ELSEVIER

Neurobiology of Aging 32 (2011) 2328.e1–2328.e9

NEUROBIOLOGY OF AGING

www.elsevier.com/locate/neuaging

TOMM40 poly-T repeat lengths, age of onset and psychosis risk in Alzheimer disease

Su Hee Chu[a], Kathryn Roeder[a], Robert E. Ferrell[b], Bernie Devlin[c], Mary Ann A. DeMichele-Sweet[c], M. Ilyas Kamboh[b,c], Oscar L. Lopez[c,d], Robert A. Sweet[c,d,e,*]

[a] *Department of Statistics, Carnegie Mellon University, Pittsburgh, PA, USA*
[b] *Department of Human Genetics, University of Pittsburgh, Pittsburgh, PA, USA*
[c] *Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA, USA*
[d] *Department of Neurology, University of Pittsburgh, Pittsburgh, PA, USA*
[e] *VISN 4 Mental Illness Research, Education and Clinical Center (MIRECC), VA Pittsburgh Healthcare System, Pittsburgh, PA, USA*

Received 1 March 2011; received in revised form 7 June 2011; accepted 17 June 2011

Abstract

Apolipoprotein E (*APOE*) ε4 alleles increase the risk for late-onset Alzheimer disease (LOAD) and decrease the age of onset. Recently, sequencing the *APOE* region in a small sample of LOAD subjects identified a variable length poly-T repeat sequence in the nearby gene, *TOMM40*, which may affect age of onset. We genotyped the *TOMM40* poly-T repeat using a novel statistical approach to refine the identification of allele length in 892 LOAD subjects and evaluated its effects on age of onset. Because psychosis in LOAD is a heritable phenotype which has shown conflicting associations with *APOE* genotype, we also evaluated the association of poly-T repeat length with psychosis. Poly-T repeat lengths had a trimodal distribution which differed between *APOE* genotype groups. After accounting for *APOE* ε4 there was no association of poly-T repeat length with age of onset. Neither *APOE* ε4 nor poly-T repeat length was associated with psychosis. Our findings do not support the association of poly-T repeat length with age of onset in LOAD. The clinical implications of this repeat length polymorphism remain to be elucidated.

Keywords: Apolipoprotein E (APOE) ε4; Late-onset Alzheimer disease (LOAD); Psychosis; *TOMM40*; Variable length poly-T repeat sequence

1. Introduction

Late-onset Alzheimer disease (LOAD) is a neurodegenerative illness with substantial heritability (Gatz et al., 1997). Similarly, and not surprisingly, as the risk of LOAD increases in a highly age-dependent manner (Ferri et al., 2005), the age of onset of LOAD is also heritable (Pedersen et al., 2001). The gene with the most strongly established relationship to LOAD risk is apolipoprotein E (*APOE*), with increased risk of LOAD found in individuals carrying 1 or 2 copies of the ε4 allele (Farrer et al., 1997). The primary effect of *APOE* ε4 alleles on LOAD risk appears to be mediated via lowering the age of onset of Alzheimer disease, with a reduction of up to 7–9 years for each ε4 allele (Reitz and Mayeux, 2009).

The *APOE* ε4 allele has also been extensively examined for association with the presence of the psychotic phenotype of LOAD (LOAD+Psychosis, LOAD+P) (Sweet et al., 2003). LOAD+P is heritable (Bacanu et al., 2005; Sweet et al., 2010), and identifies a subgroup of LOAD subjects with more severe cognitive impairment and more rapid cognitive decline (Emanuel et al., 2011; Ropacki and Jeste, 2005). Unlike age of onset, the association of *APOE* ε4 allele with

* Corresponding author at: Mail: Biomedical Science Tower, Rm W-1645, 3811 O'Hara Street, Pittsburgh, PA 15213-2593, USA; Express Mail: Biomedical Science Tower, Rm W-1645, Lothrop and Terrace Streets, Pittsburgh, PA 15213-2593, USA. Tel.: +1 412 383 8548; fax: +1 412 624 9910.
E-mail address: sweetra@upmc.edu (R.A. Sweet).

0197-4580/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.neurobiolaging.2011.06.016

psychosis has revealed inconsistent findings, with slightly more negative than positive studies, and some studies showing evidence for a protective effect (DeMichele-Sweet and Sweet, 2010). Such a pattern could result solely from type I error due to small cohorts with varying approaches to clinical characterization and analysis, however, a variable pattern of association can also arise due to a causal association with genetic variation in linkage disequilibrium with the *APOE* ε4 allele.

APOE ε4 is defined by a 2 single nucleotide polymorphism (SNP) haplotype in *APOE* exon 4. SNPs rs429358 and rs7412 each code for either arginine (C) or cysteine (T). *APOE* ε4 alleles are the CC haplotype (with TT and TC defining ε2 and ε3 alleles, respectively). Recent investigations fine-mapping the region within and surrounding *APOE* on chromosome 19 identified a set of SNPs within the nearby gene, *TOMM40*, in linkage disequilibrium with the *APOE* ε4 allele (Yu et al., 2007) and affecting *APOE* expression (Bekris et al., 2010). This finding, in part, motivated an effort to sequence the *APOE* and *TOMM40* region in subjects with Alzheimer disease (AD), in an effort to identify possible causal variants within the linked region (Roses et al., 2010). Sequencing identified a variable length poly-T repeat sequence in intron 6 of *TOMM40* that was in linkage disequilibrium with *APOE* ε4. Individuals with *APOE* ε3/ε4 genotype and long poly-T repeats (defined as ≥ 27) had significantly lower age of onset of LOAD than individuals with *APOE* ε3/ε4 genotype and short repeats (Roses et al., 2010).

Ultimately, reconciling the independent effects of *APOE* ε4 and *TOMM40* repeat length polymorphism on age of onset of LOAD will require concurrent genotyping of large numbers of subjects. To address this goal, we developed an approach to high throughput genotyping of the *TOMM40* poly-T repeat length polymorphism by starting with polymerase chain reaction (PCR) to generate an initial estimate of allele sizes and then refining these estimates with a statistical model. We evaluated the independent and joint effects of these genetic variants in a large population of 892 Caucasian individuals with LOAD, examining both the age of onset and LOAD+P phenotypes.

2. Methods

2.1. Subjects

A total of 892 Caucasian, non-Latino subjects with a final diagnosis of possible or probable AD (McKhann et al., 1984), all evaluated at the University of Pittsburgh Alzheimer Disease Research Center (ADRC), were included. All subjects were assessed as described previously (DeMichele-Sweet et al., 2011a). All data collected in this study were obtained with protocols approved by the Institutional Review Board of the University of Pittsburgh.

2.2. Assessment of psychosis

Psychosis was evaluated with the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) behavioral rating scale (Tariot et al., 1995), as described previously (DeMichele-Sweet et al., 2011a). Subjects were characterized as having no psychotic symptoms, a single psychotic symptom at only 1 time point, or multiple/recurrent psychotic symptoms, reflecting the increasing genetic loading associated with this hierarchy (Bacanu et al., 2005; Sweet et al., 2010). Finally, because the occurrence of psychosis is less frequent in the early stages of AD, subjects were required to have a Mini Mental State Examination (Folstein et al., 1975) score $\leq$20 in order to be classified as having LOAD without psychosis.

2.3. Genotyping

2.3.1. Assay

The *TOMM40* polymorphic repeat was genotyped by PCR amplification with forward primer 5′-VIC-GAGATGGGGTCTCACTATG-3' and reverse primer 5′-GTACAGGCCACAATGTG-3′, with an initial 3-minute denaturation at 95 °C, followed by 35 cycles of denaturation at 95 °C for 30 seconds, annealing at 56 °C for 30 seconds, and extension at 72 °C for 30 seconds. PCR was carried out in a final volume of 10 μL containing 10-pM primers, 200 μM dNTPs, 2 mM MgCl and 1 U of Taq polymerase (Invitrogen, Carlsbad, CA, USA). Fragments were resolved on an ABI 3730 automatic fragment analyzer, with a LIZ500 size standard, and fragment sizes were initially estimated using the output from GeneMapper v4.0 software (Applied Biosystems, Inc., Carlsbad, CA, USA).

2.3.2. Preliminary allele calls

The *TOMM40* DNA sequence has an intronic poly-T (multiple thymine base pairs) that is highly variable at a population level. Our first statistical goal was to estimate the counts of T for each of the pair of alleles carried by subjects using the intensity signals obtained from the GeneMapper (Applied Biosystems, Inc.) readout, where intensity (I) is some function of the number of times a particular length was replicated in the PCR process. The expected pattern for our method of measuring alleles would be to observe, for each allele, a maximum intensity of signal near the true count of T (N[T]), but distributed continuously with error around the true value, and PCR-based stutter around that peak that decays in intensity with increasing (integer) distance from N[T]. Thus, for an individual with 2 distinctly different poly-T alleles, a reasonable preliminary estimate can be obtained as the length associated with the maximum intensity of the 2 distinct peaks in an individual's GeneMapper (Applied Biosystems, Inc., Carlsbad, CA, USA) readout (Fig. 1). Note in Fig. 1 the PCR stutter. Measurement error is illustrated by Fig. 2. If an individual is homozygous for poly-T alleles, then the global maximum is a good estimator for both alleles.