Journal of Pharmacogenomics & Pharmacoproteomics

Editorial **Open Access**

APOE and TOMM40 in Alzheimer's Disease: A Case of Mistaken Identity, Using Precision Tools but with Loss of Accuracy

Allen D Roses*

Duke University Medical Center, Zinfandel Pharmaceuticals, Chapel Hill, NC 27517, USA

"The association of the *apolipoprotein E* (*APOE*) ε4 allele with AD is strong (odds ratios ranging from 3 to 10) and undisputed" [1]. In fact, in 2015 this remains generally true, yet this same paper falls to mention the association with the poly-T variant (rs10524523) in the TOMM40 gene (aka TOMM40'523) which encodes the channel component of the complex that transports proteins across the outer mitochondrial membrane. The same research group published in Archives of Neurology in 2012 a widely quoted editorial entitled "*TOMM40* Association with *Alzheimer's disease: Tales of APOE and Linkage Disequilibrium*" [2]. In this article, Guerreiro and Hardy made the case that although APOE and TOMM40 were adjacent to each other on chromosome 19, and in linkage disequilibrium [LD]; the association of TOMM40 was entirely due to its position in LD with APOE. To further support this argument, the authors stressed that, "since the first genome-wide association studies in AD it was clear that no other locus in the genome would present such a strong association with the disease as that on chromosome 19 around the APOE locus" [2]. What the authors failed to comprehend or consider was that the GWAS SNPs representing the APOE region were actually four TOMM40 SNPs and a 3 non-coding SNP near APOE, not either coding SNPs of APOE. The two coding SNPs that determine the differences between APOE4, APOE3, and APOE2 alleles are not present in the genotyping platforms typically used in GWAS studies. Simply stated, it was TOMM40 SNPs being measured, not APOE SNPs and therefore a case of mistaken identity. Identity is equivalent to accuracy.

Lack of precision in interpreting the GWAS data has delayed the progress of targeting medicines for AD [3,4]. In fact, in 2015, there exist significant and growing support for the genetics of TOMM40 including TOMM40'523, and a plausible hypothesis for the intraneuronal interaction of apoE3 and apoE4 fragments and the Tom40 protein complex leading to potential mode of action therapeutic targets [5-6]. In addition, TOMM40'523-APOE haplotypes are more informative for disease risk than APOE4.

The TOMM40'523 locus is not a single SNP, but a structural variation with more than 25 different poly-thymidine length variants, referred to as a simple sequence repeat [SSR]. The multiple, inherited poly-thymidine sequence lengths at the rs10524523 locus can be divided into three observable size ranges: **S**mall, less than 20 thymidine [T] length repeats; **L**arge, 20-30 Ts: and **V**ery **L**arge, more than 30 Ts. [7] All of the poly-T variant lengths are at the rs10524523 position in LD with APOE, and, when measured accurately, it is clear that the different size ranges from distinct haplotypes with different APOE alleles. Humans inherit two of the three possible allele range, or alleles [**S**, **L**, **VL**] and two of a much larger number of possible TOMM40'523-APOE haplotypes.

Using phylogenetic analysis on the complete DNA sequence of this region, it was possible to define APOE-TOMM40'523 haplotypes for each APOE allele and map the age of onset distribution by APOE4-TOMM40'523 and APOE3-TOMM40'523 haplotypes and demonstrate that each of the major haplotypes could be differentiated and related to different age of onset curves. The phylogenetic mapping of the ancestral development of a portion of the DNA that contained APOE and TOMM40 sequences clarified that the haplotypes of APOE3-S and APOE3-VL are localized to two highly distinct phylogenetic clades [7]. Using this phylogenetic approach it was possible to map all the TOMM40 SNPs and TOMM40'523 variants in LD [8]. This identified two distinctly separated clades that could be identified S and VL, that are in LD with the APOE3 or APOE2 alleles. Thus it became possible in 2010 to differentiate two different APOE3 types that could not be distinguished by APOE typing alone. This results in a higher proportion of informative individuals with TOMM40'523 sequencing that APOE4 alone. Caucasian individuals with the APOE4/4 genotype had two Long TOMM40'523 alleles. It is now possible to demonstrate that for APOE3 and APOE2 chromosomes there were two distinct haplotypes represented, allowing prognostic data for age of onset distributions from all APOE3/3 and APOE2/3 carrying individuals. Thus APOE4 distribution curves are a sub-set of the TOMM40'523 age of onset distributions, increasing the individual haplotype informative genotyping to 100% subjects, compared to 29% of subjects who carry an APOE4 allele [8].

For more than 20 years, APOE4-containing genotypes [APOE4/4, APOE3/4 and APOE2/4] were clearly associated with earlier age of onset distributions. Corder et al. identified the later effect age of onset for carriers of APOE3/3 genotypes in 1993 [9]. Saunders et al. reported the relationship of APOE4 carriage with the onset of late-onset [so-called sporadic] AD in 1993 [10]. For clarity, APOE3/3 was never AD-neutral, simply associated with later age of onset than the APOE4-containing genotypes. Over the years, clinical studies and drug trials compared the APOE4 carriers against the APOE4 non-carriers, assuming the neutrality of APOE3 and not defining the age of onset differences in the comparative groups. It has been shown in multiple clinical studies, including many therapeutic trials with APOE3/3 cases representing more than 50% of the AD prevalent population, confirming that the APOE3 variant can hardly be described as having a neutral effect. The APOE2/2 is exceedingly rare in any AD cohort [especially in neuropathologically-confirmed criteria for amyloid deposition] and APOE2/3 has the latest age of onset distribution compared to APOE3/3, APOE3/4 and APOE4/4 genotypes [8].

The relevant clinical question is to ask which measurement is better as an informative prognostic for AD: APOE4 carriers [29% of Caucasians] or APOE-TOMM40'523 haplotypes for 100% of

*Corresponding author: Allen D Roses, Jefferson-Pilot Distinguished Professor of Neurobiology and Neurology, Duke University Medical Center, Durham NC, CEO, Zinfandel Pharmaceuticals, Inc. 1450 Raleigh Road, Ste. 210, Chapel Hill, NC 27517, USA, Tel: 1-919-660-8065; E-mail: roses@zinfandelpharma.com

Caucasians. It was clear in 1995 when a panel of experts stated that presence of APOE4 was not sufficient to use as a diagnostic [11]. [That was the reason the two SNPs defining the three isoforms were excluded from the GWAS platforms so that they could not be used in clinical diagnosis without test validation as a diagnostic. In this age of "precision medicine", the biological effects have to be analyzed at least as accurately as the "Big Data" statistical analyses [12]. This has not been the case with regard to the genetics of AD. APOE measurements can be accurately ascertained and with suitable careful technique length variation in the homopolymer of Ts can also be accurately assessed [+/- 1-T differences] [13,14]. Had TOMM40'523 been discovered before APOE association, APOE4 alleles would have be considered as a subgroup of APOE-TOMM40'523 haplotypes. Thus, the biological interactions with the Tom40 protein should not be excluded from consideration as an important contributor to disease progression, as has been the case in most reviews since 2010. The genetic informative relationship between APOE and TOMM40'523 as haplotypes continues to be largely excluded from discussion and therefore wider biological considerations for potential therapies [1,4,15-17].

The penalty paid in this era of precision medicine is inaccurately interpreting the roles of TOMM40'523 and APOE in disease pathogenesis, resulting in delays testing new hypotheses while other older hypotheses fail. This is not good science. Biological data relative to pathogenic hypotheses should be enlisted early, not after an additional decade with millions of new patients with AD remains inadequately treated. Amyloid plaques are viewed by a minority of investigators as a consequence [not an etiology] of the disease process in the CNS, and this concept is greatly supported by the total lack of clinical association of the APP locus with late-onset AD. Surely the Big Data of more than 2 dozen AD GWAS studies should be analyzed appropriately for genes related to the molecular etiology of disease, for both positive and negative association with specific clinical phenotypes.

Ages of onset distributions are clinically-relevant measurements [although usually poorly documented in retrospective studies]. Age of onset is especially accurate when the criteria for onset are well-defined and subjects are studied prospectively from normal cognition to the first signs of impairment [8]. The amyloid load shows a greater density in APOE4 carrying patients, yet the genetic association of the APP gene or nearby loci with late-onset AD has been consistently negative [16]. Rather, much of the AD field insists that the amyloid cascade is the etiology for AD, and therefore its expression in brain tissue is a causal phenomenon. Similar criteria for age of onset using APOE and TOMM40'523 illustrates that both contribute to a biological effect – and the fact that they are located adjacent to each other supports the need for further interaction studies, especially with other components of the translocase of the outer mitochondrial membrane complex that regulate protein and peptide entry into the mitochondrion. This is also inferred from Figure 1, demonstrating that inheriting different TOMM40'523-S or TOMM40'523-VL haplotypes are also related to a more defined range of age of onset distributions.

The result of not considering the phylogenetic data and the age of onset distributions has led to a widespread overlooking of the biological role of TOMM40'523 in many/most comments and reviews of the genetics of AD. The confirmation of prospectively followed AD cohorts with respect to age of onset distributions takes years, and as more data has supported the role of TOMM40'523 the perils of ignoring the science has been further exposed.

This comes at a time when pursuit of the amyloid cascade hypotheses had not led to effective treatments even though tens of $billions have been invested. Target identification based on mitochondrial and other metabolic modes of pathogenesis could lead to a myriad of potential new drug development programs at a time when the pharmaceutical industry has generally decided that most central nervous disease treatments do not represent worthwhile portfolio opportunities. The nature of CNS clinical trials for new drugs entails approximately a 12-15 year development period that is usually measured against a 20 year chemical patent. The largest single cause of failure of completed trials has been for no efficacy. Another crippling cause of failure rests on the business development philosophies of pharmaceutical companies. Actually, pharmaceutical companies do not have biases, but CEOs and Clinical Development Directors do – and they generally change with management every 4-7 years on the average - usually disrupting the continuity of long, relatively expensive clinical trials. In-house developed Futility Analyses are now the second largest cause for early termination of expensive trials [17-19].

With more than three decades with the amyloid hypothesis squeezing out most new initiatives, trying to test new hypotheses in AD results is nearly impossible. The travesty is that testing a single hypothesis over 30 years has resulted in significantly diminished funding to explore equally plausible hypotheses. How does that happen? Imagine that most everyone successful in AD academics, most doing some perturbation of amyloid hypotheses, review the grant requests for innovative new hypotheses. Most are killed "with faint praise" and miss the funding line – others are simply killed before review. The intellectual density of amyloid supporters in funding positions is not supportive of peer review, but a process of conflict of interest voting. This has been commented about over the years – but still continues [20]. The only way that new ideas will be tested is when the system changes and, especially in the field of AD, when funding is designated for new hypotheses and reviews are based on sound review of data, not editorial or peer review trump cards. "Faced with the choice between changing one's mind and proving that there is no need to do so, almost everyone gets busy on the proof" [John Kenneth Galbraith].

References

1. Guerreiro R, Bras J, Toombs J, Heslegrave A, Hardy J, et al. (2015) Genetic Variants and Related Biomarkers in Sporadic Alzheimer's Disease. Curr Genet Med Rep 3: 19-25.

2. Guerreiro RJ, Hardy J (2012) TOMM40 association with Alzheimer disease: Tales of APOE and linkage disequilibrium. Arch Neurol 69: 1243-1244.

3. Jun G, Vardarajan BN, Buros J, Yu CE, Hawk MV, et al. (2012) Comprehensive search for Alzheimer disease susceptibility loci in the APOE region. Arch Neurol 69: 1270-1279.

4. Guerreiro R, Hardy J (2014) Genetics of Alzheimer's disease. Neurotherapeutics 11: 732-737.

5. Mahley RW, Huang Y (2012) Small-molecule structure correctors target abnormal protein structure and function: Structure corrector rescue of apolipoprotein E4-associated neuropathology. J Med Chem 55: 8997-9008.

6. Gottschalk WK, Lutz MW, He YT, Saunders AM, Burns DK, et al. (2014) The Broad Impact of TOM40 on Neurodegenerative Diseases in Aging. Journal of Parkinson's disease and Alzheimer's disease 1: 12.

7. Roses AD, Lutz MW, Amrine-Madsen H, Saunders AM, Crenshaw DG, et al. (2010) A TOMM40 variable-length polymorphism predicts the age of late-onset Alzheimer's disease. Pharmacogenomics J 10: 375-384.

8. Crenshaw DG, Gottschalk WK, Lutz MW, Grossman I, Saunders AM, et al. (2013) Using genetics to enable studies on the prevention of Alzheimer's disease. Clin Pharmacol Ther 93: 177-185.

9. Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, et al. (1993) Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science 261: 921-923.

10. Saunders AM, Strittmatter WJ, Schmechel D, George-Hyslop PH, Pericak-Vance MA, et al. (1993) Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. Neurology 43: 1467-1472.

11. National Institute on Aging/Alzheimer's Association Working G, Relkin N (1996) Apolipoprotein E genotyping in Alzheimer's disease. The Lancet 347: 1091-1095.

12. Zhang XD (2015) Precision Medicine, Personalized Medicine, Omics and Big Data: Concepts and Relationships. Journal of Pharmacogenomics Pharmacogenetics 6: 144.

13. Roses AD, Lutz MW, Crenshaw DG, Grossman I, Saunders AM, et al. (2013) TOMM40 and APOE: Requirements for replication studies of association with age of disease onset and enrichment of a clinical trial. Alzheimer's and Dementia: The Journal of the Alzheimer's Association 9: 132-136.

14. Roses AD, Lutz MW, Saunders AM, Goldgaber D, Saul R, et al. (2014) African-American TOMM40'523-APOE haplotypes are admixture of West African and Caucasian alleles. Alzheimer's Dement 10: 592-601.

15. Bertram L, Lill CM, Tanzi RE (2010) The Genetics of Alzheimer Disease: Back to the Future. Neuron 68: 270-281.

16. Van Cauwenberghe C, Van Broeckhoven C, Sleegers K (2015) The genetic landscape of Alzheimer disease: clinical implications and perspectives. Genet Med 117.

17. Chouraki V, Seshadri S (2014) Genetics of Alzheimer's disease. Adv Genet 87: 245-294.

18. Roses AD, St Jean PL, Ehm MG (2007) Use of whole-genome association scans in disease gene identification, drug discovery and development. IDrugs 10: 797-804.

19. Jitlal M, Khan I, Lee SM, Hackshaw A (2012) Stopping clinical trials early for futility: Retrospective analysis of several randomized clinical studies. Br J Cancer 107: 910-917.

20. Schoenfeld DA (2005) Pro/con clinical debate: It is acceptable to stop large multicenter randomized controlled trials at interim analysis for futility. Pro: Futility stopping can speed up the development of effective treatments. Crit Care 9: 34-36.