Hum Genet (2003) 113 : 258–267
DOI 10.1007/s00439-003-0960-2

ORIGINAL INVESTIGATION

Denise Harold · Timothy Peirce · Valentina Moskvina · Amanda Myers · Susan Jones · Paul Hollingworth Pamela Moore · Simon Lovestone · John Powell · Catherine Foy · Nicola Archer · Sarah Walter Amanda Edmonson · Stephen McIlroy · David Craig · Peter A. Passmore · Alison Goate · John Hardy Michael O'Donovan · Julie Williams · Malcolm Liddell · Michael J. Owen · Lesley Jones

Sequence variation in the CHAT locus shows no association with late-onset Alzheimer's disease

Received: 7 January 2003 / Accepted: 28 March 2003 / Published online: 21 May 2003

Abstract There is substantial evidence for a susceptibility gene for late-onset Alzheimer's disease (AD) on chromosome 10. One of the characteristic features of AD is the degeneration and dysfunction of the cholinergic system. The genes encoding choline acetyltransferase (ChAT) and its vesicular transporter (VAChT), *CHAT* and *SLC18A3* respectively, map to the linked region of chromosome 10 and are therefore both positional and obvious functional candidate genes for late-onset AD. We have screened both genes for sequence variants and investigated each for association with late-onset AD in up to 500 late-onset AD cases and 500 control DNAs collected in the UK. We detected a total of 17 sequence variants. Of these, 14 were in *CHAT*, comprising three non-synonymous variants (D7N in the S exon, A120T in exon 5 and L243F in exon 8), one synonymous change (H547H), nine single-nucleotide polymorphisms in intronic, untranslated or promoter regions, and a variable number of tandem repeats in intron 7. Three non-coding SNPs were detected in *SLC18A3*. None demonstrated any reproducible association with late-onset AD in our samples. Levels of linkage disequilibrium were generally low across the *CHAT* locus but two of the coding variants, D7N and A120T, proved to be in complete linkage disequilibrium.

D.H. and T.P. contributed equally to this paper

D. Harold · T. Peirce · V. Moskvina · S. Jones · P. Hollingworth
P. Moore · M. O'Donovan · J. Williams · M. Liddell · M. J. Owen
L. Jones
Department of Psychological Medicine,
University of Wales College of Medicine,
Cardiff, CF14 4XN, UK

L. Jones (✉)
Institute of Medical Genetics,
University of Wales College of Medicine,
Cardiff, CF14 4XN, UK
Tel.: +44-29-20745175, Fax: +44-29-20746551,
e-mail: jonesL1@cardiff.ac.uk

A. Myers · J. Hardy
Laboratory of Neurogenetics,
MSC 0900, Building 9, Room 1N108,
National Institute of Aging, National Institute of Health,
9000 Rockville Pike, Bethesda, MD 20892, USA

S. Lovestone · J. Powell · C. Foy · N. Archer · S. Walter
A. Edmonson
Institute of Psychiatry,
De Crespigny Park, Denmark Hill, London, SE5 8AF, UK

S. McIlroy · D. Craig · P. A. Passmore
Department of Geriatric Medicine, Queen's University Belfast,
Belfast, Northern Ireland, UK

A. Goate
Departments of Psychiatry and Neurology,
Washington University School of Medicine,
660 S Euclid, St Louis, MO 63110, USA

Introduction

Alzheimer's disease (AD), a progressive neurodegenerative disorder, accounts for more than half of all cases of dementia among people over 65 years of age (Francis et al. 1999). The neuropathological hallmarks of the disease include extracellular deposits of β-amyloid in senile plaques and the intracellular formation of neurofibrillary tangles (for a review, see Hardy 1997). Molecular genetic analyses have led to the discovery of three genes involved in early-onset autosomal dominant AD: *APP* on chromosome 21, *PSEN1* on chromosome 14 and *PSEN2* on chromosome 1 (Hardy 1997). The majority of AD cases, however, have an age at onset of over 65 years and exhibit no clear Mendelian pattern of inheritance. The only widely accepted genetic risk factor for this late-onset AD (LOAD) is the ε4 allele of the apolipoprotein E (*APOE*) gene on chromosome 19. However, variation at the *APOE* locus accounts for less than half the genetic variation in liability to AD and at least four other genes are thought to underlie the remaining risk (Daw et al. 2000; Myers and Goate 2001).

A characteristic feature of AD is the widespread degeneration and dysfunction of the basal forebrain cholinergic system (Bowen et al. 1976; Davies and Maloney 1976; Perry et al. 1977; Whitehouse et al. 1982) that is believed to contribute significantly to the cognitive symptoms associated with the disease. Choline acetyltrans-

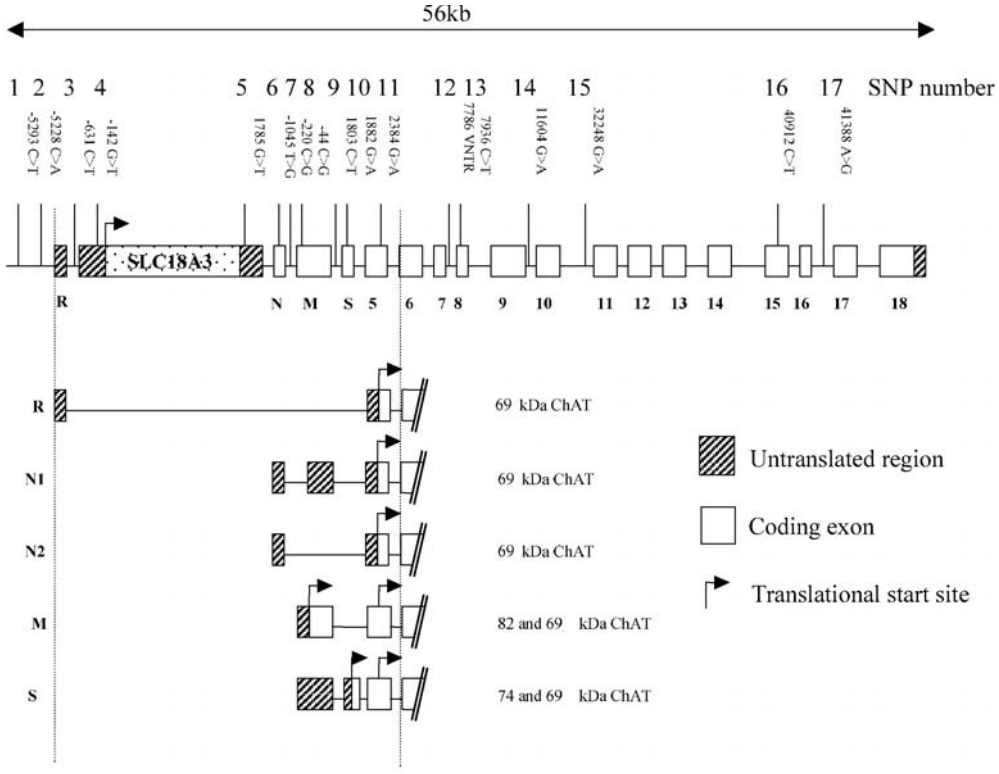

Fig. 1 Genomic structure of the *CHAT* locus showing the relationship of the *CHAT* and *SLC18A3* genes and the complex 5' structure of *CHAT* (not to scale). The five alternative ChAT encoding transcripts that are transcribed from the *CHAT* locus are shown: all variants encode the same 69-kDa ChAT protein but the M variant also encodes an 82-kDa protein, and the S variant, a 74-kDa protein (Oda et al. 1992; Touissaint et al. 1992; Chireux et al. 1995; Hahm et al. 1997; Ohno et al. 2001). For *CHAT*, nucleotide numbers start from the translational start site of the 82-kDA splice variant (*M*) of the gene. Similarly, numbering starts from the translational start site of *SLC18A3* for variants detected in this gene

ferase (ChAT) and the vesicular acetylcholine transporter (VAChT) are proteins specifically required for cholinergic neurotransmission: ChAT catalyses the biosynthesis of the neurotransmitter acetylcholine (ACh) and VAChT is responsible for the translocation of cytoplasmic ACh into synaptic vesicles. The activity of ChAT has been shown to be reduced by 50%–90% in AD patients compared with age-matched controls (Perry et al. 1978; Davies 1979) and is correlated with the depth of dementia (Wilcock et al. 1982; Bierer et al. 1995). The relatively few effective drug treatments that slow the progressive cognitive deterioration in AD have generally targeted enhancement of the cholinergic system via cholinesterase inhibitors, such as donepezil and galanthamine, which indirectly increase the synaptic concentration of ACh (Nordberg and Svensson 1998). Thus, both ChAT and VAChT are obvious functional candidates for involvement in AD.

The genes encoding ChAT (*CHAT*) and VAChT (*SLC18A3*) map to a single locus at 49.7 Mb on chromosome 10; *SLC18A3* lies within the first intron of *CHAT* (Erickson et al. 1994: Fig. 1). In a two-stage genome screen, Myers et al. (2000) found strong evidence for linkage in a region spanning approximately 44 cM from D10S1426 to D10S2327 including the *CHAT* locus. Bertram et al. (2000) also found a susceptibility locus for LOAD in this region of chromosome 10 and Ertekin-Taner et al. (2000) detected linkage by using plasma β-amyloid levels as a quantitative trait locus with a peak in the same region as that for LOAD reported by Myers et al. (2000). In addition, Mubumbila et al. (2002) have recently reported an association between a non-synonymous single-nucleotide polymorphism (SNP) in *CHAT* and LOAD. The *CHAT* locus is probably the most obvious functional candidate gene for AD in this region of linkage on chromosome 10. In order to test whether allelic variation in *CHAT* and *SLC18A3* confers susceptibility to LOAD, we have screened the genes for sequence variants and examined these in three LOAD case-control samples from the UK.

Materials and methods

Case-control samples

For the association study, the first sample comprising 131 UK Caucasian AD patients (age at onset: 72.5±6.5 years) and 118 age- and sex-matched controls (age at collection: 78.2±7.2 years) was used (UK1). Two additional samples were used for replication purposes; a UK Caucasian sample (UK2), consisting of 135 AD patients (age at onset: 74.3±6.2 years) and 135 age- and sex-matched controls (age at collection: 75.3±6.4 years), and a Northern Irish Caucasian sample (UK3), consisting of 242 AD patients (age at onset: 75.4±8.1 years) and 235 age- and sex-matched controls (age at collection: 76.7±8.4 years). All patients were diagnosed with probable AD according to NINCDS-ADRDA criteria (McKhann et al. 1984). Cognitive function of controls was assessed by using the mini mental-state examination (Folstein et al. 1975) and only those with a score of at least 28 were included in the study. None of the samples included individuals used in detecting the original linkage finding.

The mutation screening sample consisted of 14 UK Caucasian AD patients with age at onset of more than 65 years. The sample size of 14 subjects screened across all exons gives a power of 0.8 for detecting alleles with a frequency of 0.05 or above. This estimate ignores the fact that our sample is enriched for AD-susceptibility alleles and the true power is correspondingly (but unquantifiably) greater. We also examined SNP1 5293C→T that we identified from the SNP consortium database (http://snp.cshl.org/) as this lies just upstream of the most 5' fragment that we screened. High molecu-

lar weight genomic DNA was extracted from whole blood or transformed lymphoblasts following standard laboratory procedures.

Mutation detection

CHAT encodes five splice variants; M, N1, N2, R and S. *SLC18A3* is located within the first intron of the R variant of *CHAT* and is itself uninterrupted by introns (Fig. 1). The cDNA sequence for each splice variant/gene was obtained from the GenBank database at NCBI (http://www.ncbi.nlm.nih.gov). Determination of coding sequences, untranslated regions (UTRs) and intronic regions was based on alignment of the cDNA sequences with genomic clone sequences, by using BLAST sequence homology searches (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi).

Polymerase chain reaction (PCR) fragments spanning exons, UTRs and limited 5' flanking regions were designed by using Primer 3.0 (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi). Primer sequences and PCR conditions are available on request from the authors or at http://www.uwcm.ac.uk/study/medicine/psychological_medicine/pub_data/chat.htm. PCR amplification was performed under standard conditions of 1× PCR buffer (Qiagen), 1.5 mM $MgCl_2$, 250 µM dNTPs, 0.5 µM each primer, 0.6 U Hot Start *Taq* (Qiagen) and 48 ng genomic DNA in a 24-µl reaction. Cycling was conducted in an MJ Tetrad (MJ Research) with an initial denaturation of 94°C for 15 min, followed by 35 cycles of 94°C for 30 s, appropriate annealing temperature for 30 s and 72°C for 45 s with a final extension step of 72°C for 10 min. Synthesis of appropriately sized PCR products was confirmed by electrophoresis on 2% agarose gels.

Polymorphisms were identified by denaturing high-pressure liquid chromatography (DHPLC) on a Wave DNA Fragment Analysis System (Transgenomic) as previously described (Abraham et al. 2001). The 14 screening samples were amplified as described above, except the final extension in the PCR protocol was followed by denaturation at 94°C for 5 min and then cooling to 65°C over 30 min to allow heteroduplex formation. Column temperature and acetonitrile gradient were determined by using the DHPLC Melt program (http://insertion.stanford.edu/melt1.html). To ensure maximum sensitivity, in addition to the temperature suggested by the software (*n*°C), each fragment was also run at *n*+2°C. Samples showing heteroduplex formation were sequenced to identify the variant.

PCR products were purified through QIAquick columns (Qiagen) to remove unincorporated primers and dNTPs. Purified products were then bidirectionally sequenced on an ABI 377 DNA Sequencer (Applied Biosystems) with the Big Dye Terminator (v2.0) Cycle Sequencing kit (Applied Biosystems). Sequence traces were subsequently exported to Sequencher (Applied Biosystems) to characterise polymorphisms.

Genotyping

Where a natural restriction site existed that could distinguish between the two alleles of an SNP, a restriction fragment length polymorphism (RFLP) assay was devised with the original PCR primers designed for DHPLC. In some cases in which no natural restriction site existed, an artificial restriction site could be created by primer-generated mutagenesis. For each RFLP assay, the UK1 association sample was PCR-amplified and then digested with 5 U appropriate restriction enzyme (Table 1). Digested products were electrophoresed on 2.5% agarose gels.

The remaining five SNPs were typed in a multiplex primer extension assay. Extension primers were designed to be 17, 27, 37, or 47 nucleotides long and directly adjacent to the polymorphism. For each SNP, the UK1 association sample was PCR-amplified and then purified by incubation with 1 U exonuclease I and shrimp alkaline phosphatase at 37°C for 1 h. Primer extension was then performed by using the ABI SNaPshot ddNTP Primer Extension kit on the ABI 3100 Genetic Analyser. Data were analysed by Genotyper 2.5.

Table 1 Identity and assay type for sequence variants at the CHAT locus. Details of the seventeen sequence variants detected at the *CHAT* locus. All sequence numbering for *CHAT* is from the first translational start site in exon M. Amino acid numbering is from the first amino acid of the large 82-kDa isoform of CHAT (with the exception of the SNP in exon S, where numbering is from the first amino acid of the 74-kDa isoform of CHAT). Sequence numbering for *SLC18A3* is from the translational start site. SNP database numbers are given for SNPs identified previously or during the course of this study. The restriction enzymes used in RFLP are given (*PE* primer extension, *RFLP* restriction fragment length polymorphism)

Order on chromosome 10	Variant	Location	Database identity	Assay (restriction enzyme)
	CHAT			
1	–5293C→T	198 bp upstream exon R	rs733722	RFLP (*Rsa*I)
2	–5228C→A	133 bp upstream exon R		RFLP (*Nla*IV)
6	–1045T→G	Exon N (5'UTR)		RFLP[a] (*Ava*I)
7	–220C→G	40 bp upstream exon M		PE
8	–44C→G	Exon M (5'UTR)		RFLP (*Ava*I)
9	1803C→T	31 bp upstream exon S		RFLP[a] (*Mlu*I)
10	1882G→A	Exon S (D7N)	rs1880676	RFLP[a] (*Mbo*I)
11	2384G→A	Exon 5 (A120T)	Mubumbila et al. 2002	RFLP (*Bbv*I)
12	7786 VNTR	57–122 bp upstream exon 8		PCR sizing
13	7936C→T	Exon 8 (L243F)		PE
14	11604G→A	140 bp downstream exon 9	rs868750	PE
15	32248G→A	68 bp upstream exon 11	rs2377871	RFLP (*Ban*II)
16	40912C→T	Exon 15 (H547H)		PE
17	41388A→G	129bpdownstream exon 16		RFLP (*Fnu*4HI)
	SLC18A3			
3	–631C→T	336 bp upstream VAChT	rs885835	RFLP[a] (*Dra*II)
4	–142G→T	5'UTR	rs2377879	RFLP (*Bsi*HKA I)
5	1785G→T	3'UTR	rs2269338	PE

[a]An artificial restriction site was created by primer mutagenesis

Table 2 Genotypic and allelic LOAD association for the 17 sequence variants detected in *CHAT* and *SLC18A3* in the UK1 sample (*P*-values in *bold* are those that gave evidence of possible association and that were carried forward for further analysis)

	Variant	Genotype					*P*	Allele				*P*
CHAT												
1	–5293C→T	C/C		C/T		T/T	0.90	C		T		0.64
	Patients	88		35		7		211 (0.81)		49 (0.19)		
	Controls	76		34		7		186 (0.79)		48 (0.21)		
2	–5228C→A	C/C		C/A		A/A	0.66	C		A		0.66
	Patients	94		4		0		192 (0.98)		4 (0.02)		
	Controls	99		3		0		201 (0.99)		3 (0.01)		
6	–1045T→G	T/T		T/G		G/G	0.35	T		G		0.16
	Patients	75		41		6		191 (0.78)		53 (0.22)		
	Controls	59		39		10		157 (0.73)		59 (0.27)		
7	–220C→G	C/C		C/G		G/G	0.48	C		G		0.36
	Patients	95		21		2		211 (0.89)		25 (0.11)		
	Controls	95		14		2		204 (0.92)		18 (0.08)		
8	–44C→G	C/C		C/G		G/G	0.37	C		G		0.18
	Patients	80		43		6		203 (0.79)		55 (0.21)		
	Controls	65		42		10		172 (0.74)		62 (0.26)		
9	1803C→T	C/C		C/T		T/T	0.82	C		T		0.73
	Patients	50		17		1		117 (0.86)		19 (0.14)		
	Controls	67		18		2		152 (0.87)		22 (0.13)		
10	1882G→A	G/G		G/A		A/A	**0.08**	G		A		**0.09**
	Patients	34		25		9		93 (0.68)		43 (0.32)		
	Controls	49		33		3		131 (0.77)		39 (0.23)		
11	2384G→A	G/G		G/A		A/A	0.58	G		A		0.47
	Patients	69		51		11		189 (0.72)		73 (0.28)		
	Controls	65		47		6		177 (0.75)		59 (0.25)		
12	7786VNTR	1/2	2/2	2/3	2/4	3/3	0.81	1	2	3	4	0.94
	Patients	4	103	14	1	1		4	225	16	1	
	Controls	4	82	9	0	2		2	175	13	0	
13	7936C→T	C/C		C/T		T/T	0.35	C		T		0.40
	Patients	106		12		0		224 (0.95)		12 (0.05)		
	Controls	97		16		0		210 (0.93)		16 (0.07)		
14	11604G→A	G/G		G/A		A/A	**0.07**	G		A		**0.03**
	Patients	72		39		8		183 (0.77)		55 (0.23)		
	Controls	83		31		2		197 (0.85)		35 (0.15)		
15	32248G→A	G/G		G/A		A/A	0.30	G		A		0.33
	Patients	114		10		2[a]		238 (0.94)		14 (0.06)		
	Controls	99		16		1		214 (0.92)		18 (0.08)		
16	40912C→T	C/C		C/T		T/T	0.43	C		T		0.18
	Patients	101		14		1		216 (0.93)		16 (0.07)		
	Controls	89		19		2		197 (0.90)		23 (0.10)		
17	41388A→G	A/A		A/G		G/G	0.32	A		G		0.36
	Patients	39		56		35		134 (0.52)		126 (0.48)		
	Controls	26		60		32		112 (0.47)		124 (0.53)		
SLC18A3												
3	–631C→T	C/C		C/T		T/T	0.78	C		T		0.92
	Patients	70		19		3		159 (0.86)		25 (0.14)		
	Controls	72		23		2		167 (0.86)		27 (0.14)		
4	–142G→T	G/G		G/T		T/T	0.39	G		T		0.31
	Patients	121		8		0		250 (0.97)		8 (0.03)		
	Controls	102		9		1		213 (0.95)		11 (0.05)		
5	1785G→T	G/G		G/T		T/T	0.99	G		T		0.91
	Patients	91		27		1		209 (0.88)		29 (0.12)		
	Controls	88		25		1		201 (0.88)		27 (0.12)		

[a]Genotype not in HWE

The variable number of tandem repeats (VNTR) was genotyped by PCR amplification followed by visualisation on a 1.5% agarose gel.

Statistical analysis

All polymorphisms were tested for deviation from the Hardy-Weinberg equilibrium (HWE) independently in each population. χ^2 and Fisher's exact test were used to analyse SNP associations by using the Simple Interactive Statistical Analysis pages at http://home.clara.net/sisa: Fisher's exact test was used for analyses where one or more cell had a count of less than 5. The VNTR (polymorphism 12) was tested for association with LOAD by using CLUMP (Sham and Curtis 1995). Haplotypic association was tested by using EHPLUS (Xie and Ott 1993) with PMPLUS (Zhao et al. 2000) implemented to obtain empirical significance levels (Terwilliger and Ott 1994). Marker-marker linkage disequilibrium (LD) analyses were also undertaken by using HAPMAX. This program employs an EM algorithm to allow for phase unknowns. Estimated haplotype frequencies were used to calculate D' and r^2 estimates of LD. Meta-analysis was carried out by means of an inverse variance method of weighting in a fixed effect model. Heterogeneity between studies was assessed with a χ^2 test (Cooper and Hedges 1994).

Results

Polymorphisms identified

A total of 17 polymorphisms were identified in *CHAT* and *SLC18A3* (Fig. 1, Table 1), including one SNP from the SNP consortium database (http://snp.cshl.org/). All but one of the sequence variants were SNPs. Three were non-synonymous coding SNPs; an D7 N polymorphism in the S exon of *CHAT* (1882G→A), an A120T polymorphism in exon 5 of *CHAT* (2384G→A) and a L243F polymorphism in exon 8 of *CHAT* (7936C→T); one further exonic SNP was detected that did not change the protein sequence (40912C→T, H547H). Five SNPs were in the region 5' to translation in *CHAT* and two of these were in the UTRs of the alternatively spliced 5' exons M and N. The remaining four SNPs were intronic (Fig. 1). In addition, a 66-bp VNTR was also identified in intron 7 of *CHAT*, containing one to four copies of the sequence 5'-AAG GGA GGG AAG AGG AAG GAG ATG GAA GGA AGA GGG AAG GAG GGA GGG GAG GCA GAA GGG AGG GAG-3'. Three SNPs were detected in *SLC18A3*: one in the 5' upstream sequence, one in the 5' UTR and one in the 3' UTR.

Association analysis

Unless otherwise indicated, polymorphisms were in HWE. All polymorphisms identified were tested for association with LOAD in the UK1 sample of 131 cases and 118 age-matched controls (Table 2). As the sample size was relatively small, it was decided *a priori* to genotype any SNP displaying an association with $P \leq 0.10$ in the replication samples; two SNPs fulfilled these criteria: 1882G→A and 11604G→A.

For the SNP 11604G→A in intron 9 of the *CHAT* gene, the A allele appeared to be more common in LOAD patients than in controls ($P=0.034$). This SNP was genotyped in the two additional case-control samples, UK2 and UK3 (Table 3). In neither the UK2 ($P=0.181$) nor UK3 ($P=0.494$) samples was a significant difference found between allele distributions in patients and controls, suggesting that the initial positive finding is a type I error.

For the non-synonymous SNP 1882G→A in the S exon of the *CHAT* gene, there was a trend to an excess of AA homozygotes in the LOAD patients compared with controls in the UK1 sample (Table 1; $P=0.083$). The UK1 sample showed that 1882G→A and 2384G→A were in complete LD ($r^2=1$), although fewer genotypes were obtained for 1882G→A than for 2384G→A because of depletion of DNA stocks from the UK1 sample; this accounts for the discrepancy in the genotype number in Table 2. However, complete LD was also observed between these two SNPs (see below) in the UK2 sample and we therefore typed only 1882G→A in the UK3 sample. Table 4 shows that, in the UK2 sample, there was no evidence of association but some evidence was present in the UK3 sample ($P=0.032$ genotypic and 0.064 allelic). In both the UK1 and UK3 samples, there were significantly more AA homozygotes in patients than in controls ($P=0.016$) but, in the UK3 sample, this was partly attributable to the small number of AA genotypes observed in the control sample, which was out of HWE ($P=0.015$). This trend to increased AA genotypes in cases was not apparent at all in the UK2 sample. Combining all three UK samples in a meta-analysis revealed no genotypic (G/– vs A/A) or allelic association by using 2384G→A in UK1 and 1882G→A in UK2 and UK3 ($P=0.114$ genotypic and 0.373 allelic). There was no significant heterogeneity between studies ($\chi^2=3.95$, df=2, $P=0.139$ for G/– vs A/A genotypes; $\chi^2=3.49$, df=2, $P=0.175$ for G vs A).

We also genotyped the non-synonymous SNP 7936C→T in order to investigate whether there was any interaction between the amino acid changes in the two additional as-

Table 3 Association of the 11604G→A polymorphism in CHAT intron 9 with LOAD

11604G→A (CHAT13)		Genotype			*P*	Allele		*P*
		GG	GA	AA		G	A	
UK2	Patients	88	42	5	0.41	218 (0.81)	52 (0.19)	0.18
	Controls	95	33	3		223 (0.85)	39 (0.15)	
UK3	Patients	129	75	5	0.67	333 (0.80)	85 (0.20)	0.49
	Controls	130	84	8		344 (0.77)	100 (0.23)	

Table 4 Analysis of the coding polymorphisms, CHAT 10 and 13

Sample	Polymorphism		Genotype			*P genotypic*	PG/– against A/A	Allele		*P allelic*
UK2	10	1882G→A	G/G	G/A	A/A	0.70	0.80	G	A	0.44
		Patients	71	56	8			198 (0.73)	72 (0.27)	
		Controls	64	62	9			190 (0.70)	80 (0.30)	
UK3	10	1882G→A	G/G	G/A	A/A	0.03	0.01	G	A	0.06
		Patients	105	77	12			287 (0.74)	101 (0.26)	
		Controls	127	79	3[a]			333 (0.80)	85 (0.20)	
UK2	13	7936C→T	C/C	C/T	T/T	0.35	N/A	C	T	0.22
		Patients	116	18	1			250 (0.93)	20 (0.07)	
		Controls	120	13	0			253 (0.95)	13 (0.05)	
UK3	13	7936C→T	C/C	C/T	T/T	0.40	N/A	C	T	0.25
		Patients	156	25	0			337 (0.93)	25 (0.07)	
		Controls	161	34	1			356 (0.91)	36 (0.09)	

[a]Not in HWE (P=0.015)

Table 5 Genotypes and analysis for assocation with LOAD in the UK2 case-control sample for replication of UK1 haplotype association analysis

Polymorphism		Genotype			*P*	Allele (frequency)		*P*
1	–5293C→T	C/C	C/T	T/T	0.57	C	T	0.44
	Patients	65	46	6		176 (0.75)	58 (0.25)	
	Controls	69	45	3		183 (0.78)	51 (0.22)	
3	–631C→	C/C	C/T	T/T	0.68	C	T	0.66
	Patients	99	30	3		228 (0.86)	36 (0.14)	
	Controls	93	35	2		221 (0.85)	39 (0.15)	
5	1785G→T	G/G	G/T	T/T	1.00	G	T	0.97
	Patients	74	22	1		170 (0.88)	24 (0.12)	
	Controls	81	24	1		186 (0.88)	26 (0.12)	
6	–1045T→G	T/T	T/G	G/G	0.29	T	G	0.57
	Patients	72	31	9[a]		175 (0.78)	49 (0.22)	
	Controls	73	37	4		183 (0.80)	45 (0.20)	
7	–220C→G	C/C	C/G	G/G	0.31	C	G	0.23
	Patients	103	31	1		237 (0.88)	33 (0.12)	
	Controls	109	24	0		242 (0.91)	24 (0.09)	
8	–44C→G	C/C	C/G	G/G	0.34	C	G	0.55
	Patients	76	32	9[a]		184 (0.79)	50 (0.21)	
	Controls	75	36	4		186 (0.81)	44 (0.19)	
16	40912C→T	C/C	C/T	T/T	0.46	C	T	0.41
	Patients	107	25	2		239 (0.89)	29 (0.11)	
	Controls	109	23	0		241 (0.91)	23 (0.09)	

[a]Not in HWE: there is high LD between SNPs 6 and 8 (see Tables 8 and 9)

sociation samples (Table 4). No significant difference was observed in genotypic or allelic frequencies between cases and controls in any sample. When combined with the data from the 1882G→A and 2384G→A SNPs to examine the possibility of interaction between the three rare amino acid changes, no association with LOAD was seen (UK1, P=0.568; UK2, P=0.985; UK3, P=0.514)).

Haplotype association

Analysis of haplotype association with LOAD was performed for all possible two-marker haplotypes by using genotyping data generated from the UK1 sample for the 16 SNPs (Table 5). Single SNP associations that became more significant when combined in haplotypes in the UK1 sample were analysed in the UK2 sample. The other SNPs for which there was individual genotype data from the UK2 sample were also analysed in two-marker haplotypes. Several associations were identified but appeared to reflect the false-positive association of 11604G→A with LOAD in the UK1 sample (Table 6), and none of these were replicated in the UK2 sample (Table 7). Apart from the 11604G→A combinations, –631C→T and 40912C→T displayed a significant two-marker haplotype association (P=0.037) in the UK1 sample, although neither were significant by themselves. Again, this was not replicated in the UK2 sample (P=0.336).

Table 6 Two-marker SNP haplotype associations in the UK1 sample. Results where $P<0.05$ are given in *italics*

		–5228 C→A 2	–631 C→T 3	–142 G→T 4	1785 G→T 5	–1045 T→G 6	–220 C→G 7	-44 C→G 8	1803 C→T 9	1882 G→A 10	2384 G→A 11	7936 C→T 13	11604 G→A 14	32248 G→A 15	40912 C→T 16	41388 A→G 17
–5293C→T	1	0.704	0.346	0.740	0.912	0.386	0.833	0.628	0.908	0.540	0.902	0.295	*0.042*	0.517	0.313	0.846
–5228C→A	2		0.972	0.438	0.889	0.097	0.694	0.410	0.930	0.126	0.360	0.255	0.404	0.515	0.096	0.824
–631C→T	3			0.428	0.992	0.704	0.771	0.737	0.971	0.685	0.983	0.974	0.642	0.560	*0.037*	0.409
–142G→T	4				0.649	0.353	0.234	0.334	0.444	0.108	0.590	0.270	0.066	0.679	0.426	0.268
1785G→T	5					0.085	0.533	0.069	0.617	0.247	0.563	0.525	*0.011*	0.578	0.198	0.242
–1045T→G	6						0.081	0.400	0.113	0.136	0.433	0.380	*0.033*	0.442	0.221	0.352
–220C→G	7							0.061	0.913	0.413	0.488	0.191	*0.018*	0.399	0.084	0.468
–44C→G	8								0.157	0.109	0.429	0.386	*0.039*	0.373	0.320	0.387
1803C→T	9									0.215	0.247	0.066	0.344	0.963	0.542	0.565
1882G→A	10										0.113	0.197	0.216	0.206	0.550	0.207
2384G→A	11											0.532	*0.034*	0.597	0.481	0.700
7936C→T	13												0.108	0.133	0.188	0.558
11604G→A	14													0.082	0.077	0.096
32248G→A	15														0.287	0.513
40912C→T	16															0.179

Table 7 Two-marker SNP haplotype associations in the UK2 sample

		–631 C→T 3	1785 G→T 5	–1045 T→G 6	–220 C→G 7	–44 C→G 8	1882 G→A 10	7936 C→T 13	11604 G→A 14	40912 C→T 16
–5293C→T	1	0.595	0.714	0.654	0.696	0.630	0.489	0.507	0.521	0.563
–631C→T	3		0.897	0.616	0.416	0.690	0.625	0.183	0.397	0.329
1785G→T	5			0.478	0.154	0.267	0.168	0.739	0.581	0.215
–1045T→G	6				0.443	0.426	0.543	0.291	0.632	0.389
–220C→G	7					0.518	0.339	0.349	0.084	0.421
–44C→G	8						0.420	0.282	0.595	0.459
1882G→A	10							0.350	0.640	0.561
7936C→T	13								0.388	0.501
11604G→A	14									0.466

Table 8 R^2 for CHAT cases. Results where $P<0.05$ are given in *italics*

		–5228 C→A 2	–631 C→T 3	–142 G→T 4	1785 G→T 5	–1045 T→G 6	–220 C→G 7	–44 C→G 8	1803 C→T 9	1882 G→A 10	2384 G→A 11	7936 C→T 13	11604 G→A 14	32248 G→A 15	40912 C→T 16	41388 A→G 17
–5293C→T	1	0.002	0.003	0.001	0.003	0.006	0.003	0.006	0.003	0.038	0.009	0.001	0.090	0.001	0.007	0.003
–5228C→A	2		0.029	0.151	0.004	0.003	0.009	0.000	0.014	0.000	0.000	0.000	0.060	0.014	0.040	0.001
–631C→T	3			0.002	0.009	0.051	0.003	0.049	0.000	0.035	0.023	0.010	0.021	0.004	0.011	0.007
–142G→T	4				0.003	0.127	0.003	0.120	0.006	0.016	0.012	0.001	0.004	0.029	0.004	0.029
1785G→T	5					*0.375*	*0.724*	*0.371*	*0.716*	0.060	0.052	0.010	0.041	0.037	0.002	0.000
–1045T→G	6						*0.440*	*0.976*	*0.396*	0.120	0.109	0.188	0.047	0.007	0.001	0.001
–220C→G	7							*0.402*	*0.719*	0.048	0.043	0.016	0.037	0.006	0.008	0.001
–44C→G	8								*0.420*	0.114	0.105	0.194	0.043	0.007	0.001	0.001
1803C→T	9									0.006	0.009	0.102	0.048	0.013	0.018	0.009
1882G→A	10										*1.000*	0.010	0.083	0.033	0.018	0.000
2384G→A	11											0.020	0.087	0.001	0.014	0.000
7936C→T	13												0.000	0.007	0.003	0.028
11604G→A	14													0.015	0.020	0.030
32248G→A	15														*0.352*	0.013
40912C→T	16															0.070

Table 9 R^2 for CHAT controls. Results where $P<0.05$ are given in *italics*

		–5228 C→A 2	–631 C→T 3	–142 G→T 4	1785 G→T 5	–1045 T→G 6	–220 C→G 7	–44 C→G 8	1803 C→T 9	1882 G→A 10	2384 G→A 11	7936 C→T 13	11604 G→A 14	32248 G→A 15	40912 C→T 16	41388 A→G 17
–5293C→T	1	0.005	0.001	0.001	0.003	0.033	0.002	0.004	0.004	0.009	0.006	0.072	0.033	0.002	0.002	0.017
–5228C→A	2		0.056	0.114	0.000	0.000	0.001	0.000	0.014	0.000	0.000	0.021	0.001	0.000	0.003	0.000
–631C→T	3			0.011	0.025	0.066	0.017	0.067	0.017	0.032	0.010	0.016	0.030	0.051	0.098	0.069
–142G→T	4				0.000	0.163	0.001	0.135	0.010	0.025	0.017	0.004	0.008	0.025	0.001	0.000
1785G→T	5					0.091	*0.421*	0.094	*0.365*	0.011	0.005	0.000	0.050	0.018	0.053	0.048
–1045T→G	6						0.212	*0.943*	0.186	0.146	0.125	0.225	0.000	0.002	0.017	0.025
–220C→G	7							0.219	*0.745*	0.034	0.026	0.005	0.013	0.031	0.034	0.004
–44C→G	8								0.178	0.145	0.123	0.179	0.000	0.002	0.003	0.009
1803C→T	9									0.045	0.041	0.012	0.009	0.028	0.006	0.000
1882G→A	10										*1.000*	0.022	0.034	0.002	0.001	0.000
2384G→A	11											0.028	0.015	0.002	0.001	0.002
7936C→T	13												0.003	0.007	0.009	0.075
11604G→A	14													0.013	0.018	0.005
32248G→A	15														0.201	0.012
40912C→T	16															0.128

Linkage disequilibrium

Levels of LD across the 56-kb *CHAT* locus were generally low (Tables 8 and 9). LD analysis was performed separately for the UK1 case and control samples and the results were consistent between the two samples. Only two SNPs proved to be in complete LD in the UK1 sample and these were the coding SNPs, 1882G→A and 2384G→A, which are 500 bp apart; we detected no other SNPs between these two SNPs. High LD was also observed between –44C→G and –1045T→G, which are 100 bp apart. These two SNPs lie within the only region of any conserved LD in this gene, viz. that between SNPs –1045T→G and 7936C→T but, even in this region, levels of LD were generally low. Similar results were obtained for the D' estimation of LD (data not shown).

Discussion

We detected 16 and analysed 17 sequence variants within the ChAT/VAChT complex of genes. Of the 16 variants we detected six had been or were subsequently reported in SNP databases and one of the coding SNPs was independently reported by Mubumila et al. (2002; Table 1). Previous to this report, the only polymorphisms identified at the *CHAT* locus were the database SNPs and a series of rare mutations causing a recessive congenital myasthenic syndrome (Ohno et al. 2001). Three of the SNPs reported in the present study cause coding changes in *CHAT* but any effect of these coding changes on ChAT activity is unknown.

There proved to be no interaction between the coding SNPs in a situation analogous to that of *APOE* genotype. In our samples, there was no such association. The marginal evidence for association probably arises partly because of the small numbers of minor allele homozygotes in the control samples in UK1 and UK3 and because, in UK3, the control genotypes are not in HWE. Repeated genotyping of 1882G→A in this sample gave the same result, suggesting the deviation from HWE is not attributable to laboratory error. Although there are several explanations for the departure from HWE (including chance), the finding in our study might possibly reflect selection, as our control samples are all older than the case samples and the UK1 and UK3 controls are the two cohorts with the greatest average age in our study.

Mubumila et al. (2002) have found a highly significant association of the 2384G→A SNP with AD in their sample of 122 LOAD cases and 112 controls collected in France and Germany ($P<0.0005$ allelic and genotypic). Their genotype frequencies are substantially different from those observed in any of our samples, with much higher numbers of minor allele homozygotes in both cases (34.4% vs 6.7%) and controls (12.5% vs 3.9%), although this is a similar trend to that observed in UK1 and UK3, with increased numbers of minor allele homozygotes in the patients compared with controls. Both the case and control genotypes in their analysis are markedly out of HWE (cases: $P=0.0000$; controls: $P=0.0110$). We cannot determine the reasons for this but it may reflect age effects or the combination of two populations (French/German), each with different genotype frequencies or genotyping error. The significant association observed in this sample might therefore be an artifact. If we assume that the study of Mubumila et al. (2002) represents a true association, our sample has a power of greater than 0.99 to detect such an association. It is therefore surprising that we found no association in our total sample of 460 LOAD cases and 462 controls.

The two-marker haplotype association analysis revealed some significant results, most of which arose in combination with the 11604G→A SNP, which was the only significant SNP in the UK1 case control sample at the $P<0.05$ level. None of the significant results was replicated in the second sample (Tables 6 and 7) or in the combined samples, which suggests that the apparent associations were the result of chance. Thus, we conclude there is no haplotype association of the *CHAT* or *SLC18A3* genes with LOAD.

Although the *CHAT* locus is an obvious candidate as a locus functionally and positionally implicated in AD, the detailed genetic study presented here indicates that variations in *CHAT* and *SLC18A3* are unlikely to be involved in the primary pathogenesis of LOAD. The *CHAT* locus has a complicated structure and the variants that we have detected could possibly affect the regulation of these genes but this in itself may not contribute to the genetic susceptibility to LOAD. Moreover, rare *CHAT* locus polymorphisms may exist that affect susceptibility to AD but the effects on ChAT activity might be expected to be subtle, given that the abolition or reduction of activity of the ChAT enzyme is reported to cause myasthenic syndromes rather than dementia (Ohno et al. 2001). This might be related to the production of alternative transcripts in different populations of neurones, in that rare polymorphisms affecting particular splice variants of *CHAT* expressed in brain might be relevant to LOAD. However, such rare polymorphisms would not account for LOAD in more than a small proportion of the population and could not therefore account for our linkage data on chromosome 10 (Myers et al. 2000). As a final caveat, the weak LD across the gene means that we cannot exclude the possibility that there are variants (rare or common) associated with AD in regulatory elements outside the regions of the gene that we have screened and that are not in strong LD with any of the SNPs that we have genotyped. The detection of such alleles, particularly when they are of low frequency, poses formidable challenges for molecular genetic studies.

The cholinesterase inhibitors that are effective in treating early symptoms of AD increase the half-life of acetylcholine (ACh) in the synaptic cleft but a number of other gene products operate in the pathway that culminates in the release of ACh, including those that control the electrical depolarisation of the cholinergic neurone and those that control synaptic vesicle turnover in the synapse. It is also likely that the reported defects in cholinergic transmission

in AD are a secondary change downstream of the primary genetic and molecular events in the disease. Variants of genes involved in these other pathways may be important in LOAD. However, as the response to cholinesterase treatment is known to be variable and unpredictable in AD (Frances et al. 1999), it would be interesting to investigate further whether any of these sequence variants in the *CHAT* locus are associated with cholinesterase response in AD patients.

Acknowledgements We thank the Medical Research Council, UK (S.J., P.H., P.M., J.W., M.O.D., M.L., M.O. and L.J.), the Mayo Foundation and NIH/NIA (AG06786, AG16574: R.C.P. and J.H.; AG16208: A.M.G. and J.H.; AG05681: A.M.G.). F.O.Y., S.L., M.O.D., J.W., M.J.O. and L.J. are members of the Alzheimer's Research Trust's AD Research Centre Network.

References

Abraham R, Myers A, Warrant DeVrieze F, Hamshere ML, Thomas HV, Marshall H, Compton D, Spurlock G, Turic D, Hoogendoorn B, Kwon JM, Petersen RC, Tanaglos E, Norton J, Morris JC, Bullock R, Liolitsa D, Lovestone S, Hardy J, Goate A, O'Donovan M, Williams J, Owen MJ, Jones L (2001) Substantial linkage disequilibrium across the insulin degrading enzyme locus but no association with late-onset Alzheimer's disease. Hum Genet 109:646–652

Bertram L, Blacker D, Mullin K, Keeney D, Jones J, Basu S, Yhu S, McInnis MG, Go RC, Vekrellis K, Selkoe DJ, Saunders AJ, Tanzi RE (2000) Evidence for genetic linkage of Alzheimer's disease to chromosome 10q. Science 290:2302–2303

Bierer LM, Haroutunian V, Gabriel S, Knott PJ, Carlin LS, Purohit DP, Perl DP, Schmeidler J, Kanof P, Davis KL (1995) Neurochemical correlates of dementia severity in Alzheimer's disease: relative importance of the cholinergic deficits. J Neurochem 64:749–760

Bowen DM, Smith CB, White P, Davison AN (1976) Neurotransmitter-related enzymes and indices of hypoxia in senile dementia and other abiotrophies. Brain 99:459–496

Chireux MA, Le Van Thai A, Weber MJ (1995) Human choline acetyltransferase gene: localization of alternative first exons. J Neurosci Res 40:427–438

Cooper HM, Hedges LV (1994) The handbook of research synthesis. Russel Sage, New York

Davies P (1979) Neurotransmitter-related enzymes in senile dementia of the Alzheimer type. Brain Res 171:319–327

Davies P, Maloney AJ (1976) Selective loss of central cholinergic neurons in Alzheimer's disease. Lancet II:1403

Daw Warwick E, Payami H, Nemens EJ, Nochlin D, Bird TD, Schellenberg GD, Wijsman EM (2000) The number of trait loci in late-onset Alzheimer disease. Am J Hum Genet 66:196–204

Erickson JD, Varoqui H, Schafer MK-H, Modi W, Diebler M-F, Weihe E, Rand J, Eiden LE, Bonner TI, UsdinTB (1994) Functional identification of a vesicular acetylcholine transporter and its expression from a "cholinergic" gene locus. J Biol Chem 269:21929–21932

Ertekin-Taner N, Graff-Radford N, Younkin LH, Eckman C, Baker M, Adamson J, Ronald J, Blangero J, Hutton M, Younkin SG (2000) Linkage of plasma Abeta42 to a quantitative locus on chromosome 10 in late-onset Alzheimer's disease pedigrees. Science 290:2303–2304

Folstein MF, Folstein SE, McHugh PR (1975) Mini-mental state: a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 12:189–198

Francis PT, Palmer AM, Snape M, Wilcock GK (1999) The cholinergic hypothesis of Alzheimer's disease: a review of progress. J Neurol Neurosurg Psychiatry 66:137–147

Hahm SH, Chen L, Patel C, Erickson J, Bonner TI, Weihe E, Schafer MK, Eiden LE (1997) Upstream sequencing and functional characterization of the human cholinergic gene locus. J Mol Neurosci 9:223–236

Hardy J (1997) Amyloid, the presenilins and Alzheimer's disease. Trends Neurosci 20:154–159

McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM (1984) Clinical diagnosis of Alzheimer disease: report of the NINCDS-ADRDA Work Group under the auspices of the Department of Health and Human Services Task Force on Alzheimer Disease. Neurology 23:939–944

Mubumbila V, Sutter A, Ptok U, Heun R, Quirin-Stricker C (2002) Identification of a single nucleotide polymorphism in the choline acetyltransferase gene associated with Alzheimer's disease. Neurosci Lett 333:9–12

Myers AJ, Goate AM (2001) The genetics of late-onset Alzheimer's disease. Curr Opin Neurol 14:433–440

Myers A, Holmans P, Marshall H, Kwon J, Meyer D, Ramic D, Shears S, Booth J, DeVrieze FW, Crook R, Hamshere M, Abraham R, Tunstall N, Rice F, Carty S, Lillystone S, Kehoe P, Rudrasingham V, Jones L, Lovestone S, Perez-Tur J, Williams J, Owen MJ, Hardy J, Goate AM (2000) Susceptibility locus for Alzheimer's disease on chromosome 10. Science 290:2304–2305

Nordberg A, Svensson AL (1998) Cholinesterase inhibitors in the treatment of Alzheimer's disease: a comparison of tolerability and pharmacology. Drug Saf 19:465–480

Oda Y, Nakanishi I, Deguchi T (1992) A complementary DNA for human choline acetyltransferase induces two forms of enzyme with different molecular weights in cultured cells. Brain Res Mol Brain Res 16:287–294

Ohno K, Tsujino A, Brengman JM, Harper CM, Bajzer Z, Udd B, Beyring R, Robb S, Kirkham FJ, Engel AG (2001) Choline acetyltransferase mutations cause myasthenic syndrome associated with episodic apnea in humans. Proc Natl Acad Sci USA 98:2017–2022

Perry EK, Perry RH, Blessed G, Tomlinson BE (1977) Necropsy evidence of central cholinergic deficits in senile dementia. Lancet I:189

Perry EK, Tomlinson BE, Blessed G, Bergmann K, Gibson PH, Perry RH (1978) Correlation of cholinergic abnormalities with senile plaques and mental test scores in senile dementia. BMJ 2:1457–1459

Sham PC, Curtis D (1995) Monte-Carlo test for association between disease and alleles at highly polymorphic loci. Ann Hum Genet 59:97–105

Terwilliger J, Ott J (1994) Handbook of human genetic linkage. Johns Hopkins University Press, Baltimore

Touissaint JL, Geoffroy V, Schmitt M, Werner A, Garnier JM, Simoni P, Kempf J (1992) Human choline acetyltransferase (CHAT): partial gene sequence and potential control regions. Genomics 12:412–416

Wilcock GK, Esiri MM, Bowen DM, Smith CC (1982) Alzheimer's disease. Correlation of cortical choline acetyltransferase activity with the severity of dementia and histological abnormalities. J Neurol Sci 57:407–417

Whitehouse PJ, Price DL, Struble RG, Clark AW, Coyle JT, Delon MR (1982) Alzheimer's disease and senile dementia: loss of neurons in the basal forebrain. Science 215:1237–1239

Xie X, Ott J (1993) Testing linkage disequilibrium between a disease gene and marker loci. Am J Hum Genet 53:1107

Zhao H, Curtis D, Sham PC (2000) Model-free analysis and permutation tests for allelic associations. Hum Hered 50:133–139