Association of MAPT haplotypes with Alzheimer’s disease risk and MAPT brain gene expression levels

Mariet Allen1, Michaela Kachadoorian1, Zachary Quicksall1, Fanggeng Zou1, High Seng Chai2, Curtis Younkin1, Julia E Crook3, V Shane Pankratz2, Minerva M Carrasquillo1, Siddharth Krishnan1, Thuy Nguyen1, Li Ma1, Kimberly Malphrus1, Sarah Lincoln1, Gina Bisceglio1, Christopher P Kolbert1, Jin Jen4, Shubhabrata Mukherjee5, John K Kauwe6, Paul K Crane6, Jonathan L Haines7,8, Richard Mayeux9, Margaret A Pericak-Vance10, Lindsay A Farrer11, Gerard D Schellenberg12, Alzheimer’s Disease Genetics Consortium (ADGC), Joseph E Parisi13, Ronald C Petersen14, Neill R Graff-Radford15, Dennis W Dickson1, Steven G Younkin1 and Nilüfer Ertekin-Taner1,15

Abstract

Introduction: MAPT encodes for tau, the predominant component of neurofibrillary tangles that are neuropathological hallmarks of Alzheimer’s disease (AD). Genetic association of MAPT variants with late-onset AD (LOAD) risk has been inconsistent, although insufficient power and incomplete assessment of MAPT haplotypes may account for this.

Methods: We examined the association of MAPT haplotypes with LOAD risk in more than 20,000 subjects (n-cases = 9,814, n-controls = 11,550) from Mayo Clinic (n-cases = 2,052, n-controls = 3,406) and the Alzheimer’s Disease Genetics Consortium (ADGC, n-cases = 7,762, n-controls = 8,144). We also assessed associations with brain MAPT gene expression levels measured in the cerebellum (n = 197) and temporal cortex (n = 202) of LOAD subjects. Six single nucleotide polymorphisms (SNPs) which tag MAPT haplotypes with frequencies greater than 1% were evaluated.

Results: H2-haplotype tagging rs8070723-G allele associated with reduced risk of LOAD (odds ratio, OR = 0.90, 95% confidence interval, CI = 0.85-0.95, p = 5.2E-05) with consistent results in the Mayo (OR = 0.81, p = 7.0E-04) and ADGC (OR = 0.89, p = 1.26E-04) cohorts. rs3785883-A allele was also nominally significantly associated with LOAD risk (OR = 1.06, 95% CI = 1.01-1.13, p = 0.034). Haplotype analysis revealed significant global association with LOAD risk in the combined cohort (p = 0.033), with significant association of the H2 haplotype with reduced risk of LOAD as expected (p = 1.53E-04) and suggestive association with additional haplotypes. MAPT SNPs and haplotypes also associated with brain MAPT levels in the cerebellum and temporal cortex of AD subjects with the strongest associations observed for the H2 haplotype and reduced brain MAPT levels (β = −0.16 to −0.20, p = 1.0E-03 to 3.0E-03).

Conclusions: These results confirm the previously reported MAPT H2 associations with LOAD risk in two large series, that this haplotype has the strongest effect on brain MAPT expression amongst those tested and identify additional haplotypes with suggestive associations, which require replication in independent series. These biologically congruent results provide compelling evidence to screen the MAPT region for regulatory variants which confer LOAD risk by influencing its brain gene expression.
Introduction
Alzheimer's disease (AD), the most prevalent cause of dementia, is defined by two neuropathological hallmarks: senile plaques primarily composed of extracellular amyloid-beta (Aβ) deposits and intracellular neurofibrillary tangles (NFTs) comprised of hyper-phosphorylated tau protein. MAPT (micro-tubule associated protein) encodes tau and resides within a ~900 kilobase (kb) inversion polymorphism (reviewed [1]) that generates a ~1.3 megabase (Mb) region of linkage disequilibrium (LD) defined by two extended haplotypes, referred to as H1 and H2. Variants have evolved that occur on only the H1 haplotype resulting in multiple sub-haplotypes.

Both common and rare genetic variation in MAPT have been strongly implicated in primary tauopathies. Rare missense and exon 10 splicing mutations, which lead to increased levels of tau isoforms with four microtubule binding domains (aka 4-repeat or 4R tau) lead to familial frontotemporal dementia with parkinsonism linked to chromosome 17 (FTDP-17) [2,3], whereas the common MAPT H1 haplotype strongly associates with increased risk of progressive supranuclear palsy (PSP) and cortico-basal degeneration (CBD) [4-8]. A recent genome-wide association study (GWAS) of PSP risk identified MAPT as the strongest locus, with risk alleles at rs8070723 which tags the H1 haplotype and also for rs242557, which partially tags the H1c subhaplotype [8].

Despite having tauopathy as a defining lesion, reports of association between AD and genetic variants at the MAPT locus are inconsistent. While MAPT H1 [9] haplotype or H1c subhaplotype [10-13] showed association with AD risk in some studies, others failed to detect association with H1 [10,13,14], H1c [15] or other MAPT variants [16]. The sample size for most of these studies range from a few hundred to a few thousand; and the largest published study of ~17,000 subjects only evaluated the H1/H2 haplotypes but none of the H1-subhaplotypes [9].

In addition to investigations of MAPT variants with risk for tauopathies, some studies also assessed their role in gene expression. MAPT exons 2, 3, 4a, 6 8 and 10 are known to be alternatively spliced [1], there are FTDP-17 splicing mutations which increase 4R tau [2,3] and 4R tau is increased in affected brain regions in PSP and CBD [17,18]. Allele-specific gene expression studies in human brains and neuronal cell lines identified higher levels of exon 10 containing transcripts but not total MAPT associated with the H1-haplotype [19] and higher levels of exon 2- and 3-containing transcripts associated with the H2 haplotype [20]. MAPT H1c subhaplotype was associated with higher total and 4R MAPT levels in human brains [11]. A study of exon levels in multiple brain regions from humans identified higher expression levels of exon 3 associated with the H2 haplotype, but no association of MAPT levels with the H1c-subhaplotype [21]. We have previously reported association of MAPT H1-tagging and rs242557 SNPs with increased brain MAPT levels in ~400 brains from a combined cohort of subjects with AD and other brain pathologies [22]. Collectively, these findings suggest that the disease risk conferred by some MAPT variants could be due to higher total or 4R tau levels and/or that the protective effect of MAPT H2-haplotype might be secondary to an increase in N-terminal exon-containing MAPT transcripts. While these studies are informative, to date, there has not been a systematic and well-powered analysis of MAPT subhaplotypes for association with MAPT brain expression levels.

Herein, we present a comprehensive assessment of MAPT variants that tag all MAPT subhaplotypes of frequency >1% in the largest to date MAPT association study of 9,814 LOAD cases vs. 11,550 controls. Further, we evaluate association of these MAPT variants in two brain regions: the cerebellum, which is predominantly unaffected in AD and the typically affected temporal cortex from ~200 autopsied LOAD subjects. Our well-powered and complementary investigation of disease risk and gene expression provides compelling evidence for a role of transcriptional regulatory variants of MAPT in conferring LOAD risk.

Methods
Subjects and samples
Mayo clinic cohort
We evaluated LOAD risk association with MAPT variants in 2,052 LOAD cases vs. 3,406 controls from Mayo Clinic. These elderly European-American subjects were from two clinical case–control series recruited at the Mayo Clinic in Rochester, MN (RS series: 615 LOAD cases, 2,425 controls) and Jacksonville, FL (JS series: 886 LOAD cases, 981 controls), as well as 551 autopsy-confirmed LOAD subjects from the Brain Bank at Mayo Clinic Florida (Additional file 1: Table S1). All clinical subjects were evaluated by a Mayo Clinic neurologist and autopsied subjects were diagnosed by our neuropathologist (DWD). All clinical LOAD cases had probable or possible AD and all pathologic LOAD cases had definite AD according to NINCDS-ADRDA criteria [23]. All controls had a clinical dementia rating score of 0. All LOAD subjects had an age at disease diagnosis (clinical), death (autopsied) and controls at their most recent visit ≥60 years. A subset of the Mayo Clinic cohort was included in the Mayo LOAD GWAS [24] (Additional file 1: Table S1) and gene expression GWAS (eGWAS) [22] from the temporal cortex (TCX, n = 202) and cerebellum (CER, n = 197). This study was approved by the Mayo Clinic institutional review board and appropriate informed consent was obtained from all individuals.
ADGC cohort

We utilized genetic data and covariate information on the European-American subjects from the Alzheimer’s Disease Genetics Consortium (ADGC) cohort. These subjects were collected from multiple research centers and designated into the following 14 series: Adult Changes in Thought (ACT)/Electronic Medical Records and Genetics (eMERGE), National Institute on Aging (NIA) Alzheimer Disease Centers (ADCs), Alzheimer Disease Neuroimaging Initiative (ADNI), Multi-Site Collaborative Study for Genotype-Phenotype Associations in Alzheimer’s Disease (GenADA), University of Miami/Vanderbilt University/Mt. Sinai School of Medicine (UM/VU/MSSM), MIRAGE Study, Oregon Health and Science University (OHSU), NIA-LOAD, Translational Genomics Research Institute series 2 (TGEN2), Rush University Religious Orders Study/Memory and Aging Project (ROSMAP), University of Pittsburgh (UP), and Washington University (WU). Detailed descriptions of these cohorts are provided elsewhere [25,26].

The ADGC cohort included subjects from the Mayo Clinic. To avoid any overlap, all subjects from Mayo Clinic were removed from the ADGC cohort. Standard quality control (QC) measures were applied to the ADGC dataset [27] with the following cutoffs, 95% call rate per person, 1% minor allele frequency (MAF) and 95% call rate for SNP, Hardy-Weinberg equilibrium (HWE) p > 1E-06 in controls. Additionally, directly observed (not imputed) SNPs from subjects across all series were evaluated for relatedness by using KING (Kinship-based INFerence for Gwax)-Robust [27] and a single representative was chosen for each pair of individuals who were third degree relatives or closer. Similarly, one representative was chosen for each family for the MIRAGE and NIA-LOAD family based studies. All cohort genotypes were imputed to a common set of >2 million SNPs (HapMap2) by the ADGC, as described [26]. The 7,762 LOAD cases and 8,144 controls from the ADGC (Additional file 1: Table S1), which remained after the QC, were utilized for the MAPT variant associations.

RNA isolation and gene expression measurements

All samples utilized in the brain gene expression analyses in this study are a subset of the previously published Mayo Clinic expression GWAS (eGWAS) [22]. In the current study brain gene expression levels of autopsied LOAD subjects measured from the cerebellum (n = 197) and temporal cortex (n = 202) were used. RNA extraction and gene expression measurements were previously reported [22]. Briefly, total RNA was isolated from frozen post-mortem brain tissue using the Ambion RNAqueous kit according to the manufacturer’s instructions. The quantity and quality of the RNA were evaluated using the Agilent 2100 Bioanalyzer and RNA 6000 Nano Chip. Whole Genome DASL assay (WG-DASL, Illumina, San Diego, CA) was used to measure transcript levels. This platform is designed for gene expression measurements for partially degraded RNA such as is typically isolated from frozen human brains. Details of gene expression measurements, data processing and QC were already published [22]. Briefly, 15 replicate samples measured on 5–6 different plates and on 2–3 different days were included in the study for QC and also for intra-class coefficient (ICC) [28] estimations. Raw probe level mRNA expression data were exported from GenomeStudio software (Illumina Inc.) for preprocessing with background correction, variance stabilizing transformation, quantile normalization and probe filtering using the lumi package of BioConductor [29,30]. Probes with detectable signal in >75% of the samples were used in subsequent analyses. We also annotated all of the probes by comparing their positions according to NCBI Ref Seq, Build 36.3 to those of all variants within dbSNP131 and identified the list of all probes which have ≥1 variants within their sequence.

Genotyping

Six MAPT locus haplotype tagging (ht) SNPs were selected for genotyping in the Mayo Clinic cohort (Additional file 1: Figure S1, Tables S1 and S2). SNP rs8070723 was used as a proxy for the H1/H2 haplotypes defining del-In9. The remaining 5 SNPs have been previously described to tag the majority of H1 sub-haplotypes [6]. Genotypes for three SNPs (rs1467967, rs2425557 and rs8070723) for a subset of the samples were obtained from the Mayo Clinic LOAD GWAS (Additional file 1: Table S1). The remaining genotypes for these and all genotypes for an additional three SNPs (rs3785883, rs2471738 and rs75211) were obtained using Applied Biosystems’ Taqman genotyping assays. The genotypes for these six SNPs were extracted from the ADGC GWAS data [27] using PLINK [31].

Statistical analysis

MAPT single SNP association analysis with LOAD risk

All six htSNPs were tested for association with disease risk in the combined Mayo Clinic cohort, as well as individually in the JS and RS series. The same SNPs were also tested in the ADGC cohort, as well as in the ADGC + Mayo combined cohorts. All SNPs were tested for deviations from Hardy-Weinberg equilibrium (HWE) [31] in controls.

Single SNP associations with disease risk were tested assuming an additive model, using multivariable logistic regression implemented in PLINK [31] including the following covariates: Age (defined for Mayo Clinic cohort as age at diagnosis/death/last diagnosis for clinical LOAD/autopsied LOAD/controls), sex, APOE ε4 dosage and series. The analyses in the ADGC-only cohort included these covariates and also ten principal components
obtained from EIGENSTRAT [32]. Mayo Clinic-only and Mayo + ADGC analyses did not include principal components, as they were not available for many of the Mayo subjects.

**MAPT haplotype association analysis with LOAD risk**
PLINK was used to estimate haplotype frequencies using the sliding window specification with a window size of six to encompass all six of the htSNPs. Haplotype associations with LOAD risk in the Mayo Clinic series were performed with both PLINK and haplo.score [33], which revealed identical results for the single haplotype analyses. According to the score statistic approach, all possible haplotypes consistent with the observed marker genotypes are obtained, maximum likelihood estimates of the haplotype frequencies, as well as the posterior probabilities of the pairs of haplotypes for each subject are computed. These posterior probabilities are then used to compute the score statistics for the association of (ambiguous) haplotypes with LOAD risk using multivariable logistic regression analysis with inclusion of the same covariates as discussed above. Only those haplotypes with frequencies >1% in the cohorts that they were tested in were included in the association analyses.

**MAPT variant association analysis with gene expression levels**
Each of the MAPT htSNPs and the estimated haplotypes were also tested for association with gene expression levels of MAPT in the TCX and CER of LOAD subjects, as measured using three probes: ILMN_1710903, ILMN_2310814 and ILMN_2298727. These LOAD subjects were also participants in our previously published eGWAS [22]. Association analysis was carried out in PLINK using linear regression approach, whereby preprocessed probe transcript levels for the three probes in each brain region (TCX and CER) were assessed as six individual quantitative phenotypes. Covariates included in the models were age at death, sex, APOE ε4 dosage, PCR plate, RNA integrity number (RIN) and adjusted RIN², as described previously [22,34]. Only those haplotypes with frequencies >1% in the autopsy series that they were tested in were included in the association analyses.

**Results**

**Association of MAPT single SNPs with LOAD risk**
Six MAPT htSNPs were tested for association with LOAD risk in the Mayo Clinic and ADGC cohorts both individually and combined (Table 1). All SNPs had genotyping call rates ≥90% in the Mayo Cohort (~90-97%), ~83-100% in the ADGC cohort and ~85-100% in the combined cohort (Additional file 1: Table S2). MAPT rs242557 had the lowest call rate of 83% in the ADGC cohort, with all other SNPs having call rates of ≥89%. All SNPs passed the HWE cutoff of p > 1E-06 in controls, although rs242557 had HWE p < 0.05 in the Mayo Clinic, but not the ADGC controls.

There was highly significant association of H2-tagging rs8070723-G allele with reduced risk of LOAD in the Mayo Clinic cohort (odds ratio = OR = 0.81, p = 7.0E-4) with remarkably similar OR estimates in the JS and RS series (Additional file 1: Table S3) and in the independent ADGC cohort (OR = 0.89, p = 1.3E-4) (Table 1). The association in the combined Mayo + ADGC cohort for this variant was highly significant (OR = 0.90, p = 5.3E-5) and would withstand Bonferroni correction for the six tested variants but not achieve significance at a genome-wide level.

In addition, rs3785883-A allele was nominally significant in the combined Mayo + ADGC cohort (OR = 1.06, p = 0.034) with very similar OR estimates in the Mayo Clinic (OR = 1.07, 95% confidence interval = 95% CI = 0.95-1.22) and ADGC (OR = 1.08, 95% CI = 1.01-1.15) cohorts. MAPT SNP rs1467967-G allele had suggestive LOAD risk association in the combined cohort (OR = 1.05, 95% CI = 1.0-1.10, p = 0.062) with risky OR estimates in both cohorts. MAPT SNP rs242557, previously implicated in AD [11-13] is not associated with LOAD in the combined ADGC + Mayo cohort of 18,342 subjects (p = 0.974), or the individual Mayo Clinic or ADGC cohorts.

**Association of MAPT haplotypes with LOAD risk**
In the Mayo Clinic cohort of ~5,000 subjects, we identified 19 MAPT haplotypes with a frequency >1%. In this cohort, rs8070723-G allele tagged the H2 haplotype, present in 21.5% of the subjects, perfectly. Eighteen sub-haplotypes were identified on the H1 background. Three MAPT haplotypes were nominally significantly associated with LOAD risk (Table 2) and a global test for haplotypic association was also significant (p = 0.012).

As expected, the MAPT H2 haplotype was significantly associated with decreased risk for LOAD in the Mayo Clinic cohort (OR = 0.80, p = 4.1E-04). Additionally the most common sub-haplotype on the H1 background, H1b (frequency = 17.3%), was nominally significantly associated with increased risk for LOAD (OR = 1.15, p = 0.046); as was a less frequent H1 sub-haplotype J (frequency = 1.2%, OR = 1.88, p = 0.031), while three other H1 sub-haplotypes were marginally associated, also with increased LOAD risk (L, X and Y).

In the ADGC cohort, the MAPT H2 haplotype, was present in 22% of the subjects. On the H1 background, 19 sub-haplotypes were identified with a frequency of ≥1%. As with the Mayo Clinic cohort, H2 haplotype was significantly associated with reduced risk of LOAD in the ADGC cohort (OR = 0.90, p = 6.29E-04). None of the H1-subhaplotypes had significant association with LOAD risk in this cohort.
**Table 1** MAPT single SNPs association results with LOAD risk in the Mayo, ADGC and combined Mayo + ADGC cohorts

| SNP   | A1 | Mayo cohort | ADGC cohort | Mayo + ADGC |
|-------|----|-------------|-------------|-------------|
|       | N  | MAF_A | MAF_U | OR | 95% CI | P       | N   | MAF_A | MAF_U | OR | 95% CI | P       | N   | MAF_A | MAF_U | OR | 95% CI | P       |
| rs1467967 G | 4,986 | 0.335 | 0.330 | 1.10 | 0.99 - 1.22 | 0.079 | 14,365 | 0.329 | 0.327 | 1.05 | 0.99 - 1.11 | 0.062 | 19,351 | 0.330 | 0.327 | 1.05 | 1.00 - 1.10 | 0.062 |
| rs242557 A | 4,935 | 0.377 | 0.379 | 1.00 | 0.90 - 1.11 | 0.988 | 13,407 | 0.358 | 0.356 | 1.01 | 0.95 - 1.07 | 0.721 | 18,342 | 0.363 | 0.364 | 1.00 | 0.95 - 1.05 | 0.974 |
| rs3785883 A | 5,247 | 0.182 | 0.183 | 1.07 | 0.95 - 1.22 | 0.274 | 15,187 | 0.178 | 0.171 | 1.08 | 1.01 - 1.15 | 0.032 | 20,434 | 0.179 | 0.175 | 1.06 | 1.01 - 1.13 | 0.034 |
| rs2471738 T | 5,282 | 0.223 | 0.209 | 1.07 | 0.95 - 1.21 | 0.250 | 14,181 | 0.207 | 0.201 | 1.05 | 0.98 - 1.12 | 0.194 | 19,463 | 0.209 | 0.203 | 1.05 | 0.99 - 1.11 | 0.109 |
| rs8070723 G | 5,129 | 0.209 | 0.222 | 0.81 | 0.71 - 0.91 | 7.0E-04 | 15,895 | 0.211 | 0.221 | 0.89 | 0.84 - 0.94 | 1.3E-04 | 21,024 | 0.211 | 0.221 | 0.90 | 0.85 - 0.95 | 5.3E-05 |
| rs7521 A | 5,171 | 0.458 | 0.457 | 1.08 | 0.97 - 1.19 | 0.152 | 15,656 | 0.466 | 0.468 | 1.02 | 0.97 - 1.07 | 0.529 | 20,827 | 0.465 | 0.465 | 1.02 | 0.97 - 1.06 | 0.437 |

Results of multivariable logistic regression analyses are shown. CHR = chromosome, A1 = Minor Allele, N = number of subjects with genotype calls, A = Affected (LOAD subjects), U = Unaffected (Control), MAF = Minor Allele Frequency, OR = Odds Ratio, 95% CI = 95% Confidence intervals, P = p-value. Boldface values within the tables indicate significant or suggestive associations with a p-value <0.10.
Table 2 *MAPT* haplotype association results with LOAD risk in the Mayo, ADGC and combined Mayo + ADGC cohorts

| Haplotype | Alleles | Mayo cohort |          |          |          | ADGC cohort |          |          |          |          | Mayo + ADGC cohort |          |          |          |          |          |
|-----------|---------|-------------|----------|----------|----------|-------------|----------|----------|----------|----------|-------------------|----------|----------|----------|----------|----------|
|           |         | F_All | F_A | F_U | OR | P | F_All | F_A | F_U | OR | P | F_All | F_A | F_U | OR | P | F_All | F_A | F_U | OR | P |
| A (H2a)   | AGGCCG  | 0.215 | 0.205 | 0.221 | 0.80 | 4.1E-04 | 0.220 | 0.225 | 0.236 | 0.90 | 6.29E-04 | 0.221 | 0.228 | 0.240 | 0.90 | 1.53E-04 | |
| B (H1b)   | GGGCAA  | 0.173 | 0.180 | 0.169 | 1.15 | 0.046 | 0.190 | 0.194 | 0.197 | 1.05 | 0.208 | 0.185 | 0.197 | 0.196 | 1.05 | 0.089 | |
| C (H1c)   | AAGTAG  | 0.118 | 0.118 | 0.118 | 0.91 | 0.277 | 0.127 | 0.131 | 0.131 | 1.03 | 0.545 | 0.124 | 0.132 | 0.131 | 1.00 | 0.997 | |
| D (H1d)   | AGGCCA  | 0.076 | 0.071 | 0.078 | 0.99 | 0.905 | 0.075 | 0.074 | 0.079 | 0.93 | 0.195 | 0.074 | 0.077 | 0.081 | 0.91 | 0.074 | |
| E (H1e)   | AGGCCA  | 0.074 | 0.075 | 0.073 | 1.12 | 0.308 | 0.077 | 0.082 | 0.081 | 1.00 | 0.961 | 0.077 | 0.083 | 0.081 | 1.02 | 0.618 | |
| G         | GAACAA  | 0.017 | 0.014 | 0.018 | 0.91 | 0.692 | 0.012 | 0.013 | 0.013 | 0.84 | 0.256 | 0.014 | 0.014 | 0.015 | 0.85 | 0.176 | |
| H         | AGACAA  | 0.044 | 0.038 | 0.048 | 0.91 | 0.506 | 0.042 | 0.045 | 0.041 | 1.08 | 0.309 | 0.042 | 0.045 | 0.045 | 1.03 | 0.641 | |
| I         | GAGCAA  | 0.037 | 0.041 | 0.035 | 1.06 | 0.732 | 0.034 | 0.035 | 0.035 | 1.03 | 0.718 | 0.035 | 0.037 | 0.036 | 1.06 | 0.422 | |
| J         | AGCAG   | 0.012 | 0.015 | 0.010 | 1.88 | 0.031 | 0.010 | 0.011 | 0.010 | 1.20 | 0.292 | 0.011 | 0.012 | 0.010 | 1.32 | 0.049 | |
| L         | AGACAG  | 0.029 | 0.032 | 0.027 | 1.37 | 0.059 | 0.032 | 0.034 | 0.033 | 1.06 | 0.483 | 0.031 | 0.034 | 0.032 | 1.10 | 0.187 | |
| M         | GACAG   | 0.025 | 0.020 | 0.027 | 0.78 | 0.215 | 0.021 | 0.022 | 0.020 | 1.09 | 0.459 | 0.022 | 0.023 | 0.023 | 1.00 | 0.978 | |
| N         | GAGAC   | NA     | NA     | NA     | NA     | 0.11 | 0.111 | 0.111 | 1.12 | 0.527 | NA     | NA     | NA     | 0.25 | 0.25 | |
| O         | AAACAA  | 0.016 | 0.016 | 0.017 | 0.90 | 0.696 | 0.018 | 0.020 | 0.018 | 1.18 | 0.192 | 0.018 | 0.020 | 0.019 | 1.09 | 0.457 | |
| P         | GGTAG   | 0.014 | 0.014 | 0.014 | 1.31 | 0.301 | 0.013 | 0.014 | 0.014 | 1.07 | 0.628 | 0.014 | 0.015 | 0.014 | 1.06 | 0.626 | |
| R         | AAGTAG  | 0.017 | 0.017 | 0.017 | 1.01 | 0.952 | 0.012 | 0.013 | 0.011 | 1.26 | 0.162 | 0.013 | 0.014 | 0.013 | 1.21 | 0.144 | |
| U         | AAGCAG  | 0.025 | 0.028 | 0.024 | 1.14 | 0.517 | 0.025 | 0.026 | 0.025 | 1.05 | 0.667 | 0.025 | 0.027 | 0.025 | 1.04 | 0.642 | |
| V         | GGATAG  | 0.011 | 0.012 | 0.010 | 1.41 | 0.233 | 0.010 | 0.011 | 0.010 | 1.14 | 0.463 | 0.011 | 0.012 | 0.011 | 1.20 | 0.210 | |
| W         | GGGCAG  | 0.012 | <0.010 | 0.013 | 0.92 | 0.783 | NA     | NA     | NA     | NA     | 0.01 | 0.01 | 0.01 | 1.00 | 0.978 | |
| X         | GAATAG  | 0.016 | 0.018 | 0.014 | 1.59 | 0.054 | 0.013 | 0.014 | 0.014 | 1.03 | 0.817 | 0.014 | 0.015 | 0.014 | 1.16 | 0.223 | |
| Y         | AAATAG  | 0.015 | 0.017 | 0.014 | 1.64 | 0.056 | 0.013 | 0.014 | 0.013 | 1.19 | 0.306 | 0.013 | 0.015 | 0.013 | 1.22 | 0.147 | |
| Z         | GAGTAG  | NA     | NA     | NA     | NA     | 0.011 | 0.011 | 0.010 | 1.10 | 0.590 | NA     | NA     | NA     | 0.032 | 0.032 | |

Global p value: 0.0123

Results of multivariable logistic regression analyses for *MAPT* haplotypes with frequencies >1% are shown. Haplotype nomenclature is assigned as previously reported [6,35]. Alleles for the SNPs defining the haplotypes are given in the 5' to 3' order as follows: rs1467967, rs242557, rs3785883, rs2471738, rs8070723, rs7521. Haplotypes not previously observed are designated by an asterisk (*). F_All = haplotype frequency in all subjects; F_A = in affected (LOAD) and F_U = unaffected (Control) subjects. OR = Odds Ratio, P = p-value. Boldface values within the tables indicate significant or suggestive associations with a p-value <0.10.
In the combined Mayo + ADGC cohort, there was significant global haplotype association (p = 0.033). MAPT H2 haplotype had highly significant association with reduced risk of LOAD (OR = 0.90, p = 1.53E-04). MAPT J subhaplotype had nominally significant association with LOAD risk in the combined cohort (OR = 1.32, p = 0.049) with suggestive association observed for H1b and increased LOAD risk (OR = 1.05, p = 0.089) and for H1d and reduced LOAD risk (OR = 0.91, p = 0.074). H1c subhaplotype did not achieve significance in the Mayo Clinic, ADGC or Mayo + ADGC cohorts.

**Association of MAPT single SNPs and haplotypes with gene expression levels**

In our published eGWAS [22], there were three probes on the WG-DASL platform that were used to measure MAPT levels: ILMN_1710903 and ILMN_2310814 that anneal to different regions of the MAPT 3’UTR and ILMN_2298727 that anneals to exon 4a (Additional file 1: Figure S1). Given that the inclusion of exon 4a in tau transcripts in the central nervous system was not reported previously, we generated a quantitative PCR assay against this exon, and were able to successfully measure it in the human brain (data not shown). All three probes passed our QC threshold of detectability in >75% of subjects, with ILMN_1710903 and ILMN_2310814 detected in 100% of all AD brains tested in both the cerebellum (CER) and temporal cortex (TCX) and with ILMN_2298727 detectable in 98.0% of AD CER and 83.7% of AD TCX tissue. We previously estimated intraclass coefficients [28] for all gene expression probes, which represent the percentage of variance in expression between samples over total variance and which reflect the genetic component that contributes to variability in gene expression. We determined that both ILMN_2298727 and ILMN_1710903 had high ICC estimates of 87%, whereas ILMN_2310814 had a low ICC estimate of 18%. The variances of gene expression estimated from all subjects in our eGWAS of cerebellar tissue (n = 374) [22] revealed consistent findings for these three MAPT probes, with both ILMN_2298727 (0.24) and ILMN_1710903 (0.12) having variance estimates that are ~ an order of magnitude greater than that of ILMN_2310814 (0.03). We thus conclude that ILMN_2310814 is unlikely to be an informative probe.

We previously annotated all our probes for variants in their sequence [22], given the concern that such variants may result in differential binding of probes with artifactual variance in the expression levels, and therefore could result in false positive associations with genetic variants in LD with probe sequence variants [36,37]. Our annotation detected two variants within the probe sequence of ILMN_1710903 (rs67759530, rs66561280) that were also polymorphic in our autopsied AD series. ILMN_2310814 did not have any variants within its probe sequence. ILMN_2298727 annotation identified rs73314997 within its sequence, although this variant was essentially monomorphic in our eGWAS subjects [22]. Thus, of the three MAPT probes assessed in our gene expression analyses, ILMN_2310814 is unlikely to be informative and ILMN_1710903 may be prone to artifactual results. We therefore focused on ILMN_2298727 in our MAPT expression analyses (Tables 3 and 4), although we show results from all 3 MAPT probes for completeness.

Evaluation of the six MAPT SNPs revealed significant associations between ILMN_2298727 and rs1467967, rs242557, rs8070723 and rs7521. The MAPT H2 haplotype tagging rs8070723 was associated with lower MAPT levels in both CER (β = -0.16, p = 0.002) and TCX (β = -0.20, p = 4.9E-04) of LOAD subjects (Table 3), as we previously reported in this cohort [22]. The other significant variants were associated with higher MAPT levels in both brain regions. Interestingly, the same variants showed associations in the same direction with the ILMN_1710903 probe, although with higher levels of significance.

Fifteen MAPT haplotypes with frequencies >1% were identified in the autopsied LOAD subjects with complete genotypes for the 6 variants (n = 178). There was globally significant haplotype association with the TCX gene expression levels measured with ILMN_2298727 (p = 0.004) (Table 4), that may be a reflection of the significant MAPT H2 association. MAPT H2 haplotype, as expected, was associated with lower CER (β = -0.16, p = 0.003) and TCX (β = -0.20, p = 0.001) MAPT levels. MAPT H1b was marginally associated with higher TCX levels (β = 0.13, p = 0.058), I with higher CER MAPT levels (β = 0.20, p = 0.07), and L with lower TCX MAPT levels (β = -0.33, p = 0.009). Significant associations with similar directions of effect were also observed with ILMN_1710903 and MAPT H2, H1b and I haplotypes.

**Discussion**

In this largest to date evaluation of haplotypic variation at the MAPT locus in 9,814 LOAD cases and 11,550 controls, we find robust and replicable association of the MAPT H2 haplotype with reduced risk of LOAD or, equivalently, increased risk of LOAD with the MAPT H1 haplotype- in two independent cohorts from Mayo Clinic and ADGC, with similar effect size estimates. Most prior reports of haplotypic association identified LOAD risk conferred by MAPT H1c subhaplotype [10-12], which we were unable to replicate. One group identified an association between the MAPT H1 haplotype and an increased risk for amnestic mild cognitive impairment [38], which can be a prodrome to clinical AD. The only other study to evaluate MAPT in a large cohort (3,940 cases and 13,373 controls) also identified an association between the H2 haplotype and decreased LOAD risk [9]. In that study by
**Table 3** MAPT single SNPs association results with brain MAPT gene expression levels

| SNP         | Brain region | N | ILMN_1710903 BETA | ILMN_1710903 P  | ILMN_2298727 BETA | ILMN_2298727 P  | ILMN_2310814 BETA | ILMN_2310814 P  |
|-------------|--------------|---|------------------|-----------------|-------------------|-----------------|------------------|-----------------|
| rs1467967   | CER          | 166 | 0.16             | 9.3E-06         | 0.11              | 0.011           | 0.00             | 0.859           |
|             | TCX          | 171 | 0.15             | 2.6E-04         | 0.12              | 0.019           | −0.01           | 0.604           |
| rs242557    | CER          | 173 | 0.17             | 1.2E-05         | 0.08              | 0.091           | 0.01             | 0.584           |
|             | TCX          | 180 | 0.23             | 3.4E-08         | 0.10              | 0.050           | −0.02           | 0.196           |
| rs3785883   | CER          | 175 | 0.14             | 0.005           | 0.00              | 0.977           | −0.01           | 0.621           |
|             | TCX          | 181 | 0.03             | 0.548           | −0.07             | 0.266           | 0.01             | 0.706           |
| rs2471738   | CER          | 176 | 0.07             | 0.101           | 0.02              | 0.665           | 0.02             | 0.386           |
|             | TCX          | 182 | 0.11             | 0.016           | 0.09              | 0.089           | −0.01           | 0.397           |
| rs8070723   | CER          | 174 | −0.44            | 2.1E-30         | −0.16             | 0.002           | 0.01             | 0.666           |
|             | TCX          | 181 | −0.48            | 8.9E-31         | −0.20             | 4.9E-04         | 0.02             | 0.222           |
| rs7521      | CER          | 176 | 0.12             | 2.8E-04         | 0.08              | 0.048           | 0.00             | 0.897           |
|             | TCX          | 182 | 0.16             | 1.1E-05         | 0.08              | 0.084           | 0.00             | 0.976           |

Results of multivariable linear regression analyses are shown. Probes ILMN_1710903 and ILMN_2298727 anneal to the 3’UTR sequence and ILMN_2310814 targets exon 4a. CER = cerebellum, TCX = temporal cortex. Beta: Coefficient of association with the minor allele. P = p-value. Boldface values within the tables indicate significant or suggestive associations with a p-value <0.10.

Gerrish et al. [9] the H2-haplotype tagging SNP had an OR estimate of 0.89 (p = 5.20E-04), which is remarkably similar to the estimate of the H2-tagging SNP (OR = 0.90, p = 6.1E-04) and H2 haplotype (OR = 0.90, p = 1.53E-04) in our study. It should be noted that both the present study and Gerrish et al. included samples from the ADNI and TGen series. We confirmed that the MAPT H2 association retains its significance in the ADGC cohort even after removal of these two datasets (OR = 0.87, p = 6.1E-04). Thus, there is evidence of MAPT H2 association with reduced risk of LOAD in two large and independent studies. Though robust, this LOAD risk association does not achieve genome-wide significance in either study or a p value < 1.0E-7 in the recent meta-analysis of 74,046 individuals by the IGAP consortium [39]. It will be important to evaluate the IGAP dataset for availability of MAPT haplotype tagging variants and to pursue an in-depth analysis of haplotypic association at this locus.

Although the MAPT H2 haplotypic association with LOAD was clearly the strongest of the MAPT haplotypes and one that we previously reported [22], we identified additional SNPs and haplotypes with nominal significance in our study. These weaker associations would not withstand multiple testing and could represent false positives and require replication in additional series. It should be noted that some of these variants, such as rs3785883, H1b, H1d and J showed consistent direction of effect in the Mayo Clinic and ADGC cohorts. MAPT rs3785883 minor allele was previously shown to associate with higher levels of CSF tau, phospho-tau and earlier age at onset [16]. Although this prior smaller study did not identify association with LOAD risk, the biological effect of this variant which associates with increased LOAD risk in our study appears to be consistent between these two studies.

We and others previously reported association between MAPT haplotypes and brain MAPT levels [11,19-22]. In this study, we evaluated MAPT subhaplotypes for association with brain MAPT levels in two brain regions from LOAD subjects. The most robust gene expression association occurs with the H2 haplotype, as we had reported [22] (\(\beta_{\text{CER}} = −0.16\); \(p_{\text{CER}} = 0.003\); \(p_{\text{TCX}} = −0.20\), \(p_{\text{TCX}} = 0.001\)), that also has the strongest association with LOAD risk in our study. We find that this haplotype with a protective effect on LOAD associates with lower brain MAPT levels. Given multiple MAPT alternatively spliced exons leading to multiple transcripts, each with potentially different effects on function [1,20], uncovering the precise regulatory change associated with genotypic variation in this region is critical. In our study, we mainly focus on results from one probe, ILMN_2298727, that is both informative and does not have a variant in its sequence based on annotation and genotyping. This probe is expected to anneal to exon 4a, however the expression levels obtained from it can be a surrogate for total MAPT levels or levels of any of the alternatively spliced exon-containing transcripts that reside with exon 4a. Indeed, gene expression associations with this probe are consistent with those from ILMN_1710903, which should recognize all transcripts, although ILMN_1710903 is confounded by a confirmed variant within its sequence. Our findings are also congruous with prior reports of associations of H1 haplotype or H1c sub-haplotype with higher 4R [19] and or total MAPT [11] levels, as measured by alternative gene expression measurement methods.

---

**References**

[1,20] uncovering the precise regulatory change associated with genotypic variation in this region is critical.

In our study, we mainly focus on results from one probe, ILMN_2298727, that is both informative and does not have a variant in its sequence based on annotation and genotyping. This probe is expected to anneal to exon 4a, however the expression levels obtained from it can be a surrogate for total MAPT levels or levels of any of the alternatively spliced exon-containing transcripts that reside with exon 4a. Indeed, gene expression associations with this probe are consistent with those from ILMN_1710903, which should recognize all transcripts, although ILMN_1710903 is confounded by a confirmed variant within its sequence. Our findings are also congruous with prior reports of associations of H1 haplotype or H1c sub-haplotype with higher 4R [19] and or total MAPT [11] levels, as measured by alternative gene expression measurement methods.
We did not identify significant associations between the H1c subhaplotype and brain MAPT levels, though we did observe suggestive associations between both CER and TCX MAPT levels and rs242557, a variant that partially tags H1c. MAPT rs1467967 associated with significant MAPT elevations in both brain regions and a suggestive association with LOAD risk, which is biologically consistent. These and additional weaker gene expression associations with variants such as H1b, I, L and rs7521 requires further replications.

In summary, our study provides evidence of robust LOAD risk and brain MAPT level associations with MAPT H2 haplotype and nominates additional variants and subhaplotypes for further investigations in LOAD. The overall genetic contribution of MAPT variants to LOAD risk appears to be modest, in contrast to primary

| Haplotype | Alleles | Brain region | Freq | ILMN_1710903 BETA | P | ILMN_2298727 BETA | P | ILMN_2310814 BETA | P |
|-----------|---------|--------------|------|-------------------|---|-------------------|---|-------------------|---|
| A (H2a)   | AGGCCG  | CER          | 0.212| −0.45             | 8.7E-33 | −0.16             | 0.003 | 0.00      | 0.832 |
|           |         | TCX          | 0.212| −0.49             | 1.1E-31 | −0.20             | 0.001 | 0.02      | 0.308 |
| B (H1b)   | GGGCAA  | CER          | 0.180| 0.15              | 0.003 | 0.07              | 0.207 | −0.01     | 0.691 |
|           |         | TCX          | 0.180| 0.13              | 0.026 | 0.13              | 0.058 | 0.01      | 0.630 |
| C (H1c)   | AAGTAG  | CER          | 0.104| 0.11              | 0.057 | 0.02              | 0.807 | 0.02      | 0.338 |
|           |         | TCX          | 0.104| 0.17              | 0.008 | 0.05              | 0.536 | −0.02     | 0.309 |
| D (H1d)   | AAGCAA  | CER          | 0.057| 0.00              | 0.970 | 0.04              | 0.721 | 0.07      | 0.098 |
|           |         | TCX          | 0.057| 0.30              | 0.007 | 0.14              | 0.314 | −0.02     | 0.646 |
| E (H1e)   | AGGCCA  | CER          | 0.079| 0.09              | 0.273 | 0.04              | 0.621 | 0.02      | 0.554 |
|           |         | TCX          | 0.079| 0.15              | 0.079 | 0.13              | 0.212 | 0.00      | 0.964 |
| H         | AGACAA  | CER          | 0.025| 0.05              | 0.748 | −0.13             | 0.478 | 0.04      | 0.538 |
|           |         | TCX          | 0.025| 0.13              | 0.424 | −0.05             | 0.803 | −0.05     | 0.353 |
| I         | GAGCAA  | CER          | 0.059| 0.24              | 0.010 | 0.20              | 0.070 | −0.04     | 0.322 |
|           |         | TCX          | 0.059| 0.07              | 0.422 | −0.09             | 0.434 | −0.01     | 0.766 |
| L         | AGACAG  | CER          | 0.043| 0.17              | 0.062 | 0.01              | 0.921 | −0.02     | 0.634 |
|           |         | TCX          | 0.043| −0.01             | 0.954 | −0.33             | 0.009 | 0.01      | 0.841 |
| M         | GAGCAG  | CER          | 0.028| 0.26              | 0.043 | 0.01              | 0.936 | 0.01      | 0.829 |
|           |         | TCX          | 0.028| 0.16              | 0.298 | −0.23             | 0.204 | 0.00      | 0.927 |
| O         | AAACAA  | CER          | 0.037| 0.19              | 0.104 | 0.00              | 0.985 | −0.04     | 0.435 |
|           |         | TCX          | 0.037| 0.20              | 0.070 | 0.21              | 0.110 | 0.00      | 0.933 |
| P         | GGCTAG  | CER          | 0.019| 0.04              | 0.765 | 0.07              | 0.637 | 0.02      | 0.731 |
|           |         | TCX          | 0.019| 0.07              | 0.725 | 0.26              | 0.259 | −0.03     | 0.613 |
| T         | AGATAG  | CER          | 0.016| 0.01              | 0.935 | −0.24             | 0.212 | −0.07     | 0.285 |
|           |         | TCX          | 0.016| 0.00              | 0.994 | 0.47              | 0.075 | 0.02      | 0.729 |
| U         | AAGCAG  | CER          | 0.029| 0.15              | 0.251 | 0.04              | 0.786 | −0.14     | 0.008 |
|           |         | TCX          | 0.029| 0.35              | 0.008 | 0.25              | 0.112 | −0.01     | 0.836 |
| X         | GAATAG  | CER          | 0.025| 0.10              | 0.406 | 0.19              | 0.168 | 0.06      | 0.248 |
|           |         | TCX          | 0.025| −0.04             | 0.756 | 0.11              | 0.510 | 0.02      | 0.584 |
| Y*        | AAATAG  | CER          | 0.022| −0.02             | 0.884 | −0.13             | 0.458 | −0.02     | 0.761 |
|           |         | TCX          | 0.022| −0.03             | 0.841 | −0.03             | 0.882 | 0.00      | 0.933 |
| CER       | Global p|              | 3.0E-42| 0.352      | 0.271 |
| TCX       |              |              | 3.2E-36| 0.004      | 0.999 |

Haplotypes with frequencies > 1% are assessed with multivariable linear regression analysis. Alleles for the SNPs defining the haplotypes are given in the 5’ to 3’ order as follows: rs1467967, rs242557, rs3785883, rs2471738, rs8070723, rs7521. Haplotypes not previously observed are designated by an asterisk (*). CER = cerebellum, TCX = temporal cortex. Beta: Coefficient of association with the minor allele. P = p-value. Boldface values within the tables indicate significant or suggestive associations with a p-value < 0.10.
Conclusions

In summary, these findings confirm associations between MAPT H2 haplotype and both reduced risk of LOAD and lower MAPT transcript brain levels. In addition, we describe additional MAPT variants and subhaplotypes that associate with LOAD risk and/or brain MAPT levels, which require confirmation in additional series. These results highlight the importance of joint utilization of gene expression and disease risk phenotypes. Additionally, these biologically consistent findings should encourage screening efforts in the MAPT region for discovery of regulatory variants that confer LOAD risk via influencing brain levels of MAPT transcripts.

Additional file

Additional file 1: This file includes Table S1. (Demographic information of the cohorts); Table S2. (Genotype counts, call rates and Hardy Weinberg results); Table S3. (MAPT single SNP association results with LOAD risk in the individual Mayo Clinic series. Results of multivariable logistic regression analysis); Figure S1. (MAPT Refseq mRNA isoforms and SNP annotation).

Abbreviations

Aβ: Amyloid-beta; AD: Alzheimer’s disease; ADGC: Alzheimer’s disease genetics consortium; CBD: Corticobasal degeneration; CSF: Cerebrospinal fluid; CRE: Cerebellum; eGWAS. Gene expression genome-wide association study; GWAS: Genome-wide association study; HWE: Hardy-Weinberg equilibrium; ICC: Intraclass coefficient; JS: Mayo Clinic Jacksonville, FL series; LD: Linkage disequilibrium; LOAD: Late-onset Alzheimer’s disease; MAPT: Microtubule-associated protein tau; NFT: Neurofibrillary tangles; OR: Odds ratio; PSP: Progressive supranuclear palsy; QC: Quality control; RS: Mayo Clinic Rochester, MN series; SNP: Single nucleotide polymorphism; TCX: Temporal cortex.

Competing interests

RC. Petersen, M.D., Ph.D. has been a consultant to GE Healthcare and Elan Pharmaceuticals, has served on a data safety monitoring board in clinical trials sponsored by Pfizer Incorporated and Janssen Alzheimer Immunotherapy and gave a CME lecture at Novartis Incorporated. N. Graff-Radford, M.D. has served as a consultant to Codman and received grant support from Elan Pharmaceutical Research, Pfizer Pharmaceuticals, Medivation, and Forest.

Authors’ contributions

MA conceived and designed the experiments, performed the experiments, analyzed the data and wrote the paper. MK performed the experiments, ZQ performed the experiments, FZ performed the experiments, HSC analyzed the data, CFT analyzed the data, JC analyzed the data, VSP analyzed the data, MMC performed the experiments, MMC performed the experiments, SK performed the experiments, TN performed the experiments, LM performed the experiments, KGM performed the experiments, SL performed the experiments, GB performed the experiments, CPK performed the experiments, JJ performed the experiments, SM contributed reagents/materials/analysis tools, JKR contributed reagents/materials/analysis tools, PKC contributed reagents/materials/analysis tools, JLT contributed reagents/materials/analysis tools, SW contributed reagents/materials/analysis tools, MAPV contributed reagents/materials/analysis tools, LAF contributed reagents/materials/analysis tools, GO contributed reagents/materials/analysis tools, ADGC contributed reagents/materials/analysis tools, JEP contributed reagents/materials/analysis tools, RCP contributed reagents/materials/analysis tools, NKG contributed reagents/materials/analysis tools, DWD contributed reagents/materials/analysis tools, SGY contributed reagents/materials/analysis tools, NET conceived and designed the experiments, analyzed the data and wrote the paper. All authors read and approved the final manuscript.

ADGC acknowledgements

The National Institutes of Health, National Institute on Aging (NIH-NIA) supported this work through the following grants: ADGC, U01 AG039284, RC2 AG036538, NACC, U01 AG016976, NCRR, U24 AG021886; NIA LOAD, U24 AG026395, U24 AG026396 Banner Sun Health Research Institute P30 AG019610; Boston University, P30 AG013846, U01 AG10483, R01 CA129769, R01 MH082095, R01 AG17173, R01 AG025529, R01AG13319; Columbia University, P30 AG088702, R37 AG15473; Duke University, P30 AG028377, AG05128; Emory University, AG02588; Group Health Research Institute, U01 AG00671, U01 AG004610; Indiana University, P30 AG10133; Johns Hopkins University, P50 AG001146, R01 AG006688; Massachusetts General Hospital, P50 AG005134; Mayo Clinic, P50 AG0016574, U01 AG006786; Mount Sinai School of Medicine, P50 AG005138, P01 AG002219; New York University, P30 AG08051, M1R00096, U1R023093, 5R01AG12101, 5R01AG022374, 5R01AG10316, 1RC2AG026502, 1R01AG035137; Northwestern University, P30 AG013854; Oregon Health & Science University, P30 AG00817, R01 AG026916; Rush University, P30 AG010161, R01 AG019085, R01 AG15819, R01 AG17917, R01 AG03146; TGen, R01 NS059873; University of Alabama at Birmingham, P50 AG016582, UL1 RR02777; University of Arizona, R01 AG031581; University of California, Davis, P30 AG10129; University of California, Irvine, P50 AG016573, P50, P50 AG016575, P50 AG016576, P50 AG016577; University of California, Los Angeles, P50 AG016570; University of California, San Diego, P50 AG005131; University of California, San Francisco, P50 AG005130, P01 AG19724; University of Kentucky, P30 AG028383, AG05144; University of Michigan, P50 AG00671; University of Pennsylvania, P30 AG010124; University of Pittsburgh, P50 AG005133, AG030653; University of Southern California, P50 AG005142; University of Texas Southwestern, P30 AG012300; University of Miami, R01 AG027944, AG017957, R01 AG019085, AG021547, R01 AG19757, R01 AG005136; Vanderbilt University, R01 AG005138; and Washington University, P50 AG005136, P50 AG005134; and Washington University, P50 AG005136, P01 AG03991. The Kathleen Price Bryan Brain Bank at Duke University Medical Center is funded by NINDS grant # NS39764, NIMH MH06451 and by Glaxo Smith Kline. Genotyping of the TGEn2 cohort was supported by Kronos Science. The TGEn series was also funded by NIA grant AG34504 to AIM, The Banner Alzheimer’s Foundation, The John B. and Julia Cook Foundation, The Alzheimer’s Institute, the Medical Research Council, and the state of Arizona and also includes samples from the following sites: Newcastle Brain Tissue Resource (funding via the Medical Research Council, local NHS trusts and Newcastle University), MRC London Brain Bank for Neurodegenerative Diseases (funding via the Medical Research Council),South West Dementia Brain Bank (funding via numerous sources including the Higher Education Funding Council for
England (HEFCE), Alzheimer's Research Trust (ART), BRACE as well as North Bristol NHS Trust Research and Innovation Department and DeNDRoN, The Netherlands Brain Bank (funding via numerous sources including Stichting MS Research, Brain Net Europe, Hersenstichting Nederland Breinbrekend Werk, International Parkinson Fonds, Internationale Stichting Alzheimer Onderzoek), Institut de Neuropatologia, Servei Anatòmia Patològica, Universitat de Barcelona. ADNI Funding for ADNI is through the Northern California Institute for Research and Education by grants from Abbott, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenics, Johnson and Johnson, Eli Lilly and Co., Medpace, Inc., Merck and Co., Inc., Novartis AG, Pfizer Inc, F. Hoffman-La Roche, Schering-Plough, Synarc, Inc, Alzheimer's Association, Alzheimer's Drug Discovery Foundation, the Dana Foundation, and by the National Institute of Biomedical Imaging and Bioengineering and NIA grants U01 AG024904, R2C AG036315, K01 AG030514. We thank Drs. D. Stephen Snyder and Marilyn Miller from NIA who are ex-officio ADGC members. Support was also from the Alzheimer’s Association (LAF, IIRG-08-89720, MP-V, IIRG-05-14147) and the US Department of Veterans Affairs Administration, Office of Research and Development, Biomedical Laboratory Research Program. P.S.G.-H. is supported by Wellcome Trust, Howard Hughes Medical Institute, and the Canadian Institute of Health Research.

ADGC affiliations
1Department of Neurology, Johns Hopkins University, Baltimore, Maryland, 1Department of Neurology, University of Michigan, Ann Arbor, Michigan, Geriatric Research, Education and Clinical Center (GRECC), VA Ann Arbor Healthcare System (VAAHS), Ann Arbor, Michigan, Michigan Alzheimer Disease Center, Ann Arbor, Michigan, 1Department of Neurology, University of California Los Angeles, Los Angeles, California, 1Department of Psychiatry, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, Geriatric Research, Education and Clinical Center (GRECC), University of Wisconsin, Madison, Wisconsin, 1Department of Medicine, University of Wisconsin, Madison, Wisconsin, Wisconsin Alzheimer’s Institute, Madison, Wisconsin, 1Department of Medicine (Genetics Program), Boston University, Boston, Massachusetts, 1Department of Pharmacology and Neuroscience, University of North Texas Health Science Center, Fort Worth, Texas, 1Department of Human Genetics, University of Pittsburgh, Pittsburgh, Pennsylvania, 1Department of Neurological Sciences, Rush University Medical Center, Chicago, Illinois, 1Department of Behavioral Sciences, Rush University Medical Center, Chicago, Illinois, 1Civin Laboratory for Neuropathology, Banner Sun Health Research Institute, Phoenix, Arizona, 1Departments of Psychiatry, Neurology, and Psychology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, 1The John P. Hussman Institute for Human Genomics, University of Miami, Miami, Florida, 1Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami, Miami, Florida, 1National Alzheimer’s Coordinating Center, University of Washington, Seattle, Washington, 1Rush Alzheimer’s Disease Center, Rush University Medical Center, Chicago, Illinois, 1Department of Pathology, Northwestern University Feinberg School of Medicine, Chicago, Illinois, 1Cognitive Neurology and Alzheimer’s Disease Center, Northwestern University Feinberg School of Medicine, Chicago, Illinois, 1Department of Neurology, University of Washington, Seattle, Washington, 1VA Puget Sound Health Care System/GRECC, Seattle, Washington, 1Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, 1Department of Psychiatry, Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts, 1Department of Neurology, Mayo Clinic, Rochester, Minnesota, 1Swedish Medical Center, Seattle, Washington, 1Department of Neurology, University of California San Francisco, San Francisco, California, 1Department of Medicine, Duke University, Durham, North Carolina, 1Department of Neuroscience, Mount Sinai School of Medicine, New York, New York, 1Department of Psychiatry, Mount Sinai School of Medicine, New York, New York, 1Departments of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, New York, 1Department of Pathology and Immunology, Washington University, St. Louis, Missouri, 1Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, 1USF Health Byrd Alzheimer’s Institute, University of South Florida, Tampa, Florida, 1Fred Hutchinson Cancer Research Center, Seattle, Washington, 1Department of Psychiatry and Behavioral Sciences, Miller School of Medicine, University of Miami, Miami, Florida, 1Department of Pathology, University of Alabama at Birmingham, Birmingham, Alabama, 1Department of Neurology, University of Southern California, Los Angeles, California, 1Department of Neurology, University of Alabama at Birmingham, Birmingham, Alabama, 1Neurogenomics Division, Translational Genomics Research Institute, Phoenix, Arizona, 1Department of Medicine, University of Washington, Seattle, Washington, 1Department of Neurology, University of California Irvine, Irvine, California, 1Department of Psychiatry and Hope Center Program on Protein Aggregation and Neurodegeneration, Washington University School of Medicine, St. Louis, Missouri, 1Program in Translational NeuroPsychiatric Genomics, Institute for the Neurosciences, Department of Neurology & Psychiatry, Brigham and Women’s Hospital and Harvard Medical School, Boston, Massachusetts, 1Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts, 1Department of Neurology, University of California Davis, Sacramento, California, 1University of Virginia School of Medicine, Charlottesville, Virginia, 1Institute for Memory Impairments and Neurological Disorders, University of California Irvine, Irvine, California, 1Wien Center for Alzheimer’s Disease and Memory Disorders, University of California Los Angeles, Los Angeles, California, 1Department of Pathology and Laboratory Medicine, University of California Irvine, Irvine, California, 1Division of Genetics, Department of Medicine and Partners Center for Personalized Genetic Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, 1Department of Neurology, Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts, 1Center for Applied Genomics, Children’s Hospital of Philadelphia, Philadelphia, Pennsylvania, 1Department of Pathology (Neuropathology), University of Pittsburgh, Pittsburgh, Pennsylvania, 1Institute of Neurology, University College London, Queen Square, London, 1Sanders-Brown Center on Aging, Department of Molecular and Biomedical Pharmacology, University of Kentucky, Lexington, Kentucky, 1Taub Institute on Alzheimer's Disease and the Aging Brain, Department of Neurology, Columbia University, New York, New York, 1Department of Pathology, Duke University, Durham, North Carolina, 1Department of Genome Sciences, University of Washington, Seattle, Washington, 1Department of Medicine (Medical Genetics), University of Washington, Seattle, Washington, 1Sanders-Brown Center on Aging, Department of Neurology, University of Kentucky, Lexington, Kentucky, 1Department of Pathology and Laboratory Medicine, University of California Davis, Sacramento, California, 1Department of Biostatistics, Boston University, Boston, Massachusetts, 1Department of Ophthalmology, Boston University, Boston, Massachusetts, 7University of Pittsburgh Alzheimer’s Disease Research Center, Pittsburgh, Pennsylvania, 7Department of Neurology, Oregon Health & Science University, Portland, Oregon, 7Department of Neurology, Portland Veterans Affairs Medical Center, Portland, Oregon, 7Department of Pathology and Laboratory Medicine, University of California, Irvine, Irvine, California, 7Department of Neurology, Boston University, Boston, Massachusetts, 7Department of Pathology, Boston University, Boston, Massachusetts, 7Department of Neuropsychology, University of California San Francisco, San Francisco, California, 7Department of Molecular & Medical Genetics, Oregon Health & Science University, Portland, Oregon, 7Department of Epidemiology, University of Washington, Seattle, Washington, 7Department of Neurobiology and Behavior, University of California Irvine, Irvine, California, 7Group Health Research Institute, Group Health, Seattle, Washington, 7Department of Pathology, University of Washington, Seattle, Washington, 7Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine, Seattle, Washington, 7Department of Pathology, University of Michigan, Ann Arbor, Michigan, 7Department of Psychiatry, Johns Hopkins University, Baltimore, Maryland, 7Department of Preventive Medicine, University of Southern California, Los Angeles, California, 7Department of Medicine – Pulmonary, New York University, New York, New York, 7Department of Neurology, University of Miami, Miami, Florida, 7Department of Pathology, University of
ADGC disclosure statement

T.D.B. received licensing fees from and is on the speaker's bureau of Athena Diagnostics, Inc. M.R.F. receives research funding from Bristol-Myers Squibb Company, Danone Research, Elan Pharmaceuticals, Inc., Eli Lilly and Company, Novartis Pharmaceuticals Corporation, Octapharma AG, Pfizer Inc., and Sonexa Therapeutics, Inc.; Receives honoraria as scientific consultant from Accera, Inc., Astellas Pharma US Inc., Baxter, Bayer Pharmaceuticals Corporation, Bristol-Myers Squibb, Eisai Medical Research, Inc., GE Healthcare, Medavante, Medivation, Inc., Merck & Co., Inc., Novartis Pharmaceuticals Corp, Pfizer, Inc., Prana Biotechnology Ltd, QR Pharma, Inc., The sanofi-aventis Group, and Toyama Chemical Co., Ltd; and is speaker for Eisai Medical Research, Inc., Forest Laboratories, Pfizer Inc. and Novartis Pharmaceuticals Corporation. A.M.G. has research funding from AstraZeneca, Pfizer and Genentech, and has received remuneration for giving talks at Pfizer and Genentech. R.C.P. is on the Safety Monitoring Committee of Pfizer, Inc. (Wyeth) and a consultant to the Safety Monitoring Committee at Janssen Alzheimer's Immunotherapy Program (Elan), to Elan Pharmaceuticals, and to GE Healthcare. R.E.T. is a consultant to Eisai, Japan in the area of Alzheimer's genetics and a shareholder in, and consultant to Pathway Genomics, Inc, San Diego, CA.

Acknowledgements

Support for this research was provided by the National Institutes of Health grants: National Institute on Aging (RO1 AG032990 to NET and RO1 AG018023 to NRGR and SGY); National Institutes on Neurologic Diseases and Stroke (RO1 NS080820 to NET), Mayo Alzheimer's Disease Research Center: (PS0 AG016574 to RCP, DWD, NRGR, SGY, and NET), Mayo Alzheimer's Disease Patient Registry: (U01 AG006576 to RCP); National Institute on Aging (AG025711, AG017216, AG039949 to DWD). This project was also generously supported by the Robert and Conice Smith and Abigail Van Buren Alzheimer's Disease Research Program (to RCP, DWD, NRGR, and SGY), and by the Palumbo Professorship in Alzheimer's Disease Research (to SGY). MMC and NET are supported partly by GHR Foundation grants. Combining the ADGC data sets was performed by SM, and was supported by RO1 AG 029672 (to PKC), U01 HG 006375 (to Eric Larson), and U01 AG 006781 (to Eric Larson). We thank the patients and their families for their participation, without whom these studies will not have been possible.

Author details

1Department of Neuroscience, Mayo Clinic Florida, Jacksonville, FL 32224, USA
2Department of Health Sciences Research, Mayo Clinic, Rochester, MN 55905, USA
3Department of Health Sciences Research, Mayo Clinic Jacksonville, FL 32244, USA
4Medical Genome Facility, Mayo Clinic, Rochester, MN 55905, USA
5Department of Medicine, Mayo Clinic, Rochester, MN 55905, USA
6Department of Pathology and Laboratory Medicine, Mayo Clinic Florida, Jacksonville, FL 32244, USA
7Department of Neurology, Mayo Clinic, Rochester, MN 55905, USA
8Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN 55905, USA
9Department of Neurology, University of Washington, Seattle, WA 98104, WA, USA
10Department of Biology, Neuroscience, Brigham Young University, Provo, UT 84662, USA
11Department of Molecular Physiology and Biophysics, and the Vanderbilt Center for Human Genetics Research, Vanderbilt University, Nashville, TN, USA
12Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH 44106, USA
13Gertrude H. Sergievsky Center, Department of Neurology, and Taub Institute on Alzheimer's Disease and the Aging Brain, Columbia University, New York, NY, USA
14The John P. Hussman Institute for Human Genetics and Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami, Florida, USA
15Department of Biostatistics, Medicine (Genetics Program), Ophthalmology, Neurology, and Epidemiology, Boston University, Boston, MA, USA
16Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA
17Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN 55905, USA
18Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN 55905, USA
19Department of Neurology, Mayo Clinic, Rochester, MN 55905, USA
20Department of Neurology, Mayo Clinic Florida, 4500 San Pablo Road, Bisdell 3, Jacksonville, FL 32224, USA

Received: 5 March 2014 Accepted: 28 May 2014

Published: 1 July 2014
29. Du P, Kibbe WA, Lin SM: lumi: a pipeline for processing Illumina microarray. Bioinformatics 2008, 24:1547–1548.

30. Lin SM, Du P, Huber W, Kibbe WA: Model-based variance-stabilizing transformation for Illumina microarray data. Nucleic Acids Res 2008, 36:e11.

31. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC: PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007, 81:559–575.

32. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D: Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006, 38:904–909.

33. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA: Score tests for association between traits and haplotypes when linkage phase is ambiguous. Am J Hum Genet 2002, 70:425–434.

34. Allen M, Zou F, Chai HS, Younkin CS, Miles R, Nair AA, Crook JE, Pankratz VS, Carrasquillo MM, Rowley CN, Nguyen T, Ma L, Malphrus KG, Bisceglio G, Ortolaza AI, Palusak R, Middha S, Maharjan S, Georgescu C, Schultz D, Rakhshan F, Kolbert CP, Jen J, Sando SB, Aasly JO, Baricikowska M, Uitti RJ, Wszolek ZK, Ross OA, Petersen RC, et al: Glutathione S-transferase omega genes in Alzheimer and Parkinson disease risk, age-at-diagnosis and brain gene expression: an association study with mechanistic implications. Mol Neurodegener 2012, 7:13.

35. Roder HM, Hutton ML: Microtubule-associated protein tau as a therapeutic target in neurodegenerative disease. Expert Opin Ther Targets 2007, 11:435–442.

36. Stranger BE, Forrest MS, Clark AG, Minichiello MJ, Deutsch S, Lyle R, Hunt S, Kahl B, Antonarakis SE, Tavare S, Deloukas P, Dermietzelis ET: Genome-wide associations of gene expression variation in humans. PLoS Genet 2005, 1:e78.

37. Doss S, Schmidt EE, Drake TA, Luiz AJ: Cis-acting expression quantitative trait loci in mice. Genome Res 2005, 15:681–691.

38. Di Maia E, Cammarata S, Parodi ML, Borghi R, Benussi L, Galli M, Galimberti D, Ghidoni R, Ronella D, Novello C, Pollero V, Perroni L, Odetti P, Scarpini E, Binetti G, Tabaton M: The H1 haplotype of the tau gene (MAPT) is associated with mild cognitive impairment. J Alzheimers Dis 2010, 19:909–914.

39. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, Jun G, Destefano AL, Bis JC, Beecham GW, Grenier-Boley B, Russo G, Thorton-Wells TA, Jones N, Smith AV, Chouraki V, Thomas C, Ikram MA, Zelenika D, Vardarajan BN, Kamatani Y, Lin CF, Gerrish A, Schmidt H, Kunkle B, Munste D, Rujas A, Bihoreau MT, Choi SH, Reitz C, et al: Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer’s disease. Nat Genet 2013, 45:1452–1458.

doi:10.1186/alzrt268
Cite this article as: Allen et al.: Association of MAPT haplotypes with Alzheimer’s disease risk and MAPT brain gene expression levels. Alzheimer’s Research & Therapy 2014 6:39.