Identification of DeltaN isoform and polyadenylation site choice variants in molluskan *p63/p73*-like homologues

Running title: **DeltaN isoform and polyadenylation site choice variants**

***Keywords*:** p53 family, DeltaNp63/p73-like, haemocyte neoplasia, *Mytilus* species, alternate polyadenylation site choice

***Authors*:** Annette F. Muttray[1,2,*], Rachel L. Cox[3], Carol L. Reinisch[1], Susan A. Baldwin[2]

1-Laboratory of Aquatic Biomedicine, Marine Biological Laboratory, Woods Hole, MA 02543 USA

2-Department of Chemical and Biological Engineering, University of British Columbia, Vancouver, British Columbia V6T 1Z4 Canada

3-Present address: Center for Computational Biology and Bioinformatics, Columbia University, New York, NY 10032 USA

* Address of Correspondence:

Department of Chemical and Biological Engineering

University of British Columbia

2360 East Mall

Vancouver, British Columbia, V6T 1Z3 Canada

tel. +1-604-827-5760

fax +1-604-822-6003

e-mail: amuttray@chml.ubc.ca or amuttray@vcn.bc.ca

Abstract

The p53 family of transcription factors has been implicated in many vertebrate cancers. Altered p53 and p73 protein expression observed in leukemic cells of mollusks suggests that these transcription factors might be involved in invertebrate cancers as well. Here, we fully characterize the mRNA of four novel *p53*-like variants in the bivalve mollusks *Mytilus trossulus* (bay mussel) and *Mytilus edulis* (blue mussel). These species, widely used for environmental assessment, develop a haemic neoplasia (leukemia) that is frequently fatal. The correlation between expression of p53 and its close relative p73 and onset of molluskan leukemia was documented previously. We report the sequences of two distinct and novel *p63/p73*-like mRNAs, amplified by polymerase chain reaction (PCR) from both species. One of the *p63/p73*-like isoforms contains a 360 nt truncation in the 5' coding region. Based on this truncation and concomitant lack of a trans-activation (TA) domain, we designate this variant as a *DeltaNp63/p73*-like isoform: the first to be reported in an invertebrate species. In mammalian species, DeltaNp73 potently inhibits the tumor-suppressive function of p73 and p53, and its over-expression serves as a robust marker for mammalian cancer.

In addition, we report on the occurrence of alternate polyadenylation sites in the molluskan *p63/p73*: one proximal and one distal site, which differ by 1260 nt. We hypothesize that differential expression of various molluskan p63/p73-like isoforms, controlled in part by polyadenylation site choice variation, may help to interpret the apparently opposing roles of this gene in the development of cancer. Overall, this research further illustrates the utility of the molluskan model for studies involving the molecular mechanisms of oncogenesis in naturally occurring populations.

The data presented here require a revisiting of hypotheses regarding evolution of the *p53* gene family. Current hypotheses indicate that 1) the protostome gene family does not contain an

intronic promoter for DeltaN expression and 2) *p53* gene duplication did not occur in protostomes. Our characterization of DeltaN p63/73 in mussel suggests that molluskan *p53* gene family members have acquired an intronic promoter or splicing mechanism, either by invention that predates the evolutionary split of deuterostoms from protostomes, or by parallel evolution. Our data also show that *Mytilus* p53, p63/p73 and DeltaNp63/p73 are identical in their core regions with variation limited to their C- and N-terminals. This supports the notion that alternative splicing, intronic promoter usage and polyadenylation site choice may lead to expression of distinct isoforms originating from one common gene.

1. Introduction

Mussels and other shellfish are increasingly used for environmental monitoring in fresh water and marine environments due to their abundance, sessile nature and filter-feeding habits. They are susceptible to carcinogens and can develop a neoplastic disease of the haemolymph (leukemia) (Mix 1983) which occurs, at least in part, in response to environmental stressors (Reinisch et al., 1984; Farley et al., 1991; McGladdery et al., 2001; St-Jean et al., 2005). The p53 family of transcription factors, highly conserved across distantly related species, is implicated in vertebrate tumorigenesis (Concin, et al. 2004, Uramoto et al., 2004) as well as in molluskan leukemia. Specifically, p53 and p73 protein expression is altered in the clam *Mya arenaria* when affected by haemic neoplasia (Kelley, et al. 2001; Stephens, et al. 2001). Similarly, p73 protein is expressed at higher levels in leukemic haemocytes of *Mytilus edulis* (St-Jean, et al. 2005).

The functional domains of the p53 family of proteins across species are: an acidic amino-terminal transactivation domain (TA), several DNA-binding domains (DBD), and a carboxy-terminal oligomerization domain (OLIGO). In contrast to *p53*, *p63* and *p73* undergo multiple post-transcriptional processes that generate distinct classes of proteins, some of which lack the N-terminal TA domain. For *p63* and *p73*, the carboxyl terminus undergoes alternate splicing, yielding several products which may or may not contain a sterile alpha motif (SAM), a homodimerization domain (HOMO) (Jessen-Eller, et al. 2002), and a sumoylation site (SUMO) (Minty et al., 2000; Strano et al., 2001; Moll and Slade, 2004).

Well established as key elements in mammalian neuronal and epidermal development, p73 and p63, respectively, respond to DNA damage by controlling apoptosis and cell cycle arrest (Moll and Slade 2004). The NH_2-terminal lacking p63 and p73 isoforms, DeltaN, are upregulated in some cancers (Concin et al., 2004), and are known to suppress the tumor-repressive function of p53, p63 and p73 (Ishimoto et al., 2002). Here, we report on the first N-terminal variant of a

p63/p73-like cDNA from any invertebrate species. Our findings show that the molluskan bivalves *Mytilus edulis* and *Mytilus trossulus* contain an N-terminally truncated isoform that is homologous to vertebrate DeltaN *p63* and *p73*. Since the N-terminal isoform is already in use as a marker for mammalian cancer (Uramoto et al., 2004), discovery of a Delta N p53-like isoform in *Mytilus* sp. enables research into the expression levels of this variant and its potential involvement in haemic neoplasia in these species.

In addition, this report describes and characterizes 3'UTR variants of the *p63/p73*-like mRNA in these two molluskan species. We hypothesize that these variants represent tandem alternate polyadenylation sites, also identified in another bivalave mollusk *S. solidissima p63/p73* (Cox et al., 2003). Alternate polyadenylation site choice variation is widely employed and well characterized in mammalian species. Site choice variation controls gene expression in a tissue- or developmental stage-dependent manner (Edwards-Gilbert et al., 1997; Tian et al., 2005). This process should be carefully distinguished from alternative splicing, a well documented process that occurs in the p53/p63/p73 gene family. The molluskan *p63/p73* 3'UTR site choice variants described here may potentially affect mRNA stability, translatability and transport. Identification of these 3' UTR variants suggest that alternate polyadenylation of *p63/p73* may act as an important post-transcriptional regulatory event. However, further experimentation involving in vivo monitoring of these variants is required before their roles in control of *p53* family expression are known. Further, we hypothesize that the occurrence of 3'UTR control of *p63/p73* gene expression in mollusks may be an alternate mechanism for variability in protein expression that is similar to, but distinct from that which has developed in mammalian species.

A question that remains unanswered is whether the molluskan p53 and p63/p73s are coded for by multiple genes, as in mammals (Moll et al., 2001), or whether these homologues represent variants of a single ancestral gene. Based on our phylogenetic analysis of *p53*, *p63* and *p73* cDNA

sequences from diverse species and alignments of *Mytilus* p53, p63/p73 and DeltaNp63/p73 putative proteins we hypothesize that the invertebrate p53 and p63/p73 isoforms may originate from a single gene. This is in contrast to vertebrate p53, p63 and p73 isoforms where gene duplication has taken place.

2. Materials and Methods

2.1. Samples and mRNA extraction

M. trossulus were collected in November 2004 at Jericho Beach, Vancouver, British Columbia. *M. edulis* were collected in February 2005 at Buzzards Bay, Woods Hole, Massachusetts. *M. trossulus* gills were excised and immediately immersed in 1.5 ml RNA Later (Ambion Inc., Austin, TX), kept at 4^{o}C overnight, then frozen at –20^{o}C, and shipped on dry ice for RNA extraction to the Woods Hole Marine Biological Laboratory. Total RNA was extracted using Trizol reagent (Invitrogen Corp., Carlsbad, CA) following manufacturer instructions. *M. edulis* gills were excised and immediately transferred to 2 ml Trizol for total RNA extraction. RNA from each species was re-suspended in DEPC-treated water and stored at –80^{o}C. mRNA was concentrated from three pooled gill total RNA extracts using Qiagen Oligotex mRNA maxi kit (Qiagen GmbH, Hilden, Germany).

2.2. Species identification

DNA was retained during the Trizol extraction procedure from each of the six mussel gill samples and re-suspended in 0.7 to 1.4 ml of 8 mM sodium hydroxide. The internal transcribed spacer (ITS) region of the ribosomal RNA genes was amplified using a method developed by (Heath et al., 1995) with the following modifications: 1 μl of template DNA was amplified in a 25 μl-reaction containing 2 mM $MgCl_2$, 5 mM each ITS primer (Table 1) and 1 U Taq polymerase

(Invitrogen). Thermal cycling was performed as follows: 94°C for 10 min, 10 cycles of 94°C for 30", 53°C for 30", 72°C for 1', 20 cycles of 94°C for 30", 55°C for 30", 72°C for 1', and final extension for 7' at 72°C. 5 μl of the PCR reaction was transferred to a fresh reaction tube and supplemented with 7.5 μl dH_2O, 1.5 μl React 2 buffer and 1 μl *HhaI* enzyme (Invitrogen). This mix was incubated for one hour at 37°C and the reaction was heat-inactivated for 5' at 65°C. 5 μl of this enzymatic digestion and 5 μl of the remaining PCR reaction were analyzed by gel-electrophoresis (2.5 % agarose) and ethidium bromide staining.

2.3. cDNA synthesis, polymerase chain reaction (PCR) with degenerate primers, cloning and sequencing

cDNA was synthesized from 1 μg mRNA using PowerScript™ Reverse Transcriptase (BD Biosciences Clontech, Palo Alto, CA) and 500 ng oligo-$dT_{(12-18)}$ primer following the manufacturer's instructions. Initial amplification of a p73 transcript was performed on cDNA using forward primer pMe-1F, based on the *p53* sequence of *M. edulis* (Muttray et al., 2005) and degenerate reverse primer pDeg-p73R(1698-1716) designed here to match the SAM domain of aligned vertebrate and invertebrate p73 protein sequences (Figure 2A). PCR was carried out in a step-up fashion starting at an annealing temperature of 30°C, increasing by two degrees every two cycles, with a final annealing temperature of 44°C for 25 cycles. Melting at 94°C for 30 sec., annealing for 30 sec., extension at 72°C for 1.5 min., with a final extension step for 7 min. The resulting 1.6 kb-PCR product was re-amplified from a gel stab and the resulting band was purified using the High Pure PCR product purification kit (Roche Diagnostics GmbH, Mannheim, Germany), ligated into pCR®2.1-TOPO® and cloned into *E. coli* TOP10 cells using the TOPO TA cloning kit (Invitrogen Corp., Carlsbad, CA). Colonies were screened by PCR with M13 F (-20) and M13 R primers (30 cycles, annealing temperature of 55°C) and positive clones were grown

overnight in Luria Broth with ampicillin (100 μg/ml). Plasmid DNA was extracted using the Plasmid Mini kit (Qiagen, Inc., Valencia, CA), and submitted for sequencing.

2.4. Rapid amplification of cDNA ends (RACE-PCR) and final PCR

3' and 5'-RACE-ready cDNA was produced using the BD SMART^{TM} RACE cDNA amplification kit and manufacturer's instructions. Several gene-specific primers were designed based on the sequence obtained from the degenerate PCR and used in semi-nested 3' and 5' Smart RACE PCR reactions using the BD Smart RACE cDNA amplification kit (BD Biosciences Clontech) (Table 1, Figure 2A). An annealing temperature of 68 ^{o}C was used for 30 cycles. Several consistent 5'RACE and 3'RACE PCR products of different lengths were gel-purified, cloned and sequenced as described above. Resulting sequences were used to design primers for full-length amplification of putative p63/p73 and DeltaNp63/p73 variants (Table 1, Figure 2A). Full-length PCR amplification was performed using a touch-down PCR protocol, where the annealing temperature was decreased by 0.5 degrees each cycle, starting at 65 ^{o}C, with a final annealing temperature of 55 ^{o}C, for a total of 35 cycles. Primer extension was performed at 72 ^{o}C for 3 minutes.

2.5 Amino acid sequence analysis

Mytilus p63/p73-like deduced amino acid sequences were submitted to protein-protein BLAST searches. We performed pair wise multiple amino acid alignment of the deduced *Mytilus* p53, p63/p73 protein isoforms with selected species using ClustalX 1.83 (Jeanmougin et al., 1998) (residue substitution matrix Gonnet) to illustrate conserved protein domains. Gap opening penalty was set to 10; gap extension penalty at 0.1. The alignment was viewed in Boxshade 3.21 and

edited in Word (Microsoft, Redmond, WA) in order to match known highly conserved regions of the protein. The species and accession numbers are listed in Table 2.

2.6 Phylogenetic nucleic acid sequence and 3'UTR analyses

Full-length cDNA sequences of *Mytilus* and other known *p53, p63* and *p73* (Table 2) were aligned using ClustalX 1.83. Gap opening penalty was set to 15; gap extension penalty at 6.66. The phylogenetic tree was produced in ClustalX, bootstrapped 200 times and displayed using TreeView (Page, 1996). Protein sequence identities were obtained from TreeView. The 3'UTR cDNA sequences, variant A and B from *M. edulis* were aligned in ClustalX 1.83. Potential signaling motifs (PAS hexamer and upstream efficiency elements) were identified by hand.

3. Results and Discussion

3.1. Characterization of p63/p73-like isoforms in M. trossulus and edulis

The identities of the mussel species were confirmed by ITS PCR and RFLP with *Hha*I (Heath et al., 1995). Each of the three *M. trossulus* gill DNA samples showed three bands at 180, >200 and 280 nt, while each of the three *M. edulis* gill DNA samples showed a doublette at approximately 180 and 200 nt and a third band at 480 nt (Figure 1). This result compares well with results obtained previously by Heath et al. (1995) taking into account paired bands that cannot be completely resolved by standard agarose electrophoresis.

RNA extracts from the three individuals of each species were combined for mRNA extraction and RACE PCR. Four new *p53* family variants were identified per species (GenBank accession numbers in Table 2). Sequence analysis of these variants revealed them to be more similar to *p63* or *p73* than *p53* due to a significantly longer C-terminal region which contains a sterile alpha motif (SAM), a protein-protein binding domain previously identified in mammalian

p63 and p73. Thus we refer to them as *p63/p73*-like isoforms. The full-length cDNAs for *M. edulis* and *trossulus* TAp63/p73 and DeltaNp63/p73-like isoforms were identified from cDNA from the combined gill tissue RNA extracts. Figure 2B shows the PCR products by ethidium bromide-stained agarose gel electrophoresis. The size variability in the cDNAs is due to the size differential between the 5' variants (TA and DeltaN), which differ by 360 bp, and between the 3' alternate polyadenylation site choice variants (A and B), which differ by 1260 bp. The full-length cDNA sequences were translated into putative protein sequences. The DeltaN isoforms of *M. edulis* and *M. trossulus*, termed MedNp63/p73-like and MtdNp63/p73-like, respectively, are 122 amino acids (aa) shorter at the N-terminus than the TA isoforms (Mep63/p73 and Mtp63/p73) (Figure 3). This is a greater difference in length than in human (43 aa) or mouse (40 aa), but is strictly dependent on the choice of the start codon and could be as short as 58 aa. TA and DeltaN isoforms are conserved throughout the remainder of the open reading frame (ORF).

5'RACE experiments revealed that there are three possible start sites in the TA isoforms (at aa 1, 35 and 56) prior to the TA domain, while the DeltaN isoforms only have one start site initiating the ORF. The N-terminus of *Mytilus* p63/p73 may be considerably longer than in other species, depending on selection of translation start sites. None of the potential start codons have a conserved "Kozak" sequence, although aa residues 35 and 56 are preceded by an A at –3 base position and could therefore be considered stronger translation initiation sites than aa 1. The putative protein sequences are highly conserved between *M. edulis* and *M. trossulus* (99% similarity), except for three positions: Threonine in *M. edulis* is replaced with serine in *M. trossulus* at position 142, and serine and threonine are inserted at aa positions 483 and 484 in *M. edulis*. This is a similar amino acid sequence conservation to previously reported p53 sequences between the two species (Muttray et al., 2005). The complete cDNA sequences are less conserved than the protein sequences between the two species: 97 % for the coding regions of both isoforms,

and 80-82 % including the coding region and 3'UTR of both isoforms. The calculated molecular mass of the putative DeltaN protein isoform is 60 kDa, and of the TA isoform 73 kDa.

We performed a comparative analysis between the four *Mytilus* p63/p73-like deduced protein isoforms and representatives of other major lineages: *Mya arenaria* and *Spisula solidissima* p63/p73 were chosen as representatives for bivalve mollusks and likely close relatives of *Mytilus*, and two well-characterized vertebrate mammalian representatives, *Homo sapiens* and *Mus musculus* (Figure 3). The *Mytilus* p63/p73 sequences are most closely related to *Mya arenaria* p73 (57 % similarity) and *Spisula solidissima* p63/p73 (55 % similarity), and only distantly related to *H. sapiens* and *M. musculus* p73 (27-31 % similarity). These are slightly lower similarity values than was observed previously for p53 for some of these species (Muttray et al., 2005).

3.2 Analysis of p63/p73-like domains

This is the first report of an N-terminal isoform of p63 or p73 in invertebrates. The DeltaN-like isoforms in *Mytilus* lack the transcriptional activation (TA) domain of the TA-like isoforms, as has previously been described in vertebrate p63 and p73 (Figure 3). The p63/p73 TA domain is highly conserved between species and likely provides the binding site for proteins regulating p53, p63 and p73 activity. The N-terminus of the *Mytilus* p63/p73-like isoform is 55 aa longer than the N-terminus of the previously identified p53 isoform (Muttray et al., 2005) (see p53 start site in Figure 3). Following the TA domain is a highly diverse proline-rich region containing two PxxP motifs potentially involved in apoptosis and binding of SH3-containing kinases involved in signal transduction (Kelley et al., 2001). It was shown that p73 and a non-receptor tyrosine kinase involved in the response to DNA damage, c-Abl, interact at a PxxP domain between the DNA binding domains (DBDs) and oligomerization domain (OLIGO) domains in humans (reviewed in Melino, 2003). However, this PxxP is not conserved in mollusks. There are four DBDs positioned

between aa 211 and 379. These DBDs are highly conserved between species, such as 93 % similarity with *M. arenaria* and *S. solidissima* p63/p73. Based on protein BLAST searches we found 63 % and 59 % similarity with *H. sapiens* and other vertebrate p63 and p73, respectively. Based on this result, the *Mytilus* isoforms may be more p63-like than p73-like, but a final conclusion about their identities cannot be reached until their protein function is characterized. Thus far, mammalian p63 isoforms have been indicated predominately in embryonic epithelial development, but also in human leukemia (Yamaguchi et al., 2001; Nakamura et al., 2004; Bernassola et al., 2005). Using an antibody to the homology (HOMO) domain (Cox et al., 2003), it was shown that p63/p73-like proteins were more highly expressed in leukemic than in normal haemocytes of mussels (St-Jean et al., 2005).

As shown in Figure 3, the OLIGO domain is located at aa positions 425 to 457, followed by a divergent block of sequence. The C-terminus contains the sterile alpha motif (SAM), and a tetramerization domain with HOMO and sumoylation (SUMO, small ubiquitin-like modifier) domains, as was previously found in other invertebrates (Jessen-Eller et al., 2002; Cox et al., 2003). A prototypical SUMO domain was characterized in p53, p63 p73 and other proteins as the terminal lysine residue preceded by a hydrophobic amino acid and glutamic acid at position +2 (mammalian cell lines; (K*X*E/D)) (Minty et al., 2000). Here, we observe two potential SUMO sites: First, a terminal lysine 715 preceded by a hydrophobic amino acid (alanine) and glutamic acid at +2 present only in *Mytilus*, and second, a lysine 706 preceded by a hydrophilic threonine and aspartic acid at +2 present in all shown bivalves (solid border, Figure 3).

In mammalian p73, the terminal lysine motif has been shown to interact with SUMO-1, which potentiates proteasomal degradation of p73 (Minty et al., 2000). Sumoylation characteristically adds ~20 kDa to the molecular weight of the protein which was observed in SDS-PAGE on *Spisula* p63/p73 using the HOMO domain antibody (Cox et al., 2003).

An alternative SUMO motif has been identified in novel SUMO-1-interacting proteins, containing a central S*X*S (or S*X*T) triplet preceded by predominantly hydrophobic amino acids and followed by predominantly acidic amino acids (Minty et al., 2000). The serine/threonine and acidic residues essential for the interaction constitute a double CKII kinase site [(S/T)*XX*(D/E)], and thus the interaction may be regulated by phosphorylation. Interestingly, the second potential *Mytilus* alternative SUMO site (aa 703, Figure 3, dotted border) also contains the (S*X*T*XX*D) motif and may thus be relevant for an alternate sumoylation pathway in *Mytilus*.

3.3 Phylogenetic analysis of p53, p63 and p73 in invertebrates and vertebrates

The overall structure and sequence homology in the p53 family indicates that p53, p63, and p73 evolved from a common ancestor (Melino et al., 2002). This ancestor has been described as resembling a p63 containing a SAM domain (Yang et al., 2002). Our BLAST searches with the putative *Mytilus* p63/p73 also indicate that the mollusk DBDs are most similar to mammalian p63 indicating that the mollusk gene(s) are more closely related to the ancestral p63+SAM. The question remains as to whether distinct *p53, p63* and *p73* gene sequences will be characterized in invertebrate species and whether these genes will be mapped to separate chromosomes, or whether p53, p63 and/or p73 are splice variants of one gene. Gene duplication removes the pressure of natural selection from one copy of the duplicated gene and allows this copy to mutate more freely and potentially acquire new functions. If gene duplication had occurred in mollusks one would expect core regions to deviate from the ancestral p63. In fact, mollusks (cambrian, 570 M years) are evolutionary older than vertebrates (ordovizium, 500 M years) and hence the core cDNA regions should be more diverse in mollusks than in vertebrates. We performed two alignments: a multiple alignment of putative p53, p63/p73 and DeltaNp63/p73 proteins from *M. edulis* (Figure 4) and a multiple pair wise alignment and phylogenetic tree construction with cDNAs of available

vertebrate and invertebrate *p53*, *p63* and *p73* (Figure 5). *M. edulis* p53, p63/p73 and DeltaNp63/p73 proteins show identical core sequences containing DBDs II to V, NLS, and OLIGO domains (Figure 4). The three putative proteins differ only in their C- and N-terminals, suggesting that alternative splicing and intronic promoter usage could lead to expression of three different p53, p63 and p73-like isoforms from one common gene. Similarly, *Mya arenaria p53* and *p73* exhibit identical core regions (Kelley et al. 2001), and Southern blot analysis using a probe against the DBD IV region indicated the presence of only one gene (van Beneden et al. 1997). Figure 5 (second alignment) shows that invertebrate mollusk *p53* and *p63/p73* cluster in one branch according to species, while vertebrate *p53, p63,* and *p73* are on distinct branches and cluster according to isoform rather than species. Both alignments support earlier hypotheses (Yang et al., 2002; Cox et al., 2003) that invertebrate p53-like isoforms thus far identified are likely splice variants of the same gene, rather than products of distinct genes, as is the case for vertebrates where gene duplication has taken place. However, Yang *et al.* also hypothesized that the invertebrate-to-vertebrate transition led to the acquisition of a dual promoter structure that yields TA and DeltaN isoforms of p63 and p73. With the isolation of a putative DeltaN isoform of p63/p73 from *Mytilus,* we now revise this hypothesis and suggest that internal splicing or the 'invention' of an intronic promoters in the p53-family has either preceded gene duplication in phylogenetically more recent organisms and is of far more ancient origin than was previously thought, or that it is the result of parallel evolution in both vertebrates and invertebrates. An internal promoter and N-terminal splicing was also recently identified for *Drosophila* p53 (Bourdon et al., 2005) which results in two possible p53 isoforms containing either a full-length or truncated TA domain. However, the two *Drosophila* p53 isoforms did not contain the SAM domain that differentiates p53 from p63 and p73 in the vertebrate and the tentative molluskan p53 family.

The *Drosophila* (and *C. elegans*) genomes have diverged more from the ancestral genome than vertebrate lineages due to high rates of molecular evolution and secondary gene loss (Kortschak et al., 2003; Raible and Arendt, 2004). Unlike flies and nematodes, other invertebrates (including mollusks) seem to have retained more of the ancestral urbilaterian gene inventory as well (Raible and Arendt, 2004). The isolation of a DeltaN isoforms in *Mytilus* also suggests that the mollusks are well-suited model systems for the study of *p53/p63/p73*-family gene regulation, especially in regards to embryonic development and tumorigenesis as over-expression of DeltaN isoforms may contribute to the genesis of tumors by negating the tumor-suppressive activity of p53 or TAp73 (Ishimoto et al., 2002; Concin et al., 2004). Western blot analysis with antibodies against the *Spisula* p63/p73 HOMO domain showed a higher concentration of a protein less than 66 kDa in size in leukemic haemocytes in *Mytilus* when compared to normal haemocytes (St-Jean et al., 2005). Based on our results, this protein may be a DeltaN isoform (predicted 60 kDa, see above) rather than a TA isoform (predicted 73 kDa).

3.4 3'UTR analysis

Using PCR, we amplified two distinct full-length cDNA sequences (including 3'untranslated regions up to the poly A tail) for each *Mytilus* p63/p73-like isoform (Figures 2 and 3). These 3'UTR variants are similar to site choice variants identified by Cox, et al. (2003) for the *Spisula p63/p73* gene, and clearly distinct from the C-terminal splice variants reported for vertebrate p63 and p73. An alignment of the 3'UTRs (Figure 6) reveals that the *Mytilus p63/p73*-like isoform contains two alternate polyadenylation (proximal and distal) sites. This is the case for each N-terminal isoform for each species. Figure 6 represents the 3'UTR site variants for the *Mep63/p73*-like isoform only. We designate these as tandem polyadenylation site choice variants and refer to them as variant A (proximal) and variant B (distal). A sequence comparison of the

polyadenylation site variants shows minor divergence, mainly in the region of the proximal poly-A tail (Figure 6). In general, when comparing the 3'UTR sequences, we found minor species-specific differences, but no difference between isoforms within an individual species.

Well documented in mammalian genes, alternate polyadenylation site choice gives rise to different mRNAs and thus to distinct protein products, that are often associated with specific tissue expression or developmental stage (Edwards-Gilbert et al., 1997). The choice of a particular polyadenylation site is thought to involve 3' signaling elements, such as the polyadenylation signals (PAS) hexamers with the sequence AATAAA (or close variants ATTAAA, TTTAAA) that are located within 40 nucleotides of the cleavage site. In addition, phylogenetically conserved elements upstream of PAS, known as upstream efficiency elements (USEs) are known to enhance polyadenylation efficiency and mRNA processing (Moreira et al., 1998). Figure 6 shows that both site choice variants contain potential PAS signals within 30 nts of the poly-A tail, suggesting that both variants are in use, possibly under different cell conditions. One PAS hexamer in the proximal variant A falls at a greater distance than 40nts from the poly A tail. This was also shown to be the case in bivalve molluskan *p73* 3'UTR sequences, the soft shell clam *Mya arenaria* (Kelley et al., 2001) as well as the surf clam *Spisula solidissima* (Cox et al., 2003). In addition, while both proximal and distal variants contain two potential USEs, (ATTTGAA and ATTTCTTA), their proximity to the polyadenylation site in variant A (proximal) indicates that this variant may be more stable.

Recently referred to as a "molecular hotspot for pathology" (Conne et al., 2000), the 3'UTR has a major influence on the regulation of gene expression. Polyadenylation site choice can change enormously the outcome of the translated protein product in terms of its stability and size. What implications these polyadenylation site choice variants have on the expression of p53-family proteins in *Mytilus* sp. still needs to be determined.

4. Conclusions

1) Four new *p53*-family cDNA sequences, designated *TAp63/p73* A (and B) and *DeltaNp63/p73* A (and B), were identified in two closely related mussel species, *M. edulis* and *M. trossulus*. The new sequences were found to be more similar to p63 or p73 than p53 (Muttray et al., 2005) due to a C-terminal SAM domain. One sequence lacked the TA domain and was designated *DeltaNp63/p73*-like. Two polyadenylation site choice variants were identified.
2) Phylogenetic analysis of *p53*, *p63* and *p73* cDNA sequences from invertebrate and vertebrate species suggests that the putative p53 family proteins identified in *M. trossulus* and *edulis* may be coded from one gene. This is in contrast to vertebrate species where gene duplication and diversification have occurred. Molluskan p53-like isoform nomenclature is preliminary.
3) This is the first report of an invertebrate DeltaNp63/p73-like isoform. This revises earlier hypotheses and infers that the invention of a dual promoter or alternate splicing for the expression of TA and DeltaN isoforms pre-dates gene duplication within the p53 gene family in more recent lineages, or that parallel evolution has occurred.
4) As we reported previously in other molluskan species, *Mytilus TA* and *DeltaNp63/p73* contain alternate 3'UTRs that may represent functional alternate polyadenylation site choice variants. Site choice variation may have arisen as a post-transcriptional point of control for variable gene expression in this ancestral form of the *p53* gene family.

Acknowledgements

A.F.M. was supported by the Greater Vancouver Regional District, BC, Canada and a Collaborative Research and Development Grant from the Natural Science and Engineering Research Council of Canada (#CRDP J 323120 –05). R.L.C. was supported by Environmental Protection Agency grant # R82935901 and National Institute of Health grant #1R21ES012273-01.

References

Bernassola, F., Oberst, A., Melino, G. and Pandolfi, P.P. (2005). The promyelocytic leukaemia protein tumour suppressor functions as a transcriptional regulator of p63. Oncogene 24: 6982-6986.

Bourdon, J.-C., Fernandes, K., Murray-Zmijewski, F., Liu, G., Diot, A., Xirodimas, D.P., Saville, M.K. and Lane, D.P. (2005). p53 isoforms can regulate p53 transcriptional activity. Genes Dev 19: 2122-2137.

Concin, N., Becker, K., Slade, N., Erster, S., Mueller-Holzner, E., Ulmer, H., Daxenbichler, G., Zeimet, A., Zeillinger, R., Marth, C. and Moll, U. (2004). Transdominant DeltaTAp73 isoforms are frequently up-regulated in ovarian cancer. Evidence for their role as epigenetic p53 inhibitors in vivo. Cancer Res 64: 2449-2460.

Conne, B., Stutz, A. and Vassalli, J.-D. (2000). The 3' untranslated region of messenger RNA: A molecular 'hotspot' for pathology? Nat Med 6: 637-641.

Cox, R.L., Stephens, R.E. and Reinisch, C.L. (2003). p63/73 homologues in surf clam: novel signaling motifs and implications for control of expression. Gene 320: 49-58.

Edwards-Gilbert, G., Veraldi, K.L. and Milcarek, C. (1997). Alternative poly(A) site selection in complex transcription units: means to and end? Nucleic Acids Res 25: 2547-2561.

Farley, C.A., Plutschak, D.L. and Scott, R.F. (1991). Epizootiology and distribution of transmissible sarcoma in Maryland softshell clams, *Mya arenaria*, 1984-1988. Environ Health Perspect 90: 35-41.

Heath, D.D., Rawson, P.D. and Hilbish, T.J. (1995). PCR-based nuclear markers identify alien blue mussel (*Mytilus* spp.) genotypes on the west coast of Canada. Can J Fish Aquat Sci 52: 2621-2627.

Ishimoto, O., Kawahara, C., Enjo, K., Obinata, M., Nukiwa, T. and Ikawa, S. (2002). Possible oncogenic potential of deltaNp73: A newly identified isoform of human p73. Cancer Res 62: 636-641.

Jeanmougin, F., Thompson, J.D., Gouy, M., Higgins, D.G. and Gibson, T.J. (1998). Multiple sequence alignment with Clustal X. Trends Biochem Sci 23: 403-405.

Jessen-Eller, K., Kreiling, J.A., Begley, G.S., Steele, M.E., Walker, C.W., Stephens, R.E. and Reinisch, C.L. (2002). A new invertebrate member of the p53 gene family is developmentally expressed and responds to polychlorinated biphenyls. Environ Health Perspect 110: 377-85.

Kelley, M.L., Winge, P., Heaney, J.D., Stephens, R.E., Farell, J.H., Van Beneden, R.J., Reinisch, C.L., Lesser, M.P. and Walker, C.W. (2001). Expression of homologues for p53 and p73 in the softshell clam (*Mya arenaria*), a naturally-occurring model for human cancer. Oncogene 20: 748-58.

Kortschak, R.D., Samuel, G., Saint, R. and Miller, D.J. (2003). EST Analysis of the Cnidarian Acropora millepora Reveals Extensive Gene Loss and Rapid Sequence Divergence in the Model Invertebrates. Curr Biol 13: 2190-2195.

McGladdery, S., Davidson, J., Quarrar, P., MacCallum, G. and Arsenault, G. (2001). Investigation of factors asssociated with proliferation of haemic neoplasia in cultured and wild soft-shell clams (*Mya arenaria*). Charlottetown, PEI, Department of Health Management, Atlantic Veterinary College, University of Prince Edward Island**:** 15 p.

Melino, G. (2003). p73, the "assistant" guardian of the genome? Ann NY Acad Sci 1010: 9-15.

Melino, G., De Laurenzi, V. and Vousden, K.H. (2002). p73: Friend or foe in tumorigenesis. Nat Rev Cancer 2: 605-15.

Minty, A., Dumont, X., Kaghad, M. and Caput, D. (2000). Covalent modification of p73alpha by SUMO-1. J Biol Chem 275: 36316-36323.

Mix, M.C. (1983). Haemic neoplasia of bay mussels, *Mytilus edulis* L., from Oregon: occurence, prevalence, seasonality and histopathological progression. J Fish Dis 6: 239-248.

Moll, U.M., Erster, S. and Zaika, A. (2001). p53, p63, p73-solos, alliences and feuds among family members. Biochim Biophys Acta 1552: 47-59.

Moll, U.M. and Slade, N. (2004). p63 and p73: Roles in development and tumor formation. Mol Cancer Res 2: 371-386.

Moreira, A., Takagaki, Y., Brackenridge, S., Wollerton, M., Manley, J.L. and Proudfoot, N.J. (1998). The upstream sequence element of the C2 complement poly(A) signal activates mRNA 3' end formation by two distinct mechanisms. Genes Dev 12: 2522-2534.

Muttray, A.F., Cox, R.L., St-Jean, S.D., van Poppelen, P. and Reinisch, C.L. (2005). Identification and phylogenetic comparison of p53 in two distinct mussel species (*Mytilus*). Comp Biochem Physiol C Pharmacol Toxicol Endocrinol 140: 237-250.

Nakamura, H., Inokuchi, K., Yamaguchi, H. and Dan, K. (2004). Abnormalities of p51, p53, FLT3 and N-ras genes and their prognostic value in relapsed acute myeloid leukemia. Journal of Nippon Medical School 71: 270-278.

Page, R.D.M. (1996). TREEVIEW: An application to display phylogenetic trees on personal computers. Comp Appl Biosci 12: 357-358.

Raible, F. and Arendt, D. (2004). Metazoan evolution: some animals are more equal than others. Curr Biol 13: 2190-2195.

Reinisch, C.L., Charles, A.M. and Stone, A.M. (1984). Epizootic neoplasia from softshell clams collected from New Bedford Harbour. Hazardous Waste 1: 73-81.

Stephens, R.E., Walker, C.W. and Reinisch, C.L. (2001). Multiple protein differences distinguish clam leukemia cells from normal hemocytes: evidence for the involvement of p53 homologues. Comp Biochem Physiol C Pharmacol Toxicol Endocrinol 129: 329-38.

St-Jean, S.D., Stephens, R.E., Courtenay, S.C. and Reinisch, C.L. (2005). Detecting p53 family proteins in Leukemia Cells of *Mytilus edulis* from Pictou Harbour, Nova Scotia. Can J Fish Aquat Sci 62: 2055-2066.

Strano, S., Rossi, M., Fontemaggi, G., Munarriz, E., Soddu, S., Sacchi, A. and Blandino, G. (2001). From p63 to p53 across p73. FEBS Lett 490: 163-170.

Tian, B., Hu, J., Zhang, H. and Lutz, C.S. (2005). A large-scale analysis of mRNA polyadenylation of human and mouse genes. Nucl Acids Res 33: 201-212.

Van Beneden, R.J., Walker, C.W., and Laughner, E.S. (1997).Characterization of gene expression of a *p53* homologue in the soft-shell clam (*Mya arenaria*). Mol Marine Biol Biotechnol 6(2): 116-122.

Yamaguchi, H., Inokuchi, K., Sakuma, Y. and Dan, K. (2001). Mutation of the p51/p63 gene is associated with blastic crisis in chronic myelogenous leukemia. Leukemia 15: 1729-1734.

Yang, A., Kaghad, M., Caput, D. and McKeon, F. (2002). On the shoulders of giants: p63, p73 and the rise of p53. Trends Genet 18: 90-95.

Figures and Tables

Figure 1:

Species determination based on ITS PCR and RFLP for six *Mytilus* gill samples used for RNA extraction, on 2.5 % agarose. Lanes 1-6, PCR amplicons of the internal transcribed spacer (ITS) region between the 18S and 28S ribosomal RNA genes; 7, no template control; 8-13, *Hha*I digest of PCR amplicons; 1-3 and 8-10, *M. trossulus* from Jericho Beach, Vancouver, BC; 4-6 and 11-13, *M. edulis* from Buzzards Bay, Woods Hole, MA. The marker lanes contain the 100 bp ladder.

Figure 2:

Isolation of *p63/p73*-like isoforms in *Mytilus*. **A**; Primers used in the amplification strategy of *p63/p73*-like isoforms in *M. trossulus*. All four identified isoforms are shown. Functional domains (TAD, DBD, SAM, and 3'UTR) are not to scale. Corresponding primer sequences are listed in Table 1. **B**; RT-PCR amplification products of *M. trossulus* mRNA extracted from gills by gel electrophoresis on a 1 % agarose gel followed by ethidium bromide staining. PCR was carried out using the primers indicated in Table 1. For cycling parameter, see Materials and Methods.

Figure 3:

Multiple pair wise alignment of deduced amino acid sequences for TA- and DeltaNp63/p73 or p73 isoforms from *M. edulis*, *M. trossulus*, *Mya arenaria*, *Spisula solidissima*, *Homo sapiens* and *Mus musculus*. Color coding: black on white, non-homologous residues; green on white, weakly similar residues; blue on white, block of identical residues; red on gray, block of conserved residues. Residues distinguishing the two *Mytilus* species are shown on dark-gray background. TAD, DBD, NLS, OLIGO, SAM, HOMO domains and SUMO recognition sites are indicated. Manual

adjustments were made to align vertebrate and deduced invertebrate SUMO sites. The TI (tetramerization) domain is indicated by a red dashed line, and proline-rich regions are indicated by 'PxxP'. The start site of p53 in *Mytilus* is indicated with an arrow. Species and accession numbers are listed in Table 2.

Figure 4:

Multiple pair wise alignment of putative *M. edulis* p53, TAp63/p73 and DeltaNp63/p73 protein isoforms using ClustalW 1.8. Core regions are identical (*), while the C- and N-termini vary.

Figure 5:

Invertebrate p53 and p73-like orthologues cluster closely in one branch according to species, while vertebrate p53, p63 and p73 cluster in separate braches according to p53-family member. The unrooted neighbor-joined consensus tree was based on a pair wise ClustalW 1.8 alignment of nucleic acid sequences and bootstrapped 200 times. Numbers at the nodes indicate bootstrap values. The bottom scale measures genetic distances in substitutions per nucleotide. See Table 2 for species list and abbreviations.

Figure 6:

A. Schematic representation of *p63/p73*–like mRNA from *Mytilus* showing tandem polyadenylation site choice variants proximal (A) and distal (B).

B. *M. edulis p63/p73* 3' UTR regions, showing tandem polyadenylation site variants. Each sequence begins with the TGA stop site and ends with the poly A tail (in italics). Putative PAS hexamers are underlined. Putative USEs are highlighted in grey.

Table 1:

Primers used in the identification of *Mytilus* TA- and DeltaNp63/p73. “A” denotes short proximal transcripts utilizing the first alternate polyadenylation site, while “B” denotes long distal transcripts utilizing the second alternate polyadenylation site.

Table 2:

Species names, abbreviations, accession numbers of sequences generated in this publication or used for phylogenetic comparisons. Phylogenetic taxa are indicated.

Figure 1

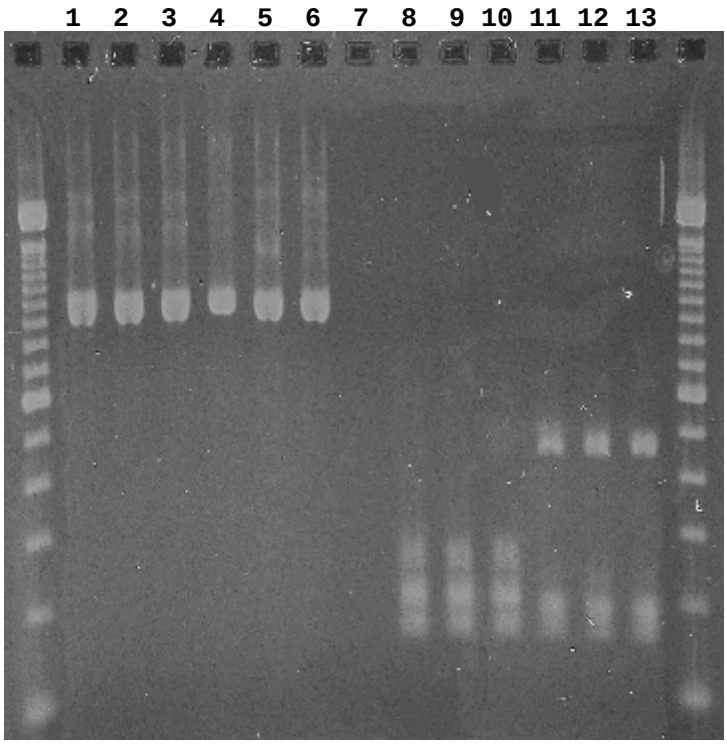
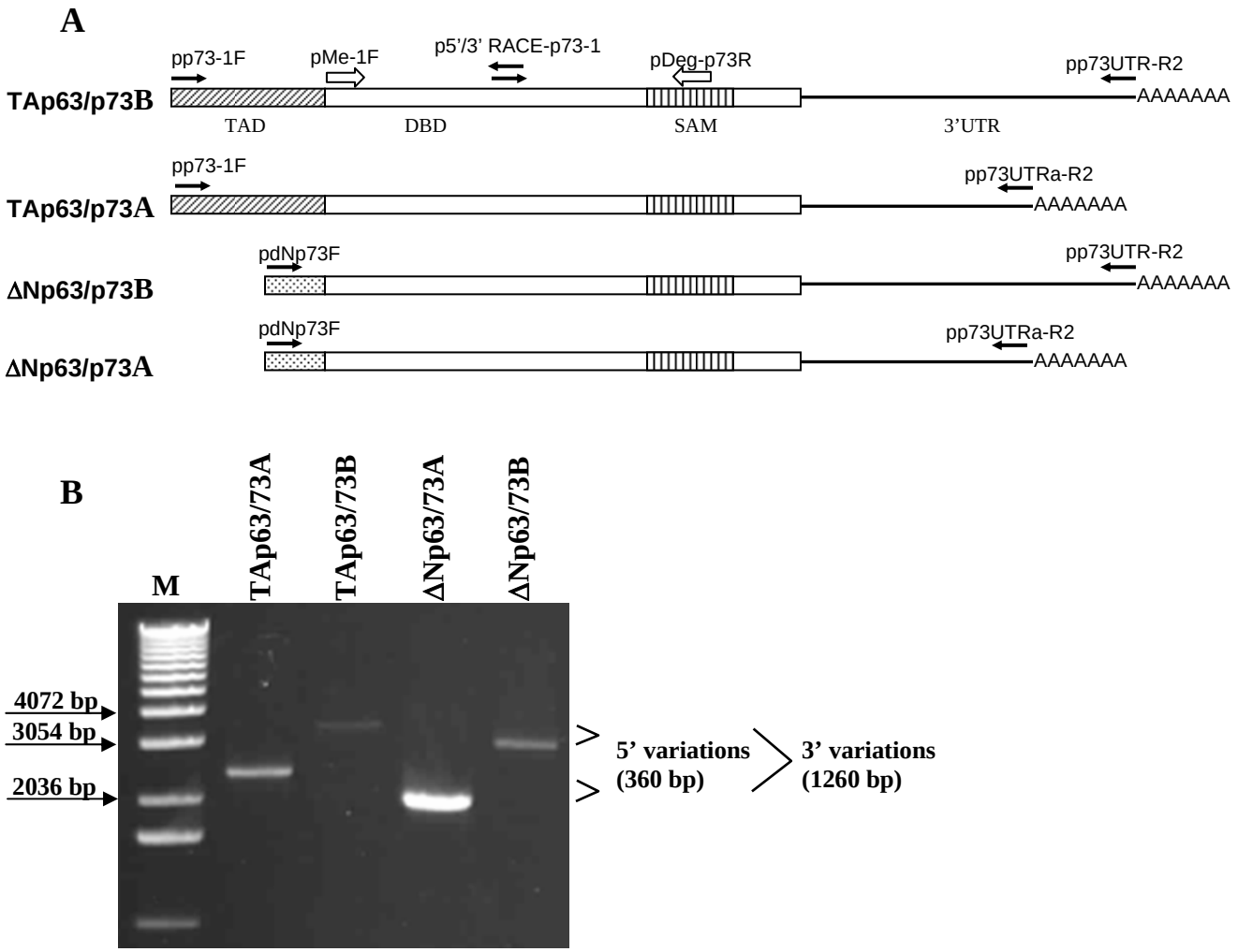
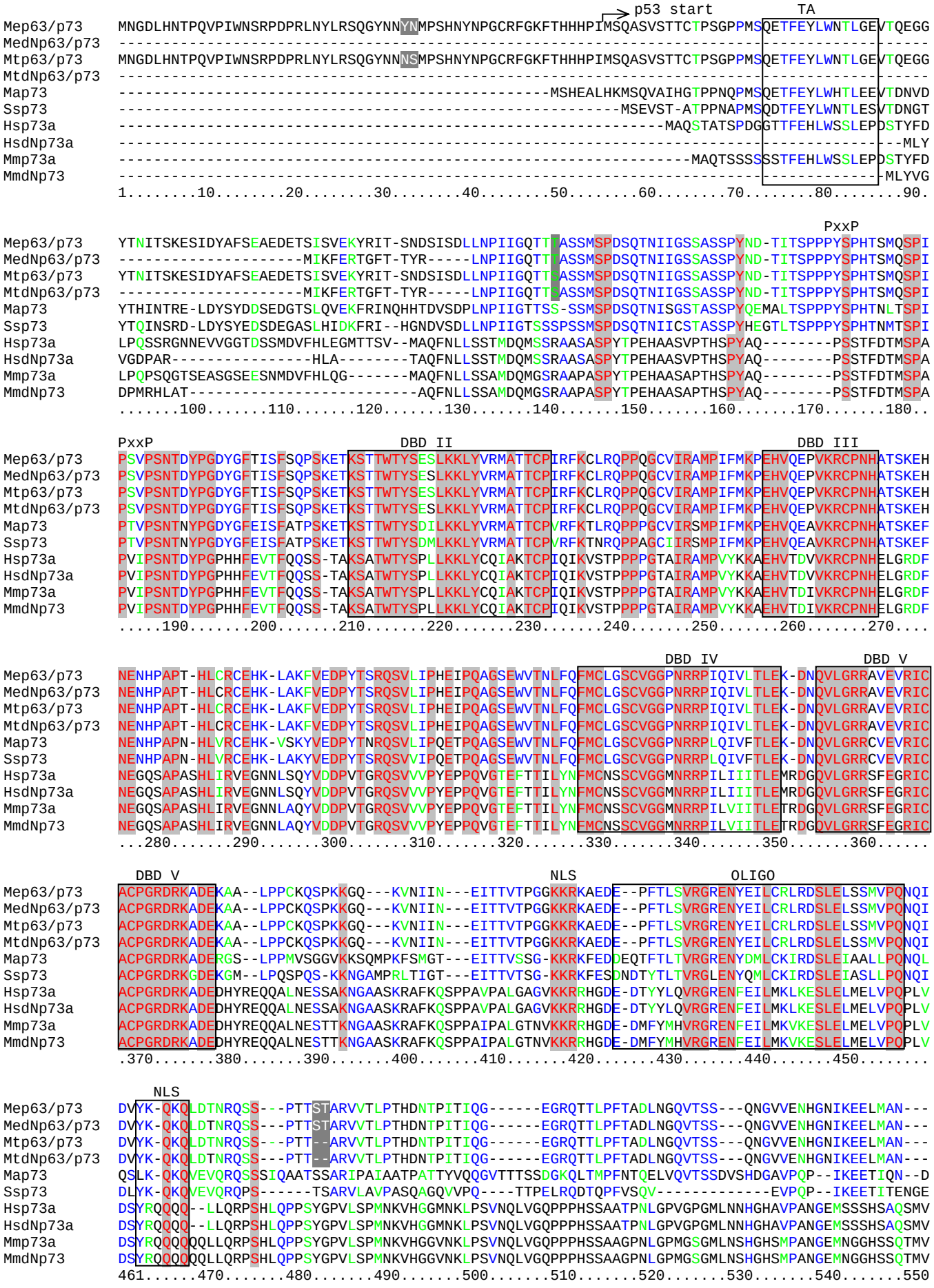


Figure 2


A
pp73-1F
pMe-1F
p5'/3' RACE-p73-1
pDeg-p73R
pp73UTR-R2
TAp63/p73B
AAAAAAA
TAD
DBD
SAM
3'UTR
pp73-1F
pp73UTRa-R2
TAp63/p73A
AAAAAAA
pdNp73F
pp73UTR-R2
ΔNp63/p73B
AAAAAAA
pdNp73F
pp73UTRa-R2
ΔNp63/p73A
AAAAAAA
B
M
TAp63/73A
TAp63/73B
ΔNp63/73A
ΔNp63/73B
4072 bp
3054 bp
2036 bp
5' variations
(360 bp)
3' variations
(1260 bp)


p53 start
TA
PxxP
PxxP
DBD II
DBD III
DBD IV
DBD V
DBD V
NLS
OLIGO
NLS
Mep63/p73
MedNp63/p73
Mtp63/p73
MtdNp63/p73
Map73
Ssp73
Hsp73a
HsdNp73a
Mmp73a
MmdNp73

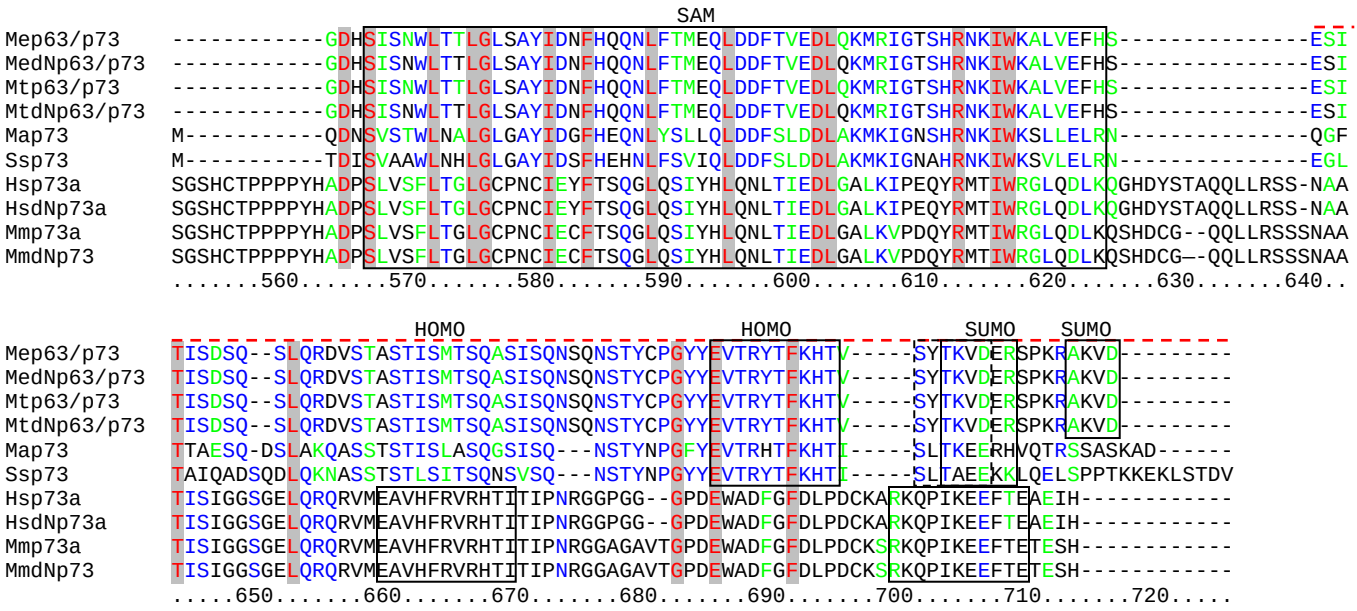

SAM
Mep63/p73
MedNp63/p73
Mtp63/p73
MtdNp63/p73
Map73
Ssp73
Hsp73a
HsdNp73a
Mmp73a
MmdNp73
HOMO
HOMO
SUMO
SUMO

Figure 3

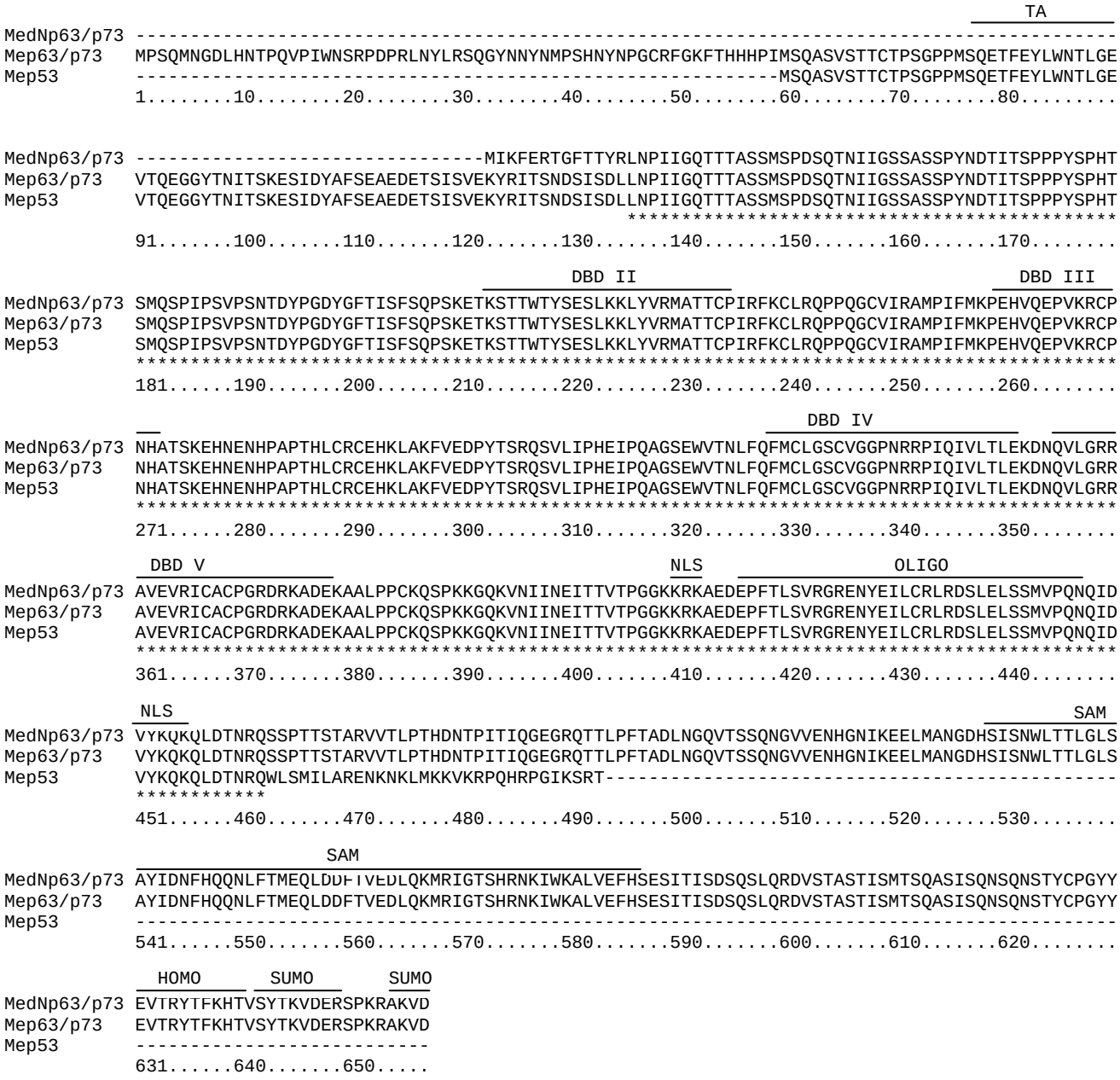

TA
MedNp63/p73
Mep63/p73
Mep53
DBD II
DBD III
DBD IV
DBD V
NLS
OLIGO
SAM
HOMO
SUMO
SUMO

Figure 4

Figure 5

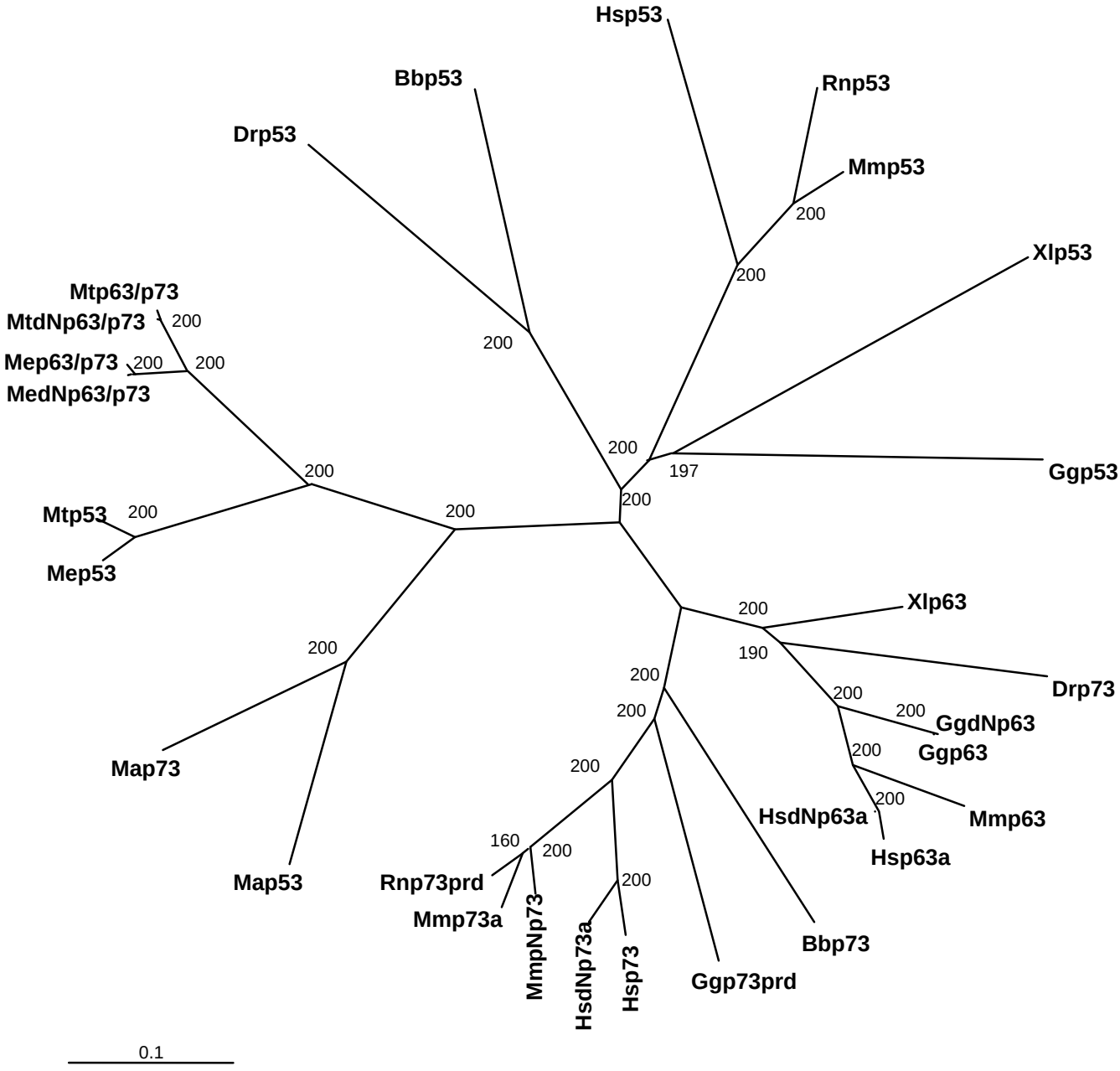
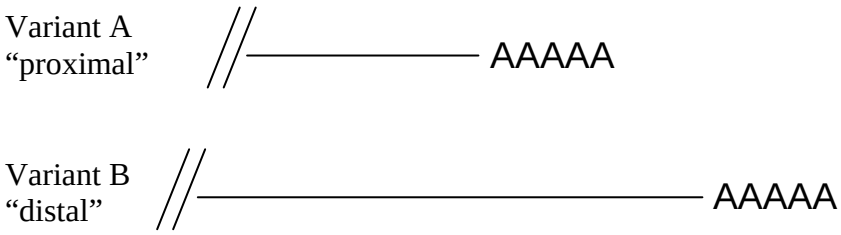

Hsp53
Bbp53
Rnp53
Drp53
Mmp53
200
Xlp53
Mtp63/p73
MtdNp63/p73
200
Mep63/p73
200
200
MedNp63/p73
200
200
197
Ggp53
200
Mtp53
200
200
Mep53
200
Xlp63
200
190
Drp73
200
200
200
GgdNp63
Ggp63
200
Map73
200
200
HsdNp63a
Mmp63
160
200
Hsp63a
Map53
Rnp73prd
200
Mmp73a
MmpNp73
HsdNp73a
Hsp73
Ggp73prd
Bbp73
0.1

Figure 6

A



B

```
Mep63/p73a TGATATGCTGAAGACTTATGCCACAACCAAACTAGGCTTGATATGATTGGTCAGTCCTTTACTGTGTAT
Mep63/p73b TGATATGCTGAAGACTTATGCCACAACCAAACTAGGCTTGATATGATTGGTCAGTCCTTTACTGTGTAT
           ************************************************************************

Mep63/p73a TGGACTGACATATATAGAGCAAATGCAATATGGTGCACAAATCAATGTCTAAATGAAGACATTGTAATC
Mep63/p73b TGGACTGACATATATAGAGCAAATGCAATATGGTGCACAAATCAATGTCTAAATGAAGACATTGTAATC
           ************************************************************************

Mep63/p73a TGATATTTTGTTAGACCGCCTGTTTTTAATGCAAATGTTAGATTCTGTCATGTGACCTGTATGGACCTA
Mep63/p73b TGATATTTTGTTAGACCGCCTGTTTTTAATGCAAATGTTAGATTCTTTTATGTGACCTGTATGGACCTA
           ******************************************** * ********************

Mep63/p73a TTATTGATTTGAAATGGCTATTTATTAAACCGATGATGATATGTAAAAATGTTGGTTTATTTCTTAGTT
Mep63/p73b TTATTGATTTGAAATGGCTATTTATTAAACCGATGATGATGCGTAAAAATGTTGGTTTATTTCTTAGTT
           ************************************** ************************

Mep63/p73a CTCTTTATCTTCAAATACTATATGTTTTCTATTGCAGTTTAAGATTAAATGGCATATCAAATTTCAAAG
Mep63/p73b CTCTTTATCTTTAACTTCTGTATGTTTTCTATTGCAGTTTAAGATTAAATGGCATATCAAATTTCAAAG
           *********** ** * ** ************************************************

Mep63/p73a CATATCCACTTATCTTAATTAAATTGCAGATTTTTTTACTGCAAAACAGGTACAAATATGTTTAAAATG
Mep63/p73b CATGACCACTTATCTTAATTAAATTGCAGATTTTTTTACTGCAAAACAGTTACAAGTATATTTAAAATG
           ***  ****************************************** ***** *** *********

Mep63/p73a TTTGTTGAAAGCCTTTGAATAAAAAAAAAAAAAAAAAAAAAAAAAA
Mep63/p73b TTTGTTGAAAGCCTTTGAATAAAAAAAAATAGCCACTGAACCAAACATTCTAGGCCTTTATGTGTAACTT
           ****************************

Mep63/p73b TTTTCAACTGTTAACTAGTTAACTGGACCATTAATAAAAATCTATTTATAACACGTAATTATGCTTCTAA
           TAATTTTAACAATAGGTGCTTAGACATAACACTTAAGATTATTATTAAAATGGCCTTTAAATTTATGTCC
           ATTCTGGGTGTTATTGTTGATGGAGAAAACAATGGAATGCCTGTTTTCTGAATGTTGTTGATTATATTAT
           ATATATATATATATGCATTTTGTCTGTTGTTCTGCTGAGATCAAGTTCAGGGTAGAATTTCTGCAATTGT
           ATTCATTTCAAGATTTCCACTAAATTCCATAGTAGGAATTTCATAATGAGGCTCAAAATTCCATAGTGAG
           GCTTAGTATGTCATTTGCAAAAGATTGTTGTCATTAGTAACTGCCATGTAGCATTTTGCAGAAGAACTAC
           TCTGAGCTAGTGTGAGTTATAACAATAATAGTGTTCAAATTTTATGATAAAACTATGTGTTCAGAATTAT
           CATGTAAACATCGAACTTTAAAAGTGTTCAACATTAAATTCATTTGCATTCATGTATATCGTTTTGTTTT
           TTTAACTGCTGATTTTGCAACAGTGCTATAAAATTGAATCAAATATTTTGCAAAATGGGGATATATTATT
           TCCCTCATAATTTCATTAATTAGTATTGTTCACCTTATGTCCTCTAAAGATGCTCATAGGTTGTTTCCAA
           GCTGAATTTTTTTGTAATCTGTTTTTAGGAAATTGGCAAAACTTGTCAAAAGAACTTCTTTTATAAAATG
           GGAATAAACTCAAATCATTTTGTGTATAATTATCAATGAAAACAACATTCACAGGTGGACATCCATTTGT
           TATAATTTTTTCATAAATTCCAGGGATGAAATTTGTTCCTCCAAGGAAATTGACTTCATAATCATTCCAT
           TGCCACCATTGTAATTCATTCTGTTTCCGTCCAAATACATGCGTGTGTTCTTTTGTTCATGCAACTACAA
           ATTANTGGCTATTTATCTGTTTAATTAAACATTACTAGTTTATTTATTTTAAACATCTCGAAAAAAAAAAA
           AAAAAAAAAAAAAAAAAA
```

Table 1:

Reaction	**Forward primer (5’->3’)**	**Reverse primer (5’->3’)**	**approximate product size (nt)**
ITS-PCR	gtttccgtaggtgaacctg	ctcgtctgatctgaggtcg	1250
Degenerate PCR	pMe-1F: atgtcacaagcttcagtttcaac	pDeg-p73R(1698-1716): ttccaDatYttRttHcKYtg	1600
3’ RACE	p3’RACE-p73-1: gtgaaggtcggcagacgacactaccg	n.a.	2500 and 1000
5’ RACE	n.a.	p5’RACE-p73-1: cggtagtgtcgtctgccgaccttcac	1800, (1600, 1400)
Full-length p63/p73 A	*M. trossulus*, pp73-2F: atgaatggggatttacacaatacacc	*M. trossulus*, pp73UTRa-R2: taaattcaaaagctttaagcaaacattttaaac	2440
	M. edulis, pp73-1F: atgccgtctcaaatgaatggg	*M. edulis*, pMep73a-R1: ggctttcaacaaacattttaaacatatttg	*2500*
Full-length p63/p73 B	pp73-1F: atgccgtctcaaatgaatggg	*M. trossulus*, pp73UTR-R2: ttcttccttagtagcaagc	3700
	pp73-1F: atgccgtctcaaatgaatggg	*M.edulis*, pMep73b-R2: ttgcatgaacaaaagaacacacg	*4000*
Full-length DeltaN p63/p73 A	pdNp73F: atgatcaaatttgagagaactgg	*M. trossulus*, pp73UTRa-R2: taaattcaaaagctttaagcaaacattttaaac	2080
	pdNp73F: atgatcaaatttgagagaactgg	*M. edulis*, pMep73a-R1: ggctttcaacaaacattttaaacatatttg	*2000*
Full-length DeltaN p63/p73 B	pdNp73F: atgatcaaatttgagagaactgg	*M. trossulus*, pp73UTR-R2: ttcttccttagtagcaagc	3340
	pdNp73F: atgatcaaatttgagagaactgg	*M.edulis*, pMep73b-R2: ttgcatgaacaaaagaacacacg	3060
Colonie screen	M13 F (-20): gtaaaacgacggccag	M13 R: caggaaacagctatgac	n.a.

Table 2:

Abbreviation	Scientific species name and cDNA designation	Accession number	Phylogenetic taxa
Mep53	*Mytilus edulis* p53 homologue	**AY579472**	Invertebrate, Mollusca
Mep63/p73	*Mytilus edulis* p63/p73 homologue, variant B (distal)	**DQ060435**	Invertebrate, Mollusca
MedNp63/p73	*Mytilus edulis* DeltaNp63/p73 homologue, variant B (distal)	**DQ060436**	Invertebrate, Mollusca
Mep63/p73	*Mytilus edulis* p63/p73 homologue, variant A (proximal)	**DQ865150**	Invertebrate, Mollusca
MedNp63/p73	*Mytilus edulis* DeltaNp63/p73 homologue, variant A (proximal)	**DQ865151**	Invertebrate, Mollusca
Mtp53	*Mytilus trossulus* p53 homologue	**AY611471**	Invertebrate, Mollusca
Mtp63/p73	*Mytilus trossulus* p63/p73 homologue, variant B (distal)	**DQ060437**	Invertebrate, Mollusca
MtdNp63/p73	*Mytilus trossulus* DeltaNp63/p73 homologue, variant B (distal)	**DQ060438**	Invertebrate, Mollusca
Mtp63/p73	*Mytilus trossulus* p63/p73 homologue, variant A (proximal)	**DQ865152**	Invertebrate, Mollusca
MtdNp63/p73	*Mytilus trossulus* DeltaNp63/p73 homologue, variant A (proximal)	**DQ865153**	Invertebrate, Mollusca
Map53	*Mya arenaria* p53 homologue	**AF253323**	Invertebrate, Mollusca
Map73	*Mya arenaria* p73 homologue	**AF253324**	Invertebrate, Mollusca
Ssp73	*Spisula solidissima* p63/p73 homologue	**AY289768**	Invertebrate, Mollusca
Drp53	*Danio rerio* p53	**AF365873**	Vertebrate, Ostariophysi
Drp73	*Danio rerio* tumor protein p73-like, transcript variant alpha 1	**NM_152986**	Vertebrate, Ostariophysi
Bbp53	*Barbus barbus* p53	**AF071570**	Vertebrate, Ostariophysi
Bbp73	*Barbus barbus* p73	**AF043641**	Vertebrate, Ostariophysi
Xlp53	*Xenopus laevis* tumor supressor protein p53	**AY221266**	Vertebrate, Amphibian
Xlp63	*Xenopus laevis* p63 DNA binding protein	**AF314148**	Vertebrate, Amphibian
Ggp53	*Gallus gallus* tumor protein p53	**NM_205264**	Vertebrate, Aves
Ggp63	*Gallus gallus* transformation related protein Trp63	**NM_204351**	Vertebrate, Aves
GgdNp63	*Gallus gallus* DeltaN p63 alpha	**AB045224**	Vertebrate, Aves
Ggp73prd	*Gallus gallus* similar to P73 alpha protein (predicted)	**XM_417545**	Vertebrate, Aves
Hsp53	*Homo sapiens* p53	**AB082923**	Vertebrate, Mammal
Hsp63a	*Homo sapiens* TA p63 alpha	**AF075430**	Vertebrate, Mammal
HspdNp63a	*Homo sapiens* DeltaN p63 alpha	**AF075431**	Vertebrate, Mammal
Hsp73	*Homo sapiens* p73 alpha	**NM_005427**	Vertebrate, Mammal
HspdNp73a	*Homo sapiens* DeltaNp73 alpha	**AB055065**	Vertebrate, Mammal
Rnp53	*Rattus norvegicus* tumor protein p53 (Tp53)	**NM_030989**	Vertebrate, Mammal
Rnp73prd	*Rattus norvegicus* transformation related protein 73 (predicted)	**XM_342992**	Vertebrate, Mammal
Mmp53	*Mus musculus* transformation associated protein p53	**X00741**	Vertebrate, Mammal
Mmp63	*Mus musculus* transformation related protein Trp63	**NM_011641**	Vertebrate, Mammal
Mmp73a	*Mus musculus* p73 alpha	**MMY9234**	Vertebrate, Mammal
MmdNp73	*Mus musculus* DeltaNp73	**MMU19235**	Vertebrate, Mammal