Original Articles

Cloning and Expression of *PARP-3 (Adprt3)* and *U3-55k*, Two Genes Closely Linked on Mouse Chromosome 9

(poly(ADP-ribose) polymerase / PARP / *Adprt* / *ADPRTL* / PARP-1 / PARP-2 / hU3-55k / Rrp9p / U3 snoRNA / U3 snoRNP / WD-repeat protein / bi-directional promoter / tissue-specific gene expression / orthologue)

P. URBÁNEK, J. PAČES, J. KRÁLOVÁ, M. DVOŘÁK, V. PAČES

Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Prague, Czech Republic

Abstract. Post-translational modification of nuclear proteins by poly(ADP-ribose) polymerase 1 (PARP-1) is involved in the regulation of DNA repair, cell death, and maintenance of genomic stability. Recently, several PARP-1 homologues have been identified constituting a family of poly(ADP-ribosyl)ating proteins. We cloned and sequenced the cDNAs of the mouse *PARP-3* (*Adprt3*) gene encoding poly(ADP-ribose) polymerase 3 and of the closely linked *U3-55k* gene coding for the U3 small nucleolar ribonucleoprotein complex-associated 55-kilodalton protein. The two genes are located in a head-to-head orientation on mouse chromosome 9 and are linked by an approximately 1.5-kb putative bi-directional promoter region. This gene arrangement is conserved between mouse and human orthologues. Three alternative non-coding 5'-end exons were found in the mouse PARP-3 mRNA. The expression patterns of *PARP-3*, *U3-55k*, *PARP-2*, and *PARP-1* genes were determined using Northern blot with mRNA from various adult mouse tissues and organs. *PARP-3* expression was found to be regulated in a tissue-specific manner. The highest expression of *PARP-3* was detected in the skeletal muscle, high to moderate levels were found in the lung, liver, kidney, ovary, spleen and heart, while thymus, small intestine and colon contained lower levels of the *PARP-3* transcripts. Notably, *PARP-3* expression was barely detectable in the whole brain and testis mRNA. In contrast to *PARP-3*, the other three genes showed ubiquitous expression with less variable mRNA levels. Interestingly, the mouse and human *PARP-2* gene has recently been shown to be connected via a bi-directional promoter with the gene for the RNase P RNA subunit (Amé et al., *J. Biol. Chem.* 276: 11092-11099, 2001). As both the U3-55k protein and the RNase P RNA are involved in the processing of precursor RNAs of the protein-synthesizing machinery (pre-rRNA and pre-tRNA, respectively), it is tempting to hypothesize that expression of some members of the two groups of genes (i.e. *PARP* vs. protein-synthesizing machinery RNA-processing genes) may be coordinately regulated under certain physiological or pathological conditions and/or in some cell types.

Mammalian poly(ADP-ribose) polymerase 1 (PARP-1) is an abundant nuclear enzyme consisting of three main domains: the N-terminal DNA-binding domain, the C-terminal PARP catalytic domain and the automodification domain, located in the central part of the protein. PARP-1 binds to DNA-strand breaks in response to DNA damage induced by ionizing radiation or alkylating agents. Binding to DNA activates its catalytic domain, leading to the synthesis of branched ADP-ribose homopolymers on various nuclear proteins, including the automodification domain of the PARP-1 molecule itself (for reviews see de Murcia and Ménissier-de Murcia, 1994; Lindahl et al., 1995; D'Amours et al., 1999; Herceg and Wang, 2001). Numerous *in vitro* and *in vivo* studies, including analysis of *PARP-1* knock-out mice and cells, have documented functions of PARP-1 in DNA repair, replication, transcription, cell proliferation and death, as well as in maintaining the genomic integrity and suppressing tumorigenesis (for reviews see de Murcia and Ménissier-de Murcia, 1994; Lindahl et al., 1995; Jeggo, 1998; Le Rhun et al., 1998; D'Amours et al., 1999; Jacobson and Jacobson, 1999; Shall and de Murcia,

Received June 28, 2002. Accepted July 17, 2002.

This work was supported by the Grant Agency of the Academy of Sciences of the Czech Republic, grant No. A5052802, and in part by the Center for Integrated Genomics and by the Grant Agency of the Czech Republic, grant No. 204/00/0554.

Corresponding author: Pavel Urbánek, Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 166 37 Prague, Czech Republic. Tel.: +420-2-20183404; fax: +420-2-24310955; e-mail: urbanek@img.cas.cz.

Abbreviations: bp – base pair(s), EST – expressed sequence tag, kb – kilobase(s), nt – nucleotide(s), 3'UTR – 3'-end untranslated region, gb:/emb:/dbj: – accession number in the GenBank/EMBL/DDBJ sequence databases, respectively, ref: – accession number in the RefSeq database of the National Center for Biotechnology Information (NCBI).

2000; Bürkle, 2001; Herceg and Wang, 2001; Smith, 2001; Tong et al., 2001).

Although *PARP-1*$^{-/-}$ mice showed genomic instability and were highly sensitive to genotoxic stress, they were viable and fertile (Wang et al., 1995; Ménissier-de Murcia et al., 1997; Wang et al., 1997; Masutani et al., 1999). This relatively mild mutant phenotype, as well as the detection of ADP-ribose polymer synthesis in the PARP-1$^{-/-}$ cells (Shieh et al., 1998) and the discovery of two PARP genes in higher plants encoding two structurally different PARP homologues (Lepiniec et al., 1995; Babiychuk et al., 1998; Mahajan and Zuo, 1998), suggested the presence of additional *PARP* genes in mouse that might functionally substitute for the lack of PARP-1 activity. Over the past three years several mammalian genes encoding proteins homologous to the PARP-1 catalytic domain were identified, constituting a family of poly(ADP-ribosyl)ating proteins (reviewed in Smith, 2001; see also Kaminker et al., 2001; Kuimov et al., 2001; Lyons et al., 2001; Ma et al., 2001; Monz et al., 2001). The least studied member of the *PARP* family is *PARP-3* (*ADPRTL3*), characterized so far only in humans (Johansson, 1999).

Here we report the cDNA sequence of the mouse *PARP-3* (*Adprt3*) gene and of the closely linked *U3-55k* gene that we identify as the mouse orthologue of the human gene encoding the hU3-55k protein (Pluk et al., 1998). hU3-55k is a WD-repeat-containing 55-kilodalton protein associated with the U3 small nucleolar ribonucleoprotein complex (U3 snoRNP) that is involved in the processing of the primary pre-ribosomal RNA transcript (Lübben et al., 1993; Pluk et al., 1998; for review on snoRNPs see Maxwell and Fournier, 1995). Its orthologues in *Saccharomyces cerevisiae* (designated Rrp9p) and *Xenopus laevis* were characterized recently (Lukowiak et al., 2000; Venema et al., 2000). The main structural feature common to these three proteins is the six (man, frog) or five (yeast) WD-repeat sequence motifs (Pluk et al., 1998; Lukowiak et al., 2000; Venema et al., 2000; for review on WD-repeat proteins see Smith et al., 1999), which are probably required for association with the U3 snoRNA (Lukowiak et al., 2000).

As the mouse *PARP-3* and *U3-55k* genes appear to be linked via an approximately 1.5-kb bi-directional promoter region, it was of interest to determine their expression patterns in various mouse tissues and organs. Here we show that *PARP-3* gene expression is tissue-specifically regulated, in contrast to the neighbouring *U3-55k* gene and also to the homologous *PARP-1* and *PARP-2* genes.

Material and Methods

Oligonucleotide primers

The sequences are written in the 5' to 3' direction, **(s)** indicates sense and **(as)** anti-sense relative to the corresponding cDNA (see below): **1 (as)**, cacccgagtgtggcatgat; **2 (as)**, agcccactggtgaggatggc; **3 (as)**, ccacattggtgccgtgccacag; **4 (s)**, ctgtggcacggcaccaatgtgg; **5 (as)**, gggcaagcagcttagaggtgaat; **6 (s)**, ctcatatacaaggagagccagtg; **7 (as)**, ggtgtcattaggtgatctaagct; **8 (s)**, caatgttctggttagggatggac; **9 (as)**, gcaccattcacaacaaccaactc; **10 (s)**, gggtagcggcaccttgaacag; **11 (as)**, aagggagagtgtcaggagccg; **12 (s)**, atgatggctctgtggccttgtg; **13 (as)**, ggaggctttaatacaaaagaggac; **14 (s)**, gaagggtcgctagatgggatc; **15 (s)**, ccagaactcaggactgttagc; **16 (s)**, tttctgctgcctcggggaacac; **17 (as)**, cctcggcagtggaccggaag; **18 (as)**, gagccctcagtctgcacagag; **19 (as)**, gctggctttgcccgcttacga; **20 (s)**, tctggaaggcgagtgctaaatg; **21 (as)**, gttctaggccttggcgctctg; **22 (s)**, ggaaaccgacacgttagcggag; **23 (as)**, ccagtacagtaataagcgtcgct.

Total RNA and poly(A)$^+$ RNA isolation

Total RNA was isolated from mouse tissues and organs using the TRIzol Reagent (Life Technologies, Inc. – Gibco BRL, Gaithersburg, MD). Whole organs of three months old mice (kidney, brain, heart, spleen, thymus, lung, ovary, testis) or their parts (liver, skeletal muscle, small intestine, colon) were used. The skeletal muscles were from hindlegs. Ovaries were dissected free of oviduct, fat pad and bursa. Testes were taken free of fat pad, epididymis and vas deferens. Poly(A)$^+$ RNA was isolated using the PolyATtract mRNA Isolation System III (Promega Corporation, Madison, WI).

cDNA and genomic DNA cloning and sequencing

Mouse PARP-3 cDNA was isolated by a combination of the rapid amplification of 5' cDNA-ends (5'-RACE) (Frohman et al., 1988) and reverse transcription – polymerase chain reaction (RT-PCR). For the 5'-RACE, three oligonucleotides (Nos. 1, 2 and 3) were designed according to the mouse expressed sequence tag (EST) (accession number gb:AI019500; Marra et al., 1999) highly similar to the human PARP-3 cDNA sequence (gb:AF083068) (Johansson, 1999). Oligonucleotide 1 was used to prime reverse transcription of mouse skeletal muscle and heart total RNA with the SuperScript II RNase H$^-$ Reverse Transcriptase (Life Technologies, Inc. – Gibco BRL, Gaithersburg, MD). Oligonucleotides 2 and 3 were used for the first and the second round of amplification, respectively. The first round consisted of initial annealing at 50°C for 2 min and extension at 72°C for 20 min, followed by 25 cycles: 40 s at 94°C, 1 min at 60°C, and 3 min at 72°C. The second round consisted of 25 cycles: 30 s at 95°C, 15 s at 66°C, and 3 min at 68°C. Both PCR amplifications were performed with the Advantage 2 Polymerase Mix (CLONTECH Laboratories, Inc., Palo Alto, CA). PCR products were fractionated by agarose gel electrophoresis, cloned with the pGEM-T Vector System (Promega Corporation, Madison, WI), and sequenced. Two cDNA forms with different 5'-end sequences were isolated initially, suggesting the existence of two alternative 5'-end exons (designated 1a and 1c) in the mouse PARP-3 mRNA. This result was confirmed by the

alignment of the cDNAs with the genomic sequence of the promoter region (Fig. 1). A third alternative 5'-end exon (designated 1b) was inferred from the alignment of the mouse EST sequence (gb:AW318917; Marra et al., 1999) and the genomic sequence (Fig. 1). The existence of the three alternative exons was further verified by RT-PCR, cloning and sequencing of the three cDNA fragments spanning over the exon 1 – exon 2 junction. The first strand cDNA for RT-PCR was prepared from skeletal muscle total RNA using the cDNA Cycle Kit (Invitrogen Corporation, Carlsbad, CA) and oligo(dT) primers. Fifty ng of the cDNA were then used as a template for PCR amplification with Taq DNA Polymerase (Sigma Chemical Company, St. Louis, MO) and primer pairs 14/17, 15/17 or 16/17 (35 cycles: 30 s at 95°C, 30 s at 60°C, and 2 min at 72°C). Primers 14, 15 and 16 are specific for the alternative exons 1a, 1b and 1c, respectively, while primer 17 is located in exon 2 common to all the three splice variants (Fig. 1).

RT-PCR was performed under the same conditions in order to clone the 3'-terminal part of the PARP-3 cDNA and the full-length U3-55k cDNA. Oligonucleotides 4 and 5 that were derived from mouse ESTs (gb:AI019500 and gb:AI183257; Marra et al., 1999) homologous to the human PARP-3 cDNA sequence were used to amplify the 3'-terminal part of the PARP-3 cDNA coding region. The 3'-end untranslated region (3'UTR) of mouse PARP-3 was amplified with the oligonucleotides 6, 7, 8 and 9 designed according to the mouse ESTs located in the 3'-end part of the PARP-3 transcript (gb:AI183257, gb:AA197532, gb:AA177323; Marra et al., 1999). Similarly, the coding region of the U3-55k cDNA was amplified with the oligonucleotides 10 and 11, and the 3'UTR of this gene with the oligonucleotides 12 and 13. These U3-55k-specific oligonucleotides were designed according to the mouse EST sequences (gb:AA409391, gb:AI447235; Ko et al., 1998; Marra et al., 1999) found by homology to the human U3-55k cDNA (emb:AJ001340) (Pluk et al., 1998). The 1454-bp genomic fragment spanning the promoter region of the mouse *PARP-3* and *U3-55k* genes was PCR-amplified with the oligonucleotides 18 and 19, derived from the second *PARP-3* exon and from the first *U3-55k* exon, respectively. All the PCR-amplified cDNA and genomic DNA fragments were fractionated by agarose gel electrophoresis, cloned and sequenced. All the nucleotide sequences originally determined in the cloned fragments were verified by direct sequencing of PCR fragments independently amplified from the cDNA (hindleg skeletal muscle) or genomic DNA (spleen) of the C57BL/6J inbred mice.

Northern blot analysis

Poly(A)$^+$ RNAs isolated from various tissues of C57BL/6J x 129/Sv F1 hybrid mice were fractionated by electrophoresis in 1% agarose-formaldehyde gel (3 μg of RNA per lane) and blotted onto the GeneScreen Hybridization Transfer Membrane (NEN Life Science Products, Inc., Boston, MA) according to the standard protocols (Sambrook et al., 1989). For each tissue, a pool of equal aliquots of RNA isolated independently from three animals of the same age and sex was used. Northern blot was hybridized overnight with the ^{32}P-labelled mouse PARP-3 cDNA probe in the ULTRAhyb Ultrasensitive Hybridization Buffer (Ambion, Inc., Austin, TX) and washed under high stringency conditions according to the manufacturer's protocol. Autoradiography was performed at –70°C using Kodak BioMax MS film and intensifying screen (Eastman Kodak Company, Rochester, NY). The membrane was then successively rehybridized with the mouse cDNA probes of U3-55k, PARP-2, and PARP-1. After each round of hybridization the probe was stripped from the membrane by boiling for 3–5 min in the solution containing 15 mM sodium chloride, 1.5 mM sodium citrate, pH 7.0 (i.e. 0.1x SSC), and 1% sodium dodecyl sulphate (SDS), and the completeness of the probe removal was checked by autoradiography.

Probes used for Northern blot hybridization

The following cloned mouse cDNA fragments were used: PARP-3 (nt 335–1316, i.e. starting from the underlined *Sma*I half-site in the sequence 5'-gggatacaggtccac-3' to the oligonucleotide 3 inclusive, gb:AF368233 – this work); U3-55k (nt 577–1461, i.e. starting at the underlined *Xho*I half-site in the sequence 5'-tcgagccaagaaggg-3' to the oligonucleotide 11 inclusive, gb:AF368232 – this work); PARP-2 (nt 43–1009, i.e. oligonucleotide 20 through oligonucleotide 21, emb:AJ007780) (Amé et al., 1999); PARP-1 (nt 22–1022, i.e. oligonucleotide 22 through oligonucleotide 23, emb:X14206) (Huppi et al., 1989). PARP-2 and PARP-1 cDNA fragments were prepared by RT-PCR amplification using the indicated oligonucleotides, cloned and verified by sequencing as described above for PARP-3 and U3-55k cDNAs. For hybridization, the cDNA fragments were labelled by nick translation with [α-^{32}P]dCTP according to Sambrook et al. (1989).

Database homology searches and sequence comparisons

The BLAST programmes (Altschul et al., 1990; Altschul et al., 1997) at the National Center for Biotechnology Information (NCBI) (Wheeler et al., 2002) (http://www.ncbi.nlm.nih.gov/BLAST/) were used to search for the PARP and U3-55k homologous sequences in the NCBI databases (nr, est, human and mouse genomic). The sequences referred to in this work and accessed through the NCBI Entrez system (Wheeler et al., 2002) originated from the following databases: GenBank (Benson et al., 2002), EMBL (Stoesser et al., 2002), DDBJ (Tateno et al., 2002) and NCBI RefSeq database (Pruitt and Maglott, 2001) (see also the

Abbreviations section). The BLAST 2 Sequences programme (Tatusova and Madden, 1999) at the NCBI (http://www.ncbi.nlm.nih.gov/blast/bl2seq/bl2.html) was used for pairwise sequence alignments. For sequence similarity comparisons in Table 1, multiple sequence alignment was performed with the Clustal W programme (Thompson et al., 1994). The percentage of identical and similar amino acids in Table 1 was obtained using the statistics-report function of the GeneDoc programme (Nicholas et al., 1997). Pairs of similar amino acid residues were defined as those that yield a greater than zero score in the BLOSUM45 scoring matrix (similarity groups: DE, NH, ST, QKR, FYW, LIVM).

Nucleotide sequence accession numbers and the new gene symbol

Sequence data from this article have been deposited with the GenBank database (Benson et al., 2002) (http://www.ncbi.nlm.nih.gov/Genbank/GenbankOverview.html) under accession numbers: AF368232 (U3-55k cDNA), AF368233 (PARP-3 cDNA including alternative exon 1a), AF368234 (*U3-55k – PARP-3* promoter region), AY046316 (PARP-3 splice variant – exon 1b), and AY046317 (PARP-3 splice variant – exon 1c). The International Committee on Standardized Genetic Nomenclature for Mice (http://www.informatics.jax.org/mgihome/nomen/) approved the symbol *Adprt3* for the *PARP-3* gene described in this paper.

Results

cDNA sequences and promoter region of PARP-3 and U3-55k genes

We cloned and sequenced the cDNA of the mouse *PARP-3* (*Adprt3*) gene. The longest open reading frame encodes the protein of 528 amino acid residues and the calculated molecular weight of 59.5 kilodaltons (kDa). The predicted mouse protein sequence is 80% identical to the human PARP-3, the only orthologue characterized so far (Johansson, 1999). We determined sequence similarity between the homologous regions of the human and mouse PARP-1, PARP-2 and PARP-3 proteins (Table 1). These sequence comparisons showed that (i) PARP-1 and PARP-2 proteins are more closely related to each other (41%–42% identity) than to the PARP-3 proteins (29%–33% identity), and (ii) both PARP-3 proteins show a slightly higher similarity to the PARP-2 proteins (33% identity) than to the PARP-1 proteins (29%–31% identity) (Table 1). A different degree of sequence conservation is observed between the pairs of orthologous proteins (human vs. mouse). Thus, human and mouse PARP-1 proteins are 94% identical, PARP-2 proteins from the two organisms show 89% identity, whereas PARP-3 proteins are the least conserved orthologues with 81% identity in the regions compared (see Table 1).

A BLAST homology search of the human sequences in the databases at the NCBI (nr, human genome), using the human PARP-3 cDNA sequence (Johansson, 1999) as query, revealed the human *PARP-3* (*ADPRTL3*) genomic locus (LocusLink LocusID: 10039; Pruitt and Maglott, 2001) within an 81.5-kb PAC clone from the 3p21.1-9 chromosomal region (gb:AC006255, *Homo sapiens* 3p21.1-9 PAC RPCI5-1087L12) and within the *Homo sapiens* chromosome 3 working draft sequence segment (ref:NT_005986) (Lander et al., 2001; Pruitt and Maglott, 2001). Interestingly, the gene encoding human U3 small nucleolar ribonucleoprotein complex-associated 55-kilodalton protein (hU3-55k) (Pluk et al., 1998) is located close to *PARP-3* in a head-to-head orientation (gb:AC006255, ref:NT_005986, LocusLink LocusID: 9136) (Pruitt and Maglott, 2001).

Table 1. Comparison of sequence similarity between the human and mouse PARP-1, PARP-2 and PARP-3 proteins

	HsPARP-1		**MmPARP-1**		**HsPARP-2**		**MmPARP-2**		**HsPARP-3**		**MmPARP-3**	
	%		%		%		%		%		%	
	i	i+s	i	i+s	i	i+s	i	i+s	i	i+s	i	i+s
HsPARP-1	100	(100)	94	(98)	42	(62)	41	(61)	30	(48)	31	(48)
MmPARP-1	94	(98)	100	(100)	41	(61)	41	(60)	29	(48)	29	(48)
HsPARP-2	42	(62)	41	(61)	100	(100)	89	(95)	33	(50)	33	(50)
MmPARP-2	41	(61)	41	(60)	89	(95)	100	(100)	33	(51)	33	(50)
HsPARP-3	30	(48)	29	(48)	33	(50)	33	(51)	100	(100)	81	(89)
MmPARP-3	31	(48)	29	(48)	33	(50)	33	(50)	81	(89)	100	(100)

Note. For each pair of the six PARP protein sequences two numbers characterizing their similarity are shown: i, percentage of identical amino acids; i+s, percentage of identical and similar amino acids (smaller numbers in parentheses). The numbers are based on the multiple sequence alignment - see Material and Methods for the details. The species names are abbreviated as follows: Hs, *Homo sapiens*; Mm, *Mus musculus*. The database accession numbers, amino acid sequence intervals used for the multiple sequence alignment and similarity determination (aa:), and references are as follows: **HsPARP-1** [gb:M32721, aa: 520-1014, (Cherney et al., 1987)]; **MmPARP-1** [emb:X14206, aa: 519-1013, (Huppi et al., 1989)]; **HsPARP-2** [emb:AJ236912, aa: 68-570, (Amé et al., 1999)]; **MmPARP-2** [emb:AJ007780, aa: 61-559, (Amé et al., 1999)]; **HsPARP-3** [gb:AF083068, aa: 35-533, (Johansson, 1999)]; **MmPARP-3** [gb:AF368233 - this work, aa: 33-528].

Consequently, the two genes are transcribed in opposite directions from an approximately 1.5-kb putative bi-directional promoter region. To investigate the organization of the two genes in the mouse genome we cloned and sequenced the cDNA encoding the mouse orthologue of the human U3-55k protein, and the mouse genomic fragment spanning the promoter region and divergently oriented 5'-end exons of the mouse *PARP-3* and *U3-55k* genes (Fig. 1). Both mouse and human U3-55k proteins consist of 475 amino acid residues, are highly conserved throughout their sequence lengths (92% identity), and have similar calculated molecular weights of 52.1 kDa and 51.8 kDa, respectively. Three kinds of protein sequence motifs were identified by Pluk et al. (1998) in the human U3-55k sequence: (i) the N-terminal putative bi-partite nuclear localization signal, (ii) the glutamic acid-rich region near the N terminus, and most significantly, (iii) five WD repeats spanning the major central part of the protein. An additional WD repeat was detected recently by two groups based on further sequence analysis and its comparison with the *Saccharomyces cerevisiae* and *Xenopus laevis* proteins (Lukowiak et al., 2000; Venema et al., 2000). All the sequence motifs are conserved in the mouse U3-55k protein at exactly the same positions as they are in the human sequence. The putative bi-partite nuclear localization signal is located between the residues 8–40, and the glutamic acid-rich stretch between the residues 64–73. The WD-repeat number and arrangement in the mouse U3-55k sequence was analysed using the Protein Sequence Analysis server at the BioMolecular Engineering Research Center, Boston University (http://bmerc-www.bu.edu/psa/) (Stultz et al., 1993; White et al., 1994; Smith et al., 1999; Yu et al., 2000). Six WD repeats were predicted in the mouse U3-55k with the following positions of the first and the last amino acid residue of the conserved WD-repeat core sequence (the so-called GH-WD-core; Smith et al., 1999): 144–174, 197–227, 239–269, 281–311, 322–351, and 374–404.

Alignment of the human PARP-3 transcript (Johansson, 1999) and GenBank EST sequences from the 5' end of the gene with the human genomic sequence of the promoter region revealed two exons located upstream of the ATG initiation codon of *PARP-3* (Fig. 1). In contrast, using a combination of 5'-RACE and RT-PCR analyses we identified three alternative 5'-end exons (designated 1a, 1b and 1c) in the corresponding genomic region of the mouse (Fig. 1). These results are in agreement with mouse EST sequences from the GenBank database (e.g. ESTs dbj:BB591382, gb:BG922859, gb:AW318917, gb:BG964545; Marra et al., 1999; Strausberg et al., 1999; Kawai et al., 2001). We note that each of the three mouse alternative exons contains an in-frame stop codon upstream of the putative ATG initiation codon, indicating that these exons are non-coding (Fig. 1).

Expression of PARP-3, U3-55k, PARP-2 and PARP-1 genes in mouse tissues

The expression patterns of *PARP-3*, *U3-55k*, *PARP-2*, and *PARP-1* genes were determined using a Northern blot with mRNA from various adult mouse tissues and organs (Fig. 2). The highest expression of *PARP-3* was detected in the skeletal muscle mRNA, as estimated from shorter autoradiogram exposures (Fig. 2 and data not shown). High to moderate levels were found in the lung, liver, kidney, ovary, spleen and heart, while thymus, small intestine and colon contained only low levels of the *PARP-3* transcripts. Notably, *PARP-3* expression was barely detectable in mRNA prepared from the brain and testis. In some tissues, e.g. in spleen, lung and small intestine, two PARP-3 mRNA species of approximately 2.7 kb and 2.4 kb were detected. Using the 3'-RACE cloning of the mouse PARP-3 cDNA (data not shown), as well as the alignment of the 3'-end untranslated region of the mouse PARP-3 cDNA (nt 1754 to 2633 in gb:AF368233) with the mouse EST sequences (e.g. ESTs gb:AA267081, dbj:BB218739, dbj:BB223095; Marra et al., 1999; Kawai et al., 2001), we localized an alternative polyadenylation site (nt 2358 in gb:AF368233) of the mouse PARP-3 transcript. The shorter mRNA species may represent transcripts terminated at this site. The estimated lengths of the two mRNA species are in good agreement with the cloned cDNA sequence (2633 bp) and its alternatively terminated variant (2358 bp).

Compared to *PARP-3*, the other three genes are expressed ubiquitously with less variability of mRNA levels among the individual tissues (Fig. 2). Particularly, U3-55k and PARP-1 transcripts are distributed in a housekeeping gene-like manner. U3-55k mRNA migrates as a single band of approximately 1.6 kb, corresponding well to the size of the cloned cDNA (1513 bp). In addition to the major mRNA species of 2.0 kb, the PARP-2 probe detected a second less prominent 2.5-kb mRNA, in agreement with the previously published expression data (Amé et al., 1999; Berghammer et al., 1999). The estimated length of the PARP-1 transcript of approximately 3.8 kb corresponds well to the previously reported size (Huppi et al., 1989; Ogura et al., 1990; Wang et al., 1995). Also, higher levels of PARP-1 mRNA in the spleen, testis and thymus (Fig. 2) are in agreement with the data of Ogura et al. (1990), Wang et al. (1995) and Schreiber et al. (2002). Overall, *PARP-3* exhibits much more variability of mRNA levels and some degree of tissue-specific expression, compared to *U3-55k*, *PARP-2*, and *PARP-1* genes.

Discussion

Mouse *PARP-3* and *U3-55k* genes are located close to each other in the head-to-head orientation and are linked by an approximately 1.5-kb putative bi-directional promoter region. Comparison with the human

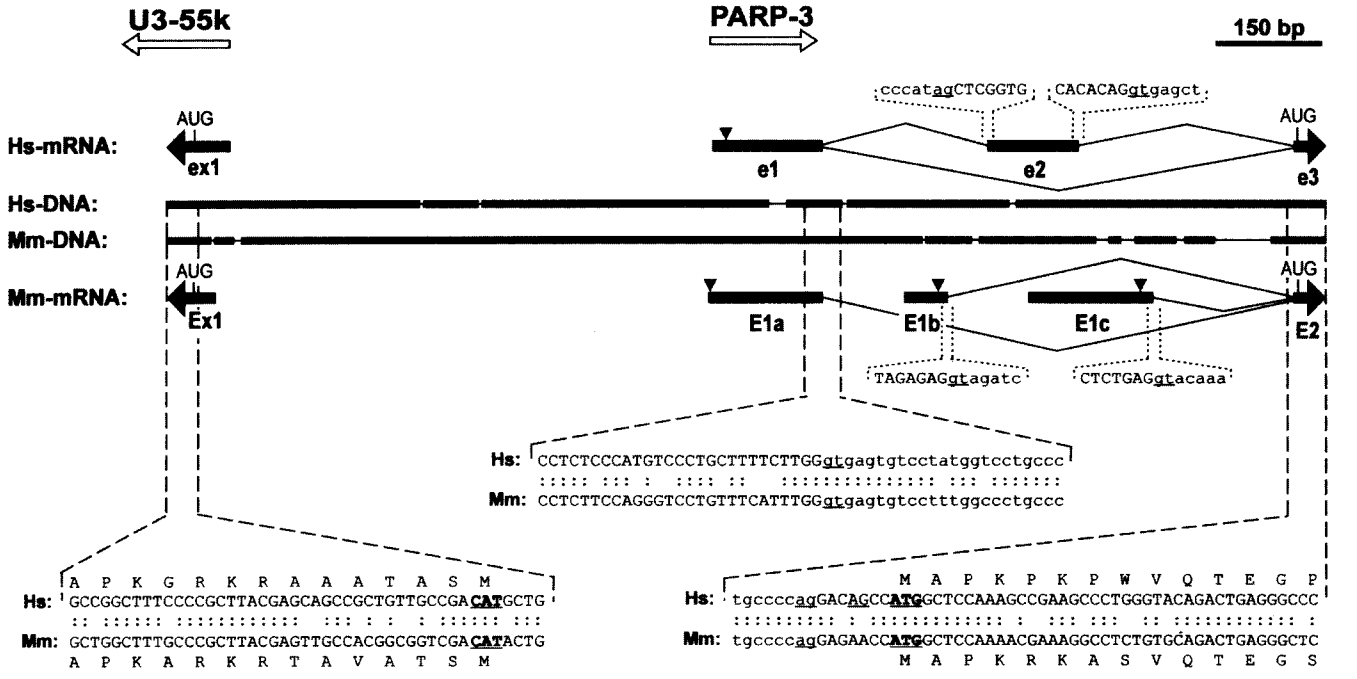

Fig. 1. Schematic comparison of the human and mouse promoter regions of *PARP-3* and *U3-55k* genes. The two parallel thick horizontal lines represent the alignment of the human (Hs-DNA; nt 52161 to 50592 [complementary strand] in gb:AC006255) and mouse (Mm-DNA; nt 1 to 1454 in gb:AF368234 – this work) genomic DNA regions encompassing the 5' ends of *PARP-3* and *U3-55k* genes. (The short interruptions of these lines filled with the thin line represent small gaps introduced into the genomic sequences in order to optimize the alignment.) The open arrows on the very top of the figure indicate the direction of transcription of the two genes. The 5'-end exons of the human and mouse U3-55k and PARP-3 transcripts are depicted as black boxes and black arrows (5'-end portions of AUG translation initiation codon-containing exons). (For simplicity, the alignment gaps are not indicated in the exon boxes.) Exon symbols: ex1, human U3-55k; Ex1, mouse U3-55k; e1, e2 and e3, human PARP-3; E1a, E1b, E1c and E2, mouse PARP-3. The human U3-55k exon ex1 is according to the reported cDNA sequence (emb:AJ001340; Pluk et al., 1998) and EST sequences (e.g. EST gb:BG825193; Strausberg et al., 1999), the mouse U3-55k exon Ex1 is according to gb:AF368232 (this work). Angled lines indicate the alternative splicing patterns of the human and mouse PARP-3 exons. The human PARP-3 splice variant e1-e2-e3 is according to EST sequences (gb:BG818516, gb:BG702085; Strausberg et al., 1999), while the splice variant e1-e3 is present in the published full-length cDNA sequence (gb:AF083068; Johansson, 1999) and several EST sequences (e.g. ESTs gb:BI554046, gb:BG913289; Strausberg et al., 1999). The mouse PARP-3 exons E1a, E1b, and E1c are alternatively spliced onto exon E2 (gb:AF368233, gb:AY046316, and gb:AY046317 – this work). We note that the splice donor site located downstream of exon E1b appears to be leaky, as a portion of exon E1b transcripts extend through this site into exon E1c and are spliced at the donor site downstream of E1c (data not shown). The vertical arrowheads above the PARP-3 exons point at the positions of the stop codons that are in frame with the downstream AUG initiation codons according to the published human and mouse cDNA sequences (Johansson, 1999, and this work, respectively). The nucleotide sequences of the exon-intron boundaries of human e2 and mouse E1b and E1c exons are shown above and below these exons, respectively, enclosed in dotted-line braces. The upper-case nucleotide symbols indicate the exon sequences, and the lower-case letters stand for the intron sequences. The conserved dinucleotides of the splice donor (gt) and splice acceptor (ag) sites are underlined. The lower part of the figure shows in a blown up detail three human (Hs) versus mouse (Mm) aligned sequence regions (enclosed in dashed line braces). Colons between the two sequences indicate identical nucleotides. The nucleotide sequences are written in the conventional 5'-to-3' left-to-right direction. *Bottom left.* Alignment of the 5' ends of the coding regions of U3-55k. Because of the leftward orientation of U3-55k transcription in this scheme, the two nucleotide sequences are given as mRNA-complementary strands and the amino acid sequences deduced from them are written, accordingly, from the right to the left. The complements of the translation initiation codons are in boldface and underlined. *Lower middle.* Conserved splice donor site of the human e1 and mouse E1a exons. The exon-intron boundary is marked as above. *Bottom right.* Alignment of the 5' ends of the human e3 and mouse E2 exons. The intron-exon boundary is marked as above. The ATG translation initiation codons are in boldface and underlined. The deduced human and mouse amino acid sequences are depicted above and below the DNA sequences, respectively. Note that there is an alternative splice acceptor site in the human exon e3, located four nucleotides upstream of the translation initiation ATG (underlined upper-case dinucleotide AG), which is used in some of the splice variant e1-e3 EST sequences (e.g. EST gb:BG913289; Strausberg et al., 1999). This alternative splice acceptor site does not appear to be conserved in the mouse sequence.

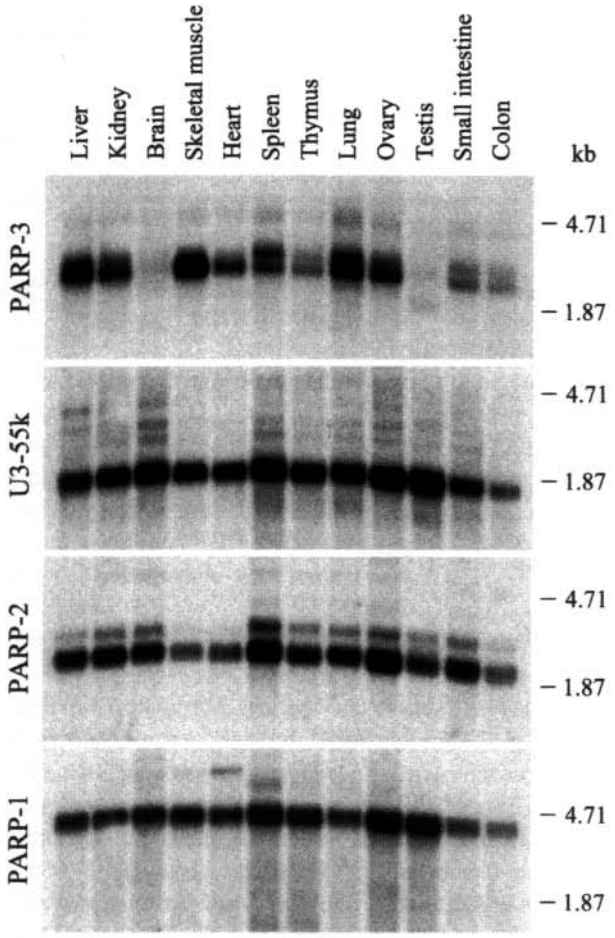

Fig. 2. Northern blot analysis of PARP-3, U3-55k, PARP-2 and PARP-1 mRNA in mouse tissues. The membrane was hybridized first with the PARP-3 probe, and then successively rehybridized with the U3-55k, PARP-2, and finally PARP-1 probes. The ubiquitously expressed U3-55k and PARP-1 mRNAs served as loading and transfer controls.

PARP-3/U3-55k genomic sequence showed that this gene organization is conserved between man and mouse (Fig. 1). Interestingly, the mouse and human *PARP-2* gene has recently been shown to be connected by a bidirectional promoter with the gene for the RNase P RNA subunit (Amé et al., 2001). Moreover, both the U3-55k protein and the RNase P RNA are involved in the processing of precursor RNAs of the protein-synthesizing machinery (pre-rRNA and pre-tRNA, respectively) (Lübben et al., 1993; Pluk et al., 1998; Lukowiak et al., 2000; Venema et al., 2000; for review on RNase P RNA see Gopalan et al., 2002). It is therefore tempting to hypothesize that there might be a functional reason for the above similarity in the promoter region organization between the *PARP* and the protein-synthesizing machinery RNA-processing genes. For example, the expression of the two groups of genes might be coordinately regulated under certain physiological or pathological conditions and/or in some cell types.

While this manuscript was in preparation, four cDNA sequences and one genomic sequence became available in the GenBank (Benson et al., 2002) and RefSeq (Pruitt and Maglott, 2001) (http://www.ncbi.nlm.nih.gov/LocusLink/refseq.html) databases of the National Center for Biotechnology Information (NCBI) under the following accession numbers and definitions: **gb:BC014703** [*Mus musculus*, similar to U3 snoRNP-associated 55-kDa protein, clone MGC:25949 IMAGE:4237895, mRNA, complete cds], **gb:BC014870** [*Mus musculus*, clone MGC:11997 IMAGE:3602116, mRNA, complete cds], **ref:XM_135106** [*Mus musculus* similar to U3 snoRNP-associated 55-kDa protein (LOC235588), mRNA], **ref:XM_135141** [*Mus musculus* similar to Poly [ADP-ribose] polymerase-3 (PARP-3) (NAD(+) ADP-ribosyltransferase-3) (Poly[ADP-ribose] synthetase-3) (pADPRT-3) (hPARP-3) (LOC235587), mRNA], **ref:NW_000356** [*Mus musculus* WGS supercontig Mm9_WIFeb01_200.]. The cDNA sequences gb:BC014870 and ref:XM_135141 correspond to our PARP-3 cDNA sequence (splice variant with the N-terminal exon 1c, i.e. gb:AY046317 – includes exon 1c, and gb:AF368233 – starting from nt 160; see also Fig. 1). The cDNA sequences gb:BC014703 and ref:XM_135106 correspond to our U3-55k cDNA sequence (gb:AF368232). The four cDNA sequences differ from the PARP-3 and U3-55k cDNAs presented in this work at several single-nucleotide positions (1 position in ref:XM_135141, 11 positions in gb:BC014870, 7 positions in gb:BC014703) and one three-nucleotide position in XM_135106. The differences may represent nucleotide polymorphisms of mouse strains. The genomic sequence ref:NW_000356 is a 7888514-bp NCBI RefSeq supercontig from mouse chromosome 9, which includes both the *PARP-3* (LocusLink LocusID: 235587) and *U3-55k* (LocusLink LocusID: 235588) genes (http://www.ncbi.nlm.nih.gov/LocusLink/) (Pruitt and Maglott, 2001) and the region corresponding to our promoter region sequence (gb:AF368234).

Mouse *PARP-3* and *U3-55k* genes are located on chromosome 9 (see above). Their chromosomal position is defined, for example, by the D9Wsu10e locus (DNA segment, Chr 9, Wayne State University 10, expressed) on chromosome 9 at 56.00 cM (Ko et al., 1998) (LocusLink LocusID: 27966, http://www.ncbi.nlm.nih.gov/LocusLink/; UniSTS: 142814, http://www.ncbi.nlm.nih.gov/genome/sts/) (Pruitt and Maglott, 2001). The D9Wsu10e marker, contained in EST gb:AA407505 (Ko et al., 1998) (LocusLink LocusID: 27966; UniSTS: 142814), is localized in the 3'-end region of the mouse U3-55k cDNA sequence (nt 1281 to 1458 in gb:AF368232 – this work). The above mouse chromosomal region is syntenic with the human chromosome 3p (Lyon and Kirby, 1996), where the human *ADPRTL3* (*PARP-3*) locus has previously been mapped to 3p21.1-p22.2 (Johansson, 1999). The syntenic chromosomal location,

together with a high degree of identity between mouse and human PARP-3 and U3-55k proteins and conserved gene arrangement in the two species (see above), identify the two genes as human and mouse orthologues.

There is less similarity between the human and mouse PARP-3 protein sequences than between the pairs of PARP-1 or PARP-2 orthologues (Table 1). Thus, in mammals, the PARP-3 proteins appear to evolve with a higher divergence rate than the PARP-1 or PARP-2 proteins. It would be interesting to see whether also the function(s) of PARP-3 tends to diverge more rapidly between man and mouse, compared to the cases of PARP-1 or PARP-2 proteins, where it could remain more conserved between the two species. This notion might be supported by the fact that the human *PARP-3* gene was shown to be ubiquitously expressed in a panel of human tissues, including brain and testis (Johansson, 1999).

Adult mouse *PARP-3* expression is regulated in a tissue-specific manner. In contrast to PARP-3, both the PARP-1 and PARP-2 mRNAs were expressed ubiquitously in the adult mouse tissues analysed, including brain and testis (Fig. 2, see also Ogura et al., 1990; Wang et al., 1995; Amé et al., 1999; Berghammer et al., 1999; Schreiber et al., 2002). We cannot, however, rule out the possibility that PARP-3 expression is restricted to a small region or few cells in these two organs.

We noted that some of the tissues (organs) where PARP-3 is not expressed or where its expression is low (i.e. brain, testis and thymus) are known to be protected by physiological barriers, entirely or in part, against the entry of some molecules from the blood (e.g. blood-brain or blood-thymus barrier) (Junqueira et al., 1995). Conversely, some other tissues that lack this barrier (e.g. liver, kidney, lung, muscle, spleen) show high levels of PARP-3 expression. In view of the role of PARP-1 in DNA repair and in the maintenance of genome integrity, we speculate that the level of the *PARP-3* gene product might in some tissues correlate with the amount of the physiological genotoxic stress the particular tissue is exposed to during organism ontogeny.

Acknowledgements

We thank Dr. Petr Pajer for valuable discussions, help and technical advice. We also thank Drs. Petr Bartůněk, Jiří Forejt and Jiří Hejnar for helpful suggestions.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389-3402.

Amé, J.-C., Rolli, V., Schreiber, V., Niedergang, C., Apiou, F., Decker, P., Muller, S., Höger, T., Ménissier-de Murcia, J., de Murcia, G. (1999) PARP-2, a novel mammalian DNA damage-dependent poly(ADP-ribose) polymerase. *J. Biol. Chem.* **274**, 17860-17868.

Amé, J.-C., Schreiber, V., Fraulob, V., Dollé, P., de Murcia, G., Niedergang, C. P. (2001) A bidirectional promoter connects the poly(ADP-ribose) polymerase 2 (PARP-2) gene to the gene for RNase P RNA. Structure and expression of the mouse PARP-2 gene. *J. Biol. Chem.* **276**, 11092-11099.

Babiychuk, E., Cottrill, P. B., Storozhenko, S., Fuangthong, M., Chen, Y., O'Farrell, M. K., Van Montagu, M., Inzé, D., Kushnir, S. (1998) Higher plants possess two structurally different poly(ADP-ribose) polymerases. *Plant J.* **15**, 635-645.

Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., Rapp, B. A., Wheeler, D. L. (2002) GenBank. *Nucleic Acids Res.* **30**, 17-20.

Berghammer, H., Ebner, M., Marksteiner, R., Auer, B. (1999) pADPRT-2: a novel mammalian polymerizing(ADP-ribosyl)transferase gene related to truncated pADPRT homologues in plants and *Caenorhabditis elegans*. *FEBS Lett.* **449**, 259-263.

Bürkle, A. (2001) Physiology and pathophysiology of poly(ADP-ribosyl)ation. *BioEssays* **23**, 795-806.

Cherney, B. W., McBride, O. W., Chen, D., Alkhatib, H., Bhatia, K., Hensley, P., Smulson, M. E. (1987) cDNA sequence, protein structure, and chromosomal location of the human gene for poly(ADP-ribose) polymerase. *Proc. Natl. Acad. Sci. USA* **84**, 8370-8374.

D'Amours, D., Desnoyers, S., D'Silva, I., Poirier, G. G. (1999) Poly(ADP-ribosyl)ation reactions in the regulation of nuclear functions. *Biochem. J.* **342**, 249-268.

de Murcia, G., Ménissier-de Murcia, J. (1994) Poly(ADP-ribose) polymerase: a molecular nick-sensor. *Trends Biochem. Sci.* **19**, 172-176.

Frohman, M. A., Dush, M. K., Martin, G. R. (1988) Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* **85**, 8998-9002.

Gopalan, V., Vioque, A., Altman, S. (2002) RNase P: variations and uses. *J. Biol. Chem.* **277**, 6759-6762.

Herceg, Z., Wang, Z.-Q. (2001) Functions of poly(ADP-ribose) polymerase (PARP) in DNA repair, genomic integrity and cell death. *Mutat. Res.* **477**, 97-110.

Huppi, K., Bhatia, K., Siwarski, D., Klinman, D., Cherney, B., Smulson, M. (1989) Sequence and organization of the mouse poly(ADP-ribose) polymerase gene. *Nucleic Acids Res.* **17**, 3387-3401.

Jacobson, M. K., Jacobson, E. L. (1999) Discovering new ADP-ribose polymer cycles: protecting the genome and more. *Trends Biochem. Sci.* **24**, 415-417.

Jeggo, P. A. (1998) DNA repair: PARP – another guardian angel? *Curr. Biol.* **8**, R49-R51.

Johansson, M. (1999) A human poly(ADP-ribose) polymerase gene family (*ADPRTL*): cDNA cloning of two novel poly(ADP-ribose) polymerase homologues. *Genomics* **57**, 442-445.

Junqueira, L. C., Carneiro, J., Kelley, R. O. (1995) *Basic Histology*. 8th edition, a LANGE medical book, Prentice-Hall International, Inc.

Kaminker, P. G., Kim, S.-H., Taylor, R. D., Zebarjadian, Y., Funk, W. D., Morin, G. B., Yaswen, P., Campisi, J. (2001) TANK2, a new TRF1-associated poly(ADP-ribose) polymerase, causes rapid induction of cell death upon overexpression. *J. Biol. Chem.* **276**, 35891-35899.

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., et al. – The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium (2001) Functional annotation of a full-length mouse cDNA collection. *Nature* **409**, 685-690.

Ko, M. S. H., Threat, T. A., Wang, X., Horton, J. H., Cui, Y., Wang, X., Pryor, E., Paris, J., Wells-Smith, J., Kitchen, J. R., Rowe, L. B., Eppig, J., Satoh, T., Brant, L., Fujiwara, H., Yotsumoto, S., Nakashima, H. (1998) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the *t*-complex and under-representation on the X chromosome. *Hum. Mol. Genet.* **7**, 1967-1978.

Kuimov, A. N., Kuprash, D. V., Petrov, V. N., Vdovichenko, K. K., Scanlan, M. J., Jongeneel, C. V., Lagarkova, M. A., Nedospasov, S. A. (2001) Cloning and characterization of TNKL, a member of tankyrase gene family. *Genes Immun.* **2**, 52-55.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. – International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921.

Lepiniec, L., Babiychuk, E., Kushnir, S., Van Montagu, M., Inzé, D. (1995) Characterization of an *Arabidopsis thaliana* cDNA homologue to animal poly(ADP-ribose) polymerase. *FEBS Lett.* **364**, 103-108.

Le Rhun, Y., Kirkland, J. B., Shah, G. M. (1998) Cellular responses to DNA damage in the absence of poly(ADP-ribose) polymerase. *Biochem. Biophys. Res. Commun.* **245**, 1-10.

Lindahl, T., Satoh, M. S., Poirier, G. G., Klungland, A. (1995) Post-translational modification of poly(ADP-ribose) polymerase induced by DNA strand breaks. *Trends Biochem. Sci.* **20**, 405-411.

Lübben, B., Marshallsay, C., Rottmann, N., Lührmann, R. (1993) Isolation of U3 snoRNP from CHO cells: a novel 55 kDa protein binds to the central part of U3 snoRNA. *Nucleic Acids Res.* **21**, 5377-5385.

Lukowiak, A. A., Granneman, S., Mattox, S. A., Speckmann, W. A., Jones, K., Pluk, H., van Venrooij, W. J., Terns, R. M., Terns, M. P. (2000) Interaction of the U3-55k protein with U3 snoRNA is mediated by the Box B/C motif of U3 and the WD repeats of U3-55k. *Nucleic Acids Res.* **28**, 3462-3471.

Lyon, M. F., Kirby, M. C. (1996) Mouse chromosome atlas. In: *Genetic Variants and Strains of the Laboratory Mouse*, Third Edition, eds. Lyon, M. F., Rastan, S., Brown, S. D. M., pp. 881-923, Oxford University Press, Oxford.

Lyons, R. J., Deane, R., Lynch, D. K., Ye, Z.–S. J., Sanderson, G. M., Eyre, H. J., Sutherland, G. R., Daly, R. J. (2001) Identification of a novel human tankyrase through its interaction with the adaptor protein Grb14. *J. Biol. Chem.* **276**, 17172-17180.

Ma, Q., Baldwin, K. T., Renzelli, A. J., McDaniel, A., Dong, L. (2001) TCDD-inducible poly(ADP-ribose) polymerase: a novel response to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin. *Biochem. Biophys. Res. Commun.* **289**, 499-506.

Mahajan, P. B., Zuo, Z. (1998) Purification and cDNA cloning of maize poly(ADP)-ribose polymerase. *Plant Physiol.* **118**, 895-905.

Marra, M., Hillier, L., Kucaba, T., Allen, M., Barstead, R., Beck, C., Blistain, A., Bonaldo, M., Bowers, Y., Bowles, L., et al. (1999) An encyclopedia of mouse genes. *Nat. Genet.* **21**, 191-194.

Masutani, M., Suzuki, H., Kamada, N., Watanabe, M., Ueda, O., Nozaki, T., Jishage, K.-I., Watanabe, T., Sugimoto, T., Nakagama, H., Ochiya, T., Sugimura, T. (1999) Poly(ADP-ribose) polymerase gene disruption conferred mice resistant to streptozotocin-induced diabetes. *Proc. Natl. Acad. Sci. USA* **96**, 2301-2304.

Maxwell, E. S., Fournier, M. J. (1995) The small nucleolar RNAs. *Annu. Rev. Biochem.* **64**, 897-934.

Ménissier-de Murcia, J., Niedergang, C., Trucco, C., Ricoul, M., Dutrillaux, B., Mark, M., Oliver, F. J., Masson, M., Dierich, A., LeMeur, M., Walztinger, C., Chambon, P., de Murcia, G. (1997) Requirement of poly(ADP-ribose) polymerase in recovery from DNA damage in mice and in cells. *Proc. Natl. Acad. Sci. USA* **94**, 7303-7307.

Monz, D., Munnia, A., Comtesse, N., Fischer, U., Steudel, W.–I., Feiden, W., Glass, B., Meese, E. U. (2001) Novel tankyrase-related gene detected with meningioma-specific sera. *Clin. Cancer Res.* **7**, 113-119.

Nicholas, K. B., Nicholas, H. B., Jr., Deerfield, D. W., II. (1997) GeneDoc: analysis and visualization of genetic variation. *EMBnet.news* **4**, (http://www.hgmp.mrc.ac.uk/embnet.news/vol4_2/genedoc.html).

Ogura, T., Takenouchi, N., Yamaguchi, M., Matsukage, A., Sugimura, T., Esumi H. (1990) Striking similarity of the distribution patterns of the poly(ADP-ribose) polymerase and DNA polymerase β among various mouse organs. *Biochem. Biophys. Res. Commun.* **172**, 377-384.

Pluk, H., Soffner, J., Lührmann, R., van Venrooij, W. J. (1998) cDNA cloning and characterization of the human U3 small nucleolar ribonucleoprotein complex-associated 55-kilodalton protein. *Mol. Cell. Biol.* **18**, 488-498.

Pruitt, K. D., Maglott, D. R. (2001) RefSeq and LocusLink: NCBI gene-centered resources. *Nucleic Acids Res.* **29**, 137-140.

Sambrook, J., Fritsch, E. F., Maniatis, T. (1989) *Molecular Cloning: a Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Schreiber, V., Amé, J.-C., Dollé, P., Schultz, I., Rinaldi, B., Fraulob, V., Ménissier-de Murcia, J., de Murcia, G. (2002) Poly(ADP-ribose) polymerase-2 (PARP-2) is required for efficient base excision DNA repair in association with PARP-1 and XRCC1. *J. Biol. Chem.* **277**, 23028–23036.

Shall S., de Murcia, G. (2000) Poly(ADP-ribose) polymerase-1: what have we learned from the deficient mouse model? *Mutat. Res.* **460**, 1-15.

Shieh, W. M., Amé, J.-C., Wilson, M. V., Wang, Z.-Q., Koh, D. W., Jacobson, M. K., Jacobson, E. L. (1998) Poly(ADP-ribose) polymerase null mouse cells synthesize ADP-ribose polymers. *J. Biol. Chem.* **273**, 30069-30072.

Smith, S. (2001) The world according to PARP. *Trends Biochem. Sci.* **26**, 174-179.

Smith, T. F., Gaitatzes, C., Saxena, K., Neer, E. J. (1999) The WD repeat: a common architecture for diverse functions. *Trends Biochem. Sci.* **24**, 181-185.

Stoesser, G., Baker, W., van den Broek, A., Camon, E., Garcia-Pastor, M., Kanz, C., Kulikova, T., Leinonen, R., Lin, Q., Lombard, V., Lopez, R., Redaschi, N., Stoehr, P., Tuli, M. A., Tzouvara, K., Vaughan, R. (2002) The EMBL Nucleotide Sequence Database. *Nucleic Acids Res.* **30**, 21–26.

Strausberg, R. L., Feingold, E. A., Klausner, R. D., Collins, F. S. (1999) The mammalian gene collection. *Science* **286**, 455-457.

Stultz, C. M., White, J. V., Smith, T. F. (1993) Structural analysis based on state-space modeling. *Protein Sci.* **2**, 305-314.

Tateno, Y., Imanishi, T., Miyazaki, S., Fukami-Kobayashi, K., Saitou, N., Sugawara, H., Gojobori, T. (2002) DNA Data Bank of Japan (DDBJ) for genome scale research in life science. *Nucleic Acids Res.* **30**, 27-30.

Tatusova, T. A., Madden, T. L. (1999) BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences. *FEMS Microbiol. Lett.* **174**, 247-250.

Thompson, J. D., Higgins, D. G., Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673-4680.

Tong, W.-M., Cortes, U., Wang, Z.-Q. (2001) Poly(ADP-ribose) polymerase: a guardian angel protecting the genome and suppressing tumorigenesis. *Biochim. Biophys. Acta* **1552**, 27-37.

Venema, J., Vos, H. R., Faber, A. W., van Venrooij, W. J., Raué, H. A. (2000) Yeast Rrp9p is an evolutionarily conserved U3 snoRNP protein essential for early pre-rRNA processing cleavages and requires box C for its association. *RNA* **6**, 1660-1671.

Wang, Z.-Q., Auer, B., Stingl, L., Berghammer, H., Haidacher, D., Schweiger, M., Wagner E. F. (1995) Mice lacking ADPRT and poly(ADP-ribosyl)ation develop normally but are susceptible to skin disease. *Genes Dev.* **9**, 509-520.

Wang, Z.-Q., Stingl, L., Morrison, C., Jantsch, M., Los, M., Schulze-Osthoff, K., Wagner, E. F. (1997) PARP is important for genomic stability but dispensable in apoptosis. *Genes Dev.* **11**, 2347-2358.

Wheeler, D. L., Church, D. M., Lash, A. E., Leipe, D. D., Madden, T. L., Pontius, J. U., Schuler, G. D., Schriml, L. M., Tatusova, T. A., Wagner, L., Rapp, B. A. (2002) Database resources of the National Center for Biotechnology Information: 2002 update. *Nucleic Acids Res.* **30**, 13-16.

White, J. V., Stultz, C. M., Smith, T. F. (1994) Protein classification by stochastic modeling and optimal filtering of amino-acid sequences. *Math. Biosci.* **119**, 35-75.

Yu, L., Gaitatzes, C., Neer, E., Smith, T. F. (2000) Thirty-plus functional families from a single motif. *Protein Sci.* **9**, 2470-2476.