Available online at www.sciencedirect.com
SCIENCE DIRECT
Gene 341 (2004) 313–322
GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES
www.elsevier.com/locate/gene

Tau gene (MAPT) sequence variation among primates

Max Holzer[a,b,*], Molly Craxton[a], Ross Jakes[a], Thomas Arendt[b], Michel Goedert[a]

[a]*MRC Laboratory of Molecular Biology, Cambridge, UK*
[b]*Department of Neuroanatomy, Paul Flechsig Institute of Brain Research, University of Leipzig, Leipzig, Germany*

Received 28 March 2004; received in revised form 16 June 2004; accepted 19 July 2004
Available online 18 September 2004
Received by F. Salvatore

Abstract

Filamentous tau deposits are a defining feature of a number of human neurodegenerative diseases. Apes and monkeys have been reported to be differentially susceptible to developing tau pathology. Despite this, only little is known about the organisation and sequence of *Tau* from nonhuman primates. Here we have sequenced *Tau* exons 1–13, including flanking intronic regions, and the region in intron 9 that contains *Saitohin* in chimpanzees, gorillas, and gibbons. Partial sequences were obtained for cynomolgus macaque and green monkey. Chimpanzee brain tau was 100% identical to human tau. Identities were 99.5% for gorilla tau and 99.0% for gibbon tau. Chimpanzee DNA was polymorphic for a repeat in intron 9, which was present in human and gorilla tau, and for the nucleotide at position +29 of the intron that follows exon 10. As was the case of the other nonhuman primates examined, chimpanzee DNA was homozygous for nucleotides used to define the H2 haplotype in human *Tau*. These differences between human and chimpanzee *Tau* may contribute to the apparent resistance of chimpanzee brain to developing tau pathology. Sequencing of *Saitohin* revealed an intact open reading frame in chimpanzee and gorilla, but not in gibbon or macaque.

Keywords: Chimpanzee; Gibbon; Gorilla; *Saitohin*; Haplotype; Splicing

1. Introduction

The tauopathies constitute a group of neurodegenerative diseases characterized by the presence of amyloid-like, filamentous deposits made of hyperphosphorylated Tau protein in nerve cells or in both nerve cells and glial cells (Lee et al., 2001; Berriman et al., 2003). They comprise Alzheimer's disease (AD), progressive supranuclear palsy (PSP), corticobasal degeneration (CBD), Pick's disease (PiD), argyrophilic grain disease (AGD), and frontotemporal dementia and parkinsonism linked to chromosome 17 (FTDP-17). In these diseases, the presence of filamentous tau deposits correlates with the loss of nerve cells.

Tau is a neuronally enriched, microtubule-associated protein, which is involved in the nucleation, elongation, and stabilization of microtubules (Lee et al., 2001). In the adult human brain, there are six isoforms of tau, produced from a single gene by alternative mRNA splicing (Goedert et al., 1989; Andreadis et al., 1992). They differ from one another by the presence or absence of a 29- or 58-amino acid insert (encoded by exon 2 and 3, respectively) in the amino-terminal half of the protein and by the inclusion, or not, of a 31-amino acid repeat, encoded by exon 10 of *Tau*, in the carboxy-terminal half of the protein. The exclusion of exon 10 leads to the production of three isoforms, each containing

Abbreviations: Aβ, beta amyloid peptide; bp, base pair(s); chi, chimpanzee; cDNA, DNA complementary to RNA; CNS, central nervous system; DTT, dithiothreitol; E, exon; EDTA, ethylenediaminetetraacetic acid; gib, gibbon; gor, gorilla; hum, human; kDa, kilodalton(s); mac, macaque; OD, optical density; PCR, polymerase chain reaction; PMSF, phenylmethylsulfonyl fluoride; RT-PCR, reverse transcription polymerase chain reaction; SDS-PAGE, sodium dodecyl sulfate polyacrylamide electrophoresis; Tris, tris(hydroxymethyl)aminomethane; UTR, untranslated region.

* Corresponding author. Department of Neuroanatomy, Paul Flechsig Institute of Brain Research, University of Leipzig, Jahnallee 59, D-04109 Leipzig, Germany. Tel.: +49 341 9725759; fax: +49 341 9725729.
E-mail address: holm@medizin.uni-leipzig.de (M. Holzer).

0378-1119/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2004.07.013

three repeats, and its inclusion leads to a further three isoforms, each containing four repeats. The repeats constitute the microtubule-binding region of tau protein. In normal adult human cerebral cortex, there are similar levels of three-repeat and four-repeat tau isoforms (Goedert and Jakes, 1990). In the developing human brain, only the shortest isoform (three repeats and no amino-terminal inserts) is expressed. The longest human tau isoform comprises exons 1–5, 7, 9, and 10–13. In the peripheral nervous system, additional tau isoforms are expressed, which contain a long N-terminal insert encoded by exon 4A (Goedert et al., 1992; Andreadis et al., 1992).

Mutations in *Tau* cause the autosomal dominantly inherited FTDP-17 (Poorkaj et al., 1998; Hutton et al., 1998; Spillantini et al., 1998). Functionally, many mutations reduce the ability of tau to promote microtubule assembly (Hasegawa et al., 1998). Moreover, an imbalance in the ratio of three-repeat to four-repeat tau that results in the relative overproduction of tau with four repeats causes FTDP-17 and leads to the formation of filaments consisting of four-repeat tau (Spillantini et al., 1997, 1998; Hutton et al., 1998). This indicates that in the human brain, a tight regulation of tau isoform ratios is essential for preventing neurodegeneration and dementia. The sporadic tauopathies PSP and CBD are also characterized by the presence of four-repeat tau-containing filaments (Lee et al., 2001). An association between PSP and CBD, and a dinucleotide repeat polymorphism in the intron between exons 9 and 10 of *Tau* has been reported (Conrad et al., 1997; Di Maria et al., 2000). Moreover, two common *Tau* haplotypes that differ only at the nucleotide level have been described (Baker et al., 1999). Homozygosity of the more common H1 allele predisposes to PSP and CBD. The intron between exons 9 and 10 of *Tau* also comprises the putative intronless gene *Saitohin* (Conrad et al., 2002). It encodes a predicted protein of 128 amino acids of unknown function, which is polymorphic at codon 7 (Q or R).

Little is known about the organisation and sequence of the *Tau* gene in the great apes and other nonhuman primates. It has been reported that the chimpanzee is relatively resistant to developing the neuropathological characteristics of AD (Gearing et al., 1994), whereas baboons, rhesus monkeys, and lemurs have been found to develop tau inclusions, as well as beta amyloid peptide (Aβ) deposits, as a function of age (Wisniewski et al., 1973; Giannakopoulos et al., 1997; Härtig et al., 2000; Schultz et al., 2000). Filamentous tau inclusions have been documented in the baboon (Schultz et al., 2000).

It remains to be established whether this differential susceptibility to developing tau inclusions is reflected in sequence differences in *Tau* between humans and nonhuman primates. Here we have addressed this question by sequencing exons 1–13 of *Tau* and their flanking intronic regions, as well as a region of intron 9 comprising *Saitohin*, in chimpanzees, gorillas, and gibbons. Partial sequence information was obtained from cynomolgus macaques and green monkeys.

2. Materials and methods

2.1. Materials

Genomic DNA, RNA, and proteins were extracted from frozen tissue of the primate species and the cell line listed in Table 1. The tissues were kindly provided by D.C. Gajdusek (National Institutes of Health, Bethesda, MD, USA; for chimp-1, gorilla, and gibbon-1) and P. Khaitovich (Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany). The COS-7 cell line was obtained from the "Deutsche Sammlung von Mikroorganismen und Zellkulturen" (Braunschweig, Germany). Importation of primate tissue to the UK followed CITES regulations.

2.2. Tissue culture

COS-7 cells were maintained in Dulbecco's modified Eagle's medium+10% foetal calf serum, supplemented with penicillin and streptomycin. The medium was changed every 2 days and the cell culture stopped after 14 days by the addition of 6 ml of Trizol per dish.

2.3. Polymerase chain reaction (PCR) and DNA sequencing

Genomic DNA was extracted using the DNeasy tissue kit (Qiagen) and included a mock-extracted sample, where water was substituted for the tissue sample. PCR primers were designed to amplify tau exons 1, 2, 3, 4, 4a, 5, 6, 7, 8, 9, 10, 11, 12, and 13, including 100–300 bp of flanking intronic regions. A DNA stretch in exon 9 comprising *Saitohin* was also amplified. A list of PCR and sequencing primers is given as Supplementary Table 1. PCR was performed in a final volume of 50 μl using TaqPCR mastermix (Qiagen), supplemented with 0.7 U of Pfu-Turbo polymerase (Stratagene), 30 ng of genomic DNA, and 50 pmol of each primer. Most reactions used a touchdown temperature profile of 5 cycles (94 °C, 5 s; 72 °C, 2 min), 5 cycles (94 °C, 5 s; 70 °C, 2 min), and 25 cycles (94 °C, 5 s; 68 °C, 2 min), with a final 10-min

Table 1
List of primate tissues and cell line

Individual	Designation	Species	Tissue
Chimpanzee-1	C8	*Pan troglodytes*	Parietal cortex
Chimpanzee-2	Herman	*Pan troglodytes*	Frontal cortex and cerebellar cortex
Chimpanzee-3	13306	*Pan troglodytes*	Muscle
Gorilla	15/08/1973	*Gorilla gorilla*	Temporal cortex
Gibbon-1	939	*Hylobates lar*	Temporal cortex
Gibbon-2	Falco	*Hylobates concolor*	Temporal cortex
Cynomolgus macaque		*Macaca fascicularis*	Sclera
Green monkey	COS-7	*Chlorocebus aethiops*	SV40-transformed kidney cell line

extension at 72 °C. Less stringent conditions were used for some reactions using gibbon and macaque DNA samples. They consisted of 5 cycles (94 °C, 5 s; 70 °C, 2 min), 5 cycles (94 °C, 5 s; 68 °C, 2 min), and 25 cycles (94 °C, 5 s; 66 °C, 15 s; 72 °C, 2 min), with a final 10-min extension at 72 °C. The PCR products were purified from 1% agarose gels using QIAquick spin columns (Qiagen) and eluted with 50 μl of water. For DNA sequencing, 10 μl of eluate was mixed with 6 μl of big dye terminator mastermix (v 3.0; PE Biosystems), 3.5 pmol of sequencing primer, 0.3 μl of 25 mM $MgCl_2$, and 3 μl of water. Cycle sequencing used a total of 27 cycles (96 °C, 30 s; 45 °C, 15 s; 60 °C, 4 min). The reactions were run on an ABI 377 DNA sequencer. Nucleotide sequences were deposited in the database with accession numbers AY574122–AY574186. Sequence alignments made use of the multiple alignment general interface (MAGI, http://www.hgmp.mrc.ac.uk) and the clustal algorithm. The human *Tau* gene sequence (accession number AC091628) was used for sequence comparison.

2.4. Reverse transcription polymerase chain reaction (RT-PCR)

RNA was extracted from the temporal cortex of chimpanzee, gorilla, gibbon, and human using the Trizol reagent (Invitrogen) and reverse-transcribed with Superscript II enzyme (Invitrogen). Alternative splicing of exons 2, 3, 6, 8, and 10 of Tau was analyzed using primers T5′ f/TE5r, TE5f/TE7r, T5f/T9r, and THF1/THR1, respectively (see Supplementary Table 1). Full-length tau transcripts were amplified using primers T5′ f/T3′ r, cloned into vector pZero2.0 (Invitrogen), and sequenced. The PCR conditions were identical to those used for genomic DNA, with the exception of the reactions using primers THF1/THR1 and Esaif/Esair, for which we used a total of 35 cycles (94 °C, 20 s; 55 °C, 20 s; 72 °C, 45 s), with a final 5-min extension at 72 °C. For DNase I treatment, 15 μg of RNA was incubated with RNase-free DNase I (190 U/ml; Qiagen) for 20 min at room temperature. RNA was purified using RNAeasy minElute cleanup columns (Qiagen) and brought to 0.5 mg/ml.

Table 2
Rate of base substitutions in different primates compared to the human MAPT sequence

Species	Nocod/ 1000 bp	Syncod/ 1000 bp	Nosyncod/ 1000 bp
chi-1	10.6	4.0	2.2
Gor	13.2	3.7	3.7
gib-1	34.0	9.2	11.4
Mac	56.7	21.1	14.1
Cos	56.7	18.5	7.1

Nocod = substitution in noncoding regions; syncod = synonymous base substitutions in coding regions; nosyncod = nonsynonymous base substitutions in coding regions.

Table 3
Comparison of base substitutions in coding and noncoding sequences in constitutively and nonconstitutively spliced exons

Base substitutions	Constitutive exon substitutions/ 1000 bp	Nonconstitutive exon substitutions/ 1000 bp
Coding sequence	18.8	40.2
Intronic flanking sequence	58.3	53.5

2.5. Exon trapping

Exon 10, together with 5′ and 3′ flanking intronic sequences, was amplified from human, chimpanzee, gorilla, and macaque DNA using primers etr10f/etr10r and cloned into the vector pSPL3 (Gibco-BRL) via the *Sac*I/*Bam*HI sites. The 3′ intronic sequence included 113 bp for all constructs, and the 5′ intronic sequence was 343 bp for human and gorilla, and 283 bp (lacking a 60-bp repeat) for chimpanzee and macaque. All constructs were verified by DNA sequencing. COS-7 cells in six-well plates were transfected with 1.2 μg of each construct. RNA was prepared 24 h after transfection and SA2L-primed cDNA produced by reverse transcription using Superscript II following the manufacturer's recommendations. Amplification was carried out for 25 cycles (94 °C, 30 s; 60 °C, 30 s; 72 °C, 60 s). Secondary PCR was performed with nested primers SD2/SA4 for 25 cycles (94 °C, 30 s; 60 °C, 30 s; 72 °C, 60 s), with a final 10-min extension at 72 °C. The relative proportions of exon 10-containing/exon 10-lacking transcripts were obtained by densitometric image analysis of ethidium bromide-stained 2% agarose gels.

2.6. Tau protein extraction and analysis

Proteins were precipitated from the phenol phase of the Trizol reagent with isopropyl alcohol, and the pellets were dissolved in 6 M guanidine hydrochloride and dialyzed against 20 mM Tris/HCl, pH 8.0, 0.1 mM EDTA, 0.1 mM DTT, and 1 mM PMSF. Dephosphorylation was done with *Escherichia coli* alkaline phosphatase for 3 h at 67 °C, following adjustment of the $MgCl_2$ concentration to 0.5 mM. For immunoblotting, samples were run using 10% sodium dodecyl sulfate polyacrylamide electrophoresis

Table 4
List of human-specific nucleotides

Location on MAPT	nt change	Location on MAPT	nt change
E1 −185	C to T	E10 −185	G to A
E4 −76	A to G	E10 −205	G to A
E6 +67	G to A	E11 −170	G to A
E6 +105	C to T	E12 +16	C to T
E7 +11	T to C	Saitohin 29 Cys10Arg	T to C
E7 +94	C to T	Saitohin +39	C to T
E8 −201	G to C		

Nucleotide (nt) change describes the conversion of the conserved bases in the nonhuman primates to the human-specific nucleotides.

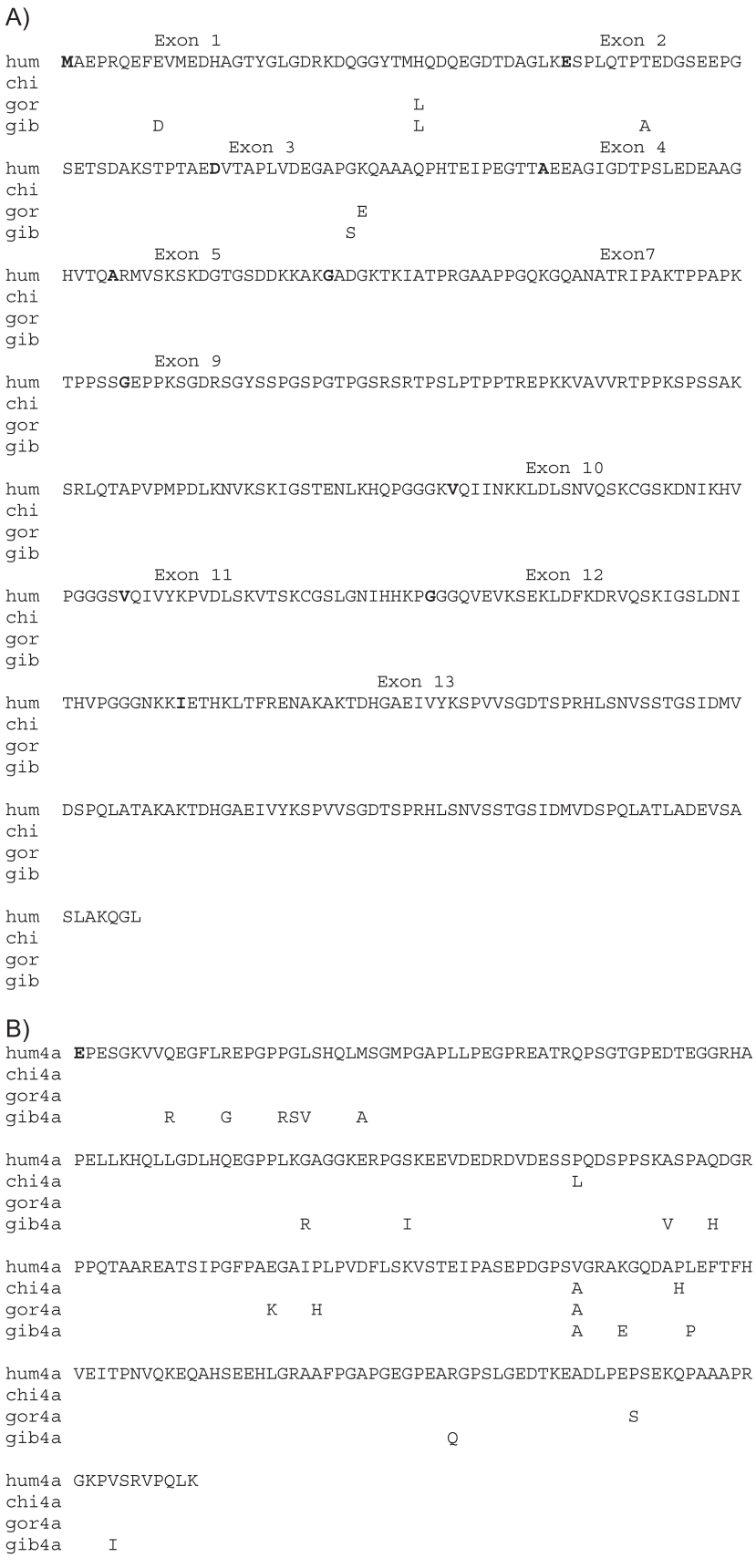

Fig. 1. Amino acid alignment of human (hum), chimpanzee (chi), gorilla (gor) and gibbon-1 (gib) exons found in central nervous system tau isoforms (A) and of exon 4a present in bigtau from the peripheral nervous system (B). The first amino acid of each exon is marked in bold.

(SDS-PAGE) and transferred to polyvinylidene difluoride membranes. After soaking in blocking solution, the membranes were incubated with the Tau-1 antibody (1:5000; Chemicon), which recognizes an unphosphorylated epitope spanning amino acids 198–205 in tau. Detection was carried out with goat anti mouse/rabbit horseradish peroxidase conjugate, prior to visualization with Supersignal ECL substrate (Pierce, Rockford).

3. Results

3.1. MAPT sequence variation at the nucleotide and protein levels

We sequenced 14 *Tau* exons (1, 2, 3, 4, 4a, 5, 6, 7, 8, 9, 10, 11, 12, and 13) with flanking intronic regions and the putative *Saitohin* gene in two chimpanzees, one gorilla, and one gibbon. We obtained partial genomic sequence from chimpanzee-3, a second gibbon, a cynomolgus macaque, and a green monkey cell line. The sequences of tau transcripts from the chimpanzee obtained by RT-PCR were identical to the assembled exonic sequences. Table 2 summarizes the nucleotide variation in the coding and noncoding regions of *Tau*. The chimpanzee sequence was most similar to the human sequence. The gorilla sequence showed an increased rate of nonsynonymous base substitutions in the coding region, whereas the gibbon sequence showed a threefold greater variation in coding and noncoding regions than the chimpanzee sequence. The estimation of sequence variation in macaque and green monkey may have been somewhat skewed, since it only relied on partial sequence information (exons 2, 3, and 8–13 for the macaque; exons 1, 2, 5, and 11–13 for the green monkey).

```
hum   gtcactccccagactgcctctgccaagtccgaaagtggaggcatccttgcgagcaagtag
chi1  gtcactccccagactgcctctgccaagtccgaaagtggaggcatccttgcgagcaagcag
chi2  gtcactccccagactgcctctgccaagtccgaaagtggaggcatccttgcgagcaagcag
chi3a gtcactccccagactgcctctgccaagtctgaaagtggaggcatccttgcgagcaagcag
chi3b gtcactccccagactgcctctgccaagtccgaaagtggaggcgtccttgcgagcaagcag
gor   gtcactccccagactgcctctgccaagtccgaaagtggaggcatccttgcgagcaagcag
gib   gtcactccccagactgcctctgccaagtccgaaagtggaggcatccttgcaagcaagcag
mac   gtcactccccagactgcctctgccaagtccgaaagtggaggcatccttgcgagcaagcag

hum   gcgggtccagggtggcgcatgtcactcatcgaaagtggaggcgtccttgcgagcaagcag
chi1  ------------------------------------------------------------
chi2  ------------------------------------------------------------
chi3a ------------------------------------------------------------
chi3b gcgggtccagggtggcgcatgtcactcatcggaagtggaggcgtccttgcgagcaagcag
gor   gcgggtccagggtggcgcatgtcactcatcgaaagtggaggcgtccttgcaagcaagcag
gib   ------------------------------------------------------------
mac   ------------------------------------------------------------

                                                        V  Q  I  I
hum   gcgggtccagggtggcgtgtcactcatcctttttttctggctaccaaagGTGCAGATAATT
chi1  gcgggtccagggtggcgtgtcactcatccttttttctggctaccaaagGTGCAGATAATT
chi2  gcgggtccagggtggcgtgtcactcatccttttttctggctaccaaagGTGCAGATAATT
chi3a gcgggtccagggtggcgtgtcactcatccttttttctggctaccaaagGTGCAGATAATT
chi3b gcgggtccagggtggcatgtcactcatccttttttctggctaccaaagGTGCAGATAATT
gor   gcgggtccagggtggcgtgtcactcatccttttttctggctaccaaagGTGCAGATAATT
gib   gcgggtccagggtggcgtgtcactcatccttttttctggctaccaaagGTGCAGATAATT
mac   gcgggtccagggtggcgtgtcactcatccttttttctggctaccaaagGTGCAGATAATT

       N  K  K  L  D  L  S  N  V  Q  S  K  C  G  S  K  D  N  I  K
hum   AATAAGAAGCTGGATCTTAGCAACGTCCAGTCCAAGTGTGGCTCAAAGGATAATATCAAA
chi1  AATAAGAAGCTGGATCTTAGCAACGTCCAGTCCAAGTGTGGCTCAAAGGATAATATCAAA
chi2  AATAAGAAGCTGGATCTTAGCAACGTCCAGTCCAAGTGTGGCTCAAAGGATAATATCAAA
chi3a AATAAGAAGCTGGATCTTAGCAACGTCCAGTCCAAGTGTGGCTCAAAGGATAATATCAAA
chi3b AATAAGAAGCTGGATCTTAGCAACGTCCAGTCCAAGTGTGGCTCAAAGGATAATATCAAA
gor   AATAAGAAGCTGGATCTTAGCAACGTCCAGTCCAAGTGTGGCTCAAAGGATAATATCAAA
gib   AATAAGAAGCTGGATCTTAGCAACGTCCAGTCCAAGTGTGGCTCAAAGGATAATATCAAA
mac   AATAAGAAGCTGGATCTTAGCAACGTCCAGTCCAAGTGTGGCTCAAAGGATAATATCAAA

       H  V  P  G  G  G  S
hum   CACGTCCCGGGAGGCGGCAGTgtgagtaccttcacacgtcccatgcgccgtgctgtggct
chi1  CACGTCCCGGGAGGCGGCAGTgtgagtaccttcacacgtcccatgcgccrtgctgtggct
chi2  CACGTCCCGGGAGGCGGCAGTgtgagtaccttcacacgtcccatgcgccgtgctgtggct
chi3a CACGTCCCGGGAGGCGGCAGTgtgagtaccttcacacgtcccatgcgccgtgctgtggct
chi3b CACGTCCCGGGAGGCGGCAGTgtgagtaccttcacacgtcccatgcgccgtgctgtggct
gor   CACGTCCCGGGAGGCGGCAGTgtgagtaccttcacacgtcccatgcgccgtgctgtggct
gib   CACGTCCCGGGAGGCGGCAGTgtgagtaccttcacacgtcccgtgcgccgtgctgtggct
mac   CACGTCCCGGGAGGCGGCAGTgtgagtaccttcacacgtcccatgcgccgtgctgtggct
```

Fig. 2. Sequence alignment of exon 10 with intronic flanking regions. Human (hum), chimpanzee (chi), gorilla (gor), gibbon-1 (gib) and macaque (mac) sequences are shown. r : g/a. The tandem repeat is underlined (first repeat single underlined; second repeat double underlined); grey shading highlights base substitutions. The +29 polymorphism in chimpanzee-1 is shown in bold. For chimpanzee-3, the two alleles are shown separately.

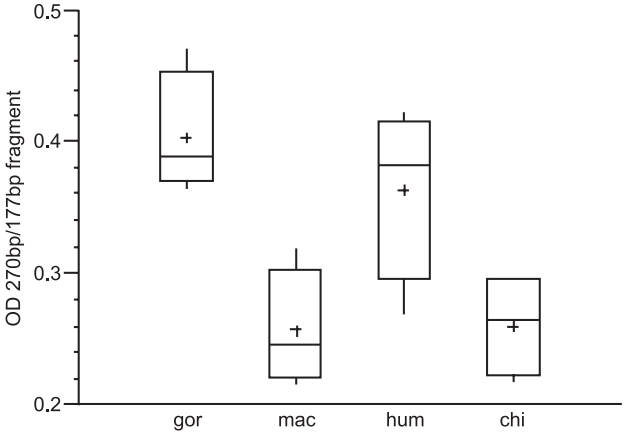

Fig. 3. Presence of the tandem repeat affects exon 10 splicing. Boxplot showing the abundance of exon 10-containing exon-trapping PCR products (270 bp) in relation to exon 10-lacking PCR products (177 bp) evaluated by densitometric analysis of ethidium bromide-stained agarose gel. Chimpanzee (chi) and macaque (mac) sequences without the tandem repeat yield less exon 10-containing PCR products than sequences from species with the tandem repeat, such as humans (hum) and gorilla (gor).

The constitutively spliced *Tau* exons 1, 4, 5, 7, 9, 11, 12, and 13 were more highly conserved than exons 2, 3, 4a, 6, 8, and 10, which are not constitutively spliced (Table 3). The flanking intronic regions of all exons had comparable nucleotide substitution rates. We examined the sequences for the presence of nucleotide changes unique to humans (Table 4). Of the seven changes, all but one resided in intronic regions. The exception was a T-to-C transition in the first nucleotide of codon 10 of *Saitohin*, resulting in a C10R change in nonhuman primates. No human-specific insertions/deletions were observed.

The predicted amino acid sequence of the longest chimpanzee brain tau isoform (441 amino acids) was identical to that of the corresponding isoform from human brain (Fig. 1A). The gorilla sequence differed in two amino acids (H32L and K87E) and the gibbon sequence in four amino acids (E9D, H32L, T52A, and G86S) from human tau. Exon 4A was less conserved. Of 251 amino acids, the chimpanzee sequence differed at three positions and the gorilla sequence at four positions from the human sequence, whereas gibbon and human sequences showed 15 differences (Fig. 1B).

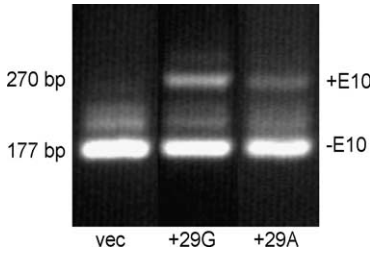

Fig. 4. The +29 A allele in the chimpanzee sequence exhibits reduced exon 10 splicing. Ethidium bromide-stained agarose gel of exon-trapping PCR products.

3.2. Analysis of the alternative splicing of exon 10 by exon trapping

Previous work has identified a 59- or 60-bp tandem repeat in intron 9 of human *Tau*, at positions −139 to −21 relative to exon 10 (Poorkaj et al., 2001a). Both copies of the repeat were present in gorilla DNA, whereas two chimpanzees had only one copy of the repeat and chimpanzee-3 was polymorphic for the tandem repeat. Gibbon and macaque DNA carried only one repeat each (Fig. 2). By exon trapping, transfection with chimpanzee and macaque constructs with one repeat gave a smaller proportion of exon 10-containing transcripts than transfection with human and gorilla constructs having two repeats (Fig. 3).

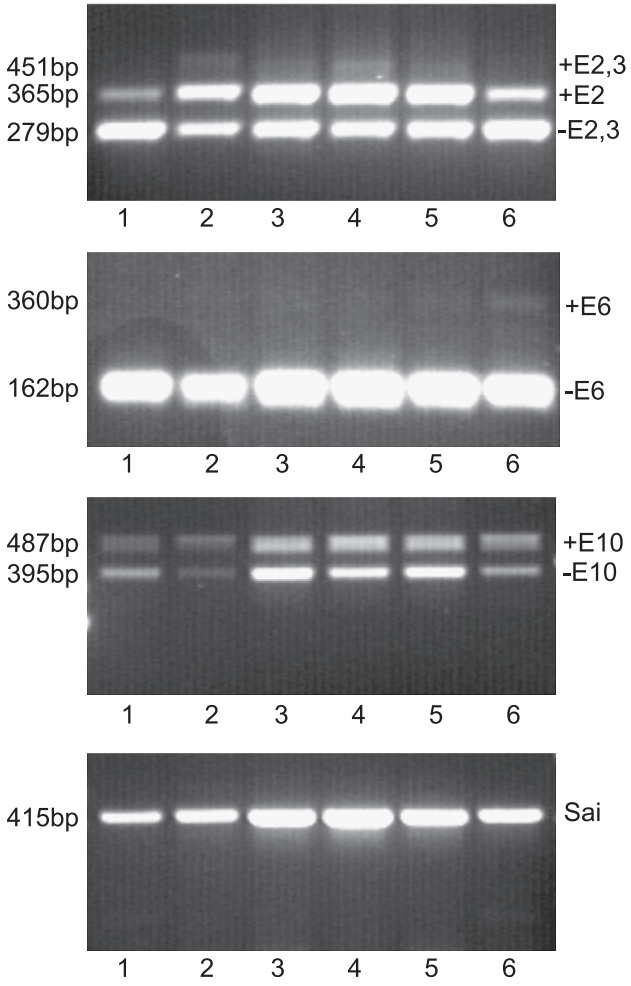

Fig. 5. RT-PCR analysis of tau isoform and saitohin expression. RNA was extracted from the temporal cortex of gorilla (lane 1), gibbon (lane 2), chimpanzee-1 (lane 3), and human (lane 6); from the cerebellum of chimpanzee-2 (lane 4); and from the frontal cortex of chimpanzee 2 (lane 5). The first panel shows the amplification of a 5′ portion of tau transcripts with primers binding in the 5′ untranslated region (UTR) and exon 5 showing alternative splicing of exons 2 and 3. The second panel shows that exon 6-containing tau transcripts in the CNS are rare in all species examined. The third panel shows differences in alternative splicing of exon 10. The fourth panel depicts Saitohin cDNA (Sai).

The intronic sequence close to the splice donor site of exon 10 was conserved between humans and nonhuman primates. The only sequence difference was found in one of the chimpanzees, which was heterozygous (g or a) for the nucleotide at position +29. The other species were homozygous for +29 g. By exon trapping, the ratio of exon 10-containing/exon 10-lacking transcripts from the chimpanzee was 0.29 for the +29 g allele and 0.13 for the +29 a allele (Fig. 4).

3.3. Analysis of the alternative splicing of exons 2, 3, 6, 8, and 10 by RT-PCR

The splicing pattern of exons 2, 3, 6, 8, and 10 in the brains of nonhuman primates was similar to that in human brain (Fig. 5). The nonhuman primates yielded the transcripts +E2+E3, +E2−E3, −E2−E3, +E10, and −E10, making it likely that they express six tau isoforms in brain. There appeared to be differences in the proportion of exon 10-

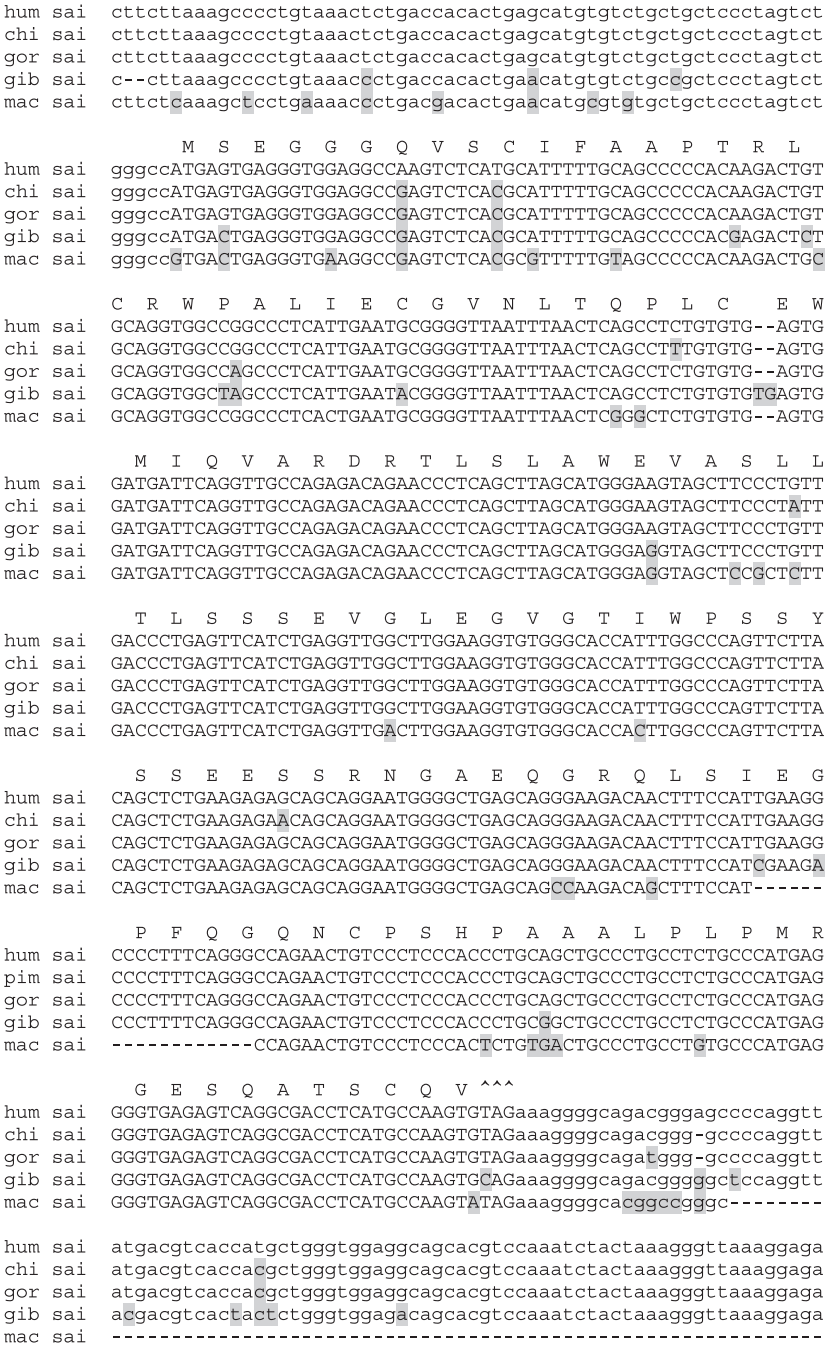

Fig. 6. Intron 9 nucleotide sequence alignment of the region containing Saitohin. The protein sequence of human saitohin is depicted above the nucleotide sequence alignment. Human (hum), chimpanzee (chi), gorilla (gor), gibbon-1 (gib) and macaque (mac) sequences are shown. Gray shading highlights base substitutions.

containing over exon 10-lacking transcripts, in that chimpanzees and gorillas expressed a higher proportion of transcripts lacking exon 10 than humans (Fig. 5). Transcripts encoding exon 6 were rare or undetectable. Transcripts encoding exon 8 were undetectable (not shown).

3.4. Saitohin is present in chimpanzee, gorilla, and orangutan, but not in gibbon or macaque

Saitohin has been described as an intronless gene that encodes a predicted protein of 128 amino acids and is nested in intron 9 of human *Tau* (Conrad et al., 2002). Codon 7 was found to be polymorphic, coding either for the more common Q or the less common R. We sequenced this putative gene in nonhuman primates and found an open reading frame encoding a predicted 128-amino acid protein in both chimpanzee and gorilla (Fig. 6). Codon 7 was homozygous for R. A C10R change relative to the human sequence was present, with the chimpanzee carrying an additional change at position 83 (S83N). In the gibbon (*Hylobates lar*), the open reading frame of saitohin was coding for a predicted protein of only amino acids, because of a 2-bp insertion that generated a stop codon. Cynomolgus macaque DNA lacked the putative start codon of saitohin, but was otherwise similar in sequence to up to 9 bp downstream of the stop codon (Fig. 6). In human brain, a saitohin transcript could be readily amplified by RT-PCR. Amplification was still possible following extensive DNase I treatment of human brain RNA, even though it was somewhat reduced compared to untreated RNA (not shown). Analysis of saitohin expression by RT-PCR showed transcripts in each primate species examined. The gibbon, with a dysfunctional open reading frame, showed comparable expression to the gorilla, with an intact open reading frame (Fig. 5).

3.5. Tau haplotype analysis in nonhuman primates

Eight single nucleotide polymorphisms, which define the H1 and H2 human *Tau* haplotypes, were analyzed (Table 5) (Supplementary Table 2). Of these polymorphisms, seven had the sequence of the H2 haplotype. No sequence variation was observed between the chimpanzee, gorilla, gibbon, macaque, and green monkey. Supplementary Table 2 lists the sequences in nonhuman primates of additional single nucleotide polymorphisms previously described in humans. In the nonhuman primates studied, these sequences were conserved and nonpolymorphic.

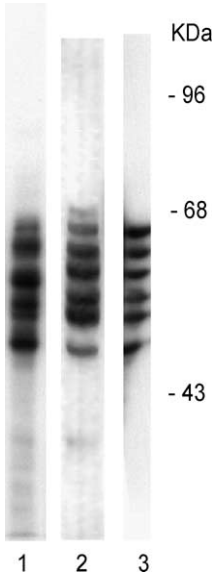

Fig. 7. Western blot of dephosphorylated brain proteins from chimpanzee (lane 1) and gibbon (lane 2), together with a mix of six recombinant human tau isoforms (lane 3). The membrane was probed with anti-tau antibody Tau-1.

3.6. Tau protein isoforms in chimpanzee and gibbon brain

By immunoblotting of dephosphorylated brain extracts with anti-tau antibody Tau-1, six bands were detected in the chimpanzee and gibbon brains; they aligned with the six recombinant human brain tau isoforms (Fig. 7). Four strong tau bands and two weaker bands were present in the chimpanzee brain, and six bands of similar intensity were found in the gibbon brain.

Table 5
Summary of primate *Tau* gene typing for H1 and H2 haplotypes

Position	Human H1/H2	Amino acid	Chimpanzee[a]	Gorilla	Gibbon[a]	Macaque	COS-7	Nonhuman primates	Reference
Exon 1 5	A/G	na	G	G	G	nd	G	H2	Baker et al.
Intron 2 +18	C/T	na	C	C	C	C	C	H1	Baker et al.
Intron 3 +9	A/G	na	G	G	G	G	nd	H2	Baker et al.
Exon 9 125	A/G	Ala/Ala	G	G	G	G	nd	H2	Baker et al.
Exon 9 209	T/C	Asn/Asn	C	C	C	C	nd	H2	Baker et al.
Saitohin 20	A/G	Gln/Arg	G	G	G	G	nd	H2	Verpillat et al.
Intron 11 +34	G/A	na	A	A	A	A	A	H2	Baker et al.
Intron 13 +34	T/C	na	C	C	C	C	C	H2	Baker et al.

na=Not applicable; nd=not determined.
[a] Two individuals were analysed per species/genus and were found to have nucleotide assignment.

4. Discussion

The predicted amino acid sequence of the longest brain isoform of chimpanzee tau was 100% identical to that of the corresponding isoform from human brain. It was identical to recently published partial sequences of chimpanzee tau (Caceres et al., 2003; Clark et al., 2003). Gorilla tau was 99.5% and gibbon tau was 99.0% identical to human tau. Six isoforms that aligned with recombinant human tau were present in chimpanzee and gibbon brain by immunoblotting. Sequence differences with human tau were located in exons 1–3, where they are unlikely to affect the ability of tau to interact with microtubules. It remains to be determined whether they can influence the ability of tau to assemble into filaments. Exon 4A from chimpanzee and gorilla was 98.8% identical to the corresponding human sequence, with 94% identity for exon 4A from gibbon. The lower degree of conservation of exon 4A is compatible with the proposed spacer function of this exon. Previous work has shown that baboon brain tau is 99% (four-repeat isoform lacking exons 2 and 3) and rhesus monkey brain tau is 98% (longest isoform) identical to human tau (Nelson et al., 1996). This contrasts with 88% identity between mouse and human tau (Poorkaj et al., 2001a), and 74% identity between chicken and human tau (Yoshida and Goedert, 2002).

From the above, it is clear that the apparent resistance of the chimpanzee to developing a filamentous tau pathology cannot be due to differences in amino acid sequence with human tau. It could result instead from differences in the nucleotide sequences that are located outside the coding region. Upstream of exon 10, human tau possesses two repeats of 59 or 60 bp (Poorkaj et al., 2001a). Gorilla DNA also carries two repeats, chimpanzee DNA can have either one or two repeats, whereas gibbon and macaque DNA have only one repeat each. By exon trapping, constructs with one repeat gave rise to a smaller proportion of exon 10-containing transcripts than constructs with both repeats.

In humans, intronic sequences located close to the splice donor site of exon 10 regulate the splicing of exon 10 (Lee et al., 2001). Mutations in this region cause FTDP-17, most probably through the destabilization of a stem loop structure (Hutton et al., 1998; Spillantini et al., 1998; Varani et al., 1999). This results in the relative overproduction of four-repeat tau (Spillantini et al., 1998). The sequences of exon 10 and the stem loop were completely conserved between humans and the nonhuman primates studied. One chimpanzee was heterozygous (g or a) at position +29 of the intron following exon 10. The other species were homozygous for +29 g. By exon trapping, +29 a constructs showed a relative increase in exon 10-lacking transcripts compared to +29 g constructs, consistent with findings using human DNA (Stanford et al., 2003). The latter study also reported a heterozygous +29 g-to-a change in a family with frontotemporal dementia, but without tau pathology. The authors suggested that a relative increase in three-repeat tau may cause disease through an ill-defined mechanism that does not involve the formation of tau filaments. However, Stanford et al. also found the +29 g-to-a change in 1 of 200 control individuals, in line with earlier works (D'Souza et al., 1999; Roks et al., 1999; Poorkaj et al., 2001b). Together with the present findings in the chimpanzee, it therefore appears likely that the +29 change is a polymorphism in both humans and chimpanzees. Analysis of a larger number of samples will be required to establish the frequency of this polymorphism in the chimpanzee population. It also remains to be seen whether chimpanzee brain expresses a higher relative proportion of three-repeat tau than human brain, as suggested by our RT-PCR experiments. This could be relevant with respect to the resistance of the chimpanzee to developing a filamentous tau pathology.

Haplotypes H1 and H2 characterize human *Tau* (Baker et al., 1999). We analysed eight single nucleotide polymorphisms that distinguish H1 from H2 in the nonhuman primates. At seven positions, nonhuman primates were homozygous for the H2 haplotype sequence. This suggests that H1, the more prevalent of the two haplotypes, may be human-specific and may have evolved as the result of selective pressure. One of these single nucleotide polymorphisms is located in the coding region of saitohin, where codon 7 encodes either Q (H1 haplotype) or R (H2 haplotype) (Conrad et al., 2002; Verpillat et al., 2002). In saitohin from chimpanzee and gorilla, this residue was R, as expected from the haplotype findings. Compared with the human sequence, chimpanzee and gorilla saitohin carried a C10R substitution, with an additional amino acid difference (S83N) in chimpanzee.

The functional relevance of saitohin is unknown. It has been reported to be coexpressed with tau in human tissues, leading to the suggestion that both proteins may function in a common pathway (Conrad et al., 2002). The present findings indicate that this is unlikely to be of general relevance, since gibbon and cynomolgus macaque did not encode intact saitohin. This contrasts with tau protein, which is conserved from chickens to humans (Goedert et al., 1989; Yoshida and Goedert, 2002). Transcription of the genomic sequence encompassing *Saitohin* in the gibbon may be a sign of noise, or may indicate a regulatory role of the noncoding RNA.

In conclusion, our results show that humans and great apes have very similar tau protein sequences, and that their differential susceptibility to developing filamentous tau inclusions may be influenced by intronic *Tau* sequences.

Acknowledgements

We wish to thank Drs. D.C. Gajdusek, P. Khaitovich, and S. Pääbo for providing primate tissues. This work was supported by the UK Medical Research Council, the European Commission (QLK6-1999-51519), and the Deutsche Forschungsgemeinschaft (HO23681-1). We gratefully acknowledge the Interdisziplinäres Zentrum für Klinische Forschung (IZKF) at the Faculty of Medicine of the

University of Leipzig (C24) and the IZKF Leipzig core unit Z03 for DNA sequencing.

Appendix A. Supplementary information

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.gene.2004.07.013.

References

Andreadis, A., Brown, W.M., Kosik, K.S., 1992. Structure and novel exons of the human *tau* gene. Biochemistry 31, 10626–10633.

Baker, M., Litvan, I., Houlden, H., Adamson, J., Dickson, D., Perez-Tur, J., Hardy, J., Lynch, T., Bigio, E., Hutton, M., 1999. Association of an extended haplotype in the *tau* gene with progressive supranuclear palsy. Hum. Mol. Genet. 8, 711–715.

Berriman, J., Serpell, L.C., Oberg, K.A., Fink, A.L., Goedert, M., Crowther, R.A., 2003. Tau filaments from human brain and from in vitro assembly of recombinant protein show cross-β structure. Proc. Natl. Acad. Sci. U. S. A. 100, 9034–9038.

Caceres, M., Lachuer, J., Zapala, M.A., Redmond, J.C., Kudo, L., Geschwind, D.H., Lockhart, D.J., Preuss, T.M., Barlow, C., 2003. Elevated gene expression levels distinguish human from non-human primate brains. Proc. Natl. Acad. Sci. U. S. A. 100, 13030–13035.

Clark, A.G., Glanowski, S., Nielsen, R., Thomas, P.D., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X., White, T.J., Sninsky, J.J., Adams, M.D., Cargill, M., 2003. Inferring nonneutral evolution from human–chimp–mouse orthologous gene trios. Science 302, 1960–1963.

Conrad, C., Andreadis, A., Trojanowski, J.Q., Dickson, D.W., Kang, D., Chen, X., Wiederholt, W., Hansen, L., Masliah, E., Thal, L.J., Katzman, R., Xia, Y., Saitoh, T., 1997. Genetic evidence for the involvement of tau in progressive supranuclear palsy. Ann. Neurol. 41, 277–281.

Conrad, C., Vianna, C., Freeman, M., Davies, P., 2002. A polymorphic gene nested within an intron of the *tau* gene: implications for Alzheimer's disease. Proc. Natl. Acad. Sci. U. S. A. 99, 7751–7756.

Di Maria, E., Tabaton, M., Vigo, T., Abbruzzese, G., Bellone, E., Donati, C., Frasson, E., Marchese, E., Montagna, P., Munoz, D.G., Pramstaller, P.P., Zanusso, G., Ajmar, F., Mandich, P., 2000. Corticobasal degeneration shares a common genetic background with progressive supranuclear palsy. Ann. Neurol. 47, 374–377.

D'Souza, I., Poorkaj, P., Hong, M., Nochlin, D., Lee, V.M.-Y., Bird, T.D., 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. U. S. A. 96, 5598–5603.

Gearing, M., Rebeck, G.W., Hyman, B.T., Tigges, J., Mirra, S.S., 1994. Neuropathology and apolipoprotein E profile of aged chimpanzees: implications for Alzheimer disease. Proc. Natl. Acad. Sci. U. S. A. 91, 9382–9386.

Giannakopoulos, P., Silhol, S., Jallageas, V., Mallet, J., Bons, N., Bouras, C., Delaere, P., 1997. Quantitative analysis of tau protein-immunoreactive accumulations and β amyloid protein deposits in the cerebral cortex of the mouse lemur, *Microcebus murinus*. Acta Neuropathol. 94, 131–139.

Goedert, M., Spillantini, M.G., Jakes, R., Rutherford, D., 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.

Goedert, M., Jakes, R., 1990. Expression of separate isoforms of human tau protein: correlation with the tau pattern in brain and effects on tubulin polymerization. EMBO J. 9, 4225–4230.

Goedert, M., Spillantini, M.G., Crowther, R.A., 1992. Cloning of a big tau microtubule-associated protein characteristic of the peripheral nervous system. Proc. Natl. Acad. Sci. U. S. A. 89, 1983–1987.

Härtig, W., Klein, C., Brauer, K., Schüppel, K.F., Arendt, T., Brückner, G., Bigl, V., 2000. Abnormally phosphorylated protein tau in the cortex of aged individuals of various mammalian orders. Acta Neuropathol. 100, 305–312.

Hasegawa, M., Smith, M.J., Goedert, M., 1998. Tau proteins with FTDP-17 mutations have a reduced ability to promote microtubule assembly. FEBS Lett. 437, 207–210.

Hutton, M., et al., 1998. Association of missense and 5′-splice-site mutations in *tau* with the inherited dementia FTDP-17. Nature 393, 702–705.

Lee, V.M.-Y., Goedert, M., Trojanowski, J.Q., 2001. Neurodegenerative tauopathies. Annu. Rev. Neurosci. 24, 1121–1159.

Nelson, P.T., Stefansson, K., Gulcher, J., Saper, C.B., 1996. Molecular evolution of tau protein: implications for Alzheimer's disease. J. Neurochem. 67, 1622–1632.

Poorkaj, P., Bird, T.D., Wijsman, E., Nemens, E., Garruto, R.M., Anderson, L., Andreadis, A., Wiederholt, W.C., Raskind, M., Schellenberg, G.D., 1998. *Tau* is a candidate gene for chromosome 17 frontotemporal dementia. Ann. Neurol. 43, 815–825.

Poorkaj, P., Kas, A., D'Souza, I., Zhou, Y., Pham, Q., Stone, M., Olson, M.V., Schellenberg, G.D., 2001a. A genomic sequence analysis of the mouse and human microtubule-associated protein tau. Mamm. Genome 12, 700–712.

Poorkaj, P., Grossman, M., Steinbart, E., Payami, H., Sadovnick, A., Nochlin, D., Tabira, T., Trojanowski, J.Q., Borson, D., Galasko, D., Reich, S., Quinn, B., Schellenberg, G.D., 2001b. Frequency of *tau* gene mutations in familial and sporadic cases of non-Alzheimer dementia. Arch. Neurol. 58, 383–387.

Roks, G., Dermaut, B., Heutink, P., Julliams, A., Backhovens, H., Van de Broeck, M., Serneels, S., Hofman, A., Van Broeckhoven, C., Van Duijn, C.M., Cruts, M., 1999. Mutation screening of the *tau* gene in patients with early-onset Alzheimer's disease. Neurosci. Lett. 277, 137–139.

Schultz, C., Dehghani, F., Hubbard, G.B., Thal, D.R., Struckhoff, G., Braak, E., Braak, H., 2000. Filamentous tau pathology in nerve cells, astrocytes and oligodendrocytes of aged baboons. J. Neuropathol. Exp. Neurol. 59, 39–52.

Spillantini, M.G., Goedert, M., Crowther, R.A., Murrell, J.R., Farlow, M.R., Ghetti, B., 1997. Familial multiple system tauopathy with presenile dementia: a disease with abundant neuronal and glial tau filaments. Proc. Natl. Acad. Sci. U. S. A. 94, 4113–4118.

Spillantini, M.G., Murrell, J.R., Goedert, M., Farlow, M.R., Klug, A., Ghetti, B., 1998. Mutation in the *tau* gene in familial multiple system tauopathy with presenile dementia. Proc. Natl. Acad. Sci. U. S. A. 95, 7737–7741.

Stanford, P.M., Shepherd, C.E., Halliday, G.M., Brooks, W.S., Schofield, P.W., Brodaty, H., Martins, R.N., Kwok, J.B.J., Schofield, P.R., 2003. Mutations in the *tau* gene that cause an increase in three repeat tau and frontotemporal dementia. Brain 126, 814–826.

Varani, L., Hasegawa, M., Spillantini, M.G., Smith, M.J., Murrell, J.R., Ghetti, B., Klug, A., Goedert, M., Varani, G., 1999. Structure of tau exon 10 splicing regulatory element RNA and destabilization by mutations of frontotemporal dementia and parkinsonism linked to chromosome 17. Proc. Natl. Acad. Sci. U. S. A. 96, 8229–8234.

Verpillat, P., Ricard, S., Hannequin, D., Dubois, B., Bou, J., Camuzat, A., Pradier, L., Frebourg, T., Brice, A., Clerget-Darpoux, F., Deleuze, J.F., Campion, D., 2002. Is the *Saitohin* gene involved in neurodegenerative diseases? Ann. Neurol. 52, 829–832.

Wisniewski, H.M., Ghetti, B., Terry, R.D., 1973. Neuritic (senile) plaques and filamentous changes in aged rhesus monkeys. J. Neuropathol. Exp. Neurol. 132, 566–584.

Yoshida, H., Goedert, M., 2002. Molecular cloning and functional characterization of chicken brain tau: isoforms with up to five tandem repeats. Biochemistry 41, 15203–15211.