

Haplotype-specific expression of exon 10 at the human *MAPT* locus

Tara M. Caffrey¹, Catharine Joachim^{2,3}, Silvia Paracchini¹, Margaret M. Esiri²
and Richard Wade-Martins^{1,*}

¹The Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK,

²Department of Clinical Neuropathology, University of Oxford, Radcliffe Infirmary, Oxford OX2 6HE, UK and ³Oxford Project to Investigate Memory and Ageing (OPTIMA), Department of Pharmacology, University of Oxford, Mansfield Road, Oxford OX1 3QT, UK

Received August 2, 2006; Revised October 1, 2006; Accepted October 31, 2006

Neurofibrillary tangles composed of exon 10+ microtubule associated protein tau (*MAPT*) deposits are the characteristic feature of the neurodegenerative diseases progressive supranuclear palsy (PSP) and corticobasal degeneration (CBD). PSP, CBD and more recently Alzheimer's disease and Parkinson's disease, are associated with the *MAPT* H1 haplotype, but the relationship between genotype and disease remains unclear. Here, we investigate the hypothesis that H1 expresses more exon 10+ *MAPT* mRNA compared to the other haplotype, H2, leading to a greater susceptibility to neurodegeneration in H1 carriers. We performed allele-specific gene expression on two H1/H2 heterozygous human neuronal cell lines, and 14 H1/H2 heterozygous control individual post-mortem brain tissue from two brain regions. In both tissue culture and post-mortem brain tissue, we show that the *MAPT* H1 haplotype expresses significantly more exon 10+ *MAPT* mRNA than H2. In post-mortem brain tissue, we show that the total level of *MAPT* expression from H1 and H2 is not significantly different, but that the H1 chromosome expresses up to 1.43-fold more exon 10+ *MAPT* mRNA than H2 in the globus pallidus, a brain region highly affected by tauopathy (maximum exon 10+ *MAPT* H1:H2 transcript ratio = 1.425, SD = 0.205, $P < 0.0001$), and up to 1.29-fold more exon 10+ *MAPT* mRNA than H2 in the frontal cortex (maximum exon 10+ *MAPT* H1:H2 transcript ratio = 1.291, SD = 0.315, $P = 0.006$). These data may explain the increased susceptibility of H1 carriers to neurodegeneration and suggest a potential mechanism between *MAPT* genetic variability and the pathogenesis of neurodegenerative disease.

INTRODUCTION

The human microtubule associated protein tau (*MAPT* or *tau*) locus lies on chromosome 17q21 and the protein is important in establishing and maintaining neuronal morphology. Six tau isoforms are expressed in the adult human central nervous system, produced by alternative splicing of exons 2, 3 and 10. Alternative splicing of exon 10 leads to a protein containing either three (3R; exon 10–) or four (4R; exon 10+) tandem repeats of a microtubule-binding motif (1). Accumulation of hyperphosphorylated intracellular tau aggregates known as neurofibrillary tangles (NFTs) is seen in Alzheimer's disease (AD), progressive supranuclear palsy (PSP), corticobasal degeneration (CBD), frontotemporal dementia (FTD) and Pick's Disease (PiD), neurodegenerative disorders

collectively referred to as tauopathies (2). In PSP and CBD, the NFTs contain an excess of 4R tau protein, hence these diseases are known as 4R-tauopathies, although there is no significant increase in soluble 4R tau protein. It has also been shown that in PSP and CBD patients, there is an increase in exon 10+ (4R) *MAPT* mRNA transcripts in regions of the brain most severely affected by neurodegeneration (3,4). In AD, the NFTs are composed of equal amounts of 4R and 3R tau, and in PiD the tau is mainly of the 3R form. Alterations in the ratio of wild-type 4R to 3R *MAPT* transcripts have been found in families with inherited FTD and parkinsonism linked to chromosome 17 (FTDP-17) (5,6) in which splice-site mutations clustered in the intronic sequences immediately following exon 10 generate transcripts constitutively including exon 10 from the mutant allele (5–7).

*To whom correspondence should be addressed. Tel: +44 01865287761; Fax: +44 01865287501; Email: richard.wade-martins@well.ox.ac.uk

More recently, a shift in 4R:3R ratios in favour of 4R *MAPT* mRNA transcripts has been found in mild cognitive impairment and AD (8,9), suggesting that disruption in 4R:3R ratios may underlie a broad range of neurodegeneration disease.

Two haplotypes exist at the *MAPT* locus, H1 and H2, of which H1 is significantly over-represented most notably in PSP and CBD (10–12). More recently, H1 variants have been found associated with AD (13) and Parkinson's disease (PD) (14). Linkage disequilibrium (LD) across the region is very high and the two haplotype blocks are correspondingly large, most recently estimated as between 1 and 2 Mb in size (15,16). The cause of the high LD has been recently shown to be the presence of a 900 kb chromosomal inversion containing the *MAPT* locus, in which the H2 *MAPT* locus is inverted relative to H1 (17). Interestingly, the inversion on the H2 chromosome has been shown to be under positive selection in Europeans (17).

It is clear, therefore, that haplotype-specific sequence variation exists at critical points across the *MAPT* locus. We are interested in the hypothesis that polymorphisms within the *MAPT* H1 haplotype sequence are responsible for subtle, higher levels of exon 10+ (4R) *MAPT* transcript expression from H1 chromosomes, leading, over time, to a greater susceptibility to neurodegenerative disease. Allele-specific, or haplotype-specific, gene expression represents a powerful tool to quantify the effect of sequence variation on transcript expression (18,19). Analysis of allele-specific gene expression within a heterozygous sample provides internal control against confounding environmental variables that would otherwise exist between samples. Here, we apply the methodology to investigate haplotype-specific differences both in total *MAPT* transcript expression, and specifically in exon 10+ (4R) *MAPT* expression. Importantly, any effect of underlying genetic polymorphic variation on gene expression and splicing is independent of disease status and can be performed either in human neuronal cell lines, or in control post-mortem brain samples, of a known genotype.

Here, we analyse *MAPT* expression in heterozygous human model neuronal cultures, and in post-mortem brain tissue, where the level of transcript expression between the H1 and H2 chromosomes can be directly compared in the same sample. We find that the H1 haplotype expresses significantly more exon 10+ (4R) *MAPT* transcript than H2. The deviation in expression from the expected 1:1 ratio is greatest in the globus pallidus, a region consistently affected by tau pathology, where the maximum observed exon 10+ (4R) *MAPT* H1:H2 transcript ratio is 1.425, (SD = 0.205, $P < 0.0001$). These data suggest a mechanism for the increased susceptibility of H1 carriers to neurodegeneration and may explain why the H2 haplotype, which has been shown to confer a protective effect against 4R tauopathy, is under positive selection in Caucasian populations.

RESULTS

Allele-specific expression assay design

Polymorphisms present in coding sequences can be used as a tag to distinguish between allelic transcripts and are required

in order to undertake allele-specific expression assays in a heterozygous sample, such as a cell line or in brain tissue. We used three previously defined (10) coding single nucleotide polymorphisms (SNPs) in the *MAPT* gene as the basis for our haplotype-specific expression assays. SNP 1 (rs17650901) lies in exon 1 and was used to compare total *MAPT* transcript expression between haplotypes. Two SNPs in exon 9, SNP 9i (rs1052553) and SNP 9ii (rs17652121), were used to specifically assay haplotype-expression of exon 10+ (4R) *MAPT* mRNA. All three of the SNPs have been previously shown to be in complete LD with the 238 bp insertion/deletion polymorphism in intron 9 used in this study to define the H1 and H2 haplotypes (10) (Fig. 1A).

The allele-specific gene expression methodology utilizes a reverse transcription (RT)-PCR reaction, followed by a primer extension reaction using either a forward or reverse extension primer, and then analysis of the base extension products by matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-ToF MS; see Materials and Methods). The ratio of H1 to H2 *MAPT* transcripts can be measured in either the total *MAPT* mRNA pool, or specifically in the exon 10+ (4R) *MAPT* mRNA pool, depending on the initial RT-PCR primers selected. Details of all the DNA oligonucleotides used in the study are shown in Figure 1 and Table 1. The methodology has significant advantages over quantitative RT-PCR (Q-RT-PCR) because Q-RT-PCR requires that certain assumptions must be made regarding the expression of control genes used to normalize the measurements.

Cell culture models

We first screened a panel of 14 genetically independent human neuronal and three commonly used non-neuronal cell lines for the presence (H1) or absence (H2) of a 238 bp insertion/deletion polymorphism in intron 9 previously used to define the H1 and the H2 haplotypes. Two neuronal lines, SK-N-F1 and SK-N-MC, were found to be H1/H2 heterozygotes, with the remainder being H1/H1 homozygotes (Table 2).

MAPT H1/H2 heterozygous cell lines SK-N-F1 and SK-N-MC were differentiated to induce a neuronal-like phenotype using a two-stage neuronal differentiation protocol originally optimized for use with SH-SY-5Y neuroblastoma cells (20). Briefly, cells plated at low density were first grown in medium supplemented with 10 μ M retinoic acid. After 5 days, the retinoic acid supplement was removed and the cells were grown in neurobasal (NB) medium containing 2 mM dibutyryl cyclic AMP plus 50 ng/ml brain derived-neurotrophic factor (BDNF) for a further 5 days (see Materials and Methods). During this differentiation phase, cells stop dividing and extend long, neuronal-like processes. Immunocytochemistry staining showed the differentiated cells now expressed as neuronal marker proteins such as neurofilament protein heavy chain (NF-H), microtubule associated protein 2 (MAP2) and tau (Fig. 2A). To investigate the suitability of the cell cultures to model a pure neuronal population, we then analysed differentiated SK-N-F1 and SK-N-MC cells for the expression of glial fibrillary acidic protein (GFAP), a marker of glial cells (Fig. 2B). Differentiated SK-N-F1 cells

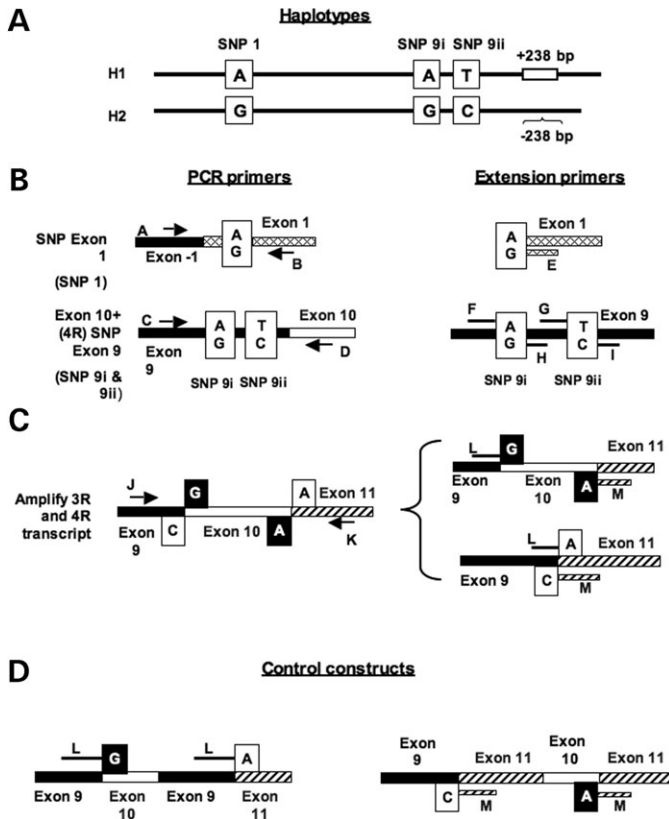


Figure 1. Primer design for MALDI-ToF MS gene expression assays at the human *MAPT* locus. The methodology requires an RT-PCR reaction, followed by a primer extension reaction and analysis by MALDI-ToF MS. Oligonucleotide identification letters refer to the sequences given in Table 1. (A) The coding polymorphisms SNP 1, SNP 9i and SNP 9ii are in complete LD with each other, and with the 238 bp insertion/deletion polymorphism in intron 9 used to define the H1/H2 haplotype. (B) The RT-PCR primers used to amplify fragments containing the coding SNPs are shown on the left, and the locations of the extension primers used for the single base extension reaction extending to the SNP are shown on the right. The locations of the coding SNPs and the genotypes of the two alternative alleles are shown in boxes, with the H1 allele shown above the H2 allele. (C) The locations of the RT-PCR primers in exons 9 and 11, and the alternative final base of the primer extension reactions extending into alternatively spliced exons, are shown. (D) Structure of artificial control constructs built containing fragments of exons 9, 10 and 11 to normalize the MALDI-ToF data to the expected 1:1 ratio of 4R:3R.

did not express GFAP, whereas SK-N-MC cells did, suggesting SK-N-F1 cells represent the better model of gene expression in human neurons.

Haplotype-specific expression in neuronal cell cultures

cDNA prepared from SK-N-F1 and SK-N-MC cells differentiated in three separate experiments was used as a template for PCR amplification to analyse allele-specific expression of *MAPT* using the haplotype-specific coding polymorphisms SNP 1, SNP 9i and SNP 9ii. Comparing total *MAPT* expression from the H1 and H2 chromosomes using the SNP in the constitutive exon 1 showed a significantly higher level of *MAPT* H1 transcript expression in SK-N-F1 ($P < 0.01$), and, to a much greater extent, in SK-N-MC ($P < 0.0001$) by one sample *t*-test (Fig. 3). Performing the initial RT-PCR reaction

using a reverse primer in exon 10 specifically sampled from the exon 10+ (4R) *MAPT* transcript pool revealed a significantly greater expression from the H1 chromosome using both SNP 9i and SNP 9ii in both cell lines ($P < 0.0001$ in all cases by one sample *t*-test) (Fig. 3). At SNP 9i and SNP 9ii, it was possible to analyse the transcript ratios present by performing the base extension reaction on the RT-PCR product using both the forward and reverse primers. Both the forward and reverse primers independently show higher expression of exon 10+ (4R) *MAPT* transcript from H1 for both exon 9 SNPs in both cell lines (data not shown).

Together, these data show that in two independent neuronal cell culture models of *MAPT* endogenous gene expression, the neurodegeneration-associated *MAPT* H1 haplotype expresses a significantly greater amount of exon 10+ (4R) *MAPT* mRNA.

MAPT expression in post-mortem brain tissue

We then extended our study to human post-mortem brain samples. Forty control cases with no evident pathological signs of neurodegenerative disease (Supplementary Material, Table S1) from the Brain Bank held by the Oxford Project to Investigate Memory and Ageing (OPTIMA) were genotyped using the H1/H2 insertion/deletion polymorphism in intron 9. Out of the 40 control brains, 14 (35%) were identified as being from H1/H2 heterozygotes, from which we obtained frontal cortex (BA46) and globus pallidus frozen tissue. The globus pallidus and frontal cortex were chosen for gene expression analysis as regions severely affected by NFTs in the 4R-tauopathies PSP and CBD, respectively.

cDNA prepared from both brain regions of the 14 controls was used as a template for PCR amplification followed by primer extension and MALDI-ToF MS analysis. Allele-specific expression analysis of total *MAPT* transcript using SNP 1 is shown in Figure 4A. We then compared haplotype-specific expression of the neurodegeneration-associated exon 10+ (4R) *MAPT* mRNA transcript using SNP 9i and SNP 9ii (Fig. 4B and C). The means and standard deviations of the ratios of *MAPT* H1:H2 transcript expression for all fourteen brains are summarized in Figure 4D.

The mean *MAPT* H1:H2 transcript ratio for the 14 brains for total *MAPT* expression showed no allelic difference in expression (frontal cortex mean *MAPT* H1:H2 transcript ratio = 1.016, SD = 0.067, $P = 0.388$ by one sample *t*-test, and for globus pallidus mean *MAPT* H1:H2 transcript ratio = 1.029, SD = 0.079, $P = 0.193$ by one sample *t*-test). However, the mean *MAPT* H1:H2 transcript ratio for the fourteen brains for exon 10+ (4R) *MAPT* mRNA transcript shows a significant allelic difference in expression towards H1. In the globus pallidus, a region of the brain known to be highly susceptible to neurodegeneration in PSP, the more common 4R-tauopathy, there is the largest deviation in mean allelic expression, with H1 contributing on an average 1.30 to 1.43-fold more exon 10+ (4R) *MAPT* mRNA than H2 [at SNP 9i the mean exon 10+ (4R) *MAPT* H1:H2 transcript ratio = 1.425, SD = 0.205, $P < 0.0001$ by one sample *t*-test; at SNP 9ii exon 10+ (4R) *MAPT* H1:H2 transcript ratio = 1.296, SD = 0.213, $P = 0.0002$ by one sample

Table 1. Sequences of DNA oligonucleotides used for quantitative analysis of haplotype-specific gene expression by MALDI-ToF MS. The oligonucleotide ID letters correspond to the oligonucleotide ID letters in Figure 1

Oligonucleotide	ID	DNA sequence
<i>MAPT</i> exon 1 SNP forward RT-PCR primer	A	CTT CTC CTC CTC CGC TGT C
<i>MAPT</i> exon 1 SNP reverse RT-PCR primer	B	CGT GAT CTT CCA TCA CTT CG
<i>MAPT</i> exon 9i and 9ii SNP forward RT-PCR primer	C	CTC CAA AAT CAG GGG ATC G
<i>MAPT</i> exon 9i and 9ii SNP reverse RT-PCR primer, specific for exon 10+ transcripts	D	CAC ACT TGG ACT GGA CGT TG
<i>MAPT</i> exon 1 SNP reverse MALDI-ToF extension primer	E	GCC ATC CTG GTT CAA AG
<i>MAPT</i> exon 9i SNP forward MALDI-ToF extension primer	F	GAG CCC AAG AAG GTG GC
<i>MAPT</i> exon 9i SNP reverse MALDI-ToF extension primer	G	TGG AGT ACG GAC CAC
<i>MAPT</i> exon 9ii SNP forward MALDI-ToF extension primer	H	CCG ATC TTG GAC TTG AC
<i>MAPT</i> exon 9ii SNP reverse MALDI-ToF extension primer	I	ATG CCA GAC CTG AAG AA
<i>MAPT</i> exon 9 forward RT-PCR primer	J	GGA GCC CAA GAA GGT GGC
<i>MAPT</i> exon 11 reverse RT-PCR primer	K	TGC TCA GGT CAA CTG GTT TG
<i>MAPT</i> exon 9 forward into exon 10 or 11 MALDI-ToF extension primer	L	GGA GGC GGG AAG GTG CA
<i>MAPT</i> exon 11 reverse into exon 9 or 10 MALDI-ToF extension primer	M	GGT TTG TAG ACT ATT TGC AC

Table 2. Genotyping summary for a panel of 14 genetically independent human neuronal and three widely used non-neuronal cell lines using the H1/H2-defining 238 bp insertion/deletion marker within intron 9. Lines SK-N-F1 and SK-N-MC were determined to be H1/H2 heterozygotes

Human cell line	Cell description	Genotype
Neuronal		
H4	Neuroglioma	H1/H1
SK-N-DZ	Neuroblastoma	H1/H1
SK-N-F1	Neuroblastoma	H1/H2
SK-N-AS	Neuroblastoma	H1/H1
BE (2)-M17	Neuroblastoma	H1/H1
SH-SY-5Y	Neuroblastoma	H1/H1
NT2	Neuronal precursor-like cells	H1/H1
HCN1A	Cortical neurons	H1/H1
HCN2	Cortical neurons	H1/H1
CHP-212	Neuroblastoma	H1/H1
LA-N-5	Neuroblastoma	H1/H1
Kelly	Neuroblastoma	H1/H1
IMR-32	Neuroblastoma	H1/H1
SK-N-MC	Neuroblastoma	H1/H2
Non-neuronal		
293 HEK	Embryonic kidney	H1/H1
HeLa	Carcinoma	H1/H1
MRC5-V2	Lung fibroblast	H1/H1

t-test]. In the frontal cortex, a region affected in the rarer CBD, there is also a highly significant difference in contribution from the two alleles, with H1 contributing on an average 1.16 to 1.29-fold more exon 10+ (4R) *MAPT* mRNA [at SNP 9i exon 10+ (4R) *MAPT* H1:H2 transcript ratio = 1.291, SD = 0.315, $P = 0.006$ by one sample *t*-test; at SNP 9ii exon 10+ (4R) *MAPT* H1:H2 transcript ratio = 1.159, SD = 0.104, $P = 0.0001$ by one sample *t*-test].

In summary, the data show first, that the strength of the *MAPT* genomic DNA promoter does not vary significantly between the H1 and H2 chromosomes in human brain tissue as assayed at SNP 1, and second, that the H1 chromosome expresses a significantly greater amount of exon 10+ (4R) *MAPT* mRNA than H2, with the greatest deviation of up to 1.43-fold found in the globus pallidus and 1.29-fold in the frontal cortex. Our data suggest it is this relative over-expression of exon 10+ (4R) *MAPT* mRNA from H1 which may contribute to the susceptibility to tauopathy. We analysed

expression of exon 10+ (4R) *MAPT* transcripts using forward and reverse primer base extension reactions at SNP 9i and 9ii to measure allelic differences. Importantly, the independent forward and reverse primer data at both SNPs in both brain regions were consistent in showing significant over-expression of exon10+ (4R) *MAPT* transcript from H1 compared to H2 (data not shown).

Recent work has uncovered sequence variation within the H1 haplotype and identified H1 sub-haplotypes associated with PSP (15,21,22), AD (13) and PD (14). We genotyped the H1/H2 brain tissue for SNPs rs1467967, rs242557/htSNP167, rs3785883, rs2471738 and rs7521 to determine the H1 sub-haplotype present (21,22). We first subdivided the brain samples into two groups ('A' allele: $n = 7$; 'G' allele: $n = 7$) based on the SNP rs242557/htSNP167 in intron 0, proposed as an SNP within a transcription-binding site responsible for modulating total *MAPT* transcript expression (22). There was no significant difference between the H1:H2 transcript ratios between the 'A' and 'G' allele samples in globus pallidus ($P = 0.533$) or frontal cortex ($P = 0.069$) by Student's *t*-test. In a second analysis, we subdivided the 14 heterozygote samples based on the presence or absence of the proposed Class II haplotype (21), which includes the two alleles conferring greatest risk to PSP, rs242557 and rs2471738. In this analysis (Class II: $n = 4$; non-Class II: $n = 10$) there was no significant difference between the levels of total *MAPT* H1:H2 transcript ratios between the Class II and non-Class II haplotype samples in the globus pallidus ($P = 0.554$) or frontal cortex ($P = 0.269$) by Student's *t*-test.

***MAPT* splice ratios in different brain regions**

These data suggest that allele-specific differences of *MAPT* exon 10 expression in the brain, particularly in the globus pallidus, contribute to the susceptibility to suffering from 4R-tauopathies. We used the MALDI-ToF MS platform to further investigate the ratio of exon 10+ (4R) to exon 10- (3R) *MAPT* transcript levels in the globus pallidus and frontal cortex using the assay shown in Figure 1B and C. We found that in 13 of the 14 (93%) control brains, the ratio of *MAPT* 4R:3R transcripts was significantly higher in the globus pallidus than the frontal cortex ($P < 0.0001$ in

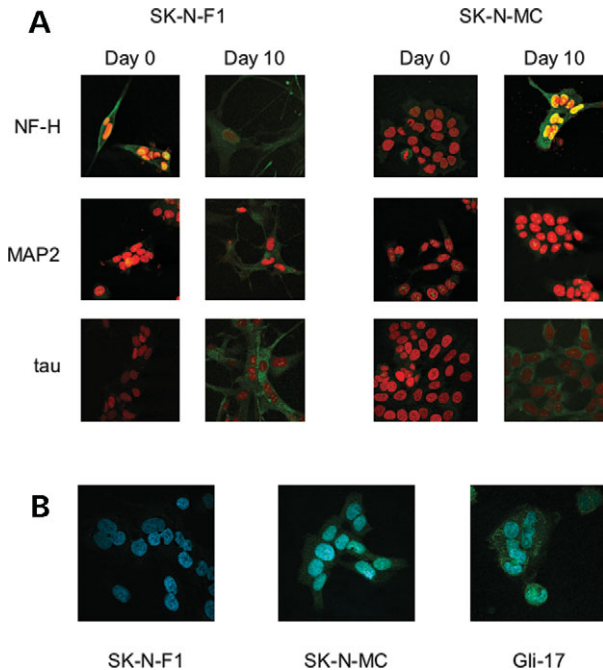


Figure 2. SK-N-F1 and SK-N-MC neuronal differentiation. (A) Immunocytochemistry analysis in undifferentiated and differentiated cells of the expression of neuronal markers NF-H, MAP2 and tau (green), counter-stained with propidium iodide nuclear stain (red). After the 10 day differentiation process both cell lines extend long, neuronal-like processes and express neuronal markers. (B) Immunocytochemistry analysis of the expression of the glial cell marker GFAP (green) counter-stained with the nuclear stain DAPI (blue). Differentiated SK-N-F1 cells do not express GFAP, whereas SK-N-MC cells do, suggesting SK-N-F1 is a better human cell culture model of neuronal gene expression.

each individual brain by one sample *t*-test), with the ratio ranging from 1.3 to 3.5 (mean = 1.652, SD = 0.6180) (Fig. 5). This higher 4R:3R *MAPT* transcript ratio in the globus pallidus compared to the frontal cortex in control brain tissue provides a further link between *MAPT* gene expression and the greater susceptibility of the globus pallidus to accumulating 4R-tau NFTs compared to the frontal cortex.

DISCUSSION

Functionally important sequence variation in the human genome lying outside coding regions is likely to be critically important in the understanding of human disease. As we rapidly accumulate DNA sequence data for the human genome, it is becoming increasingly important to utilize methodologies to help us uncover the effect of polymorphic variation on gene expression. Allele-specific, or haplotype-specific, gene expression represents a powerful tool to quantify the effect of genomic variation on gene expression.

Traditionally, candidate gene analysis has focused on coding variants. However, several recent studies examining susceptibility to neurological disorders have found strong association with disease of specific alleles or haplotypes lacking obvious changes in protein-coding sequences. In such cases, detailed investigation has focused on variation in gene expression regulated by non-coding polymorphisms.

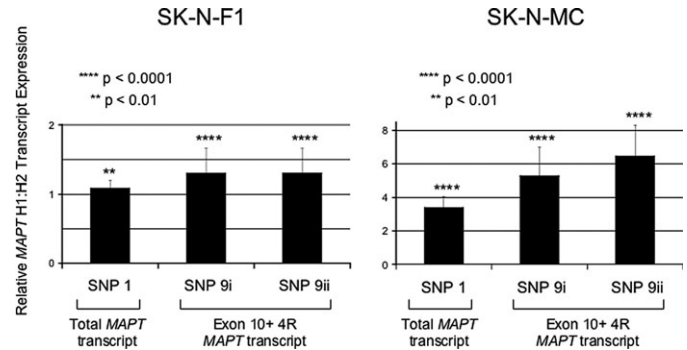


Figure 3. Allele-specific *MAPT* expression in H1/H2 heterozygous differentiated SK-N-F1 and SK-N-MC neuronal cell lines. The expression data is normalized to the ratio of each SNP present on genomic DNA to produce a standardized transcript ratio. A ratio of >1 at SNP 1 indicates more total *MAPT* transcript from H1 than H2; a ratio of >1 at SNP 9i and SNP 9ii indicates more exon 10+ (4R) *MAPT* transcript from H1 than H2. The differentiation process was independently repeated three times and each allele-specific expression analysis was performed in eight (SNP 1) or 16 (SNP 9i and SNP 9ii) replicates as described in Materials and Methods. SNP1: SK-N-F1 *MAPT* H1:H2 transcript ratio 1.082, SD 0.118, $P = 0.0025$; SK-N-MC *MAPT* H1:H2 transcript ratio 3.385, SD 0.666, $P < 0.0001$; SNP9i: SK-N-F1 exon 10+ (4R) *MAPT* H1:H2 ratio 1.294, SD 0.363, $P < 0.0001$; SK-N-MC exon 10+ (4R) *MAPT* H1:H2 transcript ratio 5.258, SD 1.724, $P < 0.0001$; SNP9ii: SK-N-F1 exon 10+ (4R) *MAPT* H1:H2 transcript ratio = 1.297, SD 0.358, $P < 0.0001$; SK-N-MC exon 10+ (4R) *MAPT* H1:H2 transcript ratio 6.442, SD 1.875, $P < 0.0001$. All *P*-values determined by one sample *t*-test.

A study investigating the genetic basis of dyslexia found association to a 77 kb region on chromosome 6p22 (23). One risk haplotype was identified within this region that contained no coding variants. Further investigation revealed ~40% relative reduction in expression of the *KIAA0319* gene from this risk haplotype (24). Allele-specific expression analysis has also been applied at two loci that have been identified as candidates conferring susceptibility to schizophrenia: catechol-O-methyltransferase (*COMT*) (25) and dystrobrevin binding protein 1 (*DTNBP1*) (26). At the *COMT* locus on chromosome 22q11, mRNA transcripts bearing an allele associated with schizophrenia showed an average of 17% relative reduction in expression (25). At the *DTNBP1* locus on chromosome 6p22, the schizophrenia-associated risk haplotype showed 21% less expression of *DTNBP1* compared to chromosomes bearing either one of two non-risk haplotypes (26). A study on promoter polymorphisms in the GABA_A receptor beta3 subunit (*GABRB3*) gene used *in vitro* luciferase reporter assays to demonstrate a 40% reduction of *GABRB3* from a promoter haplotype associated with childhood absence epilepsy (CAE) (27). Finally, a study of expression of *APOE* in human brain tissue revealed a 15–25% increase in expression of the $\epsilon 4$ allele of the *APOE* gene, the strongest risk factor for late onset AD (28). In each case discussed above, the subtle changes in relative gene expression between risk and non-risk alleles have been suggested as a functional mechanism whereby a risk haplotype confers susceptibility to disease.

The genetic association of the *MAPT* sequence variation with PSP and CBD is well-established and has been widely replicated since the first studies (11,29). The associations with PD (14,30–33) and AD (13) are more recent, but do, however, support recent suggestions that genetic variation in

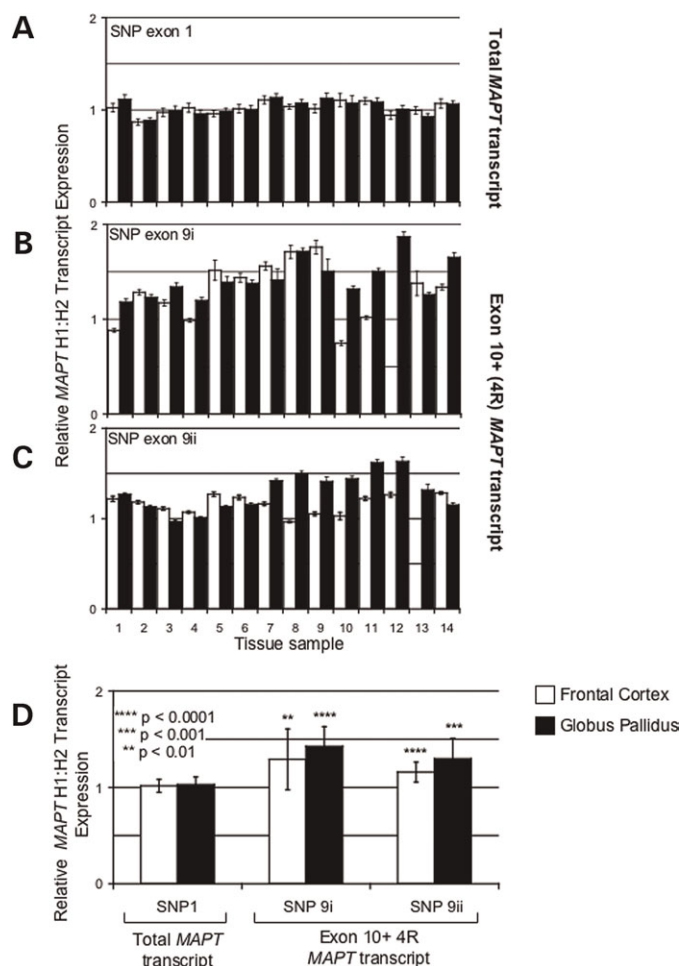


Figure 4. Allele-specific *MAPT* expression in human post-mortem brain tissue from H1/H2 heterozygous controls. The expression data is normalized to the ratio of each SNP present on genomic DNA to produce a standardized transcript ratio. A ratio of >1 indicates more H1-specific transcript than H2. (A) Total *MAPT* transcript in each of the 14 H1/H2 heterozygotes measured using SNP 1. Mean values and standard deviations from eight measurements using the reverse primer base extension reaction are shown. (B and C) Allele-specific expression in each of the 14 H1/H2 heterozygotes of exon 10+ (4R) *MAPT* transcript measured using SNP 9i (B) and SNP 9ii (C) in exon 9. Mean values and standard deviations from up to 16 measurements (eight measurements using the forward primer and eight measurements using the reverse primer) are shown. (D) Mean values and standard deviations of relative H1:H2 expression averaged from all 14 H1/H2 heterozygotes. No significant difference is seen in the ratio of total *MAPT* transcript at SNP 1 (mean *MAPT* H1:H2 transcript ratio for frontal cortex 1.016; SD 0.067; $P = 0.388$, and for globus pallidus mean *MAPT* H1:H2 transcript ratio 1.029; SD 0.079; $P = 0.193$), however, highly significant differences exist in allelic expression of exon 10+ (4R) *MAPT* as measured by both SNPs in exon 9 in globus pallidus [SNP 9i mean exon 10+ (4R) *MAPT* H1:H2 transcript ratio 1.425, SD 0.205, $P < 0.0001$; SNP 9ii exon 10+ (4R) *MAPT* H1:H2 transcript ratio 1.296, SD 0.213, $P = 0.0002$ and frontal cortex; SNP 9i exon 10+ (4R) *MAPT* H1:H2 transcript ratio 1.291, SD 0.315, $P = 0.006$; SNP exon 10+ (4R) *MAPT* 9ii H1:H2 transcript ratio 1.159, SD 0.104, $P = 0.0001$]. All P -values calculated by one sample t -test.

expression and splicing at the *MAPT* locus may be critical to a wide range of neurodegenerative disorders (34). Finally, recent findings suggesting aberrations in 4R:3R ratios in mild-cognitive impairment and AD (8,9) support the involvement of *MAPT* from the earliest stages of disease.

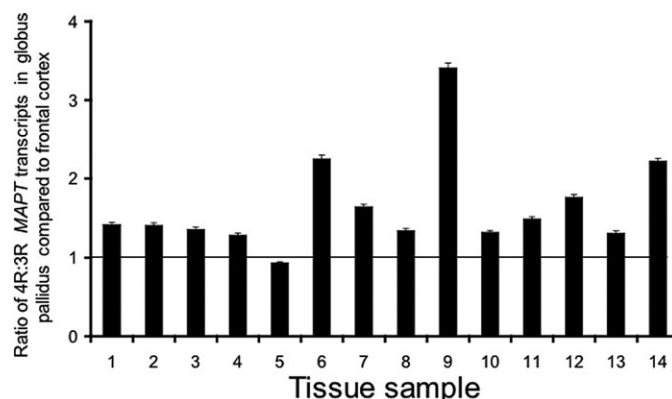


Figure 5. The *MAPT* 4R:3R transcript ratio is greater in the globus pallidus than in the frontal cortex in 13 of the 14 (93%) control brains. With only one exception, the ratio of 4R:3R transcripts was significantly higher in the globus pallidus than the frontal cortex ($P < 0.0001$ in each individual brain by one sample t -test), with the ratio of transcripts ranging from 1.3 to 3.5 (mean 1.652, SD 0.618). The frontal cortex and globus pallidus splicing data were normalized to a ratio of one using the control constructs in Figure 1C. A value of 1 indicates that the 4R:3R ratios in the globus pallidus and frontal cortex are equal. For frontal cortex and globus pallidus measurements 16 assays were performed (eight using the forward primer and eight with the reverse primer) to calculate the mean and standard deviation.

There are two potential mechanisms by which haplotype-specific expression at the *MAPT* locus may confer susceptibility to tauopathy: there may be allelic variation in total *MAPT* expression, or in expression of a disease-susceptibility variant. Our study has used allele-specific gene expression assays in H1/H2 heterozygous neuronal cell cultures and post-mortem brain tissue to test the specific hypothesis that H1 chromosomes express a greater amount of exon 10+ (4R) *MAPT* mRNA than H2 chromosomes. We show that the H1 chromosome makes a greater contribution, by up to 1.43-fold in the globus pallidus and 1.29-fold in the frontal cortex, specifically to the disease-critical exon 10+ (4R) *MAPT* transcript pool, but that total levels of *MAPT* expression are not significantly different between H1 and H2. We identify one *MAPT* H1/H2 heterozygous cell line in particular, SK-N-F1, as being an excellent model of neuronal gene expression expressing only neuronal markers post-differentiation. We show there is good consistency both within the SK-N-F1 cell line, with SNP 9i and SNP 9ii both showing a 1.3-fold over-expression of exon 10+ (4R) *MAPT* mRNA from the H1 chromosome, and between the cell line and the human post-mortem brain tissue. The close fit of the SK-N-F1 neuronal cell data and the human post-mortem tissue data suggest that the differences we observe in the brain tissue may come from haplotype-specific differences in *MAPT* transcript splicing in neurons.

We propose, therefore, that greater expression of exon 10+ (4R) *MAPT* mRNA from the H1 chromosome leads to subtle, higher levels of exon 10+ (4R) *MAPT* transcript in H1 carriers resulting in a greater susceptibility to neurodegenerative disease. The H2 chromosome expresses relatively less 4R *MAPT* mRNA, which is consistent with many studies reporting it to be a protective factor against PSP, and may explain why it is spreading under positive selection in Caucasian populations (17).

Previous studies (22,33) have used reporter gene assays to demonstrate differences in the levels of expression driven by small promoter fragments of the *MAPT* gene. The data obtained in such studies should be interpreted cautiously because the experimental design relies upon using small promoter fragments isolated from their correct genomic context. One study (33) used a ~1 kb fragment of promoter region and found that the H2 promoter was weaker than H1 in 293 kidney cells and SK-N-MC cells. Their result in SK-N-MC cells is consistent with our data showing a high relative over-expression of H1, although this may not reflect the situation *in vivo*. A second paper (22) reported results from a luciferase reporter gene study in which a small 182 bp fragment containing htSNP167 from deep within intron 0 was tested for enhancer activity. Surprisingly, the non-risk variant was found to be more transcriptionally active, although the biological relevance of this finding is unclear.

The significant advantage of our work compared to reporter gene studies using short fragments of DNA is that we are comparing expression of H1 and H2 transcripts from the actual *MAPT* genomic DNA locus in its correct chromosomal location in an internally controlled experiment. By analysing gene expression from the complete and intact *MAPT* locus, we have been able for the first time to study the effect of *MAPT* haplotype on the ratio of expression of 4R:3R transcript. We propose it is the haplotype-specific difference in exon 10+ (4R) *MAPT* expression, rather than a haplotype-specific difference in *MAPT* promoter activity that underlies the mechanism by which H1 contributes to disease susceptibility. Finally, the experimental design of a within-sample analysis between two loci of a heterozygous sample eliminates confounding environmental factors that would create a high level of noise in a between-sample analysis of post-mortem brain tissue, such as co-morbidity, age at death, cause of death and sample quality.

We find strong evidence for allelic differences in expression of exon 10+ (4R) *MAPT* mRNA between the H1 and H2 chromosomes. It is also possible that there may be more subtle differences in expression between different sub-haplotypes of H1, but that our sample size is too small to be further subdivided to detect small differences. Importantly, the SNPs defining the H1 sub-haplotypes do not include a coding SNP, which means it is not possible to perform a direct comparison using the method of analysis we describe.

We observed the most dramatic relative over-expression of exon 10+ (4R) *MAPT* mRNA from H1 in the globus pallidus, a region of the basal ganglia severely affected by NFTs in PSP, and also to a lesser extent in CBD. We showed that in 13 of the 14 brains with no overt neuropathology, the 4R:3R *MAPT* transcript ratio was higher in the globus pallidus than frontal cortex. This provides a further link at the molecular level with pathology and may explain why the globus pallidus is more susceptible to suffering a build-up of NFTs than the frontal cortex.

The human *MAPT* locus spans 130 987 bp from first to last exon. Complete sequence available for both H1 and H2 chromosomes shows that to be approximately one SNP approximately every 200 bp, and an insertion/deletion every 1 kb (www.genome.ucsc.edu). As with other complex disease loci it remains an enormous challenge to attempt to

define which of the polymorphic variants is responsible for differences in gene expression observed between haplotypes in *MAPT*. Work in our laboratory is now focused on using novel bacterial artificial chromosomes (BAC)-based genomic DNA expression systems to dissect the functional importance of sequence variation at the *MAPT* locus (35,36).

MATERIALS AND METHODS

Cell culture

The following neuronal cell lines were purchased from the European Collection of Cell Cultures (ECACC): SH-SY-5Y (ECACC no. 94030304), SK-N-AS (ECACC no. 94092302), SK-N-DZ (ECACC no. 94092305), SK-N-F1 (ECACC no. 94092304), IMR-32 (ECACC no. 86041809), SK-N-MC (ECACC no. 90022302) and KELLY (ECACC no. 922110411). The following neuronal cell lines were purchased from the American Tissue Culture Collection (ATCC): HCN1A (ATCC no. CRL10442) and HCN2 (ATCC no. CRL10742). Cell line LA-N-5 was a generous gift of R.C. Seeger (Children's Hospital, Los Angeles, CA, USA), cell line CHP-212 was a generous gift of P. Forgez (Hopital Saint-Antoine, Paris, France), cell line BE (2)-M17 was a generous gift of J. Bakowska (NIH, Bethesda, MD, USA), cell line H4 was a generous gift of M. Ingelsson (Massachusetts General Hospital, Boston, MA, USA), cell line NT2 was a generous gift of D. Pleasure (Children's Hospital of Philadelphia, Philadelphia, PA, USA), and the glial cell line Gli-17 was a kind gift of David Louis (Massachusetts General Hospital, Boston, MA, USA). Cell lines were cultured following suppliers' instructions specific to each line. DNA was extracted from cells by a standard proteinase K digestion/phenol:chloroform extraction protocol.

Cell lines SK-N-F1 and SK-N-MC were differentiated into neuronal-like cells following a two-stage protocol originally developed for SHSY5Y cells and was performed as previously described (20). Cells were plated at low density ($3-5 \times 10^3$ cells/cm²) in their recommended growth media. After 24 h, the growth media for SK-N-F1 and SK-N-MC was changed to DMEM and EMEM (EBSS), respectively, supplemented with 5% fetal calf serum, 2 mM L-glutamine, 100 U/ml penicillin and 0.1 mg/ml streptomycin and 10 μ M retinoic acid. The whole media was changed every 48 h. After 5 days, medium containing retinoic acid was removed and replaced with NB medium supplemented with B-27 (Invitrogen), 2 mM GlutaMax I (Invitrogen), 20 mM KCl, 2 mM dibutyryl cyclic AMP (Sigma) plus 50 ng/ml BDNF (Sigma). Cells were cultured in this medium for a further 5 days before staining by immunocytochemistry using Tau-1 (Chemicon), NF-H (Sigma), MAP2 (Chemicon), and GFAP (Promega) antibodies. Slides were mounted using Vectashield mounting medium with propidium iodide or Vectashield mounting medium with DAPI (Vector Laboratories). The secondary antibody Alexa Fluor 488 goat anti-mouse IgG (Molecular Probes) was used with the Tau-1 and NF-H antibodies, and Alexa Fluor 488 goat anti-rabbit IgG was used with the MAP2 and GFAP antibodies. Total RNA was extracted from SK-N-F1 and SK-N-MC cells from three independent differentiation experiments using Trizol reagent (Invitrogen)

followed by an additional purification step using an RNeasy Mini Kit (Qiagen).

Post-mortem brain material

Anonymous brain tissue from frontal cortex (BA46) and globus pallidus of 40 fresh-frozen control brains with no evident pathological signs of neurodegenerative disease (Supplementary Material, Table S1) was obtained from the Brain Bank of the OPTIMA. All OPTIMA brains are collected under ethically approved conditions with full consent of the patient and with approval of the local Ethics Committee (COREC approval number 1656). The analysis of gene expression undertaken here has also been approved by local Ethics Committee review, reference 06/Q1605/8. The mean age at death of the 14 H1/H2 heterozygous individuals studied for *MAPT* expression was 82 years (SD 8.7). DNA was extracted from frontal cortex material using a Nucleon Genomic DNA Extraction for hard tissue (Tepnel Life Sciences). Total RNA was extracted from frontal cortex and globus pallidus post-mortem brain samples cells by initially homogenizing the samples in Trizol reagent (Invitrogen) followed by a column purification using the RNeasy Mini Kit (Qiagen).

MAPT genotyping

MAPT H1/H2 genotyping was performed on brain samples and cell lines using the 238 bp insertion/deletion polymorphism in intron 9 previously shown to be an unambiguous tag defining H1 and H2. Previously described (10) PCR primers 5'-GGAAGACGTTCTCACTGATCTG-3' (forward) and 5'-AGGAGTCTGGCTTCAGTCTCTC-3' (reverse) flanking the insertion/deletion were used to genotype all cell lines and brain tissue. Typing of the H1/H2 control brain samples for SNPs rs1467967, rs242557, rs3785883, rs2471738 and rs7521 was performed as previously described (15).

Allele-specific expression analysis

RNA (1 µg) prepared from cultured cells or post-mortem brain material was reverse transcribed into cDNA using Superscript First Strand Synthesis Kit (Invitrogen) in a 20 µl reaction. For PCR amplification, 0.5 µl of cDNA was used for PCR amplification using AmpliTaq Gold DNA polymerase (Applied Biosystems) and *Pfu* Turbo DNA Polymerase (Stratagene). The specific location at the *MAPT* locus of each RT-PCR primer used is shown in Figure 1 and the complete sequences are given in Table 1.

Quantitation of haplotype-specific gene expression

DNA fragments amplified by PCR were treated by exonuclease I (New England BioLabs) to remove residual primers, followed by incubation with shrimp alkaline phosphatase (Amersham Biosciences) to dephosphorylate unincorporated deoxynucleotides. The fragments were then used as templates for a primer extension reaction using a mixture of three dideoxynucleotides and one deoxynucleotide in the reaction. The primer extension reaction products were then analysed by MALDI-ToF MS using the MassARRAY system on

Sequenom platform. Each MALDI-ToF MS assay was repeated eight times, comprising four independent RT-PCR reactions, each analysed in duplicate. Three well-characterized haplotype-specific SNPs in complete LD with the 238 bp insertion/deletion polymorphism in intron 9 (10) used to define H1 and H2 were used to identify chromosome-specific transcripts and then quantify their relative amounts (Fig. 1). For SNP 1 (rs17650901), each assay was performed eight times using the reverse extension primer. For SNP 9i (rs1052553) and SNP 9ii (rs17652121) each assay was performed eight times using both the forward and reverse extension primers as independent assays and a mean of the 16 readings was taken, with three exceptions for SNP 9i in which only the reverse extension primer was used. The MassARRAY data were processed by the Sequenom SpectroTYPER software to obtain the area under the primer extension product trace curves. Genomic DNA was used to produce a standardized ratio equivalent to equal expression from each chromosome. In the case of assays to measure ratios of 4R:3R splice variants, artificial constructs were prepared containing fragments of exons 9, 10 and 11. Each construct contains a single copy of exon junctions and was therefore used as the normalizing control (Fig. 1). Statistical analysis was performed as previously described (24) using a one sample *t*-test that calculates a *P*-value indicating the significance of the observed allelic differences in expression from the expected ratio of 1.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG Online.

ACKNOWLEDGEMENTS

This work was supported by a Research Career Development Fellowship awarded to R.W.-M. from the Wellcome Trust and by a Clarendon Fund Bursary awarded to TMC. The authors would like to thank Emma Banfield and Jennifer Taylor for advice on statistical analysis, Carolyn Sloane for sample preparation, David Smith for access to OPTIMA resources, and Clyde Francks and Anthony P. Monaco for frequent help and discussion.

Conflict of Interest statement. None declared.

REFERENCES

- Goedert, M. and Spillantini, M.G. (2000) Tau mutations in frontotemporal dementia FTDP-17 and their relevance for Alzheimer's disease. *Biochimica et Biophysica Acta*, **1502**, 110–121.
- Dickson, D. (2004) Sporadic tauopathies: Pick's disease, corticobasal degeneration, progressive supranuclear palsy and argyrophilic grain disease. In Esiri, M.M., Lee, V.M.-Y. and Trojanowski, J.Q. (eds), *The Neuropathology of Dementia*, Cambridge University Press.
- Chambers, C.B., Lee, J.M., Troncoso, J.C., Reich, S. and Muma, N.A. (1999) Overexpression of four-repeat tau mRNA isoforms in progressive supranuclear palsy but not in Alzheimer's disease. *Ann. Neurol.*, **46**, 325–332.
- Takanashi, M., Mori, H., Arima, K., Mizuno, Y. and Hattori, N. (2002) Expression patterns of tau mRNA isoforms correlate with susceptible lesions in progressive supranuclear palsy and corticobasal degeneration. *Mol. Brain Res.*, **104**, 210–219.

5. 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.
6. 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.
7. D'Souza, I., Poorkaj, P., Hong, M., Nochlin, D., Lee, V.M.-Y., Bird, T.D. and Schellenberg, G.D. (1999) Missense and silent tau gene mutations cause frontotemporal dementia with parkinsonism-chromosome 17 type, by affecting multiple alternative RNA splicing regulatory elements. *Proc. Natl. Acad. Sci. (USA)*, **96**, 5598–5603.
8. Yasojima, K., McGeer, E.G. and McGeer, P.L. (1999) Tangled areas of Alzheimer brain have upregulated levels of exon 10 containing tau mRNA. *Brain Res.*, **831**, 301–305.
9. Ginsberg, S.D., Che, S., Counts, S.E. and Mufson, E.J. (2006) Shift in the ratio of three-repeat tau and four-repeat tau mRNAs in individual cholinergic basal forebrain neurons in mild cognitive impairment and Alzheimer's disease. *J. Neurochem.*, **96**, 1401–1408.
10. 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. Gen.*, **8**, 711–715.
11. 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.
12. Houlden, H., Baker, M., Morris, H.R., MacDonald, N., Pickering-Brown, S., Adamson, J., Lees, A.J. and *al.* (2001) Corticobasal degeneration and progressive supranuclear palsy share a common tau haplotype. *Neurology*, **56**, 1702–1706.
13. Myers, A.J., Kaleem, M., Marlowe, L., Pittman, A.M., Lees, A.J., Fung, H.C., Duckworth, J., Leung, D., Gibson, A., Morris, C.M. *et al.* (2005) The H1c haplotype at the MAPT locus is associated with Alzheimer's disease. *Hum. Mol. Gen.*, **14**, 2399–2404.
14. 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.
15. Pittman, A., Myers, A.J., Duckworth, J., Bryden, L., Hanson, M., Abou-Sleiman, P., Wood, N.W., Hardy, J., Lees, A.J. and de Silva, R. (2004) The structure of the tau haplotype in controls and in progressive supranuclear palsy. *Hum. Mol. Gen.*, **13**, 1267–1274.
16. Cruts, M., Rademakers, R., Gijselinck, I., van der Zee, 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. Gen.*, **14**, 1753–1762.
17. Stefansson, H., Helgason, A., Thorleifsson, G., Steinthorsdottir, V., Masson, G., Barnard, J., Baker, A., Jonasdottir, A., Ingason, A. and Gudnadottir, V.G. (2005) A common inversion under selection in Europeans. *Nature Genet.*, **37**, 129–137.
18. Yan, H., Yuan, W., Velculescu, V.E., Vogelstein, B. and Kinzler, K.W. (2002) Allelic variation in human gene expression. *Science*, **297**, 1143.
19. Knight, J.C. (2004) Allele-specific gene expression uncovered. *Trends Genet.*, **20**, 113–116.
20. Encinas, M., Iglesias, M., Liu, Y., Wang, H., Muhasien, A., Cena, V., Gallego, C. and Comella, J.X. (2000) Sequential treatment of SH-SY5Y cells with retinoic acid and brain-derived neurotrophic factor gives rise to fully differentiated, neurotrophic factor-dependent, human neuron-like cells. *J. Neurochem.*, **75**, 991–1003.
21. Pittman, A., 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.*, **42**, 837–846.
22. Rademakers, R., Melquist, S., Cruts, M., Theuns, J., Del-Favero, J., Poorkaj, P., Baker, M., Sleegers, K., Crook, R. and De Pooter, T. (2005) High-density SNP haplotyping suggests altered regulation of tau gene expression in progressive supranuclear palsy. *Hum. Mol. Gen.*, **14**, 3281–3292.
23. Francks, C., Paracchini, S., Smith, S.D., Richardson, A.J., Scerri, T.S., Cardon, L.R., Marlow, A.J., MacPhie, I.L., Walter, J., Pennington, B.F. *et al.* (2004) A 77-kilobase region of chromosome 6p22.2 is associated with dyslexia in families from the United Kingdom and from the United States. *Am. J. Hum. Genet.*, **75**, 1046–1058.
24. Paracchini, S., Thomas, A., Castro, S., Lai, C., Paramasivam, M., Wang, Y., Keating, B.J., Taylor, J.M., Hacking, D.F., Scerri, T. *et al.* (2006) The chromosome 6p22 haplotype associated with dyslexia reduces the expression of KIAA0319, a novel gene involved in neuronal migration. *Hum. Mol. Genet.*, **15**, 1659–1666.
25. Bray, N.J., Buckland, P.R., Williams, N.M., Williams, H.J., Norton, N., Owen, M.J. and O'Donovan, M.C. (2003) A haplotype implicated in schizophrenia susceptibility is associated with reduced COMT expression in human brain. *Am. J. Hum. Genet.*, **73**, 152–161.
26. Bray, N.J., Preece, A., Williams, N.M., Moskvina, V., Buckland, P.R., Owen, M.J. and O'Donovan, M.C. (2005) Haplotypes at the dystrobrevin binding protein 1 (DTNBP1) gene locus mediate risk for schizophrenia through reduced DTNBP1 expression. *Hum. Mol. Genet.*, **14**, 1947–1954.
27. Urak, L., Feucht, M., Fathi, N., Hornik, K. and Fuchs, K. (2006) A GABRB3 promoter haplotype associated with childhood absence epilepsy impairs transcriptional activity. *Hum. Mol. Genet.*, **16**, 2533–2541.
28. Bray, N.J., Jehu, L., Moskvina, V., Buxbaum, J.D., Dracheva, S., Haroutunian, V., Williams, J., Buckland, P.R., Owen, M.J. and O'Donovan, M.C. (2004) Allelic expression of APOE in human brain: effects of epsilon status and promoter haplotypes. *Hum. Mol. Genet.*, **13**, 2885–2892.
29. Conrad, C., Andreadis, A., Trojanowski, J.Q., Dickson, D.W., Kang, D., Chen, X., Wiederholt, W., Hansen, L., Masliah, E. and Thal, L.J. *et al.* (1997) Genetic evidence for the involvement of tau in progressive supranuclear palsy. *Ann. Neurol.*, **41**, 277–281.
30. Scott, W.K., Nance, W.A., Watts, R.L., Hubble, J.P., Koller, W.C., Lyons, K., Pahwa, R., Stern, M.B., Colcher, A., Hiner, B.C. *et al.* (2001) Complete genomic screen in Parkinson's disease: evidence for multiple genes. *JAMA*, **286**, 2239–2244.
31. Martin, E.R., Scott, W.K., Nance, M.A., Watts, R.L., Hubble, J.P., Koller, W.C., Lyons, K., Pahwa, R., Stern, M.B., Colcher, A. *et al.* (2001) Association of single-nucleotide polymorphisms of the tau gene with late-onset Parkinson disease. *JAMA*, **286**, 2245–2250.
32. Pastor, P., Ezquerro, M., Munoz, E., Marti, M.J., Blesa, R., Tolosa, E. and Oliva, R. (2000) Significant association between the tau gene and A0/A0 genotype in Parkinson's disease. *Ann. Neurol.*, **47**, 242–245.
33. Kwok, J.B.J., 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.
34. Hardy, J., Pittman, A., Myers, A.J., Gwinn-Hardy, K., Fung, H.C., de Silva, R., Hutton, M. and Duckworth, J. (2005) Evidence suggesting that Homo neanderthalensis contributed the H2 MAPT haplotype to Homo sapiens. *Biochem. Soc. Trans.*, **33**, 582–585.
35. Wade-Martins, R., White, R.E., Kimura, H., Cook, P. and James, M.R. (2000) Stable correction of a genetic deficiency in human cells by an episome carrying a 115 kb genomic transgene. *Nature Biotechnol.*, **18**, 1311–1314.
36. Wade-Martins, R., Smith, E.R., Tyminski, E., Chiocca, E.A. and Saeki, Y. (2001) An infection transfer and expression system for genomic DNA loci in human and mouse cells. *Nature Biotechnol.*, **19**, 1067–1070.