Genome-Wide Association Identifies the First Risk Loci for Psychosis in Alzheimer Disease

Mary Ann A. DeMichele-Sweet, PhD,[a] Lambertus Klei, PhD,[a] Byron Creese, PhD,[d] Janet C. Harwood, PhD,[e] Elise A. Weamer, MPH,[b] Lora McClain, PhD,[a] Rebecca Sims, PhD,[e] Isabel Hernandez, MD, PhD,[f,g] Sonia Moreno-Grau, PhD,[f,g] Lluís Tárraga, MS,[f,g] Mercè Boada, MD, PhD,[f,g] Emilio Alarcón-Martín MS,[f] Sergi Valero PhD,[f,g] NIA-LOAD Family Based Study Consortium,[h] Alzheimer's Disease Genetics Consortium (ADGC),[h] Yushi Liu, PhD,[i] Basavaraj Hooli, PhD,[j] Dag Aarsland, MD,[k] Geir Selbaek, MD, Ph.D,[l] Sverre Bergh, MD,[m] Arvid Rongve, MD, PhD,[n] Ingvild Saltvedt, MD,[o] Håvard K. Skjellegrind, MD, PhD,[p] Bo Engdahl, PhD,[q] Eystein Stordal, MD,[r] Ole A. Andreassen, MD, PhD,[ae] Srdjan Djurovic, PhD,[s] Lavinia Athanasiu, PhD,[t] Davide Seripa, PhD,[u] Barbara Borroni, MD,[v] Diego Albani, MSc,[w] Gianluigi Forloni, PhD,[w] Patrizia Mecocci, MD,[x] Alessandro Serretti, MD,[y] Diana De Ronchi, MD,[y] Antonis Politis, MD,[z] AddNeuroMed Consortium,[h] Julie Williams, PhD,[e,aa] Richard Mayeux, MD,[ab] Tatiana Foroud, PhD,[ac] Agustin Ruiz, MD, PhD,[f,g] Clive Ballard, MD,[ad] Peter Holmans, PhD,[af] Oscar L. Lopez, MD,[a,b] M. Ilyas Kamboh, PhD,[c] Bernie Devlin, PhD,[a] *Robert A. Sweet, MD[a,b]

Departments of [a]Psychiatry, [b]Neurology, [c]Human Genetics, University of Pittsburgh, Pittsburgh, PA;, [d]University of Exeter Medical School, College of Medicine and Health, Exeter, UK; Norwegian, Exeter and King's College Consortium for Genetics of Neuropsychiatric Symptoms in Dementia; [e]Division of Psychological Medicine and Clinical Neuroscience, School of Medicine, Cardiff University, Cardiff, UK; [f]Research Center and Memory Clinic Fundació ACE, Institut Català de Neurociències Aplicades. Universitat Internacional de Catalunya, Barcelona, Spain; [g]CIBERNED, Network Center for Biomedical Research in Neurodegenerative Diseases, National Institute of Health Carlos III, Spain; [h]See Supplementary Acknowledgments for the contributors and their affiliations; [i]Global Statistical Science, Lilly Research Laboratories, Eli Lilly and Company, Indianapolis, IN; [j]Neurodegeneration Research, Lilly Research Laboratories, Eli

Lilly and Company, Indianapolis, IN; [k]Department of Old Age Psychiatry, Institute of Psychiatry, Psychology and Neuroscience, King's College London and Centre for Age-Related Medicine, Stavanger University Hospital; [l]Norwegian National Advisory Unit in Ageing and Health, Vestfold Hospital Trust, Tønsberg, Norway and Department Geriatric Medicine, Oslo University Hospital, Oslo, Norway, and Faculty of Medicine, University of Oslo, Oslo, Norway; [m]Research Centre of Age-related Functional Decline and Disease, Innlandet Hospital Trust, Pb 68, Ottestad 2312, Norway; [n]Department of Research and Innovation, Helse Fonna, Haugesund and Department of Clinical Medicine (K1), University of Bergen, Bergen, Norway; [o]Geriatric Department, St. Olav Hospital, University Hospital of Trondheim, Norway and Department of Neuromedicine and Movement science, Norwegian University of Science and Technology (NTNU), Norway; [p]HUNT Research Centre, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), Levanger, Norway and Levanger Hospital, Nord-Trøndelag Hospital Trust, Levanger, Norway; [q]Norwegian Institute of Public Health, Oslo, Norway; [r]Department of Mental Health, Norwegian University of Science and Technology, Trondheim, 43 Norway; [s]Department of Medical Genetics, Oslo University Hospital, Oslo, Norway and NORMENT, Department of Clinical Science, University of Bergen, Bergen, Norway; [t]NORMENT, Institute of Clinical Medicine, University of Oslo, Oslo, Norway; [u]Department of Hematology and Stem Cell Transplant, Vito Fazzi Hospital, Lecce, Italy; [v]Centre for Neurodegenerative Disorders, Department of Clinical and Experimental Sciences, University of Brescia, Italy; [w]Neuroscience Department, Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Milan, Italy; [x]Institute of Gerontology and Geriatrics, Department of Medicine and Surgery, University of Perugia, Italy; [y]Department of Biomedical and NeuroMotor Sciences, University of Bologna, Bologna, Italy; [z]1st Department of Psychiatry, Eginition Hospital, Medical School, National & Kapodistrian University of Athens, Athens, Greece; [aa]UK Dementia Research Institute @ Cardiff, School of Medicine, Cardiff University, Cardiff, UK; [ab]Departments of Neurology, Psychiatry and Epidemiology, Columbia University, New York, NY;

[ac]Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN; [ad]University of Exeter Medical School, Exeter, UK; Norwegian, Exeter and King's College Consortium for Genetics of Neuropsychiatric Symptoms in Dementia; [ae]NORMENT Centre, Institute of Clinical Medicine, University of Oslo, and Oslo University Hospital, Oslo, Norway; [af]MRC Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, UK.

For questions and correspondence, please contact: *Robert A. Sweet, MD

Mail: Biomedical Science Tower, Rm W-1645

3811 O'Hara Street, Pittsburgh, PA 15213-2593

Express Mail: Biomedical Science Tower, Rm W-1645

Lothrop and Terrace Streets, Pittsburgh, PA 15213-2593

Phone: 412-624-0064

Fax: 412-624-9910

Email: sweetra@upmc.edu

Web: www.sweetlab.pitt.edu

Running Title: GWAS of Psychosis in Alzheimer Disease

Abstract

Psychotic symptoms, defined as the occurrence of delusions or hallucinations, are frequent in Alzheimer disease (AD with psychosis, AD+P). AD+P affects ~50% of individuals with AD, identifies a subgroup with poor outcomes, and is associated with a greater degree of cognitive impairment and depressive symptoms, compared to subjects without psychosis (AD-P). Although the estimated heritability of AD+P is 61%, genetic sources of risk are unknown. We report a genome-wide meta-analysis of 12,317 AD subjects, 5,445 AD+P. Results showed common genetic variation accounted for a significant portion of heritability. Two loci, one in *ENPP6* (rs9994623, O.R. (95%CI) 1.16 (1.10, 1.22), $p=1.26x10^{-8}$) and one spanning the 3'-UTR of an alternatively spliced transcript of *SUMF1* (rs201109606, O.R. 0.65 (0.56-0.76), $p=3.24x10^{-8}$), had genome-wide significant associations with AD+P. Gene-based analysis identified a significant association with *APOE*, due to the *APOE* risk haplotype ε4. AD+P demonstrated negative genetic correlations with cognitive and educational attainment and positive genetic correlation with depressive symptoms. We previously observed a negative genetic correlation with schizophrenia; instead, we now found a stronger negative correlation with the related phenotype of bipolar disorder. Analysis of polygenic risk scores supported this genetic correlation and documented a positive genetic correlation with risk variation for AD, beyond the effect of ε4. We also document a small set of SNPs likely to affect risk for AD+P and AD or schizophrenia. These findings provide the first unbiased identification of the association of psychosis in AD with common genetic variation and provide insights into its genetic architecture.

Introduction

Psychotic symptoms, defined as the occurrence of delusions or hallucinations, constitute a phenotype within Alzheimer disease (AD+Psychosis, AD+P) that affects ~ 40% to 60% of individuals with AD and is associated with poor outcomes.[1] In comparison to AD subjects without psychosis (AD-P), AD+P subjects have greater cognitive impairments and experience more rapid declines in cognition and function that begin prior to psychosis onset.[2-9] AD+P is also often associated with increased rates of concurrent neuropsychiatric symptoms, including agitation,[10] aggression,[11,12] and depression.[5,13-15] As a consequence, AD+P is associated with increased rates of other poor outcomes, including greater distress for family and caregivers,[16] higher institutionalization rates,[17-20] worse health,[21] and increased mortality[22] compared to AD-P patients.

The AD+P phenotype is well suited for genetic studies when careful attention is paid to excluding potential phenocopies of both AD+P and AD-P. For example, we have shown that the heritability of AD+P is greatest when requiring the presence of multiple or recurrent psychotic symptoms, rather than a one-time occurrence of a single symptom.[23] Similarly, because psychotic symptoms typically emerge in the transition from mild to moderate stages of AD,[5] individuals without psychosis who are still in the early stages of disease may later manifest psychosis, and therefore, need to be excluded from analysis. Using these approaches to phenotypic characterization, we have previously reported familial aggregation of AD+P,[24] which has since been replicated in two independent cohorts.[5,25] We further estimated the heritability of the presence or absence of psychosis in AD at 61%.[23,26]

Thus, AD+P is likely to be strongly influenced by genetic variation. To date, no study has identified genome-wide significant associations with AD+P, largely due to the small sample sizes of prior studies. However, in prior reports we identified negative genetic correlation of

AD+P risk with risk for schizophrenia.[27,28] We now report a large genome-wide association meta-analysis of 12,317 AD subjects with and without psychosis. We identified two loci with genome-wide significant associations with AD+P, in *ENPP6* and *SUMF1*. In gene-based analyses, only *APOE* ($p=1.23x10^{-6}$) reached the criterion for genome-wide significance. AD+P was negatively genetically correlated with educational attainment and positively with depressive symptoms. Surprisingly, AD+P was not significantly genetically correlated with schizophrenia, but it was negatively correlated with bipolar disorder. Analysis of polygenic risk scores derived from schizophrenia (PRS_{SZ}), and bipolar disorder (PRS_{BP}) GWAS, support these genetic correlations. However, the relationship of schizophrenia risk to AD+P appears subtle, with some established risk SNPs for schizophrenia likely to confer risk for AD+P, while others confer protection.

Materials and Methods

Subjects

This study analyzed samples from 12,317 subjects diagnosed with possible, probable,[29] and when available, autopsy-confirmed definite[30] Alzheimer disease (for subject characteristics see **Table 1A**). Diagnoses were made based on diagnostic evaluations, cognitive testing, and in some cases neuropathologic assessment, conducted during subjects' participation in the following eight source programs as previously described: the Fundació ACE Barcelona Alzheimer Treatment and Research Center (ACE/GR@ACE),[31-33] a Consortium of National Institute on Aging Alzheimer Disease Centers (ADC),[34] Eli Lilly and Company (LILLY),[35,36] the Norwegian, Exeter and King's College Consortium for Genetics of Neuropsychiatric Symptoms in Dementia (NEXGENS),[37-42] the National Institute on Aging's Late Onset Alzheimer's Disease Family Study (NIA-LOAD),[5,26] the National Institute of Mental Health Genetics Initiative AD Cohort (NIMH),[24] the University of Pittsburgh Alzheimer Disease Research Center (PITT ADRC),[43,44] and the MRC genetic resource for Late-onset AD included in the Genetic and

Environmental Risk in AD Consortium (UK-Cardiff).[27,31,45] Collection of clinical data and genetic samples were approved by each source program's local Institutional Review Board or Medical Ethics Committee, as appropriate.

Characterization of Psychosis

Subjects were characterized for the presence or absence of delusions and hallucinations within the individual source programs (including their sub-studies) using the CERAD behavioral rating scale[46] (PITT ADRC and NIA-LOAD), Neuropsychiatric Inventory Questionnaire (NPI-Q,[47] NIA-LOAD, ADC, NEXGENS), NPI-Q Spanish Language Version[48] (ACE/GR@ACE), NPI[49] (UK-Cardiff, NEXGENS, LILLY), and Brief Psychiatric Rating Scale[50] (NIMH). Each of these instruments has established reliability in AD,[5,51] and we have previously used all successfully in analyses of psychosis in AD subjects.[4,5,7,23,43] AD+P was defined by the presence of persistent hallucinations or delusions throughout the course of dementia, AD-P was defined by the absence of all symptoms at all assessments. However, because psychotic symptoms typically emerge in the transition from mild to moderate stages of AD[5], individuals without psychosis, but who were still in the early stages of disease at their last assessment (CDR® Dementia Staging Instrument[52] score < 1, mini-mental state examination score[53] > 20), were considered to be at substantial risk of developing AD+P later in their course. Thus, these individuals were excluded from the analysis. We have used these approaches to characterizing and defining AD+P and AD-P in multiple studies demonstrating the heritability and association with genetic variation of the AD+P phenotype.[5,23,24,26-28,54]

For additional detail of each source program's clinical assessment methodology and demographics, see **Supplementary Material**.

Genotypes

Six of the eight program sources provided us with either blood (ACE/GR@ACE) or DNA samples (PITT ADRC, UK-Cardiff, NIA-LOAD, ADC, NIMH), all of which were processed by the Genomics Core Lab at the University of Pittsburgh. Genomic DNA was extracted from whole blood samples using the Qiamp Blood Mini kit (Qiagen, Valencia, CA). All DNA was quantitated by Pico Green (Thermo Fisher, Pittsburgh, PA) and diluted to a DNA concentration of 23ng/µl. Samples were genotyped at the Children's Hospital of Philadelphia (CHoP, Philadelphia, PA) using Illumina's Global Screening Array (Illumina, San Diego, CA). Prior to genotyping, ChoP confirmed DNA concentrations by Pico Green assay, and performed WGA on samples when necessary.

In addition to the above-mentioned blood and DNA samples, ACE/GR@ACE, LILLY, and NIA-LOAD provided us with single nucleotide polymorphism (SNP) array data. For the ADC, SNP array data was provided by the Alzheimer's Disease Genetics Consortium (ADGC). NEXGENS provided genome-wide association (GWA) statistics for the comparison of AD-P and AD+P. Additional details of the generation of SNP array data for all programs can be found in the **Supplementary Material**.

Analysis

Data from the eight program sources were processed as four cohorts (Phase 1, Phase 2, GR@ACE, and NEXGENS), based on timing of receipt of the data. Data processing, QC, and statistical analyses were uniform across three of the cohorts for which there were genotypes (Phase 1, Phase 2, GR@ACE), whereas only summary statistics were available for the fourth cohort (NEXGENS). All cohorts were analyzed separately for GWA, then statistics per SNP from these analyses were combined by meta-analysis using METAL.[55] Below we describe quality control procedures for the three genotyped cohorts. For more detail of other methods

see the **Supplementary Material**. Additional details for the NEXGENS cohort have been described previously.[56] Methods were implemented within Plink[57,58] unless otherwise noted.

Quality Control (QC) was completed by both genotype and by sample from Phase 1, Phase 2, and GR@ACE. After QC, 6,872 AD-P and 5,445 AD+P subjects, distributed across the four cohorts, remained for analysis (**Table 1B**). We determined ancestry using GemTools analysis[59] of a subset of autosomal SNPs with non-call rate < 0.001 and MAF > 0.05 for Phase 1, Phase 2, and GR@ACE. These SNPs were pruned such that, within a 50 SNP block and a 5 SNP step-size, the linkage disequilibrium $r^2 < 0.01$. For each of the three cohorts, a different subset of SNPs was chosen for ancestry analysis, and the resulting ancestry plots were used to identify the samples in the major European ancestry cluster. Analysis of NEXGENS[56] was restricted to individuals of European ancestry using genetic principal components computed by EIGENSTRAT.[60]

Genotypes were imputed using the Sanger Imputation Server,[61] the 1000 Genomes Phase3 reference panel,[62] and EAGLE2 for pre-phasing[63] for Phase 1, Phase 2, and GR@ACE. Before imputation, the genotypes were harmonized using the perl script HRC-1000G-check-bim-v4.2.5.pl. This resulted in 85,057,462 imputed or genotyped SNPs for each sample. QC of the imputed SNPs included the requirement that the INFO score for a SNP in each data set > 0.81; MAF > 0.01; and, among all European ancestry subpopulations defined by GemTools, Fst < 0.005. For NEXGENS phasing and imputation was done via the Sanger Imputation Service using the Haplotype Reference Consortium (r1.1) reference panel on all cohorts. After imputation, only SNPs with an imputation quality (INFO) score > 0.4 and MAF > 0.05 were retained.

Separate GWA analyses were performed for the Phase 1, Phase 2, and GR@ACE cohorts, to contrast AD+P versus AD-P for the 9,200,578 SNPs using the Plink option --logistic and with adjustment for the three ancestry dimensions (**Supplementary Figures S3-S5**). For chromosome X, an additional covariate for sex was included. For NEXGENS, separate logistic regressions, implemented in PLINK for each of the five NEXGENS consortium datasets (**Tables S11.1-S11.5**), was used to contrast AD+P versus AD-P for each SNP, with adjustment for the first 10 ancestry principal components. METAL software was used to conduct inverse-variance weighted fixed effects meta-analysis across the five NEXGENS datasets, applying genomic control,[55] to generate the summary statistics used in the current analysis. The four GWAS statistics (Phase 1, Phase 2, GR@ACE, NEXGENS summary), per SNP, were then meta-analyzed using METAL.

Heritability of AD+P using GenomicSEM was estimated from 1,126,265 summary statistics from our METAL analysis. Of the 7,105,229 SNPs used for GWAS, 1,126,265 matched to those available on the GenomicSEM website. Also, using genome-wide complex trait analysis (GCTA),[64] heritability was estimated from 9,031 subjects of European ancestry drawn from the Phase 1, Phase 2, and GR@ACE cohorts for which individual genotypes were available (**Table 1B**). Two eigenvectors were used to control for ancestry, 997,105 SNPs were included in the analysis.

Individuals of European ancestry from all four cohorts were used to estimate genetic correlations using LD Score[65] and LD Hub (version 1.9.3).[66] We selected phenotypes for analysis based on prior studies showing correlations with psychosis in AD (years of schooling, depressive symptoms) or genetic association with AD+P (schizophrenia), or because they are closely related with the above phenotypes. Specifically, we included intelligence, which is genetically correlated with years of schooling and bipolar disorder which is strongly genetically

correlated with both depressive symptoms and schizophrenia. Finally, we included AD as it is a necessary condition of AD±P, and two other neurodegenerative diseases, amyotrophic lateral sclerosis (ALS) and Parkinson's disease, each of which is associated with a neuropathology that may contribute to psychosis risk in AD.

We evaluated how well three different polygenic risk scores could differentiate 9,031 AD+P and AD-P subjects of European ancestry. We used the pruning and thresholding approach[67] to compute a PRS for our subjects, developed from GWAS results for AD (PRS_{AD}), [42] schizophrenia (PRS_{SZ}),[68] and bipolar disorder (PRS_{BP}),[69] separately. We used a set of GWAS p-value thresholds for SNP inclusion in each score ($5x10^{-8}$, 0.0001, 0.001, 0.01, 0.1, 0.2, 0.3, 0.4, 0.5).

Gene-based analyses were performed on the summary association statistics using the most recent version (1.08b) of MAGMA.[70] For the primary analysis, SNPs were assigned to genes if they lay within the gene boundaries (as defined by NCBI) and the MAGMA "mean" method was used to derive the gene-wide association statistic (the sum of the squared Z statistics for individual SNPs). A secondary analysis assigned SNPs to genes if they lay within 35kb upstream or 10kb downstream of the gene boundary, to capture regulatory regions.[71]

Gene set enrichment analyses were performed in MAGMA,[70] correcting for the number of SNPs in each gene, linkage disequilibrium (LD) between SNPs and LD between genes. The measure of pathway enrichment is the MAGMA "competitive" test (where the association statistic for genes in the pathway is compared to those of all other protein-coding genes).[72]

Transcriptome-wide association (TWAS) was implemented using the FUSION package[73] was used to perform a TWAS using dorsolateral prefrontal cortex expression data from the CommonMind Consortium and expression data from 13 Brain tissues from the GTEx

(Genotype-Tissue expression) consortium (v7).[74] Results were corrected for multiple testing of multiple genes within each tissue using the Bonferroni method.

See **Supplementary Methods** for additional details of QC, PRS calculation, pathway analyses and TWAS.

Results

Association Analyses

A total of 12,317 subjects, 6,872 AD-P and 5,445 AD+P, were included in this GWAS analysis (**Table 1A**). Contrasting AD-P to AD+P genotypes across the genome revealed two significant loci (**Figure 1, Supplementary Table S1**). One locus was at 4q24, mapping to an intron of *ENPP6* (best SNP rs9994623, O.R. (95%CI) 1.16 (1.10, 1.22), $p=1.26\times10^{-8}$). The other locus was at 3p26.1 (best SNP rs201109606, O.R. 0.65 (0.56-0.76), $p=3.24\times10^{-8}$). This locus spans the 3' untranslated region (3'-UTR) of an alternatively spliced variant of *SUMF1* (SUMF1-204 ENST00000448413.5). None of the SNPs showing significant association in these loci are annotated as expression quantitative trait loci (eQTL) in GTEx. Behavior of the association statistics, as assessed by probability-probability plot (**Supplementary Figure S1**), is consistent with the expectation for such analyses, and the genomic control estimate,[75] GC=1.03, shows no evidence for confounding by ancestry.

For the gene-based tests (**Supplementary Table S2**), only *APOE* ($p=1.23\times10^{-6}$) reached the criterion for genome-wide significance ($p<2.5\times10^{-6}$).[76] This association was only significant for SNPs within *APOE* itself. When the 35/10kb window around genes was used to assign SNPs, no genes reached genome-wide significance. There was substantial association signal for SNPs

in and near *APOE*, however: the smallest p-value achieved was at rs283811 (z = 5.15, p = 2.55×10^{-7}), which falls in an intron of *NECTIN2* (PVRL2 protein). The second smallest p-value occurred for rs429358 (z = 5.12, p = 2.96×10^{-7}), which is one of the two SNPs comprising the *APOE* risk haplotype ε4. These two SNPs, separated by 23,441 bp, were in modest LD (r^2 = 0.52, D' = 0.91[77]) in the 1000G CEU population sample. To determine if ε4 count could explain the AD+P association signal at this locus, we first analyzed a subset of our subjects who were characterized for the ε4 haplotype (2414 AD+P and 2509 AD-P) by logistic regression of AD+P status (yes/no) on ε4 count, after controlling for three eigenvectors for ancestry. The odds increased significantly with count of ε4 haplotypes (OR = 1.21; 95% CI: 1.11-1.31; p = 8.64×10^{-6}). Of the 537 SNPs in this locus, none achieved a p-value<0.0001 in this logistic model after controlling for ε4 count and ancestry (rs2927472 achieved the smallest p-value, 0.00053). Next, using the same subjects, we determined the LD of ε4, in terms of r^2, with 120 SNPs in the *APOE* locus, all of which had association statistic |z| > 2.0 based on the GWAS of AD+P. We then regressed the statistics for these SNPs, |z|, on their LD with ε4, yielding a strongly positive slope (b=3.14, p = 1.90×10^{-41}) and explaining 78.5% of the variance in the observed AD+P z-statistics. Thus, we conclude that the preponderance of AD+P association signal in this locus arises from ε4.

For the pathway enrichment analyses (**Supplementary Table S3**), only the Pathway Interaction Database (PID) IGF1 pathway showed significant enrichment after correction for multiple testing (p=1.17×10^{-6}, q=0.011), although it was no longer significant when the 35/10kb window was used (p=0.0469, q=0.920). Interestingly, one of the pathways found to be significantly enriched for AD risk in Kunkle et al.[78] (GO:48156, tau protein binding) showed significant enrichment (p=6.44×10^{-4} and p=2.21×10^{-3} respectively), albeit not withstanding correction for multiple testing. Given that this pathway includes APOE, the enrichment analysis was repeated excluding genes within 1Mb of APOE (a total of 70 genes), with results shown in

Supplementary Table S3. Removing the genes in the APOE region greatly reduced the significance of GO:48156 (p=0.0106), suggesting that its enrichment is mainly due to APOE.

TWAS comprised a total of 44,185 gene-tissue combinations (**Supplementary Tables S4 and S5**). No TWAS association was significant after correction for the number of tests performed in all genes and tissues combined ($p<1.13\times10^{-6}$, Bonferroni correction for 44,185 tests). Two associations were significant after Bonferroni correction for the number of genes tested in their particular tissue: VN1R108P in GTEx7 hippocampus ($p=2.94\times10^{-6}$) and FAM182B in GTEx7 cerebellum ($p=5.51\times10^{-6}$). For both genes, an increase in gene expression was associated with the presence of psychosis.

SNP-Based Heritability

While earlier studies the AD+P phenotype have shown strong clustering in families and substantial heritability, SNP-based heritability has not been estimated. We estimated it in two ways. First, by analyzing our GWAS statistics using GenomicSEM, SNP-based heritability was estimated at 0.181 ± 0.064 (Chi-square = 8.0, df=1, p = .005). An alternative approach, using the GCTA software, evaluated genotypes genome-wide to determine relationships among the samples and how they partitioned within and between AD+P and AD-P sets. This estimate was 0.312 ± 0.053 (Chi-square = 34.98, df=1, $p = 3.3\times10^{-9}$). The larger estimate probably arises due to greater information contained in estimated genetic relationships, relative to our modestly powered GWAS, although the heritability estimates are not significantly different. The GCTA analysis focused on subjects of European ancestry, genetically determined, to avoid confounding of ancestry.

Genetic Correlation, Polygenic Risk Score, and Risk SNP Analyses

Subjects of European ancestry were also used to estimate genetic correlations of AD+P with select phenotypes available from LD Hub (**Table 2**). Consistent with clinical observations, AD+P is significantly genetically correlated with “Years of Schooling” (and nearly so with the related phenotype, “Intelligence”) and with “Depressive Symptoms”. In contrast, AD+P was not significantly genetically correlated with AD (**Table 2**). Nor was AD+P significantly genetically correlated with the two other neurodegenerative disorders evaluated, ALS and Parkinson disease (**Table 2**).

We previously found a significant relationship between risk for AD+P and schizophrenia.[28] Specifically, we genotyped 94 of 128 SNPs that showed genome-wide-significance for association with schizophrenia in a sample of AD+P subjects. We constructed a predictive score for schizophrenia risk from these SNPs, then assessed whether this score predicted AD+P status in the AD sample. There was a significant negative correlation between the risk score for schizophrenia and AD+P status, which we then replicated by genotyping 60 of the 94 risk SNPs in an independent sample. Now, using SNPs from across the genome and a larger set of AD subjects, results from LD HUB show a negative, but non-significant, genetic correlation with schizophrenia, while showing a negative and significant genetic correlation with bipolar disorder (**Table 2**). In fact, no SNP with p value $<10^{-4}$ for association with psychosis in AD had a p value $<10^{-5}$ in the 108 loci associated with schizophrenia.[79]

Because bipolar disorder and schizophrenia are genetically correlated, we next asked if our original result for the 94 SNPs could be explained by an overlap of risk SNPs for schizophrenia and bipolar disorder. To do so, we tested whether the odds ratios for association of these SNPs for these disorders[69,79] were independent. They were not (**Supplementary Figure S2**); 91 of 94 SNPs had odds ratios exceeding one for both disorders, whereas 47 were expected under independence (sign test, $p = 5.8\times10^{-20}$).

Given the somewhat surprising results for the genetic correlations of AD+P with schizophrenia, bipolar disorder, and AD, we examined whether PRS scores for each of these disorders could differentiate AD+P versus AD-P subjects. In agreement with the genetic correlation, the PRS_{BP} differentiated AD+P from AD-P, whereas PRS_{SZ} showed little ability to differentiate AD+P from AD-P subjects (**Table 3**).

By contrast, the PRS_{AD} did not agree with the genetic correlation of AD and AD+P from LD Hub. PRS_{AD} significantly predicted AD+P status, in the direction of increased risk for AD+P (**Table 3**). Even when we removed the SNP representing the APOE locus, predictions remained positive (**Supplementary Table 20**). Yet the genetic correlation between our AD+P GWAS and Alzheimer's disease, as estimated in LD HUB, was negative, although non-significant. Notably, PRS_{AD} was built on results from a larger AD GWAS[42] than LD Hub uses.[45] Thus, we conjectured perhaps sample size explains the difference. However, when we computed PRS_{AD} using the GWAS data from[45], it predicts AD+P as well as PRS_{AD} from the larger GWAS, if not better (cf **Supplementary Table 21** with **Table 3**) Notably, when we analyzed the genetic correlation of the two AD studies in LD Hub, the genetic correlation was 0.9 with standard error of 0.11. Thus, it seems likely that the difference between the genetic correlation estimated by LD HUB and the results for PRS_{AD} trace to the genetic architecture of AD and differences in power for each method.

Because "uncorrelated" is not the same as "independent", we evaluated one more dimension of these data. Specifically, we evaluated the GWAS-significant SNPs (GWAS SNPs) for schizophrenia,[68] bipolar disorder,[69] and AD[42] to determine whether they also had signal in our AD+P GWAS. We approached this question in two ways. First, we queried the GWAS SNPs to determine if their p-values for AD+P were less than 10^{-4}: Eleven SNPs crossed the threshold, all

in the *APOE* locus and all associated with AD. Next, we reasoned that if some GWAS SNPs also generated risk for or protection from AD+P, whereas others did not, then those AD+P statistics would be represented by a mixture of distributions. We found support for a mixture of distributions for schizophrenia and separately for AD (**Supplementary Figures S6-S8**), while for bipolar disorder there were too few independent GWAS SNPs to have any confidence in our results (**See Supplementary Methods for details**). Curiously, while AD GWAS SNPs were consistent in their effects on AD and AD+P, effects of schizophrenia GWAS SNPs were not. Instead, risk alleles for schizophrenia could impart risk or protection for AD+P (**Supplementary Methods and Supplementary Figure S9**). In addition, using the mixture model results, we identified a small set of SNPs likely to affect risk to AD+P and either AD or schizophrenia (**Supplementary Table S6**).

Discussion

We identified evidence of genome-wide significant association with psychosis risk in AD at SNPs within *ENPP6,* and in the 3'-UTR of an alternatively spliced transcript of *SUMF1*. Exploration of multiple data sets did not reveal any current evidence linking the SNPs at these loci to variation in expression of *ENPP6, SUMF1,* or other genes. Similarly, although the alternatively spliced SUMF1-204 transcript is expressed in brain,[80] AD+P risk SNPs in the SUMF1 locus were not associated with brain expression of SUMF1-204 (S. Sieberts, Personal Communication). Nor were SNPs at these loci linked to other potential genetic mechanisms, such as variation in epigenetic modifications. However, we note that for *SUMF1*, the locus is in the 3'-UTR, a region that often serves a substantial role in regulating protein levels via post-transcriptional mechanisms.[81]

ENPP6 encodes a glycerophosphodiesterase that is highly expressed in new oligodendrocytes as they differentiate from their precursors.[82] Recent data in mice have demonstrated that differentiation of oligodendrocytes from their precursors (as indicated by increased *ENPP6*

mRNA expression) is a necessary component of early,[83] i.e. synaptic,[84] phases of new (motor) learning. ENPP6 protein can be expressed both on the myelin membrane and as a soluble form that is found extracellularly.[85,86] ENPP6 acts as a hydrolase that severs choline from substrates, including lysophosphatidylcholine, glycerophosphorylcholine, and sphingosylphosphorylcholine.[86] Of these, it has highest catalytic efficiency towards sphingosylphosphorylcholine,[85] releasing both sphingosine and phosphocholine. Sphingosine is phosphorylated to generate sphingosine-1-phosphate, which signals via the g-protein-coupled sphingosine-1-phosphate receptor (S1PR1). It is of some interest, therefore, that the S1PR1 modulator, fingolimod,[87] has been previously shown to increase excitatory synaptic transmission[88] and improve psychosis-associated behaviors in a genetic animal model of β-amyloid overproduction.[89]

The locus on chromosome 3 maps to introns spanning the 3'-UTR of an alternatively spliced transcript of *SUMF1*. *SUMF1* encodes formylglycine-generating enzyme, which serves as a master activator of lysosomal sulfatases by converting conserved cysteines to formylglycine in their active sites. As a consequence, genetic disruption of *SUMF1* leads to a multiple sulfatase deficiency syndrome.[90,91] Importantly, the transcript of *SUMF1* (SUMF1-204, ENST00000448413.5), within which our locus is located, encodes an isoform of formylglycine-generating enzyme (isoform 3, Uniprot Accession Q8NBK3-3) lacking the enzymatically active Cys341 residue.[92] The functional consequences of this change are not established, but would be anticipated to reduce or eliminate the primary enzymatic function. The function of the novel sequence that replaces the c-terminal of formylglycine-generating enzyme in isoform 3 is also not known, and BLAST of this sequence against the UNIPROT database does not identify homologous proteins. Nevertheless, speaking to the potential functional impact in AD+P, ENST00000448413.5 is detectable in cerebral cortex.[80]

Recently, an appreciation of how lysosomal storage dysfunction also leads to impaired autophagy has emerged.[93] It is thus not surprising, therefore, that selective depletion of SUMF1 in either astrocytes or neurons results in neurodegeneration.[94] How alterations in function of formylglycine-generating enzyme, due to a potential change in levels of isoform 3, may modify the course of AD through these mechanisms, and thus result in the AD+P phenotype, remains speculative. We have previously shown, however, that preservation of synaptic protein levels in the context of AD neuropathology is associated with reduced psychosis risk.[89] Thus, genetic alterations that impact degradation of synaptic proteins by the lysosome to autophagosome pathway are likely to influence risk of psychosis.

We and others[34] have previously evaluated the association of psychosis in AD with *APOE* risk haplotype ε4*,* finding inconsistent evidence of association.[34] Our current findings, obtained from by far the largest cohort to address this question, shed further light on these prior observations. We found that SNPs within *APOE* demonstrate gene-based significant association with AD+P, and that this association appears attributable to the presence of the ε4 *haplotype* itself, without a detectable contribution from other SNPs. Because the impact of ε4 on AD+P risk is not large, increasing the odds by 1.21, prior inconsistent associations with AD+P likely resulted from the much smaller sample sizes in all prior studies.

APOE ε4 has been shown to increase the accumulation of amyloid β and phosphorylated tau, and, even in the absence of Aβ overproduction, lead to reductions in dendritic markers and synaptic proteins.[95] For example, we have shown that human ε4 carriers and mice with targeted replacement of ε4 had down-regulation of numerous glutamate signaling and synaptic proteins.[96] Increased phosphotau and reduced synaptic proteins (but not altered amyloid β accumulation) have all been associated with psychosis in AD.[97] Future human and animal

model studies of ε4 that control for the contribution of other loci to the genetic risk for AD+P would be helpful in determining the relative contributions of these mechanisms to AD+P.

We previously identified, and independently replicated, an inverse association between polygenic risk for schizophrenia, defined by a limited set of schizophrenia risk SNPs[79], and risk for psychosis in AD.[28] It was thus somewhat surprising that we saw a non-significant genetic correlation between these two disorders when considering both a larger set of SNPs and a substantially enlarged cohort of AD subjects with and without psychosis. Instead, we identified a negative genetic correlation with risk for bipolar disorder, a disorder that has substantial genetic overlap with schizophrenia. Because our prior analyses relied on a subset of SNPs significantly associated with risk for schizophrenia, and this set also shows enrichment for affecting risk for bipolar disorder (**Supplementary Figure S2**), this overlap probably explains the discrepancy we now observe between our earlier results and the current results for genetic correlations. However, the lack of genetic correlation of AD+P with schizophrenia conceals an underlying complexity. We observed that schizophrenia risk SNPs evidenced significant mixture regarding AD+P risk, such that risk alleles for schizophrenia could impart risk or protection for AD+P.

In contrast, we observed a positive correlation between genetic risk for depressive symptoms and AD+P, consistent with clinical observations of co-occurrence of depressive and psychotic symptoms in AD patients,[5,13-15] and evidence that antidepressant medications may have some effect in reducing psychotic symptoms in AD.[98,99] We also observed a significant negative genetic correlation of educational level with psychosis risk in AD, and a similar pattern, but not quite significant relationship with intelligence. Greater cognitive impairment increases the risk for psychosis in AD; moreover, psychosis in AD is further associated with a more rapid rate of cognitive decline[2-9] (see also review in[100]). The current findings extend these earlier observations, by showing genetic overlap with measures that may be better construed as

indicative of cognitive reserve, as they reflect early life cognitive attainment. Cognitive reserve has long been recognized as protective against developing a degree of cognitive and functional impairment sufficient to lead to a diagnosis of AD.[101] However, somewhat counter-intuitively, once AD is diagnosed, individuals with greater cognitive reserve decline more rapidly.[101] Thus, the genetic correlations we observed may point to a biology underlying the presence of greater cognitive impairment in AD, but not the more rapid decline associated with AD+P.

The above findings are subject to several potential limitations. Although our analysis is the largest GWA study of AD+P to date, it is nevertheless modest in sample size in comparison to studies of related complex traits.[78,79] As our heritability results show, a substantial increase in sample size will identify many additional loci as having a significant association with psychosis risk in AD. Similarly, increased sample size is needed to provide the necessary power to identify genes, transcripts, genetically correlated traits, and pathways in the corresponding analyses that derive from the SNP-based associations.

One of our GWAS-significant SNPs, rs201109606, falls in a genomic region marked as simple repeats. Perhaps because the SNP is difficult to impute, it did not pass QC for some of our data. Nonetheless, the largest two of the four datasets contribute to this result and the estimates of the odds ratios for the two data sets are remarkably similar, 0.684 and 0.637. Moreover, other SNPs at this locus also support the findings (**Supplementary Table S1**), although those associations are not quite GWAS significant. Thus, this result requires replication. For additional discussion of other findings and potential limitations, see **Supplementary Material**.

Currently established treatments for psychosis in AD patients are suboptimal, perhaps reflecting in part that these treatments were not derived to prevent or reverse an identified biology of AD+P.[100] The development of effective, specific, therapeutic targets will therefore require as a

first step delineating this underlying biology. Our study provides the first unbiased evidence of association of specific genetic loci with psychosis in AD and can thus serve as an initial road map to AD+P biology. These findings, in conjunction with available functional genomic and post-mortem data, provide multiple links to mechanisms influencing synaptic function as contributors to psychosis in AD.

Acknowledgments

This study was supported by the following federal grants: AG027224 (RAS), MH116046 (RAS), MH057881 (BD), AG030653 (MIK), AG041718 (MIK), AG066468 (OLL)

A complete acknowledgements list of contributing individuals, consortia, and their grant support can be found in Supplementary Material.

Author Contributions

Each author is expected to have made substantial contributions:

1. to the conception **or** design of the work; RAS, BD, MIK, CB, OLL
2. to the acquisition, analysis, **or** interpretation of data; MAAD-S, LK, BC, JCH, EAW, LM, RS, IH, SM-G, LT, MB, EA-M, SV, YL, BH, DA, GS, SB, AR, IS, HKS, BE, ES, OAA, SD, LA, DS, BB, DA, GF, PM, AS, DDR, AP, JW, RM, TF, AR, CB, PH, OLL, MIK, BD, RAS
3. to the creation of new software used in the work; YL
4. have drafted the manuscript or substantively revised it. MAAD-S, JCH, LK, BC, LM, CB, PH, MIK, BD, RAS

Each author must have approved the submitted version (and any substantially modified version that involves the author's contribution to the study) AND to have agreed both to be personally accountable for the author's own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and the resolution documented in the literature.

Conflicts of Interest

YL and BH are currently employed by and holding stock in Eli Lilly and Company

CB reports grants and personal fees from Acadia pharmaceutical company, grants and personal fees from Lundbeck, personal fees from Roche, personal fees from Otsuka, personal fees from Biogen, personal fees from Eli Lilly, personal fees from Novo Nordisk, personal fees from AARP, grants and personal fees from Synexus, personal fees from Exciva, outside the submitted work.

OLL served as a consultant for Grifols, Inc.

IS has been investigator in the drug trial Boehringer-Ingelheim 1346.0023

OAAA is a consultant to HEALTHLYTIX, speaker honoraria from Lundbeck.

AS is or has been a consultant to or has received honoraria or grants unrelated to the present work from: Abbott, Abbvie, Angelini, Astra Zeneca, Clinical Data, Boheringer, Bristol Myers Squibb, Eli Lilly, GlaxoSmithKline, Innovapharma, Italfarmaco, Janssen, Lundbeck, Naurex, Pfizer, Polifarma, Sanofi, Servier.

MB is a consultant to GRIFOLS, BIOGEN, ROCHE, LILLY, CORTEXYME, ARACLON, MERCK. Grants La Caixa, IMI, ISCIII.

DS reports personal fees from Biogen.

References

1 Sweet, R. A., Nimgaonkar, V. L., Devlin, B. & Jeste, D. V. Psychotic symptoms in Alzheimer disease: evidence for a distinct phenotype. *Molecular Psychiatry* **8**, 383-392 (2003).

2 Ropacki, S. A. & Jeste, D. V. Epidemiology of and risk factors for psychosis of Alzheimer's disease: a review of 55 studies published from 1990 to 2003. *Am. J. Psychiatry* **162**, 2022-2030 (2005).

3 Weamer, E. A. *et al.* The relationship of excess cognitive impairment in MCI and early Alzheimer's disease to the subsequent emergence of psychosis. *Int. Psychogeriatr* **21**, 78-85 (2009).

4 Emanuel, J. E. *et al.* Trajectory of cognitive decline as a predictor of psychosis in early Alzheimer disease in the cardiovascular health study. *Am. J. Geriatr. Psychiatry* **19**, 160-168 (2011).

5 Sweet, R. A., Bennett, D. A., Graff-Radford, N. R. & Mayeux, R. Assessment and familial aggregation of psychosis in Alzheimer's disease from the National Institute on Aging Late Onset Alzheimer's Disease Family Study. *Brain* **133**, 1155-1162 (2010).

6 Seltman, H. J., Mitchell, S. & Sweet, R. A. A Bayesian model of psychosis symptom trajectory in Alzheimer's disease. *Int. J. Geriatr. Psychiatry* **31**, 204-210 (2016).

7 Sweet, R. A. *et al.* Effect of Alzheimer's disease risk genes on trajectories of cognitive function in the Cardiovascular Health Study. *Am J Psychiatry* **169**, 954-962 (2012).

8 Koppel, J. *et al.* Psychosis in Alzheimer's disease is associated with frontal metabolic impairment and accelerated decline in working memory: findings from the Alzheimer's Disease Neuroimaging Initiative. *Am. J. Geriatr. Psychiatry* **22**, 698-707 (2014).

9 Koppel, J. *et al.* Relationships between behavioral syndromes and cognitive domains in Alzheimer disease: the impact of mood and psychosis. *Am J Geriatr Psychiatry* **20**, 994-1000 (2012).

10 Gilley, D. W., Whalen, M. E., Wilson, R. S. & Bennett, D. A. Hallucinations and associated factors in Alzheimer's disease. *J. Neuropsychiatry* **3**, 371-376 (1991).

11 Gilley, D. W., Wilson, R. S., Beckett, L. A. & Evans, D. A. Psychotic symptoms and physically aggressive behavior in Alzheimer's disease. *J. Am. Geriatr. Soc* **45**, 1074-1079 (1997).

12 Sweet, R. A. *et al.* The 5-HTTPR polymorphism confers liability to a combined phenotype of psychotic and aggressive behavior in Alzheimer's disease. *International Psychogeriatrics* **13**, 401-409 (2001).

13 Wilkosz, P. A. *et al.* Prediction of psychosis onset in Alzheimer disease: the role of depression symptom severity and the HTR2A T102C polymorphism. *Am. J. Med. Genet. B Neuropsychiatr. Genet* **144B**, 1054-1062 (2007).

14 Wilkosz, P. A., Miyahara, S., Lopez, O. L., DeKosky, S. T. & Sweet, R. A. Prediction of psychosis onset in Alzheimer disease: The role of cognitive impairment, depressive symptoms, and further evidence for psychosis subtypes. *Am. J. Geriatr. Psychiatry* **14**, 352-360 (2006).

15 Lyketsos, C. G. *et al.* Neuropsychiatric disturbance in Alzheimer's disease clusters into three groups: the Cache County study. *Int. J. Geriatr. Psychiatry* **16**, 1043-1053 (2001).

16 Kaufer, D. I. *et al.* Assessing the impact of neuropsychiatric symptoms in Alzheimer's disease: the Neuropsychiatric Inventory Caregiver Distress Scale. *J. Am. Geriatr. Soc* **46**, 210-215 (1998).

17 Rabins, P. V., Mace, N. L. & Lucas, M. J. The impact of dementia on the family. *JAMA* **248**, 333-335 (1982).

18 Lopez, O. L., Wisniewski, S. R., Becker, J. T., Boller, F. & DeKosky, S. T. Psychiatric medication and abnormal behavior as predictors of progression in probable Alzheimer disease. *Arch. Neurol* **56**, 1266-1272 (1999).
19 Magni, E., Binetti, G., Bianchetti, A. & Trabucchi, M. Risk of mortality and institutionalization in demented patients with delusions. *J. Geriatr. Psychiatry Neurol* **9**, 123-126 (1996).
20 Cummings, J. L., Diaz, C., Levy, M., Binetti, G. & Litvan, I., I. Neuropsychiatric Syndromes in Neurodegenerative Disease: Frequency and Signficance. *Semin. Clin. Neuropsychiatry* **1**, 241-247 (1996).
21 Bassiony, M. M., Steinberg, M., Rosenblatt, A., Baker, A. & Lyketsos, C. G. Delusions and hallucinations in Alzheimer's disease: Prevalence and clinical correlates. *Int. J. Geriatr. Psychiatry* **15**, 99-107 (2000).
22 Wilson, R. S. *et al.* Hallucinations, cognitive decline, and death in Alzheimer's disease. *Neuroepidemiology* **26**, 68-75 (2006).
23 Bacanu, S. A. *et al.* Heritability of psychosis in Alzheimer disease. *American Journal of Geriatric Psychiatry* **13**, 624-627 (2005).
24 Sweet, R. A., Nimgaonkar, V. L., Devlin, B., Lopez, O. L. & DeKosky, S. T. Increased familial risk of the psychotic phenotype of Alzheimer disease. *Neurology* **58**, 907-911 (2002).
25 Hollingworth, P. *et al.* Increased familial risk and genomewide significant linkage for Alzheimer's disease with psychosis. *Am. J. Med. Genet. B Neuropsychiatr. Genet* **144B**, 841-848 (2007).
26 Barral, S. *et al.* Genetic variants associated with susceptibility to psychosis in late-onset Alzheimer's disease families. *Neurobiol. Aging* **36**, 3116-3116 (2015).
27 Hollingworth, P. *et al.* Genome-wide association study of Alzheimer's disease with psychotic symptoms. *Mol. Psychiatry* **17**, 1316-1327 (2012).
28 DeMichele-Sweet, M. A. A. *et al.* Genetic risk for schizophrenia and psychosis in Alzheimer disease. *Mol Psychiatry* **23**, 963-972, doi:10.1038/mp.2017.81 (2018).
29 McKhann, G. *et al.* Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology* **34**, 939-944 (1984).
30 Mirra, S. S. *et al.* The consortium to establish a registry for Alzheimer's disease (CERAD). Part II. Standardization of the neuropathologic assessment of Alzheimer's disease. *Neurology* **41**, 479-486 (1991).
31 Seshadri, S. *et al.* Genome-wide analysis of genetic loci associated with Alzheimer disease. *JAMA* **303**, 1832-1840 (2010).
32 Lambert, M. J., Hatch, D. R., Kingston, M. D. & Edwards, B. C. Zung, Beck, and Hamilton Rating Scales as measures of treatment outcome: A meta-analytic comparison. *J. Consult. Clin. Psychol* **54**, 54-59 (1986).
33 Moreno-Grau, S. *et al.* Genome-wide association analysis of dementia and its clinical endophenotypes reveal novel loci associated with Alzheimer's disease and three causality networks: The GR@ACE project. *Alzheimers Dement* **15**, 1333-1347, doi:10.1016/j.jalz.2019.06.4950 (2019).
34 DeMichele-Sweet, M. A., Lopez, O. L. & Sweet, R. A. Psychosis in Alzheimer's disease in the national Alzheimer's disease coordinating center uniform data set: clinical correlates and association with apolipoprotein e. *Int. J Alzheimers. Dis* **2011** (2011).
35 Doody, R. S. *et al.* Phase 3 trials of solanezumab for mild-to-moderate Alzheimer's disease. *N Engl J Med* **370**, 311-321, doi:10.1056/NEJMoa1312889 (2014).
36 Honig, L. S. *et al.* Trial of Solanezumab for Mild Dementia Due to Alzheimer's Disease. *N Engl J Med* **378**, 321-330, doi:10.1056/NEJMoa1705971 (2018).

37 Lovestone, S. *et al.* AddNeuroMed--the European collaboration for the discovery of novel biomarkers for Alzheimer's disease. *Ann N Y Acad Sci* **1180**, 36-46, doi:10.1111/j.1749-6632.2009.05064.x (2009).
38 Roen, I. *et al.* Resourse Use and Disease Couse in dementia - Nursing Home (REDIC-NH), a longitudinal cohort study; design and patient characteristics at admission to Norwegian nursing homes. *BMC Health Serv Res* **17**, 365, doi:10.1186/s12913-017-2289-x (2017).
39 Helvik, A. S., Engedal, K., Saltyte Benth, J. & Selbaek, G. Time from Symptom Debut to Dementia Assessment by the Specialist Healthcare Service in Norway. *Dement Geriatr Cogn Dis Extra* **8**, 117-127, doi:10.1159/000487233 (2018).
40 Eldholm, R. S. *et al.* Progression of Alzheimer's Disease: A Longitudinal Study in Norwegian Memory Clinics. *J Alzheimers Dis* **61**, 1221-1232, doi:10.3233/JAD-170436 (2018).
41 Bergh, S. *et al.* Cohort profile: the Health and Memory Study (HMS): a dementia cohort linked to the HUNT study in Norway. *Int J Epidemiol* **43**, 1759-1768, doi:10.1093/ije/dyu007 (2014).
42 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, doi:10.1038/s41588-018-0311-9 (2019).
43 DeMichele-Sweet, M. A. *et al.* No association of psychosis in Alzheimer disease with neurodegenerative pathway genes. *Neurobiol Aging* **32**, 555-511 (2011).
44 Weamer, E. A., DeMichele-Sweet, M. A., Cloonan, Y. K., Lopez, O. L. & Sweet, R. A. Incident Psychosis in Subjects With Mild Cognitive Impairment or Alzheimer's Disease. *J. Clin. Psychiatry* **77**, e1564-e1569 (2016).
45 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).
46 Tariot, P. N. *et al.* The behavior rating scale for dementia of the Consortium to Establish a Registry for Alzheimer's Disease. *Am. J. Psychiatry* **152**, 1349-1357 (1995).
47 Kaufer, D. I. *et al.* Validation of the NPI-Q, a brief clinical form of the Neuropsychiatric Inventory. *J Neuropsychiatry Clin Neurosci* **12**, 233-239 (2000).
48 Boada, M., Cejudo, J. C., Tarraga, L., Lopez, O. L. & Kaufer, D. Neuropsychiatric Inventory Questionnaire (NPI-Q): Spanish validation of a brief clinical form of the Neuropsychiatric inventory (NPI). *Neurologia* **17**, 317-323 (2002).
49 Cummings, J. L. *et al.* The neuropsychiatric inventory: comprehensive assessment of psychopathology in dementia. *Neurology* **44**, 2308-2314 (1994).
50 Overall, J. E. & Gorham, D. R. The brief psychiatric rating scale. *Psychol. Rep* **10**, 799-812 (1962).
51 Zubenko, G. S., Rosen, J., Sweet, R. A., Mulsant, B. H. & Rifai, A. H. Impact of psychiatric hospitalization on behavioral complications of Alzheimer's disease. *Am. J. Psychiatry* **149**, 1484-1491 (1992).
52 Hughes, C. P., Berg, L., Danziger, W. L., Coben, L. A. & Martin, R. L. A new clinical scale for the staging of dementia. *Br. J. Psychiatry* **140**, 566-572 (1982).
53 Folstein, M. F., Folstein, S. E. & McHugh, P. R. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res* **12**, 189-198 (1975).
54 Bacanu, S. A. *et al.* Linkage analysis of Alzheimer disease with psychosis. *Neurology* **59**, 118-120 (2002).
55 Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-2191 (2010).

56 Creese, B. *et al.* Examining the association between genetic liability for schizophrenia and psychotic symptoms in Alzheimer's disease. *Transl Psychiatry* **9**, 273, doi:10.1038/s41398-019-0592-5 (2019).
57 Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-575, doi:10.1086/519795 (2007).
58 Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7, doi:10.1186/s13742-015-0047-8 (2015).
59 Lee, A. B., Luca, D., Klei, L., Devlin, B. & Roeder, K. Discovering genetic ancestry using spectral graph theory. *Genet. Epidemiol* **34**, 51-59 (2010).
60 Wang, L., Zhang, W. & Li, Q. AssocTests: An R Package for Genetic Association Studies. *2020* **94**, 26, doi:10.18637/jss.v094.i05 (2020).
61 McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* **48**, 1279-1283, doi:10.1038/ng.3643 (2016).
62 Genomes Project, C. *et al.* A global reference for human genetic variation. *Nature* **526**, 68-74, doi:10.1038/nature15393 (2015).
63 Loh, P. R. *et al.* Reference-based phasing using the Haplotype Reference Consortium panel. *Nat Genet* **48**, 1443-1448, doi:10.1038/ng.3679 (2016).
64 Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* **88**, 76-82, doi:10.1016/j.ajhg.2010.11.011 (2011).
65 Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* **47**, 291-295, doi:10.1038/ng.3211 (2015).
66 Zheng, J. *et al.* LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* **33**, 272-279, doi:10.1093/bioinformatics/btw613 (2017).
67 Purcell, S. M. *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748-752 (2009).
68 Pardinas, A. F. *et al.* Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat Genet* **50**, 381-389, doi:10.1038/s41588-018-0059-2 (2018).
69 Stahl, E. A. *et al.* Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet* **51**, 793-803, doi:10.1038/s41588-019-0397-8 (2019).
70 de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol* **11**, e1004219, doi:10.1371/journal.pcbi.1004219 (2015).
71 Network & Pathway Analysis Subgroup of Psychiatric Genomics, C. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nat Neurosci* **18**, 199-209, doi:10.1038/nn.3922 (2015).
72 de Leeuw, C. A., Neale, B. M., Heskes, T. & Posthuma, D. The statistical properties of gene-set analysis. *Nat Rev Genet* **17**, 353-364, doi:10.1038/nrg.2016.29 (2016).
73 Gusev, A. *et al.* Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet* **48**, 245-252, doi:10.1038/ng.3506 (2016).
74 Consortium, G. T. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* **45**, 580-585, doi:10.1038/ng.2653 (2013).
75 Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55**, 997-1004 (1999).
76 Kiezun, A. *et al.* Exome sequencing and the genetic basis of complex traits. *Nat Genet* **44**, 623-630, doi:10.1038/ng.2303 (2012).

77 Devlin, B. & Risch, N. A comparison of linkage disequilibrium measures for fine-scale mapping. *Genomics* **29**, 311-322, doi:10.1006/geno.1995.9003 (1995).
78 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, doi:10.1038/s41588-019-0358-2 (2019).
79 Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421-427 (2014).
80 Gandal, M. J. *et al.* Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science* **362**, doi:10.1126/science.aat8127 (2018).
81 Tushev, G. *et al.* Alternative 3' UTRs Modify the Localization, Regulatory Potential, Stability, and Plasticity of mRNAs in Neuronal Compartments. *Neuron* **98**, 495-511.e496, doi:10.1016/j.neuron.2018.03.030 (2018).
82 Zhang, Y. *et al.* An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J Neurosci* **34**, 11929-11947, doi:10.1523/JNEUROSCI.1860-14.2014 (2014).
83 Xiao, L. *et al.* Rapid production of new oligodendrocytes is required in the earliest stages of motor-skill learning. *Nat Neurosci* **19**, 1210-1217, doi:10.1038/nn.4351 (2016).
84 Xu, T. *et al.* Rapid formation and selective stabilization of synapses for enduring motor memories. *Nature* **462**, 915-919, doi:10.1038/nature08389 (2009).
85 Greiner-Tollersrud, L., Berg, T., Stensland, H. M., Evjen, G. & Greiner-Tollersrud, O. K. Bovine brain myelin glycerophosphocholine choline phosphodiesterase is an alkaline lysosphingomyelinase of the eNPP-family, regulated by lysosomal sorting. *Neurochem Res* **38**, 300-310, doi:10.1007/s11064-012-0921-z (2013).
86 Sakagami, H. *et al.* Biochemical and molecular characterization of a novel choline-specific glycerophosphodiester phosphodiesterase belonging to the nucleotide pyrophosphatase/phosphodiesterase family. *J Biol Chem* **280**, 23084-23093, doi:10.1074/jbc.M413438200 (2005).
87 Chun, J. & Brinkmann, V. A mechanistically novel, first oral therapy for multiple sclerosis: the development of fingolimod (FTY720, Gilenya). *Discov Med* **12**, 213-228 (2011).
88 Darios, F. D. *et al.* Sphingomimetic multiple sclerosis drug FTY720 activates vesicular synaptobrevin and augments neuroendocrine secretion. *Sci Rep* **7**, 5958, doi:10.1038/s41598-017-05948-z (2017).
89 Krivinko, J., Erickson, S., MacDonald, M., Garver, M. & Sweet, R. Fingolimod Treatment Rescues Psychosis-Associated Behavioral Aberrations in Appswe/Psen1de9 Mice. *The American Journal of Geriatric Psychiatry* **26**, S144-S145, doi:10.1016/j.jagp.2018.01.175 (2018).
90 Dierks, T. *et al.* Multiple sulfatase deficiency is caused by mutations in the gene encoding the human C(alpha)-formylglycine generating enzyme. *Cell* **113**, 435-444, doi:10.1016/s0092-8674(03)00347-7 (2003).
91 Cosma, M. P. *et al.* The multiple sulfatase deficiency gene encodes an essential and limiting factor for the activity of sulfatases. *Cell* **113**, 445-456, doi:10.1016/s0092-8674(03)00348-9 (2003).
92 Holder, P. G. *et al.* Reconstitution of Formylglycine-generating Enzyme with Copper(II) for Aldehyde Tag Conversion. *J Biol Chem* **290**, 15730-15745, doi:10.1074/jbc.M115.652669 (2015).
93 Seranova, E. *et al.* Dysregulation of autophagy as a common mechanism in lysosomal storage diseases. *Essays Biochem* **61**, 733-749, doi:10.1042/EBC20170055 (2017).
94 Di Malta, C., Fryer, J. D., Settembre, C. & Ballabio, A. Astrocyte dysfunction triggers neurodegeneration in a lysosomal storage disorder. *Proc Natl Acad Sci U S A* **109**, E2334-2342, doi:10.1073/pnas.1209577109 (2012).

95 Wolfe, C. M., Fitz, N. F., Nam, K. N., Lefterov, I. & Koldamova, R. The Role of APOE and TREM2 in Alzheimer's Disease-Current Understanding and Perspectives. *Int J Mol Sci* **20**, doi:10.3390/ijms20010081 (2018).
96 Sweet, R. A. *et al.* Apolipoprotein E*4 (APOE*4) Genotype Is Associated with Altered Levels of Glutamate Signaling Proteins and Synaptic Coexpression Networks in the Prefrontal Cortex in Mild to Moderate Alzheimer Disease. *Mol. Cell Proteomics* **15**, 2252-2262 (2016).
97 Krivinko, J. M. *et al.* Synaptic Proteome Compensation and Resilience to Psychosis in Alzheimer's Disease. *Am J Psychiatry* **175**, 999-1009, doi:10.1176/appi.ajp.2018.17080858 (2018).
98 Pollock, B. G. *et al.* A double-blind comparison of citalopram and risperidone for the treatment of behavioral and psychotic symptoms associated with dementia. *Am J Geriatr Psychiatry* **15**, 942-952 (2007).
99 Pollock, B. G. *et al.* A randomized, double-blind, placebo-controlled comparison of citalopram and perphenazine for the acute treatment of psychosis and behavioral disturbances associated with dementia. *American Journal of Psychiatry* **159**, 460-465 (2002).
100 Murray, P. S., Kumar, S., DeMichele-Sweet, M. A. & Sweet, R. A. Psychosis in Alzheimer's Disease. *Biol Psychiatry* **75**, 542-552 (2014).
101 Stern, Y. Cognitive reserve in ageing and Alzheimer's disease. *Lancet Neurol* **11**, 1006-1012, doi:10.1016/s1474-4422(12)70191-6 (2012).

Figure Legends

Figure 1. SNP associations with psychosis in AD. A. Manhattan plot. The x-axis shows genomic position for autosomes and the X chromosome. The y axis shows statistical significance as −log10 (P). Each point represents an analyzed SNP. The dashed horizontal line represents the threshold for genome-wide significance ($p = 5 \times 10^{-8}$). **B-C. Zoom plots of the two genome-wide significant loci.** The x-axis shows genomic position. The left y-axis shows statistical significance as −log10 (P). Each point represents an analyzed SNP, coded by degree of linkage disequilibrium relative to the most significant SNP within the locus. Recombination rate through the region is shown on the right y axis. AD: Alzheimer Disease; LD: linkage disequilibrium; cM: centimorgans; Mb: megabase

Table 1A. Subject Characteristics.

	AD-P 6,872 (55.8%) **N (%) or Mean (SD)**	**AD+P** 5,445 (44.2%) **N (%) or Mean (SD)**	**Total** 12,317 (100.0%) **N (%) or Mean (SD)**
Female	4,008 (58.3)	3,649 (67.0)	7,657 (62.2)
Age of Onset[1]	74 (8.3)	73.4 (8.0)	73.8 (8.2)
Age at Consent	75.6 (7.0)	77.0 (6.9)	76.2 (7.0)
Age at Last Visit[1]	80.5 (8.1)	81.3 (7.7)	80.9 (7.9)
Last MMSE[1]	16.0 (6.5)	13.7 (6.8)	15.0 (6.8)
Last CDR[1]			
0.0	2 (0.0)	5 (0.1)	7 (0.1)
0.5	244 (3.8)	186 (3.8)	430 (3.8)
1.0	2,469 (38.4)	967 (19.6)	3,436 (30.2)
2.0	1,560 (24.3)	1,586 (32.2)	3,146 (27.7)
3.0	986 (15.3)	1,340 (27.2)	2,326 (20.5)
4.0	759 (11.8)	482 (9.8)	1,241 (10.9)
5.0	410 (6.4)	363 (7.4)	773 (6.8)

AD-P: Alzheimer disease without psychosis; AD+P: Alzheimer disease with psychosis; MMSE: Mini mental state exam; CDR: CDR® Dementia Staging Instrument; [1]Data not available for some subjects/source programs; see **Supplementary Tables S8-S15** for details.

Table 1B. Sample size for each cohort contributing to the meta-analysis.

GWA	**Phase 1**		**Phase 2**		**GR@ACE**		**NEXGENS**		**Total**	
Diagnosis	**AD-P**	**AD+P**	**AD-P**	**AD+P**	**AD-P**	**AD+P**	**AD-P**	**AD+P**	**AD-P**	**AD+P**
ALL	3529	3525	1045	495	1646	762	652	663	6872	5445
EUR	2665	2732	833	394	1646	762	652	663	5796	4551

GWA: Genome-wide association; AD-P: Alzheimer disease without psychosis; AD+P: Alzheimer disease with psychosis; EUR: European ancestry; GR@ACE: Genome Research at Fundacio ACE; NEXGENS: Norwegian, Exeter and King's College Consortium for Genetics of Neuropsychiatric Symptoms in Dementia

Table 2. Genetic correlations (r_g) of psychosis in Alzheimer disease with selected relevant phenotypes. Correlations were obtained from LD Hub.[66] Phenotypes in bold were chosen, a priori, based on phenotypic or genetic (schizophrenia, bipolar disorder) analyses.

Phenotype	r_g	Se	Z	p-value
Alzheimer disease	-0.374	0.321	-1.168	0.243
Amyotrophic lateral sclerosis	-0.307	0.300	-1.021	0.307
Parkinson disease	0.142	0.186	0.763	0.446
Years of schooling	-0.312	0.111	-2.816	0.005
Intelligence	-0.200	0.121	-1.650	0.099
Schizophrenia	-0.094	0.081	-1.164	0.244
Depressive symptoms	0.327	0.141	2.316	0.021
Bipolar disorder	-0.287	0.145	-1.976	0.048

Table 3. Prediction of psychosis in Alzheimer disease by polygenic risk scores built using GWAS results for Bipolar disorder, Schizophrenia, and Alzheimer disease, and the pruning and thresholding approach.

Bipolar Disorder

P Value Cut Off	N	OR	l95 OR	u95 OR	P Value	R^2
0.00000005	22	0.968	0.928	1.010	0.13	3.30×10^{-3}
0.0001	785	0.961	0.921	1.003	0.065	3.47×10^{-3}
0.001	3112	0.980	0.938	1.023	0.35	3.09×10^{-3}
0.01	14748	0.959	0.918	1.002	0.064	3.47×10^{-3}
0.1	79818	0.950	0.908	0.994	0.025	3.71×10^{-3}
0.2	134250	0.955	0.913	0.999	0.045	3.56×10^{-3}
0.3	182119	0.953	0.911	0.997	0.035	3.62×10^{-3}
0.4	225630	0.957	0.915	1.001	0.056	3.51×10^{-3}
0.5	265136	0.958	0.916	1.002	0.060	3.49×10^{-3}

Schizophrenia

P Value Cut Off	N	OR	l95.OR	u95.OR	P Value	R^2
0.00000005	329	1.051	1.008	1.096	0.021	8.02×10^{-4}
0.0001	3161	1.012	0.970	1.056	0.57	4.73×10^{-5}
0.001	8488	1.009	0.967	1.052	0.69	2.36×10^{-5}
0.01	26064	1.004	0.962	1.048	0.84	5.80×10^{-6}
0.1	99775	1.006	0.963	1.052	0.78	1.14×10^{-5}
0.2	153878	1.007	0.963	1.053	0.76	1.42×10^{-5}
0.3	198913	1.009	0.965	1.055	0.70	2.24×10^{-5}
0.4	238902	1.011	0.967	1.057	0.62	3.60×10^{-5}
0.5	274264	1.013	0.969	1.059	0.57	4.88×10^{-5}

Alzheimer Disease

P Value Cut Off	N	OR	l95.OR	u95.OR	P Value	R^2
0.00000005	63	1.088	1.043	1.135	9.75×10^{-5}	2.28×10^{-3}
0.0001	488	1.048	1.005	1.093	2.85×10^{-2}	7.18×10^{-4}
0.001	1962	1.059	1.015	1.105	7.76×10^{-3}	1.06×10^{-3}
0.01	12301	1.069	1.025	1.115	1.99×10^{-3}	1.43×10^{-3}
0.1	83722	1.095	1.050	1.142	2.30×10^{-5}	2.69×10^{-3}
0.2	144621	1.096	1.051	1.143	2.12×10^{-5}	2.71×10^{-3}
0.3	195758	1.096	1.051	1.143	2.07×10^{-5}	2.72×10^{-3}
0.4	239694	1.088	1.043	1.134	9.39×10^{-5}	2.29×10^{-3}
0.5	277645	1.088	1.043	1.135	8.80×10^{-5}	2.30×10^{-3}

P Value Cut Off: GWAS p-value threshold used for SNP included in calculating the PRS; N: number of SNPs meeting the threshold; OR: Odds ratio; l95 OR: lower limit of the 95% confidence interval for OR; u95 OR: upper limit of the 95% confidence interval for OR; R^2: partial pseudo-R^2 attributable to the PRS after adjusting for the first 2 ancestry eigenvectors.

URLs

METAL http://csg.sph.umich.edu/abecasis/metal/index.html

GemTools http://www.compgen.pitt.edu/GemTools/GemTools.htm

PLINK https://www.cog-genomics.org/plink2/

NCBI RS names https://ftp.ncbi.nih.gov/snp/redesign/latest_release/VCF/

LD-Hub: http://ldsc.broadinstitute.org/LD Hub/

Sanger Imputation Service: https://imputation.sanger.ac.uk

Psychiatric Genetics Consortium: https://www.med.unc.edu/pgc/

Imputation preparation: https://www.well.ox.ac.uk/~wrayner/tools/

CMC/AMP-AD eQTL Meta-analysis: https://www.synapse.org/#!Synapse:syn16984815

NCBI gene2go file: ftp://ftp.ncbi.nlm.nih.gov/gene/DATA/

Reactome: https://reactome.org/download-data

Gene Ontology: http://geneontology.org/docs/download-ontology/

Molecular Signatures Database: https://www.gsea-msigdb.org/gsea/msigdb/index.jsp

R package qvalue: http://github.com/jdstorey/qvalue

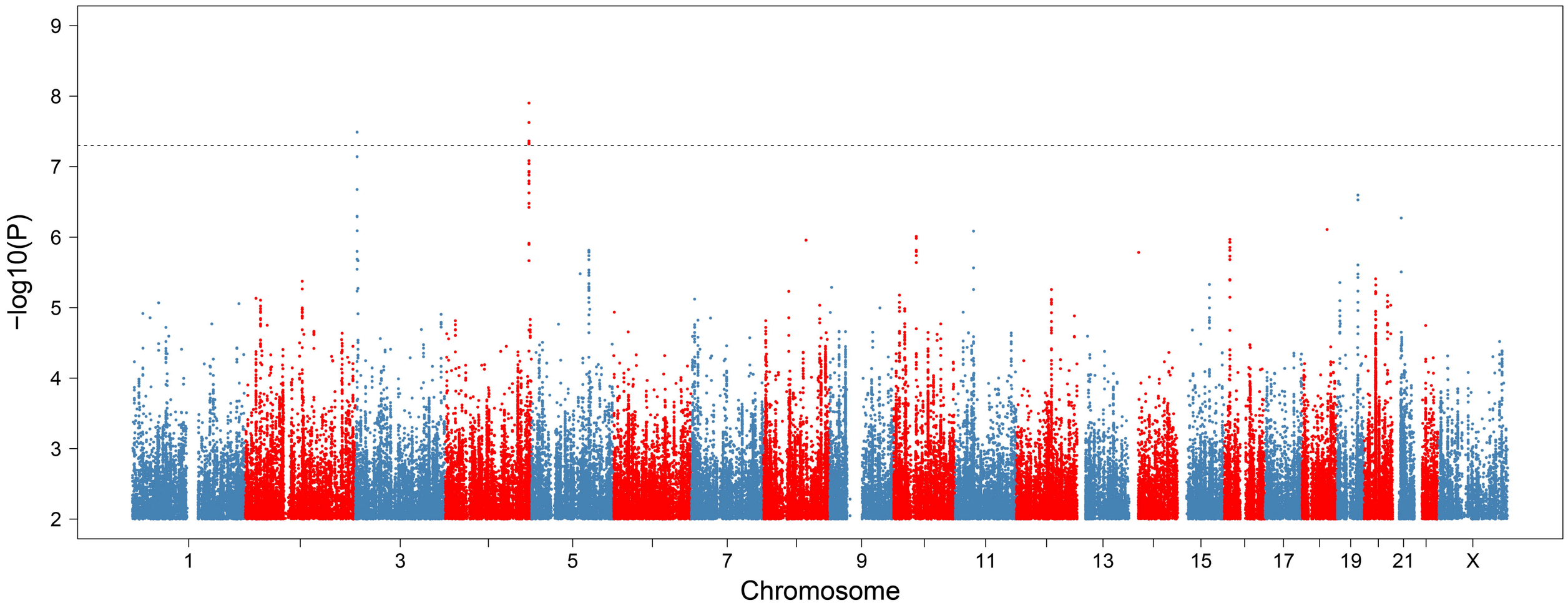
LD Score Regression (LDSC) software v1.0.0 : https://github.com/bulik/ldsc

HapMap 3 https://www.sanger.ac.uk/resources/downloads/human/hapmap3.html

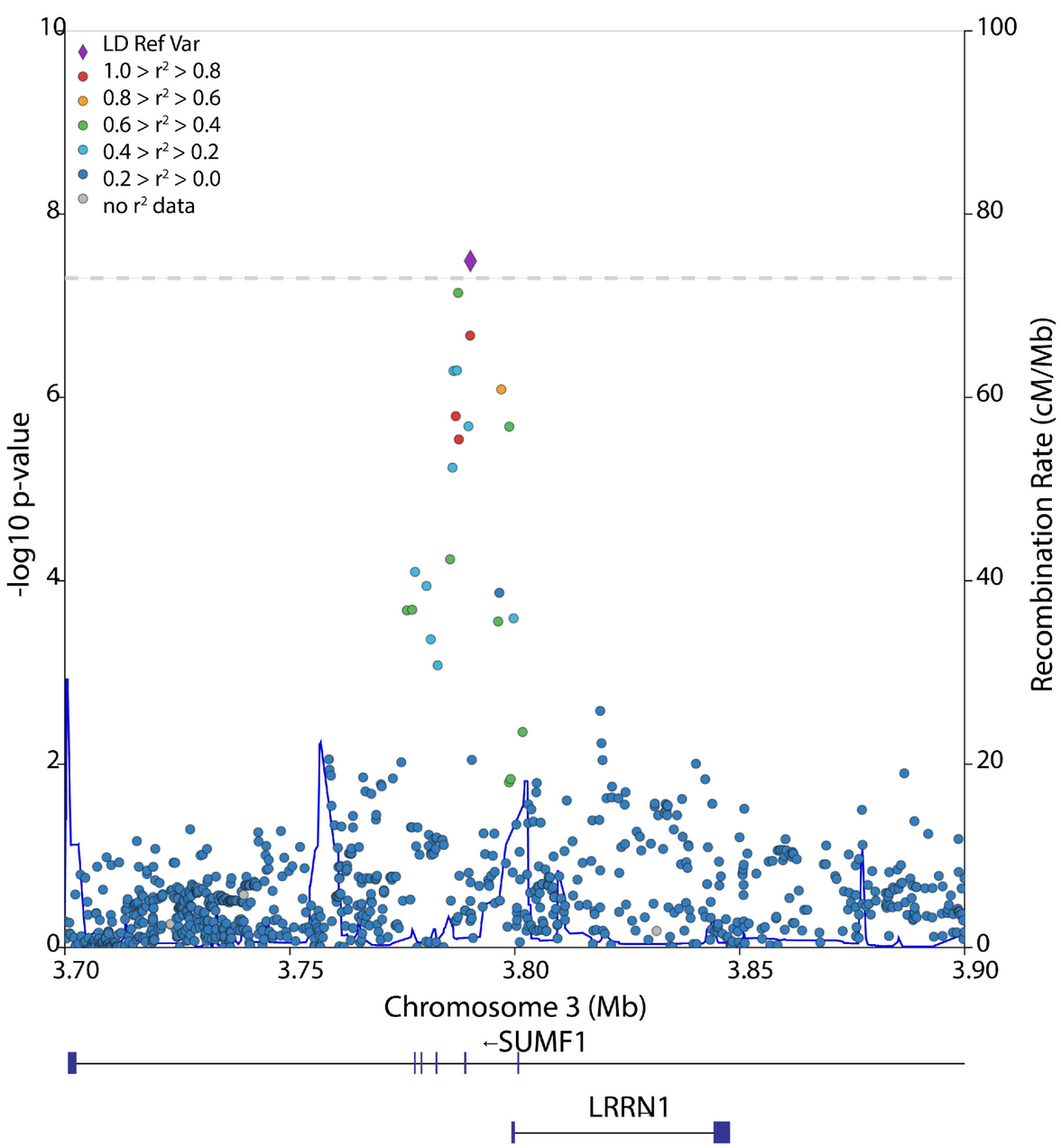
Common mind consortium: https://www.nimhgenetics.org/resources/commonmind

FUSION software http://gusevlab.org/projects/fusion/

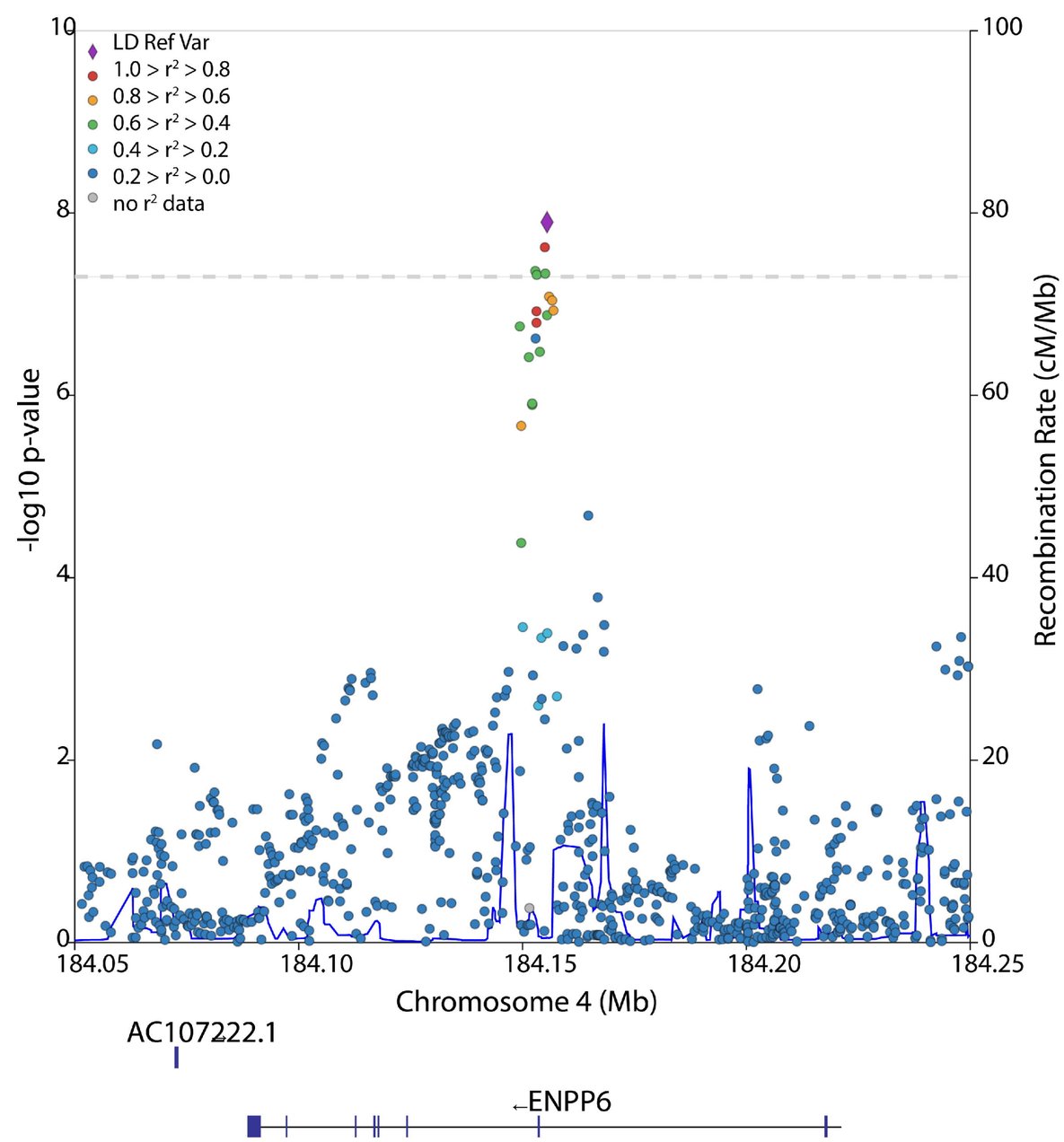
A


−log10(P)
9
8
7
6
5
4
3
2
1
3
5
7
9
11
13
15
17
19
21
X
Chromosome

B


LD Ref Var
1.0 > r² > 0.8
0.8 > r² > 0.6
0.6 > r² > 0.4
0.4 > r² > 0.2
0.2 > r² > 0.0
no r² data
-log10 p-value
Recombination Rate (cM/Mb)
3.70
3.75
3.80
3.85
3.90
Chromosome 3 (Mb)
←SUMF1
LRRN1

C


LD Ref Var
1.0 > r² > 0.8
0.8 > r² > 0.6
0.6 > r² > 0.4
0.4 > r² > 0.2
0.2 > r² > 0.0
no r² data
-log10 p-value
Recombination Rate (cM/Mb)
184.05
184.10
184.15
184.20
184.25
Chromosome 4 (Mb)
AC107222.1
←ENPP6