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,[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 Meccoci, MD,[x] Alessandro Serretti, MD,[y] Diana De Ronchi, MD,[y] Antonis Politis, MD,[z] Giulia Paroni, PhD,[u] 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; [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]Complex Structure of Geriatrics, Department of Medical Sciences, Fondazione IRCCS “Casa Sollievo della Sofferenza”, San Giovanni Rotondo (FG), 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]Institute of Psychiatry, 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 is more subtle. Some established risk SNPs for schizophrenia also 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 without the required amount of DNA were plated for whole genome amplification (WGA) and re-quantified. The above 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 additional 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 and an overview of other methods.

For more detail, see the Supplementary Material. Additional details for the NEXGENS cohort have also 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 5 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 5 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 four phenotypes. Specifically, we included intelligence, which is genetically correlated with years of schooling, 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 (Fig. 1, Supplementary Table 1). 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 1), 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 analyses (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 ($z = 5.15$, $p = 2.55\times10^{-7}$), which falls in an intron of *NECTIN2* (PVRL2 protein), the second smallest p-value occurred for rs429358 ($z = 5.12$, $p = 2.96\times10^{-7}$), 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. The odds increased significantly with count of ε4 haplotypes (OR = 1.21; 95% CI: 1.11-1.31; $p = 8.64\times10^{-6}$). 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 ($b=3.14$, $p = 1.90\times10^{-41}$) 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 ($p=1.17\times10^{-6}$, $q=0.011$), 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 ($p=6.44\times10^{-4}$ and $p=2.21\times10^{-3}$ 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 reduces 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=2.94×10-6) and FAM182B in GTEx7 cerebellum (p=5.51×10-6). 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 2), 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. Even when we removed the SNP representing the APOE locus, predictions remained positive. 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] Based on our analyses, we believe sample size explains the difference. For example, 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; likewise, when we computed two PRS from these two AD GWAS, the results were in qualitative agreement in their ability to distinguish AD+P and AD-P.

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}: 11 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 Figure S9). In addition, using the mixture model results, we identified a 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 located 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 2), 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 genetic correlation between risk for depressive symptoms and AD+P risk, 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, Bernie, Ilyas, Clive, Oscar
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, GP, 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. MAAS-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

Yushi Liu and Basavaraj Hooli are currently employed by and holding stock in Eli Lilly and Company

Dr. Ballard 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.

Oscar Lopez served as a consultant for Grifols, Inc.

Dr Saltvedt has been investigator in the drug trial Boehringer-Ingelheim 1346.0023

Ole A. Andreassen: Consultant to HEALTHLYTIX, speaker honoraria from Lundbeck.

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 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 dysequilibrium relative to the most significant SNP within the locus. Recombination rate through the region is shown on the right y axis. 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 3-10** 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

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	1.293E-01	3.304E-03
0.0001	785	0.961	0.921	1.003	6.525E-02	3.468E-03
0.001	3112	0.980	0.938	1.023	3.499E-01	3.091E-03
0.01	14748	0.959	0.918	1.002	6.404E-02	3.473E-03
0.1	79818	0.950	0.908	0.994	2.499E-02	3.711E-03
0.2	134250	0.955	0.913	0.999	4.466E-02	3.563E-03
0.3	182119	0.953	0.911	0.997	3.505E-02	3.624E-03
0.4	225630	0.957	0.915	1.001	5.550E-02	3.509E-03
0.5	265136	0.958	0.916	1.002	5.984E-02	3.490E-03

Schizophrenia

P Value Cut Off	N	OR	l95.OR	u95.OR	P Value	R^2
0.00000005	329	1.051	1.008	1.096	2.063E-02	8.018E-04
0.0001	3161	1.012	0.970	1.056	5.738E-01	4.733E-05
0.001	8488	1.009	0.967	1.052	6.910E-01	2.363E-05
0.01	26064	1.004	0.962	1.048	8.440E-01	5.795E-06
0.1	99775	1.006	0.963	1.052	7.824E-01	1.141E-05
0.2	153878	1.007	0.963	1.053	7.583E-01	1.416E-05
0.3	198913	1.009	0.965	1.055	6.985E-01	2.244E-05
0.4	238902	1.011	0.967	1.057	6.236E-01	3.602E-05
0.5	274264	1.013	0.969	1.059	5.680E-01	4.876E-05

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	9.750E-05	2.275E-03
0.0001	488	1.048	1.005	1.093	2.847E-02	7.182E-04
0.001	1962	1.059	1.015	1.105	7.760E-03	1.061E-03
0.01	12301	1.069	1.025	1.115	1.986E-03	1.432E-03
0.1	83722	1.095	1.050	1.142	2.301E-05	2.686E-03
0.2	144621	1.096	1.051	1.143	2.115E-05	2.710E-03
0.3	195758	1.096	1.051	1.143	2.073E-05	2.716E-03
0.4	239694	1.088	1.043	1.134	9.392E-05	2.285E-03
0.5	277645	1.088	1.043	1.135	8.795E-05	2.304E-03

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/~~ldhub~~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

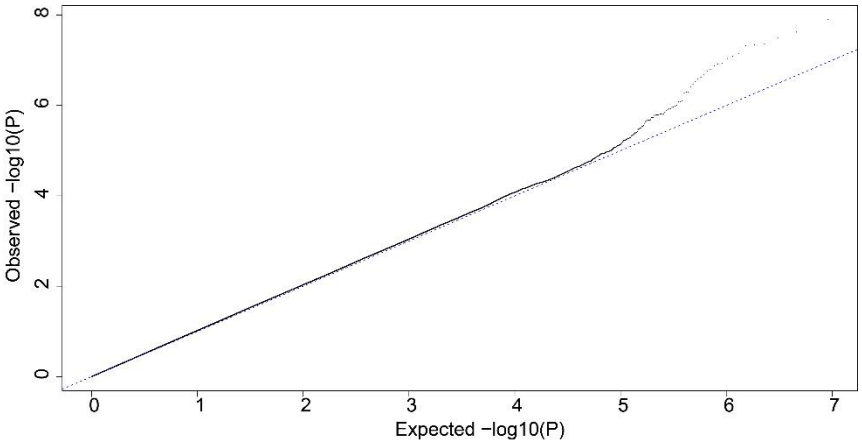
Common mind consortium: https://www.nimhgenetics.org/resources/commonmind

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

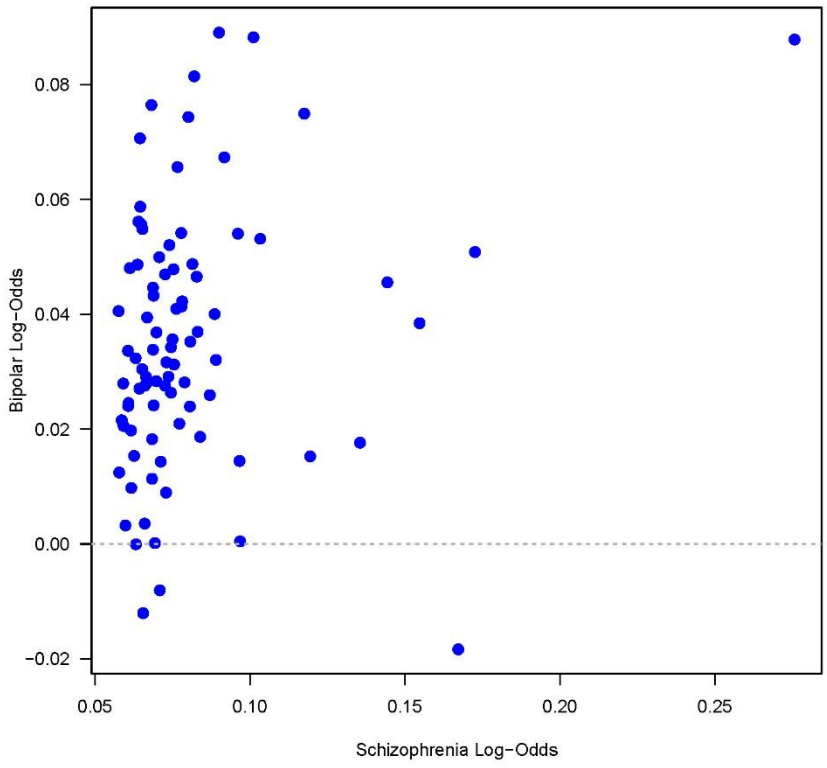
SUPPLEMENTARY RESULTS, METHODS, DISCUSSION, AND ACKNOWLEDGMENTS

TABLE OF CONTENTS

1. RESULTS- SUPPLEMENTARY FIGURES 1 & 2



Supplementary Figure 1. Observed versus expected -log$_{10}$ probabilities of statistics for genome-wide association with Alzheimer disease without versus with psychosis. Genomic control estimate,[1] GC=1.03, indicating there is no evidence for confounding by ancestry.



Supplementary Figure 2. Log of the odds ratios of 94 SNPs showing significant association with schizophrenia[2] versus their odds ratios for association with bipolar disorder.[3] Ninety-one of 94 SNPs had odds ratios exceeding one for both disorders, whereas 47 were expected under independence (sign test, $p = 5.8 \times 10^{-20}$).

2. METHODS- SUBJECT RECRUITMENT AND PSYCHOSIS CHARACTERIZATION

Overview. Subjects included in this GWA meta-analysis were provided from 8 different source programs: the Fundació ACE Barcelona Alzheimer Treatment and Research Center (ACE/GR@ACE), a consortium of National Institute on Aging Alzheimer Disease Centers (ADC), Eli Lilly and Company (LILLY), the Norwegian, Exeter and King's College Consortium for Genetics of Neuropsychiatric Symptoms in Dementia (NEXGENS), the National Institute on Aging's Late Onset Alzheimer's Disease Family Study (NIA-LOAD), the National Institute of Mental Health Genetics Initiative AD Cohort (NIMH), the University of Pittsburgh Alzheimer Disease Research Center (PITT ADRC), and the MRC Genetic Resource for Late-onset Alzheimer's disease as part of the Genetic and Environmental Risk in AD Consortium (UK-Cardiff). Collection of clinical data and genetic samples were approved by each site's local Institutional Review Board or Medical Ethics Committee, as appropriate. Methods used for selection of subjects from each program source are described in sections 2.1-2.8.

Phenotype data were provided to us by each institution. A total of 12,317 samples were used in this analysis, and their psychosis status is shown in **Supplementary Table S7**.

Supplementary Table S7. Psychosis Status of Subjects by Cohort.

Source	AD-P N (%)	AD+P N (%)	Total N (%)
ACE/GR@ACE	1,841 (67.8)	875 (32.2)	2,716 (22.1)
ADC	2,017 (67.8)	957 (32.2)	2,974 (24.1)
LILLY	1,360 (62.5)	815 (37.5)	2,175 (17.7)
NEXGENS	652 (49.6)	663 (50.4)	1,315 (10.7)
NIA-LOAD	168 (25.0)	505 (75.0)	673 (5.4)
NIMH	157 (26.5)	435 (73.5)	592 (4.8)
PITT ADRC	405 (31.7)	874 (68.3)	1,279 (10.4)
UK-Cardiff	272 (45.9)	321 (54.1)	593 (4.8)
TOTAL	**6,872 (55.7)**	**5,445 (44.3)**	**12,317 (100.0)**

AD-P: Alzheimer disease without psychosis; AD+P: Alzheimer disease with psychosis

2.1. Subject Recruitment and Psychosis Characterization – ACE/GR@ACE. Subjects in the ACE sample and the GR@ACE Stage I cohort were recruited from Fundació ACE Institut Català de Neurociències Aplicades, an Alzheimer's disease center serving the population of central Barcelona, Spain. All subjects received a structured neurological evaluation and cognitive testing as previously described.[4-6] They met criteria for a primary diagnosis of either possible or probable AD according to the National Institute on Aging and Alzheimer's Association's 2011 guidelines for defining AD,[7] as determined by a multidisciplinary working-group. Any subjects with a primary diagnosis of Dementia with Lewy bodies[8] were excluded.

Subjects were rated for psychotic symptoms on a validated Spanish language version of the Neuropsychiatric Inventory Questionnaire (NPI-Q[9]). Ratings of psychotic symptoms were performed at initial and follow up visits. The presence of psychosis at a visit was defined by either delusions or hallucinations of moderate or greater severity, or by both delusions and hallucinations of mild severity. Subjects were classified as AD+P if they were rated as having psychosis at any visit. For subjects without psychosis to be classified as AD-P, they had to have a score of 0 on the NPI-Q psychosis items at all visits and a last observed Mini Mental State Exam[10] score $\leq$ 20 or a CDR® Dementia Staging Instrument[11] score $\geq$ 1. See **Supplementary Table S8** for a summary of clinical and demographic data for these subjects.

2.2. Subject Recruitment and Psychosis Characterization – ADC. Recruitment for the ADC cohort has been described previously.[12-15] In brief, AD centers throughout the United States, each of which had received approval by their institutional review board, participated and provided data on subjects diagnosed with primary diagnoses of possible or probable AD.[7] Any subjects with a primary diagnosis of Dementia with Lewy bodies[8] were excluded. When available, subjects' autopsy data was also reviewed. Subjects were excluded if they did not have AD pathology, if AD pathology was rated as having a low likelihood of serving as the cause of dementia per NIA-Regan Institute Criteria,[16] or if the AD pathology was rated as having an intermediate likelihood of serving as the cause of dementia and other pathologies that could be the cause of dementia, such as Lewy Body or cerebrovascular disease, were present.

Phenotype data on ADC subjects were provided by the National Alzheimer's Coordinating Center (NACC) in three separate phenotype data freezes (March 2018, March 2019, and September 2019). These freezes included data for subjects' whose UDS visits were conducted between June 2005 and August 2019. Subjects were rated for psychotic symptoms on the NPI-Q[17] at initial and follow up visits. The presence of psychosis at a visit was defined by either delusions or hallucinations of moderate or greater severity, or by both delusions and hallucinations of mild severity. Subjects were classified as AD+P if they were rated as having psychosis at any visit. For subjects without psychosis to be classified as AD-P, they had to have scores of 0 on the NPI-Q psychosis items at all visits and either a last observed Mini Mental State Exam score $\leq$ 20 or a CDR® Dementia Staging Instrument score $\geq$ 1. See **Supplementary Table S9** for a summary of clinical and demographic data for these subjects.

2.3. Subject Recruitment and Psychosis Characterization – LILLY: Subjects participated in one of eight independent drug trials conducted by Eli Lilly and Company. Registration details of each of these trials is available via ClinicalTrials.gov:
LFAN: https://clinicaltrials.gov/ct2/show/NCT00594568
LFBC: https://clinicaltrials.gov/ct2/show/NCT00762411
LZAM: https://clinicaltrials.gov/ct2/show/NCT00905372
LZAN: https://clinicaltrials.gov/ct2/show/NCT00904683
LEAM: https://clinicaltrials.gov/ct2/show/NCT00051909
LEAQ: https://clinicaltrials.gov/ct2/show/NCT00843518
LYCG: https://clinicaltrials.gov/ct2/show/NCT00191009
Expedition 3 (LZAX): https://clinicaltrials.gov/ct2/show/NCT01900665.

Briefly, subjects were enrolled if they were diagnosed with probable AD.[18] Subjects were assessed for psychotic symptoms with the Neuropsychiatric Inventory (NPI[19]) The NPI was administered at timepoints and frequencies as specified for each of the included studies. The frequency and severity of each symptom were rated from 0-4 and 0-3, respectively. Frequency and severity scores at each visit were multiplied to give an overall domain score for each symptom ranging from 0-12. Individuals with a delusion or hallucination domain score $\geq$2 or with both delusion and hallucination domain scores equal to 1 at any visit were classified as AD+P. To be classified as AD-P, subjects had to have delusion and hallucination domain scores of 0 at all visits, and a last observed Mini Mental State Exam score $\leq$ 20 or a CDR® Dementia Staging Instrument score $\geq$ 1. See **Supplementary Table S10** for a summary of clinical and demographic data for these subjects.

2.4. Subject Recruitment and Psychosis Characterization –NEXGENS. NEXGENS is a collaboration between the following 5 institutions: University of Exeter, King's College London, Innlandet Hospital Trust, University of Oslo, and Stavanger University Hospital which undertakes secondary analysis of data from Alzheimer's disease patients from cohort studies and patient registries. Data for this study came from the following five program sources: **AddNeuroMed** study; Health and Memory Study in Nord-Trøndelag (**HMS**)[20]; Resource Use and Disease Couse in Dementia (**REDIC**)[21]; Norwegian registry of persons assessed for cognitive symptoms and the Progression of Alzheimer's Disease and Resource Use (**NorCog/PADR**)[22,23]; and from the following outpatient clinics in Italy and Greece (**Italian and Greek Hospital Outpatients)**: Centre for Neurodegenerative Disorders, University of Brescia; IRCCS Casa Sollievo della Sofferenza, Neuropsychiatry Clinic of the Eginition Hospital, Athens. Detail of each program source are in Sections 2.4.1 to 2.4.5.

Diagnosis of AD was performed according to ICD-10 etiological diagnosis, National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorders Association (**NINCDS-ADRDA**[18]) criteria, or clinical diagnosis by psychiatrist or geriatrician (See details of each institution described below). Individuals with a primary diagnosis of Dementia with Lewy Bodies were excluded. Any cases with a history of bipolar disorder or schizophrenia (except for the HMS study in which no information about psychiatric history was available) were excluded. Dementia severity was assessed in all cohorts by Mini Mental State Examination and the CDR® Dementia Staging Instrument. Psychotic symptoms were assessed by the NPI or the NPI-Q. For the NPI the presence of psychosis at a visit was defined by a calculated (frequency x severity) score for delusions or hallucinations $\geq$ 1 for either delusions or hallucinations. For the NPI-Q, the presence of psychosis at a visit was defined by either delusions or hallucinations of moderate or greater severity, or by both delusions and hallucinations of mild severity. For both NPI and NPI-Q, subjects were classified as AD+P if they were rated as having psychosis at any visit. Subjects were classified as AD-P if they had scores of 0 for delusions and hallucinations at all visits, and a last Mini Mental State Exam score $\leq$20 or a CDR® Dementia Staging Instrument score $\geq$ 1.

2.4.1 The Norwegian registry of persons assessed for cognitive symptoms (NorCog)/The Progression of Alzheimer's Disease and Resource use (PADR). NorCog is a national quality- and research registry collecting extensive clinical data and biobank material at outpatient clinics across Norway. Patients from NorCog included in the current study were recruited from memory clinics, geriatric-, and old-age psychiatric outpatient clinics in three of the four regional health authorities responsible for specialist healthcare in Norway. Participants underwent a comprehensive clinical assessment according to a standardized protocol including verbal interviews of the patient and proxy informant, neurocognitive testing, as well as a psychiatric, physical, and neurological examination. Diagnosis of AD was performed according to the ICD-10 criteria. The PADR study is affiliated to NorCog. PADR follows the same protocol as NorCog but it is a longitudinal observational study with assessments at the time of diagnostic workup (baseline) and follow-up after a mean of 24 months (range 16–37, 80% between 20 and 28 months). Inclusion criteria were MCI or dementia at baseline, living at home, able to give informed consent, and have a proxy informant available. For PADR, the NINCDS-ADRDA were used to diagnose AD. Diagnoses were assigned by the study researchers reviewing all available data from the baseline examination. Cerebrospinal fluid markers were used to support AD diagnosis. Psychosis in NorCog and PADR was assessed by NPI-Q. See **Supplementary Table S11.1** for clinical and demographic data.

2.4.2 Resource Use and Disease Couse in dementia - Nursing Home (REDIC-NH). REDIC-NH participants were recruited at admission to one of 47 nursing homes in 4 counties in Norway. The inclusion was from March 2012 to November 2014. Patients eligible for inclusion in the study were 65 years or older, or younger than 65 years with established dementia, with an expected stay in the NH of more than four weeks. The only exclusion criterion was a life expectancy of less than six weeks. The participants were assessed with the NPI-NH, in addition to several other assessment tools for depression, cognition, ADL function and physical disease. Assessments were carried out at baseline and every 6 months for maximum 2.5 years Based on all available information, AD according to ICD-10 was independently diagnosed by two old age psychiatrists with the possibility of consulting a third specialist, to reach a consensus. Patients were followed up with biannual assessments until death. See **Supplementary Table S11.2** for clinical and demographic data.

2.4.3 Health and Memory Study in Nord-Trøndelag (HMS). HMS participants were recruited from nursing homes in 24 municipalities in one county in middle Norway. Inclusion was between 2010 and 2011. Patients were eligible for inclusion if they had stayed in a nursing home for at least 14 days. Patients were diagnosed by the same method as in REDIC. Psychosis was assessed by NPI. No follow-up assessments were performed. See **Supplementary Table S11.3** for clinical and demographic data.

2.4.4 AddNeuroMed. The AddNeuroMed study involves six cross-European collection study sites; London (United Kingdom), Toulouse (France), Perugia (Italy), Kuopio (Finland), Lodz (Poland), and Thessaloniki (Greece) and has been described previously.[24] Participants were assessed once every three months for 1 year (5 assessments in total) on a battery of neuropsychological and psychiatric measures. Psychosis was

assessed with the NPI. Diagnosis of AD was made according to the NINCDS-ADRDA. See **Supplementary Table S11.4** for clinical and demographic data.

2.4.5 Italian & Greek Hospital Outpatients. For this sample, patient data and samples were collected through outpatient clinics at the Centre for Neurodegenerative Disorders, University of Brescia; IRCCS Casa Sollievo della Sofferenza, and the Neuropsychiatry Clinic of the Eginition Hospital, Athens. Diagnoses of AD mere made according to NINCDS-ADRDA criteria (IRCCS), ICD-10[25] criteria including MRI and CSF biomarkers (Brescia), or DSM-IV[26] criteria (Greece). Psychosis was assessed by NPI. See **Supplementary Table S11.5** for clinical and demographic data.

2.5. Subject Recruitment and Psychosis Characterization – NIA-LOAD. Recruitment for the family-based NIA-LOAD (National Institute of Aging Late Onset Alzheimer Disease) cohort has been described previously.[27,28] In brief, AD centers throughout the US participated, each of which had received approval by their institutional review board. The recruitment criteria included a family with multiple members affected with late onset AD that could provide clinical information and a biological sample for DNA extraction. The proband was required to have a diagnosis of probable[18] or definite[29] AD with onset after 60 years of age. At least one full sibling with definite, probable or possible AD and with onset after 60 years of age was also required. Finally, a third biologically related family member was required who was ≥50 years old if diagnosed as having AD or mild cognitive impairment or was ≥60 years old if unaffected. Family members were required to have had cognitive testing and clinical examination results documenting the classification as unaffected. To be included in the current study, individuals had to have a primary diagnosis of AD. Any subjects with a primary diagnosis of Dementia with Lewy bodies[8] were excluded.

NIA-LOAD subjects were rated for psychotic symptoms on the CERAD behavioral rating scale (CBRS[30]) and the NPI-Q as previously described.[31] Ratings of psychotic symptoms were performed at initial and follow up visits. The presence of psychosis as rated on the CBRS at any visit was defined as described below (see 2.7 **Subject Recruitment and Psychosis Characterization – PITT ADRC**). The presence of psychosis as rated on the NPI-Q at any visit was defined by either delusions or hallucinations of moderate or greater severity, or by both delusions and hallucinations of mild severity. Subjects were classified as AD+P if they were rated as having psychosis on either of the above scales at any visit. For subjects without psychosis to be classified as AD-P, they had to have scores of 0 on all psychosis items on both scales at all visits and a last observed CDR® Dementia Staging Instrument score $\underline{\geq}$ 1. See **Supplementary Table S12** for a summary of clinical and demographic data for these subjects.

2.6. Subject Recruitment and Psychosis Characterization – NIMH. The ascertainment and characterization of the family-based NIMH Genetics Initiative AD Cohort has been previously described.[32] All AD subjects met criteria for possible, probable,[18] or definite[29] AD. Psychotic symptoms were characterized at the time of initial evaluation, and again during follow-up evaluations, by responses to semi-structured interview questions. In a subset of subjects, this assessment was augmented by ratings on the Brief Psychiatric Rating Scale.[33] Subjects were classified as AD+P if they demonstrated either multiple delusions and/or hallucinations at any single evaluation or recurrent delusions and/or hallucinations over time, a phenotype associated with strong heritability.[34] Subjects without delusions or hallucinations at any time point, and with a last observed CDR® Dementia Staging Instrument score $\underline{\geq}$ 1, were classified as AD-P. See **Supplementary Table S13** for a summary of clinical and demographic data for these subjects.

2.7. Subject Recruitment and Psychosis Characterization – PITT ADRC. Subjects were evaluated at the University of Pittsburgh Alzheimer Disease Research Center (**PITT ADRC**), Pittsburgh, PA. All subjects were assessed at baseline with standardized neurological, neuropsychological, and psychiatric evaluations, cognitive testing, laboratory studies and brain imaging as previously described.[14,35-39] Repeat assessments were conducted annually. Subjects included in this analysis had a final clinical diagnosis of possible or probable AD.[18] Any subjects with a primary diagnosis of Dementia with Lewy bodies[8] were excluded. When available, subjects' autopsy data was also reviewed. Subjects were excluded if they did not have AD pathology, if AD pathology was rated as having a low likelihood of serving as the cause of dementia per NIA-Regan Institute Criteria,[16] or if the AD pathology was rated as having an intermediate likelihood of serving as

the cause of dementia and other pathologies that could be the cause of dementia, such as Lewy Body or cerebrovascular disease, were present.

Psychosis was evaluated with the CERAD behavioral rating scale (CBRS).[30] The CBRS was administered at initial and annual visits and in some subjects between annual visits by telephone as previously described.[27,36] Subjects were classified as AD+P if they had any hallucination or delusion symptom (CBRS item # 33-45) for 3 or more days in the previous month at any visit, as previously described.[14,35-39] To be classified as AD-P, AD subjects had to have scores of 0 for all CBRS items #33-45 and have a last observed Mini Mental State Exam score $\leq$20 or a CDR® Dementia Staging Instrument score $\geq$ 1.[14] See **Supplementary Table S14** for a summary of clinical and demographic data for these subjects.

2.8. Subject Recruitment and Psychosis Characterization – UK-Cardiff. The UK-Cardiff subjects were recruited at sites that participated in the the Medical Research Council (MRC) Genetic Resource for AD (Cardiff University; Kings College London; Cambridge University; Trinity College Dublin). Subject recruitment and clinical characterization was as previously described.[40,41] All AD cases met criteria for either probable[18] or definite[29] AD. Any subjects with a primary diagnosis of Dementia with Lewy bodies[8] were excluded.

The NPI was used to assess psychotic symptoms in all cases.[14,40] The frequency and severity of each symptom were rated from 0-4 and 0-3, respectively, and were scored to reflect the worst episode of each symptom over the lifetime of the illness. Frequency and severity scores were multiplied to give an overall domain score for each symptom ranging from 0-12. The NPI was administered at baseline and during follow up visits. Individuals with a delusion or hallucination domain score $\geq$2 or with both delusion and hallucination domain scores equal to 1 at any visit were classified as AD+P. To be classified as AD-P, subjects had to have delusion and hallucination domain scores of 0 at all visits, and a last observed Mini Mental State Exam score $\leq$ 20 or a CDR® Dementia Staging Instrument score $\geq$ 1. See **Supplementary Table S15** for a summary of clinical and demographic data for these subjects.

3. METHODS- INDIVIDUAL PROGRAM SOURCE SUBJECTS' DEMOGRAPHIC AND CLINICAL CHARACTERISTICS

Supplementary Table S8. ACE/GR@ACE Subject Characteristics

	AD-P N (%) or Mean (SD)	AD+P N (%) or Mean (SD)	Total N (%) or Mean (SD)
Sex	1,841 (67.8)	875 (32.2)	2,716 (100.0)
Male	584 (31.7)	186 (21.3)	770 (28.4)
Female	1,257 (68.3)	689 (78.7)	1,946 (71.6)
Age of Onset[1]	76.9 (7.7)	77.3 (7.3)	77.0 (7.6)
Age at last visit	82.6 (7.4)	83.7 (6.9)	82.9 (7.3)
Last MMSE[2]	14.4 (6.9)	10.9 (6.1)	13.3 (6.9)
Last CDR[3]			
1.0	180 (9.8)	20 (2.3)	200 (7.4)
2.0	183 (10.0)	69 (7.9)	252 (9.3)
3.0	353 (19.3)	103 (11.8)	456 (16.8)
4.0	729 (39.8)	363 (41.5)	1,092 (40.3)
5.0	388 (21.2)	319 (36.5)	707 (26.1)

ACE: Fundacio Alzheimer Centre Educacional/GR@ACE: Genome Research at Fundacio ACE; AD-P: Alzheimer disease without psychosis; AD+P: Alzheimer disease with psychosis; Age of Onset data not available for 37 subjects, 96 subjects were <60 (range 48-59); [2]MMSE: Mini mental status exam, data not available for 8 subjects; [3]CDR: CDR® Dementia Staging Instrument, data not available for 9 subjects.

Supplementary Table S9. ADC Subject Characteristics

	AD-P N (%) or Mean (SD)	AD+P N (%) or Mean (SD)	Total N (%) or Mean (SD)
Sex	2,017 (67.8)	957 (32.2)	2,974 (100.0)
Male	1,032 (51.2)	427 (44.6)	1,459 (49.1)
Female	985 (48.8)	530 (55.4)	1,515 (50.9)
Age of Onset[1]	72.4 (8.9)	70.9 (9.4)	71.9 (9.1)
Age at last visit	79.8 (8.9)	79.5 (9.0)	79.7 (9.0)
Last MMSE[2]	17.6 (6.9)	15.1 (7.5)	16.8 (7.2)
Last CDR			
0.5	62 (3.1)	43 (4.5)	105 (3.5)
1.0	980 (48.6)	189 (19.7)	1,169 (39.3)
2.0	618 (30.6)	318 (33.2)	936 (31.5)
3.0	357 (17.7)	407 (42.5)	764 (25.7)

ADC: Alzheimer Disease Centers; AD-P: Alzheimer disease without psychosis; AD+P: Alzheimer disease with psychosis; MMSE: Mini mental status exam; CDR: CDR® Dementia Staging Instrument. **[1]**Age of Onset not available for 10 subjects; 208 subjects were <60 (range 35-59); **[2]**Last MMSE not available for 176 subjects.

Supplementary Table S10. LILLY Subject Characteristics

	AD-P N (%) or Mean (SD)	AD+P N (%) or Mean (SD)	Total N (%) or Mean (SD)
Sex	1,360 (62.5)	815 (37.5)	2,175 (100.0)
Male	618 (45.4)	325 (39.9)	943 (43.4)
Female	742 (54.6)	490 (60.1)	1,232 (56.6)
Age of Onset[1]	71.1 (6.6)	70.8 (6.4)	71.0 (6.6)
Age at Last Visit[2]	76.9 (6.6)	76.8 (6.4)	76.9 (6.6)
Age at Consent	71.9 (6.7)	72.2 (6.5)	72.0 (6.6)
Last MMSE[3]	17.7 (4.8)	16.0 (6.1)	17.0 (5.4)
Last CDR[4]			
0.5	114 (9.8)	73 (10.6)	187 (10.1)
1.0	732 (63.2)	268 (38.8)	1,000 (54.1)
2.0	280 (24.2)	274 (39.7)	554 (30.0)
3.0	32 (2.8)	76 (11.0)	108 (5.8)

AD-P: Alzheimer disease without psychosis; AD+P: Alzheimer disease with psychosis; [1]The LILLY program source was comprised of 8 independent drug trials. Seven of the eight drug trials did not collect data for Age of Onset (N=1,281); for these trials Age at Consent, was used for screening purposes; [2]Age at Last Visit not available for 1,281 subjects; **[3]**MMSE: Mini mental state exam, data not available for 340 subjects; CDR: CDR® Dementia Staging Instrument; **[4]**CDR data not available for 326 subjects.

Supplementary Table S11.1. NEXGENS NorCog/PADR Subject Characteristics

	AD-P N (%) or Mean (SD)	AD+P N (%) or Mean (SD)	Total N (%) or Mean (SD)
Sex	165 (45.0)	202 (55.0)	367 (100.0)
Male	64 (38.8)	76 (37.6)	140 (38.1)
Female	101 (61.2)	126 (62.4)	227 (61.9)
Age at Consent[1]	76.4 (8.7)	77.3 (8.0)	76.9 (8.3)
Age at last visit	76.5 (8.7)	77.5 (8.0)	77.1 (8.3)
Last MMSE[2]	16.9 (3.2)	20.2 (4.7)	18.7 (4.4)
Last CDR[3]			
0.5	17 (12.1)	20 (11.8)	37 (11.9)
1.0	59 (41.8)	82 (48.2)	141 (45.3)
2.0	55 (39.0)	57 (33.5)	112 (36.0)
3.0	10 (7.1)	11 (6.5)	21 (6.8)

NEXGENS: Norwegian, Exeter and King's College Consortium for Genetics of Neuropsychiatric Symptoms in Dementia; NorCog: The Norwegian Registry of Persons Assessed for Cognitive Symptoms; PADR: The Progression of Alzheimer's Disease and Resource Use; AD-P: Alzheimer disease without psychosis; AD+P: Alzheimer disease with psychosis; [1]Unlike the other program sources, Age of Consent was collected for this study; MMSE: Mini mental state exam; [2]MMSE not available for 14 subjects CDR: CDR® Dementia Staging Instrument; [3]CDR not available for 56 subjects

Supplementary Table S11.2. NEXGENS REDIC-NH Subject Characteristics

	AD-P N (%) or Mean (SD)	AD+P N (%) or Mean (SD)	Total N (%) or Mean (SD)
Sex	111 (40.2)	165 (59.8)	276 (100)
Male	40 (36.0)	47 (28.5)	87 (31.5)
Female	71 (64.0)	118 (71.5)	189 (68.5)
Age at Consent[1]	86.4 (6.4)	84.4 (7.5)	85.2 (7.1)
Age at last visit	87.9 (6.7)	86.0 (7.6)	86.6 (7.4)
Last MMSE[2]	12.4 (5.3)	12.6 (7.0)	12.5 (6.3)
Last CDR[3]			
0.5	2 (1.8)	3 (1.8)	5 (1.8)
1.0	19 (17.3)	25 (15.2)	44 (16.1)
2.0	42 (38.2)	56 (34.1)	98 (35.8)
3.0	47 (42.7)	80 (48.8)	127 (46.3)

NEXGENS: Norwegian, Exeter and King's College Consortium for Genetics of Neuropsychiatric Symptoms in Dementia; REDIC-NH: Resource Use and Disease Couse in Dementia- Nursing Home; AD-P: Alzheimer disease without psychosis; AD+P: Alzheimer disease with psychosis; [1]Unlike the other program sources, Age of Consent was collected for this study; MMSE: Mini mental state exam; [2]MMSE not available for 55 subjects; CDR: CDR® Dementia Staging Instrument;[3]CDR not available for 2 subjects

Supplementary Table S11.3. NEXGENS HMS Subject Characteristics

	AD-P N (%) or Mean (SD)	AD+P N (%) or Mean (SD)	Total N (%) or Mean (SD)
Sex	91 (56.2)	71 (43.8)	162 (100.0)
Male	20 (22.0)	19 (26.8)	39 (24.1)
Female	71 (78.0)	52 (73.2)	123 (75.9)
Age at Consent[1]	86.4 (6.0)	86.7 (6.4)	86.5 (6.2)
Age at last visit	86.4 (6.0)	86.7 (6.4)	86.5 (6.2)
Last MMSE	12.4 (6.2)	12.4 (6.0)	12.4 (6.1)
Last CDR[2]			
0.5	5 (5.5)	1 (1.4)	6 (3.7)
1.0	16 (17.8)	12 (16.9)	28 (17.4)
2.0	37 (41.1)	22 (31.0)	59 (36.6)
3.0	32 (35.5)	36 (50.7)	68 (42.2)

NEXGENS: Norwegian, Exeter and King's College Consortium for Genetics of Neuropsychiatric Symptoms in Dementia; AD-P: Alzheimer disease without psychosis; AD+P: Alzheimer disease with psychosis; [1]Unlike the other program sources, Age of Consent was collected for this study; MMSE: Mini mental state exam; CDR: CDR® Dementia Staging Instrument; [2]CDR not available for 1 subject.

Supplementary Table S11.4. NEXGENS AddNeuroMed Subject Characteristics

	AD-P N (%) or Mean (SD)	AD+P N (%) or Mean (SD)	Total N (%) or Mean (SD)
Sex	65 (41.7)	91 (58.3)	156 (100)
Male	24 (36.9)	21 (23.1)	45 (28.8)
Female	41 (63.1)	70 (76.9)	111 (71.2)
Age at Consent[1]	77.0 (7.6)	77.9 (5.8)	77.5 (6.6)
Age at last visit	77.9 (7.5)	78.7 (5.8)	78.3 (6.6)
Last MMSE	15.4 (4.1)	17.6 (6.0)	16.7 (5.4)
Last CDR[2]			
0.5	0 (0)	5 (5.5)	5 (3.3)
1.0	30 (47.6)	34 (37.8)	64 (41.8)
2.0	32 (50.8)	41 (45.6)	73 (47.7)
3.0	1 (1.6)	10 (11.1)	11 (7.2)

NEXGENS: Norwegian, Exeter and King's College Consortium for Genetics of Neuropsychiatric Symptoms in Dementia; AD-P: Alzheimer disease without psychosis; AD+P: Alzheimer disease with psychosis; [1]Unlike the other program sources, Age of Consent was collected for this study; MMSE: Mini mental state exam; CDR: CDR® Dementia Staging Instrument; [2]CDR not available for 3 subjects.

Supplementary Table S11.5. NEXGENS Italian and Greece Hospital Outpatient Subject Characteristics

	AD-P N (%) or Mean (SD)	AD+P N (%) or Mean (SD)	Total N (%) or Mean (SD)
Sex	220 (62.1)	134 (37.9)	354 (100.0)
Male	87 (39.5)	54 (40.3)	141 (39.8)
Female	133 (60.5)	80 (59.7)	213 (60.2)
Age at Consent[1]	77.8 (7.7)	79.0 (6.7)	78.3 (7.4)
Age at last visit	77.8 (7.7)	79.0 (6.7)	78.3 (7.4)
Last MMSE[2]	12.6 (5.3)	11.0 (6.2)	12.0 (5.9)
Last CDR[3]			
0.0	2 (1.0)	2 (1.7)	4 (1.3)
0.5	9 (4.6)	1 (0.9)	10 (3.2)
1.0	72 (36.9)	20 (17.4)	92 (29.7)
2.0	84 (43.1)	52 (45.2)	136 (43.9)
3.0	28 (14.4)	40 (34.8)	68 (21.9)

NEXGENS: Norwegian, Exeter and King's College Consortium for Genetics of Neuropsychiatric Symptoms in Dementia; AD-P: Alzheimer disease without psychosis; AD+P: Alzheimer disease with psychosis; [1]Unlike the other program sources, Age of Consent was collected for this study; MMSE: Mini mental state exam; [2]MMSE not available for 2 subjects; CDR: CDR® Dementia Staging Instrument; ;[3]CDR not available for 44 subjects.

Supplementary Table S12. NIA-LOAD Subject Characteristics

	AD-P N (%) or Mean (SD)	AD+P N (%) or Mean (SD)	Total N (%) or Mean (SD)
Sex	168 (25.0)	505 (75.0)	673 (100.0)
Male	78 (46.4)	171 (33.9)	249 (37.0)
Female	90 (53.6)	334 (66.1)	424 (63.0)
Age of Onset[1]	74.3 (8.0)	73.6 (7.4)	73.8 (7.6)
Age at last visit[2]	82.6 (7.6)	84.9 (8.3)	84.4 (8.2)
Last CDR[3]			
0.0	0 (0.0)	3 (0.8)	3 (0.5)
0.5	0 (0.0)	17 (4.4)	17 (3.1)
1.0	66 (39.3)	61 (15.8)	127 (22.9)
2.0	39 (23.2)	88 (22.8)	127 (22.9)
3.0	63 (37.5)	216 (56.0)	279 (50.4)
4.0	0 (0.0)	1 (0.3)	1 (0.2)

NIA-LOAD: National Institute of Aging Late-onset Alzheimer disease; AD-P: Alzheimer disease without psychosis; AD+P: Alzheimer disease with psychosis; CDR: CDR® Dementia Staging Instrument. [1]AOO not available for 10 subjects, <60 for 16 subjects (range=48-58). [2]Age at last visit not available for 4 subjects. [3]CDR not available for 119 subjects.

Supplementary Table S13. NIMH Subject Characteristics

	AD-P N (%) or Mean (SD)	AD+P N (%) or Mean (SD)	Total N (%) or Mean (SD)
Sex	157 (26.5)	435 (73.5)	592 (100.0)
Male	54 (34.4)	105 (24.1)	159 (26.9)
Female	103 (65.6)	330 (75.9)	433 (73.1)
Age of Onset[1]	71.3 (8.8)	72.0 (7.7)	71.8 (8.0)
Age at Last Visit	79.7 (9.2)	81.0 (7.1)	80.6 (7.7)
Last CDR			
1.0	37 (23.6)	41 (9.4)	78 (13.2)
2.0	38 (24.2)	97 (22.3)	135 (22.8)
3.0	32 (20.4)	156 (35.9)	188 (31.8)
4.0	28 (17.8)	97 (22.3)	125 (21.1)
5.0	22 (14.0)	44 (10.1)	66 (11.1)

NIMH: National Institute of Mental Health; AD-P: Alzheimer disease without psychosis; AD+P: Alzheimer disease with psychosis; CDR: CDR® Dementia Staging Instrument. [1]Age of Onset<60 for 37 subjects (range=41-59)

Supplementary Table S14. PITT ADRC Subject Characteristics

	AD-P N (%) or Mean (SD)	AD+P N (%) or Mean (SD)	Total N (%) or Mean (SD)
Sex	405 (31.7)	874 (68.3)	1,279 (100.0)
Male	166 (41.0)	288 (33.0)	454 (35.5)
Female	239 (59.0)	586 (67.0)	825 (64.5)
Age of Onset	72.7 (6.5)	73.1 (6.4)	72.9 (6.4)
Age at last visit	78.4 (6.5)	80.4 (6.3)	79.8 (6.4)
Last MMSE[1]	17.0 (5.6)	13.2 (6.1)	14.4 (6.2)
Last CDR[2]			
0.5	34 (8.4)	21 (2.4)	55 (4.3)
1.0	225 (55.6)	198 (22.8)	423 (33.2)
2.0	123 (30.4)	459 (52.8)	582 (45.6)
3.0	21 (5.2)	171 (19.7)	192 (15.1)
4.0	2 (0.5)	21 (2.4)	23 (1.8)

PITT ADRC: University of Pittsburgh Alzheimer Disease Research Center; AD-P: Alzheimer disease without psychosis; AD+P: Alzheimer disease with psychosis; [1]MMSE: Mini mental state exam, data not available for 21 subjects; [2]CDR: CDR® Dementia Staging Instrument, data not available for 4 subjects.

Supplementary Table S15. UK-Cardiff Subject Characteristics

	AD-P N (%) or Mean (SD)	AD+P N (%) or Mean (SD)	Total N (%) or Mean (SD)
Sex	272 (45.9)	321 (54.1)	593 (100.0)
Male	97 (35.7)	77 (24.0)	174 (29.3)
Female	175 (64.3)	244 (76.0)	419 (70.7)
Age of Onset[1]	76.1 (6.5)	76.0 (7.2)	76.0 (6.9)
Age at last visit[2]	81.7 (6.7)	83.1 (6.6)	82.4 (6.7)
Last MMSE[3]	11.9 (7.4)	9.5 (8.0)	10.7 (7.8)
Last CDR[4]			
0.5	1 (1.1)	2 (1.9)	3 (1.5)
1.0	53 (57.0)	17 (16.0)	70 (35.2)
2.0	29 (31.2)	53 (50.0)	82 (41.2)
3.0	10 (10.8)	34 (32.1)	44 (22.1)

UK-Cardiff: MRC Genetic Resource for Late-onset Alzheimer's disease as part of the Genetic and Environmental Risk in AD Consortium AD-P: Alzheimer disease without psychosis; AD+P: Alzheimer disease with psychosis; MMSE: Mini mental state exam; CDR: CDR® Dementia Staging Instrument; [1]Age of Onset not

available for 1 subject; [2]Age at last visit not available for 8 subjects; [3]MMSE not available for 38 subjects; [4]CDR not available for 394 subjects because it was instituted later in the data collecting process.

4. METHODS- GENOTYPING

Overview. For subjects selected for this study based on the above described phenotypic criteria, program sources provided us with blood samples (ACE/GR@ACE), DNA samples (PITT ADRC, UK-Cardiff, NIA-LOAD, ADC, NIMH), and/or single nucleotide polymorphism array data (ACE/GR@ACE, ADC, LILLY, NIA-LOAD). Additionally, NEXGENS provided summary GWA statistics for the comparison of AD-P and AD+P. Blood and DNA samples that were supplied to us were genotyped at the Center for Applied Genomics at The Children's Hospital of Philadelphia (CHoP, Philadelphia, PA). These samples were initially 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 the DNA to 23ng/ul. Samples without the required amount of DNA were plated for whole genome amplification (WGA) and re-quantified. Prior to genotyping, CHoP confirmed DNA concentrations by Pico Green assay and performed additional WGA on samples when necessary. Supplementary Table S11 contains information of each program sources, the source of their genotypes, and the genotype arrays used by each source.

Supplementary Table S16. Program Sources, Genotype Sources, and Chip Information

Program Sources (N)	Genotype Source	Genotype Array
ACE/GR@ACE (2,716)		
• ACE	CHoP	Illumina Infinium Global Screening Array-24 v1.0
• GR@ACE	ACE	Axiom 815K Spanish Biobank Array
ADC (2,974)	ADGC wave 1-2	Illumina Human660W-Quad v1_A
	ADGC wave 3	IlluminaHumanOmniExpress-12v1_A
	ADGC wave 4-6	IlluminaHumanOmniExpress-12v1_H
	ADGC wave 7	Illumina Human OmniExpressExome-8v1.2_a
	ADGC wave 8	Illumina Human OmniExpressExome-8v1.2_a
	ADGC wave 9	Illumina Infinium Global Screening Array-24 v1.0
	ADGC wave 10	Illumina Infinium Global Screening Array-24 v1.0
LILLY (2,175)		
• LFAN, LFBC, LZAM, and LZAN trials	Lilly	Illumina 5M Genotyping Array v1.2
• LEAM, LEAQ, and LYCG trials	Lilly	Illumina 5M Genotyping Array v1.2
• Expedition 3 LZAX trial	Lilly	Illumina Infinium Global Screening Array v2.0
NEXGENS (1,315)		
• NorCog/PADR, REDIC-NH, and HMS	NEXGENS	Illumina Human OmniExpress 24 v1.0
• AddNeuroMed	NEXGENS	Illumina 610K
• Italian/Greek	NEXGENS	Illumina Global Screening Array 24 v1.0
NIA-LOAD (673)	CHoP	Illumina Infinium Global Screening Array-24 v1.0
	NIA-LOAD	Illumina OmniExpress Illumina Human610-quad v1_B Illumina 660K 700 Illumina OmniExpress Illumina 1M Illumina OmniExpress-24 v.1 Illumina Global Screening Array
NIMH (592)	CHoP	Infinium Global Screening Array-24 v1.0
PITT ADRC (1,279)	CHoP	Illumina Infinium Global Screening Array-24 v1.0
	ADGC	Illumina Human1M-Duo v3.0
	PITT ADRC Neurogenetics Core	Illumina HumanOmni1-Quad v1.1
UK-Cardiff (593)	CHoP	Infinium Global Screening Array-24 v1.0

ACE: Fundacio Alzheimer Centre Educacional/GR@ACE: Genome Research at Fundacio ACE; ADC: Alzheimer Disease Centers; NEXGENS: Norwegian, Exeter and King's College Consortium for Genetics of Neuropsychiatric Symptoms in Dementia; NIA-LOAD: National Institute of Aging Late-onset Alzheimer disease;

NIMH: National Institute of Mental Health; PITT ADRC: University of Pittsburgh Alzheimer Disease Research Center; UK-Cardiff: MRC Genetic Resource for Late-onset Alzheimer's disease as part of the Genetic and Environmental Risk in AD Consortium; CHoP: Center for Applied Genomics at The Children's Hospital of Philadelphia; ADGC: Alzheimer's Disease Genetics Consortium.

4.1. Genotyping – ACE/GR@ACE. ACE supplied us with whole blood. DNA was prepared by the Genomics Core Lab at the University of Pittsburgh and genotypes were generated by CHoP as described in the Overview, above. Other subjects were genotyped as part of the GR@ACE genomic data collection for Alzheimer's disease.[6] GR@ACE genotypes were generated using the Axiom 815K Spanish Biobank Array (Thermo Fisher). Further information on the quality control for genome-wide data has been previously described.[6]

4.2. Genotyping – ADC. ADCs submitted blood samples to the National Centralized Repository for Alzheimer's Disease and Related Dementias (**NCRAD**). As part of their participation in the Alzheimer's Disease Genetics Consortium (**ADGC**), DNA was prepared by NCRAD for genotyping and sent to the genotyping site at CHoP as previously described.[15] ADGC then provided us with genotypic data on selected subjects[15] In addition, NCRAD supplied us with DNA for subjects for whom ADGC did not have genotypic data. The DNA was then prepared by the Genomics Core Lab at the University of Pittsburgh and submitted to CHoP for genotyping as described in the Overview, above.

4.3. Genotyping – LILLY. DNA was first extracted from whole blood samples and DNA concentrations were normalized to 50ng/µl using Pico green assay to measure concentrations, followed by suspension in 10mM Tris-Cl at pH 8.5. Samples from LZAX were genotyped on the Illumina Global Screening array version 2.0 with an addition of multi-disease associated SNP panel (n~700k SNP probes in total). The rest of the trial samples were genotyped using Illumina 5M Infinium® Omni5-4 v1.2 genotyping platform as per the manufacturers recommended guidelines (n~4.5M SNP probes). Three control DNA samples from HAPMAP consortium were used on each 96 well plate to quantify genotyping assay quality across plates and batches. In addition, 5% of the total samples were used as replicates to quantify replicability across all the batches. The genotyping data went through a stringent QC pipeline, as described in[42] using SNPRelate[43] and GWASTools[44]available on Bioconductor. Genotyping QC pipeline included a 19-step process of eliminating samples and probes with spurious quality. At probe level, in cases of duplicate probes (for e.g. APOE-e4 rs7412), only the probe with the best quality was retained. SNP probes were filtered based on probe level missing rates (<95%) and in addition SNPs with heterozygosity rate outliers were excluded from the dataset (using sample call rate vs heterozygosity plot). In addition, discordant SNPs in >=3 duplicated samples were also eliminated. At the samples level, only the higher quality data from duplicated samples were retained, samples with low call rates (< 95%) and those showing conflicts between reported gender and chromosomal sex were also eliminated. In addition, sample pairs with high identity-by-state (IBS) scores were eliminated since the trial participants are unrelated.

4.4. Genotyping – NIA-LOAD. DNA was prepared by NCRAD for genotyping and sent to the genotyping site at ChoP as part of participation in the ADGC, as previously described.[15] For these subjects we utilized existing genotype data provided by NIA-LOAD. NCRAD also supplied us with DNA for subjects for whom genotypic data was unavailable from NIA-LOAD. The DNA was prepared by the Genomics Core Lab at the University of Pittsburgh and submitted to CHoP for genotyping as described in the Overview, above.

4.5. Genotyping – NIMH. NIMH provided us with DNA which was prepared by the Genomics Core Lab at the University of Pittsburgh and submitted to CHoP for genotyping as described in the Overview, above.

4.6. Genotyping – PITT ADRC. DNA for subjects for whom genotypic data was unavailable was provided by the Neurogenetics Core of the PITT ADRC or by NCRAD. The DNA was submitted to CHoP for genotyping as described in the Overview, above. For other subjects, available genotypic data was generated as previously described and provided by the Neurogenetics Core of the PITT ADRC and the ADGC.[45,46]

4.7. Genotyping –UK-Cardiff. The UK-Cardiff supplied us with DNA which was then prepared and submitted to CHoP for genotyping as described in the Overview, above.

4.8 Genotyping – NEXGENS. NEXGENS provided summary statistics for tests of AD+P versus AD-P. NEXGENS samples were genotyped on Illumina platforms, using either the OmniExpress array (NorCog/PADR, REDIC-NH, and HMS), the Illumina 610K (AddNeuroMED), or the Illumina Global Screening Array 24 v1.0 array (Italian and Greece Hospital Outpatient).

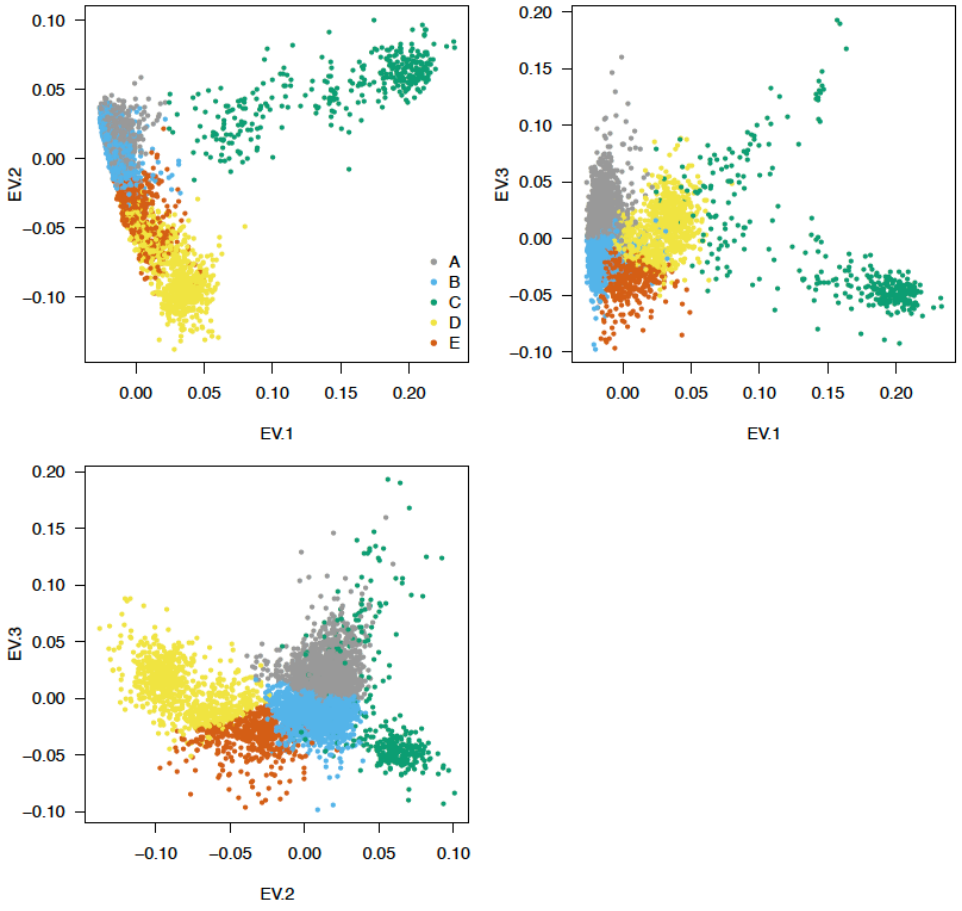
5. METHODS- QC AND ANCESTRY

5.1 QUALITY CONTROL: For QC by genotype, SNPs were removed if they had an unknown location, were monomorphic, were duplicated, did not map to an autosome or X-linked chromosome, had a non-call rate > 0.025, had a minor allele frequency MAF < 0.01, or had an exact Hardy-Weinberg p-value< 0.005 in the major European ancestry group, as defined below. For QC by sample, samples were removed if their overall SNP non-call rates > 0.025, if they showed extreme homo- or heterozygosity (<-.15 or > 0.4, respectively), if their nominal versus inferred genetic sex did not match, or if they were determined to be duplicates of other samples within or between cohorts. Raw genotype data for individual NEXGEN cohorts underwent appropriate QC steps (implemented in PLINK).[47]

5.2 ANCESTRY: Here we document batch-specific sources (Phase 1, Phase 2, GR@ACE, and NEXGENS) of samples and attributes of genetic ancestry of those samples.

5.2.1. PHASE 1

Subjects analyzed in Phase 1 were from ACE, ADC, Lilly, NIA-LOAD, NIMH, PITT ADRC and UK-Cardiff, and their genotypes were from ADGC waves 1-7, CHoP, Lilly, Pitt ADRC, and NIA-LOAD (**Supplementary Table S16**).



Supplementary Figure S3. Ancestry plot for the PHASE 1 samples. Analysis of genetic ancestry using GemTools identified three significant ancestry dimensions and divided the data up into 5 ancestry clusters (see also **Supplementary Table S17**), of which 5,396 samples in clusters A and B were of narrow European ancestry on the basis of self-identified ancestry of subjects.

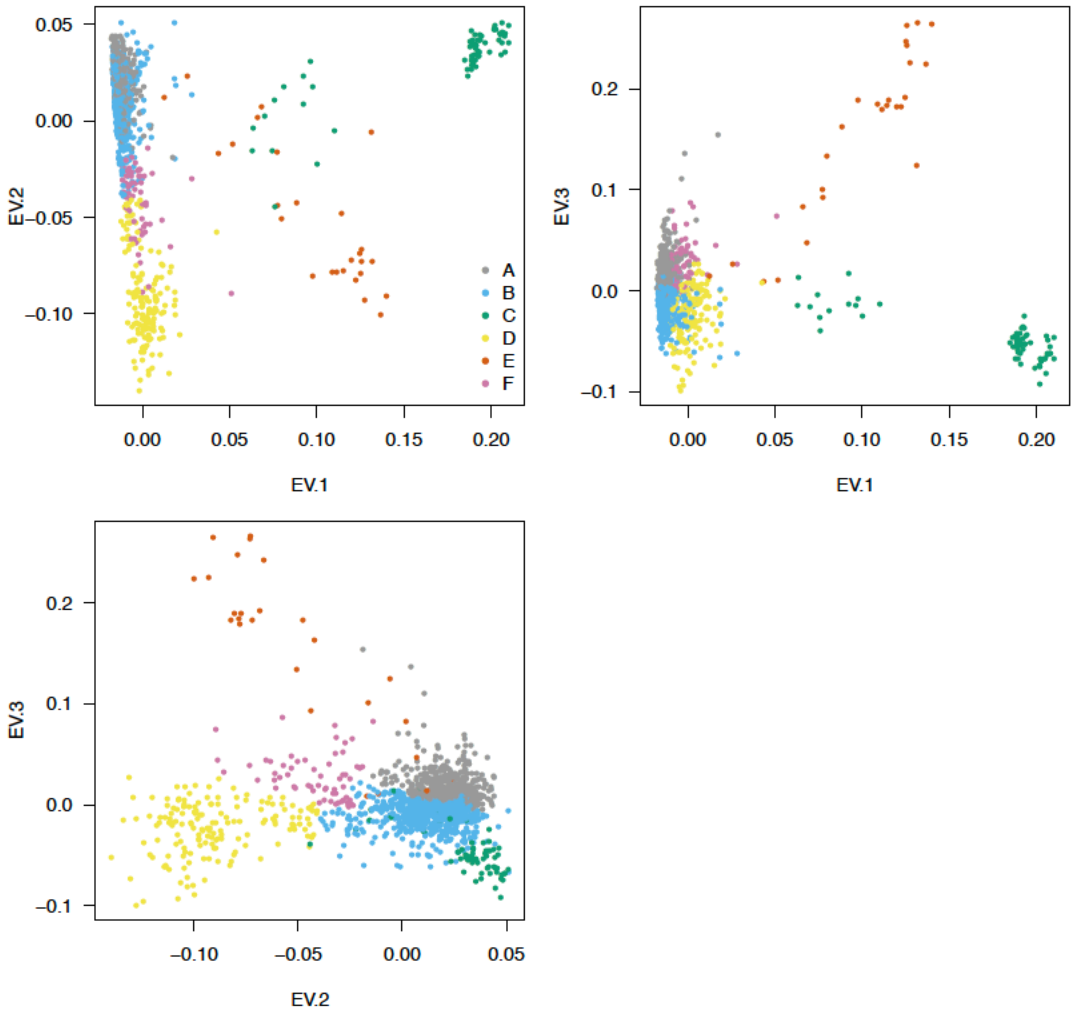
Supplementary Table S17. Distribution of PHASE 1 samples over the ancestry clusters.

	Cluster[1]						
	A	**B**	**C**	**D**	**E**	**All**	**EUR**
Psychosis Status	2,382	3,014	354	809	495	7,054	5,396
AD-P	1,152	1512	139	458	268	3,529	2,664
AD+P	1,230	1,502	215	351	227	3,525	2,732
Program Source (Genotype Source)							
ACE/GR@ACE (CHoP)	2	66	2	20	218	308	68
ADC (ADGC waves 1-7 and CHoP)	792	1,047	132	320	37	2,328	1,839
PITT ADRC (ADGC, PITT ADRC Neurogenetics Core, and CHoP)	472	436	103	191	77	1,279	908
LILLY LEAM, LEAQ, LFAN, LFBC, LYCG, LZAM, and LZAN drug trials (LILLY)	439	526	41	153	122	1,281	965
NIA-LOAD (CHoP and NIA-LOAD)	271	305	50	33	14	673	576
NIMH (CHoP)	190	264	24	90	24	592	454
UK-Cardiff (CHoP)	216	370	2	2	3	593	586

AD-P: Alzheimer disease without psychosis; AD+P: Alzheimer disease with psychosis; AA: African American;
[1] Of the subjects who self-reported ancestry: Cluster A (0.2% AA, 99.8% EUR [white]); Cluster B (0.2% AA, 99.8% EUR); Cluster C (92.2% AA, 7.8% EUR); Cluster D (0.5% AA, 99.5% EUR); Cluster E (0% AA, 100% EUR). Other self-reported ancestries represented less than 0.1% of subjects and hence were not reported here.

5.2.2. PHASE 2

Subjects analyzed in Phase 2 were from ADC and LILLY, and their corresponding genotypes were from ADCG waves 8-10 and LILLY (**Table S16**).



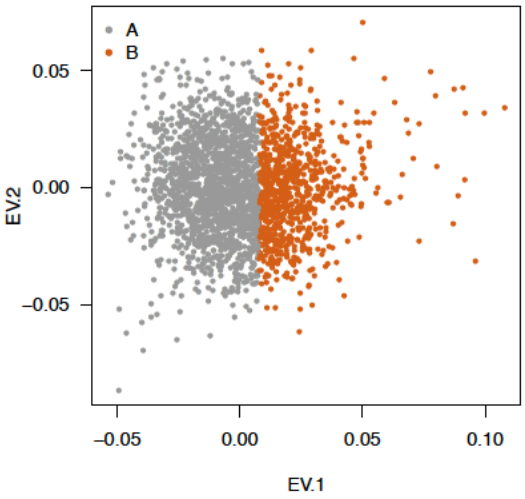
Supplementary Figure S4. Ancestry plot for the PHASE 2 samples. Analysis of genetic ancestry using GemTools identified three significant ancestry dimensions and divided the data up into 6 ancestry clusters (see also **Supplementary Table S18**), of which 1,227 samples in clusters A and B are of narrow European ancestry, as determined by genetically-based clustering with PHASE 1 samples (99% of PHASE 2 clusters A and B group together with clusters A and B of PHASE 1).

Supplementary Table S18. Distribution of PHASE 2 samples over the ancestry clusters

	Cluster							
	A	**B**	**C**	**D**	**E**	**F**	**All**	**EUR**
Psychosis Status	461	766	65	161	26	61	1,540	1,227
AD-P	310	523	42	113	15	42	1,045	833
AD+P	151	243	23	48	11	19	495	394
Program Source (Genotype Source)								
ADC (ADGC waves 8-10)	206	337	1	92	0	10	646	543
LILLY (Expedition 3 LZAX)	255	429	64	69	26	51	894	684

AD-P: Alzheimer disease without psychosis; AD+P: Alzheimer disease with psychosis

5.2.3. GR@ACE



Supplementary Figure S5. Ancestry plot for the GR@ACE samples. Analysis of GR@ACE samples' genetic ancestry using GemTools identified one significant ancestry dimension and divided the data up into 2 ancestry clusters (see also **Supplementary Table S19**). All samples are of European ancestry, as determined by genetically-based clustering with PHASE 1 and PHASE 2 samples (99.4% of GR@ACE cluster with others of European ancestry, specifically clusters A and B of PHASE 1 and PHASE 2).

Supplementary Table S19. Distribution of GR@ACE samples over the ancestry clusters.

	Cluster			
	A	**B**	**All**	**EUR**
Total	1623	785	2408	2408
AD-P	1098	548	1646	1646
AD+P	525	237	762	762
GR@ACE	1623	785	2408	2408

AD-P: Alzheimer disease without psychosis; AD+P: Alzheimer disease with psychosis

5.2.4. NEXGENS

For our meta-analysis, we received summary statistics from the NEXGENS consortium. What follows is a synopsis of the methods used by the NEXGENS consortium to produce these association summary statistics.

Raw genotype data for individual NEXGENS cohorts underwent appropriate QC steps (implemented in PLINK), as described previously.[47] SNPs with a minor allele frequency ≤5% and a Hardy Weinberg equilibrium $p < 10-5$ were excluded. The SNP and individual genotype failure threshold were set at 5%, and individuals with mean heterozygosity ±3 standard deviations were excluded. The analysis was restricted to individuals of European ancestry using genetic principal components computed by EIGENSTRAT. Related (pi-hat > 0.2) or duplicate individuals both within and between cohorts were excluded, ensuring there was no sample overlap across datasets. Phasing (EAGLE2) and imputation (PBWT) 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. Logistic regression, implemented in PLINK on each dataset individually, was used for SNP association testing with psychosis status, covarying for the first 10 ancestry principal components. The METAL software was used to conduct inverse-variance weighted fixed effects meta-analysis across the 5 datasets within the NEXGENS Consortium (**Tables S11.1-S11.5**), applying genomic control.[48]

6. METHODS- ANALYSIS

6.1. PATHWAY ANALYSIS: Gene set enrichment analyses were performed in MAGMA,[49] 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).[50] As in the gene-based analyses, we used the "mean" test statistic, and again the primary analysis assigned variants to genes if they lie within the gene boundaries, but a secondary analysis used a window of 35kb upstream and 10kb downstream to assign SNPs to genes, as in Kunkle et al.[15] We used the R package qvalue to account for multiple testing.

Pathways were defined as follows. The assignment of Gene Ontology (GO) terms to human genes was obtained from the "gene2go" file, downloaded from NCBI on March 11th 2020. "Parent" GO terms were assigned to genes using the ontology file downloaded from the Gene Ontology website on the same date. GO terms were assigned to genes based on experimental or curated evidence of a specific type, so evidence codes IEA (electronic annotation), NAS (non-traceable author statement), RCA (inferred from reviewed computational analysis) were excluded. Pathways were downloaded from the Reactome website on April 26th 2020. Biocarta, KEGG and Pathway Interaction Database (PID) pathways were downloaded from v7.1 (March 2020) of the Molecular Signatures Database. Analysis was restricted to GO terms containing between 10 and 2000 genes. No size restrictions were placed on the other gene sets, since there were many fewer of them. This resulted in a total of 10,053 gene sets for analysis.

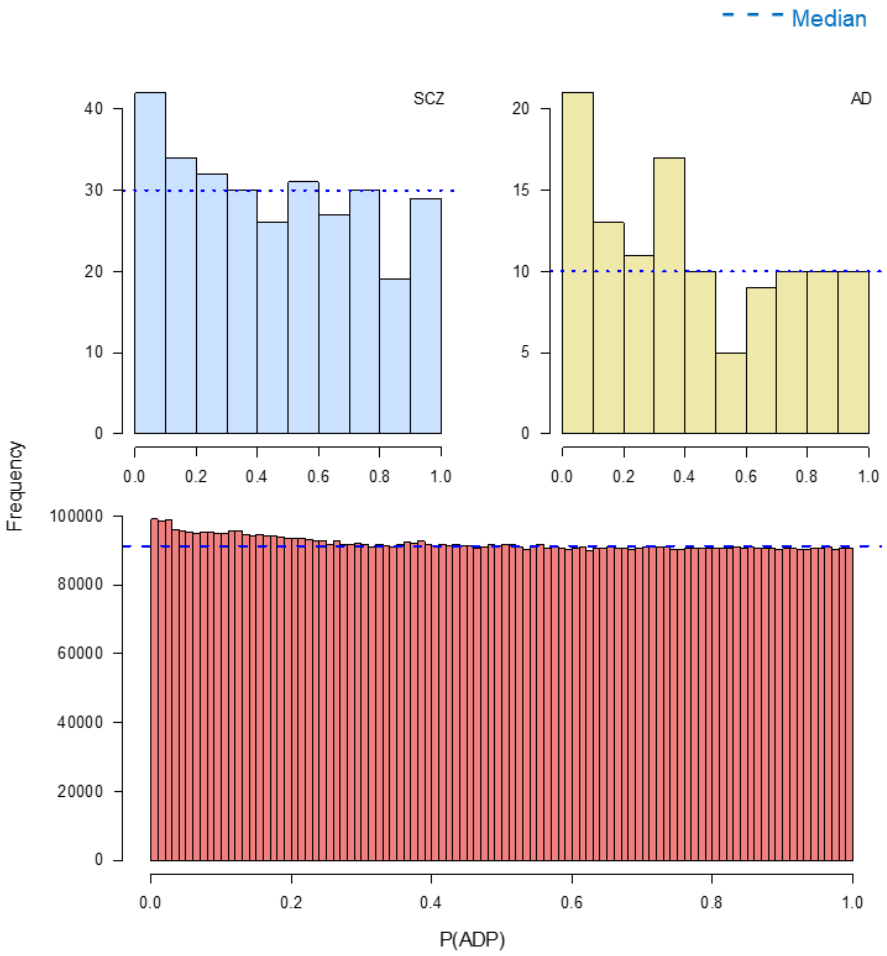
6.2. TRANSCRIPTOME-WIDE ASSOCIATION (TWAS): Using the GRCh37 hg19 genome assembly, common variants from the AD+P summary statistics were filtered using the munge_sumstats.py (v2.7.13) script from the LD Score Regression (LDSC) software and the hapmap 3.0 reference panel resulting in 1,184,883 SNPs. The FUSION package[51] 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).[52] GTEx7 and CMC expression weights and the reference panel from the 1000 Genomes European population were downloaded from the FUSION web site. Expression weights were used in TWAS for autosomal chromosomes and excluding the MHC region, with the processed AD+P summary statistics using the R script FUSION.assoc_test.R from the FUSION software.[51] Results were corrected for multiple testing of multiple genes within each tissue using the Bonferroni method.

6.3 POLYGENIC RISK SCORE: 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[53] to compute a PRS for our subjects, developed from GWAS results for AD (PRS_{AD}), [54] schizophrenia (PRS_{SZ}),[55] and bipolar disorder (PRS_{BP}),[3] 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).

As a first step in pruning SNPs for the score, we required SNPs to meet these criteria: must be included in both GWAS of interest (AD+P and other phenotype); genotypes called in all three data sets (Phase 1, Phase 2, and GR@ACE); MAF > 0.01; not an indel; difference between MAF for AD+/-P versus the MAF for the other phenotype < 0.05; if palindromic, MAF < 0.40. Also, for PRSAD, we selected only the most influential SNP (smallest GWAS p-value) within the APOE locus (chr19:45,000,000-46,000,000); and, for all three scores, we selected only the most influential SNP within the MHC region (chr6:25,000,000-34,000,000). As a next step, SNPs were clumped using Plink command --clump with a clump window of 500Kb and a maximum r2 between pairs of 0.50. The PRS were subsequently determined for each of the phenotypes using --score command with the options "center" and "sum" for p-value and for the thresholds given above. Logistic regression of AD+P status on PRS was performed using two ancestry vectors as additional covariates. Function nagelkerke in R library rcompanion was used to obtain the partial pseudo-R2 attributable to the PRS.

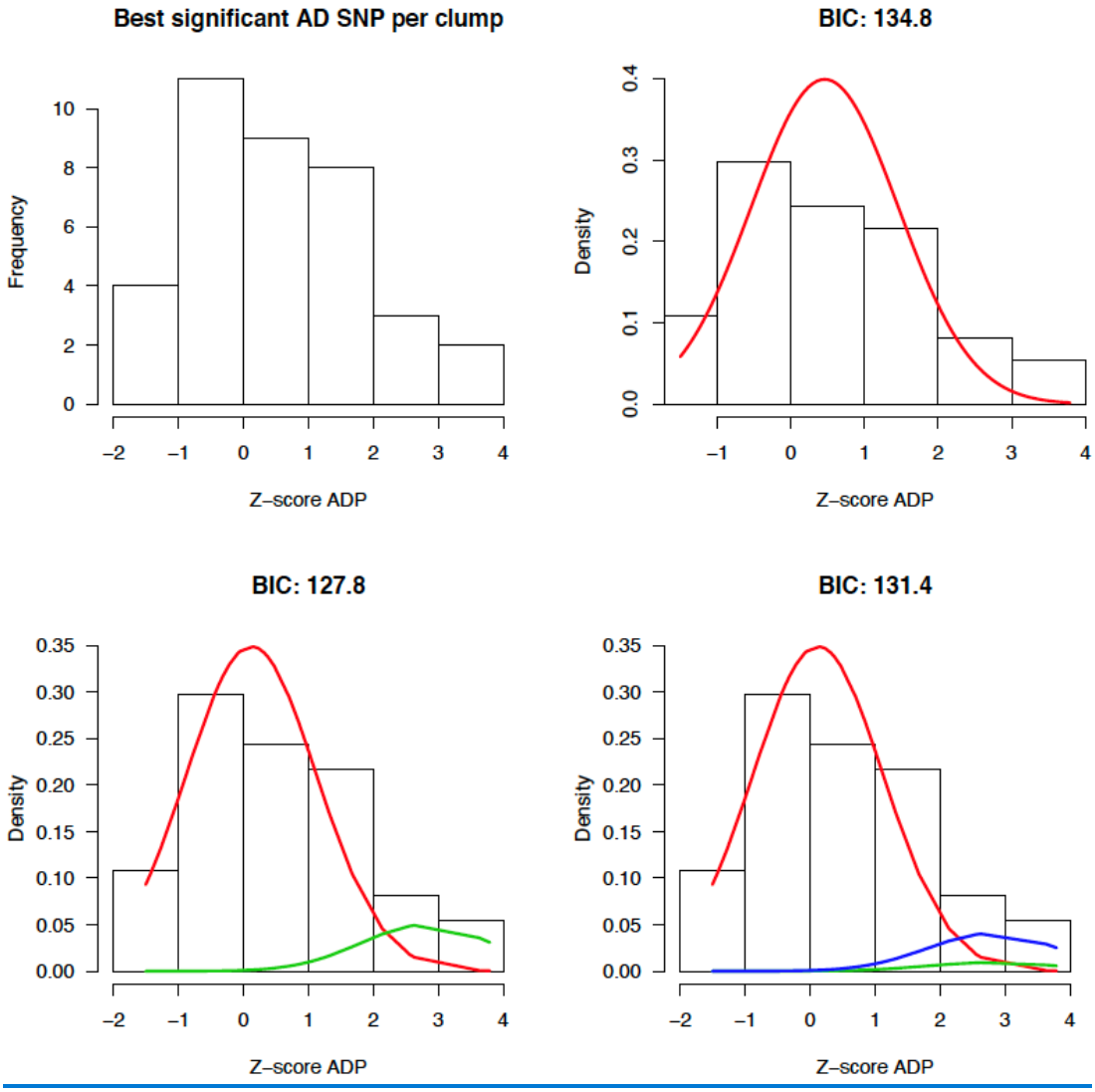
6.4. ANALYSES OF OVERLAP WITH SCHIZOPHRENIA, BIPOLAR, AND AD RISK SNPs: To evaluate whether there are consistent signals across studies and traits, we took three approaches. First, using a threshold of 5×10^{-8}, we selected GWAS-significant SNPs for schizophrenia (SCZ),[55] bipolar disorder (BPD),[3] or AD[54] and matched them to SNPs analyzed in our AD+P GWAS. Among the 18,276 significant SZ SNP, we matched 18,256 to SNPs in our AD+P GWAS. None crossed the p-value threshold of 10^{-4} for our AD+P GWAS. The same was true for the 264 significant BPD SNPs, which could be mapped to 238 SNPs from our AD+P GWAS. Among the 2394 GWAS significant AD SNPs, we matched 2284 to SNPs in our AD+P GWAS, and 11 crossed the threshold, all in the APOE region.

Next, we addressed this a different way. If there were consistent signals across studies, at least for some fraction of the SNPs (π), then we would expect the distribution of p-values from the AD+P GWAS, drawn because they had p-value < 5×10^{-8} for the complementary GWAS (SCZ, BPD, or AD), to be a mixture of two densities: fraction π, bearing true signal, would tend to have small p-values and they are often assumed to follow a Beta distribution; for the remainder, $1-\pi$, the p-values are expected to be drawn from the null distribution, a uniform on the interval (0,1). Because some SNPs are not independent, we first clumped SNPs by LD (pairwise LD r^2 < .20, using clump in PLINK with a 500Kb region) and selected the SNP, within clump, with the largest association p-value for the disease (SCZ, BPD, or AD). After clumping, if there were many SNPs falling in the same locus, such as around APOE for AD, we choose one SNP for the locus. When we fit this model to the p-values for 205, 16, and 37 selected SNPs for SCZ, BPD, and AD (**Supplementary Figure S6** for SCZ and AD, as well as the full distribution of AD+P GWAS p-values for contrast), respectively, and estimate π, we obtain $\pi_{\text{SCZ}} = 0.13$, $\pi_{\text{BPD}} = 0.39$, and $\pi_{\text{AD}} = 0.41$. The estimate for bipolar disorder, based on a small number of SNPs, is not reliable.

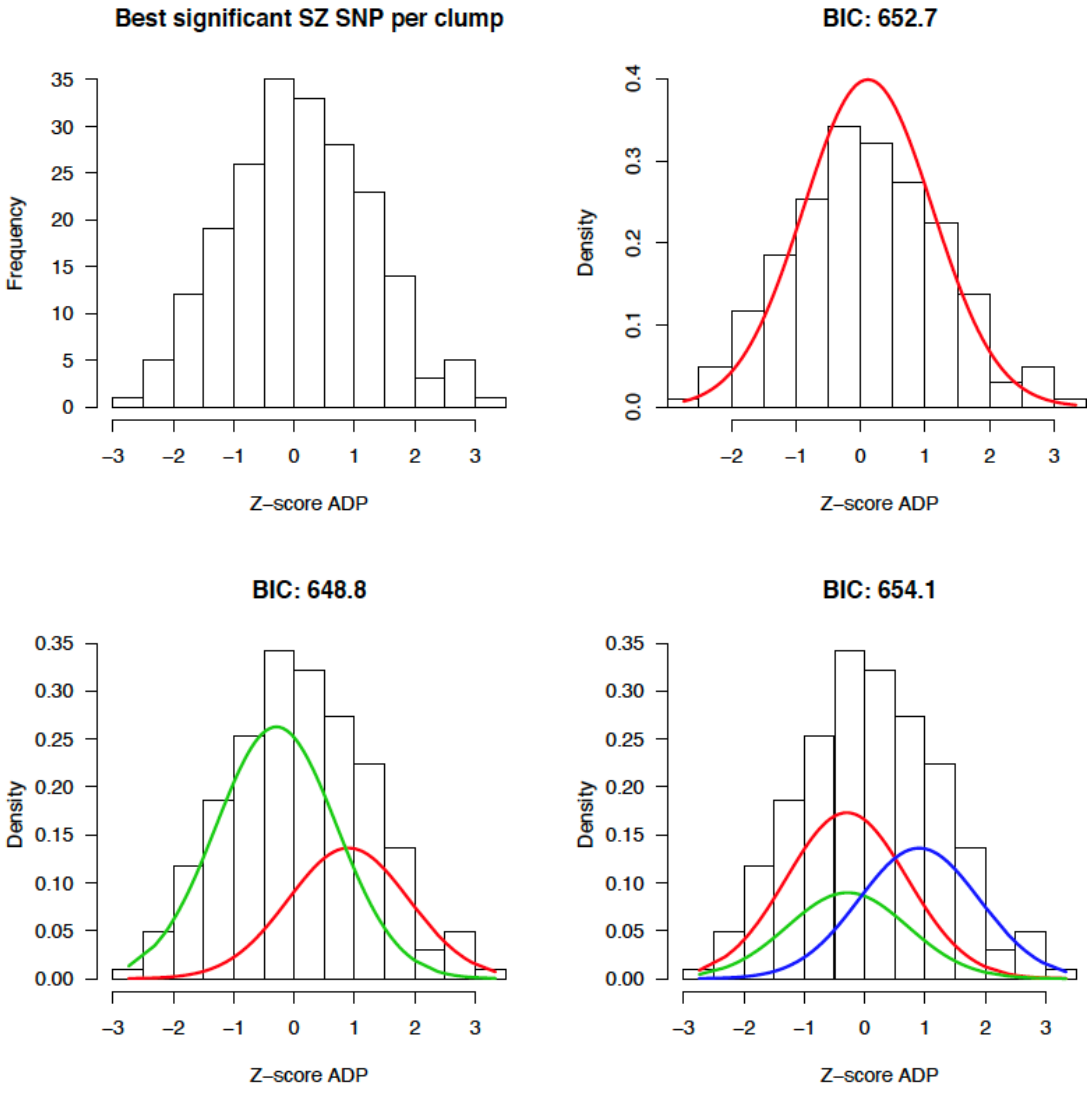


Supplementary Figure S6. Distribution of AD+P p-values for GWAS significant SNPs for SCZ and AD (top) and the entire AD+P p-value distribution (bottom). Dashed horizontal reference line is the median value of the AD+P p-values. Not shown: AD+P p-values for GWAS significant SNPs for BPD because the relatively few p-values made the mixture model analysis unreliable.

Finally, we asked if any particular GWAS-significant SNP for SCZ, BPD, and AD, or set of SNPs, also showed sufficient evidence for significance in AD+P. We approached this by using results from the clumped/thinned SNPs described above. Next, we sought evidence for a mixture of distributions of z-scores for the identified sets of SNPs. If there is evidence of a mixture, did any observations show compelling evidence for membership in one of the component distributions (PPr > 0.95)? As expected, we found evidence for a mixture for SCZ and AD, but not BPD. Both SCZ and AD z-scores were consistent with a mixture of two normal distributions (**Supplementary Figures S7-S8**). After estimating the posterior probability (PPr) of membership in these distributions, based on the parameters of the mixture model, 55 significant SNPs from the AD GWAS showed notable association for AD+P Z-scores, 21 with PPr.1 > 0.95 (left distribution) and 2 with PPr.2 > 0.95 (right distribution). For the significant SCZ SNP 71 SNP with PPr.1 > 0.95 (left distribution) and 8 with PPr.2 > 0.95 (right distribution) showed notable association with AD+P.



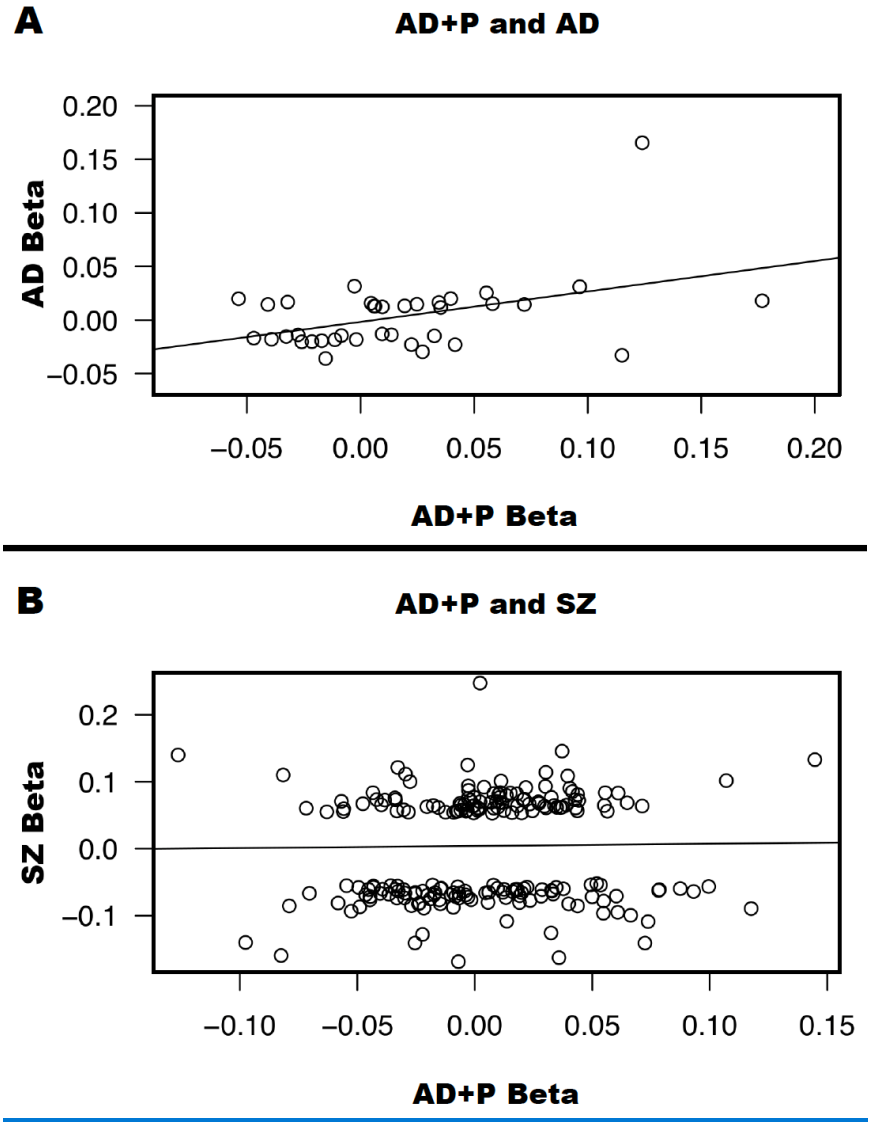
Supplementary Figure S7. Distribution of AD+P z-scores for GWAS significant SNPs for AD fit for 0, 1, 2 and 3 components of a normal. BIC is the Bayesian Information Criterion. Here, if the BIC decreases substantially with the addition of parameters, then they are needed to describe the distribution of the z-scores. In this case, the BIC decreases substantially by adding a second component to the mixture distribution, but it decreases thereafter. Hence we chose a two-component mixture as the best representation.



Supplementary Figure S8. Distribution of AD+P z-scores for GWAS significant SNPs for SCZ fit for 0, 1, 2 and 3 components of a normal. BIC is the Bayesian Information Criterion. Here, if the BIC decreases substantially with the addition of parameters, then they are needed to describe the distribution of the z-scores. In this case, the BIC decreases substantially by adding a second component to the mixture distribution, but it decreases thereafter. Hence we chose a two-component mixture as the best representation.

.

For context, we also evaluated the relationship between the estimated effects for AD and SCZ versus AD+P.



Supplementary Figure S9. Distribution of AD+P p-values for GWAS significant SNPs for SCZ and AD (top) and the entire AD+ P p-value distribution (bottom). Dashed horizontal reference line is the median value of the AD+P p-values. Not shown: AD+P p-values for GWAS significant SNPs for BPD because the relatively few p-values made the mixture model analysis unreliable.

7. SUPPLEMENTARY DISCUSSION OF RESULTS

Association Test Results

Of interest, we found some evidence for association of AD+P with rs11701 (Supplementary Table 1, 4.0×10^{-6}). rs11701 is located on chromosome 14, within two genes, *ANG* and *RNASE4*, which share promoters and some 5' exons, although each gene is spliced to a unique downstream exon containing its complete coding region. rs11701 is a synonymous coding variant in *ANG*, is intronic in *RNASE4*, and its protein products are members of the RNAse A Superfamily. Loss-of-function mutations in both *ANG* and *RNASE4* have been associated with amyotrophic lateral sclerosis/frontotemporal dementia (ALS/FTD) syndromes.[56-61] ANG preferentially hydrolyzes transfer RNA (tRNA) *in vivo*.[62] Understanding of how tRNA-derived fragments generated by ANG may contribute to neuroprotection is rapidly expanding.[63]

It would thus be parsimonious to hypothesize that less potent alterations in expression of *ANG* and *RNASE4*, or in functions of their protein products, modify neuronal function or survival, which in turn, increases the risk for psychosis in individuals with AD. Several observations are consistent with this hypothesis. For example, rs11701 has been linked to expression levels for both genes in some tissues.[64] While it is not known if mutations in *ANG* or *RNASE4* are present in AD+P, some of these mutations result in tar DNA binding protein 43 (TDP43) inclusion pathology on postmortem exam in ALS/FTD patients.[65] We have previously reported that the presence of comorbid TDP43 pathology in AD is independently associated with psychosis risk.[66]

TWAS Results

VN1R108P in GTEx7 hippocampus was TWAS significant in this analysis. VN1R108P is a member of the Vomeronasal receptors and it is classified as a pseudogene. Pseudogenes, which are homologous to protein-coding genes but have lost their coding ability, were thought to be functionless.[67] However, more recently, it has been shown that pseudogenes can be expressed and that they can have roles in gene regulation.[68,69] Such relationships have been shown previously in schizophrenia.[70] VN1R108P is not well annotated, however, it is expressed in a wide variety of tissues and could have a role in gene regulation. However, we note that the best eQTL snp in this analysis, rs4815438 is located in the ZNF337-AS1 gene which is a long non-coding RNA.

FAM182B is a long non-coding RNA (lncRNA) which is highly expressed in the cerebellum (https://www.proteinatlas.org/ENSG00000175170-FAM182B/tissue) and may be involved in the development of the foetal neocortex.[71] In addition, it has been suggested that FAM182B could be an important regulator in hepatocellular carcinoma.[72]

Long non-coding RNAs (lncRNAs) have many roles in the regulation of gene expression for example, they are involved in the regulation of the epigenome by association with chromatin modifying complexes[73] and they can affect gene expression through association with chromatin- modifying complexes.[74] FAM182B shows TWAS associations in a consistent direction in numerous brain tissues (Supplementary Table S5), suggesting that it could be an important regulator in the brain.

Definition of AD+P Phenotype

We note that we examined the association of genetic variation with a psychosis phenotype defined by the presence of one or more of multiple individual psychotic symptoms. It is possible that relevant subgroups could exist within the AD+P phenotype.[75] However, several lines of evidence support the phenotype definition used in the current study. First, our approach mirrors that used to define other psychotic disorders, e.g. schizophrenia. Second, AD+P as defined in the current study demonstrates familial aggregation and heritability.[27,31,34] Moreover, the heritability of AD+P is highest when requiring the presence of multiple or recurrent psychotic symptoms.[31,34] Nevertheless, it remains possible that any of the loci we report as significantly or suggestively associated with AD+P reflect a selective association with only a subset of psychosis symptoms.

Another potential limitation of our phenotypes, imprecision of our diagnosis of AD, seems unlikely. We did not observe any genetic correlation between AD+P risk and risk for ALS or Parkinson's disease. The lack of genetic overlap with ALS suggests that cases for whom a comorbid pathogenic process associated with ALS/FTD contributes to the manifestation of psychosis in AD account for only a small proportion of AD+P. This conclusion is consistent with our prior pathologic observations, in which the presence of TDP-43 pathology predicted only 5% of the variance in risk for AD+P.[66] Similarly, the lack of genetic correlation with Parkinson's disease indicates that our AD+P phenotype did not derive from a phenocopy due to inclusion of subjects with primary alpha-synuclein pathology, which itself is associated with psychosis.[76]

8. ACKNOWLEDGMENTS

ACE: Fundació ACE Barcelona Alzheimer Treatment and Research Center:
We thank all patients for their participation in this project. We are obliged to Trinitat Port-Carbó and her family for their support of the Fundació ACE research programs. Fundació ACE collaborates with the Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED, Spain), and is one of the participating centers of the Dementia Genetics Spanish Consortium (DEGESCO). CIBERNED is an Instituto de Salud Carlos III ISCIII Project. AR is supported by Grant PI13/02434 (Acción Estratégica en Salud, Instituto de Salud Carlos III (ISCIII), Ministerio de Economía y Competitividad, Spain) and Fundación Bancaria "La Caixa" (Barcelona, Spain).

Fundació ACE Contributors (non-authors): Asunción Lafuente MD, Octavio Rodríguez-Gómez MD, Liliana Vargas MD, Carla Abdelnour MD, Domingo Sánchez MD, Ana Mauleón MD, Maitée Rosende-Roca MD, Silvia Gil MD, PhD, Montserrat Alegret PhD, Ana Espinosa MS, Gemma Ortega PhD, Ángela Sanabria PhD, Alba Pérez MS, Marta Ibarria MS, Susana Diego MS, María José de Dios MS, Pilar Canyabate PhD, Mariola Moreno MS, Sílvia Preckler MS, Oscar Sotolongo-Grau PhD, Itziar de Rojas MS, Elvira Martín MS, Susana Ruiz NSG, Marina Tarragona NSG, Judit Serra NSG.

GR@ACE Consortium:
The Genome Research @ Fundació ACE project (GR@ACE) is supported by Fundación bancaria "La Caixa," Grifols SA, Fundació ACE, and ISCIII (Ministry of Health, Spain). We also want to thank the private sponsors who support the basic and clinical projects of our institution (Piramal AG, Laboratorios Echevarne, Araclon Biotech S.A., and Fundació ACE). We are indebted to the Trinitat Port-Carbó legacy and her family for their support of Fundació ACE research programs. Fundació ACE is a participating center in the Dementia Genetics Spanish Consortium (DEGESCO). A.R. and M.B. receive support from the European Union/EFPIA Innovative Medicines Initiative Joint Undertaking ADAPTED and MOPEAD projects (Grants No. 115975 and 115985, respectively). M.B. and A.R. are also supported by national grants PI13/02434, PI16/01861, and PI17/01474, PI19/01301 and CIBERNED grant 2019/08. Acción Estratégica en Salud is integrated into the Spanish National R + D + I Plan and funded by ISCIII (Instituto de Salud Carlos III)-Subdirección General de Evaluación and the Fondo Europeo de Desarrollo Regional (FEDER- "Una manera de Hacer Europa").
Regional (FEDER- "Una manera de Hacer Europa"). The genotyping service to generate GR@ACE GWAS data was carried out at CEGEN-PRB3-ISCIII; it is supported by grant PT17/0019, of the PE I+D+i 2013-2016, funded by ISCIII and ERDF". GR@ACE consortia would also like to thank to all researchers contributing to this project. A complete list of collaborators involved in the GR@ACE GWAS can be found at https://ciberned.es/proyectos/degesco.html.

The GR@ACE study group and affiliations:
Abdelnour C[1,2], Aguilera N[1], Alarcon-Martin E[1,3], Alegret M[1,2], Benaque A[1], Boada M[1,2], Buendia M[1], Cañabate P[1,2], Carracedo A[4,5], Corbatón A[6], de Rojas I[1], Diego S[1], Espinosa A[1,2], Gailhajenet A[1], García González P[1], Gil S[1], Guitart M[1], González Pérez A[7], Hernández I[1,2], Ibarria, M[1], Lafuente A[1], Macias J[8], Maroñas O[4], Martín E[1], Martínez MT[6], Mauleón A[1], Monté-Rubio G[1], Montrreal L[1], Moreno-Grau S[1,2], Moreno M[1], Orellana A[1], Ortega G[1,2], Pancho A[1], Pelejà E[1], Pérez-Cordon A[1], Pineda JA[8], Preckler S[1], Quintela I[3], Real LM[3,8], Rodríguez-

Gómez O[1,2], Rosende-Roca M[1], Ruiz A[1,2], Ruiz S[1,2], Sáez ME[7], Sanabria A[1,2], Santos-Santos MA[1], Serrano-Rios M[6], Sotolongo-Grau O[1], Tárraga L[1,2], Valero S[1,2], Vargas L[1].
[1]Research Center and Memory clinic Fundació ACE. Institut Català de Neurociències Aplicades. Universitat Internacional de Catalunya. Barcelona. Spain; [2]CIBERNED, Center for Networked Biomedical Research on Neurodegenerative Diseases, National Institute of Health Carlos III, Ministry of Economy and Competitiveness, Spain; [3]Dep. of Surgery, Biochemistry and Molecular Biology, School of Medicine. University of Málaga. Málaga. Spain; [4]Grupo de Medicina Xenómica, Centro Nacional de Genotipado (CEGEN-PRB3-ISCIII). Universidade de Santiago de Compostela, Santiago de Compostela, Spain; [5]Fundación Pública Galega de Medicina Xenómica- CIBERER-IDIS, Santiago de Compostela, Spain; [6]Centro de Investigación Biomédica en Red de Diabetes y Enfermedades Metabólicas Asociadas, CIBERDEM, Spain, Hospital Clínico San Carlos, Madrid, Spain; [7]CAEBI. Centro Andaluz de Estudios Bioinformáticos., Sevilla, Spain; [8]Unidad Clínica de Enfermedades Infecciosas y Microbiología. Hospital Universitario de Valme, Sevilla, Spain)

NIA-LOAD: The NIA-LOAD Family Based Study (FBS) Consortium supported the collection of samples used in this study through National Institute on Aging (NIA) grants 1U24AG056270 and 5R01AG041797. We thank contributors, including the Alzheimer's Disease Centers who collected samples used in this study, as well as patients and their families, whose help and participation made this work possible.

NIA-LOAD Investigators and affiliations:
Haines, J., Case Western Reserve University, Cleveland, OH; Farlow, M., Indiana University, Indianapolis, IN; Boeve, B., Mayo Clinic, Rochester, MN; Pericak-Vance, M. University of Miami, Miami, FL., Foroud, T., National Centralized Repository for Alzheimer's Disease, Indiana, IN; Byrd, G.S., Wake Forest University, Winston-Salem, NC; Bennett, D.A., Rush University Medical Center, Chicago, IL; Sweet, R.A., University of Pittsburgh, Pittsburgh, PA; Goate, A.M., Mount Sinai School of Medicine, New York, NY; Rosenberg, R., University of Texas Southwestern, Dallas, TX; Bird, T.D. University of Washington, Seattle, WA; Tsuang, D., University of Washington, Seattle, WA; Cruchaga, Washington School of Medicine, St. Louis, MO; Mayeux, R.P., Columbia University, New York, NY.

NCRAD: Samples included in the Pitt-ADRC, NIA-LOAD, and ADC cohorts were provided by the National Centralized Repository for Alzheimer's Disease and Related Dementias (NCRAD), which receives government support under a cooperative agreement grant (U24 AG021886) awarded by the National Institute on Aging (NIA). We thank contributors who collected samples used in this study, as well as patients and their families, whose help and participation made this work possible.

NACC: Phenotypic data for the ADC cohort were provided by NACC. The NACC database is funded by NIA/NIH Grant U01 AG016976. NACC data are contributed by the NIA funded ADCs: P30 AG019610 (PI Eric Reiman, MD), P30 AG013846 (PI Neil Kowall, MD), P50 AG008702 (PI Scott Small, MD), P50 AG025688 (PI Allan Levey, MD, PhD), P50 AG047266 (PI Todd Golde, MD, PhD), P30 AG010133 (PI Andrew Saykin, PsyD), P50 AG005146 (PI Marilyn Albert, PhD), P50 AG005134 (PI Bradley Hyman, MD, PhD), P50 AG016574 (PI Ronald Petersen, MD, PhD), P50 AG005138 (PI Mary Sano, PhD), P30 AG008051 (PI Steven Ferris, PhD), P30 AG013854 (PI M. Marsel Mesulam, MD), P30 AG008017 (PI Jeffrey Kaye, MD), P30 AG010161 (PI David Bennett, MD), P50 AG047366 (PI Victor Henderson, MD, MS), P30 AG010129 (PI Charles DeCarli, MD), P50 AG016573 (PI Frank LaFerla, PhD), P50 AG016570 (PI Marie-Francoise Chesselet, MD, PhD), P50 AG005131 (PI Douglas Galasko, MD), P50 AG023501 (PI Bruce Miller, MD), P30 AG035982 (PI Russell Swerdlow, MD) , P30 AG028383 (PI Linda Van Eldik, PhD), P30 AG010124 (PI John Trojanowski, MD, PhD), P50 AG005133 (PI Oscar Lopez, MD), P50 AG005142 (PI Helena Chui, MD), P30 AG012300 (PI Roger Rosenberg, MD), P50 AG005136 (PI Thomas Montine, MD, PhD), P50 AG033514 (PI Sanjay Asthana, MD, FRCP), P50 AG005681 (PI John Morris, MD), and P50 AG047270 (PI Stephen Strittmatter, MD, PhD)."

ADGC: The Alzheimer's Disease Genetics Consortium supported the collection of samples used in this study through National Institute on Aging (NIA) grants U01AG032984 and RC2AG036528. Data for this study were

prepared, archived, and distributed by the National Institute on Aging Alzheimer's Disease Data Storage Site (NIAGADS) at the University of Pennsylvania (U24-AG041689-01). Alzheimer's Disease Genetics Consortium members and affiliations:

Erin Abner [1], Perrie M. Adams [2], Marilyn S. Albert [3], Roger L. Albin [4-,6], Liana G. Apostolova [7-10], Steven E. Arnold [11], Sanjay Asthana [12-14], Craig S. Atwood [12-14], Clinton T. Baldwin [15], Robert C. Barber [16], Lisa L. Barnes [17-19], Sandra Barral [20-22], Thomas G. Beach [23], James T. Becker [24], Gary W. Beecham [25,26], Duane Beekly [27], David A. Bennett [17,19], Eileen H. Bigio [28,29], Thomas D. Bird [30,31], Deborah Blacker [32,33], Bradley F. Boeve [34], James D. Bowen [35], Adam Boxer [36], James R. Burke [37], Jeffrey M. Burns [38], Joseph D. Buxbaum [39-41], Nigel J. Cairns [42], Laura B. Cantwell [43], Chuanhai Cao [44], Chris S. Carlson [45], Cynthia M. Carlsson [12-14], Regina M. Carney [46], Minerva M. Carrasquillo [47], Helena C. Chui [48], Paul K. Crane [49], David H. Cribbs [50], Elizabeth A. Crocco [46], Carlos Cruchaga [51], Philip L. De Jager [52,53], Charles DeCarli [54], Malcolm Dick [55], Dennis W. Dickson [47], Rachelle S. Doody [56], Ranjan Duara [57], Nilufer Ertekin-Taner [47,58], Denis A. Evans [59], Kelley M. Faber [8], Thomas J. Fairchild [60], Kenneth B. Fallon [61], David W. Fardo [62], Martin R. Farlow [63], Lindsay A. Farrer [64-68], Steven Ferris [69], Tatiana M. Foroud [8], Matthew P. Frosch [70], Douglas R. Galasko [71], Marla Gearing [72,73], Daniel H. Geschwind [74], Bernardino Ghetti [75], John R. Gilbert [25,26], Alison M. Goate [39], Neill R. Graff-Radford [47,58], Robert C. Green [76], John H. Growdon [77], Jonathan L. Haines [78], Hakon Hakonarson [79], Ronald L. Hamilton [80], Kara L. Hamilton-Nelson [25], John Hardy [81], Lindy E. Harrell [82], Lawrence S. Honig [20], Ryan M. Huebinger [83], Matthew J. Huentelman [84], Christine M. Hulette [85], Bradley T. Hyman [77], Gail P. Jarvik [86,87], Lee-Way Jin [88], Gyungah Jun [15,64,68], M. Ilyas Kamboh [89,90], Anna Karydas [36], Mindy J. Katz [91], John S.K. Kauwe [92], Jeffrey A. Kaye [93,94], C. Dirk Keene [95], Ronald Kim [96], Neil W. Kowall[67,97], Joel H. Kramer [98], Walter A. Kukull [99], Brian W. Kunkle [25], Amanda P. Kuzma [43], Frank M. LaFerla [100], James J. Lah [101], Eric B. Larson [49,102], Yuk Ye Leung[43], James B. Leverenz [103], Allan I. Levey [101], Ge Li [31,104], Andrew P. Lieberman [105], Richard B. Lipton [91], Oscar L. Lopez [90], Kathryn L. Lunetta [64], Constantine G. Lyketsos [106], John Malamon [43], Daniel C. Marson [82], Eden R. Martin [25,26], Frank Martiniuk [107], Deborah C. Mash [108], Eliezer Masliah [71,109], Richard Mayeux [20,21], Wayne C. McCormick [49], Susan M. McCurry [110], Andrew N. McDavid [45], Stefan McDonough [111], Ann C. McKee [67,97], Marsel Mesulam [29,112], Bruce L. Miller[36], Carol A. Miller [113], Joshua W. Miller [88], Thomas J. Montine [95], John C. Morris [42,114], Shubhabrata Mukherjee [49], Amanda J. Myers [46], Adam C. Naj [43], Sid O'Bryant [115], John M. Olichney [54], Joseph E. Parisi[116], Henry L. Paulson [117], Margaret A. Pericak-Vance [25,26], Elaine Peskind [104], Ronald C. Petersen [34], Aimee Pierce [50], Wayne W. Poon [55], Huntington Potter [118], Liming Qu [43], Joseph F. Quinn [93,94], Ashok Raj[44], Murray Raskind [104], Eric M. Reiman [84,119-121], Barry Reisberg [69,122], Joan S. Reisch [123], Christiane Reitz[20-22,124], John M. Ringman [48], Erik D. Roberson [82], Ekaterina Rogaeva [125], Howard J. Rosen [36], Roger N. Rosenberg [127], Donald R. Royall [128], Mark A. Sager [13], Mary Sano [40], Andrew J. Saykin [7,8], Gerard D. Schellenberg [43], Julie A. Schneider [17,19,129], Lon S. Schneider [48,130], William W. Seeley [36], Amanda G. Smith[44], Joshua A. Sonnen [95], Salvatore Spina [75], Peter St George-Hyslop [131,132], Robert A. Stern [67], Russell H. Swerdlow [38], Rudolph E. Tanzi [77], John Q. Trojanowski [133], Juan C. Troncoso [134], Debby W. Tsuang [31,104], Otto Valladares [43], Vivianna M. Van Deerlin [133], Linda J. Van Eldik [135], Badri N. Vardarajan [20-22], Harry V. Vinters [136,137], Jean Paul Vonsattel [138], Li-San Wang [43], Sandra Weintraub [28,29], Kathleen A. Welsh-Bohmer[37,139], Kirk C. Wilhelmsen [140], Jennifer Williamson [20], Thomas S. Wingo [101], Randall L. Woltjer [141], Clinton B. Wright [142], Chuang-Kuo Wu [143], Steven G. Younkin [47], Chang-En Yu [49], Lei Yu [17,19], Yi Zhao [43]

[1]Sanders-Brown Center on Aging, College of Public Health, Department of Epidemiology, University of Kentucky, Lexington, Kentucky, [2]Department of Psychiatry, University of Texas Southwestern Medical Center, Dallas, Texas, [3]Department of Neurology, Johns Hopkins University, Baltimore, Maryland, [4]Department of Neurology, University of Michigan, Ann Arbor, Michigan, [5]Geriatric Research, Education and Clinical Center (GRECC), VA Ann Arbor Healthcare System (VAAAHS), Ann Arbor, Michigan, [6]Michigan Alzheimer Disease Center, Ann Arbor, Michigan, [7]Department of Radiology, Indiana University, Indianapolis, Indiana, [8]Department of Medical and Molecular Genetics, Indiana University, Indianapolis, Indiana, [9]Indian Alzheimer's Disease Center, Indiana University, Indianapolis, Indiana, [10]Department of Neurology, Indiana University, Indianapolis, Indiana, [11]Department of Psychiatry, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, [12]Geriatric Research, Education and Clinical Center (GRECC), University of Wisconsin, Madison, Wisconsin, [13]Department of Medicine, University of Wisconsin, Madison, Wisconsin, [14]Wisconsin Alzheimer's Disease Research Center, Madison, Wisconsin, [1] Department of Medicine (Genetics Program), Boston University, Boston, Massachusetts, [16]Department of Pharmacology and Neuroscience, University of North Texas Health Science Center, Fort Worth, Texas, [17]Department of Neurological Sciences, Rush

University Medical Center, Chicago, Illinois, [18]Department of Behavioral Sciences, Rush University Medical Center, Chicago, Illinois, [19]Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, Illinois, [20]Taub Institute on Alzheimer's Disease and the Aging Brain, Department of Neurology, Columbia University, New York, New York, [21]Gertrude H. Sergievsky Center, Columbia University, New York, New York, [22]Department of Neurology, Columbia University, New York, New York, [23]Civin Laboratory for Neuropathology, Banner Sun Health Research Institute, Phoenix, Arizona, [24]Departments of Psychiatry, Neurology, and Psychology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, [25]The John P. Hussman Institute for Human Genomics, University of Miami, Miami, Florida, [26]Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami, Miami, Florida, [27]National Alzheimer's Coordinating Center, University of Washington, Seattle, Washington, [28]Department of Pathology, Northwestern University Feinberg School of Medicine, Chicago, Illinois, [29]Cognitive Neurology and Alzheimer's Disease Center, Northwestern University Feinberg School of Medicine, Chicago, Illinois, [30]Department of Neurology, University of Washington, Seattle, Washington, [31]VA Puget Sound Health Care System/GRECC, Seattle, Washington, [32]Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, [33]Department of Psychiatry, Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts, [34]Department of Neurology, Mayo Clinic, Rochester, Minnesota, [35]Swedish Medical Center, Seattle, Washington, [36]Department of Neurology, University of California San Francisco, San Francisco, California, [37]Department of Medicine, Duke University, Durham, North Carolina, [38]University of Kansas Alzheimer's Disease Center, University of Kansas Medical Center, Kansas City, Kansas, [39]Department of Neuroscience, Mount Sinai School of Medicine, New York, New York, [40]Department of Psychiatry, Mount Sinai School of Medicine, New York, New York, [41]Departments of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, New York, [42]Department of Pathology and Immunology, Washington University, St. Louis, Missouri, [43]Penn Neurodegeneration Genomics Center, Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, [44]USF Health Byrd Alzheimer's Institute, University of South Florida, Tampa, Florida, [45]Fred Hutchinson Cancer Research Center, Seattle, Washington, [46]Department of Psychiatry and Behavioral Sciences, Miller School of Medicine, University of Miami, Miami, Florida, [47]Department of Neuroscience, Mayo Clinic, Jacksonville, Florida, [48]Department of Neurology, University of Southern California, Los Angeles, California, [49]Department of Medicine, University of Washington, Seattle, Washington, [50]Department of Neurology, University of California Irvine, Irvine, California, [51]Department of Psychiatry and Hope Center Program on Protein Aggregation and Neurodegeneration, Washington University School of Medicine, St. Louis, Missouri, [52]Program in Translational NeuroPsychiatric Genomics, Institute for the Neurosciences, Department of Neurology & Psychiatry, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, [53]Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts, [54]Department of Neurology, University of California Davis, Sacramento, California, [55]Institute for Memory Impairments and Neurological Disorders, University of California Irvine, Irvine, California, [56]Alzheimer's Disease and Memory Disorders Center, Baylor College of Medicine, Houston, Texas, [57]Wien Center for Alzheimer's Disease and Memory Disorders, Mount Sinai Medical Center, Miami Beach, Florida, [58]Department of Neurology, Mayo Clinic, Jacksonville, Florida, [59]Rush Institute for Healthy Aging, Department of Internal Medicine, Rush University Medical Center, Chicago, Illinois, [60]Office of Strategy and Measurement, University of North Texas Health Science Center, Fort Worth, Texas, [61]Department of Pathology, University of Alabama at Birmingham, Birmingham, Alabama, [62]Sanders-Brown Center on Aging, Department of Biostatistics, University of Kentucky, Lexington, Kentucky, [63]Department of Neurology, Indiana University, Indianapolis, Indiana, [64]Department of Biostatistics, Boston University, Boston, Massachusetts, [65]Department of Epidemiology, Boston University, Boston, Massachusetts, [66]Department of Medicine (Biomedical Genetics), Boston University, Boston, Massachusetts, [67]Department of Neurology, Boston University, Boston, Massachusetts, [68]Department of Ophthalmology, Boston University, Boston, Massachusetts, [69]Department of Psychiatry, New York University, New York, New York, [70]C.S. Kubik Laboratory for Neuropathology, Massachusetts General Hospital, Charlestown, Massachusetts, [71]Department of Neurosciences, University of California San Diego, La Jolla, California, [72]Department of Pathology and Laboratory Medicine, Emory University, Atlanta, Georgia, [73]Emory Alzheimer's Disease Center, Emory University, Atlanta, Georgia, [74]Neurogenetics Program, University of California Los Angeles, Los Angeles, California, [75]Department of Pathology and Laboratory Medicine, Indiana University, Indianapolis, Indiana, [76]Division of Genetics, Department of Medicine and Partners Center for Personalized Genetic Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, [77]Department of

Neurology, Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts, [78]Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, Ohio, [79]Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, [80]Department of Pathology (Neuropathology), University of Pittsburgh, Pittsburgh, Pennsylvania, [81]Institute of Neurology, University College London, Queen Square, London, United Kingdom, [82]Department of Neurology, University of Alabama at Birmingham, Birmingham, Alabama, [83]Department of Surgery, University of Texas Southwestern Medical Center, Dallas, Texas, [84]Neurogenomics Division, Translational Genomics Research Institute, Phoenix, Arizona, [85]Department of Pathology, Duke University, Durham, North Carolina, [86]Department of Genome Sciences, University of Washington, Seattle, Washington, [87]Department of Medicine (Medical Genetics), University of Washington, Seattle, Washington, [88]Department of Pathology and Laboratory Medicine, University of California Davis, Sacramento, California, [89]Department of Human Genetics, University of Pittsburgh, Pittsburgh, Pennsylvania, [90]University of Pittsburgh Alzheimer's Disease Research Center, Pittsburgh, Pennsylvania, [91]Department of Neurology, Albert Einstein College of Medicine, New York, New York, [92]Department of Biology, Brigham Young University, Provo, Utah, [93]Department of Neurology, Oregon Health & Science University, Portland, Oregon, [94]Department of Neurology, Portland Veterans Affairs Medical Center, Portland, Oregon, [95]Department of Pathology, University of Washington, Seattle, Washington, [96]Department of Pathology and Laboratory Medicine, University of California Irvine, Irvine, California, [97]Department of Pathology, Boston University, Boston, Massachusetts, [98]Department of Neuropsychology, University of California San Francisco, San Francisco, California, [99]Department of Epidemiology, University of Washington, Seattle, Washington, [100]Department of Neurobiology and Behavior, University of California Irvine, Irvine, California, [101]Department of Neurology, Emory University, Atlanta, Georgia, [102]Group Health Research Institute, Group Health, Seattle, Washington, [103]Cleveland Clinic Lou Ruvo Center for Brain Health, Cleveland Clinic, Cleveland, Ohio, [104]Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine, Seattle, Washington, [105]Department of Pathology, University of Michigan, Ann Arbor, Michigan, [106]Department of Psychiatry, Johns Hopkins University, Baltimore, Maryland, [107]Department of Medicine - Pulmonary, New York University, New York, New York, [108]Department of Neurology, University of Miami, Miami, Florida, [109]Department of Pathology, University of California San Diego, La Jolla, California, [110]School of Nursing Northwest Research Group on Aging, University of Washington, Seattle, Washington, [111]PharmaTherapeutics Clinical Research, Pfizer Worldwide Research and Development, Cambridge, Massachusetts, [112]Department of Neurology, Northwestern University Feinberg School of Medicine, Chicago, Illinois, [113]Department of Pathology, University of Southern California, Los Angeles, California, [114]Department of Neurology, Washington University, St. Louis, Missouri, [115]Internal Medicine, Division of Geriatrics, University of North Texas Health Science Center, Fort Worth, Texas, [116]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota, [117]Michigan Alzheimer's Disease Center, Department of Neurology, University of Michigan, Ann Arbor, Michigan, [118]Department of Neurology, University of Colorado School of Medicine, Aurora, Colorado, [119]Arizona Alzheimer’s Consortium, Phoenix, Arizona, [120]Banner Alzheimer's Institute, Phoenix, Arizona, [121]Department of Psychiatry, University of Arizona, Phoenix, Arizona, [122]Alzheimer's Disease Center, New York University, New York, New York, [123]Department of Clinical Sciences, University of Texas Southwestern Medical Center, Dallas, Texas, [124]Department of Epidemiology, Columbia University, New York, New York, [125] Tanz Centre for Research in Neurodegenerative Disease, University of Toronto, Toronto, Ontario, [127]Department of Neurology, University of Texas Southwestern, Dallas, Texas, [128]Departments of Psychiatry, Medicine, Family & Community Medicine, South Texas Veterans Health Administration Geriatric Research Education & Clinical Center (GRECC), UT Health Science Center at San Antonio, San Antonio, Texas, [129]Department of Pathology (Neuropathology), Rush University Medical Center, Chicago, Illinois, [130]Department of Psychiatry, University of Southern California, Los Angeles, California, [131]Tanz Centre for Research in Neurodegenerative Disease, University of Toronto, Toronto, Ontario, [132]Cambridge Institute for Medical Research and Department of Clinical Neurosciences, University of Cambridge, Cambridge, United Kingdom, [133]Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, [134]Department of Pathology, Johns Hopkins University, Baltimore, Maryland, [135]Sanders-Brown Center on Aging, Department of Anatomy and Neurobiology, University of Kentucky, Lexington, Kentucky, [136]Department of Neurology, University of California Los Angeles, Los Angeles, California, [137]Department of Pathology & Laboratory Medicine, University of California Los Angeles, Los Angeles, California, [138]Taub Institute on Alzheimer's Disease and the Aging Brain, Department of Pathology, Columbia University, New York, New York, [139]Department of Psychiatry &

Behavioral Sciences, Duke University, Durham, North Carolina, [140]Department of Genetics, University of North Carolina Chapel Hill, Chapel Hill, North Carolina, [141]Department of Pathology, Oregon Health & Science University, Portland, Oregon, [142]Evelyn F. McKnight Brain Institute, Department of Neurology, Miller School of Medicine, University of Miami, Miami, Florida, [143]Departments of Neurology, Pharmacology & Neuroscience, Texas Tech University Health Science Center, Lubbock, Texas

Lilly: We would like to, first of all, express sincere gratitude to the thousands of patients, their care-takers and the close family of patients for their selfless participation in all the clinical trials sponsored by Eli Lilly. We also thank the site investigators and personnel, the members of the trial committees, the members of the R&D teams and clinical trial teams for their contribution to the trials included in this study. We would also like to acknowledge the immense support from current and past employees of Eli Lilly; Dr. Bradley B. Miller, Dr. Laura Nisenbaum, Ms Leanne Muncie, Dr. Yupeng Li, Dr. Brian Eastwood for their role in making this collaborative work a successful endeavor.

NEXGENS: This paper represents independent research part funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. Funding for genotyping the IRCCS 1 samples was provided by the University of Exeter. IRCCS data were fully supported by "Ministero della Salute", I.R.C.C.S.Research Program, Ricerca Corrente 2018−2020, Linea n. 2"Meccanismigenetici, predizione e terapie innovative delle malattie complesse"and by the"5 × 1000"voluntary contribution to the Fondazione I.R.C.C.S. Ospedale"CasaSollievo della Sofferenza". The authors are grateful to the investigators of the AddNeuroMed consortium for providing clinical data. AddNeuroMed is funded through the EU FP6 program as part of InnoMed. The authors wish to thank the Norwegian registry of persons assessed for cognitive symptoms (NorCog) for providing access to patient data and biological material. NorCog is financed by South-Eastern Norway Regional Health Authority and Norwegian National Advisory Unit on Ageing and Health. The organization and the data collection of HMS cohort has been funded by the Norwegian Institute of Public Health, the Norwegian University of Science and Technology (NTNU), Nord-Trøndelag Hospital Trust and Innlandet Hospital Trust. The REDIC-NH study was administrated by the research center for Age-Related Functional Decline and Disease, Innlandet Hospital Trust, and was initiated by the Norwegian Health Directorate, which also provides funding for the data collection. The authors are grateful to deCODE Genetics for performing the genotyping in the Norwegian cohorts. Research reported in this publication was part supported by the National Institute on Aging of the National Institutes of Health under Award Number R01AG067015. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

NEXGENS:
Byron Creese: 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. Clive Ballard: University of Exeter Medical School, Exeter, UK; Norwegian, Exeter and King's College Consortium for Genetics of Neuropsychiatric Symptoms in Dementia. Dag Aarsland: Department of Old Age Psychiatry, Institute of Psychiatry, Psychology and Neuroscience, King's College London; Centre for Age-Related Medicine, Stavanger University Hospital, Stavanger, Norway; Norwegian, Exeter and King's College Consortium for Genetics of Neuropsychiatric Symptoms in Dementia. Geir Selbaek: Research Centre of Age-related Functional Decline and Disease, Innlandet Hospital Trust, Pb 68, Ottestad 2312, Norway; Norwegian National Advisory Unit on Ageing and Health, Vestfold Hospital Trust, Tønsberg, Norway; Faculty of Medicine, University of Oslo, Oslo, Norway; Norwegian, Exeter and King's College Consortium for Genetics of Neuropsychiatric Symptoms in Dementia. Sverre Bergh: Research Centre of Age-related Functional Decline and Disease, Innlandet Hospital Trust, Pb 68, Ottestad 2312, Norway; Norwegian National Advisory Unit on Ageing and Health, Vestfold Hospital Trust, Tønsberg, Norway; Norwegian, Exeter and King's College Consortium for Genetics of Neuropsychiatric Symptoms in Dementia. Arvid Rongve: Department of Research and Innovation, Helse Fonna, Haugesund, Norway; Department of Clinical Medicine, University of Bergen, Bergen, Norway; Norwegian, Exeter and King's College Consortium for Genetics of Neuropsychiatric Symptoms in Dementia. Ingvild Saltvedt: Geriatric Department, St. Olav hospital, University Hospital of

Trondheim, Norway; Department of Neuromedicine and Movement science, Norwegian University of Science and Technology, Trondheim, Norway. Håvard K. Skjellegrind: HUNT Research Centre, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), Levanger, NorwayBo Lars Engdahl: Norwegian Institute of Public Health. Eystein Stordal: Department of Mental Health, Norwegian University of Science and Technology, Trondheim, 43 Norway; Department of Psychiatry, Namsos Hospital, Namsos, Norway. Ole A. Andreassen: grant support-Research Council of Norway #223273, NORMENT, Institute of Clinical Medicine, University of Oslo, Oslo, Norway; NORMENT, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway. Srdjan Djurovic: Department of Clinical Medicine, University of Bergen, Bergen, Norway; NORMENT, Department of Clinical Science, University of Bergen, Bergen, Norway. Lavinia Athanasiu: NORMENT, Institute of Clinical Medicine, University of Oslo, Oslo, Norway; NORMENT, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway. Davide Seripa: Complex Structure of Geriatrics, Department of Medical Sciences, Fondazione IRCCS “Casa Sollievo della Sofferenza”, San Giovanni Rotondo (FG), Italy. Barbara Borroni: Centre for Neurodegenerative Disorders, University of Brescia, Italy. Diego Albani: Neuroscience Department, Istituto di Ricerche Farmacologiche “Mario Negri”, Milan, Italy. Patrizia Meccoci: Institute of Gerontology and Geriatrics, University of Perugia, Italy. Alessandro Serretti: Institute of Psychiatry, University of Bologna, Bologna, Italy. Giulia Paroni: Complex Structure of Geriatrics, Department of Medical Sciences, Fondazione IRCCS “Casa Sollievo della Sofferenza”, San Giovanni Rotondo (FG), Italy. On behalf of the AddNeuroMed consortium. The authors are grateful to Simon Lovestone and the AddNeuroMed consortium.

This paper represents independent research partly funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King’s College London. The views expressed are those of the author and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

National Institute of Mental Health Genetics Initiative AD Cohort (NIMH):
Data and biomaterials were collected in three projects that participated in the National Institute of Mental Health (NIMH) Alzheimer Disease Genetics Initiative. From 1991-98, the Principal Investigators and Co-Investigators were: Massachusetts General Hospital, Boston, MA, U01 MH46281, Marilyn S. Albert, Ph.D., and Deborah Blacker, M.D., Sc.D.; Johns Hopkins University, Baltimore, MD; U01 MH46290, Susan S. Bassett, Ph.D., Gary A. Chase, Ph.D., and Marshal F. Folstein, M.D.; University of Alabama, Birmingham, AL, U01 MH46373, Rodney C.P. Go, Ph.D., and Lindy E. Harrell, M.D.

Pitt ADRC:
This study was supported by the following federal grants: AG027224 (RAS), AG030653 (MIK), AG041718 (MIK), AG066468 (OLL).

MRC Genetic Research for LOAD as part of the Genetic and Environmental Risk in AD Consortium (UK-Cardiff):
Data used in the preparation of this article was obtained from the MRC Genetic Research for LOAD as part of the Genetic and Environmental Risk for Alzheimer’s disease (GERAD) Consortium (Harold et al. 2009). The MRC sample comprised up to 1419 AD cases and 1416 controls. A subset of this sample has been used in this study. Cardiff University was supported by the Wellcome Trust, Medical Research Council (MRC), Alzheimer’s Society Alzheimer’s Research UK (ARUK), Dementias Platform UK, Moondance Foundation and the Welsh Assembly Government.

UK-Cardiff contributors:
Rebecca Sims[1], Denise Harold[1], Amy Gerrish[1], Jade Chapman[1], Valentina Escott-Price[1,2], Nandini Badarinarayan[1], Richard Abraham[1], Paul Hollingworth[1], Nicola Denning[2], Charlene Thomas[1], Alun Meggy[1], Rachel Marshall[1], Lauren Luckcuck[1], Simon Lovestone[3], John Powell[4], Petroula Proitsi[4], Michelle K Lupton[4], Carol Brayne[5], David C. Rubinsztein[5], Michael Gill[6], Brian Lawlor[6], Aoibhinn Lynch[6], Peter Holmans[1], Julie Williams[1,2].

[1]Medical Research Council (MRC) Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neuroscience, School of Medicine, Cardiff University, Cardiff, UK. [2]UKDRI@Cardiff, School of Medicine, Cardiff University, Cardiff, UK. [3]Department of Psychiatry, Medical Sciences Division, University of Oxford, Oxford, UK. [3]Kings College London, Institute of Psychiatry, Department of Neuroscience, De Crespigny Park, Denmark Hill, London, UK. [4]Institute of Public Health, University of Cambridge, Cambridge, UK. [5]Cambridge Institute for Medical Research, University of Cambridge, Cambridge, UK. [6]Mercers Institute for Research on Aging, St. James Hospital and Trinity College, Dublin, Ireland.

Genetic and Environmental Risk in AD Consortium 1 (UK):

Data used in the preparation of this article was obtained from the Genetic and Environmental Risk for Alzheimer's disease (GERAD1) Consortium (Harold et al. 2009). The GERAD1 sample comprised up to 3941 AD cases and 7848 controls. A subset of this sample has been used in this study. Cardiff University was supported by the Wellcome Trust, Medical Research Council (MRC) and/or Alzheimer's Research UK (ARUK) and the Welsh Assembly Government.

GERAD1 contributors:

Denise Harold[1], Amy Gerrish[1], Jade Chapman[1], Valentina Escott-Price[1], Nandini Badarinarayan[1], Richard Abraham[1], Paul Hollingworth[1], Marian Hamshere[1], Jaspreet Singh Pahwa[1], Kimberley Dowzell[1], Amy Williams[1], Nicola Jones[1], Charlene Thomas[1], Alexandra Stretton[1], Angharad Morgan[1], Kate Williams[1], Sarah Taylor[1], Simon Lovestone[2], John Powell[3], Petroula Proitsi[3], Michelle K Lupton[3], Carol Brayne[4], David C. Rubinsztein[5], Michael Gill[6], Brian Lawlor[6], Aoibhinn Lynch[6], Kevin Morgan[7], Kristelle Brown[7], Peter Passmore[8], David Craig[8], Bernadette McGuinness[8], Janet A Johnston[8], Stephen Todd[8], Clive Holmes[9], David Mann[10], A. David Smith[11], Seth Love[12], Patrick G. Kehoe[12], John Hardy[13], Rita Guerreiro[14,15], Andrew Singleton[14], Simon Mead[16], Nick Fox[17], Martin Rossor[17], John Collinge[16], Wolfgang Maier[18], Frank Jessen[18], Reiner Heun[18], Britta Schürmann[18,19], Alfredo Ramirez[18], Tim Becker[20], Christine Herold[20], André Lacour[20], Dmitriy Drichel[20], Hendrik van den Bussche[21], Isabella Heuser[22], Johannes Kornhuber[23], Jens Wiltfang[24], Martin Dichgans[25,26], Lutz Frölich[27], Harald Hampel[28], Michael Hüll[29], Dan Rujescu[30], Alison Goate[31], John S.K. Kauwe[32], Carlos Cruchaga[33], Petra Nowotny[33], John C. Morris[33], Kevin Mayo[33], Gill Livingston[34], Nicholas J. Bass[34], Hugh Gurling[34], Andrew McQuillin[34], Rhian Gwilliam[35], Panagiotis Deloukas[35], Markus M. Nöthen[20], Peter Holmans[1], Michael O'Donovan[1], Michael J.Owen[1].

[1]Medical Research Council (MRC) Centre for Neuropsychiatric Genetics and Genomics, Neurosciences and Mental Health Research Institute, Department of Psychological Medicine and Neurology, School of Medicine, Cardiff University, Cardiff, UK. [2]Department of Psychiatry, Medical Sciences Division, University of Oxford, Oxford, UK. [3]Kings College London, Institute of Psychiatry, Department of Neuroscience, De Crespigny Park, Denmark Hill, London, UK. [4]Institute of Public Health, University of Cambridge, Cambridge, UK. [5]Cambridge Institute for Medical Research, University of Cambridge, Cambridge, UK. [6]Mercers Institute for Research on Aging, St. James Hospital and Trinity College, Dublin, Ireland. [7]Institute of Genetics, Queens Medical Centre, University of Nottingham, UK. [8]Ageing Group, Centre for Public Health, School of Medicine, Dentistry and Biomedical Sciences, Queens University Belfast, UK. [9]Division of Clinical Neurosciences, School of Medicine, University of Southampton, Southampton, UK. [10]Clinical Neuroscience Research Group, Greater Manchester Neurosciences Centre, University of Manchester, Salford, UK. [11]Oxford Project to Investigate Memory and Ageing (OPTIMA), University of Oxford, Department of Pharmacology, Mansfield Road, Oxford, UK. [12]University of Bristol Institute of Clinical Neurosciences, School of Clinical Sciences, Frenchay Hospital, Bristol, UK. [13]Department of Molecular Neuroscience and Reta Lilla Weston Laboratories, Institute of Neurology, UCL, London, UK. [14]Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, Maryland, United States of America. [15]Department of Molecular Neuroscience, Institute of Neurology, University College London, Queen Square, London WC1N 3BG, UK. [16]MRC Prion Unit, Department of Neurodegenerative Disease, UCL Institute of Neurology, London, UK. [17]Dementia Research Centre, Department of Neurodegenerative Diseases, University College London, Institute of Neurology, London, UK. [18]Department of Psychiatry, University of Bonn, Sigmund-Freud-Straβe 25, 53105 Bonn, Germany. [19]Institute for Molecular Psychiatry, University of Bonn, Bonn, Germany. [20]Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany. [21]Institute of Primary Medical Care, University Medical Center Hamburg-Eppendorf, Germany. [22]Department of Psychiatry, Charité Berlin, Germany.

[23]Department of Psychiatry, Friedrich-Alexander-University Erlangen-Nürnberg, Germany. [24]Department of Psychiatry and Psychotherapy, University Medical Center (UMG), Georg-August-University, Göttingen, Germany. [25]Institute for Stroke and Dementia Research, Klinikum der Universität München, Marchioninistr. 15, 81377, Munich, Germany. [26]Department of Neurology, Klinikum der Universität München, Marchioninistr. 15, 81377, Munich, Germany. [27]Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Germany. [28]Institute for Memory and Alzheimer's Disease & INSERM, Sorbonne Universities, Pierre and Marie Curie University, Paris, France; Institute for Brain and Spinal Cord Disorders (ICM), Department of Neurology, Hospital of Pitié-Salpétrière, Paris, France. [29]Centre for Geriatric Medicine and Section of Gerontopsychiatry and Neuropsychology, Medical School, University of Freiburg, Germany. [30]Department of Psychiatry, University of Halle, Halle, Germany. [31]Neuroscience Department, Icahn School of Medicine at Mount Sinai, New York, US. [32]Department of Biology, Brigham Young University, Provo, UT, 84602, USA. [33]Departments of Psychiatry, Neurology and Genetics, Washington University School of Medicine, St Louis, MO 63110, US. [34]Department of Mental Health Sciences, University College London, UK. [35]The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK.

9. REFERENCES

1 Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55**, 997-1004 (1999).

2 Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421-427 (2014).

3 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).

4 Seshadri, S. *et al.* Genome-wide analysis of genetic loci associated with Alzheimer disease. *JAMA* **303**, 1832-1840 (2010).

5 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).

6 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).

7 McKhann, G. M. *et al.* The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers. Dement* **7**, 263-269 (2011).

8 McKeith, I. G. *et al.* Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium. *Neurology* **65**, 1863-1872 (2005).

9 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).

10 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).

11 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).

12 Morris, J. C. *et al.* The Uniform Data Set (UDS): clinical and cognitive variables and descriptive data from Alzheimer Disease Centers. *Alzheimer Dis. Assoc. Disord* **20**, 210-216 (2006).

13 Beekly, D. L. *et al.* The National Alzheimer's Coordinating Center (NACC) database: the Uniform Data Set. *Alzheimer Dis. Assoc. Disord* **21**, 249-258 (2007).

14 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).

15 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).

16 Hyman, B. T. & Trojanowski, J. Q. Consensus recommendations for the postmortem diagnosis of Alzheimer disease from the National Institute on Aging and the Reagan Institute Working Group on diagnostic criteria for the neuropathological assessment of Alzheimer disease. *J. Neuropathol. Exp. Neurol* **56**, 1095-1097 (1997).

17 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).

18 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).

19 Cummings, J. L. *et al.* The neuropsychiatric inventory: comprehensive assessment of psychopathology in dementia. *Neurology* **44**, 2308-2314 (1994).

20 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).

21 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).

22 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).

23 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).

24 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).
25 World Health Organization. *International statistical classification of diseases and related health problems Vol. 1, Tabular list*. 10th revision edn, (World Health Organization, 1992).
26 *DSM-IV: Diagnostic and Statistical Manual of Mental Health Disorders*. Fourth edn, (American Psychiatric Association, 1994).
27 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).
28 Lee, J. H., Cheng, R., Graff-Radford, N., Foroud, T. & Mayeux, R. Analyses of the National Institute on Aging Late-Onset Alzheimer's Disease Family Study: implication of additional loci. *Arch Neurol* **65**, 1518-1526 (2008).
29 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).
30 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).
31 Barral, S. *et al.* Genetic variants associated with susceptibility to psychosis in late-onset Alzheimer's disease families. *Neurobiol. Aging* **36**, 3116-3116 (2015).
32 Blacker, D. *et al.* ApoE-4 and age at onset of Alzheimer's disease: the NIMH genetics initiative. *Neurology* **48**, 139-147 (1997).
33 Overall, J. E. & Gorham, D. R. The brief psychiatric rating scale. *Psychol. Rep* **10**, 799-812 (1962).
34 Bacanu, S. A. *et al.* Heritability of psychosis in Alzheimer disease. *American Journal of Geriatric Psychiatry* **13**, 624-627 (2005).
35 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).
36 DeMichele-Sweet, M. A. *et al.* No association of psychosis in Alzheimer disease with neurodegenerative pathway genes. *Neurobiol Aging* **32**, 555-511 (2011).
37 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).
38 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).
39 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).
40 Hollingworth, P. *et al.* Genome-wide association study of Alzheimer's disease with psychotic symptoms. *Mol. Psychiatry* **17**, 1316-1327 (2012).
41 Harold, D. *et al.* Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat Genet* **41**, 1088-1093 (2009).
42 Laurie, C. C. *et al.* Quality control and quality assurance in genotypic data for genome-wide association studies. *Genet Epidemiol* **34**, 591-602, doi:10.1002/gepi.20516 (2010).
43 Zheng, X. *et al.* A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28**, 3326-3328, doi:10.1093/bioinformatics/bts606 (2012).
44 Gogarten, S. M. *et al.* GWASTools: an R/Bioconductor package for quality control and analysis of genome-wide association studies. *Bioinformatics* **28**, 3329-3331, doi:10.1093/bioinformatics/bts610 (2012).
45 Kamboh, M. I. *et al.* Genome-wide association study of Alzheimer's disease. *Transl. Psychiatry* **2**, e117 (2012).
46 Jun, G. R. *et al.* Transethnic genome-wide scan identifies novel Alzheimer's disease loci. *Alzheimers Dement* **13**, 727-738, doi:10.1016/j.jalz.2016.12.012 (2017).
47 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).

48 Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-2191 (2010).
49 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).
50 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).
51 Gusev, A. *et al.* Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet* **48**, 245-252, doi:10.1038/ng.3506 (2016).
52 Consortium, G. T. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* **45**, 580-585, doi:10.1038/ng.2653 (2013).
53 Purcell, S. M. *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748-752 (2009).
54 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).
55 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).
56 Greenway, M. J. *et al.* A novel candidate region for ALS on chromosome 14q11.2. *Neurology* **63**, 1936-1938, doi:10.1212/01.wnl.0000144344.39103.f6 (2004).
57 Greenway, M. J. *et al.* ANG mutations segregate with familial and 'sporadic' amyotrophic lateral sclerosis. *Nat Genet* **38**, 411-413, doi:10.1038/ng1742 (2006).
58 McLaughlin, R. L. *et al.* Angiogenin levels and ANG genotypes: dysregulation in amyotrophic lateral sclerosis. *PLoS One* **5**, e15402, doi:10.1371/journal.pone.0015402 (2010).
59 Padhi, A. K. *et al.* Insights into the role of ribonuclease 4 polymorphisms in amyotrophic lateral sclerosis. *Journal of biomolecular structure & dynamics* **37**, 116-130, doi:10.1080/07391102.2017.1419147 (2019).
60 Padhi, A. K. & Gomes, J. A molecular dynamics based investigation reveals the role of rare Ribonuclease 4 variants in amyotrophic lateral sclerosis susceptibility. *Mutation research* **813**, 1-12, doi:10.1016/j.mrfmmm.2018.11.002 (2019).
61 Padhi, A. K., Narain, P. & Gomes, J. Rare Angiogenin and Ribonuclease 4 variants associated with amyotrophic lateral sclerosis exhibit loss-of-function: a comprehensive in silico study. *Metabolic brain disease* **34**, 1661-1677, doi:10.1007/s11011-019-00473-6 (2019).
62 Lyons, S. M., Fay, M. M., Akiyama, Y., Anderson, P. J. & Ivanov, P. RNA biology of angiogenin: Current state and perspectives. *RNA Biol* **14**, 171-178, doi:10.1080/15476286.2016.1272746 (2017).
63 Rashad, S., Niizuma, K. & Tominaga, T. tRNA cleavage: a new insight. *Neural Regen Res* **15**, 47-52, doi:10.4103/1673-5374.264447 (2020).
64 Westra, H. J. *et al.* Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet* **45**, 1238-1243, doi:10.1038/ng.2756 (2013).
65 Kirby, J. *et al.* Lack of unique neuropathology in amyotrophic lateral sclerosis associated with p.K54E angiogenin (ANG) mutation. *Neuropathol Appl Neurobiol* **39**, 562-571, doi:10.1111/nan.12007 (2013).
66 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).
67 Mighell, A. J., Smith, N. R., Robinson, P. A. & Markham, A. F. Vertebrate pseudogenes. *FEBS Lett* **468**, 109-114, doi:10.1016/s0014-5793(00)01199-6 (2000).
68 Cheetham, S. W., Faulkner, G. J. & Dinger, M. E. Overcoming challenges and dogmas to understand the functions of pseudogenes. *Nat Rev Genet* **21**, 191-201, doi:10.1038/s41576-019-0196-1 (2020).
69 Groen, J. N., Capraro, D. & Morris, K. V. The emerging role of pseudogene expressed non-coding RNAs in cellular functions. *Int J Biochem Cell Biol* **54**, 350-355, doi:10.1016/j.biocel.2014.05.008 (2014).
70 Bergman, O., Karry, R., Milhem, J. & Ben-Shachar, D. NDUFV2 pseudogene (NDUFV2P1) contributes to mitochondrial complex I deficits in schizophrenia. *Mol Psychiatry* **25**, 805-820, doi:10.1038/s41380-018-0309-9 (2020).
71 Florio, M. *et al.* Evolution and cell-type specificity of human-specific genes preferentially expressed in progenitors of fetal neocortex. *Elife* **7**, doi:10.7554/eLife.32332 (2018).

72 Liu, J. *et al.* Identification of key genes and long non-coding RNA associated ceRNA networks in hepatocellular carcinoma. *PeerJ* **7**, e8021, doi:10.7717/peerj.8021 (2019).
73 Khalil, A. M. *et al.* Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci U S A* **106**, 11667-11672, doi:10.1073/pnas.0904715106 (2009).
74 Tsai, M. C. *et al.* Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **329**, 689-693, doi:10.1126/science.1192002 (2010).
75 Cook, S. E. *et al.* Psychotic symptoms in Alzheimer disease: evidence for subtypes. *Am. J. Geriatr. Psychiatry* **11**, 406-413 (2003).
76 Outeiro, T. F. *et al.* Dementia with Lewy bodies: an update and outlook. *Mol Neurodegener* **14**, 5, doi:10.1186/s13024-019-0306-8 (2019).