Journal of Alzheimer's Disease 24 (2011) 247–251
DOI 10.3233/JAD-2010-101765
IOS Press

Commentary

The Importance of Being Connected

Michael W. Lutz, Donna G. Crenshaw, Ann M. Saunders and Allen D. Roses*
Deane Drug Discovery Institute, Duke University, Durham, NC, USA

Accepted 12 December 2010

In the 17 years since the original report linking the ε4 allele of the apolipoprotein E (*APOE*) gene [1] with late-onset Alzheimer's disease (LOAD), considerable effort has been expended to elucidate the role that *APOE* plays in LOAD risk and age of disease onset. LOAD is a genetically complex and heterogeneous disorder. The disease has a strong heritability that is estimated to reach nearly 80% [2], yet only a few genes that account for a proportion of the observed genetic variability have been definitively associated with the disease [3]. The only firmly established genetic risk factor for sporadic AD, and the risk factor with the largest effect size, is the ε4 allele of *APOE* [4]. Carriage of an *APOE* ε4 allele significantly increases the lifetime risk for AD, with the level of risk increasing as the ε4 allele dose increases [4, 5], and is associated with lower age of disease onset [4, 6–9]. The *APOE* gene occurs in an extended linkage region on chromosome 19 that provides an extraordinary genome wide association signal with LOAD. This association signal is generally attributed to the ε4 allele of *APOE*, but other polymorphisms in adjacent genes and the *APOE* promoter may also contribute to disease pathogenesis. In this commentary, we consider the findings of Lescai et al. [10] in the context of an extensive phylogenetic analysis of the *TOMM40-APOE* region recently reported by our group [11, 12].

In this issue, Lescai and colleagues identify a haplotype, comprised of a specific allele of an *APOE* promoter SNP (rs405509 or −219T) and the two SNPs that determine the *APOE* ε4 genotype (rs429358-rs7412 CC) that increases the risk for LOAD when the alleles are present in *cis* on the same chromosome [10]. Additionally, they show that this haplotype is significantly associated with younger age at diagnosis when compared to haplotypes that contain the G form of the −219 SNP and the *APOE* ε4 allele, which suggests a link to earlier disease onset. Notably, the authors stress the importance of the occurrence of the associated alleles in *cis*, or in phase, in their analyses [10]. Prior studies, where the phase of the alleles on the chromosome was not determined or was estimated by inference, had limited power to measure complex interactions of specific alleles at multiple polymorphic sites in the genomic interval of chromosome 19 that includes *APOE* and several other genes.

Phylogenetic analyses are widely employed in medicine to explore the evolutionary relationships between, and the emergence of mutations in, human pathogens, but these analyses are less commonly used to identify disease-associated variation in the human genome [13–16]. The branching tree representations of variations in gene sequences, like that in Fig. 1, illustrate the relationships among the sequences or the different evolutionary lineages of the chromosomes in the sampled population. In the special case of a region of high linkage disequilibrium (LD), or low recombination, mutations stay connected during evolution. This pattern of co-inheritance can provide a view of the order (or timing) for each mutation in a region of interest. The phylogeny illustrated in Fig. 1 is constructed of DNA sequences from the *APOE* LD region (50,092,405–50,101,584, NCBI version 36) on chro-

*Correspondence to: Allen D. Roses, R David Thomas Executive Training Center, One Science Drive, Suite 342, Box 90344, Durham, NC 27708, USA. Tel.: +1 919 660 8065; Fax: +1 919 681 9289; E-mail: allen.roses@duke.edu.

ISSN 1387-2877/11/$27.50

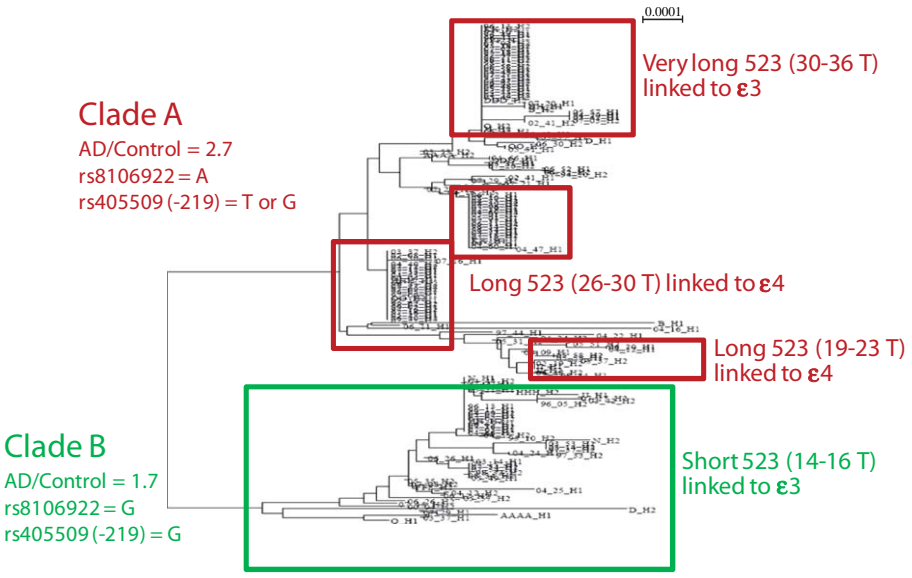

Fig. 1. The annotated phylogenetic tree shows the mapping of rs10524523 polyT lengths and *TOMM40* rs8106922 and *APOE* rs405509 (−219) SNPs. The case/control ratio for the two major clades is reported. For major clade B, which is associated with lower risk of LOAD and contains haplotypes linked to *APOE* ε3, rs10524523 lengths are uniformly short (mean = 15.9, SEM = 0.4, $n = 63$). In contrast, the polyT alleles in clade A are much longer: the average length of alleles connected to *APOE* ε4 is 27.2 (SEM = 0.29, $n = 55$, 'long') and the average length of polyT alleles connected to *APOE* ε3 is 31.57 (SEM = 0.48, $n = 90$, 'very long'). There is strong bootstrap support (973/1000) for the first major branch and moderate bootstrap support for the branches within clade A (247/1000 and 776/1000, respectively). The tree is obtained from a cohort of 105 patients (210 haplotypes) (Arizona cohort) [11].

mosome 19 from LOAD patients and controls. This genomic region includes the three *APOE* promoter polymorphisms (−219, −491 and −427) discussed by Lescai et al. [10]. Events such as mutation and recombination in an ancestral sequence, represented by a node in the tree, introduce diversity in each chromosome and the resultant lineages subsequently inherit the outcomes of these events. If an inherited mutation causes a phenotype, such as disease susceptibility, the phenotype will tend to be enriched in a lineage that arises from the genetic event [17]. Each of the leaves of the tree, as shown on the right of the Fig. 1, is composed of clusters of the most similar sequences from the sampled chromosomes. In Fig. 1, distinct lengths of a variable–length, deoxythymidine homopolymer (polyT), rs10524523 or 523, located in intron 6 of the *TOMM40* gene differentiate the leaves on this phylogeny. If a phylogeny is developed from a genomic region that is associated with disease, then each cluster of related haplotypes on the tree can be tested for association with a phenotype of interest, and the variants that distinguish each cluster can be further analyzed.

The phylogeny shown in Fig. 1 was originally published in our recent paper entitled, "A *TOMM40* variable length polymorphism predicts the age of late-onset Alzheimer's disease" [12]. The major findings of this analysis of LOAD were in Caucasians: 1) *APOE* ε4 was, almost without exception, linked to a long polyT repeat at rs10524523; 2) The ε3 allele of *APOE* was linked to either a short (<20 T residues) polyT *or* a very long (>30 T residues) polyT repeat; and 3) The 'long' (L) or 'very long' (VL) polyT alleles, which refer to homopolymer lengths greater than 20 and 30 T residues respectively, were associated with earlier age of disease onset in *APOE* ε3/4 patients (∼7 years difference) and with higher disease risk. That is, linkage to a short or very long poly-T for an *APOE* ε3 carrier resulted in later or earlier onset of LOAD, respectively, regardless of risk imparted by *APOE* ε4 [11]. This finding has now been extended to other *APOE* genotypes, including *APOE* ε3/3 patients, where it has been demonstrated that those who are homozygous for the longer alleles have disease onset an average of 9 years earlier than those who carry one short polyT [18].

We mapped the three *APOE* promoter polymorphisms (−219, −491 and −427), discussed by Lescai et al. [10] to the data for the three clinical cohorts in our 2009 publication. The three SNPs in the *APOE* promoter tend to be in linkage with specific polyT lengths. In Caucasians, the strongest linkages are between

−219 G and short polyTs and −219 T and long or very long polyTs. The phylogeny in Fig. 1 is further annotated here with the alleles of the −219 promoter SNPs. The two major conclusions of this mapping are: 1) The T allele of −219 always maps to clade A. This clade is associated with a higher risk for LOAD and an earlier age of onset for LOAD relative to clade B; and 2) the T allele of −219 always co-occurs with either a long 523 allele linked to *APOE* ε4 or a very long 523 allele linked to *APOE* ε3. The association between the two alleles of −219 and distinct lengths of 523 potentially extends the findings of Lescai et al. [10] from individuals possessing the *APOE* ε4 allele to ε3 carriers as well.

The rs8106922 SNP in intron 6 of the *TOMM40* gene remains the one variant within the 10 kilobase genomic region that best separates clades A and B in the consensus phylogenetic tree constructed for the *APOE* LD region. In the 105 subject cohort from the Arizona Alzheimer's Disease Research Center (ADCC, [12]), the G allele of rs8106922 is linked to the G allele of −219 for 100% of all haplotypes. For 82% of the clade A haplotypes, the A allele of rs8106922 is linked to the T allele of −219. The remainder of clade A haplotypes have the A allele of rs8106922 linked to −219 G. The alleles of this *TOMM40* SNP are therefore more characteristic of the two major clades. The correspondence between the indicated alleles of these two SNPs is statistically significant ($p < 0.0001$) by Fisher's exact test. Although rs8106922 distinguishes the two major clades better, −219 and several other SNPs in this region are also highly significant in terms of differentiating the clades. Interestingly, a meta-analysis of five LOAD case-control studies indicates that rs8106922 is significantly associated with disease risk, with a study-wide allelic odds ratio of 2.7 [19]. Several other SNPs in *TOMM40* are highly significantly associated with LOAD, some with allelic odds ratios in the 2.8–3.0 range, which provides further evidence for a role for this gene in LOAD [20–23].

The *APOE* promoter SNP, −219, described by Lescai et al. [10], is not significantly associated with age of LOAD onset in the *APOE* ε3/4 cohort that we analyzed for age of onset in 2009 ($p > 0.51$). However, we do identify a trend that is in accord with their findings: we calculate that the age of LOAD onset for individuals who have the TT −219 genotype is earlier (70.1 ± 1.8, $n = 15$) than the age of onset for the GT genotype (73.1 ± 1.7, $n = 12$) or GG genotype (72.0 ± 3.0, $n = 7$) in our cohort. In contrast, the 523-polyT genotype is a better age of onset classifier and is more statistically significant. That is, for *APOE* ε3/4 patients who developed LOAD after age 60, the individuals with long polyT repeats linked to *APOE* ε3 developed LOAD an average of 7 years earlier than individuals with shorter poly-T repeats linked to *APOE* ε3 (70.5 years ± 1.2 versus 77.6 years ± 2.1, $p = 0.02$, $n = 34$) [12].

TOMM40 and *APOE* are in LD on chromosome 19, with the 3′ and 5′ ends, respectively, of the genes, separated by only ∼2 kilobases [7, 24]. The proximity of the genes and the LD within the region may obscure disease risk associated with variants other than *APOE* ε4. As noted by Lescai et al. [10], the amount of ApoE protein in cerebral spinal fluid of non-demented individuals and in postmortem brain of LOAD patients is associated with SNPs within *TOMM40* and not with *APOE* SNPs, suggesting that an *APOE*-regulatory element may reside in *TOMM40* [25, 26].

In addition to the genetics, there are other data indicating that Tom40, the protein encoded by *TOMM40*, is directly involved in LOAD pathogenesis. Tom40 is the transmembrane channel subunit of the Translocase of the Outer Mitochondrial Membrane (TOM) complex. Almost all proteins involved in mitochondrial structure and function are encoded by nuclear genes and are translocated into mitochondria through this pore. Because >98% of the 1100 to 1500 known mitochondrial proteins [27] are nuclear encoded and enter mitochondria through Tom40 [28], this protein plays a critical role in mitochondrial physiology and biogenesis and therefore overall cellular health and functioning. A role for Tom40 in mitochondrial pathogenesis, proposed to be a precipitating event in Alzheimer's disease, is indicated by the finding of amyloid-β protein precursor (AβPP) lodged in the TOM channel in LOAD brain samples. This blockage results in mitochondrial dysfunction due to obstruction of the transport of nuclear-encoded cytochrome c oxidase (COX) subunits into the mitochondrion [29, 30]. Decreased COX activity gives rise to increased generation of reactive oxygen species and reduced ATP production. AβPP-induced mitochondrial toxicity is observed in samples from LOAD patients and not in samples from normal, age-matched controls, and the extent of bound AβPP is positively correlated with disease severity. Tom40 also mediates the passage of Aβ peptide, a fragment of AβPP, from the cytosol into mitochondria, where it induces mitotoxicity by inhibiting COX [31, 32]. There is, therefore, genetic and cell biological evidence that *TOMM40* contributes to LOAD risk or pathogenesis.

In summary, Lescai and colleagues [10] present important data on the role of SNPs in the *APOE* promoter and their possible interaction with variants of the

APOE gene. Moreover, the report makes a compelling case for considering the *cis* arrangement of alternative alleles in analyses of disease phenotypes, as different allelic combinations may engender different biology. Analyzing the *APOE* promoter SNPs in the context of the molecular evolution of the extended LD interval shows the linkage between specific *APOE* promoter alleles and the *APOE* ε3-very long 523 haplotype that is associated with LOAD risk and earlier disease onset, in addition to the interaction of the promoter alleles with *APOE* ε4 that Lescai et al. describe. The data presented by Lescai et al. [10], and in this Commentary, provide additional support for rs10524523 as a predictive marker for disease onset. This genetic marker provides a viable strategy for enrichment of clinical trials to test therapies that may alleviate or offset mitochondrial dysfunction and thereby delay the onset of LOAD.

DISCLOSURE STATEMENT

Authors' disclosures available online (http://www.j-alz.com/disclosures/view.php?id=708).

REFERENCES

[1] Saunders AM, Strittmatter WJ, Schmechel D, St George-Hyslop PH, Pericak-Vance MA, Joo SH, Rosi BL, Gusella JF, Crapper-MacLachlan DR, Alberts MJ, Hulette C, Crain B, Goldgaber D, Roses AD (1993) Association of apolipoprotein E allele ε4 with late-onset familial and sporadic Alzheimer's disease. *Neurology* **43**, 1467-1472.

[2] Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, Berg S, Fiske A, Pedersen NL (2006) Role of genes and environments for explaining Alzheimer disease. *Arch Gen Psychiatry* **63**, 168-174.

[3] Hollingworth P, Harold D, Jones L, Owen MJ, Williams J (2010) Alzheimer's disease genetics: current knowledge and future challenges. *Int J Geriatr Psychiatry*, doi: 10.1002/gps.2628.

[4] Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, Roses AD, Haines JL, Pericak-Vance MA (1993) Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* **261**, 921-923.

[5] Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, Myers RH, Pericak-Vance MA, Risch N, van Duijn CM (1997) Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. *JAMA* **278**, 1349-1356.

[6] Abraham R, Moskvina V, Sims R, Hollingworth P, Morgan A, Georgieva L, Dowzell K, Cichon S, Hillmer AM, O'Donovan MC, Williams J, Owen MJ, Kirov G (2008) A genome-wide association study for late-onset Alzheimer's disease using DNA pooling. *BMC Med Genomics* **1**, 44.

[7] Li H, Wetten S, Li L, Jean PLS, Upmanyu R, Surh L, Hosford D, Barnes MR, Briley JD, Borrie M, Coletta N, Delisle R, Dhalla D, Ehm MG, Feldman HH, Fornazzari L, Gauthier S, Goodgame N, Guzman D, Hammond S, Hollingworth P, Hsiung GY, Johnson J, Kelly DD, Keren R, Kertesz A, King KS, Lovestone S, Loy-English I, Matthews PM, Owen MJ, Plumpton M, Pryse-Phillips W, Prinjha RK, Richardson JC, Saunders A, Slater AJ, George-Hyslop PHS, Stinnett SW, Swartz JE, Taylor RL, Wherrett J, Williams J, Yarnall DP, Gibson RA, Irizarry MC, Middleton LT, Roses AD (2008) Candidate single-nucleotide polymorphisms from a genomewide association study of Alzheimer disease. *Arch Neurol* **65**, 45-53.

[8] Yu C-E, Seltman H, Peskind ER, Galloway N, Zhou PX, Rosenthal E, Wijsman EM, Tsuang DW, Devlin B, Schellenberg GD (2007) Comprehensive analysis of APOE and selected proximate markers for late-onset Alzheimer's disease: Patterns of linkage disequilibrium and disease/marker association. *Genomics* **89**, 655-665.

[9] Waring SC, Rosenberg RN (2008) Genome-wide association studies in Alzheimer disease. *Arch Neurol* **65**, 329-334.

[10] Lescai F, Chiamenti AM, Codemo A, Pirazzini C, D'Agostino G, Ruaro C, Ghidoni R, Benussi L, Galimberti D, Esposito F, Marchegiani F, Cardelli M, Olivieri F, Nacmias B, Sorbi S, Tagliavini F, Albani D, Boneschi FM, Binetti G, Santoro A, Forloni G, Scarpini E, Crepaldi G, Gabelli C, Franceschi C (2011) An APOE haplotype associated to decreased ε4 expression increases the risk of late onset Alzheimer's disease. *J Alzheimers Dis* **24**, 235-245.

[11] Lutz MW, Crenshaw DG, Saunders AM, Roses AD (2010) Genetic variation at a single locus and age of onset for Alzheimer's disease. *Alzheimers Dement* **6**, 125-131.

[12] Roses AD, Lutz MW, Amrine-Madsen H, Saunders AM, Crenshaw DG, Sundseth SS, Huentelman MJ, Welsh-Bohmer KA, Reiman EM (2010) A TOMM40 variable-length polymorphism predicts the age of late-onset Alzheimer's disease. *Pharmacogenomics J* **10**, 375-384.

[13] Jordan IK, Kondrashov FA, Rogozin IB, Tatusov RL, Wolf YI, Koonin EV (2001) Constant relative rate of protein evolution and detection of functional diversification among bacterial, archaeal and eukaryotic proteins. *Genome Biol* **2**, RESEARCH0053.

[14] Wan QH, Zeng CJ, Ni XW, Pan HJ, Fang SG (2009) Giant panda genomic data provide insight into the birth-and-death process of mammalian major histocompatibility complex class II genes. *PLoS One* **4**, e4147.

[15] Watts S, Vogel J, Harriman W, Itoh T, Stauss H, Goodenow R (1987) DNA sequence analysis of the C3H H-2Kk and H-2Dk loci. Evolutionary relationships to H-2 genes from four other mouse strains. *J Immunol* **139**, 3878-3885.

[16] Yu H, Zhang PC, Zhou YJ, Li GX, Pan J, Yan LP, Shi XX, Liu HL, Tong GZ (2009) Isolation and genetic characterization of avian-like H1N1 and novel ressortant H1N2 influenza viruses from pigs in China. *Biochem Biophys Res Commun* **386**, 278-283.

[17] Zöllner S, Pritchard JK (2005) Coalescent-based association mapping and fine mapping of complex trait loci. *Genetics* **169**, 1071-1092.

[18] Caselli R, Saunders A, Lutz M, Huentelman M, Reiman E, Roses A (2010) TOMM40, APOE, and age of onset of incident Alzheimer's disease. *Alzheimers Dement* **6**, s202.

[19] Grupe A, Abraham R, Li Y, Rowland C, Hollingworth P, Morgan A, Jehu L, Segurado R, Stone D, Schadt E, Karnoub M, Nowotny P, Tacey K, Catanese J, Sninsky J, Brayne C, Rubinsztein D, Gill M, Lawlor B, Lovestone S, Holmans P, O'Donovan M, Morris JC, Thal L, Goate A, Owen MJ, Williams J (2007) Evidence for novel susceptibility genes for

late-onset Alzheimer's disease from a genome-wide association study of putative functional variants. *Hum Mol Genet* **16**, 865-873.

[20] Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, Pahwa JS, Moskvina V, Dowzell K, Williams A, Jones N, Thomas C, Stretton A, Morgan AR, Lovestone S, Powell J, Proitsi P, Lupton MK, Brayne C, Rubinsztein DC, Gill M, Lawlor B, Lynch A, Morgan K, Brown KS, Passmore PA, Craig D, McGuinness B, Todd S, Holmes C, Mann D, Smith AD, Love S, Kehoe PG, Hardy J, Mead S, Fox N, Rossor M, Collinge J, Maier W, Jessen F, Schurmann B, van den Bussche H, Heuser I, Kornhuber J, Wiltfang J, Dichgans M, Frolich L, Hampel H, Hull M, Rujescu D, Goate AM, Kauwe JSK, Cruchaga C, Nowotny P, Morris JC, Mayo K, Sleegers K, Bettens K, Engelborghs S, De Deyn PP, Van Broeckhoven C, Livingston G, Bass NJ, Gurling H, McQuillin A, Gwilliam R, Deloukas P, Al-Chalabi A, Shaw CE, Tsolaki M, Singleton AB, Guerreiro R, Muhleisen TW, Nothen MM, Moebus S, Jockel K-H, Klopp N, Wichmann HE, Carrasquillo MM, Pankratz VS, Younkin SG, Holmans PA, O'Donovan M, Owen MJ, Williams J (2009) Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat Genet* **41**, 1088-1093.

[21] Potkin SG, Guffanti G, Lakatos A, Turner JA, Kruggel F, Fallon JH, Saykin AJ, Orro A, Lupoli S, Salvi E, Weiner M, Macciardi F, for the Alzheimer's Disease Neuroimaging I (2009) Hippocampal atrophy as a quantitative trait in a genome-wide association study identifying novel susceptibility genes for Alzheimer's disease. *PLoS ONE* **4**, e6501.

[22] Shen L, Kim S, Risacher SL, Nho K, Swaminathan S, West JD, Foroud T, Pankratz N, Moore JH, Sloan CD, Huentelman MJ, Craig DW, DeChairo BM, Potkin SG, Jack CR Jr., Weiner MW, Saykin AJ (2010) Whole genome association study of brain-wide imaging phenotypes for identifying quantitative trait loci in MCI and AD: A study of the ADNI cohort. *NeuroImage* **53**, 1051-1056.

[23] Takei N, Miyashita A, Tsukie T, Arai H, Asada T, Imagawa M, Shoji M, Higuchi S, Urakami K, Kimura H, Kakita A, Takahashi H, Tsuji S, Kanazawa I, Ihara Y, Odani S, Kuwano R (2009) Genetic association study on in and around the APOE in late-onset Alzheimer disease in Japanese. *Genomics* **93**, 441-448.

[24] Lai E, Riley J, Purvis I, Roses A (1998) A 4-Mb high-density single nucleotide polymorphism-based map around human APOE. *Genomics* **54**, 31-38.

[25] Bekris LM, Millard SP, Galloway NM, Vuletic S, Albers JJ, Li G, Galasko DR, DeCarli C, Farlow MR, Clark CM, Quinn JF, Kaye JA, Schellenberg GD, Tsuang D, Peskind ER, Yu C-E (2008) Multiple SNPs within and surrounding the apolipoprotein E gene influence cerebrospinal fluid apolipoprotein E protein levels. *J Alzheimers Dis* **13**, 255-266.

[26] Bekris LM, Galloway NM, Montine TJ, Schellenberg GD, Yu C-E (2010) *APOE* mRNA and protein expression in postmortem brain are modulated by an extended haplotype structure. *Am J Med Genet Part B* **153B**, 409-417.

[27] Lopez JJ, Salido GM, Pariente JA, Rosado JA (2008) Thrombin induces activation and translocation of Bid, Bax and Bak to the mitochondria in human platelets. *J Thromb Haemost* **6**, 1780-1788.

[28] Perry AJ, Rimmer KA, Mertens HD, Waller RF, Mulhern TD, Lithgow T, Gooley PR (2008) Structure, topology and function of the translocase of the outer membrane of mitochondria. *Plant Physiol Biochem* **46**, 265-274.

[29] Devi L, Prabhu BM, Galati DF, Avadhani NG, Anandatheerthavarada HK (2006) Accumulation of amyloid precursor protein in the mitochondrial import channels of human Alzheimer's disease brain is associated with mitochondrial dysfunction. *J Neurosci* **26**, 9057-9068.

[30] Anandatheerthavarada HK, Biswas G, Robin MA, Avadhani NG (2003) Mitochondrial targeting and a novel transmembrane arrest of Alzheimer's amyloid precursor protein impairs mitochondrial function in neuronal cells. *J Cell Biol* **161**, 41-54.

[31] Hansson O, Zetterberg H, Vanmechelen E, Vanderstichele H, Andreasson U, Londos E, Wallin A, Minthon L, Blennow K (2008) Evaluation of plasma Abeta(40) and Abeta(42) as predictors of conversion to Alzheimer's disease in patients with mild cognitive impairment. *Neurobiol Aging* **3**, 357-367.

[32] Manczak M, Anekonda TS, Henson E, Park BS, Quinn J, Reddy PH (2006) Mitochondria are a direct site of A beta accumulation in Alzheimer's disease neurons: implications for free radical generation and oxidative damage in disease progression. *Hum Mol Genet* **15**, 1437-1449.