ELSEVIER

Alzheimer's & Dementia ■ (2016) 1-5

Alzheimer's & Dementia

Short Report

APOE and *MS4A6A* interact with GnRH signaling in Alzheimer's disease: Enrichment of epistatic effects

Alejandro Cáceres[a,b,*], José E. Vargas[c], Juan R. González[a,b,d]

[a]*ISGlobal, Center for Research in Environmental Epidemiology (CREAL), Barcelona, Spain*
[b]*Centro de Investigación Biomédica en Red en Epidemiología y Salud Pública (CIBERESP), Madrid, Spain*
[c]*CAPES Foundation, Ministry of Education of Brazil, Brasilia, Brazil*
[d]*Department of Mathematics, Universitat Autònoma de Barcelona, Barcelona, Spain*

Abstract

Introduction: It is unknown if risk loci, identified by genome-wide association studies of late-onset Alzheimer's disease (LOAD), are linked to common molecular mechanisms through epistatic effects.
Methods: We performed genome-wide interaction studies of five risk variants for LOAD followed by enrichment analyses to find if there are pathways that simultaneously interact with more than one variant. This novel approach was applied to four independent cohorts (5393 cases and 3746 controls).
Results: We found enrichment of epistasis in gonadotropin-releasing hormone signaling with risk single-nucleotide polymorphisms in *APOE* and *MS4A6A* (P value $= 3.7 \times 10^{-5}$, P value $= 5.6 \times 10^{-6}$); vascular smooth muscle contraction pathway was also enriched in epistasis with these loci (P value $= 9.6 \times 10^{-5}$, P value $= 2.4 \times 10^{-7}$). *MS4A6A* risk variant also interacted with dilated cardiomyopathy pathway (P value $= 3.1 \times 10^{-7}$).
Discussion: In addition to *APOE*, *MS4A6A* polymorphisms should be considered in hormone trials targeting gonadotropins. Interactions of risk variants with neurovascular pathways may also be important in LOAD pathology.
© 2016 The Alzheimer's Association. Published by Elsevier Inc. All rights reserved.

Keywords: APOE; MS4A6A; PICALM; Gonadotropin; GnRH signaling; Epistasis; Alzheimer's disease; GWAS; Pathway analysis; SNP

1. Introduction

Gene–gene interactions are widely recognized as a fundamental factor in the formation of heritable traits [1]. Although single genetic variants can have great influence in the variability of specific traits, genes and their products do not act alone. Genome-wide association studies (GWASs) have successfully identified single-nucleotide polymorphisms (SNPs) associated with late-onset Alzheimer's disease (LOAD) [2]. Several SNP–SNP interactions have also been reported [3]. However, it remains to establish if risk SNPs share between themselves epistatic links to molecular mechanisms relevant for the disease.

Here, we propose to characterize the genome-wide interactions of specific risk SNPs to help identify their epistatic role within the trait. This naturally suggests the integration of genome-wide interaction associations with pathway analysis, which enables the search for pathways that interact with more than one risk variant, giving a hint on those gene sets that may couple with risk SNPs to potentiate their additive effects. In addition, such integration can help to interpret previous GWAS results as it can reveal links between risk SNPs at the molecular level. We analyzed the interactions of five risk variants of LOAD in a total 5393 cases and 3746 controls divided in four independent studies.

The authors report no conflicts of interest.
*Corresponding author. Tel.: +34 93 214 73 00; Fax: +34 93 214 73 02.
E-mail address: acaceres@creal.cat

http://dx.doi.org/10.1016/j.jalz.2016.05.009
1552-5260/© 2016 The Alzheimer's Association. Published by Elsevier Inc. All rights reserved.

ARTICLE IN PRESS

2. Methods

2.1. Data

Four studies from dbGAP [4] were analyzed. European ancestry was selected in all four studies: (1) National Institute of Aging (NIA) study (accession: phs000168.v1.p1) with 587 cases, 289 controls, and 590,247 SNPs; (2) GenADA study (phs000219.v1.p1) with 806 cases, 782 controls, and 349,252 SNPs; and (3–4) Alzheimer's Disease Genetics Consortium (phs000372.v1.p1). We kept the two genotyped batches, ADG12 and ADG3, as two distinct studies. In ADG12, we analyzed 2686 cases, 935 controls, and 592,652 SNPs and, in ADG3, 975 cases, 578 controls, and 681,273 SNPs.

We analyzed SNPs with minor allele frequency (>1%) and Hardy–Weinberg equilibrium ($Z^2 < 16$). Genome-wide principal components were calculated with Bioconductors' SNPStats package first to remove outliers (>4 SD) and afterward to adjust for stratification.

2.2. Selection of risk SNPs

Previous GWASs have identified risk SNPs of LOAD in dozens of genes [2,5–7]. We selected SNPs whose AlzGene meta-analysis was based on 10 studies or more [8] and which were genotyped or imputed with high accuracy in all the studies that we analyzed (Supplementary Table 1). We selected in total 5 SNPs, including rs429358 in *APOE*, which is the SNP that defines the *APOE* ε4 allele and which was independently genotyped in all the studies; rs744373 (*BIN1*); rs3818361 (*CR1*); rs3851179 (*PICALM*); and rs610932 (*MS4A6*A). From all selected SNPs, only rs610932 was imputed in GenADA with IMPUTE2 with a quality score of 0.998.

2.3. Genome-wide interaction study

We performed genome-wide interaction associations for the five risk SNPs selected in NIA, GenADA, ADG12, and ADG3. Fixing a risk SNP, genome-wide *P* values were obtained from the likelihood ratios, χ^2 (1), between the logistic models

y = SNP × riskSNP + SNP + riskSNP + covariates and y = SNP + riskSNP + covariates, where y was case-control status, SNP with genotypes coded (1, 2, 3) varied over the genome, and the covariates were sex, age of diagnosis if available, genome-wide principal components, and the principal components times the risk SNP. Q-Q plots were computed with SNPStats to verify correct adjustment by population stratification. All models were fitted with arm, an R-package for Bayesian regression which was robust for SNP interactions with low frequency.

2.4. Enrichment of epistatic effects for each risk SNP

For each risk SNP and study, we looked for enriched pathways with iGSEA4GWAS-v2 [9] that allows the simultaneous use of multiple data sources, like Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, KEGG diseases, and BioCarta. The method is easily adaptable to genome-wide interaction study (GWIS) results as its input is genome-wide *P* values. We mapped genes to SNPs within 100 Kb distance and performed pathway analyses for the GWAS and five GWIS within each cohort. We then identified the pathways that significantly interacted with more than one risk SNP. Fig. 1 illustrates the workflow. We set a stringent control for false-positive findings by (1) performing meta-analyses over studies at the pathway level and thus accounting for between-cohort variability and (2) setting two types of criteria for statistical significance and one criteria for repeatability. First, we computed Fisher combined probability test over the studies for each pathway uncorrected *P* value (e.g., combination of uncorrected *P* values) and set Bonferroni's significance threshold at 2.0×10^{-4} to account for the 240 pathways tested. Second, we computed Fisher test over studies for corrected *P* values (e.g., combination corrected *P* values) and set the significance threshold at .05. Third, we selected pathways with repeatable nominal significance at each and every study.

We also performed comprehensive simulations to determine the power and false discovery rate for the enrichment of epistatic effects under realistic scenarios (Supplementary Methods).

3. Results

3.1. GWAS and enrichment of individual SNP effects

Supplementary Table 2 shows the most significant results of the GWAS where a region in high linkage disequilibrium (LD) with *APOE*, including *TOMM40*, was clearly associated with LOAD for the NIA, ADG12, and ADG3 studies (Supplementary Figs. 1-4).

We performed enrichment analysis for the GWAS results of each separate study by testing a total of 240 pathways from KEGG and BioCarta (Supplementary Table 3). We observed four significant pathways for the combination of uncorrected *P* values. Although our results validated T cell receptors, neurotrophin and Wnt signaling pathways, recently found to be significantly associated with LOAD in an enrichment analysis of a GWAS meta-analysis [5], none showed full repeatability over all the four studies.

3.2. GWIS and enrichment of epistatic effects of risk loci

GWIS analyses are shown in Supplementary Table 4 and Supplementary Figs. 1-4. We observed one significant SNP–SNP interaction at genome-wide level in one cohort. rs610932 interacted SNPs within *NEDD4*, a gene previously associated with dementia severity [10].

We assessed which pathways were enriched in epistatic effects with the five risk SNPs considered. We thus conducted an enrichment of GWIS. Results are shown in Table 1. For

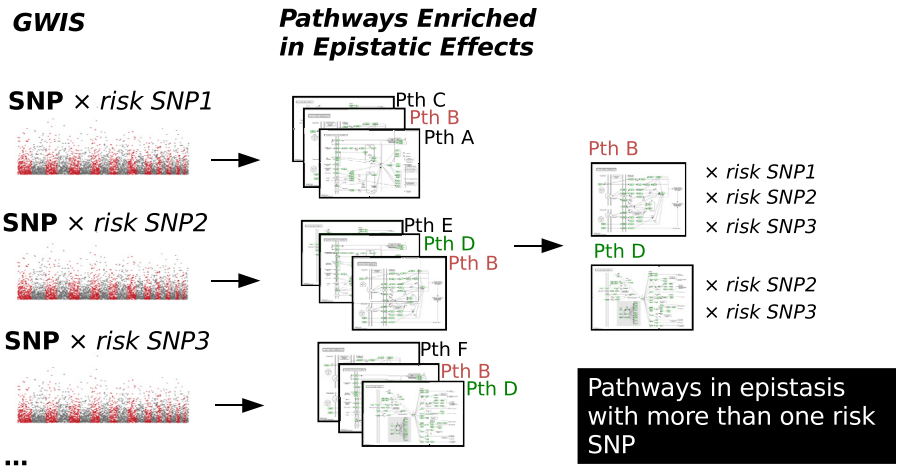

Fig. 1. Integration of genome-wide interaction study (GWIS) of risk variants with pathway analysis. We consider risk single-nucleotide polymorphisms (SNPs) that have been found and validated in previous studies. A GWIS is performed for each risk SNP followed by its corresponding enrichment analysis. Consequently, each risk SNP has an associated list of pathways enriched with epistatic effects. Pathways can have enriched epistasis with more than one risk SNP.

combined uncorrected *P* values, we found five significant pathways interacting with rs429358 (*APOE*), one with rs744373 (*BIN1*), four with rs3851179 (*CR1*), three with rs3851179 (*PICALM*), and 11 with rs610932 (*MS4A6A*). Many of the pathways have been reported with enrichment analysis of individual SNP effects [5,11], enlarging our previous list of enriched pathways for GWAS. We found three interacting pathways that consistently replicated in all four studies and were significant for the combined corrected *P* values; most notably, gonadotropin (GnRH)

Table 1
Pathways enriched in interactions with selected risk loci for LOAD

Variant (gene)	Pathway	CU *P* value	CC *P* value	ADG12	ADG3	GENADA	NIA
rs429358 (*APOE*) ×	**KEGG: GnRH signaling pathway**	**3.78E−5**	**.02**	**0.001**	**0.046**	**0.025**	**0.033**
	KEGG: Long-term potentiation	1.63E−5	.02	0.001	0.10	0.15	0.001
	KEGG: ARVC	*2.98E−5*	*.04*	*0.001*	*0.20*	*0.00*	*0.15*
	KEGG: Calcium signaling pathway	1.12E−4	.06	0.01	0.09	0.16	0.001
	KEGG: Vascular smooth muscle contraction	*9.66E−5*	*.07*	*0.001*	*0.27*	*0.22*	*0.002*
rs744373 (*BIN1*) ×	KEGG: Focal adhesion	2.33E−5	.24	0.002	0.01	0.89	0.001
rs3818361 (*CR1*) ×	KEGG: Spliceosome	1.07E−4	.03	0.04	0.003	0.004	0.31
	KEGG: Type II diabetes mellitus	1.09E−4	.07	0.02	0.002	0.12	0.02
	KEGG: Tight junction	1.07E−4	.08	0.38	0.001	0.09	0.004
	KEGG: Axon guidance	1.47E−4	.19	0.13	0.03	0.06	0.001
rs3851179 (*PICALM*) ×	KEGG: Cardiac muscle contraction	8.24E−5	.01	0.07	0.44	0.001	0.003
	KEGG: Pentose and glucuronate interconversions	1.08E−4	.03	0.52	0.005	0.001	0.05
	KEGG: ARCV	*2.66E−5*	*.08*	*0.002*	*0.04*	*0.001*	*0.30*
rs610932 (*MS4A6A*) ×	KEGG: Phosphatidylinositol signaling system	1.02E−4	.009	0.01	0.001	0.03	0.36
	KEGG: Dilated cardiomyopathy	**3.12E−7**	**.02**	**0.001**	**0.01**	**0.010**	**0.001**
	BioCarta: HDAC pathway	6.49E−6	.02	0.15	0.001	0.002	0.02
	KEGG: Vascular smooth muscle contraction	**2.40E−7**	**.03**	**0.01**	**0.008**	**0.001**	**0.001**
	KEGG: GnRH signaling pathway	**5.60E−6**	**.05**	**0.01**	**0.009**	**0.010**	**0.004**
	KEGG: Hypertrophic cardiomyopathy HCM	1.05E−4	.07	0.004	0.13	0.25	0.001
	BioCarta: NKT pathway	1.75E−4	.11	0.20	0.06	0.02	0.001
	KEGG: Galactose metabolism	8.63E−5	.12	0.005	0.01	0.79	0.002
	BioCarta: PGC1A pathway	1.69E−4	.14	0.41	0.03	0.02	0.001
	KEGG: Long-term depression	1.43E−4	.15	0.02	0.01	0.02	0.07

Abbreviation: ARVC, arrhythmogenic right ventricular cardiomyopathy

NOTE. The table shows (1) all significant associations at Bonferroni level ($<2.0 \times 10^{-4}$) for the combined uncorrected *P* values (CU *P* value); (2) combined corrected *P* values (CC *P* value), where corrected *P* values are gene-set false discovery rates as given by iGSEA4GWAS; and (3) significant associations at nominal *P* value in all the four studies in bold in the last four columns. Pathway names in bold face are significant under all the three criteria for significance. Pathway names in italics are in epistasis with more than one SNP only under CU P value.

signaling interacted with both rs429358 and rs610932. Vascular smooth muscle contraction also interacted with rs610932 and rs429358, whereas dilated cardiomyopathy interacted only with rs610932.

As expected, simulations showed that power increases with number of genes involved in the pathway, sample size, and odds ratio (OR) of interaction (Supplementary Figs. 5-9). The method is well powered to detect epistatic interactions with pathways with number of genes >46, number of subjects >1000, and interaction OR >1.1 (Supplementary Fig. 6).

4. Discussion

Hormonal processes underlying LOAD are gaining renewed interest [12,13], as GnRH function in the brain has been recently linked to aging [14]. A recent study has shown that increasing levels in plasma of luteinizing hormone, whose secretion is regulated by GnRH, increment brain amyloid burden, which depended on the presence of *APOE* ε4, i.e. rs429358 [15]. Our assessment of enriched epistasis revealed that GnRH signaling robustly interacts with rs429358 providing genomic support to this observation. Notably, LOAD associations with GnRH pathway have also been reported for genomic and gene expression data [5,16]. Our results also validate those previous findings. As leuprolide acetate, a GnRH agonist, has been recently tested in a phase 2 clinical trial of LOAD [17], our results suggest that both *APOE* and *MS4A6A* polymorphisms should be considered in the assessment of treatment response.

A genomic pathway analysis of an imaging endophenotype for LOAD showed a significant association with vascular smooth muscle contraction [11]. Brain-blood barrier damage can impair amyloid beta clearance from the brain, whereas cerebral blood flow reduction has been associated with the *APOE* ε4 allele in undemented individuals [18,19]. Our results underline the need to test other risk variants in addition to rs429358 (*APOE*), such as rs610932 (*MS4A6A*), in the evaluation these effects. We also found that the *MS4A6A* risk locus significantly interacted with dilated cardiomyopathy signaling, for which a missense mutation in *PSEN1* shows complete penetrance [20]. Stronger enrichment associations of epistasis could be obtained with a broader selection of risk SNPs.

We found a number of pathways with both repeatable interactions in all cohorts and statistical significance under multiple comparisons. Adoption of repeatability criteria can help in identifying the most consistent observations, not governed only by few strong associations. Although we adopt the conservative approach of full repeatability, on yet nominal associations, other measures of repeatability should be investigated. We also found that the enrichment of epistasis has greater power compared with enrichment of single SNP associations in cohorts smaller than 5000 (Supplementary Fig. 9) and increased the number of associated pathways (Supplementary Tables 3 and 4). Although these properties remain to be seen in other implementations of enrichment algorithms, an important gain of this novel analysis approach is that it links specific risk variants to gene pathways, helping to study the role of risk SNPs in the etiology of the disease through epistasis.

Acknowledgments

Spanish Minister of Science and Innovation supported J.R.G (MTM2011-26515). CAPES supported J.E.V. (process number: BEX 6821/14-6). We would like to thank Mariona Bustamante for useful comments.

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jalz.2016.05.009.

RESEARCH IN CONTEXT

1. Systematic review: We used the catalog of published GWAS to look for the studies that have reported the strongest SNP associations with LOAD. We also used PubMed and Google Scholar to search for articles under the terms "GnRH" and "Alzheimer's." Relevant studies involving our findings are cited.
2. Interpretation: Enrichment of epistatic effects, here implemented, can identify molecular pathways that interact with more than one risk locus. We observed that risk variants within *APOE* and *MS4A6A* converged with their interactions in GnRH signaling, a current target for pharmacologic interventions for LOAD. Evaluation of such trials should consider potential interactions with both *APOE* and *MS4A6A* polymorphisms.
3. Future directions: Deepening our understanding of the role of GnRH signaling in LOAD could inform which individuals would respond best to targeted hormone treatments. Enrichment of epistatic effects can be an important tool to interpret previous GWAS findings, helping to link risk variants at pathway level.

References

[1] Phillips PC. Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. Nat Rev Genet 2008; 9:855–67.

[2] Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C. Meta-analysis of 74,046 individuals identifies 11 new

ARTICLE IN PRESS

susceptibility loci for Alzheimer's disease. Nat Genet 2013; 45:1452–8.

[3] Combarros O, Cortina-Borja M, Smith AD, Lehmann DJ. Epistasis in sporadic Alzheimer's disease. Neurobiol Aging 2009;9:1333–13349.

[4] dBGap. Avaliable at: http://www.ncbi.nlm.nih.gov/gap. Accessed: June 21, 2016.

[5] Pérez-Palma E, Bustos BI, Villamán CF, Alarcón MA, Avila ME, Ugarte GD, et al. Overrepresentation of glutamate signalling in Alzheimer's disease: Network-based pathway enrichment using meta-analysis of genome-wide association studies. PLoS One 2014; 9:e95413.

[6] Hollingworth P, Harold D, Sims R, Gerrish A, Lambert JC, Carrasquillo MM, et al. Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. Nat Genet 2011;43:429–35.

[7] Antúnez C, Boada M, González-Pérez A, Gayán J, Ramírez-Lorca R, Marín J, et al. The membrane-spanning 4-domains, subfamily A (MS4A) gene cluster contains a common variant associated with Alzheimer's disease. Genome Med 2011;3:1–8.

[8] AlzGene. Available at: http://AlzGene.org. Accessed: June 21, 2016.

[9] Zhang K, Cui S, Chang S, Zhang L, Wang J. i-GSEA4GWAS: a web server for identification of pathways/gene sets associated with traits by applying an improved gene set enrichment analysis to genome-wide association study. Nucleic Acids Res 2010; 38:W90–5.

[10] Jahanshad N, Rajagopalan P, Hua X, Hibar DP, Nir TM, Toga AW, et al. Genome-wide scan of healthy human connectome discovers SPON1 gene variant influencing dementia severity. Proc Natl Acad Sci U S A 2013;110:4768–73.

[11] Silver M, Janousova E, Hua X, Thompson PM, Montana G, Alzheimer's Disease Neuroimaging Initiative. Identification of gene pathways implicated in Alzheimer's disease using longitudinal imaging phenotypes with sparse regression. Neuroimage 2012; 63:1681–94.

[12] Wang L, Chadwick W, Park SS, Zhou Y, Silver N, Martin B, et al. Gonadotropin-releasing hormone receptor system: modulatory role in aging and neurodegeneration. CNS Neurol Disord Drug Targets 2010;9:651–60.

[13] Verdile G, Asih PR, Barron AM, Wahjoepramono EJ, Ittner LM, Martins RN. The impact of luteinizing hormone and testosterone on beta amyloid (A-Beta) accumulation: Animal and human clinical studies. Horm Behav 2015;76:81–90.

[14] Zhang G, Li J, Purkayastha S, Tang Y, Zhang H, Yin Y, et al. Hypothalamic programming of systemic ageing involving IKK-beta, NF-kbeta and GnRH. Nature 2013;497:211–6.

[15] Verdile G, Laws SM, Henley D, Ames D, Bush AI, Ellis KA, et al. Associations between gonadotropins, testosterone and beta amyloid in men at risk of Alzheimer's disease. Mol Psychiatry 2014;19:69–75.

[16] Dutta B, Wallqvist A, Reifman J. PathNet: A tool for pathway analysis using topological information. Source Code Biol Med 2012;7:10.

[17] Bowen RL, Perry G, Xiong C, Smith MA, Atwood CS. A clinical study of lupron depot in the treatment of women with Alzheimer's disease: preservation of cognitive function in patients taking an acetylcholinesterase inhibitor and treated with high dose lupron over 48 weeks. J Alzheimers Dis 2015;44:549–60.

[18] Zlokovic BV. Neurovascular pathways to neurodegeneration in Alzheimer's disease and other disorders. Nat Rev Neurosci 2011; 12:723–38.

[19] Thambisetty M, Beason-Held L, An Y, Kraut MA, Resnick SM. APOE epsilon4 genotype and longitudinal changes in cerebral blood flow in normal aging. Arch Neurol 2010;67:93–8.

[20] Li D, Parks SB, Kushner JD, Nauman D, Burgess D, Ludwigsen S, et al. Mutations of presenilin genes in dilated cardiomyopathy and heart failure. Am J Hum Genet 2006;79:1030–9.