Published in final edited form as:
Neuron. 2013 April 24; 78(2): 256–268. doi:10.1016/j.neuron.2013.02.026.

GWAS of cerebrospinal fluid tau levels identifies novel risk variants for Alzheimer's disease

Carlos Cruchaga[1,7,*], **John S.K. Kauwe**[8,*], **Oscar Harari**[1], **Sheng Chih Jin**[1], **Yefei Cai**[1], **Celeste M. Karch**[1], **Bruno Benitez**[1], **Amanda T. Jeng**[1], **Tara Skorupa**[1], **David Carrell**[1], **Sarah Bertelsen**[1], **Matthew Bailey**[8], **David McKean**[8], **Joshua M. Shulman**[9], **Philip L. De Jager**[10,11,12], **Lori Chibnik**[10,11,12], **David A. Bennett**[13], **Steve E. Arnold**[14], **Denise Harold**[15], **Rebecca Sims**[15], **Amy Gerrish**[15], **Julie Williams**[15], **Vivianna M. Van Deerlin**[16], **Virginia M.-Y. Lee**[16], **Leslie M. Shaw**[16], **John Q. Trojanowski**[16], **Jonathan L. Haines**[17], **Richard Mayeux**[18], **Margaret A. Pericak-Vance**[19], **Lindsay A. Farrer**[20], **Gerard D. Schellenberg**[21], **Elaine R. Peskind**[22,23], **Douglas Galasko**[24], **Anne M. Fagan**[2,6,7], **David M. Holtzman**[2,5,6,7], **John C. Morris**[2,3,6], **The GERAD Consortium Alzheimer's Disease Neuroimaging Initiative (ADNI)**[†], **The Alzheimer Disease Genetic Consortium (ADGC)**, and **Alison M. Goate, PhD**[1,2,4,6,7]

[1]Department of Psychiatry, Washington University School of Medicine, St. Louis, MO 63110

[2]Department of Neurology, Washington University School of Medicine, St. Louis, MO 63110

[3]Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO 63110

[4]Department of Genetics, Washington University School of Medicine, St. Louis, MO 63110

[5]Department of Developmental Biology, Washington University School of Medicine, St. Louis, MO 63110

[6]Knight Alzheimer's Disease Research Center, Washington University School of Medicine, St. Louis, MO 63110

[7]Hope Center for Neurological Disorders, Washington University School of Medicine, St. Louis, MO 63110

[8]Department of Biology, Brigham Young University, Provo, UT, 84602

[9]Departments of Neurology and Molecular and Human Genetics, Baylor College of Medicine, and The Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital, Houston, TX

[10]Program in Translational NeuroPsychiatric Genomics, Institute for the Neurosciences Department of Neurology, Brigham and Women's Hospital, Boston, MA 02115

[11]Harvard Medical School, Boston, MA 02115

[†]Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.ucla.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.ucla.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

[*]To whom correspondence should be addressed at: Department of Psychiatry, Washington University School of Medicine, 660 South Euclid Avenue B8134, St. Louis, MO 63110. goatea@psychiatry.wustl.edu, tel. 314-362-8691, fax. 314-747-2983.
[*]Joint First Authors who contributed equally

[12]Program in Medical and Population Genetics, Broad Institute of Harvard University and M.I.T., Cambridge, MA 02142

[13]Rush Alzheimer's Disease Center and Department of Neurological Sciences, Rush University Medical Center, Chicago, IL 60062

[14]Department of Psychiatry and Neurology, University of Pennsylvania, Philadelphia, PA, USA

[15]Institute of Psychological Medicine and Clinical Neurosciences, Cardiff University, UK

[16]Department of Pathology and Laboratory Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania, USA

[17]Department of Molecular Physiology and Biophysics, Vanderbilt Center for Human Genetics Research, Vanderbilt University, Nashville, Tennessee, USA

[18]Department of Neurology, Taub Institute on Alzheimer's Disease and the Aging Brain, and Gertrude H. Sergievsky Center, Columbia University, New York, New York, USA

[19]The John P. Hussman Institute for Human Genomics, and Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami, Miami, Florida, USA

[20]Departments of Biostatistics, Medicine (Genetics Program), Ophthalmology, Epidemiology, and Neurology, Boston University, Boston, Massachusetts, USA

[21]Department of Pathology and Laboratory Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania, USA

[22]Departments of Psychiatry and Behavioral Sciences, University of Washington School of Medicine

[23]Veterans Affairs Northwest Network Mental Illness Research, Education, and Clinical Center, Seattle, WA

[24]Department of Neurosciences, University of California San Diego, La Jolla, CA

Abstract

Cerebrospinal fluid (CSF) tau, tau phosphorylated at threonine 181 (ptau) and $A\beta_{42}$ are established biomarkers for Alzheimer's Disease (AD), and have been used as quantitative traits for genetic analyses. We performed the largest genome-wide association study for cerebrospinal fluid (CSF) tau/ptau levels published to date (n=1,269), identifying three novel genome-wide significant loci for CSF tau and ptau: rs9877502 (P=4.89×10^{-9} for tau) located at 3q28 between *GEMC1* and *OSTN*, rs514716 (P=1.07×10^{-8} and P=3.22×10^{-9} for tau and ptau respectively), located at 9p24.2 within *GLIS3* and rs6922617 (P = 3.58×10^{-8} for CSF ptau) at 6p21.1 within the *TREM* gene cluster, a region recently reported to harbor rare variants that increase AD risk. In independent datasets rs9877502 showed a strong association with risk for AD, tangle pathology and global cognitive decline (P=2.67×10^{-4}, 0.039, 4.86×10^{-5} respectively) illustrating how this endophenotype-based approach can be used to identify new AD risk loci.

INTRODUCTION

AD is neuropathologically characterized by the presence of extracellular Aβ plaques and intracellular aggregates of hyperphosphorylated tau in the brain (Hardy and Selkoe, 2002). CSF $A\beta_{42}$ and tau levels have emerged as useful biomarkers for disease and endophenotypes for genetic studies of AD. CSF tau and tau phosphorylated at threonine 181 (ptau) are higher in AD cases compared with non-demented elderly controls (Shoji et al., 1998; Kawarabayashi et al., 2001; Strozyk et al., 2003; Sunderland et al., 2003; Hampel et al., 2004; Jia et al., 2005; Schoonenboom et al., 2005; Welge et al., 2009). It has been shown

that genetic variants that increase risk for AD modify CSF $A\beta_{42}$ and tau levels, including pathogenic mutations in *APP, PSEN1* and *PSEN2*, and the common variants in *APOE* (Kauwe et al., 2007; Kauwe et al., 2008; Ringman et al., 2008; Kauwe et al., 2009; Cruchaga et al., 2010). CSF ptau levels correlate with the number of neurofibrillary tangles and the load of hyperphosphorylated tau present in the brain (Buerger et al., 2006). Elevated CSF ptau levels are correlated with neuronal loss and predict cognitive decline and conversion to AD in subjects with mild cognitive impairment (de Leon et al., 2004; Buerger et al., 2006; Andersson et al., 2007). Enigmatically, CSF tau levels are normal or low in other tauopathies such as progressive supranuclear palsy, so the precise relationship between the burden of tau pathology as well as the extent of neurodegeneration and the levels of CSF tau remain to be fully clarified (Hu et al., 2011). This notwithstanding, CSF tau levels may be a useful marker to identify genetic variants implicated not only with risk for Alzheimer's disease but also age at onset (Kauwe et al., 2008) or rate of progression (Shoji et al., 1998). Previous GWAS for CSF tau, and ptau levels (Han et al., 2010; Kim et al., 2011) have been conducted in much smaller samples and have shown robust association with markers on chromosome 19 surrounding *APOE* but failed to detect additional genome-wide significant associations. We have conducted a genome-wide association study (GWAS) for CSF tau and ptau using a sample that is more than three times the size of previous studies and have successfully detected loci that show novel genome-wide significant association signals.

RESULTS

Variability in CSF tau and ptau levels explained by common variants

Before performing any analysis, we performed stringent quality control (QC) in both the genotype and the phenotype data. For the phenotype data we confirmed that the tau and ptau level followed a normal distribution after log transformation. We also performed a stepwise regression analysis to identify the covariates showing a significant association with these endophenotypes. We performed a GWAS on 1,269 unrelated individuals recruited through the Knight-ADRC at Washington University, the Alzheimer's Disease Neuroimaging Initiative, a biomarker Consortium of Alzheimer Disease Centers coordinated by University of Washington and the University of Pennsylvania (table 1, and S1). While there are differences in the absolute levels of the biomarker measurements between the different studies that likely reflect differences in the methods used for quantification (regular ELISA vs Luminex), both methods measure the same analytes, but yield different absolute levels. In addition, CSF ptau and tau levels in the different studies show similar characteristics. CSF ptau and tau levels show a 10–17 fold difference in each dataset, are normally distributed after log transformation, and have similar covariates in each dataset (see statistical analyses).

To maximize our statistical power we performed a single-stage GWAS with our combined sample (Dube et al., 2007; Rohlfs et al., 2007; Kraft and Cox, 2008). The sample includes 687 elderly non-demented individuals and 591 individuals with a clinical diagnosis of AD (table 1, and S1). We used linear regression to test the additive genetic model of each single nucleotide polymorphism (SNP) for association with CSF biomarker levels after adjustment for age, gender, site and the three principal component factors from population stratification analysis. A total of 5,815,690 imputed and genotyped SNPs were included in these analyses. The inclusion of clinical dementia rating (CDR) or case/control status did not change the results significantly. No evidence of systematic inflation of p-values was found $\lambda = 1.003$ for ptau, and 1.009 for tau). To estimate the proportion of variance in CSF tau and ptau levels explained by genetic variants we used a genome-partitioning analysis (Yang et al., 2011). Approximately 7% (ptau) and 15% (tau) of the variability in the CSF levels of these proteins are explained by variants included on the GWAS chip plus the imputed SNPs. In this study SNPs in the *APOE* region show a genome-wide significant association with CSF tau and ptau (Table 2 and 3) and explain just 0.25–0.29% of the variability in CSF tau and

ptau, suggesting that most of the genetic variability in CSF tau and ptau levels is explained by other genetic variants.

APOE variants affect CSF tau and ptau levels independently of $A\beta_{42}$

Prevailing hypotheses suggest that *APOE* ε4 exerts its pathogenic effects through an Aβ-dependent mechanism (Castellano et al., 2011). However, several SNPs in the *APOE* region were genome-wide significant with both tau and ptau (rs769449; $P = 1.96 \times 10^{-16}$ and 2.56×10^{-18}, respectively, Table 2, 4 and Figure 1). To determine whether *APOE* SNPs influence CSF tau and ptau levels independently of Aβ pathology, and disease status we performed analyses including CSF $A\beta_{42}$ levels, or CDR as covariates in a regression model. When clinical status was included as a covariate the *APOE* SNP rs769449 was still the most significant signal ($P = 1.23 \times 10^{-12}$, Table 4). When CSF $A\beta_{42}$ levels were included in the model we also found a strong, but less significant, association for rs769449 with CSF ptau levels ($P = 3.22 \times 10^{-05}$). Analyses of tau follow the same pattern (Table 4) suggesting that at least part of the tau/ptau-*APOE* association is due to the underlying association of *APOE* with $A\beta_{42}$ levels. When the sample was stratified by clinical status, rs769449 showed a strong and similar effect size in both cases (n=519; Beta: 0.067; $P=3.38\times10^{-6}$) and in controls (n=687; Beta: 0.075, $p=1.54\times10^{6}$) with CSF ptau levels (Table S2). Several studies have suggested that up to 30% of elderly non-demented control samples meet neuropathological criteria for AD (Price and Morris, 1999; Schneider et al., 2009). It has also been shown that individuals with CSF $A\beta_{42}$ levels less than 500 pg/ml in the Knight-ADRC-CSF, and 192 pg/ml in the ADNI series have evidence of Aβ deposition in the brain, as detected by PET-PIB (Fagan et al., 2006; Jagust et al., 2009). Individuals with CSF $A\beta_{42}$ levels below these thresholds could be classified as preclinical AD cases with the presumption that some evidence of fibrillar Aβ deposits would be detected (Fagan et al., 2006; Jagust et al., 2009). When we used these thresholds, rs769449 showed a significant association with CSF tau and ptau in both strata, although the effect size was almost two fold higher in individuals with high $A\beta_{42}$ levels (n=416; Beta: 0.072; $P=6.58\times10^{-5}$, for CSF tau levels) than in individuals with low $A\beta_{42}$ levels (n=478; Beta: 0.035; $P=1.83\times10^{-2}$, for CSF tau levels) (Table S2). These results indicate that the residual association of SNPs in the *APOE* region is not dependent on clinical status or the presence of fibrillar Aβ pathology and clearly suggests that DNA variants in the *APOE* gene region influence tau pathology independently of Aβ or AD disease status.

To analyze whether there is more than one independent signal in the *APOE* gene region, *APOE* genotype was included in the model as a covariate (Table 4, and additional figures on https://hopecenter.wustl.edu/data/Cruchaga_Neuron_2013). The association for the SNPs located in the *APOE* region was reduced drastically (*P*-values between 0.02 and 0.008), suggesting that most of the association in this locus is driven by *APOE* genotype.

Novel loci associated with CSF tau and ptau levels

Outside the *APOE* region, we detected genome-wide significant association with three novel loci for CSF tau, ptau or both at 3q28, 9p24.2 and 6p21.1. Several SNPs in each locus showed highly significant p-values (Figure 1). For all loci, at least one SNP was directly genotyped (Table 2) and each of the datasets contributed to the signal, showing similar effect sizes and direction (Table S3), suggesting that these are real signals and unlikely to be the result of type I error.

The strongest association for CSF tau, after *APOE*, is rs9877502 ($P = 4.98 \times 10^{-09}$), located on 3q28 between *GEMC1* and *OSTN* and the non-coding RNA *SNAR-I* (Figure 1 and 2). Fifty-five intragenic SNPs located between *SNAR-I* and *OSTN*, showed a p-value lower than 9.00×10^{-05} (additional information on https://hopecenter.wustl.edu/data/

Cruchaga_Neuron_2013). Other genes located in this region, include *IL1RAP, UTS2D* and *CCDC50*, all of which are highly expressed in the brain. Bioinformatic analyses indicate that the most significant SNP in this locus and 33 SNPs in linkage disequilibrium (LD) with rs9877502 are located in transcription factor binding sites and some of these SNPs are also part of a transcription factor matrix (table S8–10), suggesting that rs9877502 or a linked variant could influence the expression of one or more of the genes located in this region.

Rs514716, located at 9p24.2 in an intron of *GLIS3,* shows genome-wide significant association with both CSF tau and ptau levels (Figure 2). The minor allele G (MAF = 0.136) is associated with lower CSF tau (β = −0.071; $P = 1.07 \times 10^{-8}$) and ptau levels (β = −0.072; $P = 3.22 \times 10^{-9}$). Seven additional intronic SNPs show genome-wide significant association with CSF p-tau levels or p-values lower than 9.00×10^{-05} for CSF tau levels (additional information on https://hopecenter.wustl.edu/data/Cruchaga_Neuron_2013). We used the HapMap and the 1000 genome project data to identify all of the SNPs in linkage disequilibrium (LD, R^2>0.8) with rs514716. A total of nine SNPs were identified, all of them intronic. Our bioinformatic analysis indicated that none of these SNPs disrupt a core splice site, but all of them are located in a conserved region.

Finally, for CSF ptau levels, several, relatively rare SNPs (MAF= 0.06), located at 6p21.1, within the *TREM* gene cluster show genome-wide significant p-values (Figure 2). As in the case of the other genome-wide signals, at least one SNP in the region was directly genotyped (rs6922617, β = −0.094; $P = 3.58 \times 10^{-8}$, table 2), and all of the CSF series contributed to the association (table S5). In this region, there was an additional peak driven by rs6916710 (MAF=0.39; $P = 1.58 \times 10^{-4;}$ β = −0.034) located in intron 2 of *TREML2*. In a recent study, we found a rare functional variant (R47H, rs75932628) in *TREM2,* which substantially increases risk for AD (Guerreiro et al., 2012). Based on these results, we genotyped rs75932628 in the Knight-ADRC and ADNI series to test whether this variant is associated with CSF levels. *TREM2 R47H* (rs75932628) showed strong association with both CSF tau (MAF=0.01; $P = 6.9 \times 10^{-4;}$ β =0.19) and ptau levels ($P = 2.6 \times 10^{-3;}$ β =0.16). As expected the minor allele (T) of rs75932628 is associated with higher CSF tau and ptau levels. The effect size (β) for the R47H variant was twice that of rs6922617 and rs6916710 (Table 5), while the less significant p-value is explained by the lower MAF, and sample size. To determine whether the associations seen with these three SNPs represent one signal or several independent associations we analyzed the linkage disequilibrium between the SNPs and performed conditional analyses. When rs6922617, rs6916710 or rs75932628 were included as a covariate in the model the other SNPs remained significant (Table 5). In our population, none of these SNPs were in LD with each other (table S3 and additional information on https://hopecenter.wustl.edu/data/Cruchaga_Neuron_2013). Together these results suggest that these three SNPs are tagging three independent signals within the *TREM* gene cluster that influence CSF ptau levels and at least in the case of *TREM2R47H* AD risk

Conditional analysis was also performed for the other genome-wide significant loci to test whether the association signal at each locus is driven by a single effect or by multiple independent effects and to determine whether the identified loci interact with each other. For the other loci, the signal for the conditioned SNP (and other SNPs in the same locus) totally disappeared confirming that the association at each locus represents a single signal. Conditioning on the genome-wide significant SNPs did not dramatically change the signals in other parts of the genome (additional information on https://hopecenter.wustl.edu/data/Cruchaga_Neuron_2013), suggesting that there is not strong interaction between these loci and the rest of the genome.

To evaluate the specificity of these genome-wide significant loci we also examined whether the SNPs were associated with another AD biomarker, CSF Aβ42 levels. Only SNPs within the *APOE* region showed genome-wide association with CSF tau and CSF Aβ42 (rs2075650 P= 1.83×10^{-40}). For the other regions the p values for association with CSF Aβ42 were modest: 0.02 for rs9877502, 0.03, for rs514716 and for 3.6×10^{-3} rs6922617. Furthermore, the correlation between the variants that give p values $<10^{-4}$ for either phenotype was low (r^2=0.07). Together these results confirm the specificity of our results and that CSF tau/ptau and CSF Aβ42 can be used as endophenotypes to identify genetic variants that influence different facets of the AD phenotype.

Gene expression analysis

To further characterize these associations we evaluated gene expression levels in three different ways. First, we determined whether the expression levels of the identified genes are associated with case-control status. Second, we determined whether the SNPs associated with CSF tau/ptau levels also affect tau (*MAPT*) gene expression levels in brain and third, we tested whether the SNPs were associated with expression levels of the candidate genes within each locus. To do this we analyzed *MAPT, GEMC1, IL1RAP, OSTN,* and *FOXP4* gene expression using cDNA from the frontal lobes of 82 AD cases and 39 non-demented individuals obtained through the Knight-ADRC Neuropathology Core. In addition *MAPT, RFX3, SLC1A1* and *PPAPDC2* gene expression were analyzed using publically available data from 486 late onset Alzheimer's Disease cases and 279 neuropathologically clean individuals form the GSE15222 dataset (Myers et al., 2007). We found strong association for *RFX3* ($P = 1.39\times 10^{-9}$; β =0.42), *SLC1A1* ($P = 1.01\times 10^{-4}$; β =−0.28) and *PPAPDC2* ($P = 4.80\times 10^{-3}$; β =−0.35), all located in the chromosome 9 region of association, with case-control status. We also found a nominally significant association of *IL1RAP* (Chr. 3; $P = 0.04$; β =−0.18) with case-control status but not for *MAPT, GLIS3, GEMC1, OSTN or FOXP4* (table S5). None of the SNPs associated with CSF tau/ptau levels showed an association with *MAPT* gene expression levels suggesting that they impact CSF tau levels by a post-transcriptional mechanism. Rs9877502 (chr. 3) showed nominally significant association with *IL1RAP* expression ($P = 0.02$; β =−0.17), but not with other genes in the same locus: *GEMC1* ($P = 0.54$; β =−0.09), and *OSTN* ($P = 0.87$; β =−0.02, Table S5).

Impact of the novel identified loci on other AD phenotypes

Because the purpose of this endophenotype-based approach is to identify variants implicated in disease, we tested whether the most significant SNP from each locus shows association with risk for AD, tau pathology or rate of cognitive decline. For the SNP located on 3q28 between *GEMC1* and *OSTN,* each copy of the rs9877502-A allele (minor allele frequency (MAF) = 0.386) is associated with higher CSF tau levels (regression coefficient (β) = 0.052). Genotypes for rs9877502 were not available for the case-control series, but rs1316356, which is in LD with rs9877502 ($D' = 1$, $R^2 = 0.932$) showed a strong association with AD risk (β = 0.81; $P = 2.67\times 10^{-4}$). Further, in an independent analysis leveraging two prospective cohorts, the Religious Orders Study and Rush Memory and Aging Project, rs9877502 was associated with global cognitive decline (n=1,593; β = −0.014; $P = 4.6\times 10^{-5}$) and in deceased subjects this variant was associated with burden of neurofibrillary tangles at autopsy (n=651; β = 0.055; $P = 0.014$) (Table 6). Importantly, these associations showed the predicted direction of effect for these phenotypes based on the CSF tau levels: the allele associated with lower tau levels is predicted to be protective for disease risk, associated with lower tau pathology, and with slower cognitive decline.

There was also some evidence that the SNPs associated with CSF tau and ptau levels in the 6p21.1 locus are also associated with risk for AD. A rare (MAF=0.01) functional coding variant with large effect size (Odds ratio >2) for AD risk was recently reported (Guerreiro et

al., 2012). This rare SNP (TREM2-R47H, rs75932628) was also associated with CSF ptau levels at $P=2.6 \times 10^{-3}$ (table 4). For the other locus we failed to detect significant association with risk for AD, tau pathology or cognitive decline, although the direction of the effect was in the expected direction based on the CSF levels (Table 6).

Pathway analyses

We performed a pathway analysis to determine whether signals that do not achieve genome-wide significance ($p<1.0\times 10^{-04}$) are enriched for sets of biologically related genes, represented as gene ontology terms (GO) and Kyoto Encyclopedia of genes and genomes (KEGG). Gene ontology terms for lipid transport and metabolism are significant for tau and ptau (Table S6). Furthermore, the KEGG pathway "*Type II diabetes mellitus*" is also significant for ptau (enriched by *MAPK9* and *IRS2*) and tau (enriched by *MAPK9*, *IRS2* and *MAPK1*). These results and the association of genetic variants in *GLIS3*, implicated in diabetes, with CSF tau levels support previous data suggesting that diabetes could influence risk for AD.

DISCUSSION

We have previously shown that using CSF tau and ptau levels as endophenotypes it is possible to identify genetic variants implicated in AD (Kauwe et al., 2008; Kauwe et al., 2010; Cruchaga et al., 2011; Kauwe et al., 2011; Cruchaga et al., 2012). This study represents the largest GWAS for CSF tau and ptau levels performed to date. Two other GWAS using the ADNI data (N=394) have been reported previously. In these smaller studies only the *APOE* locus showed genome-wide significant association with CSF $A\beta_{42}$ and tau levels. By using a threefold larger sample size than these studies we were able to identify four independent genome-wide significant loci, including *APOE* (Table 2). We calculated that common variants tagged by SNPs on the GWAS chip explain 6.45% and 15.14% of the overall variability in CSF ptau and tau levels, respectively. The four genome-wide significant loci identified in this study explain 1.45% of CSF ptau and 1.28% of CSF tau variability (Table 3). Together these four loci explain 22% and 9% of the genetic component for CSF ptau and tau levels, respectively, indicating additional variants and genes associated with CSF tau and ptau levels may be identified in future, using larger datasets and different approaches such as whole genome sequencing.

A single stage GWAS, rather than a two stage GWAS approach using the largest series as the discovery series, with follow up of the most significant SNPs in the rest of the samples, was used to maximize power (Dube et al., 2007; Rohlfs et al., 2007; Kraft and Cox, 2008). There are several indications that the identified genome-wide significant loci are real signals and not artifacts from the analysis or type I errors. First, several SNPs in each locus show highly significant p-values (Figure 1), and at least one SNP in each locus was directly genotyped (Table 2), eliminating the possibility that the signal is the result of an imputation error. Second, each of the genome-wide significant loci is the result of a strong and consistent association in each dataset. This is especially important, because *a priori*, the absolute values for the CSF biomarker traits are significantly different between series, which could lead to the identification of false positives. The fact that the SNPs show similar effect sizes and the same direction of effect in each dataset indicates that we were able to correct for any potential series-bias and represents an internal replication of each of the associations. If we had performed a two-stage analysis we would have identified these same four loci. Finally, for three (chr. 19, APOE and 3q28 and 6p21.1) of the four genome-wide significant loci we also found that the SNPs associated with CSF levels are also associated with risk for disease, tau pathology and/or cognitive decline. Importantly, all of these associations are in the direction predicted by the CSF tau and ptau associations. The alleles associated with

lower tau and ptau levels (which would be considered protective) are associated with lower risk for AD, lower tangle counts and slower memory decline.

As in the previously published GWAS for CSF tau/ptau levels, we found that the *APOE* locus was the strongest signal for CSF tau and ptau ((Han et al., 2010; Kim et al., 2011), table 2). SNPs in this locus explain between 0.25 to 0.29% of the variability in CSF tau and ptau levels (table 3). *APOE* is a known genetic risk factor for AD and most functional studies have focused on Aβ-dependent mechanisms. To determine whether or not the association of *APOE* SNPs with CSF tau and ptau levels was dependent of Aβ pathology we performed analyses including CSF $A\beta_{42}$ levels as a covariate. We also stratified our samples by case control status and by low or high CSF $A\beta_{42}$ levels. In all of these analyses we found that the association between *APOE* SNPs and tau or ptau levels remained significant (table 4 and S2), suggesting that *APOE* may also affect tau pathology via an Aβ-independent mechanism. Several other studies support this hypothesis. APOE shows isoform specific differences in its interaction with tau *in vitro* (Gibb et al., 2000; Zhou et al., 2006) and in transgenic mice neuron-specific differences in APOE isoform proteolysis are associated with increased tau phosphorylation (Brecht et al., 2004) and pathology (Andrews-Zwilling et al., 2010). These data provide additional evidence that *APOE* could also influence risk for AD through a tau-dependent mechanism, independent of effects on Aβ. When *APOE* genotype was included as a covariate, some SNPs in the *APOE* locus showed a moderate association with CSF tau/ptau levels (rs769449; $P=9.07\times10^{-03}$), indicating that most of the association is driven by *APOE* genotype, but suggesting that there may be additional variants in this region that modify CSF tau levels and risk for AD, independently of *APOE* genotype.

SNPs within the 3q28 locus showed association with CSF tau/ptau levels and a range of AD phenotypes including AD risk in the case control dataset, tangle pathology and rate of cognitive decline providing four independent sources of evidence that variants in this region influence risk for AD through a tau-dependent mechanism. Bioinformatic analysis did not reveal any strong putative functional SNP. However, the genes located in this region (*GEMC1*, *OSTN* and the non-coding RNA *SNAR-I*, *IL6RAP, UTS2D* and *CCDC50)* are highly expressed in brain and involved in neuronal synaptogenesis (Yoshida et al., 2012). The most significant SNP in this locus and 33 SNPs in LD with rs9877502 are located in transcription factor binding sites and some of these SNPs are also part of a transcription factor matrix (additional information on https://hopecenter.wustl.edu/data/Cruchaga_Neuron_2013), suggesting that rs9877502 or a linked variant could influence the expression of one or more of the genes located in this region. Based on the results of these bioinformatic analyses we performed several gene-expression experiments. *IL1RAP* showed a nominally significant association with case-control status (P=0.04). In addition rs9877502 showed a significant association with *IL1RAP* expression in frontal cortex (P=0.02, table S12).

The lack of association with risk for AD in the ADGC GWAS for the most significant SNP in the 6p21.1 locus may reflect insufficient power because the SNP has a low minor allele frequency (MAF=0.06). This hypothesis is supported by our recent identification of a rare functional coding variant (*TREM2*- R47H, rs75932628) in the same locus which substantially increases risk for AD (Guerreiro et al., 2012), and is also associated with CSF ptau levels in the present study. Interestingly, the genome-wide significant signal (tagged by rs6922617) is not in LD with rs75932628. Conditional analyses in this region identified another independent SNP (Figure 2, Table 5), located in an intron of *TREML2* that is associated with CSF tau and ptau levels. These data suggest that in this region there are at least three independent signals modifying CSF tau levels and risk for AD. Six *TREM*-family genes (*TREM1, TREM2,* and *TREML1 to TREML4*) are located in this region suggesting

that several variants in genes with similar function may affect risk for AD in an independent manner. The genome-wide significant SNP in this locus (rs11966476; $P = 4.79 \times 10^{-8}$), is located in a regulatory element and could modify the expression of *FOXP4, TREML3, TREML4 or TREM1* (Figure 2). Unfortunately these genes were not included in the GSE15222 dataset and Taqman assays for these genes were out of the dynamic range so we were unsuccessful in analyzing expression levels in brain tissue. Despite this, data from the Allen Brain Atlas suggests that these genes are expressed in the brain. *TREM2,* was expressed at higher levels in brain tissue from AD cases compared to controls ($P = 1.35 \times 10^{-5}$), as predicted in our previous studies (Guerreiro et al., 2012).

For the 9p24.2 locus, we did not observe significant association with risk for AD. This could be because these SNPs affect another aspect of AD such as disease duration or age at onset. Alternatively, these SNPs could affect CSF clearance or protein half-life without affecting risk for AD. If this were the case, we would expect that the same locus would be associated with levels of other CSF proteins. To test this we looked at the association of all of the SNPs identified in this study at the genome-wide significance level with other CSF biomarkers. We did not observe association between these SNPs and CSF levels of either APOE or Aβ (Cruchaga et al., 2012), suggesting that these loci are specific for CSF tau levels and are not associated with CSF clearance or protein half life in general. Finally, the lack of association of these loci with AD risk could indicate that the association with this locus is a type I error. The most significant SNPs in this locus are located in intron 7 of *GLIS3*, a gene which is highly expressed in brain. However, these SNPs (rs514716) are not associated with *GLIS3* expression in our relatively small series of brain samples (82 AD cases and 39 non-demented individuals). Both common and rare variants in this gene have been associated with risk for diabetes (Barker et al., 2011; Dimitri et al., 2011). There are several studies linking AD with glucose metabolism and diabetes (Accardi et al., 2012). In fact a meta-analysis combining data from eight studies, observed an association between diabetes mellitus and increased risk for AD (OR: 1.51 95%CI=1.31–1.73) (Bertram et al., Accessed 1/26/2013). In addition our pathway analysis independently identified a diabetes pathway (path:hsa04930, *P*-value for ptau= 6.60×10^{-03}, and tau=8.00×10^{-04}, Table S6), because of an enrichment of significant SNPs in *MAPK9*, *IRS2* and *MAPK1*. Two independent analyses in this study therefore suggest that diabetes-related genes may influence CSF tau and ptau levels, and ultimately risk for AD. These data all provide supportive evidence for common variants in this locus that influence AD pathogenesis.

Finally, because SNPs identified in this study were associated with CSF tau/ptau levels, we tested whether these SNPs are also associated with *MAPT* gene expression. None of the genome-wide significant SNPs showed association with *MAPT* expression in the brain and *MAPT* expression was not associated with case-control status in our brain series, the GSE15222, or any other published work on gene expression in brain (Webster et al., 2009; Zou et al., 2012). These results suggest that the SNPs identified in this study influence CSF tau/ptau protein levels post-transcriptional mechanism. Tau protein undergoes several posttranslational modifications including acetylation, glycosylation and phosphorylation. These changes are thought to play an important role in tau-related pathogenesis (Farias et al., 2011; Marcus and Schachter, 2011). It is possible that the genes identified in this study modify tau protein levels through posttranslational modification rather than gene expression.

Together these results clearly demonstrate the utility of using these endophenotypes to identify novel AD risk variants and variants associated with the rate of decline in symptomatic AD cases. The use of these endophenotype allowed us to identify risk variants that were not identified by GWAS because either those variants did not pass the stringent multiple test correction applied in the GWAS or were not covered in the earlier studies, because of their relatively low MAF. A second advantage of this approach is that in contrast

to GWAS hits from case control studies the endophenotype predicts a specific biological hypothesis for the pathogenic effect, which can be directly tested.

In summary, we have detected four genetic loci associated with CSF levels of tau, and ptau. One of them, in *APOE*, is already known to be associated with CSF tau and $A\beta_{42}$ (Kauwe et al., 2007; Kauwe et al., 2008; Cruchaga et al., 2010; Cruchaga et al., 2011; Kauwe et al., 2011) as well as risk for AD. The other three are novel loci. The top hit for CSF tau (rs9877502; 3q28) also exhibited association with risk for AD ($P = 2.67 \times 10^{-4}$), tangle pathology ($P = 0.01$) and global memory decline ($P = 4.86 \times 10^{-5}$). SNPs in the 6q21.1 locus are in the *TREM* gene cluster close to *TREM2,* a gene in which a rare variant has recently been reported to substantially increase risk for AD (Guerreiro et al., 2012). The other genome-wide significant locus identified in this study did not show association with risk for disease, tangle pathology or memory decline. The lack of association with other AD phenotypes could be because these SNPs have a weaker impact on these phenotypes, or because they affect other aspects of AD, such as disease duration or age at onset. Alternatively, the sample size for the datasets used in the pathology and memory decline studies may not provide enough statistical power. Overall, these results illustrate how genetic studies of disease endophenotypes are an effective approach for identifying disease risk loci that is complementary to case-control association studies.

EXPERIMENTAL PROCEDURES

Subjects and phenotypes

CSF tau, ptau and $A\beta_{42}$ were measured in 1,269 individuals. 501 samples were from research participants enrolled in longitudinal studies at the Knight-ADRC, 394 in ADNI, 323 in studies at the University of Washington (UW) and 51 in studies in University of Pennsylvania (UPenn). CSF collection and $A\beta_{42}$, tau and ptau181 measurements were performed as described previously (Fagan et al., 2006). Table 1 shows the demographic data and description of the CSF biomarkers in each dataset. The samples were genotyped using Illumina chips. Cases received a diagnosis of dementia of the Alzheimer's type (DAT), using criteria equivalent to the National Institute of Neurological and Communication Disorders and Stroke-Alzheimer's Disease and Related Disorders Association for probable AD (McKhann et al., 1984). Controls received the same assessment as the cases but were non-demented. All individuals were of European descent and written consent was obtained from all participants.

While there are differences in the absolute levels of the biomarker measurements between the studies that likely reflect differences in the methods used for quantification (regular ELISA vs Luminex), ascertainment, and/or in handling of the CSF after collection, CSF ptau levels in the Knight-ADRC, ADNI, UW and UPenn samples show similar characteristics (Table S1). CSF ptau and tau show a 10 fold difference between individuals in each dataset and have similar covariates in each dataset. CSF tau and ptau.

The Religious Orders Study (ROS) and the Rush Memory and Aging Project (MAP) recruit participants without known dementia who agree to annual clinical evaluations and sign an Anatomic Gift Act donating their brains at death. The full cohort with genotype data included 1,708 subjects (817 ROS and 891 MAP). The mean age at enrollment was 78.5 years and 30.9% were male. At the last evaluation, 24.9% met clinical diagnostic criteria for AD and 21.8% had mild cognitive impairment. The summary measure of global cognitive performance was based on annual assessments of 17 neuropsychiatric tests. A nested autopsy cohort consisted of 651 deceased subjects (376 ROS and 275 MAP); mean age at death was 81.5 years and 37.6% were male. Proximate to death, 40.9% of subjects included in the autopsy cohort met clinical diagnostic criteria for AD. Bielschowsky silver stain was

used to visualize neurofibrillary tangles in tissue sections from the midfrontal, middle temporal, inferior parietal, and entorhinal cortices, and the hippocampal CA1 sector. A quantitative composite score for neurofibrillary tangle pathologic burden was created by dividing the raw counts in each region by the standard deviation of the region specific counts, and then averaging the scaled counts over the 5 brain regions to create a single standardized summary measure. Additional details of the ROS and MAP cohorts as well as the cognitive and pathologic phenotypes are described in prior publications (De Jager et al., 2012; Keenan et al., 2012)

Genotyping

The Knight-ADRC and UW samples were genotyped with the Illumina 610 or the Omniexpress chip. The ADNI samples were genotyped with the Illumina 610 chip, and the UPenn sample with the Omniexpress. Prior to association analysis, all samples and genotypes underwent stringent quality control (QC). Genotype data were cleaned by applying a minimum call rate for SNPs and individuals (98%) and minimum minor allele frequencies (0.02). SNPs not in Hardy-Weinberg equilibrium ($P< 1\times10^{-6}$) were excluded. The QC cleaning steps were applied for each genotyping array separately. We tested for unanticipated duplicates and cryptic relatedness among samples using pairwise genome-wide estimates of proportion identity-by-descent. When a pair of identical samples or a pair of samples with cryptic relatedness was identified, the sample from the Knight-ADRC or samples with a higher number of SNPs passing QC were prioritized. Eigenstrat (Price et al., 2006) was used to calculate principal component factors for each sample and confirm the ethnicity of the samples. Rs7412 and rs429358 which define the *APOE ε2/ε3/ε4* isoforms were genotyped using Taqman genotyping technology, as previously described (Koch et al., 2002; Cruchaga et al., 2009; Cruchaga et al., 2010; Kauwe et al., 2010; Cruchaga et al., 2011; Cruchaga et al., 2012).

DNA from ROS and MAP subjects was extracted from whole blood, lymphocytes or frozen postmortem brain tissue and genotyped on the Affymetrix Genechip 6.0 platform, as previously described (Keenan et al., 2012). Following standard QC procedures, imputation was performed using MACH software (version 1.0.16a) and HapMap release 22 CEU (build 36) as a reference.

Imputation in Illumina datasets

The 1000 genome data (June 2011 release) and the Beagle software were used to impute up to 6 million SNPs. SNPs with a Beagle R^2 of 0.3 or lower, a minor allele frequency (MAF) lower than 0.02, out of Hardy-Weinberg equilibrium ($p< 1\times10^{-6}$), a call rate lower than 95% or a Gprobs score lower than 0.90 were removed. A total of 5,815,690 SNPs passed the QC process. To confirm the accuracy of our imputation we genotyped 23 SNPs, included the most significant SNPs, using Sequenom. All of the SNPs, showed a concordance rate between imputed and directly genotyped calls greater than 97.9% except rs1024718 which was 93.33% (Table S7).

Statistical Analyses

Association of CSF ptau with the genetic variants was analyzed as previously reported (Cruchaga et al., 2010; Cruchaga et al., 2011; Kauwe et al., 2011). Our analysis included a total of 5,815,690 imputed and genotyped variants. CSF tau and ptau values were log transformed to approximate a normal distribution. Because the CSF biomarker levels were measured using different platforms (Innotest plate ELISA vs AlzBia3 bead-based ELISA, respectively) we were not able to combine the raw data. For the combined analyses we standardized the mean of the log transformed values from each dataset to zero. No

significant differences in the transformed and standardized CSF values for different series were found.

We used Plink to analyze the association of SNPs with CSF biomarker levels. Age, gender, site, and the three principal component factors for population structure were included as covariates. The calculated genomic inflation factor was λ=1.003, and 1.009, for tau, and ptau respectively (Supplementary figure 1). In order to determine whether the association of *APOE* with CSF tau levels was driven by case-control status we included clinical dementia rating (CDR) or CSF $A\beta_{42}$ as a covariate in the model or stratified the data by case control status. We also performed analyses including *APOE* genotype and CDR as covariates.

Association with risk for Alzheimer's disease

P-values for the most significant SNPs for the association with CSF tau and ptau were included here from the previously published GWAS for AD, consisting of 11,840 controls and 10,931 cases (Naj et al., 2011).

Genome partitioning

We used the algorithm GCTA (Genome-wide Complex Trait Analysis) to estimate the proportion of phenotypic variance explained by genome-wide and imputed SNPs (Yang et al., 2011).

Association with cognitive decline and neurofibrillary pathology

Analyses of SNP effects on global cognitive decline in ROS and MAP were performed as in prior publications (De Jager et al., 2012). Briefly, we first fit linear mixed effects models using the global cognitive summary measure in order to characterize individual paths of change, adjusted for age, sex, years of education, and their interactions with time. At least two longitudinal measures of cognition were required for inclusion in these analyses, for which data on 1,593 subjects was available. We then used these person-specific, residual cognitive decline slopes as the outcome variable in our linear regression models, with each SNP of interest coded additively relative to the minor allele, and further adjusted for study membership (ROS vs. MAP) and the first 3 principal components from population structure analysis. For analyses of neurofibrillary tangle burden, linear regression was used to relate SNPs to the pathologic summary measure, adjusting for age at death, study membership, and 3 principal components. Because the data were skewed, square-root of the scaled neurofibrillary tangle burden summary score was used in analyses.

Bioinformatics analyses

We used Pupasuite (Conde et al., 2006), the SNP Function Portal (http://brainarray.mbni.med.umich.edu/Brainarray/Database/SearchSNP/), the SNP Function annotation portal (http://brainarray.mbni.med.umich.edu/Brainarray/Database/SearchSNP/snpfunc.aspx) and the SNP and CNV Annotation Database (http://www.scandb.org) to perform the SNP annotation and to identify the putative functional SNPs.

Pathway Analysis

We applied the method ALIGATOR (Holmans et al., 2009) to identify the Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enriched by SNP with significant association. This method performs an overrepresentation analysis, evaluating the significance for each category of genes while correcting for gene size, number of SNPs genotyped per gene, overlapping genes and linkage disequilibrium between SNPs. It selects the set of genes, which are tagged by SNPs that are more significant than a specific threshold (p-values<1.0E-04). The pruning process that eliminates SNPs in linkage

disequilibrium is performed by considering only the most significant SNP among all of the SNPs that have r^2>0.2 and are within 1Mb. Moreover, we removed all of the genes that are in the *APOE* region (1Mb up/downstream) (Jones et al., 2010). The significance of each term and pathway is calculated by comparing the number of significant genes to the number of genes expected by chance. For this purpose, the algorithm generates 5,000 sets of genes, by randomly selecting SNPs until a list of n tagged genes is formed. The excess of significantly overrepresented sets of genes (Holmans et al., 2009) is calculated by applying a bootstrap method (1000 permutations).

Gene Expression analysis

Analyses of association between SNPs and gene expression was carried out using cDNA from the frontal lobes of 82 AD cases and 39 non-demented individuals obtained through the Washington University Knight-Alzheimer Disease Research Center (WU-ADRC) Neuropathology Core. Total RNA was extracted from the frontal lobe using the RNeasy mini kit (Qiagen) following the manufacturer's protocol. cDNAs were prepared from the total RNA, using the High-Capacity cDNA Archive kit (ABI). Gene expression was analyzed by real-time PCR, using an ABI-7500 real-time PCR system. Real-time PCR assays were used to quantify *MAPT, GLIS3, GEMC1, IL1RAP, OSTN, and FOXP4* cDNA levels using Taqman assays. GADPH, MAP2, AIF and GFAP were used as reference genes. Each real-time PCR run included within-plate duplicates. Real-time data were analyzed using the comparative Ct method. The Ct values of each sample were normalized with the Ct value for the housekeeping genes. We also used the GEO dataset GSE15222 (Myers et al., 2007) to analyze the association of *MAPT*, *RFX3, SLC1A1* and *PPAPDC2* genes and case-control status. None of the other genes (*GLIS3, GEMC1, IL1RAP, OSTN, FOXP4)* were found in this dataset. This dataset includes genotype and expression data from 486 late onset Alzheimer's Disease cases and 279 neuropathologically clean individuals. Association of mRNA levels with case control status or the different SNPs was carried out using ANCOVA. Stepwise regression analysis was used to identify the potential covariates (postmortem interval, age at death, site, and gender) and significant covariates were included in the analysis. SNPs were tested using an additive model with minor allele homozygotes coded as 0, heterozygotes coded as 1, and major allele homozygotes coded as 2.

ADNI material and methods

Data used in the preparation of this article were obtained from the ADNI database (www.loni.ucla.edu\ADNI). The ADNI was launched in 2003 by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, the Food and Drug Administration, private pharmaceutical companies and non-profit organizations, as a $60 million, 5-year public-private partnership. The Principal Investigator of this initiative is Michael W. Weiner, M.D. ADNI is the result of efforts of many co-investigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the U.S. and Canada. The initial goal of ADNI was to recruit 800 adults, ages 55 to 90, to participate in the research -approximately 200 cognitively normal older individuals to be followed for 3 years, 400 people with MCI to be followed for 3 years, and 200 people with early AD to be followed for 2 years." For up-to-date information see www.adni-info.org.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by grants from NIH (P30 NS069329-01, R01 AG035083, R01 AG16208, P50 AG05681, P01 AG03991, P01 AG026276, AG05136 and PO1 AG05131, U01AG032984, AG010124, R01 AG042611), AstraZeneca and the Barnes-Jewish Hospital Foundation. The authors thank the Clinical and Genetics Cores of the Knight ADRC at Washington University for clinical and cognitive assessments of the participants and for *APOE* genotypes and the Biomarker Core of the Adult Children Study at Washington University for the CSF collection and assays.

We acknowledge use of genotype data from the '610 group', part of the GERAD1 consortium, who were supported by funding from the Wellcome Trust (including GR082604MA), Medical Research Council (including G0300429), Alzheimer's Research Trust, Welsh Assembly Government, Alzheimer's Society, Ulster Garden Villages, Northern Ireland R&D Office, Royal College of Physicians/Dunhill Medical Trust, Mercer's Institute for Research on Ageing, Bristol Research into Alzheimer's and Care of the Elderly (BRACE), Charles Wolfson Charitable Trust, NIH (including PO1 AG026276, PO1 AG03991, RO1 AG16208, P50 AG05681), NIA, Barnes Jewish Hospital Foundation, Charles and Joanne Knight Alzheimer's Research Initiative of the Washington University Alzheimer's Disease Research Centre, the UCLH/UCL Biomedical Centre, Lundbeck SA, German Federal Ministry of Education and Research (BMBF): Kompetenznetz Demenzen (01GI0420), Bundesministerium für Bildung und Forschung, and Competence Network Dementia (CND) Förderkennzeichen (01GI0102, 01GI0711). Recruitment and CSF studies at University of Washington and UCSD were supported by NIH PO1 AGO5131.

Replication analysis in the Religious Orders Study and Rush Memory and Aging Project cohorts was supported by grants from the National Institutes of Health [R01 AG30146, P30 AG10161, R01 AG17917, R01 AG15819, K08 AG034290], the Illinois Department of Public Health, and the Burroughs Wellcome Fund.

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Abbott; Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Amorfix Life Sciences Ltd.; AstraZeneca; Bayer HealthCare; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is Rev March 26, 2012 coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of California, Los Angeles. This research was also supported by NIH grants P30 AG010129 and K01 AG030514.

References

Accardi G, Caruso C, Colonna-Romano G, Camarda C, Monastero R, Candore G. Can Alzheimer disease be a form of type 3 diabetes? Rejuvenation research. 2012; 15:217–221. [PubMed: 22533436]

Andersson C, Blennow K, Almkvist O, Andreasen N, Engfeldt P, Johansson SE, Lindau M, Eriksdotter-Jonhagen M. Increasing CSF phospho-tau levels during cognitive decline and progression to dementia. Neurobiol Aging. 2007

Andrews-Zwilling Y, Bien-Ly N, Xu Q, Li G, Bernardo A, Yoon SY, Zwilling D, Yan TX, Chen L, Huang Y. Apolipoprotein E4 causes age- and Tau-dependent impairment of GABAergic interneurons, leading to learning and memory deficits in mice. J Neurosci. 2010; 30:13707–13717. [PubMed: 20943911]

Barker A, Sharp SJ, Timpson NJ, Bouatia-Naji N, Warrington NM, Kanoni S, Beilin LJ, Brage S, Deloukas P, Evans DM, et al. Association of genetic Loci with glucose levels in childhood and adolescence: a meta-analysis of over 6,000 children. Diabetes. 2011; 60:1805–1812. [PubMed: 21515849]

Bertram, L.; McQueen, M.; Mullin, K.; Blacker, D.; Tanzi, R. The AlzGene Database. Alzheimer Research Forum; Available at: http://www.alzgene.org [Accessed 1/26/2013]

Brecht WJ, Harris FM, Chang S, Tesseur I, Yu GQ, Xu Q, Dee Fish J, Wyss-Coray T, Buttini M, Mucke L, et al. Neuron-specific apolipoprotein e4 proteolysis is associated with increased tau phosphorylation in brains of transgenic mice. J Neurosci. 2004; 24:2527–2534. [PubMed: 15014128]

Buerger K, Ewers M, Pirttila T, Zinkowski R, Alafuzoff I, Teipel SJ, DeBernardis J, Kerkman D, McCulloch C, Soininen H, Hampel H. CSF phosphorylated tau protein correlates with neocortical neurofibrillary pathology in Alzheimer's disease. Brain. 2006; 129:3035–3041. [PubMed: 17012293]

Castellano JM, Kim J, Stewart FR, Jiang H, DeMattos RB, Patterson BW, Fagan AM, Morris JC, Mawuenyega KG, Cruchaga C, et al. Human apoE isoforms differentially regulate brain amyloid-beta peptide clearance. Sci Transl Med. 2011; 3:89ra57.

Conde L, Vaquerizas JM, Dopazo H, Arbiza L, Reumers J, Rousseau F, Schymkowitz J, Dopazo J. PupaSuite: finding functional single nucleotide polymorphisms for large-scale genotyping purposes. Nucleic Acids Res. 2006; 34:W621–625. [PubMed: 16845085]

Cruchaga C, Fernandez-Seara MA, Seijo-Martinez M, Samaranch L, Lorenzo E, Hinrichs A, Irigoyen J, Maestro C, Prieto E, Marti-Climent JM, et al. Cortical atrophy and language network reorganization associated with a novel progranulin mutation. Cereb Cortex. 2009; 19:1751–1760. [PubMed: 19020205]

Cruchaga C, Kauwe JS, Mayo K, Spiegel N, Bertelsen S, Nowotny P, Shah AR, Abraham R, Hollingworth P, Harold D, et al. SNPs associated with cerebrospinal fluid phospho-tau levels influence rate of decline in Alzheimer's disease. PLoS Genet. 2010; 6:e1001101. pii. [PubMed: 20862329]

Cruchaga C, Kauwe JS, Nowotny P, Bales K, Pickering EH, Mayo K, Bertelsen S, Hinrichs A, Fagan AM, Holtzman DM, et al. Cerebrospinal fluid APOE levels: an endophenotype for genetic studies for Alzheimer's disease. Hum Mol Genet. 2012

Cruchaga C, Nowotny P, Kauwe JS, Ridge PG, Mayo K, Bertelsen S, Hinrichs A, Fagan AM, Holtzman DM, Morris JC, Goate AM. Association and Expression Analyses With Single-Nucleotide Polymorphisms in TOMM40 in Alzheimer Disease. Arch Neurol-Chicago. 2011; 68:1013–1019. [PubMed: 21825236]

De Jager PL, Shulman JM, Chibnik LB, Keenan BT, Raj T, Wilson RS, Yu L, Leurgans SE, Tran D, Aubin C, et al. A genome-wide scan for common variants affecting the rate of age-related cognitive decline. Neurobiol Aging. 2012; 33:1017 e1011–1015. [PubMed: 22054870]

de Leon MJ, DeSanti S, Zinkowski R, Mehta PD, Pratico D, Segal S, Clark C, Kerkman D, DeBernardis J, Li J, et al. MRI and CSF studies in the early diagnosis of Alzheimer's disease. J Intern Med. 2004; 256:205–223. [PubMed: 15324364]

Dimitri P, Warner JT, Minton JA, Patch AM, Ellard S, Hattersley AT, Barr S, Hawkes D, Wales JK, Gregory JW. Novel GLIS3 mutations demonstrate an extended multisystem phenotype. European journal of endocrinology / European Federation of Endocrine Societies. 2011; 164:437–443. [PubMed: 21139041]

Dube MP, Schmidt S, Hauser E, Darabi H, Li J, Barhdadi A, Wang X, Sha Q, Zhang Z, Wang T, et al. Multistage designs in the genomic era: providing balance in complex disease studies. Genet Epidemiol. 2007; 31(Suppl 1):S118–123. [PubMed: 18046769]

Fagan AM, Mintun MA, Mach RH, Lee SY, Dence CS, Shah AR, LaRossa GN, Spinner ML, Klunk WE, Mathis CA, et al. Inverse relation between in vivo amyloid imaging load and cerebrospinal fluid Abeta42 in humans. Ann Neurol. 2006; 59:512–519. [PubMed: 16372280]

Farias G, Cornejo A, Jimenez J, Guzman L, Maccioni RB. Mechanisms of tau self-aggregation and neurotoxicity. Curr Alzheimer Res. 2011; 8:608–614. [PubMed: 21605046]

Gibb GM, Pearce J, Betts JC, Lovestone S, Hoffmann MM, Maerz W, Blackstock WP, Anderton BH. Differential effects of apolipoprotein E isoforms on phosphorylation at specific sites on tau by glycogen synthase kinase-3 beta identified by nano-electrospray mass spectrometry. FEBS Lett. 2000; 485:99–103. [PubMed: 11094148]

Guerreiro R, Wojtas A, Bras J, Carrasquillo M, Rogaeva E, Majounie E, Cruchaga C, Sassi C, Kauwe JS, Younkin S, et al. TREM2 Variants in Alzheimer's Disease. N Engl J Med. 2012

Hampel H, Buerger K, Zinkowski R, Teipel SJ, Goernitz A, Andreasen N, Sjoegren M, DeBernardis J, Kerkman D, Ishiguro K, et al. Measurement of phosphorylated tau epitopes in the differential diagnosis of Alzheimer disease: a comparative cerebrospinal fluid study. Arch Gen Psychiatry. 2004; 61:95–102. [PubMed: 14706948]

Han MR, Schellenberg GD, Wang LS. Genome-wide association reveals genetic effects on human Abeta42 and tau protein levels in cerebrospinal fluids: a case control study. BMC neurology. 2010; 10:90. [PubMed: 20932310]

Hardy J, Selkoe DJ. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. Science. 2002; 297:353–356. [PubMed: 12130773]

Hollingworth P, Harold D, Sims R, Gerrish A, Lambert JC, Carrasquillo MM, Abraham R, Hamshere ML, Pahwa JS, Moskvina V, et al. Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. Nature Genetics. 2011; 43:429–435. [PubMed: 21460840]

Holmans P, Green EK, Pahwa JS, Ferreira MA, Purcell SM, Sklar P, Owen MJ, O'Donovan MC, Craddock N. Gene ontology analysis of GWA study data sets provides insights into the biology of bipolar disorder. Am J Hum Genet. 2009; 85:13–24. [PubMed: 19539887]

Hu WT, Trojanowski JQ, Shaw LM. Biomarkers in frontotemporal lobar degenerations--progress and challenges. Progress in neurobiology. 2011; 95:636–648. [PubMed: 21554923]

Jagust WJ, Landau SM, Shaw LM, Trojanowski JQ, Koeppe RA, Reiman EM, Foster NL, Petersen RC, Weiner MW, Price JC, Mathis CA. Relationships between biomarkers in aging and dementia. Neurology. 2009; 73:1193–1199. [PubMed: 19822868]

Jia JP, Meng R, Sun YX, Sun WJ, Ji XM, Jia LF. Cerebrospinal fluid tau, Abeta1-42 and inflammatory cytokines in patients with Alzheimer's disease and vascular dementia. Neurosci Lett. 2005; 383:12–16. [PubMed: 15936505]

Jones L, Holmans PA, Hamshere ML, Harold D, Moskvina V, Ivanov D, Pocklington A, Abraham R, Hollingworth P, Sims R, et al. Genetic evidence implicates the immune system and cholesterol metabolism in the aetiology of Alzheimer's disease. PLoS One. 2010; 5:e13950. [PubMed: 21085570]

Kauwe JS, Cruchaga C, Bertelsen S, Mayo K, Latu W, Nowotny P, Hinrichs AL, Fagan AM, Holtzman DM, Goate AM. Validating predicted biological effects of Alzheimer's disease associated SNPs using CSF biomarker levels. J Alzheimers Dis. 2010; 21:833–842. [PubMed: 20634593]

Kauwe JS, Cruchaga C, Karch CM, Sadler B, Lee M, Mayo K, Latu W, Su'a M, Fagan AM, Holtzman DM, et al. Fine mapping of genetic variants in BIN1, CLU, CR1 and PICALM for association with cerebrospinal fluid biomarkers for Alzheimer's disease. PLoS One. 2011; 6:e15918. [PubMed: 21347408]

Kauwe JS, Cruchaga C, Mayo K, Fenoglio C, Bertelsen S, Nowotny P, Galimberti D, Scarpini E, Morris JC, Fagan AM, et al. Variation in MAPT is associated with cerebrospinal fluid tau levels in the presence of amyloid-beta deposition. Proc Natl Acad Sci U S A. 2008; 105:8050–8054. [PubMed: 18541914]

Kauwe JS, Jacquart S, Chakraverty S, Wang J, Mayo K, Fagan AM, Holtzman DM, Morris JC, Goate AM. Extreme cerebrospinal fluid amyloid beta levels identify family with late-onset Alzheimer's disease presenilin 1 mutation. Ann Neurol. 2007; 61:446–453. [PubMed: 17366635]

Kauwe JS, Wang J, Mayo K, Morris JC, Fagan AM, Holtzman DM, Goate AM. Alzheimer's disease risk variants show association with cerebrospinal fluid amyloid beta. Neurogenetics. 2009; 10:13–17. [PubMed: 18813964]

Kawarabayashi T, Younkin LH, Saido TC, Shoji M, Ashe KH, Younkin SG. Age-dependent changes in brain, CSF, and plasma amyloid (beta) protein in the Tg2576 transgenic mouse model of Alzheimer's disease. J Neurosci. 2001; 21:372–381. [PubMed: 11160418]

Keenan BT, Shulman JM, Chibnik LB, Raj T, Tran D, Sabuncu MR, Allen AN, Corneveaux JJ, Hardy JA, Huentelman MJ, et al. A coding variant in CR1 interacts with APOE-epsilon4 to influence cognitive decline. Hum Mol Genet. 2012; 21:2377–2388. [PubMed: 22343410]

Kim S, Swaminathan S, Shen L, Risacher SL, Nho K, Foroud T, Shaw LM, Trojanowski JQ, Potkin SG, Huentelman MJ, et al. Genome-wide association study of CSF biomarkers Abeta1-42, t-tau, and p-tau181p in the ADNI cohort. Neurology. 2011; 76:69–79. [PubMed: 21123754]

Koch W, Ehrenhaft A, Griesser K, Pfeufer A, Muller J, Schomig A, Kastrati A. TaqMan systems for genotyping of disease-related polymorphisms present in the gene encoding apolipoprotein E. Clin Chem Lab Med. 2002; 40:1123–1131. [PubMed: 12521230]

Kraft P, Cox DG. Study designs for genome-wide association studies. Advances in genetics. 2008; 60:465–504. [PubMed: 18358330]

Marcus JN, Schachter J. Targeting post-translational modifications on tau as a therapeutic strategy for Alzheimer's disease. Journal of neurogenetics. 2011; 25:127–133. [PubMed: 22091726]

McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology. 1984; 34:939–944. [PubMed: 6610841]

Myers AJ, Gibbs JR, Webster JA, Rohrer K, Zhao A, Marlowe L, Kaleem M, Leung D, Bryden L, Nath P, et al. A survey of genetic human cortical gene expression. Nat Genet. 2007; 39:1494–1499. [PubMed: 17982457]

Naj AC, Jun G, Beecham GW, Wang LS, Vardarajan BN, Buros J, Gallins PJ, Buxbaum JD, Jarvik GP, Crane PK, et al. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. Nature Genetics. 2011; 43:436–441. [PubMed: 21460841]

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. [PubMed: 16862161]

Price JL, Morris JC. Tangles and plaques in nondemented aging and "preclinical" Alzheimer's disease. Ann Neurol. 1999; 45:358–368. [PubMed: 10072051]

Ringman JM, Younkin SG, Pratico D, Seltzer W, Cole GM, Geschwind DH, Rodriguez-Agudelo Y, Schaffer B, Fein J, Sokolow S, et al. Biochemical markers in persons with preclinical familial Alzheimer disease. Neurology. 2008; 71:85–92. [PubMed: 18509095]

Rohlfs RV, Taylor C, Mirea L, Bull SB, Corey M, Anderson AD. One-stage design is empirically more powerful than two-stage design for family-based genome-wide association studies. BMC Proc. 2007; 1(Suppl 1):S137. [PubMed: 18466480]

Schneider JA, Aggarwal NT, Barnes L, Boyle P, Bennett DA. The neuropathology of older persons with and without dementia from community versus clinic cohorts. J Alzheimers Dis. 2009; 18:691–701. [PubMed: 19749406]

Schoonenboom NS, Mulder C, Van Kamp GJ, Mehta SP, Scheltens P, Blankenstein MA, Mehta PD. Amyloid beta 38, 40, and 42 species in cerebrospinal fluid: more of the same? Ann Neurol. 2005; 58:139–142. [PubMed: 15984010]

Shoji M, Matsubara E, Kanai M, Watanabe M, Nakamura T, Tomidokoro Y, Shizuka M, Wakabayashi K, Igeta Y, Ikeda Y, et al. Combination assay of CSF tau, A beta 1-40 and A beta 1-42(43) as a biochemical marker of Alzheimer's disease. J Neurol Sci. 1998; 158:134–140. [PubMed: 9702683]

Strozyk D, Blennow K, White LR, Launer LJ. CSF Abeta 42 levels correlate with amyloid-neuropathology in a population-based autopsy study. Neurology. 2003; 60:652–656. [PubMed: 12601108]

Sunderland T, Linker G, Mirza N, Putnam KT, Friedman DL, Kimmel LH, Bergeson J, Manetti GJ, Zimmermann M, Tang B, et al. Decreased beta-amyloid1-42 and increased tau levels in cerebrospinal fluid of patients with Alzheimer disease. Jama. 2003; 289:2094–2103. [PubMed: 12709467]

Webster JA, Gibbs JR, Clarke J, Ray M, Zhang W, Holmans P, Rohrer K, Zhao A, Marlowe L, Kaleem M, et al. Genetic control of human brain transcript expression in Alzheimer disease. Am J Hum Genet. 2009; 84:445–458. [PubMed: 19361613]

Welge V, Fiege O, Lewczuk P, Mollenhauer B, Esselmann H, Klafki HW, Wolf S, Trenkwalder C, Otto M, Kornhuber J, et al. Combined CSF tau, p-tau181 and amyloid-beta 38/40/42 for diagnosing Alzheimer's disease. J Neural Transm. 2009; 116:203–212. [PubMed: 19142572]

Yang J, Manolio TA, Pasquale LR, Boerwinkle E, Caporaso N, Cunningham JM, de Andrade M, Feenstra B, Feingold E, Hayes MG, et al. Genome partitioning of genetic variation for complex traits using common SNPs. Nat Genet. 2011; 43:519–525. [PubMed: 21552263]

Yoshida T, Shiroshima T, Lee SJ, Yasumura M, Uemura T, Chen X, Iwakura Y, Mishina M. Interleukin-1 receptor accessory protein organizes neuronal synaptogenesis as a cell adhesion molecule. J Neurosci. 2012; 32:2588–2600. [PubMed: 22357843]

Zhou J, Chen J, Feng Y. Effect of truncated-ApoE4 overexpression on tau phosphorylation in cultured N2a cells. Journal of Huazhong University of Science and Technology. Medical sciences = Hua zhong ke ji da xue xue bao. Yi xue Ying De wen ban = Huazhong keji daxue xuebao. Yixue Yingdewen ban. 2006; 26:272–274.

Zou F, Chai HS, Younkin CS, Allen M, Crook J, Pankratz VS, Carrasquillo MM, Rowley CN, Nair AA, Middha S, et al. Brain expression genome-wide association study (eGWAS) identifies human disease-associated variants. PLoS Genet. 2012; 8:e1002707. [PubMed: 22685416]

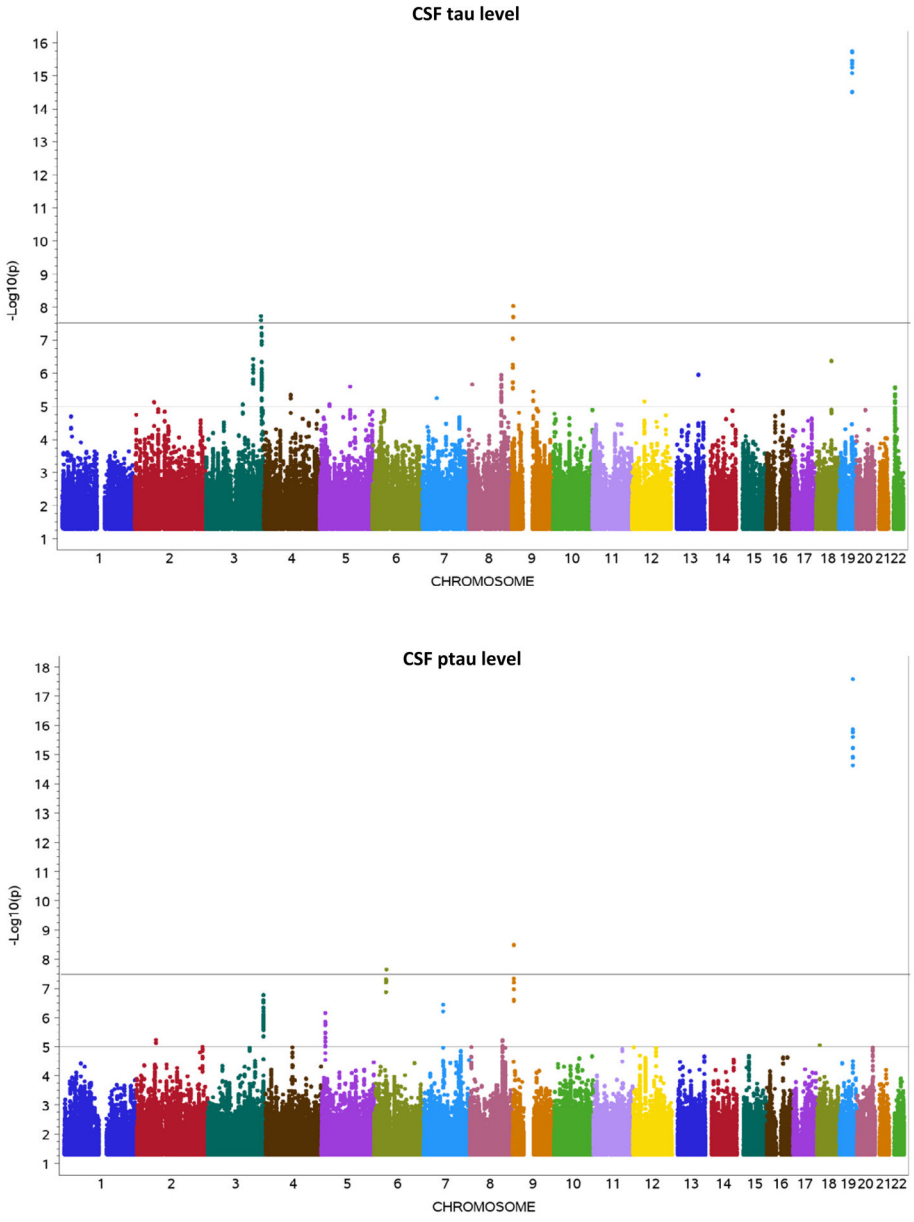

Figure 1. Genome-wide signal intensity (Manhattan) plots showing the individual *P* values (based on fixed-effects meta-analysis) against genomic position

The results for the association of CSF tau (a), and ptau (b) levels with 5,815,690 SNPs are shown. Within each chromosome, shown on the *x* axis, the results are plotted left to right from the p-terminal end. Horizontal dashed lines indicate *P* value thresholds of 1×10^{-5} and 5×10^{-8} (genome-wide significance).

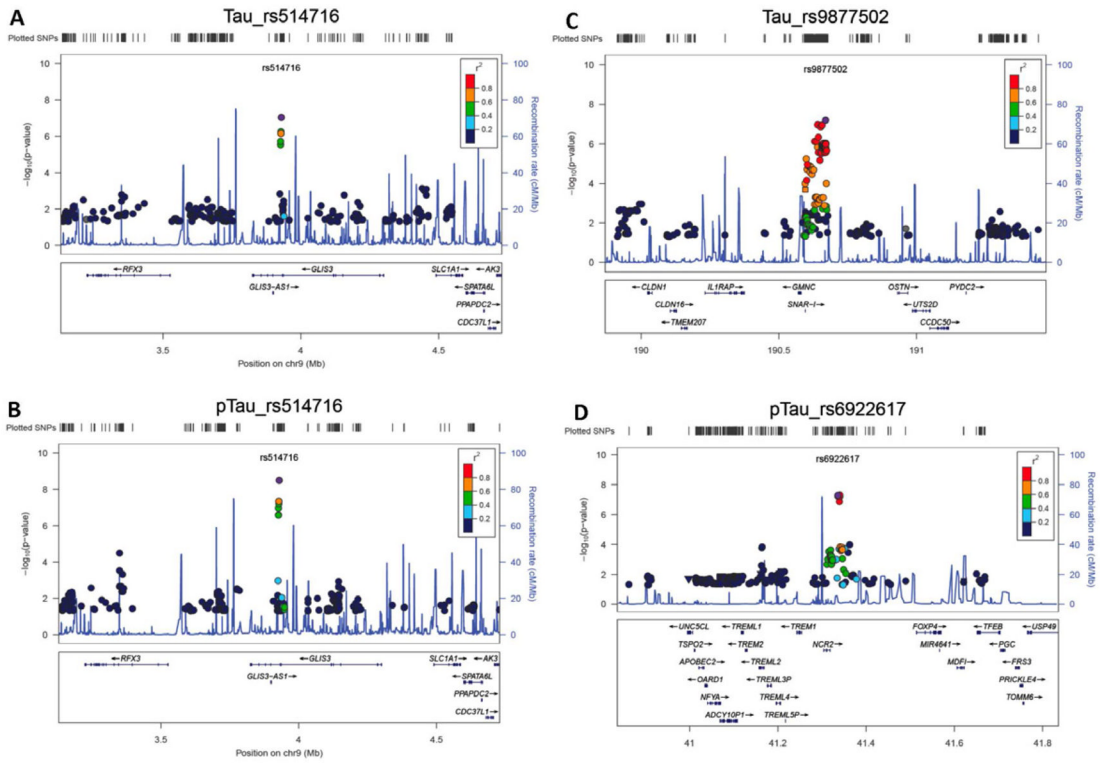

Figure 2. Regional plots for associations with CSF tau and ptau at genome-wide significance

Plots are centered on the most significant SNP at a given locus along with the combined-analysis results for SNPs in the region surrounding it (typically ± 400 kb). Symbols are colored according to the LD of the SNP with the top SNP. The light blue line represents the estimated recombination rate. Gene annotations are shown as dark green line.

Table 1

Summary of sample characteristics

	Knight-ADRC	ADNI	UW	UPenn
N	501	394	323	51
Age (years) Mean ± SD (range)	69 ± 9 (46–91)	75 ± 6.9 (55–91)	67 ± 10 (45–88)	67 ± 9 (50–86)
APOE ε4+ (%)	37	49	43	61
CDR 0 (%)	73	27	61	2
Male (%)	42	60	48	33
ptau	64 ± 35 (17–229)	34 ± 18 (7–115)	37 ± 17 (5–83)	35 ± 24 (6–116)
Tau	355 ± 229 (85–1624)	98 ± 57 (28–495)	77 ± 35 (27–204)	81 ± 60 (17–351)
$A\beta_{42}$	587± 247 (154–1293)	170 ± 56 (53–300)	216 ± 73 (64–366)	177 ± 59 (78–298)

Age at lumbar puncture (LP), percentage of males, percentage of *APOE4* allele carriers, and clinical dementia rating (CDR) at LP date for each sample. For each phenotype the mean in pg/ml with the standard deviation and range is shown. Charles F. and Joanne Knight Alzheimer's Disease Research Center at University of Washington (Knight-ADRC), Alzheimer's Disease Neuroimaging Initiative (ADNI) and for the University of Washington, Seattle (UW). Cerebrospinal Fluid (CSF). Case-control (CC).

Table 2

Genome-wide significant SNPs for CSF tau, and ptau.

CHR	SNP	MAF	Closest gene	SNP type/location	tau	ptau	$A\beta_{42}$	SNP type
19	rs769449	0.190	APOE	intron	**1.95×10^{-16}**	**2.56×10^{-18}**	**9.02×10^{-47}**	Imputed
19	rs12972970	0.193	PVRL2	intron	**8.24×10^{-16}**	**1.28×10^{-15}**	**3.74×10^{-40}**	Imputed
19	rs34342646	0.193	PVRL2	intron	**5.54×10^{-16}**	**1.17×10^{-15}**	**4.17×10^{-40}**	Imputed
19	rs34404554	0.208	TOMM40	intron	**3.58×10^{-16}**	**1.33×10^{-16}**	**1.01×10^{-39}**	Imputed
19	rs11556505	0.210	TOMM40	Synonymous	**3.48×10^{-16}**	**1.68×10^{-16}**	**1.06×10^{-39}**	Imputed
19	rs2075650	0.213	TOMM40	intron	**4.28×10^{-16}**	**5.81×10^{-16}**	**2.21×10^{-39}**	Typed
19	rs71352238	0.201	TOMM40	intergenic	**1.78×10^{-16}**	**2.46×10^{-16}**	**5.68×10^{-39}**	Imputed
19	rs12972156	0.187	PVRL2	intron	**3.03×10^{-15}**	**2.31×10^{-15}**	**6.35×10^{-37}**	Imputed
3	rs9877502	0.386	SNAR-I	intergenic	**4.98×10^{-09}**	1.68×10^{-07}	0.022	Imputed
3	rs1316356	0.362	SNAR-I	intergenic	2.80×10^{-07}	1.96×10^{-07}	0.102	Typed
9	rs514716	0.136	GLIS3	intron	**1.07×10^{-08}**	**3.22×10^{-09}**	0.026	Imputed
9	rs622536	0.114	GLIS3	intron	6.68×10^{-07}	**4.50×10^{-08}**	0.029	Imputed
9	rs59860681	0.114	GLIS3	intron	6.68×10^{-07}	**4.50×10^{-08}**	4.66×10^{-3}	Imputed
9	rs623295	0.114	GLIS3	intron	6.68×10^{-07}	**4.50×10^{-08}**	4.66×10^{-3}	Imputed
9	rs624290	0.114	GLIS3	intron	6.68×10^{-07}	**4.50×10^{-08}**	4.66×10^{-3}	Typed
6	rs6922617	0.064	NCR2	intergenic	2.55×10^{-05}	**3.58×10^{-08}**	3.69×10^{-3}	Typed
6	rs11966476	0.067	NCR2	intergenic	1.59×10^{-05}	**4.97×10^{-08}**	3.99×10^{-3}	Imputed

Standardized log-transformed CSF tau and ptau values were tested for association with the SNP in an additive model using PLINK, including age, gender, site and the third first component factors for population stratification.

The table only shows the genome-wide significant SNPs or the most significant genotyped SNPs for each locus.

All novel loci that show association with CSF tau and/or ptau show a weak association with CSF $A\beta_{42}$ levels

Table 3

Genome-partitioning analysis

	Default model			Without_*APOE*_SNPs[A]			Without_*chr 3*_SNPs[A]			Without_*chr 6*_SNPs[A]			Without_*chr 9*_SNPs[A]			Without_TopLoci[A]		
	Vp	**V(G)**	**V(G)/Vp**	**V(G)**	**V(G)/Vp**	**dif**	**V(G)**	**V(G)/Vp**	**dif**	**V(G)**	**V(G)/Vp**	**dif**	**V(G)**	**V(G)/Vp**	**dif**	**V(G)**	**V(G)/Vp**	**dif**
ptau	0.0473	0.0032	6.79%	0.0031	6.45%	0.33%	0.0028	5.93%	0.85%	0.0031	6.65%	0.14%	0.0031	5.84%	0.14%	0.0025	5.33%	1.45%
tau	0.0550	0.0083	15.14%	0.0082	14.86%	0.28%	0.0079	14.32%	0.82%	0.0083	15.07%	0.08%	0.0082	15.04%	0.96%	0.0076	13.86%	1.28%

We used the algorithm GCTA (Genome-wide Complex Trait Analysis) to estimate the proportion of phenotypic variance explained by genome-wide SNPs[16].

[A]The SNPs located in each locus with a p-value lower than 1×10^{-5} were excluded from the analysis. A total of 13 SNPs were excluded for the APOE locus, 81 for chr3, 14 for chr6 and 9 for chr 9.

Table 4

Most but not all of the CSF ptau association with *APOE* is driven by $A\beta_{42}$ levels

		Default model*		+Aβ42 in the model		+CDR in the model		+ APOE in the model	
CHR	**SNP**	**p-tau**	**tau**	**ptau**	**tau**	**ptau**	**tau**	**ptau**	**tau**
19	rs769449	**2.57E-18**	**1.95E-16**	3.22E-05	6.39E-06	**1.23E-12**	**1.04E-10**	1.29E-02	9.07E-03
19	rs34404554	**1.33E-16**	**8.24E-16**	2.91E-05	1.72E-06	**1.87E-12**	**3.19E-10**	1.43E-02	1.34E-02
19	rs11556505	**1.68E-16**	**5.54E-16**	3.86E-05	1.58E-06	**2.83E-12**	**3.00E-10**	1.77E-02	1.55E-02
19	rs71352238	**2.46E-16**	**3.58E-16**	5.33E-05	1.58E-06	**1.13E-11**	**2.04E-10**	1.37E-02	9.46E-03
19	rs2075650	**5.83E-16**	**3.48E-16**	8.82E-05	8.22E-07	**1.01E-11**	**3.88E-10**	2.24E-02	2.03E-02
19	rs34342646	**1.18E-15**	**4.28E-16**	9.36E-05	1.04E-06	**7.99E-11**	**5.87E-10**	1.25E-02	9.58E-03
19	rs12972970	**1.28E-15**	**1.78E-16**	9.96E-05	1.04E-06	**7.83E-11**	**5.94E-10**	1.25E-02	9.74E-03
19	rs12972156	**2.31E-15**	**3.03E-15**	6.72E-05	1.53E-06	**1.21E-10**	**1.10E-09**	9.26E-03	8.37E-03
3	rs9877502	1.68E-07	**4.98E-09**	5.62E-07	7.24E-07	2.47E-07	**4.39E-08**	4.60E-07	**7.77E-09**
9	rs514716	**2.99E-09**	**1.07E-08**	4.14E-07	7.43E-07	**3.76E-08**	**1.38E-08**	**2.00E-08**	2.30E-07
6	rs6922617	**3.58E-08**	2.55E-05	2.34E-06	4.03E-04	3.49E-07	1.22E-04	1.66E-06	4.61E-05

*age, gender, series and PC are included in all the analysis as covariates

Bold numbers represent p-values that pass the genome-wide significant threshold

Table 5

Association of the *TREM2* gene cluster with CSF tau and ptau levels

	MAF	n	Tau. Pvalue (beta)				Ptau. Pvalue (beta)			
			Default model	Cond. on R47H	Cond. on rs6916710	Cond. on rs6922617	Default model	Cond. on R47H	Cond. on rs6916710	Cond. on rs6922617
TREM2-R47H	0.01	815	6.9×10^{-4} (0.19)	-	1.3×10^{-3} (0.18)	5.3×10^{-4} (0.19)	2.6×10^{-3} (0.16)	-	4.7×10^{-3} (0.15)	2.0×10^{-3} (0.16)
rs6916710	0.39	815	8.2×10^{-3} (−0.03)	0.015 (−0.03)	-	9.3×10^{-3} (−0.03)	6.4×10^{-3} (−0.03)	0.011 (−0.03)	-	7.0×10^{-3} (−0.03)
rs6922617	0.06	815	4.1×10^{-5} (−0.10)	3.2×10^{-5} (−0.10)	4.6×10^{-5} (−0.09)	-	6.1×10^{-6} (−0.10)	4.9×10^{-6} (−0.10)	6.5×10^{-6} (−0.10)	-

The *TREM2 R47H* variant (rs75932628) was genotyped in the Knight-ADRC and ADNI series by Sequenom. The association of *TREM* gene cluster variant with CSF levels was performed with PLINK including, age, gender, series and principal component factors as covariates.

Table 6

Association of the top loci for CSF tau and ptau with risk for AD, tangle counts and cognitive decline

				CSF (n=1,269)		Risk for AD (n=22,771)		Tangles (n=651)		Global cognitive decline (n=1,593)	
Chr.	**Rs #**	**Phen.**	**Minor Allele**	**beta**	**p-value**	**beta**	**p-value**	**beta**	**p-value**	**beta**	**p-value**
19	**rs2075650**	tau	G	0.074	**4.29×10^{-16}**	0.985	**6.22×10^{-184}**	0.266	**9.33×10^{-10}**	−0.049	**1.22×10^{-11}**
		ptau	G	0.081	**4.92×10^{-16}**						
3	**rs9877502**	tau	A	0.052	**4.98×10^{-09}**	0.081*	2.67×10^{-04}*	0.055	0.014	−0.014	4.64×10^{-05}
		ptau	A	0.050	**1.68×10^{-07}**						
9	**rs514716**	tau	G	−0.071	**1.07×10^{-08}**	−0.022	0.518	−0.036	0.360	0.004	0.479
		ptau	G	−0.072	**2.99×10^{-09}**						
6	**rs6922617**	tau	A	−0.094	2.55×10^{-05}	−0.0174	0.677	−0.029	0.547	0.007	0.330
		ptau	A	−0.093	**3.58×10^{-08}**						

The association of the most significant SNPs for CSF tau and ptau for risk for AD, tangle count or cognitive decline is shown.

P-values and direction for risk for disease were extracted from the previously published GWAS for AD(Naj et al., 2011).

Association with tangle counts and memory decline was performed in the Religious Orders Study (ROS) and the Rush Memory and Aging Project (MAP) as previously reported(De Jager et al., 2012; Keenan et al., 2012).

* The p-value for risk for disease is for rs1316356, which is in high LD with rs9877502.

In all cases the beta is calculated in reference to the minor allele.