Lack of association between protocadherin 11-X/Y (*PCDH11X* and *PCDH11Y*) polymorphisms and Late Onset Alzheimer's Disease

Ana Miar¹, Victoria Álvarez¹, Ana I. Corao¹, Belén Alonso¹, Marta Díaz¹, Manuel Menéndez², Carmen Martínez³, Maite Calatayud⁴, Germán Morís⁴, Eliecer Coto^{1,5}

¹ Genética Molecular, Hospital Universitario Central Asturias, Oviedo, Spain.

² Neurología, Hospital A. Buylla, Mieres, Spain.

³ Neurología, Hospital Cabueñes, Gijón, Spain.

⁴ Neurología, Hospital Universitario Central Asturias, Oviedo, Spain.

⁵ Departamento de Medicina, Universidad de Oviedo, Oviedo, Spain.

Corresponce to: Dr. Eliecer Coto

Genética Molecular.

Hospital Universitario Central de Asturias

33006. Oviedo. Spain

Tel./ Fax. (34).985.27.36.57.

Email: eliecer.coto@sespa.princast.es

Keywords:

Alzheimer's disease; Protocadherin 11X/Y; DNA Polymorphisms; Genetic association.

Short title:

PCHD11X/Y polymorphisms in Alzheimer's disease

Abstract

A recent genome-wide association study (GWA) reported a significant association between single nucleotide polymorphisms (SNPs) at the *PCDH11X* gene and late-onset Alzheimer's disease (LOAD). Our research was designated to replicate this association, including non previously analyzed *PCDH11X* and *PCDH11Y* SNPs. We genotyped four *PCDH11X* and one *PCDH11Y* SNPs in a total of 420 LOAD patients and 350 healthy controls from Spain. Allele and genotype frequencies did not differ between patients and controls for the five SNPs, even after correcting by gender, age, and *APOE-ε4* status. Our data were in agreement with recent reports that failed to confirm the association between *PCDH11X* polymorphisms and LOAD, and extended the lack of association to common *PCDH11Y* variants.

Introduction

Alzheimer's disease (AD) is the most common cause of dementia in the elderly. Mutations in the genes encoding the amyloid precursor protein (*APP*), presenilin 1 (*PSEN1*), and presenilin 2 (*PSEN2*) have been associated with rare forms of early-onset familial autosomal dominant AD (1-2% of all cases of AD) (Brouwers et al., 2008). However, most of the patients have a late-onset (>65 years) disease (LOAD). Both, acquired/environmental and genetic/inherited risk factors contribute to the risk of developing LOAD. To date, the $\epsilon 4$ allele at the apolipoprotein E gene (*APOE-ε4*) is the only gene variant that has been recognized as a risk factor for LOAD in all populations (Chartier-Harlin et al., 1994; Saunders et al., 1993) (see the AlzGene database for a complete view of the gene variants that have been associated with LOAD; www.alzgene.org).

A recent genome-wide association (GWA) study reported a significant association between a protocadherin-11 X (*PCDH11X*) single nucleotide polymorphism (SNP) and LOAD (Carrasquillo et al., 2009). Although with a lower significance, a different GWA study also found some evidence of an effect of the *PCDH11X* on LOAD-risk (Lambert et al., 2009). There are two protocadherin-11 loci located in the homologue block of chromosomes X and Y: *PCDH11X* in Xq21.3 and *PCDH11Y* in Yp11.2. They belong to the cadherin superfamily and encode cell adhesion proteins that are expressed in fetal and adult brain (Blanco et al., 2000; Yoshida and Sugano, 1999). Different transcripts from both genes have been found in different brain regions, and it has been proposed that they contribute to the establishment and maintenance of specific neuronal connections (Blanco-Arias et al., 2004; Wu and Maniatis, 1999). Thus, the

association between *PCDH11X/Y* SNPs and LOAD was plausible considering the putative role of this loci in the brain.

Although the initial association between *PCDH11X* and LOAD was based on a large cohort of patients and controls, three recent reports failed to replicate this effect on LOAD-risk (Beecham et al., 2010; Lescai et al., 2010; Wu et al., 2010) (see also the AlzGene database; www.Alzgene.org). Considering these discrepancies, additional studies from different populations are necessary, and for this reason we investigated the association between LOAD and four *PCDH11X* and one *PCDH11Y* SNPs in a population from Spain.

Results

In **table 1** we summarised the main characteristics of patients and controls. *APOE-ε4* allele was significantly associated with LOAD in our population ($p < 0.0001$; OR=7.77; 95% CI=5.44-11.09). The four *PCDH11X* and the *PCDH11Y* SNPs were successfully genotyped in all the patients and controls through PCR-RFLP (**Supplementary table 1; supplementary figure 1**). Direct sequencing of several individuals representative of each genotype confirmed the accuracy of this genotyping method. In the *PCDH11X* fragment containing SNP rs4252205 (c.3035G>A; p.Arg1018Gln) we also found an intronic variant (c.3114+4T>G; SNP rs4252206) that was in complete LD with rs4252205.

In **tables 2 and 3** we summarized the genotype and allele frequencies of the *PCDH11X/Y* SNPs in the female and male LOAD and controls. Genotype frequencies did not deviate from the Hardy-Weinberg equilibrium for the four *PCDH11X* SNPs in the female groups. We did not find significant difference between patients and controls for the five *PCDH11X/Y* SNPs. We did not find significantly different onset-ages between the genotypes/alleles. *PCDH11X/Y* frequencies did not differ between *APOE-ε4* carriers and non-carriers, in patients and controls (data not shown). Logistic regression showed that none of the five SNPs was a risk factor for LOAD after including gender, age, and *APOE-ε4* status as covariates.

We analyzed the linkage disequilibrium (LD) values and haplotype frequencies for the *PCDH11X* SNPs in male and female patients and controls. We found an almost complete LD between rs5941047 and rs4568761 ($D' = 0.94$ in female, $D' = 0.98$ in male). SNP rs5984894 was also in LD with the haplotype block rs5941047/rs4568761. We did not find significantly different frequencies between haplotype frequencies in female and male patients and controls (**table 4**).

Discussion

We reported a lack of effect on LOAD-risk of the *PCDH11X* SNP rs5984894. This variant was associated to LOAD in a recent GWA (Carrasquillo et al., 2009). The association between LOAD and *PCDH11X* variants was plausible considering the expression of protocadherin 11 isoforms in several brain regions (Blanco-Arias et al., 2004). Moreover, these proteins could interact with PP1 α , a phosphatase that could promote the secretion of the amyloid precursor protein (da Cruz e Silva et al., 1995; Vintem et al., 2009). In addition, protocadherin 11 has been implicated in the regionalization and functional differentiation of brain grey matter and in regulating synaptic plasticity (Williams et al., 2006).

The *PCDH11X* rs5984894 was an intronic SNP that could have a direct effect on alternative transcription, and this could explain the association between this variant and LOAD and other neurodegenerative and psychiatric disorders. Alternatively, this SNP could be in linkage disequilibrium with other gene variants that would be the responsible for the genetic association. To investigate this, we determined the association between three additional *PCDH11X* SNPs and LOAD. None of the three *PCDH11X* variants (included a DNA variant that resulted in a missense amino acid change) were associated to LOAD in our population, even after correcting by age, gender, and *APOE- ϵ 4* status. In addition, we did not find significant different haplotype frequencies between patients and controls. Our results were in agreement with recent reports that also failed to confirm the association between SNP rs5984894 and LOAD (Beecham et al., 2010; Lescai et al., 2010; Wu et al., 2010) (**supplementary table 2**).

Taking into account our allelic frequencies for the rs5984894 A allele in women patients and controls (0.47 and 0.49, respectively) a total of 9,794 individuals should be necessary to reach a power of 80 at a $p=0.05$ (at a 1:1 patients/controls ratio). However, the difference in allele frequencies was based on a higher frequency of heterozygotes among the patients, while the frequency of AA homocygotes was the same in the two groups. By contrast, in the Carrasquillo et al. study the AA frequency was higher in LOAD women than in controls.

We also investigated the association between LOAD and a common *PCDH11Y* SNP (rs2524543). This was in complete LD with SNP rs2563389, and the two DNA polymorphisms resulted in missense amino acid changes that were conserved between species. These changes could result in functionally relevant differences between protocadherin 11 isoforms, leading to differences in susceptibility to brain diseases. However, we did not find different allele frequencies between male LOAD and controls. Two previous reports also failed to confirm an effect of these *PCDH11Y* variants and schizophrenia or other psychiatric disorders (Durand et al., 2006; Giouzeli et al., 2004).

In conclusion, we showed a lack of association between four common *PCDH11X* SNPs and LOAD-risk. In addition, a *PCDH11Y* SNP was not associated with LOAD-risk in male.

Experimental procedure

Subjects

The study included a total of 420 LOAD patients (273 women, 65%) who fulfilled the criteria for clinical probable Alzheimer's disease, according to the NINCDS-ADRDA (National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorders Association). All the patients were recruited by Neurologists from three reference Hospitals in the region of Asturias (Northern Spain, total population one million). The control group consisted of 350 healthy controls aged 60-85 years (193 women, 55%). They were individuals from the general population or healthy spouses of the patients, and although all them were apparently healthy we did not perform neuropsychological examination to exclude the existence of symptoms of cognitive impairment or any other neurodegenerative disorder in any of them. All the patients and controls were Spanish Caucasians and from the region, and gave their informed consent to participate in the study. This research was approved by the Ethical Committee of Hospital Central Asturias.

Genotyping

Three *PCDH11X* SNPs in intron 2 (SNPs rs5984894, rs5941047, and rs4568761) and a missense SNP in exon 3 (SNP rs4252205, c.3053G>A; p.Arg1018Gln; transcript id ENST00000373094, www.ensembl.org) were genotyped through restriction enzyme digestion of a polymerase chain reaction fragment (PCR-RFLP; see **supplementary table 1** for primer sequences and the PCR and RFLP conditions). To confirm the accuracy of the genotyping method, we sequenced several individuals representative of each genotype.

SNP rs2524543 was a *PCDH11Y* missense change (c.2684G>T; p.Val896Phe; transcript id ENST00000400457). Because this was in a sequence present in both *PCDH11X* and *PCDH11Y*, the PCR amplified a 433 bp fragment from the two genes. The two sequences contained an internal site for the restriction enzyme *TasI*, resulting in digestion fragments of 251 + 182 bp. The 182 bp fragment from *PCDH11Y* was digested when c.2684T was present, and resulted in two fragments of 99 + 83 bp.

Statistical analysis

The χ^2 test was used to compare allele and genotype frequencies between patients and controls, and to determine the deviation from the Hardy-Weinberg equilibrium in female patients and controls. To compare mean ages between the groups we used the Student's t-test. The logistic binary regression was used to determine the effect of the allele and genotypes on LOAD-risk including the gender, age, and *APOE-ε4* status as covariates. Power was calculated online (<http://statpages.org/proppowr.html>). Linkage disequilibrium values and haplotype frequencies were determined with Haploview 4.2 (www.broadinstitute.org/haploview). We considered a $p < 0.05$ as the level for significant differences in all the tests.

Contributions

V.A. and E.C. designated the work, analyzed the results, and wrote the manuscript. A.M., A.I.C., B.A., and M.D. performed all the experiments. G.M., M.C., C.M., and M.M. recruited the patients and obtained all the clinical and anthropometric data.

Acknowledgements

This work was supported by grants from the Spanish Fondo de Investigaciones Sanitarias-Fondos FEDER European Union (FIS-05/008 and 08/0915). A.I.C. and M.D. are the recipients of a Contrato de Apoyo a la Investigación-Fondo de Investigaciones Sanitarias.

Authors wish to thank Asociación Parkinson Asturias and Obra Social Cajastur for their support.

References

- Beecham, G.W., Naj, A.C., Gilbert, J.R., Haines, J.L., Buxbaum, J.D. Pericak-Vance, M.A., 2010. PCDH11X variation is not associated with late-onset Alzheimer disease susceptibility. *Psychiatr Genet* 20, 321-324.
- Blanco-Arias, P., Sargent, C.A. Affara, N.A., 2004. Protocadherin X (PCDHX) and Y (PCDHY) genes; multiple mRNA isoforms encoding variant signal peptides and cytoplasmic domains. *Mamm Genome* 15, 41-52.
- Blanco, P., Sargent, C.A., Boucher, C.A., Mitchell, M. Affara, N.A., 2000. Conservation of PCDHX in mammals; expression of human X/Y genes predominantly in brain. *Mamm Genome* 11, 906-914.

- Brouwers, N., Sleegers, K., Van Broeckhoven, C., 2008. Molecular genetics of Alzheimer's disease: an update. *Ann Med* 40, 562-583.
- Carrasquillo, M.M., Zou, F., Pankratz, V.S., Wilcox, S.L., Ma, L., Walker, L.P., Younkin, S.G., Younkin, C.S., Younkin, L.H., Bisceglia, G.D., Ertekin-Taner, N., Crook, J.E., Dickson, D.W., Petersen, R.C., Graff-Radford, N.R., 2009. Genetic variation in PCDH11X is associated with susceptibility to late-onset Alzheimer's disease. *Nat Genet* 41, 192-198.
- Chartier-Harlin, M.C., Parfitt, M., Legrain, S., Perez-Tur, J., Brousseau, T., Evans, A., Berr, C., Vidal, O., Roques, P., Gourlet, V. et al., 1994. Apolipoprotein E, epsilon 4 allele as a major risk factor for sporadic early and late-onset forms of Alzheimer's disease: analysis of the 19q13.2 chromosomal region. *Hum Mol Genet* 3, 569-574.
- da Cruz e Silva, E.F., da Cruz e Silva, O.A., Zaia, C.T., Greengard, P., 1995. Inhibition of protein phosphatase 1 stimulates secretion of Alzheimer amyloid precursor protein. *Mol Med* 1, 535-541.
- Durand, C.M., Kappeler, C., Betancur, C., Delorme, R., Quach, H., Goubran-Botros, H., Melke, J., Nygren, G., Chabane, N., Bellivier, F., Szoke, A., Schurhoff, F., Rastam, M., Anckarsater, H., Gillberg, C., Leboyer, M., Bourgeron, T., 2006. Expression and genetic variability of PCDH11Y, a gene specific to Homo sapiens and candidate for susceptibility to psychiatric disorders. *Am J Med Genet B Neuropsychiatr Genet* 141B, 67-70.
- Giouzele, M., Williams, N.A., Lonie, L.J., DeLisi, L.E., Crow, T.J., 2004. ProtocadherinX/Y, a candidate gene-pair for schizophrenia and schizoaffective disorder: a DHPLC investigation of genomic sequence. *Am J Med Genet B Neuropsychiatr Genet* 129B, 1-9.
- Lambert, J.C., Heath, S., Even, G., Campion, D., Sleegers, K., Hiltunen, M., Combarros, O., Zelenika, D., Bullido, M.J., Tavernier, B., Letenneur, L., Bettens, K., Berr, C., Pasquier, F., Fievet, N., Barberger-Gateau, P., Engelborghs, S., De Deyn, P., Mateo, I., Franck, A., Helisalmi, S., Porcellini, E., Hanon, O., de Pancorbo, M.M., Lendon, C., Dufouil, C., Jaillard, C., Leveillard, T., Alvarez, V., Bosco, P., Mancuso, M., Panza, F., Nacmias, B., Bossu, P., Piccardi, P., Annoni, G., Seripa, D., Galimberti, D., Hannequin, D., Licastro, F., Soininen, H., Ritchie, K., Blanche, H., Dartigues, J.F., Tzourio, C., Gut, I., Van Broeckhoven, C., Alperovitch, A., Lathrop, M., Amouyel, P., 2009. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. *Nat Genet* 41, 1094-1099.
- Lescai, F., Pirazzini, C., D'Agostino, G., Santoro, A., Ghidoni, R., Benussi, L., Galimberti, D., Federica, E., Marchegiani, F., Cardelli, M., Olivieri, F., Nacmias, B., Sorbi, S., Bagnoli, S., Tagliavini, F., Albani, D., Boneschi, F.M., Binetti, G., Forloni, G., Quadri, P., Scarpini, E., Franceschi, C., 2010. Failure to replicate an association of rs5984894 SNP in the PCDH11X gene in a collection of 1,222 Alzheimer's disease affected patients. *J Alzheimers Dis* 21, 385-388.
- Saunders, A.M., Strittmatter, W.J., Schmechel, D., George-Hyslop, P.H., Pericak-Vance, M.A., Joo, S.H., Rosi, B.L., Gusella, J.F., Crapper-MacLachlan, D.R., Alberts, M.J. et al., 1993. Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. *Neurology* 43, 1467-1472.
- Vintem, A.P., Henriques, A.G., da Cruz, E.S.O.A., da Cruz, E.S.E.F., 2009. PP1 inhibition by Aβ peptide as a potential pathological mechanism in Alzheimer's disease. *Neurotoxicol Teratol* 31, 85-88.

- Williams, N.A., Close, J.P., Giouze, M., Crow, T.J., 2006. Accelerated evolution of Protocadherin11X/Y: a candidate gene-pair for cerebral asymmetry and language. *Am J Med Genet B Neuropsychiatr Genet* 141B, 623-633.
- Wu, Q., Maniatis, T., 1999. A striking organization of a large family of human neural cadherin-like cell adhesion genes. *Cell* 97, 779-790.
- Wu, Z.C., Yu, J.T., Wang, N.D., Yu, N.N., Zhang, Q., Chen, W., Zhang, W., Zhu, Q.X., Tan, L., 2010. Lack of association between PCDH11X genetic variation and late-onset Alzheimer's disease in a Han Chinese population. *Brain Res* 1357, 152-156.
- Yoshida, K., Sugano, S., 1999. Identification of a novel protocadherin gene (PCDH11) on the human XY homology region in Xq21.3. *Genomics* 62, 540-543.

Table 1. Main characteristics of the LOAD patients and the healthy controls.

	LOAD	Controls
Total	420	350
Women	273 (65%)	193 (55%)
Mean age ($\pm$SD)	73.79 ($\pm$ 13.4)	72.85($\pm$ 6.61)
APOE ϵ4 –	237 (56%)	300 (76%)
APOE ϵ4 +	183 (44%)	50 (14%)
ϵ 34	156 (37%)	46 (13%)
ϵ 44	19 (5%)	2 (<1%)
ϵ 24	8 (2%)	2 (<1%)

Table 2. Genotype and allele frequencies of the four *PCDH11X* SNPs in the female LOAD (n=273) and controls (n=193).

		Genotypes			Alleles		
		CC	CT	TT	C	T	p-value OR (95%CI)
Rs5941047							
	Controls	31 (16%)	85 (44%)	77 (40%)	38%	62%	p=0.82
	LOAD	38 (14%)	126(46%)	109(40%)	37%	63%	OR=0.95 (0.65-1.40)
Rs4568761							
	Controls	33 (17%)	88 (46%)	72 (37%)	39%	61%	p=0.82
	LOAD	44 (16%)	131(48%)	98 (36%)	40%	60%	OR=1.04 (0.71-1.52)
Rs5984894							
	Controls	44 (23%)	95 (49%)	54 (28%)	47%	53%	p=0.67
	LOAD	63 (23%)	139(51%)	71 (26%)	49%	51%	OR=1.08 (0.75-1.56)
Rs4252205							
	Controls	187 (97%)	6 (3%)	0 (0%)	98%	2%	p=0.39
	LOAD	270(99%)	3 (1%)	0 (0%)	99.5%	<1%	OR=4.06 (0.57-28.85)

Table 3. Frequencies of the five SNPs in *PCDH11X/Y* genes in the male LOAD (n=147) and controls (n=157).

		Allelic frequencies			
Rs5941047			C	T	p-value OR (95%CI)
	Controls	N=157	60 (38%)	97 (62%)	P=0.61
	LOAD	N=147	52 (35%)	95 (65%)	OR=0.88 (0.55-1.41)
Rs4568761			C	T	
	Controls		60 (38%)	97 (62%)	P=0.61
	LOAD		52 (35%)	95 (65%)	OR=0.88 (0.55-1.41)
Rs5984894			G	A	
	Controls		74 (47%)	83 (53%)	P=0.84
	LOAD		71 (48%)	76 (52%)	OR=1.05 (0.67-1.64)
Rs4252205			G	A	
	Controls		157 (100%)	0 (0%)	
	LOAD		147 (100%)	0 (0%)	

Rs2524543 (PCDH11Y)			G	T	p-value OR (95%CI)
	Controls		66 (42%)	91 (58%)	P=0.88
	LOAD		63 (43%)	84 (57%)	OR=1.03 (0.65-1.63)

Table 4. Haplotype frequencies in LOAD patients and healthy controls. The haplotypes were defined by the three SNPS in intron 2 of *PCDH11X* (rs5984894-rs5941047-rs4568761).

Haplotype		LOAD frequencies	Controls frequencies	p-value
Female	GTT	41.7%	42.7%	0.75
	ACC	29.5%	31.8%	0.46
	ATT	16.9%	13.4%	0.14
	GCC	6.2%	6.2%	0.99
	ATC	2.5%	1.7%	0.41
	GTC	1.5%	2%	0.57
	GCT	1.1%	1.2%	0.89

Haplotype		LOAD frequencies	Controls frequencies	p-value
Male	GTT	47%	43.8%	0.46
	ACC	31.9%	37.3%	0.19
	ATT	17.9%	15.4%	0.43
	GCC	1.6%	1.9%	0.73

Supplementary material.

Lack of association between protocadherin 11-X/Y (*PCDH11X* and *PCDH11Y*) polymorphisms and Late Onset Alzheimer's Diseases, by Ana Miar et al.

Supplementary table 1. Primer sequences (forward/reverse), annealing temperatures, size of the amplified fragment (bp), and PCR-RFLP conditions for the five *PCDH11X/Y* SNPs.

SNP	Primer Sequences	Annealing	Size	Restriction enzyme	Allele sizes
Rs5941047	5' AGCAACTGACATATGGTGCATCATT	58°C	467 bp	<i>HindIII</i>	C: 169+298 T: 169+96+202
	5' TGACAGTTTGAGTTACAGAACTAGAA				
Rs4568761	5' TGGCTTGTGGTTTTGCCCATTC	58°C	419 bp	<i>AluI</i>	C: 36+383 T: 36+244+139
	5' TGGCAGTGTCTACATTTTCACCA				
Rs5984894	5' CTCAAGTATGCATTAATATGTGCAT	51°C	233 bp	<i>TaqI</i>	A: 233 G: 202+31
	5'CTAAATATTTGGATCTCAATATAATGCTT				
Rs2524543	5' TTGTCAATCTGTTTCGTGAATG	54°C	433bp	<i>TasI</i>	T: 251+99+83 G: 251+182
	5' AGATCAATAGGAAGGTCTAGTGTG				
Rs4252205	5' ACTATTTTAGCTACCTATAGGCAATG	55°C	398bp	<i>MboI</i>	G: 242+37+119 A: 242+156
	5' ATCCAACAAGTTAATTTCACTGTTA				

Supplementary table 2. Comparison of the *PCDH11X* minor allele frequencies in female LOAD and controls in our study and five published studies.

(nd: not determined in the study)

	Miar et al.	Beecham et al	Lescai et al	Carrasquillo et al	Wu et al	Lambert et al
rs5941047	C allele					
controls	0.38	0.38	nd	0.39	nd	0.38
LOAD	0.37	0.41	nd	0.44	nd	0.38
rs4568761	C allele					
controls	0.39	0.40	nd	0.42	nd	0.40
LOAD	0.40	0.43	nd	0.46	nd	0.40
rs5984894	A allele					
controls	0.47	0.46	0.52	0.46	0.85	0.48
LOAD	0.49	0.47	0.51	0.52	0.85	0.48

Supplementary figure 1. PCR-RFLP electrophoretic pattern of the five SNPs. The *PCDH11X* rs5941047, rs4568761 and rs5984894 and *PCDH11Y* rs2524543 were from male individuals. The