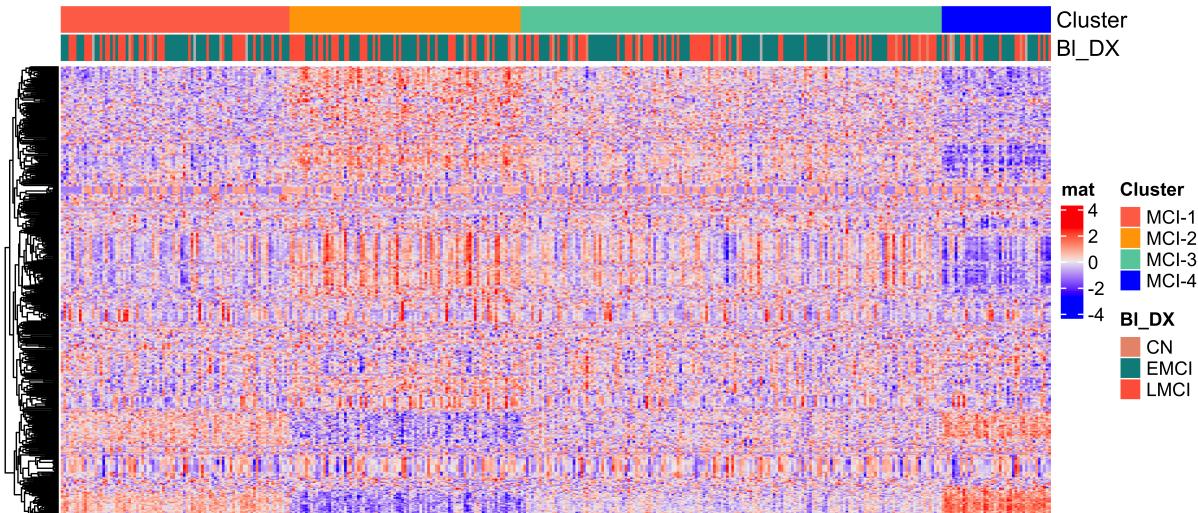
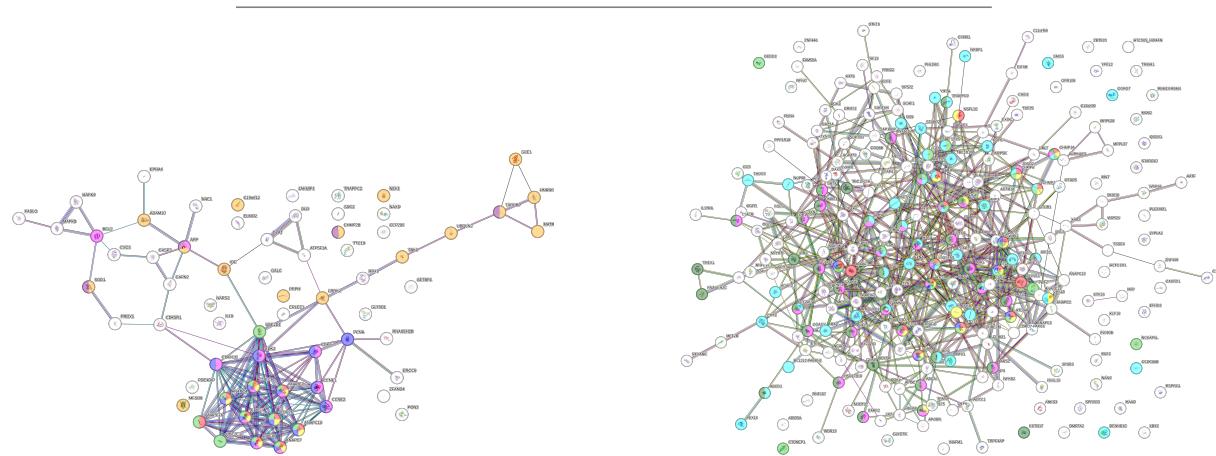


# Additional Information and Supporting Materials

## Stratifying Mild Cognitive Impairment Patients via Embedded Transcriptomic Data and Clustering Analysis



**Heatmap of MCI patient clusters transcriptomic profile.** The heatmap displays the expression levels of the 445 selected genes across four identified clusters by QUANTILE NORMALIZATION and *StdDevFilter* of 381 MCI patients. Each column represents a patient, with cluster membership indicated by the color-coded bar at the top. The number of patients in each cluster is also reported. The annotation bar immediately below denotes the baseline diagnosis (Bl\_DX) for each individual. Genes are represented by rows and are ordered according to hierarchical clustering. The number of genes grouped within each gene cluster is indicated in parentheses.



**PPI network and functional enrichment of  $g_1$  and  $g_2$  genes.**

## Clinical and Laboratory Variables Used in This Study

Variable	Description	Information Type
VISCODE2	Visit	Patient info
SubjectID	SubjectID	Patient info
DX	Diagnosis at the visit	Patient info
B1_DX	Baseline Diagnosis	Patient info
AGE	Age	Patient info
PTGENDER	Gender	Patient info
PTEDUCAT	Education	Patient info
PTETHCAT	Ethnicity	Patient info
PTRACCAT	Race	Patient info
PTMARRY	Marital	Patient info
APOE4	Number of APOEe4 alleles	Genetic
FDG	FDG-PET metaROI	Imaging
PIB	Average PIB SUVR of frontal cortex, anterior cingulate, pre-cuneus cortex, and parietal cortex	Imaging
AV45	PET AV45 ratio (cortical grey matter/whole cerebellum)	Imaging
FBB	FBB ratio (cortical grey matter/whole cerebellum)	Imaging
ABETA	CSF ABETA	CSF Biomarker
TAU	CSF TAU	CSF Biomarker
PTAU	CSF PTAU	CSF Biomarker
CDRSB	Clinical Dementia Rating Scale - Sum of Boxes	Cognitive
ADAS11	Alzheimer's Disease Assessment Scale (ADAS11)	Cognitive
ADAS13	Alzheimer's Disease Assessment Scale (ADAS13)	Cognitive
ADASQ4	Alzheimer's Disease Assessment Scale (ADASQ4)	Cognitive
MMSE	Mini Mental State Examination	Cognitive
RAVLT.immediate	Rey Auditory Verbal Learning Test	Cognitive
RAVLT.learning	Rey Auditory Verbal Learning Test	Cognitive
RAVLT.forgetting	Rey Auditory Verbal Learning Test	Cognitive
RAVLT.perc.forgetting	Rey Auditory Verbal Learning Test	Cognitive
LDELTOTAL	Logical Memory - Delayed Recall	Cognitive
DIGITSCOR	Digit Symbol Substitution	Cognitive
TRABSCOR	Trail Making Test	Cognitive
FAQ	Functional Activity Questionnaire	Cognitive
MOCA	Montreal Cognitive Assessment	Cognitive
EcogPtMem	Everyday Cognition – Patient Memory	Cognitive
EcogPtLang	Everyday Cognition – Patient Language	Cognitive
EcogPtVisspat	Everyday Cognition – Patient Visuospatial	Cognitive
EcogPtPlan	Everyday Cognition – Patient Planning	Cognitive
EcogPtOrgan	Everyday Cognition – Patient Organization	Cognitive
EcogPtDivatt	Everyday Cognition – Patient Divided Attention	Cognitive
EcogPtTotal	Everyday Cognition – Patient Total Score	Cognitive
EcogSPMem	Everyday Cognition – Study Partner Memory	Cognitive
EcogSPLang	Everyday Cognition – Study Partner Language	Cognitive
EcogSPVisspat	Everyday Cognition – Study Partner Visuospatial	Cognitive
EcogSPPlan	Everyday Cognition – Study Partner Planning	Cognitive
EcogSPOrgan	Everyday Cognition – Study Partner Organization	Cognitive
EcogSPDivatt	Everyday Cognition – Study Partner Divided Attention	Cognitive
EcogSPTotal	Everyday Cognition – Study Partner Total Score	Cognitive
FLDSTRENG	MRI Field Strength	MRI info
FSVERSION	FreeSurfer Software Version	MRI info
IMAGEUID	LONI Image ID	MRI info
Ventricles	UCSF Ventricles	Imaging
Hippocampus	UCSF Hippocampus	Imaging
WholeBrain	UCSF WholeBrain	Imaging
Entorhinal	UCSF Entorhinal	Imaging
Fusiform	UCSF Fusiform	Imaging
MidTemp	UCSF Middle temporal gyrus	Imaging
ICV	Intracerebral volume	Imaging
mPACCdigit	Preclinical Alzheimer's Cognitive Composite	Cognitive

## Supervised Gene Sets from Multiple Sources Related to Alzheimer's and Neurodegenerative Diseases

DisGeNET	Emon et al. 2020	DisGeNET + Emon et al. 2020	Human Phenotype Ontology (HPO)	DisGeNET + Emon et al. 2020+HPO	Reactome R-HSA-9675126
A2M, ACE, APOC1, ATP5A1, BACE1, BAX, BDNF, BIN1, CD2AP, CLU, CR1, EPHA1, GRN, HFE, IDE, IGF1R, IGF2, IL1B, INSR, LEP, MPO, NOS3, NPY, PLAU, SORL1	AKT1, CD33, CDK5, CRH, CYCS, DNMT1, DYRK1A, GDNF, IL18, IL6, MAPK8, MAPK9, MTHFR, NLRP3, PRKACA, SDC2, TP53, H3F3A, UCN	BCL2, CASP3, GSK3B, PICALM, TOMM40	AARS2, ABCD1, ANG, ATXN2, ATXN3, C19ORF12, CCNF, CHCHD10, COASY, COQ2, CSF1R, DAO, DCTN1, EPHA4, ERBB4, FOLR1, FUS, GUSB, HNRNPA1, HSD17B10, HTT, IDUA, MATR3, MFSD8, NAE1, NARS2, NEFH, OPTN, PANK2, PCNA, PDE8B, PFN1, PLA2G6, PON1, PON2, PON3, PPARGC1A, SLC2A3, SOD1, SPTBN1, SQSTM1, TAF15, TBK1, TREX1, TXN2, UBQLN2, UBTF, UCHL1, UNC13A, VCP, WDR45, ADAR, ANXA11, C21ORF2, CARKD, DLAT, ERCC6, FA2H, FIG4, FTL, GALC, GLE1, GLT8D1, GM2A, IDS, IFIH1, LSM11, LYST, NBN, NEK1, PRPH, RNASEH2A, RNASEH2B, RNASEH2C, SAMHD1, TANGO2, TTC19, VAPB, ZBTB20	CHMP2B, MAPT, TARDBP, TREM2	CCND1, CCNE2, CDK4, CDK6, CDKN1A, CDKN1B, DAXX, E2F1, RB1, TFDP1, ANAPC1, ANAPC10, ANAPC11, ANAPC15, ANAPC16, ANAPC2, ANAPC4, ANAPC5, ANAPC7, ATRX, CCND2, CCND3, CCNE1, CDC16, CDC23, CDC26, CDC27, CDK2, CDKN1C, E2F2, E2F3, FZR1, SKP2, TFDP2, UBE2C, UBE2D1, UBE2E1, UBE2S
Reactome R-HSA-9565265	Reactome R-HSA-8862803	KEGG Disease DB "Alzheimer"	DisGeNET + KEGG Disease DB "Alzheimer"	Disgenet + Reactome R-HSA-8862803 + KEGG Disease DB "Alzheimer"	Hubs from HENA network
OGG1	BCL2L11, CAPN1, CAPN2, CAST, CDC25B, CDC25C, CDK5R1, FASLG, FOXO3, GOLGA2, JUN, LMNA, LMNB1, PRDX1, PRDX2, SOD2, CAPNS1, CAPNS2, CDC25A, YWHAE	DLD, GBA, LRRK2, OGDH, SNCB	ABC7, ADAM10, PSEN1, PSEN2, SNCA	APP	AGAP2, GRIA1, KALRN, MLKL, PLXNA4, UBN1, ALDH1L2, C3ORF55, C9ORF172, CACNA1D, DEDD, DMBT1, DMRTA2, DNAH14, DPF1, ELMO2, ERLEC1, FAM168A, FARP2, ILDR2, JAK-M1P1, KIAA1045, KIAA1456, KIAA1731, LINC00473, LINC00598, CT45A3, LRRC14B, MBD1, MBOAT1, MUC20, NEK10, PAK6, PDE4DIP, PEAR1, SETBP1, SH3RF3, ST6GALNAC3, SYNE3, SYNGAP1, THRIB, TRAPPC2, ZBTB4, ZFAND4, ZNF391, ZNF771
DisGeNET + Emon et al. 2020 + KEGG Disease DB "Alzheimer"					APOE

**Unsupervised Gene Set selected by Z-score Normalization and MedianCoExpFilter from ADNI collections.** The genes are listed in alphabetical order.

AATF	CYHR1	MED25	SCAND1
ACO2	DCTN1	MEF2B / MEF2BNB-MEF2B	SERTAD1
ACTN4	DEDD2	MIA-RAB4B/RAB4B	SMG5
ACTR1A	DEF6	MIIP	SPRYD3
ADCK4	DENND1C	MLF2	SPSB3
ADRM1	DGKZ	MNT	STARD10
AGPAT2	DHX16	MOB2	STK16
ANKHD1-EIF4EBP3/ EIF4EBP3	DOK1	MPST	STK19
ANKS3	DVL3	MRPL28	STX5
ANXA11	EFHD2	MRPL37	STXBP2
AP1B1	EIF4H	MTX1	SUPT5H
AP2S1	EPN1	NANS	SUPT6H
AP5Z1	ETFB	NCAPH2	SYMPK
APOBR	EXOC3	NCKAP5L	TAZ
ARAF	FAM50A	NELFE	TBC1D10B
ARF3	FBXL15	NFKB2	TBC1D22A
ARHGAP4	FCHO1	NRBPI	TCF25
ARID5A	FERMT3	NSFL1C	TFEB
ARPC1B	FLJ23467*/ SCNM1	NUDT22	THOC5
ATG13	FLNA	NUP98	TLN1
ATG7	G6PD	OGFRC	TMEM175
ATG9A	GABARAP	OS9	TOP1 / TOP1P1
ATP6V0D1	GBA2	OTUB1	TRAPPC9
ATXN7L3	GLYCTK	OTUD5	TRIM41
AURKAIP1	GNAI2	PEX16	TRMU
BAX	GPR108	PFKL	TRPC4AP
BAZ2A	GPS2	PGLS	TSC2
BCKDHA	GSK3A	PHLDDB1	TSSC4
BCKDK	HAAO	PLEKHO1	TST
BCL2L2-PABPN1 / PABPN1	HCFC1R1	PLIN4	TUBB4B
BRMS1	HDLBP	PPP1R18	TWF2
C11ORF68	IDH2	PPP4C	UBA1
C12ORF10	IL17RA	PRKD2	UBA7
C15ORF39	INO80B	PRR24	UBE2M
CACFD1	INO80B-WBP1 / INO80B	PSMD4	UBL7
CAPN1	INPP5K	PYCARD	VPS18
CCDC88B	JOSD2	QSOX1	VPS72
CCS	KAT8	RAB1B	WAS
CDC73	KCTD17	RABAC1	WDR13
CHID1	KIF1C	RALBP1	WDR25
CHMP1A	KLF16	RBM14-RBM4 / RBM4	WDR46
CHMP2A	LAMTOR2	RFNG	WDTC1
CHMP6	LBX2	RFXANK	WIPI2
CLTB	LOC101928750 / NDOR1	RHOG	XAB2
COPE	LOC101930060 / LOC728138 / KIAA2013	RNF167	YIF1A
CORO7	LYPLA2	RNPEPL1	YPEL2
CORO7-PAM16 / CORO7	MAP2K2	RRAGC	ZFYVE1
CPT2	MAP3K11	RSFRY1	ZNF408
CTDNEP1	MAPK3	SAP30BP	ZNF446
CTSD	MED16	SCAF1	ZYX

### Differences in gene expression levels among patient groups

Groups	Gene	logFC	AveExpr	t	P.Value	adj.P.Val
MCI-1 Vs CN	CASP8AP2	1.06	4.17	11.64	6.84E-29	1.75E-25
	ZNF92	1.01	4.14	11.58	1.22E-28	2.73E-25
	CCNT2	1.01	4.32	11.07	1.82E-26	1.13E-23
	KIAA1033	1.07	6.89	11.07	1.85E-26	1.13E-23
	LRRC40	1.05	5.83	10.85	1.49E-25	5.76E-23
	SGTB	1.01	5.33	10.71	5.53E-25	1.52E-22
	SLC35F5	1.13	4.87	10.65	9.47E-25	2.32E-22
	FNDC3A	1.07	5.53	10.31	2.20E-23	2.86E-21
	SACS	1.14	5.34	10.31	2.23E-23	2.88E-21
	EEA1	1.11	4.63	10.12	1.27E-22	1.21E-20
MCI-1 Vs AD	NEK7	1.17	6.33	10.10	1.46E-22	1.35E-20
	TMEM55A	1.08	5.72	10.01	3.13E-22	2.51E-20
	NUDT21	1.09	6.20	10.52	3.26E-24	2.85E-21
	ZNF92	1.01	4.14	10.34	1.62E-23	8.79E-21
	CASP8AP2	1.03	4.17	9.97	4.59E-22	1.26E-19
	LRRC40	1.07	5.83	9.76	2.93E-21	5.07E-19
	ABCB10	1.06	5.79	9.74	3.48E-21	5.74E-19
	SACS	1.21	5.34	9.67	6.37E-21	9.48E-19
	TRPC1	1.01	3.21	9.55	1.83E-20	2.23E-18
MCI-1 Vs MCI-2	SLC35F5	1.00	4.88	12.60	1.18E-30	3.01E-28
MCI-2 Vs MCI-3	SACS	1.07	5.36	16.29	1.21E-45	8.12E-43
	FNDC3A	1.04	5.55	16.19	3.03E-45	1.90E-42
	PTBP2	1.05	5.26	15.08	1.22E-40	3.37E-38
	EEA1	1.03	4.67	14.15	7.79E-37	1.10E-34
	NEK7	1.10	6.37	13.45	5.33E-34	4.87E-32

### Differences in gene expression levels among patient MCI-1 and MCI-2 groups

Gene	logFC	AveExpr	t	P.Value	adj.P.Val
BZW1	1.25	7.68	28.20	2.81E-95	9.43E-92
CBFB	1.72	6.67	27.81	1.00E-93	2.87E-90
CCDC88B	-1.13	6.89	-26.24	2.00E-87	2.51E-84
FAM13B	1.02	7.61	26.21	2.72E-87	3.22E-84
N4BP2L1	1.39	7.46	26.11	6.96E-87	7.76E-84
CLINT1	1.04	7.17	26.10	7.74E-87	8.19E-84
SGPP1	1.50	5.15	26.09	8.54E-87	8.58E-84
TCEA1	1.26	7.16	25.87	6.60E-86	5.53E-83
SACS	2.05	5.36	25.83	9.73E-86	7.82E-83
FAM208B	1.26	6.70	25.33	1.02E-83	7.08E-81
SRSF11	1.10	7.29	25.28	1.77E-83	1.18E-80
BRWD1	1.10	5.07	24.98	2.98E-82	1.71E-79
FNDC3A	1.91	5.55	24.69	4.83E-81	2.42E-78
NUDT21	1.71	6.22	24.56	1.53E-80	7.30E-78
CAAP1	1.57	5.21	24.40	7.44E-80	3.18E-77
HIAT1	1.31	6.87	24.13	9.97E-79	4.01E-76
FBXL3	1.48	6.42	24.12	1.06E-78	4.16E-76
PRNP	1.39	4.64	24.10	1.30E-78	5.01E-76
PURA	1.02	7.93	24.02	2.84E-78	1.07E-75
TMEM209	1.07	4.45	23.87	1.21E-77	4.28E-75
SLC38A1	1.12	8.07	23.40	1.07E-75	3.07E-73
SLC30A7	1.65	6.15	23.39	1.14E-75	3.22E-73
CASP8AP2	1.69	4.20	23.38	1.30E-75	3.63E-73
RNF111	1.26	6.84	23.36	1.51E-75	4.10E-73
ATP11C	1.23	6.79	23.26	3.96E-75	1.05E-72
NUP54	1.17	5.19	23.25	4.62E-75	1.17E-72
ADNP	1.01	8.25	23.21	6.34E-75	1.56E-72
PPIL4	1.23	7.27	23.17	9.98E-75	2.36E-72
SLC35F5	1.96	4.88	23.13	1.42E-74	3.27E-72
CREBF	1.04	5.57	23.00	5.20E-74	1.17E-71
VPS36	1.29	6.62	22.87	1.70E-73	3.53E-71
BNIP2	1.42	7.38	22.84	2.40E-73	4.93E-71
PCNXL4	1.28	6.94	22.83	2.50E-73	5.07E-71
LRRC40	1.77	5.87	22.82	2.72E-73	5.47E-71
PTBP2	1.91	5.26	22.79	3.90E-73	7.69E-71
C2CD5	1.28	6.28	22.74	6.23E-73	1.20E-70
ARL6IP6	1.18	7.90	22.72	7.79E-73	1.46E-70
PGRMC2	1.06	5.93	22.66	1.37E-72	2.49E-70

Table continued on next page

Continued from previous page

Gene	logFC	AveExpr	t	P.Value	adj.P.Val
SRSF1	1.03	5.19	22.60	2.34E-72	4.13E-70
KIAA1033	1.79	6.91	22.58	2.87E-72	4.97E-70
RNF138	1.73	6.48	22.56	3.64E-72	6.25E-70
CHD9	1.06	4.09	22.46	9.55E-72	1.57E-69
ABCE1	1.02	5.97	22.44	1.10E-71	1.76E-69
UBA2	1.00	6.05	22.43	1.25E-71	1.98E-69
LRIF1	1.25	7.20	22.36	2.34E-71	3.59E-69
CNIH1	1.08	6.95	22.33	3.38E-71	5.06E-69
ZNRF2	1.77	5.88	22.31	3.75E-71	5.58E-69
PPP3CB	1.13	5.82	22.30	4.25E-71	6.28E-69
MICU2	1.01	7.72	22.27	6.02E-71	8.77E-69
DNAJB14	1.06	6.49	22.23	8.44E-71	1.22E-68
CSGALNACT2	1.49	7.24	22.21	9.90E-71	1.42E-68
AEBP2	1.38	4.39	22.20	1.11E-70	1.59E-68
CCNT2	1.65	4.35	22.15	1.93E-70	2.64E-68
THUMPD1	1.13	7.84	22.08	3.75E-70	5.02E-68
KLHL20	1.14	4.27	22.04	5.29E-70	6.99E-68
FAM3C	1.38	4.58	22.04	5.37E-70	7.03E-68
SLC30A5	1.08	5.85	21.94	1.45E-69	1.80E-67
B3GNT2	1.18	4.83	21.93	1.55E-69	1.90E-67
JMJD1C	1.50	6.92	21.92	1.69E-69	2.04E-67
UBR5	1.05	7.17	21.92	1.70E-69	2.04E-67
LOC100996626/PNRC2	1.15	7.05	21.90	2.07E-69	2.44E-67
EEA1	1.92	4.67	21.82	4.76E-69	5.50E-67
NUS1	1.40	5.64	21.79	6.08E-69	6.98E-67
C1ORF52	1.16	6.64	21.75	8.91E-69	1.02E-66
ZBTB33	1.02	6.67	21.72	1.22E-68	1.38E-66
AP1AR	1.49	4.97	21.70	1.46E-68	1.63E-66
OTUD1	1.03	9.19	21.63	2.82E-68	3.08E-66
USP32P1/USP32P2/USP32	1.26	8.45	21.54	6.78E-68	7.09E-66
TRAPPCL3	1.60	4.39	21.51	9.13E-68	9.31E-66
ZBTB1	1.39	5.43	21.48	1.29E-67	1.30E-65
ZNF292	1.45	6.22	21.46	1.51E-67	1.51E-65
ABCB10	1.66	5.82	21.43	2.09E-67	2.08E-65
USP32P1/USP32P2/USP32/USP6	1.27	8.81	21.40	2.67E-67	2.61E-65
TSNAX	1.11	7.38	21.35	4.28E-67	4.11E-65
TRMT1L	1.00	5.83	21.33	5.31E-67	5.01E-65
SLC5A3	1.43	6.21	21.25	1.22E-66	1.12E-64
PLEKHA3	1.08	7.20	21.22	1.64E-66	1.50E-64
ZNF644	1.32	5.97	21.21	1.76E-66	1.60E-64
PCNP	1.06	8.13	21.12	4.20E-66	3.72E-64
CMTR2	1.54	6.44	21.05	8.62E-66	7.40E-64
ANKRA2	1.29	3.28	21.03	1.01E-65	8.58E-64
MZT1	1.49	4.05	20.99	1.42E-65	1.18E-63
PPP2CA	1.25	5.15	20.99	1.46E-65	1.21E-63
DNAJC21	1.01	5.14	20.94	2.50E-65	2.04E-63
STXBP3	1.03	8.06	20.92	2.90E-65	2.35E-63
SCAI	1.40	3.97	20.91	3.25E-65	2.61E-63

**Cognitive test performance** Differences of cognitive performance metrics detected among the three MCI clusters through the pairwise Wilcoxon Rank Sum and Signed Rank Tests (\*\*\*\*  $p$ -value  $\leq 0.0001$ ; \*\*\*  $p$ -value  $\leq 0.001$ ; \*\*  $p$ -value  $\leq 0.01$ ; \*  $p$ -value  $\leq 0.05$ ; ns= not significant).

Metric	Description	Type	MCI-1 vs MCI-2	MCI-1 vs MCI-3	MCI-2 vs MCI-3
ADAS11	Alzheimer's Disease Assessment Scale-Cognitive (ADAS-Cog) 11	Cognitive	*	**	ns
ADAS11.bl	Alzheimer's Disease Assessment Scale-Cognitive (ADAS-Cog) 11	Cognitive at baseline	ns	*	ns
ADAS13	Alzheimer's Disease Assessment Scale-Cognitive (ADAS-Cog) 13	Cognitive	*	**	ns
ADAS13.bl	Alzheimer's Disease Assessment Scale-Cognitive (ADAS-Cog) 13	Cognitive at baseline	ns	*	ns
ADASQ4	Alzheimer's Disease Assessment Scale (ADAS4)	Cognitive	ns	*	ns
EcogSPLang	Everyday Cognition filled out by study partners	Cognitive	ns	*	ns
EcogSPLang.bl	Everyday Cognition filled out by study partners	Cognitive at baseline	ns	**	ns
EcogSPTotal	Everyday Cognition filled out by study partners	Cognitive	ns	*	ns
EcogSPTotal.bl	Everyday Cognition filled out by study partners	Cognitive at baseline	ns	*	ns
FAQ	Functional Activity Questionnaire	Cognitive	ns	*	**
FAQ.bl	Functional Activity Questionnaire	Cognitive at baseline	ns	*	ns
FDG	FDG-PET Meta Region of Interest	Imaging	ns	*	*
FDG.bl	FDG-PET Meta Region of Interest	Imaging at baseline	ns	**	**
MOCA	Montreal Cognitive Assessment	Cognitive	ns	**	ns
MOCA.bl	Montreal Cognitive Assessment	Cognitive at baseline	ns	**	ns
mPACCtrailsB	Preclinical Alzheimer's Cognitive Composite	Cognitive	ns	*	ns
RAVLT.immediate	RAVLT immediate assessment	Cognitive	*	****	*
RAVLT.immediate.bl	RAVLT immediate assessment	Cognitive at baseline	ns	***	*
RAVLT.learning	Rey Auditory Verbal Learning Test	Cognitive	ns	*	ns
RAVLT.learning.bl	Rey Auditory Verbal Learning Test	Cognitive at baseline	ns	**	ns
TAU	Cerebrospinal Fluid TAU	CSF Biomarker	*	ns	ns
TRABSCOR	Trail Making Test	Cognitive	ns	**	ns
TRABSCOR.bl	Trail Making Test	Cognitive at baseline	ns	**	ns
Ventricles	Ventricular volume	Imaging	ns	ns	*