Genome-wide Association Study Links *APOE*ε4 and *BACE1* Variants with Plasma Amyloid β Levels

Vincent Chouraki1,2,3,* Sven J van der Lee4,5,*
Benjamin Grenier-Boley1,2, Jeannette Simino6, Hieab Adams7, Giuseppe Tosto8,9, Charles White10,11, Natalie Terzikhan5,12, Carlos Cruchaga13, Maria J. Knol5, Shuo Li14,15, Susanna Schraen16, Megan L. Grove17, Claudia Satizabal3,15, Najaf Amin5, Claudine Berr18, Steven Younkin19,
Alzheimer's Disease Neuroimaging Initiative, Rebecca F. Gottesman20,21, Luc Buée2,22, Alexa Beiser3,14,15, David S. Knopman23, Andre Uitterlinden24, Charles DeCarli25, Jan Bressler17, Anita DeStefano3,14,15, Jean-François Dartigues26, Qiong Yang14,15, Eric Boerwinkle17,27, Christophe Tzourio26, Myriam Fornage17,28, M Arfan Ikram7, Philippe Amouyel1,2, Phil de Jager10,11,29, Chritiane Reitz8,9,30,31, Thomas H Mosley, Jr.32, Jean-Charles Lambert1,2,** Sudha Seshadri3,15,** Cornelia van Duijn5**
* contributed equally to this work
** contributed equally to this work

Information

Affiliations

1. Lille University, Inserm, Lille University Hospital, Institut Pasteur de Lille, U1167 - RID-AGE - Risk factors and molecular determinants of aging-related diseases
2. Labex Distalz, F-59000 Lille, France
3. Department of Neurology, Boston University School of Medicine, Boston, MA, USA
4. Alzheimer Center, Department of Neurology, VU University Medical Center, Neuroscience Campus Amsterdam, Amsterdam, The Netherlands
5. Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands
6. Gertrude C Ford MIND Center, Department of Data Science, John D. Bower School of Population Health, University of Mississippi Medical Center, Jackson, MS, USA
7. Departments of Epidemiology, Neurology, and Radiology and Nuclear Medicine, Erasmus Medical Center, Rotterdam, the Netherlands
8. Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University, New York, NY, USA
9. Gertrude H. Sergievsky Center, Columbia University, New York, NY, USA

10. Program in Translational NeuroPsychiatric Genomics, Institute for the Neurosciences, Departments of Neurology and Psychiatry, Brigham and Women's Hospital, Boston, Massachusetts, United States of America
11. Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts, United States of America
12. Department of Respiratory Medicine, Ghent University Hospital, Ghent, Belgium
13. Department of Psychiatry, Washington University in St. Louis, Saint Louis, Missouri, USA
14. Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA
15. The Framingham Heart Study, Framingham, MA, USA
16. Université Lille, CHU-Lille, Inserm, UF de Neurobiologie, CBPG, Lille, France
17. Human Genetics Center, Department of Epidemiology, Human Genetics, and Environmental Sciences, School of Public Health, The University of Texas Health Science Center at Houston, Houston, Texas, USA
18. INSERM U1061, University of Montpellier, Montpellier, France
19. Department of Neuroscience, Mayo Clinic, Jacksonville, Florida 32224, USA
20. Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD USA
21. Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA
22. Université de Lille, Institut National de la Santé et de la Recherche Medicale (INSERM), CHU Lille, UMR-S 1172 JPArc, Lille, France
23. Department of Neurology, Mayo Clinic College of Medicine, Rochester, MN, USA.
24. Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands
25. Department of Neurology, University of California at Davis, Davis, CA, USA
26. INSERM, UMR1219, Bordeaux University, Bordeaux Population Health Research Center, France
27. Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas, USA
28. Brown Foundation Institute of Molecular Medicine, McGovern Medical School, The University of Texas Health Science Center at Houston, Houston, Texas, USA
29. Center for Translational & Systems Neuroimmunology, Department of Neurology, Columbia University Medical Center, New York, New York, United States of America
30. Department of Neurology, Columbia University, New York, NY, USA
31. Department of Epidemiology, Columbia University, New York, NY, USA

32. Gertrude C Ford MIND Center, Department of Medicine, University of Mississippi Medical Center, Jackson, MS, USA

Corresponding authors' information

Vincent Chouraki, MD, PhD
UMR1167
Institut Pasteur de Lille,
1 rue du Professeur Calmette
59019, Lille, France
tel: +33 3 20 87 73 27
fax: +33 3 20 87 78 94

Running Title

A 1000 Genomes GWAS of plasma Aβ levels

Keywords

Plasma amyloid β levels; genetic epidemiology; genome-wide association study; *APOE*; *BACE1*; Alzheimer's disease; preclinical biomarkers; plasma biomarkers;

Contributions

Study Design: SS, VC; Study Management: ADNI, CB, CC, CR, CT, CvD, DK, EB, JFD, MAI, PA, PdJ, RG, SS, TM; Subject Recruitment: ADNI, CB, CC, CR, CT, CvD, JFD, MAI, PA, PdJ, SS, TM; Interpretation of Results: CvD, JCL, SS, VC; Drafting of the Manuscript: CvD, JCL, SS, SvdL, VC; Critical Review of the Manuscript: AB, AdS, AU, BGB, CB, CC, CdC, CR, CS, CT, CvD, CW, DK, EB, GT, HA, JB, JCL, JFD, JS, LB, MAI, MF, MG, MJK, NT, PA, PdJ, QY, RG, SL, SS, SSc, SvdL, SY, TM, VC; Statistical Methods and Analysis: AB, BGB, CW, GT, HA, JS, MJK, NT, QY, SvdL, VC; Genotyping: ADNI, CC, CR, CvD, EB, JCL, MG, PA, PdJ, SS; Aβ measurement: ADNI, CC, CR, LB, SSc, SY, TM; Bioinformatics: BGB, SvdL, VC

Abstract

Background: Amyloid β (Aβ) peptides are the products of the catalytic processing of the Aβ precursor protein (APP) by the β-secretase, BACE1 and the γ-secretase complex. Impairment of the Aβ production/clearance balance is the major pathophysiological hypothesis in Alzheimer's disease (AD). Plasma Aβ levels are easy to measure in large numbers and therefore can be used as an endophenotype to study the genetics of Aβ and its relevance to AD.

Methods: We performed genome-wide association studies (GWAS) of plasma Aβ1-40, Aβ1-42 and Aβ1-42/Aβ1-40 ratio in 12,369 non-demented participants across 8 studies, using genetic data imputed on the 1000 Genomes phase 1 version 3 reference panel. To gain further insight, we performed LD-score regression analysis of plasma Aβ-42 and Aβ-40 levels using previously published GWAS of AD and other related traits, and pathway analyses.

Results: We identified 21 variants reaching genome-wide significance across two loci. The most significant locus spanned the APOE gene, with significant associations with plasma Aβ42 levels ($p = 9.01\times10^{-13}$) and plasma Aβ42/Aβ40 ratio ($p = 6.46\times10^{-20}$). The second locus was located on chromosome 11, near the BACE1 gene ($p = 2.56\times10^{-8}$). We also observed suggestive evidence of association ($p < 1\times10^{-5}$) around genes involved in Aβ metabolism including *APP* and *PSEN2*.

Conclusion: Using plasma Aβ40 and Aβ42 levels, this GWAS was able to identify relevant and central actors of the APP metabolism in AD. Overall, this study strengthens the utility of plasma Aβ levels both as an endophenotype and a biomarker.

Introduction

Amyloid β (Aβ) peptides are the products of the catalytic processing of the Aβ precursor protein (APP) by the β-secretase, BACE1 and the γ-secretase complex. [1] Aβ peptides are mainly produced in the brain where APP and BACE1 are both highly expressed, [1] but also in circulating blood platelets [2] and in the pancreas. [1] Aβ peptides are able to self-assemble in soluble Aβ oligomers but also in insoluble fibrils that can aggregate as plaques in the brain parenchyma or in the wall of pial blood vessels where they constitute defining hallmarks of Alzheimer's disease (AD) [3] and cerebral amyloid angiopathy (CAA), [4] respectively.

There is strong evidence pointing toward a central role of Aβ peptides in the pathophysiology of AD, although the exact role remains controversial. [5] Studies have shown that a large variety of individually rare mutations in genes involved in Aβ production, including *APP*, *PSEN1* and *PSEN2*, lead to autosomal dominant early-onset forms of AD (EOAD) and to lobar hemorrhage from cerebral amyloid angiopathy. [6] Moreover, Apolipoprotein E (*APOE*) ε4, the major genetic risk factor for late-onset AD (LOAD) in the general population, [7] has been implicated in Aβ aggregation, deposition and clearance, both in brain and in blood vessels. [8]

These findings are the basis of drug discovery efforts targeting Aβ production and clearance that are currently under consideration in clinical trials, although unfortunately with negative results up to now. [9] Although these results appear to contradict the amyloid hypothesis, several explanations have been advanced. [10] One major concern is the ineffectiveness of intervention after the onset of clinical symptoms which likely result from irreversible neurodegeneration. Since the amyloid accumulation process precedes the clinical onset by decades, [11,12] early intervention in high risk groups such as Down syndrome patients and APOEε4/ε4 carriers remains to be evaluated. Another concern is the lack of knowledge concerning the precise mechanisms involving Aβ in the pathophysiology of LOAD. Indeed, except for *APOE* and *SORL1*, only a small number of genes identified by genome-wide association studies (GWAS) of LOAD [13] have been linked to APP metabolism [14,15] and Aβ-related pathways have not yet emerged in formal enrichment analyses. [16] Moreover, the genes involved in autosomal dominant forms of EOAD have not been detected in GWAS of LOAD.

To address these questions regarding the amyloid hypothesis, we and others have directly explored the genetics of Aβ through GWAS on quantitative measures of Aβ peptides, either in the cerebrospinal fluid (CSF) or in the brain, through Pittsburgh Compound B (PiB) PET scan or autopsy. [17-21] Combining the effect of AD genetic loci resulted in statistically significant effects on CSF Aβ42, suggesting that amyloid metabolism is also involved in LOAD. [21] Nevertheless, these studies are limited in their sample size due to low acceptability of lumbar puncture and brain donation and high cost of PiB PET, and therefore may lack statistical power necessary in genetic association research. Aβ peptides produced in the brain can be degraded locally or transported into the CSF and the blood stream where they can be easily detected. [22] Although the brain-derived Aβ peptides in the circulation cannot be distinguished from Aβ derived from blood platelets or pancreas, plasma Aβ levels are modestly but significantly correlated with amyloid burden in the CSF and in the brain. [23,24] Our groups

have also independently shown that plasma Aβ concentrations are prospectively associated with the future risk of developing AD, [25-28] suggesting that there is indeed a link between mechanisms controlling Aβ concentrations in plasma and AD pathophysiological processes in the brain and that circulating Aβ peptides can be used as a marker for brain amyloid metabolism. Alternatively, it has also been postulated that the relation of plasma Aβ with subsequent AD is reflecting general physiological processes in the brain, platelets and kidney, thus giving information on Aβ peptides production/clearance in each of these tissues. [29]

In this context, we set out to discover genetic determinants of circulating Aβ. We previously conducted a GWAS of plasma Aβ levels in 3,528 non-demented participants, but failed to find genome-wide significant associations. [29] The present study extends our previous work by performing a GWAS using a sample size (n=12,369) that is more than three times larger.

Methods

Study population

We included data from 12,369 European-descent participants from eight studies, the Framingham Heart Study (FHS; n=6,735), the Rotterdam study (n=1,958), the Three City Study (3C; n=1,954), the Atherosclerosis Risk in Communities Study (ARIC; n=830), the Washington Heights-Inwood Community Aging Project (WHICAP; n=193), the Epidemiological Prevention study Zoetermeer (EPOZ; n=397), the Alzheimer's Disease Neuroimaging Initiative (ADNI; n=173) and the Erasmus Rucphen Family Study (ERF; n=129). In each study, we excluded participants with prevalent dementia at the time of blood sampling used for plasma Aβ assessment (see Supplementary Methods 1 for a detailed description of each study).

Plasma Aβ assessment

Each study used different protocols for blood sampling, plasma extraction and storage and plasma Aβ assessment that have been detailed in previous publications. [25,26,28,30-32] In the FHS, Rotterdam and 3C study, plasma Aβ levels were measured at different times because of cost considerations. Various assays were used to quantify plasma Aβ1-40 and Aβ1-42 levels (see Supplementary Methods 2 for a detailed description of the protocols used in each study and Supplementary Table 1 for baseline characteristics of the study populations).

Genotyping

Each study used different genotyping platforms as previously published. [13] After applying pre-imputation variant and sample filters, genotypes were imputed using the 1000 Genomes phase 1 version 3 (all ethnicities) imputation panel and various imputation pipelines (see Supplementary Methods 3). *APOEε* genotyping was performed as part of protocols specific to each study (see Supplementary Methods 4).

Statistical analyses

Plasma Aβ levels

Plasma Aβ levels were expressed as pg/mL. In each study and for each Aβ dosage, we excluded values that were over or below 4 standard deviations around the mean. To study the variations of plasma Aβ levels in a consistent way across studies, we performed a ranked-based inverse normal transformation of plasma Aβ levels in each study. If they were significantly associated with plasma Aβ levels, this transformation was performed after adjusting for batch effect and other technical artifacts.

Genome-wide association studies

Each study performed genome-wide association studies of plasma Aβ1-40 and Aβ1-42 levels and Aβ1-42/Aβ1-40 ratio using 1000 Genomes imputed data. According to the imputation pipelines used, genetic information was available either as allele dosages or genotype probabilities. In each study, we excluded results from variants that had low imputation quality (r2 or info score < 0.3), variants with low frequency (minor allele frequency < 0.005 or minor allele count < 7) and variants that were available in small number of participant (n < 30). Association of genetic variations with plasma Aβ levels were assessed in linear regression models adjusted for sex and age at blood collection. If significantly associated with plasma Aβ levels, principal components were added in the models to account for population structure.

Genome-wide meta-analysis

Before meta-analysis, we applied a series of filters and quality check that were previously published (see Supplementary Figures 1 and 2). [33] We performed an inverse variance weighted genome-wide meta-analysis, accounting for genomic inflation factors using the METAL software. [34] Finally, we retained variants that had been meta-analyzed at least in the 3 largest available populations (FHS (n=6,735), Rotterdam Study (RS; n=1,958) and Three City Study (3C; n=1,954)). Statistical significance was defined as a p-value below 5×10-8. Signals with p-values between 1×10-5 and 5×10-8 were considered suggestive. Additional graphes and analyses were done using R v3.4.1 (Vienna, Austria).

Confirmation of the *APOEε4* signal

To confirm the APOE signal we obtained in our genome-wide meta-analysis, we reran our analysis using genotyped *APOE*ε4 and *APOE*ε2 status, adjusting for age and sex.

Annotation

Variant information were retrieved using the Feb 2009 (grch37) assembly of the human genome and dbSNP v147 in the UCSC Table Browser web tool (https://genome.ucsc.edu/cgi-bin/hgTables accessed on 2017-07-25)and the CADD database version 1.3 (http://cadd.gs.washington.edu/download accessed on 2017-07-26). [35] We considered that variants that were less than 250kb apart from one another belonged to the same locus. We then used the Ensembl Variant Effect Predictor (VEP, http://grch37.ensembl.org/info/docs/tools/vep/index.html accessed on 2017-07-25) [36] to relate those variants to nearby genes of potential interest. We also searched if those variants were also eQTL for nearby genes using data from the Genotype-Tissue Expression (GTEx) project (https://gtexportal.org/home/) [37] using the Ensembl REST API (http://rest.ensembl.org/documentation/info/species_variant, data retrieved on 2017-07-31). We corrected for multiple testing by computing False Discovery Rate (FDR), using an FDR threshold of 0.05. Finally, we cross-checked our results with previously published GWAS of AD [13] amyloid-related brain pathology, [19] and CSF Aβ42 levels. [21]

Genetic correlations

To gain further insight, we used the LD SCore software v1.0.0 [38,39] and previously published GWAS to compute genetic correlations between plasma Aβ levels and ratio and AD, [13] Parkinson's disease, [40] cognition, [41] hippocampal volume, [42] intracranial volume, [43] white matter lesions, [44] brain lobes volumes (unpublished data) and subcortical brain structures volumes. [45]

Pathway over-representation analysis

We used the ConsensusPathDB-human website (http://cpdb.molgen.mpg.de/ accessed on 2017-08-01) [46] and a curated list of genes related to Aβ [47] to check whether the genes we annotated using GTEx were over-represented in biochemical pathways (see Supplementary Table 2 for the complete list of genes). In order to detect new pathways, we performed a second pathway analysis after excluding genes from loci known for their involvement in Aβ metabolism, namely *APOE* (including *APOE*, *PVRL2*, *NKPD1*, *CTB-129P6.4*, *APOC1 and VASP*), *APP* (including APP, *AP000230.1* and *AP001596.6*), *PSEN2* (including *PSEN2* and *ADCK3*) and *BACE1* (including *BACE1*, *SIDT2*, *TAGLN*, *RP11-109L13.1*, *RNF214* and *CEP164*). Statistical significance was assessed using hypergeometric tests and corrected for multiple testing using Q-value.

Results

Genome-wide significant variants associated with plasma Aβ levels

After meta-analysis, we identified 21 variants reaching genome-wide significance across two loci (Supplementary Figures 3 to 8).

The first locus was located on chromosome 19, in the APOE gene, with significant associations with plasma Aβ1-42 levels and plasma Aβ1-42/Aβ1-40 ratio (Figures 1 and 2). For both associations, the most significant variant was rs429358, with p-values of 9.01 × 10-13 and 6.46 × 10-20 for Aβ1-42 levels and Aβ1-42/Aβ1-40 ratio, respectively (Table 1). The minor allele of this variant, which denotes the *APOE*ε4, was associated with lower plasma Aβ1-42 levels (effect size=-0.167 standard deviations (SD); 95% confidence interval (CI)=[-0.212 ; -0.121]) and lower plasma Aβ1-42/Aβ1-40 ratio (effect size=-0.212 SD; 95% CI=[-0.257 ; -0.121]; Table1 and Supplementary Figure 9). We confirmed these associations using the directly genotyped *APOE*ε4 status (Supplementary Figure 10). We did not find significant associations between genotyped *APOE*ε2 and circulating Aβ peptides levels, despite the protective effect of this variant on the risk of AD (Supplementary Figure 10).

The second genome-wide significant locus was located on chromosome 11, near the BACE1 gene that encodes the β-secretase and is involved in the initial, Aβ-producing step of APP processing (Figure 3). For the most significant variant, rs650585 (CADD score=6), the minor allele was associated with lower plasma Aβ1-40 levels (effect size=-0.073 SD; 95%CI=[-0.099 ; -0.047]; p-value=2.56 × 10-8; Table1 and Supplementary Figure 9). This variant was in moderate LD (R2=0.58) with a *BACE1* synonymous variant, rs638405 (CADD score=11), which was also associated with plasma Aβ1-40 levels (effect size=-0.071 SD, p-value=1.21 × 10-7). Both variants were associated with BACE1 expression in testis (FDR=4.52 × 10-11 and FDR=2.62 × 10-38, respectively). rs638405 was also associated with CEP164 expression in the cerebellum (FDR=2.05 × 10-6; Supplementary Table 2). Of note, this locus also contained a missense variant in the CEP164 gene that was associated with plasma Aβ1-40 levels with suggestive levels of significance (rs573455, effect size=-0.057 SD; p-value=9.16 × 10-6).

GWAS suggestive hits

Besides genome-wide significant signals at the *APOE* and *BACE1* loci, signals reaching suggestive levels of association (p<1 × 10-5) for at least one of the three plasma Aβ measures were identified for 240 variants across 73 loci.

Interestingly, we found suggestive levels of association spread across three peaks within and nearby *APP*, one of the core genes of amyloid metabolism (Figure 4). Two independent variants located within APP were suggestively associated with plasma Aβ1-40 levels: rs150707803 (effect size=-0.184 SD; p-value=2.10 × 10-6) and rs436011

(effect size=0.061 SD; p-value=3.92 × 10-6) (Figure 4A and Table 2). SNPs within this second locus, including rs436011, were associated with APP expression in the "Esophagus – Muscularis" (FDR=6.29 × 10-7 for rs436011; Supplementary Table 2). The top variant of the third locus, rs199744263, was located upstream of *APP* (Figure 4B). This variant was associated with lower Aβ1-42/Aβ1-40 ratio (effect size=-0.075 SD; p-value=2.43 × 10-6; Table 2). SNPs from this locus were associated with brain expression of *AP000230.1*, a lincRNA located directly upstream of *APP* (e.g. for rs199744263, FDR=8.33 × 10-5 for cerebellar hemisphere expression; Supplementary Table 2).

In addition to *BACE1* and *APP*, we explored genetic associations around other genes closely involved with Aβ production, namely *PSEN1* and *PSEN2* which are part of the γ-secretase complex and *ADAM10*, which encode for the α-secretase, and is involved in a competing, non-Aβ producing, processing of APP. There was a suggestive association with plasma Aβ1-40 at the *PSEN2* locus (Table 3, Supplementary Figure 11). In this locus, the top variant, rs2246221 (effect size= 0.057 SD; p-value=8.09 × 10-6) was also associated with PSEN2 expression in spleen (FDR=2.35 × 10-6), thyroid (FDR=1.77 × 10-5), lung (FDR=8.08 × 10-5), skin (FDR=1.76 × 10-2) and transformed fibroblasts (FDR=6.25 × 10-23; Supplementary Table 2). Finally, we found a suggestive association with plasma Aβ1-40 within *RGS6* which is located approximately 600kb upstream of *PSEN1* but did not find any evidence of eQTL linking the two loci (Table 3, Supplementary Figure 12).

To minimize risk of false positive signal among the remaining suggestive loci, we prioritized 12 loci containing exonic variants or variants with a CADD score higher than 10 (Supplementary Table 3). Among these variants, rs11123523 (CADD=12.2), located near the TMEM37 gene, showed the lowest p-value. This variant was associated with plasma Aβ1-40 levels (effect size=-0.074; p-value=2.80 × 10-7) and was associated with expression of a nearby gene, *SCTR*, in thyroid (FDR=3.34 × 10-10), testis (FDR=9.45 × 10-3) and tibial nerve (FDR=1.63 × 10-2). The only exonic variant was rs704, a missense variant of the VTN gene associated with plasma Aβ1-40 levels (effect size=-0.060; p-value=5.81 × 10-6).

Genetic overlap with other Aβ-related traits and diseases

In order to put those genome-wide significant variants in the context of amyloid-related pathophysiology of AD, we compared them with results obtained from CSF Aβ42, AD brain pathology and AD risk in GWAS (Table 1). The *APOE*ε4 allele was also associated with lower CSF Aβ42 levels, higher brain levels of neuritic plaques, diffuse plaques and neurofibrillary tangles and higher risk of AD. For the *BACE1* locus, we did not find genome-wide significant or suggestive associations of rs650585 with any of the aforementioned traits. Among suggestive variants, no genome-wide significant or suggestive association was found for any of the other amyloid-related traits or for AD risk.

To improve statistical power, we performed a genetic correlation analysis of GWAS results of multiple traits using LD regression. On a genome-wide scale (Figure 5), we observed a nominally significant negative genetic correlation between variants modulating plasma Aβ1-40 levels and hippocampal volume (rg=-0.64, p=0.03), and a

positive correlation between variants modulating plasma Aβ1-42 levels and white matter lesions (rg=0.42, p=0.05).

Pathway over-representation analysis

After annotation using GTEx, we obtained a list of 41 genes located nearby our suggestive and significant signals that we used to perform a pathway over-representation analysis. Using ConsensusPathDB-human and a curated list of genes related to Aβ as references, we found significant over-represention in several biochemical pathways, related to Alzheimer's disease (ConsensusPathDB Q-value=5.73 × 10-4) and generation of Aβ (curated list p-value=4.18 × 10-3; ConsensusPathDB Q-value=4.44 × 10-4; Supplementary Table 4). These results were driven by genes previously known for their involvement in Aβ metabolism (*APOE*, *BACE1*, *APP*, *PSEN2*) as the AD and Aβ pathways were no longer significant after removing these genes from the analysis (Supplementary Table 5).

Discussion

Previous genome-wide association studies of plasma Aβ40 and Aβ42 levels have failed to uncover genome-wide significant findings. In this study, we identified two genome-wide significant loci associated with plasma Aβ levels in up to 12,369 non-demented subjects of European ancestry. The top variant in the first locus, rs429358, a well-known non-synonymous variant that encodes for the APOE4 isoform, was associated with lower circulating Aβ42 levels and Aβ42/40 ratio. In the second, located near *BACE1*, rs650585 was associated with lower plasma Aβ40 levels.

The *BACE1* region encompasses several genes (*PCSK7*, *RNF214*, *BACE1*, *CEP164*) and a *BACE1* anti-sense long non-coding RNA (*BACE1-AS*). Since the β-secretase activity of BACE1 is necessary for Aβ peptide production, it is likely that *BACE1* or a local regulation of *BACE1* expression are responsible for this signal. We also found suggestive associations with plasma Aβ40 levels near *APP* and *PSEN2*, two major actors of the Aβ metabolism. APP is obviously a central element of its own metabolism and PSEN2 is a key component of the γ-secretase which processes the APP C99 fragment into Aβ peptides. [1] We speculate that the effect of the variants on the expression/biological activation of these key elements of β-amyloid processing is strong enough to allow their detection at the plasma level, despite the influence of many other simultaneous biological processes, e.g. secretion, interaction with other proteins, degradation and/or clearance. Moreover, the top variants at the *PSEN2* and *BACE1* locus were also nominally associated with Aβ42 levels in the same direction as Aβ40 levels, which is in agreement with knowledge that *PSEN2* and *BACE1* activities indifferently produce Aβ40 and Aβ42 peptides.

Conversely, the *APOE*ε4 allele had the strongest association with Aβ42 levels but was not even nominally associated with Aβ40. In line with our previous comment, this suggests that the APOE4 isoform is not involved in the early process of Aβ peptide production but in more downstream events, such as Aβ aggregation or clearance. These results might also illustrate the greater ability to aggregate of Aβ42 peptides compared to Aβ40, and the influence of APOE isoforms in the regulation of this process. [8] Interestingly, associations of *APOE*ε2 with plasma Aβ levels were not significant and effect sizes were very small. Contrary to *APOE*ε4, the effect of *APOE*ε2 on amyloid markers has been much less studied and seems to be focused on specific brain regions, which could explain why we could not detect any association. [48] This could also suggest that other, Aβ-independent, mechanisms are involved in the lower risk of AD observed in *APOE*ε2 carriers. [49]

Given the presence of suggestive associations around known APP/Aβ genes, it is likely that novel and relevant signals exist within this range of statistical significance, awaiting definite validation. In order to minimize the risk of false positive result, we used exonic location and CADD score as additional filters to screen for SNPs of interest. Among those, we observed an association between a missense mutation of the VTN gene and lower plasma Aβ40 levels. *VTN* encodes for vitronectin, a glycoprotein present in abundance in the plasma and the extracellular matrix and involved in early regulation of thrombogenesis and tissue repair. [50] Vitronectin has been associated with amyloid deposits, [51] including Aβ, both at the level of amyloid plaques in the brain [52] and near

the retinal pigment epithelium in the aging eye. [53] Modest correlation between plasma vitronectin and brain amyloid burden measured by PiB PET has been reported. [54] Vitronectin is involved in microglial activation, [55] which is relevant to AD pathophysiology. [56] Vitronectin might also be involved in small vessel disease. In a mouse model of CADASIL, an autosomal dominant disease responsible for stroke and cognitive impairment, reduction of vitronectin expression resulted in less white matter lesions. [57] Vitronectin could therefore represent a promising candidate to study mechanisms linking Aβ peptides with Aβ-related pathologies, such as AD, cerebral amyloid angiopathy and small vessel disease. Other potential genes of interest close to suggestive signals warrant further investigation. For example, *SCTR*, encoding for secretin receptor, is involved in a wide range of physiological functions, beyond the scope of this study. Interestingly though, a study reported that mice deficient for that gene display impaired hippocampal synaptic plasticity. [58]

Although research on variants associated with levels of circulating amyloid peptides is of general interest for Aβ physiology, we are interpreting the findings primarily from the perspective of brain disorders, especially AD. When cross-checking our results with other published GWAS of CSF and brain Aβ, and AD, we observed consistent results only with the *APOE*ε4 variant. As expected, this variant was associated with low plasma and CSF Aβ42, high brain Aβ and AD risk, consistent with a differential effect of APOE isoforms on Aβ aggregation and clearance from the brain. When interpreting the absence of significant association between variants in the other loci and those same traits, one should keep in mind that their effect on plasma Aβ levels was generally smaller compared to that of *APOE*ε4 so we might be underpowered to detect a significant association. It is also of interest that the LD-score regression analysis suggested a positive correlation of variants modulating plasma Aβ1-42 levels and white matter lesions and a negative correlation between plasma Aβ40 and hippocampal volume. Nevertheless, these genetic correlations were only nominally significant and await replication.

Plasma Aβ is usually considered as a poor biomarker of AD in the literature. A previous meta-analysis reported that plasma Aβ levels were not useful to make a clinical diagnosis of AD. [59] Nevertheless many of the cohorts participating in the present study have previously reported that low plasma Aβ42 and Aβ42/40 ratio levels were associated with development of AD after several years of follow-up. [25-28] These results are consistent with an early, preclinical, involvement of Aβ in AD pathophysiology and are strengthened by our present observation that APOEε4, is both associated with low plasma Aβ42 and Aβ42/40 ratio and high AD risk. Some of those studies have also reported that this association remained significant after adjusting for *APOE*ε4, [28] and we might hypothesize that variations of plasma Aβ levels are not only a side-effect of *APOE*ε4, but are also involved in AD pathophysiology. As such, plasma Aβ levels would not be only useful as a biomarker of an active amyloid process in the brain but could also be considered as a therapeutic target. In favor of this hypothesis are reports that hemodialysis or peritoneal dialysis are able to lower Aβ in the brain. [60,61] The association we observed between variants near BACE1 and plasma Aβ40 is also of interest in the light of the ongoing trials testing BACE inhibitors, even though the lack of association of these variants with AD risk should be further investigated. [62]

Our study has several strengths. First, it is, to date, the largest study of circulating amyloid peptides. This enabled us to identify known actors of Aβ metabolism and, thus,

to be optimistic about the relevance of some of our suggestive signals. Second, this study was conducted in non-demented participants and therefore is relevant for the study of early amyloid pathophysiological processes. Third, we carefully normalized the plasma Aβ data before running GWAS, thus taking into account some of the heterogeneity that has been described when using plasma Aβ levels.

Our study has also limitations worth mentioning. As stated before, the state of current knowledge makes it hard for us to extrapolate the role of these actors from the plasma compartments to the brain and further research in this area is needed. Second, the assays used in this study non-selectively measured Aβ concentrations and could not distinguish monomers from oligomers of Aβ, whether free or protein-bound. Therefore, our interpretation of the present results might differ from other studies in which assays used selectively measured monomers or oligomers of Aβ. [63] Future studies should carefully choose assays that allow measurements of each form of Aβ as this will facilitate interpretation with regard to the balance between Aβ production, aggregation and clearance. Finally we tried to prioritize signals of interest using strict criteria, thus omitting to mention other interesting signals that might be real. Therefore we hope that the unfiltered results from this study, will be helpful as a resource to the scientific community to further decipher the physiology of Aβ peptides and its links to pathophysiology of AD and other Aβ-related diseases.

In summary, our results indicate that genetic determinants of plasma Aβ40 and Aβ42 levels are close to genes known to be central actors in APP metabolism in AD. Further increasing the statistical power of plasma Aβ analyses may potentially lead us to the identification of currently unknown players in Aβ metabolism, novel hypotheses and hopefully, new preventive or therapeutic targets against Alzheimer's disease.

Figures Legends

Figure 1. Association of frequent genetic variants with plasma Aβ1-42 in the *APOE* locus

Figure 2. Association of frequent genetic variants with plasma Aβ1-42/Aβ1-40 ratio in the *APOE* locus

Figure 3. Association of frequent genetic variants with plasma Aβ1-40 in the *BACE1* locus

Figure 4. Association of frequent genetic variants with plasma Aβ1-40 (A) and Aβ1-42/Aβ1-40 ratio (B) in the *APP* locus

Figure 5. LD-score regression assessing genetic correlation between plasma Aβ levels, Alzheimer's and Parkinson's diseases, cognition, hippocampal and intracranial volumes, white matter lesions, brain lobes volumes and subcortical brain structures volumes
Notes: * denotes nominal significance ($p<0.05$), "/" denotes missing values

Tables Legends

Table 1. Associations of top variants from genome-wide significant loci with plasma Aβ levels and amyloid-related traits

Table 2. Associations of top variants from the APP locus with plasma Aβ levels and amyloid-related traits

Table 3. Associations of top SNPs from the PSEN1 and PSEN2 locus with plasma Aβ levels and amyloid-related traits

References

1 Haass, C, Kaether, C, Thinakaran, G & Sisodia S. Trafficking and Proteolytic Processing of APP. *Cold Spring Harb Perspect Med* **2**, a006270 (2012).

2 Chen, M, Inestrosa, N. C, Ross, G. S & Fernandez H. L. Platelets are the primary source of amyloid beta-peptide in human blood. *Biochem Biophys Res Commun* **213**, 96-103 (1995).

3 Hyman, B. T et al.. National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. *Alzheimers Dement* **8**, 1-13 (2012).

4 Viswanathan, A & Greenberg, S. M. Cerebral amyloid angiopathy in the elderly *Annals of Neurology* **70**, 871-880 (2011).

5 Selkoe, D. J & Hardy, J. The amyloid hypothesis of Alzheimer's disease at 25~years. *EMBO Mol Med* **8**, 595-608 (2016).

6 Cacace, R, Sleegers, K & Van Broeckhoven C. Molecular genetics of early-onset Alzheimer's disease revisited. *Alzheimers Dement* **12**, 733-748 (2016).

7 Genin, E et al.. APOE and Alzheimer disease: a major gene with semi-dominant inheritance. *Mol Psychiatry* **16**, 903-907 (2011).

8 Kanekiyo, T, Xu, H & Bu G. ApoE and Aβ in Alzheimer's Disease: Accidental Encounters or Partners? *Neuron* **81**, 740-754 (2014).

9 Hardy, J & De Strooper, B. Alzheimer's disease: where next for anti-amyloid therapies? *Brain: A Journal of Neurology* **140**, 853-855 (2017).

10 Karran, E & Hardy, J. A critique of the drug discovery and phase 3 clinical programs targeting the amyloid hypothesis for Alzheimer disease. *Ann Neurol* **76**, 185-205 (2014).

11 Bateman, R. J et al.. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. *N Engl J Med* **367**, 795-804 (2012).

12 Villemagne, V. L et al.. Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study *The Lancet. Neurology* **12**, 357-367 (2013).

13 Lambert, J. C et al.. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* **45**, 1452-1458 (2013).

14 Bali, J, Gheinani, A. H, Zurbriggen, S & Rajendran L. Role of genes linked to sporadic Alzheimer's disease risk in the production of β-amyloid peptides. *Proc Natl Acad Sci U S A* **109**, 15307-15311 (2012).

15 Chapuis, J et al.. Genome-wide, high-content siRNA screening identifies the Alzheimer's genetic risk factor FERMT2 as a major modulator of APP metabolism *Acta Neuropathologica* **133**, 955-966 (2017).

16 International Genomics of Alzheimer's Disease Consortium (IGAP), . Convergent genetic and expression data implicate immunity in Alzheimer's disease. *Alzheimers Dement* **11**, 658-671 (2015).

17 Cruchaga, C et al.. GWAS of cerebrospinal fluid tau levels identifies risk variants for Alzheimer's disease *Neuron* **78**, 256-268 (2013).

18 Ramanan, V. K et al.. APOE and BCHE as modulators of cerebral amyloid deposition: a florbetapir PET genome-wide association study *Molecular Psychiatry* **19**, 351-357 (2014).

19 Shulman, J. M et al.. Genetic susceptibility for Alzheimer disease neuritic plaque pathology. *JAMA Neurol* **70**, 1150-1157 (2013).

20 Beecham, G. W et al.. Genome-wide association meta-analysis of neuropathologic features of Alzheimer's disease and related dementias. *PLoS Genet* **10**, e1004606 (2014).

21 Deming, Y et al.. Genome-wide association study identifies four novel loci associated with Alzheimer's endophenotypes and disease modifiers *Acta Neuropathologica* **133**, 839-856 (2017).

22 Sagare, A. P, Bell, R. D & Zlokovic B. V. Neurovascular defects and faulty amyloid-β vascular clearance in Alzheimer's disease *Journal of Alzheimer's disease: JAD* **33 Suppl 1**, S87-100 (2013).

23 Toledo, J. B et al.. Factors affecting Aβ plasma levels and their utility as biomarkers in ADNI. *Acta Neuropathol* **122**, 401-413 (2011).

24 Roberts, K. F et al.. Amyloid-β efflux from the central nervous system into the plasma. *Annals of Neurology* **76**, 837-844 (2014).

25 van Oijen, M, Hofman, A, Soares, H. D, Koudstaal, P. J & Breteler M. M. B. Plasma Abeta(1-40) and Abeta(1-42) and the risk of dementia: a prospective case-cohort study. *Lancet Neurol* **5**, 655-660 (2006).

26 Lambert, J.-C et al.. Association of plasma amyloid beta with risk of dementia: the prospective Three-City Study. *Neurology* **73**, 847-853 (2009).

27 Shah, N. S et al.. Midlife blood pressure, plasma β-amyloid, and the risk for Alzheimer disease: the Honolulu Asia Aging Study. *Hypertension* **59**, 780-786 (2012).

28 Chouraki, V et al.. Plasma amyloid-β and risk of Alzheimer's disease in the Framingham Heart Study. *Alzheimers Dement* **11**, 249-257.e1 (2015).

29 Chouraki, V et al.. A genome-wide association meta-analysis of plasma A$ peptides concentrations in the elderly. *Mol Psychiatry* **19**, 1326-1335 (2014).

30 Figurski, M. J et al.. Improved protocol for measurement of plasma β-amyloid in longitudinal evaluation of Alzheimer's Disease Neuroimaging Initiative study patients. *Alzheimers Dement* **8**, 250-260 (2012).

31 Ibrahim-Verbaas, C. A et al.. Linkage analysis for plasma amyloid beta levels in persons with hypertension implicates Aβ-40 levels to presenilin 2. *Hum Genet* **131**, 1869-1876 (2012).

32 Reitz, C et al.. Association between variants in IDE-KIF11-HHEX and plasma amyloid beta levels. *Neurobiol Aging* **33**, 199.e13-199.e17 (2012).

33 Winkler, T. W et al.. Quality control and conduct of genome-wide association meta-analyses. *Nat Protoc* **9**, 1192-1212 (2014).

34 Willer, C. J, Li, Y & Abecasis G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-2191 (2010).

35 Kircher, M, Witten, D. M, Jain, P, O'Roak, B. J, Cooper, G. M & Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants *Nature Genetics* **46**, 310-315 (2014).

36 McLaren, W et al.. The Ensembl Variant Effect Predictor *Genome Biology* **17**, 122 (2016).

37 GTEx Consortium, . Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans *Science (New York, N.Y.)* **348**, 648-660 (2015).

38 Bulik-Sullivan, B et al.. An atlas of genetic correlations across human diseases and traits *Nature Genetics* **47**, 1236-1241 (2015).

39 Bulik-Sullivan, B. K et al.. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* **47**, 291-295 (2015).

40 Nalls, M. A et al.. Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nat Genet* **46**, 989-993 (2014).

41 Davies, G et al.. Genetic contributions to variation in general cognitive function: a meta-analysis of genome-wide association studies in the CHARGE consortium (N=53949). *Mol Psychiatry* **20**, 183-192 (2015).

42 Hibar, D. P et al.. Novel genetic loci associated with hippocampal volume *Nature Communications* **8**, 13624 (2017).

43 Adams, H. H. H et al.. Novel genetic loci underlying human intracranial volume identified through genome-wide association *Nature Neuroscience* **19**, 1569-1582 (2016).

44 Fornage, M et al.. Genome-wide association studies of cerebral white matter lesion burden: the CHARGE consortium. *Ann Neurol* **69**, 928-939 (2011).

45 Hibar, D. P et al.. Common genetic variants influence human subcortical brain structures. *Nature* **520**, 224-229 (2015).

46 Herwig, R, Hardt, C, Lienhard, M & Kamburov A. Analyzing and interpreting genome data at the network level with ConsensusPathDB *Nature Protocols* **11**, 1889-1907 (2016).

47 Campion, D, Pottier, C, Nicolas, G, Le Guennec, K & Rovelet-Lecrux A. Alzheimer disease: modeling an Aβ-centered biological network. *Mol Psychiatry* **21**, 861-871 (2016).

48 Grothe, M. J, Villeneuve, S, Dyrba, M, Bartrés-Faz, D, Wirth, M & Alzheimer's Disease Neuroimaging Initiative . Multimodal characterization of older APOE2 carriers reveals selective reduction of amyloid load *Neurology* **88**, 569-576 (2017).

49 Zhao, N, Liu, C.-C, Qiao, W & Bu G. Apolipoprotein E, Receptors, and Modulation of Alzheimer's Disease *Biological Psychiatry*, (2017).

50 Leavesley, D. I et al.. Vitronectin--master controller or micromanager? *IUBMB life* **65**, 807-818 (2013).

51 Winter, M, Tholey, A, Krüger, S, Schmidt, H & Röcken C. MALDI-mass spectrometry imaging identifies vitronectin as a common constituent of amyloid deposits *The Journal of Histochemistry and Cytochemistry: Official Journal of the Histochemistry Society* **63**, 772-779 (2015).

52 Akiyama, H, Kawamata, T, Dedhar, S & McGeer P. L. Immunohistochemical localization of vitronectin, its receptor and beta-3 integrin in Alzheimer brain tissue *Journal of Neuroimmunology* **32**, 19-28 (1991).

53 Thompson, R. B et al.. Identification of hydroxyapatite spherules provides new insight into subretinal pigment epithelial deposit formation in the aging eye *Proceedings of the National Academy of Sciences of the United States of America* **112**, 1565-1570 (2015).

54 Kiddle, S. J et al.. Plasma based markers of [11C] PiB-PET brain amyloid burden. *PLoS ONE* **7**, e44260 (2012).

55 Milner, R & Campbell, I. L. The extracellular matrix and cytokines regulate microglial integrin expression and activation *Journal of Immunology (Baltimore, Md.: 1950)* **170**, 3850-3858 (2003).

56 Heppner, F. L, Ransohoff, R. M & Becher B. Immune attack: the role of inflammation in Alzheimer disease. *Nat Rev Neurosci* **16**, 358-372 (2015).

57 Capone, C et al.. Reducing Timp3 or vitronectin ameliorates disease manifestations in CADASIL mice *Annals of Neurology* **79**, 387-403 (2016).

58 Nishijima, I et al.. Secretin receptor-deficient mice exhibit impaired synaptic plasticity and social behavior *Human Molecular Genetics* **15**, 3241-3250 (2006).

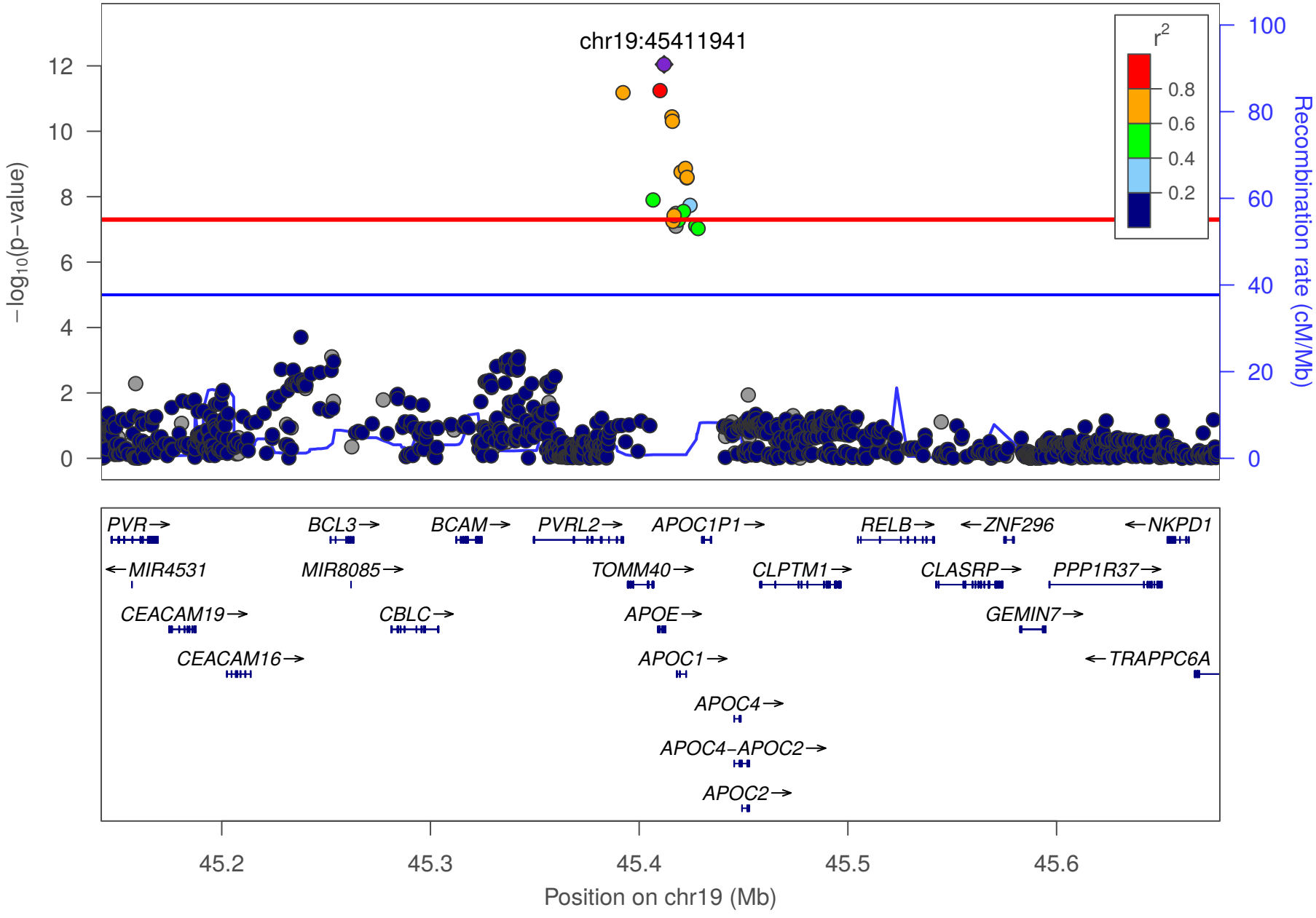
59 Olsson, B et al.. CSF and blood biomarkers for the diagnosis of Alzheimer's disease: a systematic review and meta-analysis *The Lancet. Neurology* **15**, 673-684 (2016).

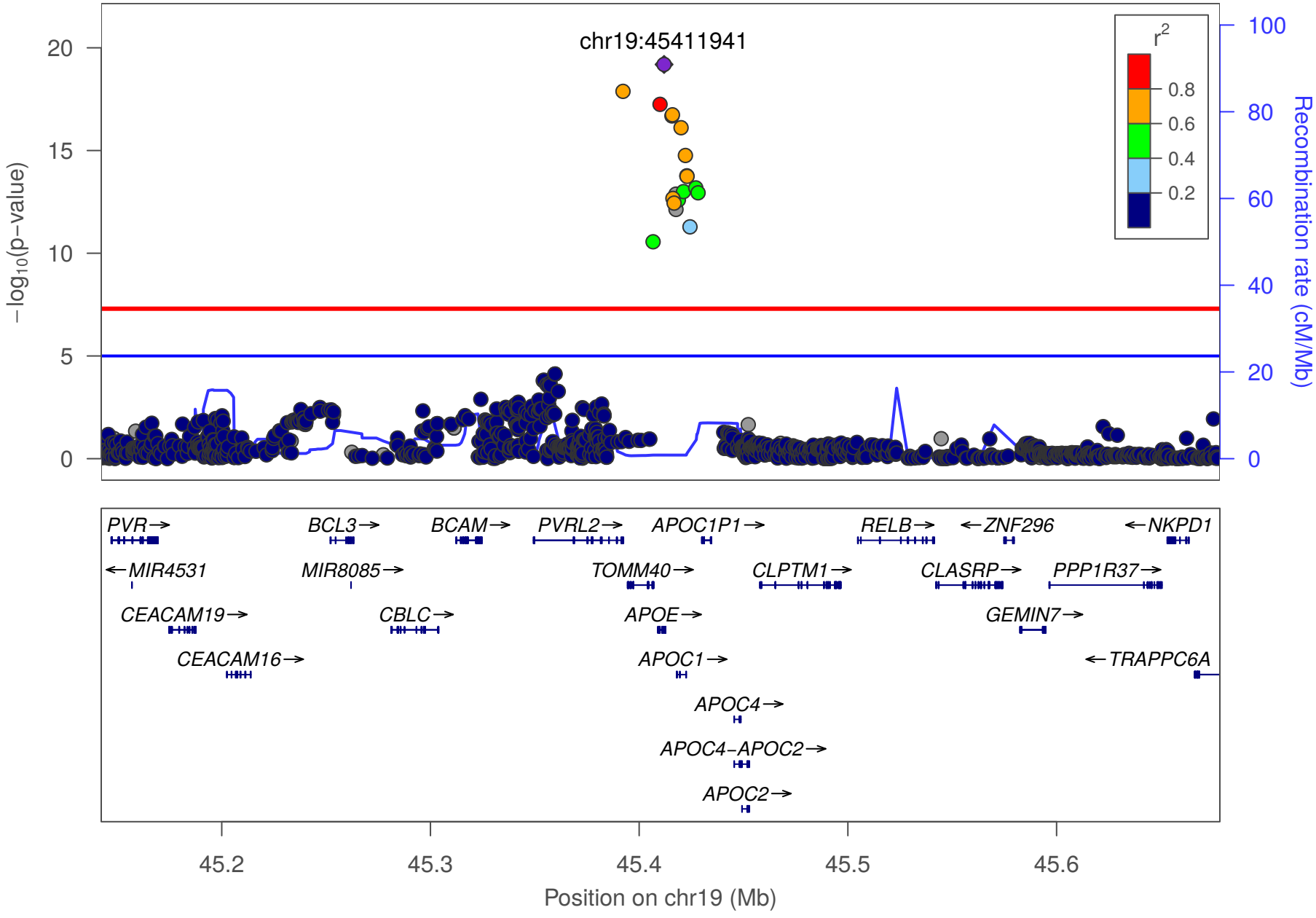
60 Sakai, K et al.. Patients that have Undergone Hemodialysis Exhibit Lower Amyloid Deposition in the Brain: Evidence Supporting a Therapeutic Strategy for Alzheimer's Disease by Removal of Blood Amyloid *Journal of Alzheimer's disease: JAD* **51**, 997-1002 (2016).

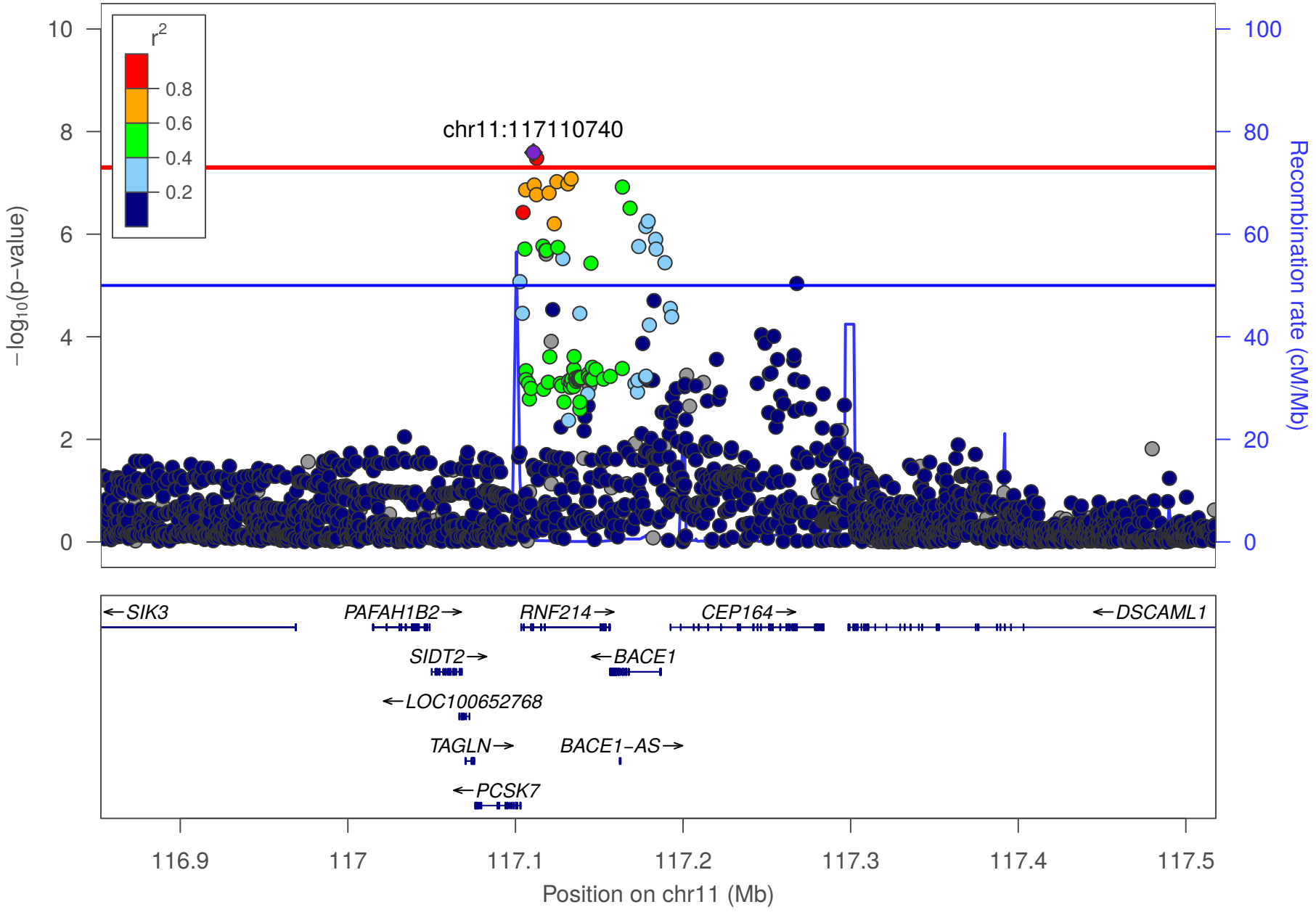
61 Jin, W.-S et al.. Peritoneal dialysis reduces amyloid-beta plasma levels in humans and attenuates Alzheimer-associated phenotypes in an APP/PS1 mouse model *Acta Neuropathologica* **134**, 207-220 (2017).

62 Mullard, A. BACE inhibitor bust in Alzheimer trial *Nature Reviews. Drug Discovery* **16**, 155 (2017).

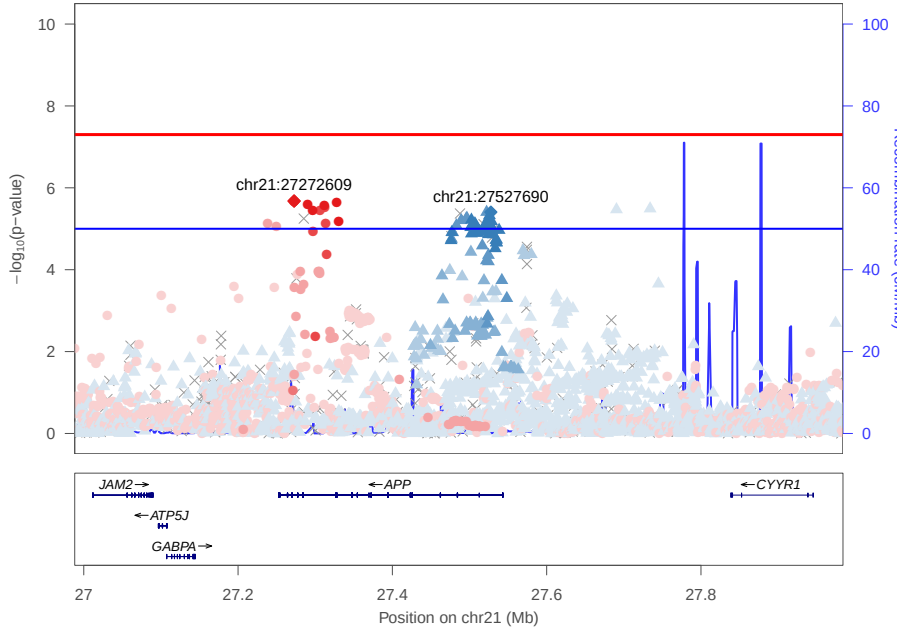
63 Fullwood, N. J, Hayashi, Y & Allsop D. Plasma amyloid-beta concentrations in Alzheimer's disease: an alternative hypothesis. *Lancet Neurol* **5**, 1000-1; author reply 1002-3 (2006).


chr19:45411941
r2
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
PVR→
BCL3→
BCAM→
PVRL2→
APOC1P1→
RELB→
←ZNF296
←NKPD1
←MIR4531
MIR8085→
TOMM40→
CLPTM1→
CLASRP→
PPP1R37→
CEACAM19→
CBLC→
APOE→
GEMIN7→
CEACAM16→
APOC1→
←TRAPPC6A
APOC4→
APOC4−APOC2→
APOC2→
45.2
45.3
45.4
45.5
45.6
Position on chr19 (Mb)

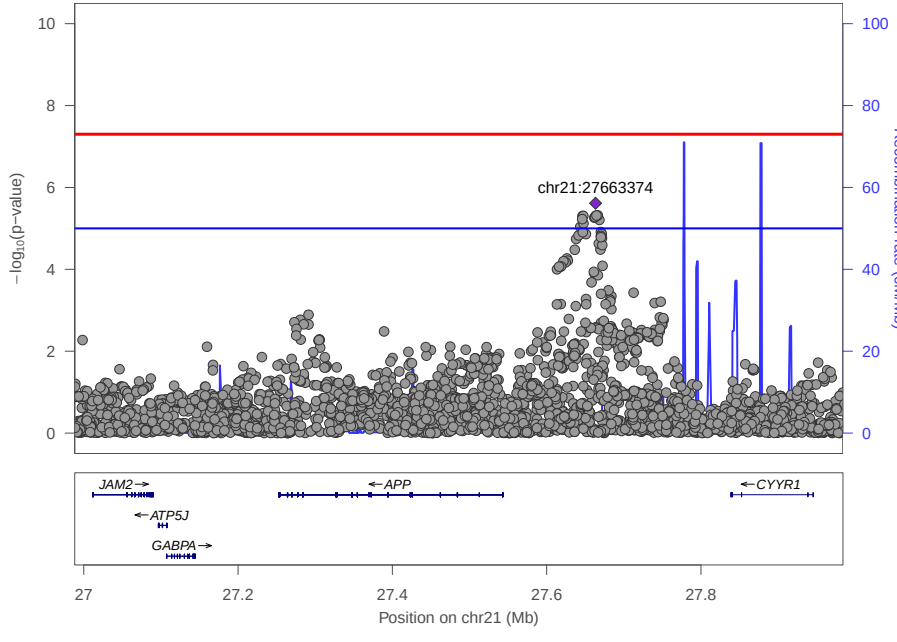

chr19:45411941
r2
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
PVR→
BCL3→
BCAM→
PVRL2→
APOC1P1→
RELB→
←ZNF296
←NKPD1
←MIR4531
MIR8085→
TOMM40→
CLPTM1→
CLASRP→
PPP1R37→
CEACAM19→
CBLC→
APOE→
GEMIN7→
CEACAM16→
APOC1→
←TRAPPC6A
APOC4→
APOC4−APOC2→
APOC2→
45.2
45.3
45.4
45.5
45.6
Position on chr19 (Mb)

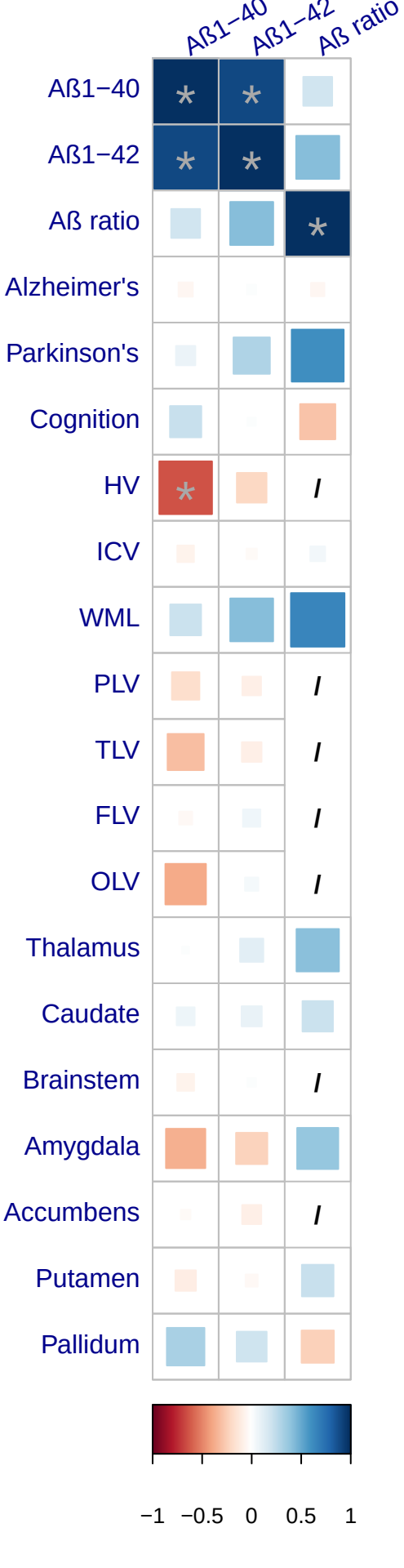

r2
0.8
0.6
0.4
0.2
chr11:117110740
−log10(p−value)
Recombination rate (cM/Mb)
10
8
6
4
2
0
100
80
60
40
20
0
←SIK3
PAFAH1B2→
RNF214→
CEP164→
←DSCAML1
SIDT2→
←BACE1
←LOC100652768
TAGLN→
BACE1−AS→
←PCSK7
116.9
117
117.1
117.2
117.3
117.4
117.5
Position on chr11 (Mb)

(A) Plasma Aβ1-40



(B) Plasma Aβ1-42/Aβ1-40 Ratio




Aß1−40
Aß1−42
Aß ratio
Aß1−40
Aß1−42
Aß ratio
Alzheimer's
Parkinson's
Cognition
HV
ICV
WML
PLV
TLV
FLV
OLV
Thalamus
Caudate
Brainstem
Amygdala
Accumbens
Putamen
Pallidum
−1
−0.5
0
0.5
1