Published in final edited form as:
Dement Geriatr Cogn Disord. 2013 ; 35(0): 340–346. doi:10.1159/000343074.

Aromatase Variants Modify Risk for Alzheimer's Disease in a Multiethnic Female Cohort

S.C. Janicki, M.D., M.P.H.[1,2,3], **N. Park, M.S.**[1,2], **R. Cheng, Ph.D.**[1,2], **N. Schupf, Ph.D, Dr.PH**[1,2,4,5], **L.N. Clark, Ph.D.**[1,6,7], and **J. H. Lee, Dr. PH**[1,2,5,6]

[1]Taub Institute for Research on Alzheimer's Disease and the Aging Brain, College of Physicians and Surgeons, Columbia Univ., NY, NY

[2]Gertrude H. Sergievsky Center, College of Physicians & Surgeons, Columbia Univ., NY, NY

[3]Department of Neurology, College of Physicians & Surgeons, Columbia Univ., NY, NY

[4]Department of Psychiatry, College of Physicians & Surgeons, Columbia Univ., NY, NY

[5]Departments of Epidemiology and Psychiatry, Columbia Univ. Medical Center, NY, NY

[6]Center for Human Genetics, College of Physicians & Surgeons, Columbia Univ., NY, NY

[7]Department of Pathology and Cell Biology, College of Physicians & Surgeons, Columbia Univ., NY, NY

Abstract

Background/Aims—Few studies of gene variants that affect estrogen activity investigate their association with risk for AD in women of different ethnicities. We investigated the influence of *CYP19* polymorphisms on risk for AD in a multiethnic cohort of women, with individual ethnicity assessed by genetic population ancestry markers (AIMs) as well as by self -identified ethnicity.

Methods—Among 1686 women participating in the Washington Heights Inwood Columbia Aging Project (WHICAP), association with risk for AD was assessed for 41 single-nucleotide

Correspondence: Sarah C. Janicki, MD, MPH; Gertrude H. Sergievsky Center, Columbia University Medical Center, P&S Box 16, 622 West 168th Street, New York, NY 10032; Tel: 212-305-9194; Fax: 212-305-2526; scj2110@columbia.edu.

Financial Disclosures
Dr. Janicki, author, is funded by NIH grant 5-T32MH020004-1, the Charles and Ann Lee Saunders Brown Fellowship Fund, and the Louis V. Gerstner Jr. Scholars Foundation. Dr. Janicki reported no biomedical financial interests or potential conflicts of interest.
Ms. Park, author, is funded by 5P01 HD035897, R01 MH084995, and R01 NS060113. Ms. Park reported no biomedical financial interests or potential conflicts of interest.
Dr. Cheng, author, is funded by NIH grants U24 AG026395, R01 AG028786, R01 NS060113, R01 MH084995, R01 AG036469, and P50 AG008702. Dr. Cheng reported no biomedical financial interests or potential conflicts of interest.
Dr. Schupf, author, is funded by NIH grants 5R01AG014673-09, R01AG037212, U01 AG023749, R01 AG028786, P01 HD35897, P50 AG08702, R01 AG034189, 1 R21CA125461-01A2, 1 R01AG034189, 1 R01AG0306040-01, 5R01AG007370-17, and receives research support from the Alzheimer's Association. Dr. Schupf has consulted for Elan/Jannsen Pharmaceuticals
Dr. Clark, author, is funded by NIH grants R01 NS060113, R01 NS073872, 5 P50 AG008702, 1RC2NS070276, 2P50NS038370-11, the NINDS Parkinson's Disease iPS Cell Line Research Consortium and P50 "Genetics of Down Syndrome", and receives research support from the Alzheimer's Association, the Parkinson's Disease Foundation and the Michael J. Fox Foundation. Dr. Clark reported no biomedical financial interests or potential conflicts of interest.
Dr. Lee, author, is funded by NIH grants R37 AG15473, U24 AG026395, 5P01 HD035897, R01 MH084995, R01 AG028786, U01 AG023749, R01 NS060113, R01 AG036469, and receives research support from the Alzheimer's Association. Dr. Lee reported no biomedical financial interests or potential conflicts of interest.

polymorphisms (SNPs) on the *CYP19* gene using multivariable logistic regression, adjusting for age, presence of an *APOE* ε4 allele, years of education, and body mass index (BMI).

Results—Risk for AD was associated with six SNPs in women of predominantly Caucasian AIMs-defined ancestry. Of these, two were also associated with decreased risk of AD in women of admixed/ Hispanic AIMs ancestry. Two separate SNPs were found to be protective in women of predominantly African AIMs-based ancestry.

Conclusions—*CYP19* polymorphisms affect risk for AD in women, and risk alleles vary by AIMs-defined ancestry. These effects are possibly due to linkage disequilibrium patterns or differences in the prevalence of comorbid risk factors mediating SNP effect on risk for AD by group.

Keywords

Alzheimer disease; Alzheimer disease and Hispanics; estrogen; Genetic polymorphisms in degenerative dementias; genetic risk factors; *CYP19*

Introduction

Estrogens are important in maintaining brain function in regions typically affected by Alzheimer's disease (AD) and variations in estrogen exposure over the lifetime may affect cognitive decline associated with AD [1,2]. However, evaluating the role of hormones and enzymes in aging and cognition is difficult since many hormone levels decline with age. It is likely that polymorphisms in genes encoding the estrogen synthesis pathway contribute to variations in lifetime hormone exposure, including age-related changes in hormone levels. Estrogens including estradiol and estrone are formed locally in the brain from the conversion of androgens by aromatase [3], a cytochrome p450 enzyme encoded by the *CYP19* gene located on chromosome 15q21.2. Several studies [4–6], but not all [7,8], have found an association between multiple single nucleotide polymorphisms (SNPs) in *CYP19* and AD. However, most studies have been conducted in Caucasian ethnic groups, and few polymorphisms have been assessed in a multiethnic cohort in which the members have all been evaluated in a consistent manner. Examination of SNPs in multiethnic groups which are evaluated without taking ancestry into account may have several limitations, including a loss of significant association due to different allele frequencies, different linkage disequilibrium patterns between ethnicities, or differences in the distribution of comorbid conditions and risk factors for AD by ethnic group. In this study, we examined the relationship between *CYP19* SNPs and the risk of AD in a multiethnic cohort of elderly women from northern Manhattan, with individual ancestry assessed by population ancestry markers as well as by self-identified ethnicity [9]. The aims of this study were to confirm previous findings of *CYP19* polymorphisms which were found to be significantly associated with risk for AD; to identify additional SNPs which confer risk for AD using a denser set of SNPs than in previous studies; and to examine whether *CYP19* variants would affect risk for AD differently in groups of women with different population ancestry. We hypothesized that genetic variants would demonstrate different patterns of association between groups with different population ancestries due to distinctive allele frequencies or linkage disequilibrium patterns between ethnic groups, as well as varying environmental factors.

Materials and Methods

Subjects

The study included 1,686 women participating in the Washington Heights Inwood Columbia Aging Project (WHICAP), a prospective study of aging and dementia among Medicare recipients age 65 years and older, residing in northern Manhattan. The population from which participants were drawn was comprised of individuals from several different countries of origin representing three broadly self – identified ethnicities (Caribbean Hispanic, African – American, and non – Hispanic White of European ancestry). The sampling strategies and recruitment outcomes of these two cohorts have been described in detail elsewhere [9].

Each subject underwent an in – person interview of health and functional ability followed by a standardized medical assessment and neuropsychological battery [10]. Assessments were conducted at 18 – 24 month intervals over a mean of 6.1 years of follow-up. AD diagnosis was based on NINCDS-ADRDA criteria. We used a conservative definition of AD in our analyses, excluding definitions of mild cognitive impairment (MCI) or isolated low neuropsychological scores in order to obtain the most robust phenotype.

Standard Protocol Approvals, Registrations, and Patient Consents

This study was reviewed and approved by the Columbia University Institutional Review Board, and written informed consent was previously obtained from all subjects.

DNA Isolation, SNP selection and Genotyping

Genomic DNA was extracted from total peripheral blood leukocytes using standard methods. We used a multistep selection process to identify candidate SNPs for genotyping. We first selected SNPs within *CYP19* that were previously reported to be associated with an increased incidence or earlier age at onset of AD in any population. We then referenced the International HapMap Project (www.hapmap.org) to select tagging SNPs in both Caucasian and African populations. To provide sufficient coverage of the gene, we selected SNPs to maintain a pairwise r^2 threshold of 0.8 in SNPs with a minimum minor allele frequency of 0.2. We obtained an average intermarker distance of approximately 3.0 kilobase pairs between SNPs, which provided good coverage of the gene as viewed on linkage disequilibrium maps (Supplementary Figures 1–3).

Forty-one *CYP19* SNPs as well as 100 ancestry informative markers (AIMs) were genotyped in a total of 1,686 samples using Illumina GoldenGate custom panels and the Illumina IScan platform. Genotyping was performed according to standard protocols (www.Illumina.com). Duplicate genotyping was performed on ten percent of samples to verify accuracy, and the concordance rate was greater than 98 percent.

Assessment of genetic population ancestry

To evaluate population stratification, we used a set of 100 unlinked ancestry informative markers (AIMs) to classify population ancestry. We selected the 100 unlinked SNPs from a panel of 650Y Illumina SNPs using a subset of subjects that had previously also had GWAS data collected. The AIMs were selected because they have allele frequencies that are

significantly different among three ethno – racial groups: non-Hispanic Whites, non-Hispanic African, and individuals of Mexican/Central American ancestry. To assess population stratification, we performed population structure analysis as implemented in the STRUCTURE program [11,12]. To anchor ancestry, we included data from Caucasians (CEPH), Yorubans (YRI) and Mexican/Central Americans from the HapMap project (Figure 1). Our self – identified White population closely aligned with the Caucasian (CEPH) samples in the HapMap dataset and our self – identified Black population clustered around the Yoruban (YRI) samples. As expected, Caribbean Hispanics clearly showed admixture of Caucasian (CEPH) and Yoruban (YRI) genetic population ancestry, and the range of admixture varied widely. We then classified participants into groups who were of predominant Caucasian ancestry as defined by the AIMs index (defined as ≥ 0.6 AIMs markers consistent with CEPH profile, n= 632) versus those who were of predominant African ancestry (defined as ≥ 0.6 AIMs markers consistent with YRI profile, n= 581). In doing so, individuals previously self-identified as Hispanic were reclassified as being of predominant Caucasian or African AIMs-defined ancestry (if their AIMs index scores were ≥ 0.6 CEPH or YRI, respectively), or admixed/ Hispanic if they did not have one predominant genetic ancestry (n= 473). Comparison of populations as defined by AIMs-defined ancestry versus self-identified ethnicity are illustrated in Supplementary Table 2.

Potential Confounders

Potential confounders included age at time of study enrollment, presence of an *APOE* ε4 allele, years of education, and body mass index (BMI). Supplementary analyses (as shown in Supplementary Figures 4–6) were also performed to evaluate the potential effects of vascular risk factor covariates on significant SNPs, and included history of diabetes mellitus and current smoking. Participants were classified according to the presence or absence of at least one *APOE* ε4 allele. Height and weight were measured at the initial evaluation to compute BMI. History of diabetes mellitus was defined as self – reported current or past history of treated or untreated diabetes.

Statistical Analyses

Prior to association analysis, we assessed whether each SNP was in Hardy Weinberg equilibrium. This analysis was performed separately within each self – identified ethnicity as well as within each AIMs – defined population of unaffected individuals using the χ^2 goodness-of-fit test in HAPLOVIEW [13]. SNPs were then evaluated in genotypic association analyses to further characterize their relationship to AD, stratifying first by AIMs – defined ancestry and then by self-reported ethnicity. We hypothesized that differences in associations between these two sets of analyses might reflect culturally – associated environmental risk factors for AD. Conversely, similarities in significant SNPs between the two analyses would demonstrate a more direct genetic effect of *CYP19* polymorphisms on risk for AD. We used multivariable logistic regression to estimate likelihood of AD by SNP genotype, adjusting for age, presence of at least one *APOE* ε4 allele, and BMI. To provide the most robust model for observing an effect of the minor allele, SNPs were analyzed using a dominant model, in which participants homozygous for the common allele were used as the reference group and the risk group included participants who were heterozygous or homozygous for the minor allele.

Results

Demographic Characteristics

Table 1 presents the demographic characteristics of our cohort. The mean age of the participants at baseline was 77.0 (± 6.7) years, and ranged from 65 to 95 years. Mean length of follow-up was 6.1 (±4.3) years. The majority of women were self – identified as Hispanic (n= 672, 39.9%) and Black (n=574, 34.0%), while 423 women were self – identified as White (25.1%). Among all participants, 511 were classified as possible or probable AD (29.5%) and 1175 as nondemented. The frequency of AD was greater in self – identified Blacks and Hispanics than in Whites (Blacks: 32.2%; Hispanics: 38.2%; Whites: 15.3%). Mean years of education differed significantly between those with and without AD (7.1 years versus 10.4 years). BMI, history of diabetes mellitus, and current smoking status did not differ significantly between individuals with or without AD. However, there were statistically significant differences between self – identified groups as well as between AIMs-defined populations in the prevalence of vascular risk factors, including diabetes mellitus and current smoking (Table 2). The proportion of women with at least one copy of the *APOE* ε4 allele was the highest in Blacks, followed by Hispanics and then by Whites (Table 2).

Genotypic Associations

For ease of discussion, we will use the numbered order of SNPs to refer to each SNP (Table 3). Among women of predominantly Caucasian AIMs-defined ancestry, two SNPs (SNPs 10 and 15—rs4775935 and rs727479, respectively), located in at the 5′ end of the *CYP19* gene, were found to be associated with decreased risk for AD, adjusting for age, BMI, and presence of an *APOE* ε4 allele. Four SNPs (27, 28, 37, and 39 – rs17647719, rs1902586, rs10163138, and rs7168331), clustered at the 3′ end of the gene, were found to be associated with increased risk for AD. Odds ratios varied between and 0.6 to 0.7 among protective SNPs (Table 3), and ranged from 1.7 to 2.6 among SNPs associated with increased risk for AD. Among women of admixed/Hispanic AIMS-defined ancestry, SNPs 10 and 15 were also found to be protective (O.R. 0.6 and 0.7, respectively). Two different SNPs (SNPs 38 and 41– rs6493495 and rs11070843) were associated with decreased risk for AD in women of predominantly African AIMs-defined ancestry, both with odds ratios of 0.7 (Table 3). To minimize the risk of false-positive findings from multiple testing, we computed empirical p-values by generating the null distribution on the basis of 1000 replicate datasets. As shown in Table 3, calculation of empirical p-values slightly attenuated the degree of significance for some genotypes (including SNP 15 among women of admixed/Hispanic AIMs-defined ancestry and SNP 38 among women of predominantly African AIMS-defined ancestry), however most remained significant.

We then repeated the analyses within strata defined by self-identified ethnicity to take into account potential role of cultural/environmental risk factors within ethno-racial groups. (Supplementary Table 1). Among self-identified whites, SNPs 10 and 15 remained significant while ORs for the remaining four SNPs that were significant among women of predominantly Caucasian AIMs-defined ancestry became attenuated and were no longer significant. However, four additional SNPs flanking the region – specifically, SNPs 5, 16,

18, and 19 – were found to be significant. Among self-identified Hispanics, two significant SNPs (10 and 16) were found in close proximity to those in the AIMs-defined ancestry groups for Hispanics and Whites. Among self-identified Blacks, two protective SNPs (SNPs 38 and 31) were no longer found to be significantly associated with risk for AD.

Haplotype Analysis

Genotypic analyses demonstrated that SNPs associated with risk for AD clustered in several distinct regions of high LD (Supplementary Figures 1–3). Strong pairwise LD between SNP loci in these blocks supported the possibility of multi-locus association at adjacent variants. We performed "sliding window" haplotype analysis within these regions as implemented in the HAPLOVIEW program using the D′ value [13], with each haplotype including two to four consecutive SNPs. While numerous haplotypes constructed from these *CYP19* SNPs were found to be significantly associated with increased or decreased risk for dementia, the most robust associations in women of predominantly Caucasian AIMs-based ancestry were haplotype A – A at SNPs 10 – 11 (O.R. 0.4, p=0.001) and haplotype A – A – C at SNPs 37 – 38 – 39 (O.R 3.17, p=0.04) (data not shown). Among women of admixed/Hispanic AIMs-defined ancestry, the most significant haplotype was G – C – A – G at SNPs 14 – 15 – 16 – 17, which was protective (O.R. 0.6, p=0.008). Among women of predominantly African AIMs-defined ancestry, only one protective haplotype, G – G at SNPs 40 – 41 was found to be significant (O.R. 0.72, p=0.04) (data not shown).

Discussion

Among 1,686 community-dwelling elderly women in a multiethnic cohort, risk for developing AD was associated with six SNPs in women of predominantly Caucasian AIMs-defined ancestry (rs4775935, rs727479, rs17647719, rs1902586, rs10163138, and rs7168331). Of these, two SNPs (rs4775935 and rs727479) were associated with decreased risk of AD in women of admixed/Hispanic AIMs ancestry. Additionally, two different SNPs (rs6493495 and rs11070843) were found to be protective in women of predominantly African AIMs-based ancestry. Use of empirical p-values slightly attenuated the degree of significance for rs727479 among women of admixed/Hispanic AIMs-defined ancestry and rs6493495 among women of predominantly African AIMS-defined ancestry; however the directionality and magnitude of effect remained the same.

Numerous papers have established that estrogen may have beneficial effects on multiple pathways that affect risk for AD. Estrogens promotes the growth and survival of cholinergic neurons[14,15] increase cholinergic activity[16], have antioxidant properties[17], and promote the nonamyloidogenic metabolism of the amyloid precursor protein[18]. Estrogens also play an important role in regulation of the vascular endothelium where they activate rapid vasodilatation, exert anti-inflammatory effects, stimulate endothelial growth and migration, and protect the vessels from atherosclerotic degeneration by elevating nitric oxide and prostaglandin levels [19,20]. Aromatase is a potentially important factor in these processes as it controls estrogen biosynthesis and is expressed in regions of the brain affected by AD [21–24]. Previous studies have demonstrated that common *CYP19* polymorphisms are associated with estradiol and androgen serum levels in premenopausal

and postmenopausal women [25–28]. As a result, *CYP19* gene variants could potentially affect risk for AD by reducing or increasing rate of conversion of androgens into estrogens, resulting in altered protection against neuronal injury or degeneration through multiple mechanisms.

Differential association of polymorphisms in a susceptibility gene for AD in groups of disparate population ancestries may occur for several reasons. First, differences in LD patterns between ethnic groups may contribute to discrepancies in genotype associations (Supplementary Figures 1 – 3). In our cohort, SNPs which were protective against AD in women of predominantly Caucasian AIMs-defined ancestry were located at the 5′ end of the gene (Supplementary Figure 1, Block 2), and those which increased risk for AD in this population clustered in two LD blocks at the 3′ end of the gene (Supplementary Figure 1, Blocks 5 – 6). Notably, these LD blocks which remained cohesive in women of predominantly Caucasian ancestry were represented by smaller sets of SNPs in women of admixed/Hispanic AIMs-defined ancestry (Supplementary Figure 2). LD blocks in this region were even more fragmented among women of predominantly African AIMs-defined ancestry (Supplementary Figure 3). This suggests that different LD patterns between *CYP19* alleles and alleles of as yet unidentified loci for susceptibility to AD between populations of different genetic ancestries may contribute to the observed variability in genotypic association.

Second, it is also possible that differences in environmental or biological risk factors among women of different genetic ancestry may play a significant role in phenotypic expression of the variants. For example, the inclusion of vascular risk factors (including history of diabetes and current smoking) attenuated the significance of SNPs in women of admixed/Hispanic (Supplementary Figure 5) or predominantly African AIMs-based ancestry (Supplementary Figure 6). In contrast, the inclusion of these vascular risk factor covariates increased the significance of several SNPs in women of predominantly Caucasian AIMs-defined ancestry, notably in SNPs located at the 3′ end of the gene (Supplementary Figure 4). The potential mediation of SNP effect by vascular risk factors is reinforced by the differential results seen in the logistic regression analyses stratified by self-identified ethnicity versus AIMs-defined ancestry. As seen in Table 2, the change in stratification had the greatest effect on sample group characteristics and size when comparing self-identified Whites versus individuals of predominantly Caucasian AIMs-based ancestry, and self-identified Black versus women of predominantly African AIMS-based ancestry. Specifically, the shift in classification of race revealed that many individuals who had genetic markers that were predominantly Caucasian or African actually identified themselves as Hispanic. These individuals also had significantly higher rates of diabetes mellitus (Supplementary Figure 7) than individuals of predominantly African AIMs who identified themselves as Black or participants of predominantly Caucasian AIMs who identified themselves as White. Notably, when these individuals were included in the AIMs-based logistic regression models for participants of predominantly Caucasian or African ancestry, several SNPs at the 3′ end of the gene became significant (Table 3), again highlighting the potential mediation of SNP effect by vascular risk factors in this part of *CYP19*. These discrepancies indicate that environmental and biological covariate risk factors may exert different modulating effects with alleles in

groups of different predominant genetic ancestries, or even within different parts of the gene.

Several previous studies have investigated the role of *CYP19* polymorphisms and AD. A study of a Finnish population found three SNPs (rs767199, rs727479, and 1065778) to be associated with increased risk for AD, as well as haplotype A1 (CACTTTGTT) [5]. Other SNPs including rs2899472, rs1008805, rs727479, and rs1143704, rs1065778 rs10046, and rs4646 have also been found to be associated with risk for AD or age at onset (some in *APOE* ε4 carriers only)[4,6] in primarily Caucasian populations. In other studies, an interaction between rs1062033 (located in the 5′-UTR region) and the genes for butyrylcholinesterase[29] and interleukin-10 [30] in increasing risk for AD were proposed. In addition, rs2899472 was found to be associated with CSF $A\beta_{1\text{-}42}$ levels in normal subjects in a GWAS investigation [3131].

We examined the majority of the SNPs cited in the above studies. While rs727479 was protective among self-defined Whites in our group, we did not find any other SNPs to be significant. Although many of these previously evaluated SNPs are clustered at the 5′ end of the gene, in LD Blocks 1 and 2 by our map (Supplementary Figure 1), many of the SNPs which we found to be significant in our participants of predominantly Caucasian AIMs-defined ancestry clustered at the 3′ end, in LD Blocks 5–7. However, as previously noted, most previous studies had been conducted in self-identified White participants who may have had different associated risk factors from our Caucasian population, which also included individuals who identified themselves as Hispanic.

Overall, our findings confirm previous studies' findings of a strong association between *CYP19* polymorphisms and risk for Alzheimer's disease among women. We also extend these studies through denser genotyping, rather than relying on imputation which can introduce false positives in multi-ethnic cohorts. From this effort, we identified additional SNPs that are associated with AD risk, and characterized how these SNPs vary among individuals of different AIMs-defined ancestries in the presence of cardiovascular risk factors. Moreover, our study illustrates the importance of controlling for population stratification as well as for environmental risk factors in association studies, as SNPs which we found to be associated with AD varied significantly between ethnic groups.

We note that most SNPs examined were intronic, and therefore may not be the critical location of the pathological variants, but may serve as markers for the critical region or may otherwise influence the expression of critical genetic markers. Further studies may characterize other genetic mechanisms that may contribute to AD, including methylation and copy number variations (CNVs). For example, SNP 27 (rs17647719) is located in a region associated with methylation, and SNP 28 (rs1902586) is located 2.6 kb from this region. Future studies with denser genotyping to achieve high resolution in all ethnic groups, along with gene expression studies, may further provide biological insights. Additional insight may also be gained through future studies conducting similar analyses in men.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported by Alzheimer's Association IIRG-08-90655, "Genetics of Estrogen and Alzheimer's Disease in a Multiethnic Cohort"(N Schupf, PI) and National Institute on Aging funded grants P01AG007232 (P.I. R. Mayeux), R01AG037212 (P.I. R Mayeux and N. Schupf).

Bibliography

1. Shughrue PJ, Lane MV, Merchenthaler I. Comparative distribution of estrogen receptor-alpha and -beta mRNA in the rat central nervous system. J Comp Neurol. 1997; 388:507–525. DOI: 10.1002/(SICI)1096-9861(19971201)388:4<507::AID-CNE1>3.0.CO;2-6 [pii]. [PubMed: 9388012]
2. McEwen BS. Invited review: Estrogens effects on the brain: Multiple sites and molecular mechanisms. J Appl Physiol. 2001; 91:2785–2801. [PubMed: 11717247]
3. Stoffel-Wagner B, Watzka M, Schramm J, Bidlingmaier F, Klingmuller D. Expression of *CYP19* (aromatase) mRNA in different areas of the human brain. J Steroid Biochem Mol Biol. 1999; 70:237–241. [PubMed: 10622413]
4. Huang R, Poduslo SE. *CYP19* haplotypes increase risk for Alzheimer's disease. J Med Genet. 2006; 43:e42. DOI: 43/8/e42 [pii] 10.1136/jmg.2005.039461. [PubMed: 16882736]
5. Iivonen S, Corder E, Lehtovirta M, Helisalmi S, Mannermaa A, Vepsalainen S, Hanninen T, Soininen H, Hiltunen M. Polymorphisms in the *CYP19* gene confer increased risk for Alzheimer disease. Neurology. 2004; 62:1170–1176. [PubMed: 15079018]
6. Butler HT, Warden DR, Hogervorst E, Ragoussis J, Smith AD, Lehmann DJ. Association of the aromatase gene with Alzheimer's disease in women. Neurosci Lett. 2010; 468:202–206. DOI: S0304-3940(09)01446-3 [pii] 10.1016/j.neulet.2009.10.089. [PubMed: 19879925]
7. Corbo RM, Gambina G, Ulizzi L, Moretto G, Scacchi R. Genetic variation of *CYP19* (aromatase) gene influences age at onset of Alzheimer's disease in women. Dement Geriatr Cogn Disord. 2009; 27:513–518. DOI: 000221832 [pii] 10.1159/000221832. [PubMed: 19478482]
8. Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE. Systematic meta-analyses of Alzheimer disease genetic association studies: The Alzgene database. Nat Genet. 2007; 39:17–23. DOI: ng1934 [pii] 10.1038/ng1934. [PubMed: 17192785]
9. Tang MX, Cross P, Andrews H, Jacobs DM, Small S, Bell K, Merchant C, Lantigua R, Costa R, Stern Y, Mayeux R. Incidence of AD in African-Americans, Caribbean Hispanics, and Caucasians in northern Manhattan. Neurology. 2001; 56:49–56. [PubMed: 11148235]
10. Stern Y, Andrews H, Pittman J, Sano M, Tatemichi T, Lantigua R, Mayeux R. Diagnosis of dementia in a heterogeneous population. Development of a neuropsychological paradigm-based diagnosis of dementia and quantified correction for the effects of education. Arch Neurol. 1992; 49:453–460. [PubMed: 1580806]
11. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000; 155:945–959. [PubMed: 10835412]
12. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. Genetics. 2003; 164:1567–1587. [PubMed: 12930761]
13. Hixson JE, Vernier DT. Restriction isotyping of human apolipoprotein e by gene amplification and cleavage with Hhai. J Lipid Res. 1990; 31:545–548. [PubMed: 2341813]
14. Goodman Y, Bruce AJ, Cheng B, Mattson MP. Estrogens attenuate and corticosterone exacerbates excitotoxicity, oxidative injury, and amyloid beta-peptide toxicity in hippocampal neurons. J Neurochem. 1996; 66:1836–1844. [PubMed: 8780008]
15. Toran-Allerand CD, Miranda RC, Bentham WD, Sohrabji F, Brown TJ, Hochberg RB, MacLusky NJ. Estrogen receptors colocalize with low-affinity nerve growth factor receptors in cholinergic

neurons of the basal forebrain. Proc Natl Acad Sci U S A. 1992; 89:4668–4672. [PubMed: 1316615]

16. Luine VN. Estradiol increases choline acetyltransferase activity in specific basal forebrain nuclei and projection areas of female rats. Exp Neurol. 1985; 89:484–490. DOI: 0014-4886(85)90108-6 [pii]. [PubMed: 2990988]
17. Behl C, Widmann M, Trapp T, Holsboer F. 17-beta estradiol protects neurons from oxidative stress-induced cell death in vitro. Biochem Biophys Res Commun. 1995; 216:473–482. DOI: S0006-291X(85)72647-2 [pii] 10.1006/bbrc.1995.2647. [PubMed: 7488136]
18. Jaffe AB, Toran-Allerand CD, Greengard P, Gandy SE. Estrogen regulates metabolism of Alzheimer amyloid beta precursor protein. J Biol Chem. 1994; 269:13065–13068. [PubMed: 8175728]
19. Gerhard M, Ganz P. How do we explain the clinical benefits of estrogen? From bedside to bench. Circulation. 1995; 92:5–8. [PubMed: 7788916]
20. Simoncini T, Mannella P, Genazzani AR. Rapid estrogen actions in the cardiovascular system. Ann N Y Acad Sci. 2006; 1089:424–430. DOI: 1089/1/424 [pii] 10.1196/annals.1386.001. [PubMed: 17261785]
21. Ishunina TA, Fischer DF, Swaab DF. Estrogen receptor alpha and its splice variants in the hippocampus in aging and Alzheimer's disease. Neurobiol Aging. 2007; 28:1670–1681. DOI: S0197-4580(06)00309-5 [pii] 10.1016/j.neurobiolaging.2006.07.024. [PubMed: 17010478]
22. Ishunina TA, van Beurden D, van der Meulen G, Unmehopa UA, Hol EM, Huitinga I, Swaab DF. Diminished aromatase immunoreactivity in the hypothalamus, but not in the basal forebrain nuclei in Alzheimer's disease. Neurobiol Aging. 2005; 26:173–194. DOI: S0197458004001575 [pii] 10.1016/j.neurobiolaging.2004.03.010. [PubMed: 15582747]
23. Stoffel-Wagner B, Watzka M, Steckelbroeck S, Schwaab R, Schramm J, Bidlingmaier F, Klingmuller D. Expression of *CYP19* (aromatase) mrna in the human temporal lobe. Biochem Biophys Res Commun. 1998; 244:768–771. DOI: S0006-291X(98)98337-1 [pii] 10.1006/bbrc.1998.8337. [PubMed: 9535740]
24. Yague JG, Munoz A, de Monasterio-Schrader P, Defelipe J, Garcia-Segura LM, Azcoitia I. Aromatase expression in the human temporal cortex. Neuroscience. 2006; 138:389–401. DOI: S0306-4522(05)01321-7 [pii] 10.1016/j.neuroscience.2005.11.054. [PubMed: 16426763]
25. Tworoger SS, Chubak J, Aiello EJ, Ulrich CM, Atkinson C, Potter JD, Yasui Y, Stapleton PL, Lampe JW, Farin FM, Stanczyk FZ, McTiernan A. Association of *CYP17*, *CYP19*, *CYP1b1*, and *COMT* polymorphisms with serum and urinary sex hormone concentrations in postmenopausal women. Cancer Epidemiol Biomarkers Prev. 2004; 13:94–101. [PubMed: 14744739]
26. Dunning AM, Dowsett M, Healey CS, Tee L, Luben RN, Folkerd E, Novik KL, Kelemen L, Ogata S, Pharoah PD, Easton DF, Day NE, Ponder BA. Polymorphisms associated with circulating sex hormone levels in postmenopausal women. J Natl Cancer Inst. 2004; 96:936–945. [PubMed: 15199113]
27. De Castro F, Moron FJ, Montoro L, Galan JJ, Real LM, Ruiz A. Re: Polymorphisms associated with circulating sex hormone levels in postmenopausal women. J Natl Cancer Inst. 2005; 97:152–153. author reply 153–154. DOI: 97/2/152-a [pii] 10.1093/jnci/dji029. [PubMed: 15657347]
28. Baghaei F, Rosmond R, Westberg L, Hellstrand M, Eriksson E, Holm G, Bjorntorp P. The *CYP19* gene and associations with androgens and abdominal obesity in premenopausal women. Obes Res. 2003; 11:578–585.10.1038/oby.2003.81 [PubMed: 12690088]
29. Combarros O, Riancho JA, Infante J, Sanudo C, Llorca J, Zarrabeitia MT, Berciano J. Interaction between *CYP19* aromatase and butyrylcholinesterase genes increases Alzheimer's disease risk. Dement Geriatr Cogn Disord. 2005; 20:153–157. DOI: 87065 [pii] 10.1159/000087065. [PubMed: 16020944]
30. Combarros O, Sanchez-Juan P, Riancho JA, Mateo I, Rodriguez-Rodriguez E, Infante J, Garcia-Gorostiaga I, Vazquez-Higuera JL, Berciano J. Aromatase and interleukin-10 genetic variants interactively modulate alzheimer's disease risk. J Neural Transm. 2008; 115:863–867.10.1007/s00702-008-0028-5 [PubMed: 18299793]

31. 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 Neurol. 2010; 10:90. DOI: 1471-2377-10-90 [pii] 10.1186/1471-2377-10-90. [PubMed: 20932310]

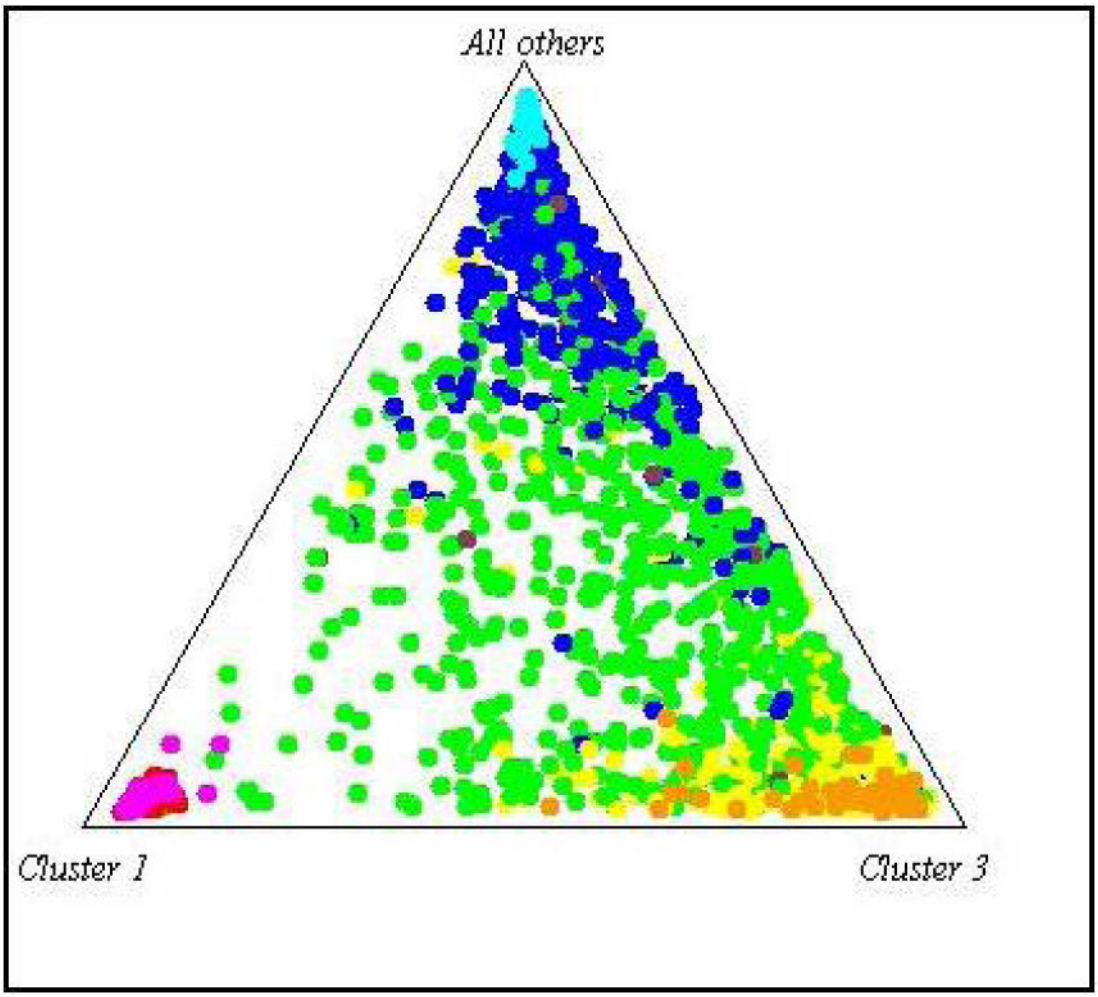

Figure 1. Plot of WHICAP participants by AIMs-defined ancestry versus HapMap populations
WHICAP participants
Yellow: Predominantly Caucasian AIMs-defined ancestry
Green: Admixed/Hispanic AIMs-defined ancestry
Blue: Predominantly African AIMs-defined ancestry
HapMap populations
Light Brown: Ancestrally homogenous Caucasian population (CEPH)
Light Blue: Ancestrally homogenous Yoruban Black population (YRI)
Red: Ancestrally homogenous East Asian population (CHJA)

Table 1

Population Characteristics

Characteristic	Non Demented	Alzheimer's Disease
Sample Size	1175	511
Age at time of enrollment (mean ± S.D.)*	75.7 (6.0)	79.9 (7.2)
Self-identified Ethnicity (n, % of ethnicity)*		
White	358 (84.6)	65 (15.3)
Hispanic	415 (61.8)	257 (38.2)
Black	389 (67.8)	185 (32.2)
Other	13 (76.5)	4 (23.5)
Body Mass Index (mean ± S.D.)	28.2 (7.9)	27.1 (5.8)
Years of education (mean ± S.D.)*	10.4 (4.6)	7.1 (4.7)
Diabetes (n, %)	180 (15.3)	104 (20.3)
Smoking (n, %)	97 (8.3)	44 (8.7)
At least one copy APOE ε4 (n, %)*	284 (24.2)	156 (30.5)

* $p < .05$

Table 2

Covariate characteristics by self-identified ethnicity and AIMs-defined ancestry

	Self-identified ethnicity			AIMs-defined ancestry		
	White	**Hispanic**	**Black**	**Predominantly Caucasian**	**Admixed/Hispanic**	**Predominantly African**
No. of individuals	423	672	574	632	473	581
At least one copy *APOE e4* (n, % of individuals within ethnicity/ AIMs-defined ancestry)*	93 (22.0)	167 (24.9)	190 (33.1)	148 (23.4)	117 (24.7)	189 (32.5)
Diabetes mellitus (n, % of individuals within ethnicity/ AIMs-defined ancestry)*	31 (7.3)	136 (20.2)	105 (18.3)	64 (10.1)	104 (22.0)	109 (18.8)
Smoking (n, % of individuals within ethnicity/AIMs-defined ancestry)*	22 (5.2)	39 (5.8)	71 (12.4)	27 (4.3)	41 (8.7)	65 (11.2)

* $p \leq 0.001$

Table 3

Odds ratios for AD by *CYP19* SNPs, stratified by AIMs – defined ancestry

					Predominantly Caucasian					Admixed/Hispanic					Predominantly African				
SNP #	**SNP**	**Position**	**Distance (bp)**	**Minor Allele**	**MAF**	***O.R.***	***95% C.I.***	***p-value***	***emp. p-value***	**MAF**	***O.R.***	***95% C.I.***	***p-value***	***emp. p-value***	**MAF**	**O.R.**	***95% C.I.***	***p-value***	***emp. p-value***
1	rs4646	49290136		A	0.3	1.2	0.9 – 1.8	0.526	0.525	0.3	0.6	0.5 – 1.0	0.073	0.080	0.3	1.1	0.8 – 1.8	0.301	0.301
2	rs10046	49290278	142	G	0.4	1.0	0.7 – 1.6	0.692	0.696	0.6	0.5	0.4 – 1.1	0.108	0.115	0.8	1.2	0.4 – 2.4	0.905	0.855
3	rs700519	49295260	4982	A	0.1	0.9	0.5 – 1.9	0.770	0.776	0.1	1.1	0.6 – 1.7	0.904	0.903	0.2	0.9	0.6 – 1.3	0.628	0.646
4	rs1143704	49297994	2734	T	0.4	0.8	0.7 – 1.3	0.599	0.593	0.4	1.0	0.6 – 1.5	0.855	0.087	0.2	0.7	0.5 – 1.1	0.152	0.913
5	rs6493488	49301214	3220	C	0.4	0.9	0.5 – 1.1	0.177	0.177	0.6	0.9	0.5 – 1.4	0.479	0.478	0.7	0.8	0.4 – 1.8	0.440	0.676
6	rs16964201	49302643	1429	A	0.1	0.9	0.5 – 1.8	0.799	0.807	0.2	1.2	0.7– 1.7	0.785	0.790	0.3	1.1	0.7 – 1.4	0.849	0.826
7	rs2899472	49303347	704	A	0.3	1.1	0.7 – 1.4	0.794	0.780	0.2	1.0	0.6 – 1.4	0.623	0.633	0.1	0.6	0.4 – 1.3	0.269	0.243
8	rs8023263	49304889	1542	C	0.5	0.9	0.6 – 1.5	0.776	0.824	0.4	0.8	0.4 – 1.2	0.225	0.229	0.3	1.0	0.5 – 2.2	0.940	0.945
9	rs4775934	49306061	1172	G	0.1	1.0	0.6 – 1.9	0.894	0.897	0.2	0.9	0.6 – 1.4	0.776	0.802	0.4	0.8	0.6 – 1.3	0.551	0.534
10	rs4775935	49306568	507	A	**0.3**	**0.7**	**0.4 – 0.9**	**0.036**	**0.041**	0.3	**0.5**	**0.4 – 0.9**	**0.032**	**0.034**	0.3	1.2	0.7 – 1.6	0.681	0.695
11	rs1065778	49307498	930	G	0.5	1.3	0.7 – 1.7	0.627	0.624	0.3	1.1	0.7 – 1.5	0.880	0/873	0.2	0.7	0.5 – 1.1	0.147	0.144
12	rs12594287	49311199	3701	A	0.1	1.4	0.9 – 2.5	0.085	0.082	0.2	0.8	0.6 – 1.4	0.644	0.626	0.2	0.9	0.6 – 1.3	0.643	0.645
13	rs700518	49316404	5205	G	0.5	1.2	0.7 – 1.8	0.572	0.589	0.3	1.0	0.7 – 1.5	0.880	0.869	0.2	0.7	0.5 – 1.1	0.200	0.198
14	rs2414096	49317071	667	A	0.5	1.1	0.7 – 1.7	0.737	0.742	0.3	1.2	0.7 – 1.6	0.713	0.720	0.2	0.9	0.5 – 1.1	0.149	0.151
15	rs727479	49321589	4518	C	**0.3**	**0.7**	**0.4 – 0.9**	**0.019**	**0.014**	0.3	**0.7**	**0.4 – 0.9**	**0.047**	**0.066**	0.2	1.3	0.8 – 1.8	0.284	0.298
16	rs10459592	49323433	1844	C	0.4	0.7	0.5 – 1.0	0.080	0.081	0.5	0.8	0.5 – 1.5	0.712	0.724	0.6	1.0	0.6 – 1.9	0.915	0.903
17	rs767199	49327679	4246	A	0.5	1.1	0.7 – 1.6	0.795	0.800	0.3	1.2	0.7 – 1.7	0.627	0.615	0.1	0.9	0.5 – 1.3	0.373	0.384
18	rs12911554	49330049	2370	G	0.4	0.9	0.5 – 1.2	0.202	0.203	0.5	0.8	0.4 – 1.1	0.093	0.109	0.5	1.0	0.7 – 1.6	0.919	0.902
19	rs7172156	49333590	3541	A	0.4	0.8	0.5 – 1.1	0.165	0.159	0.4	1.1	0.7 – 1.5	0.913	0.907	0.3	0.8	0.6 – 1.3	0.556	0.559
20	rs2008691	49335352	1762	G	0.2	1.3	0.8 – 1.8	0.402	0.414	0.3	0.9	0.5 – 1.2	0.212	0.206	0.4	1.2	0.7 – 1.4	0.841	0.837
21	rs11856927	49335997	645	C	0.4	1.4	0.8 – 1.9	0.280	0.297	0.4	1.0	0.6 – 1.5	0.840	0.839	0.5	0.6	0.4 – 1.0	0.053	0.059
22	rs1008805	49336891	894	G	0.4	0.8	0.5 – 1.2	0.236	0.203	0.3	1.2	0.7 – 1.7	0.626	0.637	0.2	1.0	0.7 – 1.5	0.829	0.801
23	rs6493494	49337127	236	A	0.4	1.4	0.9 – 2.1	0.164	0.174	0.3	1.2	0.8 – 1.9	0.246	0.263	0.3	1.0	0.7 – 1.4	0.771	0.767
24	rs10519299	49338638	1511	C	0.5	1.6	0.9 – 2.3	0.092	0.092	0.4	1.3	0.8 – 1.9	0.333	0.352	0.3	0.9	0.6 – 1.3	0.693	0.693
25	rs12050767	49344549	5911	G	0.5	1.5	0.9 – 2.3	0.096	0.109	0.4	1.1	0.7 – 1.9	0.497	0.505	0.4	0.8	0.6 – 1.3	0.472	0.472

					Predominantly Caucasian					Admixed/Hispanic					Predominantly African				
SNP #	**SNP**	**Position**	**Distance (bp)**	**Minor Allele**	**MAF**	***O.R.***	***95% C.I.***	***p-value***	***emp. p-value***	**MAF**	***O.R.***	***95% C.I.***	***p-value***	***emp. p-value***	**MAF**	**O.R.**	***95% C.I.***	***p-value***	***emp. p-value***
26	rs749292	49346023	1474	A	0.5	1.4	0.9 – 2.3	0.093	0.089	0.4	1.0	0.7 – 1.7	0.651	0.680	0.5	0.9	0.6 – 1.4	0.592	0.620
27	rs17647719	49355496	9473	G	**0.1**	**1.8**	**1.1 – 2.8**	**0.037**	**0.038**	0.2	1.0	0.7 – 1.5	0.904	0.896	0.3	1.0	0.7 – 1.5	0.854	0.869
28	rs1902586	49358145	2649	A	**0.1**	**1.7**	**1.1 – 2.8**	**0.033**	**0.026**	0.3	1.3	0.8 – 1.8	0.337	0.313	0.5	0.8	0.6 – 1.4	0.659	0.647
29	rs16953058	49359552	1407	A	0.1	1.6	0.7 – 4.0	0.230	0.249	0.1	0.8	0.5 – 1.4	0.544	0.531	0.1	0.9	0.6 – 1.4	0.723	0.734
30	rs936306	49366890	7338	A	0.2	1.4	0.9 – 2.0	0.166	0.168	0.4	0.9	0.6 – 1.4	0.725	0.723	0.5	0.7	0.5 – 1.2	0.295	0.310
31	rs12594203	49370853	3963	T	0.1	1.5	0.9 – 2.4	0.158	0.175	0.2	1.2	0.8 – 1.9	0.263	0.242	0.4	0.8	0.6 – 1.2	0.355	0.356
32	rs2470176	49371231	378	G	0.2	1.4	0.9 – 2.0	0.159	0.157	0.6	1.1	0.7 – 1.5	0.936	0.934	0.6	0.7	0.5 – 1.3	0.305	0.322
33	rs2470152	49382264	11033	A	0.5	0.7	0.5 – 1.2	0.253	0.230	0.5	0.6	0.4 – 1.1	0.127	0.128	0.5	0.9	0.6 – 1.5	0.800	0.832
34	rs3751592	49393870	11606	G	0.3	1.3	0.9 – 2.0	0.098	0.095	0.4	1.2	0.9 – 1.9	0.232	0.229	0.3	1.4	0.9 – 1.8	0.207	0.210
35	rs1004984	49400821	6951	A	0.4	1.4	0.9 – 2.1	0.113	0.099	0.5	1.0	0.7 – 1.6	0.873	0.880	0.5	0.8	0.5 – 1.1	0.192	0.193
36	rs1902585	49401198	377	G	0.5	0.9	0.5 – 1.1	0.184	0.181	0.3	1.1	0.7 – 1.6	0.842	0.850	0.1	1.1	0.8 – 1.8	0.419	0.424
37	rs10163138	49401416	218	A	**0.1**	**2.6**	**1.1 – 6.3**	**0.044**	**0.026**	0.1	0.8	0.5 – 1.4	0.568	0.554	0.3	1.1	0.8 – 1.6	0.587	0.606
38	rs6493495	49405594	4178	G	0.1	1.3	0.8 – 1.9	0.317	0.305	0.2	1.1	0.7 – 1.6	0.781	0.796	0.3	**0.7**	**0.4 – 0.9**	**0.042**	**0.053**
39	rs7168331	49408275	2681	G	**0.4**	**1.5**	**1.1 – 2.5**	**0.028**	**0.025**	0.5	1.0	0.6 – 1.6	0.913	0.912	0.6	0.8	0.6 – 1.4	0.646	0.615
40	rs1870049	49412515	4240	G	0.2	1.2	1.0 – 2.2	0.083	0.072	0.4	1.2	0.7 – 1.6	0.736	0.706	0.4	0.9	0.6 – 1.3	0.510	0.545
41	rs11070843	49413843	1328	G	0.1	1.3	0.8 – 2.0	0.251	0.233	0.2	1.1	0.7 – 1.7	0.632	0.629	0.2	**0.6**	**0.4 – 0.9**	**0.027**	**0.038**

Notes:

- ORs were computed using a dominant model in which individuals with at least 1 copy of the minor allele were coded as having a risk allele
- Analyses are adjusted for age, BMI, years of education, and presence of at least one copy of an APOE ε4 allele
- MAF: Minor allele frequency
- Emp. p-value: To minimize the risk of false-positive findings from multiple testing, we computed empirical p-values by generating the null distribution on the basis of 1000 replicate datasets.