

RESEARCH

Open Access



Admixture mapping reveals the association between Native American ancestry at 3q13.11 and reduced risk of Alzheimer's disease in Caribbean Hispanics

Andréa R. V. R. Horimoto¹ , Diane Xue² , Timothy A. Thornton^{1,2}  and Elizabeth E. Blue^{2,3*} 

Abstract

Background: Genetic studies have primarily been conducted in European ancestry populations, identifying dozens of loci associated with late-onset Alzheimer's disease (AD). However, much of AD's heritability remains unexplained; as the prevalence of AD varies across populations, the genetic architecture of the disease may also vary by population with the presence of novel variants or loci.

Methods: We conducted genome-wide analyses of AD in a sample of 2565 Caribbean Hispanics to better understand the genetic contribution to AD in this population. Statistical analysis included both admixture mapping and association testing. Evidence for differential gene expression within regions of interest was collected from independent transcriptomic studies comparing AD cases and controls in samples with primarily European ancestry.

Results: Our genome-wide association study of AD identified no loci reaching genome-wide significance. However, a genome-wide admixture mapping analysis that tests for association between a haplotype's ancestral origin and AD status detected a genome-wide significant association with chromosome 3q13.11 (103.7–107.7Mb, $P = 8.76E-07$), driven by a protective effect conferred by the Native American ancestry (OR = 0.58, 95%CI = 0.47–0.73). Within this region, two variants were significantly associated with AD after accounting for the number of independent tests (rs12494162, $P = 2.33E-06$; rs1731642, $P = 6.36E-05$). The significant admixture mapping signal is composed of 15 haplotype blocks spanning 5 protein-coding genes (*ALCAM*, *BBX*, *CBLB*, *CCDC54*, *CD47*) and four brain-derived topologically associated domains, and includes markers significantly associated with the expression of *ALCAM*, *BBX*, *CBLB*, and *CD47* in the brain. *ALCAM* and *BBX* were also significantly differentially expressed in the brain between AD cases and controls with European ancestry.

Conclusion: These results provide multiethnic evidence for a relationship between AD and multiple genes at 3q13.11 and illustrate the utility of leveraging genetic ancestry diversity via admixture mapping for new insights into AD.

Keywords: Admixture mapping, Alzheimer's disease, Local ancestry, Genetic architecture, Genome scan, Multi-omics

* Correspondence: em27@uw.edu

²Institute for Public Health Genetics, University of Washington, Seattle, WA, USA

³Division of Medical Genetics, University of Washington, BOX 357720, Seattle, WA 98195-7720, USA

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

Late-onset Alzheimer's disease (AD) is a leading cause of death in the USA, affecting approximately 1 in 10 Americans over the age of 65 years, with prevalence expected to double by 2050 [1]. Heritability estimates for AD range from 58 to 79% [2]; despite this strong genetic component, much of the underlying genetic variance remains to be explained [3]. Although the *APOE* $\epsilon 4$ allele is the strongest common genetic risk factor for AD [4, 5], dozens of loci have been associated with AD via genome-wide association studies (GWAS) [6, 7]. However, similar to other complex diseases, the vast majority of genetic discoveries for AD have been from GWAS performed in samples with predominantly European ancestry [8]. Much less is known about AD genetics across diverse populations, and in particular African Americans and Hispanic/Latino Americans who have increased risk of AD compared to European Americans [9–11].

Admixture mapping is a powerful alternative approach to GWAS for gene-mapping in recently admixed populations. Unlike widely used GWAS approaches that treat genetic ancestry differences as potential confounders in the analysis, admixture mapping leverages genetic ancestry differences [12–15]. With admixture mapping, regions of the genome with unusual local ancestry patterns relative to genome-wide averages are tested for association with a phenotype [16]. Admixture mapping is most powerful when both disease risk and trait-locus allele frequencies differ across groups, and it can be viewed as a complement to GWAS [17, 18].

Here, we performed genome scans of AD using both GWAS and admixture mapping approaches to identify regions associated with AD in Caribbean Hispanics, an admixed population with European, Native American, and African ancestry [19, 20]. Admixture mapping identified a genome-wide significant association between AD and Native American ancestry on 3q13.11, while GWAS identified no loci reaching genome-wide significance. Transcriptomic studies in samples with European ancestry nominate *ALCAM* and *BBX* as candidate protein-coding genes within the significant admixture mapping signal on 3q13.11, supported by the association between genetic variation and gene expression levels as well as differential expression between AD cases and controls. These results underline the power and challenges of leveraging genetic ancestry differences for new insight into the genetic architecture of late-onset AD in multiethnic populations.

Methods

Data

Genotype and phenotype data for 3067 participants in the Columbia University Study of Caribbean Hispanics

and Late Onset Alzheimer's Disease (CU Hispanics) were downloaded through dbGaP (Study Accession number phs000496.v1.p1), described in detail elsewhere [21]. The CU Hispanics study recruited subjects using both familial AD (22%) and sporadic case-control (78%) ascertainment. Subjects were excluded if they had any missing data for sex, AD status, *APOE* $\epsilon 2/\epsilon 3/\epsilon 4$ genotypes, and either age-at-onset of dementia or age-at-last-evaluation.

European (from Utah) and African (Yorubans) samples from HapMap 3 [22] and Native American samples (Colombians, Pima, and Maya) from the Human Genome Diversity Project [23, 24] were used as reference populations. The reference datasets were merged using PLINK (v1.07) [25], resulting in 598,470 common autosomal single-nucleotide polymorphisms (SNPs). Genome coordinates were updated to build NCBI37/hg19 using LiftOver [26] to match the CU Hispanic data. The reference and CU Hispanics datasets were merged, randomly removing reference samples to balance ancestral population representation. Variants with a genotype missing rate > 5%, samples missing > 5% of genotypes, and 502 duplicated samples were excluded. Heterozygosity analysis identified 43 CU Hispanic outliers for both the F coefficient (> 0.12 , mean 0.02 ± 0.03) and heterozygosity rate (< 0.28 , mean 0.32 ± 0.01), consistent with previous reports and pedigree documentation of consanguinity [27]. As both our admixture mapping and association tests adjust for genetic relatedness, keeping these samples had minimal impact on results (Additional file 1). The final combined dataset included 294,252 SNPs and 2754 samples: 2565 CU Hispanics plus 63 samples from each reference population. The overall genotyping rate was 0.993.

Genetic relatedness matrix

A genetic relatedness matrix (GRM) was estimated in a recursive manner using the PC-AiR and PC-Relate functions within the GENESIS R package [28–30]. The final combined data set was included in these analyses to improve inference of population structure. PC-AiR partitions subjects into unrelated and related sets based on kinship estimates from KING-robust [31], performs principal components (PC) analysis on the set of unrelated subjects, then projects PC values for the related set. PC-Relate adjusted the GRM for the first four PCs derived by PC-AiR, and the PC-AiR and PC-Relate steps were repeated using this adjusted GRM. The final GRM contains kinship coefficients that are robust to the population structure within our sample.

Ancestry proportions

As suggested by an established pipeline [32], the CU Hispanic and reference samples were phased jointly

using ShapeIt2 [33] and 1000 Genomes phase 3 haplotypes [34] as reference. Local ancestry estimation was performed using RFMix (v1.5.4) [35]. Local ancestry values were averaged to estimate global European, African, and Native American ancestry proportions.

Admixture mapping

Admixture mapping was performed using a logistic mixed model for the AD phenotype, in which all European, African, and Native American ancestries were tested simultaneously. Admixture mapping was conducted using the GENESIS R package [30] available in Bioconductor [36]. We fit mixed models under the null hypothesis of no genetic association, adjusting for global ancestry proportions and *APOE* ϵ 2 and ϵ 4 allele dosages as fixed effects and the GRM as a random effect. The association between each admixture linkage disequilibrium (LD) block and the null model was evaluated by a score test. Recent admixture, as such observed in Hispanic/Latinos, creates long-range LD which dramatically reduces the number of independent tests in an admixture mapping genome scan, leading to a less-severe multiple testing correction. Genome-wide significance was defined as $P < 5E-05$ and suggestive evidence for significance was defined as $P < 0.001$, as suggested by previous studies of Hispanic populations [37, 38]. We evaluated the suitability of these significance thresholds by extending the method proposed by Shriner et al. 2011 [39] for three ancestral populations. We estimated the effective number of tests for each ancestral population by fitting autoregressive models to the vectors of African, European, and Native American local ancestry dosages in our sample (African: 251.1, European: 210.3, Native American: 281.2) and defined the final effective number of tests as the sum of the two largest values. This Bonferroni-corrected significance threshold of $P < 9.39E-05$ is slightly less conservative than our original threshold, suggesting it is well-suited for this sample.

Secondary admixture mapping analyses considered the effect of each reference group separately to identify which ancestral population was driving the significant signals. The coefficients of each lead SNP in the most significant LD-block were estimated, taking the allelic dosage of the ancestry driving the signal into account. Manhattan plots were prepared using the qqman R package [40], while regional association plots were generated using LocusZoom [41]. Additional sensitivity analyses assessed the robustness of our findings to age and sex covariate adjustment.

Association testing

SNPs and samples were submitted to the data cleaning procedures described above without the inclusion of reference samples, leaving 931,670 SNPs and 2565 CU

Hispanic samples. We conducted the association testing for AD using a logistic mixed model implemented in the GENESIS R package [30]. Using the fitted null model described above, we tested the association between each SNP and the phenotype with a score test. Genome-wide significance was defined as $P < 5E-08$. Region-specific thresholds within the 3q13.11 locus for significant ($P < 6.74E-05$) and suggestive ($P < 1.35E-05$) evidence for association were adjusted for the effective number of tests, estimated by Genetic Type I error calculator [42].

Locus interpretation and gene prioritization

Conditional admixture mapping analyses were performed, applying the original model with further adjustment for allele dosage at SNPs of interest, individually and jointly. LD was estimated by both r^2 and D' using PLINK [25] in a set of 1349 unrelated CU Hispanics. LD plots based on the correlation statistic D' by reference population were prepared using Haploview [43]. The Ensembl Variant Effect Predictor (v99 [44]) toolset generated SNP-level annotations within regions of interest.

The Accelerating Medical Partnerships for Alzheimer's Disease (AMP-AD) project has provided a publicly available repository of multi-omic data aimed at finding genetic targets for AD therapeutics. We extracted significant *cis* expression quantitative trait loci (*cis*-eQTLs) from a recent AMP-AD study [45] (<https://www.synapse.org/#!Synapse:syn17015233>), representing RNA-sequencing data generated on brain samples from the Mayo study, Religious Orders Study, Rush Memory and Aging studies, and Mount Sinai Brain Bank study generated across 7 tissues types: cerebellum ($N = 261$), temporal cortex ($N = 262$), dorsolateral prefrontal cortex ($N = 573$), inferior gyrus ($N = 230$), superior temporal gyrus ($N = 225$), frontal pole ($N = 260$), and parahippocampal gyrus ($N = 225$). We extracted evidence for differential gene expression in post-mortem brain tissues between those affected by AD and controls from another AMP-AD study [46] (<https://www.synapse.org/#!Synapse:syn11914606>), restricted to the meta-analysis results from the random effects model. A false discovery rate (FDR) cutoff of < 0.05 provided by the AMP-AD studies was applied to both the differential gene expression and eQTL results.

The genome is organized into topologically associated domains (TADs) in three-dimensional space, where genes within the same TAD are more likely to be regulated by common *cis*-regulatory elements and transcription factors. Genes within the same TAD as the haplotypes associated with AD were extracted from the 3D Genome Browser [47] and the human dorsolateral prefrontal cortex data (DLPFC) [48], again using the study-specific FDR < 0.05 as the significance threshold.

Genetic variation and patterns of LD vary across populations, and ideally colocalization analyses should use association and eQTL results representing the same population; unfortunately, large eQTL studies of Caribbean Hispanic populations are unavailable. Colocalization analyses comparing our admixture mapping or association studies are restricted to comparisons with the AMP-AD eQTLs representing samples with primarily European ancestry, which may identify relationships between eQTLs and AD risk shared between these populations [49, 50]. Approximate Bayes factor colocalization was performed using the Coloc package in R (v3.2-1) [50], which computes five posterior probabilities: PP0 = no association with either trait; PP1 = association with trait 1 but not trait 2; PP2 = association with trait 2 but not trait 1; PP3 = association with both traits, two independent causal SNPs; and PP4 = association with trait 1 and trait 2, one causal SNP shared for both traits. The LocusCompareR package in R (v1.0.0) [51] illustrated the correlation between admixture mapping or association results and eQTL data.

Results

Sample description

The CU Hispanics data represented 2565 subjects, where the 1174 cases were affected either by familial AD (22%) or sporadic AD (78%). Age-at-onset ranged from 44 to 100 years while the censoring age among the unaffected controls ranged from 35 to 100 years. The mean age and sex are similar across cases and controls (Table 1). The frequency of the protective *APOE* ϵ 2 allele [52] is approximately 35% lower among cases, while the well-established risk allele ϵ 4 [4, 5] is almost twice as common among cases than controls (Table 1). Global average ancestry proportion estimates vary widely across samples, from nearly zero to 0.99 per reference population (Fig. 1). Average ancestry proportions are 0.58 ± 0.17 European, 0.33 ± 0.19 African, and 0.09 ± 0.08 Native American ancestry.

Admixture mapping and GWAS

We identified a genome-wide significant association between AD and local ancestry at 3q13.11 ($P < 5E-05$; Table 2, Fig. 2). The 3q13.11 signal is supported by

Table 1 Sample description

Sample	N	Females (%)	Age (years)	ϵ 2 (%)	ϵ 4 (%)
Affected	1174	64.1	74.9 (9.4)	4.4	28.0
Unaffected	1391	67.7	72.5 (8.5)	6.7	14.5
Total	2565	66.1	73.6 (9.0)	5.6	20.7

Definitions: N = sample size, Age = mean \pm standard deviation of age-at-onset of dementia (affected) or age-at-last-evaluation (unaffected), ϵ 2 = frequency of the *APOE* ϵ 2 allele, ϵ 4 = frequency of the *APOE* ϵ 4 allele

significant evidence of association across multiple LD-blocks (103.7 to 107.7Mb, min. $P = 8.76E-07$; Table 2), where the lead SNP rs10933849 is a common variant across the ancestral populations (alternate frequency: 0.56, 0.84, and 0.61 for 1000 Genomes phase 3 Africans, Europeans, and Native Americans, respectively). This region spans five protein-coding genes: *ALCAM*, *CBLB*, *BBX*, *CCDC54*, and *CD47*. Secondary analyses indicated that Native American ancestry at the lead SNP of each LD-block was associated with a protective effect against AD risk (OR 0.58–0.66; $P < 3.24E-04$; Additional file 2). Greater correlation is observed between 15 SNPs tagging the LD blocks within the 3q13.11 locus in the Native American reference data than in the European or African data (Additional file 3), providing further evidence that the admixture mapping signal between AD and 3q13.11 is driven by a Native American haplotype.

Suggestive evidence of association between local ancestry and AD was observed at six additional loci: 2q22.2, 6q22.31, 8q24.22, 9p21.3, 14q12, and 19p13.3 ($P < 0.001$; Fig. 2, Table 2). LD-block-specific results for significant and suggestive associations with AD are provided in Additional file 2. Two LD-blocks with European background were responsible for the suggestive signal at 2q22.2, intersecting the gene *LRP1B*. The suggestive signal on 6q22.31 is driven by the Native American ancestry and was captured by a single LD-block within the *TRDN* gene. On 8q24.22, we observed three LD-blocks with Native American background driving the signal which spans the *ZFAT* gene. Two LD-blocks spanning the *DMRTA1* gene were responsible for the signal on 9p21.3, driven by the Native American ancestry. The signal on 14q12, driven by the African ancestry, was captured by five LD-blocks implicating *ARHGAP5* and *AKAP6*. Nine LD-blocks within a 1.3Mb region were responsible for the signal on 19p13.3 driven by African ancestry, implicating *ABCA7* and dozens of other genes (Table 2). Sensitivity analyses revealed the admixture mapping results are robust to the inclusion of age and sex as covariates (Additional file 4). In contrast, traditional GWAS for AD did not identify any loci reaching genome-wide significance ($P < 5E-08$; Additional file 5).

Locus interpretation and gene prioritization

Targeted association testing within the 3q13.11 locus found two SNPs significantly associated with AD (rs12494162, $P = 2.33E-06$; rs1731642, $P = 6.36E-05$), and 22 SNPs with suggestive evidence of association with AD ($P < 1.35E-03$; Table 3). The first SNP, rs12494162, falls within an intron of lncRNA *DUBR*, while rs1731642 is an intergenic variant. These two SNPs, rs12494162 and rs1731642, are not in LD within our data ($r^2 = 0.003$; $D' = 0.17$) and may represent independent association signals. This is consistent with

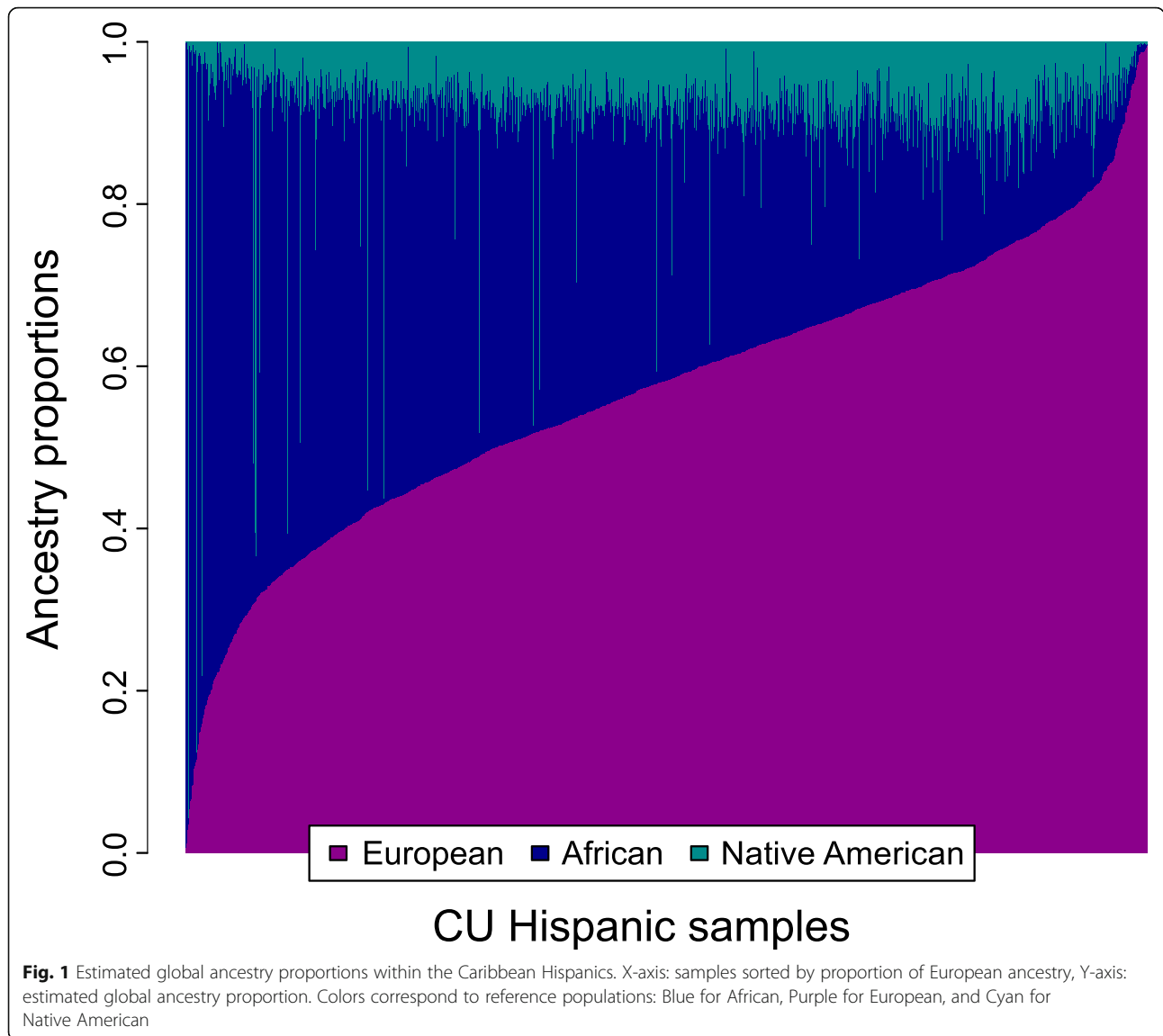


Table 2 Evidence of association between local ancestry and Alzheimer’s disease in the Caribbean Hispanics

Chr	Position	rsID	Gene	Consequence	Ancestry	OR (95% CI)	P
2q22.2	142,486,253–143,387,612	rs13024316	<i>LRP1B</i>	Intron	EUR	1.28 (1.13–1.45)	6.82E–04
3q13.11	103,747,624–107,725,831	rs10933849	Intergenic	Intergenic	NAM	0.58 (0.47–0.73)	8.76E–07
6q22.31	123,548,997–123,838,033	rs6940177	<i>TRDN</i>	Intron	NAM	1.44 (1.19–1.75)	9.54E–04
8q24.22	135,308,849–135,856,404	rs4308771	RP11-513H8.1	Intron	NAM	1.36 (1.12–1.65)	5.41E–04
9p21.3	22,207,037–22,870,294	rs4977586	Intergenic	Intergenic	NAM	0.70 (0.56–0.87)	4.56E–04
14q12	32,485,703–33,033,695	rs1952961	RP11-187E13.2	Intron	AFR	0.81 (0.71–0.93)	7.24E–04
19p13.3	266,034–1,505,874	rs3787017	<i>PALM</i>	Intron	AFR	1.29 (1.14–1.47)	4.26E–04

Values are given for the SNP with the smallest P value per locus. Definitions: Chr = chromosome, Position = base pair position in NCBI37/hg19 genome build, Lead SNP = single-nucleotide polymorphism with the smallest P value within each LD-block, Ancestry = African (AFR), European (EUR), and Native American (NAM), OR = odds ratio, genome-wide significant evidence for association = $P < 5E-05$, suggestive evidence for association = $P < 0.001$. The block with significant evidence for association is highlighted in bold font. All intronic variants are canonical transcripts. Results for each block associated with Alzheimer’s disease at each locus are provided in Additional file 2

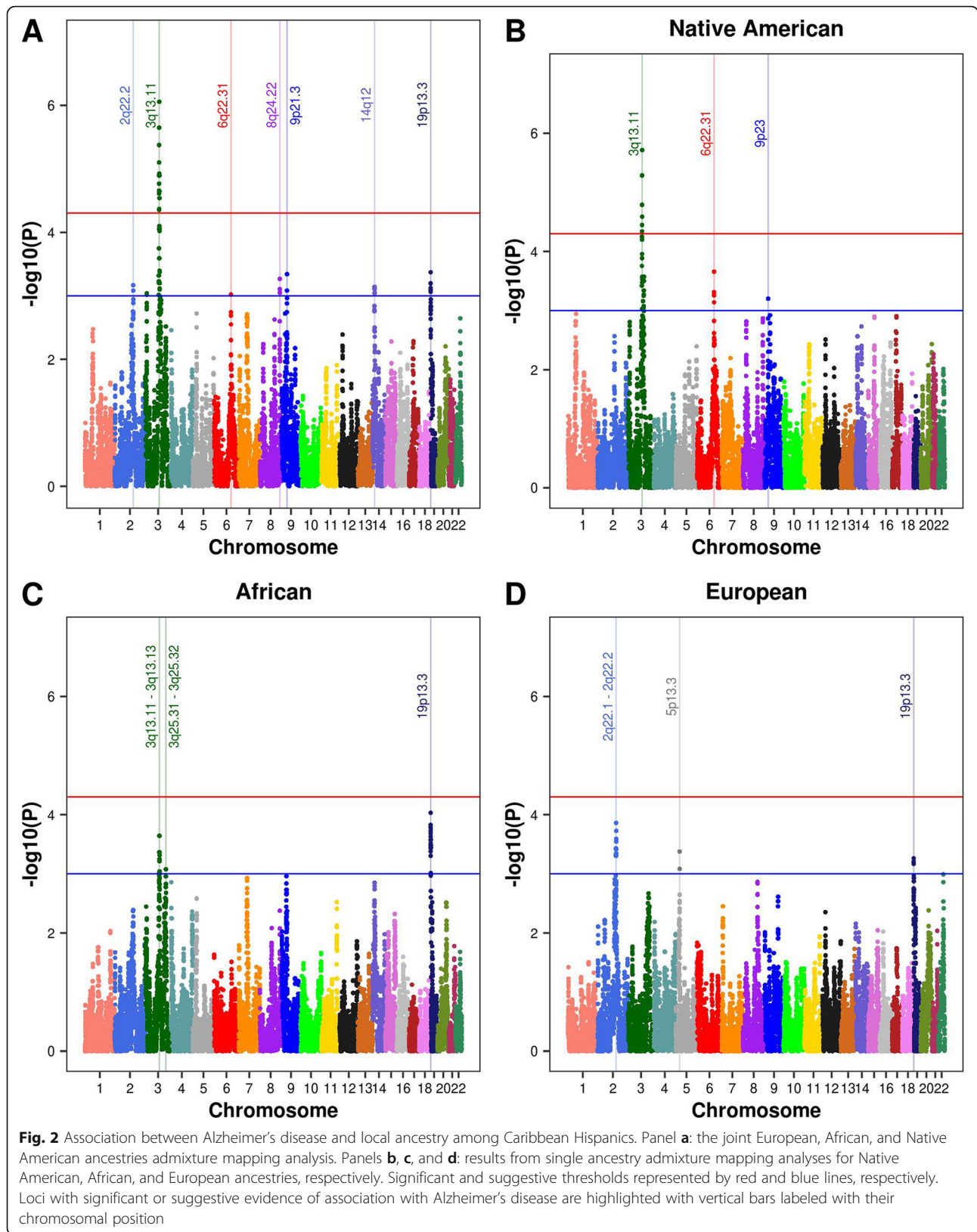


Table 3 Variants with significant or suggestive evidence for association with Alzheimer's disease within 3q13.11

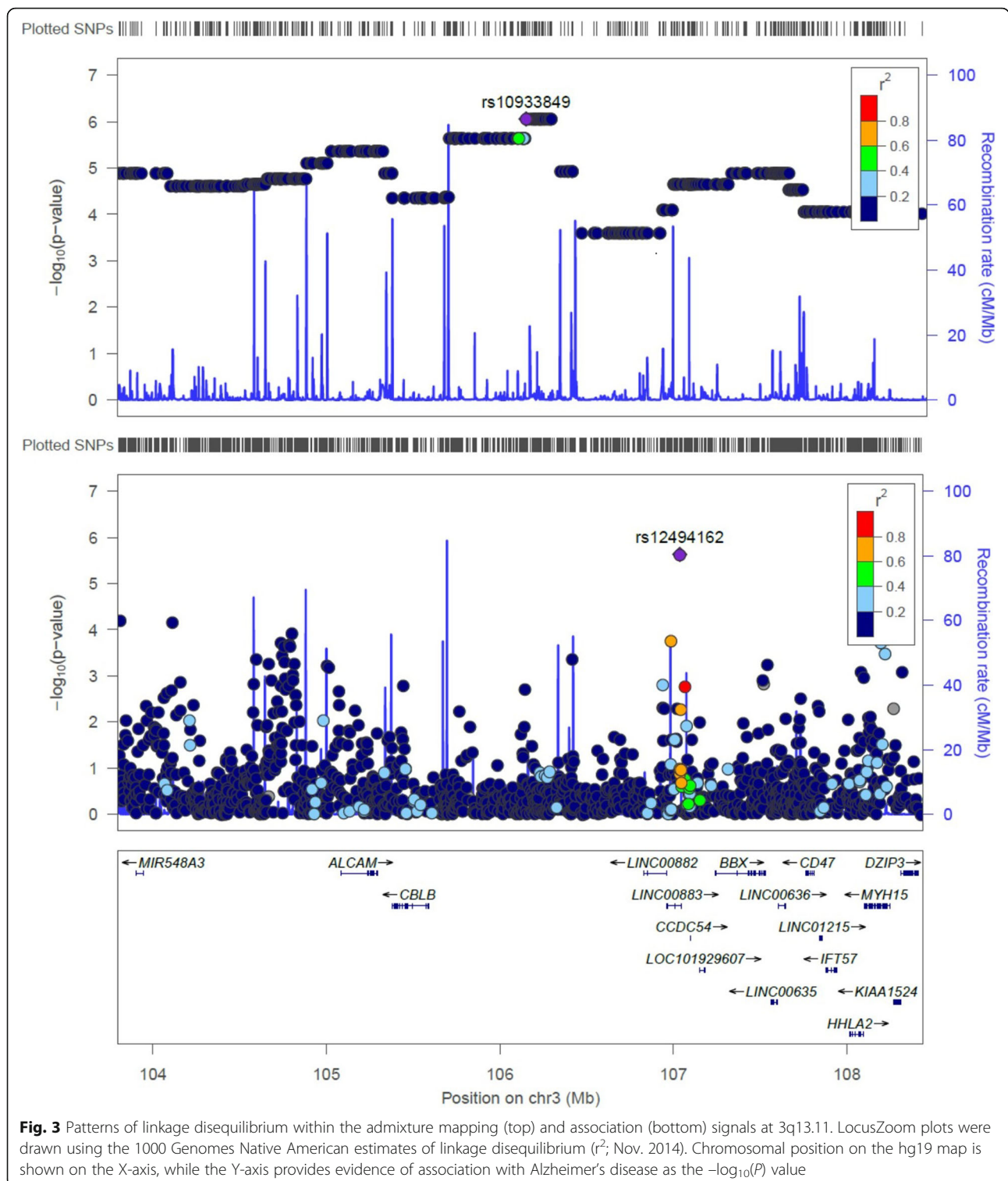
Chr	SNP	Position	Allele	Gene	Consequence	OR (95% CI)	P
3	rs1731642	103,811,750	G	38kb 5' of AC016970.1	intergenic	1.33 (1.17—1.53)	6.36E-05
3	rs9848147	104,111,719	C	338kb 5' of AC016970.1	intergenic	1.65 (1.30—2.08)	7.02E-05
3	rs2673478	104,158,244	G	385kb 5' of AC016970.1	intergenic	1.22 (1.09—1.38)	1.33E-03
3	rs16850638	104,594,467	A	491kb 5' of <i>ALCAM</i>	intergenic	0.69 (0.57—0.84)	4.24E-04
3	rs12492893	104,667,061	A	419kb 5' of <i>ALCAM</i>	intergenic	1.23 (1.10—1.38)	5.45E-04
3	rs9288795	104,683,753	G	402kb 5' of <i>ALCAM</i>	intergenic	0.83 (0.74—0.92)	1.21E-03
3	rs16850772	104,736,837	A	349kb 5' of <i>ALCAM</i>	intergenic	1.26 (1.12—1.41)	1.91E-04
3	rs9883825	104,745,091	G	341kb 5' of <i>ALCAM</i>	intergenic	0.80 (0.71—0.90)	3.55E-04
3	rs13325696	104,755,481	A	330kb 5' of <i>ALCAM</i>	intergenic	1.28 (1.13—1.45)	2.29E-04
3	rs1503079	104,759,032	A	327kb 5' of <i>ALCAM</i>	intergenic	0.81 (0.72—0.91)	4.92E-04
3	rs1566720	104,761,408	A	324kb 5' of <i>ALCAM</i>	intergenic	0.82 (0.73—0.92)	1.13E-03
3	rs1587707	104,786,133	C	300kb 5' of <i>ALCAM</i>	intergenic	0.80 (0.71—0.89)	2.06E-04
3	rs2895295	104,790,415	C	295kb 5' of <i>ALCAM</i>	intergenic	0.80 (0.71—0.90)	2.32E-04
3	rs10933809	104,800,350	A	285kb 5' of <i>ALCAM</i>	intergenic	0.79 (0.70—0.89)	1.19E-04
3	rs1503089	104,805,417	G	280kb 5' of <i>ALCAM</i>	intergenic	1.23 (1.09—1.39)	9.28E-04
3	rs1503075	104,806,853	G	279kb 5' of <i>ALCAM</i>	intergenic	0.81 (0.72—0.92)	1.23E-03
3	rs1503158	104,815,105	G	271kb 5' of <i>ALCAM</i>	intergenic	0.80 (0.71—0.90)	5.52E-04
3	rs13322578	105,002,637	A	83kb 5' of <i>ALCAM</i>	intergenic	0.77 (0.66—0.89)	5.90E-04
3	rs9816851	105,012,805	G	73kb 5' of <i>ALCAM</i>	intergenic	0.77 (0.67—0.89)	6.64E-04
3	rs9860520	106,416,700	A	9.1kb 3' of <i>Y_RNA</i>	intergenic	0.79 (0.70—0.90)	4.28E-4
3	rs12489299	106,985,680	A	<i>DUBR</i>	intron	0.69 (0.57—0.83)	1.70E-04
3	rs12494162	107,036,379	A	<i>DUBR</i>	Intron	0.69 (0.60—0.80)	2.33E-06
3	rs7615167	107,514,134	A	<i>BBX</i>	Intron	1.56 (1.20—2.02)	1.23E-03
3	rs4855772	107,540,375	G	10kb 3' of <i>BBX</i>	intergenic	1.80 (1.31—2.49)	5.81E-04

Significance thresholds are based on the effective number of independent tests. Abbreviations: Chr = chromosome, SNP = single-nucleotide polymorphism identifier, Position = chromosome 3 position on the GRCh37/hg19 map, Allele = effect allele, OR = odds ratio, significant evidence of association = $P < 6.7E-05$, suggestive evidence of association = $P < 1.4E-03$. Significant results are highlighted in bold font

LocusZoom plots of the admixture mapping and association signals at 3q13.11 using 1000 Genomes Native American estimates of LD (Fig. 3). The lead SNP rs12494162 is in LD with several other SNPs with evidence of association with AD, as expected. In contrast, the lead SNP from the admixture mapping analysis has modest evidence of LD with other SNPs on haplotypes associated with AD, as the admixture mapping signal is driven by differences in ancestry proportions rather than specific genotypes at the locus.

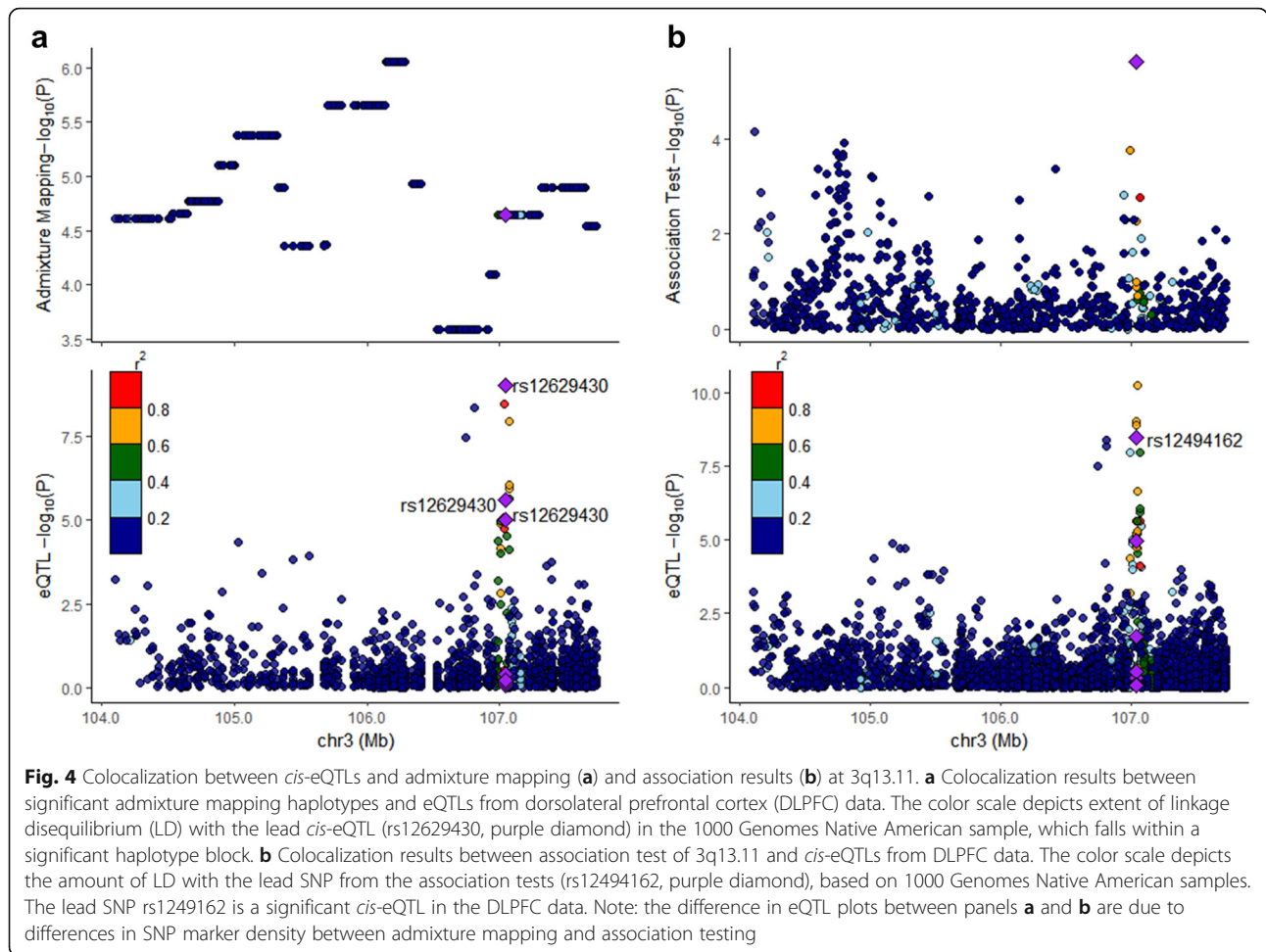
Conditional admixture mapping analyses including both rs12494162 and rs1731642 as covariates eliminated the signal at 3q13.11 ($P = 0.01$), while analyses conditioned on either SNP alone only weakened the signals (Additional file 6), suggesting that admixture mapping and GWAS approaches may be tagging the same underlying variant. We assessed evidence of colocalization between eQTLs identified in DLPCF samples from subjects with primarily European ancestry and the admixture mapping and association

signals at 3q13.11, as comparable studies representing Native Americans are unavailable. These analyses can only identify shared genetic architecture between eQTLs shared across populations and our admixture mapping or association results, which may represent fewer than half of eQTLs [49]. The leading eQTL within the 3q13.11 locus falls within a haplotype significantly associated with AD in the admixture mapping analysis (Fig. 4a): rs12629430 is significantly associated with the expression of lincRNA *DUBR* ($Z = -4.47$, $FDR = 4.9E-04$), lincRNA RP11.446H18.1 ($Z = -6.22$, $FDR = 1.0E-07$), and lincRNA RP11-446H18.6 ($Z = -4.77$, $FDR = 1.4E-04$). Colocalization analyses of the eQTL and admixture mapping signals did not reject the null hypothesis ($PP0 = 0.9550$). In contrast, the lead SNP from our targeted association testing within 3q13.11, rs12494162, is also an eQTL significantly associated with the expression of lincRNA RP11.446H18.1 ($Z = -6.01$, $FDR = 3.4E-07$), and lincRNA RP11-446H18.6 ($Z = -4.44$, $FDR = 5.4E-04$)



(Fig. 4b). Colocalization analyses are not significant, but suggest association with both AD and gene expression here and weakly favor the model of independent SNPs driving these association ($PP_3 = 0.5070$) rather than one shared SNP ($PP_4 = 0.4130$).

Within the 3q13.11 locus (chr3:103,747,624-107,725,831), we prioritized candidate protein-coding genes which fell either within one of the 15 LD-blocks associated with AD or within an intersecting TAD using the following features in transcriptomic studies



representing European ancestry: (1) genes in which expression in brain tissue is significantly associated with *cis*-eQTL within the region of interest and (2) genes which are differentially expressed in the brain between AD cases and controls. The 3q13.11 region of interest spans five protein-coding genes: *ALCAM*, *BBX*, *CBLB*, *CCDC54*, and *CD47* while four additional genes fall within the same TAD as *BBX*: *IFT57*, *HHLA2*, *MYH15*, and *KIAA1524* (Additional file 7). A

recent transcriptomic study [45] of 1694 brain samples identified 369 significant *cis*-eQTLs for *IFT57*, 182 significant *cis*-eQTLs for *ALCAM*, 118 significant *cis*-eQTLs for *CBLB*, 47 significant *cis*-eQTLs for *CD47*, 22 significant *cis*-eQTLs for *BBX*, and 6 significant *cis*-eQTLs for *MYH15* (FDR < 0.05). The strongest *cis*-eQTL per gene is reported in Table 4, with all *cis*-eQTLs reported in Additional file 8. Another transcriptomic study [46] including an overlapping

Table 4 Significant *cis* expression quantitative loci (eQTLs) for candidate genes within 3q13.11 region of interest

Position	SNP	Gene	Z	β	Allele	P	FDR
105,108,867	rs6797043	<i>ALCAM</i>	-4.89	-0.3429	C	1.29E-06	8.24E-05
107,583,197	rs9875001	<i>BBX</i>	3.72	0.2620	T	2.17E-04	7.43E-03
105,488,134	rs139969708	<i>CBLB</i>	4.01	0.4411	T	7.03E-05	2.85E-03
107,550,819	rs1908324	<i>CD47</i>	3.83	0.2512	T	1.45E-04	5.29E-03
107,917,824	rs9857584	<i>IFT57</i>	6.91	0.6714	C	1.33E-11	1.89E-09
107,823,224	rs80003826	<i>MYH15</i>	-3.92	-0.4498	A	9.85E-05	3.80E-03

Results are filtered to only include those with false discovery rate (FDR) < 0.05, restricted to the SNP with the smallest FDR value. Results for all significant eQTLs are presented in Additional file 8 and the original data can be found at <https://www.synapse.org/#!Synapse:syn17015233>. Legend: Position = GRCh37 position of the variant on chromosome 3, Z = Z statistic, β = estimated effect size, allele = effect allele, FDR = false discovery rate

sample set of 2114 brain samples representing 478 cases and 300 controls identified significant evidence for differential gene expression in AD for both *ALCAM* ($Z = 2.75$, $FDR = 2.76E-02$) and *BBX* ($Z = 3.73$, $FDR = 1.84E-03$) (Additional file 9). While variation in the 3q13.11 region is associated with expression levels of *ALCAM*, *BBX*, *CBLB*, *CD47*, *IFT57*, and *MYH15* in the brain, only *ALCAM* and *BBX* were significantly differentially expressed between AD cases and controls.

Discussion

Admixture mapping of AD within a Caribbean Hispanic sample identified one genome-wide significant signal on 3q13.11 ($P = 8.76E-07$) and six unique suggestive signals at 2q22.2, 6q22.31, 8q24.22, 9p21.3, 14q12, and 19p13.3. The admixture mapping signal on 3q13.11 spanned 15 haplotype blocks, where the Native American ancestry is associated with reduced risk of AD. Association between the Native American ancestry and reduced risk of AD has previously been reported [53, 54]. Suggestive evidence of association between the 3q13.11 locus and AD has recently been reported in an African American GWAS involving nearly three times the sample size as our study [55], demonstrating the effectiveness of the admixture mapping approach as a complement to GWAS.

While admixture mapping provides insights into the genetic basis of disease in multiethnic populations, integration of AD transcriptomics allowed us to nominate candidate genes within 3q13.11. *ALCAM* and *BBX*, the genes with significant evidence for both brain eQTLs and differential expression between AD cases and controls, both have robust support in the literature for a functional relationship to AD. Proteomic studies suggest *ALCAM*, which plays a role in neuron-neuron adhesion and neurite growth networks, is dysregulated during the progression of AD [56]. *ALCAM* is also involved in blood-brain barrier disruption and T cell-dependent neurodegeneration [57], biological pathways implicated in the progression of AD [58]. Furthermore, *ALCAM* is a target gene of miR-142 which is significantly upregulated in the AD brain [59, 60]. *BBX* is differentially expressed in the entorhinal cortex and hippocampus and appears to play a role in the crosstalk between the peripheral blood and the central nervous system [61]. Multiple studies have shown that *BBX* is differentially expressed in the AD brain [61, 62], while another implicated *BBX* as a candidate Master Regulator responsible for AD progression [63].

Each of the loci harboring suggestive admixture mapping signals have also been previously associated with AD risk and/or pathology. The signal on 19p13.3 is driven by African ancestry and spans *ABCA7*, a gene in

which coding changes have been associated with risk of AD in both African American and European American samples [64–67]. *LRP1B* within the 2q22.2 signal has been implicated in the production and presentation of amyloid beta ($A\beta$) [68], while multiple *LRP1B* haplotypes are associated with risk of developing AD in studies representing European Americans [69] and Caribbean Hispanics [70]. Variants on 14q13.1 near *NPAS3* have been associated with AD biomarkers [71] and general cognitive function [72]. Variants in *ZFAT* on 8q24.22 have been associated with extreme longevity [73] and cerebrospinal fluid *tau/A\beta42* levels [74]. Within 6q22.31, *TRDN* variants have implicated in cerebral $A\beta$ deposition in *APOE* $\epsilon 4$ non-carriers [75] and rate of cognitive decline in AD [76]. Finally, 9p21.3 has previously been linked to AD risk [77], and variants within the region have been associated with both vascular dementia and AD [78].

Limitations

Our study has several limitations. Admixture mapping identifies regions associated with a given trait which must then be fine-mapped to identify the underlying risk variants. Colocalization analysis is not well powered in our study due to the poor representation of non-European populations in large eQTL data sets, as the genetic architecture of eQTLs can be ancestry specific [49, 79]. Fine-mapping analyses of whole-genome sequence data collected in this sample may allow the detection of the variants responsible for the admixture mapping signals. Publicly available datasets comparable in size or Native American ancestry proportions suitable for replication analyses are not available. Ongoing efforts, including AMP-AD and the Alzheimer's Disease Sequencing Project, will provide data which may assist these efforts in the future.

Conclusions

Most AD GWAS have represented samples with European ancestry, and alternative strategies may detect additional genetic variants influencing AD in multiethnic populations. Caribbean Hispanics, despite being more likely to be diagnosed with AD [80, 81], have been underrepresented in AD genetics studies [82]. We illustrated the power of admixture mapping for detecting loci associated with AD in a Caribbean Hispanic sample, provided robust evidence for this association, and nominated several candidate genes with orthogonal functional and statistical evidence for a role in AD. Further investigation of these loci and nominated genes could lead to a better understanding of the genetic heterogeneity of AD in populations with significant Native American ancestry.

Abbreviations

A β : Amyloid beta; AD: Alzheimer's disease; AMP-AD: Accelerating Medical Partnerships for Alzheimer's Disease; CU Hispanics: Participants in the Columbia University Study of Caribbean Hispanics and Late Onset Alzheimer's Disease study; DLPFC: Dorsolateral prefrontal cortex; eQTL: Expression quantitative trait locus; FDR: False discovery rate; GRM: Genetic relatedness matrix; GWAS: Genome-wide association study; LD: Linkage disequilibrium; PC: Principal components; SNP: Single-nucleotide polymorphism; TAD: Topologically associated domain

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13195-021-00866-9>.

Additional file 1: Joint Photographic Group format .jpg file. Admixture mapping of Alzheimer's disease in the Caribbean Hispanics excluding the heterozygosity outlier samples. Description: Joint European, African and Native American ancestries admixture mapping analysis, with chromosomal position on hg19 on the X-axis and $-\log_{10}(P)$ values on the Y-axis. Significant and suggestive thresholds represented by red and blue lines, respectively. Loci with significant or suggestive evidence of association with Alzheimer's disease are highlighted with vertical bars labeled with the chromosomal position of the peak.

Additional file 2: File format: Microsoft Word .docx file. Title: Linkage disequilibrium blocks within regions with either genome-wide significant or suggestive evidence for association between local ancestry and Alzheimer's disease. Description: Regions reaching at least suggestive evidence of association with Alzheimer's disease are defined and the evidence for that association is summarized. Abbreviations: Chr = chromosome, Position = physical positions based on GRCh37/hg19 map, SNP = single nucleotide polymorphism, Effect size (95%CI): odds ratio for AD followed by the 95% confidence interval.

Additional file 3: File format: Portable graphics format .png file. Title: Linkage disequilibrium patterns at 3q13.11 by ancestry. Description: Each panel illustrates the amount of linkage disequilibrium (D) between pairs of markers in the 3q13.11 locus using different reference populations drawn from the 1000 Genomes data (Nov 2014).

Additional file 4: File format: Joint Photographic Group format .jpg file. Title: Admixture mapping of Alzheimer's disease in the Caribbean Hispanics adjusted for age and sex. Description: Joint European, African and Native American ancestries admixture mapping analysis, with chromosomal position on hg19 on the X-axis and $-\log_{10}(P)$ values on the Y-axis. Significant and suggestive thresholds represented by red and blue lines, respectively. Loci with significant or suggestive evidence of association with Alzheimer's disease are highlighted with vertical bars labeled with the chromosomal position of the peak.

Additional file 5: File format: Portable graphics format .png file. Title: Genome-wide association testing results for Alzheimer's disease. Description: Alzheimer's disease status was tested for association with genotypes using a logistic regression model, adjusting for global ancestry proportions and APOE $\epsilon 2$ and $\epsilon 4$ allele dosages as fixed effects and the genetic relatedness matrix as a random effect. Genomic position on the hg19 map are provided on the X-axis and $-\log_{10}(P)$ values on the Y-axis. The dotted horizontal line corresponds to a genome-wide significance threshold of 5E-08.

Additional file 6: File format: Portable graphics format .png file. Title: Conditional admixture mapping results at 3q13.11. Description: Each panel represents an admixture mapping analysis in the 3q13.11 locus, conditioned on the two single nucleotide polymorphisms (SNPs) associated with Alzheimer's disease. The first panel shows the admixture mapping results with both SNPs are included in the analysis model, while the latter two adjust for only the named SNP. The X-axis represents the genomic position on chromosome 3 and the Y-axis represents $-\log_{10}(P)$ values. The horizontal red line represents region-wide significance, while the blue line represents suggestive evidence of association. Green dots represent the locus reaching genome-wide significance in the original admixture mapping analysis.

Additional file 7: File format: Portable graphics format .png file. Title: Local context of admixture mapping and association signals at 3q13.11. Description: The top panel illustrates the admixture mapping testing for association with Alzheimer's disease status model. The second panel provides the association testing results for AD. The third panel provides the position of genes within the region of interest. The fourth panel illustrates the topologically associated domains (TADs; blue and gold bars) indicated by Hi-C experiments in DLPFC, where the heat map in magenta indicates the number of sequencing reads aligning to a pair of physical positions. Red horizontal lines represent the genome-wide significance threshold ($P < 5E-05$) and a blue line at the suggestive threshold ($P < 0.001$) used for admixture mapping. Sequence positions are aligned to the GRCh37/hg19 genome reference and are represented by the X-axis.

Additional file 8: File format: Microsoft Word .docx file. Title: Significant *cis* expression quantitative loci (eQTLs) for candidate genes within 3q13.11 region of interest. Description: Evidence that genotypes at a marker are associated with gene expression values. Results are filtered to only include those with false discovery rate < 0.05 . Data are stored in the Synapse repository, Synapse ID: syn17015233, <https://www.synapse.org/#Synapse:syn17015233>. Legend: Chr = chromosome, Position = GRCh37 position of the variant on chromosome 3, Z = Z statistic, FDR = false discovery rate, β = estimated effect size, A1 = allele 1, A2 = allele 2, A2freq = observed frequency of the A2 allele, Aup = allele associated with increased expression of the gene.

Additional file 9: File format: Microsoft Word .docx file. Title: Evidence of differential gene expression between those with and without Alzheimer's disease for 3q13.11 candidate genes. Description: Data are stored in the Synapse repository, Synapse ID: syn11914606, <https://www.synapse.org/#Synapse:syn11914606>, file meta.anlz.ad_cntrl.tsv, which included no values for CCDC54 or HHLA2. Legend: z.fixed = Z statistic for the fixed effects model, p.fixed = p-value for fixed effect model, z.random = Z statistic for the random effects model, p.random = p-value for the random effects model, fdr.fixed = false discovery rate from the fixed effects model, fdr.random = false discovery rate from random effects model. Results reaching the significance level of fdr.random < 0.05 are highlighted in bold.

Acknowledgements

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 the National Institutes of Health, National Institute on Aging (NIA) [grant number R01AG059737]. 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, R01AG015473, R56AG051876]. We acknowledge the EFIGA study participants and the EFIGA research and support staff for their contributions to this study.

Authors' contributions

All authors made substantial contributions to the design of the work, interpretation of data, and critically revised the manuscript. EEB acquired the data, and ARVH and DX analyzed the data and drafted the work. All authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Funding

Research funding was provided by the National Institutes of Health, National Institute on Aging (NIA) [grant number R01AG059737].

Availability of data and materials

The genome scan data representing the Caribbean Hispanics are available through an application to the database of Genotypes and phenotypes (dbGaP; study ID: phs000496.v1.p1). These data are consented for general research use and are available through an application to dbGaP (https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000496.v1.p1). Analysis pipelines were based upon the GENESIS package in the R programming language (<https://bioconductor.org/packages/release/bioc/html/GENESIS.html>) and a published ancestry inference pipeline (https://github.com/armartin/ancestry_pipeline).

Declarations

Consent to publication

Not applicable. See Availability of data and material.

Ethics approval and consent to participate

This study was reviewed and approved by the University of Washington Institutional Review Board (STUDY ID: STUDY00000240). Data analysis was restricted to publically available data consented for general research use.

Competing interests

On behalf of all authors, the corresponding author states that there is no conflict of interest.

Author details

¹Department of Biostatistics, University of Washington, Seattle, WA, USA. ²Institute for Public Health Genetics, University of Washington, Seattle, WA, USA. ³Division of Medical Genetics, University of Washington, BOX 357720, Seattle, WA 98195-7720, USA.

Received: 4 February 2021 Accepted: 20 June 2021

Published online: 03 July 2021

References

- 2020 Alzheimer's disease facts and figures. *Alzheimers Dement*. 2020.
- 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(2):168–74. <https://doi.org/10.1001/archpsyc.63.2.168>.
- Ridge PG, Hoyt KB, Boehme K, Mukherjee S, Crane PK, Haines JL, et al. Assessment of the genetic variance of late-onset Alzheimer's disease. *Neurobiol Aging*. 2016;41:200 e13–20.
- 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(5123):921–3. <https://doi.org/10.1126/science.8346443>.
- Strittmatter WJ, Saunders AM, Schmechel D, Pericak-Vance M, Enghild J, Salvesen GS, et al. Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proc Natl Acad Sci U S A*. 1993;90(5):1977–81. <https://doi.org/10.1073/pnas.90.5.1977>.
- Andrews SJ, Fulton-Howard B, Goate A. Interpretation of risk loci from genome-wide association studies of Alzheimer's disease. *Lancet Neurol*. 2020;19(4):326–35. [https://doi.org/10.1016/S1474-4422\(19\)30435-1](https://doi.org/10.1016/S1474-4422(19)30435-1).
- Bellenguez C, Grenier-Boley B, Lambert JC. Genetics of Alzheimer's disease: where we are, and where we are going. *Curr Opin Neurobiol*. 2020;61:40–8. <https://doi.org/10.1016/j.conb.2019.11.024>.
- Popejoy AB, Fullerton SM. Genomics is failing on diversity. *Nature*. 2016; 538(7624):161–4. <https://doi.org/10.1038/538161a>.
- Reitz C, Mayeux R. Genetics of Alzheimer's disease in Caribbean Hispanic and African American populations. *Biol Psychiatry*. 2014;75(7):534–41. <https://doi.org/10.1016/j.biopsych.2013.06.003>.
- 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(3):233–43. <https://doi.org/10.1016/j.jalz.2015.02.012>.
- 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(7):727–38. <https://doi.org/10.1016/j.jalz.2016.12.012>.
- McKeigue PM. Prospects for admixture mapping of complex traits. *Am J Hum Genet*. 2005;76(1):1–7. <https://doi.org/10.1086/426949>.
- Gouveia MH, Cesar CC, Santolalla ML, Anna HPS, Scliar MO, Leal TP, et al. Genetics of cognitive trajectory in Brazilians: 15 years of follow-up from the Bambui-Epigen Cohort Study of Aging. *Sci Rep*. 2019;9(1):18085. <https://doi.org/10.1038/s41598-019-53988-4>.
- Wang H, Cade BE, Sofer T, Sands SA, Chen H, Browning SR, et al. Admixture mapping identifies novel loci for obstructive sleep apnea in Hispanic/Latino Americans. *Hum Mol Genet*. 2019;28(4):675–87. <https://doi.org/10.1093/hmg/ddy387>.
- Ziyatdinov A, Parker MM, Vaysse A, Beaty TH, Kraft P, Cho MH, et al. Mixed-model admixture mapping identifies smoking-dependent loci of lung function in African Americans. *Eur J Hum Genet*. 2020;28(5):656–68. <https://doi.org/10.1038/s41431-019-0545-8>.
- Shriner D. Overview of admixture mapping. *Curr Protoc Hum Genet*. 2013; Chapter 1:Unit 1 23.
- Chakraborty R, Weiss KM. Admixture as a tool for finding linked genes and detecting that difference from allelic association between loci. *Proc Natl Acad Sci U S A*. 1988;85(23):9119–23. <https://doi.org/10.1073/pnas.85.23.9119>.
- Zhu X, Cooper RS, Elston RC. Linkage analysis of a complex disease through use of admixed populations. *Am J Hum Genet*. 2004;74(6):1136–53. <https://doi.org/10.1086/421329>.
- Benn-Torres J, Bonilla C, Robbins CM, Waterman L, Moses TY, Hernandez W, et al. Admixture and population stratification in African Caribbean populations. *Ann Hum Genet*. 2008;72(Pt 1):90–8. <https://doi.org/10.1111/j.1469-1809.2007.00398.x>.
- Blue EE, Horimoto A, Mukherjee S, Wijsman EM, Thornton TA. Local ancestry at APOE modifies Alzheimer's disease risk in Caribbean Hispanics. *Alzheimers Dement*. 2019;15(12):1524–32. <https://doi.org/10.1016/j.jalz.2019.07.016>.
- Tosto G, Fu H, Vardarajan BN, Lee JH, Cheng R, Reyes-Dumeyer D, et al. F-box/LRR-repeat protein 7 is genetically associated with Alzheimer's disease. *Ann Clin Transl Neurol*. 2015;2(8):810–20. <https://doi.org/10.1002/acn3.223>.
- 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(7311):52–8. <https://doi.org/10.1038/nature09298>.
- Cann HM, de Toma C, Cazes L, Legrand MF, Morel V, Piouffre L, et al. A human genome diversity cell line panel. *Science*. 2002;296(5566):261–2. <https://doi.org/10.1126/science.296.5566.261b>.
- Cavalli-Sforza LL. The Human Genome Diversity Project: past, present and future. *Nat Rev Genet*. 2005;6(4):333–40. <https://doi.org/10.1038/nrg1579>.
- 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(3):559–75. <https://doi.org/10.1086/519795>.
- Kuhn RM, Haussler D, Kent WJ. The UCSC genome browser and associated tools. *Brief Bioinform*. 2013;14(2):144–61. <https://doi.org/10.1093/bib/bbs038>.
- Vardarajan BN, Schaid DJ, Reitz C, Lantigua R, Medrano M, Jimenez-Velazquez IZ, et al. Inbreeding among Caribbean Hispanics from the Dominican Republic and its effects on risk of Alzheimer disease. *Genet Med*. 2015;17(8):639–43. <https://doi.org/10.1038/gim.2014.161>.
- Conomos MP, Reiner AP, Weir BS, Thornton TA. Model-free estimation of recent genetic relatedness. *Am J Hum Genet*. 2016;98(1):127–48. <https://doi.org/10.1016/j.ajhg.2015.11.022>.
- 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(4):276–93. <https://doi.org/10.1002/gepi.21896>.
- Gogarten SM, Sofer T, Chen H, Yu C, Brody JA, Thornton TA, et al. Genetic association testing using the GENESIS R/Bioconductor package. *Bioinformatics*. 2019;35(24):5346–8. <https://doi.org/10.1093/bioinformatics/btz567>.
- Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. *Bioinformatics*. 2010;26(22):2867–73. <https://doi.org/10.1093/bioinformatics/btq559>.
- Martin ER, Tunc I, Liu Z, Slifer SH, Beecham AH, Beecham GW. Properties of global- and local-ancestry adjustments in genetic association tests in admixed populations. *Genet Epidemiol*. 2018;42(2):214–29. <https://doi.org/10.1002/gepi.22103>.
- Delaneau O, Coulonges C, Zagury JF. Shape-IT: new rapid and accurate algorithm for haplotype inference. *BMC Bioinformatics*. 2008;9(1):540. <https://doi.org/10.1186/1471-2105-9-540>.
- 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. *Nature*. 2015;526(7571):68–74.
- 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(2):278–88. <https://doi.org/10.1016/j.ajhg.2013.06.020>.
- Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS, et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods*. 2015;12(2):115–21. <https://doi.org/10.1038/nmeth.3252>.

37. Brown LA, Sofer T, Stulp AM, Baier LJ, Kramer HJ, Masindova I, et al. Admixture mapping identifies an Amerindian ancestry locus associated with albuminuria in hispanics in the United States. *J Am Soc Nephrol*. 2017;28(7):2211–20. <https://doi.org/10.1681/ASN.2016091010>.
38. Grinde KE, Brown LA, Reiner AP, Thornton TA, Browning SR. Genome-wide significance thresholds for admixture mapping studies. *Am J Hum Genet*. 2019;104(3):454–65. <https://doi.org/10.1016/j.ajhg.2019.01.008>.
39. Shriner D, Adeyemo A, Rotimi CN. Joint ancestry and association testing in admixed individuals. *PLoS Comput Biol*. 2011;7(12):e1002325. <https://doi.org/10.1371/journal.pcbi.1002325>.
40. Turner SD. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *J Open Source Software*. 2018;3(25):731.
41. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Glied TP, et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics*. 2010;26(18):2336–7. <https://doi.org/10.1093/bioinformatics/btq419>.
42. Li MX, Yeung JM, Cherny SS, Sham PC. Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Hum Genet*. 2012;131(5):747–56. <https://doi.org/10.1007/s00439-011-1118-2>.
43. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005;21(2):263–5. <https://doi.org/10.1093/bioinformatics/bth457>.
44. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, et al. The Ensembl variant effect predictor. *Genome Biol*. 2016;17(1):122. <https://doi.org/10.1186/s13059-016-0974-4>.
45. Sieberts SK, Perumal TM, Carrasquillo MM, Allen M, Reddy JS, Hoffman GE, et al. Large eQTL meta-analysis reveals differing patterns between cerebral cortical and cerebellar brain regions. *Sci Data*. 2020;7(1):340. <https://doi.org/10.1038/s41597-020-00642-8>.
46. Logsdon BA, Perumal TM, Swarup V, Wang M, Funk C, Gaiteri C, et al. Meta-analysis of the human brain transcriptome identifies heterogeneity across human AD coexpression modules robust to sample collection and methodological approach. 2019.
47. Wang Y, Song F, Zhang B, Zhang L, Xu J, Kuang D, et al. The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. *Genome Biol*. 2018;19(1):151. <https://doi.org/10.1186/s13059-018-1519-9>.
48. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*. 2012;485(7398):376–80. <https://doi.org/10.1038/nature11082>.
49. Shang L, Smith JA, Zhao W, Kho M, Turner ST, Mosley TH, et al. Genetic architecture of gene expression in European and African Americans: an eQTL mapping study in GENOA. *Am J Hum Genet*. 2020;106(4):496–512. <https://doi.org/10.1016/j.ajhg.2020.03.002>.
50. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet*. 2014;10(5):e1004383. <https://doi.org/10.1371/journal.pgen.1004383>.
51. Liu B, Gludemans MJ, Rao AS, Ingelsson E, Montgomery SB. Abundant associations with gene expression complicate GWAS follow-up. *Nat Genet*. 2019;51(5):768–9. <https://doi.org/10.1038/s41588-019-0404-0>.
52. Benjamin R, Leake A, McArthur FK, Ince PG, Candy JM, Edwardson JA, et al. Protective effect of apoE epsilon 2 in Alzheimer's disease. *Lancet*. 1994;344(8920):473–4. [https://doi.org/10.1016/S0140-6736\(94\)91804-X](https://doi.org/10.1016/S0140-6736(94)91804-X).
53. 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(5):311–7. <https://doi.org/10.1159/000339672>.
54. Moreno DJ, Ruiz S, Rios A, Lopera F, Ostos H, Via M, et al. Association of GWAS top genes with late-onset Alzheimer's disease in Colombian population. *Am J Alzheimers Dis Other Dement*. 2017;32(1):27–35. <https://doi.org/10.1177/1533317516679303>.
55. Kunkle BW, Schmidt M, Klein HU, Naj AC, Hamilton-Nelson KL, Larson EB, et al. Novel Alzheimer disease risk loci and pathways in African American individuals using the african genome resources panel: a meta-analysis. *JAMA Neurol*. 2021;78(1):102–13. <https://doi.org/10.1001/jamaneurol.2020.3536>.
56. Zelaya MV, Perez-Valderrama E, de Morentin XM, Tunon T, Ferrer I, Luquin MR, et al. Olfactory bulb proteome dynamics during the progression of sporadic Alzheimer's disease: identification of common and distinct olfactory targets across Alzheimer-related co-pathologies. *Oncotarget*. 2015;6(37):39437–56. <https://doi.org/10.18632/oncotarget.6254>.
57. Gonzalez H, Elgueta D, Montoya A, Pacheco R. Neuroimmune regulation of microglial activity involved in neuroinflammation and neurodegenerative diseases. *J Neuroimmunol*. 2014;274(1–2):1–13. <https://doi.org/10.1016/j.jneuroim.2014.07.012>.
58. Zenaro E, Piacentino G, Constantin G. The blood-brain barrier in Alzheimer's disease. *Neurobiol Dis*. 2017;107:41–56. <https://doi.org/10.1016/j.nbd.2016.07.007>.
59. Lau P, Bossers K, Janky R, Salta E, Frigerio CS, Barbash S, et al. Alteration of the microRNA network during the progression of Alzheimer's disease. *EMBO Mol Med*. 2013;5(10):1613–34. <https://doi.org/10.1002/emmm.201201974>.
60. Ghanbari M, Munshi ST, Ma B, Lendemeijer B, Bansal S, Adams HH, et al. A functional variant in the miR-142 promoter modulating its expression and conferring risk of Alzheimer disease. *Hum Mutat*. 2019;40(11):2131–45. <https://doi.org/10.1002/humu.23872>.
61. Pang X, Zhao Y, Wang J, Zhou Q, Xu L, Kang, et al. The bioinformatic analysis of the dysregulated genes and microRNAs in entorhinal cortex, hippocampus, and blood for Alzheimer's disease. *Biomed Res Int*. 2017;2017:9084507.
62. Wang Y, Wang Z. Identification of dysregulated genes and pathways of different brain regions in Alzheimer's disease. *Int J Neurosci*. 2020;130(11):1082–94. <https://doi.org/10.1080/00207454.2020.1720677>.
63. Aubry S, Shin W, Cray JF, Lefort R, Qureshi YH, Lefebvre C, et al. Assembly and interrogation of Alzheimer's disease genetic networks reveal novel regulators of progression. *Plos One*. 2015;10(3):e0120352. <https://doi.org/10.1371/journal.pone.0120352>.
64. Hollingworth P, Harold D, Sims R, Gerrish A, Lambert JC, Carrasquillo MM, et al. Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. *Nat Genet*. 2011;43(5):429–35. <https://doi.org/10.1038/ng.803>.
65. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet*. 2013;45(12):1452–8. <https://doi.org/10.1038/ng.2802>.
66. Reitz C, Jun G, Naj A, Rajbhandary R, Vardarajan BN, Wang LS, et al. Variants in the ATP-binding cassette transporter (ABCA7), apolipoprotein E 4, and the risk of late-onset Alzheimer disease in African Americans. *JAMA*. 2013;309(14):1483–92. <https://doi.org/10.1001/jama.2013.2973>.
67. De Rooek A, Van Broeckhoven C, Sleegers K. The role of ABCA7 in Alzheimer's disease: evidence from genomics, transcriptomics and methylomics. *Acta Neuropathol*. 2019;138(2):201–20. <https://doi.org/10.1007/s00401-019-01994-1>.
68. Cam JA, Zerbinatti CV, Knisely JM, Hecimovic S, Li Y, Bu G. The low density lipoprotein receptor-related protein 1B retains beta-amyloid precursor protein at the cell surface and reduces amyloid-beta peptide production. *J Biol Chem*. 2004;279(28):29639–46. <https://doi.org/10.1074/jbc.M313893200>.
69. Poduslo SE, Huang R, Spiro A 3rd. A genome screen of successful aging without cognitive decline identifies LRP1B by haplotype analysis. *Am J Med Genet B Neuropsychiatr Genet*. 2010;153B(1):114–9. <https://doi.org/10.1002/ajmg.b.30963>.
70. Shang Z, Lv H, Zhang M, Duan L, Wang S, Li J, et al. Genome-wide haplotype association study identify TNFRSF1A, CASP7, LRP1B, CDH1 and TG genes associated with Alzheimer's disease in Caribbean Hispanic individuals. *Oncotarget*. 2015;6(40):42504–14. <https://doi.org/10.18632/oncotarget.6391>.
71. Chouraki V, De Bruijn RF, Chapuis J, Bis JC, Reitz C, Schraen S, et al. A genome-wide association meta-analysis of plasma Abeta peptides concentrations in the elderly. *Mol Psychiatry*. 2014;19(12):1326–35. <https://doi.org/10.1038/mp.2013.185>.
72. Trampush JW, Yang ML, Yu J, Knowles E, Davies G, Liewald DC, et al. GWAS meta-analysis reveals novel loci and genetic correlates for general cognitive function: a report from the COGENT consortium. *Mol Psychiatry*. 2017;22(3):336–45. <https://doi.org/10.1038/mp.2016.244>.
73. Miller JB, Ward E, Staley LA, Stevens J, Teerlink CC, Tavana JP, et al. Identification and genomic analysis of pedigrees with exceptional longevity identifies candidate rare variants. *Neurobiol Dis*. 2020;143:104972. <https://doi.org/10.1016/j.nbd.2020.104972>.
74. Maxwell TJ, Corcoran C, Del-Aguila JL, Budde JP, Deming Y, Cruchaga C, et al. Genome-wide association study for variants that modulate relationships between cerebrospinal fluid amyloid-beta 42, tau, and p-tau levels. *Alzheimers Res Ther*. 2018;10(1):86. <https://doi.org/10.1186/s13195-018-0410-y>.

75. Li QS, Parrado AR, Samtani MN, Narayan VA. Alzheimer's disease neuroimaging I. Variations in the FRA10AC1 fragile site and 15q21 are associated with cerebrospinal fluid Aβ₁₋₄₂ level. *Plos One*. 2015;10(8): e0134000.
76. Sherva R, Gross A, Mukherjee S, Koesterer R, Amouyel P, Bellenguez C, et al. Genome-wide association study of rate of cognitive decline in Alzheimer's disease patients identifies novel genes and pathways. *Alzheimers Dement*. 2020;16(8):1134–45. <https://doi.org/10.1002/alz.12106>.
77. Zuchner S, Gilbert JR, Martin ER, Leon-Guerrero CR, Xu PT, Browning C, et al. Linkage and association study of late-onset Alzheimer disease families linked to 9p21.3. *Ann Hum Genet*. 2008;72(Pt 6):725–31. <https://doi.org/10.1111/j.1469-1809.2008.00474.x>.
78. Emanuele E, Lista S, Ghidoni R, Binetti G, Cereda C, Benussi L, et al. Chromosome 9p21.3 genotype is associated with vascular dementia and Alzheimer's disease. *Neurobiol Aging*. 2011;32(7):1231–5. <https://doi.org/10.1016/j.neurobiolaging.2009.07.003>.
79. Gay NR, Gludemans M, Antonio ML, Abell NS, Balliu B, Park Y, et al. Impact of admixture and ancestry on eQTL analysis and GWAS colocalization in GTEx. *Genome Biol*. 2020;21(1):233. <https://doi.org/10.1186/s13059-020-02113-0>.
80. Tang MX, Cross P, Andrews H, Jacobs DM, Small S, Bell K, et al. Incidence of AD in African-Americans, Caribbean Hispanics, and Caucasians in northern Manhattan. *Neurology*. 2001;56(1):49–56. <https://doi.org/10.1212/WNL.56.1.49>.
81. Reitz C, Mayeux R. Alzheimer disease: epidemiology, diagnostic criteria, risk factors and biomarkers. *Biochem Pharmacol*. 2014;88(4):640–51. <https://doi.org/10.1016/j.bcp.2013.12.024>.
82. Vega IE, Cabrera LY, Wygant CM, Velez-Ortiz D, Counts SE. Alzheimer's disease in the Latino community: intersection of genetics and social determinants of health. *J Alzheimers Dis*. 2017;58(4):979–92. <https://doi.org/10.3233/JAD-161261>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

