Neurobiology of Aging 38 (2016) 141–150

Contents lists available at ScienceDirect

Neurobiology of Aging

journal homepage: www.elsevier.com/locate/neuaging

Discovery of gene-gene interactions across multiple independent data sets of late onset Alzheimer disease from the Alzheimer Disease Genetics Consortium

Timothy J. Hohman [a], William S. Bush [b], Lan Jiang [c], Kristin D. Brown-Gentry [c], Eric S. Torstenson [c], Scott M. Dudek [c], Shubhabrata Mukherjee [d], Adam Naj [e], Brian W. Kunkle [f], Marylyn D. Ritchie [g], Eden R. Martin [f,h], Gerard D. Schellenberg [i], Richard Mayeux [j], Lindsay A. Farrer [k,l,m,n,o], Margaret A. Pericak-Vance [f,p], Jonathan L. Haines [b], Tricia A. Thornton-Wells [q,*], for the Alzheimer's Disease Genetics Consortium

[a] *Vanderbilt Memory & Alzheimer's Center, Department of Neurology, Vanderbilt University Medical Center, Nashville, TN, USA*
[b] *Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH, USA*
[c] *Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN, USA*
[d] *Department of Medicine, University of Washington, Seattle, WA, USA*
[e] *Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, PA, USA*
[f] *Dr. John T. Macdonald Foundation Department of Human Genetics and John P. Hussman Institute for Human Genomics, Miller School of Medicine, University of Miami, Miami, FL, USA*
[g] *Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA, USA*
[h] *Department of Public Health Sciences, Miller School of Medicine, University of Miami, Miami, FL, USA*
[i] *Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA*
[j] *Gertrude H. Sergievsky Center, Department of Neurology and the Taub Institute for Research on Alzheimer's Disease and the Aging Brain, College of Physicians and Surgeons, Columbia University, New York, NY, USA*
[k] *Department of Medicine (Biomedical Genetics), Boston University, Boston, MA, USA*
[l] *Department of Neurology, Boston University, Boston, MA, USA*
[m] *Department of Ophthalmology, Boston University, Boston, MA, USA*
[n] *Department of Epidemiology, Boston University, Boston, MA, USA*
[o] *Department of Biostatistics, Boston University, Boston, MA, USA*
[p] *Department of Neurology, Miller School of Medicine, University of Miami, Miami, FL, USA*
[q] *Vanderbilt Genetics Institute, Department of Molecular Physiology & Biophysics, Vanderbilt University Medical Center, Nashville, TN, USA*

ARTICLE INFO

Article history:
Received 30 March 2015
Received in revised form 28 October 2015
Accepted 28 October 2015
Available online 6 November 2015

Keywords:
Gene-gene interactions
Epistasis
Alzheimer disease
Biofilter

ABSTRACT

Late-onset Alzheimer disease (AD) has a complex genetic etiology, involving locus heterogeneity, polygenic inheritance, and gene-gene interactions; however, the investigation of interactions in recent genome-wide association studies has been limited. We used a biological knowledge-driven approach to evaluate gene-gene interactions for consistency across 13 data sets from the Alzheimer Disease Genetics Consortium. Fifteen single nucleotide polymorphism (SNP)-SNP pairs within 3 gene-gene combinations were identified: *SIRT1* × *ABCB1*, *PSAP* × *PEBP4*, and *GRIN2B* × *ADRA1A*. In addition, we extend a previously identified interaction from an endophenotype analysis between *RYR3* × *CACNA1C*. Finally, post hoc gene expression analyses of the implicated SNPs further implicate *SIRT1* and *ABCB1*, and implicate *CDH23* which was most recently identified as an AD risk locus in an epigenetic analysis of AD. The observed interactions in this article highlight ways in which genotypic variation related to disease may depend on the genetic context in which it occurs. Further, our results highlight the utility of evaluating genetic interactions to explain additional variance in AD risk and identify novel molecular mechanisms of AD pathogenesis.

* Corresponding author at: Department of Molecular Physiology & Biophysics, Vanderbilt University, Translational Medicine, Novartis Institutes of Biomedical Research 45 Sidney Street, 1202L Boston, MA 02139, USA. Tel.: +1 617 871 8112; fax: 617 871 8508.
E-mail address: triciathorntonwells@gmail.com (T.A. Thornton-Wells).

0197-4580/$ – see front matter © 2016 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neurobiolaging.2015.10.031

1. Introduction

Alzheimer disease (AD) has a strong yet complex genetic etiology and has already demonstrated allelic and locus heterogeneity and polygenic inheritance. It is possible that additional complexity, including gene-gene interactions, is also involved in the etiology of AD. Although rare mutations in multiple genes can affect early-onset AD, only common variation in *APOE* has a large effect on the more common late-onset form of AD (LOAD). Recent genome-wide association studies in LOAD have identified up to 21 additional novel genetic loci for AD, including genes from multiple pathways, such as beta-amyloid processing and clearance, calcium signaling, and extracellular matrix (Lambert et al., 2013; Naj et al., 2011). Other than *APOE*, the identified genetic loci have very modest effects, and in total the known genetic influences in LOAD still explain only about 33% of the broad-sense heritability (Ridge et al., 2013), which has been estimated to be 60%–80% (Gatz et al., 2006; So et al., 2011). One possible source of additional heritability is gene-gene interactions. Known loci could further influence disease risk through interactions with each other, as well as with other as yet unknown genetic factors. Also, novel loci with no detectable independent main effect on LOAD risk could interact with each other to significantly increase risk.

To date, the investigation of gene-gene interactions in LOAD has been pursued almost exclusively using a hypothesis-driven, candidate gene approach. Arosio et al. (2004) reported an interaction between variants in the proinflammatory cytokine genes IL6 and IL10, and Mateo et al. (2006) reported an interaction between the dopamine beta-hydroxylase gene and each of the 2 cytokine genes *IL1A* and *IL6* (Arosio et al., 2004; Mateo et al., 2006). The Epistasis Project was able to replicate both of these findings in LOAD (Combarros et al., 2010). Interactions between variants in the transferrin gene (*TF*) and the hemochromatosis gene (*HFE*) also have been identified and replicated in multiple cohorts for association with LOAD (Kauwe et al., 2010; Robson et al., 2004). An interaction between the insulin gene (*INS*) and the peroxisome proliferator–activated receptor alpha gene (*PPARα*) has been reported in Northern but not Southern Europeans (Heun et al., 2012; Kolsch et al., 2012). Risk for LOAD and vascular dementia reportedly vary according to the interaction of genotypes in the *MTHFR* and *IL6* genes (Mansoori et al., 2012).

Even in hypothesis-free genome-wide association studies (GWAS) of AD, when testing of gene-gene interactions has been incorporated, it has been restricted to interactions between *APOE* and other risk loci with known main effect associations. Belbin et al. (2011) investigated interactions among 21 LOAD candidate and confirmed risk genes, including *APOE*, *BIN1*, *CLU*, *CR1*, and *PICALM* but failed to detect any interactions with disease status or age-at-onset that were significant after correction for multiple testing (Belbin et al., 2011). Similarly, Carrasquillo et al. (2011) failed to identify significant interactions between variants in *BIN1* and other LOAD risk genes, including *APOE*, *CLU*, *CR1*, and *PICALM* (Carrasquillo et al., 2011).

In this study, we aimed to identify novel gene-gene interactions that demonstrated association with LOAD across multiple independent data sets. We used a network-based approach to discovery, using prior biological knowledge about LOAD candidate genes—the pathways in which they participate and the genes with which they are related or are known to interact—to guide initial selection of gene-gene models for investigation (Bush et al., 2009). We also used a meta-analysis approach by which we could evaluate the consistency of each identified single nucleotide polymorphism (SNP) × SNP interaction across the 13 independent data sources while correcting for the total number of comparisons evaluated. Finally, we performed a comprehensive analysis of 2 gene-gene pairs that were previously identified in projects by our research group leveraging endophenotypes of AD to validate the observed effects in case-control data sets.

2. Materials and methods

2.1. Data sets and quality control procedures

Study data consisted of subjects from 13 data sets available through the Alzheimer's Disease Genetics Consortium, including the following: the Adult Changes in Thought; the National Institute on Aging Alzheimer Disease Centers (ADC1, ADC2, ADC3); the Alzheimer's Disease Neuroimaging Initiative (ADNI); Oregon Health & Science University (OHSU); Rush University Religious Orders Study/Memory and Aging Project (ROS/MAP); Translational Genomics Research Institute series 2 (TGEN2); University of Miami/Vanderbilt University/Mt. Sinai School of Medicine (UM/VU/MSSM); and Washington University (WashU). All subjects were recruited under protocols approved by the appropriate Institutional Review Boards.

After quality control, the combined data set included samples from 7758 LOAD cases and 6724 cognitively normal elder (CNE) controls. For most of the cohorts, LOAD cases met National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria for probable or definite LOAD with age at onset greater than 60 years, and clinically-defined CNEs had a documented Mini-Mental State Exam, Cognitive Abilities Screening Instrument, or Modified Mini Mental State Exam (3MS) score in the normal range. The only exceptions were TGEN2 and ADNI. The TGEN2 data set comprised clinically- and neuropathologically characterized brain donors, 668 with LOAD and 365 CNEs without dementia or significant LOAD pathology. The samples were obtained from 21 different National Institute on Aging-support LOAD Center brain banks and from the Miami Brain Bank as previously described (Caselli et al., 2007; Liang et al., 2011; Reiman et al., 2007; Webster et al., 2009). Additional samples from other brain banks in the United States, United Kingdom, and the Netherlands were obtained in the same manner. The ADNI data set comprised 268 LOAD cases and 173 CNEs with neuroimaging support for diagnosis. In the ADNI cohort, LOAD subjects were between the ages of 55–90 years, had an Mini-Mental State Exam score of 20–26 inclusive, met NINCDS/ADRDA criteria for probable LOAD (McKhann et al., 1984), and had a magnetic resonance image consistent with the diagnosis of LOAD at the most recent follow-up. Table 1 presents descriptive statistics for each of the data sets.

2.2. Genotyping

Samples were genotyped at different stages of recruitment on the Affymetrix 6 (UM/VU/MSSM), Affymetrix 1M (TGEN2), Illumina 610 (ADNI, OHSU, UM/VU/MSSM), Illumina 660 (Adult Changes in Thought, ADC1, ADC2, WashU), Illumina OmniExpress (ADC3), and Illumina IM (ROS/MAP, UM/VU/MSSM). Each data set was independently imputed using IMPUTE2 with 1000 Genomes Phase 2 samples of European ancestry. Because we were primarily interested in discovering novel gene-gene interactions and not those that modify risk of the major LOAD gene, *APOE*, we excluded SNPs within 50 kb of *APOE*.

2.3. Quality control procedures

Quality control procedures were applied to each data set separately. Genotype data were cleaned by applying a 98% threshold for genotyping efficiency and a minimum minor allele frequency (MAF)