

Cluster Analysis of Risk Factor Genetic Polymorphisms in Alzheimer's Disease

C. N. Randall · D. Strasburger · J. Prozonc · S. N. Morris · A. D. Winkie ·
G. R. Parker · D. Cheng · E. M. Fennell · I. Lanham · N. Vakil ·
J. Huang · H. Cathcart · R. Huang · S. E. Poduslo

Accepted: 4 February 2008 / Published online: 29 February 2008
© Springer Science+Business Media, LLC 2008

Abstract Multiple genetic variants may contribute to the risk of developing Alzheimer's disease. We have analyzed polymorphisms in 9 genes to determine whether particular combinations would contribute to this risk. The genes were APOE, LDLr, CST3, CTSD, TNF, BACE1, MAPT, STH, eNOS, and TFCP2. Three risk groups for the disease were identified. Risk group I was younger, was heterozygous for the CST3 (GA), CTSD2936 (AG), TNF -308 (AG) genetic variants. Risk group II was older, was homozygous for the -427 APOE promoter polymorphism (TT), and heterozygous for the MAPT deletion and for the STH variant (QR). Group III had both the youngest and oldest subjects, were heterozygous for the -863 (AC) and -1031 (CT) TNF promoter polymorphisms. All three groups carried the

APOE 4 allele and were heterozygous for both BACE1 polymorphisms. The control groups were carriers of the APOE 3 allele and were homozygous for the BACE1 genetic variants.

Keywords Alzheimer's disease · Genetic variants · Cluster analysis · Risk factors

Introduction

Alzheimer's disease is characterized by a progressive decline in cognition and distinct neuropathology. The disease is complex. Many genes may interact to cause the disease or act as risk factors. The most prevalent recognized risk factor for the disease is the E4 allele of the APOE gene which may account for 40–50% of the risk for late-onset Alzheimer's disease. Many genetic polymorphisms have been implicated as risk factors. We had previously analyzed polymorphisms in APOE, APOC1, low density lipoprotein receptor (LDLr on 19p13.3, binds APOE), cystatin SN (CST3 on 20p11.2; a cathepsin proteinase inhibitor, colocalizing with the A β peptide in plaques), and cathepsin D (CTSD on 11p15.5, lysosomal protease found in neuritic plaques) by cluster analysis [1]. To this list we have added beta-site APP-cleaving enzyme 1 (BACE1 on 11q23, β -secretase), tumor necrosis factor (TNF on 6p21.3, inflammatory cytokine), microtubule-associated protein tau (MAPT on 17q21.1, maintains neuronal morphology), saitoihin (STH 17q21.1, a gene nested within the MAPT gene), nitric oxide synthase 3 (NOS3 on 7q36, potential mediator of inflammation and vascular activity), and transcription factor CP2 (TFCP2 on 12q13, interaction with proteins involved with inflammation). All of these gene products are involved with the processing of APP, cellular response to injury, interaction with A β in

C. N. Randall, S. N. Morris, A. D. Winkie and G. R. Parker—STAR students.

C. N. Randall, D. Strasburger, J. Prozonc, S. N. Morris, A. D. Winkie, G. R. Parker, D. Cheng and E. M. Fennell contributed equally to this study.

Special issue article in honor of Dr. George DeVries.

C. N. Randall · D. Strasburger · J. Prozonc ·
S. N. Morris · A. D. Winkie · G. R. Parker · D. Cheng ·
E. M. Fennell · I. Lanham · N. Vakil · J. Huang · H. Cathcart ·
R. Huang · S. E. Poduslo (✉)
Institute of Molecular Medicine and Genetics (IMMAG),
Medical College of Georgia, 1120 15th Street,
Augusta, GA 30912, USA
e-mail: spoduslo@mail.mcg.edu

S. E. Poduslo
Department of Neurology, Medical College of Georgia, Augusta,
GA, USA

S. E. Poduslo
VA Medical Center Augusta, Augusta, GA, USA

plaques, interaction with microtubules in neurofilamentary tangles, neuroinflammation, or interaction with other risk factors for the disease. Each genetic variant has been inconsistently associated with the disease. Identification of risk sets of factors contributing to the disease will aid in determining those at risk for the disease.

Experimental Procedure

Subjects

There were 181 patients with Alzheimer's disease (128 female and 53 male) and 119 (62 female and 57 male) control subjects. The clinical diagnosis of probable Alzheimer's disease was made according to NINCDS-ADRDA criteria [2], following a review of the medical records, documenting a progressive decline in cognition and appropriate blood tests to rule out other medical conditions, including thyroid and vitamin B12 deficiencies. We also included CT or MRI imaging of the brains of patients which showed cortical atrophy, but no evidence of strokes or tumors. The patients were of European descent. The spouses of the patients and siblings had a similar age, ethnic background, and exposure to environmental factors to control for unmeasured risk factors other than age and race. All subjects or authorized representatives for the patients gave written consent for the study, in accordance with the institutional review board guidelines.

Genotyping

DNA was isolated from blood samples with a proteinase K digestion and phenol/chloroform extraction. Genotypes were performed as described: APOE [3]; CST3 [4]; LDLr [5]; APOE promoter polymorphisms [6]; BACE1 [7]; TNF [8, 9]; CTSD [10]; MAPT [11]; STH [12]; NOS3 [13, 14]; TFCP2 [15] (Table 1).

Cluster Analysis

To investigate all the variables simultaneously without multiple comparisons, the latent classification statistical model, the Grade of Membership (GoM, developed at the Center for Demographic Studies at Duke University) was used [1, 16–18]. The data is represented by model-based groups, defined by frequencies of responses for the variables. Individuals are not assigned to a group, but are assigned a membership score for each group, based on the degree of similarity. Information on age and genetic variants were used as internal variables to define the pure types while sex and APOE were external variables. The data for the genetic variants listed, the reference age (either age-of-

onset or age at enrollment), gender, and disease status were analyzed simultaneously. Missing or limited information and small sample sizes can be used without specifying a particular genetic model.

Results

Five groups were identified. Groups I, II, and III were the Alzheimer's patients; groups IV and V were the control subjects (Table 2).

Group I had the youngest age-of-onset (60–69 years) and were 85% female; 63% carried an APOE 4 allele. Some 88% of this group carried the APOE promoter polymorphism TC at position –427. They were heterozygous for the CST3 and CTSD-2936 genetic variants (GA and AG respectively). The question relevance factor (QRF or the relevance of a variable to a pure type, with a lower limit of zero) for these variants were 1.85 and 1.25, suggesting that they were influential in defining the group. A QRF of one is neutral. This group was heterozygous for the TNF promoter polymorphism at position –308 (AG). They were also heterozygous for the two BACE genetic variants (CG and CT).

Group II had their age-of-onset as 70–80 years; 64% carried the APOE 4 allele. This group was homozygous for the APOE promoter polymorphism at position –427 (TT). Group II was also heterozygous for the MAPT deletion and the STH genetic variant. The QRFs for these two genetic variants were 1.86 and 1.9 which defined the group.

Group III was the youngest and the oldest group with age-of-onset as <65 years and 70–85 years. There were more males (38%) in this group. Some 77% carried an APOE 4 allele. Most were heterozygous for two of the TNF promoter polymorphisms at –863 and –1031 (AC and CT, respectively) with a QRF for the –1031 polymorphism of 1.53. They were also heterozygous for the BACE1 rs7083 polymorphism (CT).

Group IV had the age at enrollment of 70–85 years and were not affected with the disease. They were 100% carriers of the APOE 3 allele. They were also homozygous for the two BACE1 polymorphisms (GG and TT).

Group V had the age of enrollment between 70–80 years and also were not affected with the disease. They were also 100% carriers of the APOE 3 allele. They were homozygous for the BACE1 genetic variant, rs7083 (CC) with a QRF of 1.7.

Discussion

No genetic model is specified with the GoM software; maximum likelihoods estimate the model parameters.

Table 1 Genetic variants

Gene	Chromosome	Forward primer	Reverse primer	Product	Res Enz	Fragments	References
BACE1	11q23	rs638405	5'-ATCTCCCTGACCGCTCT-3'	321 bp	<i>HphI</i>	C 321 bp G 177 and 144 bp	Ref. [7]
BACE1	11q23	rs7083	5-TAACACAGGTCCCAATTCTCTCA-3	157 bp	<i>RsaI</i>	C 157 bp T 132 bp	http://www.ncbi.nlm.nih.gov
eNOS	7q35	Glu298Asp	5'-AACCCCTCTGGCCCACT-3'	200 bp	<i>MboI</i>	Asp 101 and 99 bp	Ref. [14]
eNOS	7q35	27bp repeat	5'-CTATGGTAGTGCCTTGGCTGGAGG-3'	4–6 repeats		167, 194, 221 bp	Ref. [13]
TNF α	6p21.3	-238 (rs361525)	5'-ATCTGGAGGAAGCGGTAGTG-3'	152 bp	<i>MspI</i>	G 133 bp and A 152 bp	Ref. [8]
		-308 (rs1800629)	5'-ATCTGGAGGAAGCGGTAGTG-3'	222 bp	<i>NcoI</i>	A 222 bp and G 206 bp	Ref. [8]
		-857	5'-AGTATGGGGACCCCGTTAA-3'	274 bp	<i>HincII</i>	T 174 bp and C 154 bp	Ref. [9]
		-863	5'-GCTCTGAGGAATGGTTACAG-3'	120 bp	<i>BsiI</i>	A 106 bp and C 120 bp	Ref. [9]
		-1031	5'-CAAAGGAGAACTGAGAGGA-3'	209 bp	<i>FokI</i>	T 187 bp and C 209 bp	Ref. [9]
STH	17q21	tau between exons 9 and 10	5'-CCC TGT AAA CTC TGA CCA CAC-3	226 bp	<i>HinfI</i>	Q 171 and 55 bp	Ref. [12]
MAPT	17q21	deletion	5'-GGAAGACGTTCTACTGATCTG-3'	dimucleotide repeat		R 97, 74, and 55 bp	Refs. [11, 21]
TFCP2	12q13	G/A	5'-TGGAGTCCAGTGGCGTGATC	287 bp	<i>BspI286I</i>	A 287 and G 202 bp	Ref. [15]
LDLR8	19p13.2	Ala/Thr	5'-ATG TCG ACC AAG CCT CTT TCT CTC TCT TC -3'	193 bp	<i>SnaI</i>	G 193 and A 144 bp	Ref. [22]
LDLR13	19p13.2	T/C	5'-CAG CCT GGG CAA CAA AAG TGA AA -3'	422 bp	<i>AvaII</i>	T 422, C 256, and C 166 bp	Ref. [5]
CST3	20p11.2	Ala/Thr	5'-GCGGTCTCTCTATCTAGC	500 bp	<i>SSTII</i>	A 500, G 357, G 143 bp	Ref. [23]
CTSD	11p15.5	Ala224Val	5'-GTGACAGGCAGGAGTTTGGT	250 bp	<i>MwoI</i>	T 250, C 168, C 82 bp	Ref. [10]
CTSD	11p15.5	rs2292963	5'-CATCCCTGCAGTTTCAGAAAGG 3'	225 bp	43 bp SNP		http://www.ncbi.nlm.nih.gov
APOE-491	19q13.2	A/T	5'-CAA GGT CAC ACA GGC AAC- 3'	228 bp	<i>DraI</i>	T 228, A 209 bp	Refs. [6, 24]
APOE-427	19q13.2	C/T	5'-TGT TGG CCA GGC TGG TTT TAA- 3'	228 bp	<i>AluI</i>	C 228, T 144, T 84 bp	Refs. [6, 24]
			5'-CAA GGT CAC ACA GGC AAC- 3'				
			5'-TGT TGG CCA GGC TGG TTT TAA- 3'				
			5'-TGT TGG CCA GGC TGG TTT TAA- 3'				

Table 2 Alzheimer's disease risk groups

Attributes	I	II	III	IV	V	H
AD	100	100	100	0	0	0.69
Age						0.71
<65	0	19	43	0	0	
60–69	100	0	0	0	0	
70–80	0	81	27	56	100	
>80	0	0	30	44	0	
Female	85	71	62	46	66	0.04
Male	15	29	38	54	34	
APOE						0.26
E23	7	0	0	24	0	
E33	29	36	19	51	64	
E24	15	0	4	0	0	
E34	15	56	51	25	36	
E44	33	8	26	0	0	
APOE-491						0.47
AA	70	46	100	0	100	
AT	30	54	0	77	0	
TT	0	0	0	23	0	
APOE-427						0.31
TT	12	100	100	83	100	
TC	88	0	0	17	0	
CC	0	0	0	0	0	
LDLr8						0.2
GG	100	100	100	50	98	
AG	0	0	0	50	0	
AA	0	0	0	0	2	
LDLr13						0.35
TT	0	44	0	76	0	
TC	62	33	61	24	77	
CC	38	23	39	0	23	
CST3						0.58
GG	0	87	100	68	100	
GA	100	0	0	0	0	
AA	0	13	0	32	0	
CTSD-224						0.2
CC	100	43	100	100	69	
CT	0	57	0	0	31	
TT	0	0	0	0	0	
CTSD-2963						0.48
AA	0	0	0	15	0	
AG	100	0	0	0	49	
GG	0	100	100	85	51	
TNF-238						0.17
AA	0	0	0	0	0	
AG	0	0	45	0	0	
GG	100	100	56	100	100	
TNF-308						0.5
AA	0	3	0	0	15	

Table 2 continued

Attributes	I	II	III	IV	V	H
AG	100	0	0	0	0	
GG	0	97	100	100	85	
TNF-857						0.15
CC	55	88	100	71	92	
CT	45	12	0	29	0	
TT	0	0	0	0	8	
TNF-863						0.53
AA	0	0	19	0	0	
AC	0	0	81	0	0	
CC	100	100	0	100	100	
TNF-1031						0.57
CC	0	0	20	0	0	
CT	0	0	80	0	0	
TT	100	100	0	100	100	
BACE1 rs638405						0.61
CC	0	0	0	0	65	
CG	100	59	67	0	35	
GG	0	41	33	100	0	
BACE1 rs7083						0.98
CC	0	0	0	0	100	
CT	100	75	100	0	0	
TT	0	25	0	100	0	
MAPTdel						0.53
DD	0	5	6	11	0	
DF	0	95	0	0	0	
FF	100	0	94	89	100	
STH						0.56
QQ	100	0	94	92	100	
QR	0	100	0	0	0	
RR	0	0	6	8	0	
eNOS Glu298Asp						0.23
GG	46	52	63	0	60	
GA	34	48	37	74	40	
AA	20	0	0	26	0	
eNOS repeats						0.18
44repeat	0	0	13	0	0	
45repeat	0	39	0	47	17	
55repeat	100	61	87	53	83	
TFCP2						0.17
AA	0	0	0	0	4	
AG	46	0	0	0	26	
GG	54	100	100	100	70	

Moreover, multiple loci can be considered together to determine risk. It is interesting that the addition of additional genetic variants lowered the informativeness (H, the entropy score which provides the information content for the variable; values close to the lower limit of zero provide

no information) of APOE 4 from our previous study (1.17 previously to 0.26) [1]. This suggests that the risk of APOE 4 may be reduced when other genetic variants are present. All three groups were 63–77% APOE 4. Also of interest is that many of the loci did not exhibit any risk for the disease, indicating that in this population of patients, the genetic variants are not risk factors.

According to the *alzgene* website (www.alzforum.org/res/com/gen/alzgene/), of the genetic variants that we evaluated, only APOE, CST3, CTSD, and TNF (–1031) were significant by meta analysis.

Risk factors for Group I which was the youngest group were CST3:GA and CTSD:AG which is similar to our previous findings [1]. However, TNF(-308):AG and BACE1 (rs7083):CT were also relevant. Risk factors for Group II which was intermediate in age-of-onset had the most prevalent risk factors of the MAPT deletion:DF and the STH:QR. That for Group III was the promoter polymorphisms in TNF at –1031 (CT). Being homozygous for the BACE1 polymorphisms may be protective for the control subjects in Groups IV and V.

The BACE1 protein product initiates A β cleavage; increased levels can increase A β production [19]. The protein products for both the CTSD and the CST3 genes and the APOE gene are found in amyloid plaques [1]. Whether the APOE 4 promoter polymorphism interacted with the heterozygous genetic variants to increase the deposition into amyloid plaques in younger patients leading to glial activation and increased TNF production is of considerable interest.

In addition to neuritic plaques, neurofibrillary tangles are present neuropathologically in brains of Alzheimer's patients. The tangles are composed of paired helical filaments of hyperphosphorylated tau proteins. Previous studies have shown that the H1 haplotype of MAPT is in linkage disequilibrium with the Q allele of STH [20]. Our studies show that the heterozygous forms of both genes define Group II with the age-of-onset between 70 and 80 years; this group is 64% APOE 4. Thus this group may have a propensity to develop more neurofibrillary tangles.

The group with the oldest age-of-onset of the disease was Group III (70–85 years). The most significant genetic variant was the TNF promoter polymorphism (–1031 CT) that may enhance transcription, indicating that inflammation may be the major risk factor in this age group.

Being homozygous for the BACE1 genetic variants and carrying the APOE 3 allele may be the most protective against the disease phenotype.

Our results suggest that genetic variants involved with plaque deposition may be more relevant in the younger group. Neurofibrillary tangles may be more relevant for the intermediate in age, group II, while inflammation is more relevant for the oldest group.

This approach of using cluster analysis does not specify a genetic model. Maximum likelihoods are used to estimate model parameters. Risk sets can be identified even though the sample size is small as GoM uses the L_1 rather than the L_2 criterion (where the deviations are not squared). Moreover, multiple comparisons are avoided as the data are analyzed simultaneously. Thus the use of GoM cluster analysis is useful to study genetic interactions which may be further proven in the biological context.

Acknowledgments The authors thank the Texas and Georgia families for participating in the DNA Bank. DNA was also obtained from the National Cell Repository for Alzheimer's disease, supported by the cooperative agreement grant U24AG21886. The research was supported by MCG startup funds, the summer STAR program, and a VA Merit Award.

References

1. Corder EH, Huang R, Cathcart HM, Lanham IS, Parker GR, Cheng D, Smith S, Poduslo SE (2006) Membership in genetic groups predicts Alzheimer disease. *Rejuvenation Res* 9:89–93
2. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan E (1984) Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA work group under the auspices of the Department of Health and Human Services task force on Alzheimer's disease. *Neurology* 34:939–944
3. Poduslo SE, Neal M, Herring K, Shelly J (1998) The apolipoprotein C1A allele as a risk factor for Alzheimer's disease. *Neurochem Res* 23:361–367
4. Cathcart HM, Huang R, Lanham IS, Corder EH, Poduslo SE (2005) Cystatin C as a risk factor for Alzheimer disease. *Neurology* 64:755–757
5. Cheng D, Huang R, Lanham IS, Cathcart HM, Howard M, Corder EH, Poduslo SE (2005) Functional interaction between APOE 4 and LDL receptor isoforms in Alzheimer's disease. *J Med Genet* 42:129–131
6. Parker GR, Cathcart HM, Huang R, Lanham IS, Corder EH, Poduslo SE (2005) Apolipoprotein gene E4 allele promoter polymorphisms as risk factors for Alzheimer's disease. *Psychiatric Genet* 15:271–275
7. Nowotny P, Kwon JM, Chakraverty S, Nowotny V, Morris JC, Goate AM (2001) Association studies using novel polymorphisms in BACE 1 and BACE 2. *NeuroReport* 12:1799–1802
8. Day CP, Grove J, Daly AK, Stewart MW, Avery PJ, Walker M (1998) Tumour necrosis factor-alpha gene promoter polymorphism and decreased insulin resistance. *Diabetologia* 41:430–434
9. Ma SL, Tang NLS, Lam LCW, Chiu HFK (2004) Association between tumor necrosis factor- α promoter polymorphism and Alzheimer's disease. *Neurology* 62:307–309
10. Papassotiropoulos A, Bagli M, Kurz A, Kornhuber J, Forstl H, Maier W, Pauls J, Lautenschlager N, Heun R (2000) A genetic variation of cathepsin D is a risk factor for Alzheimer's disease. *Ann Neurol* 47:399–403
11. Baker M, Litvan I, Houlden H, Adamson J, Dickson D, Perez-Tur J, Hardy J, Lynch T, Bigio E, Hutton M (1999) Association of an extended haplotype in the tau gene with progressive supranuclear palsy. *Hum Mol Genet* 8:711–715
12. Conrad C, Vianna C, Freeman M, Davies P (2002) A polymorphic gene nested within an intron of the tau gene: Implications for Alzheimer's disease. *Proc Natl Acad Sci USA* 99:7751–7756

13. Alvarez R, Alvarez V, Lahoz CH, Martinez C, Pena J, Sanchez JM, Guisasaola LM, Salas-Puig J, Moris G, Vidal JA, Ribacoba R, Menes BB, Ura D, Coto E (1999) Angiotensin converting enzyme and endothelial nitric oxide synthase DNA polymorphisms and late onset Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 67:733–736
14. Dahiyat M, Cumming A, Harrington C, Wischik C, Xuereb J, Corrigan F, Breen G, Shaw D, St Clair D (1999) Association between Alzheimer's disease and the NOS3 gene. *Ann Neurol* 46:664–667
15. Lambert J-C, Goumidi L, Wavrant-De Vrieze F, Frigard B, Harris JM, Cummings A, Coates J, Pasquier F, Cottel D, Gaillac M, St Clair D, Mann DMA, Hardy J, Lendon CL, Amoutel P, Chartier-Harlin M-C (2000) The transcriptional factor LBP-1C/CP2/LSF gene on chromosome 12 is a genetic determinant of Alzheimer's disease. *Hum Mol Genet* 9:2275–2280
16. Manton KG, Woodbury MA, Tolley HD (1994) Statistical applications using fuzzy sets. John Wiley & Sons, New York
17. Woodbury MA, Manton KG (1982) A new procedure for analysis of medical classification. *Methods Inf Med* 21:210–220
18. Gold DT, Woodbury MA, George LK (1990) Relationship classifications using grade of membership analysis: a typology of sibling relationships in later life. *J Gerontol* 45:S43–S51
19. Li Y, Zhou W, Tong Y, He G, Song W (2006) Control of APP processing and Abeta generation level by BACE1 enzymatic activity and transcription. *FASEB J* 20:285–292
20. Conrad C, Vianna C, Schultz C, Thal DR, Ghebremedhin E, Lenz J, Braak H, Davies P (2004) Molecular evolution and genetics of the Saitohin gene and tau haplotype in Alzheimer's disease and argyrophilic grain disease. *J Neurochem* 89:179–188
21. De Silva R, Hardy J, Crook J, Khan N, Graham EA, Morris CM, Wood NW, Lees AJ (2002) The tau locus is not significantly associated with pathologically confirmed sporadic Parkinson's disease. *Neurosci Lett* 330:201–203
22. Retz W, Thome J, Durany N, Harsanyi A, Retz-Junginger P, Kornhuber J, Riederer P, Rosler M (2001) Potential genetic markers of sporadic Alzheimer's dementia. *Psychiatr Genet* 11:115–122
23. Balbin M, Grubb A, Abrahamson M (1993) An Ala/Thr variation in the coding region of the human cystatin C gene (CST3) detected as a SstII polymorphism. *Hum Genet* 92:206–207
24. Artiga MJ, Bullido MJ, Sastre I, Recuero M, Garcia MA, Aldudo J, Vazquez J, Valdivieso F (1998) Allelic polymorphisms in the transcriptional regulatory region of apolipoprotein E gene. *FEBS Lett* 421:105–108