HGG Advances

ARTICLE

The genetic architecture of Alzheimer disease risk in the Ohio and Indiana Amish

Michael D. Osterman,[1,2,*] Yeunjoo E. Song,[1] Larry D. Adams,[3] Renee A. Laux,[1] Laura J. Caywood,[3] Michael B. Prough,[3] Jason E. Clouse,[3] Sharlene D. Herington,[3] Susan H. Slifer,[3] Audrey Lynn,[1] M. Denise Fuzzell,[1] Sarada L. Fuzzell,[1] Sherri D. Hochstetler,[1] Kristy Miskimen,[1] Leighanne R. Main,[2,4] Daniel A. Dorfsman,[3,5] Paula Ogrocki,[6,7] Alan J. Lerner,[6,7] Jairo Ramos,[3] Jeffery M. Vance,[3,5] Michael L. Cuccaro,[3,5] William K. Scott,[3,5] Margaret A. Pericak-Vance,[3,5] and Jonathan L. Haines[1,2,*]

Summary

Alzheimer disease (AD) is the most common type of dementia and is currently estimated to affect 6.2 million Americans. It ranks as the sixth leading cause of death in the United States, and the proportion of deaths due to AD has been increasing since 2000, while the proportion of many other leading causes of deaths have decreased or remained constant. The risk for AD is multifactorial, including genetic and environmental risk factors. Although *APOE ε4* remains the largest genetic risk factor for AD, more than 26 other loci have been associated with AD risk. Here, we recruited Amish adults from Ohio and Indiana to investigate AD risk and protective genetic effects. As a founder population that typically practices endogamy, variants that are rare in the general population may be of a higher frequency in the Amish population. Since the Amish have a slightly lower incidence and later age of onset of disease, they represent an excellent and unique population for research on protective genetic variants. We compared AD risk in the Amish and to a non-Amish population through *APOE* genotype, a non-*APOE* genetic risk score of genome-wide significant variants, and a non-*APOE* polygenic risk score considering all of the variants. Our results highlight the lesser relative impact of *APOE* and differing genetic architecture of AD risk in the Amish compared to a non-Amish, general European ancestry population.

Introduction

Alzheimer disease (AD), the most common type of dementia, is the sixth leading cause of death in the United States and occurs in over 35% of individuals age 85 and older.[1,2] It is currently estimated that 6.2 million Americans are living with AD.[1] Deaths attributable to AD increased by 146.2% from 2000–2018, whereas other leading causes of death remained constant or decreased.[1] This burden of AD is expected to increase due to increased longevity and decreased fertility, known as population aging.[1,3,4] The cost of managing AD will continue to increase, with an expected annual global cost surpassing $50 billion by 2050.[1,5,6] People living with AD also suffer from severe degradation of their quality of life, including reduced independence and being at higher risk of somatic and psychiatric comorbidities.[7–9] Improved understanding of AD risk and subsequent improvements to screening, prediction, and prevention efforts are needed to reduce these burdens. As current medications only marginally and temporarily delay the progression and lessen the severity of AD, its growing prevalence serves as an imperative issue.

Risk for AD is multifactorial, including genetic and environmental risk factors.[8,10,11] While only 2%–5% of all cases of AD are strongly familial (e.g., resulting from high penetrance mutations),[12] the overall heritability of late-onset AD is estimated to be as high as 70% based on twin studies and genome-wide association studies (GWASs); however, such estimates can vary by population and environment.[13–15] Genetic risk for AD is complex, including more than 26 independent associated loci spanning diverse population groups.[16–18] Despite this large number of loci, the currently confirmed loci associated with AD risk account for only a small proportion of the overall heritability of AD.[15,16] Increased sample sizes and diversity of study populations will help GWASs to elucidate the remainder of the heritability.

The largest genetic influence on late-onset AD is conferred by the apolipoprotein E (*APOE*) locus[19] on chromosome 19. The *APOE ε4* allele confers 3- to 15-fold increased risk for those holding either 1 or 2 copies of the risk allele compared to those holding no risk alleles, while the *APOE ε2* allele confers significant protection from AD.[19–21] This association between AD and *APOE* has been replicated across many different and diverse populations.[22–24]

One such population is the Amish, who are descendants of German and Swiss Anabaptist immigrants who settled

[1]Department of Population and Quantitative Health Sciences, Case Western Reserve University School of Medicine, Cleveland, OH, USA; [2]Cleveland Institute for Computational Biology, Case Western Reserve University, Cleveland, OH, USA; [3]John P. Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami, FL, USA; [4]Department of Genetics and Genome Sciences, Case Western Reserve University School of Medicine, Cleveland, OH, USA; [5]Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami Miller School of Medicine, Miami, FL, USA; [6]Department of Neurology, University Hospitals Cleveland Medical Center, Cleveland, OH, USA; [7]Department of Neurology, Case Western Reserve University School of Medicine, Cleveland, OH, USA

*Correspondence: mdo17@case.edu (M.D.O.), jlh213@case.edu (J.L.H.)

https://doi.org/10.1016/j.xhgg.2022.100114.

in the United States during the 18th and 19th centuries. Communities currently living in Holmes County, Ohio and Elkhart and LaGrange Counties, Indiana are mostly descendants from the German Palatinate, while the communities in Adams County, Indiana largely descend from Swiss Anabaptist immigrants.[25–27] Subsequent cultural and religious isolation has restricted the introduction of new genetic variation,[25,28] leading the genetic variation within the Amish to be derived from the more general European gene pool. Due to the presence of genetic drift and the Amish practice of endogamy, it is possible that Amish allele frequencies and non-Amish, European-ancestry allele frequencies may be quite different.[29] It is for these reasons that the Amish can serve as an ideal candidate for genetic research. Study of the Amish can allow for detection and consideration of effects that may not otherwise be captured in studies of the general population.[30] The combination of these factors provides an ideal situation for the investigation of genetic variation that influences complex traits, including AD.

The Amish have a unique etiology of AD, as a slightly lower prevalence of AD has been reported within Amish populations, even after accounting for the effect of a lower frequency of the *APOE ε4* risk allele.[31–33] An improved understanding of what protective or other risk-bearing variants the Amish may be enriched for could prove helpful in improving the general understanding of the genetic risk of AD.

We recruited adults from Amish families living in Holmes County, Ohio and Elkhart, LaGrange, and Adams Counties, Indiana for studies of dementia. Our current focus is recruiting individuals who are cognitively unimpaired (CU) relative to age-normed benchmarks but at elevated risk for developing AD based on family history. We characterized this population and compared it with a non-Amish European-ancestry population living in the US for age, *APOE* genotype, and both a genetic risk score (GRS) using genome-wide significant variants from the Jansen et al. (2019)[17] genome-wide meta-analysis and a polygenic risk score (PRS) spanning the entire genome. A PRS typically offers additional predictive ability due to the inclusion of loci of small effect or loci that do not reach genome-wide significance criteria.[34–36]

Material and methods

Subjects

Individuals included in this study have been recruited over the past 20 years for multiple studies of AD or dementia,[31,37–39] age-related macular degeneration,[40–42] and successful aging.[43–45] For all of these studies, the primary criteria for enrollment included being age 50 or older, being part of the Amish community, and being of Amish descent. Recruitment primarily included community-based home visits. All of the individuals were screened for cognitive status. For the present study, individuals were included if they were CU based on cognitive screening and were age 75 and older. For our dementia studies, we prioritized the inclusion of individuals with at least one family member with possible or probable AD. Participants were recruited from Amish families living in Holmes County, Ohio and Elkhart, LaGrange, and Adams Counties, Indiana. Research complied with the Health Insurance Portability and Accountability Act and the Declaration of Helsinki. Informed consent was obtained from participants and the study was approved by the appropriate institutional review board (IRB).

Cognitive screening

At the time of enrollment, individuals were cognitively screened using a combination of the Modified Mini-Mental State (3MS) education-adjusted examination (all individuals),[46] the AD8 Dementia Screening Interview,[47] the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) Word List Memory Task,[48] and the Trail Making test (for the dementia studies).[49] Individuals were classified as CU or cognitively impaired (CI) based on established cutoffs.[46,49] Individuals initially classified as CI were further evaluated by a clinical adjudication board, comprising neurologists and neuropsychologists, to further classify them as having mild cognitive impairment (MCI), AD, cognitive impairment with no dementia (CIND), or having an unclear status. Based on these screening results, individuals were ultimately classified as unaffected if they were CU or affected if they were considered to have AD or other dementia. Individuals classified as CI but not having AD or other dementia were excluded from analysis.

Genotyping

At the time of enrollment, 30 mL of blood were collected from all of the participants for use in direct DNA extraction and storage of plasma. Genotype data were collected using an Illumina Expanded Multi-Ethnic Genotyping Array[50] with custom content (MEGAex+3k) or an Illumina Global Screening Array[51] (GSA). The MEGAex chip includes over 2 million markers, whereas the GSA chip includes a base quantity of 660,000 markers. When performing chip genotyping, we also included customized content of up to 6,000 variants to the MEGAex chip, including over 1,100 novel variants that have already been identified from our previous Amish whole-exome sequencing (WES) and whole-genome sequencing (WGS) studies and other associated variants from GWAS and the National Institute on Aging's Alzheimer's Disease Sequencing Project[52,53] (ADSP) studies that are not already on the chip. After genotype data were attained, imputation was performed based on a Haplotype Reference Consortium (HRC) panel.[54,55] We investigated genetic relationships of individuals within the overall study population by calculating kinship coefficients using KING version 2.26.[56] Furthermore, we compared the average genetic relationship across subpopulations based on recruitment site and cognitive status.

Quality control and imputation

Quality control (QC) was performed independently on the MEGAex+3k and GSA genotyping chip sets before merging. For SNPs, this included removal for excess genotype missingness, exclusion of monomorphic and duplicate SNPs, severe deviations from Hardy-Weinberg equilibrium among common SNPs, and Mendelian errors. For samples, this included removal for inconsistent genetic and self-reported sex, and overall genotyping completeness $< 5\%$. Imputation was performed using the Michigan Imputation Server[55] and the HRC reference set. The MEGAex+3k and GSA datasets were imputed separately, and each was submitted using the GRCh38 build for autosomes and

hg19 for the X chromosome. The reference population for HRC was European and the phasing was done using the Eagle option. Each dataset underwent QC separately after imputation (as described above and in the supplemental methods), including filtering by INFO score with a separate threshold for common and rare (minor allele frequency [MAF] < 0.01) SNPs. The two separately imputed sets were then merged after QC into one set of 2,096 samples using overlapping SNPs contained in both. The final imputed dataset contained 8,311,803 SNPs. Of these, 759,280 were rare and the remaining 7,552,523 were common. Full details regarding the QC and imputation process can be found in the supplemental information.

Non-Amish comparison group

We compared the Amish population to an existing source of non-Amish, European-ancestry individuals living within the United States. The ascertainment for this population has been described elsewhere.[57,58] Briefly, it included individuals ascertained by the University of Miami John P. Hussman Institute for Human Genomics and the Case Western Reserve University Department of Population & Quantitative Health Sciences. Recruitment was primarily through clinics, but also included home visits. After standard quality control, a total of 2,470 adults were included, with an approximate 1:1 case-control ratio. Case status was based on clinical diagnosis and was confirmed by autopsy when possible. These included using the same neurocognitive battery of tests as in the Amish. Diagnoses were further evaluated by two independent neurologists in the absence of clinical or autopsy diagnosis. The first onset of symptoms reported by the patient, his or her informant, or otherwise extracted from medical records was recorded as age at onset. Control subjects were considered from those who received a score of 27 or higher on a 3MS examination and were at least 60 years of age. Other phenotype information includes sex, age of examination, and age of onset in cases.

Comparisons in genetic risk of AD

The Amish population and comparison group were initially compared for distributions of sex, age, and cognitive status. Comparisons by genetic risk loci were performed in subsets of the combined Amish and non-Amish data after exclusion of individuals younger than age 75 years old to account for the age-related incidence of AD[32,59,60] in addition to differences in age distribution between the Amish data and the non-Amish comparison data. Only SNPs that passed QC and had information available after imputation in both the Amish and non-Amish groups were considered for subsequent analysis. Dosage information was considered for imputed SNPs with an INFO score of 0.9 or greater.

A GRS was generated using 22 genome-wide significant variants (Table 1), excluding *APOE* variants, as reported in the recent Jansen et al. (2019)[17] genetic meta-analysis. The GRS was constructed using PRSice-2[61] and goodness of fit was assessed in R version 3.5.1.[62] For ease of interpretation, the mean and standard deviation of the GRS were scaled to zero and one, respectively.

A non-*APOE* PRS was generated using a pruning and thresholding approach in PRSice-2[61] and the best-fit PRS model, in terms of correlation coefficient R^2, across the combined Amish and non-Amish dataset was used. All SNPs from the Jansen et al. (2019)[17] meta-analysis were included for PRS construction, except for those within 500 kb of either main *APOE* SNP (rs429358 and rs7412). The parameters for clumping in the construction of PRSs included a 500-kb window centered on each index SNP and an r^2 threshold of 0.1. A PRS was calculated and fit for each threshold beginning at $p = 5 \times 10^{-8}$ and increasing in steps of 5×10^{-5} until the p value threshold of 0.5 was reached. Then, an additional PRS was calculated using a p value threshold of 1.0 (all variants after clumping). A best-fit PRS was chosen in combined data after applying across different potential p value thresholds of included index SNPs. For ease of interpretation, the mean and standard deviation of the PRS were scaled to zero and one, respectively.

Distributions of the GRS and PRS were compared across the populations and by AD or other dementia case status. GRS and PRS models were compared with an *APOE*-only model, a covariate-only (sex and age) model, and a combined *APOE* and covariate model. *APOE* genotype was considered as a count of *ε2* and *ε4* alleles. Interaction terms including either *APOE ε2* and *ε4* allele and Amish group membership were created in a combined Amish and non-Amish analysis to investigate the potential for differential effects of *APOE* within the Amish. Additional models were constructed including GRS and PRS to investigate the overall predictive ability of the risk scores with and without the presence of the other variables. The predictive value of the constructed models was assessed by area under the receiver operating characteristic (ROC) curve (AUC).

Results

Characteristics of study population

After quality control and assurance, the genotype information of 2,096 Amish individuals was available for analysis. The mean and median ages of the population were 75.17 and 79 years, respectively, with a range of 21–110 years old (Figure S1). Of these, 1,965 had a cognitive examination performed. This dataset included 1,146 females and 819 males (Table S1). A total of 1,367 were classified after consensus expert review as CU, 385 were CI, 18 had MCI, and 326 were unclear or missing. Among the 385 with CI, 152 individuals (7.3% of the total sample) were considered to have possible or probable AD or other dementia. The remainder were considered as having CIND.

The non-Amish study population originally contained 2,470 individuals with 1,449 females and 1,021 males (Table S1). The mean and median ages of the non-Amish participants were 75.09 and 75 years (range: 60–100 years old; Figure S1). This included 1,126 CU controls and 1,177 cases with AD or other dementia.

After exclusion of individuals younger than age 75 years (Figure S2) and those not classified as affected or unaffected, 1,091 Amish participants remained. The mean and median ages of this Amish subset were 82.97 and 82 years, respectively (Table 2). This subgroup of the Amish contained 954 unaffected (CU) individuals and 137 affected (AD or other dementia) individuals. The mean and median age of the non-Amish group after exclusion of individuals younger than age 75 years were 80.83 and 80 (Figure S2). This non-Amish subgroup contained 416 CU controls and 544 AD or other dementia cases.

After exclusion of individuals younger than age 75 years, a lower prevalence of *APOE ε4* alleles and a higher prevalence of *ε2* alleles were observed in affected Amish

Table 1. List of SNPs included in non-*APOE* genetic risk score based on genome-wide significant variants

Chr	BP (GRCh37)	SNP	Nearest gene	A1	A2	Effect estimate (β)	MAF-affected Amish	MAF-CU Amish	MAF-non-Amish cases	MAF-non-Amish controls
1	161155392	rs4575098	*ADAMTS4*	A	G	0.016	0.248	0.283	0.230	0.255
1	207796828	rs2093760	*CR1*	A	G	0.024	0.259	0.238	0.202	0.181
2	127891427	rs4663105	*BIN1*	C	A	0.031	0.423	0.394	0.447	0.374
4	11026028	rs6448453	*CLNK*	A	G	0.015	0.376	0.327	0.291	0.303
4	117232235	rs7657553	*HS3ST1*	A	G	0.005	0.307	0.331	0.282	0.261
6	47432637	rs9381563	*CD2AP*	C	T	0.014	0.416	0.434	0.373	0.353
7	99971834	rs1859788	*ZCWPW1*	A	G	−0.018	0.307	0.294	0.340	0.351
7	143108158	rs7810606	*EPHA1*	C	T	−0.015	0.369	0.374	0.478	0.480
8	27464929	rs4236673	*CLU/PTK2B*	A	G	−0.020	0.347	0.386	0.364	0.397
10	11717397	rs11257238	*ECHDC3*	C	T	0.013	0.339	0.362	0.366	0.350
11	59958380	rs2081545	*MS4A6A*	A	C	−0.018	0.387	0.451	0.393	0.382
11	85776544	rs867611	*PICALM*	G	A	−0.020	0.493	0.492	0.285	0.310
11	121435587	rs11218343	*SORL1*	C	T	−0.036	0.062	0.051	0.031	0.054
14	92938855	rs12590654	*SLC24A4*	A	G	−0.015	0.343	0.363	0.342	0.332
15	59022615	rs442495	*ADAM10*	C	T	−0.014	0.285	0.281	0.314	0.331
15	63569902	rs117618017	*APH1B*	T	C	0.018	0.153	0.146	0.074	0.052
16	31133100	rs59735493	*KAT8*	A	G	−0.013	0.219	0.195	0.299	0.287
17	5138980	rs113260531	*SCIMP*	A	G	0.020	0.099	0.109	0.124	0.121
17	47450775	rs28394864	*ABI3*	A	G	0.012	0.438	0.491	0.472	0.476
19	46241841	rs76320948	*AC074212.3*	T	C	0.035	0.055	0.040	0.027	0.019
19	51727962	rs3865444	*CD33*	A	C	−0.014	0.332	0.287	0.291	0.323
20	54998544	rs6014724	*CASS4*	G	A	−0.023	0.058	0.097	0.076	0.092

The genetic risk score was calculated as a sum of the product of the effect estimate, β, and the number of variants for each individual across each SNP. Effect estimates from Jansen et al. (2019)[17] were used. Chr, chromosome; BP, base pair; GRCh37, Genome Reference Consortium Build 37; A1, effect (minor) allele; A2, reference allele; MAF, minor allele frequency; CU, cognitively unimpaired for age-normed benchmarks.

individuals than in non-Amish individuals (Table 3). The unaffected Amish have a similar distribution of *APOE* genotype to that of the non-Amish controls, except for a lower prevalence of the *APOE ε2|ε3* genotype.

Relatedness of Amish study participants

The average kinship coefficient across all individuals in the Amish study population was 0.0037, which is equivalent to being related between third and fourth cousins. The average kinship coefficient across subpopulations stratified by primary study site and cognitive status were similar (Table S2).

GRS

The non-*APOE* GRS was constructed using effect estimates from 22 genome-wide significant SNPs from the Jansen et al. (2019)[17] genetic meta-analysis (Table 1). These 22 variants were chosen due to having variant information available in both the Amish and non-Amish populations. After non-*APOE* GRS construction, we observed, in general, less variance among the Amish GRS, regardless of affection status, than among the non-Amish comparison group (Figure 1). Although the mean GRS was greater for the affected Amish than in the unaffected Amish, this difference is not statistically significant ($p = 0.07$). The GRS was able to distinguish

Table 2. Demographic information and cognitive status of Amish study population and non-Amish comparison group after exclusion of individuals younger than age 75 years

Trait	Amish n (%)	Non-Amish n (%)
Female	636 (58.3)	563 (58.6)
CU	954 (87.4)	416 (43.3)
AD or other dementia	137 (12.6)	544 (56.7)

CU, cognitively unimpaired; AD, Alzheimer disease.

Table 3. *APOE* distribution by population and AD case status of individuals age 75 years and older with known *APOE* genotype

APOE genotype	Affected Amish n (%)	Non-Amish cases n (%)	p value	Unaffected Amish n (%)	Non-Amish controls n (%)	p value
ε2|ε2	0 (0.0)	0 (0.0)	–	1 (0.1)	3 (0.7)	0.052
ε2|ε3	14 (10.2)	21 (3.9)	0.003[a]	84 (9.0)	52 (12.7)	0.038[a]
ε2|ε4	2 (1.5)	13 (2.4)	0.503	15 (1.6)	12 (2.9)	0.112
ε3|ε3	70 (51.1)	201 (37.1)	0.003[a]	632 (67.4)	264 (64.2)	0.258
ε3|ε4	42 (30.7)	240 (44.3)	0.004[a]	195 (20.8)	78 (19.0)	0.447
ε4|ε4	9 (6.6)	67 (12.4)	0.055	11 (1.2)	2 (0.5)	0.234
Total *ε2*	16 (5.8)	34 (3.1)	–	101 (5.4)	70 (8.5)	–
Total *ε3*	196 (71.5)	663 (61.2)	–	1,543 (82.2)	658 (80.0)	–
Total *ε4*	62 (22.6)	387 (35.7)	–	232 (12.4)	94 (11.4)	–

A p value for two-sample population proportion *Z* score test in individuals age 75 years and older is provided for comparisons between affected Amish versus non-Amish cases in addition to unaffected Amish versus non-Amish controls.
[a]A significant difference in proportion at $\alpha = 0.05$.

between the non-Amish case group members and the non-Amish control group members ($p = 6.46 \times 10^{-5}$).

PRS

The best-fit non-*APOE* PRS was constructed using all SNPs (p value cutoff of 1.0) after clumping was applied. We observed that the values of the non-*APOE* PRS in the Amish individuals affected by AD or other dementia were modestly lower than in the non-Amish case group (Figure 2). The values of the PRS in the non-Amish controls were generally lower than that of the unaffected Amish. Overall, the difference in PRS values between the Amish affected and unaffected is much smaller than between the non-Amish case and control group members. The PRS was unable to distinguish between affection status in the Amish ($p = 0.38$), but it was able to distinguish between case status in the non-Amish population ($p < 2.2 \times 10^{-16}$).

Regression models

We evaluated the association of the GRS, PRS, sex, age, and *APOE* allele count with the primary outcome of AD or other dementia by building a series of logistic regression models, after stratification by source population (Table 4). Age was associated ($p < 0.05$) with the primary outcome across all models in both source populations. Sex was associated ($p < 0.05$) with AD or other dementia in the non-Amish models but not the Amish models.

We found that in univariate models, the *APOE ε4* allele count was associated ($p = 2.8 \times 10^{-6}$) with affection status in the Amish but that *APOE ε2* allele count was not ($p = 0.45$), whereas both *APOE ε2* ($p < 1.7 \times 10^{-4}$) and *APOE ε4* allele ($p < 2.2 \times 10^{-16}$) counts were associated with case status in the non-Amish (Table S3). A combined Amish and non-Amish analysis yielded that both an *APOE ε2* × Amish ($p = 3.2 \times 10^{-4}$) and an *APOE ε4* × Amish ($p = 8.9 \times 10^{-10}$) interaction term were associated with affection/case status in a model also including sex and age covariates. The implied *APOE ε2* odds ratio (OR) from this unstratified analysis model was 1.02 in the Amish and 0.68 in the non-Amish. The implied *APOE ε4* OR was 1.12 in the Amish and 1.35 in the non-Amish. In the stratified multivariate models, there is increased risk associated with the *APOE ε4* allele count across both populations but only significant protective effects of *APOE ε2* allele count within the non-Amish population (Table 4). We identified similar findings when investigating *APOE* effects as a factor of *APOE* genotype (Table S4).

The GRS and PRS were associated ($p < 0.05$) with the primary outcome across all of the models tested, including PRS in the non-Amish populations (Table 4). However, the GRS and PRS were not significantly associated with the primary outcome in the Amish population.

We also evaluated goodness of fit through AUC across each of these models (Table 5). We determined that the AUC of the sex and age-only (covariate) model is larger in the Amish (0.69) than the non-Amish population (0.60). By contrast, we determined that the AUC for an *APOE* genotype-only model is larger in the non-Amish population (0.71) than in the Amish population (0.59). A higher AUC was observed for the GRS and PRS models in the non-Amish population than in the Amish population both in the GRS- and PRS-only models as well as in the full models with covariates.

Discussion

This study characterized and evaluated the genetic risk for AD in an Amish population and compared it to a non-Amish population of predominantly European ancestry. We demonstrated that there are inherent differences in the underlying genetic risk structure for AD in the Amish. These may be through either different or undetected loci compared to a more general European ancestry population. This warrants further investigation to elucidate pathways involved in AD risk because the Amish are a subpopulation of European immigrants that have practiced endogamy since arriving in the United States. Thus, they share some genetic risk with the general European ancestry population but also harbor unique genetic risk.

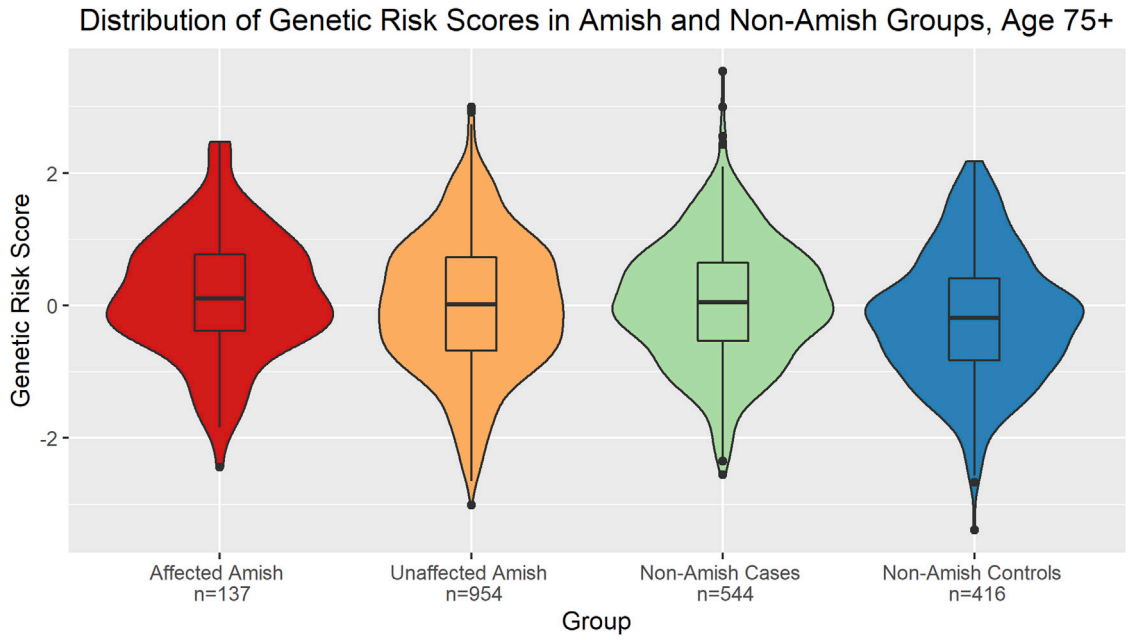

Figure 1. Violin plot and boxplot of distribution of non-*APOE* genetic risk scores by Amish and Alzheimer disease status
Genetic risk scores were constructed using only genome-wide significant SNPs, excluding *APOE* variants. Only individuals age 75 years or older were included. GRS was able to distinguish ($p = 6.46 \times 10^{-5}$) between the non-Amish case and non-Amish control group members in addition to the Amish affected and non-Amish control group members ($p = 1.79 \times 10^{-4}$) but not between the affected Amish and unaffected Amish individuals ($p = 0.072$).

The results demonstrate not only that there is less variation in *APOE* genotype within the Amish but also that the *APOE* genotype may not play as large of a role in the development of AD or other dementia as within a typical European ancestry population. Our results with this larger, updated dataset confirm prior findings that *APOE* has a smaller effect on AD risk in the Amish population than in a non-Amish population,[32] possibly due to the lower prevalence of *APOE ε4* in the Amish population. There is, however, evidence that the effect of *APOE* may differ between the Amish and non-Amish (Table 4). This is further supported by the statistical significance of the *APOE* allele count × Amish group membership interaction term in a model including sex and age covariates. Other studies

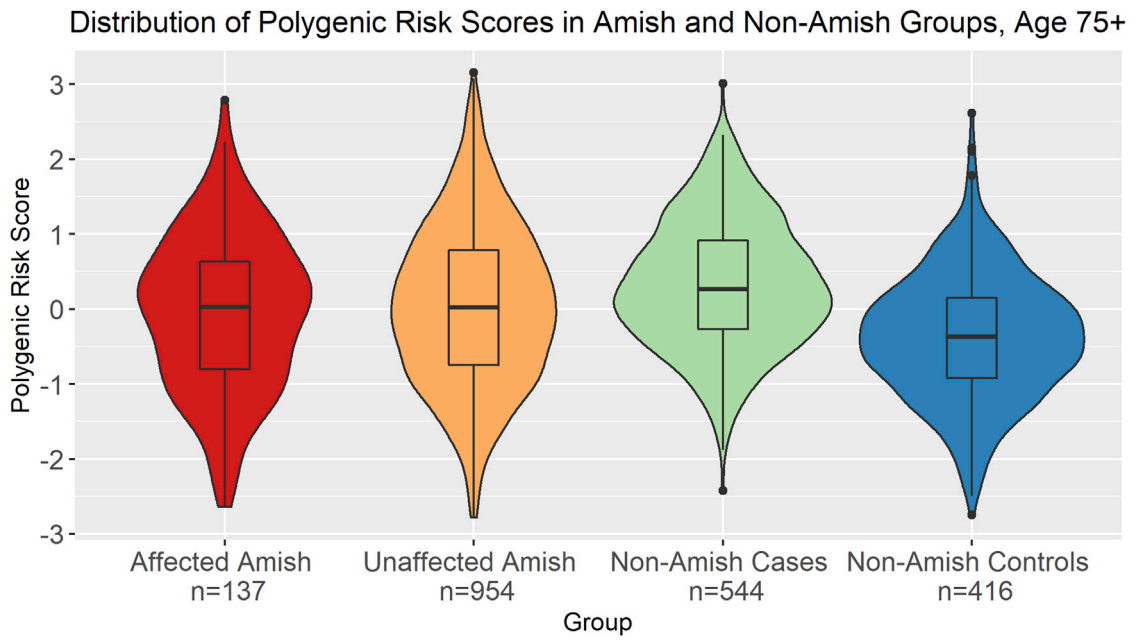

Figure 2. Violin plot and boxplot of distribution of non-*APOE* polygenic risk scores by Amish and Alzheimer disease status
Polygenic risk scores were constructed using a pruning and thresholding approach on all variants, excluding those within 500 kb of either *APOE* SNP. Only individuals age 75 years or older were included.

Table 4. Effect estimates of predictors of AD status across multivariate models with count of *APOE* alleles as covariate, separated by source population

Trait	Amish full GRS model (95% CI)	Non-Amish full GRS model (95% CI)	Amish full PRS model (95% CI)	Non-Amish full PRS model (95% CI)
Female sex	0.986 (0.949–1.025)	1.073 (1.014–1.136)[a]	0.987 (0.949–1.025)	1.059 (1.003–1.117)[a]
Age	1.019 (1.015–1.023)[a]	1.023 (1.017–1.030)[a]	1.019 (1.015–1.023)[a]	1.020 (1.014–1.026)[a]
APOE ε2 allele count	1.020 (0.960–1.084)	0.845 (0.774–0.922)[a]	1.016 (0.956–1.080)	0.845 (0.778–0.918)[a]
APOE ε4 allele count	1.114 (1.071–1.159)[a]	1.348 (1.289–1.410)[a]	1.115 (1.072–1.160)[a]	1.305 (1.250–1.362)[a]
GRS or PRS	1.015 (0.996–1.034)	1.052 (1.022–1.084)[a]	0.997 (0.979–1.025)	1.183 (1.149–1.218)[a]

Models include sex, age, count of *APOE ε2* and *ε4* alleles, and non-*APOE* genome-wide significant genetic risk score (GRS) or non-*APOE* polygenic risk score (PRS). Estimates are presented as odds ratios (OR = e^{β}). Effect size is considered for 1-year change in age, per copy of each *APOE* allele, and per 1 standard deviation change in GRS or PRS.
[a]Significance at $\alpha = 0.05$.

have found extensive evidence of the variance of *APOE* effect across ancestry groups and potential interactions of nearby genes with *APOE*.[18,21,63,64] As a founder population descending from a subpopulation of European ancestry, there is the potential to determine how the risk conferred by *APOE* and surrounding regions may differ from a more general non-Hispanic white population.

The non-*APOE* GRS and non-*APOE* PRS (Table 4) have only moderate predictive value on their own, but in addition to covariates, they do provide a meaningful increase in predictability in a logistic regression model for case/affection status. We determined that, based on a GRS of genome-wide significant SNPs from a recent meta-analysis of GWASs,[17] there exists more variation of genetic risk in a non-Amish population than in an Amish population. When extending to a PRS analysis, this phenomenon is much more prominent. The PRS model also added distinguishing ability in AD or other dementia status in the non-Amish population. We determined that a non-*APOE* GRS and a non-*APOE* PRS do not seem to differ greatly in their predictive ability of affection status in the Amish, suggesting that risk scores created using effect size weights derived from non-Amish European samples may not accurately predict risk in the Amish, especially among variants that do not meet the criteria for genome-wide significance in the European population. This is somewhat similar to previous findings[31] of GRSs that included *APOE* but highlights that *APOE* still plays an important role in AD prediction in the Amish. By extending these reuslts to consider the ability of the AD PRS to distinguish between states of cognitive impairment, we determined that there was a significant difference in mean PRS distribution between the CI and CU Amish (Figure S3).

In predicting the primary outcome of AD or other dementia, our results suggest that of the factors considered here, age is the most crucial risk factor in the Amish population, whereas *APOE* and PRS have greater importance in the non-Amish population. We observe much worse predictive ability when using a PRS that includes SNPs that do not meet genome-wide significance criteria when applied to the Amish population compared to the non-Amish population, suggesting that the underlying genetic architecture for AD risk is dissimilar to that of a general European ancestry population, especially among SNPs that do not meet the criteria for genome-wide significance in the non-Amish population. This result is also supported by an exploratory post hoc PRS analysis in only the Amish subjects in whom we found that the genome-wide significant threshold of $p = 5 \times 10^{-8}$ (the GRS threshold) was the best-fit non-*APOE* PRS after a pruning and thresholding approach similar to the main analysis.

The lower predictive ability in the Amish for GRS and PRS comprising known AD risk factors suggests that the genetic risk profile in the Amish is significantly different. When combining this with information that the Amish have a lower prevalence of cognitive impairment and dementia,[31,32,65] it may be possible that their genetic architecture is protective of these outcomes in a way that using risk estimates from a general European ancestry population cannot explain. Further study of population prevalence and non-genetic, behavioral, and environmental risk factors within Amish and non-Amish populations may help us to understand whether potential protection of the Amish from cognitive impairment is exclusively due to differences in genetic architecture or a combination of genetic architecture and other risk factors. Our results add to mounting evidence that there is genetic risk for AD in the Amish that is not captured by genetic risk scores derived from non-Amish populations.[31,32,37]

We conclude that there are evident differences in the genetic architecture for AD risk in the Amish compared to a non-Amish European ancestry population, especially in terms of *APOE* genotype frequency, PRS distribution, and their conferred risk. Future genomic studies including the Amish should consider using effect estimates from an Amish analysis to determine whether there are substantial differences in predictive ability than are seen after PRS construction using effect estimates from a non-Amish population. Identification of why the Amish appear to be relatively protected from AD and cognitive impairment, in general, warrants further study to determine whether the protection is granted by protective loci, differential effects of known loci, non-genetic lifestyle factors of the Amish, or a combination of these factors. Further study of this population will allow us to identify risk factors enriched in the

Table 5. Goodness of fit of predictive models by sex, age, *APOE* allele count, genetic risk score, and polygenic risk score

Group	COV	*APOE*	COV + *APOE*	GRS	GRS + *APOE*	GRS + COV	GRS + COV + *APOE*	PRS	PRS + *APOE*	PRS + COV	PRS + COV + *APOE*
Amish	0.69	0.59	0.74	0.54	0.61	0.70	0.75	0.52	0.60	0.69	0.74
Non-Amish	0.60	0.71	0.77	0.58	0.74	0.63	0.77	0.72	0.81	0.74	0.83

For each constructed logistic regression model, the area under the curve of a receiver operating characteristic curve is presented. The outcome of interest in each model is probable or confirmed AD or other dementia. COV, sex and age covariates; GRS, genetic risk score including only genome-wide significant single nucleotide variants, excluding *APOE* variants; PRS, polygenic risk score using a pruning and thresholding approach, excluding single nucleotide polymorphisms within 500 kb of *APOE* variants.

Amish that may enhance previously identified pathways[66,67] important in the development of AD and identify additional pathways or mechanisms that contribute to or protect against cognitive decline. By extending this cohort through new recruitment and longitudinal follow-up, the power to identify both novel risk and protective genetic loci and potential predictors of progression from normal cognition to AD will be increased. Additional studies such as these will allow for better detection of rare effects and better understanding of the differences in the genetic risk of AD between the Amish and non-Amish populations.

Data and code availability

Data are available through The National Institute on Aging Genetics of Alzheimer's Disease Data Storage Site (NIAGADS; https://www.niagads.org/). Code used in the analysis is available through GitHub at https://github.com/mdosterman/HGGA_Genetic_Architecture_Amish.

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.xhgg.2022.100114.

Acknowledgments

We thank the Amish families for their willing participation in our study. We used the Anabaptist Genealogy Database and Swiss Anabaptist Genealogy Association to help determine relationships. This study is supported by National Institutes of Health/ National Institute on Aging grant no. AG058066 (to J.L.H., M.A.P.-V., and W.K.S.). We acknowledge the resources provided by the Department of Population and Quantitative Health Sciences at the Case Western Reserve University School of Medicine and the John P. Hussman Institute for Human Genomics at the University of Miami Miller School of Medicine.

Declaration of interests

The authors declare no competing interests.

Received: December 20, 2021
Accepted: April 22, 2022

References

1. (2019). Alzheimer's disease facts and figures (2019). Alzheimers Dement. *15*, 321–387.
2. Hebert, L.E., Weuve, J., Scherr, P.A., and Evans, D.A. (2013). Alzheimer disease in the United States (2010-2050) estimated using the 2010 census. Neurology *80*, 1778–1783. https://doi.org/10.1212/wnl.0b013e31828726f5.
3. Anderson, G.F., and Hussey, P.S. (2000). Population aging: a comparison among industrialized countries. Health Aff. (Millwood). *19*, 191–203. https://doi.org/10.1377/hlthaff.19.3.191.
4. de Meijer, C., Wouterse, B., Polder, J., and Koopmanschap, M. (2013). The effect of population aging on health expenditure growth: a critical review. Eur. J. Ageing *10*, 353–361. https://doi.org/10.1007/s10433-013-0280-x.
5. Deb, A., Thornton, J.D., Sambamoorthi, U., and Innes, K. (2017). Direct and indirect cost of managing alzheimer's disease and related dementias in the United States. Expert Rev. Pharmacoecon. Outcomes Res. *17*, 189–202. https://doi.org/10.1080/14737167.2017.1313118.
6. Langa, K.M., Larson, E.B., Wallace, R.B., Fendrick, A.M., Foster, N.L., Kabeto, M.U., Weir, D.R., Willis, R.J., and Herzog, A.R. (2004). Out-of-Pocket health care expenditures among older Americans with dementia. Alzheimer Dis. Assoc. Disord. *18*, 90–98. https://doi.org/10.1097/01.wad.0000126620.73791.3e.
7. Martyr, A., Nelis, S.M., Quinn, C., Wu, Y.T., Lamont, R.A., Henderson, C., Clarke, R., Hindle, J.V., Thom, J.M., Jones, I.R., et al. (2018). Living well with dementia: a systematic review and correlational meta-analysis of factors associated with quality of life, well-being and life satisfaction in people with dementia. Psychol. Med. *48*, 2130–2139. https://doi.org/10.1017/s0033291718000405.
8. Medeiros, R., Newcombe, E.A., Huat, T.J., Valmas, N., Camats-Perna, J., and Silva, M.L. (2018). Inflammation: the link between comorbidities, genetics, and Alzheimer's disease. J. Neuroinflammation. *15*, 276. https://doi.org/10.1186/s12974-018-1313-3.
9. Doraiswamy, P.M., Leon, J., Cummings, J.L., Marin, D., and Neumann, P.J. (2002). Prevalence and impact of medical comorbidity in Alzheimer's disease. J. Gerontol. A Biol. Sci. Med. Sci. *57*, M173–M177. https://doi.org/10.1093/gerona/57.3.m173.
10. Xu, W., Tan, L., Wang, H.F., Jiang, T., Tan, M.S., Tan, L., Zhao, Q.F., Li, J.Q., Wang, J., and Yu, J.T. (2015). Meta-analysis of modifiable risk factors for Alzheimer's disease. J. Neurol. Neurosurg. Psychiatry *86*, 1299–1306. https://doi.org/10.1136/jnnp-2015-310548.
11. Hersi, M., Irvine, B., Gupta, P., Gomes, J., Birkett, N., and Krewski, D. (2017). Risk factors associated with the onset and progression of Alzheimer's disease: a systematic review of the evidence. Neurotoxicology *61*, 143–187. https://doi.org/10.1016/j.neuro.2017.03.006.
12. Williamson, J., Goldman, J., and Marder, K.S. (2009). Genetic aspects of alzheimer disease. Neurologist *15*, 80–86.

13. Gatz, M., Pedersen, N.L., Berg, S., Johansson, B., Johansson, K., Mortimer, J.A., Posner, S.F., Viitanen, M., Winblad, B., and Ahlbom, A. (1997). Heritability for Alzheimer's disease: the study of dementia in Swedish twins. J. Gerontol. A Biol. Sci. Med. Sci. *52*, 117–125. https://doi.org/10.1093/gerona/52a.2.m117.
14. Cuyvers, E., and Sleegers, K. (2016). Genetic variations underlying Alzheimer's disease: evidence from genome-wide association studies and beyond. Lancet Neurol. *15*, 857–868. https://doi.org/10.1016/s1474-4422(16)00127-7.
15. Ridge, P.G., Mukherjee, S., Crane, P.K., and Kauwe, J.S.K. (2013). Alzheimer's disease: analyzing the missing heritability. PLoS One *8*, e79771. https://doi.org/10.1371/journal.pone.0079771.
16. Kunkle, B.W., Grenier-Boley, B., Sims, R., Bis, J.C., Damotte, V., Naj, A.C., Boland, A., Vronskaya, M., van der Lee, S.J., Amlie-Wolf, A., et al. (2019). Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. Nat. Genet. *51*, 414–430. https://doi.org/10.1038/s41588-019-0358-2.
17. Jansen, I.E., Savage, J.E., Watanabe, K., Bryois, J., Williams, D.M., Steinberg, S., Sealock, J., Karlsson, I.K., Hägg, S., Athanasiu, L., et al. (2019). Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. Nat. Genet. *51*, 404–413. https://doi.org/10.1038/s41588-018-0311-9.
18. Kunkle, B.W., Schmidt, M., Klein, H.U., Naj, A.C., Hamilton-Nelson, K.L., Larson, E.B., Evans, D.A., De Jager, P.L., Crane, P.K., Buxbaum, J.D., et al.; Writing Group for the Alzheimer's Disease Genetics Consortium ADGC (2021). Novel alzheimer disease risk loci and pathways in african American individuals using the african genome resources panel: a meta-analysis. JAMA Neurol. *78*, 102–113. https://doi.org/10.1001/jamaneurol.2020.3536.
19. Corder, E.H., Saunders, A.M., Strittmatter, W.J., Schmechel, D.E., Gaskell, P.C., Small, G.W., Roses, A.D., Haines, J.L., and Pericak-Vance, M.A. (1993). Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science *261*, 921–923. https://doi.org/10.1126/science.8346443.
20. Corder, E.H., Saunders, A.M., Risch, N.J., Strittmatter, W.J., Schmechel, D.E., Gaskell, P.C., Rimmler, J.B., Locke, P.A., Conneally, P.M., Schmader, K.E., et al. (1994). Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. Nat. Genet. *7*, 180–184. https://doi.org/10.1038/ng0694-180.
21. Rajabli, F., Feliciano, B.E., Celis, K., Hamilton-Nelson, K.L., Whitehead, P.L., Adams, L.D., Bussies, P.L., Manrique, C.P., Rodriguez, A., Rodriguez, V., et al. (2018). Ancestral origin of ApoE ε4 Alzheimer disease risk in Puerto Rican and African American populations. PLoS Genet. *14*, e1007791. https://doi.org/10.1371/journal.pgen.1007791.
22. Liu, Y., Yu, J.T., Wang, H.F., Han, P.R., Tan, C.C., Wang, C., Meng, X.F., Risacher, S.L., Saykin, A.J., and Tan, L. (2015). APOE genotype and neuroimaging markers of Alzheimer's disease: systematic review and meta-analysis. J. Neurol. Neurosurg. Psychiatry *86*, 127–134. https://doi.org/10.1136/jnnp-2014-307719.
23. O'Donoghue, M.C., Murphy, S.E., Zamboni, G., Nobre, A.C., and Mackay, C.E. (2018). APOE genotype and cognition in healthy individuals at risk of Alzheimer's disease: a review. Cortex *104*, 103–123. https://doi.org/10.1016/j.cortex.2018.03.025.
24. Tzioras, M., Davies, C., Newman, A., Jackson, R., and Spires-Jones, T. (2019). Invited Review: APOE at the interface of inflammation, neurodegeneration and pathological protein spread in Alzheimer's disease. Neuropathol. Appl. Neurobiol. *45*, 327–346. https://doi.org/10.1111/nan.12529.
25. Hostetler, J.A. (1993). Amish Society (JHU Press).
26. Jackson, C.E., Symon, W.E., Pruden, E.L., Kaehr, I.M., and Mann, J.D. (1968). Consanguinity and blood group distribution in an Amish Isolate. Am. J. Hum. Genet. *20*, 522–527.
27. Agarwala, R., Biesecker, L.G., Tomlin, J.F., and Schaffer, A.A. (1999). Towards a complete North American Anabaptist genealogy: a systematic approach to merging partially overlapping genealogy resources. Am. J. Med. Genet. *86*, 156–161. https://doi.org/10.1002/(sici)1096-8628(19990910)86:2<156::aid-ajmg13>3.0.co;2-5.
28. Kraybill, D.B. (2001). The Riddle of Amish Culture (JHU Press).
29. Hou, L., Faraci, G., Chen, D.T.W., Kassem, L., Schulze, T.G., Shugart, Y.Y., and McMahon, F.J. (2013). Amish revisited: next-generation sequencing studies of psychiatric disorders among the Plain people. Trends Genet. *29*, 412–418. https://doi.org/10.1016/j.tig.2013.01.007.
30. Hatzikotoulas, K., Gilly, A., and Zeggini, E. (2014). Using population isolates in genetic association studies. Brief. Funct. Genomics *13*, 371–377. https://doi.org/10.1093/bfgp/elu022.
31. D'Aoust, L.N., Cummings, A.C., Laux, R., Fuzzell, D., Caywood, L., Reinhart-Mercer, L., Scott, W.K., Pericak-Vance, M.A., and Haines, J.L. (2015). Examination of candidate exonic variants for association to alzheimer disease in the Amish. PLoS One *10*, e0118043. https://doi.org/10.1371/journal.pone.0118043.
32. Pericak-Vance, M.A., Johnson, C.C., Rimmler, J.B., Saunders, A.M., Robinson, L.C., D'Hondr, E.G., Jackson, C.E., and Haines, J.L. (1996). Alzheimer's disease and apolipoprotein e-4 allele in an amish population. Ann. Neurol. *39*, 700–704. https://doi.org/10.1002/ana.410390605.
33. Johnson, C.C., Rybicki, B.A., Brown, G., and Jackson, C.E. (1993). Prevalence of dementia in the Amish: a three county survey. Am. J. Epidemiol. *138*.
34. Igo, R.P., Kinzy, T.G., and Cooke Bailey, J.N. (2019). Genetic risk scores. Curr. Protoc. Hum. Genet. *104*, e95. https://doi.org/10.1002/cphg.95.
35. Osterman, M.D., Kinzy, T.G., and Bailey, J.N.C. (2021). Polygenic risk scores. Curr. Protoc. *1*, e126. https://doi.org/10.1002/cpz1.126.
36. Choi, S.W., Mak, T.S.H., and O'Reilly, P.F. (2020). Tutorial: a guide to performing polygenic risk score analyses. Nat. Protoc. *15*, 2759–2772. https://doi.org/10.1038/s41596-020-0353-1.
37. Cummings, A.C., Jiang, L., Velez Edwards, D.R., McCauley, J.L., Laux, R., McFarland, L.L., Fuzzell, D., Knebusch, C., Caywood, L., Reinhart-Mercer, L., et al. (2012). Genome-wide association and linkage study in the amish detects a novel candidate late-onset alzheimer disease gene. Ann. Hum. Genet. *76*, 342–351. https://doi.org/10.1111/j.1469-1809.2012.00721.x.
38. Ashley-Koch, A.E., Shao, Y., Rimmler, J.B., Gaskell, P.C., Welsh-Bohmer, K.A., Jackson, C.E., Scott, W.K., Haines, J.L., and Pericak-Vance, M.A. (2005). An autosomal genomic screen for dementia in an extended Amish family. Neurosci. Lett. *379*, 199–204. https://doi.org/10.1016/j.neulet.2004.12.065.
39. van der Walt, J.M., Scott, W.K., Slifer, S., Gaskell, P.C., Martin, E.R., Welsh-Bohmer, K., Creason, M., Crunk, A., Fuzzell, D., McFarland, L., et al. (2005). Maternal lineages and alzheimer disease risk in the old order amish. Hum. Genet. *118*, 115–122. https://doi.org/10.1007/s00439-005-0032-x.
40. Hoffman, J.D., Cooke Bailey, J.N., D'Aoust, L., Cade, W., Ayala-Haedo, J., Fuzzell, D., Laux, R., Adams, L.D., Reinhart-Mercer, L., Caywood, L., et al. (2014). Rare complement factor H

variant associated with age-related macular degeneration in the Amish. Investig. Ophthalmol. Vis. Sci. *55*, 4455–4460. https://doi.org/10.1167/iovs.13-13684.

41. Waksmunski, A.R., Igo, R.P., Song, Y.E., Cooke Bailey, J.N., Laux, R., Fuzzell, D., Fuzzell, S., Adams, L.D., Caywood, L., Prough, M., et al. (2019). Rare variants and loci for age-related macular degeneration in the Ohio and Indiana Amish. Hum. Genet. *138*, 1171–1182. https://doi.org/10.1007/s00439-019-02050-4.
42. Sardell, R.J., Nittala, M.G., Adams, L.D., Laux, R.A., Cooke Bailey, J.N., Fuzzell, D., Fuzzell, S., Reinhart-Mercer, L., Caywood, L.J., Horst, V., et al. (2016). Heritability of choroidal thickness in the amish. Ophthalmology *123*, 2537–2544. https://doi.org/10.1016/j.ophtha.2016.09.001.
43. Courtenay, M.D., Gilbert, J.R., Jiang, L., Cummings, A.C., Gallins, P.J., Caywood, L., Reinhart-Mercer, L., Fuzzell, D., Knebusch, C., Laux, R., et al. Mitochondrial Haplogroup X is associated with successful aging in the Amish.Hum. Genet. 131 201-208
44. Edwards, D.R.V., Gilbert, J.R., Hicks, J.E., Myers, J.L., Jiang, L., Cummings, A.C., Guo, S., Gallins, P.J., Konidari, I., Caywood, L., et al. (2013). Linkage and association of successful aging to the 6q25 region in large Amish kindreds. Age (Omaha) *35*, 1467–1477. https://doi.org/10.1007/s11357-012-9447-1.
45. Edwards, D.R.V., Gilbert, J.R., Jiang, L., Gallins, P.J., Caywood, L., Creason, M., Fuzzell, D., Knebusch, C., Jackson, C.E., Pericak-Vance, M.A., et al. (2011). Successful aging shows linkage to chromosomes 6, 7, and 14 in the amish. Ann. Hum. Genet. *75*, 516–528. https://doi.org/10.1111/j.1469-1809.2011.00658.x.
46. Chui, T.E.L. (1987). Modified Mini mental state examination (3MS). J. Clin. Psychiatry *48*, 314–318.
47. Galvin, J.E., Roe, C.M., and Morris, J.C. (2007). Evaluation of cognitive impairment in older adults: combining brief informant and performance measures. Arch. Neurol. *64*, 718–724. https://doi.org/10.1001/archneur.64.5.718.
48. Welsh, K.A., Butters, N., Mohs, R.C., Beekly, D., Edland, S., Fillenbaum, G., and Heyman, A. (1994). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part V. A normative study of the neuropsychological battery. Neurology *44*, 609. https://doi.org/10.1212/wnl.44.4.609.
49. Tombaugh, T.N. (2004). Trail Making Test A and B: normative data stratified by age and education. Arch. Clin. Neuropsychol. *19*, 203–214. https://doi.org/10.1016/s0887-6177(03)00039-8.
50. Multi-Ethnic Genotyping Array Consortium. https://www.illumina.com/science/consortia/human-consortia/multi-ethnic-genotyping-consortium.html.
51. Global Screening Array Consortium. https://www.illumina.com/science/consortia/human-consortia/global-screening-consortium.html.
52. Beecham, G.W., Bis, J.C., Martin, E.R., Choi, S.H., DeStefano, A.L., Van Duijn, C.M., Fornage, M., Gabriel, S.B., Koboldt, D.C., Larson, D.E., et al. (2017). The alzheimer's disease sequencing Project: study design and sample selection. Neurol. Genet. *3*, e194. https://doi.org/10.1212/nxg.0000000000000194.
53. Kuzma, A., Valladares, O., Cweibel, R., Greenfest-Allen, E., Childress, D.M., Malamon, J., Gangadharan, P., Zhao, Y., Qu, L., Leung, Y.Y., et al. (2016). NIAGADS: the NIA genetics of alzheimer's disease data storage site. Alzheimers Dement. *17*, e052258.
54. Loh, P.-R., Danecek, P., Palamara, P.F., Fuchsberger, C., A Reshef, Y., K Finucane, H., Schoenherr, S., Forer, L., Mccarthy, S., Abecasis, G.R., et al. (2016). Reference-based phasing using the Haplotype reference Consortium panel. Nat. Genet. *48*, 1443–1448. https://doi.org/10.1038/ng.3679.
55. Das, S., Forer, L., Schönherr, S., Sidore, C., Locke, A.E., Kwong, A., Vrieze, S.I., Chew, E.Y., Levy, S., McGue, M., et al. (2016). Next-generation genotype imputation service and methods. Nat. Genet. *48*, 1284–1287. https://doi.org/10.1038/ng.3656.
56. Manichaikul, A., Mychaleckyj, J.C., Rich, S.S., Daly, K., Sale, M., and Chen, W.M. (2010). Robust relationship inference in genome-wide association studies. Bioinformatics *26*, 2867–2873. https://doi.org/10.1093/bioinformatics/btq559.
57. Liang, X., Slifer, M., Martin, E.R., Schnetz-Boutaud, N., Bartlett, J., Anderson, B., Züchner, S., Gwirtsman, H., Gilbert, J.R., Pericak-Vance, M.A., et al. (2009). Genomic convergence to identify candidate genes for Alzheimer disease on chromosome 10. Hum. Mutat. *30*, 463–471. https://doi.org/10.1002/humu.20953.
58. Kohli, M.A., John-Williams, K., Rajbhandary, R., Naj, A., Whitehead, P., Hamilton, K., Carney, R.M., Wright, C., Crocco, E., Gwirtzman, H.E., et al. (2013). Repeat expansions in the C9ORF72 gene contribute to Alzheimer's disease in Caucasians. Neurobiol. Aging *34*. 1519.e5–e12. https://doi.org/10.1016/j.neurobiolaging.2012.10.003.
59. Jacobs, D., Sano, M., Marder, K., Bell, K., Bylsma, F., Lafleche, G., Albert, M., Brandt, J., and Stern, Y. (1994). Age at onset of Alzheimer's disease. Neurology *44*, 1215. https://doi.org/10.1212/wnl.44.7.1215.
60. Huff, F.J., Growdon, J.H., Corkin, S., and Rosen, T.J. (1987). Age at onset and rate of progression of alzheimer's disease. J. Am. Geriatr. Soc. *35*, 27–30. https://doi.org/10.1111/j.1532-5415.1987.tb01315.x.
61. Choi, S.W., and O'Reilly, P.F. (2019). PRSice-2: polygenic Risk Score software for biobank-scale data. Gigascience *8*, 1–6. https://doi.org/10.1093/gigascience/giz082.
62. R Development Core Team, R (2011). R: A Language and Environment for Statistical Computing.
63. Jun, G.R., Chung, J., Mez, J., Barber, R., Beecham, G.W., Bennett, D.A., Buxbaum, J.D., Byrd, G.S., Carrasquillo, M.M., Crane, P.K., et al. (2017). Transethnic genome-wide scan identifies novel Alzheimer's disease loci. Alzheimers Dement. *13*, 727–738. https://doi.org/10.1016/j.jalz.2016.12.012.
64. Bussies, P.L., Rajabli, F., Griswold, A., Dorfsman, D.A., Whitehead, P., Adams, L.D., Mena, P.R., Cuccaro, M., Haines, J.L., GS, B., et al. (2020). Use of local genetic ancestry to assess TOMM40-523' and risk for Alzheimer disease. Neurol. Genet. *6*, e404. https://doi.org/10.1212/nxg.0000000000000404.
65. Johnson, C.C., Rybicki, B.A., Brown, G., D'Hondt, E., Herpolsheimer, B., Roth, D., and Jackson, C.E. (1997). Cognitive impairment in the Amish: a four county survey. Int. J. Epidemiol. *26*, 387–394. https://doi.org/10.1093/ije/26.2.387.
66. Karch, C.M., Cruchaga, C., and Goate, A.M. (2014). Alzheimer's disease genetics: from the bench to the clinic. Neuron *83*, 11–26. https://doi.org/10.1016/j.neuron.2014.05.041.
67. Dumitrescu, L., Mahoney, E.R., Mukherjee, S., Lee, M.L., Bush, W.S., Engelman, C.D., Lu, Q., Fardo, D.W., Trittschuh, E.H., Mez, J., et al. (2020). Genetic variants and functional pathways associated with resilience to Alzheimer's disease. Brain *143*, 2561–2575. https://doi.org/10.1093/brain/awaa209.