APOE stratified genome-wide association studies provide novel insights into the genetic etiology of Alzheimers's disease.

Jesper Qvist Thomassen*§[1], Hampton Leonard*[2], Brittany Ulms*[3], Benjamin Grenier-Boley*[4], Sami Heikkinen*[5], Pablo Garcia-González*[6-8], Atahualapa Castillo-Morales**[9], Masataka Kikuchi**[10], Jungsoo Gim**[11-14], Han Cao**[15,16], Fahri Küçükali**[17,18], Najaf Amin**[3], Dabin Yoon**[11,12], Itziar de Rojas[6,7,19], Pilar Alvarez Jerez[2], Victoria Alvarez[20,21], Beatrice Arosio[22,23], Céline Bellenguez[4], Sverre Bergh[24,25], Kimberley Billingsley[2], Cornelis Blauwendraat[26], Merce Boada[6,7], Barbara Borroni[27,28], Paola Bossù[29], María J. Bullido[7,30,31], Antonio Daniele[32,33], Ángel Carracedo[34,35], Alexandre de Mendonça[36], Mark Cookson[2], Jürgen Deckert[37], Martin Dichgans[38-40], Srdjan Djurovic[41,42], Oriol Dols-Icardo[7,43], Carole Dufouil[44], Emrah Düzel[45,46], Valentina Escott-Price[47], Tormod Fladby[48,49], Laura Fratiglioni[50,51], Amy K.Y. Fu[15,16,52,53], Daniela Galimberti[54,55], Jose Maria García-Alberca[2,7], Vilmantas Giedraitis[56], Guillermo Garcia-Ribas[57], Caroline Graff[50,51], Timo Grimmer[58], Edna Grünblatt[59-61], OIivier Hanon[62], Lucrezia Hausner[63], Stefanie Heilmann-Hemibach[64], Jakub Hort[65,66], Frank Jessen[67-69], Kendall Jensen[70], Caroline Jonson[2], Yoontae Kim[12,13], Nicole Kuznetsov[2], Ville Leinonen[71,72], Anssi Lipponen[5], Jiao Luo[1,73], Mary Makarious[2], Henna Martiskainen[5], Carlo Masullo[74], Patrizia Mecocci[75,76], Shima Mehrabian[77], Pablo Mir[7,78,79], Akinori Miyashita[10], Susanne Moebus[80], Kin Y. Mok[15,16,81], Laura Molina Porcel[82,83], Fermin Moreno[7,84,85], Benedetta Nacmias[86,87], Lucilla Parnetti[88], Pau Pastor[89,90], Jordi Pérez-Tur[7,91], Oliver Peters[92,93], Yolande A.L. Pijnenburg[94,95], Gerard Piñol-Ripoll[96,97], Julius Popp[98-100], Innocenzo Rainero[101], Luis M Real[102,103], Steffi Riedel-Heller[104], Eloy Rodriguez-Rodriguez[7,105], Arvid Rongve[106], Giacomina Rossi[107], Jose Luis Royo[103], Dan Rujescu[108,109], Ingvild Saltvedt[110], María Eugenia Sáez[111], Raquel Sánchez-Valle[112], Florentino Sanchez-Garcia[113], Nicolai Sandau[1], Nikolaos Scarmeas[114,115], Katja Scheffler[110,116], Norbert Scherbaum[117], Anja Schneider[68,118], Geir Selbæk[119-121], Davide Seripa[122], Vincenzo Solfrizzi[123,124], Marco Spallazzi[125], Alessio Squassina[126], Eystein Stordal[127], Niccoló Tesi[95,128,129], Lucio Tremolizzo[130,131], Kumar P. Tripathi[132,133], Wiesje M. van der Flier[94,95,134], Julie Williams[9,135], Jens Wiltfang[136-138], Dag Aarsland[139-141], Andrew B Singleton[2], Philippe Amouyel[4], Stéphanie Debette[44], Gael Nicolas[142], Sven van der Lee[94,95,128], Henne Holstege[94,95,128,143,144], Maria Victoria Fernandez[6], Patrick Gavin Kehoe[145], Kristel Sleegers[17,18], Martin Ingelsson[56,146,147], Roberta Ghidoni[28], Ole A. Andreassen[42,148,149], Peter A. Holmans[135], Pascual Sánchez-Juan[7,150], Rebecca Sims[135], Nancy Y. Ip[15,16,52,53], Kun Ho Lee[11,13,14,151], Takeshi Ikeuchi[10], Alfredo Ramirez[68,69,118,132,152], Agustin Ruiz[6,7,153], Mikko Hiltunen#[5], Jean-Charles Lambert#[4], Cornelia van Duijn#[3,154], Mike Nalls#§[2], Ruth Frikke-Schmidt#§[1,155]

*Shared co-first authors, **shared co-second authors, #shared last authors, §shared corresponding authors.

1. Department of Clinical Biochemistry, Copenhagen University Hospital - Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark.
2. Center for Alzheimer's and Related Dementias, National Institute on Aging, Bethesda, MD, USA.
3. Nuffield Department of Population Health, University of Oxford, Oxford, UK.
4. Université de Lille, Inserm, CHU Lille, Institut Pasteur de Lille, Lille, France.
5. Institute of Biomedicine, University of Eastern Finland, Yliopistoranta 1E, 70211 Kuopio, Finland

6. Ace Alzheimer Center Barcelona, Universitat Internacional de Catalunya (UIC), Barcelona, Spain.
7. Networking Research Center on Neurodegenerative Diseases (CIBERNED), Instituto de Salud Carlos III, Madrid, Spain.
8. PhD Program in Biotechnology, Faculty of Pharmacy and Food Sciences, University of Barcelona, 08028 Barcelona, Spain.
9. UK Dementia Research Institute, School of Medicine, Cardiff University, Wales, UK.
10. Department of Molecular Genetics, Brain Research Institute, Niigata University, Japan.
11. Department of Biomedical Science, Chosun University, Gwangju, 61452, Republic of Korea.
12. Well-aging Medicare Institute & CSU G-LAMP Project Group, Chosun University, Gwangju, 61452, Republic of Korea.
13. BK FOUR Department of Integrative Biological Sciences, Chosun University, Gwangju, 61452, Republic of Korea.
14. Gwangju Alzheimer's & Related Dementia (GARD) Cohort Research Center, Chosun University, Gwangju, 61452, Republic of Korea.
15. Division of Life Science, State Key Laboratory of Molecular Neuroscience and Daniel and Mayce Yu Molecular Neuroscience Center, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China.
16. Hong Kong Center for Neurodegenerative Diseases, InnoHK, Hong Kong, China.
17. VIB Center for Molecular Neurology, VIB, Antwerp, Belgium.
18. Department of Biomedical Sciences, University of Antwerp, Antwerp, Belgium.
19. Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, Esch-sur-Alzette, Luxembourg.
20. Laboratorio de Genética. Hospital Universitario Central de Asturias, Oviedo, Spain.
21. Instituto de Investigación Sanitaria del Principado de Asturias (ISPA).
22. Department of Clinical Sciences and Community Health, University of Milan, Milan, Italy.
23. Geriatric Unit, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy.
24. The Research Centre for Age-related Functional Decline and Disease, Innlandet Hospital Trust, Ottestad, Norway.
25. Norwegian National Centre for Aging and Health, Vestfold Health Trust, Tønsberg, Norway.
26. Coalition for Aligning Science, 5425 Wisconsin Ave #600, Chevy Chase , MD 20815, USA.
27. Department of Clinical and Experimental Sciences, University of Brescia, Italy.
28. Molecular Markers Laboratory, IRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy.
29. Laboratory of Experimental Neuropsychobiology, Clinical Neuroscience and Neurorehabilitation Department, IRCCS Santa Lucia Foundation, Rome, Italy.
30. Centro de Biología Molecular Severo Ochoa (UAM-CSIC).
31. Instituto de Investigacion Sanitaria 'Hospital la Paz' (IdIPaz), Madrid, Spain.
32. Department of Neuroscience, Università Cattolica del Sacro Cuore, Rome, Italy.
33. Neurology Unit, IRCCS Fondazione Policlinico Universitario A. Gemelli, Rome, Italy.
34. Grupo de Medicina Xenómica, CIBERER, CIMUS. Universidade de Santiago de Compostela, Santiago de Compostela, Spain.
35. Fundación Pública Galega de Medicina Xenómica- IDIS, Santiago de Compostela, Spain.
36. Faculty of Medicine, University of Lisbon, Portugal.
37. Department of Psychiatry, Psychosomatics and Psychotherapy, Center of Mental Health, University Hospital of Würzburg, Germany.
38. Institute for Stroke and Dementia Research (ISD), University Hospital, LMU Munich, Munich, Germany.
39. German Center for Neurodegenerative Diseases (DZNE), Munich, Germany.
40. Munich Cluster for Systems Neurology (SyNergy), Munich, Germany.
41. Department of Medical Genetics, Oslo University Hospital, Oslo, Norway
42. Centre for Precision Psychiatry, Institute of Clinical Medicine, University of Oslo, Oslo, Norway.
43. Department of Neurology, Institut de Recerca Sant Pau, Hospital de la Santa Creu i Sant Pau, Universitat Autònoma de Barcelona, Barcelona, Spain.

44. University of Bordeaux, Inserm Bordeaux Population Health Research Center U1219, Bordeaux, France.
45. German Center for Neurodegenerative Diseases (DZNE), Magdeburg, Germany.
46. Institute of Cognitive Neurology and Dementia Research (IKND), Otto-von-Guericke University, Magdeburg, Germany.
47. UKDRI@ Cardiff, School of Medicine, Cardiff University, Cardiff, UK.
48. Institute of Clinical Medicine, Campus Ahus, University of Oslo, Oslo, Norway.
49. Department of Neurology, Akershus University Hospital, Lørenskog, Norway.
50. Karolinska University Hospital, Theme inflammation and Aging, Sweden.
51. Karolinska Institutet, Center for Alzheimer Research, Department NVS, Division of ARC, Stockholm, Sweden.
52. Guangdong Provincial Key Laboratory of Brain Science, Disease and Drug Development, HKUST Shenzhen Research Institute, Shenzhen, Guangdong, 518057, China.
53. SIAT-HKUST Joint Laboratory for Brain Science, Shenzhen, Guangdong, 518055, China.
54. Neurodegenerative Diseases Unit, Fondazione IRCCS Ca' Granda, Ospedale Policlinico, Milan, Italy.
55. Dept. of Biomedical, Surgical and Dental Sciences, University of Milan, Milan, Italy.
56. Department of Public Health and Caring Sciences, Clinical Geriatrics, Uppsala University, Uppsala, Sweden.
57. Hospital Universitario Ramon y Cajal, IRYCIS, Madrid.
58. Center for Cognitive Disorders, Department of Psychiatry and Psychotherapy, Technical University of Munich, School of Medicine, Munich, Germany.
59. Department of Child and Adolescent Psychiatry and Psychotherapy, University Hospital of Psychiatry Zurich, University of Zurich, Zurich, Switzerland.
60. Neuroscience Center Zurich, University of Zurich and ETH Zurich, Switzerland.
61. Zurich Center for Integrative Human Physiology, University of Zurich, Switzerland.
62. Université Paris Cité, EA 4468, APHP, Hospital Broca, Memory Resource and Research Centre of de Paris-Broca-Ile de France, Paris, France.
63. Department of Geriatric Psychiatry, Central Institute for Mental Health Mannheim, Faculty Mannheim, University of Heidelberg, Germany.
64. Institute of Human Genetics, University of Bonn, School of Medicine & University Hospital Bonn, Bonn, Germany.
65. Memory Clinic, Department of Neurology, Charles University, Second Faculty of Medicine and Motol University Hospital, Czech Republic.
66. International Clinical Research Center, St. Anne's University Hospital Brno, Brno, Czech Republic.
67. Department of Psychiatry and Psychotherapy, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany.
68. German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany.
69. Cologne Excellence Cluster on Cellular Stress Responses in Aging-Associated Diseases (CECAD), University of Cologne, Cologne, Germany.
70. Center for Noninvasive Diagnostics, TGen, Phoenix, AZ, USA.
71. Department of Neurosurgery, Kuopio University Hospital, Puijonlaaksontie 2 FI-70210 Kuopio, Finland.
72. Institute of Clinical Medicine-Neurosurgery, University of Eastern Finland, Kuopio, Finland.
73. Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Science, University of Copenhagen, Blegdamsvej 3, 2200 Copenhagen.
74. Institute of Neurology, Catholic University of the Sacred Heart, Rome, Italy.
75. Division of Gerontology and Geriatrics, Department of Medicine and Surgery, University of Perugia, Italy.
76. Division of Clinical Geriatrics, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden.
77. Clinic of Neurology, UH "Alexandrovska", Medical University - Sofia, Sofia, Bulgaria.

78. Unidad de Trastornos del Movimiento, Servicio de Neurología y Neurofisiología. Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Seville, Spain.
79. Departamento de Medicina, Facultad de Medicina, Universidad de Sevilla, Seville, Spain.
80. Institute for Urban Public Health, University Hospital of University Duisburg-Essen, Essen, Germany.
81. Department of Molecular Neuroscience, UCL Institute of Neurology, London WC1N 3BG, U.K.
82. Neurological Tissue Bank - Biobanc- Hospital Clinic -IDIBAPS, Barcelona, Spain.
83. Alzheimer's disease and other cognitive disorders Unit. Neurology Department, Hospital Clinic , Barcelona, Spain.
84. Department of Neurology. Hospital Universitario Donostia. San Sebastian, Spain.
85. Neurosciences Area. Instituto Biodonostia. San Sebastian, Spain.
86. Department of Neuroscience, Psychology, Drug Research and Child Health University of Florence, Florence, Italy.
87. IRCCS Fondazione Don Carlo Gnocchi, Florence, Italy.
88. Centre for Memory Disturbances, Lab of Clinical Neurochemistry, Section of Neurology, University of Perugia, Perugia, Italy.
89. Fundació Docència i Recerca MútuaTerrassa, Terrassa, Barcelona, Spain.
90. Memory Disorders Unit, Department of Neurology, Hospital Universitari Mutua de Terrassa, Terrassa, Barcelona, Spain.
91. Unitat de Genètica Molecular, Institut de Biomedicina de València-CSIC, Valencia, Spain.
92. German Center for Neurodegenerative Diseases (DZNE), Berlin, Germany.
93. Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Institute of Psychiatry and Psychotherapy, Hindenburgdamm 30, 12203 Berlin, Germany.
94. Alzheimer Center Amsterdam, Neurology, Vrije Universiteit Amsterdam, Amsterdam UMC location VUmc, Amsterdam, The Netherlands.
95. Amsterdam Neuroscience, Neurodegeneration, Amsterdam, The Netherlands.
96. Unitat Trastorns Cognitius, Hospital Universitari Santa Maria de Lleida, Lleida, Spain.
97. Institut de Recerca Biomedica de Lleida (IRBLLeida), Lleida, Spain.
98. Old Age Psychiatry, Department of Psychiatry, Lausanne University Hospital, Lausanne, Switzerland.
99. Department of Geriatric Psychiatry, University Hospital of Psychiatry Zürich, Zürich, Switzerland.
100. Institute for Regenerative Medicine, University of Zürich, Switzerland.
101. Department of Neuroscience "Rita Levi Montalcini", University of Torino, Torino, Italy.
102. Unidad Clínica de Enfermedades Infecciosas y Microbiología. Hospital Universitario de Valme, Sevilla, Spain.
103. Depatamento de Especialidades Quirúrgicas, Bioquímica e Inmunología. Facultad de Medicina. Universidad de Málaga. Málaga, Spain.
104. Institute of Social Medicine, Occupational Health and Public Health, University of Leipzig, 04103 Leipzig, Germany.
105. Neurology Service, Marqués de Valdecilla University Hospital (University of Cantabria and IDIVAL), Santander, Spain.
106. Haugesund Hospital, Department of Research and Innovation, Haugesund, Norway.
107. Unit of Neurology V - Neuropathology, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy.
108. Department of Psychiatry and Psychotherapy, Medical University Vienna, Vienna, Austria.
109. Martin-Luther-University Halle-Wittenberg, University Clinic and Outpatient Clinic for Psychiatry, Psychotherapy and Psychosomatics, Halle (Saale), Germany.
110. Department of Neuromedicine and Movement Science, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), Department of Geriatrics, St. Olav's Hospital, Trondheim University Hospital, Trondheim, Norway.
111. CAEBI, Centro Andaluz de Estudios Bioinformáticos, Sevilla, Spain.

112. Alzheimer's disease and other cognitive disorders unit. Service of Neurology. Hospital Clínic of Barcelona. Institut d'Investigacions Biomèdiques August Pi i Sunyer, University of Barcelona, Barcelona, Spain.
113. Department of Immunology, Hospital Universitario Doctor Negrín, Las Palmas de Gran Canaria, Las Palmas, Spain.
114. Taub Institute for Research in Alzheimer's Disease and the Aging Brain, The Gertrude H. Sergievsky Center, Depatment of Neurology, Columbia University, New York, NY.
115. 1st Department of Neurology, Aiginition Hospital, National and Kapodistrian University of Athens, Medical School, Greece.
116. Department of Neurology and Clinical Neurophysiology, Trondheim University Hospital , Trondheim, Norway.
117. Department of Psychiatry and Psychotherapy, LVR-Klinikum Essen, University of Duisburg-Essen, Germany.
118. Department for Cognitive Disorders and Old Age Psychiatry, University Hospital Bonn, Bonn, Germany.
119. Norwegian National Advisory Unit on Aging and Health, Vestfold Health Trust, Tønsberg, Norway.
120. Institute of Health and Society, Faculty of Medicine, University of Oslo, Oslo, Norway.
121. Centre for Old Age Psychiatric Research, Innlandet Hospital Trust, Ottestad, Norway.
122. Department of Hematology and Stem Cell Transplant, Vito Fazzi Hospital, Lecce, Italy.
123. Interdisciplinary Department of Medicine, Geriatric Medicine and Memory Unit, University of Bari "A. Moro, Bari, Italy.
124. Academic Division "C. Frugoni" & Hospital Division of Internal and Geriatric Medicine, Policlinico Hospital, Bari, Italy.
125. Department of Medicine and Surgery, Unit of Neurology, University-Hospital of Parma, Parma, Italy.
126. Department of Biomedical Sciences, Section of Neuroscience and Clinical Pharmacology, University of Cagliari, Cagliari, Italy.
127. Namsos Hospital, Namsos, Norway.
128. Genomics of Neurodegenerative Diseases and Aging, Human Genetics, Vrije Universiteit Amsterdam, Amsterdam UMC location VUmc, Amsterdam, The Netherlands.
129. Delft Bioinformatics Lab, Delft University of Technology, Delft, The Netherlands.
130. Neurology Unit, IRCCS "San Gerardo dei Tintori", Monza, Italy.
131. School of Medicine and Surgery, University of Milano-Bicocca, Monza, Italy.
132. Division of Neurogenetics and Molecular Psychiatry, Department of Psychiatry and Psychotherapy, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany.
133. Estudios en Neurociencias y Sistemas Complejos (ENyS) CONICET-HEC-UNAJ.
134. Epidemiology and Data Science, Vrije Universiteit Amsterdam, Amsterdam UMC location VUmc, Amsterdam, The Netherlands.
135. Division of Psychological Medicine and Clinical Neuroscience, School of Medicine, Cardiff University, Wales, UK.
136. Department of Psychiatry and Psychotherapy, University Medical Center Goettingen, Goettingen, Germany.
137. German Center for Neurodegenerative Diseases (DZNE), Goettingen, Germany.
138. Medical Science Department, iBiMED, Aveiro, Portugal.
139. Department of Old Age Psychiatry, Division of Academic Psychiatry, Institute of Psychiatry, Psychology and Neuroscience, King's College London, United Kingdom.
140. Wolfson Centre for Age Related Diseases, Division of Neuroscience of the Institute of Psychiatry, Psychology and Neuroscience, King's College London, United Kingdom.
141. Stavanger University Hospital, Stavanger, Norway.
142. Univ Rouen Normandie, Normandie Univ, Inserm U1245 and CHU Rouen, Department of Genetics and Reference Center for Developmental Disorders, Rouen, France.
143. VIB Center for Brain and Disease Research, ON5 Herestraat 49 - box 602, 3000 Leuven, Belgium.
144. Department of Neurosciences, Leuven Brain Institute, KU Leuven, Leuven, Belgium.

145. Bristol Medical School, University of Bristol, Bristol, UK.
146. Krembil Brain Institute, University Health Network, Toronto, Ontario, Canada.
147. Tanz Centre for Research in Neurodegenerative Diseases, Departments of Medicine and Laboratory Medicine & Pathobiology, University of Toronto, Toronto, Ontario, Canada.
148. Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway.
149. KG Jebsen Centre for Neurodevelopmental disorders, University of Oslo, Oslo, Norway.
150. Alzheimer's Centre Reina Sofia-CIEN Foundation, Madrid, Spain.
151. Department of Neural Development and Disease, Korea Brain Research Institute, Daegu, 41062, Repuiblic of Korea.
152. Department of Psychiatry & Glenn Biggs Institute for Alzheimer's and Neurodegenerative Diseases, San Antonio, TX, USA.
153. Biggs Institute for Alzheimer's and Neurodegenerative Diseases, University of Texas Health Science Center, San Antonio, Texas, USA.
154. Department of Epidemiology, ErasmusMC, Rotterdam, The Netherlands.
155. Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3B, 2100 Copenhagen, Denmark.

Among the more than 90 identified genetic risk loci for late-onset Alzheimer's disease (AD) and related dementias, the apolipoprotein E gene (*APOE*) ε2/ε3/ε4 polymorphisms remains the longstanding benchmark for genetic disease risk with a consistently large effect across studies[1-10]. Despite this massive signal, the exact mechanisms for how ε4 increases and for how ε2 decreases dementia risk is not well-understood. Importantly, recent trials of anti-amyloid therapies suggest less efficacy and higher risks of severe side effects in ε4 carriers[11-13], hampering the treatment of those with the highest unmet need. To improve our understanding of the genetic architecture of AD in the context of its main genetic driver, we performed genome-wide association studies (GWASs) stratified by ε4 and ε2 carrier status. Such insights may help to understand and overcome side effects, to impact clinical trial enrolment strategies, and to create the scientific basis for targeted mechanism-driven therapies in neurodegenerative diseases.

(Introductory paragraph wordcount: 149 (max 150 words))

The present work is the largest meta-analysis GWAS attempt to provide the most informative overview of the genetics of AD according to *APOE* ε2/ε3/ε4 stratification, bringing together European, Asian, Asian-American, African-American, and admixed American ancestry cohorts based on clinically diagnosed AD. An overview of the included consortia and cohorts, as well as the analysis strategy is given in Figure 1. Individuals were grouped in ε22+ε32, ε33 and ε44+ε43 strata to maximize statistical power and individuals with the ε42 genotype were excluded (Supplementary Tables 1-2). In the ε22+ε32 stratum, the meta-analysis was based on 2,606 AD cases, 70,388 controls and 10,336,311 variants (Supplementary Table 3, Supplementary Figure 1). In the ε33 stratum, the meta-analysis was based on 24,033 AD cases, 363,161 controls and 17,127,662 variants (Supplementary Table 4, Supplementary Figure 1). Finally, in the ε44+ε43 stratum, the meta-analysis was based on 29,122 AD cases, 164,206 controls and 14,672,059 variants (Supplementary Table 5, Supplementary Figure 1; Supplementary Tables 6 and Supplementary Figure 1 and Supplementary Figure 2 for strata ε44).

For the ε22+ε32 stratum no signals reached a genome-wide significance level of $<2.5\times10^{-8}$ (Supplementary Figure 3). In total, 25 loci reached a genome-wide significance level in strata ε33 or ε44+ε33 only or in both (Figure 2, Table 1, Supplementary Figures 4-24). For the nine loci found in both strata, they are well known genetic risk loci associated with AD: *CR1*, *BIN1*, *HLA*, *TREM2*, *CLU*, *MS4A64*, *PICALM*, *APH1B*, *ABCA7*. Nine loci were exclusively observed in the ε44+ε43 stratum, 6 are known AD loci: *PILRA*, *SORL1*, *ADAM10*, *ACE*, *LILRA5*, *CASS4* and 3 loci are novel AD loci: *HP1BP3*, *PTPRC, DDHD1.* Notably, *DDHD1* is close to the *FERMT2* locus, which is recognized as a genetic risk factor for AD. However, using conditional testing, we found that the *DDHD1* signal is independent of *FERMT2* (Figure 2, Supplementary Table 7). Finally, among the 7 loci only reaching genome wide significance level in the ε33 stratum, 4 are known as AD risk loci: *TMEM106B, SHARPIN, GRN, MAPT,* and 3 loci are novel: *SCL50A1, NPAS3, CHST9.*

Of note, we also performed a meta-analysis restricted to the ε44 carriers. Only one well-established locus (*BIN1*) was observed in the ε44 stratum including 5,814 AD cases, 14,415 controls and 9,723,486 variants (Supplementary Table 6, Supplementary Figure 1 and Supplementary Figure 2). Forest plots across cohorts and strata are shown in Supplementary Figures 25-53. In addition, we applied clumping procedures and conditional testing to define potential independent signals within each locus detected in the ε33 and ε44+ε43 strata. This approach detected 4 loci presenting two independent signals (*HLA-DRA*, *TREM2*, *CLU*, *APH1B*, Supplementary Table 7), details are specified in legend to Table 1.

The present meta-analyses do not allow us to fully determine whether there is a significant difference between the signals observed in the ε33 and ε44+ε43 strata due to sample and statistical power variations. To formally test for SNP x *APOE* interactions, we employed two complementary approaches. First, we performed interaction tests based on stratified effect comparisons using additive and dominant models with summary statistics from the different *APOE* strata, testing only variants that were significant in one or both strata. Second, we conducted a genome-wide SNP x *APOE* ε4 carrier interaction GWAS using a formal interaction term in a unified model.

In the stratified effect comparison testing dominant *APOE* ε4 effects (meta-analysis of differences between ε33 and ε44+ε43 in each cohort), we found 7 significant stratified effect differences, 3 signals where the effect sizes were attenuated with the presence of an ε4 allele (*SLC50A1*, *TMEM106B*, *NPAS3*) (Table 2, Figure 3) and 3 signals where the effect sizes were augmented with the presence of an ε4 allele (*BIN1*, *HLA-DRA -1*, *CLU*, *DDHD1*) (Table 2, Figure 4). Forest plots of the effect differences across cohorts are shown in Supplementary Figures 54-55. In the additive mixed-effect model we additionally identified *SHARPIN* as interacting with *APOE* ε4, where the effect size was attenuated with an increasing number of ε4 alleles (Table 2, Figure 3). Stratified effect difference sensitivity analyses are shown in Supplementary Table 8.

In the genome-wide SNP x *APOE* ε4 carrier interaction GWAS, two new loci (*SMYD2* and *ANK3*) reached genome-wide significance in the SNP term (comparable to *APOE* ε33 strata) (Supplementary Figure 56, Supplementary Table 9. The *APOE* ε4 carrier term was highly significant ($p<1x10^{-300}$) and SNP independent. In the interaction term (SNP x *APOE* ε4 carrier) only the *APOE* locus was significant. However, the *DDHD1* locus reached an interaction p-value of $1.62x10^{-6}$. The loci with an interaction p-value below $1x10^{-5}$ are shown in Supplementary Table 10 and the results from the interaction GWAS for the 29 regions of interest are shown in Supplementary Table 11.

We evaluated the 29 regions of interest in additional cohorts of East Asian (EAS) ancestry, representing Japanese (JADNI, CL, NP; EAS-JPN), Chinese (HKS; EAS-CHN), and Korean (GARD; EAS-KOR) populations, as well as in Asian American (ADSP-AAC), African American (ADSP-AFR), and in admixed American (ADSP-AMR) multi-ancestry populations (Supplementary Table 12). Despite several limitations, i.e. difference in linkage structure between cohorts of different ancestries, limitation in statistical power and lead variants being different from the causal variants, similar signals could be observed for several variants (Supplementary Figures 57-83). The meta-analyzed results for *HLA-DRA-1* and *DDHD1* were similar in the East Asian cohorts compared with the European cohorts (Figure 5). A summary of the main interaction results are presented in Figure 6.

For lead variant rs10131116 in *DDHD1* a significant eQTL associated with decreased *DDHD1* expression was observed in the ROSMAP dorsolateral prefrontal cortex (β=-0.098, eQTL $p=1.67\times10^{-5}$, n=560)[6] and in the GTEx (v10) brain-putamen (basal-ganglia) (β=-0.23, eQTL $p=3.6\times10^{-5}$, n=253)[14]. We tested whether rs10131116 was associated with lower *DDHD1* expression according to *APOE* strata, and nominal p-values were 9×10^{-4} for ε33 and 7×10^{-3} for ε44+ε43 (Supplementary Table 13, Supplementary Figure 84). Further, a gene-based analysis

confirmed several of the significant loci from the main stratified GWAS analysis (Figure 6). Also, in the ε44+ε43 stratum the known AD genes *EPHA1*, *TCPN1* and *BLNK* and in the ε33 stratum the *SLC24A4*, *INPP5D* and *SH2B2* (novel but close to the known loci *SPDYE3/PILRA/TMEM225B*) reached significance. *ACE* and *SNX1* were significant in both the ε44+ε43 and ε33 strata, whereas *PRAMEF1* was significant in the ε44 stratum and *GFRA1* in the ε22 stratum (Figure 6, Supplementary Tables 14-17). Next, we performed a pathway enrichment analysis on the *APOE* stratified GWAS results from the eight European studies (Supplementary Tables 18 and 19). After correction for multiple testing in each stratum ($q<0.05$), 12 and 33 pathways reached statistical significance for the *APOE* ε33 and ε44+ε43 strata, respectively. Overall, pathways related to the complement and immune systems were overrepresented in the *APOE* ε44+ε44 stratum compared to the ε33 stratum, whereas amyloid and neurofibrillary tangle biology was highlighted in both strata. No pathway analysis reached statistical significance for the ε22+ε32 nor for the ε43 stratum. Lastly, we performed a Summary-data-based Mendelian Randomization to test for potential effects of expression on AD that are shared by a causal variant for both *APOE* ε4 carriers (ε44+ε43) and non-carriers (ε33). Four genes passed our significance threshold and HEIDI filtering: *STAG3* (*PILRA* locus) in the ε44+ε43 stratum (Cortex) and *LRRC37A*, *ARL17B* and *LRRC37A2* (*MAPT* locus) in the ε33 stratum (multiple brain regions) (Supplementary Table 20).

Discussion

By conducting a comprehensive series of *APOE* stratified GWAS analyses, we identified a number of biologically plausible genomic signals that modify the effect of the strongest genetic AD risk variant to date - the *APOE* ε4 allele. As these findings suggest differential biological pathways dependent on the *APOE* strata, they may provide inspiration to how we use genetics in designing randomized clinical trials of future AD medicines.

New genetic signals

HP1BP3, *SLC50A1*, *PTPRC*, *NPAS3*, *DDHD1*, and *CHST9* from the stratified variant based GWAS analysis, *SMYD2* from the interaction GWAS and *PRAMEF1* and *GFRA1* from the gene-based analysis are novel genomic signals for AD risk, only appearing when stratified by *APOE* carrier status, and supported by significant stratified effect comparison tests for *SLC50A1*, *NPASS*, and *DDHD1*, discussed in detail in paragraphs below. *HP1BP3* encodes heterochromatin protein 1 binding protein 3 and is a regulator of cell cycle progression[15]. *PTPRC* encodes protein tyrosine phosphatase receptor type C also known as CD45 that increasingly is understood to play a role in the innate immune system[16,17]. C*HST9* encodes carbohydrate sulfotransferase 9, an enzyme that transfers sulphate to the 4-position of GalNAc. GalNAc4ST-1 and -2 transcripts are highly expressed in the pituitary gland and trachea[19,20]. *SMYD2* encodes a Protein-lysine N-methyltransferase that methylates both histones and non-histone-proteins[21]. *PRAMEF1* encodes a protein **PRAME family member 1** and has been associated with cancer[22]. *GFRA1* encodes for GDNF family receptor alpha-1 which is a receptor for both Glial cell line-derived neurotrophic factor (GDNF) and neurturin (NTN); both potent neurotrophic factors and key regulators of neuron survival and differentiation[23]. GDNF family receptor alpha-1 has been linked to the restoration of AD neuron survival[24].

Genetic variants interacting significantly with APOE carrier status with attenuated effect size in ε4 carriers

SLC50A1 is a novel AD signal that only emerges in the *APOE* ε33 stratum, and encodes solute carrier family 50 member 1, which is a sugar transporter for intercellular exchange and nutrition of pathogens[25]. The previously identified signal, *TMEM106B* [6] encodes the lysosomal type II transmembrane protein 106B, and residues of the protein have recently been shown to be amyloidogenic in an age dependent manner and in several neurodegenerative diseases including AD [26]. The presently identified lead hit is in the regulatory 3'UTR part of the gene and is in full LD (r^2 =0.99) with the previously reported rs1990622 variant – a variant that is associated with reduced expression of *TMEM106B*[27,28] and with earlier age-at-onset of frontotemporal lobar degeneration in *GRN* mutation carriers. Both the *TMEM106B* and *GRN* signals are sufficiently strong to reach genome-wide significance level in our previous overall GWAS[6], however the present *APOE* stratified analyses illustrate that these signals only manifest in the *APOE* ε33 context, although only *TMEM106B* reached statistical significance in the stratified effect comparison test. Interestingly, *TMEM106B* and *GRN* were recently associated only with non-AD pathology in a comprehensive GWAS of multiple neuropathology endophenotypes of dementia[29]. Further aspects of pathophysiology are discussed in the Supplementary Note. *NPAS3* encodes a neuronal transcription factor implicated in several neuropsychiatric conditions[30] and is reported to have a regulatory function on the expression of reelin[31]. In adults, reelin binds to the ApoE-Receptor2 (apoER2) and the very low-density lipoprotein receptor (VLDLR) modulating AMPA and NMDA activity in the post-synaptic region, affecting APP processing and tau hyperphosphorylation, and competes with apoE in receptor binding[32]. Further, *SHARPIN* variants have been shown to affect NF-kB signalling in the nervous system, a central mediator of inflammatory and immune responses,

and apoE is suggested to interact with NF-kB signalling[33,34]. Additionally, we observed consistent directionality in all European cohorts, even though the interaction test did not reach statistical significance, suggesting that the effect of *MAPT* is attenuated in an ε4 context in agreement with a previous report[35]. Finally, Forest plots of the effect differences for the significant stratified effect comparison tests show similar directionality across cohorts.

Genetic variants interacting significantly with APOE carrier status with augmented effect size in ε4 carriers

The *HLA* region on chromosome 6 is highly complex. The present data add an extra layer to this complexity since we observed that one *HLA* locus associates with increased risk of AD in ε4 carriers, but not in ε3 carriers, while another independent *HLA* locus associates with increased risk of AD in both strata. Further, in the present study we confirmed *CLU* as one of the strongest genomic signals for AD, and documented for the first time that this signal was substantially stronger in *APOE* ε4 carriers compared to ε33 carriers. We also identified a new independent signal within the *FERMT2* locus, where the nearest gene is *DDHD1* (distance 241,028 bp) which encodes a member of the Phospholipase A1 family important in lipid and phospholipid metabolism[23]. The fact that the rs10131116 *DDHD1* variant in the present GWAS was associated with a decreased risk of AD specifically in *APOE* ε4 carriers together with the recent identification of the rs10131116 as an eQTL associated with decreased *DDHD1* gene expression[6,36], highlights *DDHD1* as an interesting focus for drug discovery. *DDHD1* is also in the same biological pathway as a genome-wide significant known signal (*PLCG2*)[6]. Recently *DDHD2* – a gene known to be involved in hereditary spastic paraplegia - was reported to provide flux of saturated fatty acids for neuronal energy and function[37], supporting a potential importance of this class of genes in neuronal functioning. Importantly, the *HLA* and *DDHD1* signals were similar in European and Asian cohorts,

despite differences in statistical power. Finally, Forest plots of the effect differences also here show similar directionality across cohorts.

Conclusion

By performing the to date largest *APOE* stratified GWAS, we have identified novel as well as well-established AD loci, where the effect is manifested specifically in an *APOE* ε4 carrier or in an ε33 genotype context. These findings are supported by pathway analysis, highlighting distinct *APOE* carrier status dependent biological mechanisms, and may have the potential to change our current understanding of the pathogenesis of AD. Based on our present findings and based on the observed less efficacy and higher risks of severe side effects in ε4 carriers in recent trials of anti-amyloid therapies, it may now be timely to consider including *APOE* stratification a priori in future randomized clinical trials of emerging AD medicines.

(Main text wordcount: 2213 (max 2000))

Methods

Populations

We used the following European ancestry consortia/biobanks: European Alzheimer's Disease & Dementia BioBank (EADB), FinnGen, European Alzheimer's Disease Initiative (EADI), Bonn, Genome Research at Fundacio ACE (Gr@CE), Genetic and Environmental Risk in Alzheimer's Disease (GERAD), Alzheimer's Disease Sequencing Project (ADSP), and UK Biobank (UKB). Additionally, we evaluated 33 regions of interest in cohorts of East Asian (EAS) ancestry, representing Japanese (JADNI, CL, NP; EAS-JPN), Chinese (HKS; EAS-CHN), and Korean (GARD; EAS-KOR) populations, as well as in Asian American (ADSP-AAC), African American (ADSP-AFR), and in native admixed American (ADSP-AMR) multi-ancestry populations. *APOE* genotype was determined by the imputed data using rs7412 and rs429358. Where directly genotyped data was available, samples with a mismatch between the imputed and genotyped *APOE* genotype were excluded. FinnGen R11 was used, excluding samples from the ADGEN study as they are embedded in the EADB. AD cases were defined by diagnosis or by use of AD medication (ATC code: N06D). Individuals diagnosed with other forms of dementia were excluded from the controls, and all individuals with an AD diagnosis were included independent of other dementia diagnoses. In UKB only those defined as white British were included, and AD was defined as being diagnosed with AD from electronic medical records (EMR). No information regarding proxies was used in the AD definition or as exclusion criteria. In all datasets, controls younger than 60 years were excluded to better balance the age distributions in the cases and controls and to avoid the inclusion of young controls. All cases were kept. Written informed consent was obtained from study participants or, for those with substantial cognitive impairment, a caregiver, legal guardian, or other proxy. Study protocols for all cohorts were reviewed and approved by the appropriate institutional review boards.

Quality control and imputation

A standard quality control was performed on the samples and variants in all datasets. The samples were imputed using the TOPMed except for ADSP and FinnGen (Supplementary Table 2). FinnGen was imputed with a Finnish whole-genome sequencing (WGS) reference panel (SiSu v4). Ancestry estimates and QC for UKB and ADSP were done using GenoTools[38]

GWAS analysis

Test of the association between clinical AD status and autosomal genetic variants were conducted separately in each cohort by means of logistic regression or mixed models using an additive genetic model. Three software implementations were used, SNPTEST 2.5.6[39], PLINK2[40] and REGENIE[41] and adjusted for sex, age, genomic principle components (gPCs) and genotyping centers/batches when necessary (Supplementary Table 2). gPCs were calculated in the full cohort and not in the individual strata to avoid generating possible ascertainment bias effects. Sensitivity analyses were carried out in some datasets to check that adjusting for age did not introduce spurious findings. In SNPTEST we analyzed the genotype probabilities using the newml method. In PLINK2 and REGENIE dosages were used combined with the glm regression (Firth regression if failed convergence) in PLINK2 and Firth regression in REGENIE. For each dataset we filtered out variants with (a) missing data on the effect size, standard error or p-value, (b) an absolute effect above 5, (c) an imputation quality below 0.3, and (d) variants not fulfilling $2\times\min(N_{cases},N_{controls})\times MAF\times info > 5$ (an unbalanced MAC-info score), where info is imputation quality. A fixed-effect meta-analysis using an inverse-variance weighted as implemented in METAL v2020-05-05 was performed combining the results from each dataset. Variants were excluded if a heterogeneity p-value was below 5×10^{-8} or if variants did not pass quality control in at

least two of the three major datasets (EADB-TOPMed, FinnGen, UKB). The genomic inflation factor was computed with a median approach after exclusion of the *APOE* region (44-46 Mb on chromosome 19 in GRCh38) both for variants with MAF>1% and in the entire dataset. Manhattan plots were made using topr package[42] (v.2.0.2) in R (v. 4.3.3).

Definition of loci

Around each variant with a p-value below 2.5×10^{-8}, a region of ±500kb was defined per fixed-effect meta-analysis. We assumed the individual stratified GWASs as separate families of analyses when setting the p-value threshold. Most of the variants will be highly correlated across the strata except for those variants that interact with the *APOE* genotype which is expected to be a small minority of the total tested variants. Hence, if one was to consider all GWAS tests as belonging to the same family it would lead to a large increase in risk of Type II errors without much gain in controlling for Type I errors. However, to be conservative in setting a p-value threshold for the stratified GWAS analysis, we set a stricter Bonferroni correction of $p<5\times10^{-8}/2=2.5\times10^{-8}$ to take into account the addition of the two main strata (ε33 and ε43+ε44).We used the PLINK2 clumping procedure to define independent hits in each region. This procedure is iterative, starting with the variant with the lowest p-value in the respective region (the lead SNP). All variants within loci and in linkage disequilibrium (LD) with the lead variant (r^2 higher that 0.01) are assigned to the clump belonging to the lead SNP. If any variant with a p-value below $2.5\text{x}10^{-8}$ is unassigned to a clump in the respective region, the variant with the lowest p-value is found among the remaining variants and the clumping is repeated until all variants have been assigned a clump. LD in the EADB TOPMed imputed dataset was computed using high quality imputed dosages (imputation info>0.8). The clumping procedure was run in both the ε44+ε43 and ε33 GWAS results and the results for the respective loci were compared. Loci plots were generated using locuszoomr (v. 0.3.5) in R (v.

4.3.3). Forest plots of variant effects across *APOE* strata were generated using forestplotter (v. 1.1.2) in R (v. 4.3.3). The independence of several signals within a locus was tested by conditional meta-analysis in the five largest cohorts (EADB TOPMed, FINNGEN, UK Biobank, GR@CE, ADSP). and meta-analyzed.

Stratified effect comparison tests

Two different tests of stratified effect comparisons between autosomal variants and the *APOE* strata were performed using summary statistics results from the different *APOE* strata. P-value threshold for significance was determined using a Bonferroni correction for 29 regions of interest (including the nine loci in common in both strata, the nine and seven loci found in the ε44+ε43 and ε33 strata respectively and the four additional signals found in four of the 25 loci; p<0.05/29 regions of interest=0.0017). The first test analysed the effect difference between the ε44+ε43 and ε33 *APOE* strata. The effect difference was calculated in each cohort separately ($\Delta\beta_i = \beta_{i,\varepsilon44+\varepsilon43} - \beta_{i,\varepsilon33}$, i is the cohort, β is the autosomal effect estimated in the stratified GWASs), with the SE of the effect difference calculated as the square root of sum of squares of SE for the effects ($SE_{\Delta\beta_i} = \sqrt{SE^2_{i,\varepsilon44+\varepsilon43} + SE^2_{i,\varepsilon33}}$). The effect differences were combined across studies in a fixed effect meta-analysis with an inverse-variance weighted approach (METAL v2020-05-05 software). The test was referred to as the dominant test because it tests if the presence of an ε4-allele changes the effect of the autosomal variant (independent of number of ε4-alleles). Forest plots for the calculated effect difference in each cohort was provided to access the robustness of the stratified effect comparison test across the cohorts. The second test was a fixed effect model estimating the effect from the number of ε4 alleles: $\beta_{ij} = \gamma_0 + \gamma_{\varepsilon4}x_{ij} + \gamma_{cohort}C_{ij} + \varepsilon_{ij}$, where i is the cohort, j is the strata (ε33, ε43, ε44), β_{ij} is the autosomal effect, x_{ij} is the number of ε4-alleles (0,1,2), C_{ij} is the cohort, γ_0 is the intercept, $\gamma_{\varepsilon4}$ is the effect of one ε4 allele, γ_{cohort} is the cohort effect and ε_{ij} the error term. The

second model was referred to as the additive model and was estimated using R (v.4.3.3) and Metafor package (v4.6.0).

We performed mixed effect sensitivity interaction models to test if the stratified effect comparisons results were prone to between-strata relatedness bias, which could be a potential bias in the FinnGen cohort. The models were specified as following. Dominant mixed effect model: $\beta_{ik} = \gamma_0 + \gamma_k x_{ik} + u_i + \varepsilon_{ik}$, where i is the cohort, k is the strata (ε33, ε44+ε43) x_{ik} is the strata (ε33, ε44+ε43, categorical), γ_k is the strata effect ($\gamma_{\varepsilon43+\varepsilon44}$ is the reference) and u_i is the random effect**.** Additive mixed model: $\beta_{ij} = \gamma_0 + \gamma_{\varepsilon4} x_{ij} + u_i + \varepsilon_{ij}$, with the same notation as above. For the calculation of effect difference both the ε33 and ε44+ε43 effects should be available for a cohort to be included in the analysis. In the other models we included all available data, e.g. the ε44 GWAS in the Bonn cohort was not performed due to power, but the data from the Bonn ε33, ε43 and ε44+ε43 GWASs were included if the SNPs passed QC.

Genome-wide interaction testing

Interaction testing between *APOE* ε4 carrier status and autosomal genetic variants for association with clinical AD was conducted separately in the five largest cohorts (EADB TOPMed, FinnGen, UK Biobank, GR@CE, ADSP) using logistic regression or mixed models with an additive genetic model. We implemented a gene-environment (GxE) interaction framework, treating *APOE* ε4 carrier status as the exposure modifier. The aim of the analysis was to mimic as closely as possible the main analysis contrasting the ε33 and ε44+ε43 strata. A full multi-allelic gene-gene interaction GWAS across all *APOE* genotypes was regarded as outside the scope of this project.

APOE ε4 carrier status was defined as 0 for individuals in the *APOE* ε33 strata and 1 for individuals in the *APOE* ε44+ε43 stratum. All other strata were excluded from the analysis. The model used for analysis was AD ~ SNP + *APOE* ε4 carrier + SNP×*APOE* ε4 carrier + age + sex + cohort-specific

covariates. The software and the specific adjustments for the different cohorts is detailed in Supplementary Table 2. Results from the individual cohorts were QCed as described in the stratified GWAS and the SNP and SNP×*APOE*_ε4_carrier estimates were meta-analysed separately using an inverse-variance weighted fixed-effect meta-analysis as implemented in METAL v2020-05-05. The same significance threshold ($p<2.5\times10^{-8}$) was used in the interaction model as in the stratified GWAS meta-analysis, as SNP and interaction terms constitute independent tests. In this setup the reference level in the interaction model corresponds to the reference level in the *APOE* ε33 strata GWAS and hence $\beta_{SNP}^{interaction} = \beta_{SNP}^{\varepsilon 33\ strata}$. Likewise the estimated effect for a SNP in the *APOE* ε44+ε43 strata corresponds to $\beta_{SNP}^{\varepsilon 44+\varepsilon 43\ strata} = \beta_{SNP}^{interaction} + \beta_{SNP\ x\ APOE_\varepsilon 4_carrier}^{interaction}$ but with the reference value shifted by the *APOE*_ε4_carrier estimate $\beta_{APOE_\varepsilon 4_carrier}^{interaction}$, which is independent of the individual SNP.

Pathway analysis

Pathway analyses were performed on each meta-analysis stratum result separately using FUMA v1.6.1[43]. As requested by FUMA, all variants were annotated with an rsID using VEP release 112 and then lifted to the GRCh37 assembly using Picard LiftoverVcf tool (v3.1.1)[44]. Variants having no rsID or failing the lift to the GRCh37 assembly were removed from the analysis. Remaining variants were then uploaded to FUMA and the pathway analysis was performed by MAGMA v1.08 [45] using two different windows to assign a variant to a gene: 0kb (main analysis) and a window of 35kb upstream and 10kb downstream (second analysis). To account for multiple testing, we computed a false discovery rate (FDR, Benjamini-Hochberg) based on the number of genes included in the analysis. Pathways having a q-value<0.05 in either of the two windows were considered significant.

eQTL analysis

We performed an *APOE* stratified cis-eQTL mapping analysis (*APOE* ε33 (n=342) and *APOE* ε44+ε43 (n=130)) in the ROSMAP dorsolateral prefrontal cortex (DLPFC; n=560) cohort to investigate association of the lead variant in the *DDHD1* locus (*rs10131116*) with the RNA expression of nearby genes in a 1 Mb window around the variant, following the methodology described as before[6]. For the *APOE* stratified cis-eQTL mapping, the genetic principal components (gPCs) and the gene expression Probabilistic Estimation of Expression Residuals (PEER) factors were calculated within the respective strata separately; and sex, first 3 gPCs, and PEER factors (first 45 for *APOE* ε33 and first 15 for *APOE* ε44+ε43) were included as covariates.

Summary data-based Mendelian Randomization

Summary data-based Mendelian Randomization (SMR) was performed using the SMR software developed and maintained by the Yang Lab, with default parameters[46,47] using summary data from the *APOE* ε33 and *APOE* ε44+ε43 strata GWAS. We used cis-eQTL data from 5 of the 7 regions in the MetaBrain Consortium dataset (cerebellum, cortex, basal ganglia, hippocampus, and spinal cord)[48]. We applied a significance threshold of pSMR_multi <6.12E-06, which corresponds to the Bonferroni-corrected value at $\alpha = 0.05$ for 8,166 unique genes tested across all regions. Additionally, we filtered at a HEIDI p-value threshold of pHEIDI >0.01 to remove associations with inferred pleiotropy and only kept results where the number of SNPs included in the HEIDI tests was greater than 3 (nsnp_HEIDI >3).

Data availability

Summary statistics will be made available upon publication through the European Bioinformatics Institute GWAS Catalog (https://www.ebi.ac.uk/gwas/).

Code availability

The software used is referenced in the Online Methods and the applied codes will be made available on https://github.com/JesperQT/APOEstratGWAS.

Acknowledgements

We thank the many study participants, researchers and staff for collecting and contributing to the data, the high-performance computing service at the University of Lille and the staff at CEA-CNRGH for their help with sample preparation and genotyping and excellent technical assistance. This work was funded by a grant (EADB) from the EU Joint Programme – Neurodegenerative Disease Research. INSERM UMR1167 is also funded by the INSERM, Institut Pasteur de Lille, Lille Métropole Communauté Urbaine and French government's LABEX DISTALZ program (development of innovative strategies for a transdisciplinary approach to AD). This work was further funded by grants from the Lundbeck Foundation (R278-2018-804), the Danish Heart Foundation, and the Research Fund at Sygeforsikringen Danmark (2021-0245). Najaf Amin is funded by National Institute on Aging (NIH) and Oxford-GSK Institute of Molecular and Computational Medicine (IMCM). Cornelia M van Duijn is supported by the US National Institute on Aging (NIH), NovoNordisk, the Oxford-GSK Institute of Molecular and Computational Medicine (IMCM), Centre of Artificial Intelligence for Precision Medicines (CAIPM) of the University of Oxford and King Abdul Aziz University, Alzheimer Research UK (ARUK), UK National Institute for Health and Care Research (NIHR) Oxford Research Center (BRC), ZonMW (Delta Dementie) and Alzheimer Nederland. This research was also supported in part by the

Intramural Research Program of the NIH, National Institute of Aging (NIA), National Institutes of Health, Department of Health and Human Services; project number ZO1 AG000534, as well as the National Institute of Neurological Disorders and Stroke. This work utilized the computational resources of the NIH STRIDES Initiative (https://cloud.hoh.gov) through the Other Transaction agreement – Azure: OT2OD032100, Google Cloud Platform: OT2OD027060, Amazon Web Services: OT2OD027852. This work utilized the computational resources of the NIH HPC Biowulf cluster (https://hpc.noh.gov). Part of this research has been conducted using the UK Biobank Resource under Application Number 33601, and this work uses data provided by patients and collected by the National Health Service as part of their care and support. H Leonard is collaborator on UK Biobank Application Number 33601 and had access to the data under this approved project. Full consortium acknowledgements and funding are in the Supplementary Note.

Author contributions

Conception and design: JQ Thomassen, H Leonard, B Ulms, B Grenier-Boley, JC Lambert, C van Duijn, M Nalls, R Frikke-Schmidt.

Data analysis: JQ Thomassen, H Leonard, B Ulms, B Grenier-Boley, S Heikkinen, P Garcia, A Castillo-Morales, M Kikuchi, J Gim, H Cao, F Küçükali, N Amin, D Yoon.

Sample contribution: All authors.

Interpretation of data: JQ Thomassen, H Leonard, B Ulms, B Grenier-Boley, F Küçükali, N Amin, JC Lambert, C van Duijn, M Nalls, R Frikke-Schmidt.

Core writing group: JQ Thomassen, H Leonard, B Grenier-Boley, JC Lambert, C van Duijn, M Nalls, R Frikke-Schmidt.

Competing interests

C van Duijn is currently the Research Director Brain Health of the Health Data Research UK (HDR UK) and the UK Dementia Research Institute (UK DRI), working in partnership with Dementias Platform UK (DPUK)

Some authors' participation in this project was part of a competitive contract awarded to DataTecnica LLC by the National Institutes of Health to support open science research. M Nalls also owns stock in Character Bio Inc. and Neuron23 Inc.

Additional information

Correspondence and requests for materials should be addressed to Jesper Qvist Thomassen, Mike Nalls, and Ruth Frikke-Schmidt.

References

1. Lambert, J.-C. *et al.* Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nature Genetics* **45**, 1452-1458 (2013).
2. Jansen, I.E. *et al.* Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat Genet* **51**, 404-413 (2019).
3. Kunkle, B.W. *et al.* Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Abeta, tau, immunity and lipid processing. *Nat Genet* **51**, 414-430 (2019).
4. Wightman, D.P. *et al.* A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease. *Nat Genet* **53**, 1276-1282 (2021).
5. de Rojas, I. *et al.* Common variants in Alzheimer's disease: Novel association of six genetic variants with AD and risk stratification by polygenic risk scores. *medRxiv* (2020).
6. Bellenguez, C. *et al.* New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat Genet* **54**, 412-436 (2022).
7. Lambert, J.C., Ramirez, A., Grenier-Boley, B. & Bellenguez, C. Step by step: towards a better understanding of the genetic architecture of Alzheimer's disease. *Mol Psychiatry* **28**, 2716-2727 (2023).
8. Corder, E. *et al.* Gene Dose of Apolipoprotein E type 4 Allele and the Risk of Alzheimer's Disease in Late Onset Families. *Science* **261**, 921-3 (1992).
9. Lake, J. *et al.* Multi-ancestry meta-analysis and fine-mapping in Alzheimer's disease. *Mol Psychiatry* **28**, 3121-3132 (2023).
10. Fernandez-Calle, R. *et al.* APOE in the bullseye of neurodegenerative diseases: impact of the APOE genotype in Alzheimer's disease pathology and brain diseases. *Mol Neurodegener* **17**, 62 (2022).
11. Mintun, M.A. *et al.* Donanemab in Early Alzheimer's Disease. *N Engl J Med* **384**, 1691-1704 (2021).
12. van Dyck, C.H. *et al.* Lecanemab in Early Alzheimer's Disease. *N Engl J Med* **388**, 9-21 (2023).
13. Sperling, R.A. *et al.* Trial of Solanezumab in Preclinical Alzheimer's Disease. *N Engl J Med* **389**, 1096-1107 (2023).
14. GTEx Portal (https://www.gtexportal.org/home/snp/rs10131116).
15. Dutta, B. *et al.* Profiling of the Chromatin-associated Proteome Identifies HP1BP3 as a Novel Regulator of Cell Cycle Progression. *Mol Cell Proteomics* **13**, 2183-97 (2014).
16. Al Barashdi, M.A., Ali, A., McMullin, M.F. & Mills, K. Protein tyrosine phosphatase receptor type C (PTPRC or CD45). *J Clin Pathol* **74**, 548-552 (2021).
17. Kaplan, R. *et al.* Cloning of Three Human Tyrosine Phosphatases Reveals a Multigene Family of Receptor-Linked Protein-Tyrosine-Phosphatases Expressed in Brain. *Proc Natl Acad Sci U S A* **87**, 7000-7004.
18. Hoeng, J.C. *et al.* Identification of new human cadherin genes using a combination of protein motif search and gene finding methods. *J Mol Biol* **337**, 307-17 (2004).
19. Kang, H.-G., Evers, M.R., Xia, G., Baenziger, J.U. & Schachner, M. Molecular Cloning and Expression of anN-Acetylgalactosamine-4-O-sulfotransferase That Transfers Sulfate to Terminal and Non-terminal β1,4-LinkedN-Acetylgalactosamine. *Journal of Biological Chemistry* **276**, 10861-10869 (2001).
20. Hiraoka, N., Misra, A., Belot, F., Hindsgaul, O. & Fukuda, M. Molecular cloning and expression of two distinct human N-acetylgalactosamine 4-O-sulfotransferases that transfer sulfate to GalNAcβ1-4GlcNAcβ1-R in both N- and O-glycans. *Glycobiology* **11**, 495-504 (2001).
21. Yi, X., Jiang, X.J. & Fang, Z.M. Histone methyltransferase SMYD2: ubiquitous regulator of disease. *Clin Epigenetics* **11**, 112 (2019).
22. https://www.ncbi.nlm.nih.gov/gene/65121.

23. The Human Protein Atlas (proteinatlas.org).
24. Konishi, Y. *et al.* Deficiency of GDNF Receptor GFRalpha1 in Alzheimer's Neurons Results in Neuronal Death. *J Neurosci* **34**, 13127-38 (2014).
25. Chen, L.Q. *et al.* Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature* **468**, 527-32 (2010).
26. Schweighauser, M. *et al.* Age-dependent formation of TMEM106B amyloid filaments in human brains. *Nature* **605**, 310-314 (2022).
27. Cruchaga, C. *et al.* Association of TMEM106B gene polymorphism with age at onset in granulin mutation carriers and plasma granulin protein levels. *Arch Neurol* **68**, 581-6 (2011).
28. Van Deerlin, V.M. *et al.* Common variants at 7p21 are associated with frontotemporal lobar degeneration with TDP-43 inclusions. *Nat Genet* **42**, 234-9 (2010).
29. Shade, L.M.P. *et al.* GWAS of multiple neuropathology endophenotypes identifies new risk loci and provides insights into the genetic risk of dementia. *Nat Genet* **56**, 2407-2421 (2024).
30. Michaelson, J.J. *et al.* Neuronal PAS Domain Proteins 1 and 3 Are Master Regulators of Neuropsychiatric Risk Genes. *Biol Psychiatry* **82**, 213-223 (2017).
31. Erbel-Sieler, C. *et al.* Behavioral and regulatory abnormalities in mice deficient in the NPAS1 and NPAS3 transcription factors. *Proc Natl Acad Sci U S A* **101**, 13648-53 (2004).
32. Stranahan, A.M., Erion, J.R. & Wosiski-Kuhn, M. Reelin signaling in development, maintenance, and plasticity of neural networks. *Ageing Res Rev* **12**, 815-22 (2013).
33. Asanomi, Y. *et al.* A rare functional variant of SHARPIN attenuates the inflammatory response and associates with increased risk of late-onset Alzheimer's disease. *Mol Med* **25**, 20 (2019).
34. Arnaud, L. *et al.* APOE4 drives inflammation in human astrocytes via TAGLN3 repression and NF-kappaB activation. *Cell Rep* **40**, 111200 (2022).
35. Jun, G. *et al.* A novel Alzheimer disease locus located near the gene encoding tau protein. *Mol Psychiatry* **21**, 108-17 (2016).
36. <final supplementary data Bellenguez et al Nature Genetics.pdf>.
37. Saber, S.H. *et al.* DDHD2 provides a flux of saturated fatty acids for neuronal energy and function. *Nat Metab* **7**, 2117-2141 (2025).
38. Vitale, D. *et al.* GenoTools: an open-source Python package for efficient genotype data quality control and analysis. *G3 (Bethesda)* **15**(2025).
39. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* **39**, 906-13 (2007).
40. Chang, C.C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**(2015).
41. Mbatchou, J. *et al.* Computationally efficient whole-genome regression for quantitative and binary traits. *Nat Genet* **53**, 1097-1103 (2021).
42. Juliusdottir, T. topr: an R package for viewing and annotating genetic association results. *BMC Bioinformatics* **24**, 268 (2023).
43. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* **8**, 1826 (2017).
44. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol* **17**, 122 (2016).
45. de Leeuw, C.A., Mooij, J.M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol* **11**, e1004219 (2015).
46. Wu, Y. *et al.* Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. *Nat Commun* **9**, 918 (2018).
47. Zhu, Z. *et al.* Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet* **48**, 481-7 (2016).
48. de Klein, N. *et al.* Brain expression quantitative trait locus and network analyses reveal downstream effects and putative drivers for brain-related diseases. *Nat Genet* **55**, 377-388 (2023).

topmed_id [a]	Chromosome	Position [b]	Ref	Eff	Rsid [c]	Loci [d]	Gene [d]	EAF [e]	*APOE* ε33 OR (95% CI) [f]	*APOE* ε33 P-value g	*APOE* ε33 I^2 [h]	*APOE* ε44+ε43 EAF [e]	*APOE* ε44+ε43 OR (95% CI) [f]	*APOE* ε44+ε43 P-value [g]	*APOE* ε44+ε43 I^2 [h]
chr1:20745474:C:T	1	20745474	C	T	rs2274119	HP1BP3	HP1BP3	0.074	1.03 (0.99-1.08)	0.165	11,1	0.072	**1.15 (1.10-1.21)**	**1.33e-08**	0
chr1:155135691:G:A	1	155135691	G	A	rs12726330	SLC50A1	SLC50A1	0.037	**1.22 (1.14-1.31)**	**8.17e-09**	0	0.038	1.00 (0.93-1.07)	0.937	10.3
chr1:198710886:G:A	1	198710886	G	A	rs12733073	PTPRC	PTPRC	0.010	1.20 (1.07-1.35)	1.81e-3	0	0.010	**1.46 (1.29-1.65)**	**1.63e-09**	38.1
chr1:207577223:T:C	1	207577223	T	C	rs679515	CR1	CR1	0.81	**0.91 (0.88-0.93)**	**8.30e-12**	33,8	0.70	**0.87 (0.84-0.90)**	**1.44e-19**	55.2
chr2:127135234:C:T	2	127135234	C	T	rs6733839	BIN1	BIN1	0.38	**1.14 (1.11-1.16)**	**2.30e-26**	64,6	0.38	**1.20 (1.17-1.23)**	**9.88e-46**	58.3
chr6:32411770:C:T	6	32411770	C	T	rs17208902	HLA-DRA [i]	HLA-DRA -1	0.25	1.03 (1.00-1.06)	0.0409	0	0.26	**1.11 (1.08-1.14)**	**4.82e-13**	10.2
chr6:32464090:G:T	6	32464090	G	T	rs9268888	HLA-DRA [i]	HLA-DRA -2	0.54	**0.93 (0.91-0.96)**	**4.57e-09**	0	0.54	0.94 (0.92-0.96)	7.60e-07	54.3
chr6:41161469:C:T	6	41161469	C	T	rs143332484	TREM2 [i]	TREM2 -1	0.011	1.29 (1.16-1.44)	3.27e-06	15.9	0.011	**1.43 (1.27-1.62)**	**9.11e-09**	7.4
chr6:41161514:C:T	6	41161514	C	T	rs75932628	TREM2 [i]	TREM2 -2	0.0029	**2.78 (2.12-3.66)**	**2.37e-13**	0	0.0032	**2.18 (1.66-2.85)**	**1.30e-08**	42.2
chr7:12242825:T:C	7	12242825	T	C	rs7805419	TMEM106B	TMEM106B	0.38	**0.91 (0.89-0.93)**	**4.14e-15**	22.7	0.37	0.97 (0.94-0.99)	8.98e-3	46.4
chr7:100386466:T:C	7	100386466	T	C	rs2906657	PILRA	PILRA	0.30	0.94 (0.91-0.96)	4.71e-07	0	0.30	**0.92 (0.89-0.94)**	**3.39e-10**	44.1
chr8:27362470:C:T	8	27362470	C	T	rs73223431	CLU [j]	PTK2B	0.37	1.06 (1.04-1.09)	7.41e-07	43.8	0.37	**1.10 (1.07-1.13)**	**4.30e-14**	0
chr8:27610986:C:A	8	27610986	C	A	rs867230	CLU [j]	CLU	0.57	**1.08 (1.06-1.11)**	**5.32e-11**	21.9	0.57	**1.15 (1.12-1.18)**	**6.21e-27**	0
chr8:144103704:G:A	8	144103704	G	A	rs34173062	SHARPIN	SHARPIN	0.075	**1.19 (1.14-1.25)**	**7.77e-14**	4.7	0.070	1.08 (1.02-1.14)	4.62e-3	20.2
chr11:60173126:T:A	11	60173126	T	A	rs7232	MS4A6A	MS4A6A	0.34	**0.91 (0.89-0.93)**	**7.67e-14**	0	0.33	**0.89 (0.86-0.91)**	**2.62e-19**	44.9
chr11:86113817:A:G	11	86113817	A	G	rs659023	PICALM[k]	PICALM	0.60	1.07 (1.04-1.10)	4.03e-08	43.8	0.58	**1.11 (1.09-1.14)**	**2.27e-16**	11.6
chr11:86141937:G:A	11	86141937	G	A	rs10792831	PICALM[k]	PICALM	0.54	1.09 (1.06-1.11)	**4.42e-12**	41.2	0.56	1.09 (1.06-1.11)	**2.75e-11**	0
chr11:121564878:T:C	11	121564878	T	C	rs11218343	SORL1	SORL1	0.034	0.89 (0.83-0.95)	2.47e-4	49.6	0.032	**0.81 (0.75-0.87)**	**5.60e-09**	0

topmed_id [a]	Chromosome	Position [b]	Ref	Eff	Rsid [c]	Loci [d]	Gene [d]	EAF [e]	*APOE* ε33 OR (95% CI) [f]	*APOE* ε33 P-value g	*APOE* ε33 I^2 [h]	*APOE* ε44+ε43 EAF [e]	*APOE* ε44+ε43 OR (95% CI) [f]	*APOE* ε44+ε43 P-value [g]	*APOE* ε44+ε43 I^2 [h]
chr14:33428905:G:C	14	33428905	G	C	rs187023552	NPAS3	NPAS3	0.016	**1.42 (1.26-1.61)**	**1.04e-08**	0	0.015	0.98 (0.86-1.11)	0.722	82.6
chr14:53394351:T:C	14	53394351	T	C	rs10131116	FERMT2	DDHD1	0.37	1.02 (1.00-1.05)	0.0416	43.0	0.37	**0.93 (0.90-0.95)**	**6.61e-09**	0
chr15:58790588:T:G	15	58790588	T	G	rs347116	ADAM10	ADAM10	0.39	0.96 (0.94-0.98)	9.04e-4	0	0.40	**0.93 (0.90-0.95)**	**1.11e-08**	0
chr15:63279621:C:T	15	63279621	C	T	rs75763893	APH1B [j]	APH1B	0.13	**1.12 (1.08-1.15)**	**2.22e-10**	40.5	0.13	**1.14 (1.10-1.18)**	**3.87e-12**	32.6
chr15:63407216:C:T	15	63407216	C	T	rs181364771	APH1B [j]	LINC02568	0.028	**1.27 (1.18-1.37)**	**1.34e-09**	30	0.029	1.18 (1.09-1.28)	7.69e-05	21.8
chr17:44352876:C:T	17	44352876	C	T	rs5848	GRN	GRN	0.32	**1.09 (1.07-1.12)**	**1.74e-12**	0	0.32	1.05 (1.03-1.08)	1.24e-4	0
chr17:46111701:A:G	17	46111701	A	G	rs7225002	MAPT	MAPT H2	0.38	**0.93 (0.91-0.95)**	**4.88e-10**	36.4	0.39	0.97 (0.95-1.00)	0.0320	31.1
chr17:63470201:G:A	17	63470201	G	A	rs8077276	ACE	ACE	0.58	1.05 (1.03-1.08)	1.10e-05	0	0.58	**1.09 (1.07-1.12)**	**3.45e-12**	14.7
chr18:27352028:C:A	18	27352028	C	A	rs544488330	CHST9	CHST9	0.0025	**2.21 (1.70-2.88)**	**3.69e-09**	70.1	0.0019	1.32 (0.85-2.06)	0.220	4.8
chr19:1050875:A:G	19	1050875	A	G	rs12151021	ABCA7	ABCA7	0.64	**0.91 (0.89-0.94)**	**1.47e-12**	70.7	0.63	**0.91 (0.89-0.94)**	**6.75e-11**	16.6
chr19:54304006:C:T	19	54304006	C	T	rs1761453	LILRA5	LILRA5	0.45	0.97 (0.95-0.99)	9.65e-3	48.8	0.45	**0.93 (0.91-0.95)**	**8.53e-09**	0
chr20:56449045:G:A	20	56449045	G	A	rs113221226	CASS4	CASS4	0.069	0.91 (0.87-0.96)	1.88e-4	37.8	0.063	**0.84 (0.79-0.89)**	**4.34e-10**	0

Table 1: Genome wide significant hits

[a)] Topmed R2 identifier [b)] GRCh38 assembly. [c)] Reference single-nucleotide polymorphism (SNP) (rs) numbers, according to dbSNP build 156 [d)] Nearest protein-coding or long intergenic non-protein coding RNA according to Ensembl release 111. [e)] Effect allele frequency [f)] Odds ratio (OR) and 95% confidence intervals (CI) calculated with respect to the effect allele. [g)] Two-sided raw P-values derived from a fixed-effect meta-analysis. P-value <2.5e-8 is considered genome wide significant due to multiple GWAS strata tested. [h)] Heterogeneity I^2 statistics. [i)] In the *HLA-DRA* and *TREM2* loci different independent variants reached genome wide significance level in the two strata. [j)] In the *CLU* and *APH1B* loci a second independent variant reached genome wide significance in one of the strata. [k)] In the *PICALM* locus two different variants in the two strata reached genome wide significance and had the lowest p-value in the locus. However, the two variants were in linkage disequivilibrium ($r^2 > 0.4$).

				Dominant interaction – effect difference					Additive interaction – per ε4-allele				
topmed_id [a)]	**rsid** [b)]	**Loci**	**Gene** [c)]	**ΔBeta** [d)]	**SE**	**P-value** [e)]	**I^2** [f)]	**Het p-value** [g)]	**Beta** [h)]	**SE**	**P-value**	**I^2** [f)]	**Het p-value** [g)]
chr1:20745474:C:T	rs2274119	HP1BP3	HP1BP3	0.092	0.034	7.2e-3	0	0.97	0.062	0.027	0.025	23	0.18
chr1:155135691:G:A	**rs12726330**	**SLC50A1**	**SLC50A1**	**-0.19**	**0.051**	**1.6e-4**	0	0.80	**-0.13**	**0.040**	**0.0013**	35	0.098
chr1:198710886:G:A	rs12733073	PTPRC	PTPRC	0.20	0.089	0.026	19	0.28	0.13	0.073	0.086	0.58	0.44
chr1:207577223:T:C	rs679515	CR1	CR1	-0.045	0.022	0.038	46	0.072	-0.021	0.017	0.226	52	4.1e-3
chr2:127135234:C:T	**rs6733839**	**BIN1**	**BIN1**	**0.064**	**0.018**	**3.7e-4**	0	0.73	0.045	0.014	1.5e-3	58	**6.5e-4**
chr6:32411770:C:T	**rs17208902**	**HLA-DRA**	**HLA-DRA -1**	**0.078**	**0.020**	**1.1e-4**	0	0.59	**0.057**	**0.016**	**3.6e-4**	0	0.60
chr6:32464090:G:T	rs9268888	HLA-DRA	HLA-DRA -2	0.0089	0.018	0.61	0	0.55	-0.0027	0.014	0.85	24	0.16
chr6:41161469:C:T	rs143332484	TREM2	TREM2 -1	0.098	0.085	0.25	0	0.74	0.087	0.070	0.22	7.5	0.37
chr6:41161514:C:T	rs75932628	TREM2	TREM2 -2	-0.29	0.20	0.16	8.3	0.36	-0.14	0.22	0.51	0	0.53
chr7:12242825:T:C	**rs7805419**	**TMEM106B**	**TMEM106B**	**0.065**	**0.018**	**3.1e-4**	27	0.21	**0.046**	**0.014**	**1.1e-3**	49	7.9e-3
chr7:100386466:T:C	rs2906657	PILRA	PILRA	-0.015	0.019	0.43	0	0.88	-0.014	0.015	0.34	40	0.036
chr8:27362470:C:T	rs73223431	CLU	PTK2B	0.036	0.018	0.042	0	0.58	0.041	0.014	0.0042	0.17	0.46
chr8:27610986:C:A	**rs867230**	**CLU**	**CLU**	**0.060**	**0.018**	**8.0e-4**	0	0.99	**0.051**	**0.014**	**2.8e-4**	0	0.49
chr8:144103704:G:A	**rs34173062**	**SHARPIN**	**SHARPIN**	-0.10	0.037	4.9e-3	47	0.080	**-0.098**	**0.029**	**7.8e-4**	15	0.29
chr11:60173126:T:A	rs7232	MS4A6A	MS4A6A	-0.030	0.019	0.11	44	0.087	-0.027	0.015	0.071	6.7	0.37
chr11:86113817:A:G	rs659023	PICALM	PICALM	0.045	0.018	0.015	13.	0.33	0.039	0.015	0.0071	46	0.020
chr11:121564878:T:C	rs11218343	SORL1	SORL1	-0.079	0.050	0.12	33.	0.16	-0.096	0.041	0.019	18	0.25
chr14:33428905:G:C	**rs187023552**	**NPAS3**	**NPAS3**	**-0.37**	**0.091**	**4.2e-05**	46	0.17	**-0.26**	**0.073**	**0.00032**	65	0.034
chr14:53394351:T:C	**rs10131116**	**FERMT2**	**DDHD1**	**-0.10**	**0.018**	**1.1e-08**	18.	0.29	**-0.069**	**0.014**	**1.2e-06**	32	0.088
chr15:58790588:T:G	rs347116	ADAM10	ADAM10	-0.034	0.018	0.068	0	0.74	-0.023	0.015	0.11	0	0.62
chr15:63279621:C:T	rs75763893	APH1B	APH1B	0.017	0.026	0.51	0	0.64	0.017	0.021	0.41	12	0.30
chr15:63407216:C:T	rs181364771	APH1B	LINC02568	-0.082	0.058	0.16	21	0.27	-0.067	0.047	0.15	30.	0.12
chr17:44352876:C:T	rs5848	GRN	GRN	-0.041	0.019	0.030	0	0.86	-0.034	0.015	0.021	0	0.98

				Dominant interaction – effect difference					Additive interaction – per ε4-allele				
topmed_id [a)]	**rsid** [b)]	**Loci**	**Gene** [c)]	**ΔBeta** [d)]	**SE**	**P-value** [e)]	**I^2** [f)]	**Het p-value** [g)]	**Beta** [h)]	**SE**	**P-value**	**I^2** [f)]	**Het p-value** [g)]
chr17:46111701:A:G	rs7225002	MAPT	MAPT H2	0.045	0.018	0.012	64	0.0075	0.032	0.014	0.025	34	0.066
chr17:63470201:G:A	rs8077276	ACE	ACE	0.031	0.018	0.088	0	0.82	0.028	0.014	0.048	0	0.82
chr19:1050875:A:G	rs12151021	ABCA7	ABCA7	-0.012	0.019	0.52	0	0.89	0.012	0.015	0.43	55	**1.6e-3**
chr19:54304006:C:T	rs1761453	LILRA5	LILRA5	-0.036	0.018	0.046	0	0.61	-0.032	0.014	0.023	0	0.62
chr20:56449045:G:A	rs113221226	CASS4	CASS4	-0.098	0.039	0.011	3.7	0.40	-0.058	0.031	0.061	2.9	0.42

Table 2: Interaction testing results

Results of interaction testing with a dominant and an additive fixed effect model on the summary data of the individual cohorts. The dominant model was a difference of effect analysis while the additive model included the number of ε4-allelles. Significant interactions (p-value < 0.0017) are in bold with the significant p-values underlined. [a)] Topmed R2 identifier [b)] Reference single-nucleotide polymorphism (SNP) (rs) numbers, according to dbSNP build 156 [c)] Nearest protein-coding or long intereneric non-coding RNA according to Ensembl release 111. [d)] ΔBeta is calculated as β_{43+44}- β_{33}. [e)] Two-sided raw P- values were derived from a fixed-effect meta-analysis. [f)] I^2: I-statistics, residual heterogeneity of the unaccounted variability [g)] Het P-value: Test for residual heterogeneity. [h)] effect per one ε4-allele.

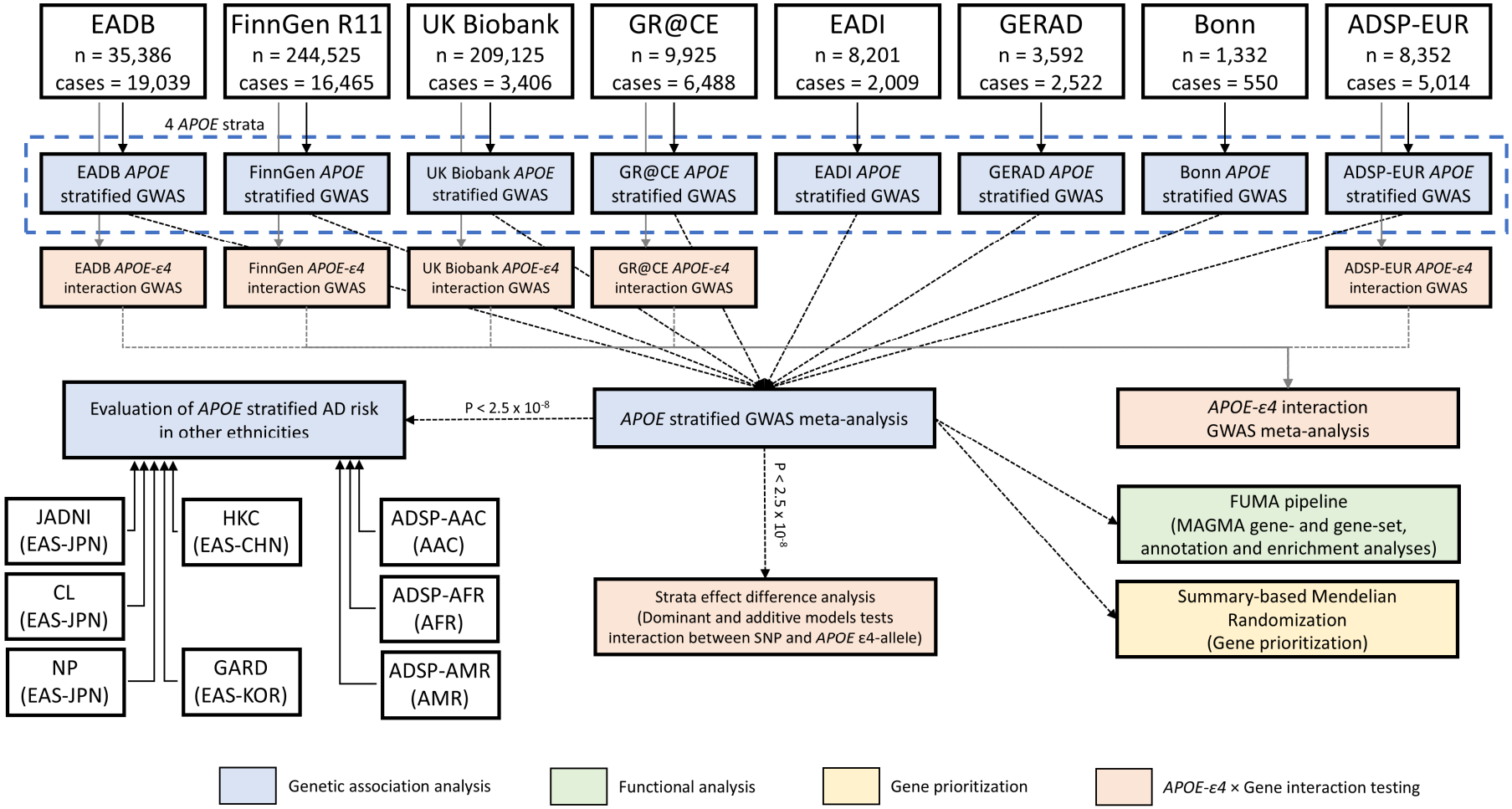

Figure 1. Flowchart of *APOE* stratified GWAS analysis plan.

Eight European, five East Asian and three American cohorts were included. Genetic association analysis (genome-wide association study (GWAS)) was performed in four *APOE* strata (ε33, ε44, ε43+ε44, ε22+ε32). Genome-wide significant signals was tested with functional analysis, genetic characterization, and gene-gene interaction analysis.

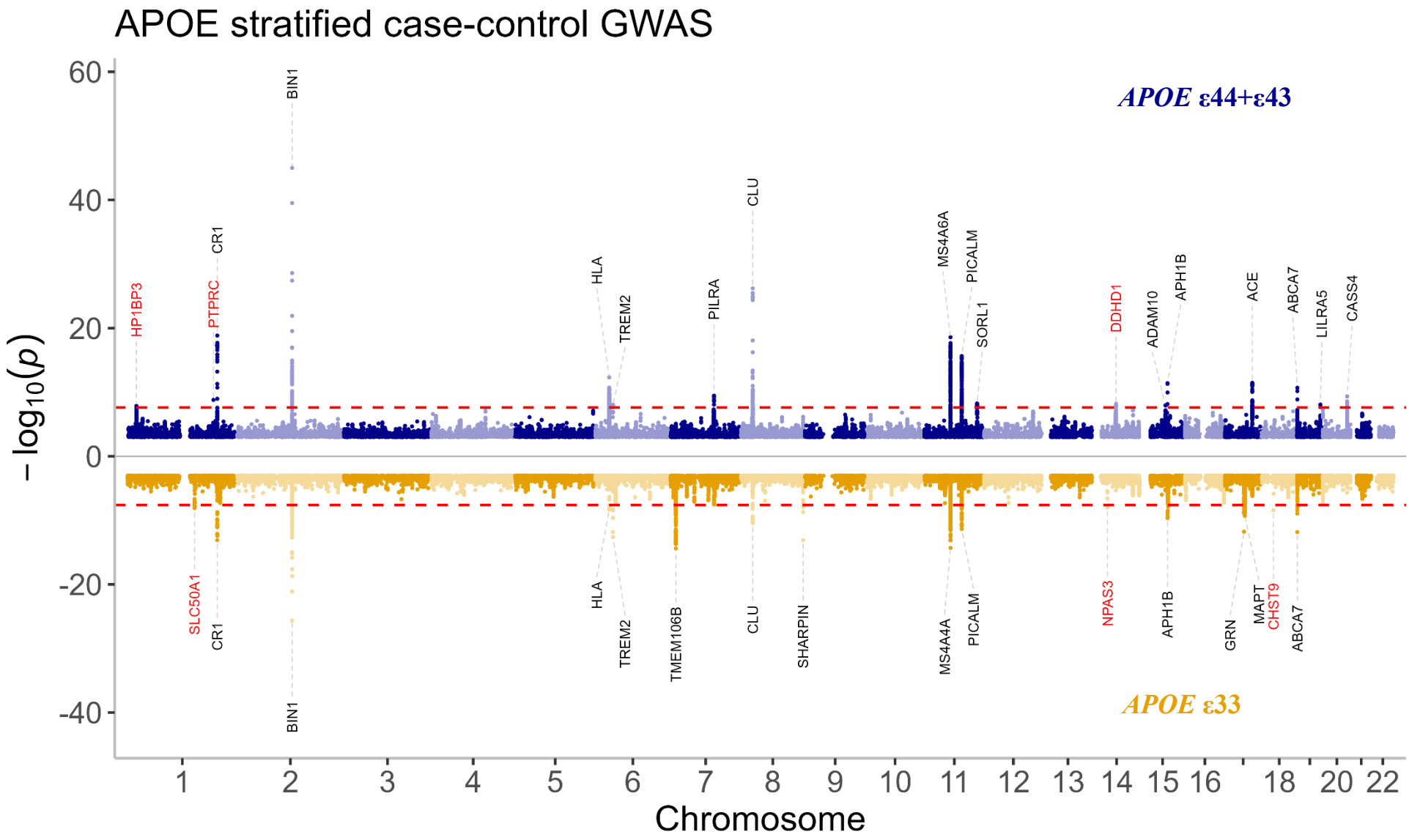

Figure 2: Miami plot of the *APOE* ε33 and ε44+ε43 strata.

The Manhattan plot for the *APOE* ε44+ε43 strata is shown in blue in the upper part of the figure and for the *APOE* ε33 strata in orange in the lower part of the figure. Genome wide significant loci are annotated with the nearest gene (known loci in black and new in red). Two-sided raw P-values were derived from a fixed-effect meta-analysis. The red dashed lines show the genome-wide significant level or two GWAS studies ($P=2.5\times10^{-8}$). *APOE*: Apolipoprotein E gene.

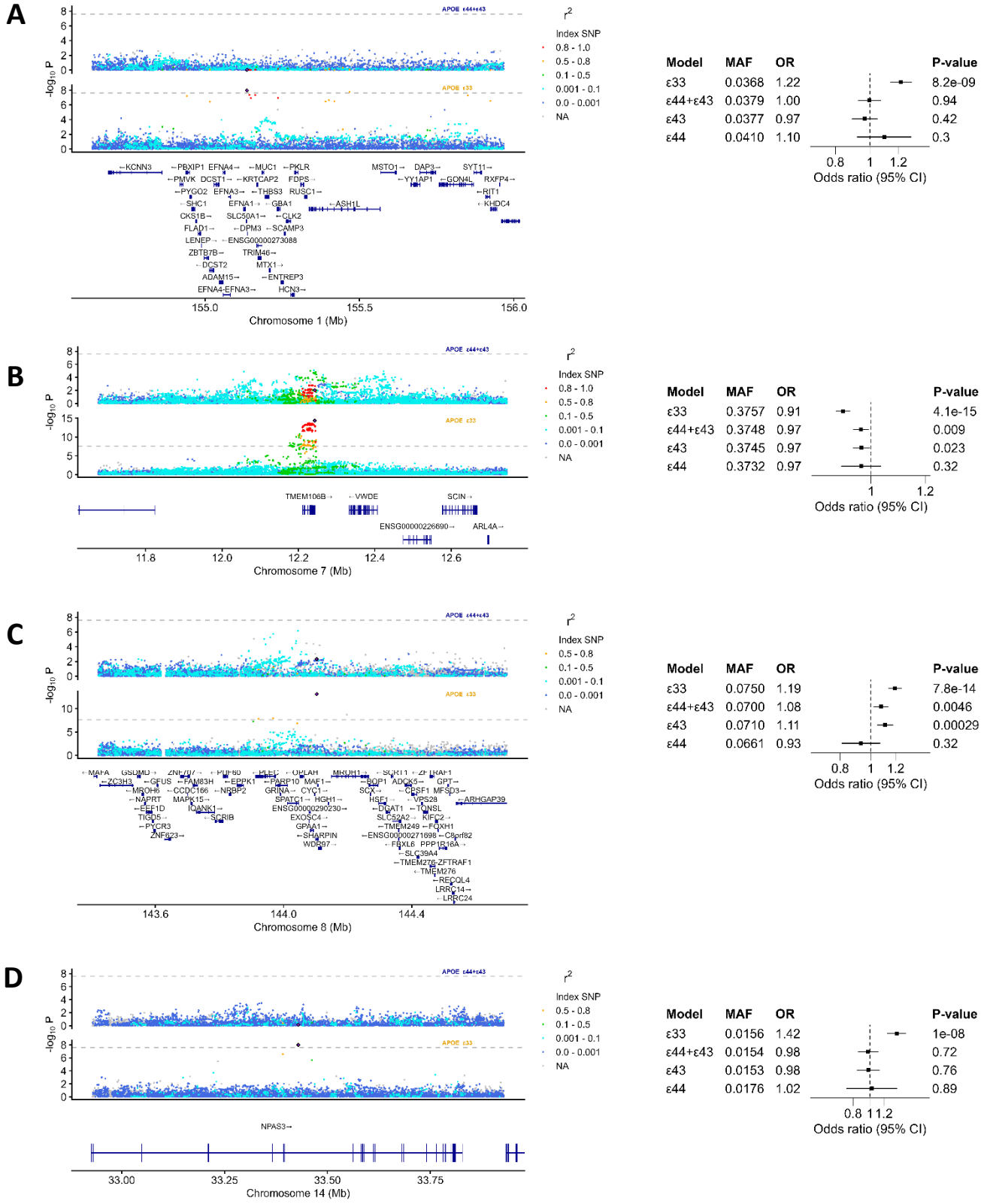

Figure 3: Loci and forest plots for *SLC50A1*, *TMEM106B*, *SHARPIN* and *NPAS3* where the effect attenuates with the *APOE* ε4 allele.

Loci plots for *SLC50A1* (A), *TMEM106B* (B), *SHARPIN* (C), *NPAS3* (D), in *APOE* ε33 and *APOE* ε44+ε43 strata. Forest plots for the lead SNPs in the four *APOE* strata ε33, ε44+ε43, ε43 and ε44. In the forest plot each *APOE* strata is shown to visualize potential dominant or additive interaction.

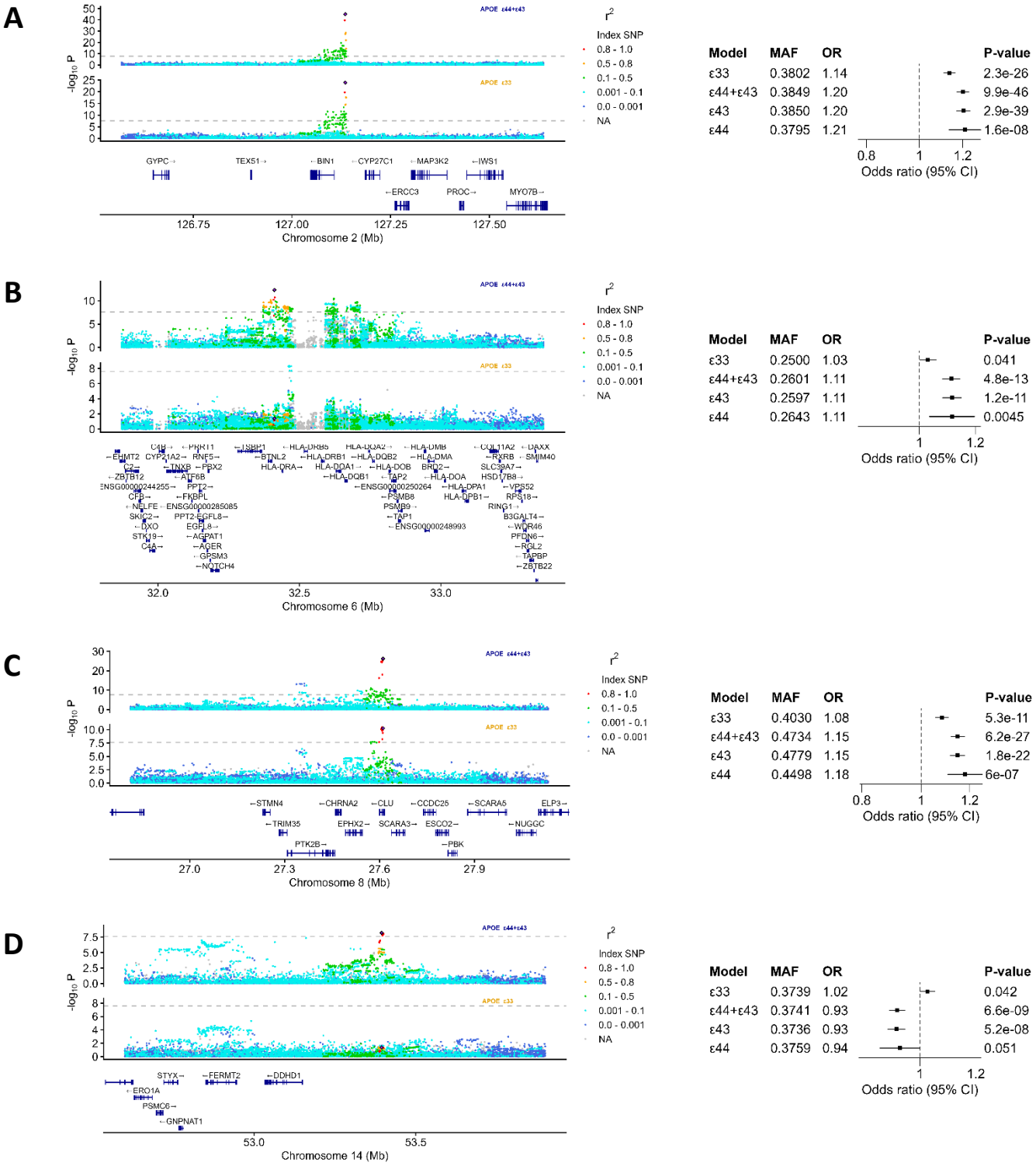

Figure 4: Loci and forest plots for the loci (*BIN1, HLA-DRA -1*, *CLU*, *DDHD1)* where the effect is augmented with the *APOE* ε4 allele.

Loci plots for *BIN1* (A), *HLA-DRA -1 (B)*, *CLU (C)*, *DDHD1 (D)* in *APOE* ε33 and *APOE* ε44+ε43 strata. Forest plots for the lead SNPs in the four *APOE* strata ε33, ε44+ε43, ε43 and ε44. In the forest plot each *APOE* strata is shown to visualize potential dominant or additive interaction.

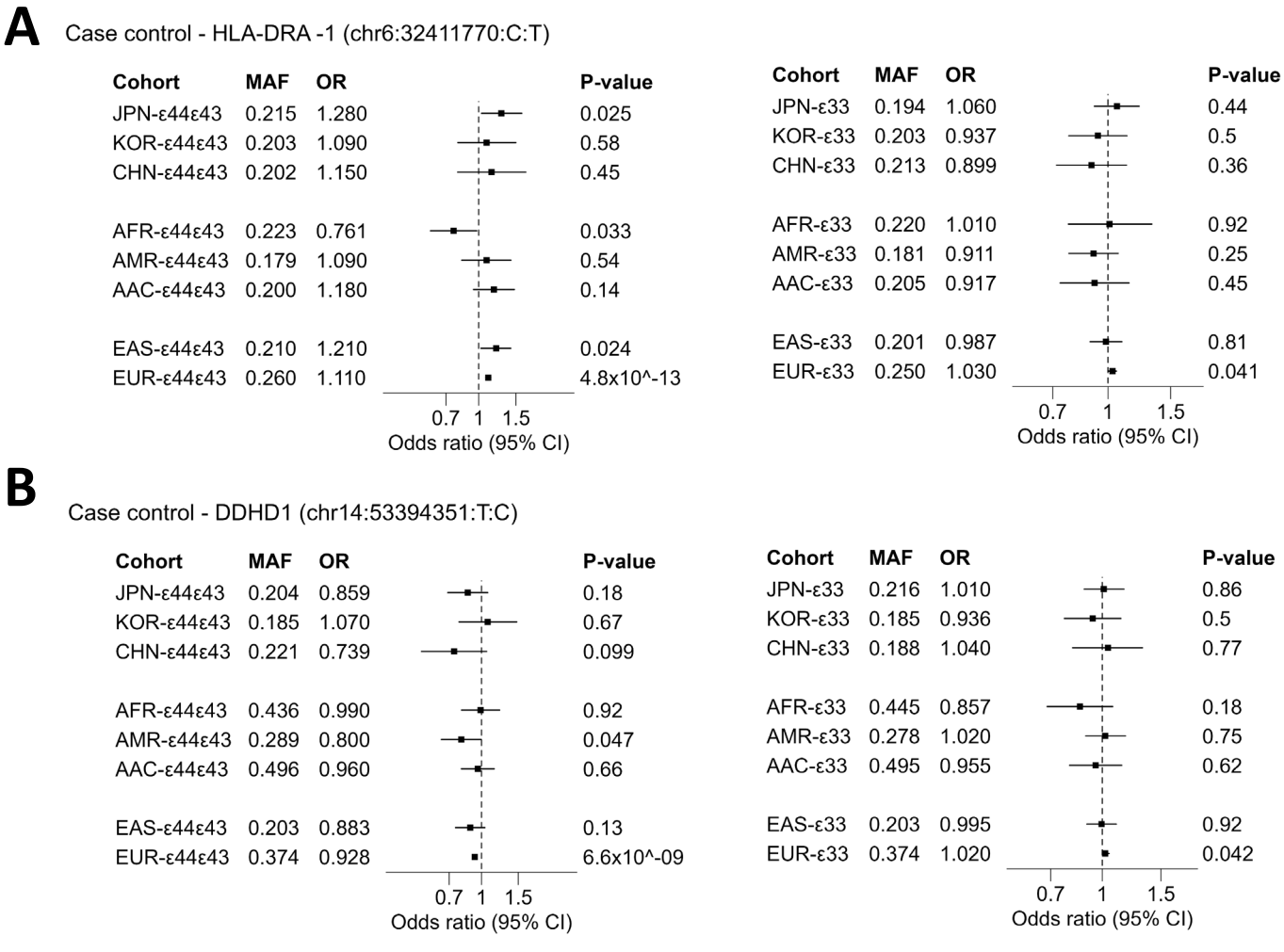

Figure 5: Multi-ancestry evaluation

A: Multi-ancestry results for lead variant in *HLA-DRA* locus. B: Multi-ancestry results for lead variant in *DDHD1* locus. AAC: Asian American, AFR: African American, AMR: Admixed American, CHN: Hong-Kong Chinese, JPN: Japanese, KOR: Korean, EAS: East Asian ancestry meta-analysis, EUR: European ancestry meta-analysis.

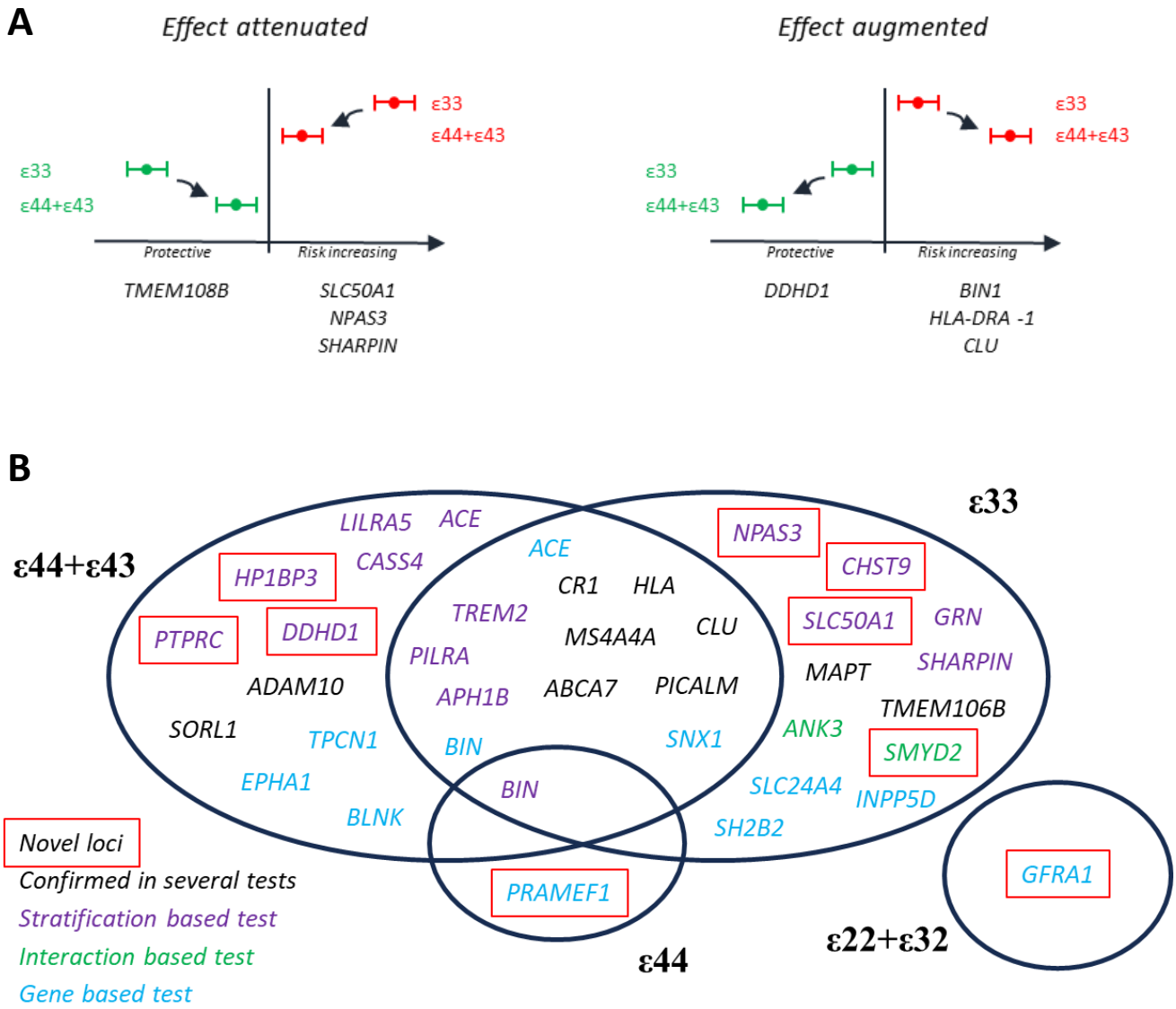

Figure 6: Summary of findings

A) Summary of interaction findings. Left is illustrating attenuated protective and attenuated risk increasing interaction effects for the *APOE*-ε4 allele. Right is illustrating the similar augmented protective and augmented risk increasing interaction effects for the *APOE*-ε4 allele. Loci exhibiting the illustrated effect are noted below. The vertical line illustrates an odds ratio of 1 (no effect).

B) Venn-diagram of the genome-wide significant loci associated with AD, showing the overlap between *APOE*-strata and between variant based testing (stratified or interaction GWAS) and gene-based testing (MAGMA).