Well-travelled pathways into neurodegeneration

All *MAPT* out?

Selina Wray and John Hardy (Department of Molecular Neuroscience and Rita Lilla Weston Laboratories, UCL Institute of Neurology)

Neurodegenerative disorders are characterized pathologically by insoluble protein aggregates in the neurons of affected brain regions. A common theme in neurodegenerative disease which has emerged is that mutations in the genes encoding the abnormally deposited protein cause autosomal-dominant inherited forms of disease. More recently, it has become apparent that common genetic variance at the same loci that cause Mendelian disease predispose risk to sporadic disease, generally by altering the expression levels of wild-type protein. The microtubule-associated protein tau (*MAPT*) is a classical example of this principle, and in this article, we discuss the biology and genetics of *MAPT* in disease, and compare and contrast *MAPT* with other loci implicated in neurodegeneration.

Key words: *Alzheimer's disease, Mendelian disease, Parkinson's disease, tau, tauopathy*

The tauopathies

Hyperphosphorylated insoluble tangles composed of the microtubule-associated protein tau are a pathological hallmark of a range of disorders termed the tauopathies. The tauopathies are clinically diverse and include Alzheimer's disease (AD), progressive supranuclear palsy (PSP), corticobasal degeneration (CBD) and frontotemporal dementia with parkinsonism linked to chromosome 17 (FTDP-17). Tau is a predominantly neuronal protein whose major function is to bind to and stabilize microtubules. Alternative splicing of a single gene generates six protein isoforms of tau that differ by the presence of zero, one or two N-terminal repeats and either three (3R) or four (4R) C-terminal microtubule-binding repeats which allow tau to interact with the tubulin cytoskeleton[1,2]. This process is regulated by phosphorylation, where phosphorylation at specific residues favours the detachment of tau from the microtubules[3]. However, in disease, the transient nature of this phosphorylation is lost, and hyperphosphorylated insoluble aggregates of tau are observed in neurons and glia of affected brain regions[4].

Alzheimer's disease

AD is the most common and best characterized of the tauopathies, where the two main pathological hallmarks are the neurofibrillary tangles composed of paired-helical filaments of hyperphosphorylated insoluble tau and extracellular plaques composed of amyloid β-peptide (Aβ). Aβ is derived from the proteolytic processing of the amyloid precursor protein (APP) by the γ-secretase complex. For many years, debate centred on what was most important – tau or amyloid – and a major breakthrough in AD research came with the finding that mutations in the *APP* gene and the presenilin 1 and 2 genes (which form part of the γ-secretase complex) cause early-onset familial AD. These findings led to the development of the amyloid cascade hypothesis, where amyloid processing sits upstream of tau in the AD pathogenesis[5]. Although the discovery of familial AD mutations provided conclusive proof of the importance of APP processing in disease, and showed that AD is in fact a secondary tauopathy, several lines of evidence continued to support the notion that tau is intrinsically linked with the neurodegenerative process in AD. Pathological correlation studies have shown that the extent of tau pathology correlates highly with the degree of dementia and the extent of neuronal loss in AD[6,7]. *In vitro* studies have shown that primary cultures of neurons from tau-knockout mice are resistant to Aβ toxicity, and *in vivo* work crossing transgenic mice overexpressing the human *APP* gene with tau-knockout mice resulted in a reduction of Aβ-induced cognitive impairments[8,9]. Taken together, this evidence strongly suggests that, although Aβ production and deposition is the initiating factor in the AD pathological cascade, tau is necessary for amyloid toxicity and resultant neurodegeneration.

Mutations in *MAPT* cause frontotemporal dementia with parkinsonism

Understanding the links between tau phosphorylation, aggregation and neuronal death has been the focus of much research over the last decade. However, it is research unravelling the genetics of *MAPT* and its association with disease that has provided conclusive evidence that tau is essential for neuronal integrity and that tau dysfunction is central to the pathology of several neurodegenerative diseases. Frontotemporal lobar degeneration (FTLD) is the second most common form of dementia after AD.

Several families with autosomal-dominantly inherited frontotemporal dementia with parkinsonism (FTDP) were identified with linkage to chromosome 17q21, and in 1998, the causal link between tau dysfunction and neurodegeneration was confirmed with the discovery of missense and splice-site mutations *MAPT* leading to FTDP-17 (Figure 1)[7,10]. To date, more than 40 mutations in *MAPT* have been described (www.molgen.ua.ac.be/ADMutations/default.cfm?MT=1&ML=1&Page=MutByGene). Missense mutations are clustered within the microtubule-binding repeat region of the protein (although several N-terminal mutations have been described) and disrupt the ability of tau to bind to microtubules and/or increase the propensity of tau to aggregate. Exon 10 of the tau gene codes for the extra microtubule-binding repeat seen in 4R tau isoforms and splice mutations in exon 10 and the flanking introns increase exon 10 inclusion, thereby leading to an increase in the amount of 4R tau isoforms. In the adult human central nervous system, the levels of 3R and 4R tau isoforms are approximately equal and so the discovery of mutations causing aberrant tau splicing suggests that the correct balance of tau isoforms is necessary for neuronal integrity[11]. This isoform imbalance is also observed in the sporadic tauopathies PSP and CBD. *MAPT* mutations directly influence the resultant tau pathology: for example, coding mutations affecting all six tau isoforms lead to paired-helical filaments similar to those seen in AD brain, whereas splicing mutations increasing 4R tau isoforms generally lead to straight filaments similar to those seen in the 4R tauopathies PSP and CBD[12].

The complex genetic architecture of the *MAPT* locus and its relevance to disease

The sporadic tauopathies PSP and CBD are characterized pathologically by the deposition of mainly 4R tau isoforms, but this occurs in the absence of tau-splicing mutations. Attempts to understand the biological basis of PSP and CBD have led to extensive dissection of the genetic architecture of the *MAPT* locus, revealing tau to be an exquisite example of how common genetic variation at the same loci underlying autosomal-dominant forms of disease can confer risk of sporadic disorders.

The tau gene sits in the largest known block of linkage disequilibrium in the human genome, spanning a region of ~1.8 Mb. This region of high-linkage disequilibrium was created by a chromosomal rearrangement around 3 million years ago, which resulted in a 900 kb inversion-polymorphism and created two haplotypes in Caucasian populations: H1 and H2 (Figure 2)[13]. Investigation of the single nucleotide polymorphisms (SNPs) within *MAPT* and their association with PSP revealed an extended haplotype encompassing the whole *MAPT* locus, of which the more common H1 variant is over-represented in PSP, a finding which has been replicated in several populations[14].

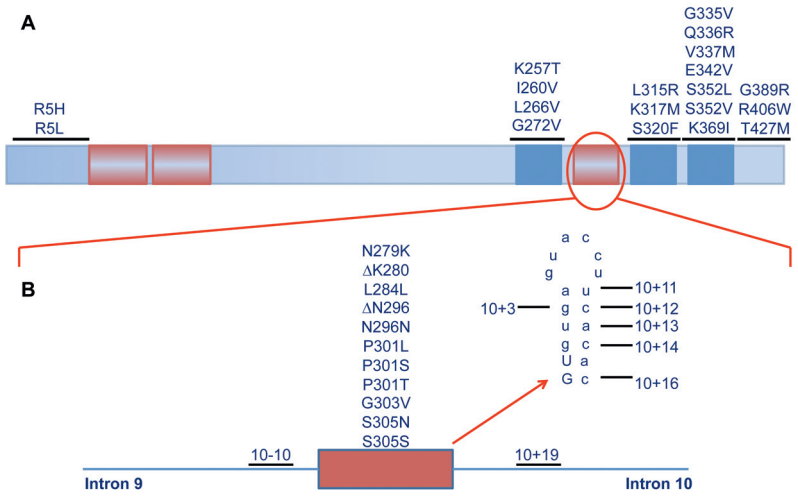

Figure 1. Missense and splice-site mutations in *MAPT*. (**a**) The longest tau isoform (2N4R) is shown with alternatively spliced exons indicated in red. The microtubule-binding repeats at the C-terminus of the molecule are indicated. Coding mutations in exons 1, 9, 11, 12 and 13 of *MAPT* are shown. (**b**). Coding and silent missense mutations in exon 10 are shown, together with splice mutations in the flanking introns. Several of these splice mutations are thought to disrupt a predicted stem–loop structure in *MAPT* pre-mRNA.

The H1 haplotype has accumulated further diversity, whereas the H2 haplotype has remained largely invariant. SNPs existing only on the H1 background have allowed the division of H1 into the sub-haplotypes H1a, H1b and H1c[15]. Of these, the H1c haplotype shows the strongest association with PSP in both UK and US cohorts[16]. A strong association of the H1 haplotype is also seen with the 4R tauopathies CBD and argyrophilic grain disease (AGD), as well as a weaker association that has been reported for AD[16].

Too much of a good thing?

Although genetic studies have provided compelling evidence for the intrinsic role of tau in the pathogenesis of the tauopathies, it is necessary to understand the functional consequences of common genetic variation in order to elucidate the underlying biology of disease. The H1 and H2 haplotypes both produce tau protein with the same coding sequence, so it is likely that genetic variance contributes to disease risk by altering tau expression and/or tau splicing.

As PSP and CBD are both characterized pathologically by excess 4R tau, it is logical to hypothesize that risk variants may increase the both total tau production and the production of exon 10-containing isoforms. The strongest association of the H1c haplotype with PSP has been mapped to the promoter region of *MAPT*[7,17]. *In vitro* reporter assays have shown that the H1c promoter region exhibits increased transcriptional activity in comparison with H2, and allele-specific expression assays in H1/H2 heterozygote neuroblastoma cell lines have showed that the H1 haplotype produces more exon 10-containing transcripts[18]. These findings are supported by work carried out in post-mortem brain tissue, where highest levels of total and 4R *MAPT* transcripts were produced from the H1c allele[17].

The concept that increased expression levels of proteins with normal sequence can initiate the disease process is not a new one[19]. Down's syndrome individuals develop AD-like plaques and tangles consistent with increased APP production due to trisomy 21 (where the *APP* gene is situated). Duplications and triplications of the α-synuclein gene (*SNCA*),

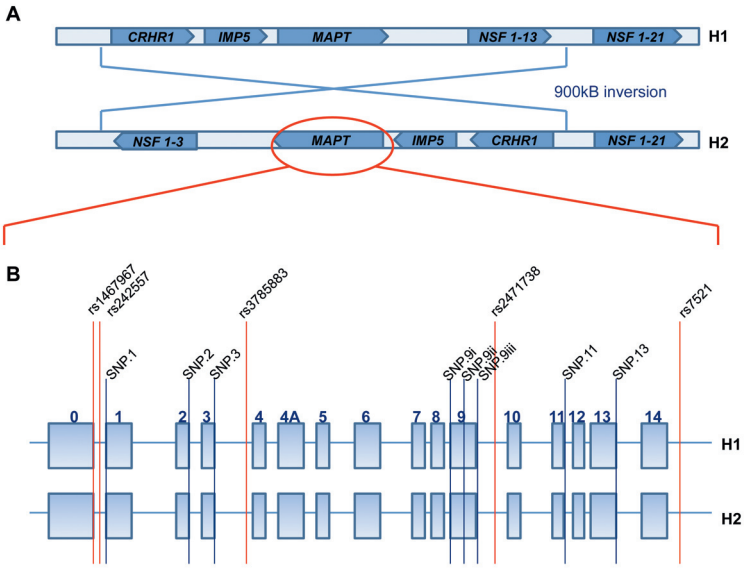

Figure 2. The complex genetic architecture of the *MAPT* locus. (**a**) *MAPT* sits in the largest block of linkage disequilibrium in the human genome. A chromosomal rearrangement resulted in a 900 kb inversion-polymorphism and produced two haplotypes, H1 and H2, where H2 is inverted with respect to H1. (**b**) H1 and H2 haplotypes are distinguished by eight tagging SNPs (SNP.1, SNP.2, SNP.3, SNP.9i, SNP.9ii, SNP.9iii, SNP.11 and SNP.13) and a 238 bp deletion in intron 9 present only on the H2 background. The H1 haplotype has accumulated further diversity and sub-haplotypes of *MAPT* have been identified, characterized by the SNPs rs1467967, rs242557, rs3785883, rs2471738 and rs7521 (indicated in red). rs242557, situated in the regulatory region of M*APT*, has been shown to be most strongly associated with PSP and the other 4R tauopathies. Adapted from Caffrey and Wade-Martins, 2007[31].

the major component of the Lewy body pathology in Parkinson's disease (PD), cause autosomal-dominant disease, and gene dosage is directly correlated with disease severity.

It is worth emphasizing that, although a clear functional basis for the association between *MAPT* variation and increased disease risk has been established, the genetic risk variant is common in the population. This means that other factors, genetic or environmental, must act in concert with *MAPT* to influence disease risk in the sporadic tauopathies, and these remain to be determined.

The usual suspects in unexpected places

Although PD is a synucleinopathy and not a tauopathy, the identification of *MAPT* mutations with clinically diagnosed parkinsonism prompted several groups to examine the association of *MAPT* haplotypes with PD risk. Several groups reported a small, but significant, positive association between *MAPT* and PD, but this was more controversial than previous association studies with PSP. Genome-wide association studies (GWAS) provide an unbiased means for the identification of common genetic variants associated with disease susceptibility and two recently published GWAS identified variation in *MAPT* as a major risk factor for PD, confirming the previous candidate-gene association studies[20,21]. Intriguingly, fine-mapping of the risk haplotype for PD has shown that, although it is part of the H1 clade, it is not the same variant that confers risk for PSP[22]. Increased 4R tau levels have been detected in PD brain, which suggests that the different variants could act through the same functional mechanism of increased tau expression[22].

In addition to *SNCA* mutations that cause autosomal dominant PD, variation in α-synuclein is also a major risk factor for sporadic PD. It has been suggested that *SNCA* and *MAPT* may act synergistically to influence disease risk[23]. Understanding the interactions between synuclein and tau, and how they contribute to PD pathogenesis, will be an important goal of future research.

The secondary tauopathies are disorders where tau pathology exists but is not the primary pathological insult. An intriguing example of a secondary tauopathy is myotonic dystrophy type I (DM1), a multisystemic disorder caused by a CTG repeat expansion in the gene encoding myotonic dystrophy protein kinase (*DMPK*). One suggested disease mechanism is that the CUG repeats of *DMPK* mRNA could sequester RNA-binding proteins, thus leading to a generic disruption in transcript processing[24]. Altered splicing of tau exons 2, 6 and 10 of the tau gene is seen in DM1 brain and this is accompanied by insoluble tau tangles, consistent with the idea that aberrant tau splicing can predispose to tau aggregation[25]. Tau pathology has also been described in dementia pugilistica, multiple sclerosis, amyotrophic lateral sclerosis (ALS) and Niemann–Pick disease Type C; this list is not exhaustive and demonstrates the clinical diversity of disorders with tau pathology.

An obvious question when considering the secondary tauopathies is whether coincidental tau pathology has any influence on cell death and disease progression. With this in mind, it is important to remember that AD is actually a secondary tauopathy, and despite tau pathology being downstream of amyloid, tau deposition is not without consequence and remains centrally linked to neuronal demise. The contribution of genetic variation at the *MAPT* locus to the secondary tauopathies remains to be determined.

TDP-43 and the dark matter of genetics

Mutations in the progranulin gene (*PGRN*) were identified as the cause of ubiquitin-positive tau-negative chromosome-17 linked FTLD (FTLD-U)[26,27]. The main component of the pathological lesions in FTLD-U is not progranulin, but the trans-active response DNA protein 43 (TDP-43)[28]. Mutations in the TDP-43 gene (*TARDP*) causing ALS, a disorder closely related to FTLD-U, have been subsequently identified[29]. This is consistent with the theme of mutations in the proteins that are the main pathological component of the sporadic disease causing inherited forms of the disorder. However, the story subsequently diverges as, to date, no *TARDP* common variants pre-disposing to disease have been found either in candidate gene studies or GWAS in ALS. This could be explained in part by population diversity (as we have seen previously for *MAPT*) and common risk variants may yet be found in different populations. Another explanation is that rare variants of intermediate effect exist in *TARDP* that are yet to be discovered. Rare variants would not be captured by GWAS, which only assess common variation, and they could contribute substantially to risk without showing Mendelian linkage.

Although they almost certainly exist and their effects can be detected, technically, rare variants are much harder to capture and as such they have recently been termed the 'dark matter' of genetics. Rare variant capture will be an important step in understanding the missing heritability of common disorders and an excellent review of potential strategies is given by Manolio et al.[30].

Conclusions

The importance of tau in neurodegeneration was overlooked until the discovery of mutations in *MAPT* that cause autosomal-dominant disease. Since then, *MAPT* has emerged as a notable example of a common mechanism underlying clinically diverse disorders, where mutations and common variation at the same loci cause autosomal-dominant disease, and increase risk to sporadic disease respectively. Moreover, in disorders characterized by insoluble protein deposition, it is the genes encoding the aggregated protein that are genetically associated with disease. However, our understanding of the genetics of neurodegenerative disease is by no means complete, and the next challenge facing neurogeneticists is the complete capture of common and rare variants that predispose to disease risk. This will enable the complete functional dissection of the mechanisms of protein deposition and neuronal death in inherited and sporadic neurodegenerative disorders. ■

Selina Wray graduated from the University of Nottingham in 2004 with an MSci in Biochemistry and Biological Chemistry. She was awarded her PhD in 2008 from the MRC Centre for Neurodegeneration Research, King's College London, where she first started researching the role of tau in neurodegeneration. Selina is currently based in the Department of Molecular Neuroscience, UCL Institute of Neurology, where she has been awarded an Alzheimer's Research Trust fellowship to continue her research into the molecular mechanisms of dementia. email: S.Wray@ion.ucl.ac.uk

John Hardy is currently the Chair of Molecular Biology of Neurological disease at UCL Institute of Neurology. He started researching the genetics of Alzheimer's disease as an assistant professor at St Mary's Hospital, Imperial College London, prior to taking up posts at the University of South Florida and the Mayo Clinic, Jacksonville, FL, USA. John was Head of Neurogenetics at the National Institute of Ageing, Bethesda, MD, USA, from 2001, before returning to London in 2007. Research in John's group continues to focus on the genetic basis of neurological disease, and in 2009, John was elected Fellow of the Royal Society in recognition of his achievements in this area. email: J.Hardy@ion.ucl.ac.uk

References

1. Andreadis, A., Brown, W.M. and Kosik, K.S. (1992) Biochemistry **31**, 10626–10633
2. Kosik, K.S., Orecchio, L.D., Bakalis, S. and Neve, R.L. (1989) Neuron **2**, 1389–1397
3. Lindwall, G. and Cole, R.D. (1984) J. Biol. Chem. **259**, 5301–5305
4. Grundke–Iqbal, I., Iqbal, K., Tung, Y.C., Quinlan, M., Wisniewski, H.M. and Binder, L.I. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 4913–4917
5. Hardy, J.A. and Higgins, G.A. (1992) Science **256**, 184–185
6. Alafuzoff, I., Iqbal, K., Friden, H., Adolfsson, R. and Winblad, B. (1987) Acta Neuropathol. **74**, 209–225
7. Poorkaj, P., Bird, T.D., Wijsman, E. et al. (1998) Ann. Neurol. **43**, 815–825
8. Rapoport, M., Dawson, H.N., Binder, L.I., Vitek, M.P. and Ferreira, A. (2002) Proc. Natl. Acad. Sci. U.S.A **99**, 6364–6369
9. Roberson, E.D., Scearce-Levie, K., Palop, J.J. et al. (2007) Science **316**, 750–754
10. Hutton, M., Lendon, C.L., Rizzu, P. et al. (1998) Nature **393**, 702–705
11. Goedert, M., Spillantini, M.G., Potier, M.C., Ulrich, J. and Crowther, R.A. (1989) EMBO J. **8**, 393–399
12. Gasparini, L., Terni, B. and Spillantini, M.G. (2007) Neurodegener. Dis. **4**, 236–253
13. Stefansson, H., Helgason, A., Thorleifsson, G. et al. (2005) Nat. Genet. **37**, 129–137
14. Baker, M., Litvan, I., Houlden, H. et al. (1999) Hum. Mol. Genet. **8**, 711–715
15. Pittman, A.M., Myers, A.J., Duckworth, J. et al. (2004) Hum. Mol. Genet. **13**, 1267–1274
16. Pittman, A.M., Fung, H.-C. and de Silva, R. (2006) Hum. Mol. Genet. **15**, R188–R195
17. Myers, A.J., Pittman, A.M., Zhao, A.S. et al. (2007) Neurobiol. Dis. **25**, 561–570
18. Caffrey, T.M., Joachim, C., Paracchini, S., Esiri, M.M. and Wade-Martins, R. (2006) Hum. Mol. Genet. **15**, 3529–3537
19. Hardy, J. (2005) Biochem. Soc. Trans. **33**, 578–581
20. Satake, W., Nakabayashi, Y., Mizuta, I. et al. (2009) Nat. Genet. **41**, 1303–1307
21. Simón-Sánchez, J., Schulte, C., Bras, J.M. et al. (2009) Nat. Genet. **41**, 1308–1312
22. Tobin, J.E., Latourelle, J.C., Lew, M.F. et al. (2008) Neurology **71**, 28–34
23. Goris, A., Williams-Gray, C.H., Clark, G.R., et al. (2007) Ann. Neurol. **62**, 145–153
24. Miller, J.W., Urbinati, C.R., Teng-Umnuay, P. et al. (2000) EMBO J. **19**, 4439–4448
25. Dhaenens, C.M., Schraen-Maschke, S., Tran, H. et al. (2008) Exp. Neurol. **210**, 467–478
26. Baker, M., Mackenzie, I.R., Pickering-Brown, S.M. et al. (2006) Nature **442**, 916–919
27. Cruts, M., Gijselinck, I., van der Zee, J. et al. (2006) Nature **442**, 920–924
28. Neumann, M., Sampathu, D.M., Kwong, L.K. et al. (2006) Science **314**, 130–133
29. Sreedharan, J., Blair, I.P., Tripathi, V.B. et al. (2008) Science **319**, 1668–1672
30. Manolio, T.A., Collins, F.S., Cox, N.J., et al. (2009) Nature **461**, 747–753
31. Caffrey, T.M. and Wade-Martins, R. (2007) Neurobiol. Dis. **27**, 1–10