Gene-Based Analysis in HRC Imputed Genome Wide Association Data Identifies Three Novel Genes for Alzheimer's Disease

Emily Baker[1,2], Rebecca Sims[1], Ganna Leonenko[1], Aura Frizzati[1], Janet C. Harwood[1], Detelina Grozeva[1], GERAD/PERADES[+], CHARGE[+], ADGC[+], EADI[+], IGAP consortia[+], Kevin Morgan[3], Peter Passmore[4], Clive Holmes[5], John Powell[6,7], Carol Brayne[8], Michael Gill[9,10], Simon Mead[11], Paola Bossù[12], Gianfranco Spalletta[13], Alison M. Goate[14,15], Carlos Cruchaga[16,15], Wolfgang Maier[17], Reinhard Heun[18], Frank Jessen[17,19], Oliver Peters[20,21], Martin Dichgans[22,23,24], Lutz Frölich[25], Alfredo Ramirez[19,26], Lesley Jones[1], John Hardy[27], Dobril Ivanov[2], Matthew Hill[2], Peter Holmans[1], Nicholas D. Allen[2], B. Paul Morgan[2], Sudha Seshadri[28], Gerard D. Schellenberg[29], Philippe Amouyel[30], Julie Williams[1,2*], Valentina Escott-Price[1,2*]

1 Medical Research Council Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, UK
2 UK Dementia Research Institute at Cardiff University, Cardiff, UK,
3 Human Genetics, School of Life Sciences, Life Sciences Building A27, University Park, University of Nottingham, Nottingham, NG7 2RD, UK
4 Centre for Public Health, School of Medicine, Dentistry and Biomedical Sciences, Queens University, Belfast, UK
5 Division of Clinical Neurosciences, School of Medicine, University of Southampton, Southampton, UK
6 Department of Basic and Clinical Neuroscience, Institute of Psychiatry, Psychology and Neuroscience, Kings College London, London, UK
7 Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Herston, Queensland, Australia
8 Institute of Public Health, University of Cambridge, Cambridge, UK
9 Mercer's Institute for Research on Ageing, St. James' Hospital, Dublin, Ireland
10 James Hospital and Trinity College, Dublin, Ireland
11 MRC Prion Unit at UCL, Institute of Prion Diseases, London, UK
12 Experimental Neuropsychobiology Laboratory, IRCCS Santa Lucia Foundation, Department of Clinical and Behavioral Neurology, Rome, Italy
13 Laboratory of Neuropsychiatry, IRCCS Santa Lucia Foundation, Rome, Italy
14 Icahn School of Medicine at Mount Sinai, New York, New York, USA
15 Hope Center Program on Protein Aggregation and Neurodegeneration, Washington University School of Medicine, St Louis, Missouri, USA
16 Department of Psychiatry, Washington University School of Medicine, St Louis, Missouri, USA
17 German Centre for Neurodegenerative Diseases (DZNE), 53127 Bonn, Germany
18 Department of Psychiatry and Psychotherapy, University of Bonn, 53127, Bonn, Germany
19 Department of Psychiatry and Psychotherapy, University of Cologne, 50937 Cologne, Germany
20 Department of Psychiatry and Psychotherapy, Charité Berlin, Berlin, Germany
21 German Center for Neurodegenerative Diseases (DZNE), Berlin, Germany.
22 Institute for Stroke and Dementia Research, Klinikum der Universität München, Munich, Germany

23 German Center for Neurodegenerative Diseases (DZNE, Munich), Munich, 80336, Germany
24 Munich Cluster for Systems Neurology (SyNergy), Munich, Germany
25 Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Heidelberg, Germany.
26 Department for Neurodegenerative Diseases and Geriatric Psychiatry, University Hospital Bonn, Bonn, Germany
27 Department of Molecular Neuroscience, UCL, Institute of Neurology, London, UK
28 Department of Neurology, Boston University School of Medicine, Boston, Massachusetts, USA
29 Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA
30 Univ. Lille, Inserm, CHU Lille University Hospital, Institut Pasteur de Lille, LabEx DISTALZ-UMR1167 - RID-AGE - Risk factors and molecular determinants of aging-related, F-59000 Lille, France.

* Corresponding authors:
E-mail: EscottPriceV@cardiff.ac.uk (VEP)
E-mail: WilliamsJ@cardiff.ac.uk (JW)

+ For more information on these consortia and their full membership lists, please see the Acknowledgements and S1 Appendix.

Abstract

Late onset Alzheimer's disease is the most common form of dementia for which about 30 susceptibility loci have been reported. The aim of the current study is to identify novel genes associated with Alzheimer's disease using the largest up-to-date reference single nucleotide polymorphism (SNP) panel, the most accurate imputation software and a novel gene-based analysis approach which tests for patterns of association within genes, in the powerful genome-wide association dataset of the International Genomics of Alzheimer's Project Consortium, comprising over 7 million genotypes from 17,008 Alzheimer's cases and 37,154 controls. In addition to earlier reported genes, we detected three novel gene-wide significant loci $PPARGC1A$ ($p = 2.2 \times 10^{-6}$), $RORA$ ($p = 7.4 \times 10^{-7}$) and $ZNF423$ ($p = 2.1 \times 10^{-6}$). $PPARGC1A$ and $RORA$ are involved in circadian rhythm; circadian disturbances are one of the earliest symptoms of Alzheimer's disease. $PPARGC1A$ is additionally linked to energy metabolism and the generation of amyloid beta plaques. $RORA$ is involved in a variety of functions apart from circadian rhythm, such as cholesterol metabolism and inflammation. The $ZNF423$ gene resides in an Alzheimer's disease-specific protein network and is likely involved with centrosomes and DNA damage repair.

Introduction

Late Onset Alzheimer's disease (LOAD) is a devastating neurodegenerative condition with significant genetic heritability [1]. The apolipoprotein E ($APOE$) gene is the strongest genetic risk factor for LOAD [2]. Subsequently, more genes were found to be associated with AD development. The Genetic and Environmental Risk in Alzheimer's Disease (GERAD) Consortium published a Genome-Wide Association Study (GWAS) that identified novel variants in CLU and $PICALM$ which were associated with AD [3]. Concurrently, the European Alzheimer's Disease Initiative (EADI) identified an association between the $CR1$ and CLU loci and AD [4]. Subsequent publications by

GERAD, the Alzheimer's Disease Genetic Consortium (ADGC) and Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium identified a further 5 novel loci [5] [6] [7]. The International Genomics of Alzheimer's Project (IGAP) [4] Consortium is an amalgamation of these four different genetic groups (GERAD, EADI, ADGC and CHARGE). Meta-analysis of the 4 GWAS datasets determined 11 novel variants associated with AD. A gene-based analysis has been undertaken in the IGAP AD data using Brown's method [8]. This approach determined two additional novel genes; *TP53INP1* and *IGHV1-67* [9]. Additionally, low frequency risk variants have been identified through next generation sequencing (*TREM2*) [10] and a whole-exome association study (*PLCG2*, *TREM2* and *ABI3* [11]).

Gene-based analysis is an alternative to GWAS analyses, which considers the association of an individual single nucleotide polymorphism (SNP) with disease. Gene-based analyses provide more power due to the aggregate effect of multiple SNPs being larger than that of individual SNPs. For example, determining the association of genes rather than SNPs, is beneficial since genes are more robust across different populations, this is due to the linkage disequilibrium (LD) between SNPs resulting in different SNPs being associated in different populations [12]. Gene-based analyses are being widely used in the field and as expected, are able to identify novel genes or pathways associated with disease. Pathways clustering in eight areas of biology have been found to be associated with AD using the ALIGATOR [13] algorithm [14] [15].

The aim of the current study is to identify novel genes associated with AD using the largest up-to-date reference SNP panel, the most accurate imputation software and a novel gene-based analysis approach. In this study, we used the GERAD data [3] which have been imputed using the latest Haplotype Reference Consortium data (HRC). Polygenic Linkage disequilibrium- Adjusted Risk Score (POLARIS) [16] is a powerful gene-based method which produces a risk score per person per gene, adjusts for LD between SNPs and informs the analysis with summary statistics from an external data set. POLARIS, unlike standard Polygenic Risk Score (PRS) does not require data to be pruned for LD prior to analysis, so it is able to incorporate information from a larger number of SNPs. We employed the POLARIS approach [16] and using the individual genotypes in the GERAD imputed data, produced the risk score for each individual for every gene considered. The IGAP [4] SNP summary statistic data, where the individuals from GERAD were excluded, were used to generate the gene-based PRS.

Results

For the imputed GERAD data, using a window around the gene of 35kb upstream and 10kb downstream [17], SNPs are assigned to 18,087 genes which are plotted on a Manhattan plot in Fig 1. The 12 gene-wide significant genes from this analysis are shown in Table 1, where gene-wide significance is defined as $p < 2.5 \times 10^{-6}$ [18]. A large number of genes reside on chromosome 19 and these are likely influenced by the large effect of *APOE*. Three novel genes have been identified from this analysis: *PPARGC1A*, *RORA* and *ZNF423*. *PPARGC1A* (peroxisome proliferator-activated receptor gamma co-activator 1alpha) is a master regulator that mainly regulates energy metabolism [19] [20]. It has been linked to the generation of amyloid beta plaques [21] and circadian rhythm [22]. *RORA* (Retinoic acid receptor-related orphan receptor alpha) is involved in a variety of functions such as circadian rhythm, cholesterol metabolism and inflammation [23]. Its expression is also upregulated in the AD hippocampus [24]. The *ZNF423* gene resides in an AD-specific protein network which also includes other AD-related genes such as *APOE*, *CLU*, *ABCA7*, *TREM2* etc. [25]. *ZNF423* is likely involved with centrosomes and DNA damage repair [26]. The *SCARA3* gene overlaps *CLU* which has previously been identified as being associated

with AD [3] [4]. The POLARIS gene-based results for the genes previously identified as being associated with AD are seen in S1 Table, these genes contain genome-wide significant SNPs ($p < 5 \times 10^{-8}$). Table 1 additionally shows the POLARIS gene-based results conditioned on the *APOE* gene, this is done by including the POLARIS *APOE* gene risk score into the logistic regression model. *ZNF423* conditioned on *APOE* no longer reaches gene-wide significance, but *PPARGC1A* and *RORA* remain significant, suggesting an association independent of *APOE*. In addition, *BCAM*, *PVRL2* and *APOC4-APOC2* on chromosome 19 remain gene-wide significant, even after adjusting for *APOE* in the model, suggesting a potential signal beyond *APOE*. To investigate this, we additionally conditioned on *BCAM* (the most significant chromosome 19 gene after adjusting for *APOE*) to determine whether this explains the remaining effect. Results are shown in S2 Table; when conditioning on *APOE* and *BCAM* the only remaining gene-wide significant gene on chromosome 19 is *APOC4-APOC2*, suggesting that the majority of signals on chromosome 19 are explained by *APOE* and *BCAM*. We were unable to condition on *APOE* genotype since these are not available for all subjects, so removal of an association may be due to reduced sample size.

Fig 1. Manhattan Plot for the POLARIS Gene-Based Analysis in Imputed GERAD Data Using a Gene Window 35kb Upstream and 10kb Downstream.

Table 1. Gene-Wide Significant Genes from POLARIS Gene-based Analysis in GERAD Imputed Data Using a Gene Window (35kb Upstream and 10kb Downstream)

Chr	Gene	No. of SNPs	POLARIS			POLARIS, conditioned on *APOE*		
			Beta	SE	P-value	Beta	SE	P-value
4	*PPARGC1A*	480	0.877	0.1851	2.2×10^{-6}	0.920	0.1885	1.0×10^{-6}
8	*SCARA3* (*CLU*)	240	0.526	0.1064	7.8×10^{-7}	0.537	0.1090	8.3×10^{-7}
15	*RORA*	1813	0.334	0.0674	7.4×10^{-7}	0.338	0.0688	9.1×10^{-7}
16	*ZNF423*	1056	0.551	0.1163	2.1×10^{-6}	0.541	0.1187	5.1×10^{-6}
19	*BCL3*	88	0.377	0.0674	4.2×10^{-9}	0.291	0.0656	8.8×10^{-6}
19	*CBLC*	50	0.605	0.1161	1.8×10^{-7}	0.455	0.1183	0.00012
19	*BCAM*	71	0.556	0.0543	1.4×10^{-24}	0.492	0.0555	7.6×10^{-19}
19	*PVRL2*	160	0.546	0.0299	9.4×10^{-75}	0.430	0.0491	2.0×10^{-18}
19	*TOMM40*	108	0.500	0.0298	3.4×10^{-63}	0.334	0.0891	0.00018
19	*APOE*	55	0.520	0.0315	4.4×10^{-61}	NA	NA	NA
19	*APOC1*	34	0.475	0.0315	1.5×10^{-51}	-0.249	0.1031	0.01575
19	*APOC4-APOC2*	62	0.615	0.0871	1.6×10^{-12}	0.419	0.0892	2.5×10^{-6}

In order to narrow down the disease-associated SNPs for each of these novel genes, we investigated the gene expression patterns using the BRAINEAC [27] database from the UK Brain Expression Consortium. For the *PPARGC1A* gene the SNP rs67436520, which is downstream of the *PPARGC1A* gene, has the best cis-expression quantitative trait loci (eQTL) p-value of 3.3×10^{-4}, this is expressed in the hippocampus. The best cis-eQTL p-value in the *RORA* gene is 1.5×10^{-4}, this is for SNP rs113223478 which is 78.5kb upstream of the gene, between the *NARG2* and *ANXA2* genes, and is expressed in the substantia nigra. This SNP will not be included in the POLARIS score, however, it could be tagged by SNPs included in the score. Finally, SNP rs2270396 has the best cis-eQTL p-value in the *ZNF423* gene with a p-value of 3.0×10^{-5} and is expressed in the frontal cortex. These SNPs were checked in RegulomeDB [28] and Variant Effect Predictor [29]. They are all intergenic variants that are not in any well-defined regulatory region of the genome and do not overlap the best risk SNPs, so it is difficult

to predict how these SNPs may affect the regulation of the expression of these genes. 89

Discussion 90

A gene-based analysis was performed using the individual genotypes in the GERAD 91
imputed data and the summary statistics from IGAP data excluding the GERAD 92
subjects was used to inform the analysis. This analysis expands a gene window around 93
the gene, 35kb upstream and 10kb downstream, which is likely to include 94
transcriptional regulatory elements in the gene [17] and thus contain SNPs influencing 95
gene expression. Three novel genes were found to be associated with AD using the 96
POLARIS method. The novel genes are *PPARGC1A*, *RORA* and *ZNF423*, all of which 97
have credible biological relevance to AD. These results are already adjusted for LD 98
between SNPs in the gene, using the POLARIS methodology. Most of the genes 99
identified before in IGAP data [4] [9] were also identified by POLARIS as statistically 100
significant, however, since previous results are based on IGAP stages 1 and 2, POLARIS 101
p-values were slightly larger. 102

We investigated disease-associated SNPs using expression patterns, which 103
highlighted individual SNPs. A limitation of this analysis is that the POLARIS score 104
tests the aggregated risk across the gene and is unlikely due to a single SNP. 105

The product of the *PPARGC1A* gene, PGC-1α (Peroxisome proliferator-activated 106
receptor gamma coactivator 1-alpha) is part of the PGC-1 family of transcriptional 107
coactivators that mainly regulate mitochondrial biogenesis to in turn regulate the 108
cellular energy metabolism [19]. It is also involved in other cellular and physiological 109
functions, including the response to a variety of cellular and external stimuli, cellular 110
glucose homeostasis, circadian rhythm, and the regulation of neuronal apoptosis. 111

The regulation of this gene is complex; it has multiple isoforms and alternative 112
promoters [30] and gene expression is regulated by a variety of stimuli, including 113
cytokines, insulin, exercise and the cold [31]. PGC-1α can induce ribosomal 114
transcription under stress conditions such as oxidative stress and exercise [32]. 115

Previous animal model work has shown that overexpression of hPGC-1α in APP23 116
mice improved spatial and recognition memory, along with a significant reduction of Aβ 117
deposition [21]. Furthermore, hPGC-1α overexpression also reduced the levels of 118
proinflammatory cytokines and microglial activation [21] [33]. This suggests a direct 119
link with recent genetic evidence of microglia-mediated innate immune response 120
involvement in AD [11]. In addition, activation of PGC-1α by EKR and p38 inhibitors 121
have been shown to improve spatial and learning memory in Aβ-injected rats [34]. 122
PPARGC1A has also been implicated in the pathogenesis of other neurodegenerative 123
disorders, namely Huntington's and Parkinson's diseases [35]. It has been shown that 124
mutated Huntingtin represses PGC-1α, affecting mitochondrial function, hence 125
ribosomal biogenesis may be affected in Huntington's disease [36]. There is a brain 126
specific promoter 587kb upstream of human *PPARGC1A* [37], which is located in a 127
genomic region associated with age of onset of Huntington's disease and relevant here is 128
that hippocampal PGC-1α expression is decreased in the AD brain [38]. A randomised 129
controlled trial of a PPAR-γ agonist, pioglitazone, found improved cognition and 130
regional cerebral blood flow in patients with mild AD [39]. 131

RORA (Retinoic acid receptor-related orphan receptor alpha) is a nuclear hormone 132
receptor with diverse cellular roles [40], for example in immunity, cerebellum 133
development [41], lipid metabolism [42], circadian rhythms and inflammation [23]. 134
RORA regulates its target genes by binding to the ROR response elements (RORE) in 135
the gene regulatory region [43]. It has been shown to regulate more than 3,000 genes in 136
human monocytic and endothelial cell lines [44]. It has a role in the regulation of the 137
BDNF pathway and its expression is upregulated in AD hippocampus [24]. *RORA* and 138

PPARGC1A are close biological partners, with PGC-1α regulating the expression of a number of clock genes through the coactivation of the ROR family of orphan nuclear receptors [45]. *RORA* has been shown to be linked to other genes previously implicated in AD [25] and also has been implicated in a large number of neuropsychiatric disorders, such as post-traumatic stress disorder [46] [47] and autism [48]. Furthermore, *RORA* trans-activates IL-6 and is thought to be neuro-protective in astrocytes and anti-inflammatory in peripheral tissues [49]. The two genes, *RORA* and *PPARGC1A* that we report here provide further evidence of the involvement of inflammation in the pathogenesis of AD.

Finally, *ZNF423* is a nuclear protein that belongs to the Kruppel-like C2H2 zinc finger proteins. *ZNF423* directs bone morphogenetic protein (BMP)-dependent signalling activity and aberrant forms impede B cell differentiation [50]. Furthermore, elevated gene-expression of *ZNF423* has been shown to occur in patients with systemic lupus erythematosus, pointing to an impaired function of B cells in human mesenchymal stem cells [51]. *ZNF423* resides in an AD-specific protein network [25]. *ZNF423* is likely involved with centrosomes and DNA damage repair [26]. It is downregulated in human neuroblastoma and glioma [52] [53] and also has a role in breast cancer [54]. Previously, it also has been shown that missense and LoF variants are likely to be pathogenic for abnormality of brain morphology, Joubert syndrome and Nephronophthisis with autosomal dominant or autosomal recessive inheritance (`www.omim.org`, `https://www.ncbi.nlm.nih.gov/clinvar/`). These disorders present with a range of phenotypic characteristics, with the central nervous system being affected too (more specifically the cerebellar vermis). In nur12 mouse model (with introduced nonsense mutation in exon 4 of the mouse Zfp423 gene), Alcaraz et al. [55] observed loss of the corpus callosum, reduction of hippocampus, and a malformation of the cerebellum reminiscent of patients with Dandy-Walker syndrome. Within the cerebellum, Zfp423 was observed to be expressed in both ventricular and external germinal zones. Loss of Zfp423 was also observed to lead to diminished proliferation by granule cell precursors in the external germinal layer and abnormal differentiation and migration of ventricular zone-derived neurons and Bergmann glia [55].

Conclusion

POLARIS is a gene-based analysis which produces a genetic risk score per individual per gene, whilst adjusting for LD between SNPs in the gene. This methodology was applied to the latest HRC imputation of the GERAD data, and the summary statistics from IGAP (excluding GERAD subjects) were used as weights in the score. This led to the identification of 3 novel genes associated with AD; these genes are *PPARGC1A*, *RORA* and *ZNF423*. There is evidence that these genes are credible candidates in AD, with *PPARGC1A* and *RORA* being linked to circadian rhythm, *PPARGC1A* is implicated in energy metabolism and the generation of amyloid plaques, *RORA* is linked to cholesterol metabolism and inflammation and *ZNF423* is likely involved in DNA damage repair and resides in an AD-specific protein network.

Materials and methods

The Haplotype Reference Consortium (HRC), version r1.1 2016, was used to impute GERAD genotype data on the Michigan Imputation Server [56], which to date, allows the most accurate imputation of genetic variants. Imputed genotype probabilities (also known as dosages) were converted to the most probable genotype with a probability threshold of 0.9 or greater. SNPs were removed if: their imputation INFO-score< 0.4,

minor allele frequency (MAF)< 0.01, missingness of genotypes≥ 0.05 or HWE$< 10^{-6}$. A total of 6,119,694 variants were retained. To correct for population structure and genotyping differences, all analyses were adjusted for age, gender and the top 3 principal components.

POLARIS was applied to this GERAD (3,332 cases, 9,832 controls; see S3 Table for cohort details) imputed data, using the IGAP [4] data (17,008 cases, 37,154 controls) excluding GERAD subjects (IGAPnoGERAD) as an external dataset to derive weights from the best powered data set avaliable. The IGAP data was imputed using a previous reference panel (1000 genomes, Dec 2010 release). There were 3,169,839 SNPs in common between imputed GERAD and IGAP summary statistics data. The GERAD imputed data contain individual genotypes for every SNP, enabling the production of a risk score per person per gene, and the IGAPnoGERAD data contains effect sizes for every SNP, which are used to weight the risk score. A gene-based risk score was produced for every individual in the GERAD data.

POLARIS adjusts for LD between SNPs and therefore, the SNPs were not pruned for LD and the entire data were used in this analysis. POLARIS adjusts for LD by using spectral decomposition of the correlation matrix between SNPs. Such a matrix was derived for each gene using the individual genotypes from the GERAD imputed data. It was ensured that SNPs had consistent reference alleles across both independent datasets; IGAPnoGERAD and imputed GERAD. If alleles in IGAPnoGERAD were coded in the opposite direction to those in GERAD, the summary effect size for the SNP was inverted. SNPs with alleles AT, TA, CG or GC were excluded.

SNPs were assigned to genes using GENCODE (v19) gene models [57]. Only genes with known gene status and those marked as protein coding were used. A gene window containing SNPs which were within 35kb upstream and 10kb downstream of the gene was considered. This window was used since it is likely to contain transcriptional regulatory elements [17]. SNPs which belong to multiple genes were assigned to all those genes. In the HRC imputed GERAD data, 2,296,690 SNPs were assigned to 18,087 genes.

A POLARIS score was produced for each of these genes, and the overall association of the gene with AD is determined using a logistic regression model, adjusting for population covariates, age and sex.

Acknowledgments

References

1. Gatz M, Reynolds CA, Fratiglioni L, et al. Role of genes and environments for explaining alzheimer disease. Archives of General Psychiatry. 2006;63(2):168–174.

2. Strittmatter WJ, Saunders AM, Schmechel D, Pericak-Vance M, et al. Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. Proc Natl Acad Sci U S A. 1993;90(5):1977–81.

3. Harold D, Abraham R, Hollingworth P, Sims R, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat Genet. 2009;41(10):1088–93.

4. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet. 2013;45(12):1452–8.

5. Hollingworth P, Harold D, Sims R, Gerrish A, et al. Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. Nat Genet. 2011;43(5):429–35.

6. Naj AC, Jun G, Beecham GW, Wang LS, et al. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. Nat Genet. 2011;43(5):436–41.

7. Seshadri S, Fitzpatrick AL, Ikram MA, DeStefano AL, et al. Genome-wide analysis of genetic loci associated with Alzheimer disease. JAMA. 2010;303(18):1832–40.

8. Brown MB. A Method for Combining Non-Independent, One-Sided Tests of Significance. Biometrics. 1975;31(4):987–992.

9. Escott-Price V, Bellenguez C, Wang LS, Choi SH, et al. Gene-wide analysis detects two new susceptibility genes for Alzheimer's disease. PLoS One. 2014;9(6):e94661.

10. Guerreiro R, Wojtas A, Bras J, Carrasquillo M, et al. TREM2 variants in Alzheimer's disease. N Engl J Med. 2013;368(2):117–27.

11. Sims R, van der Lee SJ, Naj AC, Bellenguez C, et al. Rare coding variants in PLCG2, ABI3, and TREM2 implicate microglial-mediated innate immunity in Alzheimer's disease. Nature Genetics. 2017;49:1373.

12. Li MX, Gui HS, Kwan JS, Sham PC. GATES: a rapid and powerful gene-based association test using extended Simes procedure. Am J Hum Genet. 2011;88(3):283–93.

13. Holmans P, Green EK, Pahwa JS, Ferreira MA, et al. Gene ontology analysis of GWA study data sets provides insights into the biology of bipolar disorder. Am J Hum Genet. 2009;85(1):13–24.

14. Jones L, Holmans PA, Hamshere ML, Harold D, et al. Genetic evidence implicates the immune system and cholesterol metabolism in the aetiology of Alzheimer's disease. PLoS One. 2010;5(11):e13950.

15. Jones L, Lambert JC, Wang LS, Choi SH, et al. Convergent genetic and expression data implicate immunity in Alzheimer's disease. Alzheimers Dement. 2015;11(6):658–71.

16. Baker E, Schmidt KM, Sims R, O'Donovan MC, et al. POLARIS: Polygenic LD-adjusted risk score approach for set-based analysis of GWAS data. Genetic Epidemiology. 2018;42(4):366–377.

17. O'Dushlaine C, Rossin L, Lee PH, Duncan L, et al. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. Nat Neurosci. 2015;18(2):199–209.

18. Kiezun A, Garimella K, Do R, Stitziel NO, et al. Exome sequencing and the genetic basis of complex traits. Nat Genet. 2012;44(6):623–30.

19. Luo C, Widlund HR, Puigserver P. PGC-1 Coactivators: Shepherding the Mitochondrial Biogenesis of Tumors. Trends Cancer. 2016;2(10):619–631.

20. Scarpulla RC, Vega RB, Kelly DP. Transcriptional integration of mitochondrial biogenesis. Trends Endocrinol Metab. 2012;23(9):459–66.

21. Katsouri L, Lim YM, Blondrath K, Eleftheriadou I, et al. PPARgamma-coactivator-1alpha gene transfer reduces neuronal loss and amyloid-beta generation by reducing beta-secretase in an Alzheimer's disease model. Proc Natl Acad Sci U S A. 2016;113(43):12292–12297.

22. Hood S, Amir S. Neurodegeneration and the Circadian Clock. Front Aging Neurosci. 2017;9:170.

23. Jetten AM. Retinoid-related orphan receptors (RORs): critical roles in development, immunity, circadian rhythm, and cellular metabolism. Nucl Recept Signal. 2009;7:e003.

24. Acquaah-Mensah GK, Agu N, Khan T, Gardner A. A regulatory role for the insulin- and BDNF-linked RORA in the hippocampus: implications for Alzheimer's disease. J Alzheimers Dis. 2015;44(3):827–38.

25. Hu YS, Xin J, Hu Y, Zhang L, et al. Analyzing the genes related to Alzheimer's disease via a network and pathway-based approach. Alzheimers Res Ther. 2017;9(1):29.

26. Chaki M, Airik R, Ghosh AK, Giles RH, et al. Exome capture reveals ZNF423 and CEP164 mutations, linking renal ciliopathies to DNA damage response signaling. Cell. 2012;150(3):533–48.

27. Trabzuni D, Ryten M, Walker R, Smith C, et al. Quality control parameters on a large dataset of regionally dissected human control brains for whole genome expression studies. J Neurochem. 2011;119(2):275–82.

28. Boyle AP, Hong EL, Hariharan M, Cheng Y, et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. 2012;22(9):1790–7.

29. McLaren W, Gil L, Hunt SE, Riat HS, et al. The Ensembl Variant Effect Predictor. Genome Biol. 2016;17(1):122.

30. Martinez-Redondo V, Pettersson AT, Ruas JL. The hitchhiker's guide to PGC-1alpha isoform structure and biological functions. Diabetologia. 2015;58(9):1969–77.

31. Fernandez-Marcos PJ, Auwerx J. Regulation of PGC-1alpha, a nodal regulator of mitochondrial biogenesis. Am J Clin Nutr. 2011;93(4):884S–90.

32. Jesse S, Bayer H, Alupei MC, Zugel M, et al. Ribosomal transcription is regulated by PGC-1alpha and disturbed in Huntington's disease. Sci Rep. 2017;7(1):8513.

33. Nijland PG, Witte ME, van het Hof B, van der Pol S, et al. Astroglial PGC-1alpha increases mitochondrial antioxidant capacity and suppresses inflammation: implications for multiple sclerosis. Acta Neuropathol Commun. 2014;2:170.

34. Ashabi G, Ramin M, Azizi P, Taslimi Z, et al. ERK and p38 inhibitors attenuate memory deficits and increase CREB phosphorylation and PGC-1alpha levels in Abeta-injected rats. Behav Brain Res. 2012;232(1):165–73.

35. Tsunemi T, Spada ARL. PGC-1alpha at the intersection of bioenergetics regulation and neuron function: from Huntington's disease to Parkinson's disease and beyond. Prog Neurobiol. 2012;97(2):142–51.

36. Cui L, Jeong H, Borovecki F, Parkhurst CN, et al. Transcriptional repression of PGC-1alpha by mutant huntingtin leads to mitochondrial dysfunction and neurodegeneration. Cell. 2006;127(1):59–69.

37. Soyal SM, Felder TK, Auer S, Hahne P, et al. A greatly extended PPARGC1A genomic locus encodes several new brain-specific isoforms and influences Huntington disease age of onset. Hum Mol Genet. 2012;21(15):3461–73.

38. Qin W, Haroutunian V, Katsel P, Cardozo CP, et al. PGC-1alpha expression decreases in the Alzheimer disease brain as a function of dementia. Arch Neurol. 2009;66(3):352–61.

39. Sato T, Hanyu H, Hirao K, Kanetaka H, et al. Efficacy of PPAR-gamma agonist pioglitazone in mild Alzheimer disease. Neurobiol Aging. 2011;32(9):1626–33.

40. Cook DN, Kang HS, Jetten AM. Retinoic Acid-Related Orphan Receptors (RORs): Regulatory Functions in Immunity, Development, Circadian Rhythm, and Metabolism. Nucl Receptor Res. 2015;2.

41. Vogel MW, Sinclair M, Qiu D, Fan H. Purkinje cell fate in staggerer mutants: Agenesis versus cell death. Journal of Neurobiology. 2000;42(3):323–337.

42. Kim K, Boo K, Yu YS, Oh SK, et al. RORalpha controls hepatic lipid homeostasis via negative regulation of PPARgamma transcriptional network. Nat Commun. 2017;8(1):162.

43. Chauvet C, Vanhoutteghem A, Duhem C, Saint-Auret G, et al. Control of gene expression by the retinoic acid-related orphan receptor alpha in HepG2 human hepatoma cells. PLoS One. 2011;6(7):e22545.

44. Gulec C, Coban N, Ozsait-Selcuk B, Sirma-Ekmekci S, et al. Identification of potential target genes of ROR-alpha in THP1 and HUVEC cell lines. Exp Cell Res. 2017;353(1):6–15.

45. Liu C, Li S, Liu T, Borjigin J, et al. Transcriptional coactivator PGC-1α integrates the mammalian clock and energy metabolism. Nature. 2007;447:477.

46. Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, et al. Identification and analysis of functional elements in 1human genome by the ENCODE pilot project. Nature. 2007;447(7146):799–816.

47. Miller MW, Wolf EJ, Logue MW, Baldwin CT. The retinoid-related orphan receptor alpha (RORA) gene and fear-related psychopathology. J Affect Disord. 2013;151(2):702–8.

48. Sayad A, Noroozi R, Omrani MD, Taheri M, et al. Retinoic acid-related orphan receptor alpha (RORA) variants are associated with autism spectrum disorder. Metabolic Brain Disease. 2017;32(5):1595–1601.

49. Journiac N, Jolly S, Jarvis C, Gautheron V, et al. The nuclear receptor RORα exerts a bi-directional regulation of IL-6 in resting and reactive astrocytes. Proceedings of the National Academy of Sciences of the United States of America. 2009;106(50):21365–21370.

50. Harder L, Eschenburg G, Zech A, Kriebitzsch N, et al. Aberrant ZNF423 impedes B cell differentiation and is linked to adverse outcome of ETV6-RUNX1 negative B precursor acute lymphoblastic leukemia. J Exp Med. 2013;210(11):2289–304.

51. Feng X, Che N, Liu Y, Chen H, et al. Restored immunosuppressive effect of mesenchymal stem cells on B cells after olfactory 1/early B cell factor-associated zinc-finger protein down-regulation in patients with systemic lupus erythematosus. Arthritis Rheumatol. 2014;66(12):3413–23.

52. Huang S, Laoukili J, Epping MT, Koster J, et al. ZNF423 is critically required for retinoic acid-induced differentiation and is a marker of neuroblastoma outcome. Cancer Cell. 2009;15(4):328–40.

53. Signaroldi E, Laise P, Cristofanon S, Brancaccio A, et al. Polycomb dysregulation in gliomagenesis targets a Zfp423-dependent differentiation network. Nat Commun. 2016;7:10753.

54. Bond HM, Scicchitano S, Chiarella E, Amodio N, et al. ZNF423: A New Player in Estrogen Receptor-Positive Breast Cancer. Front Endocrinol (Lausanne). 2018;9:255.

55. Alcaraz WA, Gold DA, Raponi E, Gent PM, et al. Zfp423 controls proliferation and differentiation of neural precursors in cerebellar vermis formation. Proc Natl Acad Sci U S A. 2006;103(51):19424–9.

56. Das S, Forer L, Schonherr S, Sidore C, et al. Next-generation genotype imputation service and methods. Nat Genet. 2016;48(10):1284–1287.

57. Harrow J, Frankish A, Gonzalez JM, Tapanari E, et al. GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res. 2012;22(9):1760–74.

Supporting information

S1 Appendix. Authors who contributed to the generation of original study data for GERAD, ADGC, CHARGE and EADI.

S1 Table. POLARIS Gene-Based Results for GWAS Associated Genes

S2 Table. Gene-Wide Significant Genes from POLARIS Gene-based Analysis Conditioned on *APOE* and *BCAM* and best IGAP SNP and p-value

S3 Table. **GERAD Cohort Descriptives and Sample Size**