European Journal of Human Genetics (2001) 9, 437–444

www.nature.com/ejhg

ARTICLE

Lack of replication of association findings in complex disease: an analysis of 15 polymorphisms in prior candidate genes for sporadic Alzheimer's disease

Jonathan A Prince*[,1], Lars Feuk[1], Sarah L Sawyer[1], Johan Gottfries[5], Anne Ricksten[2], Katarina Nägga[3], Nenad Bogdanovic[4], Kaj Blennow[2] and Anthony J Brookes[1]

[1]*Center for Genomics Research, Karolinska Institute, Stockholm, Sweden;* [2]*Department of Clinical Neuroscience and Transfusion Medicine, University of Göteborg, Sahlgren's University Hospital, Sweden;* [3]*Department of Geriatric Medicine, Linköping University Hospital, Linköping, Sweden;* [4]*Department of Clinical Neuroscience and Family Medicine, Section of Geriatric Medicine, Karolinska Institute, Huddinge University Hospital, Stockholm, Sweden;* [5]*Department of Medicinal Chemistry, AstraZeneca, Mölndal, Sweden*

There is considerable enthusiasm for the prospect of using common polymorphisms (primarily single nucleotide polymorphisms; SNPs) in candidate genes to unravel the genetics of complex disease. This approach has generated a number of findings of loci which are significantly associated with sporadic Alzheimer's disease (AD). In the present study, a total of 15 genes of interest were chosen from among the previously published reports of significant association in AD. Genotyping was performed on polymorphisms within those genes (14 SNPs and one deletion) using Dynamic Allele Specific Hybridization (DASH) in 204 Swedish patients with sporadic late-onset AD and 186 Swedish control subjects. The genes chosen for analysis were; low-density lipoprotein receptor-related protein (LRP1), angiotensin converting enzyme (DCP1), alpha-2-macroglobulin (A2M), bleomycin hydrolase (BLMH), dihydrolipoyl S-succinyltransferase (DLST), tumour necrosis factor receptor superfamily member 6 (TNFRSF6), nitric oxide synthase (NOS3), presenilin 1 (PSEN1), presenilin 2 (PSEN2), butyrylcholinesterase (BCHE), Fe65 (APBB1), oestrogen receptor alpha (ESR1), cathepsin D (CTSD), methylenetetrahydrofolate reductase (MTHFR), and interleukin 1A (IL1A). We found no strong evidence of association for any of these loci with AD in this population. While the possibility exists that the genes analysed are involved in AD (ie they have weak effects and/or are population specific), results reinforce the need for extensive replication studies if we are to be successful in defining true risk factors in complex diseases. *European Journal of Human Genetics* (2001) **9**, 437–444.

Keywords: SNP; DASH; polymorphism; Alzheimer's; sporadic; association

Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disorder of the elderly with an annual incidence of approximately 1% in 65–69 year-olds increasing to about 8% in 85–95 year olds.[1] Over the past decade, there has been considerable effort towards unravelling the genetics of AD. A minor fraction of AD cases can be attributed to the autosomal dominant inheritance of mutations in either amyloid precursor protein (APP),[2] presenilin 1 (PSEN1),[3] or presenilin 2 (PSEN2).[4,5] Together these may account for as much as 70% of all early-onset AD cases,[6] but this represents only a small fraction of all AD cases. The vast majority of individuals afflicted with AD suffer from the non-familial (sporadic) form of the disorder, which is primarily evident in late-onset cases. To date, the only well-replicated genetic risk factor for sporadic AD is possession of the e4 allele of the gene encoding apolipoprotein E (APOE).[7–9] Estimates are that APOE contributes to anywhere between 30–70% of the total

*Correspondence: Dr JA Prince, Center for Genomics Research, Karolinska Institute, Theorells väg 3, S-171 77 Stockholm, Sweden. Tel: +46 8 7286274; Fax: +46 8 331547; E-mail: Jonathan.Prince@cgr.ki.se
Received 21 November 2000; revised 9 February 2001; accepted 7 March 2001

genetic variance in AD.[10,11] Thus, a large proportion of the genetic variance in AD must be determined by as yet unidentified loci. It has been suggested that there may exist as many as four additional loci with an effect equal or greater to that of APOE.[11]

There are numerous reports of positive association findings in sporadic AD. The dominant approach has been to employ polymorphisms within candidate genes from biological pathways assumed to be involved in the pathogenesis of AD. Other than for APOE, significant association findings have been made for; alpha-2-macroglobulin (A2M),[12] interleukin 1A (IL1a),[13,14] angiotensin converting enzyme (DCP1),[15] nitric oxide synthase (NOS3),[16] PSEN1,[17] PSEN2,[18] butyrylcholinesterase (BCHE),[19] dihydrolipoamide succinyltransferase (DLST),[20] Fe65 (APBB1),[21] oestrogen receptor-alpha (ESR1),[22] cathepsin D (CTSD),[23] low-density lipoprotein receptor-related protein (LRP1),[24] bleomycin hydrolase (BLMH)[25] and tumour necrosis factor receptor superfamily member 6 (TNFRSF6).[26]

The chance of identifying risk alleles in complex disease (eg APOE-e4 in AD) is primarily dependent upon the size of the effect of a particular gene in that disease and to the size of the population studied. On the basis that most association studies are performed on small to moderate sized single populations, the only way to confirm the importance of a risk allele in a disease is by replication. Attempts at replication have been performed on some of these genes, but results have in general been negative. The importance of independent replication has also recently been emphasised in response to the original finding of association with A2M and AD.[12,27] There are a number of factors which may contribute to the inability to reproduce association findings in a complex disease. The results of the original studies may be false positives due to population stratification or lack of proper correction for multiple testing. In turn, results in follow up studies may be false negatives due to lack of power and/or differences in ethnic background. In the present study, we have addressed the question of whether or not any of the aforementioned genes are risk factors for sporadic AD in a Swedish population. Considering the strength of the effects reported in previous studies and given the size of our population, this study had 80% power (α=0.05) for detecting association in any of these genes if they had similar effects in this population. Data for the APOE-e4 allele have been obtained earlier for this material[28] and are therefore included for comparative purposes in this study.

Materials and Methods

Clinical materials

All patients and controls were of Swedish descent. The AD group consisted of 204 patients, 79 men and 125 women. The control group consisted of 186 individuals, 78 men and 108 women. None of the AD patients had a family history of dementia.

Of the 204 AD cases, 111 had a clinical and 93 a neuropathological diagnosis. All clinically diagnosed AD patients underwent a thorough investigation, which included a medical history, physical, neurological and psychiatric examination, screening laboratory tests, ECG, X-ray of the chest, EEG, and computerised tomography (CT) of the brain. The diagnosis of 'probable AD' was made according to the NINCDS-ADRDA criteria.[24] All neuropathologically diagnosed AD patients also fulfilled the clinical NINCDS-criteria for probable AD[29] and met the neuropathological CERAD criteria for definitive AD.[30,31]

Of the 186 controls, 76 were healthy volunteers, while 108 were autopsy cases. The healthy volunteers were individuals without history, symptoms or signs of psychiatric or neurological disease, malignant disease, or systemic disorders. Cognitive status was examined using MMSE,[32] and individuals with scores below 28 were not included as controls. The autopsy control group consisted of patients who had died from cardiac disease or malignant disease. Their medical records revealed no history of dementia, or psychiatric or neurological diseases. Post-mortem examination revealed no macroscopic infarcts, and upon histopathological examination low numbers of senile plaques and neurofibrillary tangles (NFTs) were found. All autopsy cases (AD and control) were matched by age at death and all clinically diagnosed AD cases and healthy volunteers were matched by age-at-onset/age-at-exam respectively.

PCR and choice of polymorphisms

PCR primers were designed using Oligo version 5.0 software (Molecular Biology Insights, CO, USA) to amplify target genomic DNA fragments of 47–60 bp which contained the polymorphism of interest centered within the product. Details of the PCR primer sequences and the sequences of the probes employed for genotyping (see below) are shown in Table 1. In 13 of the chosen genes, the polymorphisms originally reported to be associated were genotyped and in two genes, different polymorphisms were used. These two latter cases were for DCP1 and PSEN2. For DCP1, a surrogate was used because of limitations in our genotyping technology for scoring large deletions (for which the original report was made). In the case of PSEN2, the original report was for a 3′-UTR variant and we chose to focus on an intronic variant. The DCP1 variation used was less than 200 bp from the common ALU insertion/deletion and the studied polymorphism for PSEN2 was less than 1 kb from the originally published variation. All oligonucleotides (primers and probes) were ordered from Interactiva GmbH (Germany) and were supplied after HPLC purification.

PCR was performed in 96-well microtiter plates (Corning Costar, NY, USA) on a TouchDown™ temperature cycling apparatus (Hybaid, UK). Reactions (25 μl total volume) consisted of 2–5 ng genomic DNA, 20 pmol of a regular oligonucleotide primer and 4 pmol of a 5′-biotinylated primer, 0.75u AmpliTaq Gold (PE Biosystems, CA, USA),

10% dimethylsulphoxide, 1× Perkin-Elmer PCR-buffer (including 1.5 mM $MgCl_2$) and 0.2 mM each of dGTP, dATP, dTTP and dCTP. Thermal cycling employed the following conditions: 1× (10 min at 94°C), 40× (15 s at 94°C, 30 s at

Table 1 Details of DASH assays

Gene Symbol	*SNP ID*	*Oligo Name*	*Sequence (5′-3′)*	*Original Reference/Comments*
PSEN1	SNP000062972	Forward	CCTCCGTGGACCGACCGCCAA	(van Duijin *et al.*, 1999): Found in early-onset AD
		Reverse	b-CGTTGTCGTCATTTGCGGCTCTGG	Several polymorphisms spanning the PSEN 1 gene
		Probe 1	GCTCTGGCGTCGTTGGC	were analysed
		Probe 2	GCTCTGGCATCGTTGGC	
PSEN2	SNP000005166	Forward	b-CTGTCGTCAGTACACGGTCCCT	(Brookes *et al.*, 1997): Found in early-onset AD.
		Reverse	CAGCAGTGGGATGGCATTGGGT	Original SNP in 3′UTR.
		Probe 1	CATTGGGTATTAGGCAG	An SNP within intron 10 was scored in our study.
		Probe 2	CATTGGGTGTTAGGCAG	
DLST	SNP000002340	Forward	GGTGAAGGAAACTGGGGTGGCT	(Nakano *et al.*, 1997): Found in early and late onset AD.
		Reverse	b-TCGGGCCTAAAACAGACTTCAT	Haplotypes were used from 2 SNPs. These SNPs are
		Probe 1	GGTGGCTTAATATTTCA	68 bp apart. Only one of the variants was used in our
		Probe 2	GGTGGCTTGATATTTCA	study.
LRP1	SNP000002603	Forward	TCTGCAATGTGGTCTAGGACTG	Kang *et al.*, 1997): Late-onset AD. Over-representation of
		Reverse	b-TCGGTAGTGTGGCCCCTCATCT	CC homozygotes found in AD.
		Probe 1	GCATGGACGGCTCAG	
		Probe 2	GCATGGATGGCTCAG	
DCP1	SNP000006569	Forward	CAGATCTGACGAATGTTATGGC	(Kehoe *et al.*, 1999): Found in late-onset AD. Original
		Reverse	b-ATAACAGGTCTTCATATTTCC	finding made using common ins/del. Enrichment
		Probe 1	ATGGCCACATCCCGGAA	of insertion of AD found. A surrogate SNP was used in our
		Probe 2	ATGGCCACGTCCCGGAA	study.
IL1a	SNP000003482	Forward	TTAATAATAGTAACCAGGCAAC	(Grimaldi *et al.*, 2000; Nicoll *et al.*, 2000) Found in both
		Reverse	b-TTTTTACATATGAGCCTTCAAT	early and late onset AD. TT homozygotes enriched in AD.
		Probe 1	AGGCAACATCATTGAAG	
		Probe 2	AGGCAACACCATTGAAG	
BCHE	SNP000011184	Forward	ATTTTACAGGAAATATTGATGA	(Lehmann *et al.*, 1997): Late-onset AD. Found in whole set
		Reverse	b-CTGCTTTCCACTCCCATTCT	and in e4 positive individuals. Enrichment of AA
		Probe 1	TTGATGAAGCAGAATGG	homozygotes found in AD.
		Probe 2	TTGATGAAACAGAATGG	
ESR1	SNP000002389	Forward	GTGCCTTGTTGGATGCTGAGCC	(Brandi *et al.*, 1999): Late-onset AD. Found using haplotypes
		Reverse	b-TAGGATCATAATCGGAATAGAG	derived from 2 SNPs. These SNPs are 50 bp apart. Only one
		Probe 1	GAGCCCCCCATACTCTA	of the variants was used in our study.
		Probe 2	GAGCCCCCGATACTCTA	
APBB1	SNP000005020	Forward	TGGCCCCATGGGAGGTAGGCTG	(Hu *et al.*, 1998): Late-onset cases. 3 bp deletion scored
		Reverse	b-GGAGACAAGCAGGGCAGTCACT	which was in complete LD with the SNP scored in our study.
		Probe 1	TAGGCTGGAGAGTGACT	
		Probe 2	TAGGCTGGGGAGTGACT	
BLMH	SNP000003199	Forward	TAGCTGTGTTAGAGCATGAACC	(Montoya *et al.*, 1998): No age of onset specified.
		Reverse	b-CATAGGGTCCAATGCTGGCAGG	Enrichment of GG homozygotes in AD found in whole set
		Probe 1	AACCCATTATCCTGCCA	and in e4 negative individuals.
		Probe 2	AACCCATTGTCCTGCCA	
CTSD	SNP000008899	Forward	CCCCGTCTCAAAGTACTCCCAG	(Papassotiropoulos *et al.*, 2000): Primarily late-onset cases.
		Reverse	b-GCACCTCAGTCACGACTGGCAC	Enrichment of CT heterozygotes found in AD.
		Probe 1	CTCCCAGGCGGTGCCAG	
		Probe 2	CTCCCAGGTGGTGCCAG	
NOS3	SNP000002600	Forward	CACTGCTGCTTCAGGTCCCAGA	(Dahiyat *et al.*, 1999): Found in late-onset AD but not
		Reverse	b-CGGGTGGCAGTAGGAAGAGTTC	early-onset. An excess of GG homozygotes was found in AD.
		Probe 1	CAGATGATCCCCCAG	
		Probe 2	CAGATGAGCCCCCAG	
TNFSFR6	SNP000005172	Forward	TCATATGGTTAACTGTCCATTCC	(Feuk *et al.*, 2000): Early-onset AD. Enrichment of GG
		Reverse	b-GCAACATGAGAGGCTCACAGAC	homozygotes found in AD.
		Probe 1	ATTCCAGGAACGTCT	
		Probe 2	ATTCCAGAAACGTCT	
MTHFR	SNP000002390	Forward	b-ACTTGAATGAGAAGGTGTCTGC	(Frost *et al.*, 1995): This has only been suggested to be a
		Reverse	AAAGCTGCGTGATGATGAAATC	candidate for AD. There is as yet no convincing genotyping
		Probe 1	TGAAATCGGCTCCCGCA	data showing association in AD.
		Probe 2	TGAAATCGACTCCCGCA	
A2M	IND000007566	Forward	TATGGTAATGTCTTCTTCCTCAC	(Blacker *et al.*, 1998): Inheritance of the deletion found to
		Reverse	b-CCGCTTCCCATTACATCTGACTC	confer increased susceptibility to AD. A family based strategy
		Probe 1	CACTCACCATAGAGT	was used.
		Probe 2	CCTCACTCAGAGTCAG	

SNP IDs are according to (http://hgbase.cgr.ki.se/). Sequences shown are forward and reverse PCR primers and DASH probes for both allelic variants of listed SNPs. The presence of a 'b' denotes the attachment of a biotin group at the 5′ end of a sequence.

annealing temperature (51, 55, or 58°C)). Before being used for genotyping (see below), 5 μl of the completed PCR reaction was examined upon a 2.0% low-melt agarose gel to confirm successful amplification. In large-scale experiments, only a few random representative samples were tested in this way.

Dynamic allele-specific hybridization (DASH)

DASH assays were conducted in a similar fashion to the method described previously,[33,34] with details as follows. Two probes (shown in Table 1) were designed for each SNP, representing both allelic sequences complementary to the biotinylated strand of the PCR product. Biotinylated PCR products (20–25 μl) were diluted to 50 μl with HEN buffer (0.1 M HEPES, 10 mM EDTA, 50 mM NaCl, pH 8.0). Of this, 25 μl was added to one well of a 96-well streptavidin coated microtiter plate (Omni-plate, Hybaid, UK) and surface binding was allowed to proceed for a minimum of 1 h at room temperature. The solution was then removed and 50 μl of 0.1 M NaOH was added at room temperature to cause bound PCR products to denature for 5 min. The NaOH solution was discarded, thus removing the unbiotinylated PCR product strand, and the well was rinsed with a further aliquot of 50 μl 0.1 M NaOH. A 50 μl solution containing HEN, Sybr Green I dye at a 1 : 10 000 dilution (Molecular Probes, OR, USA), and 30 pmol of one of the two allele specific oligonucleotide probes was added. The microtiter plate was sealed with a plastic cover and rapidly heated to 85°C, and then cooled to 25°C over 3–4 min. This allowed the probe to hybridise to the bound PCR product strand, regardless of which alleles were present. The solution was replaced with HEN containing Sybr Green I dye at a 1 : 10 000 dilution. Plates were placed in a DASH instrument (Hybaid, UK) and dye fluorescence (proportional to the amount of duplex DNA present) was monitored while heating from 35 to 85°C at a rate of 0.3°C/s. Probe-duplex denaturation was determined by following the fluorescence decrease as temperature was increased. The assay procedure, beginning with the NaOH wash step, was repeated on the bound samples using the alternative allele specific probe. Genotypes were scored from fluorescence curves as previously described.[33] In this way, observed peaks in the negative first derivative plot of change in fluorescence with respect to temperature were taken to indicate the observed melting temperature (Tm).

Statistical analysis

Statistical differences in genotype and allele frequencies between cases and controls were ascertained using χ^2 and Fisher's exact tests. ANOVA and logistic regression were used to test for relationships between NFT scores and genotypes in autopsy confirmed samples as well as to assess differences in age of onset with respect to genotype in the whole AD set. These tests were performed using StatView version 5.0 (Abacus Concepts).

Principal Components Analysis (PCA) and Partial Least Squares (PLS) were used to reveal underlying structure in the whole sample of genotypes. These are data compression methods that summarise data tables using bilinearisation of a given data matrix.[35] In the case of PLS the compressed descriptor matrix is regressed to a similarly compressed response matrix whereafter the correlation between the object vectors, derived from the respective matrices, is optimised.[36] The PCA is normally a straightforward tool for revealing major variance structure and clustering.[37] In cases where the clustering is subtle however, one can introduce retrospective knowledge of expected clustering, eg clinical diagnosis, to enhance the power of the PCA by regressing it to a discrete variable comprising such information, ie PLS discriminant analysis (PLS-DA).[38] The PLS-DA does not bias the description matrix as such but changes the angle from which the data is observed to the most favourable for the purpose of the study, as provided by the discrete regression variable. In the present study 204 AD patients and 186 controls were included and the 16 SNPs were coded 0, 1 or 2 for the occurrence of the least frequent allele. Gender was also included in the analysis. The regressor was a dummy variable indicating AD as 1 and control as 0. Model validation was made by the default tool kit provided by the software (SIMCA-P, 8.0) mainly driven by the cross-validation procedure set to dividing the total data-set into seven groups and leaving one group out for prediction.[39]

Results

We have attempted to replicate a large number of association studies which each point to the involvement of a specific candidate gene in the aetiology of sporadic AD. To begin this project, we first determined whether or not our sample would have sufficient power to detect the previously reported associations. The reported odds ratios for previous studies ranged from 2.04–12.8. Using the frequencies of the risk factors for which those odds ratios were obtained as a basis, we estimated that our sample had over 80% power to detect all of the previously reported effects (α=0.05). The association between APOE-e4 and AD is also highly significant in our material (for possession of the e4 allele; OR=5.2, CI=3.35–8.04, $P<0.0001$)[28] suggesting sufficient power to detect associations in loci with much smaller effects. The χ^2 derived significances for the genotype frequency differences between cases and controls are shown in Table 2. In general, genotype and allele frequencies were similar to those of the original publications in which association was reported. There were no significant deviations from Hardy-Weinberg equilibrium for any of the studied polymorphisms.

Among all of the tests that were performed in this study, only two positive findings were made. We found a significant increase in the frequency of heterozygotes for the BCHE polymorphism in controls compared to AD patients when

Table 2 Distribution of genotypes for AD patients and controls

Gene	*Genotypes*			*Significance*
DCP1	*AA*	*AG*	GG	
AD	58/202 (28.7%)	99/202 (49.0%)	45/202 (22.2%)	*P*=0.1814
Control	34/167 (20.3%)	92/167 (55.0%)	41/167 (24.5%)	
BCHE	*AA*	*AG*	*GG*	
AD	11/201 (5.4%)	62/201 (30.8%)	128/201 (63.7%)	*P*=0.4449
Control	5/166 (3.0%)	57/166 (34.3%)	104/166 (62.6%)	
TNFRSF6	*GG*	*AG*	*AA*	
AD	41/204 (20.0%)	100/204 (49.0%)	63/204 (30.8%)	*P*=0.2675
Control	40/179 (22.3%)	97/179 (54.1%)	42/179 (23.5%)	
NOS3	*TT*	*GT*	*GG*	
AD	23/202 (11.3%)	75/202 (37.1%)	104/202 (51.5%)	*P*=0.1315
Control	13/180 (7.2%)	83/180 (46.1%)	84/180 (46.7%)	
BLMH	*AA*	*AG*	*GG*	
AD	94/204 (46.1%)	85/204 (41.7%)	25/204 (12.3%)	*P*=0.1643
Control	67/175 (38.3%)	90/175 (51.4%)	18/175 (10.3%)	
CTSD	*CC*	*CT*	*TT*	
AD	177/204 (86.8%)	27/204 (13.2%)	0/204 (0%)	*P*=0.8645
Control	152/174 (87.4%)	22/174 (12.6%)	0/174 (0%)	
DLST	*AA*	*AG*	*GG*	
AD	59/204 (28.9%)	106/204 (52.0%)	39/204 (19.0%)	*P*=0.1893
Control	65/174 (37.4%)	83/174 (47.7%)	26/174 (14.9%)	
ESR1	*CC*	*CG*	*GG*	
AD	114/202 (55.9%)	76/204 (37.3%)	14/204 (6.9%)	*P*=0.1536
Control	111/173 (64.2%)	56/173 (32.4%)	6/173 (3.5%)	
APBB1	*AA*	*AG*	*GG*	
AD	152/204 (74.5%)	45/204 (22.1%)	7/204 (3.4%)	*P*=0.7229
Control	125/176 (71.0%)	45/176 (25.6%)	6/176 (3.4%)	
IL1a	*CC*	*CT*	*TT*	
AD	89/198 (44.9%)	89/198 (44.9%)	20/198 (10.1%)	*P*=0.264
Control	93/175 (53.1%)	65/175 (37.1%)	17/175 (9.7%)	
LRP1	*CC*	*CT*	*TT*	
AD	155/204 (76.0%)	47/204 (23.0%)	2/204 (1%)	*P*=0.2262
Control	124/171 (72.5%)	41/171 (24.0%)	6/171 (3.5%)	
PSEN2	*CC*	*CT*	*TT*	
AD	13/202 (6.4%)	64/202 (31.7%)	125/202 (61.9%)	*P*=0.4107
Control	6/175 (3.4%)	58/175 (33.1%)	111/175 (63.4%)	
MTHFR	*CC*	*CT*	*TT*	
AD	98/204 (48.0%)	90/204 (44.1%)	16/204 (7.8%)	*P*=0.3206
Control	83/172 (48.2%)	68/172 (39.5%)	21/172 (12.2%)	
PSEN1	*CC*	*CT*	*TT*	
AD	173/204 (84.8%)	31/2045 (15.2%)	0/204 (0%)	*P*=0.2895
Control	150/176 (85.2%)	24/176 (13.6%)	2/176 (1.1%)	
A2M	*ins/in*	*ins/del*	*del/del*	
AD	139/200 (69.5%)	56/200 (28.0%)	5/200 (2.5%)	*P*=0.3624
Control	137/183 (74.9%)	40/183 (21.9%)	6/183 (3.3%)	

Significance was determined by χ^2 or Fisher's exact tests where applicable. Presented *P* values are uncorrected for multiple testing.

conditioned for individuals who lacked the APOE-e4 allele (*P*=0.003). This result differed from the association which was originally found for this variant in which a significant enrichment of AA-homozygotes was found in APOE-e4 carriers.[19] We also found evidence for an effect of the TNFRSF6 AA-genotype on age of onset in AD. The average age of onset in AA-homozygotes at this variant site was 4.5 years less than for GG-homozygotes (mean $\pm$ SD: 73.8 years $\pm$ 6.7 for GG-homozygotes; 69.3 years $\pm$ 9.7 for AA-homozygotes; $P<0.05$). This finding also differed from the original finding of association at this locus.[26] We interpret both of these results (TNFRSF6 and BCHE) with caution given the number of tests that were performed in the whole study as well as lack of agreement with the original studies in which association was found. Application of a Bonferroni correction to both the BCHE and TNFRSF6 results eliminates the significance of the findings. There were no other significant findings in any of the other loci in terms of age of onset in this study.

With the exception of the findings for TNFRSF6 and BCHE, we found no other evidence of genotype or allele association in any of the 15 studied loci in this AD sample. We also failed to find any other significant association after conditioning all genotypes for APOE-e4 status (ie carrier or non-carrier of the

e4 allele). No relationship between the studied polymorphisms and NFT scores in the autopsy confirmed subjects was found. The loss of a few samples in both cases and controls seen in Table 2 were due to PCR failures. This was probably a result of low quantities of genomic DNA in those wells as the same samples tended to amplify poorly in all assays.

Initial modelling with PCA revealed a low degree of structure in the data. To enhance the search for clusters PLS-DA was applied and also this method failed to find significant principal components as judged by cross validation. Such outcome indicates that the major variation of the data had little in common with the regressor, namely the diagnostic group (AD or control). Inspection of the score plot for the two first principal components revealed an enrichment of AD patients in one end coinciding with the effect of the APOE-e4 allele (see loadings as plotted in Figure 1). All other genotypes and gender comprised a random distribution compared to the regressors AD *vs* controls. In addition to linear PLS-DA, we also investigated possible non-linearity by introducing all possible square and cross-terms to the modelling. However, this did not improve the clustering power of the present data set.

Discussion

The genetics of sporadic AD is complex, involving the interaction of a number of genes and the environment. It has been suggested that there may be as many as five independent loci which have a large effect on the development of sporadic AD.[11] The most convincing evidence for an allelic association in AD is in the APOE locus.[7] The search for other loci of large effect in AD over the past 8 years however has been dizzying, with a large number of positive findings for which replication studies have generated negative results as well as a number of studies in which replication has not yet been attempted. In the present study, we have extracted the most significant findings in sporadic AD from the literature and performed association studies on either those reported polymorphisms (13 cases) or on polymorphisms within the genes of interest but in close proximity to the polymorphisms for which association was reported (two cases). We found no strong evidence for genetic association among the 15 tested variants in this AD sample. While we cannot exclude a biological effect for any of the studied genes, our data suggest that the effect would have to be small (OR < 1.8) in order to avoid detection.

How should we reconcile our inability to replicate any of the original association findings? It is important to emphasise that these results by no means entail that the original findings were wrong. Our study differed in several respects to the original reports. First and foremost, we cannot exclude the possibility that our population differs from the initially studied populations in terms of allelic or non-allelic heterogeneity. If the gene effects previously reported are specific to the studied populations then we would invariably be biased in our results using a different population. The range of

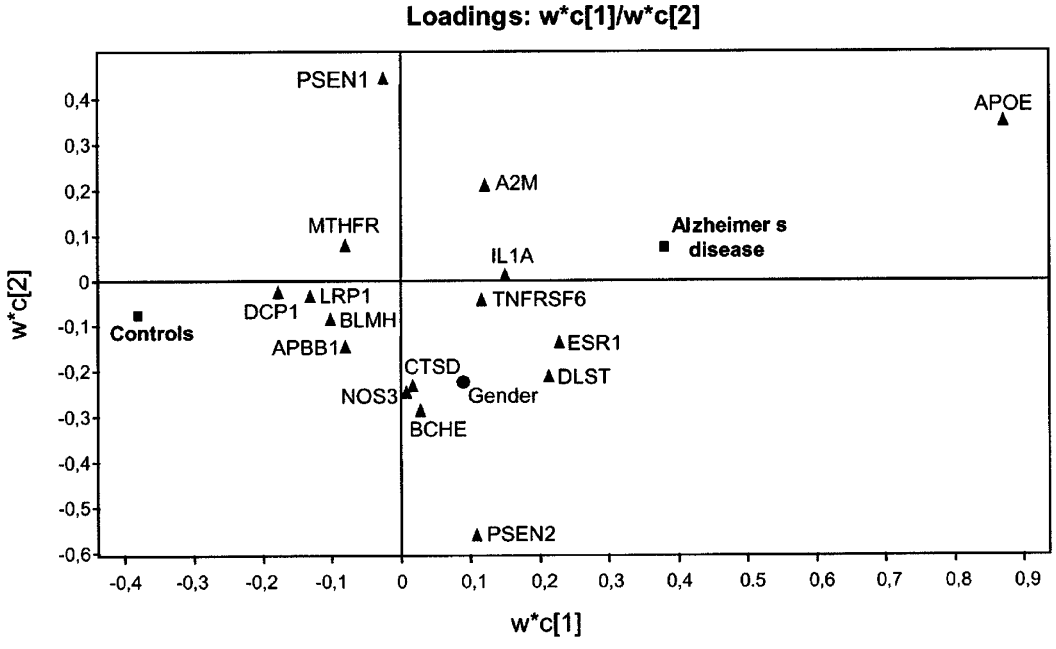

Figure 1 The initial genotyping variable correlations to the compressed and diagnosis biased genotype vectors (w1 and w2), indicating random correlation for all genotypes except the ApoE4 allele which correlated with the AD phenotype. The direction of the discriminant indicator (c1 and c2) was superimposed for clarity of direction (ie AD *vs* controls).

populations used in previous studies was large (primarily from the US and Europe), but none employed Swedish or Nordic individuals. The possibility also exists that the studied polymorphisms may be in linkage disequilibrium (LD) with a true pathogenic variant, and that the extent of LD may have been higher in the original populations. Performing association studies using numerous polymorphisms within the same gene has been suggested as a possible solution to this[40] and may also be the only solution to avoiding the problems of allelic heterogeneity. In two of the 15 genes (DCP1 and PSEN2), we used polymorphisms that were in close proximity to those originally reported. There is debate about the extent to which LD will allow for the detection of pathogenic variants via surrogate polymorphisms, but the suggestion is that useful association may extend as far as 50 kb.[41] We attempted to use the closest available variants, but this may have nonetheless led to our inability to replicate those findings.

In a few of the original studies, the most significant findings were made in early onset cases. The majority of our cases had an age of onset over 65 years of age. We did attempt to stratify our study into <65 years and $\geqslant 65$ years for age of onset, and while results were negative, this causes perhaps too great a reduction in power since the number of individuals in the former group is quite small (<40). It is thus possible that some of the original findings may have been made due to enhanced gene effects in early onset cases.

To investigate if putative combinations of SNPs might result in an additive risk for AD, we also performed multivariate analysis. For this we employed PLS-DA, which maximises the ability to reveal clustering in a given data set with an indication of which genotypes correlate to the AD-diagnosis. In the initial linear model the only correlating SNP was the APOE-e4 allele. However, a linear model would not reveal whether or not certain combinations of alleles that might be associated with increased risk for AD (eg the exponential increase in risk for cardiovascular disease by smoking, high cholesterol and obesity). Therefore, a non-linear PLS-DA was undertaken that included all possible cross and square combinations of the alleles. This analysis also failed to find any further association to AD except for the APOE-e4 allele.

In conclusion, results suggest that with the exception of possession of the APOE-e4 allele, none of the other investigated SNPs contribute substantially to the development of AD in this sample. While we cannot completely exclude the potential influence of these genes in AD (dependent upon their effect size), the fact that we were unable to replicate previous findings illustrates the fact that there are still issues that need to be resolved if association studies are to be successful in identifying genetic risk factors in complex disease. This includes the need for larger and perhaps better defined case-control populations, stricter statistics (with an adequate correction for multiple testing) and most importantly, because there may be considerable population differences, a forum for the publication of negative results.

Acknowledgements

Generous financial support has been provided by; The Swedish Brain Fund, and the AMF Jubileum Fund, Tore Nilsons Foundation for Medical Research, Loo Ostermans Foundation, The Swedish Old Servants Foundation (Gamla Tjänarinnor), and Carl Tryggers Foundation for Scientific Research.

References

1 Hebert LE, Scherr PA, Beckett LA *et al*: Age-specific incidence of Alzheimer's disease in a community population. *JAMA* 1995; **273**: 1354–1359.
2 Goate A, Chartier-Harlin M-C, Mullan M *et al*: Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature* 1991; **349**: 704–706.
3 Sherrington R, Rogaev EI, Liang Y *et al*: Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature* 1995; **375**: 754–760.
4 Rogaev EI, Sherrington R, Rogaeva EA *et al*: Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene. *Nature* 1995; **376**: 775–778.
5 Levy-Lahad E, Wasco W, Poorkaj P *et al*: Candidate gene for the chromosome 1 familial Alzheimer's disease locus. *Science* 1995; **269**: 973–977.
6 Campion D, Dumanchin C, Hannequin D *et al*: Early-onset autosomal dominant Alzheimer disease: prevalence, genetic heterogeneity, and mutation spectrum. *Am J Hum Genet* 1999; **65**: 664–670.
7 Strittmatter WJ, Saunders AM, Schmechel D *et al*: Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proc Natl Acad Sci* 1993; **90**: 1977–1981.
8 Saunders AM, Strittmatter WJ, Schmechel D *et al*: Association of apolipoprotein E allele epsilon 4 with late-onset Alzheimer's disease. *Neurology* 1993; **43**: 1467–1472.
9 Roses AD, Saunders AM: APOE is a major susceptibility gene for Alzheimer's disease. *Curr Opin Biotechnol* 1994; **5**: 663–667.
10 Slooter AJ, Cruts M, Kalmijn S *et al*: Risk estimates of dementia by apolipoprotein E genotypes from a population-based incidence study: the Rotterdam Study. *Arch Neurol* 1998; **55**: 964–968.
11 Daw EW, Payami H, Nemens EJ *et al*: The number of trait loci in late-onset Alzheimer's disease. *Am J Hum Genet* 2000; **66**: 196–204.
12 Blacker D, Wilcox MA, Laird NM *et al*: Alpha-2 macroglobulin is genetically associated with Alzheimer's disease. *Nat Genet* 1998; **19**: 357–360.
13 Nicoll JA, Mrak RE, Graham DI *et al*: Association of interleukin-1 gene polymorphisms with Alzheimer's disease. *Ann Neurol* 2000; **47**: 365–368.
14 Grimaldi LM, Casadei VM, Ferri C *et al*: Association of early-onset Alzheimer's disease with an interleukin-1alpha gene polymorphism. *Ann Neurol* 2000; **47**: 361–365.
15 Kehoe PG, Russ C, McIlory S *et al*: Variation in DCP1, encoding ACE, is associated with susceptibility to Alzheimer's disease. *Nat Genet* 1999; **21**: 71–72.
16 Dahiyat M, Cumming A, Harrington C *et al*: Association between Alzheimer's disease and the NOS3 gene. *Ann Neurol* 1999; **46**: 664–667.
17 van Duijn CM, Cruts M, Theuns J *et al*: Genetic association of the presenilin-1 regulatory region with early-onset Alzheimer's disease in a population-based sample. *Eur J Hum Genet* 1999; **7**: 801–806.
18 Brookes AJ, Howell WM, Woodburn K, Johnstone EC, Carothers A: Presenilin-I, presenilin-II, and VLDL-R associations in early onset Alzheimer's disease. *Lancet* 1997; **350**: 336–337.

19 Lehmann DJ, Johnson C, Smith AD: Synergy between the genes for butyrylcholinesterase K variant and apolipoprotein E4 in late-onset confirmed Alzheimer's disease. *Hum Mol Genet* 1997; **6**: 1933 – 1936.
20 Nakano K, Ohta S, Nishimaki K, Miki T, Matuda S: Alzheimer's disease and DLST genotype. *Lancet* 1997; **350**: 1367 – 1368.
21 Hu Q, Kukull WA, Bressler SL *et al*: The human FE65 gene: genomic structure and an intronic biallelic polymorphism associated with sporadic dementia of the Alzheimer type. *Hum Genet* 1998; **103**: 295 – 303.
22 Brandi ML, Becherini L, Gennari L *et al*: Association of the estrogen receptor alpha gene polymorphisms with sporadic Alzheimer's disease. *Biochem Biophys Res Commun* 1999; **265**: 335 – 338.
23 Papassotiropoulos A, Bagli M, Kurz A *et al*: A genetic variation of cathepsin D is a major risk factor for Alzheimer's disease. *Ann Neurol* 2000; **47**: 399 – 403.
24 Kang DE, Saitoh T, Chen X *et al*: Genetic association of the low-density lipoprotein receptor-related protein gene (LRP), an apolipoprotein E receptor, with late-onset Alzheimer's disease. *Neurology* 1997; **49**: 56 – 61.
25 Montoya SE, Aston CE, DeKosky ST, Kamboh MI, Lazo JS, Ferrell RE: Bleomycin hydrolase is associated with risk of sporadic Alzheimer's disease. *Nat Genet* 1998; **18**: 211 – 212.
26 Feuk L, Prince JA, Emahazion T, Breen G, StClair D, Brookes AJ: Apolipoprotein-E dependent role for the FAS receptor in early onset Alzheimer's disease: positive association finding for a polymorphism in the *TNFRSF6* gene. *Human Genetics* 2000; **107**: 391 – 396.
27 Cohen B: Freely associating. *Nat Genet* 1999; **22**: 1 – 2.
28 Blennow K, Ricksten A, Prince JA *et al*: No association between the alpha2-macroglobulin (A2M) deletion and Alzheimer's disease, and no change in A2M mRNA, protein, or protein expression. *J Neural Transm* 2000; **107**: 1065 – 1079.
29 McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM: Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of department of health and human services task force on Alzheimer's disease. *Neurology* 1984; **34**: 939 – 944.
30 Mirra SS, Heyman A, McKeel D: The consortium to establish a registry for Alzheimer's disease (CERAD). Part II. Standardization of the neuropathologic assessment of Alzheimer's disease. *Neurology* 1991; **41**: 479 – 486.
31 Bogdanovic N, Morris JH: Diagnostic criteria for Alzheimer's disease in multicentre brain banking; in *Neuropathological Diagnostic Criteria for Brain Banking*. Biomedical and Health Research Series, IOS Press, pp 20 – 29.
32 Folstein M, Folstein S, McHugh P: ''Mini-Mental State'' A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975; **12**: 189 – 198.
33 Howell WM, Jobs M, Gyllensten U, Brookes AJ: Dynamic allele-specific hybridization. A new method for scoring single nucleotide polymorphisms. *Nat Biotechnol* 1999; **17**: 87 – 88.
34 Prince JA, Feuk L, Emahazion T, Jobs M, Howell WM, Brookes AJ. Robust and accurate single nucleotide polymorphism genotyping by dynamic allele specific hybridization (DASH): design criteria and assay validation. *Genome Res* 2000; **11**: 152 – 162.
35 Jackson JEA: *A user's guide to principal components*. Wiley, NY, 1991.
36 Wold S, Ruhe A, Wold H, Dunn III WJ: The collinearity problem in linear regression. The partial least squares approach to generalised inverses. *J Sci Stat Comput* 1984; **5**: 735 – 743.
37 Linusson A, Gottfries J, Lindgren F, Wold S: Statistical molecular design of building blocks for combinatorial chemistry. *J Med Chem* 2000; **43**: 1320 – 1328.
38 Gottfries J, Blennow K, Wallin A, Gottfries CG: Diagnosis of dementias using partial least squares discriminant analysis. *Dementia* 1995; **6**: 83 – 88.
39 Wold S: Cross-validatory estimation of the number of components in factor and principal component models. *Technometrics* 1978; **20**: 397 – 405.
40 Martin ER, Lai EH, Gilbert JR *et al*: SNPing away at complex diseases: analysis of single-nucleotide polymorphisms around APOE in Alzheimer's disease. *Am J Hum Genet* 2000; **67**: 383 – 394.
41 Collins A, Lonjou C, Morton NE: Genetic epidemiology of single-nucleotide polymorphisms. *Proc Natl Acad Sci* 1999; **96**: 15173 – 15177.