MAPT haplotypes associate with Alzheimer's disease risk and brain gene expression

Mariet Allen[1], Michaela Kachadoorian[1], Zachary Quicksall[1], Fanggeng Zou[1], High Seng Chai[2], Curtis Younkin[1], Julia Crook[1], V. Shane Pankratz[2], Minerva Carrasquillo[1], Siddharth Krishnan[1], Thuy Nguyen[1], Li Ma[1], Kimberly Malphrus[1], Sarah Lincoln[1], Gina Bisceglio[1], Christopher Kolbert[2], Jin Jen[2], Ronald Petersen[2], Neill Graff-Radford[1], Dennis Dickson[1], Steven Younkin[1] and Nilufer Ertekin-Taner[1]. 1:Mayo Clinic, Jacksonville, FL, United States and 2:Mayo Clinic, Rochester, MN, United States.

Abstract

MAPT encodes for tau, the constituent of neurofibrillary tangles, a hallmark of Alzheimer's disease (AD) neuropathology. Common variation at the *MAPT* locus associate with risk for tauopathies, although the evidence for genetic association with late-onset AD (LOAD) is inconsistent. This is likely due to insufficient power and/or lack of detailed investigations of haplotypic variability at this locus. In this study we investigated associations of *MAPT* haplotypes with LOAD risk in a large case-control cohort and with brain expression levels of *MAPT*. We genotyped six *MAPT* haplotype-tagging SNPs in >5,400 subjects. We measured *MAPT* expression levels in the cerebellum and temporal cortex of ~200 autopsied AD subjects.

MAPT H2 haplotype was significantly associated with decreased risk of LOAD and decreased brain *MAPT* expression. The most common H1 sub-haplotype (H1b) had nominally significant association with increased risk of LOAD and a global test for the 19 haplotypes was also significant, with a number of additional significant or suggestive sub-haplotypes. Our results strongly suggest that *MAPT* has regulatory variants which confer LOAD risk by influencing its brain expression. These findings require replication in additional cohorts.

Aims

1) Test *MAPT* sub-haplotype associations with risk of LOAD in a large case-control series.

2) Test *MAPT* sub-haplotype associations with *MAPT* brain gene expression levels.

Methods I

Subjects: Elderly Caucasian subjects from two clinical case-control series (JS and RS) were utilized in this study. A third set of LOAD samples (AUT) included in this study were selected from the Mayo Clinic, Jacksonville brain bank (Table 1). For gene expression measures: RNA was isolated from Cerebellum (CER) and Temporal Cortex (TCX) brain tissue selected from the Mayo Clinic Brain Bank (Dr. Dennis Dickson). CER: 197 pathologic AD cases. TCX: 198 pathologic AD cases, as described [1].

SNPs: Six *MAPT* locus haplotype tagging (ht) SNPs were selected for genotyping; SNP rs8070723 was used as a proxy for the H1/H2 defining del-In9, the remaining 5 SNPs have been previously described to tag the majority of intra-H1 variability [2]

Methods II

Genotyping: Genotypes for three SNPs (rs1467967, rs242557 and rs8070723) were obtained, for a subset of the samples, from the Mayo Clinic LOAD GWAS. The remaining genotypes for these and an additional three SNPs (rs3785883, rs2471738 and rs7521) were obtained using Applied Biosystems® Taqman genotyping assays (Table 1).

Transcriptome Measurements: Whole-genome DASL (24,526 probes, NCBI Ref Seq, Build 36.2). Designed for partially degraded fresh frozen and FFPE tissues.

Statistical Analysis: AD association: Logistic regression analysis in PLINK corrected for Age, gender, series and ApoEε4 dose. Gene Expression Association: Multivariable linear regression analysis in PLINK, corrected for age at death, gender, RIN and *APOE* ε4 genotype, as previously described [1].

Results

- Significant association with reduced risk of LOAD was found for the H1/H2 differentiating SNP, rs8070723, in both the JS and RS series independently This finding was confirmed in the combined series (Table 2) where the p-value (7.0E-04) would maintain significance even after conservative Bonferonni correction for multiple tests.
- Eighteen sub-haplotypes were identified on the H1 background with a frequency of ≥1%. Of the total 19 haplotypes identified, 3 were nominally significant associated with risk for LOAD, (Table 3) and a global test for association was also significant (p = 0.012).
- A total of 15 *MAPT* haplotypes (freq >1%) were identified in the set of Autopsy LOAD samples included in this analysis, We identified significant association of the H2 haplotype with reduced *MAPT* expression levels (Table 4)

Table 1

Series	Cases			Elderly Controls		
	N	% Female	Age (SD)	N	% Female	Age (SD)
JS	886	62%	77.7 (6.5)	981	58%	76.5 (7.4)
RS	615	61%	80.0 (7.8)	2425	54%	78.3 (5.6)
AUT	551	59%	81.2 (8.6)	na	na	na
ALL	2052	61%	79.3 (7.6)	3406	55%	77.8 (6.2)

Table 1. Demographic information for series studied. JS = Jacksonville Series, RS = Rochester Series, AUT = Autopsy AD's. N=Number of Subjects, SD-Standard Deviation.

Figure 1

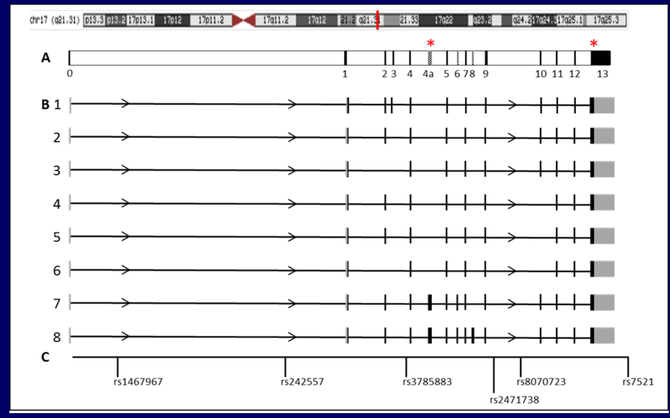

Figure 1. Location of *MAPT*, depiction of Refseq mRNA isoforms and SNP annotation. (A) The position of *MAPT* on Chromosome 17 is shown with a red line. The structure of the gene is shown with exons (0 -13) in solid black or striped boxes and introns in white. Red asterix (*) indicate exons to which DASL probes bind. (B) Eight mRNA transcripts are shown with grey and black boxes; black boxes represent coding exons and grey boxes represent un-translated regions. (C) The relative positions of the six haplotype tagging SNPs genotyped in this study are shown.

Table 2

CHR	SNP	BP	A1	N_A	N_U	MAF_A	MAF_U	OR	L95	U95	P
17	rs1467967	41342006	G	1868	3118	0.335	0.330	1.10	0.99	1.22	0.079
17	rs242557	41375573	A	1802	3133	0.377	0.379	1.00	0.90	1.11	0.988
17	rs3785883	41410269	A	1954	3293	0.182	0.183	1.07	0.95	1.22	0.274
17	rs2471738	41431900	T	1980	3302	0.223	0.209	1.07	0.95	1.21	0.250
17	**rs8070723**	**41436901**	**G**	**1903**	**3226**	**0.209**	**0.222**	**0.81**	**0.71**	**0.91**	**7.0E-04**
17	rs7521	41461242	A	1921	3250	0.458	0.457	1.08	0.97	1.19	0.152

Table 2. Single SNPs association results for 6 SNPs in the combined dataset of JS, RS and AUT cases vs. JS and RS controls. A1 = Minor Allele, A = Affected LOAD patients, U = Control subjects, MAF = Minor Allele Frequency, OR = Odds Ratio, L95 and U95 = Lower and Upper 95% Confidence intervals, P = p-value

References

1. Zou et al., Brain expression genome-wide association study (eGWAS) identifies human disease-associated variants. PLoS Genet. 2012;8(6):e1002707. doi: 10.1371/journal.pgen.1002707.
2. Pittman, A.M., et al., Linkage disequilibrium fine mapping and haplotype association analysis of the tau gene in progressive supranuclear palsy and corticobasal degeneration. J Med Genet, 2005. 42(11): p. 837-46.

Table 3

Haplotype	Alleles	F_All	F_A	F_U	OR	P
A (H2a)	AGGCGG	0.215	0.205	0.221	0.80	**4.1E-04**
B (H1b)	GGGCAA	0.173	0.180	0.169	1.15	**0.046**
C (H1c)	AAGTAG	0.118	0.118	0.118	0.91	0.277
D (H1d)	AAGCAA	0.076	0.071	0.078	0.99	0.905
E (H1e)	AGGCAA	0.074	0.075	0.073	1.12	0.308
H	AGACAA	0.044	0.038	0.048	0.91	0.506
I	GAGCAA	0.037	0.041	0.035	1.06	0.732
L	AGACAG	0.029	0.032	0.027	1.37	**0.059**
U	AAGCAG	0.025	0.028	0.024	1.14	0.517
M	GAGCAG	0.025	0.020	0.027	0.78	0.215
R	AGGTAG	0.017	0.017	0.017	1.01	0.952
G	GAACAA	0.017	0.014	0.018	0.91	0.692
O	AAACAA	0.016	0.016	0.017	0.90	0.696
x	GAATAG	0.016	0.018	0.014	1.59	**0.054**
y*	AAATAG	0.015	0.017	0.014	1.64	**0.056**
P	GGGTAG	0.014	0.014	0.014	1.31	0.301
J	AGGCAG	0.012	0.015	0.010	1.88	**0.031**
w	GGGCAG	0.012	<0.010	0.013	0.92	0.783
v	GGATAG	0.011	0.012	0.010	1.41	0.233
Global						**0.0123**

Table 3. Association of *MAPT* haplotypes with LOAD. A total of 19 *MAPT* haplotypes (freq ≥ 1%) were identified in the combined Mayo LOAD series, Haplotype nomenclature is assigned as previously reported [2]. One haplotype was identified that was not reported by Pittman *et al*, here it assigned as y*. F = haplotype frequency, All = All subjects, A = LOAD patients, U = Control subjects, OR = Odds Ratio, P= p-value.

Table 4

				ILMN_1710903		ILMN_2298727	
Haplotype	Alleles	Brain Region	Freq	BETA	P	BETA	P
A (H2a)	AGGCGG	CER	0.212	-0.45	**8.7E-33**	-0.16	**0.003**
		TCX		-0.49	**1.1E-31**	-0.20	**0.001**
B (H1b)	GGGCAA	CER	0.180	0.15	**0.003**	0.07	0.207
		TCX		0.13	**0.026**	0.13	0.058
C (H1c)	AAGTAG	CER	0.104	0.11	0.057	0.02	0.807
		TCX		0.17	**0.008**	0.05	0.536
E (H1e)	AGGCAA	CER	0.079	0.09	0.273	0.04	0.621
		TCX		0.15	0.079	0.13	0.212
I (H1i)	GAGCAA	CER	0.059	0.24	**0.010**	0.20	0.070
		TCX		0.07	0.422	-0.09	0.434
D (H1d)	AAGCAA	CER	0.057	0.00	0.970	0.04	0.721
		TCX		0.30	**0.007**	0.14	0.314
Global		CER			**3.0E-42**		0.352
		TCX			**3.2E-36**		**0.004**

Table 4. Association of the most common *MAPT* haplotypes (Freq > 5%) with *MAPT* brain gene expression. Haplotype nomenclature is assigned as previously reported [2].

Conclusions

- To the best of our knowledge, this is the largest study to evaluate *MAPT* sub-haplotypes for association with risk of LOAD, to date, and clearly shows a protective effect from the H2 haplotype and suggestive risk association with some H1 subhaplotypes.
- The results presented are consistent with the hypothesis that reduced *MAPT* expression affords protection from tauopathies such as LOAD; we find association of the H2 haplotype with reduced risk of LOAD and *MAPT* gene expression.