

Brain Expression Genome-Wide Association Study (eGWAS) Identifies Human Disease-Associated Variants

Fanggeng Zou^{1,9}, High Seng Chai^{2,9}, Curtis S. Younkin^{1,9}, Mariet Allen^{1,9}, Julia Crook³, V. Shane Pankratz², Minerva M. Carrasquillo¹, Christopher N. Rowley¹, Asha A. Nair², Sumit Middha², Sooraj Maharjan², Thuy Nguyen¹, Li Ma¹, Kimberly G. Malphrus¹, Ryan Palusak¹, Sarah Lincoln¹, Gina Bisceglia¹, Constantin Georgescu¹, Naomi Kouri¹, Christopher P. Kolbert⁴, Jin Jen⁴, Jonathan L. Haines⁵, Richard Mayeux⁶, Margaret A. Pericak-Vance⁷, Lindsay A. Farrer⁸, Gerard D. Schellenberg⁹, Alzheimer's Disease Genetics Consortium¹¹, Ronald C. Petersen¹⁰, Neill R. Graff-Radford¹¹, Dennis W. Dickson¹, Steven G. Younkin¹, Nilüfer Ertekin-Taner^{1,11*}

1 Department of Neuroscience, Mayo Clinic, Jacksonville, Florida, United States of America, **2** Department of Biostatistics, Mayo Clinic, Rochester, Minnesota, United States of America, **3** Department of Biostatistics, Mayo Clinic, Jacksonville, Florida, United States of America, **4** Microarray Core, Mayo Clinic, Rochester, Minnesota, United States of America, **5** Department of Molecular Physiology and Biophysics and Vanderbilt Center for Human Genetics Research, Vanderbilt University, Nashville, Tennessee, United States of America, **6** Gertrude H. Sergievsky Center, Department of Neurology, and Taub Institute on Alzheimer's Disease and the Aging Brain, Columbia University, New York, New York, United States of America, **7** The John P. Hussman Institute for Human Genomics and Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami, Miami, Florida, United States of America, **8** Departments of Biostatistics, Medicine (Genetics Program), Ophthalmology, Neurology, and Epidemiology, Boston University, Boston, Massachusetts, United States of America, **9** Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, United States of America, **10** Department of Neurology, Mayo Clinic, Rochester, Minnesota, United States of America, **11** Department of Neurology, Mayo Clinic, Jacksonville, Florida, United States of America

Abstract

Genetic variants that modify brain gene expression may also influence risk for human diseases. We measured expression levels of 24,526 transcripts in brain samples from the cerebellum and temporal cortex of autopsied subjects with Alzheimer's disease (AD, cerebellar $n = 197$, temporal cortex $n = 202$) and with other brain pathologies (non-AD, cerebellar $n = 177$, temporal cortex $n = 197$). We conducted an expression genome-wide association study (eGWAS) using 213,528 *cis*SNPs within ± 100 kb of the tested transcripts. We identified 2,980 cerebellar *cis*SNP/transcript level associations (2,596 unique *cis*SNPs) significant in both ADs and non-ADs ($q < 0.05$, $p = 7.70 \times 10^{-5}$ – 1.67×10^{-82}). Of these, 2,089 were also significant in the temporal cortex ($p = 1.85 \times 10^{-5}$ – 1.70×10^{-141}). The top cerebellar *cis*SNPs had 2.4-fold enrichment for human disease-associated variants ($p < 10^{-6}$). We identified novel *cis*SNP/transcript associations for human disease-associated variants, including progressive supranuclear palsy *SLC12A2*/rs11568563, Parkinson's disease (PD) *MMRN1*/rs6532197, Paget's disease *OPTN*/rs1561570; and we confirmed others, including PD *MAPT*/rs242557, systemic lupus erythematosus and ulcerative colitis *IRF5*/rs4728142, and type 1 diabetes mellitus *RPS26*/rs1701704. In our eGWAS, there was 2.9–3.3 fold enrichment ($p < 10^{-6}$) of significant *cis*SNPs with suggestive AD-risk association ($p < 10^{-3}$) in the Alzheimer's Disease Genetics Consortium GWAS. These results demonstrate the significant contributions of genetic factors to human brain gene expression, which are reliably detected across different brain regions and pathologies. The significant enrichment of brain *cis*SNPs among disease-associated variants advocates gene expression changes as a mechanism for many central nervous system (CNS) and non-CNS diseases. Combined assessment of expression and disease GWAS may provide complementary information in discovery of human disease variants with functional implications. Our findings have implications for the design and interpretation of eGWAS in general and the use of brain expression quantitative trait loci in the study of human disease genetics.

Citation: Zou F, Chai HS, Younkin CS, Allen M, Crook J, et al. (2012) Brain Expression Genome-Wide Association Study (eGWAS) Identifies Human Disease-Associated Variants. *PLoS Genet* 8(6): e1002707. doi:10.1371/journal.pgen.1002707

Editor: Greg Gibson, Georgia Institute of Technology, United States of America

Received: October 14, 2011; **Accepted:** March 27, 2012; **Published:** June 7, 2012

Copyright: © 2012 Zou et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: Support for this research was provided by the National Institutes of Health grants: National Institute on Aging (R01 032990 to NET and R01 AG018023 to NR Graff-Radford and SG Younkin); Mayo Alzheimer's Disease Research Center: (P50 AG016574 to RC Petersen, DW Dickson, NR Graff-Radford, SG Younkin, and N Ertekin-Taner); Mayo Alzheimer's Disease Patient Registry: (U01 AG006576 to RC Petersen); National Institute on Aging (AG025711, AG017216, AG003949 to DW Dickson). This project was also generously supported by the Robert and Clarice Smith and Abigail Van Buren Alzheimer's Disease Research Program (to RC Petersen, DW Dickson, NR Graff-Radford, and SG Younkin), and by the Palumbo Professorship in Alzheimer's Disease Research (to SG Younkin). N Ertekin-Taner is the recipient of National Institutes of Health (KL2 RR024151) and Siragusa Foundation grants. ADGC Authors: The National Institutes of Health, National Institute on Aging (NIH-NIA) supported this work through the following grants: ADGC, U01 AG032984, RC2 AG036528; NACC, U01 AG016976; NCRAD, U24 AG021886; NIA LOAD, U24 AG026395, U24 AG026390; Banner Sun Health Research Institute, U24 NS072026, P30 AG19610 Arizona Department of Health Services contract 211002, Arizona Biomedical Research Commission contracts 4001, 0011, 05-901 and 100, Michael J. Fox Foundation for Parkinson's Research; Boston University, P30 AG013846, R01 HG02213, K24 AG027841, U01 AG10483, R01 CA129769, R01 MH080295, R01 AG009029, R01 AG017173, R01 AG025259; Columbia University, P50 AG008702, R37 AG015473; Duke University, P30 AG028377; Emory University, AG025688; Group Health U01 AG06781; Indiana University, P30 AG10133; Johns Hopkins University, P50 AG005146, R01 AG020688; Massachusetts General Hospital, P50 AG005134; Mayo Clinic, P50 AG016574; Mount Sinai School of Medicine, P50 AG005138, P01 AG002219; New York University, P30 AG08051, MO1RR00096, and UL1 RR029893; Northwestern University, P30 AG013854; Oregon Health and Science University, P30 AG008017, R01 AG026916; Rush University, P30 AG010161, R01 AG019085, R01 AG15819, R01 AG17917, R01 AG30146; University of Alabama at Birmingham, P50 AG016582, UL1RR02777; University of Arizona/TGEN, P30 AG019610, R01 AG031581, R01 NS059873; University of California, Davis, P30 AG010129; University of California, Irvine, P50 AG016573, P50, P50 AG016575, P50 AG016576, P50 AG016577; University of California, Los Angeles, P50

AG016570; University of California, San Diego, P50 AG005131; University of California, San Francisco, P50 AG023501, P01 AG019724; University of Kentucky, P30 AG028383; University of Michigan, P50 AG008671; University of Pennsylvania, P30 AG010124, AG17586; University of Pittsburgh, P50 AG005133, AG030653; University of Southern California, P50 AG005142; University of Texas Southwestern, P30 AG012300; University of Miami, R01 AG027944, AG010491, AG027944, AG021547, AG019757; University of Washington, P50 AG005136, UO1 HG004610; Vanderbilt University, R01 AG019085; and Washington University, P50 AG005681, P01 AG03991. ADNI Funding for ADNI is through the Northern California Institute for Research and Education by grants from Abbott, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly, Medpace, Merck, Novartis AG, Pfizer Inc, F. Hoffman-La Roche, Schering-Plough, Synarc, Alzheimer's Association, Alzheimer's Drug Discovery Foundation, the Dana Foundation, and by the National Institute of Biomedical Imaging and Bioengineering and NIA grants U01 AG024904, RC2 AG036535, K01 AG030514. We thank Creighton Phelps, Marcelle Morrison-Bogorad, and Marilyn Miller from NIA who are ex-officio ADGC members. Support was also from the Alzheimer's Association (LAF, IIRG-08-89720; MP-V, IIRG-05-14147) and the Veterans Affairs Administration. P.S.G.-H. is supported by Wellcome Trust, Howard Hughes Medical Institute, and the Canadian Institute of Health Research. Additional acknowledgements for TGen data: Many data and biomaterials were collected from several National Institute on Aging (NIA) and National Alzheimer's Coordinating Center (NACC, grant #U01 AG016976) funded sites. Amanda J. Myers, PhD (University of Miami, Department of Psychiatry) and John A. Hardy, PhD (Reta Lila Weston Institute, University College London) collected and prepared the series. Marcelle Morrison-Bogorad, PhD., Tony Phelps, PhD, and Walter Kukull PhD are thanked for helping to co-ordinate this collection. The directors, pathologist and technicians involved include: National Institute on Aging: Ruth Seemann, John Hopkins Alzheimer's Disease Research Center (NIA grant # AG05146); Juan C. Troncoso, MD, Dr. Olga Pletnikova, University of California, Los Angeles (NIA grant # P50 AG16570); Harry Vinters, MD, Justine Pomakian, The Kathleen Price Bryan Brain Bank, Duke University Medical Center (NIA grant #AG05128, NINDS grant # NS39764, NIMH MH60451 also funded by Glaxo Smith Kline); Christine Hulette, MD, Director, John F. Ervin, Stanford University; Dikran Horoupian, MD, Ahmad Salehi, MD, PhD, New York Brain Bank, Taub Institute, Columbia University (NYBB); Jean Paul Vonsattel, MD, Katerina Mancevska, Massachusetts Alzheimer's Disease Research Center (P50 AG005134); E. Tessa Hedley-Whyte, MD, MP Frosch, MD, Karlotta Fitch, University of Michigan (NIH grant P50-AG08671); Dr. Roger Albin, Lisa Bain, Eszter Gombosi, University of Kentucky (NIH #AG05144); William Markesbery, MD, Sonya Anderson, Mayo Clinic, Jacksonville; Dennis W. Dickson, MD, Natalie Thomas, University of Southern California; Carol A. Miller, MD, Jenny Tang, M.S., Dimitri Diaz, Washington University, St Louis Alzheimer's Disease Research Center (NIH #P50AG05681); Dan McKeel, MD, John C. Morris, MD, Eugene Johnson, Jr., PhD, Virginia Buckles, PhD, Deborah Carter, University of Washington, Seattle (NIH #P50 AG05136); Thomas Montine, MD, PhD, Aimee Schantz, MEd., University of Pennsylvania School of Medicine, Alzheimer's Disease Research Center; John Q. Trojanowski, MD, Virginia M Lee, MD, Vivianna Van Deerlin, MD, Terry Schuck, Boston University Alzheimer's Disease Research Center (NIH grant P30-AG13846); Ann C. McKee, Carol Kubilus Sun Health Research Institute, Arizona (NIA #P30 AG19610); Joseph Rogers, PhD, Thomas G. Beach, MD, PhD, Lucia I. Sue Emory University; Bruce H. Wainer, MD, PhD, Marla Gearing, PhD, University of Texas, Southwestern Medical School; Charles L. White, III, M.D., Roger Rosenberg, Marilyn Howell, Joan Reisch, University of California, Davis; William Ellis, MD, Mary Ann Jarvis, Rush University Medical Center, Rush Alzheimer's Disease Research Center (NIH #AG10161); David A. Bennett, M.D. Julie A. Schneider, MD, MS, Karen Skish, MS, PA (ASCP)MT, Wayne T Longman, University of Miami/NPF Brain Endowment Bank; Deborah C. Mash, MD, Margaret J Basile, Mitsuko Tanaka Oregon Health and Science University; Randy Wotljer, PhD. Additional tissues include samples from the following sites: Newcastle Brain Tissue Resource (funding via the Medical Research Council, local NHS trusts and Newcastle University); C. M. Morris, MD, Ian G. McKeith, Robert H. Perry MRC London Brain Bank for Neurodegenerative Diseases (funding via the Medical Research Council); Simon Lovestone, MD PhD, Safa Al-Sarraj, MD, Claire Troakes, South West Dementia Brain Bank (funding via numerous sources including the Higher Education Funding Council for England (HEFCE), Alzheimer's Research Trust (ART), BRACE as well as North Bristol NHS Trust Research and Innovation Department and DeNDRoN); Seth Love, MD, Patrick Kehoe, PhD, Laura Palmer, The Netherlands Brain Bank (funding via numerous sources including Stichting MS Research, Brain Net Europe, Hersenstichting Nederland Breinbrekend Werk, Internationale Parkinson Fonds, Internationale Stichting Alzheimer Onderzoek); Inge Huitinga, MD, Marleen Rademaker, Michiel Kooreman, Institut de Neuropatologia, Servei Anatomia Patologica, Universitat de Barcelona; Isidre Ferrer Abizanda, MD, PhD, Susana Casas Boluda.

Competing Interests: NR Graff-Radford has served as a consultant to Codman and received grant support from Elan Pharmaceutical Research, Pfizer Pharmaceuticals, Medivation, and Forrest. RC Petersen has been a consultant to GE Healthcare and Elan Pharmaceuticals, has served on a data safety monitoring committee for Pfizer and Janssen Alzheimer Immunotherapy, and has provided a CME lecture for Novartis. ADGC Authors: T.D.B. received licensing fees from and is on the speaker's bureau of Athena Diagnostics. M.R.F. receives research funding from BristolMyersSquibb Company, Danone Research, Elan Pharmaceuticals, Eli Lilly, Novartis Pharmaceuticals, OctaPharma AG, Pfizer, and Sonexa Therapeutics; receives honoraria as scientific consultant from Accera, Astellas Pharma US Inc., Baxter, Bayer Pharmaceuticals Corporation, BristolMyersSquibb, Eisai Medical Research, GE Healthcare, Medavante, Medivation, Merck, Novartis Pharmaceuticals, Pfizer, Prana Biotechnology, QR Pharma, The sanofi-aventis Group, and Toyama Chemical; and is speaker for Eisai Medical Research, Forest Laboratories, Pfizer, and Novartis Pharmaceuticals. A.M.G. has research funding from AstraZeneca, Pfizer and Genentech, and has received remuneration for giving talks at Pfizer and Genentech. R.C.P. is on the Safety Monitoring Committee of Pfizer (Wyeth) and a consultant to the Safety Monitoring Committee at Janssen Alzheimer's Immunotherapy Program (Elan), to Elan Pharmaceuticals, and to GE Healthcare. R.E.T. is a consultant to Eisai, Japan, in the area of Alzheimer's genetics and a shareholder in, and consultant to, Pathway Genomics, San Diego, CA.

* E-mail: taner.nilufer@mayo.edu

† These authors contributed equally to this work.

‡ Membership of the Alzheimer's Disease Genetics Consortium (ADGC) is provided in the Acknowledgments.

Introduction

Expression quantitative trait loci (eQTL) are genomic loci that influence levels of gene transcripts and can be mapped by genetic linkage in families or eGWAS in unrelated populations [1]. eQTLs are distinct from other complex trait loci, because they directly identify the target gene, since the transcript trait is a reflection of the mRNA level from a single gene. Furthermore, eQTLs imply regulation of gene expression as the mechanism of action for the underlying variants. Recently, few studies identified an enrichment of eQTLs from lymphocytes [2] and lymphoblasts [3] amongst human complex disease and trait GWAS loci, suggesting that eQTLs may be useful in mapping human disease-associated variants.

Most human eQTL mapping studies to date assessed immortalized lymphoblastoid cell lines [4,5,6,7,8,9,10,11,12] and family-based samples from the CEPH [4,5,6,7,8,13] (Centre d'Etude du polymorphisme humain) or HapMap [10,11,14,15] repositories. Multiple other small and large scale eQTL studies investigated other tissues and populations including lymphocytes [16], monocytes [17],

T-cells [18], fibroblasts [18], skin [19], subcutaneous and omental adipose tissue [20,21], bone [22], liver [23] and brain [24,25].

Despite the assumption that brain eQTLs would also influence human diseases and traits, there are no systematic gene mapping studies for human diseases that utilize brain gene expression phenotypes. Furthermore, the brain region most relevant for such studies and the influence of brain pathology on eQTL mapping studies are largely unknown. To address these issues, we performed an eQTL using cerebellar tissue from 197 subjects with Alzheimer's disease (AD) neuropathology and 177 with other pathologies (non-AD). We validated the results in a different brain region using temporal cortex samples from 202 ADs and 197 non-ADs (Supplementary Tables 1 and 2 in Dataset S1), 85% of whom overlapped with the cerebellar group. We evaluated significant *cis*SNPs from our study for association with human diseases/traits using a GWAS catalog [26]. We also assessed our significant eGWAS *cis*SNPs for association with two central nervous system (CNS) diseases, progressive supranuclear palsy (PSP) [27] and AD risk [28], using two recent GWAS for these diseases.

Author Summary

Genetic variants that regulate gene expression levels can also influence human disease risk. Discovery of genomic loci that alter brain gene expression levels (brain expression quantitative trait loci = eQTLs) can be instrumental in the identification of genetic risk underlying both central nervous system (CNS) and non-CNS diseases. To systematically assess the role of brain eQTLs in human disease and to evaluate the influence of brain region and pathology in eQTL mapping, we performed an expression genome-wide association study (eGWAS) in 773 brain samples from the cerebellum and temporal cortex of ~200 autopsied subjects with Alzheimer's disease (AD) and ~200 with other brain pathologies (non-AD). We identified ~3,000 significant associations between *cis*SNPs near ~700 genes and their cerebellar transcript levels, which replicate in ADs and non-ADs. More than 2,000 of these associations were reproducible in the temporal cortex. The top *cis*SNPs are enriched for both CNS and non-CNS disease-associated variants. We identified novel and confirmed previous *cis*SNP/transcript associations for many disease loci, suggesting gene expression regulation as their mechanism of action. These findings demonstrate the reproducibility of the eQTL approach across different brain regions and pathologies, and advocate the combined use of gene expression and disease GWAS for identification and functional characterization of human disease-associated variants.

Our results demonstrate the power of the brain eQTL approach in the identification and characterization of many human CNS and non-CNS disease-associated variants. This study also highlights the remarkable reproducibility of human eQTLs across different brain regions and pathologies, which has implications for the design of eGWAS in general. Combined assessment of eQTLs and disease risk loci can be instrumental in mapping disease genes with regulatory variants.

Results

Brain eGWAS

Levels of 24,526 transcripts for 18,401 genes were measured in 773 brain samples from the cerebellum and temporal cortex of ~200 ADs and ~200 non-ADs, using WG-DASL assays. Nearly 70% of all probes could be detected in >75% of the samples tested. All autopsied subjects were genotyped for 313,330 single nucleotide polymorphisms (SNPs) from Illumina HumanHap300-Duo Genotyping BeadChips, as part of the Mayo AD GWAS [29]. An eGWAS testing association of transcript levels with *cis*SNPs was performed using multivariable linear regression correcting for *APOE* ϵ 4 dosage, age at death, gender and multiple technical variables. False discovery rate (FDR)-based *q* values [30] (*q*) were used for corrections of multiple testing.

To achieve internal replication, we first analyzed the ADs and non-ADs separately. In our cerebellar eGWAS, at $q < 0.05$, there were 5,271 significant *cis*SNP/transcript associations (1,156 unique genes) in the AD, 4,450 (1,022 unique genes) in the non-AD and 10,281 (1,875 unique genes) in the combined datasets. Q-Q plots suggested a clear excess of significant results (Figure 1, Figure S1a–S1d). 2,980 *cis*SNP/transcript associations (2,596 unique *cis*SNPs, 686 unique genes) were significant at $q < 0.05$ in both ADs and non-ADs (Table 1, Supplementary Table 3 in Dataset S1, Figure S2). The direction and magnitude of associations in both groups demonstrate remarkable similarities (Pearson's correlation coefficient = 0.98, $p < 0.0001$). The box plots depicted for some of these top associations (Figure S3a–S3c) demonstrate this replication in ADs and non-ADs. Most associations have an additive or dominant pattern with respect to the minor allele.

To assess the genetic component contributing to gene expression variability, we estimated intraclass correlation coefficients (ICC) [31] in the 15 samples measured in replicate on 5–6 different plates and 2–3 different days. Between-subject variance accounted for a median of 60% of total probe expression variance (Supplementary Table 4 in Dataset S1; Figure S4). The 746 probes for the top 2,980 cerebellar *cis*SNP associations had higher between-subject variance (median = 78%).

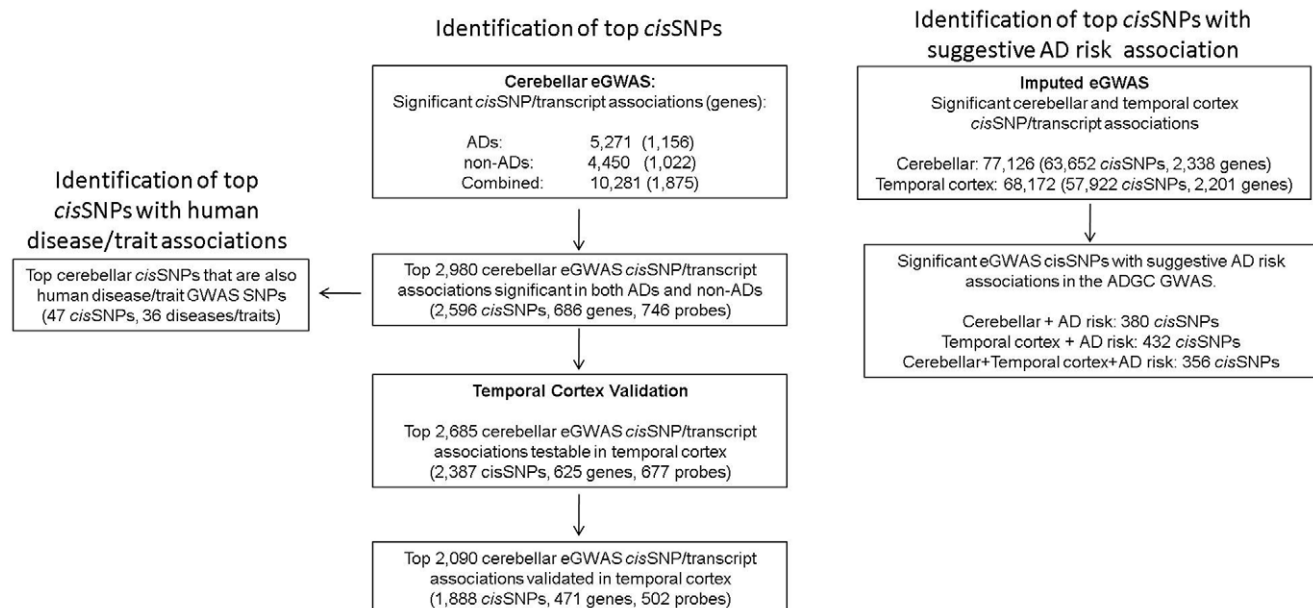


Figure 1. Summary of brain eGWAS and human disease associations.

doi:10.1371/journal.pgen.1002707.g001

Using multivariable linear regression, we next estimated the percent variation in cerebellar probe expression levels due to the “best” *cis*SNP for each transcript after accounting for technical and biological covariates. We found that the “best” *cis*SNP explained a median of ~3% of the expression variation. For the top 746 probes, the “best” *cis*SNPs accounted for a median of 18% of the expression variance (Table 2, Supplementary Table 5 in Dataset S1).

The top 2,980 cerebellar eGWAS associations were followed up in the temporal cortex validation study. We found that 2,685 top cerebellar *cis*SNP/transcript associations could be tested in the temporal cortex (2,387 unique *cis*SNPs, 677 unique probes and 625 unique genes) (Figure 1, Table 3, Supplementary Table 6 in Dataset S1). A total of 2,089 of these (1,888 unique *cis*SNPs, 502 unique probes and 471 unique genes) were significant after study-wide Bonferroni corrections, many of which had effect sizes showing remarkable similarity to those from the cerebellar eGWAS (Pearson’s correlation coefficient = 0.94, $p < 0.0001$).

The top cerebellar eGWAS results were also compared to published liver [23] and brain [24,25] eGWAS and overlap was identified for 4–11% of the top transcripts from these published studies (Text S1). Using HapMap2 genotypes, all transcripts and association threshold $p < 1.0E-4$ in our eGWAS, we determined that 24–32% of the top transcripts from the published eGWAS overlapped with ours.

We used the cerebellar eGWAS as the discovery analysis and the temporal cortex eGWAS as the validation; since our goal is to identify significant *cis*SNP associations while minimizing any confounding factors due to pathology and given the fact that half of our subjects had pathologic AD, in which cerebellum is relatively unaffected whereas temporal cortex is one of the first affected brain regions. Nonetheless, we have also used temporal cortex as the discovery set and cerebellum as the validation, with remarkably similar results (Text S1, Supplementary Tables 7 and 8 in Dataset S1).

Enrichment of brain *cis*SNPs among human disease-associated SNPs

To examine whether the brain eGWAS approach identified variants implicated in human diseases/traits, we linked the 2,596 top cerebellar eGWAS *cis*SNPs to the “Catalog of Published GWAS” [26], which compiles weekly search results from all published GWAS of $\geq 100,000$ SNPs where associations of $p \leq 1.0E-05$ are reported. We identified 47 *cis*SNPs that were also associated with 36 diseases/traits (Table 4, Supplementary Table 9 in Dataset S1). This represents a 2.4-fold enrichment of significant cerebellar *cis*SNPs amongst disease/trait associated SNPs, which is significant ($p < 10^{-6}$) based on simulations adjusted for minor allele frequencies [3] (Text S1).

Among the 36 diseases/traits associating with top cerebellar *cis*SNPs were central nervous system (CNS)-related conditions including Parkinson’s disease (PD), Moyamoya disease, cognitive performance and attention-deficit hyperactivity disorder (ADHD). We both identified novel *cis*SNP/transcript associations and confirmed some previously reported ones. We found novel associations between rs6532197, which confers increased risk of PD [32], and higher brain levels of *MMRNI* (cerebellar eGWAS $p = p_{\text{Cer}} = 4.86 \times 10^{-12}$; temporal cortex eGWAS $p = p_{\text{TCx}} = 4.57 \times 10^{-9}$). *MMRNI* encodes for multimerin and was found to be in a region of duplication/triplication with *SNCA* (encoding α -synuclein), a well-established risk gene in PD [33]. We found no significant *cis*SNP/*SNCA* level associations. These results suggest that *MMRNI* may deserve further investigations as an additional PD risk gene.

Another example of a *cis*SNP which associates with human disease risk is rs8070723, the minor allele of which is associated with reduced risk of PD [32] and reduced brain *MAPT* levels ($p_{\text{Cer}} = 3.36 \times 10^{-7}$ – 7.02×10^{-69} ; $p_{\text{TCx}} = 9.03 \times 10^{-4}$ – 8.61×10^{-44}). Rs11012 minor allele, which confers increased risk of PD [34], showed association with lower brain *LRRC37A4* levels ($p_{\text{Cer}} = 1.69 \times 10^{-33}$; $p_{\text{TCx}} = 3.378E^{-20}$). *MAPT* region variants were previously identified to associate with brain levels of *MAPT* and *LRRC37A4* in neurologically normal subjects [27,32], in a *MAPT* haplotype H1/H2-dependent manner [27]. Indeed, rs8070723 is in tight linkage disequilibrium with rs1052553 ($r^2 = 0.95$, $D' = 0.97$), the major allele of which marks the *MAPT*-H1 haplotype and associates with higher brain *MAPT* levels [24].

Many top cerebellar *cis*SNPs also associate with non-CNS diseases/traits (Supplementary Table 9 in Dataset S1). *IRF5* *cis*SNP rs4728142 is associated with both cerebellar *IRF5* levels and risk of systemic lupus erythematosus (SLE) [35]. Previously, *IRF* variants were shown to influence *IRF* splicing and expression as well as SLE risk [36,37]. Interestingly, rs4728142 is also associated with ulcerative colitis (UC) [38] given both *IRF5* and *TNPO3* are reported as candidate genes. Given its influence on *IRF5*, but not *TNPO3* expression levels, rs4728142 most likely marks *IRF5*, but not *TNPO3* as the candidate UC risk gene.

Our approach to identify human disease-associated SNPs amongst the 2,596 top cerebellar eGWAS *cis*SNPs may be overly conservative, given our selection criteria to only include transcripts that are detectable in >75% of the subjects and only those *cis*SNPs that are significant in both independent cohorts (ADs and non-ADs). Furthermore, given that our eGWAS genotyping platform consisted of ~300 K SNPs, it is plausible that transcript associations with SNPs from the “Catalog of Published GWAS” [26] may be missed if those SNPs did not exist in our platform. To address these issues, we repeated the cerebellar and temporal cortex eGWAS, without restrictions for transcript detection rates and using genotypes imputed to HapMap2 (>2 million SNPs). Comparison of the eGWAS associations with $p < 1.0E-4$ to the “Catalog of Published GWAS” identified 392 unique cerebellar *cis*SNPs that also associate with 189 human diseases/traits; and 339 such temporal cortex *cis*SNPs associating with 167 diseases/traits (Text S1, Supplementary Tables 10 and 11 in Dataset S1). Amongst the associations identified by this less stringent approach were those for brain levels of *CLU* [39,40], *CR1* [40] and *GAB2* [41] which were identified as risk loci in GWAS of Alzheimer’s disease.

We also performed comparisons of the eGWAS results from the ADs and non-ADs separately to determine whether there were any results unique to these diagnostic groups (Text S1, Supplementary Tables 12, 13, 14, 15 in Dataset S1). Although 13–25% of the disease/trait associations were with *cis*SNPs that were unique to ADs or non-ADs, all but a few of these could also be identified in the combined analysis of all subjects. There were only 2–7 human diseases/traits with *cis*SNP associations that were detectable just in ADs or non-ADs, but not the combined group.

Of these unique *cis*SNP, those that associate with cerebellar levels of *C9orf72* in non-ADs are interesting, as these variants were previously identified in GWAS of amyotrophic lateral sclerosis (ALS), where *C9orf72* was one of the candidate genes at the disease locus [42,43]. This gene was recently identified as the most common cause of familial ALS, with a repeat expansion leading to loss of an alternatively spliced transcript [44,45]. These results further support the utility of the combined eGWAS and disease GWAS approaches in the potential identification of disease genes with modified transcript levels as the plausible disease mechanism.

Table 1. Examples of top cerebellar eGWAS cisSNP/transcript associations.

CHR	SNP	PROBE	SYMBOL	NON-AD				AD				ALL			
				P	P _{Bonf}	Q	Beta	P	P _{Bonf}	Q	Beta	P	P _{Bonf}	Q	Beta
11	rs11552421	ILMN_1651745	TMEM25	1.44E-68	6.37E-63	1.14E-58	2.69	1.67E-82	7.41E-77	7.26E-71	2.65	3.50E-153	1.55E-147	3.10E-138	2.67
5	rs3776455	ILMN_1718932	MTRR	6.70E-64	2.97E-58	1.90E-54	1.30	4.06E-70	1.80E-64	7.41E-60	1.12	6.37E-133	2.83E-127	7.34E-120	1.20
6	rs1096699	ILMN_1694711	MAD2L1BP	6.09E-56	2.70E-50	4.86E-47	1.31	6.54E-60	2.90E-54	5.35E-50	1.22	1.41E-116	6.25E-111	5.44E-104	1.26
12	rs10843881	ILMN_2345908	DDX11	7.87E-54	3.49E-48	3.15E-45	-1.06	5.59E-58	2.48E-52	9.92E-49	-1.06	1.48E-112	6.57E-107	2.39E-100	-1.06
12	rs2708389	ILMN_1730477	TAS2R43	9.30E-59	4.12E-53	1.29E-49	1.13	5.26E-53	2.34E-47	2.23E-44	1.01	2.80E-112	1.24E-106	2.39E-100	1.07
12	rs1971762	ILMN_1807798	ATP5G2	1.52E-44	6.73E-39	4.49E-37	0.96	4.76E-59	2.11E-53	2.11E-49	1.04	1.10E-104	4.87E-99	1.64E-93	1.01
21	rs2838859	ILMN_2376667	POFUT2	1.19E-48	5.28E-43	8.06E-41	-0.81	1.17E-53	5.20E-48	6.24E-45	-0.89	2.86E-103	1.27E-97	2.73E-92	-0.86
8	rs3802266	ILMN_2184966	ZHX2	8.18E-44	3.63E-38	2.17E-36	-1.89	2.13E-56	9.47E-51	2.64E-47	-1.99	4.35E-101	1.93E-95	3.06E-90	-1.94
19	rs260462	ILMN_2052079	ZNF544	7.19E-54	3.19E-48	3.15E-45	-1.21	1.75E-47	7.79E-42	2.60E-39	-1.12	8.26E-101	3.67E-95	4.38E-90	-1.16
17	rs2525574	ILMN_2369018	EV12A	2.04E-48	9.05E-43	1.08E-40	1.81	8.82E-49	3.92E-43	1.61E-40	1.79	1.52E-99	6.73E-94	4.88E-89	1.81
7	rs6953367	ILMN_1723984	PIR1B	1.58E-52	6.99E-47	4.02E-44	1.21	8.75E-49	3.89E-43	1.66E-40	0.99	4.80E-97	2.13E-91	1.23E-86	1.08
7	rs7313	ILMN_2400759	CPVL	2.40E-42	1.06E-36	5.15E-35	-0.88	2.13E-49	9.48E-44	5.07E-41	-1.01	5.00E-92	2.22E-86	5.96E-82	-0.96
11	rs7124057	ILMN_2262288	EEF1G	6.28E-50	2.78E-44	6.10E-42	-1.03	1.54E-41	6.83E-36	4.95E-34	-0.74	2.64E-91	1.17E-85	2.67E-81	-0.90
10	rs2182513	ILMN_1689177	PPAPDC1A	4.60E-35	2.04E-29	2.09E-28	0.87	1.25E-52	5.55E-47	4.30E-44	0.82	1.08E-87	4.79E-82	6.93E-78	0.84
6	rs2975033	ILMN_2130441	HLA-H	8.19E-42	3.63E-36	1.60E-34	1.53	8.13E-42	3.61E-36	2.83E-34	1.39	3.53E-84	1.57E-78	1.73E-74	1.45
5	rs11738432	ILMN_2093720	THG1L	5.52E-45	2.45E-39	1.84E-37	-0.94	3.78E-39	1.68E-33	9.18E-32	-0.81	8.06E-84	3.58E-78	3.49E-74	-0.86
15	rs12912744	ILMN_1691772	ZSCAN29	1.25E-35	5.53E-30	7.32E-29	-0.52	4.47E-47	1.99E-41	5.78E-39	-0.59	3.47E-83	1.54E-77	1.33E-73	-0.56
22	rs136564	ILMN_1809147	FAM118A	1.79E-35	7.93E-30	8.98E-29	0.88	1.58E-44	7.02E-39	7.45E-37	0.82	8.44E-81	3.74E-75	2.57E-71	0.85
17	rs8070454	ILMN_2388272	MED24	1.77E-33	7.82E-28	5.89E-27	-1.98	7.10E-46	3.15E-40	7.11E-38	-2.05	3.51E-80	1.56E-74	9.58E-71	-2.03
20	rs2387976	ILMN_2064132	NANP	3.88E-38	1.72E-32	3.69E-31	-0.57	1.42E-40	6.29E-35	3.68E-33	-0.51	9.58E-79	4.25E-73	1.93E-69	-0.54
7	rs10239340	ILMN_2312606	IRF5	5.84E-33	2.59E-27	1.66E-26	0.81	7.47E-44	3.32E-38	3.05E-36	0.82	1.52E-76	6.73E-71	2.31E-67	0.81
6	rs2395943	ILMN_1683279	PEX6	7.06E-35	3.13E-29	2.93E-28	-0.39	1.90E-41	8.43E-36	5.68E-34	-0.40	3.63E-76	1.61E-70	5.06E-67	-0.39
13	rs1046028	ILMN_2390162	PHF11	1.08E-34	4.79E-29	4.28E-28	0.37	1.26E-40	5.61E-35	3.52E-33	0.33	1.07E-75	4.73E-70	1.36E-66	0.35
17	rs1981997	ILMN_1710903	MAPT	1.89E-37	8.38E-32	1.48E-30	-0.51	2.55E-33	1.13E-27	1.20E-26	-0.44	4.16E-71	1.85E-65	3.88E-62	-0.48
9	rs2240913	ILMN_1811048	GPR107	4.16E-34	1.84E-28	1.51E-27	0.62	3.74E-36	1.66E-30	4.81E-29	0.69	2.76E-69	1.23E-63	2.39E-60	0.66
16	rs11649236	ILMN_2201966	N4BP1	1.99E-39	8.82E-34	2.59E-32	-1.70	2.71E-32	1.20E-26	1.10E-25	-1.48	5.69E-69	2.53E-63	4.37E-60	-1.57
10	rs7909832	ILMN_1795336	PTER	3.28E-33	1.45E-27	1.01E-26	-0.70	6.43E-35	2.86E-29	5.23E-28	-0.81	1.70E-68	7.54E-63	1.04E-59	-0.75
5	rs11134054	ILMN_1789419	EXOC3	1.45E-35	6.43E-30	8.07E-29	-0.54	1.07E-33	4.77E-28	5.40E-27	-0.59	4.40E-68	1.95E-62	2.52E-59	-0.57
19	rs10405576	ILMN_2074477	GPR4	2.49E-32	1.10E-26	6.57E-26	-1.35	4.86E-34	2.16E-28	3.04E-27	-1.27	5.76E-68	2.56E-62	3.09E-59	-1.32
6	rs1988374	ILMN_2075334	HIST1H4C	1.98E-29	8.78E-24	2.23E-23	-0.42	8.82E-38	3.92E-32	1.60E-30	-0.44	4.76E-67	2.11E-61	2.05E-58	-0.43
21	rs6517526	ILMN_2296011	BRWD1	3.00E-30	1.33E-24	4.50E-24	-0.71	6.10E-37	2.71E-31	9.69E-30	-0.88	6.31E-67	2.80E-61	2.51E-58	-0.80
22	rs737950	ILMN_1697286	SF3A1	2.77E-30	1.23E-24	4.36E-24	0.42	3.78E-34	1.68E-28	2.52E-27	0.45	2.05E-66	9.11E-61	7.72E-58	0.44
1	rs1763601	ILMN_2364072	CLCNKA	1.06E-36	4.69E-31	7.35E-30	1.30	8.25E-31	3.66E-25	1.99E-24	1.14	3.49E-66	1.55E-60	1.24E-57	1.22
12	rs10878255	ILMN_2183938	LEMD3	1.25E-29	5.54E-24	1.55E-23	-0.68	9.20E-36	4.09E-30	9.31E-29	-0.56	2.61E-65	1.16E-59	7.91E-57	-0.62
19	rs4806187	ILMN_1655637	UPK1A	6.89E-27	3.05E-21	4.54E-21	-0.99	5.73E-38	2.55E-32	1.15E-30	-1.05	2.82E-65	1.25E-59	8.11E-57	-1.04
11	rs2848630	ILMN_1765332	TIMM10	4.66E-36	2.07E-30	3.05E-29	-0.62	1.18E-28	5.23E-23	1.84E-22	-0.42	3.40E-65	1.51E-59	9.32E-57	-0.53
17	rs962800	ILMN_2401641	ALDH3A2	7.84E-28	3.48E-22	6.56E-22	0.44	4.61E-38	2.05E-32	9.84E-31	0.49	3.31E-64	1.47E-58	8.21E-56	0.46

There are 2,980 cerebellar cisSNP/transcript associations with $q < 0.05$ both in the ADs and non-ADs. Some of these top associations are shown. Only one cisSNP/transcript pair is selected for depiction. The chromosome (CHR), SNP, Probe, Gene Symbol (SYMBOL) of these associations are depicted. The uncorrected (P), Bonferroni-corrected (P_{Bonf}) P, q values, and Beta coefficients of association are shown for the Non-ADs, ADs and combined (All) analyses. Regression coefficients are based on the SNP minor allele using an additive model. doi:10.1371/journal.pgen.1002707.t001

Table 2. Variance of cerebellar probe expression levels due to technical, biological, and *cis*SNP effects.

Probe	Symbol	Raw_Variance	R2technical	addR2covariates	adjR2covariates	addR2best-SNP	adjR2best-SNP	Best-cis-SNP
ILMN_1651745	TMEM25	1.894	0.08	0.004	0.005	0.784	0.852	rs11552421
ILMN_1718932	MTRR	0.864	0.05	0.004	0.005	0.768	0.809	rs3776455
ILMN_1694711	MAD2L1BP	1.169	0.154	0.007	0.008	0.642	0.763	rs1096699
ILMN_1730477	TAS2R43	1.212	0.372	0.001	0.002	0.477	0.76	rs2708389
ILMN_2345908	DDX11	0.798	0.029	0.003	0.003	0.732	0.755	rs4031375
ILMN_1807798	ATP5G2	0.722	0.156	0.009	0.011	0.61	0.722	rs1971762
ILMN_2052079	ZNF544	0.983	0.055	0.003	0.003	0.675	0.714	rs260462
ILMN_2369018	EVI2A	2.208	0.043	0.005	0.005	0.68	0.711	rs2525574
ILMN_2184966	ZHX2	2.816	0.11	0.012	0.013	0.629	0.706	rs3802266
ILMN_2376667	POFUT2	0.503	0.076	0.032	0.035	0.649	0.703	rs2838859
ILMN_1723984	PILRB	0.671	0.253	0.006	0.008	0.519	0.695	rs6955367
ILMN_2400759	CPVL	0.644	0.062	0.013	0.014	0.631	0.672	rs7313
ILMN_2262288	EEF1G	0.615	0.107	0.025	0.028	0.592	0.663	rs7124057
ILMN_1689177	PPAPDC1A	0.519	0.138	0.006	0.007	0.57	0.661	rs2182513
ILMN_2093720	THG1L	0.621	0.208	0.003	0.004	0.512	0.647	rs11738432
ILMN_1691772	ZSCAN29	0.19	0.14	0.011	0.013	0.548	0.638	rs12912744
ILMN_1809147	FAM118A	0.282	0.07	0.002	0.002	0.589	0.634	rs104664
ILMN_2130441	HLA-H	1.386	0.092	0.033	0.037	0.573	0.628	rs2975033
ILMN_2388272	MED24	3.036	0.043	0.007	0.007	0.597	0.624	rs8070454
ILMN_2064132	NANP	0.283	0.194	0.005	0.006	0.502	0.62	rs2387976
ILMN_1683279	PEX6	0.164	0.245	0.006	0.008	0.46	0.609	rs2395943
ILMN_2312606	IRF5	0.662	0.181	0.005	0.006	0.496	0.606	rs10239340
ILMN_2390162	PHF11	0.117	0.165	0.015	0.018	0.501	0.601	rs1046028
ILMN_1710903	MAPT	0.118	0.052	0.01	0.011	0.552	0.584	rs1981997
ILMN_1811048	GPR107	0.519	0.354	0.002	0.004	0.374	0.58	rs2240913
ILMN_2201966	N4BP1	1.97	0.029	0.005	0.006	0.555	0.572	rs11649236
ILMN_1795336	PTER	0.591	0.143	0.012	0.014	0.486	0.567	rs7909832
ILMN_1789419	EXOC3	0.214	0.097	0.012	0.014	0.508	0.563	rs11134054
ILMN_2364072	CLCNKA	1.356	0.044	0.009	0.01	0.535	0.562	rs1763601
ILMN_2075334	HIST1H4C	0.213	0.296	0.007	0.01	0.395	0.561	rs198834
ILMN_2296011	BRWD1	0.821	0.413	0.005	0.009	0.328	0.559	rs6517526
ILMN_2074477	GPR4	1.342	0.058	0.016	0.017	0.526	0.558	rs10405576
ILMN_1697286	SF3A1	0.168	0.359	0.009	0.014	0.356	0.555	rs737950
ILMN_1765332	TIMM10	0.164	0.079	0.005	0.006	0.507	0.551	rs2848630
ILMN_2401641	ALDH3A2	0.22	0.142	0.012	0.014	0.472	0.543	rs2108971
ILMN_2183938	LEMD3	0.335	0.084	0.026	0.029	0.495	0.541	rs10878255
ILMN_2198408	MFF	0.153	0.26	0.002	0.003	0.399	0.54	rs7560053
ILMN_1728199	POLE	0.113	0.175	0.006	0.007	0.446	0.538	rs4883627
ILMN_1655637	UPK1A	1.044	0.131	0.023	0.026	0.467	0.538	rs4806187
ILMN_2209027	RPS26	0.238	0.146	0.002	0.002	0.457	0.535	rs10876864

Results from some of the top probes are depicted. Only one probe is selected per gene for depiction. R2technical = variance due to technical variables only (i.e. plates, RIN). addR2covariates = added proportion of variance due to biological covariates (i.e. age, sex, ApoE4 dose), adjR2covariates = addR2covariates adjusted for technical variance, addR2best-SNP = proportion of variance due to the best *cis*SNP, adjR2best-SNP = addR2best-SNP adjusted for technical variance.
doi:10.1371/journal.pgen.1002707.t002

Identification of brain *cis*SNPs among PSP GWAS loci

In a recent PSP GWAS [27], four loci near *MAPT*, *STX6*, *EIF2AK3*, and *MOBP* conferred significant risk, in addition to three suggestive loci at *1q41 intergenic locus*, *BMS1* and *SLCO1A2*. We assessed these seven strongest PSP risk loci in our eGWAS in the ADs, non-ADs and combined datasets, as well as the PSP

subset of non-ADs (Table 5, Supplementary Table 16 in Dataset S1). We found novel, significant rs11568563 minor allele associations with reduced brain *SLCO1A2* levels ($p_{\text{Cer}} = 2.33 \times 10^{-8}$; $p_{\text{TCx}} = 4.36 \times 10^{-2} - 9.14 \times 10^{-13}$), which confers increased PSP risk [27]. *SLCO1A2* encodes solute carrier organic anion transporter family member 1a2 and is a drug transporter into the CNS [46].

Table 3. Validation of top cerebellar *cis*SNP/transcript associations in the temporal cortex.

CHR	SNP	PROBE	SYMBOL	Cerebellar eGWAS				Temporal Cortex Validation		
				ALL_P	ALL_P _{Bonf}	ALL_Q	ALL_BETA	ALL_P	ALL_P _{Bonf-study}	ALL_BETA
5	rs3776455	ILMN_1718932	MTRR	6.37E-133	2.83E-127	7.34E-120	1.20	5.15E-133	1.38E-129	1.23
12	rs10843881	ILMN_2345908	DDX11	1.48E-112	6.57E-107	2.39E-100	-1.06	1.81E-113	4.87E-110	-1.16
12	rs1971762	ILMN_1807798	ATP5G2	1.10E-104	4.87E-99	1.64E-93	1.01	7.83E-107	2.10E-103	1.15
21	rs2838859	ILMN_2376667	POFUT2	2.86E-103	1.27E-97	2.73E-92	-0.86	4.15E-115	1.11E-111	-0.84
8	rs3802266	ILMN_2184966	ZHX2	4.35E-101	1.93E-95	3.06E-90	-1.94	1.11E-104	2.97E-101	-2.35
19	rs260462	ILMN_2052079	ZNF544	8.26E-101	3.67E-95	4.38E-90	-1.16	7.00E-129	1.88E-125	-1.30
17	rs2525574	ILMN_2369018	EVI2A	1.52E-99	6.73E-94	4.88E-89	1.81	4.56E-89	1.22E-85	1.67
7	rs6955367	ILMN_1723984	PILRB	4.80E-97	2.13E-91	1.23E-86	1.08	1.70E-141	4.56E-138	1.56
7	rs7313	ILMN_2400759	CPVL	5.00E-92	2.22E-86	5.96E-82	-0.96	5.76E-71	1.55E-67	-1.34
11	rs7124057	ILMN_2262288	EEF1G	2.64E-91	1.17E-85	2.67E-81	-0.90	1.49E-91	4.01E-88	-0.70
10	rs2182513	ILMN_1689177	PPAPDC1A	1.08E-87	4.79E-82	6.93E-78	0.84	8.07E-01	NS	0.01
5	rs11738432	ILMN_2093720	THG1L	8.06E-84	3.58E-78	3.49E-74	-0.86	1.16E-100	3.12E-97	-1.09
15	rs12912744	ILMN_1691772	ZSCAN29	3.47E-83	1.54E-77	1.33E-73	-0.56	6.18E-96	1.66E-92	-0.63
22	rs136564	ILMN_1809147	FAM118A	8.44E-81	3.74E-75	2.57E-71	0.85	7.87E-57	2.11E-53	0.68
17	rs8070454	ILMN_2388272	MED24	3.51E-80	1.56E-74	9.58E-71	-2.03	3.97E-85	1.06E-81	-2.16
20	rs2387976	ILMN_2064132	NANP	9.58E-79	4.25E-73	1.93E-69	-0.54	1.97E-87	5.29E-84	-0.65
7	rs10239340	ILMN_2312606	IRF5	1.52E-76	6.73E-71	2.31E-67	0.81	3.76E-90	1.01E-86	0.92
6	rs2395943	ILMN_1683279	PEX6	3.63E-76	1.61E-70	5.06E-67	-0.39	1.64E-69	4.40E-66	-0.46
13	rs1046028	ILMN_2390162	PHF11	1.07E-75	4.73E-70	1.36E-66	0.35	5.38E-83	1.44E-79	0.38
17	rs1981997	ILMN_1710903	MAPT	4.16E-71	1.85E-65	3.88E-62	-0.48	2.42E-44	6.48E-41	-0.51
9	rs2240913	ILMN_1811048	GPR107	2.76E-69	1.23E-63	2.39E-60	0.66	2.11E-80	5.66E-77	0.95
16	rs11649236	ILMN_2201966	N4BP1	5.69E-69	2.53E-63	4.37E-60	-1.57	1.95E-85	5.25E-82	-1.44
10	rs7909832	ILMN_1795336	PTER	1.70E-68	7.54E-63	1.04E-59	-0.75	1.24E-84	3.32E-81	-1.00
5	rs11134054	ILMN_1789419	EXOC3	4.40E-68	1.95E-62	2.52E-59	-0.57	1.51E-78	4.06E-75	-0.60
19	rs10405576	ILMN_2074477	GPR4	5.76E-68	2.56E-62	3.09E-59	-1.32	5.20E-62	1.39E-58	-1.53
6	rs198834	ILMN_2075334	HIST1H4C	4.76E-67	2.11E-61	2.05E-58	-0.43	1.25E-51	3.35E-48	-0.52
22	rs737950	ILMN_1697286	SF3A1	2.05E-66	9.11E-61	7.72E-58	0.44	1.01E-71	2.71E-68	0.46
1	rs1763601	ILMN_2364072	CLCNKA	3.49E-66	1.55E-60	1.24E-57	1.22	6.22E-50	1.67E-46	1.02
12	rs10878255	ILMN_2183938	LEMD3	2.61E-65	1.16E-59	7.91E-57	-0.62	1.73E-69	4.65E-66	-0.78
19	rs4806187	ILMN_1655637	UPK1A	2.82E-65	1.25E-59	8.11E-57	-1.04	3.82E-39	1.02E-35	-0.72
11	rs2848630	ILMN_1765332	TIMM10	3.40E-65	1.51E-59	9.32E-57	-0.53	2.04E-68	5.48E-65	-0.45
17	rs962800	ILMN_2401641	ALDH3A2	3.31E-64	1.47E-58	8.21E-56	0.46	4.26E-76	1.14E-72	0.44
2	rs7560053	ILMN_2198408	MFF	6.41E-63	2.85E-57	1.38E-54	0.37	5.20E-71	1.40E-67	0.30
12	rs10876864	ILMN_2209027	RPS26	6.36E-62	2.82E-56	1.06E-53	0.49	3.01E-64	8.07E-61	0.57
12	rs4883627	ILMN_1728199	POLE	8.61E-62	3.82E-56	1.36E-53	0.33	3.95E-19	1.06E-15	0.22
6	rs2191651	ILMN_1694100	PRIM2	1.55E-61	6.88E-56	2.26E-53	-0.59	5.06E-57	1.36E-53	-0.57
10	rs9527	ILMN_2151056	C10orf32	2.90E-61	1.29E-55	3.90E-53	-0.46	1.27E-65	3.42E-62	-0.45

Of the 2,980 top *cis*SNP/transcript associations, 2,685 existed in the temporal cortex replication study. Some of these top associations are shown. Only one *cis*SNP/transcript pair is selected for depiction. The chromosome (CHR), SNP, Probe, Gene Symbol (SYMBOL) of these associations are depicted. The uncorrected (P), genome-wide (P_{Bonf}) and study-wide Bonferroni-corrected (P_{Bonf-study}) P values, Beta coefficient of association are shown for the combined (All) analyses in the cerebellar eGWAS and the temporal cortex replication study. Regression coefficients are based on the SNP minor allele using an additive model.
doi:10.1371/journal.pgen.1002707.t003

Fine-mapping of the *SLCO1A2* region revealed rs11568563 to be the strongest *cis*SNP influencing brain levels of this gene (Figure S5). This SNP was also identified as the top PSP-associating variant at this locus [27]. All other *cis*SNPs that associate with brain *SLCO1A2* levels have weaker effects that appear to be due to their LD with rs11568563, which is a missense coding mutation within *SLCO1A2*. Whether rs11568563 is merely tagging the functional variant(s)

regulating levels of *SLCO1A2* or coding changes also influence expressed transcript levels require further investigations. Additionally, *MAPT*/rs242557 minor allele increased PSP risk [27] and brain *MAPT* levels ($p_{\text{Cer}} = 9.78 \times 10^{-3}$ – 8.8×10^{-13} , $p_{\text{TCx}} = 1.1 \times 10^{-8}$). *MAPT*/rs8070723 minor allele associated with lower brain *MAPT* levels in our eGWAS, decreased PSP risk [27], similar to a PD GWAS [32]. We also found nominally significant increases in brain

Table 4. Examples of top cerebellar eGWAS *cis*SNPs also associated with complex diseases/traits.

Disease GWAS												
Cerebellar eGWAS												
CHR	SNP	MinorAllele	PROBE	SYMBOL	ALL_P	ALL_BETA	PUBMED ID	Disease/Trait	Reported Gene(s)	Strongest SNP-Risk Allele	P-Value	OR or beta
7	rs4132601	C	ILMN_1676575	IKZF1	1.63.E-21	-0.36	19684604	Acute lymphoblastic leukemia (childhood)	IKZF1	rs4132601-C	1.00E-19	1.69
19	rs260461	A	ILMN_2052079	ZNF544	3.10.E-16	0.73	18821565	Attention deficit hyperactivity disorder	ZNF544	rs260461-?	8.00E-06	NR
11	rs10838738	G	ILMN_2195462	C10TNF4	1.48.E-09	-0.18	19079261	Body mass index	MTCH2	rs10838738-G	5.00E-09	0.07
2	rs13015714	C	ILMN_1781700	IL18R1	8.30.E-24	0.55	18311140	Celiac disease	IL1RL1,IL18R1,IL18RAP, SLC9A4	rs13015714-C	4.00E-09	1.28
2	rs13015714	C	ILMN_2313672	IL1RL1	1.22.E-13	0.84	18311140	Celiac disease	IL1RL1,IL18R1,IL18RAP, SLC9A4	rs13015714-C	4.00E-09	1.28
7	rs2252521	A	ILMN_2400759	CPVL	1.86.E-18	0.55	19734545	Cognitive performance	CPVL	rs2252521-?	5.00E-06	NR
14	rs4444235	G	ILMN_1740900	BMP4	3.20.E-18	-0.27	19011631	Colorectal cancer	BMP4	rs4444235-C	8.00E-10	1.11
15	rs748404	G	ILMN_1691772	ZSCAN29	2.29.E-40	-0.42	19654303	Lung cancer	TGM5	rs748404-?	1.00E-06	1.15
6	rs499818	A	ILMN_1661622	TBC1D7	1.32.E-15	0.17	17903304	Major CVD	Intergenic	rs499818-?	7.00E-06	NR
17	rs6565681	A	ILMN_1749722	RNF213	2.10.E-22	-0.38	21048783	Moyamoya disease	RNF213	rs6565681-A	2.00E-08	4.82
6	rs2517713	C	ILMN_2130441	HLA-H	8.27.E-15	-0.66	19664746	Nasopharyngeal carcinoma	HLA-A	rs2517713-A	4.00E-20	1.88
6	rs3129055	G	ILMN_2203729	HCG4	1.90.E-14	-0.7	19664746	Nasopharyngeal carcinoma	HLA-F	rs3129055-G	7.00E-11	1.51
10	rs1561570	G	ILMN_2381899	OPTN	1.51.E-11	0.07	20436471	Paget's disease	OPTN	rs1561570-?	6.00E-13	1.54
17	rs8070723	G	ILMN_1710903	MAPT	7.02.E-69	-0.47	21044948	Parkinson's disease	MAPT	rs8070723-?	7.00E-12	1.3
17	rs11012	A	ILMN_2393693	LRR37A4	1.69.E-33	-0.52	20070850	Parkinson's disease	PLEKHM1, MAPT, IMP5	rs11012-T	6.00E-08	1.43
17	rs11012	A	ILMN_2286783	LRR37A4	1.64.E-37	-0.65	20070850	Parkinson's disease	PLEKHM1, MAPT, IMP5	rs11012-T	6.00E-08	1.43
4	rs6532197	G	ILMN_1660114	MMRN1	4.86.E-12	0.52	19915575	Parkinson's disease	MMRN1	rs6532197-G	1.00E-07	1.32
11	rs538147	A	ILMN_1772208	CCDC88B	3.92.E-22	-0.27	21399635	Primary biliary cirrhosis	RP56KA4	rs538147-G	2.00E-10	1.23
2	rs13385191	G	ILMN_1660275	C2orf43	7.01.E-22	-0.26	20676098	Prostate cancer	C2orf43	rs13385191-G	8.00E-08	1.15
5	rs27524	A	ILMN_2336220	ERAP1	9.60.E-17	0.3	20953190	Psoriasis	ERAP1	rs27524-A	3.00E-11	1.13
20	rs1008953	A	ILMN_1756590	SY51	4.15.E-10	-0.09	20953189	Psoriasis	SDC4	rs1008953-C	1.00E-07	1.14
7	rs4728142	A	ILMN_2312606	IRF5	2.30.E-24	-0.52	19838193	Systemic lupus erythematosus	IRF5	rs4728142-A	8.00E-19	1.43
6	rs2301271	A	ILMN_1700428	HLA-DOB	1.27.E-11	0.38	21408207	Systemic lupus erythematosus	HLA-DQA2	rs2301271-T	2.00E-12	1.47
11	rs4963128	A	ILMN_2349061	IRF7	6.84.E-21	-0.27	18204446	Systemic lupus erythematosus	KIAA1542	rs4963128-?	3.00E-10	1.28
11	rs4963128	A	ILMN_2349061	IRF7	6.84.E-21	-0.27	21408207	Systemic lupus erythematosus	KIAA1542	rs4963128-?	4.00E-06	1.33
12	rs1701704	C	ILMN_2209027	RPS26	1.31.E-39	0.42	18198356	Type 1 diabetes	RAB5B, SUOX, IKZF4, ERBB3, CDK2	rs1701704-C	9.00E-10	1.25
12	rs3764021	A	ILMN_1782729	CLECL1	7.67.E-22	-0.39	17554300	Type 1 diabetes	NR	rs3764021-C	5.00E-08	1.57
15	rs8042680	A	ILMN_2395932	UNC45A	1.84.E-13	-0.26	20581827	Type 2 diabetes	PRC1	rs8042680-A	2.00E-10	1.07
7	rs4728142	A	ILMN_2312606	IRF5	2.30.E-24	-0.52	21297633	Ulcerative colitis	IRF5, TNPO3	rs4728142-A	2.00E-08	1.07
9	rs4077515	A	ILMN_1811301	INPP5E	2.74.E-13	-0.12	20228799	Ulcerative colitis	CARD9	rs4077515-C	5.00E-08	1.14

The top 2,980 cerebellar eGWAS *cis*SNPs were compared to the "Catalog of Published GWAS" (www.genome.gov/gwastudies). Some of the resulting 60 common associations are reported. The chromosome (CHR), SNP, eGWAS Minor Allele, Probe, Gene Symbol (SYMBOL) of these associations are depicted. The uncorrected (P) and Beta coefficient of associations are shown for the combined (All) analyses of the cerebellar eGWAS. Regression coefficients are based on the SNP minor allele using an additive model. The information for the complex disease/trait GWAS was downloaded from their website accessed on 04/23/2011. The disease/trait associated SNPs shown are the strongest SNPs depicted in the disease/trait GWAS. The associating allele (Strongest SNP-Risk Allele), p-value, OR or beta for the strongest disease/trait SNPs are shown. doi:10.1371/journal.pgen.1002707.t004

MOBP levels ($p_{\text{Cer}} = 2.13 \times 10^{-2} - 1.71 \times 10^{-7}$; $p_{\text{TCx}} = 1.55 \times 10^{-2} - 1.57 \times 10^{-6}$) with rs1768208, which increases PSP risk [27].

The recent PSP GWAS by Hoglinger et al. [27] included eQTL analysis for the significant loci using brain expression levels from 387 subjects without clinical neurologic diseases. In addition to associations between *MAPT* locus *cis*SNPs with brain *MAPT* and *LRRC37A4* levels, they also detected signals for the nearby *ARL17A* and *PLEKHM1* genes, neither of which were detectable in our eGWAS. They also identified *cis*SNP associations with brain *MOBP* levels but even stronger influence on the nearby *SLC25A38* levels. We did not identify significant *cis*SNP/*SLC25A38* brain expression associations. Although some of the significant probes for *MOBP* and *MAPT* harbor variants within their probe sequence, which may potentially confound associations with expression levels, these genes had other significant probes without any sequence variants (Text S1).

Most non-AD subjects in our study had pathologic diagnosis of PSP ($n_{\text{Cer}} = 98$, $n_{\text{TCx}} = 107$, Supplementary Table 2 in Dataset S1). We assessed the 2,980 top cerebellar *cis*SNP/transcript associations in this subset, and found that most results were consistent with the ADs (Supplementary Tables 17 and 18 in Dataset S1).

Enrichment of brain *cis*SNPs in the AD GWAS from ADGC

To investigate whether any of the significant brain *cis*SNPs may influence risk of AD, we compared our eGWAS results to the AD risk associations from the large AD GWAS conducted by ADGC [28]. We obtained results of meta-analyses for the ADGC Stage 1+2 cohort (11,840 LOAD vs. 10,931 controls) [28] and investigated those SNPs with suggestive AD risk association in this dataset ($p_{\text{meta}} < 10^{-3}$). To ensure uniform comparison between our eGWAS and the ADGC GWAS, we assessed results from >2 million SNPs for each study using SNPs genome-wide imputed to HapMap phase 2 (release 22). There were 77,126 cerebellar (63,652 unique SNPs, 2,338 unique genes) and 68,172 temporal cortex (57,922 unique SNPs and 2,201 unique genes) *cis*SNP/transcript associations significant at $q < 0.05$ representing a clear excess (Figure S6). There were 380 *cis*SNPs that were significant for the cerebellar transcript associations and also had suggestive AD risk associations (2.9-fold enrichment), 432 such temporal cortex *cis*SNPs (3.3-fold enrichment) and 356 *cis*SNPs significant in both the cerebellum and temporal cortex (2.7-fold enrichment, $p < 10^{-6}$ for all three analyses) (Figure 1, Supplementary Tables 19 and 20 in Dataset S1).

MAPT and *LRRC37A4* *cis*SNPs, implicated in PSP [27] and PD [32] GWAS and which significantly influenced brain levels of these genes also had suggestive AD risk associations ($p_{\text{meta}} = 8.82 \times 10^{-4} - 1.53 \times 10^{-3}$). *Cis*SNP alleles associating with lower brain *MAPT* levels were associated with lower AD risk, similar to PD [32] and PSP [27] GWAS, which may suggest a common mechanism for these neurodegenerative diseases. *ABCA7*, identified recently as a novel LOAD risk locus [28,47], had significant cerebellar *cis*SNPs. Further investigations of the other genes with evidence of brain transcript and AD risk association is warranted to understand their role in AD (Text S1).

To ensure that we did not miss any associations due to the stringent eGWAS criteria that we applied, we repeated the analyses using no restrictions for transcript detection rates and eGWAS p value threshold of $p < 1.0E-4$. We also investigated *cis*SNPs identified in AD and non-AD brains, both separately, and jointly, given that some *cis*SNP associations may be unique to one group. We compared these eGWAS results to the ADGC GWAS as described above (Supplementary Tables 21, 22, 23, 24, 25, 26 in Dataset S1). Using cerebellar and temporal cortex eGWAS from

all subjects, 561 and 488 unique transcripts with *cis*SNPs that yield suggestive AD risk associations were identified, respectively. There were 259–312 such transcripts identified in each AD or non-AD eGWAS, with >50% overlap between the two diagnostic groups' results, although many of these results could be identified in the eGWAS of combined samples. About 7–10% of the transcripts could only be identified in just ADs or non-ADs, but not the combined eGWAS. Amongst such unique transcripts were *CLU* and *BIN1*, which reside at the LOAD GWAS loci [39,40,48] and associate with *cis*SNPs in the cerebellum of non-ADs. Detailed analyses of the *CLU* locus *cis*SNP/transcript associations are *in press* [49].

Discussion

In a large eQTL study on 773 brain samples from ~400 autopsied subjects, we demonstrate significant contribution of genetic factors to human brain gene expression, reliably detected across different brain regions and pathologies. There is significant enrichment of brain *cis*SNPs amongst disease-associated variants, advocating gene expression changes as a mechanism for the first time for certain genes implicated in human diseases, including PSP (*SLCO1A2*), PD (*MMRN1*), Paget's disease (*OPTN*) while replicating others (e.g. PD/*MAPT*, SLE/UC/*IRF5*). *MAPT* *cis*SNPs associating with PSP, PD and AD risk highlight potential common mechanisms for these neurodegenerative diseases.

The reported results have several important implications for the genetics of human brain gene expression: First, despite technical challenges of gene expression measurements in post-mortem brain tissue [50], ~70% of the transcriptome can be reliably detected in >75% of the subjects across two brain regions and different disease pathologies. Second, although there is significant contribution from technical covariates, genetic factors account for a substantial proportion of the variance in brain gene expression levels. We estimate that genetic factors explain an average 3% (range: 0–85%) of the variance in human cerebellar gene expression overall, and 18% (range: 8–85%) of the variance for the top *cis*-regulated transcripts. These estimates show remarkable similarity to those from other eQTL studies, such as a large, family-based lymphocyte eQTL, where *cis* eQTLs had an overall median effect size of 1.8% and significant eQTLs accounted for 24.6% of the variance in expression [16]. Similarly, significant *cis*SNPs explained 2–90% of expression variance in a liver eGWAS [23].

Third, there is remarkable replication of significant *cis*SNP associations across different brain regions and underlying tissue pathologies. Indeed, the 2,980 top cerebellar *cis*SNP/transcript associations represent 58% and 68% of all significant associations in the ADs and non-ADs. Since >50% of the non-ADs were comprised of subjects with PSP, we also conducted a separate analysis of this pathologically distinct group of non-ADs and again determined that many of the top *cis*SNPs were also significant in the PSPs despite the small sample size ($n = 98$). Importantly, most of the *cis*SNPs had highly similar effect sizes in the ADs, non-ADs and PSP subset of non-ADs. Furthermore, 78% of the top cerebellar *cis*-associations were also significant in the temporal cortex. Cerebellum is a relatively unaffected region in AD, whereas temporal cortex is typically one of the first areas to harbor neuropathology [51]. It is not inherently evident whether the unaffected or affected tissue regions would be most suitable for eQTL studies. Whereas unaffected regions would have the advantage of minimizing confounding on expression measurements from pathology (such as inflammation and cell death), affected regions may be more relevant for disease-associated

Table 5. PSP GWAS cisSNP/transcript associations in the cerebellar and temporal cortex.

CHR	SNP	PROBE	Symbol	ALL CER		PSP CER		AD CER		ALL TCX		PSP TCX		AD TCX		PSP GWAS ^b			
				BETA	P	BETA	P	BETA	P	BETA	P	BETA	P	BETA	P	BETA	P	Gene	OR
17	rs8070723	ILMN_1710903	MAPT	-0.473	7.02E-69	-0.428	1.38E-08	-0.429	1.49E-31	-0.498	8.61E-44	-0.276	9.05E-02	-0.495	3.12E-37	MAPT	5.11	4.43-5.91	1.5 × 10 ⁻¹¹⁸
12	rs11568563	ILMN_2381020	SLCO1A2	-0.397	2.33E-08	-0.371	3.62E-03	-0.438	1.98E-05	-0.730	9.14E-18	-0.644	3.89E-05	-0.715	1.23E-08	SLCO1A2	0.68	0.59-0.78	7.0 × 10 ⁻⁸
17	rs242557	ILMN_1710903	MAPT	0.183	8.80E-13	0.087	1.66E-02	0.179	2.39E-06	0.181	1.10E-08	-0.033	6.65E-01	0.251	4.03E-10	MAPT	0.51	0.48-0.55	2.7 × 10 ⁻⁷¹
3	rs1768208	ILMN_2298464	MOBP	0.363	1.71E-07	0.342	1.24E-02	0.334	3.70E-04	0.325	1.57E-06	0.342	2.33E-02	0.305	1.23E-03	MOBP	0.73	0.67-0.78	5.3 × 10 ⁻¹⁷
3	rs1768208	ILMN_2414962	MOBP	0.243	1.17E-04	0.220	8.02E-02	0.204	1.25E-02	0.180	1.76E-05	0.238	1.74E-02	0.137	8.18E-03	MOBP	0.73	0.67-0.78	5.3 × 10 ⁻¹⁷
17	rs8070723	ILMN_2298727	MAPT	-0.173	3.36E-07	-0.063	5.91E-01	-0.161	8.40E-04	-0.137	9.03E-04	-0.073	6.15E-01	-0.211	8.67E-05	MAPT	5.11	4.43-5.91	1.5 × 10 ⁻¹¹⁸
3	rs1768208	ILMN_1768947	MOBP	0.143	7.57E-03	0.079	4.55E-01	0.159	1.72E-02	0.107	3.03E-03	0.199	1.83E-02	0.067	1.38E-01	MOBP	0.73	0.67-0.78	5.3 × 10 ⁻¹⁷
1	rs1411478	ILMN_2157951	STX6	-0.005	6.41E-01	-0.031	1.32E-01	-0.002	8.97E-01	0.027	1.17E-02	-0.003	8.90E-01	0.041	8.50E-03	STX6	0.79	0.73-0.84	3.5 × 10 ⁻¹¹
3	rs1768208	ILMN_1750271	MOBP	0.074	2.13E-02	0.064	2.84E-01	0.073	8.86E-02	0.046	1.55E-02	0.082	1.03E-01	0.045	4.29E-02	MOBP	0.73	0.67-0.78	5.3 × 10 ⁻¹⁷
12	rs11568563	ILMN_1656097	SLCO1A2	-0.077	2.41E-01	-0.027	8.14E-01	-0.117	2.12E-01	-0.117	4.36E-02	-0.061	6.08E-01	-0.064	4.22E-01	SLCO1A2	0.68	0.59-0.78	7.0 × 10 ⁻⁸
17	rs242557	ILMN_2310814	MAPT	0.002	8.27E-01	-0.032	9.13E-02	0.008	6.19E-01	-0.026	1.09E-01	-0.054	3.49E-01	-0.022	9.33E-02	MAPT	0.51	0.48-0.55	2.7 × 10 ⁻⁷¹
17	rs8070723	ILMN_2200636	KIAA1267	0.021	1.14E-01	0.046	2.87E-01	0.036	4.04E-02	-0.016	1.70E-01	-0.025	5.80E-01	-0.026	1.02E-01	MAPT	5.11	4.43-5.91	1.5 × 10 ⁻¹¹⁸
17	rs8070723	ILMN_1665311	STH	0.044	1.79E-02	0.114	7.56E-02	0.026	3.05E-01	0.051	1.86E-01	0.188	2.59E-01	-0.022	6.24E-01	MAPT	5.11	4.43-5.91	1.5 × 10 ⁻¹¹⁸
17	rs242557	ILMN_2298727	MAPT	0.071	9.78E-03	0.086	9.40E-02	0.075	8.56E-02	0.041	2.17E-01	-0.054	4.26E-01	0.110	2.26E-02	MAPT	0.51	0.48-0.55	2.7 × 10 ⁻⁷¹
1	rs1411478	ILMN_2167416	MR1	0.045	4.27E-02	0.092	4.48E-02	0.015	6.02E-01	0.029	2.24E-01	-0.002	9.66E-01	0.077	1.20E-02	STX6	0.79	0.73-0.84	3.5 × 10 ⁻¹¹
17	rs8070723	ILMN_2310814	MAPT	-0.007	5.77E-01	-0.028	5.12E-01	0.007	6.94E-01	0.018	3.65E-01	0.039	7.52E-01	0.024	1.03E-01	MAPT	5.11	4.43-5.91	1.5 × 10 ⁻¹¹⁸
17	rs242557	ILMN_2200636	KIAA1267	-0.012	2.59E-01	-0.022	2.56E-01	-0.019	2.23E-01	0.008	3.83E-01	0.030	1.50E-01	-0.003	8.08E-01	MAPT	0.51	0.48-0.55	2.7 × 10 ⁻⁷¹
17	rs242557	ILMN_1665311	STH	-0.027	6.59E-02	-0.020	4.88E-01	-0.017	4.42E-01	-0.024	4.28E-01	-0.082	2.91E-01	0.027	4.89E-01	MAPT	0.51	0.48-0.55	2.7 × 10 ⁻⁷¹
1	rs1411478	ILMN_1721833	IER5	-0.041	2.12E-01	-0.074	2.56E-01	NA	NA	0.016	4.76E-01	0.025	5.73E-01	0.000	9.95E-01	STX6	0.79	0.73-0.84	3.5 × 10 ⁻¹¹
10	rs2142991	ILMN_1772713	BMS1	0.001	9.36E-01	-0.007	8.18E-01	0.006	7.26E-01	0.003	8.02E-01	0.028	2.30E-01	-0.014	4.56E-01	BMS1	1.3	1.18-1.44	3.2 × 10 ⁻⁷
3	rs1768208	ILMN_1781231	SLC25A38	0.004	7.76E-01	-0.020	4.33E-01	0.025	1.47E-01	0.002	9.02E-01	0.001	9.78E-01	0.019	2.84E-01	MOBP	0.73	0.67-0.78	5.3 × 10 ⁻¹⁷
2	rs7571971	ILMN_1714809	RPIA	-0.035	7.03E-02	-0.071	9.57E-02	NA	NA	0.002	9.21E-01	0.027	5.19E-01	0.008	7.68E-01	EIF2AK3	0.75	0.70-0.82	4.2 × 10 ⁻¹³
3	rs1768208	ILMN_2411723	RPSA	-0.025	6.92E-02	-0.057	3.77E-02	-0.034	8.72E-02	0.000	9.66E-01	-0.008	7.29E-01	-0.001	9.41E-01	MOBP	0.73	0.67-0.78	5.3 × 10 ⁻¹⁷

Seven SNPs near six PSP candidate risk genes from a recent PSP GWAS [27] were tested for transcript associations in *cis* in our cerebellar (CER) eGWAS and temporal cortex replications. The chromosome (CHR), SNP, Probe, Gene Symbol (SYMBOL) of these associations are depicted. The uncorrected (P) and Beta coefficient of association are shown for the combined (ALL), PSP, AD analyses in the cerebellar eGWAS and the temporal cortex replication study. Regression coefficients are based on the SNP minor allele using an additive model. NA; Not available. a. Though a SNP is present in the sequence of probe ILMN_2298727 (rs73314997), this is essentially monomorphic in the eGWAS subjects. b. The odds ratios (OR), confidence intervals and P values of disease risk association from the PSP GWAS are also shown using results from Supplementary Table 2 in Dataset S1 of this study [27]. doi:10.1371/journal.pgen.1002707.t005

eQTL mapping. The substantial overlap in significant *cis*SNP associations between different brain regions and disease types in our study implies that sample size may be the most critical element of successful eQTL mapping. In other words, analysis of expression data collected in different tissue regions and diseases, provided there is careful statistical control, could greatly enhance power to detect eQTLs. Nevertheless, there may be important eQTLs that are specific to brain region and disease.

It is not obvious whether the *cis*SNP that display similar effects in different brain regions and different disease types would have relevance to human disease. The top 2,596 cerebellar *cis*SNPs that are significant in both ADs and non-ADs, and many of which are also significant in the temporal cortex are also enriched for variants implicated in human disease, including CNS disease, such as PD and PSP. Thus, the fourth implication of our study is that it may be possible to map disease-associated variants using eQTL studies conducted in unaffected tissue or unaffected subjects. In addition to providing a general characterization of the genetics of brain gene expression, this study successfully replicated many previously published *cis*SNP associations, such as rs8070723/*MAPT* level, rs11012/*LRRC37A4* level associations, both of which were implicated in PD. We found novel brain expression level associations for transcripts implicated in disease, including rs11568563 association with *SLCO1A2*, recently identified in a PSP GWAS. The disease-associating *cis*SNP associations identified in this study were not restricted to CNS diseases, but also included non-CNS diseases, such as SLE, where we replicated the previously published rs4728142/*IRF5* level associations.

These findings imply that many disease-associated *cis*SNPs can influence gene expression independently of tissue/region/pathology, and be mapped reliably in tissue which is unaffected, not disease-related or from unaffected subjects. Indeed, our findings are consistent with a study of lymphoblastoid cell lines from subjects affected and unaffected with asthma, where Dixon et al. [12] found no differences between asthmatics and non-asthmatics. Furthermore, they detected significant transcript level associations with SNPs that also associate with asthma. Emilsson et al. [20] performed eQTL mapping in both blood and adipose tissue and determined that >50% of significant adipose tissue *cis*SNPs were also significant in blood. This is similar to the overlap we detected for cerebellum and temporal cortex, though two brain regions are more likely to have similar eQTL profiles than two different tissues.

Although many *cis*SNP effects can be detected in many different tissue types and disease conditions as shown here and by others [12,20], there conceivably exist expression variants which exert their effects in a tissue or disease-specific manner. For example in the eQTL comparing blood and adipose tissue, Emilsson et al. [20] also found that more transcripts from adipose tissue had significant correlations with obesity-related traits. In reality, both scenarios may be at play, such that some expression variants have more ubiquitous effects, whereas others may need tissue/cell/region/disease specific factors to exert their influence on gene expression. Indeed, many of the CNS disease related *cis*SNP associations in our brain eGWAS could not be identified in our comparison to a liver eGWAS [23] or an existing database for a LCL eGWAS [12], suggesting that disease-relevant tissue may be necessary to detect effects of certain *cis*SNPs, and highlighting the value of this brain eGWAS for CNS traits/conditions.

Despite the enrichment of our samples with tissue from AD subjects and our use of both cerebellar and temporal cortex tissue, we did not identify strong transcript associations for some of the top genes recently implicated in AD risk in large LOAD GWAS studies [28,39,40,47,48]. This could be because the AD risk

variants in these genes exert their effects via mechanisms other than influencing transcript levels, namely changes in protein conformation. If so, even the negative results from an eGWAS could be informative in guiding the future deep-sequencing efforts which should focus on coding rather than non-coding, functional regions. Alternative explanations include technical shortcomings, such as inability to measure all transcript species, measurements of global rather than cell-specific gene expression, not including all tested disease-associated variants in our genotyping platform. We also need to consider that the top genes nearest the strongest variants from the LOAD GWAS may not be actual disease genes. These loci require further investigations to account for this possibility. Additionally, our criteria for selection of the top *cis*SNPs, requiring significance in both ADs and non-ADs, might be too stringent, thereby leading to some false negative results. Finally, it may be possible to identify additional disease-related expression variants by focusing on those that have differential influence in disease vs. non-disease tissue, although this was not a focus of analysis in this study. Given that our non-AD tissue also consisted of subjects with other neurodegenerative diseases, there may be more similarities with the AD tissue, making it more difficult to detect variants with differential disease-related expression-associations in our current study. Nevertheless, we did find associations with *cis*SNPs for *ABCA7*, a novel AD risk locus gene [28,47] and *MAPT* [52,53], [24,54] implicated in AD.

It is important to emphasize that although the identification of transcript level associations provides another layer of confidence for disease-associating variants and genes, it is entirely possible that a variant in an LD region encompassing multiple genes, could be marking a functional disease variant in one gene and an expression variant in another gene. Thus, although highly useful in conjunction with disease association studies, eGWAS should be seen as a guide rather than ultimate evidence in disease-mapping efforts. Similarly, absence of eGWAS associations for a disease-associated variant should not be seen as contradictory evidence, but rather raise the possibility of alternative functional mechanisms for that variant.

Despite the wealth of information our study provides, we acknowledge several shortcomings. First, our non-ADs were not normal controls but often had other brain pathologies. It will be necessary to seek replication of these findings or novel *cis*SNP/transcript associations in normal brain tissue, as well. Second, we only focused on single SNP associations. The preliminary observations from our eGWAS findings suggest that multiple independent variants may affect brain expression levels of some genes, whereas others might be under the influence of a single strong variant. Finally, like any association study, it is not clear whether the *cis*SNPs identified in our eGWAS are themselves the functional SNPs or simply in LD with un-genotyped regulatory variants. Future studies focusing on analysis of haplotypes, SNPxSNP interactions, novel variant discovery and functional *in-vitro* studies testing effects of multiple variants are required to dissect the genetic variation underlying brain gene expression levels.

In summary, this cerebellar eGWAS study and the temporal cortex validations provide insight about the genetics of brain gene expression, a framework to guide future studies with respect to tissue/region/disease choice in eQTL studies, examples about the utility of this approach in gene mapping, replication of some known transcript associations and evidence for novel transcript associations in human disease. Combined eGWAS-disease GWAS approach may provide complementary information in mapping human disease and enable identification of functional variants that may not be possible by either approach alone.

The complete set of results from the brain eGWAS can be accessed at the National Institute on Aging Genetics of Alzheimer's Disease Data Storage (NIAGADS) website at <http://alois.med.upenn.edu/niagads/>. Questions about the dataset can be addressed to the corresponding author of this manuscript (taner.nilufer@mayo.edu).

Methods

Subjects

All subjects were participants in the published Mayo LOAD GWAS [29] as part of the autopsy-based series (AUT_60–80). All subjects had neuropathologic evaluation by DWD. All ADs had definite diagnosis according to the NINCDS-ADRDA criteria [55] and had Braak scores of ≥ 4.0 . All non-ADs had Braak scores of ≤ 2.5 , and many had brain pathology unrelated to AD (Supplementary Tables 1 and 2 in Dataset S1). Three-hundred forty subjects had measurements in both cerebellum and temporal cortex. This study was approved by the appropriate institutional review board.

Expression genome-wide association study (eGWAS)

RNA extraction and gene expression measurements. Total RNA was extracted from frozen brain samples using the Ambion RNAqueous kit according to the manufacturer's instructions. The quantity and quality of the RNA samples were determined by the Agilent 2100 Bioanalyzer using the Agilent RNA 6000 Nano Chip. Transcript levels were measured using the Whole Genome DASL assay (Illumina, San Diego, CA). Probe annotations were done based on NCBI Ref Seq, Build 36.2. The RNA samples were randomized across the chips and plates using a stratified approach to ensure balance with respect to diagnosis, age, gender, RINs and APOE genotype. Replicate samples were utilized for QC and also for intra-class coefficient (ICC) estimations. Raw probe level mRNA expression data were exported from GenomeStudio software (Illumina Inc.) for preprocessing with background correction, variance stabilizing transformation, quantile normalization and probe filtering using the lumi package of BioConductor [56,57] (Text S1). A probe with detectable signal in $>75\%$ of the samples was regarded as informative and used in subsequent analyses, although we also did supplementary analyses without imposing any restrictions based on probe detection levels. The number of informative probes differed slightly between the AD, non-AD and combined groups (Figure S7).

Genome-wide genotyping. Genotypes were generated using Illumina's HumanHap300-Duo Genotyping BeadChips and analyzed with an Illumina BeadLab Station (Illumina, San Diego, CA) at the Mayo Clinic Genotyping Shared Resource according to the manufacturer's protocols. The LOAD GWAS QC methods were previously published [29] (Text S1).

Statistical methods for eGWAS. Linear regression analysis to test for *cis*SNP/transcript associations were done in PLINK [58]. Preprocessed probe transcript levels were utilized as endophenotypes. Each probe was assessed separately, even though one gene may have multiple probes. *Cis*SNPs localized to ± 100 kb flanking region of the gene targeted by the probe of interest, mapped according to NCBI Build 36, were assessed for transcript level associations, using an additive model, with the minor allele dosage (0, 1, 2) as the independent variable, and APOE $\epsilon 4$ dosage (0, 1, 2), age at death, gender, PCR plate, RIN, (RIN-RINmean)² as covariates. The cerebellum and temporal cortex expression levels were analyzed separately. The ADs and non-ADs were analyzed both separately and jointly. The joint analyses included diagnosis as an additional covariate (AD = 1, non-AD = 0). We also ran analyses including the top 10 eigenvectors from EIGENSTRAT, and

compared eGWAS results to those excluding the eigenvectors (Text S1, Figure S8, Supplementary Table 27 in Dataset S1) [59].

Q values used for multiple testing corrections are based on false-discovery rates [30] and were corrected for genomic inflation of significance (Text S1). In addition, permutations ($p_{\text{perm-WY}}$) and Bonferroni adjustment were used for comparison of correction strategies. Permutation p values were obtained by shuffling the endophenotype, while maintaining the covariate structure, 10,000 times and applying the Westfall and Young [60] resampling-style stepdown approach to account for correlations between probes.

Variance of gene expression

To assess the genetic contribution to the variance in human cerebellar gene expression, we first determined between-subject variance, as a percentage of the total variance in probe expression, using ICC [31] for 15 samples measured in replicate on 5–6 different plates and 2–3 different days.

Using multivariable linear regression models, we then calculated the proportion of variance in cerebellar gene expression levels that were explained by technical effects (PCR plate, RIN, (RIN-RINmean)²), biological covariates (APOE $\epsilon 4$ dosage, age at death, gender) and the “best” *cis*SNP for each probe. These analyses were carried out on the combined dataset consisting of cerebellar expression measurements from 374 subjects and 15,283 probes with at least one *cis*SNP (Text S1).

Replication of top cerebellar eGWAS hits in the temporal cortex

We identified 2,980 *cis*SNP/transcript associations (2,596 unique SNPs, 746 unique probes and 686 unique genes) that achieved genome-wide significance within both the ADs and non-ADs analyses with q values < 0.05 . All 2,980 *cis*SNP/transcript associations achieved genome-wide significance with $q < 0.05$ and $p_{\text{Bonf}} < 0.05$ in the combined ADs+non-ADs analysis. We sought validation of these hits in the temporal cortex of 399 subjects who had WG-DASL whole transcriptome measurements and whole-genome genotypes. RNA extractions, QC, WG-DASL measurements, transcript level detections and association analyses were performed for these temporal cortex samples, in the same manner as that for the cerebellar samples. After appropriate QC, 2,685 of the 2,980 top cerebellar *cis*SNP/transcript associations remained detectable among the temporal cortex results (2,387 unique SNPs, 677 unique probes and 625 unique genes).

Comparison of cerebellar eGWAS results with other published complex disease and trait GWAS

To determine whether the cerebellar eGWAS captured variants implicated in complex diseases/traits, we compared the top 2,980 cerebellar eGWAS *cis*SNPs with the top disease/trait associated SNPs in the “Catalog of Published GWAS” [26], curated by the National Human Genome Research Institute (www.genome.gov/gwastudies). This catalog compiles weekly search results from all published GWAS of $\geq 100,000$ SNPs where associations of $p \leq 1.0E-05$ are reported. The catalog accessed on 04/23/2011 had 5,272 entries. We restricted our search to those entries where the “SNPs” column had only one SNP with an rs number. Thus, haplotypes and variants without rs numbers were excluded. There were 5,101 entries after this exclusion, comprised of 4,248 unique SNPs and 433 unique diseases. One SNP may associate with > 1 disease/trait and each disease/trait may have ≥ 1 associating SNP. This list was linked to the 2,980 top cerebellar *cis*SNPs by common rs numbers.

To assess whether the number of observed *cis*SNPs that have both significant cerebellar eGWAS and disease/trait associations represent a significant enrichment, we performed simulations while adjusting for *cis*SNP minor allele frequencies, as previously reported [3]. We performed 1 million simulations and adjusted for the minor allele frequencies of all the tested *cis*SNPs in 10 bins from 0–0.05 to 0.45–0.50. Using the total number of *cis*SNPs that are both transcript and disease/trait associating for each simulation, we obtained an empirical p value and an estimate of fold-enrichment.

Cerebellar eGWAS results were also compared to other published eGWAS results from a human liver [23] and two human brain [24,25] studies. The methods and results are depicted in Text S1.

Alzheimer's Disease Genetics Consortium (ADGC) meta-analyses

To determine whether any of the *cis*SNPs significant at $q < 0.05$ influenced risk of AD, we obtained meta-analyses results from the ADGC [28]. The cohorts that are assessed by ADGC, as well as the methodological details of the meta-analyses are described in detail in a recent publication [28]. Briefly, the meta-analyses of the ADGC dataset results reported here (Supplementary Tables 17 and 18 in Dataset S1) are generated from the combined analyses of stage 1 and stage 2 cohorts (Text S1), with detailed descriptions provided elsewhere [28]. Stage 1 cohorts are comprised of 8,309 LOAD cases and 7,366 cognitively normal elder controls. Stage 2 has 3,531 LOAD vs. 3,565 control subjects. Each cohort was tested for AD risk association using a logistic regression approach, assuming an additive model and adjusting for age, sex, APOE $\epsilon 4$ dosage and principal components from EIGENSTRAT [59]. The meta-analyses results were generated using the inverse variance method implemented in the software package METAL [61].

Supporting Information

Dataset S1 This file includes Supplementary Tables 1–27. The individual supplementary table legends are included in the first tab of this file.
(XLS)

Figure S1 Q-Q-Plots: Q-Q plots of observed (y-axis) versus expected (X-axis) $-\log(p)$ values of association for all *cis*SNP/transcript associations in the combined cerebellar 374 samples obtained before (a,b) and after (c,d) inflation-adjustments. Q-Q plots for all data points (a, c), as well as those that are in the lower, left hand corner (b,d) are shown. The data in b and d account reflect the association results, where there should be no deviations from the expected (i.e. null hypothesis of no association).
(PDF)

Figure S2 Venn diagram of significant cerebellar *cis*SNP/transcript associations: Q values < 0.05 in the ADs, non-ADs and combined (All) analyses. Notably, 2,980 *cis*-SNP/transcript associations are significant both in the ADs and non-ADs.
(PDF)

Figure S3 Box Plots of some top *cis*SNP/transcript associations in the non-AD (a), AD (b) and combined groups (c): The SNP genotypes are shown on the X-axis with the genotype counts in parentheses. Variance stabilizing transformed (VST) expression levels are on the Y-axis. The bottom and top of a box represent the lower and upper quartiles, respectively. The band near the middle of the box is the median. The ends of the whiskers depict the most extreme observations still within 1.5 inter quartile range of the

corresponding quartile. Any data not included between the whiskers are plotted as dots.
(PDF)

Figure S4 Histogram of intra-class coefficients (ICC) for the cerebellar probe expressions. Using 15 replicate samples, ICC, which is the between-subject variance, as a percentage of the total variance in probe expression, was estimated for 17,121 probes.
(JPG)

Figure S5 Data plots of SNPs tested for association with expression levels of *SLCO1A2* in the Temporal Cortex and Cerebellum. Forty-six SNPs were tested for association of *SLCO1A2* levels in the Cerebellum (Blue lines) and Temporal Cortex (Pink lines). P-values were transformed using $-\log_{10}$ and are plotted against the position of each SNP along the chromosome (Kbp). Genes found within the locus boundaries are shown from the UCSC genome browser (<http://genome.ucsc.edu/>). The LD across the locus is represented by a plot generated with Haploview, using data from the Mayo GWAS. The top eSNP in this study, rs11568563, is highlighted on the p-value plot by red squares and a red box around the SNP in the list of rs numbers. This is also the top PSP-associating SNP at this locus in Hoglinger et al. (*Nat Genet*, 2011) [27].
(PDF)

Figure S6 Q-Q-Plots for cerebellar and temporal cortex *cis*SNP/transcript associations with the HapMap phase 2 imputed genotypes: Q-Q plots of observed (y-axis) versus expected (X-axis) $-\log(p)$ values of association for all *cis*SNP/transcript associations in the combined dataset obtained before (a) and after (b) genomic inflation-adjustments, as discussed in the text. Also shown are the Q-Q plots for the temporal cortex associations in the combined dataset obtained before (c) and after (d) inflation-adjustments.
(PDF)

Figure S7 Venn diagram of detectable cerebellar probes. Venn diagram of cerebellar probes detectable in $\geq 75\%$ of subjects in the AD (AD), non-AD (CON) and combined (All) analyses. Notably, 13,349 probes were detectable in all 374 subjects.
(PDF)

Figure S8 Scatterplots of $-\log_{10}$ p values for eGWAS associations with and without inclusion of eigenvectors. Transformed P-values of a) Cerebellar and b) Temporal Cortex eGWAS *cis*SNP/transcript associations from models including (y-axis) and excluding (x-axis) the top 10 eigenvectors are plotted. A linear line demonstrating the null hypothesis of no deviation of the results between the two datasets is also shown. The results are displayed for those SNPs with a Hardy-Weinberg P-value $> 1.0E-07$ and a probe detection threshold $> 75\%$.
(JPG)

Text S1 Supplementary Results, Methods and References.
(DOC)

Acknowledgments

ADGC Authors

Liana G. Apostolova, MD,¹ Steven E. Arnold, MD,² Clinton T. Baldwin, PhD,³ Robert Barber, PhD,⁴ Michael M. Barmada, PhD,⁵ Thomas Beach, MD, PhD,⁶ Gary W. Beecham, PhD,^{7,8} Duane Beekly, BS,⁹ David A. Bennett, MD,^{10,11} Eileen H. Bigio, MD,¹² Thomas D. Bird, MD,¹³ Deborah Blacker, MD,^{14,15} Bradley F. Boeve, MD,¹⁶ James D. Bowen, MD,¹⁷ Adam Boxer, MD, PhD,¹⁸ James R. Burke, MD, PhD,¹⁹ Jacqueline Buros, BS,³ Joseph D. Buxbaum, PhD,^{20,21,22} Nigel J. Cairns, PhD, FRCPATH,²³ Laura B. Cantwell, MPH,²⁴ Chuanhai Cao, PhD,²⁵ Chris S. Carlson, PhD,²⁶ Regina M. Carney, MD,²⁷ Steven L. Carroll,

MD, PhD,²⁸ Helena C. Chui, MD,²⁹ David G. Clark, MD,³⁰ Jason Carneveaux, BS,³¹ Carl W. Cotman, PhD,³² Paul K. Crane, MD, MPH,³³ Carlos Cruchaga, PhD,³⁴ Jeffrey L. Cummings, MD,¹ Philip L. De Jager, MD, PhD,^{35,36} Charles DeCarli, MD,³⁷ Steven T. DeKosky, MD,³⁸ F. Yesim Demirci, MD,⁵ Ramon Diaz-Arrastia, MD, PhD,³⁹ Malcolm Dick, PhD,³² Beth A. Dombroski, PhD,²⁴ Ranjan Duara, MD,⁴⁰ William G. Ellis, MD,⁴¹ Denis Evans, MD,⁴² Kelley M. Faber, MS,¹¹⁴ Kenneth B. Fallon, MD,²⁸ Martin R. Farlow, MD,⁴³ Steven Ferris, PhD,⁴⁴ Tatiana M. Foroud, PhD,¹¹⁴ Matthew P. Frosch, MD, PhD,⁴⁵ Douglas R. Galasko, MD,⁴⁶ Paul J. Gallins, MS,⁷ Mary Ganguli, MD,⁴⁷ Marla Gearing, PhD,^{48,49} Daniel H. Geschwind, MD, PhD,⁵⁰ Bernardino Ghetti, MD,⁵¹ John R. Gilbert, PhD,^{7,8} Sid Gilman, MD, FRCP,⁵² Bruno Giordani, PhD,⁵³ Jonathan D. Glass, MD,⁵⁴ Alison M. Goate, D.Phil,³⁴ Robert C. Green, MD,^{3,55,56} John H. Growdon, MD,⁵⁷ Hakon Hakonarson, MD, PhD,⁵⁸ Ronald L. Hamilton, MD,⁵⁹ John Hardy, PhD,⁶⁰ Lindy E. Harrell, MD, PhD,³⁰ Elizabeth Head, PhD,⁶¹ Lawrence S. Honig, MD, PhD,⁶² Matthew J. Huentelman, PhD,³¹ Christine M. Hulette, MD,⁶³ Bradley T. Hyman, MD, PhD,⁵⁷ Gail P. Jarvik, MD, PhD,^{64,65} Gregory A. Jicha, MD, PhD,⁶⁶ Lee-Way Jin, MD, PhD,⁴¹ Nancy Johnson, PhD,⁶⁷ Gyungah Jun, PhD,^{68,69} M. Ilyas Kamboh, PhD,^{5,70} Jason Karlawish, MD,⁷¹ Anna Karydas, BA,¹⁸ John S.K. Kauwe, PhD,⁷² Jeffrey A. Kaye, MD,^{73,74} Ronald Kim, MD,⁷⁵ Edward H. Koo, MD,⁴⁶ Neil W. Kowall, MD,^{55,76} Patricia Kramer, PhD,^{77,73} Walter A. Kukull, PhD,⁷⁸ James J. Lah, MD, PhD,⁵⁴ Eric B. Larson, MD, MPH,⁷⁹ Allan I. Levey, MD, PhD,⁵⁴ Andrew P. Lieberman, MD, PhD,⁸⁰ Oscar L. Lopez, MD,⁷⁰ Kathryn L. Lunetta, PhD,⁶⁸ Wendy J. Mack, PhD,⁸¹ Daniel C. Marson, JD, PhD,³⁰ Eden R. Martin, PhD,^{7,8} Frank Martiniuk, PhD,⁸² Deborah C. Mash, PhD,⁸³ Eliezer Masliah, MD,^{46,84} Wayne C. McCormick, MD, MPH,³³ Susan M. McCurry, PhD,⁸⁵ Andrew N. McDavid, BA,²⁶ Ann C. McKee, MD,^{55,76} Marsel Mesulam, MD,^{86,87} Bruce L. Miller, MD,¹⁸ Carol A. Miller, MD,⁸⁸ Joshua W. Miller, PhD,⁴¹ Thomas J. Montine, MD, PhD,⁸⁹ John C. Morris, MD,^{23,90} Amanda J. Myers, PhD,⁹¹ Adam C. Naj, PhD,⁷ Petra Nowotny, PhD,³⁴ Joseph E. Parisi, MD,^{92,93} Daniel P. Perl, MD,⁹⁴ Elaine Peskind, MD,⁹⁵ Wayne W. Poon, PhD,³² Huntington Potter, PhD,²⁵ Joseph F. Quinn, MD,⁷³ Ashok Raj, MD,²⁵ Ruchita A. Rajbhandary, MPH,⁷ Murray Raskind, MD,⁹⁵ Eric M. Reiman, MD,^{31,96,97,98} Barry Reisberg, MD,^{44,99} Christiane Reitz, MD, PhD,^{100,101,102} John M. Ringman, MD,¹ Erik D. Roberson, MD, PhD,³⁰ Ekaterina Rogaeva, PhD,¹⁰³ Roger N. Rosenberg, MD,³⁹ Mary Sano, PhD,²¹ Andrew J. Saykin, PsyD,^{33,104} Julie A. Schneider, MD,^{105,10} Lon S. Schneider, MD,^{29,106} William Seeley, MD,¹⁸ Michael L. Shelanski, MD, PhD,¹⁰⁷ Michael A. Slifer, MD, PhD,^{7,8} Charles D. Smith, MD,⁶⁶ Joshua A. Sonnen, MD,⁸⁹ Salvatore Spina, MD,⁵¹ Peter St George-Hyslop, MD, FRCP,^{103,108} Robert A. Stern, PhD,⁵⁵ Rudolph E. Tanzi, PhD,⁵⁷ John Q. Trojanowski, MD, PhD,²⁴ Juan C. Troncoso, MD,¹⁰⁹ Debby W. Tsuang, MD,⁹⁵ Viviana M. Van Deerlin, MD, PhD,²⁴ Badri Narayan Vardarajan, MS,³ Harry V. Vinters, MD,^{1,110} Jean Paul Vonsattel, MD,¹¹¹ Li-San Wang, PhD,²⁴ Sandra Weintraub, PhD,^{86,87} Kathleen A. Welsh-Bohmer, PhD,^{19,112} Jennifer Williamson, MS,⁶² Randall L. Woltjer, MD, PhD¹¹³

ADGC Author Affiliations

¹Department of Neurology, University of California Los Angeles, Los Angeles, California; ²Department of Psychiatry, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania; ³Department of Medicine (Genetics Program), Boston University, Boston, Massachusetts; ⁴Department of Pharmacology and Neuroscience, University of Texas Southwestern, Fort Worth, Texas; ⁵Department of Human Genetics, University of Pittsburgh, Pittsburgh, Pennsylvania; ⁶Civin Laboratory for Neuropathology, Banner Sun Health Research Institute, Phoenix, Arizona; ⁷The John P. Hussman Institute for Human Genomics, University of Miami, Miami, Florida; ⁸Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami, Miami, Florida; ⁹National Alzheimer's Coordinating Center, University of Washington, Seattle, Washington; ¹⁰Department of Neurological Sciences, Rush University Medical Center, Chicago, Illinois; ¹¹Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, Illinois; ¹²Department of Pathology, Northwestern University, Chicago, Illinois; ¹³Department of Neurology, University of Washington, Seattle, Washington; ¹⁴Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts; ¹⁵Department of Psychiatry, Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts; ¹⁶Department of Neurology, Mayo Clinic, Rochester, Minnesota; ¹⁷Swedish Medical Center, Seattle, Washington; ¹⁸Department of Neurology, University of California San Francisco, San Francisco, California; ¹⁹Department of Medicine, Duke

University, Durham, North Carolina; ²⁰Department of Neuroscience, Mount Sinai School of Medicine, New York, New York; ²¹Department of Psychiatry, Mount Sinai School of Medicine, New York, New York; ²²Departments of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, New York; ²³Department of Pathology and Immunology, Washington University, St. Louis, Missouri; ²⁴Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania; ²⁵Byrd Alzheimer Institute, University of Southern Florida Health, Tampa, Florida; ²⁶Fred Hutchinson Cancer Research Center, Seattle, Washington; ²⁷Department of Psychiatry, Vanderbilt University, Nashville, Tennessee; ²⁸Department of Pathology, University of Alabama at Birmingham, Birmingham, Alabama; ²⁹Department of Neurology, University of Southern California, Los Angeles, California; ³⁰Department of Neurology, University of Alabama at Birmingham, Birmingham, Alabama; ³¹Neurogenomics Division, Translational Genomics Research Institute, Phoenix, Arizona; ³²Institute for Memory Impairments and Neurological Disorders, University of California Irvine, Irvine, California; ³³Department of Medicine, University of Washington, Seattle, Washington; ³⁴Department of Psychiatry and Hope Center Program on Protein Aggregation and Neurodegeneration, Washington University School of Medicine, St. Louis, Missouri; ³⁵Program in Translational NeuroPsychiatric Genomics, Department of Neurology, Brigham and Women's Hospital, Boston, Massachusetts; ³⁶Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts; ³⁷Department of Neurology, University of California Davis, Sacramento, California; ³⁸University of Virginia School of Medicine, Charlottesville, Virginia; ³⁹Department of Neurology, University of Texas Southwestern, Dallas, Texas; ⁴⁰Wien Center for Alzheimer's Disease and Memory Disorders, Mount Sinai Medical Center, Miami Beach, Florida; ⁴¹Department of Pathology and Laboratory Medicine, University of California Davis, Sacramento, California; ⁴²Rush Institute for Healthy Aging, Department of Internal Medicine, Rush University Medical Center, Chicago, Illinois; ⁴³Department of Neurology, Indiana University, Indianapolis, Indiana; ⁴⁴Department of Psychiatry, New York University, New York, New York; ⁴⁵C.S. Kubik Laboratory for Neuropathology, Massachusetts General Hospital, Charlestown, Massachusetts; ⁴⁶Department of Neurosciences, University of California San Diego, La Jolla, California; ⁴⁷Department of Psychiatry, University of Pittsburgh, Pittsburgh, Pennsylvania; ⁴⁸Department of Pathology and Laboratory Medicine, Emory University, Atlanta, Georgia; ⁴⁹Emory Alzheimer's Disease Center, Emory University, Atlanta, Georgia; ⁵⁰Neurogenetics Program, University of California Los Angeles, Los Angeles, California; ⁵¹Department of Pathology and Laboratory Medicine, Indiana University, Indianapolis, Indiana; ⁵²Department of Neurology, University of Michigan, Ann Arbor, Michigan; ⁵³Department of Psychiatry, University of Michigan, Ann Arbor, Michigan; ⁵⁴Department of Neurology, Emory University, Atlanta, Georgia; ⁵⁵Department of Neurology, Boston University, Boston, Massachusetts; ⁵⁶Department of Epidemiology, Boston University, Boston, Massachusetts; ⁵⁷Department of Neurology, Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts; ⁵⁸Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania; ⁵⁹Department of Pathology (Neuropathology), University of Pittsburgh, Pittsburgh, Pennsylvania; ⁶⁰Institute of Neurology, University College London, Queen Square, London; ⁶¹Department of Molecular and Biomedical Pharmacology, University of California Irvine, Irvine, California; ⁶²Taub Institute on Alzheimer's Disease and the Aging Brain, Department of Neurology, Columbia University, New York, New York; ⁶³Department of Pathology, Duke University, Durham, North Carolina; ⁶⁴Department of Genome Sciences, University of Washington, Seattle, Washington; ⁶⁵Department of Medicine (Medical Genetics), University of Washington, Seattle, Washington; ⁶⁶Department of Neurology, University of Kentucky, Lexington, Kentucky; ⁶⁷Department of Psychiatry and Behavioral Sciences, Northwestern University, Chicago, Illinois; ⁶⁸Department of Biostatistics, Boston University, Boston, Massachusetts; ⁶⁹Department of Ophthalmology, Boston University, Boston, Massachusetts; ⁷⁰University of Pittsburgh Alzheimer's Disease Research Center, Pittsburgh, Pennsylvania; ⁷¹Department of Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania; ⁷²Department of Biology, Brigham Young University, Provo, Utah; ⁷³Department of Neurology, Oregon Health & Science University, Portland, Oregon; ⁷⁴Department of Biomedical Engineering, Oregon Health & Science University, Portland, Oregon; ⁷⁵Department of Pathology and Laboratory Medicine, University

of California Irvine, Irvine, California; ⁷⁶Department of Pathology, Boston University, Boston, Massachusetts; ⁷⁷Department of Molecular & Medical Genetics, Oregon Health & Science University, Portland, Oregon; ⁷⁸Department of Epidemiology, University of Washington, Seattle, Washington; ⁷⁹Group Health Research Institute, Seattle, Washington; ⁸⁰Department of Pathology, University of Michigan, Ann Arbor, Michigan; ⁸¹Department of Preventive Medicine, University of Southern California, Los Angeles, California; ⁸²Department of Medicine - Pulmonary, New York University, New York, New York; ⁸³Department of Neurology, University of Miami, Miami, Florida; ⁸⁴Department of Pathology, University of California San Diego, La Jolla, California; ⁸⁵School of Nursing Northwest Research Group on Aging, University of Washington, Seattle, Washington; ⁸⁶Alzheimer's Disease Center, Northwestern University, Chicago, Illinois; ⁸⁷Cognitive Neurology, Northwestern University, Chicago, Illinois; ⁸⁸Department of Pathology, University of Southern California, Los Angeles, California; ⁸⁹Department of Pathology, University of Washington, Seattle, Washington; ⁹⁰Department of Neurology, Washington University, St. Louis, Missouri; ⁹¹Department of Psychiatry & Behavioral Sciences, University of Miami, Miami, Florida; ⁹²Department of Anatomic Pathology, Mayo Clinic, Rochester, Minnesota; ⁹³Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota; ⁹⁴Department of Pathology, Mount Sinai School of Medicine, New York, New York; ⁹⁵Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, Washington; ⁹⁶Department of Psychiatry, University of Arizona, Phoenix, Arizona; ⁹⁷Arizona Alzheimer's Consortium, Phoenix, Arizona; ⁹⁸Banner Alzheimer's Institute, Phoenix, Arizona; ⁹⁹Alzheimer's Disease Center, New York University, New York, New York; ¹⁰⁰Taub Institute on Alzheimer's Disease and the Aging Brain, Columbia University, New York, New York; ¹⁰¹Gertrude H. Sergievsky Center, Columbia University, New York, New York; ¹⁰²Department of Neurology, Columbia University, New York, New York; ¹⁰³Tanz Centre for Research in Neurodegenerative Disease,

University of Toronto, Toronto, Ontario; ¹⁰⁴Department of Radiology and Imaging Sciences, Indiana University, Indianapolis, Indiana; ¹⁰⁵Department of Pathology (Neuropathology), Rush University Medical Center, Chicago, Illinois; ¹⁰⁶Department of Psychiatry, University of Southern California, Los Angeles, California; ¹⁰⁷Department of Pathology, Columbia University, New York, New York; ¹⁰⁸Cambridge Institute for Medical Research and Department of Clinical Neurosciences, University of Cambridge, Cambridge, Massachusetts; ¹⁰⁹Department of Pathology, Johns Hopkins University, Baltimore, Maryland; ¹¹⁰Department of Pathology & Laboratory Medicine, University of California Los Angeles, Los Angeles, California; ¹¹¹Taub Institute on Alzheimer's Disease and the Aging Brain, Department of Pathology, Columbia University, New York, New York; ¹¹²Department of Psychiatry & Behavioral Sciences, Duke University, Durham, North Carolina; ¹¹³Department of Pathology, Oregon Health & Science University, Portland, Oregon; ¹¹⁴Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, Indiana.

Author Contributions

Conceived and designed the experiments: SG Younkin, N Ertekin-Taner. Performed the experiments: F Zou, MM Carrasquillo, CN Rowley, T Nguyen, L Ma, KG Malphrus, R Palusak, S Lincoln, G Bisceglia, N Kouri, CP Kolbert, J Jen. Analyzed the data: F Zou, HS Chai, CS Younkin, J Crook, VS Pankratz, M Allen, MM Carrasquillo, AA Nair, S Middha, S Maharjan, C Gorgescu, SG Younkin, N Ertekin-Taner. Contributed reagents/materials/analysis tools: Alzheimer's Disease Genetics Consortium, JL Haines, R Mayeux, MA Pericak-Vance, LA Farrer, GD Schellenberg, RC Petersen, NR Graff-Radford, DW Dickson, SG Younkin. Wrote the paper: F Zou, HS Chai, CS Younkin, J Crook, VS Pankratz, M Allen, SG Younkin, N Ertekin-Taner.

References

- Ertekin-Taner N (2011) Gene expression endophenotypes: a novel approach for gene discovery in Alzheimer's disease. *Mol Neurodegener* 6: 31.
- Murphy A, Chu JH, Xu M, Carey VJ, Lazarus R, et al. (2010) Mapping of numerous disease-associated expression polymorphisms in primary peripheral blood CD4+ lymphocytes. *Hum Mol Genet* 19: 4745–4757.
- Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, et al. (2010) Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. *PLoS Genet* 6: e1000888. doi:10.1371/journal.pgen.1000888.
- Schadt EE, Monks SA, Drake TA, Lusk AJ, Che N, et al. (2003) Genetics of gene expression surveyed in maize, mouse and man. *Nature* 422: 297–302.
- Cheung VG, Conlin LK, Weber TM, Arcaro M, Jen KY, et al. (2003) Natural variation in human gene expression assessed in lymphoblastoid cells. *Nat Genet* 33: 422–425.
- Morley M, Molony CM, Weber TM, Devlin JL, Ewens KG, et al. (2004) Genetic analysis of genome-wide variation in human gene expression. *Nature* 430: 743–747.
- Monks SA, Leonardson A, Zhu H, Cundiff P, Pietrusiak P, et al. (2004) Genetic inheritance of gene expression in human cell lines. *Am J Hum Genet* 75: 1094–1105.
- Cheung VG, Spielman RS, Ewens KG, Weber TM, Morley M, et al. (2005) Mapping determinants of human gene expression by regional and genome-wide association. *Nature* 437: 1363–1369.
- Stranger BE, Forrest MS, Clark AG, Minichiello MJ, Deutsch S, et al. (2005) Genome-wide associations of gene expression variation in humans. *PLoS Genet* 1: e78. doi:10.1371/journal.pgen.0010078.
- Stranger BE, Nica AC, Forrest MS, Dimas A, Bird CP, et al. (2007) Population genomics of human gene expression. *Nat Genet* 39: 1217–1224.
- Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, et al. (2007) Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* 315: 848–853.
- Dixon AL, Liang L, Moffatt MF, Chen W, Heath S, et al. (2007) A genome-wide association study of global gene expression. *Nat Genet* 39: 1202–1207.
- Dausset J, Cann H, Cohen D, Lathrop M, Lalouel JM, et al. (1990) Centre d'étude du polymorphisme humain (CEPH): collaborative genetic mapping of the human genome. *Genomics* 6: 575–577.
- (2005) A haplotype map of the human genome. *Nature* 437: 1299–1320.
- Stranger BE, Forrest MS, Clark AG, Minichiello MJ, Deutsch S, et al. (2005) Genome-wide associations of gene expression variation in humans. *PLoS Genet* 1: e78. doi:10.1371/journal.pgen.0010078.
- Goring HH, Curran JE, Johnson MP, Dyer TD, Charlesworth J, et al. (2007) Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. *Nat Genet* 39: 1208–1216.
- Rotival M, Zeller T, Wild PS, Mouché S, Szymczak S, et al. (2011) Integrating genome-wide genetic variations and monocyte expression data reveals trans-regulated gene modules in humans. *PLoS Genet* 7: e1002367. doi:10.1371/journal.pgen.1002367.
- Dimas AS, Deutsch S, Stranger BE, Montgomery SB, Borel C, et al. (2009) Common regulatory variation impacts gene expression in a cell type-dependent manner. *Science* 325: 1246–1250.
- Ding J, Gudjonsson JE, Liang L, Stuart PE, Li Y, et al. (2010) Gene expression in skin and lymphoblastoid cells: Refined statistical method reveals extensive overlap in cis-eQTL signals. *Am J Hum Genet* 87: 779–789.
- Emilsson V, Thorleifsson G, Zhang B, Leonardson AS, Zink F, et al. (2008) Genetics of gene expression and its effect on disease. *Nature* 452: 423–428.
- Zhong H, Beaulaurier J, Lum PY, Molony C, Yang X, et al. (2010) Liver and adipose expression associated SNPs are enriched for association to type 2 diabetes. *PLoS Genet* 6: e1000932. doi:10.1371/journal.pgen.1000932.
- Grundberg E, Adoue V, Kwan T, Ge B, Duan QL, et al. (2011) Global analysis of the impact of environmental perturbation on cis-regulation of gene expression. *PLoS Genet* 7: e1001279. doi:10.1371/journal.pgen.1001279.
- Schadt EE, Molony C, Chudin E, Hao K, Yang X, et al. (2008) Mapping the genetic architecture of gene expression in human liver. *PLoS Biol* 6: e107. doi:10.1371/journal.pbio.0060107.
- Myers AJ, Gibbs JR, Webster JA, Rohrer K, Zhao A, et al. (2007) A survey of genetic human cortical gene expression. *Nat Genet* 39: 1494–1499.
- Webster JA, Gibbs JR, Clarke J, Ray M, Zhang W, et al. (2009) Genetic control of human brain transcript expression in Alzheimer disease. *Am J Hum Genet* 84: 445–458.
- Hindorf LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, et al. (2009) Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A* 106: 9362–9367.
- Hoglinger GU, Melhem NM, Dickson DW, Sleiman PM, Wang LS, et al. (2011) Identification of common variants influencing risk of the tauopathy progressive supranuclear palsy. *Nat Genet* 43: 699–705.
- Naj AC, Jun G, Beecham GW, Wang LS, Vardarajan BN, et al. (2011) Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. *Nat Genet* 43: 436–441.
- Carrasquillo MM, Zou F, Pankratz VS, Wilcox SL, Ma L, et al. (2009) Genetic variation in PCDH11X is associated with susceptibility to late-onset Alzheimer's disease. *Nat Genet* 41: 192–198.
- Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A* 100: 9440–9445.
- Barry WT, Kernagis DN, Dressman HK, Griffin RJ, Hunter JD, et al. (2010) Intratumor heterogeneity and precision of microarray-based predictors of breast cancer biology and clinical outcome. *J Clin Oncol* 28: 2198–2206.

32. Simon-Sanchez J, Schulte C, Bras JM, Sharma M, Gibbs JR, et al. (2009) Genome-wide association study reveals genetic risk underlying Parkinson's disease. *Nat Genet* 41: 1308–1312.
33. Fuchs J, Nilsson C, Kachergus J, Munz M, Larsson EM, et al. (2007) Phenotypic variation in a large Swedish pedigree due to SNCA duplication and triplication. *Neurology* 68: 916–922.
34. Edwards TL, Scott WK, Almonte C, Burt A, Powell EH, et al. (2010) Genome-wide association study confirms SNPs in SNCA and the MAPT region as common risk factors for Parkinson disease. *Ann Hum Genet* 74: 97–109.
35. Han JW, Zheng HF, Cui Y, Sun LD, Ye DQ, et al. (2009) Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. *Nat Genet* 41: 1234–1237.
36. Graham RR, Kozyrev SV, Baechler EC, Reddy MV, Plenge RM, et al. (2006) A common haplotype of interferon regulatory factor 5 (IRF5) regulates splicing and expression and is associated with increased risk of systemic lupus erythematosus. *Nat Genet* 38: 550–555.
37. Graham RR, Kyogoku C, Sigurdsson S, Vlasova IA, Davies LR, et al. (2007) Three functional variants of IFN regulatory factor 5 (IRF5) define risk and protective haplotypes for human lupus. *Proc Natl Acad Sci U S A* 104: 6758–6763.
38. Anderson CA, Boucher G, Lees CW, Franke A, D'Amato M, et al. (2011) Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet* 43: 246–252.
39. Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, et al. (2009) Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat Genet*.
40. Lambert JC, Heath S, Even G, Campion D, Sleegers K, et al. (2009) Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. *Nat Genet* 41: 1094–1099.
41. Reiman EM, Webster JA, Myers AJ, Hardy J, Dunckley T, et al. (2007) GAB2 alleles modify Alzheimer's risk in APOE epsilon4 carriers. *Neuron* 54: 713–720.
42. van Es MA, Veldink JH, Saris CG, Blauw HM, van Vught PW, et al. (2009) Genome-wide association study identifies 19p13.3 (UNC13A) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. *Nat Genet* 41: 1083–1087.
43. Laaksovirta H, Peuralinna T, Schymick JC, Scholz SW, Lai SL, et al. (2010) Chromosome 9p21 in amyotrophic lateral sclerosis in Finland: a genome-wide association study. *Lancet Neurol* 9: 978–985.
44. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, et al. (2011) Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron* 72: 245–256.
45. Renton AE, Majounie E, Waite A, Simon-Sanchez J, Rollinson S, et al. (2011) A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* 72: 257–268.
46. Lee W, Glaeser H, Smith LH, Roberts RL, Moeckel GW, et al. (2005) Polymorphisms in human organic anion-transporting polypeptide 1A2 (OATP1A2): implications for altered drug disposition and central nervous system drug entry. *J Biol Chem* 280: 9610–9617.
47. Hollingworth P, Harold D, Sims R, Gerrish A, Lambert JC, et al. (2011) Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. *Nat Genet* 43: 429–435.
48. Seshadri S, Fitzpatrick AL, Ikram MA, DeStefano AL, Gudnason V, et al. (2010) Genome-wide analysis of genetic loci associated with Alzheimer disease. *JAMA* 303: 1832–1840.
49. Allen M, Zou F, Chai HS, Younkin CS, Crook J, et al. (2012) Novel late-onset Alzheimer's disease loci variants associate with brain gene expression. *Neurology* in press.
50. Atz M, Walsh D, Cartagena P, Li J, Evans S, et al. (2007) Methodological considerations for gene expression profiling of human brain. *J Neurosci Methods* 163: 295–309.
51. Braak H, Braak E (1991) Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol (Berl)* 82: 239–259.
52. Caffrey TM, Joachim C, Paracchini S, Esiri MM, Wade-Martins R (2006) Haplotype-specific expression of exon 10 at the human MAPT locus. *Hum Mol Genet* 15: 3529–3537.
53. Caffrey TM, Joachim C, Wade-Martins R (2008) Haplotype-specific expression of the N-terminal exons 2 and 3 at the human MAPT locus. *Neurobiol Aging* 29: 1923–1929.
54. Myers AJ, Pittman AM, Zhao AS, Rohrer K, Kaleem M, et al. (2007) The MAPT H1c risk haplotype is associated with increased expression of tau and especially of 4 repeat containing transcripts. *Neurobiol Dis* 25: 561–570.
55. McKhann G, Drachman D, Folstein M, Katzman R, Price D, et al. (1984) 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.
56. Du P, Kibbe WA, Lin SM (2008) lumi: a pipeline for processing Illumina microarray. *Bioinformatics* 24: 1547–1548.
57. Lin SM, Du P, Huber W, Kibbe WA (2008) Model-based variance-stabilizing transformation for Illumina microarray data. *Nucleic Acids Res* 36: e11.
58. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81: 559–575.
59. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, et al. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38: 904–909.
60. Westfall PH, Zaykin DV, Young SS (2002) Multiple tests for genetic effects in association studies. *Methods Mol Biol* 184: 143–168.
61. Willer CJ, Li Y, Abecasis GR (2010) METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 26: 2190–2191.