Association of glyceraldehyde-3-phosphate dehydrogenase locus variant with late-onset Alzheimer's disease

Mariet Allen[1,2], Claire Cox[1], Olivia Belbin[1], Li Ma[1], Gina D. Bisceglio[1], Samantha L. Wilcox[1], Chanley C. Howell[1],Talisha A. Hunter[1], Oliver Culley[1], Louise P. Walker[1], Minerva M. Carrasquillo[1], Dennis W. Dickson[1], Ronald C. Petersen, M.D[3], Neill R. Graff-Radford[1,3], · Steven G. Younkin[1], Nilüfer Ertekin-Taner[1,3].

1Department of Neuroscience, Mayo Clinic College of Medicine, Jacksonville, FL, USA, 2Human Genetics Unit, Institute of Genetics and Molecular Medicine, Medical Research Council, Edinburgh, United Kingdom, 3Department of Neurology, Mayo Clinic College of Medicine, Rochester, MN, USA.

Background: Previous studies have implicated variants in glyceraldehyde-3-phosphate dehydrogenase gene (*GAPDH*) and its paralogues in late-onset Alzheimer's disease (LOAD), although the strength and direction of association have not been consistent. Our objective is to explore the associations between *GAPDH*, its paralogues and LOAD.
Methods: We genotyped three previously reported SNPs (rs3741916 in *GAPDH*, rs2029721 in *pGAPD* and rs4806173 in *GAPDHS*) and 22 additional SNPs, at the *GAPDH* and *GAPDHS* loci, in three case-control series collectively composed of 2112 cases and 3808 controls. We tested these SNPs for association with LOAD using multivariable logistic regression analysis adjusting for age, gender and ApoE. We performed meta-analysis of the most significant rs3741916 SNP using previously published series and ours.
Results: *GAPDH* variant rs3741916, which resides in the 5'UTR of this gene, showed the strongest evidence of association with LOAD (p value = 0.003). None of the other SNPs were as significant. The minor G allele showed a protective effect in our combined series (OR = 0.87, 95% confidence interval (CI) = 0.79-0.96). This result is consistent with all of the published follow-up series but is in the opposite direction to three of the four series from the original report. Combined meta-analysis of all published series with available data and ours suggests presence of heterogeneity (Breslow-Day p < 0.0001). Meta-analysis of the 4 follow-up series with available data, including ours, revealed a significant protective effect for the minor G allele of rs3741916 (OR = 0.85, 95% CI = 0.76-0.96, p value = 0.0094).
Conclusions: Our results provide supportive evidence for the presence of LOAD risk variants in the vicinity of *GAPDH*. The most promising variant (rs3741916) is unlikely to be the functional variant given opposing effects in different series. To provide greater understanding of the AD association at this heterogeneous locus, we intend to genotype this SNP in additional case-control series to further increase our sample size. Identification of the functional variant(s) in this region likely awaits deep variant discovery efforts.

Summary

- In 2004 Li et al [1] reported association of SNPs at the *GAPDH* locus and it paralogues with LOAD.
- We focused on three of the SNPs reported by Li et al that formed multi-locus genotypes in their study (*GAPDH* rs3741916, *pGAPD* rs2029721 and *GAPDHS* rs4806173).
- Analysis of individual SNPs revealed significant association for the *GAPDH* SNP rs3741916 (aka rs1136666) in our combined Mayo series (Table1) , we were unable to replicate the multi-locus association in our series (data not shown).
- Two follow-up studies were published in 2006 (Lin et al[2]) and 2008 (Lee at al[3]) that included one or more of these SNPs. Each of the three studies reported association of these SNPs using different statistical tests.
- In Table 2, we summarize the results of these publications, for the three SNPs our study has focused on, and report the results in our series based upon the statistical methods used by others [1,2,3].
- Analysis of additional SNPs (Figure 1) at the *GAPDH* locus failed to identify SNPs with more significant association than rs3741916 (data not shown).
- Meta-analysis of all published series with available data[1,3], plus our Mayo series, reveals heterogeneity at the *GAPDH* locus (Fig 2a).
- Meta analysis excluding the initial study (Li et al) ,reveals significant association of the minor allele of rs3741916 with decreased risk for LOAD.

SNP (Locus)	Chr	Base Position	Strata	Series	Cs N (MAF)	Cn N(MAF)	OR (95%CI)	p-value
rs3741916 (*GAPDH*) aka rs1136666	12	6,514,252	**All**	JS	859 (0.26)	974 (0.28)	0.90 (0.77: 1.05)	0.159
				RS	611 (0.27)	2411 (0.28)	0.93 (0.81: 1.08)	0.361
				AUT	585 (0.23)	354 (0.27)	**0.78 (0.61: 1.00)**	**0.047**
				JS/RS/AUT	2055 (0.25)	3739 (0.28)	**0.87 (0.79: 0.96)**	**0.003**
rs2029721 (*pGAPD*)	12	61,435,611	**AAE/D ≤78**	JS	590 (0.40)	637 (0.36)	**1.39 (1.07: 1.81)**	**0.013**
				RS	556 (0.37)	1356 (0.38)	0.94 (0.74: 1.21)	0.638
				AUT	574 (0.37)	354 (0.36)	1.20 (0.90: 1.60)	0.225
				JS/RS/AUT	1720 (0.38)	2347 (0.37)	1.14 (0.98: 1.31)	0.090
rs4806173 (*GAPDHS*)	19	40,716,765	**All**	JS	591 (0.39)	642 (0.38)	1.03 (0.87: 1.23)	0.732
				RS	556 (0.39)	1372 (0.38)	0.99 (0.86: 1.15)	0.923
				AUT	586 (0.38)	358 (0.38)	0.87 (0.70: 1.09)	0.222
				JS/RS/AUT	1733 (0.38)	2372 (0.38)	0.98 (0.89: 1.08)	0.688

Table 1. The results of logistic regression analysis under an additive model in the Mayo Clinic Series. Chr = Chromosome Cs = AD case, Cn = Control subject. N = number of subjects in series that have genotype data. MAF = minor allele frequency. OR = Odds Ratio, CI = Confidence Interval. . Logistic regression includes Age*, Gender and presence of an ApoE4 allele included as covariates, *Age and AAE/D refers to Age at Diagnosis, Examination or Death. Nominally significant p-values (<0.05) are highlighted in bold.

SNP (Locus)	Chr	Position (bp)	Strata	Series	Cs N (MAF)	Cn N (MAF)	Allelic OR (95%CI)	Allelic p-value	Logistic Regression (Dominant Model) OR (95%CI)	Logistic Regression (Dominant Model) p-value
A rs3741916 (*GAPDH*) aka rs1136666	12	6,514,252	ApoE4· (Li *et al*, 2004)	**Li *et al*: W/UC/UK**	(0.297)	(0.247)	**1.27(1.06:1.53)**	**0.008**	nr	nr
				JS	311 (0.25)	702 (0.27)	0.92 (0.74: 1.15)	0.478	0.95 (0.75: 1.28)	0.874
				RS	276 (0.28)	1838 (0.28)	1.00 (0.82: 1.23)	0.960	1.12 (0.86: 1.45)	0.406
				AUT	274 (0.22)	232 (0.28)	**0.73 (0.54: 0.99)**	**0.042**	**0.67 (0.46: 0.97)**	**0.034**
				JS/RS/AUT	819 (0.25)	2814 (0.28)	0.88 (0.78: 1.00)	0.051	0.92 (0.79: 1.08)	0.313
B rs3741916 (*GAPDH*) aka rs1136666	12	6,514,252	AAE/D* < mean (Lin *et al*, 2006)	**Lin *et al* C-C series**	nr	nr	nr	nr	**0.39 (0.21:0.70)**	**0.002**
				JS	417 (0.27)	563 (0.28)	0.95 (0.78: 1.17)	0.683	0.98 (0.74: 1.29)	0.870
				RS	229 (0.28)	1200 (0.29)	0.93 (0.74: 1.16)	0.537	0.95 (0.70: 1.30)	0.765
				AUT	209 (0.24)	225 (0.27)	0.83 (0.60: 1.13)	0.243	0.85 (0.56: 1.28)	0.435
				JS/RS/AUT	855 (0.26)	1988 (0.29)	0.89 (0.79: 1.02)	0.089	0.92 (0.77: 1.10)	0.349
C rs3741916 (*GAPDH*) aka rs1136666	12	6,514,252	All (Lee *et al*, 2008)	**Lee *et al* C-C series**	(0.14)	(0.21)	nr	**0.027**	nr	0.054
				JS	859 (0.26)	974 (0.28)	0.90 (0.78: 1.05)	0.178	0.92 (0.75: 1.12)	0.399
				RS	611 (0.27)	2411 (0.28)	0.91 (0.79: 1.05)	0.212	0.95 (0.78: 1.14)	0.554
				AUT	585 (0.23)	354 (0.27)	0.81 (0.65: 1.00)	0.053	0.77 (0.57: 1.04)	0.086
				JS/RS/AUT	2055 (0.25)	3739 (0.28)	**0.86 (0.79: 0.94)**	**8.0E-04**	**0.88 (0.78: 0.99)**	**0.033**
D rs2029721 (*pGAPD*)	12	61,435,611	AAE/D* > mean (Li *et al*, 2004 & Lin *et al* 2006)	**Li *et al*: W/UC/UK**	(0.308)	(0.352)	0.80 (0.68:0.97)	**0.018**	nr	nr
				Lin *et al*: C-C series	nr	nr	nr	**0.004**	0.55 (0.32:0.96)	**0.036**
				JS	320 (0.39)	309 (0.36)	1.13 (0.89: 1.42)	0.322	1.16 (0.82: 1.63)	0.404
				RS	336 (0.36)	719 (0.38)	0.94 (0.78: 1.14)	0.530	0.96 (0.73: 1.27)	0.766
				AUT	369 (0.36)	130 (0.34)	0.95 (0.70: 1.31)	0.759	1.31 (0.84: 2.05)	0.237
				JS/RS/AUT	1025 (0.37)	1158 (0.37)	1.00 (0.88:1.13	0.999	1.05 (0.87: 1.26)	0.642
E rs4806173 (*GAPDHS*)	19	40,716,765	AAE/D* < mean (Li *et al*, 2004)	**Li *et al*: W/UC/UK**	(0.326)	(0.42)	**0.66 (0.55:0.80)**	**3.0E-04**	nr	nr
				JS	270 (0.39)	330 (0.40)	0.97 (0.76: 1.23)	0.812	0.95 (0.65: 1.38)	0.785
				RS	220 (0.39)	644 (0.38)	1.05 (0.83: 1.32)	0.691	1.00 (0.70: 1.41)	0.977
				AUT	209 (0.38)	229 (0.36)	1.09 (0.82: 1.45)	0.575	1.16 (0.77: 1.75)	0.490
				JS/RS/AUT	699 (0.39)	1203 (0.38)	1.03 (0.89: 1.18)	0.729	1.00 (0.81: 1.23)	0.960

Table 2. Replication analysis of previous reports: Chr = Chromosome, Cs = AD case, Cn = Control subject. N = number of subjects in series that have genotype data. MAF = minor allele frequency. OR = Odds Ratio, CI = Confidence Interval. Allelic association tested using chi-squared test with no covariates. Logistic regression uses Age*, Gender and presence of an APOE4 allele as covariates, *Age and AAE/D refers to Age at Diagnosis, Examination or Death. The mean age differs between the different publications. The mean age is 78 in our combined series. Nominally significant p-values (<0.05) are highlighted in bold, nr = Not reported.

References

1. Li Y, Nowotny P, Holmans P, et al. Association of late-onset Alzheimer's disease with genetic variation in multiple members of the GAPD gene family. Proc Natl Acad Sci U S A 2004;101:15688-15693.
2. Lin PI, Martin ER, Bronson PG, et al. Exploring the association of glyceraldehyde-3-phosphate dehydrogenase gene and Alzheimer disease. Neurology 2006;67:64-68.
3. Lee JH, Cheng R, Rogaeva E, et al. Further examination of the candidate genes in chromosome 12p13 locus for late-onset Alzheimer disease. Neurogenetics 2008;9:127-138.

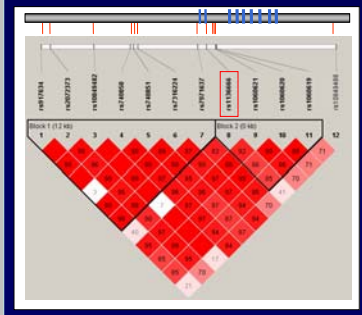

Figure 1. Linkage disequilibrium in the combined Mayo Clinic series at the *GAPDH* locus. LD was estimated and haplotype blocks were defined using the "Solid Spine" method implemented in HAPLOVIEW. Darker shades of red indicate increasing strength of LD (D'). Exons are represented with blue boxes and SNPs are represented with red lines. *GAPDH* SNP rs3741916 (aka rs1136666) is highlighted by a red box.

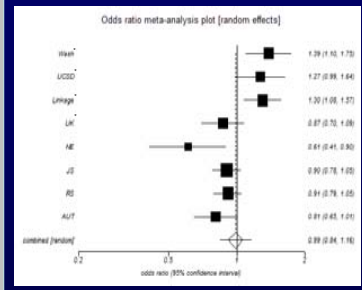

Figure 2a. *GAPDH* rs3741916 (aka rs1136666) Meta Analysis of all series with reported counts and frequencies. Breslow-Day p-value <0.0001.
Lin et al data indicates significant association of the minor allele of this SNP with decreased risk for LOAD (OR=0.39). This data is not included in meta-analysis due to unreported counts or frequencies.

* indicates series reported in the original study (Li et al).

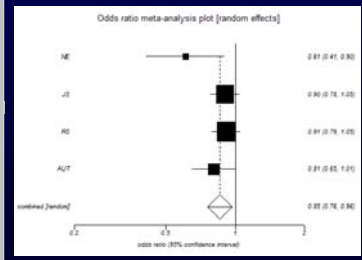

Figure 2b. Meta Analysis of all follow-up series with known counts and frequencies. Breslow Day p-value = 0.1967, Combined series p-value = 0.0094.