



Published in final edited form as:

Psychiatr Genet. 2010 December ; 20(6): 321–324. doi:10.1097/YPG.0b013e32833b635d.

PCDH11X variation is not associated with late-onset Alzheimer disease susceptibility

Gary W Beecham^{1,*}, Adam Naj¹, John R Gilbert¹, Jonathan L Haines², Joseph D Buxbaum³, and Margaret A Pericak-Vance¹

¹ John P. Hussman Institute for Human Genomics and the Dr. John T. Macdonald Foundation Department of Human Genetics, Miller School of Medicine, University of Miami, Miami FL

² Center for Human Genetics research, Vanderbilt University, Nashville TN

³ Laboratory of Molecular Neuropsychiatry, Mount Sinai School of Medicine, New York, NY

SUMMARY

A recent genome-wide association study and follow-up shows significant association with the *protocadherin 11 X-linked* (PCDH11X) gene. Carrasquillo et al. (2009) show statistical association with four PCDH11X polymorphisms (rs5984894, rs2573905, rs5941047, rs4568761) in five of seven cohorts. The combined analysis of 2,356 cases and 2,384 controls showed the strongest association with a p-value of 2.2×10^{-7} with an allele specific odds ratio of 1.30 (95% CI, 1.18–1.43) at the rs5984894 polymorphism. We tested for association at these four SNPs in two independent datasets and then performed a joint analysis. Though we had adequate power to detect effects sizes with the reported odds ratios, we did not detect association between LOAD and the PCDH11X polymorphisms in our dataset of 889 cases and 850 controls, indicating that the PCDH11X association, if not a false positive, is not as strong or generalized as previously hypothesized.

Keywords

Alzheimer disease; genetic association study; PCDH11X

INTRODUCTION

Late onset Alzheimer disease (LOAD [MIM 104300]) is the leading cause of dementia in the elderly and has a complex etiology, with a strong genetic risk component. A recent genome-wide association study implicated the *protocadherin 11 X-linked* (PCDH11X) gene as a risk locus for LOAD Carrasquillo et al. (2009). Their combined dataset contained 2,356 cases and 2,384 controls, and detected association at the rs5984894 SNP (p-value of 2.2×10^{-7} , odds ratio = 1.3 [C.I., 1.18–1.43]). To validate this result, we examined the rs5984894 SNP and three others (rs2573905, rs5941047, and rs4568761) in two independent Alzheimer datasets and tested for the association between the four SNPs and LOAD.

*Corresponding Author: Gary Beecham, John P. Hussman Institute for Human Genomics, 1501 NW 10th Avenue, BRB 417, M860, Miami, FL 33136, gbeecham@med.miami.edu, Telephone: 305-243-7965, Fax: 305-243-2396.

MATERIALS AND METHODS

Sample Ascertainment

The Beecham sample set is derived from the Collaborative Alzheimer Project (CAP, the John P. Hussman Institute for Human Genomics at the Miller School of Medicine University of Miami and the Center for Human Genetics Research at Vanderbilt University Medical Center) as described in Beecham et al. (2009). The Naj sample set is derived from additional samples from CAP, samples from the National Cell Repository for Alzheimer's Disease (NCRAD), and an autopsy series from Mount Sinai Medical (Naj et al., 2009). The autopsy series consisted of 306 cases and 81 controls recently deceased who had affection status verified through both clinical review and autopsy. Written informed consent was obtained for all participants, in agreement with protocols approved by the institutional review board at each contributing center. Each LOAD affected individual met the NINCDS-ADRDA criteria for possible, probable or definite AD and had an age at onset (AAO) greater than 60 years of age McKhann et al. (1984). All cognitive controls were examined; none showed signs of dementia in clinical history or upon interview and each had a Mini-Mental State Exam (MMSE) score > 27 or a Modified Mini-Mental State (3MS) Exam score > 87 Teng and Chui (1987). No cases and only 81 controls were obtained through NCRAD, so overlap with the Carrasquillo dataset, if any, is minimal.

After quality control filters the Beecham sample set contains a total of 988 individuals (492 LOAD cases 496 cognitive controls), and the Naj sample set contains a total of 755 samples (399 cases and 356 controls). Both sample sets are described in Table 1.

Genotyping

For both Beecham and Naj sample sets, DNA was extracted using Puregene chemistry (QIAGEN, Germantown, MD, USA). For the Beecham samples, we performed genotyping using the Illumina Beadstation and the Illumina HumanHap 550 beadchip, following the recommended conditions. For the Naj samples, we used the Illumina HumanHap 1M beadchip. Genotyping efficiency was greater than 99%, and quality assurance was achieved by the inclusion of two CEPH controls that were genotyped multiple times. The lab was blinded to affection status and quality-control samples. Sample and SNP quality control measures are discussed in Beecham et al. (2009). All samples were subject to a 98% efficiency threshold and were tested for gender consistency and relatedness. All SNPs were subject to efficiency, MAF, and Hardy-Weinberg thresholds. Heterozygous genotypes for males were set to missing. The four SNPs that were part of the primary analysis (rs5984894, rs2573905, rs5941047, and rs4568761) each had over 99% genotyping efficiency, though the rs2573905 SNP was only genotyped on the 1M chip in the Naj sample set.

Statistical Analysis

We tested for population substructure using Eigenstrat software Price et al. (2006) under the default settings (including five iterations of outlier removal). Hardy-Weinberg equilibrium was tested using an exact test in PLINK Purcell et al. (2007). We tested for association using logistic regression in PLINK. For the rs5984894 SNP we used both an additive model (Table 2) and a genotypic model (Table 3). For the genotypic models, males and females were tested separately (male hemizygotes for the major allele were the referent in the male-only analysis; female homozygotes for the major allele were the referent in the female-only analysis). For rs2573905, rs5941047, and rs4568761 we used the additive model (Table 4). Sex was included as a covariate in the additive model and the top three principle components were included to correct for any population substructure. Power was calculated using the Genetic Power Calculator Purcell et al. (2003).

RESULTS & DISCUSSION

The PCDH11X polymorphisms were not associated with late-onset Alzheimer disease in our datasets. The rs5984894 SNP was not associated with LOAD in either dataset and had a p -value=0.80 in the combined analysis, with an odds ratio of 1.02 (Table 1; 95% CI: 0.88–1.18). With this sample size, we have over 93% power to detect the reported odds ratio of 1.30 at the α =0.10 level. To test for sex or genotype specific effects we used logistic regression and a genotypic model on males and females separately (Table 2). There was no association detected in either the hemizygous males (p =0.59), female heterozygotes (p =0.90), or the female homozygotes (p =0.95). Genotypes were obtained on 1,739 of 1,743 samples (efficiency = 99.8%), there was no Hardy-Weinberg disequilibrium (case p -value = 0.25, control p -value = 0.29), and there were very few heterozygous males (1.3%). Minor allele frequency of the rs5984894 SNP for controls was comparable to that found by Carrasquillo et al, implying there were no major differences in genotyping between the different platforms. There were three additional SNPs highlighted in the Carrasquillo manuscript: rs2573905, rs5941047, and rs4568761. None of these SNPs were associated with LOAD (Table 3; p -values=0.96, 0.35, 0.41 respectively), though the rs2573905 SNP was not genotyped in the Beecham et al dataset. Finally, we performed association analysis across the gene (Figure 1). None of the 80 SNPs tested in the Naj dataset were associated with LOAD. Three of the 23 SNPs tested in the Beecham dataset showed nominal association (rs370928, rs453810, rs117393; p -value = 0.03, 0.04, 0.03), though none would survive a multiple testing correction. These SNPs were not in strong LD with those reported in Carrasquillo et al ($r^2 < 0.30$).

We do not replicate the PCDH11X association signal. We have 93% power to detect the reported odds ratio (1.30) at an α =0.10 level and high quality genotyping, yet do not have statistical association. This suggests that any effect of PCDH11X on Alzheimer risk in our population is less than OR=1.30. Given that we have 80% power to detect an effect as small as an OR of 1.22, any effect of the PCDH11X polymorphisms is likely smaller than this. It is of note that we do not have adequate power to detect effects in the lower range of the reported OR confidence interval (1.18 to 1.43). Follow-up studies often fail to detect association because initial association results tend to over-estimate genetic effects sizes. This “Winner’s curse” effect can make replication difficult Goring et al. (2001); Kraft (2008).

It is also possible that loci other than those tested confer risk (allelic heterogeneity) or that polymorphisms may confer risk only in individuals with particular genetic backgrounds or environments (gene-gene or gene-environment interaction). Although our datasets seem quite comparable, if the Carrasquillo datasets were enriched for the susceptibility backgrounds or environments and our cohorts were not, then they would detect association and we would not. Though PCDH11X is an intriguing candidate gene, it must be further studied to determine its role, if any, in the etiology of Alzheimer disease.

Acknowledgments

We thank the AD patients and their families for participating in our study. This work was supported by grants from the National Institutes of Health: National Institute on Aging (AG20135, AG19757, AG010491, AG0005138, AG002219), National Institute of Neurological Disorders and Stroke (NS31153), the Alzheimer Association, and the Louis D. Scientific Award of the Institut de France. A subset of the participants was ascertained while Margaret A. Pericak-Vance was a faculty member at Duke University.

References

Beecham GW, Martin ER, Li YJ, Slifer MA, Gilbert JR, Haines JL, Pericak-Vance MA. Genome-wide association study implicates a chromosome 12 risk locus for late-onset Alzheimer disease. *Am J Hum Genet* 2009;84:35–43. [PubMed: 19118814]

- Carrasquillo MM, Zou F, Pankratz VS, Wilcox SL, Ma L, Walker LP, Younkin SG, Younkin CS, Younkin LH, Bisceglia GD, Ertekin-Taner N, Crook JE, Dickson DW, Petersen RC, Graff-Radford NR, Younkin SG. Genetic variation in PCDH11X is associated with susceptibility to late-onset Alzheimer's disease. *Nat Genet* 2009;41(2):192–8. [PubMed: 19136949]
- Goring HH, Terilliger JD, Blangero J. Large upward bias in estimation of locus-specific effects from genome-wide scans. *Am J Hum Genet* 2001;69:1149–1161. [PubMed: 11590550]
- Kraft P. Curses --winner's and otherwise--in genetic epidemiology. *Epidemiology* 2008;19(5):649–51. [PubMed: 18703928]
- McKhann G, Drachman D, Folstein M. Clinical diagnosis of Alzheimer's disease: Report of the NINCDSADRDA Work Group under the auspices of the Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 1984;34:939–944. [PubMed: 6610841]
- Naj AC, Beecham GW, Martin ER, Gallins PJ, Powell EH, Konidari I, Whitehead PL, Cai G, Kajiwara Y, Haroutunian H, Scott WK, Vance J, Slifer MA, Gwirstman HE, Gilbert JR, Haines JL, Pericak-Vance MA. Dementia Revealed: Novel Chromosome 6 Locus for Late-Onset Alzheimer Disease Provides Genetic Evidence for Folate-Pathway Abnormalities. Submitted. *PLoS: Genetics*. 2009
- Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006;38:904–909. [PubMed: 16862161]
- Purcell S, Cherny SS, Sham PC. Genetic Power Calculator: Design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 2003;19:149–150. [PubMed: 12499305]
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC. PLINK: a toolset for whole-genome association and population-based linkage analysis. *Am J Hum Genet* 2007;81:559–575. [PubMed: 17701901]
- Teng EL, Chui HC. The modified Mini-Mental state (3MS) examination. *J Clin Psychiatry* 1987;48:314–318. [PubMed: 3611032]

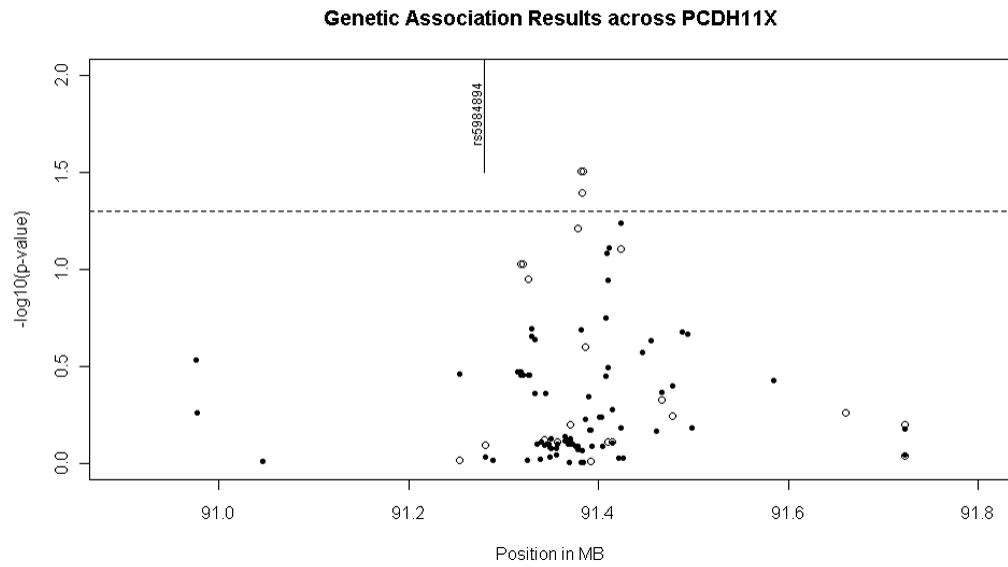


FIGURE 1. Plot of the PCDH11X region; open circles are from the Beecham et al dataset, filled-in circles are from the Naj dataset. The dashed line represents the $p\text{-value}=0.05$ line.

TABLE 1

Description of Sample sets

	Beecham dataset	Naj dataset
Sample Size	988	755
Cases	492	399
Controls	496	356
Age at Onset (cases)	72.9 (+/- 6.6)	81.7 (+/- 10.2)
Age at Exam (controls)	74.3 (+/- 6.5)	74.1 (+/- 9.0)
Percent female		
Cases	61%	68%
Controls	63%	59%

TABLE 2

Descriptive statistics and additive association results for SNP rs5984894.

Series	N		MAF*		HWE		P-value	OR** (95% CI)
	Cases	Controls	Cases	Controls	Cases	Controls		
Beecham et al	491	496	0.47	0.46	0.07	0.25	0.81	1.02 (0.85–1.24)
Naj et al	398	354	0.51	0.48	0.81	0.99	0.90	0.98 (0.78–1.25)
Combined	889	850	0.49	0.47	0.25	0.29	0.80	1.02 (0.88–1.18)
Carrasquillo et al Combined	2,356	2,384	0.52	0.46	0.47	0.82	2.2×10^{-7}	1.30 (1.18–1.43)

* The "A" allele is the minor allele (positive strand).

** Odds ratios are calculated with the major allele "G" as the referent.

TABLE 3

Logistic regression results for rs5984894 genotypic model.

Series	Male Hemizygous		Female heterozygotes		Female homozygotes	
	OR* (95% CI)	P-value	OR* (95% CI)	P-value	OR* (95% CI)	P-value
Beecham et al	0.92 (0.61–1.39)	0.70	0.98 (0.68–1.41)	0.90	1.11 (0.72–1.71)	0.64
Naj et al	1.20 (0.78–1.85)	0.40	0.95 (0.63–1.43)	0.81	0.97 (0.65–1.64)	0.90
Combined	1.09 (0.80–1.49)	0.59	1.02 (0.76–1.36)	0.90	0.99 (0.71–1.38)	0.95
Carrasquillo et al Combined	1.18 (0.99–1.41)	0.07	1.26 (1.05–1.51)	0.01	1.75 (1.42–2.16)	2.0×10^{-7}

* Odds ratios are calculated with the major allele "C" as the referent (positive strand).

TABLE 4

Descriptive and association statistics for rs2573905, rs5941047, rs4568761.

SNP	Minor Allele*	N		MAF		HWE		Carrasquillo et al combined		Beecham & Naj combined	
		Cases	Controls	Cases	Controls	Cases	Controls	P-value	OR** (95% CI)	P-value	OR** (95% CI)
rs2573905	T	398	353	0.50	0.47	0.76	0.60	1.6×10^{-7}	1.29 (1.17–1.42)	0.96	0.99 (0.78–1.26)
rs5941047	C	886	852	0.41	0.38	0.64	0.17	8.0×10^{-5}	1.21 (1.10–1.34)	0.35	1.08 (0.92–1.26)
rs4568761	C	891	852	0.43	0.40	0.88	0.58	0.001	1.17 (1.07–1.29)	0.41	1.07 (0.92–1.24)

* Alleles shown for positive strand.

** Odds ratios are calculated with the major allele as the referent.