

Alzheimer's Disease: Analyzing the Missing Heritability

Perry G. Ridge^{1,2}, Shubhabrata Mukherjee³, Paul K. Crane³, John S. K. Kauwe^{1*}, Alzheimer's Disease Genetics Consortium¹

1 Department of Biology, Brigham Young University, Provo, Utah, United States of America, **2** ARUP Institute for Clinical and Experimental Pathology, Salt Lake City, Utah, United States of America, **3** Department of Medicine, University of Washington, Seattle, Washington, United States of America

Abstract

Alzheimer's disease (AD) is a complex disorder influenced by environmental and genetic factors. Recent work has identified 11 AD markers in 10 loci. We used Genome-wide Complex Trait Analysis to analyze >2 million SNPs for 10,922 individuals from the Alzheimer's Disease Genetics Consortium to assess the phenotypic variance explained first by known late-onset AD loci, and then by all SNPs in the Alzheimer's Disease Genetics Consortium dataset. In all, 33% of total phenotypic variance is explained by all common SNPs. *APOE* alone explained 6% and other known markers 2%, meaning more than 25% of phenotypic variance remains unexplained by known markers, but is tagged by common SNPs included on genotyping arrays or imputed with HapMap genotypes. Novel AD markers that explain large amounts of phenotypic variance are likely to be rare and unidentifiable using genome-wide association studies. Based on our findings and the current direction of human genetics research, we suggest specific study designs for future studies to identify the remaining heritability of Alzheimer's disease.

Citation: Ridge PG, Mukherjee S, Crane PK, Kauwe JSK, (2013) Alzheimer's Disease: Analyzing the Missing Heritability. PLoS ONE 8(11): e79771. doi: 10.1371/journal.pone.0079771

Editor: Hemant K. Paudel, McGill University Department of Neurology and Neurosurgery, Canada

Received: July 17, 2013; **Accepted:** September 27, 2013; **Published:** November 7, 2013

Copyright: © 2013 Ridge 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 project was provided by the National Institutes of Health (R01AG042611), the Alzheimer's Association (MNIRG-11-205368) the Charleston Conference on Alzheimer's disease and the Brigham Young University Gerontology Program. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Contributors to the genetic analysis data included Principal Investigators on projects that were individually funded by NIA, other NIH institutes, or private entities. 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 P30 AG019610; Boston University, P30 AG013846, U01 AG10483, R01 CA129769, R01 MH080295, R01 AG017173, R01 AG025259, R01AG33193; Columbia University, P50 AG008702, R37 AG015473; Duke University, P30 AG028377, AG05128; Emory University, AG025688; Group Health Research Institute, U01 AG06781, U01 HG004610; 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, UL1 RR029893, 5R01AG012101, 5R01AG022374, 5R01AG013616, 1RC2AG036502, 1R01AG035137; Northwestern University, P30 AG013854; Oregon Health & Science University, P30 AG008017, R01 AG026916; Rush University, P30 AG010161, R01 AG019085, R01 AG15819, R01 AG17917, R01 AG30146; TGen, R01 NS059873; University of Alabama at Birmingham, P50 AG016582, UL1RR02777; University of Arizona, R01 AG031581; 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, AG05144; University of Michigan, P50 AG008671; University of Pennsylvania, P30 AG010124; 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; Vanderbilt University, R01 AG019085; and Washington University, P50 AG005681, P01 AG03991. The Kathleen Price Bryan Brain Bank at Duke University Medical Center is funded by NINDS grant # NS39764, NIMH MH60451 and by Glaxo Smith Kline. Genotyping of the TGEN2 cohort was supported by Kronos Science. The TGen series was also funded by NIA grant AG034504 to AJM, The Banner Alzheimer's Foundation, The Johnnie B. Byrd Sr. Alzheimer's Institute, the Medical Research Council, and the state of Arizona and also includes samples from the following sites: Newcastle Brain Tissue Resource (funding via the Medical Research Council, local NHS trusts and Newcastle University), MRC London Brain Bank for Neurodegenerative Diseases (funding via the Medical Research Council), 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), 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), Institut de Neuropatologia, Servei Anatomia Patologica, Universitat de Barcelona. 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 Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly and Co., Medpace, Inc., Merck and Co., Inc., Novartis AG, Pfizer Inc, F. Hoffman-La Roche, Schering-Plough, Synarc, Inc., 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. The authors thank Drs. D. Stephen Snyder 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 US Department of Veterans Affairs Administration, Office of Research and Development, Biomedical Laboratory Research Program. P.S.G.-H. is supported by Wellcome Trust, Howard Hughes Medical Institute, and the Canadian Institute of Health Research.

Competing interests: The authors have declared that no competing interests exist.

* E-mail: kauwe@byu.edu

¶ Membership of the Alzheimer's Disease Genetics Consortium is provided in the Acknowledgments.

Introduction

Alzheimer's disease (AD) is the most common form of dementia. Worldwide estimates of prevalence vary, with estimates of 24 to 35 million people affected [1-3]. Combined with an aging population, prevalence is expected to increase to 1 in 85 people by 2050 [2].

AD is a heterogeneous disease caused by a combination of environmental and genetic factors. The most important risk factor for Alzheimer's disease is age [1,4]. Environmental risk factors include hypertension, estrogen supplements [5], smoking [6,7], stroke, heart disease, depression, arthritis, and diabetes [8]. In addition, certain lifestyle choices appear to decrease the risk of AD: exercise [9], intellectual stimulation [10], and maintaining a Mediterranean diet (including fish) [11,12].

The genetics of AD are complex. Several genes are known to harbor either causative or risk variants for AD. There are two primary types of AD as defined by age. The first is early-onset, or familial AD, and the second type is late-onset AD (LOAD), sometimes termed sporadic AD. Three genes, *APP* [13], *PSEN1* [14], and *PSEN2* [15] are known to harbor many highly penetrant, autosomal dominantly-inherited variants, which lead to early-onset AD but account for only a small fraction of total AD cases.

LOAD accounts for 99% of AD cases and is caused by a more complex underlying genetic architecture. Genome-wide association studies (GWAS) have identified 10 different loci associated with AD (Table 1). Recent applications of next-generation sequencing (NGS) have suggested rare variants play important role and have large effects in the etiology of AD [16-18]. Identifying additional variants will provide information that is integral to the development, evaluation and application of effective therapeutic strategies for AD. Lee et al. [19] used 3,333 cases and 3,924 controls, including 2,699 population-based controls to estimate that common genetic variants account for 24% of variance in AD. They also estimated the contribution of APOE using several proxy SNPs, with varying degrees of LD, with the APOE $\epsilon 4$ allele to estimate the APOE effect at approximately 4%. Here we evaluate the variance in AD status explained by common SNPs and along with all recently identified AD genes, including direct genotyping of the APOE $\epsilon 2$ and $\epsilon 4$ alleles, in 5,708 AD cases and 5,214 clinically ascertained controls. We also suggest strategies for identifying the remaining AD genes.

Methods

Dataset

We used the Alzheimer's Disease Genetics Consortium (ADGC) dataset described in Naj et al. [20] for our analyses. Samples were genotyped using Affymetrix and Illumina SNP chips. Quality control of the imputed data was performed as described by Naj et al. 2011 [20]. Briefly, markers with a minor allele frequency of less than 1% and deviation from HWE

Table 1. Late-onset Alzheimer's disease associated genes/variants.

Variant	Gene	Abbreviation	Odds Ratio
rs429358	Apolipoprotein E ($\epsilon 4$ allele)[58]	APOE	3.685
rs7412	Apolipoprotein E ($\epsilon 2$ allele)[59]	APOE	0.621
rs744373	Bridging Integrator 1[32]	BIN1	1.166
rs11136000	Clusterin[33,60]	CLU	0.879
rs3764650	ATP-binding cassette, sub-family A (ABC1), member 7[31]	ABCA7	1.229
rs3818361	Complement component (3b/4b) receptor 1 (Knops blood group)[60]	CR1	1.174
rs3851179	Phosphatidylinositol binding clathrin assembly protein[33]	PICALM	0.879
rs610932	Membrane-spanning 4-domains, subfamily A, member 6A[31]	MS4A6A	0.904
rs3865444	CD33 molecule[20,31]	CD33	0.893
rs670139	Membrane-spanning 4-domains, subfamily A, member 4E[31]	MS4A4E	1.079
rs9296559	CD2-associated protein[20,31]	CD2AP	1.117

The dbSNP identification number, gene name, gene abbreviation, and odds ratio for each of the top variants from the Alzgene.org meta-analyses. The SNP in CD2AP, rs9349407, is not present in this sample, so we used rs9296559 instead as a proxy. These two SNPs are close together (1,108 base pairs apart) and in very high LD ($r^2=1$).

doi: 10.1371/journal.pone.0079771.t001

where $P < 10^{-6}$ were removed. To have a common set of SNPs across all samples, imputation to HapMap phase 2 (release 22) [21] was performed using MaCH [22] and strand ambiguous SNPs were removed, resulting in a rectangular dataset with 2,042,114 SNPs. Only SNPs imputed with $R^2 \geq 0.50$ were included in the dataset. We added an additional two SNPs, rs7412 and rs429358, corresponding to APOE $\epsilon 2$ and $\epsilon 4$, respectively.

We used a compiled dataset of directly genotyped SNPs common to all 15 studies to assess cryptic relatedness and calculate principal components to account for population-specific variations in allele distribution. We excluded strand ambiguous SNPs, resulting in a rectangular dataset with 21,880 directly observed (not imputed) SNPs in common across all the studies. We filtered SNPs with pairwise LD ($r^2 < 0.20$), resulting in a dataset with 17,054 SNPs. We used both PLINK [23] and KING-ROBUST [24] for relatedness analysis. KING-ROBUST provided unbiased kinship coefficient estimates for related individuals in our dataset. We excluded up to 3rd degree relatives (kinship ≥ 0.0442) for a final dataset containing 19,692 individuals.

Of the 19,692 individuals in the original dataset we analyzed a subset of 10,922 individuals who had complete data for the 11 markers listed in Table 1, AD case-control status, age, sex, and 10 principal components from the population stratification analysis (missingness rates for each of the covariates and

case-control status are reported in Table S1). Basic demographic information for the 10,922 individuals in the subset of the dataset used in this study is presented in Table 2. We collected chromosome length and number of genes per chromosome from the Vega database [25].

Genetic Analyses

We used Genome-wide Complex Trait Analysis (GCTA)[26], a tool that implements the methods described in Yang et al. [27], Lee et al. [28], and Yang et al. [29] to estimate the phenotypic variance explained by known AD genes and tagged by SNPs on the SNP arrays. Briefly, GCTA uses a mixed linear model and treats the effects of SNPs as random effects, effectively testing all the SNPs together for effect (in contrast to GWAS, which considers each SNP individually). We used age, sex, and 10 principal components as covariates. For the analyses in which we examined unexplained phenotypic variance, we also controlled for the 11 known AD markers (Table 1). The 11 known AD markers are the AlzGene.org top hits and are the markers with replicable evidence for association with AD. Each of these markers is present in our dataset except rs9349407 in CD2AP. As proxy we used rs9296559, which is in very high LD with rs9349407 ($r^2=1$). We specified a population prevalence of LOAD at 0.13 [30].

Ethics Statement

All study procedures were approved by the Institutional Review Boards of Brigham Young University and the University of Washington.

Results

We estimated the variance in AD case-control status focusing first on the 11 known AD markers (Table 1). Together these markers account for 7.8% (standard error 0.03) of the phenotypic variance (Table 3). Next, we estimated the explained phenotypic variance for each chromosome (Figure 1). Chromosome 19 accounts for the highest proportion of phenotypic variance.

In all, the 2,042,116 SNPs in the HapMap imputed ADGC dataset explain 33.1% of phenotypic variance (genetic variance of 0.0711, standard error 0.0072). The APOE $\epsilon 2$ and $\epsilon 4$ alleles account for 5.9% (standard error 0.03) of the phenotypic variance (Table 3). The other 9 known high frequency SNPs identified in GWAS explain an additional 1.9% (standard error 0.03)(Table 3). After controlling for these 11 markers, an additional 25.3% of the total phenotypic variance (genetic variance of 0.046, standard error 0.006) is explained with as-yet unidentified variants (Table 3). The remaining phenotypic variance explained by each chromosome after controlling for the 11 known markers is shown in Figure 2. SNPs on chromosomes 1, 4, 5 and 17 account for the largest percentage of remaining unexplained phenotypic variance compared to other chromosomes, each accounting for more than 2% (Figure 2). Chromosomes 9, 14, and 21 account for the least (<0.0001% each); however, there is unexplained phenotypic variance on all the autosomes. There is no relationship between explained variance and chromosome

Table 2. Demographic information for individuals in the analysis dataset.

	Count	Cases / Controls	Age (Cases / Controls)
Male	4489	2378/2111	74.8 (73.6 / 76.1)
Female	6433	3330/3103	75.3 (74.9 / 75.7)
Total	10922	5708/5214	75.1 (74.3 / 75.9)

We report total individuals, sex, case-control status, and average age for the 10,922 individuals analyzed in this report.

doi: 10.1371/journal.pone.0079771.t002

Table 3. Summary of genetic and phenotypic variance measurements.

SNP Set	Genetic Variance (Standard Error)	% Phenotypic Variance (Standard Error)
All SNPs	0.071 (0.0072)	33.12% (0.033)
APOE $\epsilon 2/\epsilon 4$	0.020 (0.0066)	5.92% (0.033)
All known AD markers (Table 1)	0.025 (0.0066)	7.78% (0.033)
All SNPs excluding known AD markers	0.046 (0.0060)	25.34% (0.033)

In this table we summarize our results, showing genetic and percent phenotypic variance for four different subsets of SNPs: all SNPs in the ADGC dataset, the two APOE alleles ($\epsilon 2$ and $\epsilon 4$), all known AD markers (as listed in Table 1), and all SNPs excluding those in Table 1.

doi: 10.1371/journal.pone.0079771.t003

length (p-value = 0.8), or number of genes per chromosome (p-value = 0.7).

Discussion

A clear understanding of the genetic architecture of Alzheimer's disease provides the foundation of information needed to cure this terrible disease. While many large GWAS for AD have been performed and several replicable loci have been identified (as referenced in Table 1), relatively little phenotypic variance is explained by these variants. Our estimates of phenotype variance explained by common genetic variants and by the APOE locus are higher than those of Lee et al. [19]. We estimated total phenotypic variance explained by common SNPs to be ~33%. In contrast Lee et al. [19] estimated ~24%. In our study we used genotyped and HapMap imputed SNPs, whereas Lee et al. [19] used only directly genotyped SNPs. Inclusion of imputed SNPs improved heritability estimates and suggests that imputed SNPs should be included in such studies. In addition to using imputed variants, our dataset was larger and our controls were clinically ascertained. Differences in the estimates for APOE (~6% in this study compared to ~4% in Lee et al. [19]) could be due to these same characteristics as well as the direct genotyping of the APOE $\epsilon 2$ and $\epsilon 4$ alleles in our samples as opposed to the use of proxy markers. Regardless, both studies provide evidence that a considerable amount of variance in AD is explained by

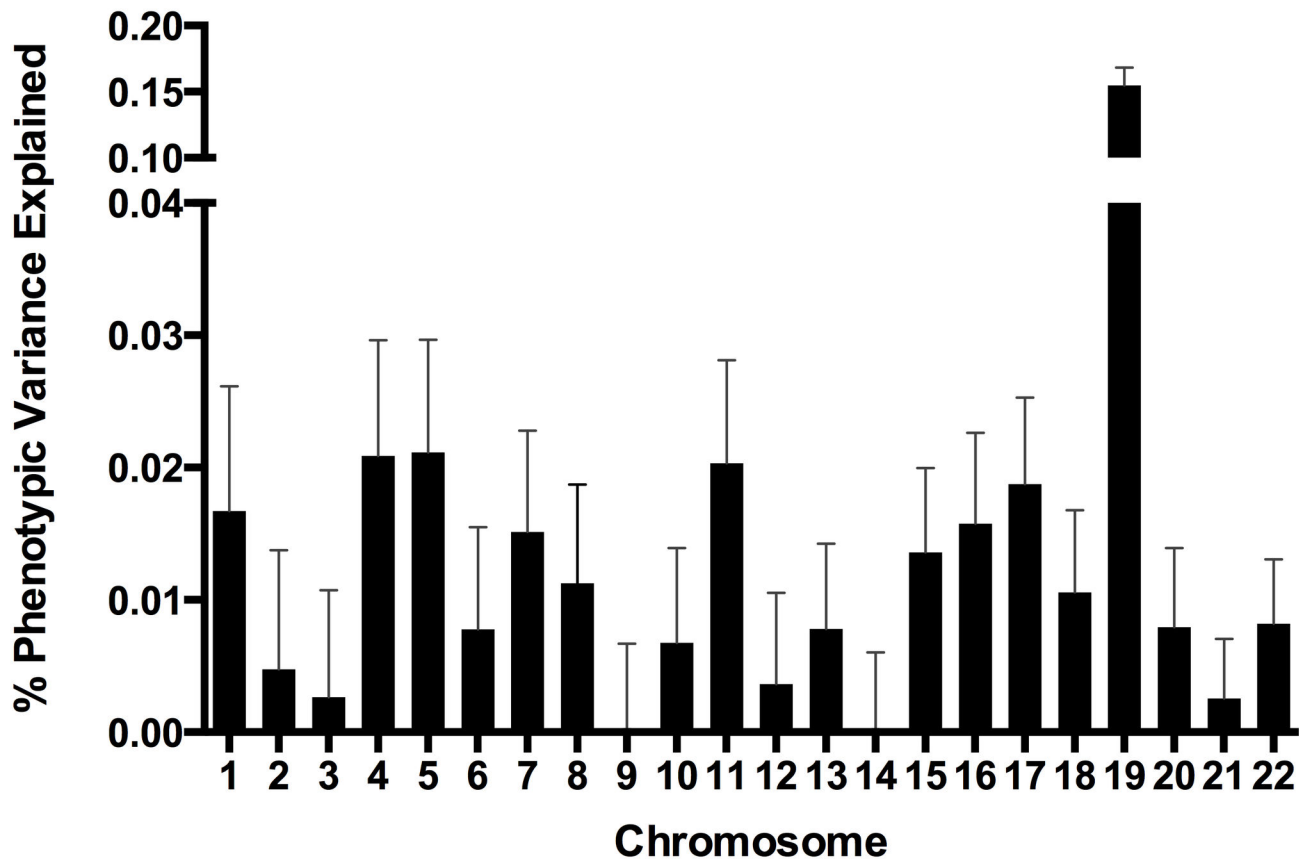


Figure 1. Unexplained Alzheimer's disease variance, by chromosome. In this figure we show phenotypic variance, by chromosome, explained by all SNPs. Error bars correspond to standard error.

doi: 10.1371/journal.pone.0079771.g001

yet unidentified genetic variation. Together these results show that there is clearly much work to be done if we are to solve the genetic architecture of AD. GWAS with sample sizes performed to date are able to identify common variants with moderate to small effect sizes. Results of GWAS in AD and other conditions suggest there may be a large number of such variants and that additional variants of this type can be identified by increasing sample sizes. However, additional loci detected with the GWAS strategy will likely have effects either similar to or smaller than SNPs already identified. The range of SNPs identifiable by current GWAS [20,31-36] is marked on Figure 3 by the large box bordered by dots (the GWAS search space), with recent GWAS hits inside the labeled oval. The GWAS being conducted by the International Genomics of Alzheimer's Project represents a substantial increase in sample size and will undoubtedly identify additional common loci with small effects on AD risk. Nevertheless, it is unlikely that many common variants of even modest effect size remain to be identified.

There are still many AD variants that remain to be identified, however, and these variants exist on every autosome (Figure 2). Variants with large effects are almost certainly present in very low frequencies or they would have been identified in

GWAS. While such variants are unlikely to be detected using traditional GWAS due to limitations of r^2 based "tagging" for alleles with different frequencies [37] the current analysis allows for high D' values between common alleles and rare variants of large effect to contribute to the explained variance. These rare variants of large effect appear in the smaller box bordered with dashes in Figure 3. To date, identified alleles of this type have clear functional effects and large effect sizes compared to associated alleles from GWAS. Detecting rare variants of large effect requires different experimental designs than GWAS such as sequencing causal loci. Exome chip array studies target known variation in coding regions, even those of very low frequency; this may prove a promising and economical approach. However, accurately genotyping variants of less than 1-2% using these arrays is quite challenging, and for variants that are present below these frequencies other approaches are required.

Two seemingly contradictory hypotheses exist about the architecture of complex disease: the common disease/common variant hypothesis and the multiple rare variant hypothesis. In the first, many common variants of small effect size collectively explain disease risk, while in the second, rare variants, some with large effect and high penetrance, explain disease risk.

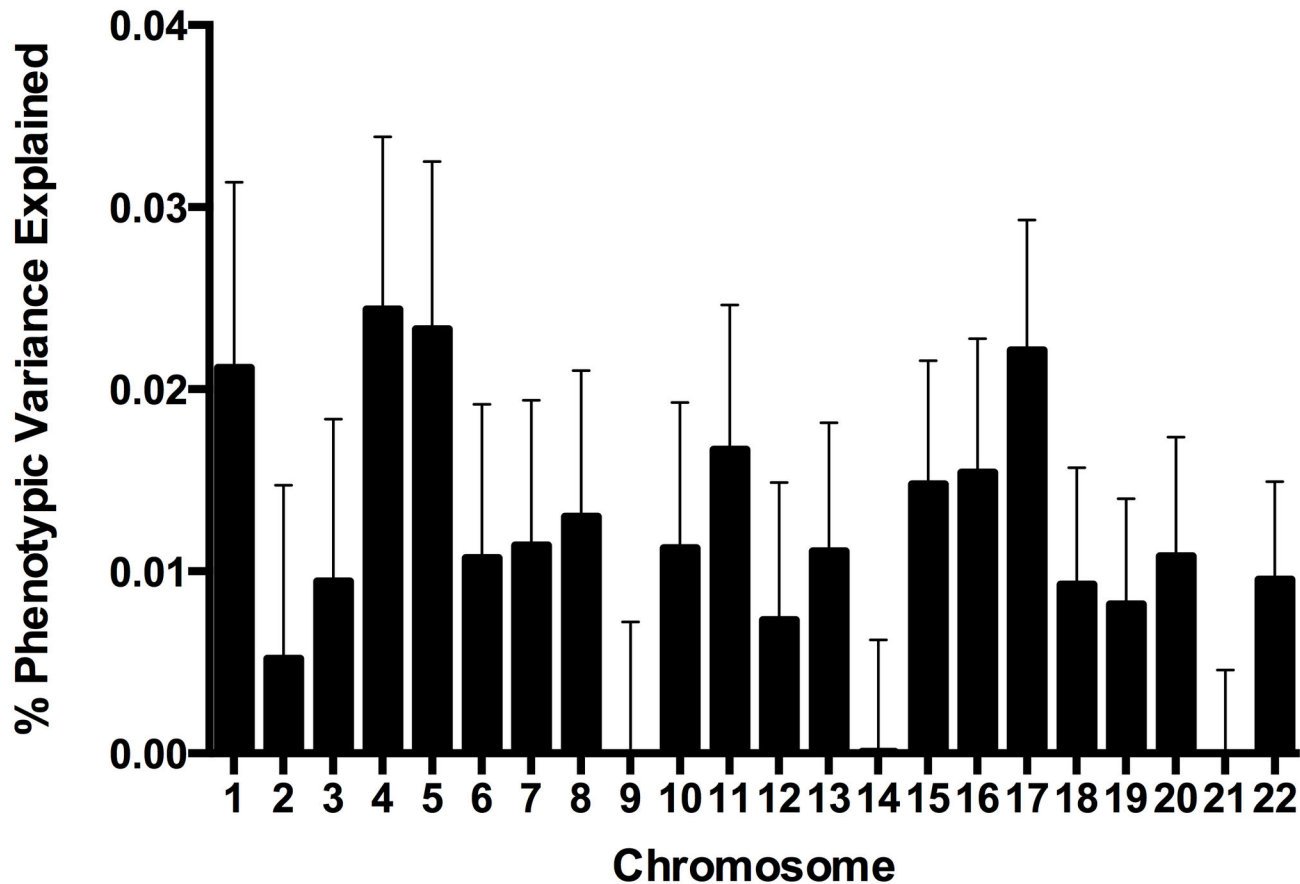


Figure 2. Unexplained Alzheimer's disease variance, by chromosome, excluding known Alzheimer's disease markers. This figure is the same as Figure 1 except we have removed variance explained by known Alzheimer's disease markers. Error bars correspond to standard error.

doi: 10.1371/journal.pone.0079771.g002

However, as suggested by Singleton et al. [38] these two hypotheses are not mutually exclusive and the genetics of complex diseases are likely a hybrid of the two. Singleton et al. [38] suggests that both common and rare variants that increase or decrease disease risk are likely to be found in the same loci and coined the phrase “pleomorphic risk loci”. To date, AD genetics research has largely focused on common variants that influence disease risk, likely due to technological and financial constraints. However, the advent of next generation sequencing (NGS) and falling costs of this technology have made it possible to expand AD research to include searching for rare variants. Recently, this technology was used to identify a functional variant that protects against Alzheimer's disease in the amyloid precursor protein (*APP*)[17]. Additionally, two groups recently used NGS to identify additional, likely functional, variants associated with AD in the triggering receptor expressed on myeloid 2 (*TREM2*) gene [16,18]. The *TREM2* variant is present in about 1% of the general population and has a high odds ratio (2.9 to 5.1 depending on the dataset). Likewise, the *APP* variant is extremely rare (frequency of 0.00038), but confers a large protective effect on

carriers. Larger scale applications of this technology and careful study design are likely to identify additional variants and further explain the remaining phenotypic variance in AD.

Family-based studies are also an effective application of NGS. These studies require carefully ascertained families and accurate pedigree data and can be used to identify high effect, low frequency variants (located in the box with longer dashes in Figure 3). Family-based studies are especially powerful because large effect, low frequency disease-causing (or disease-modifying) sequence features, some of which may be unique to a single family, are likely to segregate, at least partially, with disease status. These approaches have not yet been extensively applied in AD research. Nevertheless, family-based studies utilizing large-scale genome or exome sequencing have recently been used to identify disease-causing variants in several Mendelian [39-41] and complex disorders [42,43].

It is also possible that gene-gene interactions account for much of the unexplained variance in AD status [44]. These interactions are widespread and common [45,46] and approaches to understand the effects of epistatic interactions

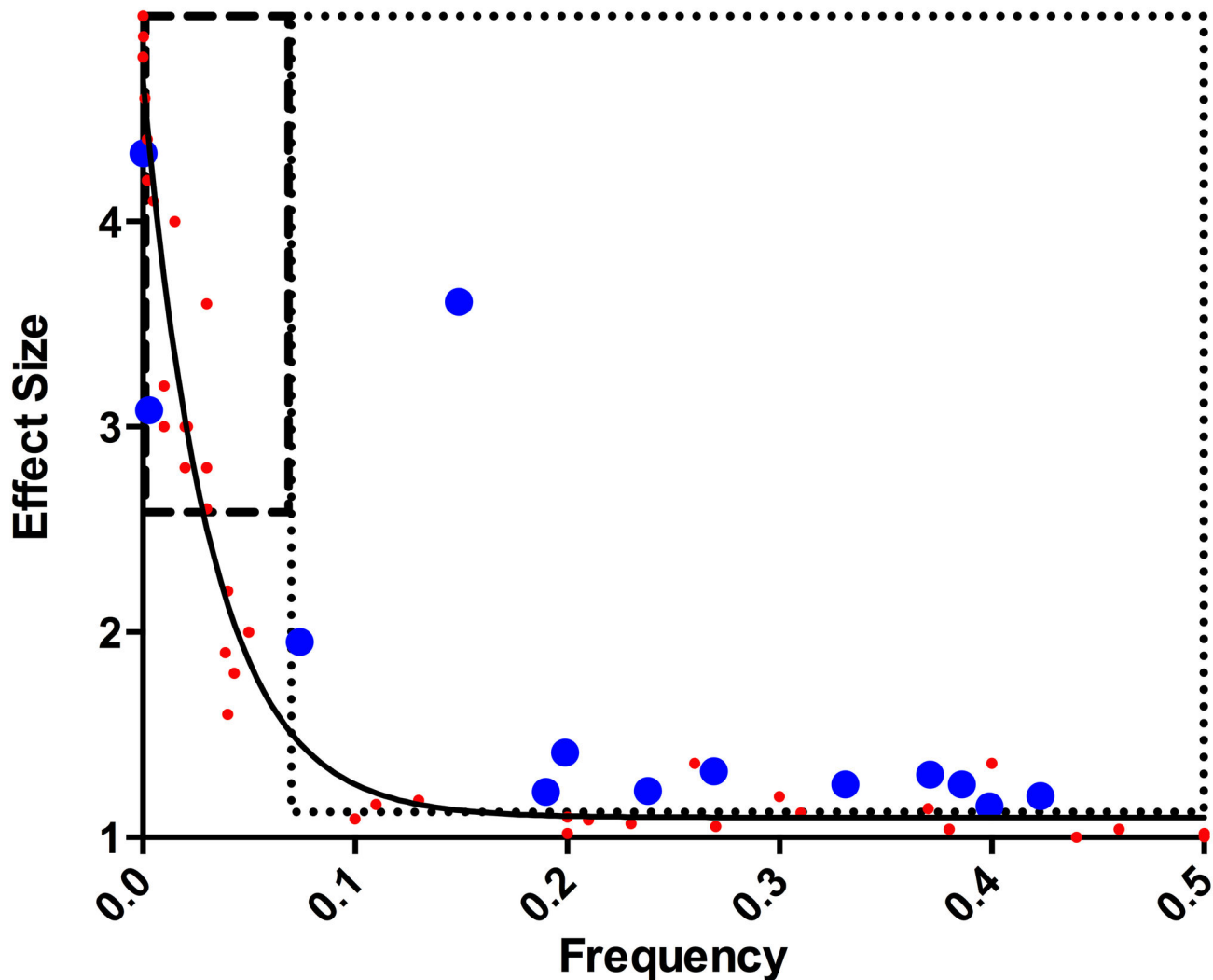


Figure 3. Variant search space. Real and hypothetical variants are graphed by effect size (y-axis) and population frequency (x-axis). Known Alzheimer's disease SNPs are blue circles and hypothetical SNPs are red circles. The large box on the right outlined with dots, is the GWAS search space and the smaller box on the left, outlined with dashes, is the next-generation sequencing search space. Known Alzheimer's disease SNPs are those found in Table 1 as well as APP and TREM2, which are both labeled on the graph.

doi: 10.1371/journal.pone.0079771.g003

exist and continue to mature [44,47]. Several interesting candidate interactions have been identified and Ebbert et al. 2013 (accepted) recently demonstrated that allowing interactions improves the diagnostic utility of the known AD markers. Unfortunately, the complexity of this problem and the extremely large samples sizes required to perform agnostic screens for gene-gene interactions make it very difficult to conduct effective screens for these effects.

AD is a highly complex disease with substantial genetic and environmental components. Our results suggest that genetic variance accounts for ~30% of phenotypic variance, but over 75% of this phenotypic variance remains unexplained by currently identified AD genes. Future AD genetics research must leverage larger samples and novel technologies such as

NGS to identify rare, high penetrant variation and gene-gene interactions that are likely to explain the remaining genetic and phenotypic variance in AD.

Genetic research in AD has followed roughly the same model as the study of other complex diseases; largely focusing on the identification of common variants of modest effect using association studies. Scientist in many disease fields have successfully identified numerous associated variants (this is a small representative sample [48-57]). The transition from a focus on common variants to a focus on the identification of low frequency variants is now underway. These rare, functionally relevant markers are often more easily characterized than common variants of small effect. This will lead to strong and testable hypotheses for the development of therapeutics, thus

accelerating the progress toward effective prevention and treatment.

Supporting Information

Table S1. Missingness rates for covariates and case-control status. The Alzheimer's Disease Genetics Consortium dataset consists of 19,692 total individuals. We removed any individuals missing any of the covariates (listed here) or case-control status (included in this table). (DOCX)

Acknowledgements

We would like to thank Mark T.W. Ebbert for his insightful comments on the manuscript.

Alzheimer's Disease Genetics Consortium

Biological samples and associated phenotypic data used in primary data analysis were stored at the Principal Investigator's institutions, and at the National Cell Repository for Alzheimer's Disease (NCRAD), at the NIA Genetics of Alzheimer's Disease Data Storage Site (NIAGADS) at the University of Pennsylvania, and the NIA Alzheimer's Disease Genetics Consortium Data Storage Site at the University of Pennsylvania.

The members of the Alzheimer's Disease Genetics Consortium are: Marilyn S. Albert¹, Roger L. Albin²⁻⁴, Liana G. Apostolova⁵, Steven E. Arnold⁶, Clinton T. Baldwin⁷, Robert Barber⁸, Michael M. Barmada⁹, Lisa L. Barnes^{10, 11}, Thomas G. Beach¹², Gary W. Beecham^{13, 14}, Duane Beekly¹⁵, David A. Bennett^{10, 16}, Eileen H. Bigio¹⁷, Thomas D. Bird¹⁸, Deborah Blacker^{19,20}, Bradley F. Boeve²¹, James D. Bowen²², Adam Boxer²³, James R. Burke²⁴, Joseph D. Buxbaum^{25, 26, 27}, Nigel J. Cairns²⁸, Laura B. Cantwell²⁹, Chuanhai Cao³⁰, Chris S. Carlson³¹, Regina M. Carney¹³, Minerva M. Carrasquillo³³, Steven L. Carroll³⁴, Helena C. Chui³⁵, David G. Clark³⁶, Jason Comeveaux³⁷, Paul K. Crane³⁸, David H. Cribbs³⁹, Elizabeth A. Crocco⁴⁰, Carlos Cruchaga⁴¹, Philip L. De Jager^{42,43}, Charles DeCarli⁴⁴, Steven T. DeKosky⁴⁵, F. Yesim Demirci⁹, Malcolm Dick⁴⁶, Dennis W. Dickson³³, Ranjan Duara⁴⁷, Nilufer Ertekin-Taner^{33,48}, Denis Evans⁴⁹, Kelley M. Faber⁵⁰, Kenneth B. Fallon³⁴, Martin R. Farlow⁵¹, Lindsay A. Farrer^{7,52,76,77,83}, Steven Ferris⁵³, Tatiana M. Foroud⁵⁰, Matthew P. Frosch⁵⁴, Douglas R. Galasko⁵⁵, Mary Ganguli⁵⁶, Marla Gearing^{57,58}, Daniel H. Geschwind⁵⁹, Bernardino Ghetti⁶⁰, John R. Gilbert^{13,14}, Sid Gilman², Jonathan D. Glass⁶¹, Alison M. Goate⁴¹, Neill R. Graff-Radford^{33,48}, Robert C. Green⁶², John H. Growdon⁶³, Jonathan L. Haines^{64, 65}, Hakon Hakonarson⁶⁶, Kara L. Hamilton-Nelson¹³, Ronald L. Hamilton⁶⁷, John Hardy⁶⁸, Lindy E. Harrell³⁶, Elizabeth Head⁶⁹, Lawrence S. Honig⁷⁰, Matthew J. Huentelman³⁷, Christine M. Hulette⁷¹, Bradley T. Hyman⁶³, Gail P. Jarvik^{72,73}, Gregory A. Jicha⁷⁴, Lee-Way Jin⁷⁵, Gyungah Jun^{7,76,77}, M. Ilyas Kambh^{9,78}, Anna Karydas²³, John S.K. Kauwe⁷⁹, Jeffrey A. Kaye^{80,81}, Ronald Kim⁸², Edward H. Koo⁵⁵, Neil W. Kowall^{83,84}, Joel H. Kramer⁸⁵, Patricia Kramer^{80,86}, Walter A. Kukull⁸⁷, Frank M. LaFerla⁸⁸, James J. Lah⁹¹, Eric B. Larson^{38,89}, James B. Leverenz⁹⁰, Allan I. Levey⁶¹, Ge Li⁹¹, Andrew P. Lieberman⁹², Chiao-Feng Lin²⁹, Oscar L. Lopez⁷⁸, Kathryn L. Lunetta⁷⁶,

Constantine G. Lyketsos⁹³, Wendy J. Mack⁹⁴, Daniel C. Marson³⁶, Eden R. Martin^{13,14}, Frank Martiniuk⁹⁵, Deborah C. Mash⁹⁶, Eliezer Masliah^{55,97}, Richard Mayeux^{70, 109, 110}, Wayne C. McCormick³⁸, Susan M. McCurry⁹⁸, Andrew N. McDavid³¹, Ann C. McKee^{83,84}, Marsel Mesulam⁹⁹, Bruce L. Miller²³, Carol A. Miller¹⁰⁰, Joshua W. Miller⁷⁵, Thomas J. Montine⁹⁰, John C. Morris^{28, 101}, Jill R. Murrell^{50, 60}, Amanda J. Myers⁴⁰, Adam C. Naj¹³, John M. Olichney⁴⁴, Vernon S. Pankratz¹⁰², Joseph E. Parisi^{103,104}, Margaret A. Pericak-Vance^{13, 14}, Elaine Peskind⁹¹, Ronald C. Petersen²¹, Aimee Pierce³⁹, Wayne W. Poon⁴⁶, Huntington Potter³⁰, Joseph F. Quinn⁸⁰, Ashok Raj³⁰, Murray Raskind⁹¹, Eric M. Reiman^{37,105-107}, Barry Reisberg^{53,108}, Christiane Reitz^{70,109,110}, John M. Ringman⁵, Erik D. Roberson³⁶, Ekaterina Rogaeva¹¹¹, Howard J. Rosen²³, Roger N. Rosenberg¹¹², Mary Sano²⁶, Andrew J. Saykin^{50,113}, Gerard D. Schellenberg²⁹, Julie A. Schneider^{10,114}, Lon S. Schneider^{35,115}, William W. Seeley²³, Amanda G. Smith³⁰, Joshua A. Sonnen⁹⁰, Salvatore Spina⁶⁰, Peter St George-Hyslop^{111,116}, Robert A. Stern⁸³, Rudolph E. Tanzi⁶³, John Q. Trojanowski²⁹, Juan C. Troncoso¹¹⁷, Debby W. Tsuang⁹¹, Otto Valladares²⁹, Viviana M. Van Deerlin²⁹, Linda J. Van Eldik¹¹⁸, Badri N. Vardarajan⁷, Harry V. Vinters^{5,119}, Jean Paul Vonsattel¹²⁰, Li-San Wang²⁹, Sandra Weintraub⁹⁹, Kathleen A. Welsh-Bohmer^{24, 121}, Jennifer Williamson⁷⁰, Randall L. Woljter¹²², Clinton B. Wright¹²³, Steven G. Younkin³³, Chang-En Yu³⁸, Lei Yu¹⁰

¹Department of Neurology, Johns Hopkins University, Baltimore, Maryland, ²Department of Neurology, University of Michigan, Ann Arbor, Michigan, ³Geriatric Research, Education and Clinical Center (GRECC), VA Ann Arbor Healthcare System (VAAHS), Ann Arbor, Michigan, ⁴Michigan Alzheimer Disease Center, Ann Arbor, Michigan, ⁵Department of Neurology, University of California Los Angeles, Los Angeles, California, ⁶Department of Psychiatry, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, ⁷Department of Medicine (Genetics Program), Boston University, Boston, Massachusetts, ⁸Department of Pharmacology and Neuroscience, University of North Texas Health Science Center, Fort Worth, Texas, ⁹Department of Human Genetics, University of Pittsburgh, Pittsburgh, Pennsylvania, ¹⁰Department of Neurological Sciences, Rush University Medical Center, Chicago, Illinois, ¹¹Department of Behavioral Sciences, Rush University Medical Center, Chicago, Illinois, ¹²Civin Laboratory for Neuropathology, Banner Sun Health Research Institute, Phoenix, Arizona, ¹³The John P. Hussman Institute for Human Genomics, University of Miami, Miami, Florida, ¹⁴Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami, Miami, Florida, ¹⁵National Alzheimer's Coordinating Center, University of Washington, Seattle, Washington, ¹⁶Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, Illinois, ¹⁷Department of Pathology, Northwestern University, Chicago, Illinois, ¹⁸Department of Neurology, University of Washington, Seattle, Washington, ¹⁹Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, ²⁰Department of Psychiatry, Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts, ²¹Department of Neurology, Mayo Clinic, Rochester, Minnesota, ²²Swedish Medical Center, Seattle,

Washington, ²³Department of Neurology, University of California San Francisco, San Francisco, California, ²⁴Department of Medicine, Duke University, Durham, North Carolina, ²⁵Department of Neuroscience, Mount Sinai School of Medicine, New York, New York, ²⁶Department of Psychiatry, Mount Sinai School of Medicine, New York, New York, ²⁷Departments of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, New York, ²⁸Department of Pathology and Immunology, Washington University, St. Louis, Missouri, ²⁹Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, ³⁰USF Health Byrd Alzheimer's Institute, University of South Florida, Tampa, Florida, ³¹Fred Hutchinson Cancer Research Center, Seattle, Washington, ³²Department of Psychiatry, Vanderbilt University, Nashville, Tennessee, ³³Department of Neuroscience, Mayo Clinic, Jacksonville, Florida, ³⁴Department of Pathology, University of Alabama at Birmingham, Birmingham, Alabama, ³⁵Department of Neurology, University of Southern California, Los Angeles, California, ³⁶Department of Neurology, University of Alabama at Birmingham, Birmingham, Alabama, ³⁷Neurogenomics Division, Translational Genomics Research Institute, Phoenix, Arizona, ³⁸Department of Medicine, University of Washington, Seattle, Washington, ³⁹Department of Neurology, University of California Irvine, Irvine, California, ⁴⁰Department of Psychiatry and Behavioral Sciences, Miller School of Medicine, University of Miami, Miami, Florida, ⁴¹Department of Psychiatry and Hope Center Program on Protein Aggregation and Neurodegeneration, Washington University School of Medicine, St. Louis, Missouri, ⁴²Program in Translational NeuroPsychiatric Genomics, Institute for the Neurosciences, Department of Neurology & Psychiatry, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, ⁴³Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts, ⁴⁴Department of Neurology, University of California Davis, Sacramento, California, ⁴⁵University of Virginia School of Medicine, Charlottesville, Virginia, ⁴⁶Institute for Memory Impairments and Neurological Disorders, University of California Irvine, Irvine, California, ⁴⁷Wien Center for Alzheimer's Disease and Memory Disorders, Mount Sinai Medical Center, Miami Beach, Florida, ⁴⁸Department of Neurology, Mayo Clinic, Jacksonville, Florida, ⁴⁹Rush Institute for Healthy Aging, Department of Internal Medicine, Rush University Medical Center, Chicago, Illinois, ⁵⁰Department of Medical and Molecular Genetics, Indiana University, Indianapolis, Indiana, ⁵¹Department of Neurology, Indiana University, Indianapolis, Indiana, ⁵²Department of Epidemiology, Boston University, Boston, Massachusetts, ⁵³Department of Psychiatry, New York University, New York, New York, ⁵⁴C.S. Kubik Laboratory for Neuropathology, Massachusetts General Hospital, Charlestown, Massachusetts, ⁵⁵Department of Neurosciences, University of California San Diego, La Jolla, California, ⁵⁶Department of Psychiatry, University of Pittsburgh, Pittsburgh, Pennsylvania, ⁵⁷Department of Pathology and Laboratory Medicine, Emory University, Atlanta, Georgia, ⁵⁸Emory Alzheimer's Disease Center, Emory University, Atlanta, Georgia, ⁵⁹Neurogenetics Program, University of California Los Angeles, Los Angeles, California, ⁶⁰Department of Pathology and Laboratory Medicine, Indiana University, Indianapolis, Indiana, ⁶¹Department of Neurology, Emory University, Atlanta, Georgia, ⁶²Division of Genetics, Department of Medicine and Partners Center for Personalized Genetic Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, ⁶³Department of Neurology, Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts, ⁶⁴Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, Tennessee, ⁶⁵Vanderbilt Center for Human Genetics Research, Vanderbilt University, Nashville, Tennessee, ⁶⁶Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, ⁶⁷Department of Pathology (Neuropathology), University of Pittsburgh, Pittsburgh, Pennsylvania, ⁶⁸Institute of Neurology, University College London, Queen Square, London, ⁶⁹Sanders-Brown Center on Aging, Department of Molecular and Biomedical Pharmacology, University of Kentucky, Lexington, Kentucky, ⁷⁰Taub Institute on Alzheimer's Disease and the Aging Brain, Department of Neurology, Columbia University, New York, New York, ⁷¹Department of Pathology, Duke University, Durham, North Carolina, ⁷²Department of Genome Sciences, University of Washington, Seattle, Washington, ⁷³Department of Medicine (Medical Genetics), University of Washington, Seattle, Washington, ⁷⁴Sanders-Brown Center on Aging, Department of Neurology, University of Kentucky, Lexington, Kentucky, ⁷⁵Department of Pathology and Laboratory Medicine, University of California Davis, Sacramento, California, ⁷⁶Department of Biostatistics, Boston University, Boston, Massachusetts, ⁷⁷Department of Ophthalmology, Boston University, Boston, Massachusetts, ⁷⁸University of Pittsburgh Alzheimer's Disease Research Center, Pittsburgh, Pennsylvania, ⁷⁹Department of Biology, Brigham Young University, Provo, Utah, ⁸⁰Department of Neurology, Oregon Health & Science University, Portland, Oregon, ⁸¹Department of Neurology, Portland Veterans Affairs Medical Center, Portland, Oregon, ⁸²Department of Pathology and Laboratory Medicine, University of California Irvine, Irvine, California, ⁸³Department of Neurology, Boston University, Boston, Massachusetts, ⁸⁴Department of Pathology, Boston University, Boston, Massachusetts, ⁸⁵Department of Neuropsychology, University of California San Francisco, San Francisco, California, ⁸⁶Department of Molecular & Medical Genetics, Oregon Health & Science University, Portland, Oregon, ⁸⁷Department of Epidemiology, University of Washington, Seattle, Washington, ⁸⁸Department of Neurobiology and Behavior, University of California Irvine, Irvine, California, ⁸⁹Group Health Research Institute, Group Health, Seattle, Washington, ⁹⁰Department of Pathology, University of Washington, Seattle, Washington, ⁹¹Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, Washington, ⁹²Department of Pathology, University of Michigan, Ann Arbor, Michigan, ⁹³Department of Psychiatry, Johns Hopkins University, Baltimore, Maryland, ⁹⁴Department of Preventive Medicine, University of Southern California, Los Angeles, California, ⁹⁵Department of Medicine - Pulmonary, New York University, New York, New York, ⁹⁶Department of Neurology, University of Miami, Miami, Florida,

⁹⁷Department of Pathology, University of California San Diego, La Jolla, California, ⁹⁸School of Nursing Northwest Research Group on Aging, University of Washington, Seattle, Washington, ⁹⁹Cognitive Neurology and Alzheimer's Disease Center, Northwestern University, Chicago, Illinois, ¹⁰⁰Department of Pathology, University of Southern California, Los Angeles, California, ¹⁰¹Department of Neurology, Washington University, St. Louis, Missouri, ¹⁰²Department of Biostatistics, Mayo Clinic, Rochester, Minnesota, ¹⁰³Department of Anatomic Pathology, Mayo Clinic, Rochester, Minnesota, ¹⁰⁴Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota, ¹⁰⁵Arizona Alzheimer's Consortium, Phoenix, Arizona, ¹⁰⁶Department of Psychiatry, University of Arizona, Phoenix, Arizona, ¹⁰⁷Banner Alzheimer's Institute, Phoenix, Arizona, ¹⁰⁸Alzheimer's Disease Center, New York University, New York, New York, ¹⁰⁹Gertrude H. Sergievsky Center, Columbia University, New York, New York, ¹¹⁰Department of Neurology, Columbia University, New York, New York, ¹¹¹Tanz Centre for Research in Neurodegenerative Disease, University of Toronto, Toronto, Ontario, ¹¹²Department of Neurology, University of Texas Southwestern, Dallas, Texas, ¹¹³Department of Radiology and Imaging Sciences, Indiana University, Indianapolis, Indiana, ¹¹⁴Department of Pathology (Neuropathology), Rush University Medical Center, Chicago,

Illinois, ¹¹⁵Department of Psychiatry, University of Southern California, Los Angeles, California, ¹¹⁶Cambridge Institute for Medical Research and Department of Clinical Neurosciences, University of Cambridge, Cambridge, ¹¹⁷Department of Pathology, Johns Hopkins University, Baltimore, Maryland, ¹¹⁸Sanders-Brown Center on Aging, Department of Anatomy and Neurobiology, University of Kentucky, Lexington, Kentucky, ¹¹⁹Department of Pathology & Laboratory Medicine, University of California Los Angeles, Los Angeles, California, ¹²⁰Taub Institute on Alzheimer's Disease and the Aging Brain, Department of Pathology, Columbia University, New York, New York, ¹²¹Department of Psychiatry & Behavioral Sciences, Duke University, Durham, North Carolina, ¹²²Department of Pathology, Oregon Health & Science University, Portland, Oregon, ¹²³Evelyn F. McKnight Brain Institute, Department of Neurology, Miller School of Medicine, University of Miami, Miami, Florida

Author Contributions

Conceived and designed the experiments: PGR JSKK. Performed the experiments: PGR. Analyzed the data: PGR SM PC JSKK. Contributed reagents/materials/analysis tools: SM ADGC. Wrote the manuscript: PGR JSKK.

References

1. Querfurth HW, LaFerla FM (2010) Alzheimer's disease. *N Engl J Med* 362: 329-344. doi:10.1056/NEJMra0909142. PubMed: 20107219.
2. Brookmeyer R, Johnson E, Ziegler-Graham K, Arrighi HM (2007) Forecasting the global burden of Alzheimer's disease. *Alzheimers Dement* 3: 186-191. doi:10.1016/j.jalz.2007.04.381. PubMed: 19595937.
3. Ballard C, Gauthier S, Corbett A, Brayne C, Aarsland D et al. (2011) Alzheimer's disease. *Lancet*.
4. Herrup K (2010) Reimagining Alzheimer's disease—an age-based hypothesis. *J Neurosci* 30: 16755-16762. doi:10.1523/JNEUROSCI.4521-10.2010. PubMed: 21159946.
5. Patterson C, Feightner JW, Garcia A, Hsiung GY, MacKnight C et al. (2008) Diagnosis and treatment of dementia: 1. Risk assessment and primary prevention of Alzheimer disease. *CMAJ* 178: 548-556. doi: 10.1503/cmaj.070796. PubMed: 18299540.
6. Cataldo JK, Prochaska JJ, Glantz SA (2010) Cigarette smoking is a risk factor for Alzheimer's Disease: an analysis controlling for tobacco industry affiliation. *J Alzheimers Dis* 19: 465-480. PubMed: 20110594.
7. Almeida OP, Hulse GK, Lawrence D, Flicker L (2002) Smoking as a risk factor for Alzheimer's disease: contrasting evidence from a systematic review of case-control and cohort studies. *Addiction* 97: 15-28. doi:10.1046/j.1360-0443.2002.00016.x. PubMed: 11895267.
8. Lindsay J, Laurin D, Verreault R, Hébert R, Helliwell B et al. (2002) Risk factors for Alzheimer's disease: a prospective analysis from the Canadian Study of Health and Aging. *Am J Epidemiol* 156: 445-453. doi:10.1093/aje/kwf074. PubMed: 12196314.
9. Podewils LJ, Guallar E, Kuller LH, Fried LP, Lopez OL et al. (2005) Physical activity, APOE genotype, and dementia risk: findings from the Cardiovascular Health Cognition Study. *Am J Epidemiol* 161: 639-651. doi:10.1093/aje/kwi092. PubMed: 15781953.
10. Wang HX, Karp A, Winblad B, Fratiglioni L (2002) Late-life engagement in social and leisure activities is associated with a decreased risk of dementia: a longitudinal study from the Kungsholmen project. *Am J Epidemiol* 155: 1081-1087. doi:10.1093/aje/k155.12.1081. PubMed: 12048221.
11. Scarmeas N, Stern Y, Mayeux R, Luchsinger JA (2006) Mediterranean diet, Alzheimer disease, and vascular mediation. *Arch Neurol* 63: 1709-1717. doi:10.1001/archneur.63.12.noc60109. PubMed: 17030648.
12. Patterson C, Feightner J, Garcia A, MacKnight C (2007) General risk factors for dementia: a systematic evidence review. *Alzheimers Dement* 3: 341-347. doi:10.1016/j.jalz.2007.07.001. PubMed: 19595956.
13. Goate A, Chartier-Harlin MC, Mullan M, Brown J, Crawford F et al. (1991) Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature* 349: 704-706. doi:10.1038/349704a0. PubMed: 1671712.
14. Sherrington R, Rogaeve EI, Liang Y, Rogaeve EA, Levesque G et al. (1995) Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature* 375: 754-760. doi: 10.1038/375754a0. PubMed: 7596406.
15. Levy-Lahad E, Wasco W, Poorkaj P, Romano DM, Oshima J et al. (1995) Candidate gene for the chromosome 1 familial Alzheimer's disease locus. *Science* 269: 973-977. doi:10.1126/science.7638622. PubMed: 7638622.
16. Guerreiro R, Wojtas A, Bras J, Carrasquillo M, Rogaeve E et al. (2012) TREM2 Variants in Alzheimer's Disease. *N Engl J Med*, 368: 117-27. PubMed: 23150934.
17. Jonsson T, Atwal JK, Steinberg S, Snaedal J, Jonsson PV et al. (2012) A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. *Nature* 488: 96-99. doi:10.1038/nature11283. PubMed: 22801501.
18. Jonsson T, Stefansson H, Ph DS, Jonsdottir I, Jonsson PV et al. (2012) Variant of TREM2 Associated with the Risk of Alzheimer's Disease. *N Engl J Med*, 368: 107-16. PubMed: 23150908. PubMed: 23150908.
19. Lee SH, Harold D, Nyholt DR, Goddard ME, Zondervan KT et al. (2013) Estimation and partitioning of polygenic variation captured by common SNPs for Alzheimer's disease, multiple sclerosis and endometriosis. *Hum Mol Genet* 22: 832-841. doi:10.1093/hmg/dd5491. PubMed: 23193196.
20. 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. doi:10.1038/ng.801. PubMed: 21460841.
21. International HapMap Consortium (2003) The International HapMap Project. *Nature* 426: 789-796. doi:10.1038/nature02168. PubMed: 14685227.
22. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR (2010) MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol* 34: 816-834. doi:10.1002/gepi.20533. PubMed: 21058334.
23. 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. doi:10.1086/519795. PubMed: 17701901.

24. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M et al. (2010) Robust relationship inference in genome-wide association studies. *Bioinformatics* 26: 2867-2873. doi:10.1093/bioinformatics/btq559. PubMed: 20926424.
25. Wilming LG, Gilbert JG, Howe K, Trevanion S, Hubbard T et al. (2008) The vertebrate genome annotation (Vega) database. *Nucleic Acids Res* 36: D753-D760. PubMed: 18003653.
26. Yang J, Lee SH, Goddard ME, Visscher PM (2011) GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 88: 76-82. doi: 10.1016/j.ajhg.2010.11.011. PubMed: 21167468.
27. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK et al. (2010) Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* 42: 565-569. doi:10.1038/ng.608. PubMed: 20562875.
28. Lee SH, Wray NR, Goddard ME, Visscher PM (2011) Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet* 88: 294-305. doi:10.1016/j.ajhg.2011.02.002. PubMed: 21376301.
29. Yang J, Manolio TA, Pasquale LR, Boerwinkle E, Caporaso N et al. (2011) Genome partitioning of genetic variation for complex traits using common SNPs. *Nat Genet* 43: 519-525. doi:10.1038/ng.823. PubMed: 21552263.
30. Association; As (2012) Alzheimer's Association Annual Report: Alzheimer's disease Facts and Figures.
31. 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. doi:10.1038/ng.803. PubMed: 21460840.
32. Biffi A, Anderson CD, Desikan RS, Sabuncu M, Cortellini L et al. (2010) Genetic variation and neuroimaging measures in Alzheimer disease. *Arch Neurol* 67: 677-685. doi:10.1001/archneurol.2010.108. PubMed: 20558387.
33. 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* 41: 1088-1093. doi: 10.1038/ng.440. PubMed: 19734902.
34. Hu X, Pickering E, Liu YC, Hall S, Fournier H et al. (2011) Meta-analysis for genome-wide association study identifies multiple variants at the BIN1 locus associated with late-onset Alzheimer's disease. *PLOS ONE* 6: e16616. doi:10.1371/journal.pone.0016616. PubMed: 21390209.
35. Carrasquillo MM, Belbin O, Hunter TA, Ma L, Bisceglia GD et al. (2010) Replication of CLU, CR1, and PICALM associations with alzheimer disease. *Arch Neurol* 67: 961-964. doi:10.1001/archneurol.2010.147. PubMed: 20554627.
36. Comevaux JJ, Myers AJ, Allen AN, Pruzin JJ, Ramirez M et al. (2010) Association of CR1, CLU and PICALM with Alzheimer's disease in a cohort of clinically characterized and neuropathologically verified individuals. *Hum Mol Genet* 19: 3295-3301. doi:10.1093/hmg/ddq221. PubMed: 20534741.
37. Wray NR (2005) Allele frequencies and the r^2 measure of linkage disequilibrium: impact on design and interpretation of association studies. *Twin Res Hum Genet Off J International Society For Twin Studies* 8: 87-94. doi:10.1375/twin.8.2.87. PubMed: 15901470.
38. Singleton A, Hardy J (2011) A generalizable hypothesis for the genetic architecture of disease: pleomorphic risk loci. *Hum Mol Genet* 20: R158-R162. doi:10.1093/hmg/ddr358. PubMed: 21875901.
39. Rope AF, Wang K, Evjenth R, Xing J, Johnston JJ et al. (2011) Using VAAST to identify an X-linked disorder resulting in lethality in male infants due to N-terminal acetyltransferase deficiency. *Am J Hum Genet* 89: 28-43. doi:10.1016/j.ajhg.2011.05.017. PubMed: 21700266.
40. Choi M, Scholl UI, Ji W, Liu T, Tikhonova IR et al. (2009) Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci U S A* 106: 19096-19101. doi:10.1073/pnas.0910672106. PubMed: 19861545.
41. Roach JC, Glusman G, Smit AF, Huff CD, Hubley R et al. (2010) Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science* 328: 636-639. doi:10.1126/science.1186802. PubMed: 20220176.
42. Krebs CE, Paisan-Ruiz C (2012) The use of next-generation sequencing in movement disorders. *Front Genet* 3: 75. PubMed: 22593763.
43. Kilpinen H, Barrett JC (2013) How next-generation sequencing is transforming complex disease genetics. *Trends Genet* 29: 23-30. doi:10.1016/j.tig.2012.10.001. PubMed: 23103023.
44. Turton JC, Bullock J, Medway C, Shi H, Brown K et al. (2011) Investigating statistical epistasis in complex disorders. *J Alzheimers Dis* 25: 635-644. PubMed: 21483092.
45. Mackay TF (2004) The genetic architecture of quantitative traits: lessons from *Drosophila*. *Curr Opin Genet Dev* 14: 253-257. doi: 10.1016/j.gde.2004.04.003. PubMed: 15172667.
46. Shao H, Burrage LC, Sinasac DS, Hill AE, Ernest SR et al. (2008) Genetic architecture of complex traits: large phenotypic effects and pervasive epistasis. *Proc Natl Acad Sci U S A* 105: 19910-19914. doi: 10.1073/pnas.0810388105. PubMed: 19066216.
47. Combarros O, Cortina-Borja M, Smith AD, Lehmann DJ (2009) Epistasis in sporadic Alzheimer's disease. *Neurobiol Aging* 30: 1333-1349. doi:10.1016/j.neurobiolaging.2007.11.027. PubMed: 18206267.
48. Ferreira MA, Matheson MC, Duffy DL, Marks GB, Hui J et al. (2011) Identification of IL6R and chromosome 11q13.5 as risk loci for asthma. *Lancet* 378: 1006-1014. doi:10.1016/S0140-6736(11)60874-X. PubMed: 21907864.
49. Wain LV, Verwoert GC, O'Reilly PF, Shi G, Johnson T et al. (2011) Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. *Nat Genet* 43: 1005-1011. doi: 10.1038/ng.922. PubMed: 21909110.
50. Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP et al. (2012) Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 491: 119-124. doi:10.1038/nature11582. PubMed: 23128233.
51. Lill CM, Roehr JT, McQueen MB, Kavvoura FK, Bagade S et al. (2012) Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: The PDGene database. *PLOS Genet* 8: e1002548. PubMed: 22438815.
52. Wen W, Cho YS, Zheng W, Dorajoo R, Kato N et al. (2012) Meta-analysis identifies common variants associated with body mass index in east Asians. *Nat Genet* 44: 307-311. doi:10.1038/ng.1087. PubMed: 22344219.
53. Haiman CA, Chen GK, Vachon CM, Canzian F, Dunning A et al. (2011) A common variant at the TERT-CLPTM1L locus is associated with estrogen receptor-negative breast cancer. *Nat Genet* 43: 1210-1214. doi:10.1038/ng.985. PubMed: 22037553.
54. Psychiatric GWAS Consortium Bipolar Disorder Working Group (2011) Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet* 43: 977-983. doi:10.1038/ng.943. PubMed: 21926972.
55. Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD et al. (2011) Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* 478: 103-109. doi:10.1038/nature10405. PubMed: 21909115.
56. Bradfield JP, Taal HR, Timpson NJ, Scherag A, Lecoer C et al. (2012) A genome-wide association meta-analysis identifies new childhood obesity loci. *Nat Genet* 44: 526-531. doi:10.1038/ng.2247. PubMed: 22484627.
57. Jenuth JP (2000) The NCBI. Publicly available tools and resources on the Web. *Methods Mol Biol* 132: 301-312. PubMed: 10547843.
58. Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC et al. (1993) Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* 261: 921-923. doi: 10.1126/science.8346443. PubMed: 8346443.
59. Corder EH, Saunders AM, Risch NJ, Strittmatter WJ, Schmechel DE et al. (1994) Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. *Nat Genet* 7: 180-184. doi:10.1038/ng0694-180. PubMed: 7920638.
60. 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. doi: 10.1038/ng.439. PubMed: 19734903.