The Journal of Experimental Biology 210, 3461-3472
Published by The Company of Biologists 2007
doi:10.1242/jeb.009183

FoxK1 splice variants show developmental stage-specific plasticity of expression with temperature in the tiger pufferfish

Jorge M. O. Fernandes[1,2], Matthew G. MacKenzie[1,3], James R. Kinghorn[1] and Ian A. Johnston[1,*]

[1]*School of Biology, University of St Andrews, St Andrews, Fife KY16 8LB, UK,* [2]*Department of Fisheries and Natural Sciences, Bodø University College, No-8049 Bodø, Norway and* [3]*School of Life Sciences Research, University of Dundee, Dundee DD1 5EH, UK*

*Author for correspondence (e-mail: iaj@st-and.ac.uk)

Accepted 24 July 2007

Summary

FoxK1 is a member of the highly conserved forkhead/winged helix (Fox) family of transcription factors and it is known to play a key role in mammalian muscle development and myogenic stem cell function. The tiger pufferfish (*Takifugu rubripes*) orthologue of mammalian *FoxK1* (*TFoxK1*) has seven exons and is located in a region of conserved synteny between pufferfish and mouse. *TFoxK1* is expressed as three alternative transcripts: *TFoxK1-α*, *TFoxK1-γ* and *TFoxK1-δ*. *TFoxK1-α* is the orthologue of mouse *FoxK1-α*, coding for a putative protein of 558 residues that contains the forkhead and forkhead-associated domains typical of Fox proteins and shares 53% global identity with its mammalian homologue. *TFoxK1-γ* and *TFoxK1-δ* arise from intron retention events and these transcripts translate into the same 344-amino acid protein with a truncated forkhead domain. Neither are orthologues of mouse *FoxK1-β*. In adult fish, the *TFoxK1* splice variants were differentially expressed between fast and slow myotomal muscle, as well as other tissues, and the FoxK1-α protein was expressed in myogenic progenitor cells of fast myotomal muscle. During embryonic development, *TFoxK1* was transiently expressed in the developing somites, heart, brain and eye. The relative expression of *TFoxK1-α* and the other two alternative transcripts varied with the incubation temperature regime for equivalent embryonic stages and the differences were particularly marked at later developmental stages. The developmental expression pattern of *TFoxK1* and its localisation to mononuclear myogenic progenitor cells in adult fast muscle indicate that it may play an essential role in myogenesis in *T. rubripes*.

Key words: forkhead box/winged helix, myocyte nuclear factor (MNF), myogenesis, thermal plasticity, alternative splicing.

Introduction

Myogenesis is a complex process involving the commitment of pluripotent stems cells to a myogenic fate, orchestrated by the MyoD family of basic helix-loop-helix transcription factors *Myf5*, *MyoD* (also known as *Myod1*; Mouse Genome Informatics) and *MRF4* (also known as *Myf6*; Mouse Genome Informatics) (Kassar-Duchossoy et al., 2004; Rudnicki and Jaenisch, 1995). Muscle progenitor cells (MPCs) represent a self-renewing population, the progeny of which undergo further cell division before exiting the cell cycle and entering a programme of terminal differentiation that involves the upregulation of the MyoD family members *Myog* and *MRF4*, leading to the expression of muscle-specific proteins (Andres and Walsh, 1996; Kassar-Duchossoy et al., 2004). In the mouse, proliferating myoblasts withdraw from the cell cycle and fuse to form multinucleated myotubes leading to primary and secondary muscle fibres in the embryonic and foetal stages, respectively (Kelly and Rubinstein, 1994). Although muscle fibre number changes little after birth (Rowe and Goldspink, 1969), the MPCs provide additional myonuclei as the muscle fibres in the growing animal increase in length and diameter (Moss and Leblond, 1971). MPCs are also involved in nuclear turnover in adult muscle and have a role in repair from injury and adaptive responses to exercise (reviewed by Chen and Goldhamer, 2003). The satellite cells found beneath the basal lamina of adult skeletal muscle represent myogenic progenitors that are mitotically quiescent (Schultz, 1996). Quiescent and activated MPCs express paired box protein 7 (*Pax7*) (Seale et al., 2000), the expression of which is downregulated once the cells express *MRF4* and *Myog* and enter the programme of terminal differentiation (Chen and Goldhamer, 2003).

The FoxK1 protein is a member of the forkhead/winged helix family (Fox) of transcription factors, which constitute a diverse group that display a remarkable diversity and play crucial functions in several biological processes, including development and oncogenesis (Carlsson and Mahlapuu, 2002). All members of the Fox family are characterised by the presence of a forkhead (FH) domain, a 110-residue DNA binding region that consists of three α-helices and three β-strands flanked by two wing-like loops, resulting in a three-dimensional structure that resembles the wings of a butterfly (Clark et al., 1993). The structure of FoxK1 is rather unusual amongst Fox proteins, in that one of the typical wings is replaced by an 8-residue C-terminal α-helix (Chuang et al., 2002). FoxK1 also contains a

forkhead-associated (FHA) domain, which is a phosphopeptide recognition region (Durocher and Jackson, 2002). Since the identification of the homeotic gene *forkhead* in *Drosophila* (Weigel et al., 1989) more than 150 members of the Fox family have been found in taxa as diverse as yeast and mammals; this family is now divided into 17 subgroups designated A to Q (Kaestner et al., 2000). Relatively few members of the Fox family have been characterised with respect to their function and target genes. During mouse embryogenesis FoxK1 (also known as Foxk1; Mouse Genome Informatics) was detected transiently in the developing myotome, limb precursors, heart tube and certain regions of the brain (Garry et al., 1997). Knockout mice with a functional null allele at the *FoxK1* locus showed impaired satellite cell function, which resulted in delayed and incomplete skeletal muscle regeneration following injury (Garry et al., 2000). The two alternatively spliced isoforms of *FoxK1* (*FoxK1-α* and *FoxK1-β*) were found to have reciprocal expression patterns during muscle regeneration, indicating that they might exert opposite effects on their target genes (Garry et al., 2000). *FoxK1-β* expression predominates in quiescent satellite cells, whereas *FoxK1-α* is the main isoform expressed in proliferating myoblasts derived from activated satellite cells (Garry et al., 2000). *FoxK1* is one of the few known markers of quiescent satellite cells in mammalian muscle (Garry et al., 1997).

In teleosts, at least three phases of myogenesis can be distinguished in fast myotomal muscle: an embryonic phase, stratified hyperplasia from distinct germinal zones and mosaic hyperplasia (Johnston, 2006). The majority of muscle fibres are formed by mosaic hyperplasia in the larval, juvenile and early adult stages involving the activation of MPCs throughout the myotome (Rowlerson and Veggetti, 2001). The continuation of myotube production in adult fish reflects the large increase in body mass between the larvae and the final body size (Johnston, 2006). Teleosts are ectothermic and the outcome of the myogenic programme is profoundly affected by epigenetic factors, particularly embryonic temperature (Johnston and Hall, 2004). For example, in Atlantic herring (*Clupea harengus* L.) heterochronic shifts were observed with respect to the rostral to caudal progression of myofibril assembly and the outgrowth of primary motor neurons, which started at earlier somite stages as temperature was increased (Johnston et al., 1995). Embryonic temperature has been shown to influence the number of MPCs and muscle fibre recruitment patterns in herring (Johnston, 1993; Johnston et al., 1998), Atlantic salmon (*Salmo salar* L.) (Johnston et al., 2000a; Johnston et al., 2000b) and European sea bass (*Dicentrarchus labrax* L.) (Alami-Durante et al., 2007), indicating that temperature experienced during early development has a lasting influence on adult muscle phenotype. In Atlantic salmon, the maximum fibre number in seawater stages varied by up to 20% according to the temperature regime experienced during early development (Johnston et al., 2003).

The availability of a draft genome sequence of the tiger pufferfish (*Takifugu rubripes* Temminck and Schlegel) (Aparicio et al., 2002) provides an excellent opportunity for investigating the molecular basis of developmental plasticity of myogenesis in teleosts. The aim of the present study was to characterise the *FoxK1* gene and its splice variants in *T. rubripes* and to test the hypothesis that its expression with respect to embryonic stage was a function of developmental temperature.

Materials and methods

Animals and sample collection

Two juvenile tiger puffer fish *Takifugu rubripes* Temminck and Schlegel, with an approximate mass of 160 g, were bred in captivity at the Fisheries Laboratory (University of Tokyo, Maisaka, Shizuoka Prefecture, Japan). Two wild-caught adult specimens, weighing ~1.4 kg, were purchased from the local fish market at Maisaka (Shizuoka Prefecture, Japan). Fish were humanely killed according to the British Home Office guidelines by over-anaesthesia in a solution of 0.2 mmol l^{-1} 3-aminobenzoic acid ethyl ester (Sigma, Gillingham, Dorset, UK) buffered with sodium bicarbonate (Sigma). Samples of fast and slow myotomal muscle, heart, liver, skin, brain and gonads (adult fish only) were dissected and stored in RNAlater (Ambion/Applied Biosystems, Warrington, Lancashire, UK) for subsequent RNA extraction. *T. rubripes* eggs were purchased from a commercial source (Nisshin Marinetech Co., Yokohama, Japan). The eggs from a single female were fertilised at 17°C using the sperm of two males and, after approximately 4 h, they were transferred to the Fisheries Laboratory. Embryos were split into three temperature groups and incubated at either 15°C, 18°C or 21°C (within ±0.5°C). After hatching, the temperature of all tanks was gradually increased to 18°C and larvae were reared for approximately 2 months. Samples of embryos and larvae collected throughout development were preserved in RNA later.

RNA extraction and cDNA synthesis

Tissue samples and eggs (0.1 g) were placed in 1 ml Tri reagent (Sigma) and homogenised with FastRNA Pro Green beads (Qbiogene Inc., Cambridge, UK) using the FastPrep Instrument (Qbiogene) for 40 s at a speed setting of 6.0. Total RNA was isolated according to the manufacturer's protocol. Following DNase treatment (Turbo DNA-free; Ambion) to remove any potential genomic DNA contamination, RNA quality was verified by electrophoresis on a 1% (m/v) agarose (Bioline, London, UK) gel under denaturing conditions. Total RNA was then quantified with the fluorescent nucleic acid stain RiboGreen (Molecular Probes/Invitrogen, Paisley, UK), according to the instructions provided by the manufacturer. First-strand cDNA was synthesised from 1 μg of total RNA using a RETROscript kit (Ambion), according to the recommended method. A 1:1 mixture of random decamers and oligo$(dT)_{18}$ was used as first-strand primers for cDNA synthesis. Following denaturation of the RNA by incubation at 85°C, the reverse transcription was performed at 50°C. A negative control lacking reverse transcriptase was included.

FoxK1 cDNA cloning

The protein sequences available for mouse FoxK1 (NP_951031 and NP_034942) were used for TBLASTN similarity searches against the third assembly of the *T. rubripes* genome (available at http://www.ensembl.org/Fugu_rubripes/index.html), using a BLOSUM80 matrix, a word size of four and a maximum expected value cut-off equal to 1×10^{-5}. This prediction of the

Table 1. *List of primer pairs used to amplify FoxK1 from Takifugu rubripes*

Primer pair	Forward primer (5′→3′)	Reverse primer (5′→3′)	Amplicon size (bp)
cDNA-FoxK1	GGACGACACTGGAGCAAGA	GGGGCTCATTTTTGATCAAG	1676
ISH-FoxK1-α	GAGTCCAAGCCACCTTATTC	GGGGCTCATTTTTGATCAAG	823
ISH-FoxK1-γ	GAGTCCAAGCCACCTTATTC	CTGGAGCGCTCTGGGAGTAC	703
ISH-FoxK1-δ	GAGTCCAAGCCACCTTATTC	CTGGAGCGCTCTGGGAGTAC	940
qPCR-FoxK1-α	GCTGGCAGAACTCCATCAGACAC	TCGCTGCCGCCTTTTCCTG	158
qPCR-FoxK1-γ	CAGAACTCCATCAGACACAACC	CAGGGGCGCTTCTGGAGG	208
qPCR-FoxK1-δ	CGTGACTCTGCCGGACTCTG	CTGACCAAGTAGCACAGCAAGATC	174

cDNA-FoxK1: sequences of the primers used to initially amplify the *T. rubripes* orthologue of *FoxK1*. ISH primer pairs: sequences used to subclone the three *TFoxK1* splice variants for *in situ* hybridisation. qPCR primer pairs: sequences used to quantify expression of the alternative *TFoxK1* transcripts by real-time PCR.

T. rubripes orthologue of *FoxK1* was refined manually and specific primers were designed. Sequences of the primers (Invitrogen, Paisley, UK) used for cDNA amplification (cDNA-FoxK1) are shown in Table 1. *FoxK1* was amplified by PCR using cDNA obtained from 2-month post-hatch larvae. The 25 μl reaction mixtures for PCR amplification contained 1 μl cDNA template, 40 nmol of each primer, 0.1 μmol l^{-1} dNTPs, 1× PCR buffer (Amersham, Amersham, Buckinghamshire, UK) and 1 i.u. Taq DNA polymerase (Amersham). Amplification reactions were performed on a Genius thermocycler (Techne, Duxford, Cambridgeshire, UK) as follows: initial denaturation at 95°C for 3 min, 35 cycles of denaturation for 30 s at 95°C, annealing at 56°C for 30 s and extension at 72°C for 1 min and one final extension for 10 min at 72°C. PCR products were analysed by electrophoresis on a 1.2% agarose gel in modified Tris–acetate–EDTA buffer (Millipore, Billerica, MA, USA) and extracted from the gel using the Montage gel nebuliser system (Millipore). The purified PCR products were ligated to a pCR4-TOPO T/A vector (Invitrogen), which was then used to transform chemically competent TOP10 *Escherichia coli* cells (Invitrogen).

DNA sequencing

Sequencing reactions of the plasmid clones were performed in both directions with T3 or T7 primers and the DNA was sequenced with an ABI PRISM 377 DNA Sequencer (Applied Biosystems, Warrington, UK) at the Dundee Sequencing Service (University of Dundee, UK).

Sequence analyses

BLASTX similarity searches of the *T. rubripes FoxK1* sequences were performed against the complete non-redundant GenBank database (http://www.ncbi.nlm.nih.gov/BLAST/) using the default parameters. Following translation of the *T. rubripes FoxK1* nucleotide sequences using DNAman (Lynnon Biosoft, Quebec, Canada), the Conserved Domain Database and Search Service (v2.04) at NCBI was used to identify conserved domains in the predicted protein sequences (Marchler-Bauer and Bryant, 2004). The putative FoxK1 proteins from *T. rubripes* were aligned with their orthologues from mouse (NP_951031, NP_034942) and zebrafish (NP_956196) using ClustalW (Thompson et al., 1994) on the BioEdit sequence alignment editor (Hall, 1999). A sequence identity matrix between these sequences was also obtained with BioEdit (Hall, 1999). For genomic sequence analyses, data were obtained from the current Ensembl assemblies of the *T. rubripes*, zebrafish and mouse genomes (http://www.ensembl.org/). Genomic organisation of *FoxK1* was determined by comparison of cDNA and genomic sequences using the alignment program Spidey (Wheelan et al., 2001). The structure of donor and acceptor splice sites in *T. rubripes FoxK1* was analysed using the Splice Site Prediction by Neural Network (Reese et al., 1997). The MartView data mining tool (http://www.ensembl.org/Multi/martview) was used to identify the genes present in a 100 kb region either side of the *T. rubripes FoxK1* locus and the corresponding orthologues in mouse and zebrafish.

Whole-mount in situ hybridisation

The three splice variants of *T. rubripes FoxK1* were subcloned using the primers listed in Table 1 and the method described above, in order to obtain cDNA clones of suitable size. *FoxK1-α*, *FoxK1-γ* and *FoxK1-δ* DNA templates for probe synthesis were obtained from the corresponding pCR4-TOPO plasmids by PCR using standard M13 primers and the thermocycling conditions described above. T7 and T3 RNA polymerases (Roche, East Sussex, UK) were used to synthesise digoxigenin (DIG)-labelled RNA probes by *in vitro* transcription, according to the manufacturer's protocol. Sense probes were used as negative controls. Whole-mount *in situ* hybridisation with DIG-labelled probes for *FoxK1-α*, *FoxK1-γ* and *FoxK1-δ* was performed essentially as described previously (Fernandes et al., 2006), with minor modifications. DIG-labelled probes for detection of *FoxK1-α* shared 44% and 46% identity with those for *FoxK1-γ* and *FoxK1-δ*, respectively, whereas *FoxK1-γ* and *FoxK1-δ* DIG-labelled probes were 74% identical. These differences in probe sequences should permit the specific detection of each splice variant of FoxK1. For each selected developmental stage, five *T. rubripes* embryos, reared at 18°C, were used. Optimal permeabilisation of *T. rubripes* embryos was achieved by incubation at 20°C with 20 μg ml^{-1} proteinase K (Roche) for 5, 10 and 15 min for pre-somite, segmentation and post-somitogenesis stages, respectively. Whole embryos and flat-mounted embryos were observed under a binocular microscope (Leica MZ7.5, Milton Keynes, UK) and Leitz DMRB microscope (Leica) with DIC optics, respectively, and images were acquired with a Nikon Coolpix 4500 digital camera (Surrey, UK).

Quantitative real-time PCR (qPCR)

T. rubripes embryos incubated at 15°C, 18°C or 21°C were collected at different developmental stages. Tissues from juvenile and adult stages were collected as previously described. Total RNA extraction and cDNA synthesis were performed as described above. PrimerSelect software (DNAStar Inc., Madison, USA) was used to design specific *FoxK1-α*, *FoxK1-γ* and *FoxK1-δ* primer pairs (Table 1). Quantitative real-time PCR (qPCR) was carried out using an ABI Prism 7000 instrument (Applied Biosystems) with SYBR Green reagents (QuantiTect SYBR Green PCR, Qiagen, Crawley, West Sussex, UK), as recommended by the manufacturer. The 25 μl reaction mixtures contained 1 μl cDNA (diluted 1:5), 0.4 μmol l^{-1} each primer and 1× QuantiTect SYBR Green PCR master mix. PCR amplification of target genes was performed in duplicate using the following thermal profile: initial activation at 95°C for 15 min followed by 40 cycles of 15 s at 94°C, 30 s at 56°C and 30 s at 72°C. After each run, a dissociation protocol with a gradient from 60°C to 90°C was used to ascertain the specificity of the primers. ROX was used as passive dye for normalisation of SYBR Green fluorescence. *RNA polymerase II* was used as internal standard, since it was a more stable housekeeping gene than 18S rRNA or *elongation factor 1α*. The primer pair used to amplify 171 bp from the large subunit of *RNA polymerase II* was: 5′-CAGCCCAGATGAACTTAAACGG-3′ (forward) and 5′-CCAGGACACTCTGTCATGTTGC-3′ (reverse). Threshold cycle values (C_T) were determined with the 7000 System Sequence Detection Software (Applied Biosystems) using an arbitrary threshold of 1 and a baseline set between 6 and 15 cycles. Standard curves for each gene were obtained by amplifying fivefold serial dilutions (ranging from 1:5 to 1:625) of a reference mixture containing equal amounts of cDNA from each sample. These standard curves were used to estimate the PCR efficiency of each amplicon, using the Relative Expression Sofware Tool (REST) (Pfaffl et al., 2002). C_T values were converted into relative expression levels according to the mathematical model proposed by Pfaffl (Pfaffl, 2001). Statistical analysis of *TFoxK1* expression during development at different incubation temperatures was performed on SPSS 12.0 (SPSS Inc., Chicago, USA), using a general linear model with stage, splice variant and temperature as fixed factors. The Bonferroni test was used for *post-hoc* multiple comparisons between categories. Differences in tissue distribution of the three *TFoxK1* splice variants were investigated by two-way ANOVA with Holm–Sidak *post-hoc* tests using the SigmaStat statistical package (Systat software, London, UK). In all instances significance levels were set at $P<0.05$.

Antibody production and immunohistochemistry

Anti-FoxK1 polyclonal antibodies were prepared against a synthetic peptide designed from the putative translation of *T. rubripes FoxK1-α*. The antibody's epitope was located in a conserved region near the carboxyl terminus with a high antigenic index, as determined from the antigenicity plot of TFoxK1-α (http://bioinformatics.org/JaMBW/3/1/7/index.html). The peptide antigen Tyr-Arg-Tyr-Ser-Gln-Ser-Ala-Pro-Gly-Ser-Pro-Val-Ser-Ala-Gln-Pro-Val-Ile-Met was commercially synthesised and coupled to keyhole limpet haemocyanin. This hapten-carrier conjugate was used to immunise rabbits for antisera production, following a standard immunisation schedule (Cambridge Research Biochemicals, Durham, UK).

Myonuclei expressing FoxK1-α were identified by immunohistochemistry using the polyclonal antibody described above. Transverse sections of fast myotomal muscle from adult *T. rubripes* were stained for FoxK1-α according to the method described by Johnston et al. (Johnston et al., 2004). The primary anti-FoxK1-α antibody was used at a final concentration of 1:1500, whereas the anti-rabbit IgG-biotin conjugate was diluted 1:800.

Results

Identification of the FoxK1 gene

The best match to mouse *FoxK1* was found on scaffold 40 of the FUGU 4.0 *T. rubripes* genome assembly (P=4.1 e^{-152}). Two other significant matches with P values of 8.6 e^{-144} and 4.0 e^{-55} were also identified on scaffolds 443 and 847, respectively, but the predicted transcripts did not correspond to *FoxK1*. Unexpectedly, PCR amplification using primers designed from the sequence predicted on scaffold 40 yielded three specific products of different sizes. These cDNA sequences had open reading frames of 1676, 1868 and 2105 bp, which shared an overall 79.5% identity. BLASTX searches revealed that the three *T. rubripes* transcripts were most similar to mouse *FoxK1*, and were therefore designated *TFoxK1*. *TFoxK1-α* (GenBank: AY566278) coded for a putative protein of 558 residues that contains the FH domain typical of Fox proteins (Fig. 1). *TFoxK1-γ* (GenBank: AY566280) and *TFoxK1-δ* (GenBank: AY566279) translated into the same 344-residue protein with a partially truncated FH domain (Fig. 1). The first 332 amino acids of TFoxK1-γ and TFoxK1-δ were identical to those of TFoxK1-α, as was the FHA domain (Fig. 1). The C-terminal residues VGPFWLKLNALQ present only in TFoxK1-γ and TFoxK1-δ (Fig. 1) were encoded by the cDNA sequence 5′-GTGGGCCCATTCTGGCTGAAACTTAATGCTTTGCAA-3′. TFoxK1 showed a high degree of global conservation with the other vertebrate FoxK1 proteins (Fig. 1). TFoxK1-α shared 53% identity with its homologues in mouse and zebrafish. Within the FHA domain, TFoxK1-α showed 74% and 84% identity to the mouse and zebrafish sequences, respectively. The FH domain was particularly well conserved, displaying 92% identity between TFoxK1-α and the mouse and zebrafish FoxK1 proteins. Zebrafish and mouse FoxK1 shared 84% and 99% of their residues in the FHA and FH domains, respectively.

Genomic organisation and synteny analysis of FoxK1

FoxK1 genes in *T. rubripes* and zebrafish spanned 6.4 kb and 30.9 kb, respectively, and the difference in size was attributable to larger introns in the zebrafish gene (Fig. 2). Introns I, II and VIII in zebrafish *FoxK1* were particularly long, at 4.7 kb, 10.5 kb and 3.7 kb, respectively. The structure of the *TFoxK1* gene (partial coding sequence) consisted of seven exons and six introns, whereas the zebrafish orthologue had nine exons and eight introns, including a 463 bp untranslated region (UTR) (Fig. 2). The complete coding sequence of zebrafish *FoxK1* (*DFoxK1*; also known as *foxk1* – Zfin) was slightly larger (1920 bp) and shared 58% identity with *TFoxK1*. The splice sites of exon–intron boundaries for the first six exons were conserved between *TFoxK1* and zebrafish *FoxK1*. However, no

evolutionary conserved intronic regions (i.e. over 60% identity over 100 bp) were detected between the *TFoxK1* and zebrafish *FoxK1* genes. Comparison between the *TFoxK1* cDNA and genomic sequences clearly revealed that *TFoxK1-α*, *TFoxK1-γ* and *TFoxK1-δ* arise from alternative splicing events. Is it unlikely that the alternative transcripts identified in this study represent RT-PCR artefacts, since identical splice variants were identified in repeated amplifications from independent samples and splice variant-specific primers were used for the real-time PCR assays. The alternative transcript *TFoxK1-δ* is the largest, for it contains both introns IV and V, whereas the splice variant γ resulted from retention of intron IV (Fig. 2). Most of the donor and acceptor splice sites of *TFoxK1* had good matches to the consensus sequences and were considered strong splice sites (Table 2). However, donor sites 3 and 4 had significantly lower scores than the mean donor score for all exons (0.86) and the acceptor score for exon 2 was significantly lower than the mean acceptor value of 0.90 (Table 2).

FoxK1 is located on scaffold 40, chromosome 3, and chromosome 5 band G2, of the *T. rubripes*, zebrafish and

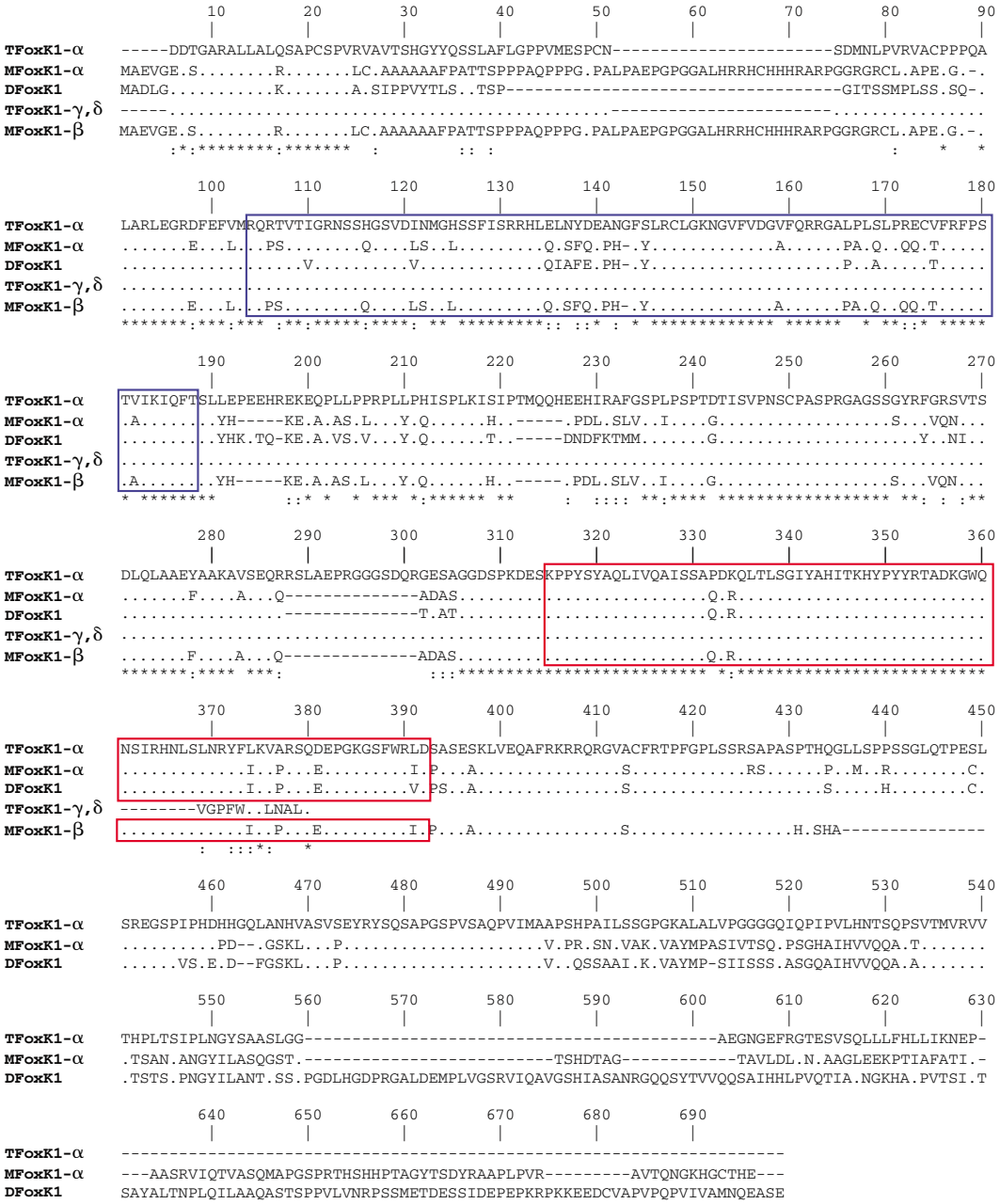

Fig. 1. Multiple sequence alignment of FoxK1 protein sequences from *T. rubripes* (TFoxK1) and their orthologues from mouse (MFoxK1) and zebrafish (DFoxK1). The accession numbers are as follows: TFoxK1-α (GenPept: AAS68037), TFoxK1-γ (GenPept: AAS68039), TFoxK1-δ (GenPept: AAS68038), MFoxK1-α (GenPept: NP_951031), MFoxK1-β (GenPept: NP_034942) and DFoxK1-α (GenPept: NP_956196). Amino acid residues identical to TFoxK1-α are represented by a dot, gaps are indicated by a dash, global identity is marked by an asterisk and conserved substitutions are represented by a colon. The forkhead and forkhead-associated domains are indicated by red and blue rectangles, respectively.

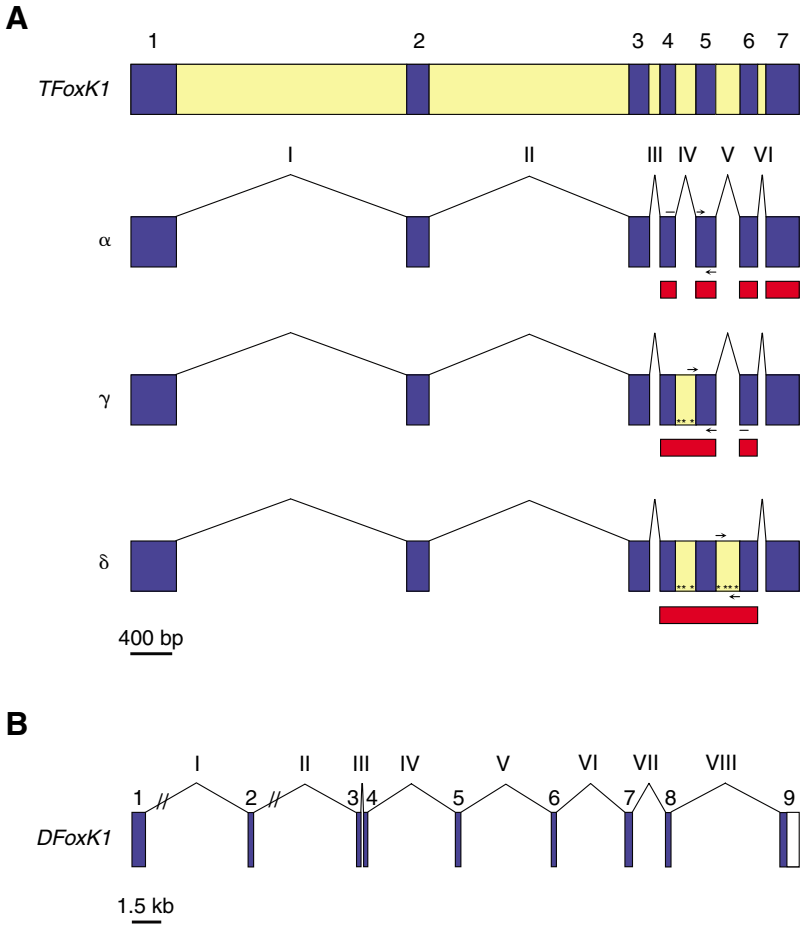

Fig. 2. Structure of the *FoxK1* transcripts from *T. rubripes* and zebrafish. (A) Gene structure of *FoxK1* from *T. rubripes* (*TFoxK1*) and schematic representation of its three alternatively spliced transcripts (α, γ and δ). Exons (blue boxes) are indicated by Arabic numbers and introns labelled with Roman numerals; introns, yellow boxes; stop codons, asterisks. The positions of the primers used for qPCR are indicated by arrows (not to scale). Red bars represent the location of the riboprobes used for *in situ* hybridisation. Scale bar, 400 bp. (B) Exon–intron structure of zebrafish *FoxK1* (*DFoxK1*; GenBank: NM_199902). Scale bar, 1.5 kb. Introns I and II of *DFoxk1* are not depicted to scale; the size of these introns is 4.7 kb and 10.5 kb, respectively. The last exon comprises a 463 kb untranslated region (white rectangle).

mouse genomes, respectively. Comparative mapping of the genes surrounding *TFoxK1* showed that it lies within a region of conserved synteny between *T. rubripes*, zebrafish and mouse, thus confirming that it is the true orthologue of *FoxK1* in these species (Fig. 3A,B). The genes in the vicinity of *TFoxK1* consisted of *abcg1* (white protein homologue), *arpc1a* and *arpc1b* (actin-related protein 2/3 complex subunits a and b), *mmd2* (monocyte to macrophage differentiation factor 2), *slipr* (scaffolding protein SLIPR), *cyp3a* (cytochrome P450), *bat4* (G5 protein), *sdk1* (sidekick homologue 1) and two predicted genes coding for hypothetical proteins, herein designated *hyp1* and *hyp2* (Fig. 3). With the exception of *bat4*, which was located on chromosome 17, the mouse orthologues were present in a 3.3 Mb syntenic region (Fig. 3A). Local gene inversions could be observed in the region surrounding *FoxK1* in mouse. Synteny conservation was not as prominent between *T. rubripes* and zebrafish, since most of the *T. rubripes* genes had orthologues on zebrafish chromosome 1 (Fig. 3B). *hyp1* was duplicated in zebrafish and one of its copies was inverted. Only *arpc1a* and *arpc1b* were present on the same chromosomal segment as *FoxK1*, and these genes were inverted in relation to the *T. rubripes* orthologues. *T. rubripes cyp3a* and *hyp2* did not have orthologues in either mouse or zebrafish in this region.

Transient expression of TFoxK1 during embryonic development

Expression of *TFoxK1-α* in *T. rubripes* embryos incubated at 18°C was not detected during gastrulation (Fig. 4A) but it showed a dramatic increase at the early stages of the segmentation period [63 h post-fertilisation (h.p.f.) 2–6 somites; Fig. 4B]. At this stage, *TFoxK1-α* was expressed in the cephalic region and developing somites. At the 5- to 6-somite stage, *TFoxK1-α* expression was found in pairs of symmetrical bands on either side of the notochord (Fig. 4B). Lower levels of *TFoxK1-α* transcripts were also detected in the pre-somitic paraxial mesoderm. Rostrocaudal progression of the staining

Table 2. *Exon–intron boundary sequences and splice site strength parameters of T. rubripes FoxK1-α*

Exon	Size (bp)	Donor (3′ boundary) Sequence	Score	Acceptor (5′ boundary) Sequence	Score
1	435	CCAGAGA**gt**aagtgc	0.99		
2	216	CCATCAG**gt**gagaag	0.99	ttttct**ag**GTGTGTG	0.99
3	199	CCCAAAG**gt**gaagtc	0.54*	tccttc**ag**CGTTCCC	0.63*
4	147	CTGGCAG**gt**gggccc	0.64*	ttttcc**ag**GATGAGT	0.99
5	194	CCTCCAG**gt**gagtgg	0.99	cttttc**ag**AACTCCA	0.89
6	173	GCTCCAG**gt**gagtct	1.00	ttctac**ag**AAGCGCC	0.99
7	312			ctccgt**ag**GATCTCC	0.92

Exonic and intronic sequences are represented in upper and lower case, respectively. Exon–intron boundaries are highlighted in bold. *Splice site scores that are significantly lower than the mean value (P=0.02).

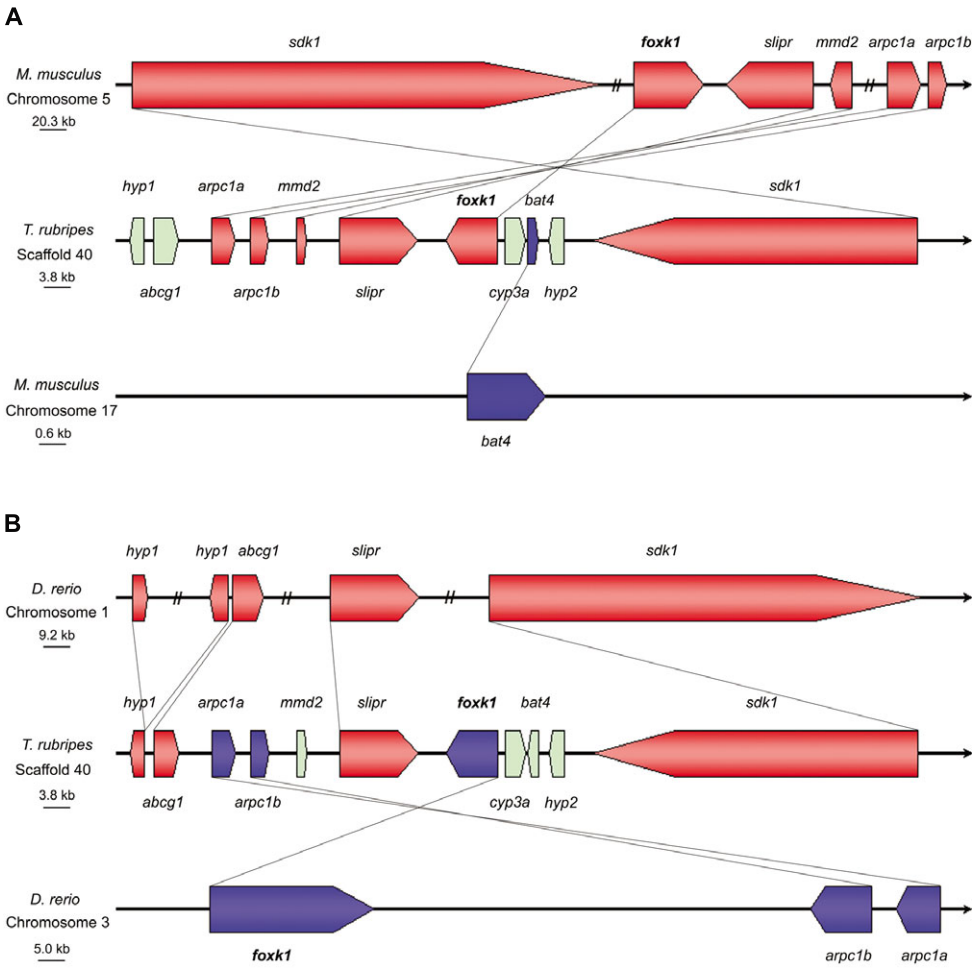

Fig. 3. Partial synteny map of the genomic neighbourhood of *FoxK1* in *T. rubripes*, zebrafish and mouse. Synteny between *TFoxK1* located in scaffold 40 and the corresponding regions in chromosomes 5 and 17 of mouse (A) and chromosomes 1 and 3 of zebrafish (B). Genes are represented by pointed coloured bars. Broken lines drawn between chromosomes indicate homology between individual genes. Genes that are coloured light green have no known orthologues in the genomic regions shown.

throughout segmentation was not observed. *TFoxK1-α* expression was downregulated as segmentation progressed and approximately 13 h later (76 h.p.f., 10- to 12-somite stage) it was limited to the optic vesicles, the developing midbrain and a subset of cells flanking the notochord (Fig. 4C). Approximately midway through somitogenesis (16-somite stage, 86 h.p.f.), *TFoxK1-α* transcripts could be detected in the lens and retina of the developing eyes, in the tubular heart and in the rhombo-mesencephalic fissure that marks the boundary between the midbrain and the hindbrain (Fig. 4D,E). *TFoxK1-α* expression declined rapidly after this stage and by the end of the segmentation period it was no longer detectable. The spatial and temporal distribution of *TFoxK1-γ* and *TFoxK1-δ* was also investigated by whole-mount in situ hybridisation. No qualitative differences in the developmental expression patterns of *FoxK1-α*, *TFoxK1-γ* and *TFoxK1-δ* were apparent (data not shown).

Differential expression of TFoxK1 splice variants with embryonic temperature

The expression levels of the three *TFoxK1* splice variants during development of embryos incubated at 15, 18 or 21°C were determined by real-time PCR. Data were represented

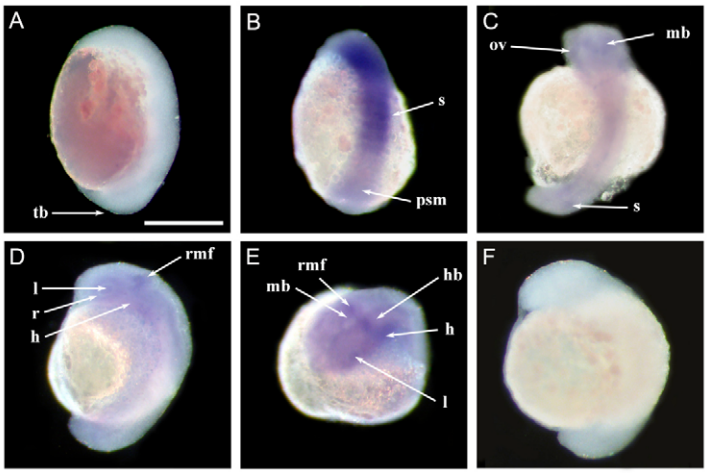

Fig. 4. Spatial expression pattern of *TFoxK1-α* in *T. rubripes* embryos during development at 18°C. *TFoxK1-α* transcripts were detected by whole-mount *in situ* hybridisation with a DIG-labelled cRNA probe. The different developmental stages are (A) 57 h.p.f., bud stage of gastrulation, (B) 63 h.p.f., 5-somite stage, (C) 76 h.p.f., 10-somite stage, (D,E) 86 h.p.f., 16-somite stage. A sense (negative) control of a 10-somite stage embryo is shown in F. Anterior is top. Scale bar, 0.5 mm. h, tubular heart; hb, hindbrain; l, lens; mb, midbrain; ov, optic vesicle; psm, pre-somitic mesoderm; r, retina; rmf, rhombo-mesencephalic fissure; s, somites; tb, tail bud.

as ratios in relation to the expression level of *TFoxK1-α* during mid-gastrulation at 18°C, following normalisation with RNA polymerase II as internal standard. At each developmental stage there were differences in the relative proportions of *TFoxK1-α*, *TFoxK1-γ* and *TFoxK1-δ* (Fig. 5) and the splicing pattern changed throughout development ($P<0.001$). *TFoxK1-δ* was the least expressed of the three alternative transcripts at any stage and temperature (Fig. 5C). The effect of temperature on expression of *TFoxK1* splice variants depended on the developmental stage ($P=0.04$). *TFoxK1-α* and *TFoxK1-γ* had similar expression patterns and their highest transcript levels at 15°C were observed at the onset of somitogenesis. By contrast, *TFoxK1-α* and *TFoxK1-γ* expression at 21°C was higher at the end of the segmentation period and maximal at the hatching stage. Differences in relative amounts of *TFoxK1-γ* and *TFoxK1-α* were greater at 21°C than 18°C or 15°C, and particularly striking at later developmental stages (Fig. 5A,B). At the hatching stage, the expression ratios between *TFoxK1-γ* and *TFoxK1-α* were 0.9, 1.2 and 2.4 at 15, 18 and 21°C, respectively.

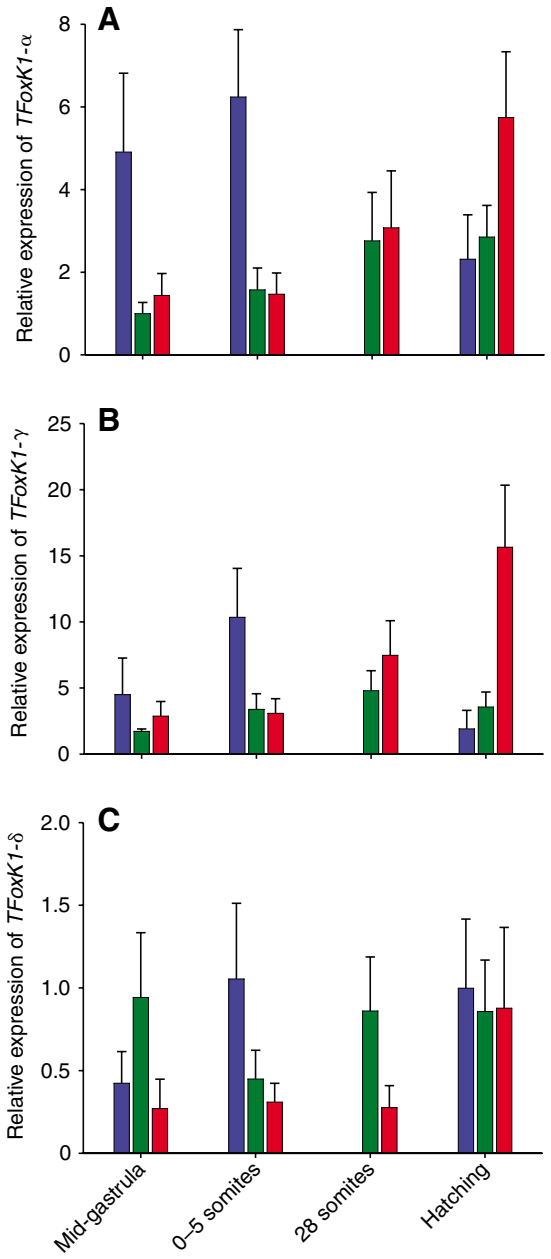

Fig. 5. Influence of embryonic temperature on expression of *TFoxK1* splice variants. Relative expression of *TFoxK1-α, TFoxK1-γ* and *TFoxK1-δ* during development of *T. rubripes* embryos at 15°C (blue bars), 18°C (green bars) or 21°C (red bars). The different developmental stages are indicated on the bottom *x* axis. Unfortunately, the 28-somite-stage samples from embryos incubated at 15°C were not available, owing to technical difficulties. For each data point, *TFoxK1* mRNA levels measured by qPCR were normalised (using RNA polymerase II as standard) and expressed as ratios in relation to the expression level of *TFoxK1-α* during mid-gastrulation at 18°C. Values are means ± s.e.m. ($N=3$).

Tissue distribution of TFoxK1 splice variants in adult fish

Analysis of variance of qPCR results showed that the three *TFoxK1* splice variants were differentially expressed in different tissues of *T. rubripes*, but no significant differences in expression levels were found between juvenile and adult fish, representing growth stages in which myotube production was active or inhibited, respectively (Fernandes et al., 2005). The three *TFoxK1* alternative transcripts were expressed in all tissues examined, including cardiac muscle (H) and fast (WM) and slow (RM) myotomal muscle (Fig. 6). *TFoxK1-α* and *TFoxK1-γ* were found to be the most abundant splice variants, with expression levels 10- to 20-fold higher than those of *TFoxK1-δ*. In order to investigate the localisation of TFoxK1-α protein in adult fast myotomal muscle a specific antibody was constructed. TFoxK1-α protein was exclusively expressed in mononuclear cells (arrowheads, Fig. 6 inset) corresponding to the MPCs, which represented 3–5% of the total myonuclei (data not shown).

Discussion

In the present study we have characterised the *T. rubripes* orthologue of mouse FoxK1, a transcription factor that is crucial for myogenic progenitor cell function in mouse (Garry et al., 2000). To the best of our knowledge, this is the first report of a non-mammalian *FoxK1* gene. TFoxK1-α is homologous to mouse FoxK1-α and zebrafish FoxK1-α, and contains the two domains characteristic of the Fox family of proteins: the FHA domain, which is involved in phospho-dependent protein–protein interactions, and the FH region that is essential for DNA binding.

The amino-terminal, proline-rich region of the transcriptional activation domain of mouse FoxK1 (Bassel-Duby et al., 1994) is absent in TFoxK1. Both *TFoxK1-γ* and *TFoxK1-δ* code for the same, truncated isoform of TFoxK1, which differs from TFoxK1-α at the 12 carboxyl-terminal residues. The putative translation product of these transcripts contains an incomplete helix H3 in the FH domain and, therefore, it is likely to have a limited DNA binding ability, since helix H3 is crucial for DNA binding (Gajiwala and Burley, 2000) and some of the residues that are involved in its insertion into the major groove of DNA are absent. Hence, it seems that neither TFoxK1-γ nor TFoxK1-δ correspond to mouse FoxK1-β, which is the shorter of two isoforms of FoxK1. Mouse FoxK1-β comprises only 409 amino

acid residues, compared with 617 in FoxK1-α, and differs from FoxK1-α by only six residues at its carboxyl terminus. Nevertheless, mouse FoxK1-β binds DNA with high affinity *in vitro* and can function as a transcriptional repressor in transient transfection assays of C2C12 myogenic cells (Yang et al., 1997). A member of the forkhead/winged helix family of transcription factors has also been recently identified in *Xenopus laevis* and termed *XFoxK1* (Pohl and Knochel, 2004). Despite its name, this gene does not correspond to mouse *FoxK1*; in fact, it seems to be the orthologue of mammalian *interleukin-2 enhancer binding factor* (*FoxK2*).

The 6.4 kb *TFoxK1* gene is composed of seven exons and six small introns located on scaffold 40 of the FUGU 4.0 *T. rubripes* genome assembly. A comparison between mouse *FoxK1* and *TFoxK1* gene structures was not completed owing to extensive gaps in this region of the mouse genome sequence. The human orthologue of *FoxK1*, consisting of nine exons, has been identified using computer-based searches and mapped to chromosome 7p22.1 (Katoh, 2004). Zebrafish *FoxK1* was more similar to the mammalian gene, as it also contained nine exons and eight introns. Nonetheless, the splice sites of the first six exon–intron boundaries were conserved between zebrafish *FoxK1* and *TFoxK1*, and they shared 65.8% identity at the nucleotide level within these 6 exons. The introns of zebrafish *FoxK1* were substantially larger than the corresponding ones in *TFoxK1*, particularly intron II, which was five times larger than in pufferfish. The short intronic and intergenic regions are a characteristic feature of the compact genome of the *Tetraodontidae* family (Aparicio et al., 2002). The *TFoxK1* gene was linked to *abcg1*, *arpc1a*, *arpc1b*, *mmd2*, *slipr*, *cyp3a*, *bat4* and *sdk1* genes on scaffold 40 of the *T. rubripes* genome assembly. The first two genes found downstream of *TFoxK1* are the cytochrome P450 enzyme *cyp3a* and the HLA-B-associated transcript 4 (*bat4*), the function of which has yet to be determined. It is possible that these neighbouring genes might share common regulatory elements, since they are in such close proximity. Interestingly, there is a predicted gene (*hyp2*) found downstream of *TFoxK1*, which does not have orthologues in mouse or zebrafish and codes for a putative peptide of 437 residues that contains a cytochrome *c* haeme-binding site and a C2H2-type zinc-finger domain. Synteny analysis revealed that *TFoxK1* was the true orthologue of zebrafish and mouse *FoxK1*. In spite of the longer evolutionary distance between mouse and pufferfish, the nature and order of genes in the *TFoxK1* chromosomal segment were more similar between these two species than between the two fish species. The murine orthologue of *TFoxK1* was present in a 3.3 Mb syntenic area of chromosome 5G2 that seems to have been subjected to local gene inversion. This region contains genes with varied functions, including the cell adhesion protein sidekick 1 (*sdk1*) that is involved in axonal guidance (Yamagata et al., 2002), the scaffolding protein SLIPR (*slipr*) that is thought to link the receptor protein tyrosine phosphatase β with its substrates at the plasma membrane (Adamsky et al., 2003), the monocyte to macrophage differentiation factor 2 (*mmd2*) whose function is unknown, and the actin-related protein 2/3 complex subunits a and b (*arpc1a* and *arpc1b*), which are involved in actin cytoskeleton organisation and biogenesis (Gournier et al., 2001). The murine orthologues of the white protein homologue (*abcg1*), *cyp3a* and the hypothetical genes *hyp1* and *hyp2* were absent in this segment of chromosome 5G2. Synteny conservation between *T. rubripes* and zebrafish was disrupted as a result of chromosomal translocations and inversions. Only the *arpc1a* and *arpc1b* genes were found in the vicinity of zebrafish *FoxK1* on chromosome 3.

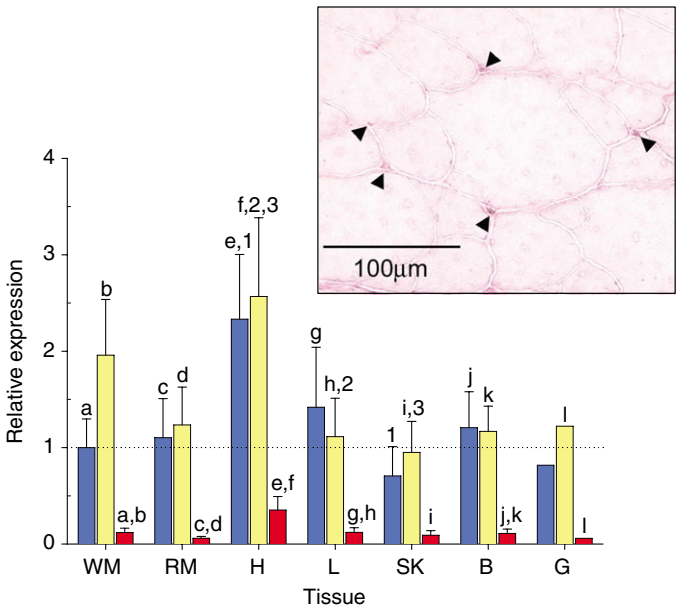

Fig. 6. Tissue distribution of the three alternatively spliced transcripts of *FoxK1* in adult *T. rubripes*. Relative expression levels of *TFoxK1-α* (blue bars), *TFoxK1-γ* (yellow bars) or *TFoxK1-δ* (red bars) in fast muscle (WM), slow muscle (RM), heart (H), liver (L), skin (SK), brain (B) and gonads (G) are represented as ratios against the expression level of *TFoxK1-α* in fast muscle. RNA polymerase II was used as standard. Values are means ± s.e.m. (*N*=4). Statistically significant differences (*P*<0.05) between splice variants within a tissue are labelled with the same letter; significant differences in expression of an isoform across different tissues are indicated by the same number. (Inset) Expression of FoxK1-α in fast myotomal muscle of adult fish. Transverse sections of fast muscle were immunostained for FoxK1-α, a molecular marker of muscle progenitor cells (MPCs). FoxK1-α-positive nuclei (indicated by arrowheads) occupy a sublaminar position that is typical of MPCs.

Alternative pre-mRNA splicing plays a crucial role in regulating gene function by generating a large number of mRNA transcripts and protein isoforms from a limited number of genes. The variant transcripts generated by alternative splicing can have changes in coding sequence, premature stop codons or alterations in the 5′ or 3′ UTRs. The effects of alternative splicing vary from a complete loss of function to acquisition of a new one, to complex and subtle changes in protein function (Stamm et al., 2005). Importantly, alternative splicing can also modulate transcript levels by targeting mRNAs for degradation by nonsense-mediated decay (NMD) (Lewis et al., 2003). One of the main types of alternative splicing is the retention of introns that would normally be excised. This seems to be a frequent phenomenon in the human transcriptome, as demonstrated by a large scale analysis of intron retention in 21 106 known human genes, which revealed that 14.8% of these genes exhibited intron retention events, mostly located in UTRs

(Galante et al., 2004). *TFoxK1* is expressed as three alternatively spliced transcripts, two of which had non-excised introns: *TFoxK1-δ* contained introns IV and V, whereas *TFoxK1-γ* only retained intron IV. The poor splicing efficiency of intron IV might be explained, at least in part, by the relatively weak donor site of exon 4 and acceptor site of exon 5, both of which had consensus scores lower than the average scores for all *TFoxK1* exons. The strong consensus splice sites of exons 5 (donor) and 6 (acceptor), indicate that retention of intron V was rather unexpected. However, it is well known that splice site selection depends not only on the 5′ and 3′ splice sites but also on branch points and exonic–intronic sequence enhancer and silencer elements (Stamm et al., 2005). Retention of intron IV results in the introduction of a premature stop codon, which would result in a shorter variant of FoxK1. However, it is improbable that this truncated isoform be translated, since *TFoxK1-γ* and *TFoxK1-δ* have translation termination sites 760 and 525 nucleotides, respectively, upstream from the 3′-most exon–exon junction. Hence, *TFoxK1-γ* and *TFoxK1-δ* splice variants are likely to be degraded by NMD (Conti and Izaurralde, 2005) and this might be a mechanism of regulating *TFoxK1* transcript levels with important physiological consequences. Indeed, recent studies of gene expression profiling in NMD-deficient human cells revealed that NMD is a crucial post-transcriptional event that regulates expression of a significant number of transcripts involved in a broad range of cellular processes (Mendell et al., 2004).

No significant differences were observed in the temporal and spatial expression patterns of the three *TFoxK1* splice variants, except that the intensity of *TFoxK1-δ* staining was considerably weaker, indicating that *TFoxK1-δ* is expressed at lower levels than the two other transcripts. The low abundance of *TFoxK1-δ* compared with *TFoxK1-α* or *TFoxK1-γ* might be related to the presence of strong consensus splice sequences around intron V, which would tend to promote its excision. The earliest detectable expression of *TFoxK1-α* coincided with the onset of somitogenesis (Fig. 4B) and the activation of the myogenic regulatory factor *Myog* (Fernandes et al., 2006). In contrast to *Myog* (Fernandes et al., 2006), *TFoxK1-α* transcripts were present in the pre-somitic mesoderm during the early stages of segmentation. *FoxK1* expression in mouse has been shown to be independent of *Myog*, since its spatial expression pattern in embryos bearing null mutations in the *Myog* gene was indistinguishable from that in the wild type (Garry et al., 1997). In the somites, *TFoxK1-α* was transiently expressed in a rostrocaudal gradient and markedly downregulated as the somites matured. This pattern was broadly similar to that observed in mouse, as determined by immunofluorescence using a polyclonal antibody that did not distinguish between the two mouse FoxK1 isoforms. The murine gene was expressed concomitantly with *Myf5* in a rostrocaudal gradient in the somites of the developing myotome, including the cells migrating to the limb buds (Garry et al., 1997). Whilst *Myog* expression persisted in the somites of *T. rubripes* embryos until the end of segmentation (Fernandes et al., 2006), *TFoxK1-α* expression was ephemeral and by mid-segmentation it was limited to immature somites and pre-somitic mesoderm (Fig. 4C). Besides being expressed in the developing skeletal musculature, *TFoxK1-α* was also expressed in the heart tube, as reported in mouse (Garry et al., 1997). No transcripts were detected in the pectoral fin bud primordia, which contrasts with the expression of *Myog* in *T. rubripes* (Fernandes et al., 2006). *TFoxK1-α* was expressed in the developing nervous system from an early stage of ontogeny in the head rudiment and developing midbrain (Fig. 4B,C). Transcripts of *TFoxK1-α* were particularly abundant in the rhombo-mesencephalic fissure (Fig. 4D,E), around the region where the cerebellar primordium is formed (Candal et al., 2005). Murine *FoxK1* was also expressed in the developing brain (Garry et al., 1997) and in selected cortical and dopaminergic areas of the adult brain, including the piriform cortex and the Purkinje cell layer of the cerebellum (Wijchers et al., 2006). Taken together with our data, these results indicate that *FoxK1* may play a conserved role in maintenance of developing and adult neurons in vertebrates. Additionally, *TFoxK1-α* was expressed in the optic vesicles, lens and retina of *T. rubripes* embryos, which suggests that *TFoxK1* may have a novel function in eye development.

TFoxK1-α was present in skeletal muscle, heart, brain and liver, as has been reported in mouse (Yang et al., 1997). In addition, it was expressed in the skin and gonads of adult fish. The function of *TFoxK1* in adult tissues other than skeletal muscle has not been elucidated to date. No significant differences in *TFoxK1* transcript levels were found between juvenile and adult fish. Similarly, the muscleblind-like genes *mbnl2a* and *mbnl3* were equally expressed in fast muscle of fish that were actively recruiting fibres by mosaic hyperplasia and in adult fish that had stopped producing new myotubes (Fernandes et al., 2007; Fernandes et al., 2005). Differences in the abundance levels of the three *TFoxK1* splice variants were detected in the different tissues of *T. rubripes*: *TFoxK1-δ* was expressed at lower levels than *TFoxK1-α* and *TFoxK1-γ* in all tissues studied and, in general, the latter two splice variants were transcribed at similar levels. The presence of *TFoxK1* in multiple tissues suggests that it might regulate cell cycle progression and transcription, not only in muscle but also in other cell types. The α and β alternatively spliced transcripts of *FoxK1* are molecular markers of proliferating and quiescent satellite cells, respectively (Garry et al., 2000; Garry et al., 1997). In *T. rubripes*, TFoxK1-α was also detected in mononuclear cells in fast myotomal muscle of adult fish, which correspond to myogenic cells (Johnston, 2006). The function of mammalian FoxK1-α as a transcriptional activator or repressor remains uncharacterised and nothing is known about its potential binding partners or the downstream target genes that are regulated by FoxK1-α, other than the cyclin-dependent kinase inhibitor $p21^{CIP}$ (Hawke et al., 2003).

Variation in the number of muscle fibres and in some cases myogenic progenitor cell numbers with embryonic temperature in the free swimming larval stages is a widespread phenomenon amongst teleosts and has been documented in a range of unrelated species from temperate and tropical environments including cod (Hall and Johnston, 2003), Atlantic salmon (Johnston et al., 2000b), European sea bass (Alami-Durante et al., 2007) and zebrafish (Hung-Tai Lee and Ian A. Johnston, unpublished results). The tiger pufferfish also shows plasticity of myogenesis with embryonic temperature, since the number of fast muscle fibres in newly hatched larvae is significantly less at 21°C than 15°C or 18°C (Ian A. Johnston, unpublished results).

Developmental plasticity in embryonic myogenesis could potentially involve any one of a number of steps, including commitment of stem cells to the myogenic lineage, the number of divisions of myogenic progenitors prior to cell cycle exit, apoptosis, migration and fusion events. In *T. rubripes*, the rate of decline of *Myog* from the formation of the first to the last somite-pair was greater at 21°C than 15°C, with intermediate rates of decrease at 18°C (Fernandes et al., 2006). These results lead to the hypothesis that lower embryonic temperatures delay and prolong muscle differentiation, at least in part involving changes in *Myog* expression (Fernandes et al., 2006). The real-time PCR analysis of *TFoxK1* transcripts demonstrated that the expression of this gene also varies with respect to embryonic temperature for equivalent developmental stages. Moreover, the relative amounts of the three *TFoxK1* transcripts were affected by incubation temperature, indicating differential regulation of alternative splicing induced by temperature. Since *TFoxK1* is likely to be involved in myogenesis and is expressed in MPCs, it is a good candidate gene for playing a key role in the temperature-induced phenotypic plasticity of muscle development observed in *T. rubripes*.

List of abbreviations

C_T	threshold cycle
DIG	digoxigenin
FH	forkhead
FHA	forkhead-associated
Fox	forkhead/winged helix family
MPC	muscle progenitor cell
Myog	myogenin
NMD	nonsense-mediated decay
QPCR	quantitative real-time PCR
TFoxK1	*T. rubripes* forkhead box protein K1
UTR	untranslated region

We would like to thank Professor Yuzuru Suzuki, Dr Kiyoshi Kikuchi, Dr Hiroaki Suetake and Mr Naoki Misuno at the Graduate School of Agricultural and Life Sciences (University of Tokyo, Japan) for their invaluable help during the collection of *T. rubripes* embryos. We thank Mrs Marguerite Abercromby (University of St Andrews, UK) for her assistance with the immunohistochemistry experiments. We are grateful to Ms Shelby Steele (University of Ottawa, Canada) for her critical reading of the manuscript. This work was funded by an Environmental Genomics Grant (NER/S/2001/00250) from the Natural Environment Research Council, UK.

References

Adamsky, K., Arnold, K., Sabanay, H. and Peles, E. (2003). Junctional protein MAGI-3 interacts with receptor tyrosine phosphatase beta (RPTP beta) and tyrosine-phosphorylated proteins. *J. Cell Sci.* **116**, 1279-1289.

Alami-Durante, H., Olive, N. and Rouel, M. (2007). Early thermal history significantly affects the seasonal hyperplastic process occurring in the myotomal white muscle of *Dicentrarchus labrax* juveniles. *Cell Tissue Res.* **327**, 553-570.

Andres, V. and Walsh, K. (1996). Myogenin expression, cell cycle withdrawal, and phenotypic differentiation are temporally separable events that precede cell fusion upon myogenesis. *J. Cell Biol.* **132**, 657-666.

Aparicio, S., Chapman, J., Stupka, E., Putnam, N., Chia, J. M., Dehal, P., Christoffels, A., Rash, S., Hoon, S., Smit, A. et al. (2002). Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. *Science* **297**, 1301-1310.

Bassel-Duby, R., Hernandez, M. D., Yang, Q., Rochelle, J. M., Seldin, M. F. and Williams, R. S. (1994). Myocyte nuclear factor, a novel winged-helix transcription factor under both developmental and neural regulation in striated myocytes. *Mol. Cell. Biol.* **14**, 4596-4605.

Candal, E., Anadon, R., Bourrat, F. and Rodriguez-Moldes, I. (2005). Cell proliferation in the developing and adult hindbrain and midbrain of trout and medaka (teleosts): a segmental approach. *Brain Res. Dev. Brain Res.* **160**, 157-175.

Carlsson, P. and Mahlapuu, M. (2002). Forkhead transcription factors: key players in development and metabolism. *Dev. Biol.* **250**, 1-23.

Chen, J. C. and Goldhamer, D. J. (2003). Skeletal muscle stem cells. *Reprod. Biol. Endocrinol.* **1**, 101.

Chuang, W. J., Yeh, I. J., Hsieh, Y. H., Liu, P. P., Chen, S. W. and Jeng, W. Y. (2002). 1H, 15N and 13C resonance assignments for the DNA-binding domain of myocyte nuclear factor (Foxk1). *J. Biomol. NMR* **24**, 75-76.

Clark, K. L., Halay, E. D., Lai, E. and Burley, S. K. (1993). Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature* **364**, 412-420.

Conti, E. and Izaurralde, E. (2005). Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. *Curr. Opin. Cell Biol.* **17**, 316-325.

Durocher, D. and Jackson, S. P. (2002). The FHA domain. *FEBS Lett.* **513**, 58-66.

Fernandes, J. M., Mackenzie, M. G., Elgar, G., Suzuki, Y., Watabe, S., Kinghorn, J. R. and Johnston, I. A. (2005). A genomic approach to reveal novel genes associated with myotube formation in the model teleost, *Takifugu rubripes*. *Physiol. Genomics* **22**, 327-338.

Fernandes, J. M., MacKenzie, M. G., Wright, P. A., Steele, S. L., Suzuki, Y., Kinghorn, J. R. and Johnston, I. A. (2006). Myogenin in model puffer fish species: comparative genomic analysis and thermal plasticity of expression during early development. *Comp. Biochem. Physiol.* **1D**, 35-45.

Fernandes, J. M., Kinghorn, J. R. and Johnston, I. A. (2007). Characterization of two paralogous muscleblind-like genes from the tiger pufferfish (*Takifugu rubripes*). *Comp. Biochem. Physiol.* **146B**, 180-186.

Gajiwala, K. S. and Burley, S. K. (2000). Winged helix proteins. *Curr. Opin. Struct. Biol.* **10**, 110-116.

Galante, P. A., Sakabe, N. J., Kirschbaum-Slager, N. and de Souza, S. J. (2004). Detection and evaluation of intron retention events in the human transcriptome. *RNA* **10**, 757-765.

Garry, D. J., Yang, Q., Bassel-Duby, R. and Williams, R. S. (1997). Persistent expression of MNF identifies myogenic stem cells in postnatal muscles. *Dev. Biol.* **188**, 280-294.

Garry, D. J., Meeson, A., Elterman, J., Zhao, Y., Yang, P., Bassel-Duby, R. and Williams, R. S. (2000). Myogenic stem cell function is impaired in mice lacking the forkhead/winged helix protein MNF. *Proc. Natl. Acad. Sci. USA* **97**, 5416-5421.

Gournier, H., Goley, E. D., Niederstrasser, H., Trinh, T. and Welch, M. D. (2001). Reconstitution of human Arp2/3 complex reveals critical roles of individual subunits in complex structure and activity. *Mol. Cell* **8**, 1041-1052.

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids. Symp. Ser.* **41**, 95-98.

Hall, T. E. and Johnston, I. A. (2003). Temperature and developmental plasticity during embryogenesis in the Atlantic cod *Gadus morhua* L. *Mar. Biol.* **142**, 833-840.

Hawke, T. J., Jiang, N. and Garry, D. J. (2003). Absence of p21CIP rescues myogenic progenitor cell proliferative and regenerative capacity in Foxk1 null mice. *J. Biol. Chem.* **278**, 4015-4020.

Johnston, I. A. (1993). Temperature influences muscle differentiation and the relative timing of organogenesis in herring (*Clupea harengus*) larvae. *Mar. Biol.* **116**, 363-379.

Johnston, I. A. (2006). Environment and plasticity of myogenesis in teleost fish. *J. Exp. Biol.* **209**, 2249-2264.

Johnston, I. A. and Hall, T. E. (2004). Mechanisms of muscle development and responses to temperature change in fish larvae. In *The Development of Form and Function in Fishes and the Question of Larval Adaptation*. Vol. 40 (ed. J. J. Govoni), pp. 85-116. Bethesda, MD: American Fisheries Society.

Johnston, I. A., Vieira, V. V. and Abercromby, M. (1995). Temperature and myogenesis in embryos of the Atlantic herring *Clupea harengus*. *J. Exp. Biol.* **198**, 1389-1403.

Johnston, I. A., Cole, N., Abercromby, M. and Vieira, V. L. (1998). Embryonic temperature modulates muscle growth characteristics in larval and juvenile herring. *J. Exp. Biol.* **201**, 623-646.

Johnston, I. A., McLay, H. A., Abercromby, M. and Robins, D. (2000a). Early thermal experience has different effects on growth and muscle fibre

recruitment in spring- and autumn-running Atlantic salmon populations. *J. Exp. Biol.* **203**, 2553-2564.

Johnston, I. A., McLay, H. A., Abercromby, M. and Robins, D. (2000b). Phenotypic plasticity of early myogenesis and satellite cell numbers in atlantic salmon spawning in upland and lowland tributaries of a river system. *J. Exp. Biol.* **203**, 2539-2552.

Johnston, I. A., Manthri, S., Alderson, R., Smart, A., Campbell, P., Nickell, D., Robertson, B., Paxton, C. G. and Burt, M. L. (2003). Freshwater environment affects growth rate and muscle fibre recruitment in seawater stages of Atlantic salmon (*Salmo salar* L.). *J. Exp. Biol.* **206**, 1337-1351.

Johnston, I. A., Abercromby, M., Vieira, V. L., Sigursteindottir, R. J., Kristjansson, B. K., Sibthorpe, D. and Skulason, S. (2004). Rapid evolution of muscle fibre number in post-glacial populations of Arctic charr *Salvelinus alpinus*. *J. Exp. Biol.* **207**, 4343-4360.

Kaestner, K. H., Knochel, W. and Martinez, D. E. (2000). Unified nomenclature for the winged helix/forkhead transcription factors. *Genes Dev.* **14**, 142-146.

Kassar-Duchossoy, L., Gayraud-Morel, B., Gomes, D., Rocancourt, D., Buckingham, M., Shinin, V. and Tajbakhsh, S. (2004). Mrf4 determines skeletal muscle identity in Myf5:Myod double-mutant mice. *Nature* **431**, 466-471.

Katoh, M. (2004). Identification and characterization of human FOXK1 gene in silico. *Int. J. Mol. Med.* **14**, 127-132.

Kelly, A. M. and Rubinstein, N. A. (1994). The diversity of muscle fiber types and its origin during development. In *Myology: Basic and Clinical* (ed. A. G. Engel and C. Franzini-Armstrong), pp. 119-133. New York: McGraw-Hill.

Lewis, B. P., Green, R. E. and Brenner, S. E. (2003). Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc. Natl. Acad. Sci. USA* **100**, 189-192.

Marchler-Bauer, A. and Bryant, S. H. (2004). CD-Search: protein domain annotations on the fly. *Nucleic Acids Res.* **32**, W327-W331.

Mendell, J. T., Sharifi, N. A., Meyers, J. L., Martinez-Murillo, F. and Dietz, H. C. (2004). Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. *Nat. Genet.* **36**, 1073-1078.

Moss, F. P. and Leblond, C. P. (1971). Satellite cells as the source of nuclei in muscles of growing rats. *Anat. Rec.* **170**, 421-435.

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **29**, e45.

Pfaffl, M. W., Horgan, G. W. and Dempfle, L. (2002). Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res.* **30**, e36.

Pohl, B. S. and Knochel, W. (2004). Isolation and developmental expression of *Xenopus* FoxJ1 and FoxK1. *Dev. Genes Evol.* **214**, 200-205.

Reese, M. G., Eeckman, F. H., Kulp, D. and Haussler, D. (1997). Improved splice site detection in Genie. *J. Comput. Biol.* **4**, 311-323.

Rowe, R. W. and Goldspink, G. (1969). Muscle fibre growth in five different muscles in both sexes of mice. *J. Anat.* **104**, 519-530.

Rowlerson, A. and Veggetti, A. (2001). Cellular mechanisms of post-embryonic muscle growth in aquaculture species. In *Muscle Development and Growth*. Vol. 18 (ed. I. A. Johnston), pp. 103-140. San Diego: Academic Press.

Rudnicki, M. A. and Jaenisch, R. (1995). The MyoD family of transcription factors and skeletal myogenesis. *BioEssays* **17**, 203-209.

Schultz, E. (1996). Satellite cell proliferative compartments in growing skeletal muscles. *Dev. Biol.* **175**, 84-94.

Seale, P., Sabourin, L. A., Girgis-Gabardo, A., Mansouri, A., Gruss, P. and Rudnicki, M. A. (2000). Pax7 is required for the specification of myogenic satellite cells. *Cell* **102**, 777-786.

Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D., Thanaraj, T. A. and Soreq, H. (2005). Function of alternative splicing. *Gene* **344**, 1-20.

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

Weigel, D., Jurgens, G., Kuttner, F., Seifert, E. and Jackle, H. (1989). The homeotic gene fork head encodes a nuclear protein and is expressed in the terminal regions of the *Drosophila* embryo. *Cell* **57**, 645-658.

Wheelan, S. J., Church, D. M. and Ostell, J. M. (2001). Spidey: a tool for mRNA-to-genomic alignments. *Genome Res.* **11**, 1952-1957.

Wijchers, P. J., Hoekman, M. F., Burbach, J. P. and Smidt, M. P. (2006). Identification of forkhead transcription factors in cortical and dopaminergic areas of the adult murine brain. *Brain Res.* **1068**, 23-33.

Yamagata, M., Weiner, J. A. and Sanes, J. R. (2002). Sidekicks: synaptic adhesion molecules that promote lamina-specific connectivity in the retina. *Cell* **110**, 649-660.

Yang, Q., Bassel-Duby, R. and Williams, R. S. (1997). Transient expression of a winged-helix protein, MNF-beta, during myogenesis. *Mol. Cell. Biol.* **17**, 5236-5243.