Human Molecular Genetics, 2005, Vol. 14, No. 21 ***3281–3292***
doi:10.1093/hmg/ddi361
Advance Access published on September 29, 2005

High-density SNP haplotyping suggests altered regulation of *tau* gene expression in progressive supranuclear palsy

Rosa Rademakers[1,†], Stacey Melquist[4,†], Marc Cruts[1], Jessie Theuns[1], Jurgen Del-Favero[1], Parvoneh Poorkaj[7,8,9,10], Matt Baker[4], Kristel Sleegers[1], Richard Crook[4], Tim De Pooter[1], Samira Bel Kacem[1], Jennifer Adamson[4], Dirk Van den Bossche[1], Marleen Van den Broeck[1], Jennifer Gass[4], Ellen Corsmit[1], Peter De Rijk[1], Natalie Thomas[5], Sebastiaan Engelborghs[2,3], Michael Heckman[6], Irene Litvan[11], Julia Crook[6], Peter P. De Deyn[2,3], Dennis Dickson[5], Gerard D. Schellenberg[7,8,9,10], Christine Van Broeckhoven[1,*,‡] and Michael L. Hutton[4,‡]

[1]Department of Molecular Genetics, Flanders Interuniversity Institute for Biotechnology, [2]Division of Neurology, Middelheim Hospital, [3]Laboratory of Neurochemistry and Behavior, Institute Born-Bunge, University of Antwerp, Belgium, [4]Department of Neuroscience, [5]Department of Pathology and [6]Biostatistics Unit, Mayo Clinic, Jacksonville, FL, USA and [7]Department of Medicine, [8]Department of Neurology, [9]Department of Pharmacology and [10]Geriatric Research Education and Clinical Center, Seattle Veterans Affairs Medical Center, University of Washington, Seattle, WA, USA and [11]University of Louisville School of Medicine, Louisville, KY, USA

Two extended haplotypes exist across the tau gene—H1 and H2—with H1 consistently associated with increased risk of progressive supranuclear palsy (PSP). Using 15 haplotype tagging SNPs (htSNPs), capturing >95% of *MAPT* haplotype diversity, we performed association analysis in a US sample of 274 predominantly pathologically confirmed PSP patients and 424 matched control individuals. We found that PSP risk is associated with one of two major ancestral H1 haplotypes, H1B, increasing from 14% in control individuals to 22% in PSP patients ($P < 0.001$). In young PSP patients, the H1B risk could be localized to a 22 kb regulatory region in intron 0 ($P < 0.001$) and could be fully explained by one SNP, htSNP167, creating a LBP-1c/LSF/CP2 site, shown to regulate the expression of genes in other neurodegenerative disorders. Luciferase reporter data indicated that the 182 bp conserved regulatory region, in which htSNP167 is located, is transcriptionally active with both alleles differentially influencing expression. Further, we replicated the htSNP167 association in a second, independently ascertained US PSP patient–control sample. However, the htSNP association showed that H1 risk alone could not explain the overall differences in H1 and H2 frequencies in PSP patients and control individuals. Thus, risk variants on different H1 htSNP haplotypes and protective variants on H2 contribute to population risk for PSP.

INTRODUCTION

Progressive supranuclear palsy (PSP) is the second most common form of parkinsonism after Parkinson disease (PD) (1). Distinctive clinical features include vertical supranuclear gaze and pseudobulbar palsy, early postural instability and progressive subcortical dementia (2). Abundant subcortical neurofibrillary tangles (NFTs) and neuropil threads of hyperphosphorylated tau protein characterize PSP neuropathologically. The deposition of hyperphosphorylated tau in neurons as fibrillar structures of varying morphology is also seen in

*To whom correspondence should be addressed at: Department of Molecular Genetics (VIB8), Neurodegenerative Brain Diseases Research Group, University of Antwerp, Building V – Room 0.10, Universiteitsplein 1, B-2610 Antwerpen, Belgium. Tel: +32 32651001; Fax: +32 32651012; Email: christine.vanbroeckhoven@ua.ac.be
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.
‡The authors wish it to be known that, in their opinion, the last two authors should be regarded as joint Last Authors.

© *The Author 2005. Published by Oxford University Press. All rights reserved.*
For Permissions, please email: journals.permissions@oxfordjournals.org

Downloaded from https://academic.oup.com/hmg/article/14/21/3281/2355838 by guest on 19 April 2024

Table 1. *MAPT* H1–H2 extended haplotype association

Individuals	*n*	Alleles (%)		Genotypes (%)			H1/H1 versus H1/H2 + H2/H2		
		H1	H2	H1/H1	H1/H2	H2/H2	OR	95% CI	*P*-value
Control individuals	424	680 (80.2)	168 (19.8)	277 (65.3)	126 (29.7)	21 (5.0)			
PSP patients	274	513 (93.6)	35 (6.4)	241 (88.0)	31 (11.3)	2 (0.7)	3.9	2.6–5.9	<0.001
PSP young	106	201 (94.8)	11 (5.2)	96 (90.6)	9 (8.5)	1 (0.9)	5.1	2.6–10.1	<0.001
PSP old	130	239 (91.9)	21 (8.1)	110 (84.6)	19 (14.6)	1 (0.8)	2.9	1.7–4.9	<0.001
PSP-pure	87	164 (94.3)	10 (5.7)	77 (88.5)	10 (11.5)	0 (0.0)	4.1	2.1–8.1	<0.001
PSP-plus	157	291 (92.7)	23 (7.3)	136 (86.6)	19 (14.6)	2 (1.3)	3.4	2.1–5.7	<0.001

PSP young, age at death ≤75; PSP old, age at death ≥76.

other neuropathological disorders including Alzheimer's disease (AD), frontotemporal dementia (FTD), Pick's disease, corticobasal degeneration (CBD) and argyrophilic grain disease, collectively known as tauopathies.

Tau is a microtubule (MT)-associated protein that promotes MT assembly, increases MT stability and plays a role in maintaining neuronal integrity and axonal transport (3,4). Tau binding to MT is mediated by imperfect repeats of 31 or 32 amino acid residues encoded by exons 9–12 of the tau gene (*MAPT*, MT-associated protein tau) located at 17q21.31. In adult human brain, six major tau isoforms are generated by alternative mRNA splicing (5). Alternative splicing of exon 10 results in two major classes of tau isoforms with three repeat (3R) or four repeat (4R) MT binding domains. In PSP, selective deposition of 4R tau isoforms was observed, although 3R and 4R levels in soluble tau were normal suggesting that alternative splicing of exon 10 is not altered in PSP (6). Identification of *MAPT* mutations in rare familial forms of FTD with parkinsonism linked to chromosome 17 (FTDP-17) pointed to a central role of tau pathology in neurodegeneration (7–9). Forty different *MAPT* mutations have been identified (http://www.molgen.ua.ac.be/FTDMutations) in patients that were clinically and pathologically characterized by variable neurodegenerative conditions including atypical PSP (10,11). In addition to the highly penetrant *MAPT* mutations, common genetic variations have been identified in *MAPT*, which constituted two major *MAPT* haplotypes (H1 and H2) in a 1.3 Mb genomic region (12–15). Association of the common haplotype H1 (~80%) with increased PSP risk has been consistently replicated in different Caucasian populations worldwide (12,16–19). Significant H1 associations were also reported in two CBD populations (20,21) and in a genetically homogeneous PD population from Norway (22,23). Recently, Stefansson *et al*. (24) and we (25) showed that *MAPT* is flanked by inverted low copy repeats that likely mediated inversion of the 570 kb *MAPT* segment, explaining the extended LD and suppressed H1/H2 recombination in the *MAPT* region.

To investigate whether specific genetic variants in *MAPT* on the H1 haplotype are responsible for the increased PSP risk, we analyzed a highly informative panel of 15 haplotype tagging SNPs (htSNPs) for genetic association in a large US PSP patient–control sample. To optimize our chances to finemap the associated genetic susceptibility in *MAPT*, we stratified the large US sample for age and PSP pathology. The association was subsequently replicated in a second independently ascertained US PSP patient–control sample.

RESULTS

H1–H2 extended haplotype association

We genotyped htSNP16 in the US patient–control sample and analyzed association with the *MAPT* extended haplotypes H1 and H2 in the overall sample as well as in the age and pathology strata (Table 1). In the overall sample, highly significant association was calculated for H1/H1 versus H1/H2 + H2/H2 with an odds ratios (OR) for PSP of 3.9 (95% CI 2.6–5.9). The strongest H1/H1 risk association was obtained in the subgroups of young PSP (OR = 5.1; 95% CI 2.6–10.1) and PSP-pure (OR = 4.1; 95% CI 2.1–8.1) patients.

Single H1 htSNP associations

We genotyped the 14 H1 htSNPs in the four LD blocks along *MAPT* in the US patient–control sample (Fig. 1A). In the overall sample, significant allelic associations were seen with five htSNPs with the strongest association observed with htSNP167 in block 2 resulting from a significant 11.6% increase of the allele A (OR = 1.6; 95% CI 1.2–2.0; $P < 0.001$) (Table 2). Significant associations were also observed with three htSNPs located within *MAPT*—htSNP364, htSNP510 and htSNP519—and htSNP563 localized 3′ of exon 13/14 (Fig. 1A). Genotypic association was only significant for htSNP167 (OR A/A versus A/G + G/G = 1.7; 95% CI 1.1–2.5; $P = 0.008$) (data not shown). Association analyses in the age-strata showed allelic association in patients with age at death ≤75 only with htSNP167 in LD block 2, whereas in patients with age at death ≥76, the same five htSNPs as in the overall analysis gave significant association (Table 2). The htSNP167 association was strongest in young PSP patients and resulted from an increased allele A frequency of 15% (age at death ≤75: OR = 2.2; 95% CI 1.6–3.1; $P < 0.001$ and age at death ≥76: OR = 1.5; 95% CI 1.1–2.0; $P = 0.005$). In addition, logistic regression analysis in the overall sample showed evidence of an interaction of age with the allele A of htSNP167 with the OR approximately doubling with a reduction of 10 years in age (OR = 2.0; 95% CI 1.3–3.0; $P = 0.002$). Allelic association with htSNP167 was also stronger in the PSP-pure patients ($P < 0.001$) compared with the PSP-plus patients (Table 2). As expected

Downloaded from https://academic.oup.com/hmg/article/14/21/3281/2355838 by guest on 19 April 2024

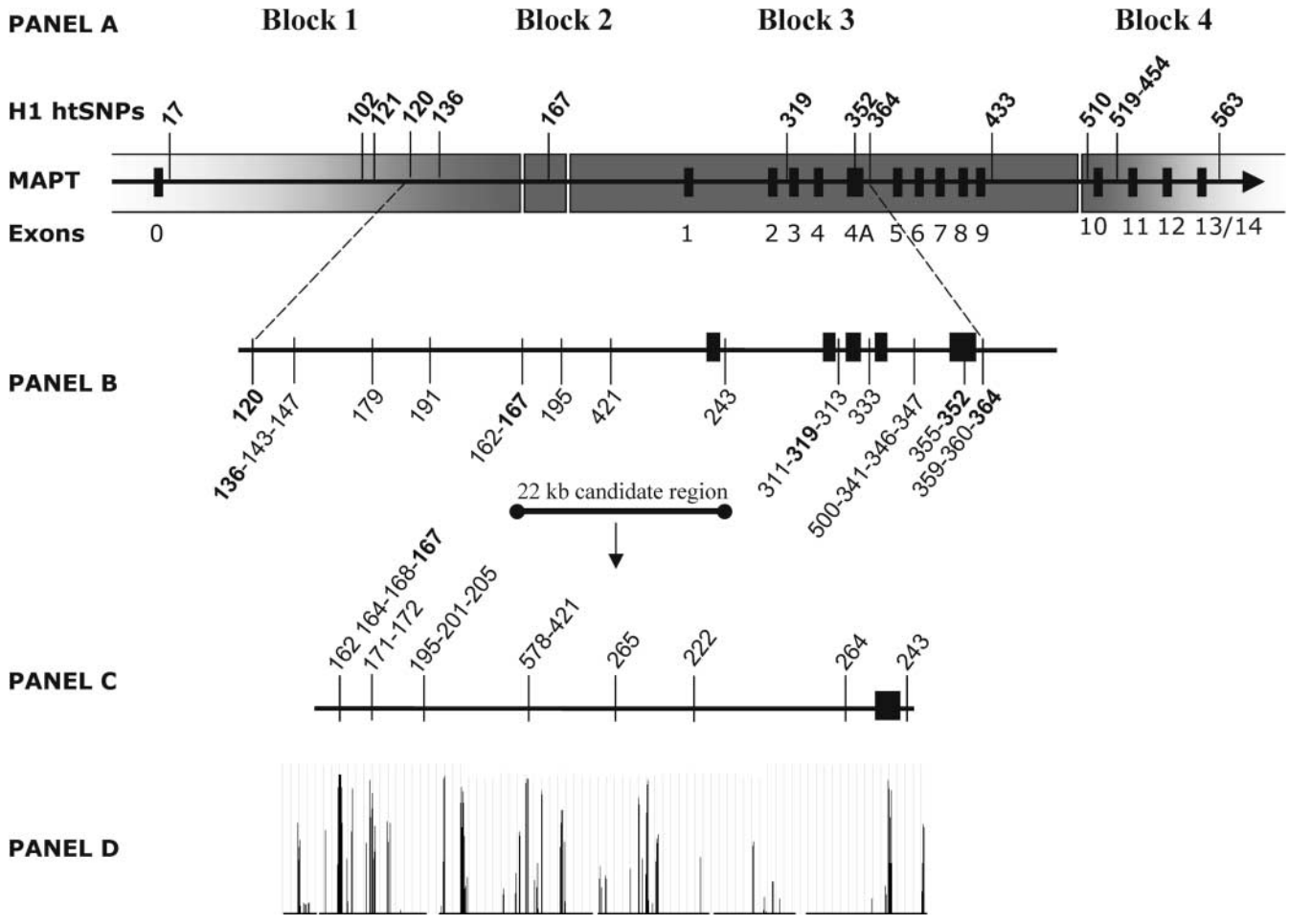

Figure 1. Scheme of *MAPT* genomic structure. *MAPT* exons (black boxes) are not drawn to scale, H1 SNPs located <1 kb apart in the genomic sequence are shown together at one position and htSNPs are in bold. **(A)** Fourteen htSNPs respective to their LD blocks (gray boxes). **(B)** Frequent H1 SNPs located within the 63 kb region selected for sliding window haplotypes analysis. **(C)** Rare SNPs located within the 22 kb candidate region, of which SNP164 and SNP205 were observed only in the PSP patients (Supplementary Material, Table S1). (**D**) Conservation plot extracted from the UCSC genome browser. Peaks represent sequences with $>50\%$ conservation in chimpanzee, dog, mouse and rat.

from the high proportion (63%) of patients with an age at death ≤ 75 included in the PSP-pure group and the high proportion (65%) of patients with an age at death ≥ 76 in the PSP-plus group, the genetic findings in the PSP-pure and PSP-plus groups largely mirrored the association results obtained in the young and old PSP patients. Further stratified association analyses were therefore only performed in the young and old PSP groups.

H1 htSNP subhaplotype associations

We estimated htSNP subhaplotype frequencies in LD blocks 1, 3 and 4 in the US patient–control sample; LD block 2 with only one htSNP is represented by allelic frequencies of htSNP167. In the overall sample, htSNP subhaplotype distributions were significantly different in blocks 1 and 4 (Table 3). In LD block 1, the association was primarily accounted for by a 9% increase of H1-B1B (CGTGT) ($P_{\text{H1-B1B}} = 0.002$), and in LD block 4 by an 8% increase of H1-B4C (GCTG) ($P_{\text{H1-B4C}} = 0.002$). Although not significant, we also observed a 5% increased frequency of H1-B3B (GCGG) in LD block 3. Moreover, highly significant association was observed for the ancestral H1 subhaplotype, H1B (CGTGT-A-GCGG-GCTG), increasing from 14% in control individuals to 22% in PSP patients ($P = 0.002$). Haplotype association analysis in the age strata indicated that htSNP subhaplotype distributions for LD blocks 1, 3 and 4 were not associated within the subgroup of young PSP patients (Table 4). In the old PSP patients, htSNP subhaplotype distributions in blocks 1, 3 and 4 were significantly different only in LD block 4 ($P = 0.046$). This association was primarily accounted for by an 8% increase of H1-B4C (GCTG) ($P_{\text{H1-B4C}} = 0.01$). In addition, the ancestral H1 subhaplotype H1B (CGTGT-A-GCGG-GCTG) showed a strong association ($P = 0.002$) in old PSP patients.

High-density SNP mapping

To finemap the H1 genetic risk for PSP detected with htSNP167, we selected a genomic region of 63 kb between htSNPs SNP120 and SNP364 extending into LD blocks 1 and 3 by at least 20 kb (Fig. 1B). In this region that spanned

Downloaded from https://academic.oup.com/hmg/article/14/21/3281/2355838 by guest on 19 April 2024

Table 2. Single H1 htSNP allelic associations

		Control individuals[a] ($n = 680$)		Patients[a] Overall ($n = 513$)			Age at death ≤75 years ($n = 201$)			Age at death ≥76 years ($n = 239$)			PSP-pure ($n = 164$)			PSP-plus ($n = 291$)		
SNP	Allele	*n*	%	*n*	%	*P*-value	*n*	%	*P*-value	*n*	%	*P*-value	*n*	%	*P*-value	*n*	%	*P*-value
17	C	549	80.7	420	81.9	0.62	145	84.3	0.28	200	83.7	0.31	137	83.5	0.41	242	83.2	0.37
	G	131	19.3	93	18.1		27	15.7		39	16.3		27	16.5		49	16.8	
102	T	293	43.7	208	40.5	0.28	76	43.4	0.92	99	41.4	0.55	65	39.6	0.34	125	43.0	0.84
	G	378	56.3	305	59.5		99	56.6		140	58.6		99	60.4		166	57.0	
121	T	459	67.5	369	71.9	0.10	120	69.4	0.65	177	74.1	0.06	118	72.0	0.27	211	72.5	0.12
	C	221	32.5	144	28.1		53	30.6		62	25.9		46	28.0		80	27.5	
120	G	585	86.0	450	87.7	0.39	145	83.8	0.45	211	88.3	0.38	141	86.0	0.99	254	87.3	0.6
	A	95	14.0	63	12.3		28	16.2		28	11.7		23	14.0		37	12.7	
136	T	477	70.1	374	72.9	0.30	120	69.4	0.83	174	72.8	0.44	118	72.0	0.65	209	71.8	0.6
	A	203	29.9	139	27.1		53	30.6		65	27.2		46	28.0		82	28.2	
167	G	373	54.9	222	43.3	<0.001	66	37.7	<0.001	106	44.4	0.005	62	37.8	<0.001	129	44.3	0.003
	A	307	45.1	291	56.7		109	62.3		133	55.6		102	62.2		162	55.7	
319	G	332	48.8	269	52.4	0.22	83	47.7	0.77	126	52.7	0.30	87	53.0	0.33	147	50.5	0.63
	A	348	51.2	244	47.6		91	52.3		113	47.3		77	47.0		144	49.5	
352	C	561	85.0	427	84.6	0.83	135	80.4	0.14	203	85.7	0.81	129	80.6	0.17	249	86.2	0.64
	T	99	15.0	78	15.4		33	19.6		34	14.3		31	19.4		40	13.8	
364	A	342	50.3	219	42.7	0.01	75	43.1	0.09	99	41.4	0.02	72	43.9	0.14	123	42.3	0.02
	G	338	49.7	294	57.3		99	56.9		140	58.6		92	56.1		168	57.7	
433	G	603	88.7	461	89.9	0.51	157	90.8	0.44	217	90.8	0.36	146	89.0	0.9	265	91.1	0.27
	A	77	11.3	52	10.1		16	9.2		22	9.2		18	11.0		26	8.9	
510	G	333	49.0	289	56.3	0.01	87	50.0	0.83	137	57.3	0.03	86	52.4	0.43	163	56.0	0.04
	A	347	51.0	224	43.7		87	50.0		102	42.7		78	47.6		128	44.0	
519	T	546	80.3	375	73.1	0.003	127	73.4	0.05	173	72.4	0.01	124	75.6	0.18	207	71.1	0.002
	C	134	19.7	138	26.9		46	26.6		66	27.6		40	24.4		84	28.9	
454	T	500	75.4	381	76.8	0.58	121	74.7	0.84	183	77.9	0.45	127	78.4	0.43	215	77.3	0.53
	A	163	24.6	115	23.2		41	25.3		52	22.1		35	21.6		63	22.7	
563	G	268	39.4	236	46.0	0.02	72	41.1	0.65	113	47.3	0.03	69	42.1	0.53	137	47.1	0.03
	A	412	60.6	227	54.0		103	84.3		126	52.7		95	57.9		154	52.9	

[a]*n* corresponds to number of alleles included in each analysis.

40 kb of *MAPT* intron 0 and exons 1 to 4A, we genotyped 11 extra H1 SNPs identified during the genomic *MAPT* sequencing in the young PSP patients and control individuals. Allelic associations were highly significant for htSNP167 (OR = 1.8; 95% CI 1.3–2.5; $P_{\text{allelic}} < 0.001$) and SNP421 (OR = 2.0; 95% CI 1.4–2.7; $P_{\text{allelic}} < 0.001$) (Table 4). Genotypic associations for htSNP167 and SNP421 in the H1/H1 individuals were $P = 0.002$ with ORs (A/A versus A/G + G/G) of 2.2 (95% CI 1.3–3.7) for htSNP167 and 2.3 (95% CI 1.4–3.9) for SNP421 (data not shown). Haplotype association analysis using sliding windows of three SNPs was significant within a region of seven SNPs (SNP191–htSNP319) with the highest significance observed for htSNP167–SNP195–SNP421 ($P < 0.001$) (Fig. 2). Sliding window analysis using two-SNP haplotypes showed significant association within a 22 kb interval (SNP162–SNP243) when the haplotype included either htSNP167 or SNP421. The most significant association was observed when both htSNP167 and SNP421, which showed a high degree of LD with $D' = 0.97$, were combined ($P < 0.001$) (data not shown).

Transcriptional effect of htSNP167A>G

Of the SNPs in the 22 kb candidate interval, including 10 rare SNPs (minor allele frequency < 5%) that we previously identified in our resequencing effort (Fig. 1C), 11 SNPs altered one or more putative transcription factor binding sites (Table 5). However, only htSNP167 was located in a region that was highly conserved compared with chimpanzee, dog, mouse and rat (Fig. 1D). The degree of conservation in this 182 bp region (91.5%) was higher than that of *MAPT* exon 1 and suggested the presence of functionally important *cis*-elements. *In silico* analysis showed that the allele A of htSNP167 abolished putative binding sites for the CP2-erythrocyte Factor related to drosophila Elf1 (LBP-1c/LSF/CP2, short CP2) (core/matrix similarity 0.83/0.843) and for the Se-Cys tRNA gene transcription-activating factor (STAF.01) (core/matrix similarity 1/0.805) (Table 5).

To test whether the 182 bp fragment surrounding htSNP167 contained functional elements, we used a luciferase reporter gene analysis. We cloned the 182 bp fragment of the *MAPT* regulatory region in intron 0 upstream of the minimal SV40 early promoter and obtained significantly enhanced transcription in transiently transfected non-neuronal (HEK293, $P < 0.001$) and neuronal (N2a, $P < 0.001$) cells (Fig. 3A). Both htSNP167 alleles, increased expression with allele G showing a small (16%) but significant ($P < 0.001$) higher increase in activity than allele A in N2a cells. When we used a fragment of the *MAPT* H1 proximal promoter, the luciferase expression in mouse and human neuroblastoma

Downloaded from https://academic.oup.com/hmg/article/14/21/3281/2355838 by guest on 19 April 2024

Table 3. H1 htSNP subhaplotype associations overall and in age-strata

LD Block	Haplotype[a]	htSNP alleles	Control individuals[b]	Patients[b]					
				Overall		Age at death ≤75 years		Age at death ≥76 years	
			%	%	*P*-value[c]	%	*P*-value[c]	%	*P*-value[c]
Block 1			($n = 668$)	($n = 512$)	0.043	($n = 200$)	0.66	($n = 238$)	0.18
	H1-B1A	CTTGA	26	26	0.47	28		26	
	H1-B1B	CGTGT	23	32	0.002	28		31	
	H1-B1C	CTTGT	13	10	0.26	11		11	
	H1-B1D	CGCAT	11	9	0.22	10		9	
	H1-B1E	GGCGT	12	9	0.21	8		9	
Block 2			($n = 680$)	($n = 513$)	<0.001	($n = 201$)	<0.001	($n = 239$)	0.005
	H1-B2A	G	55	43		40		44	
	H1-B2B	A	45	57		60		56	
Block 3			($n = 656$)	($n = 504$)	0.15	($n = 196$)	0.74	($n = 236$)	0.28
	H1-B3A	ACAG	30	23		26		24	
	H1-B3B	GCGG	24	29		27		30	
	H1-B3C	ACGG	8	10		8		11	
	H1-B3D	ATGG	12	14		15		13	
	H1-B3E	GCAA	7	7		8		6	
	H1-B3F	GCAG	12	12		10		12	
Block 4			($n = 662$)	($n = 496$)	0.037	($n = 190$)	0.37	($n = 234$)	0.046
	H1-B4A	ATTA	44	37	0.018	38		38	0.078
	H1-B4B	GTAG	16	16	0.80	15		14	0.54
	H1-B4C	GCTG	19	27	0.002	26		27	0.010
	H1-B4D	GTTA	8	9	0.69	10		7	0.74
	H1-B4E	ATAA	7	6	0.35	8		5	0.26
Block 1-2-3-4			($n = 632$)	($n = 490$)	0.003	($n = 188$)	0.12	($n = 232$)	0.008
	H1A	CTTGA-G-ACAG-ATTA	20	16	0.038	18		17	0.25
	H1B	CGTGT-A-GCGG-GCTG	14	22	0.002	20		24	0.002

[a]In each population, haplotypes with an estimated overall frequency of <5% were not included in the association analysis.
[b]*n* corresponds to numbers of H1 haplotypes included in each analysis.
[c]Value in first row is the global *P*-value; others are haplotype-specific *P*-values, shown only if the global *P*-value is <0.05.

cells was significantly higher than in the SV40 promoter ($P < 0.001$) (Fig. 3B). When we cloned the 182 bp regulatory fragment upstream of the H1 promoter, luciferase activity of both htSNP167 alleles differed again by 31% in N2a cells and 27% in SHSY-5Y cells ($P < 0.001$, Fig. 3B).

Replication of htSNP167 association

We replicated the htSNP167 association in an independent US PSP patient–control sample that showed significant association when examining the extended *MAPT* haplotypes H1 and H2, with an OR for H1/H1 versus H1/H2 + H2/H2 of 5.0 (95% CI 3.0–8.4). Significant allelic association was calculated with htSNP167 resulting from a 9% increase of allele A (OR = 1.4; 95% CI 1.1–1.9; $P = 0.01$) (Table 6). When only young, pathologically confirmed PSP patients were considered ($n = 30$), the allele A frequency of htSNP167 further increased to 58.6% compared with 46.3% in control individuals; however, this finding did not reach significance because of the small number of patients in this subgroup ($P = 0.07$). When patient and control groups of both US association samples were compared, similar htSNP167 allele distributions were obtained ($P_{\text{homogeneity}} = 0.52$), and a highly significant association of PSP with htSNP167 was obtained in the overall combined US patient–control samples ($P < 0.001$).

DISCUSSION

In this study, we were able to refine the genetic association of PSP within *MAPT* by detailed association analysis in one of the largest known collections of pathologically confirmed PSP patients worldwide, using a highly informative htSNP panel capturing > 95% of *MAPT* haplotype diversity. Previous attempts to finemap PSP risk in *MAPT* had been less successful because of lack of power in the experimental patient–control samples, and absence of a detailed high-density SNP map and full knowledge of the LD substructure of *MAPT*.

We showed that PSP risk in the extended US PSP patient–control sample originated on one of the two major ancestral *MAPT* H1 subhaplotypes, H1B ($P < 0.001$), which had a frequency of 14% in US control individuals comparable to what we previously observed in other Caucasian control samples (25,26) (unpublished data). In the overall sample, the highest significance was obtained in LD block 2 spanning ~10 kb of regulatory sequences in intron 0 of *MAPT* and tagged by htSNP167 ($P < 0.001$). In the age and pathology stratified

Downloaded from https://academic.oup.com/hmg/article/14/21/3281/2355838 by guest on 19 April 2024

Table 4. H1 SNP finemapping of young PSP risk

SNP	Allele	US sample				
		Control individuals ($n = 680$)		Patients ≤ 75 years ($n = 201$)		
		n	%	*n*	%	*P*-value
120	G	585	86.0	170	84.6	0.61
	A	95	14.0	31	15.4	
136	T	477	70.1	141	70.1	1.0
	A	203	29.9	60	29.9	
143	G	290	43.0	98	49.0	0.14
	A	384	57.0	102	21.0	
147	T	376	57.4	117	58.8	0.73
	A	279	42.6	82	41.2	
191	A	435	67.2	118	59.3	0.04
	G	212	32.8	81	40.7	
162	A	335	49.5	104	51.7	0.57
	G	342	50.5	97	48.3	
167	G	373	54.9	80	39.8	<0.001
	A	307	45.1	121	60.2	
195	G	459	68.6	150	75.4	0.07
	A	210	31.4	49	24.6	
421	G	359	55.2	73	38.6	<0.001
	A	291	44.8	116	61.4	
243	C	511	77.1	154	77.4	0.93
	T	152	22.9	45	22.6	
319	G	332	48.8	100	49.8	0.82
	A	348	51.2	101	50.2	
313	A	501	75.8	145	78.0	0.54
	T	160	24.2	41	22.0	
346	G	388	58.2	110	55.3	0.47
	A	279	41.8	89	44.7	
347	A	416	63.9	138	69.3	0.16
	T	235	36.1	61	30.7	
355	A	621	94.2	185	93.9	0.86
	G	38	5.8	12	6.1	
352	C	561	85.1	162	82.2	0.32
	T	98	14.9	35	17.8	
364	A	342	50.4	91	45.3	0.20
	G	337	49.6	110	54.7	

htSNPs in bold.

analyses, we also observed highly significant genotypic association with htSNP167 in young and PSP-pure patients ($P = 0.002$), which remained significant when we performed a Bonferroni correction to adjust for multiple testing. As the stratified data were highly similar in the young and PSP-pure patients, overlapping by 63%, further stratified analyses were performed only in the age strata. In young PSP patients, only 5% of the allele A-tagged htSNP167 subhaplotypes corresponded with the ancestral H1B, whereas all others represented different recombined derivatives of H1B explaining the finemapping to LD block 2. In contrast, in the old PSP patients, the htSNP167 allele A association was explained nearly entirely by association to H1B ($P \leq 0.002$). Our data in the US sample suggested that the genetic H1 risk variants for PSP in intron 0 are old mutations that appeared on H1B and have been separated over time by recombination in flanking regions that define the current LD block structure in *MAPT*. In young patients, the larger genetic component of the disease or greater genetic homogeneity allowed finemapping of these functional variants to *MAPT* intron 0. In contrast, in the old patients, the genetic risk variants could not be finemapped possibly because these PSP subgroups are indeed more genetically heterogeneous with other weaker modifiers existing on H1B beyond htSNP167.

Next we finemapped the young PSP risk to a region of 22 kb in LD block 2 surrounding htSNP167 and located upstream of exon 1, ~48 kb downstream of the major transcriptional start site (27). Significant association with young PSP was observed only with subhaplotypes that included htSNP167 or SNP421. When haplotype frequencies composed of htSNP167 and SNP421 were estimated in young PSP patients, a 15% increase of haplotype A–A was observed in the US sample. In an attempt to replicate the significant association in the 5′ regulatory region of *MAPT*, we analyzed htSNP167 in a second, independently ascertained US PSP patient–control sample that predominantly comprised clinical diagnosed PSP patients and identified significant allelic association in the replication sample ($P = 0.01$) and in the combined samples ($P < 0.001$).

Recently, Pittman *et al*. (28) published association data from a PSP patient–control sample that overlapped with the extended US sample we used in our study. They identified two common ancestral H1 subhaplotypes, B and C, which could correspond to, respectively, H1A and H1B that we identified on the basis of *MAPT* resequencing and LD mapping (25) (unpublished data). They identified a 56.3 kb interval on subhaplotype C spanning a region starting in exon 0 to intron 9. We were able to significantly reduce the candidate region, because we analyzed the data in the most genetic and homogenous group of young PSP patients and made optimal use of our high-density SNP map. Nevertheless, their data is likely to be consistent with ours, as htSNP167 (rs242557) on subhaplotype C also gave the highest significant association in their study ($P < 0.001$). Also of interest is that in a Norwegian PD sample an H1 subhaplotype containing SNP421 (rs242562) defined the PD-associated risk (23). In addition, we recently identified a protective two-SNP haplotypes including htSNP167 spanning 1 kb in *MAPT* intron 0 in a Belgian late-onset PD sample (age at onset ≤ 76) (26). Together these data support a role for altered regulation of transcription due to specific genetic variants in intron 0 of *MAPT* underlying the risk in different neurodegenerative brain diseases characterized by tauopathy.

The most parsimonious explanation for the PSP association, we observed in our analyses, is that a functional variant influences the regulation of transcription of *MAPT* resulting in increased risk for disease. *In silico* analysis further indicated that htSNP167 was located in a highly conserved 182 bp sequence (Fig. 2D). htSNP167A>G created an LBP-1c/LSF/CP2 binding site that belongs to a family of transcription factors regulating genes involved in development and cell fate determination and has previously been shown to regulate the expression of genes associated with neurodegenerative disorders (29–31). The absence of the LBP-1c/LSF/CP2 site in carriers of the htSNP167 allele A may affect *MAPT* expression overall or under specific circumstances (e.g. stress), which could in turn modify the risk for developing PSP.

Initial reporter gene analyses showed that the 182 bp fragment was able to enhance SV40 promoter driven luciferase activity confirming that it contained functional *cis*-elements. Also an artificial construct, in which the 182 bp fragment

Downloaded from https://academic.oup.com/hmg/article/14/21/3281/2355838 by guest on 19 April 2024

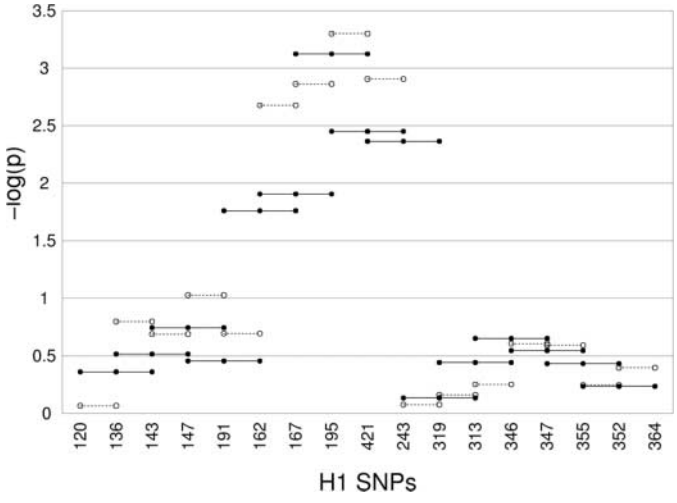

Figure 2. Sliding window SNP haplotype analyses. The H1 SNPs used are located in a 63 kb region spanning 40 kb of *MAPT* intron 0 and exons 1–4A. $-\log_{10}(P)$ values are indicated with full lines (three SNP haplotypes) and dashed lines (two SNP haplotypes).

was cloned upstream of an active *MAPT* H1 promoter fragment, showed that the htSNP167 alleles A and G had differential effects on H1 promoter driven luciferase activity in neuronal cells. The observation that allele G rather than the risk allele A increased promoter activity may seem counterintuitive but should not be considered as biologically relevant. In our construct, the 182 bp regulatory sequence was cloned upstream of the H1 promoter to test its effect on H1 expression dependent on the htSNP167 allele. An experimental approach using a larger genomic construct of the *MAPT* regulatory region would be needed to better mimic the *in vivo* situation.

In summary, we have shown that the allele A of the functional htSNP167, located on the ancestral *MAPT* H1B haplotype, is associated with PSP in two independent US PSP patient–control populations. In our combined US sample, the risk for PSP associated with htSNP167 is predominantly accounted for by an increase in the frequency of the homozygous genotype A/A, suggesting a recessive or a gene dose-dependent mode of inheritance. However, htSNP167 alone does not completely account for the PSP association with the extended *MAPT* haplotype H1 in the US population. This is clear, as removal of individuals carrying the htSNP167 genotype A/A from the analysis did not completely eliminate the association based on H1 and H2 frequencies in the remaining US patients and control individuals (OR reduced from 4.6 to 4.2). This implies that in addition to high-risk H1 subhaplotypes (e.g. tagged by htSNP167 allele A), H2 itself behaves as a protective haplotype that accounts for a major component of the H1–H2 association with PSP in the US population. The protective effect of H2 may reflect the functional impact of one or more H2-specific genetic variants and indeed the H2 tau promoter has recently been reported to show reduced activity in reporter gene assays (32). Reduced tau expression may well explain the protective effect of H2 in PSP, as lowered total tau levels might slow the process of aggregation, NFT formation and thus prevent or delay PSP onset. In addition, as *MAPT* H1 and H2 are inverted (24,25), it is tempting to speculate that the inversion is involved in conferring a protective effect to H2, for example by altering the higher order structure in this region leading to differential expression of H1 and H2. However, importantly, the observed *MAPT* inversion does not contradict our identification of htSNP167 allele A as a high risk H1 variant because in this study, we finemapped PSP risk on H1. Future studies based on our findings in the US PSP populations should help determining the underlying functional basis for the association between PSP and *MAPT* haplotypes. This will in turn have important implications for understanding the etiology of PSP and other neurodegenerative tauopathies and for identifying potential therapeutic targets.

MATERIALS AND METHODS

PSP patient–control samples

PSP society US sample. We selected 244 patients, collected between 1993 and 2004, with a primary pathological diagnosis of PSP and with frozen material available, from the Mayo Clinic (Jacksonville) brain bank. The majority of the patients ($n = 208$) were obtained through the Society of PSP brain bank. Patients from the following studies were also included in the analysis: Mayo Clinic Alzheimer Disease Center ($n = 4$), Mayo Clinic PD Center ($n = 12$), Alzheimer Disease and Related Disorders Association ($n = 5$) and the Florida AD Initiative ($n = 15$). Two additional pathologically confirmed PSP patients were obtained from the Northwestern University ADC. Pathological assessment of the patients was as previously described (33). The mean age at death in the complete sample was 75.6 ± 7.7 (range 53–98). In addition, a total of 30 clinically defined patients of PSP were also included in the analysis. Twenty one patients were collected from 1998 to 2002 at the Henry L. Jackson Foundation/National Institutes of Health. Additional clinical patients were recruited from the Mayo Clinic (Jacksonville) Movement Disorders Clinic ($n = 4$), Mayo Clinic (Scottsdale) ($n = 2$), the University of Minnesota ($n = 1$) and the University of British Columbia ($n = 2$). The mean age at collection for the clinical patients was 68.1 ± 7.4 (range 47–83). A total of 424 age and gender matched cognitively normal control individuals were obtained through the Normal and Pathological Aging protocol at the Mayo Clinic (34,35). All patients and control individuals were Caucasian, from the USA and Canada.

In the stratified analyses, we used pathologically confirmed PSP patients ($n = 244$) that comprised two groups, i.e. patients with pathological findings consistent with PSP-alone ('PSP-pure', $n = 87$) and those whose primary pathological diagnosis was PSP but who also had a major secondary pathology such as AD-like amyloid plaques ('PSP-plus', $n = 157$). In addition, we stratified the PSP patient group by age; we chose an age at death of 75 as cut-off age to roughly match numbers in the young PSP (≤ 75, $n = 106$) and PSP-pure patient groups.

Tertiary referral center US replication sample. A total of 48 patients, collected between 2004 and 2005, with a primary

Downloaded from https://academic.oup.com/hmg/article/14/21/3281/2355838 by guest on 19 April 2024

Table 5. MatInspector analysis of SNPs in the 22 kb candidate region

SNP	Numbering relative to AC091628.2	Alteration	Family/matrix	Further information	OT	CS	MS
164	g.55612T>C		No match				
168	g.55736G>A	+	FAST/FAST1.01	FAST-1 SMAD interacting protein	0.81	0.85	0.814
167	g.55776A>G	+	XSEC/STAF.01	Se-Cys tRNA gene transcription activating factor	0.77	1	0.805
		D	EKLF/EKLF.01	Erythroid krueppel like factor (EKLF)	0.89	1	0.925/0.896
		+	CP2F/CP2.02	LBP-1c (leader-binding protein-1c), LSF (late SV40 factor), CP2, SEF (SAA3 enhancer factor)	0.84	0.833	0.843
171	g.56921A>G	D	NOLF/OLF1.01	Olfactory neuron-specific factor	0.82	1	0.847/0.827
		+	BARB/BARBIE.01	Barbiturate-inducible element	0.88	1	0.883
172	g.57112C>T		No match				
195	g.58024G>A	+	MEF3/MEF3.01	MEF3 binding site, present in skeletal muscle-specific transcriptional enhancers	0.83	1	0.923
201	g.58227G/>A	+	CP2F/CP2.01	CP2	0.9	1	0.915
		D	OAZF/ROAZ.01	Rat C2H2 Zn finger protein involved in olfactory neuronal differentiation	0.73	1	0.84/0.878
		−	ETSF/CETS1P54.01	c-Ets-1(p54)	0.94	1	0.941
205	g.58382C>T	D	PERO/PPARA.01	PPAR/RXR heterodimers	0.7	0.807	0.708/0.717
		D	AP4R/TH1E47.01	Thing1/E47 heterodimer, TH1 bHLH member specific expression in a variety of embryonic tissues	0.93	1	0.931/0.959
		−	PAX3/PAX3.01	Pax-3 paired domain protein, expressed in embryogenesis, mutated in Waardenburg Syndrome	0.76	0.78	0.775
578	g.62554A>G	D	HOXF/EN1.01	Homeobox protein engrailed (en-1)	0.77	1	0.844/0.818
		+	MZF1/MZF1.01	Myeloid zinc finger protein MZF1	0.98	1	0.985
421	g.62802A>G	+	DEAF/NUDR.01	NUDR (nuclear DEAF-1 related transcriptional regulator protein)	0.73	1	0.731
		+	NFKB/NFKAPPAB65.01	NF-kappaB (p65)	0.87	0.826	0.879
		+	NFKB/NFKAPPAB50.01	NF-kappaB (p50)	0.83	1	0.891
265	g.66199A>G	−	PAX8/PAX8.01	PAX 2/5/8 binding site	0.88	1	0.885
222	g.68665C>T	+	MYT1/MYT1L.01	Myelin transcription factor 1-like, neuronal C2HC zinc finger factor 1	0.92	1	0.93
264	g.74334G>A	−	ETSF/ELK1.01	Elk-1	0.82	0.866	0.844
		+	OCT1/OCT1.05	Octamer-binding factor 1	0.9	1	0.914

htSNP in bold. OT, optimized threshold similarity; CS, core similarity and MS, matrix similarity; +, transcription factor binding site (TFBS) created; – TFBS abolished; D, TFBS altered.

pathologic diagnosis of PSP and with frozen material available were selected from the Mayo Clinic (Jacksonville) brain bank. The majority of the patients were obtained through the Society of PSP brain bank. Additional patients with a pathologic diagnosis of PSP were obtained from the Harvard Brain Tissue Resource Center ($n = 20$) and the University of Texas Southwestern Medical Center in Dallas ($n = 8$). The mean age at death for the pathologically confirmed patients was 74.9 ± 7.7 (range 46–93) including 30 young PSP patients (age at death ≤ 75) and 29 old PSP patients (age at death ≥ 76). In addition, a total of 102 patients with a clinical diagnosis of PSP were included in the analysis. Sixty-four clinical patients were collected from movement disorder clinics at the University of Kansas, the Toronto Hospital and Buffalo Veterans Affairs Medical Center. Additional clinical patients were recruited from the Mayo Clinic (Jacksonville) Department of Neurology ($n = 34$), the University of Louisville School of Medicine ($n = 1$), Mayo Clinic (Scottsdale) ($n = 2$) and the University of Medicine and Dentistry of New Jersey/Robert Wood Johnson Medical School ($n = 1$). The mean age at collection for the clinical sample was 71.2 ± 6.3 (range 56–86). A total of 396 age and gender matched cognitively normal control individuals were obtained from the Memory Disorders Clinic at Mayo Clinic (Jacksonville) ($n = 216$) and from the Normal and Pathological Aging protocol at the Mayo Clinic ($n = 180$) (34,35). All patients and control individuals were Caucasian from the USA and Canada.

MAPT genomic sequencing

We performed an extensive *MAPT* genomic sequencing in 23 Caucasian Dutch-speaking individuals (25) (unpublished data). This large-scale *MAPT* sequencing effort served multiple goals: (i) genomic mutation analysis in related individuals, i.e. four FTDU-17 patients linked to 17q21 and three unaffected non-segregating siblings (25), (ii) high-density SNP detection in unrelated individuals, i.e. eight PSP and two CBD patients and (iii) six community control individuals. The ten sporadic tauopathy patients were recruited in a prospective Belgian study of neurodegenerative and vascular dementia (36), were diagnosed based on NINCDS-PSP criteria and had a mean onset age 67.6 ± 7.8. The 23 individuals together represented 43 unrelated chromosomes of which 36 carried H1 and seven H2.

Downloaded from https://academic.oup.com/hmg/article/14/21/3281/2355838 by guest on 19 April 2024

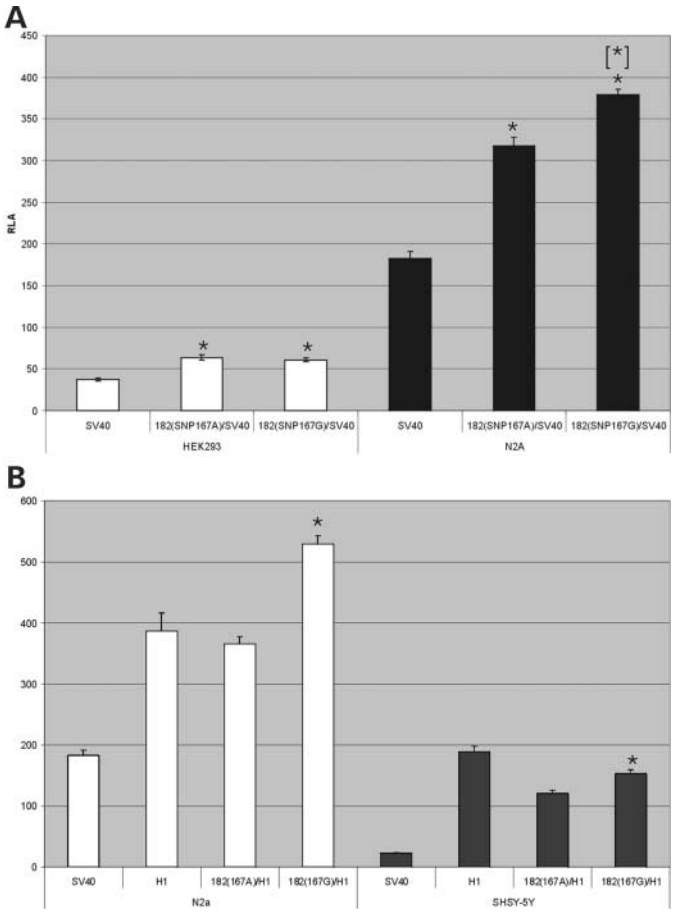

Figure 3. Luciferase reporter gene analyses. Bars represent relative luciferase activity of at least 12 independent transient transfection experiments and its standard error of the mean (SEM). (**A**) Effect of the 182 bp conserved sequence in the *MAPT* regulatory region of intron 0 containing htSNP167 allele A or G on SV40 driven luciferase activity. White bars correspond to data obtained in HEK293 cells and black bars in N2A cells. * represents the difference compared with SV40 expression $P < 0.0001$. [*] represents the difference between alleles A and G at htSNP167 $P < 0.001$. (**B**) The effect of the 182 bp regulatory fragment on MAPT H1 promoter driven luciferase activity. White bars correspond to data obtained in N2a cells and black bars in SHSY-5Y cells. * represents the difference between alleles A and G $P < 0.001$.

We sequenced a genomic region of 138.5 kb of *MAPT* (GenBank accession number NC_000017.9 from position 41323746–41462318), including 3.9 kb upstream of exon 0 and 3.7 kb downstream exons 13/14, which were conserved in rodent. PCR primers were developed to cover the complete genomic sequence with the exclusion of long stretches of interspersed repeat elements using SNPbox (37). Standard PCR reactions were performed and amplification products were purified, sequenced in both directions using the BigDye Terminator Cycle Sequencing v3.1 (Applied Biosystems) and analyzed on an ABI3730 DNA analyzer (Applied Biosystems). Sequences were analyzed using NovoSNP (38) and by visual inspection of the sequence reads. No causal mutations were detected in the FTDU-17 patients (25). In the ten sporadic tauopathy patients, we identified 30 patient-only genetic variants of which 14 were absent in 102 unrelated community control individuals (Supplementary Material, Table S1). Subsequent analysis of the 14 SNPs in 80 unrelated US PSP patients identified SNP163 (g.54661A/T) and SNP461 (g.127153_127154delCT) in one patient each.

In total 577 genetic variants were observed in the *MAPT* genomic sequence of which 424 variants (388 SNPs and 36 indels) were assigned to H2 based on their absence in H1/H1 individuals ($n = 16$) and their heterozygosity in H1/H2 individuals ($n = 7$). The remaining 153 were designated H1 variants (140 SNPs and 13 indels). For H2 genotyping, we selected SNP16 (g.8117G>A; numbering for this SNP and all genetic variants relative to AC091628.2) located in intron 0 as ht SNP. Validation of SNP16 in 91 Belgian PD patients showed complete LD of SNP16 with two previously published H2 SNPs, SNP9i (g.109929A>G) and SNP9ii (g.110013T>C) (12).

LD substructure of *MAPT*

Pairwise measures of linkage disequilibrium (LD) were calculated using Ldmax (http://www.sph.umich.edu/csg/abecasis/GOLD/docs/ldmax.html). In the LD mapping of *MAPT*, we used only the 85 common H1 variants with rare allele frequencies of ≥5%. Further, we excluded 42 H1 variants from further analyses as they showed complete LD with at least one other H1 variant in the pairwise analyses. As we used pyrosequencing as SNP genotyping method, we preferred SNPs to indels. The remaining 43 H1 variants were genotyped in 51 control triads ascertained in Flanders. One SNP, SNP162, was excluded from the LD analysis as in 7.9% of control individuals, the rare H1 allele was observed on H2. We defined four LD blocks that spanned ∼45 kb including exon 0 and the 5′ part of intron 0 (block 1), ∼10 kb in intron 0 (block 2), ∼63 kb of the 3′ end of intron 0 and exons 1–9 (block 3) and ∼20 kb comprising exons 10–13/14 (block 4) (Fig. 1A). In each LD block, we estimated common H1 subhaplotypes (frequency ≥2%) in the parental chromosomes in which subhaplotypes could be set unambiguously and obtained nine subhaplotypes in block 1 ($n = 17$), seven in block 3 ($n = 18$) and six in block 4 ($n = 6$) representing 87, 85 and 96%, respectively, of the H1 subhaplotype diversity. Block 2 contained only one common H1 SNP, SNP167. Using the computer program Tag‘n’Tell (http://snp.cgb.ki.se/tagntell/), we selected 14 htSNPs—five in block 1, one in block 2 and four in blocks 3 and 4 each (Fig. 1A)—tagging together 94% of *MAPT* H1 haplotype diversity. Together with H2 SNP16, the htSNP panel captured 95% of *MAPT* haplotype diversity. In the US control chromosomes ($n = 632$), the two ancestral *MAPT* H1 haplotypes, H1A (CTTGA-G-ACAG-ATTA) and H1B (CGTGT-A-GCGG-GCTG), had frequencies of 20 and 14%, respectively, similar to those we previously observed in Belgian ($n = 291$, H1A = 19%; H1B = 14%), French ($n = 412$, H1A = 22%; H1B = 15%) and Spanish ($n = 84$, H1A = 19%; H1B = 13%) control samples (25,26) (unpublished data).

SNP genotyping

MAPT H1 and H2 haplotyping was performed by htSNP16 (g.8117G>A) with the allele G tagging H1 and the allele A tagging H2. htSNP16 genotyping was performed by pyrosequencing on a PSQTM pyrosequencer using 5′-bio-CCT

Downloaded from https://academic.oup.com/hmg/article/14/21/3281/2355838 by guest on 19 April 2024

Table 6. Allelic association of htSNP167 in the US replication sample

htSNP	Allele	Control individuals ($n = 559$)		Patients Overall ($n = 329$)			Age at death $\leq$75 years ($n = 58$)		
		n	%	*n*	%	*P*	*n*	%	*P*
167	G	300	53.7	147	44.7	0.01	24	41.4	0.07
	A	259	46.3	182	55.3		34	58.6	

CTG TCG ACT ATC AGG TAA GC-3′ and 5′-CTT GCT CAG GTC CCT TTC C-3′ as PCR primers and 5′-CGG CGC ACG AAG C-3′ as reverse sequencing primer. Pyrosequencing assays were also developed for all H1 htSNPs except for htSNP352 (g.97065C>T) and for 10 additional H1 SNPs selected for high-density SNP LD mapping (Supplementary Material, Table S2). htSNP352 and SNP355 (g.96899A/G) were genotyped by direct sequencing. For all known H1 SNPs that were analyzed in this study, rs-numbers are included in Supplementary Material, Table S2.

Statistical analyses

Hardy Weinberg equilibrium (HWE) was calculated for the all htSNPs using the HWE program and all patient and control populations were in HWE (39). For single H1 htSNP association analysis, two different strategies were applied to compensate for the significant differences in relative numbers of H1 chromosomes between PSP patients and control individuals. For allelic single htSNP association, we used all H1 htSNP haplotypes including H1/H1 + H1/H2 individuals as all htSNPs were invariable on H2. Genotypic single htSNP associations were assessed using Genepop (http://wbiomed.curtin.edu.au/genepop/) in subpopulations of H1/H1 patients and control individuals. To investigate H1 haplotype associations, we used the score method of Schaid *et al*. (40). Only haplotypes with an estimated overall frequency of $\geq$5% were considered for these analyses. The *P*-values reported are those based on asymptotic assumptions. These assumptions were verified also by calculating simulation *P*-values based on 1000 random permutations of patient and control labels and results were found to be consistent. The H1 htSNP haplotypes of H1/H2 individuals were included by separating out the H1 haplotypes and randomly pairing these up to create a set of pseudo-H1/H1 individuals to add to the pool of H1/H1 individuals. Although overall results were not found to be sensitive to this random pairing, this process was repeated 11 times for each analysis, and the results of the analysis giving the median overall *P*-value for haplotype associations are reported. In the age-stratified analysis, the control group was used as a whole, because none of the single htSNPs allelic distributions was significantly different when comparing young versus old control individuals grouped on the basis of median age of 75 at inclusion. The level of significance was defined as $P < 0.05$. OR are shown with their 95% CI.

Sequence predictions of functional elements

Transcription factor binding sites were identified in the 22 kb candidate region from 54720 to 76720 in AC091628.2 with MatInspector (41) using a core similarity cut-off value of 0.75 and an optimized matrix similarity threshold. Conserved sequences and TFBS were detected using the VISTA tools (http://www-gsd.lbl.gov/vista/index.shtml). MatInspector was applied to investigate the effect of SNP167A>G on putative transcription factor binding sites using a core similarity cut-off value of 0.75 and an optimized matrix similarity threshold.

Luciferase reporter constructs

To test whether the 182 bp conserved regulatory region, located 47946–48127 bp downstream of the major *MAPT* transcription initiation site (27) in intron 0 of *MAPT*, contained regulatory elements, we amplified genomic DNA of an H1/H1 individual using primers 5′-ggggtaccccTTGATGATGCATG GACCTCTCT-3′ and 5′-tccccccgggggggaAAGCAAAA GAAAGACTGTGGAAGG-3′. The H1/H1 individual was selected for homozygosity except htSNP167A>G using genomic sequencing. The primers contained a *Kpn* I and an *Xma* I, respectively, site for cloning of the 182 bp fragment upstream of the SV40 early promoter of the firefly luciferase gene in the pGL3-Promoter vector (Promega). Similar experiments were performed using, instead of the SV40 promoter, the H1 MAPT proximal promoter, which we defined as a 1.1 kb genomic fragment of −911/+190 including exon 0 (positions relative to the major *MAPT* transcription start site at nt 7805 in AC091628.2). The promoter fragment was obtained by PCR amplification of genomic DNA of an H1/H2 individual using primers 5′-tccccccgggggggaCTGCCC TAAACCCCCCTACA-3′ and 5′-ggaagatcttccCAGCGGAG GAGGAGAAGGT-3′, and sequencing indicated that it did not comprise H1 SNPs. The primers contained an *Xma* I and a *Bgl* II site, respectively, for cloning upstream of the firefly luciferase gene in the pGL3-Basic vector (Promega). Both htSNP167 alleles of the 182 bp regulatory fragment were subcloned upstream of the *MAPT* H1 promoter using the *Kpn* I and *Xma* I restriction sites. The sequence integrity of all inserts was confirmed using vector-specific primers.

Eukaryotic cell culture and transient transfection

Human SHSY-5Y and mouse N2a neuroblastoma cells were propagated in a minimal essential medium with Earle's salt, 10% fetal bovine serum, 2 mM L-glutamine, 200 IU/ml of penicillin, 200 g/ml of streptomycin and 0.1 mM non-essential amino acids (Invitrogen). Human embryonic kidney cells (HEK293) were propagated in OptiMem (Invitrogen) with 10% fetal bovine serum, 200 IU/ml of penicillin and 200 g/ml of streptomycin. For transient transfection SHSY-5Y, N2a and HEK293 cells were seeded in 24-well tissue culture

Downloaded from https://academic.oup.com/hmg/article/14/21/3281/2355838 by guest on 19 April 2024

dishes at 0.65×10^6, 1.2×10^5 and 7.5×10^5 cells/well, respectively, and allowed to recover for 24 h. Cells were co-transfected with 32 ng (HEK293, N2a) or 80 ng (SHSY-5Y) of pRL-TK plasmid containing the herpes simplex virus thymidine kinase promoter upstream of the renilla luciferase gene (Promega) and 800 ng of either one of the *MAPT* constructs or one of the control plasmids, using Lipofectamine 2000 (2.4 μl, Invitrogen). Empty pGL3-basic vector was used as a negative control, pGL3-Promoter plasmid as a positive control.

Luciferase activity measurements

Transfected cells were cultured for 24–30 h, washed with 1 ml phosphate-buffered saline (Invitrogen) and lysed with Passive lysis buffer (Promega). Firefly luciferase activities (LA_F) and renilla luciferase activities (LA_R) were measured sequentially using a Dual-Luciferase reporter assay system and a Veritas™ (Promega) and/or VICTOR Light (Perkin Elmer, Boston, MA, USA) Microplate Luminometer w/Dual Reagent Injectors. To correct for transfection efficiency and DNA uptake, the relative luciferase activity (RLA) was calculated as $RLA = LA_F/LA_R$. The Mann–Whitney *U* test, a non-parametric analog to the unpaired *t*-test, was used to compare the RLA produced by the different transfectants.

ACKNOWLEDGEMENTS

The authors are grateful to the patients and their families for their cooperation in this project, to Eileen Bigio, Neill Graff-Radford, Lawrence Golbe, Donald Calne, Paul Tuite, Francesch Valldeoriola, María José Marti, Adriana Cardozo, Victor Obach and Matilde Calopa for providing samples and to Sally Serneels for help with *MAPT* genotyping. This research was funded by the Special Research Fund of the University of Antwerp (BOF-UA), the Fund for Scientific Research Flanders (FWO-F), the Interuniversity Attraction Poles program P5/19 of the Belgian Federal Science Policy Office (POD) and the International Alzheimer Research Foundation (IARF); the EU contract LSHM-CT-2003-503330 (APOPIS); the Progressive Supranuclear Palsy Association (PSP-Europe) (P.P.); the Society for Progressive Supranuclear Palsy (PSP-USA) (A.G., E.T. and P.P.), the Harvard Brain Tissue Resource Center (PHS #MH/NS 31862), the NIA (PO1 #AG 17216 to M.H.) and the Mayo Foundation, USA. R.R., M.C. and J.T. are post-doctoral fellows of the FWO-F, Belgium; S.M. was funded by a NRSA post-doctoral fellowship #AG 24030, USA.

Conflict of Interest statement. The authors declare that they have no competing financial interests.

REFERENCES

1. Bower, J.H., Maraganore, D.M., McDonnell, S.K. and Rocca, W.A. (1997) Incidence of progressive supranuclear palsy and multiple system atrophy in Olmsted County, Minnesota, 1976 to 1990. *Neurology*, **49**, 1284–1288.
2. Litvan, I., Agid, Y., Calne, D., Campbell, G., Dubois, B., Duvoisin, R.C., Goetz, C.G., Golbe, L.I., Grafman, J., Growdon, J.H. *et al*. (1996) Clinical research criteria for the diagnosis of progressive supranuclear palsy (Steele–Richardson–Olszewski syndrome): report of the NINDS-SPSP international workshop. *Neurology*, **47**, 1–9.
3. Ebneth, A., Godemann, R., Stamer, K., Illenberger, S., Trinczek, B. and Mandelkow, E. (1998) Overexpression of tau protein inhibits kinesin-dependent trafficking of vesicles, mitochondria, and endoplasmic retziculum: implications for Alzheimer's disease. *J. Cell. Biol.*, **143**, 777–794.
4. Hirokawa, N. (1994) Microtubule organization and dynamics dependent on microtubule-associated proteins. *Curr. Opin. Cell Biol.*, **6**, 74–81.
5. Goedert, M., Spillantini, M.G., Jakes, R., Rutherford, D. and Crowther, R.A. (1989) Multiple isoforms of human microtubule-associated protein tau: sequences and localization in neurofibrillary tangles of Alzheimer's disease. *Neuron*, **3**, 519–526.
6. Liu, W.K., Le, T.V., Adamson, J., Baker, M., Cookson, N., Hardy, J., Hutton, M., Yen, S.H. and Dickson, D.W. (2001) Relationship of the extended tau haplotype to tau biochemistry and neuropathology in progressive supranuclear palsy. *Ann. Neurol.*, **50**, 494–502.
7. Hutton, M., Lendon, C.L., Rizzu, P., Baker, M., Froelich, S., Houlden, H., Pickering-Brown, S., Chakraverty, S., Isaacs, A., Grover, A. *et al*. (1998) Association of missense and 5′-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature*, **393**, 702–705.
8. Poorkaj, P., Bird, T.D., Wijsman, E., Nemens, E., Garruto, R.M., Anderson, L., Andreadis, A., Wiederholt, W.C., Raskind, M. and Schellenberg, G.D. (1998) Tau is a candidate gene for chromosome 17 frontotemporal dementia. *Ann. Neurol.*, **43**, 815–825.
9. Spillantini, M.G., Murrell, J.R., Goedert, M., Farlow, M.R., Klug, A. and Ghetti, B. (1998) Mutation in the *tau* gene in familial multiple system tauopathy with presenile dementia. *Proc. Natl Acad. Sci. USA*, **95**, 7737–7741.
10. Pastor, P., Pastor, E., Carnero, C., Vela, R., Garcia, T., Amer, G., Tolosa, E. and Oliva, R. (2001) Familial atypical progressive supranuclear palsy associated with homozygosity for the delN296 mutation in the *tau* gene. *Ann. Neurol.*, **49**, 263–267.
11. Rademakers, R., Cruts, M. and Van Broeckhoven, C. (2004) The role of tau (MAPT) in frontotemporal dementia and related tauopathies. *Hum. Mutat.*, **24**, 277–295.
12. Baker, M., Litvan, I., Houlden, H., Adamson, J., Dickson, D., Perez-Tur, J., Hardy, J., Lynch, T., Bigio, E. and Hutton, M. (1999) Association of an extended haplotype in the *tau* gene with progressive supranuclear palsy. *Hum. Mol. Genet.*, **8**, 711–715.
13. Oliveira, S.A., Scott, W.K., Zhang, F., Stajich, J.M., Fujiwara, K., Hauser, M., Scott, B.L., Pericak-Vance, M.A., Vance, J.M. and Martin, E.R. (2004) Linkage disequilibrium and haplotype tagging polymorphisms in the Tau H1 haplotype. *Neurogenetics*, **5**, 147–155.
14. Pastor, P., Ezquerra, M., Perez, J.C., Chakraverty, S., Norton, J., Racette, B.A., McKeel, D., Perlmutter, J.S., Tolosa, E. and Goate, A.M. (2004) Novel haplotypes in 17q21 are associated with progressive supranuclear palsy. *Ann. Neurol.*, **56**, 249–258.
15. Pittman, A.M., Myers, A.J., Duckworth, J., Bryden, L., Hanson, M., Abou-Sleiman, P., Wood, N.W., Hardy, J., Lees, A. and de Silva, R. (2004) The structure of the tau haplotype in controls and in progressive supranuclear palsy. *Hum. Mol. Genet.*, **13**, 1267–1274.
16. Conrad, C., Andreadis, A., Trojanowski, J.Q., Dickson, D.W., Kang, D., Chen, X., Wiederholt, W., Hansen, L., Masliah, E., Thal, L.J. *et al*. (1997) Genetic evidence for the involvement of tau in progressive supranuclear palsy. *Ann. Neurol.*, **41**, 277–281.
17. Higgins, J.J., Golbe, L.I., De Biase, A., Jankovic, J., Factor, S.A. and Adler, R.L. (2000) An extended 5′-tau susceptibility haplotype in progressive supranuclear palsy. *Neurology*, **55**, 1364–1367.
18. Morris, H.R., Janssen, J.C., Bandmann, O., Daniel, S.E., Rossor, M.N., Lees, A.J. and Wood, N.W. (1999) The *tau* gene A0 polymorphism in progressive supranuclear palsy and related neurodegenerative diseases. *J. Neurol. Neurosurg. Psychiatry*, **66**, 665–667.
19. Oliva, R., Tolosa, E., Ezquerra, M., Molinuevo, J.L., Valldeoriola, F., Burguera, J., Calopa, M., Villa, M. and Ballesta, F. (1998) Significant changes in the tau A0 and A3 alleles in progressive supranuclear palsy and improved genotyping by silver detection. *Arch. Neurol.*, **55**, 1122–1124.
20. Di Maria, E., Tabaton, M., Vigo, T., Abbruzzese, G., Bellone, E., Donati, C., Frasson, E., Marchese, R., Montagna, P., Munoz, D.G. *et al*. (2000) Corticobasal degeneration shares a common genetic background with progressive supranuclear palsy. *Ann. Neurol.*, **47**, 374–377.

Downloaded from https://academic.oup.com/hmg/article/14/21/3281/2355838 by guest on 19 April 2024

21. Houlden, H., Baker, M., Morris, H.R., MacDonald, N., Pickering-Brown, S., Adamson, J., Lees, A.J., Rossor, M.N., Quinn, N.P., Kertesz, A. *et al*. (2001) Corticobasal degeneration and progressive supranuclear palsy share a common tau haplotype. *Neurology*, **56**, 1702–1706.
22. Farrer, M., Skipper, L., Berg, M., Bisceglio, G., Hanson, M., Hardy, J., Adam, A., Gwinn-Hardy, K. and Aasly, J. (2002) The tau H1 haplotype is associated with Parkinson's disease in the Norwegian population. *Neurosci. Lett.*, **322**, 83–86.
23. Skipper, L., Wilkes, K., Toft, M., Baker, M., Lincoln, S., Hulihan, M., Ross, O.A., Hutton, M., Aasly, J. and Farrer, M. (2004) Linkage disequilibrium and association of MAPT H1 in Parkinson disease. *Am. J. Hum. Genet.*, **75**, 669–677.
24. Stefansson, H., Helgason, A., Thorleifsson, G., Steinthorsdottir, V., Masson, G., Barnard, J., Baker, A., Jonasdottir, A., Ingason, A., Gudnadottir, V.G. *et al*. (2005) A common inversion under selection in Europeans. *Nat. Genet.*, **37**, 129–137.
25. Cruts, M., Rademakers, R., Gijselinck, I., van der, Z.J., Dermaut, B., De Pooter, T., De Rijk, P., Del Favero, J. and Van Broeckhoven, C. (2005) Genomic architecture of human 17q21 linked to frontotemporal dementia uncovers a highly homologous family of low copy repeats in the tau region. *Hum. Mol. Genet*, **14**, 1753–1762.
26. Rademakers, R., Pals, P., Lohmann, E., Durr, A., Engelborghs, S., De Pooter, T., Van den Broeck, M., Camuzat, A., Nuytemans, K., Pickut, B.A., Cruts, M., De Deyn, P.P., Cras, P., Brice, A. and Van Broeckhoven, C. High-density SNP haplotyping in two independent Caucasian PD association samples implicates tau splicing in early-onset Parkinson's disease susceptibility. *Hum. Mol. Genet*, in review.
27. Andreadis, A., Wagner, B.K., Broderick, J.A. and Kosik, K.S. (1996) A tau promoter region without neuronal specificity. *J. Neurochem.*, **66**, 2257–2263.
28. Pittman, A.M., Myers, A.J., Abou-Sleiman, P., Fung, H.C., Kaleem, M., Marlowe, L., Duckworth, J., Leung, D., Williams, D., Kilford, L. *et al*. (2005) Linkage disequilibrium fine-mapping and haplotype association analysis of the *tau* gene in progressive supranuclear palsy and corticobasal degeneration. *J. Med. Genet*, in press.
29. Bing, Z., Reddy, S.A., Ren, Y., Qin, J. and Liao, W.S. (1999) Purification and characterization of the serum amyloid A3 enhancer factor. *J. Biol. Chem.*, **274**, 24649–24656.
30. Kang, H.C., Chae, J.H., Kim, B.S., Han, S.Y., Kim, S.H., Auh, C.K., Yang, S.I. and Kim, C.G. (2004) Transcription factor CP2 is involved in activating mBMP4 in mouse mesenchymal stem cells. *Mol. Cells*, **17**, 454–461.
31. Lau, K.F., Miller, C.C., Anderton, B.H. and Shaw, P.C. (1999) Molecular cloning and characterization of the human glycogen synthase kinase-3beta promoter. *Genomics*, **60**, 121–128.
32. Kwok, J.B., Teber, E.T., Loy, C., Hallupp, M., Nicholson, G., Mellick, G.D., Buchanan, D.D., Silburn, P.A. and Schofield, P.R. (2004) Tau haplotypes regulate transcription and are associated with Parkinson's disease. *Ann. Neurol.*, **55**, 329–334.
33. Josephs, K.A. and Dickson, D.W. (2003) Diagnostic accuracy of progressive supranuclear palsy in the Society for Progressive Supranuclear Palsy brain bank. *Mov. Disord.*, **18**, 1018–1026.
34. Caselli, R.J., Osborne, D., Reiman, E.M., Hentz, J.G., Barbieri, C.J., Saunders, A.M., Hardy, J., Graff-Radford, N.R., Hall, G.R. and Alexander, G.E. (2001) Preclinical cognitive decline in late middle-aged asymptomatic apolipoprotein E-e4/4 homozygotes: a replication study. *J. Neurol. Sci.*, **189**, 93–98.
35. Caselli, R.J., Hentz, J.G., Osborne, D., Graff-Radford, N.R., Barbieri, C.J., Alexander, G.E., Hall, G.R., Reiman, E.M., Hardy, J. and Saunders, A.M. (2002) Apolipoprotein E and intellectual achievement. *J. Am. Geriatr. Soc.*, **50**, 49–54.
36. Engelborghs, S., Dermaut, B., Goeman, J., Saerens, J., Marien, P., Pickut, B.A., Van den, B.M., Serneels, S., Cruts, M., Van Broeckhoven, C. *et al*. (2003) Prospective Belgian study of neurodegenerative and vascular dementia: APOE genotype effects. *J. Neurol. Neurosurg. Psychiatry*, **74**, 1148–1151.
37. Weckx, S., De Rijk, P., Van Broeckhoven, C. and Del Favero, J. (2004) SNPbox, a modular software package for large scale primer design. *Bioinformatics*, **21**, 385–387.
38. Weckx, S., Del Favero, J., Rademakers, R., Claes, L., Cruts, M., De Jonghe, P., Van Broeckhoven, C. and De Rijk, P. (2005) novoSNP, a novel computational tool for sequence variation discovery. *Genome Res.*, **15**, 436–442.
39. Terwilliger, J. and Ott, J. (1994) *Handbook of Human Genetic Linkage*. John Hopkins University Press, Baltimore, MD.
40. Schaid, D.J., Rowland, C.M., Tines, D.E., Jacobson, R.M. and Poland, G.A. (2002) Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am. J. Hum. Genet.*, **70**, 425–434.
41. Quandt, K., Frech, K., Karas, H., Wingender, E. and Werner, T. (1995) MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res.*, **23**, 4878–4884.

Downloaded from https://academic.oup.com/hmg/article/14/21/3281/2355838 by guest on 19 April 2024