HHS Public Access

Author manuscript

Alzheimers Dement. Author manuscript; available in PMC 2020 December 01.

Published in final edited form as:

Alzheimers Dement. 2019 December ; 15(12): 1524–1532. doi:10.1016/j.jalz.2019.07.016.

Local ancestry at *APOE* modifies Alzheimer's disease risk in Caribbean Hispanics

Elizabeth E. Blue[1], **Andréa R.V.R Horimoto**[2,3], **Shubhabrata Mukherjee**[4], **Ellen M. Wijsman**[1,2,5], **Timothy A. Thornton**[2,6]

[1]Division of Medical Genetics, University of Washington, Seattle, WA, USA

[2]Department of Biostatistics, University of Washington, Seattle, WA, USA

[3]Laboratory of Genetics and Molecular Cardiology, Heart Institute, University of Sao Paulo Medical School, Sao Paulo, Brazil

[4]Department of Medicine, University of Washington, Seattle, WA, USA

[5]Department of Genome Sciences, University of Washington, Seattle, WA, USA

[6]Department of Statistics, University of Washington, Seattle, WA

Abstract

INTRODUCTION—Although the relationship between *APOE* and Alzheimer's disease (AD) is well established in populations of European descent, the effects of *APOE* and ancestry on AD risk in diverse populations is not well understood.

METHODS—Logistic mixed model regression and survival analyses were performed in a sample of 3067 Caribbean Hispanics and 3028 individuals of European descent to assess the effects of *APOE* genotype, local ancestry, and genome-wide ancestry on AD risk and age-of-onset.

RESULTS—Among the Caribbean Hispanics, individuals with African-derived ancestry at *APOE* had 39% lower odds of AD compared to individuals with European-derived *APOE*, after adjusting for *APOE* genotype, age, and genome-wide ancestry. While *APOE* E2 and E4 effects on AD risk and age-of-onset were significant in the Caribbean Hispanics, they were substantially attenuated compared to European ancestry individuals.

DISCUSSION—These results suggest that additional genetic variation in the *APOE* region influences AD risk beyond *APOE* E2/E3/E4.

Correspondence: Elizabeth E. Blue, Division of Medical Genetics, University of Washington, BOX 357220, Seattle WA 98195-7720. Phone: 206-685-4666, Fax: 206-543-3050, em27@uw.edu, Timothy A. Thornton, Department of Biostatistics, University of Washington, Seattle, WA, USA. Phone: 206-543-8004, Fax: 206-543-3286, tathornt@uw.edu.

Alzheimer's Disease Neuroimaging Initiative (ADNI): Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.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.usc.edu/wpcontent/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Keywords

Alzheimer's disease; age-at-onset; *APOE*; admixture; ancestry

1. Background

Alzheimer's disease (AD; MIM 104300) is a progressive dementing disease, affecting >10% of Americans 65 years of age or older [1]. Risk of AD is influenced by both genetic and environmental factors [2], although heritability estimates suggest as much as 75% of the phenotypic variance can be attributed to genetics [3]. Rare variants in *APP* [4], *PSEN1* [5], and *PSEN2* [6, 7] cause autosomal dominant early-onset AD, and dozens of loci have been associated with risk or age-at-onset (AAO) of AD [2]. *APOE* has long stood at the forefront of common genetic risk factors for AD [8], with the E4 isoform providing the most consistent genetic association with the strongest effect on AD risk [2].

More research into the genetics of AD is necessary to better understand individual-level risks and the pathogenesis of AD across diverse populations [9]. Similar to most genome-wide association studies (GWAS) [10, 11], the vast majority of genetic studies of AD have primarily focused on samples with European ancestry [12, 13]. However, the risk of AD differs across populations. African Americans and Hispanics/Latinos, for example, are more likely to develop AD or other dementias than non-Hispanic whites [14]. Both African Americans and Hispanics/Latinos have admixed African, European, and Native American ancestry, and previous studies suggest Native American ancestry is protective against AD [15, 16] while other studies identified associations with proportional African ancestry and increased risk [9, 16, 17]. Differences in AD risk by ancestry within and between these populations may be driven by clinical and genetic differences, but there is also evidence that environmental factors, such as educational attainment and socioeconomic status, may explain some of the observed differences in risk of AD [2].

While *APOE* is a well-established risk locus for AD in populations with predominantly European ancestry, the relationship between *APOE* on AD risk in populations with diverse ancestries is not well understood. There is conflicting evidence regarding associations between *APOE* and AD in studies of both Hispanic/Latino [12, 16–19] and African American [12, 18–20] populations. Significant associations between *APOE* and AD risk have been observed in genetic studies of African Americans and Hispanics, typically in cohorts with large sample sizes, but with effect sizes that are smaller relative to studies of European ancestry populations [12, 18, 20]. Disentangling the relationship between *APOE*, AD, and ancestry continues to be challenging, in part because *APOE* allele frequencies vary considerably world-wide. For example, the frequency of the two missense variants rs429358 and rs7412 defining the *APOE* E2/E3/E4 isoforms are clinal in Europe [21], and are more common in African populations and less common among Latinos [22].

Here, we investigate the relationship between *APOE*, ancestry, and AD in Caribbean Hispanics. In a large cohort of 3067 Caribbean Hispanics, we first infer both local ancestry at *APOE* and genome-wide ancestry, and then estimate the effects of *APOE* alleles and ancestry on both AD risk and age-of-onset. To provide perspective, we compare the results

from the large Caribbean Hispanic cohort to estimated effects of *APOE* alleles on AD risk and AAO in a similarly sized cohort of European ancestry (N = 3028), a smaller sample of Caribbean Hispanics (N = 408), and a small longitudinal cohort-based European ancestry study (N=909). Although we find that *APOE* E2 alleles were protective and E4 alleles increased risk of AD in both the European and Caribbean Hispanic cohorts, the magnitude of these effects was consistently weaker in the Hispanic data sets and were robust to adjustment for European ancestry genome-wide. Using local ancestry estimates at the *APOE* region in Caribbean Hispanics, we assessed risk of AD as well as estimated and compared effects of *APOE* E3 and E4 alleles on phased African or European haplotypic backgrounds. In addition to detecting significant associations with genome-wide ancestry, logistic regression models also revealed that local ancestry at the *APOE* region is a significant predictor of AD risk in the Caribbean Hispanics. These results suggest that the weaker estimated effects of *APOE* on AD risk previously observed in studies of admixed populations are not driven exclusively by small sample sizes, ascertainment, or global ancestry proportions, but are also influenced by ancestry-specific genetic variation at or near the *APOE* locus.

2. Methods

2.1. Data sets

The data for this study represents a subset of the National Institute on Aging's Late-Onset Alzheimer's Disease and the National Cell Repository for Alzheimer's Disease Family study (NIALOAD; Database of Genotypes and Phenotypes, dbGaP, study phs000168.v2.p2) [23], the Columbia University Study of Caribbean Hispanics with Familial and Sporadic Late Onset AD (CU Hispanics; dbGaP study phs000496.v1.p1)[24], and subjects of European ancestry from the Alzheimer's Disease Neuroimaging Initiative (ADNI) study (adni.loni.usc.edu)[25]. Informed consent was obtained as described in the original studies. Individuals with northwestern European ancestry from the NIALOAD study were selected for the NIALOAD Europeans data set, while Hispanic individuals ascertained from site 4H were selected for the NIALOAD Hispanics data set. AD diagnoses met National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) and the Alzheimer's Disease and Related Disorders Association (ADRDA) criteria [26] and controls were screened for cognitive impairment using neurological and/or neuropsychological exams [23–25]. The NIALOAD study recruited subjects from families with an affected sibling pair with AAO of AD ≥ 60 years. The ADNI study used a longitudinal clinic-based research cohort study, with unrelated cases and controls. The CU Hispanics study recruited subjects using both familial AD (22%) and sporadic case-control (78%) ascertainment designs. We focus on the relationship between *APOE*, ancestry, and AD, and did not exclude the small proportion of subjects that may have pathogenic variants at other AD loci [27].

These data were filtered to include only those who had no missing data for sex, AD status, *APOE* genotypes, and either age-at-onset of dementia, age-at-last-evaluation (NIALOAD Europeans), or age when sampled (NIALOAD Hispanics). Age-at-last-evaluation was missing for all NIALOAD Hispanic individuals. Individuals with a difference of 10 years or more between age-at-onset of dementia and age-at-diagnosis of dementia were also excluded

from this study because of concerns about the accuracy of age-at-onset data estimated many years later. Similarly, the CU Hispanics data set was filtered to include only those who had no missing data for sex, AD status, *APOE* genotypes, and either age-at-onset of dementia or age-at-last-evaluation.

Public data from HapMap phase 3 (ftp://ftp.ncbi.nlm.nih.gov/hapmap/genotypes/2010-05_phaseIII) [28] and the Human Genome Diversity Project (HGDP; http://www.hagsc.org/hgdp/files.html) [29, 30] were downloaded for ancestry estimation analyses. Files were downloaded in PLINK [31] file format. Genotypes at 1,457,897 markers for 165 Europeans (CEU) and 203 Africans (YRI) were pulled from the HapMap phase III data, and a random sample of 63 Europeans and 63 Africans extracted. The HapMap genotypes were aligned to the forward strand of the build 36/hg18 human reference genome. From the HGDP genotype data, the genotypes at 660,918 markers aligned to the build 36/hg18 build of the human reference genome were extracted for the 63 samples with Pima, Mayan, or Colombian ancestry. We combined the HapMap and HGDP samples by merging genotypes at the 598,470 autosomal markers in PLINK. The marker positions were updated to build 37/hg19 using the LiftOver tool (http://genome.ucsc.edu/cgibin/hgLiftOver); 129 markers were excluded due to missing hg19 positions. These data were merged with the inference data set (NIALOAD Hispanics or CU Hispanics), the marker set reduced to the shared panel missing <5% data per marker, leaving 359,441 markers and 3,283 subjects for ancestry inference in the CU Hispanics and 498,677 markers and 730 subjects in the NIALOAD Hispanics analyses.

2.2. Statistical analyses

Cox proportional hazards analyses and Kaplan-Meier curves were calculated using the survival package in R [32–34]. Cox proportional hazards analyses of the complete data set included a null model, a model adjusting for the count of *APOE* E2 alleles, a model adjusting for the count of *APOE* E4 alleles, and a model adjusting for both the count of *APOE* E2 and E4 alleles. Kaplan-Meier curves were constructed for visualization of *APOE* effects using all samples with complete covariate data: 3,028 NIALOAD Europeans, 408 NIALOAD Hispanics, and 3,067 CU Hispanics. However, Cox proportional hazards analyses were restricted to one randomly-selected subject per family with complete covariate data to ensure validity of the associations: 1,503 NIALOAD Europeans, 148 NIALOAD Hispanics, and 2,492 CU Hispanics. Model fit was tested using a likelihood ratio test in R.

Ancestry proportions were estimated by combining 1000 genomes reference haplotypes (https://mathgen.stats.ox.ac.uk/impute/1000GP_Phase3.html) [35], Shapeit2 haplotype phasing software [36, 37], and RFMix software [38] with the HapMap and HGDP reference data and NIALOAD or CU Hispanics inference data using components of a published ancestry pipeline (https://github.com/armartin/ancestry_pipeline). Local ancestry proportions across markers were averaged per individual to estimate global ancestry proportions. Local ancestry proportions near *APOE* were defined as the average proportion in the interval spanning the *APOE* gene plus a flanking marker, chr19: 45,401,666–45,414,451 (hg19).

In the CU Hispanics, we identified subjects with predominantly African or European ancestry at *APOE* (probability ≥ 80%) and who were also homozygous for either the *APOE* E3 or E4 allele and performed Cox proportional hazards analyses using one randomly-selected subject per family within each ancestry group, which resulted in 766 CU Hispanics. We similarly obtained a sample of 784 NIALOAD Europeans that were homozygous for either the *APOE* E3 or E4 allele by randomly selecting one subject per family. NIALOAD Hispanics were excluded from these analyses due to small sample sizes. For the CU Hispanics, the Cox proportional hazards model adjusted for *APOE* E4/E4 genotype, % global Native American ancestry, % global African ancestry, and an indicator of African ancestry at the *APOE* locus. Case-control analyses were performed with all 880 CU Hispanics, including relatives, with predominantly African or European ancestry at *APOE* and homozygous *APOE* E3 or E4 genotypes using the logistic mixed model implementation in the GENESIS package [39] in R, which incorporates a genetic relatedness matrix (GRM) to adjust for relatedness within the sample. Principal components analyses were performed using PC-AiR [40] to account for relatedness between individuals, while an orthogonal GRM was estimated using PC-Relate [41] to adjust kinship estimates for the first four principal components (PCs). Similar analyses of Native American ancestry near *APOE* or *APOE* E2 homozygotes were excluded due to small sample sizes. Similar logistic mixed model case-control analyses were performed with all 1456 NIALOAD Europeans, including relatives, homozygous *APOE* E3 or E4 genotypes using GENESIS. For the ADNI data, logistic regression was performed in R with pre-calculated PCs, AAO/censoring age, affectation status, and *APOE* genotype data included as predictors.

3. Results

3.1. Sample description

Table 1 summarizes the sample sizes, distribution in age-at-onset (affected) or censoring age (at-risk), and frequencies of the *APOE* E2, E3, and E4 alleles across data sets. The NIALOAD Europeans and CU Hispanics have comparable sample sizes, while the NIALOAD Hispanics and ADNI Europeans are much smaller and have a higher proportion of affected individuals. The average age-at-onset and censoring age is slightly younger among the NIALOAD Europeans and older among the ADNI Europeans, while age-at-onset has a wider range among the CU Hispanics (70 years) when compared to the other data sets (34–56 years). Most data sets contain a small fraction of cases with age-at-onset ≤ 60 years: 4.0% of the NIALOAD Europeans, 6.5% of the NIALOAD Hispanics, and 7.5% of the CU Hispanics; sensitivity analyses excluding these subjects showed no impact on survival or association analyses. The CU Hispanics also have a lower frequency of the *APOE* E4 allele (21% vs. 31–32%) and higher frequency of the *APOE* E3 allele (74% vs. 62–64%) relative to the other three data sets.

Global ancestry proportions among the Hispanic data sets show a complex admixture of European, African, and Native American ancestry (Figure1, Supplemental Table 1). Both Hispanic data sets exhibit little Native American ancestry on average, with most individuals exhibiting primarily European ancestry with African admixture. On average, the NIALOAD Hispanics have more European and less African ancestry than the larger CU Hispanic

sample and appear to have two major clusters: most samples have predominately European ancestry, while a second cluster exhibits comparable European and African proportions. In contrast, the CU Hispanics exhibit more variation, with some individuals having predominately European, African, or Native American ancestry.

3.2. *Survival analyses: effects of* APOE *E2 and E4 alleles across populations*

In both the CU Hispanics and the NIALOAD Europeans, *APOE* E2 allele is associated with increased age-at-onset (AAO) of AD, while the E4 allele is associated with reduced AAO of AD (Figure 2, Supplemental Figure 1). However, we observe *APOE* E2 and E4 have weaker effects on the hazard of AD in the CU Hispanics when compared to the NIALOAD Europeans. While the effect of *APOE* E2 on the hazard of AD among the NIALOAD Europeans is highly significant (hazard ratio, HR, = 0.28, 95% CI = 0.19 – 0.40), the *APOE* E2 effect is substantially attenuated in the CU Hispanics (HR = 0.66, 95% CI = 0.54 – 0.81) and is not significantly associated with AAO in the much smaller NIALOAD Hispanic data set (HR = 0.59, 95% CI = 0.32 – 1.11). Similarly, while the effect of *APOE* E4 count on the hazard of AD is highly significant among the NIALOAD Europeans (HR = 3.44, 95% CI = 3.08 – 3.83), the effect is significant but attenuated in the CU Hispanics (HR = 1.98, 95% CI = 1.80 – 2.19), and is only marginally significant in the NIALOAD Hispanics (HR = 1.39, 95% CI = 1.02 – 1.89). Including both the count of *APOE* E2 and E4 alleles as covariates significantly improved the Cox proportional hazards models in each data set, relative to the null model: NIALOAD Europeans (X^2 = 479.11, df=2, p < 2e-16), NIALOAD Hispanics (X^2 = 6.43, df = 2, p = 0.0402), and CU Hispanics (X^2 = 167.57, df = 2, p < 2e-16). Adjusting the Cox proportional hazards models for global European ancestry proportions did not affect estimates of *APOE* effects on the hazard of AD, and the covariate itself was not associated with the hazard of AD (Supplemental Table 2). These results suggest that the relatively larger effect of *APOE* alleles on AAO of AD among the NIALOAD Europeans is not simply a function of sample size, as the NIALOAD European and CU Hispanic data sets are of comparable size and the two Hispanic data sets give similar effect sizes with very different sample sizes and ascertainment strategies. Nor are these differences explained by European admixture proportions. Instead, these results suggest that early association tests between *APOE* and AD in recently admixed populations were likely hindered by both small sample sizes and relatively smaller *APOE*-related effect sizes. Survival curves of AAO of AD by *APOE* genotype, rather than count of E2 or E4 alleles, showed similar patterns of weaker *APOE* effects in the Hispanic versus the European samples and suggest there may be different genotype-level relationships between *APOE* and AAO of AD across populations (Supplemental Figure 2).

3.3. Local ancestry modifies the relationship between APOE and AD

We further explored the relationship between ancestry and *APOE* effects on AD risk by comparing *APOE* E3 and *APOE* E4 homozygotes with either predominantly European or African derived ancestry at *APOE* within a subset of unrelated CU Hispanics. This effectively phases the *APOE* alleles onto haplotypes of either African or European origin and avoids assuming either an additive or a multiplicative effect of either allele on AD risk. There were only a few *APOE* E2 haplotypes and Native American derived haplotypes within the CU Hispanics, so they were not considered in this analysis.

Survival analyses of the CU Hispanics revealed significant effects for *APOE* E4/E4 genotype and global ancestry proportions on the hazard of AD (N = 766, Table 2). As expected, the *APOE* E4/E4 genotype significantly increased the hazard of AD (HR = 3.89, 95% CI = 2.69 – 5.63). However, as the proportion of non-European ancestry genome-wide increased, so did the hazard of AD. A one percentage point increase in genome-wide average Native American ancestry increased the hazard of AD by 3% (HR = 1.03, 95% CI = 1.01 – 1.06), and for a one percentage point increase in genome-wide average African ancestry genome-wide, the hazard of AD increased by 1% (HR = 1.01, 95% CI = 1.00 – 1.02).

Interestingly, we find that local ancestry at *APOE* in the CU Hispanics is a significant predictor of AD risk after adjustment for *APOE* genotype, global ancestry, and age (Table 3). A logistic mixed model association analysis including relatives (N = 880) found *APOE* E4/E4 genotype dramatically increased the odds of AD (odds ratio, OR, = 8.59, 95% CI = 4.49 – 16.43), while a one percentage point increase in genome-wide average Native American ancestry significantly increased the odds of AD by 3% (OR = 1.03, 95% CI = 1.00 – 1.06). We also found that the odds of AD increases by 2% for each one-year increase in age (OR = 1.02, 95% CI = 1.01 – 1.04). Interestingly, individuals with predominantly African ancestry at *APOE* had significantly reduced risk of AD, even after adjustment for *APOE* genotype, age, and global ancestry (OR = 0.61, 95% CI = 0.38 – 0.97), thus implicating ancestral origin of an *APOE* allele as an important predictor for risk of AD. This result suggests that local ancestry at the *APOE* locus is tagging additional genetic variation in the region that influences AD risk beyond E2 and E4, and these variants likely differ in frequency between European and African populations.

To further investigate the relationship between local *APOE* ancestry and background genetic variation, we performed similar association testing for AD risk among the NIALOAD European *APOE* E3 or E4 homozygotes (N = 1456; Supplemental Table 3) and the CU Hispanics with predominantly European ancestry at *APOE* (N = 700; Supplemental Table 4). Both analyses adjusted for population structure, age in years, an *APOE* E4/E4 effect, and relatedness. We found the *APOE* E4/E4 effect was approximately three times larger in the NIALOAD Europeans (OR = 15.89, 95% CI = 11.66 – 21.66) than in the CU Hispanics with European-derived *APOE* alleles (OR = 5.60, 95% CI = 2.13 – 14.69). These results suggest genetic variation outside the *APOE* region, environmental factors, or both influence the estimated relationship between *APOE* genotype and risk of AD.

3.4. Comparing APOE effects in European and Caribbean Hispanic Individuals ascertained for familial versus sporadic AD

As the ascertainment of the NIALOAD Europeans and CU Hispanics differed, we compared *APOE* effects across both European and CU Hispanic data sets ascertained for familial AD versus sporadic AD. Among the Europeans, we see that the estimated *APOE* E4/E4 effect is about 20% larger in the unrelated ADNI sample ascertained for sporadic AD (N = 478, OR = 19.03, 95% CI = 8.37 – 43.29) than the NIALOAD Europeans (Supplemental Table 3). We see a similar pattern in the CU Hispanics, where the estimated *APOE* E4/E4 effect is twice as big in the sample ascertained for sporadic AD (N = 626, OR = 9.58, 95% CI = 3.84 –

23.90) versus familial AD (N = 254, OR = 4.59, 95% CI = 1.72 – 12.26) (Supplemental Table 5). These results suggest that *APOE* E4/E4 effects may be larger in samples ascertained for sporadic AD versus familial AD, regardless of ancestry. We again see greater *APOE* E4/E4 effect size estimates in the European versus the CU Hispanics, where the estimated OR is approximately three times larger for the familial AD set and twice as large for the sporadic AD set. These results provide further evidence of differential *APOE* effects by ancestry on AD risk.

4. Discussion

This study has shown that the estimated effect of *APOE* E2/E3/E4 genotypes on the risk of AD differs between European and Caribbean Hispanic data sets, and that those differences are influenced by local ancestry. Even when comparing data sets of similar size, the effects of *APOE* E2 and E4 alleles on risk were weaker in the Caribbean Hispanics than the NIALOAD Europeans. This weaker effect size for *APOE* alleles on the hazard of AD was robust to adjustment for global European ancestry. Among Hispanics with predominantly African or European ancestry at *APOE* and homozygous E3 or E4 genotypes ("ancestry-phased" alleles), survival analyses found *APOE* genotype and global ancestry significantly influenced the hazard of AD, while a logistic model adjusting for *APOE* genotype, global ancestry, and age revealed the association between local ancestry at *APOE* and AD risk. The estimated increase in risk of AD among *APOE* E4 homozygotes versus *APOE* E3 homozygotes was approximately three times larger in the NIALOAD Europeans than among the CU Hispanics with European-derived *APOE* alleles. Taken together, our study has shown the effects of *APOE* E3 and E4 alleles on risk of AD differ depending on their local ancestry, or haplotypic background, in addition to genetic or ancestry differences across genome.

Even after adjustment for *APOE* E2/E3/E4 genotype, local ancestry is associated with risk of AD, which suggests that additional genetic variation in the *APOE* region influences AD risk. *APOE* E2/E3/E4 alleles occur on haplotypes defined by two missense variants, which themselves contain variation influencing DNA sequence and gene regulation. The sequence context of the *APOE* E4 allele varies considerably across populations [42, 43], and resequencing of *APOE* has identified ancestry-specific coding changes [43, 44]. Sequence variation around *APOE* can also significantly affect methylation patterns [45], suggesting that both the protein composition and gene expression can vary among individuals with the same *APOE* E2/E3/E4 genotype. Similar to the work by Zaitlan et al. [46], who found that heritability estimates based upon local ancestry captured more of the genetic variance than a genotyped marker panel, we believe that local ancestry in the *APOE* region is likely tagging the aggregated effects of both common and rare genetic variation influencing AD risk that varies in frequency across populations. If this is indeed the case, our logistic regression results for the Caribbean Hispanics samples suggest that there may be additional coding or regulatory variation influencing AD risk that vary in frequency between African and European populations in the *APOE* region beyond the E2/E3/E4 alleles.

Our study cautions against applying results from association testing in one population to estimate personalized risk in members of another population. Polygenic risk scores are used

to estimate individual level risk using genotypes at genome-scan markers and published effect size estimates for each allele. Application of effect sizes estimated in subjects with European ancestry to other populations creates directional inconsistencies and biased risk scores [47]. Population-specific functional alleles or patterns of linkage disequilibrium may explain these discrepancies in part, as the genome-scan marker is less able to tag the underlying risk variant [47]. We have shown that even within the same population, the effect sizes for *APOE* alleles vary depending on local ancestry, as supported by recent work in other Hispanic and African American samples[48]. Others have shown similar results at other AD-associated loci in African-American samples [9], although they did not investigate the *APOE* region. This suggests that local ancestry, as well as population-specific sequence variation, allele frequencies, and patterns of linkage disequilibrium should be considered when predicting genetic risk of AD in populations with non-European ancestry.

Additional studies in larger and more diverse data sets are needed to evaluate the complex relationship between ancestry, *APOE*, and AD. Our logistic models restricted to samples homozygous for both local ancestry and *APOE* genotype allowed us to focus on the relationship between local ancestry, *APOE* genotype, and AD, but do not represent the full, complex relationships between *APOE* genotype, ancestry, and risk of AD across Caribbean Hispanic populations. Furthermore, ancestry proportions vary widely within and between admixed population, such as Hispanic [49, 50] and African American [9, 51] populations, as we observed in our two Caribbean Hispanic data sets. Future genetic studies across diverse populations have great potential to provide new insights into the range of consequences of the well-established *APOE* E2/E3/E4 isoforms and their interactions with regulatory variants.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements/Conflicts/Funding Sources

We thank the study participants and their families for their contribution to this research. The authors declare no conflicts of interest. Research funding was provided by National Institute on Aging (NIA) [grant numbers R00AG040184, R01AG059737, P50AG005136]. The Genetic Studies of Alzheimer's disease in Caribbean Hispanics (EFIGA, dbGaP Study Accession phs000496.v1.p1) were funded by the NIA [grant numbers 5R37AG015473, RF1AG015473, R56AG051876]. We acknowledge the EFIGA study participants and the EFIGA research and support staff for their contributions to this study. The National Institute on Aging Late-Onset Alzheimer's Disease study (dbGaP Study Accession phs000168.v2.p2) was also supported by the NIA [grant numbers U24AG026395, R01AG041797]. ADNI data collection and genotyping were supported by U01 AG024904 (Weiner M, PI). Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd. and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. 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 coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

References

[1]. Hebert LE, Weuve J, Scherr PA, Evans DA. Alzheimer disease in the United States (2010–2050) estimated using the 2010 census. Neurology. 2013;80:1778–83. [PubMed: 23390181]

[2]. Lane CA, Hardy J, Schott JM. Alzheimer's disease. Eur J Neurol. 2017.

[3]. Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, Berg S, et al. Role of genes and environments for explaining Alzheimer disease. Arch Gen Psychiatry. 2006;63:168–74. [PubMed: 16461860]

[4]. Goate A, Chartier-Harlin MC, Mullan M, Brown J, Crawford F, Fidani L, et al. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. Nature. 1991;349:704–6. [PubMed: 1671712]

[5]. Sherrington R, Rogaev EI, Liang Y, Rogaeva EA, Levesque G, Ikeda M, et al. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. Nature. 1995;375:754–60. [PubMed: 7596406]

[6]. Levy-Lahad E, Wasco W, Poorkaj P, Romano DM, Oshima J, Pettingell WH, et al. Candidate gene for the chromosome 1 familial Alzheimer's disease locus. Science. 1995;269:973–7. [PubMed: 7638622]

[7]. Levy-Lahad E, Wijsman EM, Nemens E, Anderson L, Goddard KA, Weber JL, et al. A familial Alzheimer's disease locus on chromosome 1. Science. 1995;269:970–3. [PubMed: 7638621]

[8]. Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science. 1993;261:921–3. [PubMed: 8346443]

[9]. Hohman TJ, Cooke-Bailey JN, Reitz C, Jun G, Naj A, Beecham GW, et al. Global and local ancestry in African-Americans: Implications for Alzheimer's disease risk. Alzheimers Dement. 2016;12:233–43. [PubMed: 26092349]

[10]. Popejoy AB, Fullerton SM. Genomics is failing on diversity. Nature. 2016;538:161–4. [PubMed: 27734877]

[11]. MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E, et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Res. 2017;45:D896–D901. [PubMed: 27899670]

[12]. Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. JAMA. 1997;278:1349–56. [PubMed: 9343467]

[13]. Jun GR, Chung J, Mez J, Barber R, Beecham GW, Bennett DA, et al. Transethnic genome-wide scan identifies novel Alzheimer's disease loci. Alzheimers Dement. 2017;13:727–38. [PubMed: 28183528]

[14]. Chen HY, Panegyres PK. The Role of Ethnicity in Alzheimer's Disease: Findings From The C-PATH Online Data Repository. J Alzheimers Dis. 2016;51:515–23. [PubMed: 26890783]

[15]. Benedet AL, Moraes CF, Camargos EF, Oliveira LF, Souza VC, Lins TC, et al. Amerindian genetic ancestry protects against Alzheimer's disease. Dement Geriatr Cogn Disord. 2012;33:311–7. [PubMed: 22759767]

[16]. Moreno DJ, Pino S, Rios A, Lopera F, Ostos H, Via M, et al. Genetic Ancestry and Susceptibility to Late-Onset Alzheimer Disease (LOAD) in the Admixed Colombian Population. Alzheimer Dis Assoc Disord. 2017;31:225–31. [PubMed: 28369008]

[17]. Teruel BM, Rodriguez JJ, McKeigue P, Mesa TT, Fuentes E, Cepero AA, et al. Interactions between genetic admixture, ethnic identity, APOE genotype and dementia prevalence in an admixed Cuban sample; a cross-sectional population survey and nested case-control study. BMC Med Genet. 2011;12:43. [PubMed: 21435264]

[18]. Tang MX, Stern Y, Marder K, Bell K, Gurland B, Lantigua R, et al. The APOE-epsilon4 allele and the risk of Alzheimer disease among African Americans, whites, and Hispanics. JAMA. 1998;279:751–5. [PubMed: 9508150]

[19]. Tang MX, Maestre G, Tsai WY, Liu XH, Feng L, Chung WY, et al. Relative risk of Alzheimer disease and age-at-onset distributions, based on APOE genotypes among elderly African Americans, Caucasians, and Hispanics in New York City. Am J Hum Genet. 1996;58:574–84. [PubMed: 8644717]

[20]. Sahota A, Yang M, Gao S, Hui SL, Baiyewu O, Gureje O, et al. Apolipoprotein E-associated risk for Alzheimer's disease in the African-American population is genotype dependent. Ann Neurol. 1997;42:659–61. [PubMed: 9382479]

[21]. Corbo RM, Scacchi R. Apolipoprotein E (APOE) allele distribution in the world. Is APOE*4 a 'thrifty' allele? Ann Hum Genet. 1999;63:301–10. [PubMed: 10738542]

[22]. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016;536:285–91. [PubMed: 27535533]

[23]. Lee JH, Cheng R, Graff-Radford N, Foroud T, Mayeux R, National Institute on Aging Late-Onset Alzheimer's Disease Family Study G. Analyses of the National Institute on Aging Late-Onset Alzheimer's Disease Family Study: implication of additional loci. Arch Neurol. 2008;65:1518–26. [PubMed: 19001172]

[24]. Blue EE, Yu CE, Thornton TA, Chapman NH, Kernfeld E, Jiang N, et al. Variants regulating ZBTB4 are associated with age-at-onset of Alzheimer's disease. Genes Brain Behav. 2018;17:e12429. [PubMed: 29045054]

[25]. Saykin AJ, Shen L, Yao X, Kim S, Nho K, Risacher SL, et al. Genetic studies of quantitative MCI and AD phenotypes in ADNI: Progress, opportunities, and plans. Alzheimers Dement. 2015;11:792–814. [PubMed: 26194313]

[26]. Blacker D, Albert MS, Bassett SS, Go RC, Harrell LE, Folstein MF. Reliability and validity of NINCDS-ADRDA criteria for Alzheimer's disease. The National Institute of Mental Health Genetics Initiative. Arch Neurol. 1994;51:1198–204. [PubMed: 7986174]

[27]. Cruchaga C, Haller G, Chakraverty S, Mayo K, Vallania FL, Mitra RD, et al. Rare variants in APP, PSEN1 and PSEN2 increase risk for AD in late-onset Alzheimer's disease families. PLoS One. 2012;7:e31039. [PubMed: 22312439]

[28]. International HapMap C, Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs RA, et al. Integrating common and rare genetic variation in diverse human populations. Nature. 2010;467:52–8. [PubMed: 20811451]

[29]. Jakobsson M, Scholz SW, Scheet P, Gibbs JR, VanLiere JM, Fung HC, et al. Genotype, haplotype and copy-number variation in worldwide human populations. Nature. 2008;451:998–1003. [PubMed: 18288195]

[30]. Li JZ, Absher DM, Tang H, Southwick AM, Casto AM, Ramachandran S, et al. Worldwide human relationships inferred from genome-wide patterns of variation. Science. 2008;319:1100–4. [PubMed: 18292342]

[31]. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559–75. [PubMed: 17701901]

[32]. Therneau T A Package for Survival Analysis in S. version. 2.38 ed2015.

[33]. Terry M Therneau PMG. Modeling Survival Data: Extending the Cox Model. New York: Springer; 2000.

[34]. Team RC. R: A language and environment for statistical computing. Vienna, Austria: the R Foundation for Statistical Computing; 2011.

[35]. Delaneau O, Marchini J, Genomes Project C, Genomes Project C. Integrating sequence and array data to create an improved 1000 Genomes Project haplotype reference panel. Nat Commun. 2014;5:3934. [PubMed: 25653097]

[36]. Delaneau O, Howie B, Cox AJ, Zagury JF, Marchini J. Haplotype estimation using sequencing reads. Am J Hum Genet. 2013;93:687–96. [PubMed: 24094745]

[37]. O'Connell J, Gurdasani D, Delaneau O, Pirastu N, Ulivi S, Cocca M, et al. A general approach for haplotype phasing across the full spectrum of relatedness. PLoS Genet. 2014;10:e1004234. [PubMed: 24743097]

[38]. Maples BK, Gravel S, Kenny EE, Bustamante CD. RFMix: a discriminative modeling approach for rapid and robust local-ancestry inference. Am J Hum Genet. 2013;93:278–88. [PubMed: 23910464]

[39]. Conomos MP, Thornton T. GENESIS: GENetic EStimation and Inference in Structured samples (GENESIS): Statistical methods for analyzing genetic data from samples with population structure and/or relatedness. R package version 2.4.0. 2016.

[40]. Conomos MP, Miller MB, Thornton TA. Robust inference of population structure for ancestry prediction and correction of stratification in the presence of relatedness. Genet Epidemiol. 2015;39:276–93. [PubMed: 25810074]

[41]. Conomos MP, Reiner AP, Weir BS, Thornton TA. Model-free Estimation of Recent Genetic Relatedness. Am J Hum Genet. 2016;98:127–48. [PubMed: 26748516]

[42]. Babenko VN, Afonnikov DA, Ignatieva EV, Klimov AV, Gusev FE, Rogaev EI. Haplotype analysis of APOE intragenic SNPs. BMC Neurosci. 2018;19:16. [PubMed: 29745836]

[43]. Fullerton SM, Clark AG, Weiss KM, Nickerson DA, Taylor SL, Stengard JH, et al. Apolipoprotein E variation at the sequence haplotype level: implications for the origin and maintenance of a major human polymorphism. Am J Hum Genet. 2000;67:881–900. [PubMed: 10986041]

[44]. Radwan ZH, Wang X, Waqar F, Pirim D, Niemsiri V, Hokanson JE, et al. Comprehensive evaluation of the association of APOE genetic variation with plasma lipoprotein traits in U.S. whites and African blacks. PLoS One. 2014;9:e114618. [PubMed: 25502880]

[45]. Liu J, Zhao W, Ware EB, Turner ST, Mosley TH, Smith JA. DNA methylation in the APOE genomic region is associated with cognitive function in African Americans. BMC Med Genomics. 2018;11:43. [PubMed: 29739406]

[46]. Zaitlen N, Pasaniuc B, Sankararaman S, Bhatia G, Zhang J, Gusev A, et al. Leveraging population admixture to characterize the heritability of complex traits. Nat Genet. 2014;46:1356–62. [PubMed: 25383972]

[47]. Kim MS, Patel KP, Teng AK, Berens AJ, Lachance J. Genetic disease risks can be misestimated across global populations. Genome Biol. 2018;19:179. [PubMed: 30424772]

[48]. Rajabli F, Feliciano BE, Celis K, Hamilton-Nelson KL, Whitehead PL, Adams LD, et al. Ancestral origin of ApoE epsilon4 Alzheimer disease risk in Puerto Rican and African American populations. PLoS Genet. 2018;14:e1007791. [PubMed: 30517106]

[49]. Levran O, Awolesi O, Shen PH, Adelson M, Kreek MJ. Estimating ancestral proportions in a multi-ethnic US sample: implications for studies of admixed populations. Hum Genomics. 2012;6:2. [PubMed: 23244743]

[50]. Halder I, Yang BZ, Kranzler HR, Stein MB, Shriver MD, Gelernter J. Measurement of admixture proportions and description of admixture structure in different U.S. populations. Hum Mutat. 2009;30:1299–309. [PubMed: 19572378]

[51]. Mathias RA, Taub MA, Gignoux CR, Fu W, Musharoff S, O'Connor TD, et al. A continuum of admixture in the Western Hemisphere revealed by the African Diaspora genome. Nat Commun. 2016;7:12522 [PubMed: 27725671]

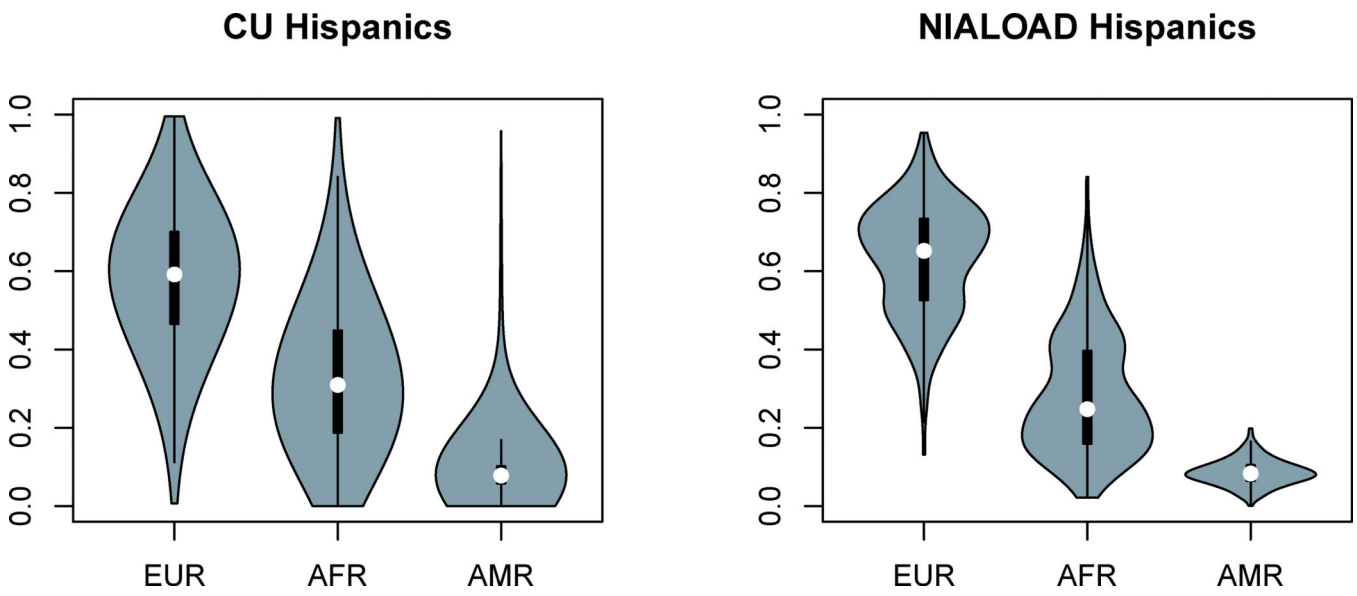

Figure 1. Violin plots of ancestry proportions (vertical axis) among the Hispanic data sets.
EUR: European ancestry proportion, AFR: African ancestry proportion, AMR: Native American ancestry proportion, white dot: median, black bar: interquartile range, black line: 95% confidence interval, grey curves: kernel density plot.

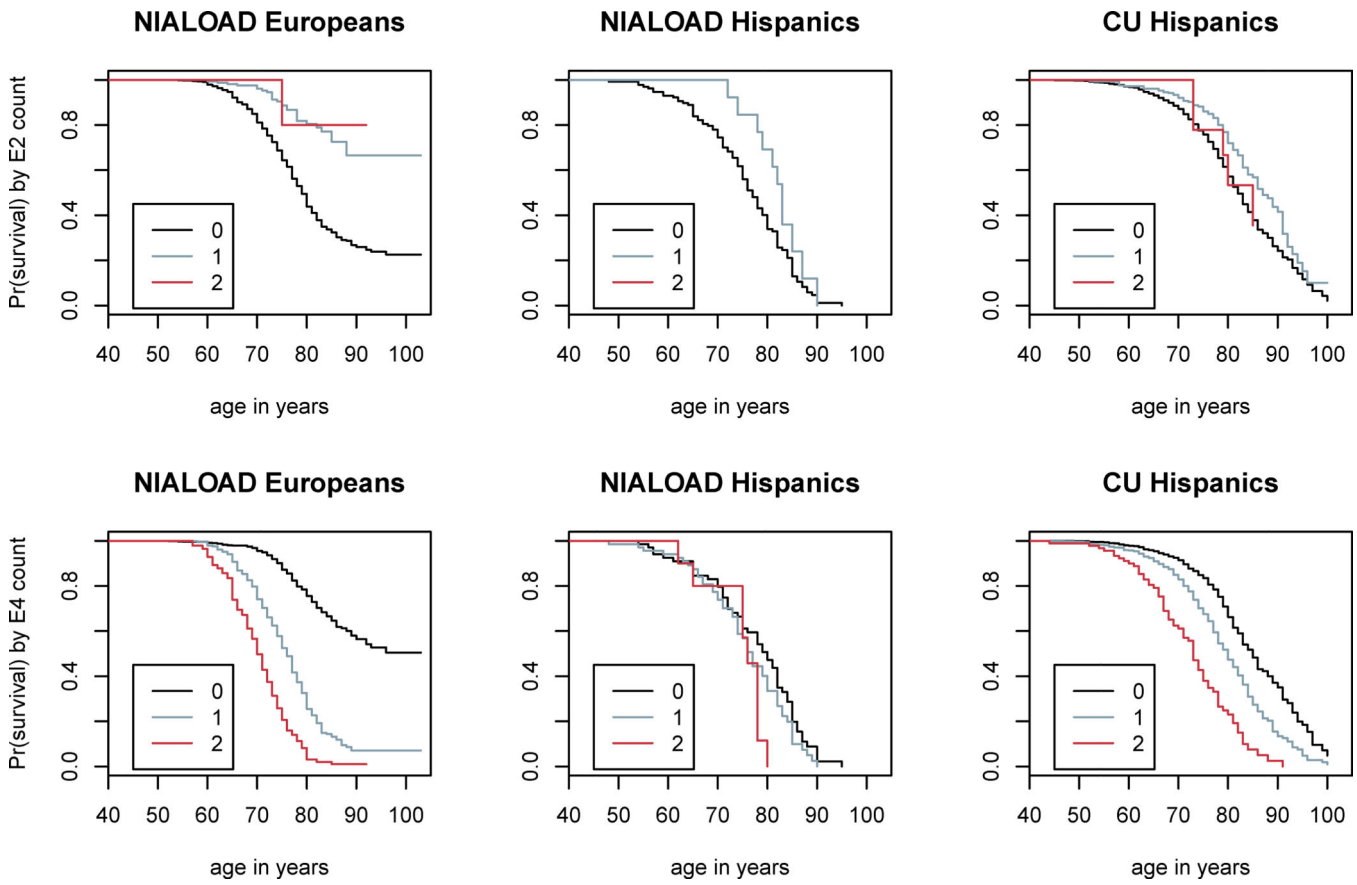

Figure2. Kaplan-Meier survival curves by *APOE* allele counts.

Pr(survival): probability of not being diagnosed with Alzheimer's disease. Kaplan-Meier survival analyses were restricted samples with complete covariate data: 3,028 NIALOAD Europeans, 408 NIALOAD Hispanics, and 3,067 CU Hispanics.

Table 1.

Description of the data.

	NIALOAD Europeans	NIALOAD Hispanics	CU Hispanics	ADNI Europeans
N_{total}	3028	408	3067	909
$N_{affected}$	1238	292	1329	587
$N_{markers}$	592,126	595,243	904,966	NA
$Age_{affected}$	73.28 [48–96]	74.82 [42–98]	74.84 [30–100]	76.72 [60–94]
$Age_{at\text{-}risk}$	69.37 [42–103]	67.95 [50–95]	73.17 [35–100]	74.11 [60–90]
%Female	62%	64%	66%	43%
APOE E2 freq	5%	5%	6%	4%
APOE E3 freq	62%	64%	74%	64%
APOE E4 freq	32%	31%	21%	31%

Table 2.

Survival analysis for Alzheimer's disease among subjects homozygous for either African- or European-derived *APOE* E3 or E4 alleles in the CU Hispanics.

Analysis model adjusted for global ancestry, local ancestry at *APOE*, and *APOE* genotype (N = 766, 283 events). Analysis restricted to *APOE* E3 or E4 homozygotes with predominantly African or European ancestry at *APOE*, one subject per family. 100 bootstrap replicates of sampling of one subject per family were used to provide percentiles for covariate effects. AMR: Native American, AFR: African, local AFR ancestry: indicates African ancestry (probability > 0.80) at the *APOE* locus, HR: hazard ratio, SE: standard error.

Covariate	HR	2.5%ile (HR)	97.5%ile (HR)
APOE E4/E4 genotype	3.8894	2.6878	5.6280
% global AMR ancestry	1.0333	1.0104	1.0570
% global AFR ancestry	1.0094	1.0017	1.0170
local AFR ancestry at *APOE*	0.7251	0.4893	1.0740

Table 3.

Association testing Alzheimer's disease among subjects homozygous for either African- or European-derived *APOE* E3 or E4 alleles in the CU Hispanics.

Analysis model adjusted for global ancestry, local ancestry at *APOE*, *APOE* genotype, and age in years (N = 880, includes relatives). Analysis restricted to *APOE* E3 or E4 homozygotes with predominantly African or European ancestry at *APOE* (probability > 0.80). AMR: Native American, AFR: African, local AFR ancestry: indicates African ancestry (probability > 0.80) at the *APOE* locus, OR: odds ratio, SE: standard error.

Covariate	β	SE (β)	OR	2.5%ile (OR)	97.5%ile (OR)
APOE E4/E4 genotype	2.1510	0.3308	8.5935	4.4935	16.4343
% global AMR ancestry	0.0329	0.0153	1.0334	1.0029	1.0649
% global AFR ancestry	0.0042	0.0048	1.0042	0.9948	1.0137
local AFR ancestry at *APOE*	–0.5012	0.2410	0.6058	0.3777	0.9716
Age (years)	0.0242	0.0083	1.0245	1.0080	1.0413