Alzheimer's & Dementia 12 (2016) 2-10

Alzheimer's & Dementia

Featured Article

Genome-wide linkage analyses of non-Hispanic white families identify novel loci for familial late-onset Alzheimer's disease

Brian W. Kunkle[a], James Jaworski[a], Sandra Barral[b,c,d], Badri Vardarajan[b,c,f,g], Gary W. Beecham[a], Eden R. Martin[a], Laura S. Cantwell[h], Amanda Partch[h], Thomas D. Bird[i,j], Wendy H. Raskind[j,k], Anita L. DeStefano[l], Regina M. Carney[a,m], Michael Cuccaro[a,n], Jeffrey M. Vance[a,n], Lindsay A. Farrer[k,o,p,q,r], Alison M. Goate[f,g], Tatiana Foroud[s], Richard P. Mayeux[b,c,d,e,t], Gerard D. Schellenberg[h], Jonathan L. Haines[u], Margaret A. Pericak-Vance[a,n,*]

[a]*John P. Hussman Institute for Human Genomics, Miller School of Medicine, University of Miami, Miami, FL, USA*
[b]*The Taub Institute of Research on Alzheimer's Disease, College of Physicians and Surgeons, Columbia University, New York, NY, USA*
[c]*The Gertrude H. Sergievsky Center, College of Physicians and Surgeons, Columbia University, New York, NY, USA*
[d]*Department of Neurology, College of Physicians and Surgeons, Columbia University, New York, NY, USA*
[e]*Department of Psychiatry, College of Physicians and Surgeons, Columbia University, New York, NY, USA*
[f]*Department of Psychiatry, Washington University School of Medicine, St. Louis, MO, USA*
[g]*Hope Center for Neurological Disorders, Department of Neurology, Washington University School of Medicine, St. Louis, MO, USA*
[h]*Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA*
[i]*Department of Neurology, University of Washington, Seattle, WA, USA*
[j]*Department of Medicine, University of Washington, Seattle, WA, USA*
[k]*Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, USA*
[l]*Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA*
[m]*Department of Psychiatry and Behavioral Sciences, Miller School of Medicine, University of Miami, Miami, FL, USA*
[n]*Dr. John T. Macdonald Foundation Department of Human Genetics, Miller School of Medicine, University of Miami, Miami, FL, USA*
[o]*Department of Medicine (Biomedical Genetics), Boston University School of Medicine and Public Health, MA, USA*
[p]*Department of Neurology, Boston University School of Medicine and Public Health, MA, USA*
[q]*Department of Ophthalmology, Boston University School of Medicine and Public Health, MA, USA*
[r]*Department of Epidemiology, Boston University School of Public Health, MA, USA*
[s]*Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA*
[t]*The Department of Epidemiology, School of Public Health, Columbia University, New York, NY, USA*
[u]*Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH, USA*

Abstract

Introduction: Few high penetrance variants that explain risk in late-onset Alzheimer's disease (LOAD) families have been found.

Methods: We performed genome-wide linkage and identity-by-descent (IBD) analyses on 41 non-Hispanic white families exhibiting likely dominant inheritance of LOAD, and having no mutations at known familial Alzheimer's disease (AD) loci, and a low burden of *APOE* ε4 alleles.

Results: Two-point parametric linkage analysis identified 14 significantly linked regions, including three novel linkage regions for LOAD (5q32, 11q12.2–11q14.1, and 14q13.3), one of which replicates a genome-wide association LOAD locus, the *MS4A6A-MS4A4E* gene cluster at 11q12.2. Five of the 14 regions (3q25.31, 4q34.1, 8q22.3, 11q12.2–14.1, and 19q13.41) are supported by strong multipoint results (logarithm of odds [LOD*] ≥1.5). Nonparametric multipoint analyses produced an additional significant locus at 14q32.2 (LOD* = 4.18). The 1-LOD confidence interval for this region contains one gene, *C14orf177*, and the microRNA *Mir_320*, whereas IBD analyses

None of the authors has a conflict of interest.

*Corresponding author. Tel.: +1-877-686-6444; Fax: +1-305-243-2396.

E-mail address: mpericak@med.miami.edu

http://dx.doi.org/10.1016/j.jalz.2015.05.020

implicates an additional gene *BCL11B*, a regulator of brain-derived neurotrophic signaling, a pathway associated with pathogenesis of several neurodegenerative diseases.
Discussion: Examination of these regions after whole-genome sequencing may identify highly penetrant variants for familial LOAD.

Keywords: Non-Hispanic white; Late-onset Alzheimer's disease; Linkage; High penetrance; Identity by descent; Familial; Genetics

1. Background

Although more than two dozen loci that contribute to late-onset Alzheimer disease (LOAD) have been identified [1], few genes with highly penetrant rare variants (e.g., *APP*, *PSEN1*, and *PSEN2* in early-onset familial Alzheimer's disease [AD] [2]) that explain risk in families heavily burdened with LOAD have been found. It is likely that rare variants contribute to complex disease, however [3], and recent reports implicating rare variants in *PLD3*, *APP*, and *TREM2* [4–7] support their involvement in both sporadic and familial LOAD. Identification of additional rare mutations driving genetic risk in familial LOAD will help in defining new pathways for therapeutic and preventive treatments.

Linkage analysis in large multiplex pedigrees is a robust approach for identifying disease loci in the presence of allelic heterogeneity, and thus can be valuable for targeting regions for sequencing studies [8]. To identify genomic regions likely to contain rare (minor allele frequency [MAF] $\leq$0.01) and low-frequency ($0.01 \geq \text{MAF} \leq 0.05$) LOAD risk and possibly protective genetic variants, a large number of well-characterized families were screened for inclusion in a linkage scan. The selected extended families are uniquely suited for discovery of genomic regions containing high penetrant AD variants. We performed extensive parametric two-point and nonparametric multipoint linkage analysis on 385 individuals in 41 non-Hispanic white (NHW) families. Loci identified through this study can help prioritize regions of the genome for analyses of whole-exome or whole-genome sequence data from NHW LOAD families or case-control cohorts.

2. Methods

2.1. Study samples

The 42 NHW families selected for linkage analyses are from five collections assembled by investigators at the University of Pennsylvania (eight families), the University of Miami (12 families), Case Western University (one family), the National Institute on Aging Late-Onset Alzheimer's Disease (NIALOAD) family study (17 families), and the National Cell Repository for Alzheimer's Disease (four families). Detailed descriptions of the ascertainment and evaluation of subjects in these cohorts have been provided elsewhere [9–11]. To maximize the probability of detecting segregating novel rare variants, we developed several selection criteria including (1) having four or more affected individuals with genomic DNA samples; (2) exhibiting likely dominant inheritance of LOAD; (3) free of known mutations at established Alzheimer's Disease/Frontotemporal Dementia Mendelian loci (*APP, PSEN1, PSEN2, MAPT*, or *GRN*); and (4) reduced representation of the *APOE* ε4 allele. Criterion number 4 included prioritized selection of families with *APOE* ε2/ε2, ε2/ε3, and ε3/ε3 affected individuals (requiring at least one affected family member without any *APOE* ε4 allele and any affected individuals with a single *APOE* ε4 must have age at onset [AAO] <72). In the 41 NHW families ultimately analyzed, 385 individuals (3–11 cases per pedigree) had genotyping data available and were included in the present analyses. Of those families, 75.6% (31 of 41) have at least one autopsy confirmed LOAD case (Table 1).

2.2. Genotyping and quality control procedures

Genome-wide single-nucleotide polymorphism (SNP) genotyping was performed on several different platforms across the study cohorts, including the Illumina HumanHap550, Illumina 1M, HumanOmniExpress, HumanOmniExpress Exome, and HumanOmni2.5 arrays. A call rate threshold of 98% was applied, and the data were then merged to form a final linkage data set for analysis. SNPs were only included in the analysis if they were present in at least 60% of samples; 319,409 SNPs were selected for analysis and aligned to the Rutgers Map v.3 [12]. Among this group of SNPs, 26,959 were excluded because the MAF was less than 0.05 and/or the genotype distribution differed significantly ($P < 10^{-6}$ in controls) from Hardy-Weinberg equilibrium. An additional 919 SNPs not present in the HapMap CEPH (Utah residents with ancestry from northern and western Europe) (CEU) data set were removed, reducing the number of SNPs available for analysis to 291,531 SNPs. More than three-fourths of these SNPs (77%; n = 225,250 SNPs) were present in ≥90% of samples. Checks for relatedness, Mendelian inconsistencies, and gender based on X-chromosome heterozygosity were performed using PLINK [13]. One sample was dropped due to Mendelian inconsistencies, and one duplicate sample