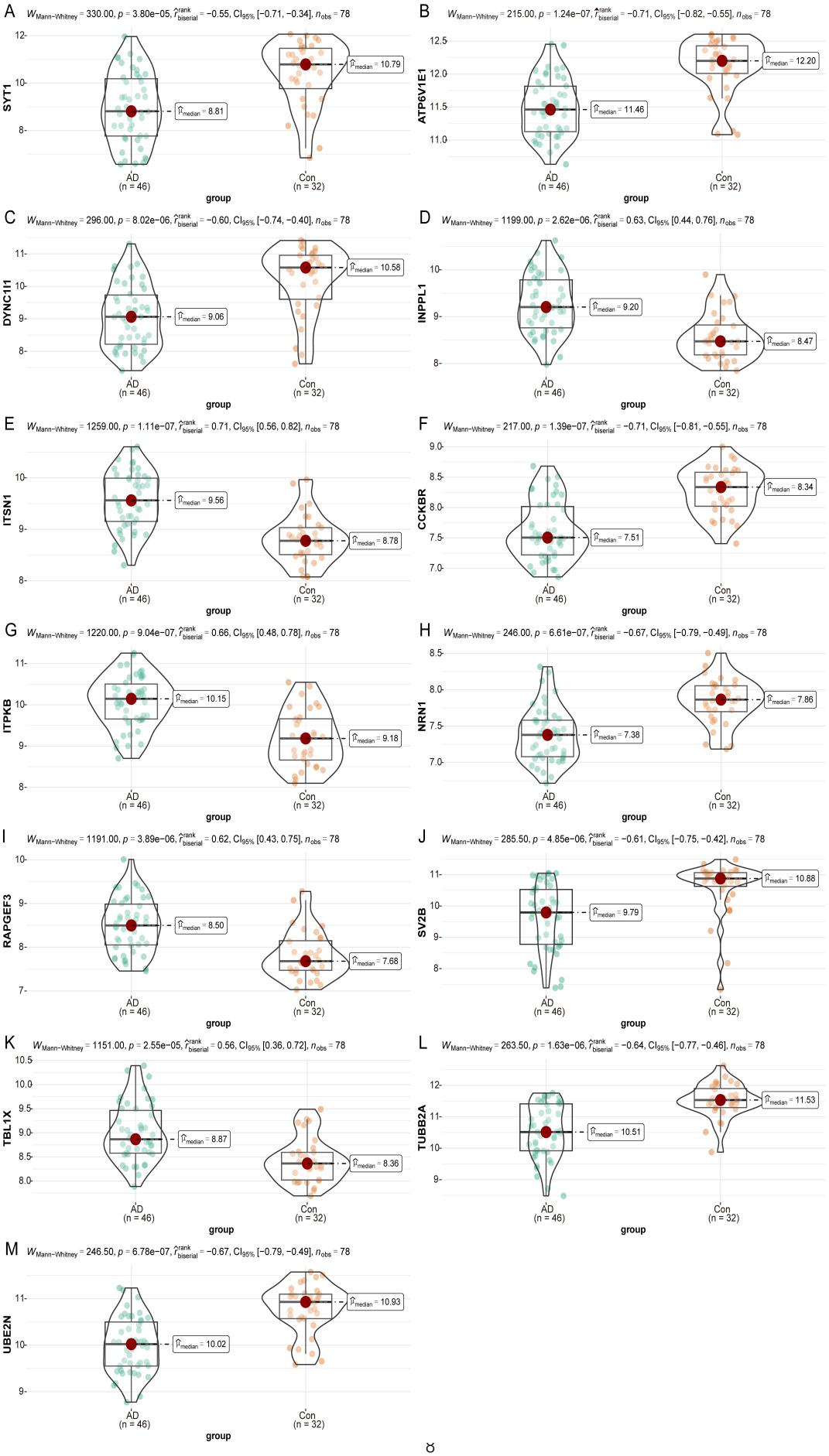
A
$W_{Mann-Whitney}$ = 330.00, p = 3.80e−05, $\hat{r}^{rank}_{biserial}$ = −0.55, $CI_{95\%}$ [−0.71, −0.34], n_{obs} = 78
SYT1
$\hat{\pi}_{median}$ = 8.81
$\hat{\pi}_{median}$ = 10.79
AD (n = 46)
Con (n = 32)
group
B
$W_{Mann-Whitney}$ = 215.00, p = 1.24e−07, $\hat{r}^{rank}_{biserial}$ = −0.71, $CI_{95\%}$ [−0.82, −0.55], n_{obs} = 78
ATP6V1E1
$\hat{\pi}_{median}$ = 11.46
$\hat{\pi}_{median}$ = 12.20
AD (n = 46)
Con (n = 32)
group
C
$W_{Mann-Whitney}$ = 296.00, p = 8.02e−06, $\hat{r}^{rank}_{biserial}$ = −0.60, $CI_{95\%}$ [−0.74, −0.40], n_{obs} = 78
DYNC1I1
$\hat{\pi}_{median}$ = 9.06
$\hat{\pi}_{median}$ = 10.58
AD (n = 46)
Con (n = 32)
group
D
$W_{Mann-Whitney}$ = 1199.00, p = 2.62e−06, $\hat{r}^{rank}_{biserial}$ = 0.63, $CI_{95\%}$ [0.44, 0.76], n_{obs} = 78
INPPL1
$\hat{\pi}_{median}$ = 9.20
$\hat{\pi}_{median}$ = 8.47
AD (n = 46)
Con (n = 32)
group
E
$W_{Mann-Whitney}$ = 1259.00, p = 1.11e−07, $\hat{r}^{rank}_{biserial}$ = 0.71, $CI_{95\%}$ [0.56, 0.82], n_{obs} = 78
ITSN1
$\hat{\pi}_{median}$ = 9.56
$\hat{\pi}_{median}$ = 8.78
AD (n = 46)
Con (n = 32)
group
F
$W_{Mann-Whitney}$ = 217.00, p = 1.39e−07, $\hat{r}^{rank}_{biserial}$ = −0.71, $CI_{95\%}$ [−0.81, −0.55], n_{obs} = 78
CCKBR
$\hat{\pi}_{median}$ = 7.51
$\hat{\pi}_{median}$ = 8.34
AD (n = 46)
Con (n = 32)
group
G
$W_{Mann-Whitney}$ = 1220.00, p = 9.04e−07, $\hat{r}^{rank}_{biserial}$ = 0.66, $CI_{95\%}$ [0.48, 0.78], n_{obs} = 78
ITPKB
$\hat{\pi}_{median}$ = 10.15
$\hat{\pi}_{median}$ = 9.18
AD (n = 46)
Con (n = 32)
group
H
$W_{Mann-Whitney}$ = 246.00, p = 6.61e−07, $\hat{r}^{rank}_{biserial}$ = −0.67, $CI_{95\%}$ [−0.79, −0.49], n_{obs} = 78
NRN1
$\hat{\pi}_{median}$ = 7.38
$\hat{\pi}_{median}$ = 7.86
AD (n = 46)
Con (n = 32)
group
I
$W_{Mann-Whitney}$ = 1191.00, p = 3.89e−06, $\hat{r}^{rank}_{biserial}$ = 0.62, $CI_{95\%}$ [0.43, 0.75], n_{obs} = 78
RAPGEF3
$\hat{\pi}_{median}$ = 8.50
$\hat{\pi}_{median}$ = 7.68
AD (n = 46)
Con (n = 32)
group
J
$W_{Mann-Whitney}$ = 285.50, p = 4.85e−06, $\hat{r}^{rank}_{biserial}$ = −0.61, $CI_{95\%}$ [−0.75, −0.42], n_{obs} = 78
SV2B
$\hat{\pi}_{median}$ = 9.79
$\hat{\pi}_{median}$ = 10.88
AD (n = 46)
Con (n = 32)
group
K
$W_{Mann-Whitney}$ = 1151.00, p = 2.55e−05, $\hat{r}^{rank}_{biserial}$ = 0.56, $CI_{95\%}$ [0.36, 0.72], n_{obs} = 78
TBL1X
$\hat{\pi}_{median}$ = 8.87
$\hat{\pi}_{median}$ = 8.36
AD (n = 46)
Con (n = 32)
group
L
$W_{Mann-Whitney}$ = 263.50, p = 1.63e−06, $\hat{r}^{rank}_{biserial}$ = −0.64, $CI_{95\%}$ [−0.77, −0.46], n_{obs} = 78
TUBB2A
$\hat{\pi}_{median}$ = 10.51
$\hat{\pi}_{median}$ = 11.53
AD (n = 46)
Con (n = 32)
group
M
$W_{Mann-Whitney}$ = 246.50, p = 6.78e−07, $\hat{r}^{rank}_{biserial}$ = −0.67, $CI_{95\%}$ [−0.79, −0.49], n_{obs} = 78
UBE2N
$\hat{\pi}_{median}$ = 10.02
$\hat{\pi}_{median}$ = 10.93
AD (n = 46)
Con (n = 32)
group

Fig. S4. Expression of characterized genes in the validation set. (A-M) Feature gene expression in the GSE109887 dataset. **(A)** SYT1 in AD was significantly lower than that in the control group, p = 3.80e-05. **(B)** Compared to the control, ATP6V1E1 was significantly lower in the AD group, p = 1.24e-07. **(C)** Compared to the control, DYNC1I1 was significantly lower in the AD group, p = 8.02e-06. **(D)** INPPL1 in AD was significantly elevated compared to the control group, p = 2.62e-06. **(E)** ITSN1 in AD was significantly elevated compared to controls, p = 1.11e-07. **(F)** CCKBR was significantly decreased in AD compared with the control group, p = 1.39e-07. **(G)** ITPKB in AD was significantly higher than that in the control group, p = 9.04e-07. **(H)** NRN1 in AD was significantly lower than that in the control group, p = 6.61e-07. **(I)** RAPGEF3 in AD was significantly elevated compared to controls, p = 3.89e-06. **(J)** SV2B was significantly decreased in AD compared to controls, p = 4.85e-06. **(K)** TBL1X in AD was significantly higher than that in the control group, p = 2.55e-05. **(L)** TUBB2A in AD was significantly lower than that in the control group, p = 1.63e-06. **(M)** UBE2N in AD was significantly lower than that in the control group, p = 6.78e-07.

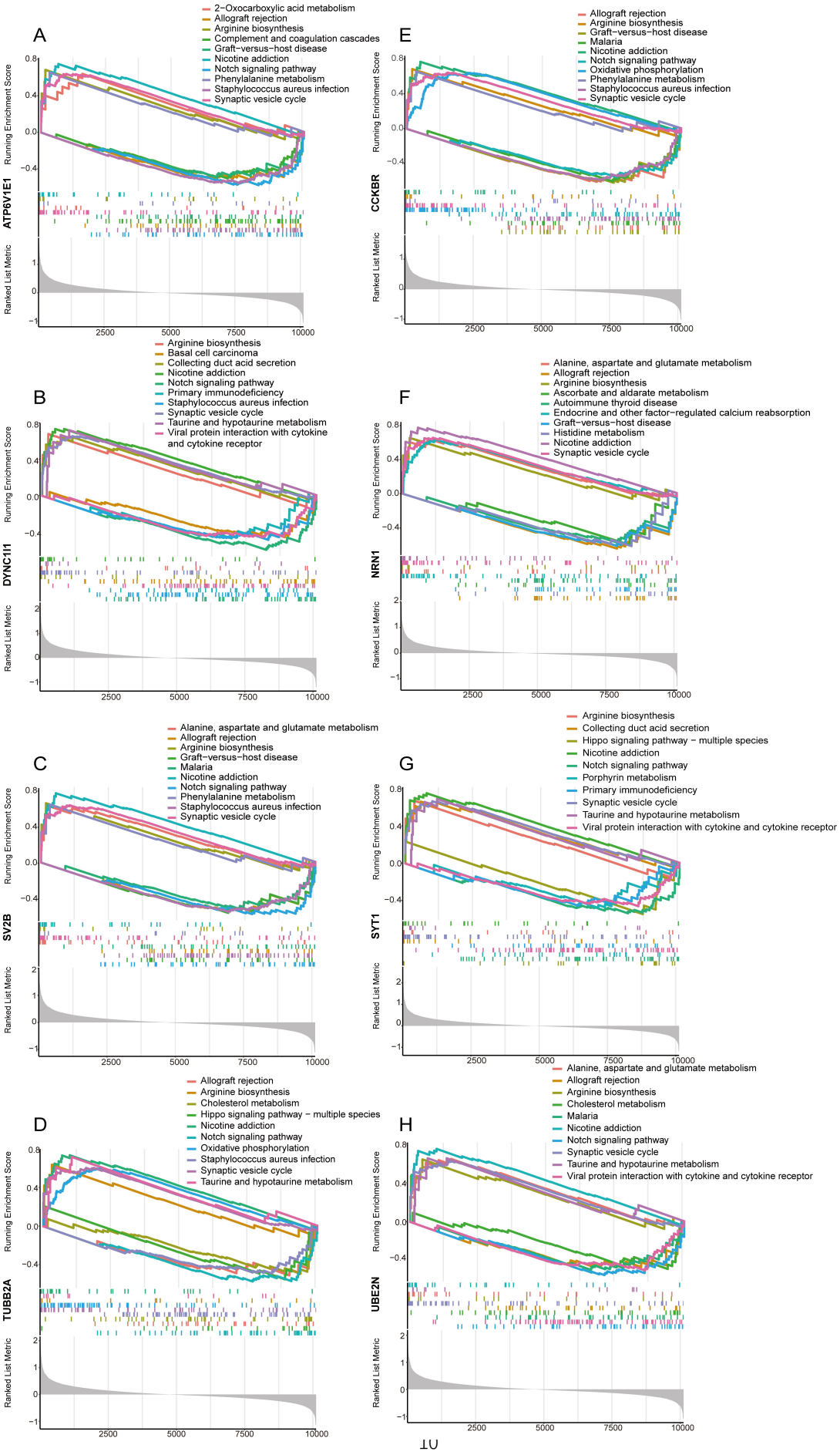
A
2−Oxocarboxylic acid metabolism
Allograft rejection
Arginine biosynthesis
Complement and coagulation cascades
Graft−versus−host disease
Nicotine addiction
Notch signaling pathway
Phenylalanine metabolism
Staphylococcus aureus infection
Synaptic vesicle cycle
ATP6V1E1
B
Arginine biosynthesis
Basal cell carcinoma
Collecting duct acid secretion
Nicotine addiction
Notch signaling pathway
Primary immunodeficiency
Staphylococcus aureus infection
Synaptic vesicle cycle
Taurine and hypotaurine metabolism
Viral protein interaction with cytokine and cytokine receptor
DYNC1I1
C
Alanine, aspartate and glutamate metabolism
Allograft rejection
Arginine biosynthesis
Graft−versus−host disease
Malaria
Nicotine addiction
Notch signaling pathway
Phenylalanine metabolism
Staphylococcus aureus infection
Synaptic vesicle cycle
SV2B
D
Allograft rejection
Arginine biosynthesis
Cholesterol metabolism
Hippo signaling pathway − multiple species
Nicotine addiction
Notch signaling pathway
Oxidative phosphorylation
Staphylococcus aureus infection
Synaptic vesicle cycle
Taurine and hypotaurine metabolism
TUBB2A
E
Allograft rejection
Arginine biosynthesis
Graft−versus−host disease
Malaria
Nicotine addiction
Notch signaling pathway
Oxidative phosphorylation
Phenylalanine metabolism
Staphylococcus aureus infection
Synaptic vesicle cycle
CCKBR
F
Alanine, aspartate and glutamate metabolism
Allograft rejection
Arginine biosynthesis
Ascorbate and aldarate metabolism
Autoimmune thyroid disease
Endocrine and other factor−regulated calcium reabsorption
Graft−versus−host disease
Histidine metabolism
Nicotine addiction
Synaptic vesicle cycle
NRN1
G
Arginine biosynthesis
Collecting duct acid secretion
Hippo signaling pathway − multiple species
Nicotine addiction
Notch signaling pathway
Porphyrin metabolism
Primary immunodeficiency
Synaptic vesicle cycle
Taurine and hypotaurine metabolism
Viral protein interaction with cytokine and cytokine receptor
SYT1
H
Alanine, aspartate and glutamate metabolism
Allograft rejection
Arginine biosynthesis
Cholesterol metabolism
Malaria
Nicotine addiction
Notch signaling pathway
Synaptic vesicle cycle
Taurine and hypotaurine metabolism
Viral protein interaction with cytokine and cytokine receptor
UBE2N
Running Enrichment Score
Ranked List Metric
2500
5000
7500
10000

Fig. S5. GSEA analysis of characteristic genes. Based on the median value of each gene, genes were divided into high expression and low expression in AD data. The differential expression genes were analyzed by GSEA. Each group displayed 10 pathways, with the most significantly upregulated 5 pathways and the most significantly downregulated 5 pathways. **(A)** GSEA analysis of ATP6V1E1. **(B)** GSEA analysis of DYNC1I1. **(C)** GSEA analysis of SV2B. **(D)** GSEA analysis of TUBB2A. **(E)** GSEA analysis of CCKBR. **(F)** GSEA analysis of NRN1. **(G)** GSEA analysis of SYT1. **(H)** GSEA analysis of UBE2N.

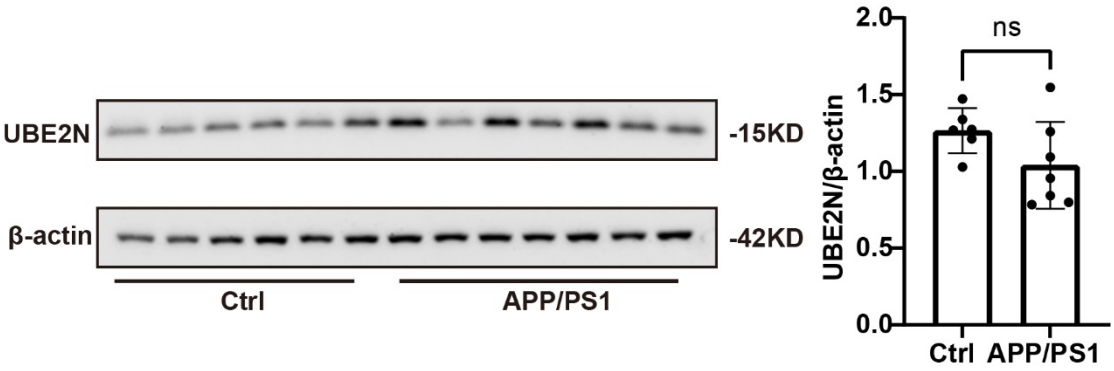

Fig. S6. UBE2N expression in the temporal cortex of control (n=6) and APP/PS1 mice (n = 7).

1 Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics: a journal of integrative biology* **16**, 284-287 (2012).

2 Ogata, H. *et al.* KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic acids research* **27**, 29-34 (1999).