Alzheimer's & Dementia 13 (2017) 493-497

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.

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

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