

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Gene xx (2005) xxx–xxx

**GENE**  
 SECTION  
 EVOLUTIONARY GENOMICS
[www.elsevier.com/locate/gene](http://www.elsevier.com/locate/gene)

## The prion protein gene: Identifying regulatory signals using marsupial sequence

Marko Premzl<sup>a,b,\*</sup>, Margaret Delbridge<sup>a</sup>, Jill E. Greedy<sup>b</sup>, Peter Wilson<sup>c</sup>, Matthew Johnson<sup>c</sup>, John Davis<sup>c</sup>, Elizabeth Kuczek<sup>c</sup>, Jennifer A. Marshall Graves<sup>a</sup>

<sup>a</sup>Comparative Genomics Group, Research School of Biological Sciences, Australian National University, P.O. Box 475, Canberra, ACT 2601, Australia

<sup>b</sup>Computational Proteomics Group, The John Curtin School of Medical Research, Australian National University, P.O. Box 334, Canberra, ACT 2601, Australia

<sup>c</sup>Australian Genome Research Facility, Gehrman Laboratories, University of Queensland, St Lucia, QLD 4072, Australia

Received 12 August 2004; received in revised form 11 November 2004; accepted 26 November 2004

Received by T. Gojobori

### Abstract

The function of the prion protein gene (*PRNP*) and its normal product PrP<sup>C</sup> is elusive. We used comparative genomics as a strategy to understand the normal function of *PRNP*. As the reliability of comparisons increases with the number of species and increased evolutionary distance, we isolated and sequenced a 66.5 kb BAC containing the *PRNP* gene from a distantly related mammal, the model Australian marsupial *Macropus eugenii* (tammar wallaby). Marsupials are separated from eutherians such as human and mouse by roughly 180 million years of independent evolution. We found that tammar *PRNP*, like human *PRNP*, has two exons. Prion proteins encoded by the tammar wallaby and a distantly related marsupial, *Monodelphis domestica* (Brazilian opossum) *PRNP* contain proximal PrP repeats with a distinct, marsupial-specific composition and a variable number. Comparisons of tammar wallaby *PRNP* with *PRNPs* from human, mouse, bovine and ovine allowed us to identify non-coding gene regions conserved across the marsupial–eutherian evolutionary distance, which are candidates for regulatory regions. In the *PRNP* 3' UTR we found a conserved signal for nuclear-specific polyadenylation and the putative cytoplasmic polyadenylation element (CPE), indicating that post-transcriptional control of *PRNP* mRNA activity is important. Phylogenetic footprinting revealed conserved potential binding sites for the MZF-1 transcription factor in both upstream promoter and intron/intron 1, and for the MEF2, MyT1, Oct-1 and NFAT transcription factors in the intron(s). The presence of a conserved NFAT-binding site and CPE indicates involvement of PrP<sup>C</sup> in signal transduction and synaptic plasticity. © 2004 Elsevier B.V. All rights reserved.

**Keywords:** Prion protein gene; Prion protein; Comparative genomics; 3' untranslated region; Phylogenetic footprinting; Transcription factors; Signal transduction; Synaptic plasticity

**Abbreviations:** *PRNP*, prion protein gene; PrP, prion protein; PrP<sup>C</sup>, normal isoform of prion protein; PrP<sup>Sc</sup>, pathogenic isoform of prion protein; *SPRN*, Shadow of prion protein gene; Sho, Shadow protein; ORF, open reading frame; PCR, polymerase chain reaction; RT-PCR, reverse transcriptase PCR; MY, million years; BAC, bacterial artificial chromosome; FISH, fluorescent in situ hybridisation; UTR, untranslated region; CPE, cytoplasmic polyadenylation element; CPEB, cytoplasmic polyadenylation element binding protein; MZF-1, Myeloid zinc finger-1; Oct-1, Octamer 1; MEF2, Myocyte enhancer factor-2; MyT1, Myelin transcription factor 1; NFAT, Nuclear factor of activated T-cells.

\* Corresponding author. Division of Molecular Bioscience, The John Curtin School of Medical Research, Australian National University, Canberra, Australia.

E-mail address: [Marko.Premzl@anu.edu.au](mailto:Marko.Premzl@anu.edu.au) (M. Premzl).

0378-1119/\$ - see front matter © 2004 Elsevier B.V. All rights reserved.  
doi:10.1016/j.gene.2004.11.049

## 1. Introduction

The prion protein gene *PRNP* is best known for its role in prion diseases, but its normal function remains elusive. Its product, prion protein (PrP), has the ability to fold into a dynamic, physiological conformation (PrP<sup>C</sup>), and into a compact, pathogenic conformation (PrP<sup>Sc</sup>) that causes prion diseases (Prusiner, 1998).

There are several hypotheses about the normal role of PrP<sup>C</sup> consistent with its localization on the cell membrane. Among other alternatives, PrP<sup>C</sup> could be a signal transduction protein, as its activation in vitro triggers a signalling pathway for which the terminal targets in both neuronal and non-neuronal cells are the MAP kinases ERK1/2 (Schneider et al., 2003).

Mammalian *PRNP* is a housekeeping gene. It has been characterized in several eutherian species: hamster (Li and Bolton, 1997), human, sheep, mouse (Lee et al., 1998) and bovine (Hills et al., 2001). These analyses have identified, as conserved features of eutherian *PRNP* promoters, their GC richness and a lack of TATA box.

However, there are some differences in gene structure and regulation of gene expression among species. *PRNP* genes contain three exons, with exons 1 and 2 encoding the 5' UTR region of mRNA in mouse, sheep (Lee et al., 1998), rat (Saeki et al., 1996) and bovine (Hills et al., 2001), but only two exons in human, the first of which encodes the 5' UTR region (Lee et al., 1998). Two or three exons are transcribed alternatively in the mRNA encoded by Syrian hamster *PRNP* (Li and Bolton, 1997). Complete ORF and 3' UTR region are encoded by the 3' terminal *PRNP* exon. There is a single transcription start site in all eutherian *PRNPs* known except for rodent species (mouse, rat, hamster), which have multiple transcription start sites.

Next, regulatory signals controlling expression of the mammalian *PRNP* gene have not yet been resolved by *PRNP* gene expression studies (Table S1). Experiments on rat, mouse, bovine and human *PRNP* have variously ascribed regulatory roles for the elements in promoter and intron/intron 1. The problems are that functional regions differ between species and between the cell lines used in different experiments. Of note here is that the regulatory elements that determine physiological expression of *Prnp* reside in the upstream promoter and both introns in mouse (Fischer et al., 1996), but no regulatory elements were described in the intron 2 by present.

Comparative genomics is a powerful tool for understanding biological function. Such analysis is particularly useful for detecting conserved regulatory sequences (e.g. transcription factor-binding sites), a process known as phylogenetic footprinting (Blanchette and Tompa, 2002). However, availability of genomic DNA sequence from representatives of major mammal and vertebrate lineages limit such studies.

Reliability of the cross-species analysis increases with number of species compared and the optimal evolutionary distance between species depends on the biological question

addressed. We therefore isolated and characterized the *PRNP* gene from a distantly related mammal in order to identify features of *PRNP* gene structure and discover potential regulatory elements. Marsupial mammals diverged from eutherian (“placental”) mammals about 180 million years ago, so provide important middle ground (~180 MY) between intra-eutherian (~80 MY) and mammal-bird (~310 MY) comparisons (Wakefield and Graves, 2003). Comparisons of marsupial and eutherian genes and chromosomes have provided many important and unexpected insights, for instance in identifying the mammalian testis—determining gene and understanding sex chromosome evolution, and have also lead to discovery of new human genes (reviewed by Graves and Westerman, 2002).

We used the model Australian marsupial *Macropus eugenii* (tammar wallaby). Here we report the isolation and characterization of a BAC harbouring the tammar *PRNP* gene, and compare it with the *PRNP* genes from species in which prion diseases occur naturally (human, bovine, ovine) or experimentally (mouse). This comparison across the eutherian–marsupial distance enabled us to infer the evolution and dynamics of the mammalian *PRNP* gene and identify conserved non-coding gene segments representing potential regulatory elements.

## 2. Materials and methods

### 2.1. Cloning of tammar wallaby *PRNP* cDNA

Using genomic DNA as a template we performed PCRs in a total volume of 50 µl containing 1× PCR buffer (10 mM Tris–HCl, 50 mM KCl, 1.5 mM MgCl<sub>2</sub>; Roche), 200 µM dNTPs (Roche), 1–2 U of the Taq polymerase (Roche), 200 ng of template, and 200 pmol of degenerate primers G-Forward and G-Reverse, respectively (Table S2). After 2 min of denaturation at 94 °C, we ran 35 amplification cycles using a touch-down protocol as follows: 1 min of denaturation at 94 °C, 1 min of annealing with temperatures ranging from 59 °C to 52 °C, and 1 min of extension at 72 °C. Finally, we extended the PCR products further during 10 min at 72 °C. The 214 bp PCR product was cloned using the pGEM-T Easy (Promega) cloning kit. The plasmids harbouring cloned fragment were templates for sequencing reactions using the standard primers T3 and T7 and BigDye Terminator v3.1 Cycle sequencing kit (Applied Biosystems). Products of the sequencing reaction were run on an ABI3730 DNA sequencer (Applied Biosystems).

Next, we screened a random primed pouch young tammar wallaby cDNA library as template using PCR. This library was made previously using the ZAPII vector (Clontech) and a female (day 0) and a male (day 0) tammar wallaby pouch young as mRNA sources. Its titre was 10<sup>9</sup> pfu/ml. After two extension steps in the reverse direction, and after four extension steps in the forward direction, we

extended the initial 215 bp to a final length of 1978 bp. Each extension step consisted of two rounds of amplification using semi-nested PCR. In round one, we used 1 µl of cDNA library as template with 50 pmol of upstream specific primer and 50 pmol of universal primer (T3 or T7). In round two, we used 1 µl of the first round PCR reaction as template with 50 pmol of downstream specific primer and 50 pmol of universal primer. The composition of the PCR reaction mixture was as above. The initial denaturation step was 7 min at 94 °C when the cDNA library was used as template, or 1 min at 94 °C when the PCR reaction was used as template, followed by 35 amplification cycles of 1 min denaturation at 94 °C, 1 min of annealing at 65 °C, and 2 min of extension at 72 °C. Finally, we extended PCR products for 10 min at 72 °C. The PCR products amplified in second round of each extension step were cloned using either the TOPO cloning kit (Invitrogen) or pGEM-T Easy (Promega) cloning kit, and sequenced as above. After each extension step we merged sequences, extended cDNA sequence, and designed new PCR primers.

## 2.2. Cloning of 3' end of *Monodelphis domestica* *PRNP* cDNA

We first amplified, cloned and sequenced a specific region of Brazilian opossum *PRNP* ORF using genomic DNA as a template and the degenerate primers G-Forward and G-Reverse as above. In order to amplify the 3' end of Brazilian opossum *PRNP* cDNA, we screened a poly-T primed *Monodelphis* testis cDNA library made previously using the predigested Lambda ZAP II/*Eco*RI/*CIAP* vector (Stratagene) as template using PCR. In the first step of asymmetric PCR, we used 1 µl of library as template, 50 pmol of primer MdF1 (Table S2) and 2.5 pmol of primer T7 with the same PCR conditions as above. In the second PCR step, we used 0.1 µl of the first step PCR reaction as template and 50 pmol of MdF1 and T7 primer each following the same reaction conditions as above. Finally, we cloned and sequenced the PCR products as above. The sequence (575 bp) contained part of the *PRNP* ORF between 1 and 360 bp (ORF coordinates 451–810 bp) and the mRNA 3' UTR region between 361 and 575 bp.

Using the tammar wallaby *PRNP* ORF nucleotide sequence as search query and the local BLAST service, we searched the *Monodelphis* Whole Genome Shotgun trace database (NCBI) and identified traces **334046496** and **346363827** corresponding to the Brazilian opossum PrP ORF regions between 1–470 bp and 1–333 bp, respectively. We combined this information with our experimental data and assembled a complete ORF. The ORF is 810 bp.

## 2.3. RT-PCR determination of tammar wallaby *PRNP* expression

We extracted RNA from snap-frozen tammar wallaby cerebral cortex, testis, mammary gland, liver, kidney, and

stomach tissue samples, respectively, using the RNeasy Mini Kit (Qiagen). We then used 1 µg of RNA as template and random hexamer primers in order to synthesize the first-strand cDNA by using the Expand Reverse Transcriptase kit (Roche), following the manufacturer's instructions. Next, we designed PCR primers spanning the tammar wallaby *PRNP* intron (5'RT-F and 5'RT-R, Table S2), and used 5 µl of the first-strand cDNA as template for the PCRs with 25 pmol of each primer. After 5 min of initial denaturation at 94 °C, we ran 30 amplification cycles with 45 s of denaturation at 94 °C, 30 s of annealing at 54.5 °C, and 30 s of extension at 72 °C. We further extended PCR products for 10 min at 72 °C. Composition of the PCR reaction mixture was as above. The 185 bp PCR products were cloned and sequenced as above to confirm their identity.

## 2.4. Screening of tammar wallaby genomic DNA BAC library

We screened the tammar wallaby bacterial artificial chromosome (BAC) library (Victorian Institute of Animal Science, Sankovic et al., unpublished) to identify a clone harbouring the *PRNP* gene. Our probe was the specific 214 bp product from tammar wallaby *PRNP* amplified as above. We first labelled 25 ng of the probe with <sup>32</sup>P-dATP using the Megaprime™ DNA labelling kit (Amersham) and removed unincorporated <sup>32</sup>P-dATPs using the ProbeQuant™ G-50 Micro Columns (Amersham). Next, we prehybridized Hybond-N nylon membranes (Amersham) containing complete BAC library overnight at 65 °C in a modified Church and Gilbert buffer (0.25 M NaHPO<sub>4</sub>, 5 mM EDTA, 7% SDS, pH 7) containing 100 µg/ml of denatured salmon sperm DNA and 1% of bovine serum albumin (MP Biomedicals). We then added probe and hybridised it to the membranes at 65 °C overnight. After rinsing filters in 2× SSC and washing them in 2× SSC/0.1% SDS, and in 1× SSC/0.1% SDS, we exposed them to an X-ray film (Kodak) for 1 week.

## 2.5. PCR using BAC DNA

We used the BAC clone DNA as template (100 ng per PCR reaction) and 50 pmol of each G-Forward and G-Reverse primers (Table S2) as above to test whether BAC clone contains *PRNP* gene. We then cloned and sequenced the 214 bp PCR products as above.

## 2.6. Southern blotting

We cut the BAC clone DNA and genomic DNA, respectively, with restriction enzymes *Sal*I (Roche), *Xba*I (Roche), and both *Sal*I and *Xba*I, respectively. Restriction digests were electrophoresed on a 1% UltraPure™ Agarose (Invitrogen) gel and transferred to the Hybond-N+ (Amersham) nylon membranes according to the standard protocol (Sambrook et al., 1989). We used the same probe as above,

and we also labelled it, hybridised it to membranes, and later on washed the membranes as above. Finally, we exposed the membranes to the X-ray film (Kodak) overnight.

### 2.7. Fluorescent *in situ* hybridisation

We labelled the BAC clone DNA with biotin-14-dATP by nick translation following the BIONICK™ Labelling System kit (Invitrogen) protocol. We then co-precipitated 10 µg of labelled BAC DNA with 10 µg of sheared tammar wallaby genomic DNA and 10 µg of sheared salmon sperm DNA overnight at  $-20^{\circ}\text{C}$ , dissolved these precipitates in 50% formamide, 10% dextran sulphate and  $2\times$  SSC, denatured them at  $80^{\circ}\text{C}$ , pre-annealed them at  $37^{\circ}\text{C}$ , and finally applied them to pre-treated slides containing tammar wallaby chromosome metaphase spreads. After 3 days in moist chamber at  $37^{\circ}\text{C}$ , we washed the slides in buffers containing 50% formamide/ $2\times$  SSC,  $2\times$  SSC and  $0.1\times$  SSC, and then we blocked them in 5% BSA (MP Biomedicals) with 0.1% Tween 20 (SIGMA). Next, we applied FITC-conjugated avidin and Cy3-conjugated anti-digoxin antibodies dissolved in 1% BSA (MP Biomedicals), 0.1% Tween 20 (SIGMA) and  $4\times$  SSC to the slides during 30 min at  $37^{\circ}\text{C}$ . After post-antibody washes we stained the slides with DAPI (1 µg/ml in  $2\times$  SSC) and mounted them using the VECTASHIELD (VECTOR) mounting medium. We observed and recorded images using an Axioplan 2 microscope (Zeiss) and camera (Diagnostic Instruments).

### 2.8. Analysis of evolutionary distances

We inferred evolutionary distances in a set of aligned PRPs using the MEGA2 program (Kumar et al., 2001; <http://www.megasoftware.net>) run locally under Windows. Evolutionary distances were defined as a number of amino acid substitutions between a pair of sequences. Using the Complete-Delete option to remove alignment gaps, we calculated number of valid common sites and number of sites different between two sequences. We then determined percent of identity between a pair of sequences manually.

### 2.9. Sequencing of BAC clone

We picked one positive BAC colony and grew it for 6 h in LB medium containing chloramphenicol. Then we inoculated 500 ml of YT broth with 1 ml of culture and grew it for 18 h. Next, we used the Qiagen Large Construct Kit to prepare the BAC DNA. We sheared 5 µg of BAC DNA to a 2–4 kb range using the Hydroshear (GeneMachines), and then we end-repaired it with the DNA Terminator End-Repair kit (Lucigen) according to the manufacturer's instructions. Next, we purified sheared fragments using the QiaExII (Qiagen), and quantitated them on agarose gel using the Mass Ladder markers (Fermantas). We then cloned the sheared and end-filled DNA fragments (0.5 µg/ligation) into a pSMART-HC vector system (Luci-

gen) according to the manufacturer's instructions. We finally prepared plasmids as templates for sequencing reactions, sequenced them using the BigDye Terminator v3.1 Cycle sequencing kit (Applied Biosystems), and analysed sequencing reaction products using the ABI3730 DNA sequencer (Applied Biosystems). We placed all sequences in the Phred/Phrap/Consed System (<http://www.phrap.org>) for assembly and finishing.

### 2.10. Transposable element content analysis

We used the slow speed option of the RepeatMasker program as a free web service (<http://ftp.genome.washington.edu/RM/RepeatMasker.html>) to determine the relative content of interspersed repeats, small RNA, satellites, simple repeats and low complexity sequence in our nucleotide sequences. First, we analysed the transposable element content in the tammar wallaby, human, mouse, bovine and ovine *PRNP* genes. Then we determined the distribution of transposable element in the tammar wallaby, human, mouse, bovine and ovine genomic sequences harbouring the *PRNP* gene described below.

### 2.11. Cross-species comparisons

We used the VISTA web server (Mayor et al., 2000; <http://www-gsd.lbl.gov/VISTA/>) to align *in silico* genomic sequences harbouring the *PRNP* gene. The 66.5 kb BAC sequence contained the *PRNP* gene (30055–50234 bp) together with flanking sequences. The human (chr20: 4558866–4650555 bp; Ensembl v16.33.1) and mouse (Chr2F3: 132052342–132080812 bp; Ensembl v22.32b.1) genomic DNA sequences contain complete proximal and distal intergenic regions. *PRNP* in the bovine genomic sequence (AJ298878; NCBI) encompasses 49430–69659 bp. We merged two overlapping ovine genomic sequences (U67922 and AY184242; NCBI) into a single 46955 bp contig with the *PRNP* lying between 5666 and 26295 bp. This sequence contains the complete distal intergenic region (20961 bp). Transposable elements (but not simple repeats and low complexity DNA) in genomic sequences were masked using the RepeatMasker program as above. We used the human sequence and its annotation from the Ensembl genome browser as the base sequence. Pairwise sequence comparisons were calculated with a threshold of 75% identity in a 100 bp window. The minimum identity shown in the VISTA plot is 40%.

### 2.12. Phylogenetic footprinting and transcription factor-binding site analysis

We performed phylogenetic footprinting of the tammar wallaby, human, mouse, bovine and ovine genomic sequences described above using the FootPrinter program (Blanchette and Tompa, 2002; <http://abstract.cs.washington.edu/~blanchem/FootPrinterWeb/FootPrinterInput.pl>) run

locally under operating system Red Hat Linux ver. 2.4.18-3 (<http://www.redhat.com>). We used a conservative approach in this analysis accepting only motifs detected in all five species. We allowed 2 bp mismatch for the 13 bp and 12 bp motifs (parsimony score 2), 1 bp mismatch for 11 bp and 10 bp motifs (parsimony score 1), and no mismatch for 9 bp and 8 bp motifs (parsimony score 0). Before phylogenetic footprinting, transposable elements and exons were masked and excised from the sequences. Analyses of the upstream *PRNP* promoter and *PRNP* intron(s) were run separately. Due to the limited size of the proximal intergenic sheep sequence (see above), we limited analysis of the upstream promoter to a final sequence size of approximately 5 kb. Because of the large intron sizes, we restricted the search for motifs in the intron(s) to a length of 10% of the human *PRNP* intron size (approximately 1270 bp), using an option `subregion_size 1270`, as suggested by Blanchette and Tompa (2002). We used the FootPrinter options `triplet_filtering` and `post_filtering` in all analyses.

We analysed potential transcription factor-binding sites in the same genomic sequences using the MatInspector program available as a web service (<http://www.genomatix.de/cgi-bin/eldorado/main.pl>). This program searches the query sequence for transcription factor-binding sites deposited in the TRANSFAC database. Again, we used a conservative approach in this analysis: motifs denoting transcription factor-binding sites were identified with core similarity 1 and matrix similarity score above the optimised matrix similarity score.

### 2.13. Databases

We downloaded human and mouse genomic sequences from the Ensembl genome browsers (<http://www.ensembl.org>). The bovine and ovine genomic sequences were downloaded from the NCBI database (<http://www.ncbi.nlm.nih.gov>).

### 2.14. Software

For basic handling of nucleotide and protein sequences we used the Vector NTI program (InforMax, Frederick MD, USA). Primers for PCR reactions were designed using the MacVector 7.0 program (Oxford Molecular Group 2000). We used the Cameleon v3.14 (Oxford Molecular 1995, now owned by Accelrys) to align amino acid sequences and the CHROMA program (<http://www.lg.ndirect.co.uk/chroma/index.htm>) to prepare the alignment figure for publication. Signal peptide cleavage sites were predicted using the SignalP program (<http://www.cbs.dtu.dk/services/SignalP/>) and GPI-anchor sites were predicted by using the bigPI-predictor program ([http://mendel.imp.univie.ac.at/gpi/gpi\\_server.html](http://mendel.imp.univie.ac.at/gpi/gpi_server.html)) available as web services. We used the NIX interactive program to annotate the tammar BAC sequence (<http://menu.hgmp.mrc.ac.uk/menu-bin/Nix/Nix.pl>). Figures were drawn using Microsoft PowerPoint and Adobe Illustrator.

## 3. Results

### 3.1. Isolation and initial characterization of tammar wallaby *PRNP*

#### 3.1.1. Cloning of tammar wallaby *PRNP* cDNA and 3' end of *M. domestica* *PRNP* cDNA

We cloned the tammar wallaby cDNA using modified strategy of Simonic et al., 1997 (Fig. 1). First, using the degenerate primers and genomic DNA as template, we amplified, cloned and sequenced a short specific fragment of the *PRNP* ORF. In the second step, we used this sequence to design specific nested primers to screen the cDNA library. After each of six semi-nested PCR experiments, the sequence of the cDNA was further extended. The overlapping sequences were finally merged into a single cDNA.

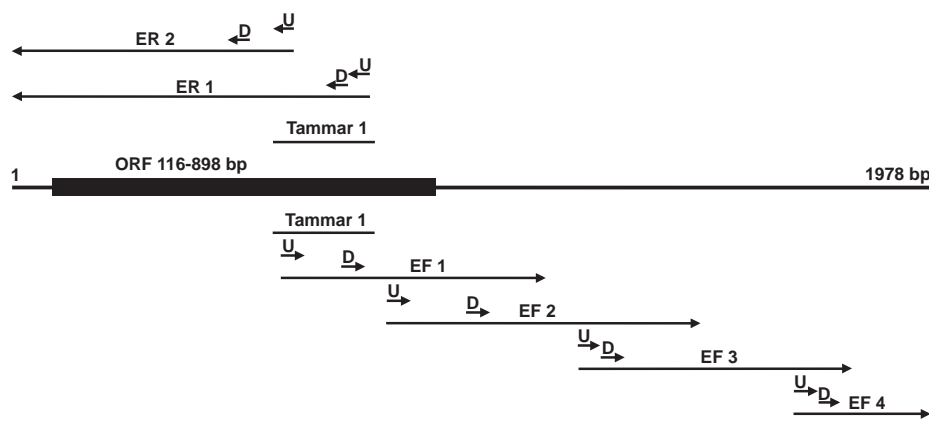


Fig. 1. Cloning of cDNA encoded by tammar wallaby *PRNP*. cDNA is shown by horizontal bar. Numbers denote sizes of cDNA and ORF in bp. Large arrows depict cDNA fragments that were amplified from random primed pouch young cDNA library; small arrows indicate positions of primers for semi-nested PCR. Tammar 1, region of ORF cloned by using genomic DNA as template and degenerate primers G-Forward and G-Reverse; ER, extension step in reverse direction; EF; extension step in forward direction; U, upstream primer for semi-nested PCR; D, downstream primer for semi-nested PCR; ORF, open reading frame.

The 1978 bp cDNA harbours the complete ORF for PrP. The 783 bp ORF (positions 116–898 bp in the cDNA sequence) translates into a protein of 260 amino acids. The nucleotide (**AY659988**) and protein (**AAT68002**) sequences were deposited in the NCBI.

Using the same strategy we first cloned a short specific region of the Brazilian opossum *PRNP* ORF. Using this sequence to design a specific primer, we used asymmetric PCR to screen the Brazilian opossum testis poly-A primed cDNA library and amplify a 3' end of cDNA (575 bp) harbouring the partial ORF and complete 3' UTR of the cDNA. The complete ORF (810 bp) encoding a protein of 269 residues was assembled by combining our experimental data with the Monodelphis Whole Genome Shotgun traces deposited in the NCBI database. The partial cDNA (**AY659989**) and ORF (**BK005535**) nucleotide sequences were deposited in the GenBank database.

### 3.1.2. Expression of tammar wallaby *PRNP*

We used RT-PCR with the intron-spanning primers to show expression of the tammar wallaby *PRNP* in cerebral cortex, testis, mammary gland, liver, kidney and stomach (Fig. S1).

### 3.1.3. Isolation of BAC clone containing tammar wallaby *PRNP*

We used the amplified segment of the tammar *PRNP* coding region also as a probe to screen the tammar wallaby DNA BAC library and detected a single BAC clone. To test whether its DNA insert harbours the *PRNP* gene, we extracted and purified the BAC DNA and used it as template for the PCR. Then we cloned and sequenced the amplified product and confirmed its identity. Results of Southern blotting using tammar BAC and genomic DNA also confirmed the presence of *PRNP* in the BAC clone DNA insert (not shown).

We then sequenced the entire BAC insert (66512 bp) as described in Materials and methods. The nucleotide (**AY659987**) and protein (**AAT68001**) sequences were deposited in the NCBI.

### 3.1.4. Location of tammar wallaby *PRNP*

We determined the physical location of the tammar wallaby *PRNP* using fluorescent in situ hybridisation (FISH). The whole BAC clone DNA was labelled with biotin-14-dCTP by nick translation, then hybridised to tammar wallaby metaphase chromosome spreads. The *PRNP* BAC mapped to the distal end of long arm of tammar wallaby chromosome 1 (Fig. S2).

## 3.2. Analysis of tammar wallaby prion protein

We aligned the tammar wallaby PrP of 260 amino acids (**AAT68002**) and Brazilian opossum PrP of 269 amino acids (translated from **BK005535**) with the human (**NP\_898902**), mouse (**NP\_035300**), bovine (**Q01880**), ovine

(**AAR14248**), possum (**AAA61833**), chicken (**P27177**), turtle (**CAB81568**) and frog PrP (**CAC86159**) sequences (Fig. 2).

We provisionally divided the PrP sequence into several regions with distinctive amino acid character: signal peptides (proximal indicating PrPs are extracellular proteins and distal denoting GPI-anchor addition site), basic region(s), repeats, hydrophobic region, and C-terminal region (Premzl et al., 2003).

There is an extensive conservation in the middle stretch of PrP extending from the second basic segment (W 99 in human sequence) to the C-terminal region (Y 169 in human PrP) including the whole hydrophobic region. However, this amino acid alignment revealed differences in the PrP conservation as well.

We inferred evolutionary distances between the tammar wallaby PrP and other PrPs from our alignment using the MEGA2 program to calculate number of sites at which two sequences compared are different in the full-length protein and the four regions above (Table 1).

The full-length amino acid sequence of tammar wallaby PrP shows the highest identity with the marsupial and eutherian PrPs, less with the chicken and turtle PrPs, and least with the frog PrP.

The hydrophobic region in the middle of the protein is the most conserved region of PrP. On the other hand, the repeat region shows good conservation among the mammalian PrPs but much less with the chicken and turtle PrP, and none with the frog PrP. PrPs of eutherian mammals contain five octa- or nona-repeats (PH/QGGG/T(G)WGQ), and the tammar wallaby PrP contains four nona- or deca-repeats (PQGGGTNWGQ, PHPGGSSWGQ, PHAGGSNWGQ, PHGGSNWGQ). Whereas the possum PrP also contain four nona- or deca-repeats (PQGGGTNWGQ, PHPGGSNWGQ, PHPGGSSWGQ, PHGGSNWGQ), there are five deca-repeats in the Brazilian opossum PrP (PQGGGTNWGQ, PHAGGSNWGQ, PRPGGSNWGQ and duplicated PHPGGSNWGQ). This compares with seven hexa-repeats (PRQPGY, PHNPGY, PQNPGY) and one hepta-repeat (PGWGQGY) in the chicken PrP, and eight hexa-repeats (PSNPGY, PQNPGY, PRNPSY, PHNPAY, PPNPAY, PPNPGY, PHNPSY, PQNPGY) in the turtle PrP. There are no repeats in the *Xenopus laevis* PrP.

The basic region is less conserved among the mammalian PrPs than the repeats. However, when the chicken and turtle PrPs are compared with the tammar wallaby PrP, it is better conserved than the repeat region.

In the C-terminal region, the two cysteine residues that form a disulphide bond are conserved, as well as the two glycosylation consensus sites common to all PrPs (but not the third glycosylation consensus site of chicken PrP). All residues comprising the binding site for protein X (Prusiner, 1998) are conserved in the marsupial sequences: Q 175, Q 179, I 222 and Q 226 in the tammar PrP. A peculiar polythreonine stretch (T 195–T 200 in the tammar PrP) flanked by basic residues is also conserved in the



Table 1  
Evolutionary distances between tammar wallaby PrP and human, mouse, bovine, ovine, possum, opossum, chicken, turtle and *Xenopus laevis* PrPs

Tammar wallaby	N	Human		Mouse		Bovine		Ovine		Possum		Opossum		Chicken		Turtle		Xenopus							
		L	d	%	L	d	%	L	d	%	L	d	%	L	d	%	L	d	%	L	d	%			
Full-length	260	242	52	79	242	151	80	244	46	81	258	15	94	248	28	89	211	112	47	211	113	46	184	111	40
Basic	55	45	10	78	45	9	80	45	9	80	53	2	96	46	6	87	34	15	56	36	16	56	43	19	56
Repeats	39	33	4	88	32	4	87	33	4	88	36	1	97	39	4	90	24	14	42	24	16	33	–	–	–
Hydrophobic	20	20	2	90	20	0	100	20	0	100	20	0	100	20	0	100	20	2	90	20	2	90	16	3	81
C-terminal	99	98	23	77	99	23	77	98	15	85	99	6	94	96	9	91	86	46	47	83	46	45	78	49	37

N, number of residues in tammar wallaby PrP; L, number of valid common sites in pairwise alignments; d, number of sites different; %, percent identity.

mammalian PrPs. There are several insertions in the chicken, turtle and frog PrPs compared with the mammalian proteins.

### 3.3. Annotation of tammar wallaby BAC containing PRNP

First, we aligned the cDNA nucleotide sequence with the BAC DNA insert sequence using the Vector NTI program package. Apart from the PRNP gene (30055–50234 bp), the tammar wallaby BAC genomic fragment harbours 30054 bp of the proximal intergenic sequence, and 16278 bp of the distal intergenic sequence. The PRNP exon 1 is between 30055–30159 bp, the exon 2 is between 48361–50234 bp and the ORF lies between 48371–49153 bp. We note exon 2 is incomplete as a conserved nuclear-specific polyadenylation signal is further downstream (see Section 3.3.3).

Next, in order to assess basic features of the BAC DNA insert, we employed a number of programs included in the NIX interactive tool to analyse the gene content, distribution of CpG islands, GC content and transposable element content.

BLASTN program analysis using data from the EMBL database indicated a significant match ( $E=1e-130$ ) between tammar sequence (48371–48634 bp) and the marsupial possum *Trichosurus vulpecula* PRNP (**P51780**). We predicted three CpG islands using GrailEXP analysis. The CpG island stretching between 30050–30278 bp corresponds to the promoter and first exon of the PRNP gene. Analysis with RepeatMasker showed that 29% of the sequence is derived from the interspersed repeats, simple repeats, and low complexity sequence, and that the overall GC content of the BAC DNA insert is 41%.

#### 3.3.1. Organization of PRNP genes

We determined the features of the tammar wallaby PRNP gene and compared the sizes and nucleotide composition of exons and intron with those of the human, mouse, bovine and ovine PRNP gene (Table S3).

Tammar wallaby (like human) PRNP has only two exons, compared with the three exons in the mouse, bovine, and ovine gene. The 3' terminal exon includes complete coding sequence in all these PRNPs. The GC content of the first exon is higher than the GC content of the other exons and introns.

We detected only one transcription start site in the tammar wallaby PRNP.

#### 3.3.2. Repeat content of the tammar wallaby PRNP

We analysed the distribution of transposable elements in the tammar wallaby PRNP using the RepeatMasker program. The distribution of repeats in tammar wallaby PRNP was compared with those within the human, mouse, bovine and ovine PRNP gene (Table 2).

This analysis revealed that the repeat classes within the PRNP gene are different in all the species. Whereas the

Table 2

Summary of transposable element content in tammar wallaby, human, mouse, bovine and ovine *PRNP* gene

Species	Length (bp)	SINE (%)	LINE (%)	LTR (%)	DNA (%)	Total (%)
Tammar	20180	13.1	6.8	0	3.3	23.2
Human	15241	4.6	40.7	0	0.9	46.2
Mouse	28471	5.6	4	46.6	0	56.2
Bovine	20230	5	15.2	4.9	6.1	31.2
Ovine	20630	4.2	13.6	6.4	5.9	30.1

SINE, short interspersed element; LINE, long interspersed element; LTR, retrotransposons; DNA, DNA transposons.

SINEs are the most abundant in the tammar wallaby *PRNP*, the human gene has accumulated mostly LINES. The mouse *Prnp* contains more LTRs than the other *PRNPs*, and the bovine and ovine *PRNPs* harbour relatively more DNA transposons.

### 3.3.3. Cross-species comparisons of genomic sequences harbouring *PRNP*

We aligned the human, mouse, bovine, ovine and tammar wallaby genomic sequences harbouring *PRNP* using the VISTA alignment tool (Fig. 3A and B).

The VISTA plot indicates conserved regions in the *PRNP* gene and adjacent genomic sequences. Only two regions in the 92 kb alignment, both in the *PRNP* 3' terminal exon, are conserved in all the species. The first conserved region lies in the coding sequence and the

second is in the terminal 3' untranslated region (3' UTR) of the *PRNP* gene.

We aligned genomic sequences corresponding to the conserved 3' UTR region (human 71242–71354 bp, mouse 78558–78670 bp, bovine 69539–69647 bp, ovine 26192–26300 bp, tammar wallaby 50846–50970 bp) using the VectorNTI program package (Fig. 3C). This alignment indicated in all the *PRNPs* conservation of the sequence ATTAAA that is a signal essential for nuclear-specific polyadenylation (Mendez and Richter, 2001). In the sheep sequence, we also identified a general consensus sequence (TTTTTAT) of cytoplasmic polyadenylation element (CPE) that is a potential binding site for the cytoplasmic polyadenylation element binding (CPEB) protein required for maturation-specific polyadenylation (Mendez and Richter, 2001; Si et al., 2003). In the other eutherian sequences aligned with the CPE in sheep we found one mismatch (TTTCTAT), and there are two mismatches in the tammar sequence (TTTATAG; or one mismatch in the sequence ATTTTAT lying 2 bp upstream between 50844–50968 bp). These two elements are also conserved in the 3' UTR of our Brazilian opossum cDNA sequence (AY659989). The nuclear-specific polyadenylation signal lies between 537–542 bp and the motif TTTATAG lies between 462 and 468 bp (or GTTTTAT lying 2 bp proximally between 460–466 bp). Next, 18 bp of the sequence immediately downstream to the putative CPE is conserved in the eutherian sequences with only two mismatches in the marsupial sequences. The whole 26 bp

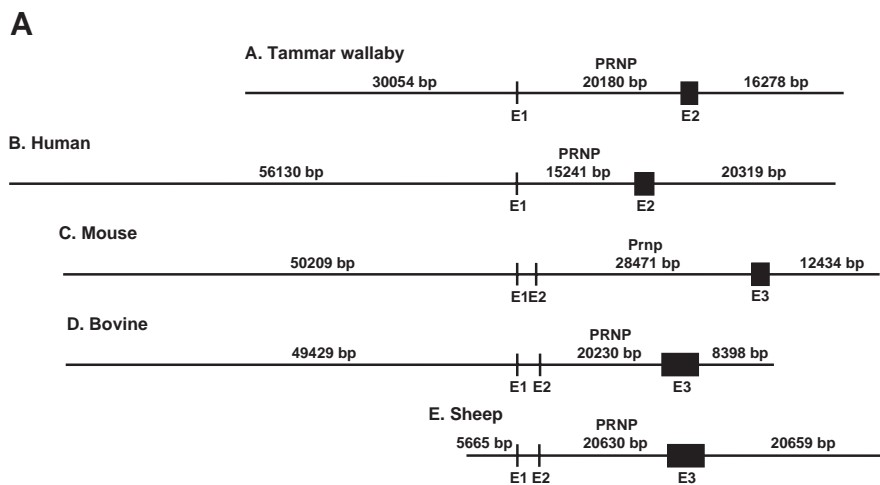


Fig. 3. Cross-species analysis. (A) Overview of genomic DNA sequences used in comparisons. Numbers denote sizes of *PRNP* gene and sequences proximal and distal to *PRNP* gene in bp. E1, exon 1; E2, exon 2; E3, exon 3. (B) VISTA plot indicates sequence homology in pairwise genomic DNA sequence alignments between: 1, human and mouse; 2, human and bovine; 3, human and sheep; 4, human and tammar wallaby. *PRNP*, prion protein gene. For human (base) sequence, direction of gene transcription is indicated by horizontal arrow, blue rectangles denote coding exons, light blue rectangles indicate 5' and 3' UTRs. Peaks in pairwise alignments illustrating sequence homology are depicted relative to human sequence (horizontal axis); percents of homology (40–100%) are indicated on vertical axis. Peaks in conserved non-coding sequence (CNS), UTR, and coding sequence fitting experimental cut-off for conservation (75% over 100 bp) are coloured pink, light blue, and blue, respectively. Asterisks denote gene regions conserved in all species. (C) Aligned genomic sequences corresponding to the conserved region in *PRNP* 3' UTR and to the *Monodelphis domestica PRNP* cDNA 3' UTR. Asterisk denotes consensus nuclear-specific polyadenylation signal site (ATTAAA). CPE, cytoplasmic polyadenylation element; general consensus CPE sequence in sheep (TTTTTAT) is marked by rectangle. Bold, putative CPE and conserved sequences immediately proximally and distally, and conserved nuclear-specific polyadenylation signal site.

sequence, including also one nucleotide immediately upstream to the putative CPE, is the best conserved region in this alignment.

Next, we determined the nucleotide composition of the aligned sequences using the Vector NTI software. The aligned human, mouse, bovine, ovine, tammar wallaby and

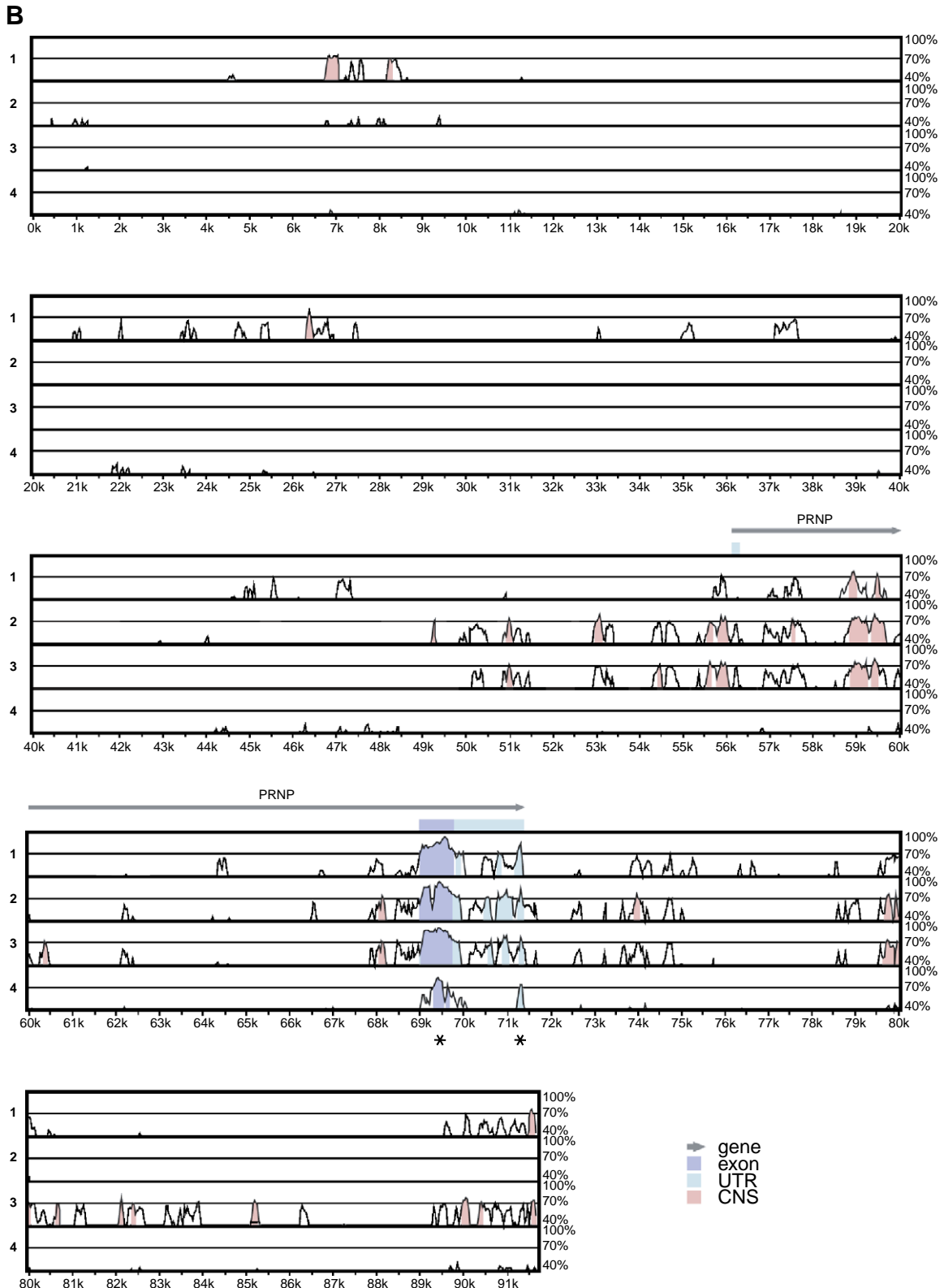


Fig. 3 (continued).

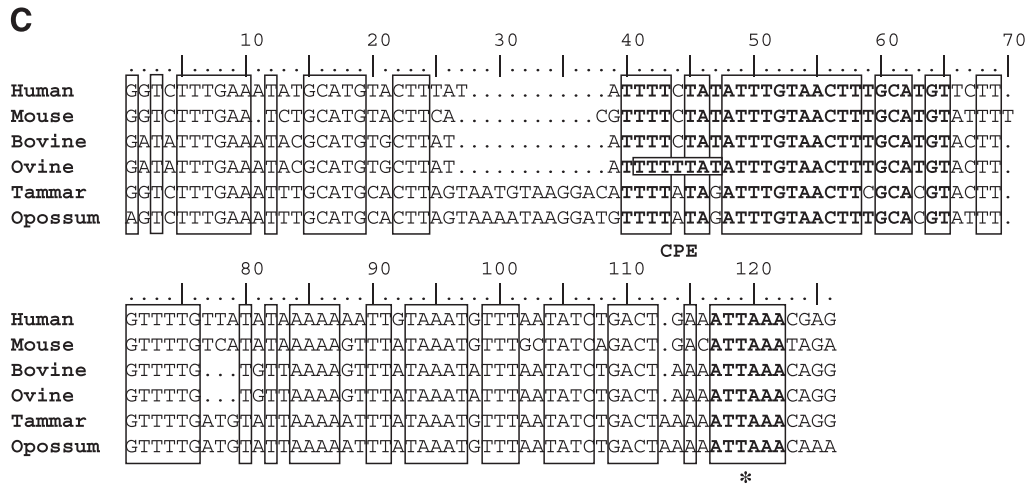


Fig. 3 (continued).

Brazilian opossum *PRNP* 3' UTR sequences are all AT rich (77%, 73%, 77%, 78%, 74% and 78%, respectively).

The VISTA plot showed no conservation in the exon 1 sequence (human sequence 56131–56235 bp) either across all the species, nor among the eutherians. A region in the upstream promoter (approximately 500 bp upstream to exon 1) is conserved in the human, bovine and ovine genes, but is not as well conserved in the mouse *Prnp* (Lee et al., 1998) and no conservation is detected in the tammar wallaby gene.

Conservation of the human and other eutherian “exon 2”/exon 2 sequences (58919–59017 bp in human) is evident in the VISTA plot, as well as conservation of the sequences within next distal 500 bp. There is no conservation of this sequence within the tammar *PRNP* gene.

### 3.3.4. Phylogenetic footprinting of *PRNP*

The disadvantage of the cross species genomic sequence alignments is that conserved elements shorter than the imposed experimental cut-offs are not recognised. In order to detect short conserved sequence fragments that could be

potential regulatory elements in the *PRNP* gene, we ran the FootPrinter program analysis. This software discriminates short conserved sequence motifs from the non-functional background noise across a set of orthologous gene sequences. In parallel to phylogenetic footprinting, we also analysed the same genomic sequences using the MatInspector program to find the potential binding sites for transcription factors. Then we matched manually the results of these two analyses.

Using FootPrinter we detected 51 short conserved motifs that may be potential regulatory elements: 18 in the promoter and 33 in the intron(s) of *PRNP*. Specifically, we found in the promoter three 12 bp motifs, four 10 bp motifs, one 9 bp motif, and ten 8 bp motifs (not shown). In the intron(s), we identified two 13 bp motifs, three 12 bp motifs, two 11 bp motifs, two 10 bp motifs, and twenty four 8 bp motifs (not shown).

Among the 51 motifs, we found manually seven motifs that were each predicted by MatInspector to bind the same transcription factor in all five species. One such motif was in

Table 3

Motifs harbouring conserved transcription factor-binding site

Motif	TF, strand, re-value	Human <sup>a</sup>	Mouse <sup>a</sup>	Bovine <sup>a</sup>	Ovine <sup>a</sup>	Tammar wallaby <sup>a</sup>
1. CCCT <b>CCCC</b>	MZF-1 – 4.03	–1209/–1202	–4238/–4231 –6450/–6443	–899/–892	–894/–887	–674/–667
2. GGG <b>AGGGG</b>	MZF-1 + 4.03	578/582	2376/2383	901/908	890/897	1629/1636
3. TTT <b>TTTAA</b>	MEF2 – 0.66	2727/2734	4614/4621 4775/4782	1090/1097 2338/2345 8049/8056	1084/1091 5621/5628	4960/4967 5152/5159 11157/11164
4. TTT <b>TAAAA</b> Ta/g	MEF2 + 0.66	3122/3132 6255/6265	2553/2563	2362/2372	2347/2357	5033/5043
5. <b>TGCATATT</b>	Oct-1 – 0.03	3321/3328	2162/2169	1703/1710	1699/1706	4110/4117
6. <b>AAGTTTTG</b>	MyT1 + 0.88	4267/4274	7461/7468	6799/6806	6847/6854	4850/4857
7. <b>ATTTTCCA</b>	NFAT – 2.06	6065/6072	4805/4812	4336/4343	6750/6757	5378/5385

Motif, conserved genomic DNA nucleotide sequence identified by phylogenetic footprinting; TF, transcription factor; strand, motif orientation with respect to *PRNP* gene; re-value, TF-binding site matrix expectation value for number of matches per 1000 bp of random DNA sequence; <sup>a</sup>, motif coordinates are relative to transcription start; bold and underlined, core TF-binding sequence.

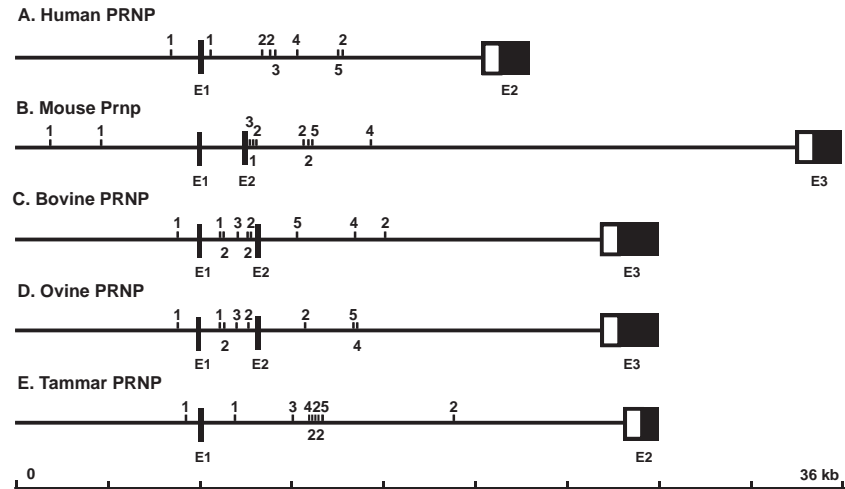


Fig. 4. Result of combined phylogenetic footprinting and transcription factor-binding site analyses. Numbers denote short sequence motifs predicted to bind same transcription factor each in all five species: 1, MZF-1-binding motif; 2, MEF2-binding motif; 3, Oct-1-binding motif; 4, MyT1-binding motif; 5, NFAT-binding motif. Figure is approximately to scale as shown by ruler. E1, exon 1; E2, exon 2; E3, exon 3; white box, ORF.

the promoter and six of these were in the intron(s) (Table 3, Fig. 4).

One conserved motif in the proximal promoter and one conserved motif in the intron/intron 1 were predicted to bind Myeloid zinc finger-1 (MZF-1) transcription factor. In the intron(s) solely we identified two conserved motifs predicted to bind Myocyte enhancer factor-2 (MEF2), and also motifs predicted to bind Octamer 1 (Oct-1), Myelin transcription factor 1 (MyT1), and Nuclear factor of activated T-cells (NFAT) transcription factor family, respectively. These putative regulatory signals lie in a similar order and at comparable relative distances across five genomic sequences.

The four short motifs conserved in the promoters of rodents, human, ovine and bovine *PRNP* of unknown functional significance (Lee et al., 1998) were not found in the tammar wallaby sequence.

#### 4. Discussion

Comparison of the tammar *PRNP* and adjacent sequences with sequences from the eutherian orthologues permits new insights into the evolution and regulatory organization of the mammalian *PRNP* gene.

##### 4.1. *PRNP* gene organization

Human *PRNP* organization has been regarded as somewhat atypical (Lee et al., 1998). However, we show here that two, rather than three, exons are present also in tammar *PRNP*.

##### 4.2. Evolution of PrP

Evolution has shaped regions in PrP differently. At one extreme, there is a strong conservation of the middle

hydrophobic region implying its functional and/or structural importance (Wopfner et al., 1999; van Rheede et al., 2003). This region is also essential for the conformational promiscuity of PrP (Prusiner, 1998). The only other mammalian protein known that contains similar sequence to this region is the Shadow protein encoded by the *SPRN* gene (Premzl et al., 2003). Remarkably, the broader conserved PrP region bordered by W 99 and Y 169 in the human PrP sequence corresponds exactly to a segment of PrP<sup>C</sup> that interacts with PrP<sup>Sc</sup> (Prusiner, 1998).

At the other extreme, the composition of repeat region exhibits variability. Whereas PrP repeats differ across vertebrate lineages, they are conserved within lineages (Wopfner et al., 1999; van Rheede et al., 2003), suggesting that they may have evolved lineage-specific roles, for example, binding of different ligands. Indeed, our comparison among the tammar wallaby, possum and Brazilian opossum PrP repeats showed that the marsupial PrP repeats (Windl et al., 1995) have conserved distinct character as well. We also demonstrate for the first time variability in the number of marsupial proximal repeats further indicating dynamic character of this PrP region.

Spontaneous prion formation can occur in any mammal expressing PrP (Legname et al., 2004), and conservation between the marsupial and eutherian PrPs may suggest that the marsupial PrPs have the same pathogenic potential.

##### 4.3. Transposable elements in *PRNP*

Eutherian *PRNPs* accumulate transposable elements extensively and in an independent fashion, as shown by Lee et al. (1998). The result of our RepeatMasker analysis also showed different *PRNP* dynamics in all five species. (We note, however, that caution is required in interpreta-

tion of our results as the marsupial, and to some extent ruminant, transposable elements are likely to be under-represented in the RepeatMasker libraries and this may result in experimental bias.) In stark contrast, the *SPRN* gene, a mammalian paralogue of *PRNP*, harbours no transposable elements at all (Premzl et al., 2004).

The Bov-B, Bov-tA and Mariner transposable elements are integrated in the sheep and bovine *PRNP* 3' UTR (Lee et al., 1998). It has been shown that recently expanded gene classes have their transcripts enriched with transposable elements (van de Lagemaat et al., 2003). Therefore, in the genes with mammalian-specific function such as response to external stimuli, stress response, and defence transposable elements are likely to be over-represented. Conversely, they are excluded from highly conserved genes. Consistent with our analysis of PrP evolution above, this suggests that the mammalian *PRNP* gene, or at least regions of PrP (e.g. repeats), may have evolved mammalian-specific role. The function of the paralogue *SPRN*, on the other hand, could be more conserved.

#### 4.4. Cross-species comparison of *PRNP*

In mRNAs, 3' UTRs are known to contain *cis*-acting regulatory elements that modulate post-transcriptional mRNA activities. We showed that the nuclear-specific polyadenylation signal site sequence in *PRNP* is conserved mammalian-wide, as already known for the eutherian sequences (Lee et al., 1998). This sequence and the maturation-specific polyadenylation signal CPE consensus sequence in sheep are the only *cis*-acting signals we recognised in the conserved *PRNP* 3' UTR. Whereas the other aligned sequences are different from the general consensus CPE sequence in sheep, we note that the size and content of CPEs vary as well as their number and relative distance from the nuclear-specific polyadenylation signal site (Mendez and Richter, 2001). These two regulatory signals together modulate the post-transcriptional regulation of activities of mRNAs encoding structural and regulatory factors required for maintenance of the long-term forms of synaptic plasticity (Si et al., 2003). Although we cannot draw conclusions upon computational analysis solely, we note that the activation of PrP<sup>C</sup> in vitro does upregulate the activity of the MAPKs ERK1 and ERK2, modulators of expression of genes involved in synaptic plasticity (Schneider et al., 2003). Further, an analysis of patterns of gene expression during the development of the mouse hippocampus showed clustering of *Prnp* with the genes involved in synaptic function, signal transduction, control of transcription and translation, glucose and oxidative metabolism and membrane regulation of ionic concentration, indicating involvement of PrP<sup>C</sup> in synaptic plasticity (Mody et al., 2001).

We conclude that conservation across the mammalian-wide evolutionary range indicates conserved function for the *PRNP* 3' UTR.

#### 4.5. Phylogenetic footprinting and regulatory organization of *PRNP*

The common transcription factor-binding sites were found in either eight bp or nine bp conserved motifs. The motifs were found in the promoter and intron(s), in line with previous observations (Fischer et al., 1996). An eight bp motif will occur randomly once in 65536 bp, and a nine bp motif will occur randomly once in 262144 bp. Conservation and similar arrangement of such motifs in a number of species separated by approximately 180 MY of independent evolution suggests their functional relevance.

MZF-1 regulates cellular proliferation but it is also a negative regulator of expression of the mouse extracellular superoxide dismutase (EC-SOD or SOD3) (Zelko and Folz, 2003). Some features of the mammalian EC-SOD-coding gene such as two or three exons with the entire coding sequence within the 3' terminal exon, lack of the TATA box and ubiquitous expression are in common with those of *PRNP*.

The neuron-specific cofactors such as MEF2 drive high level expression of neuronal genes acting co-ordinately with the ubiquitous transcription factors (Okamoto et al., 2002).

Oct-1 is a ubiquitous transcription factor but also controls tissue-specific genes including several in brain or lymphatic cells (reviewed by Sytina and Pankratova, 2003).

Myt1 is essential for control of expression of genes involved in terminal neuronal differentiation (Bellefroid et al., 1996).

Transcription factors from the NFAT family may influence various forms of synaptic plasticity. NFATc4 or NFAT3 is essential for regulation of gene expression during the brain-derived neurotropic factor-mediated synaptic plasticity (Groth and Mermelstein, 2003). Expression of genes involved in the long-term facilitation of synaptic plasticity is induced in the nucleus and the mRNA activities are regulated locally (Si et al., 2003). This analysis indicates conserved non-coding regions in the *PRNP* gene that could meet molecular requirements for these modes of gene and transcript regulation. The conserved NFAT-binding site may be a signal for induction of the mammalian *PRNP* gene transcription in the nucleus, and the conserved 3' UTR could mediate post-transcriptional regulation of mRNA activity.

The NFAT proteins have an essential role in the immune response (reviewed in Rao et al., 1997). Apart from the neuronal cells, PrP<sup>C</sup> activation in vitro increased the activity of the MAPKs ERK1 and ERK2 in the lymphoid and neuroendocrine cells as well (Schneider et al., 2003). The conserved NFAT transcription factor-binding site could, therefore, also mediate *PRNP*'s contribution to activation of lymphoid cells.

#### 4.6. Conclusion

We demonstrate here utility of the marsupial sequence in analysis of the human disease-related *PRNP* gene. We

compared it with eutherian sequences and found a number of conserved non-coding regions representing potential regulatory elements. Mammalian-wide sequence conservation suggests functional relevance for these gene regions. Our findings shed new light on the normal function of mammalian *PRNP* gene supporting the signal transduction hypothesis. The *PRNP* gene could be involved in signalling cascades that contribute to cell–cell interactions and act anti-apoptotically. Yet, in the heterogenous set of cells expressing PrP<sup>C</sup>, this will contribute to a number of different phenotypes such as the synaptic plasticity and activation of lymphoid cells.

### Acknowledgements

JEG and JAMG are supported by the ANU IAS block grant. We thank Prof. Frances Shannon, The John Curtin School of Medical Research, ANU, for valuable discussion. We thank Dr. Frank Grützner, Research School of Biological Sciences, ANU, for assistance with the FISH experiments. Dr. Denise Carvalho-Silva, Research School of Biological Sciences, ANU, provided the *Monodelphis domestica* testis cDNA library. We used the Biomolecular Research Facility of The John Curtin School of Medical Research, ANU, to sequence the PCR products. The tammar wallaby BAC was sequenced at the Australian Genome Research Facility, Brisbane.

### Appendix A. Supplementary information

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.gene.2004.11.049](https://doi.org/10.1016/j.gene.2004.11.049).

### References

- Bellefroid, E.J., et al., 1996. X-MyT1, a *Xenopus* C2HC-type zinc finger protein with a regulatory function in neuronal differentiation. *Cell* 87, 1191–1202.
- Blanchette, M., Tompa, M., 2002. Discovery of regulatory elements by a computational method for phylogenetic footprinting. *Genome Res.* 12, 739–748.
- Fischer, M., et al., 1996. Prion protein (PrP) with amino-proximal deletions restoring susceptibility of PrP knockout mice to scrapie. *EMBO J.* 15, 1255–1264.
- Graves, J.A.M., Westerman, M., 2002. Marsupial genetics and genomics. *Trends Genet.* 18, 517–521.
- Groth, R.D., Mermelstein, P.G., 2003. Brain-derived neurotrophic factor activation of NFAT (nuclear factor of activated T-cells)-dependent transcription: a role for the transcription factor NFATc4 in neurotrophin-mediated gene expression. *J. Neurosci.* 23, 8125–8134.
- Hills, D., Comincini, S., Schlaepfer, J., Dolf, G., Ferretti, L., Williams, J.L., 2001. Complete genomic sequence of the bovine prion gene (PRNP) and polymorphism in its promoter region. *Anim. Genet.* 32, 231–232.
- Kumar, S., Tamura, K., Jakobsen, I.B., Nei, M., 2001. MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* 17, 1244–1245.
- Lee, I.Y., et al., 1998. Complete genomic sequence and analysis of the prion protein gene region from three mammalian species. *Genome Res.* 8, 1022–1037.
- Legname, G., Baskakov, I.V., Nguyen, H.O., Reisner, D., Cohen, F.E., DeArmond, S.J., Prusiner, S.B., 2004. Synthetic mammalian prions. *Science* 305, 673–676.
- Li, G., Bolton, D.C., 1997. A novel hamster prion protein mRNA contains an extra exon: increased expression in scrapie. *Brain Res.* 751, 265–274.
- Mayor, C., et al., 2000. VISTA: visualizing global DNA sequence alignments of arbitrary length. *Bioinformatics* 16, 1046–1047.
- Mendez, R., Richter, J.D., 2001. Translational control by CPEB: a means to the end. *Nat. Rev., Mol. Cell Biol.* 2, 521–529.
- Mody, M., et al., 2001. Genome-wide gene expression profiles of the developing mouse hippocampus. *Proc. Natl. Acad. Sci. U. S. A.* 98, 8862–8867.
- Okamoto, S., Sherman, K., Bai, G., Lipton, S.A., 2002. Effect of the ubiquitous transcription factors, SP1 and MAZ, on NMDA receptor subunit type 1 (NR1) expression during neuronal differentiation. *Brain Res. Mol. Brain Res.* 107, 89–96.
- Premzl, M., Sangiorgio, L., Strumbo, B., Marshall Graves, J.A.M., Simonic, T., Gready, J.E., 2003. Shadoo, a new protein highly conserved from fish to mammals and with similarity to prion protein. *Gene* 314, 89–102.
- Premzl, M., Gready, J.E., Jermini, L.S., Simonic, T., Marshall Graves, J.A.M., 2004. Evolution of vertebrate genes related to Prion and Shadoo proteins—clues from comparative genomic analysis. *Mol. Biol. Evol.* 21, 2210–2231.
- Prusiner, S.B., 1998. Prions. *Proc. Natl. Acad. Sci. U. S. A.* 95, 13363–13383.
- Rao, A., Luo, C., Hogan, P.G., 1997. Transcription factors of the NFAT family: regulation and function. *Annu. Rev. Immunol.* 15, 707–747.
- Saeki, K., Matsumoto, Y., Matsumoto, Y., Onodera, T., 1996. Identification of a promoter region in the rat prion protein gene. *Biochem. Biophys. Res. Commun.* 219, 47–52.
- Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Schneider, B., Mutel, V., Pietri, M., Ermonval, M., Mouillet-Richard, S., Kellermann, O., 2003. NADPH oxidase and extracellular regulated kinases 1/2 are targets of prion protein signaling in neuronal and nonneuronal cells. *Proc. Natl. Acad. Sci. U. S. A.* 100, 13326–13331.
- Si, K., Giustetto, M., Etkin, A., Hsu, R., Janisiewicz, A.M., Miniaci, M.C., Kim, J.H., Zhu, H., Kandel, E.R., 2003. A neuronal isoform of CPEB regulates local protein synthesis and stabilizes synapse-specific long-term facilitation in *Aplysia*. *Cell* 115, 893–904.
- Simonic, T., et al., 1997. cDNA cloning and expression of the flavoprotein D-aspartate oxidase from bovine kidney cortex. *Biochem. J.* 322, 729–735.
- Sytina, E.V., Pankratova, E.V., 2003. Transcription factor Oct-1: plasticity and multiplicity of functions. *Mol. Biol.* 37, 637–648.
- van de Lagemaat, L.N., Landry, J.R., Mager, D.L., Medstrand, P., 2003. Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *Trends Genet.* 19, 530–536.
- van Rheede, T., Smolenaars, M.M., Madsen, O., de Jong, W.W., 2003. Molecular evolution of the mammalian prion protein. *Mol. Biol. Evol.* 20, 111–121.
- Wakefield, M.J., Graves, J.A., 2003. The kangaroo genome. Leaps and bounds in comparative genomics. *EMBO Rep.* 4, 143–147.
- Windl, O., Dempster, M., Estibeiro, P., Lathe, R., 1995. A candidate marsupial PrP gene reveals two domains conserved in mammalian PrP proteins. *Gene* 159, 181–186.
- Wopfner, F., et al., 1999. Analysis of 27 mammalian and 9 avian PrPs reveals high conservation of flexible regions of the prion protein. *J. Mol. Biol.* 289, 1163–1178.
- Zelko, I.N., Folz, R.J., 2003. Myeloid zinc finger (MZF)-like, Kruppel-like and Ets families of transcription factors determine the cell-specific expression of mouse extracellular superoxide dismutase. *Biochem. J.* 369, 375–386.