The Genetic Architecture of Alzheimer Disease in the Mid-Western U.S. Amish

JL Haines[1,2], LN D'Aoust[3], R Laux[1], D Fuzzell[1], L Caywood[4], L Reinhart-Mercer[4], WK Scott[4], MA Pericak-Vance[4]

[1]Department of Epidemiology & Biostatistics, Case Western Reserve University, Cleveland, OH, USA
[2]Institute for Computational Biology, Case Western Reserve University, Cleveland, OH, USA
[3]Vanderbilt University School of Medicine, Nashville, TN, USA
[4]Hussman Institute for Human Genomics, University of Miami, Miller School of Medicine, Miami, FL, USA

INTRODUCTION

Objectives: There is a strong genetic influence on late onset Alzheimer disease (LOAD). Large international efforts have identified numerous susceptibility loci for LOAD and yet the majority of the genetic risk remains unidentified. We are examining LOAD in the Amish (an isolated religious group of Swiss-German origin) of Indiana and Ohio because of their cultural and genetic isolation from the general population and their more homogeneous life style.

Methods: Individuals scoring <87 on the 3MS were given a detailed neuropsychological battery; those scoring ≥87 were considered cognitively normal. We performed genome-wide SNP linkage and association studies on 921 individuals (109 with LOAD) using the Modified Quasi-Likelihood Score (MQLS) test and Merlin, respectively. We determined the overall burden of known LOAD loci by calculating a genetic risk score (GRS). Whole exome sequencing (WES) was performed on 166 cases and controls.

Results: Affected Amish individuals have a significantly higher GRS score than unaffected individuals ($P= 4.5 \times 10^{-3}$); but a much lower burden than an equivalent general population dataset of affected individuals ($P= 1.0 \times 10^{-7}$). Linkage analysis identified four likely loci that do not overlap with the known LOAD loci. WES identified a number of novel rare variants that are nominally associated with LOAD, but none reach Bonferroni-corrected significance.

Conclusions: While the Amish clearly carry some of the same risk loci as the general population, the GRS, linkage, and sequencing data strongly indicate that there are different loci that also play a significant role in LOAD in the Amish.

METHODS

We previously used the Affymetrix V 6.0 chip to perform a GWAS for LOAD (Cummings et al., 2012). We performed substantial quality control procedures and tested for linkage and association. Linkage was tested using Merlin. Given the extremely large interconnected Amish pedigree, we used PEDCUT to generate smaller subpedigrees suitable for linkage analysis. Association analysis was carried out using MQLS (developed by Thornton and McPeek), adjusting for relatedness using previously calculated kinship coefficients. Both linkage and association were performed using data from our total data set of 1,252 individuals. After quality control, we had 903 individuals available for analysis.

We selected 53 affected individuals and 65 unaffected individuals for whole exome sequencing (WES) from the total of 1,252 individuals. Exome sequencing used the Agilent SureSelect Human All Exon 50Mb capture kit and the Illumina HiSeq 2000. Raw sequence reads were aligned using BWA, duplicate reads were removed using Picard tools, and base recalibration and local realignment were performed using the Genome Analysis Toolkit (GATK). Variants were called by GATK's Unified Genotyper and recalibrated by VQSR. Basic quality control consisted of removing samples and variants with high levels of missingness and samples with possible gender errors and low concordance rates with previous genotyping. These processes resulted in 62,897 variants in or very near genes for analysis.

For the rare variants, we focused our analyses on two classes of genes: the first including 19 genes previously implicated in LOAD through GWAS or carrying early-onset mutations; the second including genes in our four previously identified significant genetic linkage regions.

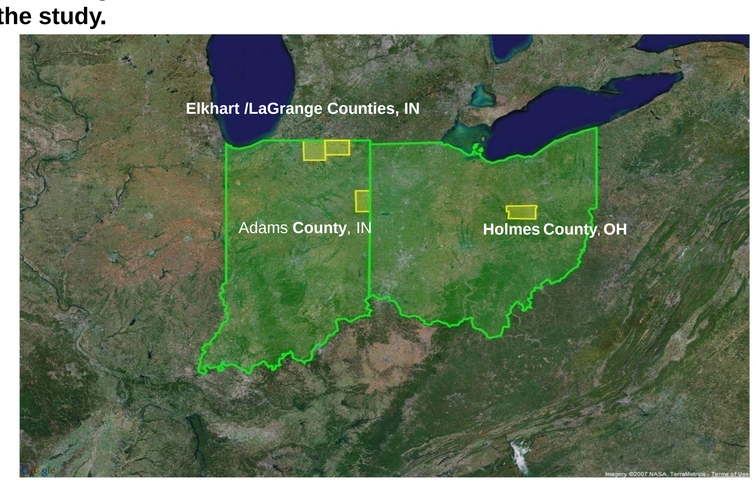
In addition, we specifically genotyped the 22 known LOAD associated loci using the Sequenom platform. 18 of the 22 loci passed quality control and were included in the genetic risk score analysis (GRS). 126 Amish cases (average age-at-onset=78) and 503 cognitively normal Amish controls (average age-at-exam=79) were used in this analysis.

For comparison, we used a set of 473 Caucasian cases (average age-at-onset=74) and 498 cognitively normal Caucasian controls (average age-at-exam=74) ascertained from the general U.S. population (in Tennessee, North Carolina, and Florida).

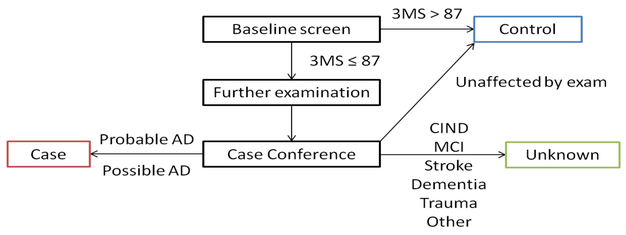
GRS scores were calculated by summing up the number of minor alleles at each locus weighted by the associated odds ratio (from Lambert et al., 2013). Logistic regression was used to compare the GRS across the datasets.

DATASET

We have been working with the Amish communities in central Ohio and northern Indiana for over 15 years. In this time, we have gained considerable experience in the complicated aspects of ascertaining individuals within this community. From public directories and referrals from previously enrolled participants, individuals over the age of 80 were identified. Over 30% of the Amish populations over the age of 65 have been contacted and 87% of these individuals have consented to participate in the study.

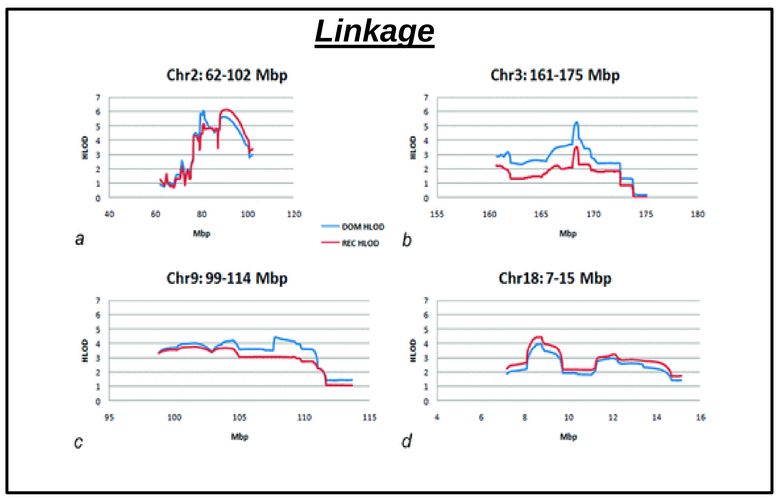


Door-to-door interviews were conducted and a baseline screen was performed using the Modified-Mini Mental Status (3MS) exam. Scores on this exam and findings from further cognitive testing were used during case conferences to generate a consensus diagnosis following the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) and the Alzheimer's Disease and Related Disorders Association (ADRDA) criteria (see figure).

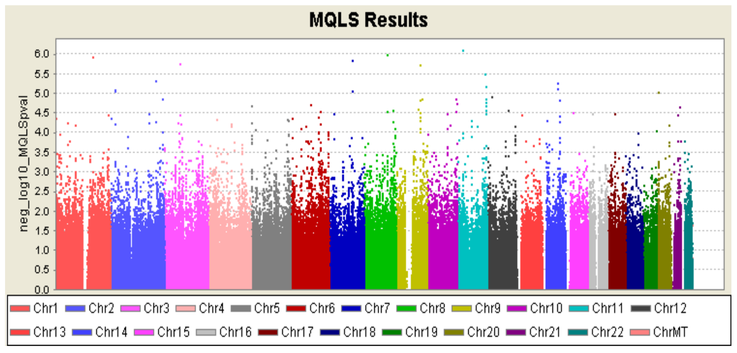


RESULTS

Linkage

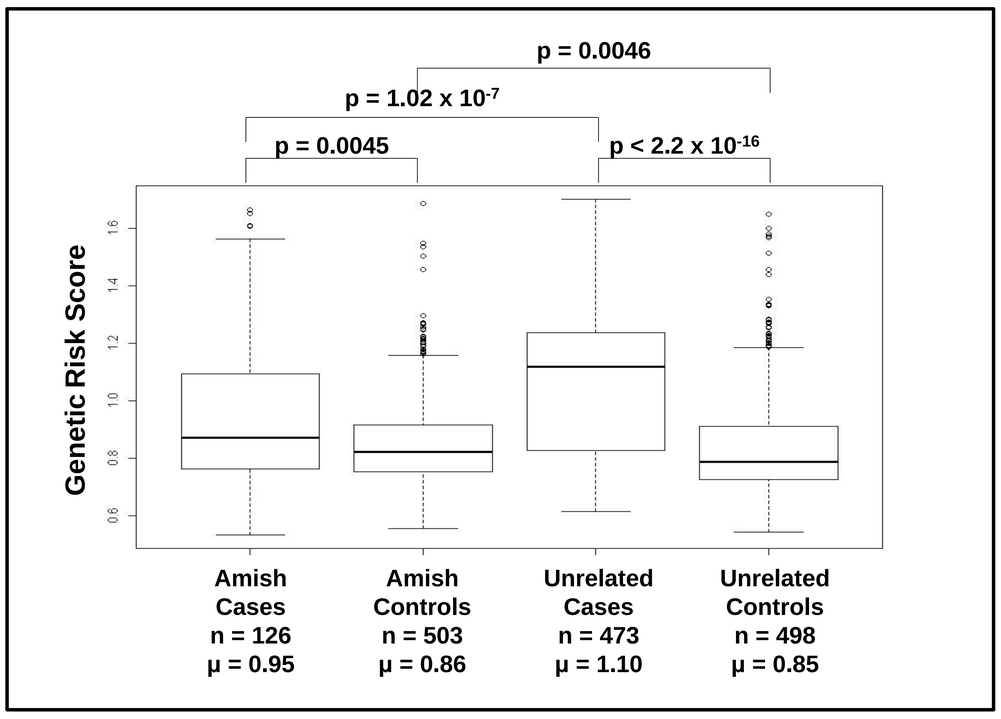


Association



Rare Variants in Known Genes

Gene	Location	Variants	dbSNP	ESP	1000G	Novel
ABCA7	19p13.3	20	19	18	18	1
APOE	19q13.2	0	0	0	0	0
APP	21q21.3	1	1	1	0	0
BIN1	2q14	3	3	3	3	0
CD2AP	6p12	1	1	1	1	0
CD33	19q13.3	1	1	1	1	0
CLU	8p21-p12	2	2	2	2	0
CR1	1q32	9	9	9	9	0
EPHA1	7q34	3	3	3	3	0
MS4A	11q12.2	45	42	41	35	1
PICALM	11q14	2	2	2	2	0
PSEN1	14q24.3	1	1	1	1	0
PSEN2	1q31-q42	2	1	1	1	1
SORL1	11q23.2-q24.2	15	14	14	14	1
TREM2	6p21.1	1	1	1	1	1



SUMMARY

- Known loci explain a small portion of LOAD in the Amish
- Novel loci are likely in the Amish
- Genetic architecture of LOAD in the Amish is complex

ACKNOWLEDGEMENTS

- NIH: AG019085, T32 GM07347
- Anabaptist Genealogy Database (Dr. Leslie Biesecker, Principal Investigator).
- Amish families for their kind participation