ELSEVIER

Comparative Biochemistry and Physiology, Part C 140 (2005) 237–250

CBP

www.elsevier.com/locate/cbpc

Identification and phylogenetic comparison of *p53* in two distinct mussel species (*Mytilus*)

Annette F. Muttray[a,*,1], Rachel L. Cox[b,1], Sylvie St-Jean[c], Paul van Poppelen[d], Carol L. Reinisch[b], Susan A. Baldwin[a]

[a]*Department of Chemical and Biological Engineering, University of British Columbia, 2216 Main Mall, Vancouver, BC, Canada V6T 1Z4*
[b]*Laboratory of Aquatic Biomedicine, Marine Biological Laboratory, 7 MBL Street, Woods Hole, MA 02543, United States*
[c]*National Water Research Institute, Environment Canada, 867 Lakeshore Rd., Burlington, ON, Canada L7R 4A6*
[d]*Greater Vancouver Regional District, 4330 Kingsway, Burnaby, BC, Canada V5H 4G8*

Received 21 September 2004; received in revised form 9 February 2005; accepted 10 February 2005

Abstract

The extent to which humans and wildlife are exposed to anthropogenic challenges is an important focus of environmental research. Potential use of *p53* gene family marker(s) for aquatic environmental effects monitoring is the long-term goal of this research. The *p53* gene is a tumor suppressor gene that is fundamental in cell cycle control and apoptosis. It is mutated or differentially expressed in about 50% of all human cancers and *p53* family members are differentially expressed in leukemic clams. Here, we report the identification and characterization of the *p53* gene in two species of *Mytilus*, *Mytilus edulis* and *Mytilus trossulus*, using RT-PCR with degenerate and specific primers to conserved regions of the gene. The *Mytilus* p53 proteins are 99.8% identical and closely related to clam (*Mya*) p53. In particular, the 3′ untranslated regions were examined to gain understanding of potential post-transcriptional regulatory pathways of *p53* expression. We found nuclear and cytoplasmic polyadenylation elements, adenylate/uridylate-rich elements, and a K-box motif previously identified in other, unrelated genes. We also identified a new motif in the *p53* 3′UTR which is highly conserved across vertebrate and invertebrate species. Differences between the *p53* genes of the two *Mytilus* species may be part of genetic determinants underlying variation in leukemia prevalence and/or development, but this requires further investigation. In conclusion, the conserved regions in these *p53* paralogues may represent potential control points in gene expression. This information provides a critical first step in the evaluation of *p53* expression as a potential marker for environmental assessment.

Keywords: Environmental effects monitoring; Leukemia; Mussel; *Mytilus*; *p53* gene family; Phylogenetic footprint analysis; Toxicogenomics; 3′ untranslated region

1. Introduction

Environmental effects monitoring often includes in situ biological information provided by fish and benthic invertebrate species to assess the health status of their environment (Environment Canada, 2003). The marine bivalve mollusc, mussel *Mytilus* sp., is widely used by Mussel Watch Programs (O'Connor and Beliaeff, 1995), for monitoring pulp and paper mill effluents (Salazar and Salazar, 1997; St-Jean et al., 2003) and, more recently, in municipal effluents effects monitoring (St-Jean et al., in press). The sessile nature of this species facilitates the establishment of cause-and-effect relationships in time and space (Widdows et al., 1995) and minimizes the possible confounding factor associated with the use of migratory species and those using larger areas for foraging.

One sublethal monitoring endpoint currently under investigation by our laboratories is haemic neoplasia (leukemia), an ultimately fatal condition well documented in clams and mussels. Leukemia is characterized by

* Corresponding author. Tel.: +1 604 827 5760; fax: +1 604 822 5407.
E-mail address: amuttray@vcn.bc.ca (A.F. Muttray).
[1] Both authors contributed equally to this publication.

1532-0456/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.cca.2005.02.011

continuously dividing malignant cells in the haemolymph of shellfish and is thought to be caused by anthropogenic substances (pesticides, PCBs), abnormal temperatures, viral transmission (McGladdery et al., 2001) and/or genetic background and seasonality (Elston et al., 1992). Other authors (Krishnakumar et al., 1999) were unable to link chemical exposure to onset of haemic neoplasia. We and others have demonstrated that the *p53* gene and its family member, *p73*, are implicated in the onset of molluscan leukemia (Barker et al., 1997; Kelley et al., 2001; Stephens et al., 2001). In the softshell clam *Mya arenaria*, mutations occur in the *p53* transcript of leukemic vs. normal haemocytes (Barker et al., 1997). In addition, transcription of the *p53* gene family member *p73* is up-regulated in leukemic haemocytes of adult *M. arenaria (*Kelley et al., 2001) and p53 family member protein expression is altered in the transition of normal haemocytes to leukemia cells (Stephens et al., 2001). In leukemic *M. arenaria* haemocytes, p53 is located in the cytoplasm rather than in the nucleus as it is for normal cells, indicating that p53 is functionally altered in leukemic haemocytes (Kelley et al., 2001). These observations suggested that *p53* family genes and other genes regulating their expression in *Mytilus* spp. could be used as potential early-warning biomarkers for evaluation of anthropogenic impacts affecting DNA structure or function. As a start, the goal of the research presented here was to identify and characterize the *p53* gene in two species of *Mytilus*, *Mytilus edulis* and *Mytilus trossulus*, both of which are currently used for bio-monitoring.

The p53s, a well-characterized family of transcription regulators, act to promote expression of genes controlling DNA editing and repair, apoptosis and carcinogenesis. The *p53* tumor suppressor gene was first discovered as a suspected oncogene by three independent research groups in 1979 (DeLeo et al., 1979; Lane and Crawford, 1979; Linzer and Levine, 1979). Many studies illustrate the importance of the *p53* gene, as it is either mutated or inactivated in over 50% of human cancers (O'Brate and Giannakakou, 2003). p53 has been termed the "gatekeeper of the genome" as well as a "network hub" (Vogelstein et al., 2000) because of its central role in the molecular networks that decide the fate of cellular life and death. As a transcriptional activator, p53 is normally inactive or rapidly degraded, but becomes activated upon damage to DNA by radiation or chemical treatments, hypoxia, or activation of oncogenes. Thus, p53 prevents cells from passing on the wrong DNA message potentially turning these cells into malignant tumors.

The first molluscan *p53* to be identified originated from the squid *Loligo forbesi* (Ishioka et al., 1995) but the squid sequence is more similar to *p63/73,* close homologues of *p53.* Kelley et al. (2001) isolated *p53* and *p73* from the bivalve mollusc *M. arenaria* and a partial sequence for a *p53* is available for the oyster *Crassostrea rhizophorae* (Genbank accession number AY442309). Bhaskaran et al. (1999) described *p53* from various fish species and suggested the use of *p53* to study mutagenesis in fish and genotoxins in the aquatic environment.

The p53 protein is comprised of several conserved regions. The N-terminal transactivation domain provides the generic transactivation function and the binding site for MDM2, the main regulator of p53 stability. A proline-rich region has complex roles as a protein-binding site and a specific regulator of apoptosis (Courtois et al., 2004, and references therein). The C-terminus includes several domains involved in oligomerization and regulation of the specific DNA-binding activity. It also possesses a non-specific DNA binding activity thought to be involved in nonspecific p53-mediated DNA repair and in DNA/RNA reannealing (Wolkowicz and Rotter, 1997). The regulation of p53 function is tightly controlled through several mechanisms including *p53* transcription and translation, protein stability and post-translational modification, as well as p53 location in the cell nucleus or cytoplasm (O'Brate and Giannakakou, 2003, and references therein). Several authors have used various monoclonal and polyclonal antibodies raised against human and clam (*Spisula*, *Mya*) p53 family members to study the expression of the p53 family and to distinguish leukemic cells from normal cells (Kelley et al., 2001; Stephens et al., 2001; Jessen-Eller et al., 2002; Cox et al., 2003).

We isolated haemocytes from two species of bivalve mollusc, *Mytilus* spp., and used these as a source for identification of *p53* homologues. Here we report the identification, phylogenetic characterization and footprint analysis of distinct *p53* homologues from *M. edulis* and *M. trossulus*. These two species, albeit similar morphologically, differ significantly at the physiological level, such as gamete incompatibility (Rawson et al., 2003), different growth patterns (Penny et al., 2002), temporal separation and duration of spawning in Atlantic mussels, and total egg production and size (Toro et al., 2002). Work carried out on the Pacific coast (BC) by our research team on caged mussels has confirmed the reported differences in growth and reproductive cycles between the species in the Pacific region. While spawning in *M. edulis* is observed to occur in April, spawning in *M. trossulus* at the same location (cage) is observed from late June to early August. Here we report distinct variability in the primary gene structure of the *p53* paralogues from these two species, a difference which may help understand species-specific susceptibilities to environmental challenges.

Phylogenetic comparisons as well as footprint analyses of the *Mytilus* spp. *p53* sequences revealed that unique and highly conserved sequence sites occur in the 3′ untranslated regions (3′UTRs). Our previous reports suggest that the occurrence of *cis*-acting signaling sites in the 3′UTRs of the *p53* gene family members in clams may control gene expression (Cox et al., 2003). Further investigations are required to determine the potential role of these sites in post-transcriptional signaling or regulation of gene expression. This work makes use of comparative

data to identify potential regulatory mechanisms controlling *p53* gene expression. This represents an important first step toward the use of *p53* gene family expression as a marker for leukemia and a valid environmental assessment monitor.

2. Materials and methods

2.1. Organisms

Certified *M. edulis* (imported from Prince Edward Island) were obtained from Island Scallops, Qualicum, British Columbia, and deployed in cages in March 2003 at various locations in the Vancouver Harbour, BC. *M. edulis* for this study were collected in October 2003 from one of the cages. *M. trossulus* was collected in March 2004 at Jericho Beach, Vancouver Harbour, British Columbia. In a previous report, Jericho Beach (Locarno) mussels were found to be 98% *M. trossulus* based on shell morphometry measurements (Mallet, 2003). Mussels used for the isolation of *p53* were subsequently identified and confirmed by the same method by Mallet Research Services, NS (McDonald et al., 1991; Mallet and Carver, 1995).

2.2. Total RNA extraction

Haemolymph was withdrawn from the posterior adductor muscle area using a dry syringe with 22.5-gauge needle. A drop of haemolymph was deposited on a microscope slide and haemocytes were allowed to adhere to the glass surface for 5 min at room temperature. Samples were then examined for potential parasite and mantle fluid contamination by microbial source or by tissue such as gamete, and moribund individuals were discarded. Estimate of percentage of neoplastic vs. normal haemocytes was determined using phase contrast microscopy (McGladdery et al., 2001). mRNA was extracted from haemocytes of an *M. edulis* animal in transitional phase (Farley et al., 1991) and an *M. trossulus* animal in normal phase using the Trizol reagent and protocol (Invitrogen Life Technologies, Mississauga, ON). Haemolymph was centrifuged at 3000 rpm using a microfuge and carefully resuspended in 0.5 ml chilled Trizol. The suspension was snap-frozen in a dry-ice/ethanol bath and stored at −80 °C until further extraction.

2.3. RT-PCR

First strand cDNA synthesis was carried out on approximately 5 μg of total RNA extract with oligo-dT primers following the guidelines for PowerScript™ Reverse Transcriptase (BD Biosciences Clontech, Mississauga, ON). Initial sequences were obtained for *M. edulis* by degenerate PCR using primer design that was based on *Mya* p53 protein DNA binding region (forward primer DegF2, reverse primer DegR2K, Table 1, kindly provided by Charles Walker). This resulted in the amplification of an initial 320-bp product which was further extended to the 5′ and 3′ region by RACE PCR (BD Biosciences Clontech). Final full-length clones for both species for the coding region were obtained with forward primer pMe-28F and reverse primer pMe1323R (Table 1), and for the 3′ untranslated region (3′UTR) by 3′RACE PCR with forward primer 3′RACE829. Step-down PCR cycling conditions for all reactions were as follows: Initial melting at 95 °C for 1 min, 5 cycles of 94 °C for 30 s, 68 °C for 45 s, 72 °C for 3 min; 5 cycles of 94 °C for 30 s, 66 °C for 45 s, 72 °C for 3 min, and so on to 5 cycles of 94 °C for 30 s, 62 °C for 45 s, 72 °C for 3 min; and finally 20 cycles of 94 °C for 30 s, 60 °C for 45 s, 72 °C for 3 min; followed by a final extension at 72 °C for 10 min. Clones were obtained by TA cloning into plasmid vector pCR2.1 and transformation of *Escherichia coli* INVαF' or TOP10F' (Invitrogen Life Technologies). PCR products for the coding regions were cloned directly, while PCR products from the 3′RACE PCR were gel purified from 0.7% to 1.0% agarose TAE gels using the Roche High Pure PCR Product Purification kit (Roche Applied Science, Laval, QC) or the QiaQuick Gel Extraction kit (Qiagen Mississauga, ON) using manufacturers' instructions. Plasmids with the correct length insert based on PCR with universal forward and reverse M13 primers and gel electrophoresis were then submitted for sequencing to the Nucleic Acid and Protein Sequencing facility, University of British Columbia. All clones were sequenced in both directions using the universal M13 primers. The number of clones sequenced for the final sequences were as follows: three clones each for *M. edulis* and *M. trossulus* coding region, 10 clones for *M. edulis* 3′UTR, and one clone for *M. trossulus* 3′UTR.

Table 1
Primers used in the identification of *p53* in *Mytilus* sp.

Name	Sequence 5'→3'
Deg F2	gtIaaRMgItgYccIaaYcaK
Deg R2K	NggRcaNgcRcaDatNcKNacYtc
pMe-28F	tggaaagttcactcatcatcacc
pMe1323R	atatatcctcaatgttcctgaacc
3'RACE829	catgtgtaggaggaccaaacagaaggcc

Capital letters indicate degenerate residues: I, inosine; R, A/G; M, A/C; Y, C/T; K, G/T; N, A/T/G/C; D, A/G/T.

2.4. Amino acid sequence analysis

All *Mytilus* sequences were submitted to Discontiguous Mega BLAST searches. We performed pairwise amino acid alignment of the deduced *Mytilus* p53 proteins with selected species using Align X (residue substitution matrix Blosum) to illustrate conserved protein domains. Gap opening penalty was set to 10; gap extension penalty at 0.05; gap separation penalty at 8. The alignment was edited in order to match highly conserved regions of the protein. A second pairwise multiple alignment with 32 other known p53

proteins was performed using ClustalX 1.83 (residue substitution matrix Gonnet) (Jeanmougin et al., 1998). The species and accession numbers are listed in Table 2. Gap opening penalty was set to 10; gap extension penalty at 0.1. The phylogenetic tree was produced in ClustalX, bootstrapped 200 times and displayed using TreeView (Page, 1996). Protein sequence identities were obtained from ClustalX alignment.

2.5. Nucleic acid sequence analyses

3′UTR sequences were submitted to UTRdb (Pesole et al., 2002) for analysis. 3′UTR sequences were aligned in ClustalX 1.83 and known regulatory elements were edited by hand. We also submitted a wide range of *p53* 3′UTRs to a phylogenetic footprint analysis using the FootPrinter 2.1 Web server available at http://bio.cs.washington.edu/software.html. Parameters were set to default, except for motif size and maximum number of mutations, which were set according to the diversity of the species examined. For instance, for a broad diversity search we allowed for two mutations in a small motif of 6–10, while for a narrow diversity search, we allowed no mutation and a longer motif size of 8–10.

3. Results and discussion

3.1. Identification and characterization of Mytilus p53 cDNA

The nucleic acid sequence of the cDNA for *p53* was determined from two different mussel species, *M. edulis* and *M. trossulus,* and deposited in Genbank as AY579472 and AY611471, respectively. The total length of the cDNAs are 2288 nucleotides (nt) for both species, with an open reading frame (ORF) of 1302 nt, predicting a protein of 434 amino acids in length with a calculated molecular mass of 50 kDa, based on an average amino acid weight of 115 Da. This is within the range of previously reported p53 proteins. The nucleic acid sequences of the ORFs of *M. edulis* and *M. trossulus* are 96.5% identical. However, predicted amino acid (aa) sequences for the two species are 99.8% identical, indicating slightly different codon usage by the two species.

3.2. Comparative analysis of p53 amino acid homologues

We performed a comparative analysis between the two *Mytilus* p53 sequences and representatives of other major

Table 2
List of species, accession numbers, and abbreviations used for the phylogenetic analysis of p53 proteins in Fig. 2

Common name	Scientific name	Abbreviation	Accession number	Phylogenetic taxa
Entamoeba	*Entamoeba histolytica*	Ehp53	AJ489250	Eukaryota, Entamoeba
Blue mussel	*Mytilus edulis*	Mep53	AY579472	Invertebrate, Mollusca
Bay mussel	*Mytilus trossulus*	Mtp53	AY611471	Invertebrate, Mollusca
Softshell clam	*Mya arenaria*	Map53	AF253323	Invertebrate, Mollusca
Squid	*Loligo forbesi*	Lfp73	U43595	Invertebrate, Mollusca
Fruit fly	*Drosophila melanogaster*	Dmp53	AF263722	Invertebrate, Arthropoda
Potato beetle	*Leptinotarsa decemlineata*	Ldp53	BD250011	Invertebrate, Arthropoda
Flour beetle	*Tribolium castaneum*	Tcp53	BD250012	Invertebrate, Arthropoda
Nematode	*Caenorhabditis elegans*	Cep-1	AF440800	Invertebrate, Nematoda
Swordtail	*Xiphophorus helleri*	Xhp53	AF043946	Vertebrate, Neoteleostei
Medaka	*Oryzias latipes*	Olp53	U57306	Vertebrate, Neoteleostei
Flounder	*Platichthys flesus*	Pfp53	Y08919	Vertebrate, Neoteleostei
Congo puffer	*Tetraodon miurus*	Tmp53	AF071571	Vertebrate, Neoteleostei
Barbel	*Barbus barbus*	Bbp53	AF071570	Vertebrate, Ostariophysi
Zebrafish	*Danio rerio*	Drp53	AF365873	Vertebrate, Ostariophysi
Channel catfish	*Ictalurus punctatus*	Ipp53	AF074967	Vertebrate, Ostariophysi
Rainbow trout	*Oncorhynchus mykiss*	Omp53	M75145	Vertebrate, Ostariophysi
Clawed frog	*Xenopus laevis*	Xlp53	X77546	Vertebrate, Amphibian
Chicken	*Gallus gallus*	Ggp53	NM_205264	Vertebrate, Aves
Human	*Homo sapiens*	Hsp53	AB082923	Vertebrate, Mammal
Dog	*Canis familiaris*	Cfp53	AB020761	Vertebrate, Mammal
Cattle	*Bos taurus*	Btp53	NM_174201	Vertebrate, Mammal
Vervet monkey	*Chlorocebus aethiops*	Cap53	X16384	Vertebrate, Mammal
Chinese hamster	*Cricetulus griseus*	Cgp53	U50395	Vertebrate, Mammal
Guinea pig	*Cavia porcellus*	Cpp53	AJ009673	Vertebrate, Mammal
Cat	*Felis catus*	Fcp53	D26608	Vertebrate, Mammal
Natal rat	*Mastomys natalensis*	Mnp53part	U48617	Vertebrate, Mammal
Sheep	*Ovis aries*	Oap53	X81705	Vertebrate, Mammal
European rabbit	*Oryctolagus cuniculus*	Ocp53	X90592	Vertebrate, Mammal
Rat	*Rattus norvegicus*	Rnp53	NM_030989	Vertebrate, Mammal
Mouse	*Mus musculus*	Mmp53	X00741	Vertebrate, Mammal
Pig	*Sus scrofa*	Ssp53	AF124298	Vertebrate, Mammal
Beluga whale	*Delphinapterus leucas*	Dlp53	AF475081	Vertebrate, Mammal

lineages: *M. arenaria*, a representative for bivalve molluscs and likely a close relative of *Mytilus*, and *Drosophila melanogaster*, a second representative for invertebrates, and four species of vertebrates, *Xenopus laevis*, two fish, *Danio rerio* and *Barbus barbus*, and one well-characterized mammalian representative, *Homo sapiens* (Table 2 and Fig. 1).

Overall, amino acid sequence identities between *M. edulis* p53 and the other aligned species, based on ClustalX identity tables, were as follows: *M. arenaria* Map53, 69%;

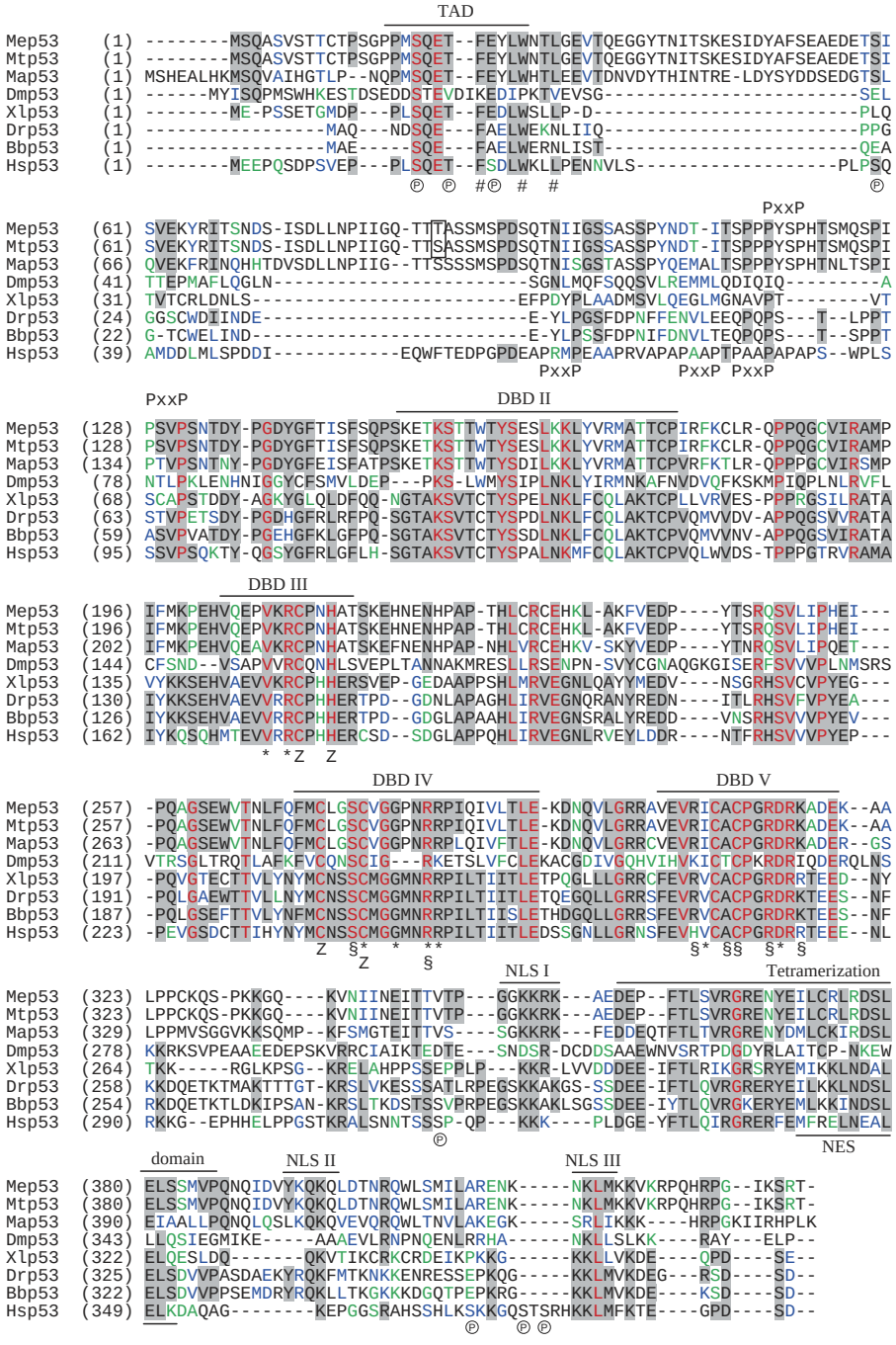


Fig. 1. Multiple pairwise alignment of amino acid sequences for p53 homologues from *Mytilus edulis* (Mep53), *Mytilus trossulus* (Mtp53), *Mya arenaria* (Map53), *Drosophila melanogaster* (Dmp53), *Xenopus laevis* (Xlp53), *Danio rerio* (Drp53), *Barbus barbus* (Bbp53), and *Homo sapiens* (Hsp53). Color coding: black on white, non-homologous residues; green on white, weakly similar residues; blue on white, block of similar residues; black on gray, conserved residues; red on gray background, identical residues. Abbreviations: TAD, transcriptional activation domain (or DNA binding domain I); DBD II-V, DNA binding domains II-V; NLS I and II, nuclear localization domain I and II; NES, nuclear export domain; PxxP, proline-rich domains in shellfish species; Ⓟ, putative conserved phosphorylation sites; #, key MDM2 binding sites; §, residues where Hsp53 binds to DNA; *, Hsp53 mutational hot-spots; Z, residues involved in zinc binding. The open box indicates the single amino acid change between Mep53 and Mtp53.

D. melanogaster Dmp53, 18%; *X. laevis* Xlp53, 36%; *D. rerio* Drp53, 38%; *B. barbus* Bbp53, 38%; and *H. sapiens* Hsp53, 31%. *Drosophila* p53 has the least identity with *Mytilus* p53. As predicted, *Mya* p53 is the closest known relative to *Mytilus* p53. However, no N-terminal extension was detected in *Mytilus* when compared to *Mya*. In *M. arenaria*, a N-terminal extension was detected 8 aa distance upstream from the conventional start site which leads to a potential second start site for protein translation (Kelley et al., 2001).

3.3. Mytilus p53 DNA binding domains

Mytilus p53s have five DNA binding domains (DBDs, including the transcriptional activation domain, TAD), which are also found in *Mya* and other species. DBD II to V are located between amino acid residues 150 and 319 in both Mytilidae. These domains are highly conserved between species. Overall predicted protein identities of these four DBDs of *Mytilus* p53 are 88% when compared to *Mya* p53 but only 63% when compared to *Homo* p53. The transcriptional activation domain (TAD) provides the binding site for proteins that regulate p53 expression, DNA editing and repair, and apoptosis. The TAD is highly conserved across species (except *D. melanogaster*). The Mep53 TAD is 100% identical with Map53 and 75% identical with human p53. Negative feedback regulation of p53 activity is driven by a protein called MDM2, and positive regulation of p53-dependent tumor suppression is driven by the transcriptional co-activator p300 (Shimizu and Hupp, 2003, and papers therein). MDM2 functions as an ubiquitin ligase in the nucleus and, upon binding to p53 TAD, destines p53 for export to the cytoplasm and subsequent degradation. It is involved in an autoregulatory feedback loop that maintains p53 at basal level in healthy non-stressed cells. Defects in the pathways of MDM2 regulation of p53 are common in tumors that retain wild-type p53 (Vousden and Lu, 2002). The crucial MDM2 contact sites in human p53 are F19, W23 and L26, (shown in Fig. 1 with # signs). These are also found in Mep53 and Mtp53. Key p300 contact sites in human p53 are overlapping with the MDM2 contact sites and are less conserved between species. Only three out of the six identified p300 binding sites are conserved between human and *Mytilus* p53. Different kinases, such as ATM (ataxia telangiectasia mutated) or CHK2 (checkpoint kinase 2) can modify p53 at specific amino acids within the TAD region. Ser15, Thr18 and Ser20 have been identified as important phosphorylation sites (Shimizu and Hupp, 2003, and papers therein). As shown in Fig. 1 (with P) Ser15 and Thr18 are highly conserved between species, including Mep53 and Mtp53, while the residue at position 20 is less conserved (fish, *Mya* and *Mytilus* display a glutamate at this position). It has been found that a number of human tumors are associated with cytoplasmic localization of p53. As mentioned above, MDM2 regulates p53 activity by ubiquitination, followed by nuclear export and degradation. MDM2 overexpression was frequently found in advanced leukemic cells of human patients which have a p53-null phenotype due to rapid p53 degradation (Konikova and Kusenda, 2003), but also in relation with low p53 levels or conformational mutants of p53 as measured with antibody PAb240. Similarly, leukemic clam haemocytes show localization of Map53 and Map73 in the cytoplasm and not in the nucleus (Kelley et al., 2001). Since these MDM2 binding sites are also conserved in *Mytilus*, this particular mechanism of leukemia may also be conserved in this species.

The region between the TAD and DBD II is the most divergent region in p53 and contains a number of proline residues. It appears that the molluscs (including *L. forbesi* p53, sequence not shown) have a much longer region just before the proline-rich region than other organisms (Fig. 1), the function of which is not known. The proline-rich region of the mussel p53 contains two PXXP motifs at positions 116 and 128. This is very similar to Map53, which also has only two PXXP motifs, while human p53 contains three such motifs and additional proline residues. These motifs are entirely absent from *Drosophila* and *Xenopus* p53 and are present at non-homologous sequence positions in other organisms. This proline-rich region is involved in apoptosis and may bind SH3-containing kinases involved in signal transduction (Kelley et al., 2001). This region also contains the only amino acid position which is different between the two mussel species: Ser86 and Thr86 for *M. edulis* and *M. trossulus*, respectively. It is unlikely that this variation will have an effect on p53 function and regulation as it is located in a less-conserved region of the protein.

Mussel p53 DBD II is 92% identical with *Mya* p53, but only 69% identical with human p53. DBD III shares 92 % identity with *Mya* and 81% identity with human p53. It contains two zinc binding sites which are fully conserved in all presented species. Similarly, two mutational hot-spots (Val 173 and Cys 175 in Hsp53) are 100% conserved (May and May, 1999). Mep53 DBD IV is 91% identical with *Mya* and 73% identical with human p53. Mutational hot-spots are based on the frequency of point mutations found in human cancers. Point mutations with the highest frequencies have been termed hot-spots. The three mutational hotspots contained in DBD IV are highly conserved, with the exception of *Drosophila* p53. Zinc and DNA binding sites are 100% conserved between all presented species. Known residues for mutational hot-spots and DNA binding sites for neighboring DBD V are slightly less conserved between species. Overall identities for this region between *Mytilus* and *Mya* are 100%, and between *Mytilus* and *Homo* 93%. Analysis of distribution of mutations in human p53 shows that they are essentially clustered in the central region of the protein and in particular within the four DNA binding domains (Soussi and May, 1996). Conservation of human mutational hot-spots in other species not only points to the essential nature of these amino acids for p53 function but

also raises the possibility that these residues may be involved in cancers in other species.

Mytilus p53 has a tripartite nuclear localization signal (NLS I, II and III) which is similar to human p53. In our alignment, *Mya* p53 also shows a tripartite NLS, although Kelley et al. (2001) indicated that *Mya* may only have two NLS. All NLS are rich in lysine residues. It is known that lysine K320 in Hsp53 NLS I, which is highly conserved between species, is acetylated to enhance p53 stability and p53-specific DNA binding with p53-regulated proteins (Liu et al., 2000). Nuclear import of p53 is enabled by its NLS while nuclear export is enabled by its nuclear export signal (NES) which is located within the tetramerization domain (see below). In humans, when DNA is damaged, p53 gets imported into the nucleus via its NLS and undergoes tetramerization, binds and activates DNA-damage response genes (O'Brate and Giannakakou, 2003, and papers therein). The tetramer state of p53 masks the nuclear export

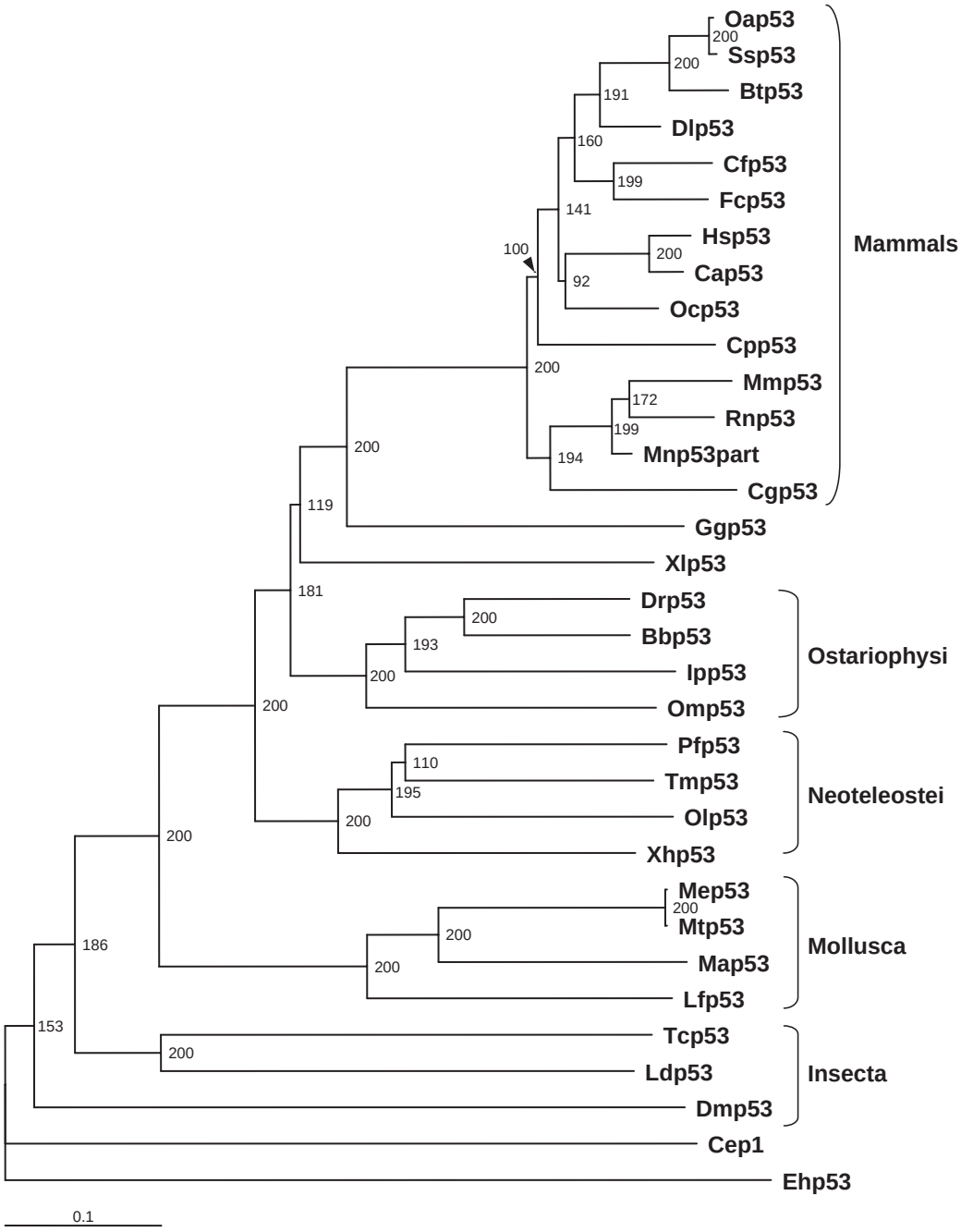


Fig. 2. Phylogenetic relationship between p53 proteins of diverse species indicating separate lineages for mollusca and annelida. See Table 2 for species list and abbreviations. The neighbor-joined consensus tree was based on a pairwise ClustalX alignment, bootstrapped 200 times and rooted with *Entamoeba* p53 as an outgroup. Numbers at the nodes indicate bootstrap values. The bottom scale measures genetic distances in substitutions per nucleotide. The clustering of p53s in phylogenetic groups (Mammals, Ostariophysi, Neoteleostei, Mollusca and Insecta) is indicated with brackets.

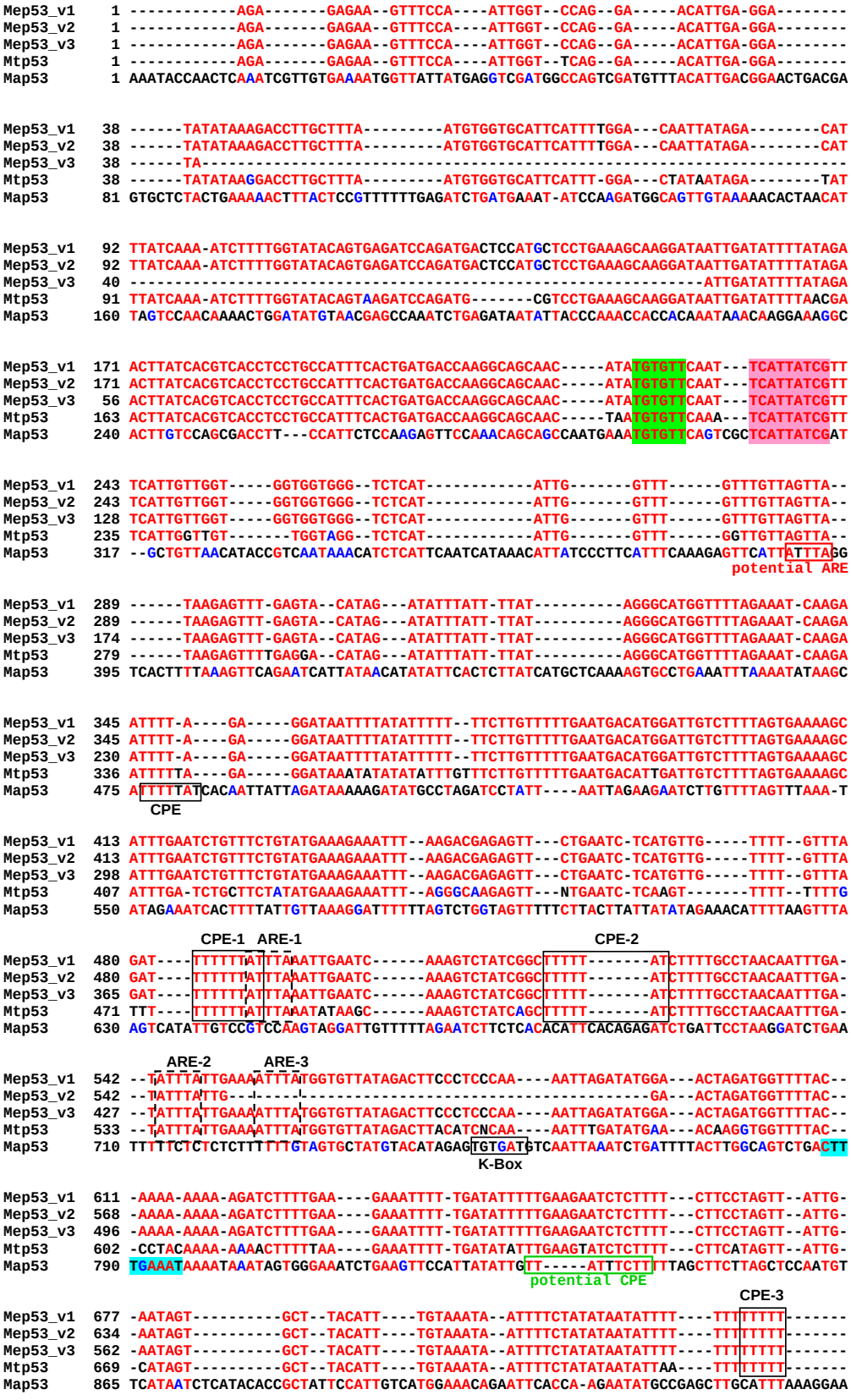


Fig. 3. Clustal X alignment of the 3′UTR of the *p53* gene of three variants of *M. edulis*, *Mep53*_v1, *Mep53*_v2 (AY735341), *Mep53*_v3 (AY735340), *M. trossulus* (Mtp53) and *M. arenaria* (Map53). Color coding: red, identical residues; blue, similar residues; black, non-similar residues. Abbreviations: CPE, cytoplasmic polyadenylation signal; ARE, adenylate/uridylate-rich elements. Conserved known regulatory regions are indicated by open and closed boxes. Colored backgrounds refer to potential conserved motifs identified in Fig. 4.

Fig. 3 (*continued*).

signal thereby preventing export to the cytoplasm. As mentioned above, p53 functions as a tetrameric protein. The tetramerization domain is 59% identical with *Mya* p53 and 53% identical with human p53 tetramerization domain. Glycine 334 (in humans) is crucial for stability of p53 and conserved throughout all presented lineages. Leucines 348 and 350 (in humans) are crucial for NES. They are conserved in *Mytilus*, but interestingly not in *Mya* where the second leucine is substituted by isoleucine. Positions 341 and 344 (in humans) are critical for oligomerization and are conserved non-polar residues, mostly leucines. The ability to form tetramers allows p53 to behave in a dominant-negative fashion (Hofseth et al., 2004). Our identification of highly conserved regions within the functional domain of *Mytilus* p53 suggests that its overall function is conserved. These observations, taken with previous data (Kelley et al., 2001) implicating p53 family members in the onset of leukemia, further suggest the utility of *Mytilus p53* gene expression as an indicator for early onset leukemia and an important molecular target of genotoxins in the environment.

3.4. Phylogenetic analysis

The recent finding of a *p53* homologue in *Entamoeba histolytica* (Mendosa et al., 2003) suggests that the *p53* gene family is of ancient origin. We performed a phylogenetic analysis of many to-date available p53 amino acid sequences to (a) confirm *Mytilus* p53 sequence relationships with other invertebrates, and (b) clarify the phylogenetic relationship of the *D. melanogaster* p53 homologue in relation to other invertebrates and vertebrates (Fig. 2, Table 2). We set entamoeba Ehp53 as the outgroup as it is likely the most distantly related p53 known thus far. We found that *Mytilus* p53s cluster with other known mollusc invertebrate p53s, especially with the other known bivalve p53s, as would be expected. The family Mytilidae dates back to the Jurassic or perhaps even Devonian times (Soot-Ryen, 1969) (ca. 400 million years ago). The tree topology supports the now-accepted view that the molluscan lineage is not part of a common lineage with the annelids as was suspected previously (Wilmer, 1990). It is remarkable that p53 of *D. melanogaster* (Dmp53) is monophyletic and does not cluster with the other insect p53 and suggests that *p53* of *D. melanogaster* may have undergone more recent mutations, for instance, loss of the MDM2 binding region. Because of its role as guardian of the genome as well as its recent identification in entamoeba, it may be possible to use *p53* and its relatives *p63* and *p73* (Yang et al., 2002) as central regulatory genes for phylogenetic studies as more sequences become available.

3.5. Analysis of the 3′ untranslated region of Mytilus p53 cDNA

Untranslated regions at the 3′ end of the mRNA contain signals for mRNA translation, polyadenylation, stability and subcellular localization, and play therefore an important role in gene regulation and expression at the post-transcriptional level (Pesole et al., 2002). We obtained the 3′UTR by 3′RACE PCR starting at the conserved DBD V. We sequenced 10 clones in *M. edulis* and subsequently one clone in *M. trossulus*. The total length, starting after the stop codon TGA and ending at the start of the polyA tail, is 917 nt in *M. edulis*, and 874 nt in *M. trossulus*, and fall well within the range of 3′UTR lengths (maximum 9142 nt, minimum 15 nt, average 444.5 nt) reported for invertebrate species (Pesole et al., 2001) (Fig. 3).

Mya p53 3′UTR contains 2027 nt and is therefore comparatively longer than the *Mytilus p53* 3′UTRs. The sequences vary widely between the mussels and the clam 3′UTRs, probably because these regions are under less evolutionary constraints than the coding regions (Conne et al., 2000; Grzybowska et al., 2001). Sequence alignment of the three consensus sequences (*Mep53* v1, v2, v3) obtained by manual alignment from 10 *M. edulis* clones revealed various deletions in the 3′UTR of *p53* for this species (Fig. 3). One clone had a deletion of 114 nt starting at position 40 (termed version 3), while three other clones had a deletion of 42 nt starting at position 551 (termed version 2) of the

full-length 3′UTR. Six of the ten clones were full-length (termed version 1). *M. trossulus p53* 3′UTR was most similar to the full-length version of the *M. edulis p53* 3′UTR (v1) with a minor deletion of 7 nt at 128 and a deletion of 31 nt at position 817. Regions bordering these deletions were not found to be reverse complement to each other, and it is therefore unlikely that these deletions are artifacts of polymerase slippage. It is currently unknown whether these deleted (or inserted) regions affect expression of the transcript or function of the protein.

While post-translational regulation and cellular localization of p53 have received considerable attention (for a review, see O'Brate and Giannakakou, 2003), post-transcriptional regulation of p53 activity has perhaps been undervalued in recent years. We identified the following regulatory regions based on current literature in bivalve *p53* 3′UTR:

(1) Tandem nuclear polyadenylation sites "AATAAA" are located 20 nt upstream of the polyA tail in *M. edulis*. A single polyadenylation site is located 14 nt upstream of the polyA tail in *M. trossulus*. This sequence is required for proper poly(A) tail addition and mRNA stabilization in the nucleus and binds to a complex of four polypeptides, the polyadenylation specificity factor (CPSF) (Verrotti et al., 1996). mRNAs initially receive their poly(A) tails of approximately 250 nucleotides in length in the nucleus of the cells. Upon entering the cytoplasm, poly(A) is removed in most cells.

(2) We located three cytoplasmic polyadenylation elements (CPEs) in a region starting at 482 nt after the stop codon (Fig. 3). CPEs are AU-rich regions and are generally defined as "$T_{4–6}$AT" (Pesole and Liuni, 1999, see also UTRsite at http://bighost.ba.itb.cnr.it/srs6bin/wgetz?-e+[UTRSITE-ID:*]). Located near the nuclear polyadenylation element, the CPEs are evolutionarily conserved sites, known to regulate translational activation by elongation of the poly(A) tail in the cytoplasm of the cell. Generally, poly(A) elongation confers translational activation while deadenylation promotes translational silencing (Richter, 1999). CPEs have mostly been studied in mouse, *Xenopus* and *Drosophila* oocyte maturation and early development. Detailed mutagenesis experiments established that minimal perturbations of *Xenopus* CPEs can abolish their function (Verrotti et al., 1996, and references therein). Sequence comparison across species is difficult because of the AT-richness of the region and because of a "substantial context and position effect on CPE function" (Verrotti et al., 1996), which is illustrated by the alignment of the *Mytilus* and *Mya* 3′UTRs (Fig. 3). Whether the deletion of regions of the *Mep53*_v2 and the *Mtp53* 3′UTR located within the CPE region has an effect on cytoplasmic polyadenylation and translational activation is entirely speculative at this point. In *Xenopus*, CPEs are recognized by CPE binding proteins which, together with the CPSF, may form a core cytoplasmic polyadenylation apparatus that is conserved across species (Verrotti et al., 1996). Recent evidence suggests that cytoplasmic polyadenylation not only plays a role in mRNA activation in early development, but also in synaptic memory (or plasticity) after stimulation in the brain (Richter, 2001) and in cell cycle control upon recognition of DNA damage in yeasts (Read and Norbury, 2002). CPE-dependent polyadenylation has also been demonstrated in human MCF7 breast cancer cell line, in a CPE-containing 3′UTR fragment of cyclin B mRNA (Groisman et al., 2002).

(3) The 3′UTR of *p53* also contains six adenylate/uridylate-rich elements (AREs) which are usually defined by the pentamer AUUUA or the nonamer UUAUUUA(U/A)(U/A), and have previously been found in labile mRNAs for regulatory proteins such as proto-oncoproteins, growth factors and their receptors, inflammatory mediators and cytokines (Grzybowska et al., 2001). The AREs primary function is to target mRNAs for selective degradation. However, ARE-mediated decay is itself regulated: under stress conditions, cell stimulation, or during oncogenic transformation, ARE-containing mRNAs are stabilized. Their regulatory functions are expressed through the specific binding of proteins which can modify transcript stability. The mussel sequences contain five consecutive ARE sequences in comparison to the clam *p53* sequence which contains only one (Fig. 3). As already concluded by Read and Norbury (2002), the further characterization of cytoplasmic polyadenylation implicates this mechanism of translational control in the regulation of increasingly diverse cellular processes, but seemingly most often in cellular responses to stress, DNA damage, replication block, and normal cell cycle events. It may therefore be of no surprise to find CPEs and AREs in *p53* as well as in its relatives, *p63/73* (Cox et al., 2003). One of the main advantages of translational control is that it enables rapid changes in gene expression without requiring gene transcription or mRNA transport (Read and Norbury, 2002).

In an attempt to identify novel signaling elements in the 3′UTR of *Mytilus p53*, we submitted the sequences to the UTRscan program at http://bighost.area.ba.itb.cnr.it/BIG/UTRScan/ (Pesole and Liuni, 1999) and identified a K-box motif as indicated in Fig. 3. The K-Box motif (ATGTGATA) occurs 179 nt upstream of the poly-A tail in all *Mytilus p53* and potentially in *Mya p53*. The more conserved "TGTGAT" motif occurs in *Mya p53*, 380 nt upstream of the poly(A) tail and may be a second K-box in this species. We also searched other aligned species for the

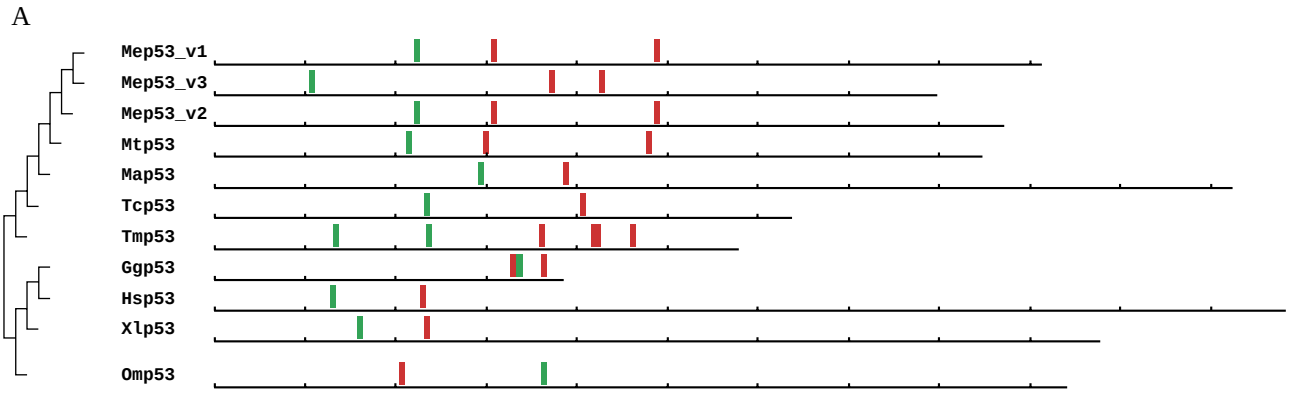
conserved "TGTGAT" K-box motif and found it in *D. rerio p53* 3′UTR, at 327 nt upstream from the poly-A tail. We did not find it in *Drosophila p53* or in *Loligo p53*. This is the first identification of the K-box motif in any species as a potential site of 3′UTR transcriptional regulation in the *p53* gene family. The K-box was originally identified in *Drosophila* Notch signaling proteins E(spl)-C (Lai et al., 1998) where it is loosely associated with a CAAC motif, not present in *Mytilus*, *Mya* and *Danio p53* 3′UTR. Lai et al. (1998) found that the presence of the K-box resulted in a decreased level of mRNA as well as protein in developing *Drosophila* embryos. Mutation or deletion of the K-box motif resulted in over-expression of the reporter construct in all developing tissues. These K-box motifs may be recognized by micro-RNAs, which mediate translational inhibition in *Drosophila* and *C. elegans* (Lai et al., 2004, and references therein). We hypothesize that the K-box motif found in *Mytilus* and *Mya p53* 3′UTR may play a role in regulation of p53 levels in addition to or in concert with other 3′UTR translational elements as well as the MDM2 positive regulatory feedback loop. Phylogenetic conservation of K-boxes in the *p53* gene family remains to be investigated.

The *p53* 3′UTRs of the two *Mytilus* species examined are 93% identical based mostly on point mutations and two regions missing from *M. trossulus* when compared to the full-length *M. edulis* 3′UTR. The point mutation at position 857 eliminates one of the two nuclear polyadenylation sites. However, this may be due to polymerase error and needs to be confirmed in subsequent sequences. The deleted regions do not affect any of the conserved functional regions, but would affect spacing between the regions.

Whether any of the observed differences result in differential expression of the p53 protein in the two species is unclear and may also be due to other factors up- and downstream of *p53* transcription/translation within the regulatory pathways. A widely used antibody against p53 protein (Ab-1, clone Pab421) detects no 53-kDa protein in *M. edulis* normal and leukemic haemocytes but does cross-react with a 120-kDa variant (St-Jean et al., in press). *p53* may be transcribed into pre-mRNA and spliced into mature mRNA at levels high enough to be detected by RT-PCR, but is either not translated into protein or the protein may be degraded or modified at such a rate that it cannot be detected by Western blotting. Interestingly, it was shown that there is an inverse relationship between mRNA levels and protein levels of a p120 (now believed to be a post-translationally modified p63/73 homologue) during embryonic development of the surf clam *Spisula solidissima* (Jessen-Eller et al., 2002). Also, Conne et al. (2000) found that p53 protein is often undetectable in acute myelogenous leukemia (AML) cells in human. This raises the possibility that p53 half-life is altered in AML cells and/or that *p53* gene expression is translationally regulated. The identification of translational regulatory elements in the 3′UTR of *p53* corroborate this hypothesis but would require further detailed investigations to confirm cellular localization, stability and translation of *p53* mRNA.

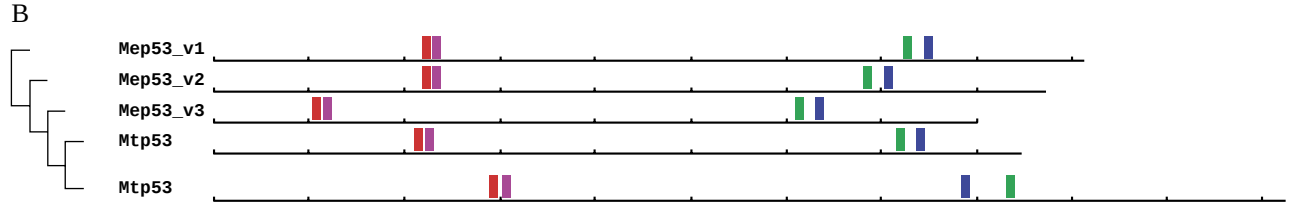
3.6. Phylogenetic footprint analysis of the 3′ UTR

If the *p53* family is indeed of ancient origin, it is possible that functional elements have been conserved in the untranslated regions. Selective pressure would cause functional elements to evolve at a slower rate than that of nonfunctional sequences. Phylogenetic footprinting deduces novel regulatory elements by considering orthologous regions of a single gene from several species (Blanchette and Tompa, 2002). We ran a series of phylogenetic footprint analyses using FootPrinter 2.1 (Blanchette and Tompa, 2003) for a very diverse group of *p53* sequences to potentially identify novel regulatory elements conserved across species. The analysis included *p53* 3′UTR sequences from *Mytilus* spp., clam *M. arenaria*, flour beetle *Tribolium castaneum*, puffer fish *Tetraodon miurus*, trout *Oncorhynchus mykiss*, chicken *Gallus gallus*, frog *X. laevis*, and *H. sapiens* (see Table 2 for accession numbers). The 3′UTR for *Cep-1* and *Ehp53* were not included. We searched for conserved elements of lengths 6–10 in consecutive runs with a parsimony score (= number of mutations) of 2. Only the smallest motif length, 6, returned two conserved motifs across this wide range of species. Fig. 4A gives the location and sequence of the two motifs: tg(t/c)(g/t)tt (green) and tattta (red). The first motif may be a previously unidentified regulatory element in the 3′UTR of *p53* and is also indicated as green background in Fig. 3. Of all the sequences analyzed, only *T. miurus* has a duplication of this motif. The second motif identified by FootPrinter overlaps with the CPE/ARE region identified in Fig. 3 for the *Mytilus* sp. (484, *Mep53*_v1). FootPrinter identified a similar region in *Map53* at position 386 which was not previously identified. The ClustalX alignment (Fig. 3) failed to align the CPEs found in *Mytilus* with potential CPEs in *M. arenaria*, likely due to low sequence conservation and AT-richness of the region. We used the FootPrinter program on the *Mytilidae* and *M. arenaria* to look for highly conserved potential regulatory elements within the mollusc bivalve family. For this, we increased the motif length to 9 and allowed no mutations, because the evolutionary distance is short between *Mytilus* and *Mya*. Two tandem motifs were identified (Fig. 4B): motif 1 (pink) is an unidentified conserved motif adjacent to the highly conserved motif identified in the previous footprint analysis (and is also indicated with a pink background in Fig. 3); motif 2 (green) coincides with the CPE region identified for *Mytilus* sp. in Fig. 3 (722, *Mep53*_v1); motif 3 coincides with the highly conserved motif identified in the previous footprint analysis (and indicated with a green background in Fig. 3); and motif 4 (blue) is an unidentified conserved motif. Motif 2 can be interpreted as a potential CPE in *M. arenaria* at location 833 (green frame in Fig. 3). Our analysis found previously known regulatory elements as well as highly conserved motifs with



Motif #1	Parsimony score: 2.00	
species	position	
MEP53_V1	221	tgtgtt
MEP53_V3	106	tgtgtt
MEP53_V2	221	tgtgtt
MTP53	213	tgtgtt
MAP53	292	tgtgtt
TCP53	233	tgtgtt
TMP53	132	tgtgtt
TMP53	235	tgtgtt
GGP53	335	tgcttt
HSP53	130	tgcttt
XLP53	160	tgcgtt
OMP53	362	tgtgtt

Motif #2	Parsimony score: 2.00	
species	position	
MEP53_V1	308	tattta
MEP53_V1	487	tattta
MEP53_V3	372	tattta
MEP53_V3	426	tattta
MEP53_V2	308	tattta
MEP53_V2	487	tattta
MTP53	299	tattta
MTP53	478	tattta
MAP53	386	tattta
TCP53	406	tattta
TMP53	359	tatttt
TMP53	418	ttttta
TMP53	422	tatgta
TMP53	460	tattca
GGP53	328	ttttta
GGP53	363	ttttta
HSP53	229	ttttta
XLP53	234	tattta
OMP53	205	tattta



Motif #1	Parsimony score: 0.00	
species	position	
MEP53_V1	231	tcattatcg
MEP53_V2	231	tcattatcg
MEP53_V3	116	tcattatcg
MTP53	223	tcattatcg
MAP53	305	tcattatcg

Motif #2	Parsimony score: 0.00	
species	position	
MEP53_V1	725	ttatttctt
MEP53_V2	682	ttatttctt
MEP53_V3	610	ttatttctt
MTP53	716	ttatttctt
MAP53	833	ttatttctt

Motif #3	Parsimony score: 0.00	
species	position	
MEP53_V1	220	atgtgttca
MEP53_V2	220	atgtgttca
MEP53_V3	105	atgtgttca
MTP53	212	atgtgttca
MAP53	291	atgtgttca

Motif #4	Parsimony score: 0.00	
species	position	
MEP53_V1	747	ctttgaaat
MEP53_V2	704	ctttgaaat
MEP53_V3	632	ctttgaaat
MTP53	738	ctttgaaat
MAP53	786	ctttgaaat

Fig. 4. Output for the phylogenetic footprint analysis. The graphic panel shows the phylogenetic tree relating the sequences. Each horizontal line is labeled with the name of the gene (see Table 2 for species identification) and represents the entire 3′UTR sequence. The colored bars above the lines indicate the position of discovered motifs. The bar colors correspond to the font colors in the table below. This table shows the exact sequences and positions of each motif. (A) *p53* 3′UTRs from a diverse range of species. (B) *p53* 3′UTRs from the bivalve species.

unknown function. We found that the FootPrinter results are highly sensitive to the input criteria such as species selection, motif length and parsimony score. During the analysis, we chose a low parsimony score to ensure that the motifs reported were well conserved. Therefore, we may have missed potential functional motifs with a higher degree of variability.

The data presented here contribute to a rapidly emerging story that describes potential transcriptional control of expression in the *p53* gene family. Further analysis of phylogenetically diverse sequences will contribute to this model. Variability in the 3′UTR of molluscan *p53* family members was previously identified by Kelley et al. (2001) who found that unlike in mammals, the molluscan *p73* and

p53 have almost complete identity of the core sequences and suggest that the *p73* is a 3′ gene variant of the *p53* gene, and that a divergence in gene function occurred early in evolution. In addition, Cox et al. (2003) found that two unique polyadenylation site variants may control expression of the *p73* gene in another molluscan species, *S. solidissima*. Taken together, these observations and the data presented here serve to demonstrate the complexities of *p53* gene family regulation and the insights gained by analysis of non-mammalian species.

4. Conclusions

In order for both species of *Mytilus*, *M. edulis* and *M. trossulus*, to be used for coastal or marine environmental effects monitoring, their responses to similar challenges must be quantified. In addition, differences in the natural prevalence, cyclical nature and progression of haemic neoplasia between the two species must be determined to enable differentiation between natural phenomena and anthropogenic effects. We identified and compared the *p53* sequences in both species as a first step in the development of a rapid and repeatable screening tool for haemic neoplasia in Mytilidae. The *p53* cDNA sequences, designated *Mep53* and *Mtp53*, were identified in *M. edulis* and *M. trossulus*, respectively, and found to be 96.5% similar to each other. The coding regions of the *p53* cDNA contain highly conserved regions similar to most identified *p53*. Putative p53 proteins of the two *Mytilus* species are 99.8% similar to each other. The 3′ non-coding region contains a number of known regulatory sequences: adenylate/uridylate-rich elements, cytoplasmic polyadenylation sites, nuclear polyadenylation sites and a K-box motif, which have not been identified previously in a *p53*. Further studies are required to ascertain whether *p53* is differentially expressed or mutated in leukemic haemocytes of *Mytilus*.

Acknowledgements

The authors acknowledge Melissa Kelley and Raymond Stephens for their expertise and review of the manuscript. AFM was supported by the Greater Vancouver Regional District, BC, Canada, and RLC was supported by STAR grant R82935901 from the Environmental Protection Agency (USA). We also thank André Mallet for his assistance in *Mytilus* species identification, as well as three anonymous reviewers.

References

Barker, C.M., Calvert, R.J., Walker, C.W., Reinisch, C.L., 1997. Detection of mutant p53 in clam leukemia cells. Exp. Cell Res. 232, 240–245.

Bhaskaran, A., May, D., Rand-Weaver, M., Tyler, C.R., 1999. Fish p53 as a possible biomarker for genotoxins in the aquatic environment. Environ. Mol. Mutagen. 33, 177–184.

Blanchette, M., Tompa, M., 2002. Discovery of regulatory elements by computational method for phylogenetic footprinting. Genome Res. 12, 739–748.

Blanchette, M., Tompa, M., 2003. Footprinter: a program designed for phylogenetic footprinting. Nucleic Acids Res. 31, 3840–3842.

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

Courtois, S., de Fromentel, C., Hainaut, P., 2004. p53 protein variants: structural and functional similarities with p63 and p73 isoforms. Oncogene 23, 631–638.

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

DeLeo, A.B., Jay, G., Appella, E., Dubois, G.C., Law, L.W., Old, L.J., 1979. Detection of a transformation-related antigen in chemically induced sarcomas and other transformed cells of the mouse. Proc. Natl. Acad. Sci. U. S. A. 76, 2420–2424.

Elston, R.A., Moore, J.D., Brooks, K., 1992. Disseminated neoplasia in bivalve molluscs. Rev. Aquat. Sci. 6, 405–466.

Environment Canada, 2003. National assessment of pulp and paper environmental effects monitoring data: a report synopsis. NWRI Scientific Assessment Report Series No. 2. National Water Research Institute, Burlington, ON. 28 p.

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

Groisman, I., Jung, M.Y., Sarkissian, M., Cao, Q., Richter, J.D., 2002. Translational control of the embryonic cell cycle. Cell 109, 473–483.

Grzybowska, E., Wilczynska, A., Siedlecki, J., 2001. Breakthroughs and views: regulatory functions of 3′UTRs. Biochem. Biophys. Res. Commun. 288, 291–295.

Hofseth, L., Hussain, S., Harris, C.C., 2004. p53: 25 years after its discovery. Trends Pharmacol. Sci. 25, 117–181.

Ishioka, C., Englert, C., Winge, P., Yan, Y.X., Engelstein, M., Friend, S.H., 1995. Mutational analysis of the carboxy-terminal portion of p53 using both yeast and mammalian cell assays in vivo. Oncogene 10, 1485–1492.

Jeanmougin, F., Thompson, J.D., Gouy, M., Higgins, D.G., 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., 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, 377–385.

Kelley, M.L., Winge, P., Heaney, J.D., Stephens, R.E., Farell, J.H., Van Beneden, R.J., Reinisch, C.L., Lesser, M.P., 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, 748–758.

Konikova, E., Kusenda, J., 2003. Altered expression of p53 and MDM2 proteins in hematological malignancies. Neoplasma 50, 31–40.

Krishnakumar, P.K., Casillas, E., Snider, R.G., Kagley, A.N., Varanasi, U., 1999. Environmental contaminants and the prevalence of hemic neoplasia (leukemia) in the common mussel (*Mytilus edulis complex*) from Puget Sound, Washington, U.S.A. J. Invertebr. Pathol. 73, 135–146.

Lai, E., Burks, C., Posakony, J., 1998. The K box, a conserved 3′UTR sequence motif, negatively regulates accumulation of *Enhancer of split* complex transcripts. Development 125, 4077–4088.

Lai, E., Wiel, C., Rubin, G., 2004. Complementary miRNA pairs suggest a regulatory role for miRNA:miRNA duplexes. RNA 10, 171–175.

Lane, D.P., Crawford, L.V., 1979. T antigen is bound to a host protein in SV40-transformed cells. Nature 278, 261–263.

Linzer, D.I., Levine, A.J., 1979. Characterization of a 54K dalton cellular SV40 tumor antigen present in SV40-transformed cells and uninfected embryonal carcinoma cells. Cell 17, 43–52.

Liu, Y., Colosimo, A., Yang, X., Liao, D., 2000. Adenovirus E1B 55-kilodalton oncoprotein inhibits p53 acetylation by PCAF. Mol. Cell Biol. 20, 5540–5553.

Mallet, A., 2003. Species classification of Canadian West coast mussels based on shell morphometry measurements. Mallet Research Services, Dartmouth, NS, Report prepared for the Greater Vancouver Regional District, p. 9.

Mallet, A., Carver, C., 1995. Comparative growth and survival patterns of *Mytilus trossulus* and *Mytilus edulis* in Atlantic Canada. Can. J. Fish. Aquat. Sci. 52, 1873–1880.

May, P., May, E., 1999. Twenty years of p53 research: structural and functional aspects of the p53 protein. Oncogene 18, 7621–7636.

McDonald, J., Seed, R., Koehn, R., 1991. Allozyme and morphometric characters of three species of *Mytilus* in the northern and southern hemispheres. Mar. Biol. (Berlin) 111, 323–335.

McGladdery, S.E., Reinisch, C.L., MacCallum, G.S., Stephens, R.E., Walker, C.L., Davidson, J.T., 2001. Haemic neoplasia in soft-shell clams (*Mya arenaria*): recent outbreaks in Atlantic Canada and discovery of a p53 gene homologue associated with the condition. Bull. Aquacult. Assoc. Canada 101, 19–26.

Mendosa, L., Orozco, E., Rodriguez, M., Garcia-Rivera, G., Sanchez, T., Garcia, E., Gariglio, P., 2003. Ehp53, an *Entamoeba histolytica* protein, ancestor of the mammalian tumor suppressor p53. Microbiology 149, 885–893.

O'Brate, A., Giannakakou, P., 2003. The importance of p53 location: nuclear or cytoplasmic zip code? Drug Resist. Updat. 6, 313–322.

O'Connor, T., Beliaeff, B., 1995. Recent trends in coastal environmental quality results from the mussel watch project. Silver Spring, Md. (1305 East–West Highway, Silver Spring 20910-3281) US Dept. of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, Office of Ocean Resources Conservation and Assessment, Coastal Monitoring and Bioeffects Assessment Division.

Page, R.D.M., 1996. TREEVIEW: an application to display phylogenetic trees on personal computers. Comp. Appl. Biosci. 12, 357–358.

Penny, R.W., Hart, M.J., Templeman, N., 2002. Comparative growth of cultured blue mussels, *Mytilus edulis* and *Mytilus trossulus* and their hybrids, in naturally occurring mixed-species stock. Aquacult. Res. 33, 693–702.

Pesole, G., Liuni, S., 1999. Internet resources for the functional analysis of 5′ and 3′ untranslated regions of eukaryotic mRNA. Trends Genet. 15, 378.

Pesole, G., Mignone, F., Gissi, C., Grillo, G., Licciulli, F., Liuni, S., 2001. Structural and functional features of eukaryotic mRNA untranslated regions. Gene 276, 73–81.

Pesole, G., Liuni, S., Grillo, G., Licciulli, F., Mignone, F., Gissi, C., Saccone, C., 2002. UTRdb and UTRsite: specialized databases of sequences and functional elements of 5′ and 3′ untranslated regions of eukaryotic mRNAs. Nucleic Acids Res. 30, 335–340.

Rawson, P.D., Slaughter, C., Yund, P.O., 2003. Patterns of gamete incompatibility between the blue mussels *Mytilus edulis* and *Mytilus trossulus*. Mar. Biol. 143, 317–325.

Read, R., Norbury, C., 2002. Roles for cytoplasmic polyadenylation in cell cycle regulation. J. Cell. Biochem. 87, 258–265.

Richter, J., 1999. Cytoplasmic polyadenylation in development and beyond. Microbiol. Mol. Biol. Rev. 63, 446–456.

Richter, J., 2001. Think globally, translate locally: what mitotic spindles and neuronal synapses have in common. Proc. Natl. Acad. Sci. U. S. A. 98, 7069–7071.

Salazar, M., Salazar, S., 1997. Using caged bivalves to characterize exposure and effects associated with pulp and paper mill effluents. Water Sci. Technol. 35, 213–220.

Shimizu, H., Hupp, T., 2003. Intrasteric regulation of MDM2. Trends Biochem. Sci. 28, 346–349.

Soot-Ryen, T., 1969. Family mytilidar rafinesque 1815. In: Moore, R. (Ed.), Treatise on Invertebrate Palaeontology, Part N. Mollusca 6. Geological Society of America/University of Kansas Press, pp. 271–280.

Soussi, T., May, P., 1996. Structural aspects of the p53 protein in relation to gene evolution: a second look. J. Mol. Biol. 260, 623–637.

Stephens, R.E., Walker, C.W., Reinisch, C.L., 2001. Multiple protein differences distinguish clam leukemia cells from normal hemocytes: evidence for the involvement of p53 homologues. Comp. Biochem. Biophys. C 129, 329–338.

St-Jean, S.D., Courtenay, S.C., Parker, R.W., 2003. Immunomodulation in blue mussels (*Mytilus edulis*) exposed to a pulp and paper mill effluent in Eastern Canada. Water Qual. Res. J. Canada 38, 647–666.

St-Jean, S.D., Stephens, R.E., Courtenay, S.C., Reinisch, C.L., in press. Detecting p53 family proteins in haemocytic leukemia cells of *Mytilus edulis* from Pictou Harbour, Nova Scotia. Can. J. Fish. Aquat. Sci.

Toro, J., Thompson, R., Innes, D., 2002. Reproductive isolation and reproductive output in two sympatric mussel species (*Mytilus edulis*, *M. trossulus*) and their hybrids from Newfoundland. Mar. Biol. 141, 897–909.

Verrotti, A.C., Thompson, S.R., Wreden, C., Strickland, S., Wickens, M., 1996. Evolutionary conservation of sequence elements controlling cytoplasmic polyadenylation. Proc. Natl. Acad. Sci. U. S. A. 93, 9027–9032.

Vogelstein, B., Lane, D., Levine, A.J., 2000. Surfing the p53 network. Nature 408, 307–310.

Vousden, K., Lu, X., 2002. Live or let die: the cell's response to p53. Nat. Rev. Cancer 2, 594–604.

Widdows, J., Donkin, P., Brinsley, M., Evans, S., Salkeld, P., Franklin, A., Law, R., Waldock, M., 1995. Scope for growth and contaminant levels in North Sea mussels *Mytilus edulis*. Mar. Ecol. Prog. Ser. 127, 131–148.

Wilmer, P., 1990. Invertebrate Relationships: Patterns in Animal Evolution. Cambridge Univ. Press, Cambridge.

Wolkowicz, R., Rotter, V., 1997. The DNA binding regulatory domain of p53: see the C. Pathol. Biol. (Paris) 45, 785–796.

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