HHS Public Access

Author manuscript

Alzheimers Dement. Author manuscript; available in PMC 2024 February 01.

Published in final edited form as:

Alzheimers Dement. 2023 February ; 19(2): 611–620. doi:10.1002/alz.12686.

Genetic variants in the *SHISA6* gene are associated with delayed cognitive impairment in two family datasets

Jairo Ramos[1], **Laura J. Caywood**[1], **Michael B. Prough**[1], **Jason E. Clouse**[1], **Sharlene D. Herington**[1], **Susan H. Slifer**[1], **M. Denise Fuzzell**[2], **Sarada L. Fuzzell**[2], **Sherri D. Hochstetler**[2], **Kristy L. Miskimen**[2], **Leighanne R. Main**[2], **Michael D. Osterman**[2,5], **Andrew F. Zaman**[1], **Patrice L. Whitehead**[1], **Larry D. Adams**[1], **Renee A. Laux**[2], **Yeunjoo E. Song**[2], **Tatiana M. Foroud**[6], **Richard P. Mayeux**[7,8,9], **Peter St. George-Hyslop**[10], **Paula K. Ogrocki**[3], **Alan J. Lerner**[3], **Jeffery M. Vance**[1,4], **Michael L. Cuccaro**[1,4], **Jonathan L. Haines**[2,5], **Margaret A. Pericak-Vance**[1,4], **William K. Scott**[1,4]

[1]John P. Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami, FL, USA

[2]Case Western Reserve University School of Medicine, Cleveland, OH, USA

[3]University Hospitals Cleveland Medical Center, Cleveland, OH, USA

[4]The Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami Miller School of Medicine, Miami, FL, USA

[5]Cleveland Institute for Computational Biology, Case Western Reserve University, Cleveland, OH, USA

[6]Indiana Alzheimer's Disease Center, Indiana University School of Medicine, Indianapolis, IN, USA

[7]Taub Institute on Alzheimer's Disease and the Aging Brain, Department of Neurology, Columbia University, New York, NY, USA

[8]Gertrude H. Sergievsky Center, Columbia University, New York, NY, USA

[9]Department of Neurology, Columbia University, New York, NY, USA

[10]University of Toronto, Toronto, ON, Canada.

Abstract

INTRODUCTION: Studies of cognitive impairment (CI) in Amish communities have identified sibships containing CI and cognitively unimpaired (CU) individuals. We hypothesize that CU individuals may carry protective alleles delaying age at onset (AAO) of CI.

Corresponding author: William K. Scott, Address: John P. Hussman Institute for Human Genomics, University of Miami Miller School of Medicine. 1501 NW 10th Avenue. Room 414. Miami, FL 33136. w.scott@med.miami.edu.

Conflict of Interest

Jairo Ramos, Laura J. Caywood, Michael B. Prough, Jason E. Clouse, Sharlene D. Herington, Susan H. Slifer, M. Denise Fuzzell, Sarada L. Fuzzell, Sherri D. Hochstetler, Kristy L. Miskimen, Leighanne R. Main, Michael D. Osterman, Andrew F. Zaman, Patrice L. Whitehead, Larry D. Adams, Renee A. Laux, Yeunjoo E. Song and Paula K. Ogrocki, have no conflict of interest to report.

METHODS: 1,522 individuals screened for CI were genotyped. The outcome studied was AAO for CI individuals or age at last normal exam for CU individuals. Cox mixed-effects models examined association between age and single nucleotide variants (SNV).

RESULTS: Three SNVs were significantly associated ($p<5 \times 10^{-8}$) with AAO on chromosomes 6 (rs14538074; Hazard Ratio (HR)=3.35), 9 (rs534551495; HR=2.82), and 17 (rs146729640; HR=6.38). The chromosome 17 association was replicated in the independent NIA-LOAD dataset.

DISCUSSION: The replicated genome-wide significant association with AAO on chromosome 17 is located in the *SHISA6* gene, which is involved in post-synaptic transmission in the hippocampus and is a biologically plausible candidate gene for Alzheimer's disease (AD).

Keywords

cognitive impairment; age at onset; genome wide association study; Cox mixed-effects models; *SHISA6* gene; Alzheimer's disease; protective variants

1. BACKGROUND

Dementia, which includes Alzheimer disease (AD), cerebrovascular, Lewy body, and mixed dementias [1], has a tremendous effect on the entire family, since the largest proportion, at least 60 % in the U.S., of caregivers are spouses, children and children-in-law [2]. Dementia also carries a heavy emotional and financial toll, as people can live for many years cognitively impaired, requiring high levels of care with annual estimated costs in the U.S. in 2010 between $159 billion and $215 billion [1,3]. Understanding the factors that influence the age at which cognitive impairment begins is foundational in developing approaches toward reducing the prevalence and progression of disease. Much effort has been made to identify genetic factors that increase the risk and result in earlier onset of dementia, and dozens of risk variants have been identified. The search for protective variants, including those that delay age at onset (AAO), has been less common. Studies of time to onset of dementia symptoms, employing survival analysis approaches, offer a potentially powerful means of identifying protective variants in age related neurodegenerative diseases that shorten the duration of cognitive impairment (CI) by delaying onset of symptoms. In complex diseases such as AD and other neurodegenerative disorders that cause dementia, the identification of both risk and protective variants is essential to elucidate the molecular mechanisms involved in disease development and progression. This is a crucial step in the discovery of new therapeutic targets [4]. In this study, we aim to identify protective variants associated with delay onset of dementia using survival analysis.

The complex genetic architecture of dementia and AD has complicated the search for both risk and protective variants. Reproducibility of genetic association is influenced by locus heterogeneity, smaller sample sizes, and population differences. Some of these limitations are addressed by large-scale consortium approaches to genome-wide association studies (GWAS) [5]. A second strategy to improve power and reduce heterogeneity is to study a genetically homogenous population such as the Amish [6]. The Amish in the United States live in relatively genetically isolated communities [6]; members of these communities have participated in longitudinal studies of aging, cognition, and cognitive impairment for over

two decades [6,7,8,9,10,11,12] which have identified sibships with both individuals with CI and siblings that are cognitively unimpaired (CU). Our hypothesis is that individuals that are CU carry genetic variants that delay AAO of CI despite having an increased risk due to advanced age (over age 75) and family history of CI (including 18.1% with siblings having CI).

In this study we performed a GWAS of AAO of CI in 1,522 people from the Amish communities living in Ohio and Indiana. Genome-wide significant results were examined for replication in an independent sample of 3,350 individuals from the NIA-LOAD family study of AD. A replicated association on chromosome 17 was identified, supporting the hypothesis that genetic variants delay the development of cognitive impairment in individuals who are CU despite elevated risk due to age and family history of dementia.

2. METHODS

2.1 Study Populations and Clinical Assessment

A description of the sample of 1,522 Amish used in this study is provided in Table 1. All participants were of European descent, with an average age of 73 at the last examination (range: 43 to 99). Women were 58.3% of the sample. Individuals were residents of Amish communities in Ohio (Holmes County) and Indiana (Adams, Elkhart, and LaGrange counties) and recruited as previously described by Pericak-Vance et al. [7], Hahs et al. [6], Edwards et al. [8], Edwards et al. [10], and Ramos et al. [12]. This is part of a population-based study of aging and age-related phenotypes such as AD performed by researchers from the University of Miami and Case Western Reserve University after Institutional Review Board (IRB) review and approval.

The cohort selected for this research come from three main studies. For the first study, we enrolled individuals from 1994 to 2017 in family-based studies of dementia and AD. Participants over 65 were initially identified by a population-based survey in 1991–1993 [13] and then by expansion of families from these index cases. Time-to-CI was established retrospectively through informant interview, or when that was not reliable, at the examination date where the individual was identified as impaired. Family members who were cognitively unimpaired (CU) were censored at the most recent examination date. Most of these individuals were examined only once. For the second study, we conducted a community-based study of cognitive and physical function in individuals over age 80 from 2003 to 2010. For people with CI, time-to-CI was established retrospectively through informant interview, or if that could not be reliably established, at the age at which the examination showed cognitive impairment. For CU, they were censored at the most recent examination date. Some CU individuals were examined a second time 2–3 years later. For the third study, from 2017 to date, we contacted all siblings of people with CI identified in the two studies above and examined those 76 and older. We also contacted individuals in the community in the same age group who were not siblings. Some of these individuals were participants in the first two studies, and thus this visit was a re-examination 3–5 years from the last visit. For people with CI, time-to-CI was established retrospectively through informant interview, or if that could not be reliably established, at the age at which the

examination showed cognitive impairment. For CU, they were censored at the most recent examination date.

All Amish participants were evaluated for cognitive function at least once using the Modified Mini-Mental State exam (3MS) [14]. DNA was extracted from a blood sample obtained at this baseline examination, which also included an assessment of family history of dementia or AD, a medical history of common chronic diseases, and basic sociodemographic information. The average number of clinical examinations was 1.4 (range: 1 to 5).

Individuals with education-adjusted 3MS scores below 87 at baseline or at a follow-up visit were classified as cognitively impaired (CI). Those scoring 87 and above at all visits were considered cognitively unimpaired (CU). AAO of CI was determined as the age at the examination when the participant was observed impaired for the first time. For individuals with CI at baseline, this may overestimate AAO of CI. For survival analysis, the outcome was defined as AAO of CI for those with cognitive impairment, and age at last examination for those classified as CU. The outcome of interest here was *CI-free time,* rather than the development of dementia, and therefore we chose a broad definition of CI (including AD and other types of dementia) as the event, rather than narrower definitions of AD. CI and CU classification was based on 3MS score as mentioned above. This allows the identification of loci protective against a wider range of dementia phenotypes, at the possible expense of greater heterogeneity of effects and lower power if the loci are specific to AD. We therefore attempted to replicate our findings in a second family dataset with individuals meeting this more specific AD case definition and their unaffected relatives.

For external replication we used the family-based National Institute on Aging Genetics Initiative for Late-Onset Alzheimer's Disease (NIA-LOAD) dataset [15,16] composed of 3,350 subjects (1,785 late onset AD (LOAD) cases and 1,565 controls neurologically evaluated as unaffected). Subjects were evaluated using standard research criteria for the diagnosis of LOAD as established in McKhann et al. [17], including the Clinical Dementia Rating score [18]. Cases were defined as those individuals with definite or probable LOAD diagnosis as per guidelines in McKhann et al. [17], while AAO was established as the age at which the family first observed memory problems or the age at first exam where cognitive impairment was noted. For controls the age used was the age at last examination when the absence of dementia was determined [17,19]. The demographics of the sample are provided in Table 2. More details about the sample are described in Lee et al. [19] and Kunkle et al. [5].

Written informed consent was obtained from study participants or, for those with substantial cognitive impairment, from a caregiver, legal guardian, or other proxy, and the study protocols for all populations were reviewed and approved by the appropriate Institutional review boards (IRB's).

2.2 Genotyping

DNA extracted from peripheral blood samples was used for genotyping on the Illumina Infinium Global Screening (GSA) or Expanded Multi-Ethnic Genotyping (MEGAex) Array.

After pre-imputation QC, which included removing markers missing more than 2% of the data, monomorphic markers, markers that had MAF < 0.05 and markers that failed Hardy Weinberg exact test at P value < 10^{-6}, the MEGAex final set before imputation was 552,801 markers and the GSA was 279,931. The two datasets were imputed separately using the Haplotype Reference Consortium (HRC) reference panel. Post-Imputation QC parameters were MAF >= 0.01, r^2 >=0.4; MAF < 0.01, r^2 >=0.8. Subsequently, the overlapping markers that passed QC from both sets were merged for a total of over 7 million common and rare imputed variants. Hardy-Weinberg Equilibrium (HWE) was tested for all SNPs and variants significantly out of HWE (P value < 10^{-6}) were removed before the analysis. Supplementary Table 1 summarizes the imputed SNPs retained for analysis by chromosome.

The NIA-LOAD dataset was genotyped using the Illumina 610 array containing 592,532 SNPs. For genotyping details see Wijsman et al. [15]. After genotype chip QC, the dataset was imputed to the 1,000 Genomes Project using IMPUTE2. The total number of imputed SNPs before QC was 39,127,740. For imputation details see Kunkle et al. [4]. We concentrated the replication analysis on chromosomes 6, 9 and 17 where we had found the three genome-wide significant results in the Amish cohort. Supplementary Table 2 shows the number of imputed SNPs analyzed on each chromosome.

2.3 Statistical Analysis

The Cox proportional hazards model is widely used in the analysis of time-to-event phenotypes (such as AAO) and mixed-effects models are commonly used in family-based genetic studies to adjust for relatedness of individuals. We used Coxmeg [20] an R package for conducting GWAS of age-at-onset traits in related individuals using Cox mixed-effects models (CMEMs). The CU survival time outcome was defined as AAO of CI (cases) or age at last normal cognitive examination (CU (censored) controls) in the Amish and NIA-LOAD (replication) cohorts. Familial relationships were controlled by including a random effect specified by the genetic relationship matrix (GRM) estimated from the genotype data in both Amish and NIA-LOAD datasets. In both datasets sex was included as a covariate. Individual SNPs, coded additively as imputed dosages of the less common allele, were added to this base model to test for association with CU survival time. The strength of this association is estimated by the hazard ratio (HR). The qqman package in R was used to summarize the significance of the association of each SNP with CU survival time in a Manhattan plot of p-values from the chi-square test of association. In the Amish dataset, genome-wide significant results were determined using the traditional GWAS threshold P < 5×10^{-8}, and suggestive results were highlighted using the threshold P < 1×10^{-5}. LocusZoom [21] was used to generate regional association plots and Linkage Disequilibrium analysis (LD), based on 1000 Genomes, of all common variants located in a 1Mb interval centered on SNPs with genome-wide significant associations. Three genome-wide significant results were identified, and markers in a 1-Megabase (Mb) region centered on each peak SNP were selected for analysis in the NIA-LOAD replication dataset, using identical statistical models to the Amish discovery dataset. To assess potential interaction of genome-wide significant effects with *APOE* genotype, stratified analysis was conducted in Amish and NIA-LOAD samples subdivided into APOE-4 carriers and non-carriers. Meta-analysis of hazard ratios from the stratified analysis was also performed using the R package meta [22].

3. RESULTS

3.1 Demographic statistics and genotype quality control for Amish dataset

The demographic characteristics of the Amish dataset are shown in Table 1. A total of 362 subjects were cognitively impaired (CI). Of these, 58% were women. The remaining 1,160 were cognitively unimpaired (CU). Of these, 58.4% were women. Age at last examination was 81.6 for the CI group and 70.6 for the CU group. The mean age at first exam below 3MS cutoff point for the CI group was 81.2 years.

The total number of imputed variants after QC, including rare variants, was 7,819,581 (Supplementary Table 1). Tests of HWE identified 3,630 variants significantly deviating from HWE ($p < 10^{-6}$). After removing these variants, 7,815,951 were retained for GWAS analysis.

3.2 GWAS of CU survival time

The GWAS of CU survival time used a Cox mixed-effects model with each imputed variant and sex as a fixed-effects covariate, with the kinship matrix included as a random effect. The results for all autosomes are summarized in the Manhattan plot in Figure 1. Three common variants (minor allele frequency (MAF) > 0.01) were significantly associated ($p < 5 \times 10^{-8}$) with CI-free survival time on chromosomes 6 (rs145348074), 9 (rs534551495), and 17 (rs146729640). Table 3 lists for each associated region the p value, hazard ratio, minimum allele frequency (MAF), and overlapping or closest gene and genes located within +/− 500 thousand base pairs. The QQ plot of GWAS p-values of CU survival time, as well as the genomic inflation factor (lambda) statistic (λ=1.01) suggest that the significance of these GWAS results is not inflated (Figure 2). To determine if there were broad regions of association surrounding these peak markers, we examined regional association plots (Supplemental Figures 1-3), identified variants with suggestive evidence of association ($p < 10^{-5}$, Supplemental Table 3), and examined linkage disequilibrium patterns in the region to assess whether there was evidence for multiple association signals.

3.3 Regional plot and local linkage disequilibrium (LD) analysis for each genome wide variant

As summarized in Table 4, regional association analysis of 1-Megabase (Mb) regions on chromosomes 6 and 9 identified three additional variants in moderate to strong LD (r^2>0.4) with the peak markers and suggestive evidence of association with AAO ($p < 1 \times 10^{-5}$). No additional markers were identified in the chromosome 17 region. Regional association plots are provided in Supplemental Figures 1, 2 and 3.

3.4 Replication analysis of the three genome-wide significant results in the independent NIA-LOAD dataset.

The demographic characteristics of the NIA-LOAD replication dataset are shown in Table 2. There are 1,785 people with AD (65% women) and 1,565 CU controls (60% women). AAO was 73.1 for cases and the age at exam for the control group was 73.9. The replication analyses were conducted in the 1Mb interval centered on each peak variant from the Amish

GWAS. For chromosome 6, variants numbered 14,532; for chromosome 9, variants were 13,602; for chromosome 17 variants were 17,000 (Supplementary Table 2).

Analysis of CU survival time in the NIA-LOAD dataset detected nominal evidence of association in the chromosome 17 interval, at the same peak marker as the Amish GWAS (rs146729640; p=0.02, Table 5). Results for the other two regions were not significant. While the result on chromosome 17 did not exceed the Bonferroni-adjusted significance threshold (p=0.05/17,000 variants tested = 3×10^{-6}), the replication of the result at the *exact same variant* with nominal significance suggests that the observed association is reproducible.

APOE-4 dose was strongly associated with CU survival time in the Amish (HR=1.61, 95% CI (1.28 – 2.02); p=5.08×10–5). Because of this strong main effect, we tested for interaction between *APOE-4* dose and the top SNP in the GWAS. Significant interactions (p<0.05) were evaluated by stratifying the dataset by *APOE-4* carrier status to determine if these SNPs had different effects by *APOE-4* genotype.

3.5 Stratified Analysis of Chromosome 17 regions by APOE-4 carrier status

To explore the possibility that *APOE* genotype might influences these results, the discovery and replication datasets were stratified by *APOE-4* carrier status and the Cox mixed-effects models were run in each subset. In both the Amish (Table 6) and NIA-LOAD (Table 7) datasets, the effect (hazard ratio) at rs146729640 was stronger in the subset of individuals that did not carry the *APOE-4* allele. This was also confirmed in the meta-analysis combining the results of the non *APOE-4* carriers' groups in the Amish and NIA-LOAD (Table 8; HR= 3.56, 95 %CI= 2.19–5.78).

4. DISCUSSION

This study identified a reproducible association of CU survival time with an intronic polymorphism of the *SHISA6* gene on chromosome 17, where the more frequent allele "G" of rs146729640 is associated with later AAO of CI and the minor allele "A" was strongly associated with earlier AAO of CI (p=5.6×10^{-9}) acting as a risk factor with a strong hazard ratio (HR= 6.38, 95 %CI= 3.42–11.90). The "A" allele frequency in our cohort was 0.012, twice the frequency of 0.006 reported by The Allele Frequency Aggregator (ALFA) database (https://www.ncbi.nlm.nih.gov/snp/docs/gsr/alfa/) for European population.

Large-scale genome-wide association studies (GWAS) of disease-free survival offer a powerful method of identifying factors that delay the onset of disease, manifesting in later AAO. Such discoveries might illuminate mechanisms underlying disease progression of age-related diseases such as CI and AD. In the present study we performed a GWAS of 7.8 M imputed variants, examining association with CU survival time in 1,520 Amish individuals longitudinally evaluated for CI. Three genome-wide significant results were examined for replication in the independent NIA-LOAD family sample of 1785 individuals with AD and 1565 CU controls, and the association at rs146729640 (chromosome 17) was replicated at a nominal significance level in the NIA-LOAD dataset. We note that the outcome of interest in both datasets, time free of cognitive impairment, was the same, despite differences in the

definition used to identify CI. The association at rs146729640 in both datasets despite these differences is encouraging. This intronic polymorphism is located within the *SHISA6* gene, an excellent candidate gene for AD.

SHISA6 has previously been implicated in AD pathogenesis in human proteomics and transgenic mouse model studies. The *SHISA6* is among the 60 proteins, out of 4,582 analyzed, found significantly altered by increasing age in the human hippocampus. *SHISA6* was found downregulated in this study, which was carried out using hippocampal samples from 16 Chinese brain tissue donors free of neurological diseases across four age groups using the youngest one (22–49) as reference: 22–49, 50–69, 70–89, and >90 [23]. Additionally, several experiments *in vivo* using *APOE* transgenic mice have revealed that the *Shisa6* gene is differentially expressed (downregulated $\log_2$ fold change = - 0.53) in the entorhinal cortex (EC) of aged *APOE* mice (*APOE3/4 vs. APOE3/3*) [24]. Based on these findings, we performed a stratified analysis and found that the strength of association at rs145348074 (*SHISA6* gene) was increased in the non-*APOE4* carriers in both the Amish and the NIA-LOAD replication data sets. Notably, the EC is one of the first regions to be affected by AD pathology in humans. *Shisa6* is also involved in maintenance of high-frequency synaptic transmission between neurons in the mouse hippocampus by regulating AMPA-type glutamate receptors (AMPAR) [25].

We note that the family-based design and use of cognitive-impairment free survival time as the outcome introduces a few potential limitations to the study and interpretation of results. One potential limitation in our study is that for people in the Amish cohort who were CI at baseline (prevalent cases), the AAO of CI we used (age at exam) was probably later than the true AAO, as symptoms may have started earlier and been unrecognized or unreported prior to enrollment in our study. This could have biased our results towards the null; therefore, the true association may be stronger than the current estimate. A second potential limitation is that CI is a multifactorial condition driven by a mixture of genetic, vascular, metabolic, and lifestyle-related factors [26]. Heterogeneity in types of CI (MCI, AD, other dementia) could lead to spurious associations (in direction or strength of association) due to Simpson's paradox. However, based on the estimation that the most well-known cause of MCI (Mild Cognitive Impairment)- around 50 % in people age 65 and older- is AD [27], and since the results were replicated in the NIA-LOAD data set, where cases had been clinically diagnosed with LOAD [5, 19], we are reassured that the detected association at SNP rs146729640 is not likely to be due to such bias.

The identification of variants that modify the AOO in AD will reveal potential pathways for intervention in the progression of the symptoms of the disease. Only a few studies have been conducted to identify genetic variants associated with AAO in AD. Lalli et al. [28] identified the variant rs9909184 located near the CCL11 gene on chromosome 17 associated with delayed onset of mild-cognitive impairment and dementia using a cohort of 72 subjects affected with early-onset familial AD. Our study identified the variant rs146729640, that may be associated with AAO of LOAD, located on the same chromosome although in a different region. The discovery of AAO modifying factors is essential to develop targets for early interventions aimed to prevent the onset of the disease.

GWAS studies are susceptible to false positives results due to the large number of statistical tests. Therefore, an internally replicated design that focuses on signals that are independently detected in a second dataset provides stronger evidence of association. While our GWAS was conducted in a family-based study of 1,520 Amish, we tested our three genome-wide significant findings in a second family dataset (NIA-LOAD) with twice the sample size. The replicated result in the *SHISA6* gene is bolstered by its biological plausibility; this gene has been implicated in transgenic mouse models of AD and human proteomic studies of the aging hippocampus. The effect in the NIA-LOAD is somewhat smaller, but still statistically significant. These results illustrate the utility of using isolated populations to facilitate replicable gene discoveries in complex traits. These specific results suggest that *SHISA6* might modulate the rate of cognitive decline, and studies to elucidate the specific effect and mechanism underlying this reproducible association are warranted.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We thank the Amish families for allowing us into their communities and for participating in our study. We also acknowledge the contributions of the Anabaptist Genealogy Database and Swiss Anabaptist Genealogy Association. This study is supported by National Institutes of Health / National Institute of Aging, grant 1RF1AG058066 (to Jonathan L. Haines, Margaret Pericak-Vance, and William K. Scott) and grant U24 AG056270 (to Tatiana M. Foroud, Alison M. Goate and Richard P. Mayeux). Alan J. Lerner and Jeffrey M. Vance also received support from NIH for this manuscript. Finally, we acknowledge the resources provided by the Department of Population and Quantitative Health Sciences, School of Medicine at Case Western Reserve University and the John P. Hussman Institute for Human Genomics at University of Miami, Miller School of Medicine.

Disclosure Statement

The following co-authors report these activities in the last 36 months:

Tatiana M. Foroud have received grants from NIH, MJFF and NCAA-DOD, Richard P. Mayeux from NIA, Peter St. George-Hyslop from the Alzheimer's Association, Alan J. Lerner from NINDS and NIA, Jeffery M. Vance from NIH, Michael L. Cuccaro from NIA, Jonathan L. Haines from NIH, Margaret A. Pericak-Vance from NIH and William K. Scott from NIH, NIA and the US Department of Veterans Affairs. All payments were made to their institutions.

Peter St. George-Hyslop received royalties, paid to the University of Toronto, on patents relating to the use of the presenilin gene sequences for the discovery of mutations causing Alzheimer disease in at risk and symptomatic subjects. Alan J Lerner received royalties from Elsevier for editing a book on neurology.

Tatiana M. Foroud received consulting fees for serving on advisory committees from the University of Pennsylvania, Mayo Clinic and University of Washington. William K. Scott received seminar speaker honorarium from the University of Chicago.

REFERENCES

1. Hale JM, Schneider DC, Mehta NK, & Myrskyla M (2020). Cognitive impairment in the U.S.: Lifetime risk, age at onset, and years impaired. SSM - Population Health, 11, 100577. [PubMed: 32300635]
2. Brodaty H, & Donkin M (2009). Family caregivers of people with dementia. Dialogues in Clinical Neuroscience, 11(2), 217–228. [PubMed: 19585957]
3. Hurd MD, Martorell P, Delavande A, Mullen KJ, & Langa KM (2013). Monetary costs of dementia in the United States. The New England Journal of Medicine, 368(14), 1326–1334. [PubMed: 23550670]

4. Floris M, Olla S, Schlessinger D, & Cucca F (2018). Genetic-driven druggable target identification and validation. Trends in Genetics : TIG, 34(7), 558–570. [PubMed: 29803319]
5. Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, et al. (2019). Author correction: Genetic meta-analysis of diagnosed alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. Nature Genetics, 51(9), 1423–1424. [PubMed: 31417202]
6. Hahs DW, McCauley JL, Crunk AE, McFarland LL, Gaskell PC, Jiang L, et al. (2006). A genome-wide linkage analysis of dementia in the amish. American Journal of Medical Genetics.Part B, Neuropsychiatric Genetics: The Official Publication of the International Society of Psychiatric Genetics, 141B(2), 160–166. [PubMed: 16389594]
7. Pericak-Vance MA, Johnson CC, Rimmler JB, Saunders AM, Robinson LC, D'Hondt EG, et al. (1996). Alzheimer's disease and apolipoprotein E-4 allele in an amish population. Annals of Neurology, 39(6), 700–704. [PubMed: 8651641]
8. Edwards DR, Gilbert JR, Jiang L, Gallins PJ, Caywood L, Creason M, et al. (2011). Successful aging shows linkage to chromosomes 6, 7, and 14 in the amish. Annals of Human Genetics, 75(4), 516–528. [PubMed: 21668908]
9. Cummings AC, Jiang L, Velez Edwards DR, McCauley JL, Laux R, McFarland LL, et al. (2012). Genome-wide association and linkage study in the amish detects a novel candidate late-onset alzheimer disease gene. Annals of Human Genetics, 76(5), 342–351. [PubMed: 22881374]
10. Edwards DR, Gilbert JR, Hicks JE, Myers JL, Jiang L, Cummings AC, et al. (2013). Linkage and association of successful aging to the 6q25 region in large amish kindreds. Age (Dordrecht, Netherlands), 35(4), 1467–1477. [PubMed: 22773346]
11. D'Aoust LN, Cummings AC, Laux R, Fuzzell D, Caywood L, Reinhart-Mercer L, et al. (2015). Examination of candidate exonic variants for association to alzheimer disease in the amish. PloS One, 10(2), e0118043. [PubMed: 25668194]
12. Ramos J, Chowdhury AR, Caywood LJ, Prough M, Denise Fuzzell M, Fuzzell S, et al. (2021). Lower levels of education are associated with cognitive impairment in the old order amish. Journal of Alzheimer's Disease: JAD, 79(1), 451–458.
13. Johnson CC, Rybicki BA, Brown G, D'Hondt E, Herpolsheimer B, Roth D, et al. (1997). Cognitive impairment in the Amish: A four county survey. International Journal of Epidemiology, 26(2), 387–394. [PubMed: 9169175]
14. Teng EL, & Chui HC (1987). The modified mini-mental state (3MS) examination. The Journal of Clinical Psychiatry, 48(8), 314–318. [PubMed: 3611032]
15. Wijsman EM, Pankratz ND, Choi Y, Rothstein JH, Faber KM, Cheng R, et al. (2011). Genome-wide association of familial late-onset alzheimer's disease replicates BIN1 and CLU and nominates CUGBP2 in interaction with APOE. PLoS Genetics, 7(2), e1001308. [PubMed: 21379329]
16. Tosto G, Bird TD, Tsuang D, Bennett DA, Boeve BF, Cruchaga C, et al. (2017). Polygenic risk scores in familial alzheimer disease. Neurology, 88(12), 1180–1186. [PubMed: 28213371]
17. McKhann G, Drachman D, Folstein M, Katzman R, Price D, & Stadlan EM (1984). 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, 34(7), 939–944. [PubMed: 6610841]
18. Hughes CP, Berg L, Danziger WL, Coben LA, & Martin RL (1982). A new clinical scale for the staging of dementia. The British Journal of Psychiatry: The Journal of Mental Science, 140, 566–572. [PubMed: 7104545]
19. Lee JH, Cheng R, Graff-Radford N, Foroud T, Mayeux R, & National Institute on Aging Late-Onset Alzheimer's Disease Family Study Group. (2008). Analyses of the national institute on aging late-onset alzheimer's disease family study: Implication of additional loci. Archives of Neurology, 65(11), 1518–1526. [PubMed: 19001172]
20. He L, & Kulminski AM (2020). Fast algorithms for conducting large-scale GWAS of age-at-onset traits using cox mixed-effects models. Genetics, 215(1), 41–58. [PubMed: 32132097]
21. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. (2010). LocusZoom: Regional visualization of genome-wide association scan results. Bioinformatics (Oxford, England), 26(18), 2336–2337. [PubMed: 20634204]

22. Balduzzi S, Rucker G, & Schwarzer G (2019). How to perform a meta-analysis with R: A practical tutorial. Evidence-Based Mental Health, 22(4), 153–160. [PubMed: 31563865]

23. Xu B, Gao Y, Zhan S, Xiong F, Qiu W, Qian X, et al. (2016). Quantitative protein profiling of hippocampus during human aging. Neurobiology of Aging, 39, 46–56. [PubMed: 26923401]

24. Area-Gomez E, Larrea D, Pera M, Agrawal RR, Guilfoyle DN, Pirhaji L, et al. (2020). APOE4 is associated with differential regional vulnerability to bioenergetic deficits in aged APOE mice. Scientific Reports, 10(1), 4277–020-61142–8.

25. Klaassen RV, Stroeder J, Coussen F, Hafner AS, Petersen JD, Renancio C, et al. (2016). Shisa6 traps AMPA receptors at postsynaptic sites and prevents their desensitization during synaptic activity. Nature Communications, 7, 10682.

26. Rosenberg A, Mangialasche F, Ngandu T, Solomon A, & Kivipelto M (2020). Multidomain interventions to prevent cognitive impairment, alzheimer's disease, and dementia: From FINGER to world-wide FINGERS. The Journal of Prevention of Alzheimer's Disease, 7(1), 29–36.

27. -2021 alzheimer's disease facts and figures.(2021). Alzheimer's & Dementia : The Journal of the Alzheimer's Association, 17(3), 327–406.

28. Lalli MA, Bettcher BM, Arcila ML, Garcia G, Guzman C, Madrigal L, et al. (2015). Whole-genome sequencing suggests a chemokine gene cluster that modifies age at onset in familial alzheimer's disease. Molecular Psychiatry, 20(11), 1294–1300. [PubMed: 26324103]

HIGHLIGHTS

- Replicated genome-wide association of CI-free survival time in the *SHISA6* gene.
- *SHISA6* influences post-synaptic transmission.
- *SHISA6* expression declines with age in hippocampus.
- *SHISA6* is downregulated in entorhinal cortex of aged APOE4 transgenic mice.
- *SHISA6* is thus a replicated, biologically plausible candidate gene for AAO of CI.

RESEARCH IN CONTEXT

1. **Systematic Review**: We reviewed the literature, using traditional (e.g., PubMed) sources for longitudinal studies of aging and cognitive impairment. Several reports of sibships with CI and unaffected (CU) individuals paralleled our observations in the Amish. We reviewed publications on the *SHISA6* gene and found influences on AD pathogenesis in humans and animal models.

2. **Interpretation**: In 1,522 individuals screened for CI and genotyped, we found three single nucleotide variants (SNV) associated with age at onset (AAO) of CI. One association on chromosome 17 (rs146729640) in the *SHISA6* gene was replicated in the NIA-LOAD family study dataset (N=3,350). *SHISA6,* involved in post-synaptic transmission in the hippocampus, is a biologically plausible candidate gene for AD that might play a role in the rate of cognitive decline.

3. **Future Directions:** Future research should examine whether genetic variation influences *SHISA6* gene expression in brain tissue from AD and control donors.

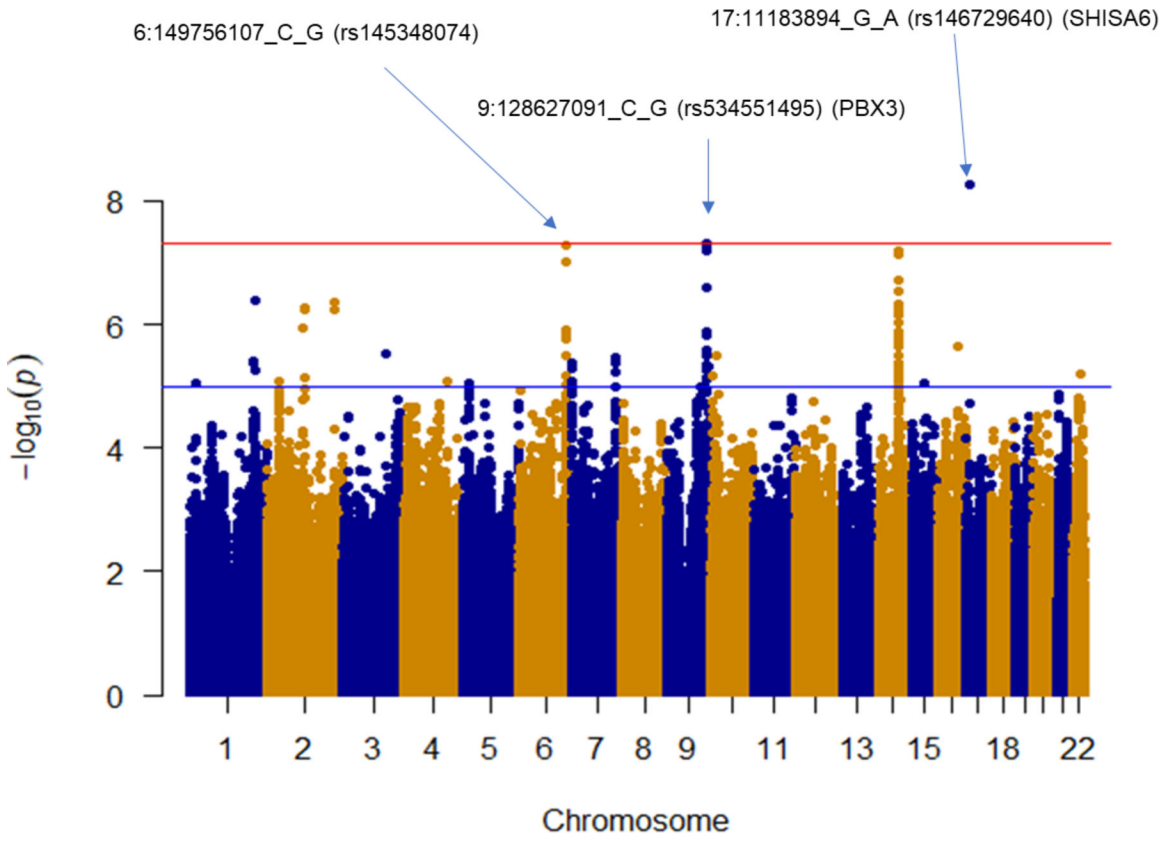

Figure 1. Manhattan plot showing results of GWAS of CU survival time using Cox mixed-effects model:

Red line indicates the threshold for genome wide significance ($P < 5 \times 10^{-8}$) while the blue line represents the suggestive threshold ($P < 1 \times 10^{-5}$). Three variants reached genome wide significance: 6:149756107_C_G (rs145348074), 9:128627091_C_G (rs534551495), and 17:11183894_G_A (rs146729640). Mapping is based on GRCh37/hg19 assembly.

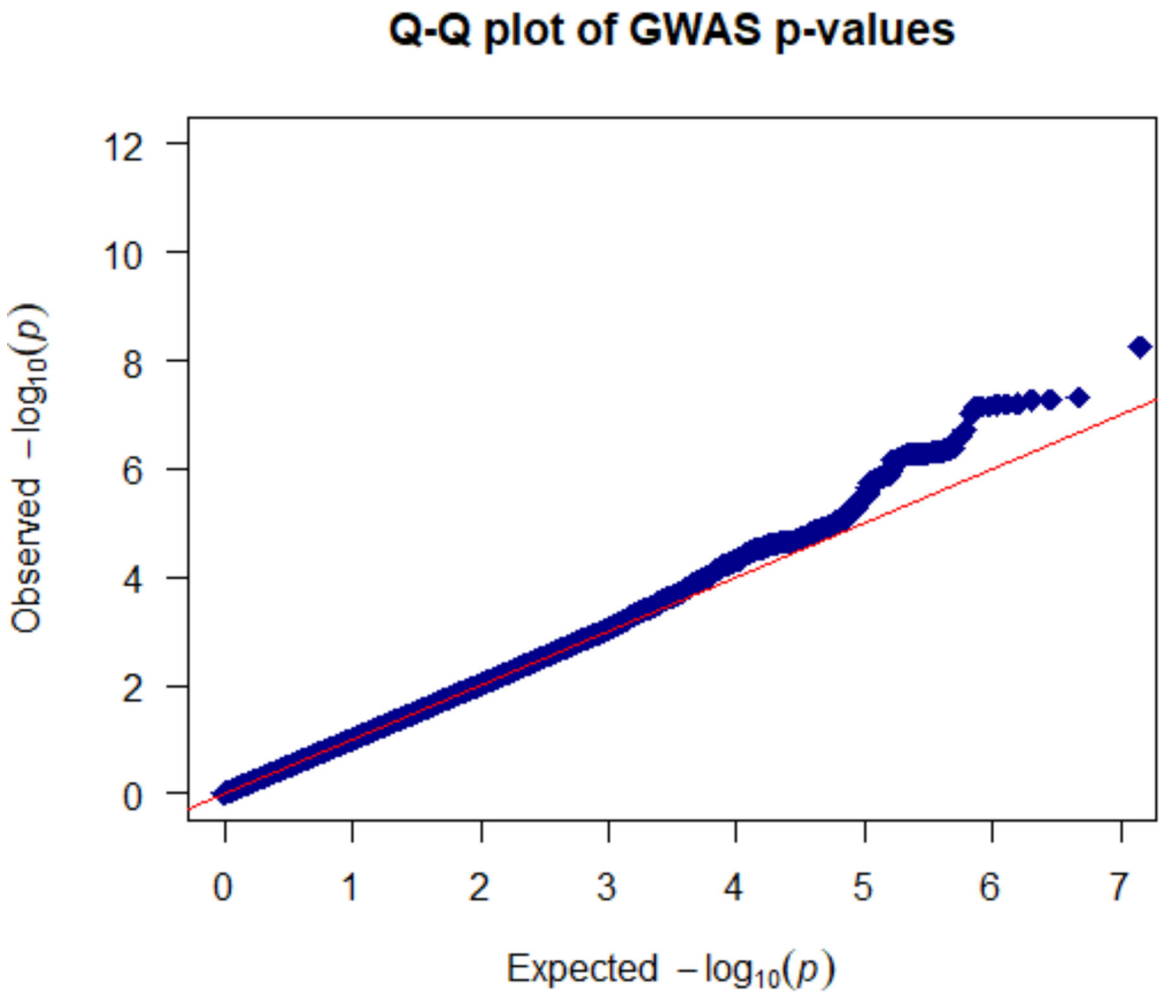

Figure 2. Quantile – quantile (QQ) plot showing the results of GWAS p-values of CU survival time in the Amish cohort using Cox mixed-effects model:

This figure clearly shows that GWAS is not inflated. The genome-wide inflation factor (λ) was 1.01

Table 1.

Demographic characteristics of the 1,522 Amish individuals included in the study. About 24% of the sample was cognitively impaired at the last examination. Cognitively impaired individuals were older at the last examination, and the two groups had similar proportions of men and women.

Column1	COGNITIVELY IMPAIRED	COGNITIVELY UNIMPAIRED	TOTAL
N	362	1160	1522
MALE	152 (42.0 %)	483 (41.6 %)	635 (41.7 %)
FEMALE	210 (58.0 %)	677 (58.4 %)	887 (58.3 %)
AGE at the last exam	81.6	70.6	73.23
AAO	81.2	NA	81.2
APOE4 carriers	105	290	395
non-APOE4 carriers	257	870	1127
AGE (AAO for CI and Age at exam for CU) APOE4 carriers	80.1	70.1	72.7
AGE (AAO for CI and Age at exam for CU) non-APOE4 carriers	81.6	70.7	73.2

Table 2.

Demographic characteristics of NIA-LOAD data set used for replication. About 53% of the sample was clinically diagnosed with LOAD. Age at onset for AD cases was similar to age at exam for controls. The group of AD has a higher proportion of women (65 vs 60%).

	AD	CONTROLS	TOTAL
N	1,785	1,565	3,350
MALE	623 (34.9 %)	623 (39.8 %)	1,246 (37.2 %)
FEMALE	1,162 (65.1 %)	942 (60.2 %)	2,104 (62.8 %)
AGE (AAO for AD and Age at exam for controls)	73.1	73.9	73.5
APOE4 carriers	1,326	572	1,898
non-APOE4 carriers	459	993	1,452
AGE (AAO for AD and Age at exam for controls) APOE4 carriers	72.5	71.5	72.2
AGE (AAO for AD and Age at exam for controls) non-APOE4 carriers	74.8	75.2	75.1

Table 3.

Three common variants were significantly associated ($p < 5 \times 10^{-8}$) with CU survival time in the Amish dataset. The strongest result was on chromosome 17, followed by chromosomes 6 and 9. Variants are annotated as overlapping a gene if located within the gene boundaries. Other genes within a 1Mb interval centered on the associated variant are also listed.

Location and base pair change (build hg19)	SNP (rs)	MAF	Variant type	Hazard Ratio	P-value	Overlapping Genes	Genes within +/− 500Kb
6:149756107_C_G	rs145348074	0.0237382	Intergenic	3.35	5.26E-08	Intergenic variant. Closest gene: *ZC3H12D*	*UST, UST-AS1, LOC105378047, TAB2, TAB2-AS1, SUMO4, ZC3H12D, PPIL4, GINM1, KATNA1, LATS1, LOC645967, NUP43, PCMT1, LRP11, RAET1E-AS1, RAET1E, RAET1G, LOC105378052*
9:128627091_C_G	rs534551495	0.0379333	Intron	2.82	4.83E-08	*PBX3*	*GAPVD1, MAPKAP1, LOC51145, PBX3, LOC10192911, MVB12B*
17:11183894_G_A	rs146729640	0.0119911	Intron	6.38	5.59E-09	*SHISA6*	*TMEM220-AS1, DNAH9, TMEM238L, PIRT, SHISA6,*

Table 4.

Results of local linkage disequilibrium (LD) analysis of genome wide significant variants. Variants rs141138977 and rs140984313 were found in LD ($1.0 > r^2 \geq 0.8$) with rs534551495, the genome wide signal found on chromosome 9. Variant rs143831028 was in LD ($0.6 > r^2 \geq 0.4$) with rs145348074, the genome wide signal found on chromosome 6. None of the variants was found in LD ($r^2 \geq 0.4$) with rs146729640, the genome wide signal found on chromosome 17.

Genome wide Variant	Variant (rs)	MAF	P value	Genes	Variants in LD (+/−500Kb)					
					Variant	Variant (rs)	r^2	MAF	P value	Genes
6:149756107_C_G	**rs145348074**	0.024	5.26E-08	Intergenic variant. Closest gene: ***ZC3H12D***	**6:149622242_A_T**	**rs143831028**	**0.6 > r2 ≥ 0.4**	0.023	1.00E-7	***TAB2***
					6:149688198_G_A	rs192019712	0.4 > r2 ≥ 0.2	0.015	1.29E-6	*TAB2*
					6:149898888_T_A	rs188874345	0.4 > r2 ≥ 0.2	0.016	1.35E-6	*GINM1*
					6:149907017_C_T	rs530694742	0.4 > r2 ≥ 0.2	0.016	1.44E-6	*GINM1*
					6:149959451_T_C	rs142372487	0.4 > r2 ≥ 0.2	0.016	1.27E-6	*KATNA1*
					6:150057902_G_A	rs541233509	0.4 > r2 ≥ 0.2	0.016	1.58E-6	*NUP43*
9:128627091_C_G	**rs534551495**	0.038	4.83E-08	***PBX3***	**9:128627225_G_T**	**rs141138977**	**1.0 > r2 ≥ 0.8**	0.038	5.45E-08	***PBX3***
					9:128319404_G_A	**rs140984313**	**1.0 > r2 ≥ 0.8**	0.038	6.46E-08	***MAPKAP1***
					9:128481411_A_T	rs113929904	0.4 > r2 ≥ 0.2	0.046	2.50E-7	*Closest gene: MAPKAP1*
					9:128175053_G_A	rs149607027	0.4 > r2 ≥ 0.2	0.042	1.33E-6	*Closest gene: MAPKAP1*

Genome wide Variant	Variant (rs)	MAF	P value	Genes	Variants in LD (+/−500Kb)					
					Variant	Variant (rs)	r^2	MAF	P value	Genes
17:11183894_G_A	**rs146729640**	0.012	5.59E-09	***SHISA6***	17:11306227_G_A	rs111837116	0.4 > r2 ≥ 0.2	0.017	4.72E-04	SHISA6

Table 5.

Results of the Cox mixed effect model for replication using NIA-LOAD data set show nominal association (p = 0.02) in the chromosome 17 variant (rs146729640).

Chr	Variant	Variant (rs)	Data set	# of subjects	MAF	Hazard ratio	P-value
17	17:11183894_G_A	rs146729640	Amish	1520	0.012	6.38	5.59E-09
			NIA-LOAD	3350	0.012	1.49	0.0217
9	9:128627091_C_G	rs534551495	Amish	1520	0.038	2.82	4.83E-08
			NIA-LOAD	3350	0.006	0.98	0.9356
6	6:149756107_C_G	rs145348074	Amish	1520	0.024	3.35	5.26E-08
			NIA-LOAD	3350	0.013	1.16	0.4164

Table 6.

Stratified survival analysis of rs146729640 variant in the Amish dataset. The strength of association was increased in the non- *APOE4* carriers.

***APOE4* stratification Amish**	**SNP (chr:pos)**	**SNP (rs)**	**N (CI)**	**N (CU)**	**Hazard Ratio**	**Hazard Ratio 95% CI**	**P-value**	**Gene**
No	17:11183894_G_A	rs146729640	362	1,160	6.38	3.42 ; 11.90	5.59E-09	*SHISA6*
APOE4 carriers	17:11183894_G_A	rs146729640	105	290	2.93	1.16 ; 7.40	2.32E-02	SHISA6
non-*APOE4* carriers	17:11183894_G_A	rs146729640	257	870	8.96	4.25 ; 18.91	8.44E-09	SHISA6

Table 7.

Stratified survival analysis of rs146729640 variant in the NIA-LOAD dataset. Similar to the Amish discovery set, the strength of association was increased in the non-*APOE4* carriers.

***APOE4* stratification NIA-LOAD**	**SNP (chr:pos)**	**SNP (rs)**	**N (AD)**	**N (controls)**	**Hazard Ratio**	**Hazard Ratio 95% CI**	**P-value**	**Gene**
No	17:11183894_G_A	rs146729640	1,785	1,565	1.49	1.06 ; 2.08	0.0217	*SHISA6*
APOE4 carriers	17:11183894_G_A	rs146729640	1,326	572	1.35	0.95 ; 1.93	0.0950	SHISA6
non-*APOE4* carriers	17:11183894_G_A	rs146729640	459	993	1.82	0.96 ; 3.43	0.0658	SHISA6

Table 8.

Meta-analysis of rs146729640 variant in the Amish and NIA-LOAD datasets. Similar to the individual results of the discovery and replication studies, the strength of association was increased in the non-*APOE4* carriers

***APOE4* stratification**	**SNP (rs)**	**Studies combined**	**Number of observations**	**Meta-Analysis results**			
				Model	**Hazard Ratio**	**Hazard Ratio 95 % CI**	**P-value**
NO	rs146729640	Amish and NIA-LOAD	4,872	Fixed effect	2.07	1.54; 2.78	<0.0001
				Random effects	3.00	0.72; 12.50	*0.1308*
APOE4 carriers	rs146729640	Amish and NIA-LOAD	2,293	Fixed effect	1.50	1.072; 2.09	*0.0178*
				Random effects	1.76	0.86; 3.60	0.1219
non-*APOE4* carriers	rs146729640	Amish and NIA-LOAD	2,579	Fixed effect	3.56	2.19; 5.78	<0.0001
				Random effects	3.99	0.83; 19.04	0.0831