Prdm1a and miR-499 act sequentially to restrict Sox6 activity to the fast-twitch muscle lineage in the zebrafish embryo

XinGang Wang¹, Yosuke Ono¹, Swee Chuan Tan¹, Ruth JinFen Chai¹,*, Caroline Parkin¹,† and Philip W. Ingham¹,2,‡

SUMMARY
Sox6 has been proposed to play a conserved role in vertebrate skeletal muscle fibre type specification. In zebrafish, sox6 transcription is repressed in slow-twitch progenitors by the Prdm1a transcription factor. Here we identify sox6 cis-regulatory sequences that drive fast-twitch-specific expression in a Prdm1a-dependent manner. We show that sox6 transcription subsequently becomes derepressed in slow-twitch fibres, whereas Sox6 protein remains restricted to fast-twitch fibres. We find that translational repression of sox6 is mediated by miR-499, the slow-twitch-specific expression of which is in turn controlled by Prdm1a, forming a regulatory loop that initiates and maintains the slow-twitch muscle lineage.

KEY WORDS: Sox6, Translational control, miR-499 (mir499), Muscle fibre type, Zebrafish, Prdm1

INTRODUCTION
During skeletal myogenesis in vertebrates, mesodermally derived muscle progenitor cells give rise to myoblasts that in turn differentiate into either slow-twitch or fast-twitch muscle fibres. In zebrafish embryos, slow-twitch fibres derive from paraxial mesodermal cells that lie immediately adjacent to the notochord, the so-called adaxial cells (Devoto et al., 1996; Blagden et al., 1997). The slow-twitch differentiation pathway is initiated in adaxial cells by the transcription factor Prdm1a (also known as Blimp1), expression of which is activated by notochord-derived Sonic hedgehog (Shh) protein (Roy et al., 2001; Baxendale et al., 2004). In the absence of Prdm1a activity, adaxial cells adventitiously activate transcription of fast-twitch-specific genes, such as those encoding the fast myosin heavy chains and fast troponin, as well as the transcription factor Sox6, the ectopic expression of which downregulates transcription of the slow-twitch-specific genes, including slow myosin heavy chain 1 (smhc1), slow troponin (stnnc; tnnc1b – Zebrafish Information Network) and prox1 in prdm1a mutants (von Hofsten et al., 2008). Reciprocally, loss of Sox6 function in fast-twitch fibres results in ectopic expression of at least one slow-twitch-specific gene, stnnc, leading to the suggestion that Sox6 plays a key role in promoting fast-twitch-specific fate by repressing the slow-twitch differentiation pathway (von Hofsten et al., 2008).

A similar role was previously proposed for Sox6 during mammalian myogenesis, based on the excess of fibres with slow-twitch character that differentiate in Sox6 mutant mice (Hagiwara et al., 2005; Hagiwara et al., 2007). Consistent with this, overexpression of Sox6 has been shown to repress transcription of the slow-twitch-specific β-MHC and troponin genes and to cause a concomitant increase in slow-twitch and decrease in fast-twitch muscle fibres in transgenic mice (van Rooij et al., 2009), and very recent studies have demonstrated that Sox6 binds to conserved cis-regulatory elements in slow-twitch fibre genes to repress their transcription in adult fast-twitch muscle (Quiat et al., 2011). Little is known about how the lineage-specific activity of Sox6 is regulated in mammals; however, a role for the microRNA miR-499 has been suggested, based upon the finding that the 3′UTRs of human, mouse and rat Sox6 contain between 1 and 4 consensus recognition sites for the miR-499 SEED sequence and the demonstration that miR-499 reduces Sox6 mRNA levels both in vitro and in vivo (McCarthy et al., 2009; van Rooij et al., 2009; Bell et al., 2010).

Here, through the in vivo analysis of sox6 regulation using transgenic zebrafish carrying sox6 reporter gene constructs, we identify 5′ cis-acting sequences that are sufficient to drive spatially restricted transcription in the fast-twitch muscle progenitors and 3′ cis-acting sequences that mediate translational repression in slow-twitch fibres. Our findings reveal a prdm1a/mir499/sox6 gene regulatory network that underlies the establishment and maintenance of the slow-twitch muscle fibre lineage in response to Shh signalling.

MATERIALS AND METHODS
Zebrafish strains and husbandry
Adult fish were maintained on a 14-hour light/10-hour dark cycle at 28°C in the AVA (Singapore) certificated IMCB Zebrafish Facility. Previously described zebrafish strains used were prdm1anrd (Hernandez-Lagunas et al., 2005), Tg(actin:GAL4) (Scheer and Campos-Ortega, 1999), Tg(PACprdm1:GFP)i106 and Tg(BACsmyhc1:gfp)i108 (Elworthy et al., 2008).

Constructs and transgenics
The sox6 transcription start site (TSS) was identified by 5′ RACE (Ambion) following the manufacturer’s protocol. Fragments of differing size were amplified by PCR from BAC DNA (BXS537340.15) and cloned into a Gateway 5′ entry vector p5E-MCS (Kwan et al., 2007). Reporter constructs were generated by site-specific recombination with pME-EGFP or pME-mCherry, p3E-xv40 and pDestTol2pA. The 3 kb sox6 3′ UTR was identified by 3′ RACE, amplified from genomic DNA and cloned into the sox6:EGFP tol2 vector replacing SV40 poly(A). Sequences identified in

¹Institute of Molecular and Cell Biology, 61 Biopolis Drive, Proteos, Singapore 138673, Singapore. ²Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, Singapore 117543, Singapore.

‡Author for correspondence (pingham@imcb.a-star.edu.sg)

*Present address: School of Anatomy and Human Biology, The University of Western Australia, Perth 6009, WA, Australia
†Present address: MRC Centre for Developmental and Biomedical Genetics, Firth Court, Western Bank, Sheffield S10 2TN, UK
"Author for correspondence (pingham@imcb.a-star.edu.sg)

Accepted 10 August 2011
this analysis have been deposited in GenBank under accession numbers JN216841, JN216842. The *smyhc1-lyn-tolTomato* construct utilised a 9.7 kb *smyhc1* promoter (Elworthy et al., 2008) to drive *tdTomato* fused to the minimal myristoyl-palmitoyl-membrane anchor sequence of Lyn kinase (Pyenta et al., 2001). To overexpress miR-499, its precursor (91 bp) was cloned into a *UAS:EGFP* stem-loop vector. All tol2 constructs were co-injected with tol2 mRNA into one-cell stage embryos to produce transgenic lines or for transient expression analysis.

**Site-directed mutagenesis**

Mutation of cis-regulatory elements was performed using the QuikChange Site-Directed Mutagenesis Kit (Stratagene) according to the manufacturer’s protocol, except that iProof DNA polymerase (BioRad) was utilised. The core sequence of the Prdm1a binding site GAAA was mutated into CTCT.

The miR-499 SEED recognition site TCTTA in the *prdm1anrd* promoter utilized for cloning of the precursor (91 bp) was co-injected with tol2 mRNA into one-cell stage embryos to produce transgenic lines or for transient expression analysis.

**Generation and purification of anti-Sox6 antibody**

A short fragment of *sox6* cDNA 5’ to the HMG box was cloned into the His tag expression vector pET-21b. The encoded fusion protein was used to immunise rabbits by Absebra-Animab (Beijing). The resulting serum was affinity purified using the His-tagged Sox6 following standard protocols (Nishino et al., 1994).

**In situ hybridisation and immunofluorescence**

DIG-labelled miR-499 antisense locked nucleic acid (LNA) probe was purchased from Exiqon. Standard in situ hybridisation (ISH) was performed as described (Oxtoy and Jowett, 1993; Strähle et al., 1994). Fluorescence ISH was performed on sections using anti-DIG peroxidase and fluorescence substrate Cy5 tyramide signal amplification (TSA, Perkin Elmer) according to Brand and Holley (Brand and Holley, 2009). Antibody staining was performed as previously described (Elworthy et al., 2008) using mAb F59 (anti-myosin heavy chain 1; DSHB) 1:50; anti-zebrafish *sox6* 1:500; and anti-EGFP (Abcam) 1:5000. Secondary antibodies comprised Rhodamine-coupled anti-mouse (Invitrogen) at 1:500 and FITC-coupled anti-rabbit (Invitrogen) at 1:1000. Images were acquired with an Olympus Fluoview confocal microscope using Olympus FV10-ASW software and analysed using ImageJ software (NIH).

** Morpholino sequences and injections**

MO1-*myod1* (5’-ATATCCGACAACCTCCATCTTTTTTTG-3’) and MO1-*myf5* (5’-TACGTCCATGATTGGTTTGGTGTTG-3’) were co-injected at 0.1 pm each per embryo (Lin et al., 2006). MO2-prdm1a (5’-TGGTGTCATACCTCTTTGGAGTCTG-3’) was injected at 0.2 pm per embryo (Baxendale et al., 2004). Morpholinos were supplied by Gene Tools (Philomath, OR, USA).

**RESULTS AND DISCUSSION**

**Spatially regulated sox6 transcription is controlled by myogenic regulatory factors and Prdm1a**

Transcription of *sox6* in the zebrafish embryo is clearly detectable by in situ hybridisation (ISH) at the 10–somite stage in the lateral region of each somite, which is the location of fast-twitch muscle progenitors; however, no transcripts are detectable in the adaxial cells, which are the slow-twitch muscle progenitors (Fig. 1A). The levels of *sox6* transcript reach their peak in the somites 1 day post-fertilisation (dpf) and gradually diminish thereafter, although persisting at least until 6 dpf (see below). Transcripts are also detectable in the otic vesicle (Fig. 1E) and in the head muscle, pectoral fin, kidney and in specific regions of the brain and retina (see Fig. S1 in the supplementary material).

As previously shown, *sox6* transcription is derepressed in adaxial cells in the absence of Prdm1a function (von Hofsten et al., 2008). Ectopic transcripts are first readily detectable in adaxial cells at the 10-somite stage in embryos homozygous for the *prdm1a* null allele (Fig. 1B); at the same stage, expression of *stnck* is markedly reduced (Fig. 1C,D) and continues to decline as embryogenesis proceeds, whereas the signal due to ectopic *sox6* expression persists and increases in intensity (not shown).

The myogenic regulatory factors (MRFs) Myod1 and Myf5 are essential for the commitment of cells to the myogenic differentiation programme (reviewed by Pownall et al., 2002). Previous studies have shown that simultaneous morpholino-mediated knockdown of both genes results in the loss of myogenic gene expression in the somites (Hinits et al., 2009). Little, if any, *sox6* transcript was detected in the somites of *myod1;myf5* double-morphant embryos as compared with wild-type embryos (Fig. 1E,F); expression in the otic vesicle, which serves as a positive internal control, was unaffected. Morpholino-mediated knockdown of the *Mef2* transcription factors (Hinits and Hughes, 2007), by contrast, had no effect on *sox6* transcription (see Fig. S2 in the supplementary material).

**Identification of cis-acting sox6 regulatory elements**

To define the cis-acting sequences that mediate the spatiotemporal regulation of *sox6*, we first assayed the activity of sequences from the 5’ end of the gene in transgenic embryos. A 10.5 kb DNA fragment that includes 7.6 kb upstream of the transcription start site (TSS) and 2.9 kb extending downstream as far as the third exon in the 5’UTR region (see Fig. 2A) was cloned into an *EGFP* stem-loop vector. Co-injection of this construct with tol2 transposase mRNA resulted in strong transient expression of *EGFP* in the skeletal muscle fibres (see Fig. S1D in the supplementary material). Embryos from a stable line, *Tg(10.5 kb sox6:EGFP)*, showed expression specifically in the fast-twitch fibres of both trunk and head muscle, the intermediate cell mass, as well as in the otic vesicle (Fig. 2B, see Fig. S1E-G in the supplementary material).

Interspecific sequence alignments revealed that this fragment contains a number of regions conserved between teleosts (conserved non-coding elements or CNEs), some of which are also conserved in mammals (highly conserved non-coding
To test their functional significance, we generated deletion derivatives of the 10.5 kb fragment lacking two or three of the upstream CNEs and assayed their activities in stable transgenic lines. \(Tg(5.7 \text{ kb } \text{sox6:EGFP}_{sv40})_{i253}\) embryos showed an essentially identical pattern of expression to that of \(Tg(10.5 \text{ kb } \text{sox6:EGFP}_{sv40})_{i252}\), indicating that the two most distal elements that are conserved between fish species are dispensable for expression in the myotome (Fig. 2C). \(Tg(4.4 \text{ kb } \text{sox6:EGFP}_{sv40})_{i255}\) embryos, by contrast, showed much lower and highly mosaic EGFP expression, albeit restricted to fast-twitch fibres, as well as ectopic expression in the notochord (Fig. 2E), indicating a crucial requirement for the more proximal HCNE1, which is highly conserved between zebrafish and mammals. Addition of a 0.5 kb fragment of this HCNE, which includes the single Myod1 consensus binding site upstream of the TSS, to generate \(Tg(4.9 \text{ kb } \text{sox6:EGFP}_{sv40})_{i254}\), was sufficient to restore the wild-type expression pattern driven by the longer (10.5 and 5.7 kb) fragments (Fig. 2D).

To confirm that these \textit{sox6} cis-regulatory elements respond to Prdm1a repression, we generated an additional line, \(Tg(5.7 \text{ kb } \text{sox6:mCherry}_{sv40})_{i256}\), in which the 5.7 kb fragment drives mCherry expression and crossed this to fish carrying the previously described \(Tg(PACprdm1:GFP)_{i106}\) line that drives GFP expression specifically in the slow-twitch lineage (Elworthy et al., 2008). As expected, embryos derived from this cross show mutually exclusive patterns of EGFP and mCherry expression in the myotome. Knockdown of Prdm1a activity by morpholino injection, however, resulted in ectopic mCherry expression in EGFP-expressing cells. Thus, transcription driven by the 5.7 kb fragment is repressed by Prdm1a. Notably, we identified nine consensus binding sites for Prdm1a within the 5.7 kb sequence. To test whether Prdm1a-dependent repression is mediated by these binding sites, lines carrying a modified 5.7 kb:EGFP reporter, in which all nine sites had been mutated by site-directed mutagenesis, were generated. Surprisingly, no difference in expression was seen between the wild-type and the mutated reporter constructs (data not shown). Thus, Prdm1a-mediated repression is independent of the consensus binding sites flanking the TSS, suggesting that it might be an indirect effect.

**Highly conserved sequences within the second intron of sox6 are sufficient for fast-twitch-specific transcription**

The interspecies sequence alignment revealed a second, short (219 bp) HCNE within the second intron, the deletion of which abolishes the activity of the 5.7 kb:EGFP reporter in stable lines (not shown). To determine whether HCNE2 acts as a muscle enhancer, we cloned it into the EGFP vector with a \(\beta\)-\textit{globin} minimal promoter. A stable transgenic line, \(Tg(sox6
HCNE2:mCherry)<sup>257</sup>, carrying this construct showed EGFP expression exclusively in the myotome, which unexpectedly was restricted to fast-twitch fibres (Fig. 2H). As with the Tg(5.7 kb sox6:mCherry)<sup>236</sup> reporter, Tg(sox6HCNE2:mCherry)<sup>257</sup> was ectopically expressed in slow muscle cells in pldm1a morphant embryos (Fig. 2I). Analysis of the HCNE sequence using MatInspector software (Cartharius et al., 2005) identified a single Prdm1a consensus binding sequence, as well as sites for various transcription factors including Mef2, Srf, Dmrt2 and Maf (see Fig. S4 in the supplementary material). Mutation of any of these sites resulted in complete loss of reporter gene expression in stable lines (data not shown).

**Transcription of sox6 is derepressed in slow-twitch fibres at postembryonic stages**

Although expression of the Tg(5.7 kb sox6:EGFP<sup>sv40</sup>)<sup>253</sup> reporter gene is initially restricted to fast-twitch fibres (Fig. 3A), this restriction starts to break down by 4 dpf (Fig. 3B), and, by 6 dpf, most slow-twitch fibres are also EGFP positive (Fig. 3C). To confirm that this expression profile reflects that of the endogenous gene, we assayed sox6 mRNA distribution in Tg(BACsmyhc1:gfp)<sup>108</sup> larvae using fluorescent ISH. At 1 dpf, the red sox6 signal generated by tyramide Cy5 substrate was localised to fast-twitch fibre cells and excluded from the EGFP-expressing slow-twitch fibres (Fig. 3D). At 6 dpf, by contrast, sox6 transcript could be detected in all the slow muscle cells (Fig. 3F). Despite this accumulation of endogenous sox6 mRNA, no Sox6 protein could be detected in slow-twitch fibres, even by 6 dpf (Fig. 3E,G).

The absence of Sox6 protein implies that sox6 mRNA is subject to post-transcriptional regulation in slow-twitch fibres. Previous studies have implicated Mir499, a microRNA gene residing in the nineteenth intron of Myh7b, in regulating Sox6 transcript levels in mammalian myofibres, but not in direct repression of Sox6 translation (McCarthy et al., 2009; van Rooij et al., 2009; Bell et al., 2010). MicroRNAs are known to regulate both mRNA stability and translation by targeting sequences in the 3′ UTRs of mRNAs (Huntzinger and Izaurralde, 2011). Analysis of the sox6 3′ UTR in the miRBase database (http://www.mirbase.org) identified two recognition sequences for miR-499; a further three potential target sites were identified by manual alignment with the miR-499 SEED sequence alone (see Fig. S5 in the supplementary material). None of these is a perfect match, implying that they might mediate translational repression rather than transcript degradation (Huntzinger and Izaurralde, 2011).

**Repression of Sox6 activity in slow-twitch fibres is maintained by miR-499**

To determine the functional significance of these miR-499 target sequences, we replaced the SV40 3′ UTR sequences in Tg(5.7 kb sox6:EGFP<sup>sv40</sup>)<sup>253</sup> with the endogenous sox6 3′ UTR and used this construct to generate the stable line Tg(5.7 kb sox65.7:EGFPsox63′ UTR)<sup>258</sup>. As in the original transgenic line, expression of EGFP driven by this modified transgene was restricted to fast-twitch fibres from 48 hours post-fertilisation (hpf). In contrast to the original SV40 3′ UTR line (Fig. 3H), however, expression remained restricted to the fast-twitch fibres in larvae after 4 dpf (Fig. 3I), consistent with the miR-499 sites mediating translational repression of the reporter. To confirm this, we mutated all five potential miR-499 target sites and used this construct to generate the stable line Tg(5.7 kb sox6:EGFPsox663′ UTR-M<sup>499</sup>)<sup>259</sup>. Larvae from this line similarly showed EGFP expression restricted to the fast-twitch fibres at 2 dpf, but by 6 dpf most slow muscle also showed ectopic EGFP expression (Fig. 3J). We also
generated a 1 kb internal deletion of the sox6 3′UTR that removes all five miR-499 target sites and used this to generate a stable Tg(5.7 kb sox6:EGFPsox63/H11032 line. Such transgenic larvae also showed ectopic expression of EGFP in slow-twitch fibres by 6 dpf; interestingly, the EGFP expression levels within the slow-twitch fibres were significantly higher than in the fast-twitch fibres (Fig. 3K).

Our findings imply that lineage-restricted expression of miR-499 is sufficient to suppress sox6 translation in slow-twitch fibres. To explore this hypothesis, we generated probes for zebrafish miR-499 and its host gene myh7b (GenBank accession JN216840), and used these to analyse the spatial distribution of their transcripts by ISH. As predicted, both transcripts were found to accumulate specifically in slow muscle fibres (Fig. 4B-E). To confirm that expression of miR-499 is sufficient to repress sox6 translation, we next forced its expression in fast-twitch fibres by injecting a muscle-specific actin:GAL4 driver line: such ectopic expression of myh7b and its host gene by exon skipping.

Fig. 4. Prdm1a promotes the spatially restricted expression of myh7b and mir499, which is sufficient to repress sox6 translation. (A) Exon-intron structure of the myh7b gene showing the location of mir499 in intron 18. (B-E) Transverse sections of wild-type embryos hybridised with probes for myh7b or mir499 transcripts, both of which are restricted to the superficially located slow-twitch fibres and the medially located muscle pioneers, a subset of the slow-twitch lineage. (F-H) Sagittal optical sections of a 24 hpf Tg(actin:GAL4) embryo injected with an EGFP-mir499 construct under UAS control: fibres expressing the construct (f, green) show significant reduction in nuclear accumulation of Sox6 protein (G, red) as revealed by the merged image (H); nuclei are stained with DAPI (blue). (I, J) Flat mounts of 10-somite stage wild-type and prdm1a<sup>mutant</sup> mutant embryos hybridised with probe for myh7b, showing strong downregulation in the absence of Prdm1a function. (K, L) Gene regulatory network underlying the specification and maintenance of slow-twitch and fast-twitch muscle lineages in zebrafish. Scale bars: 20 µm.

Our data reveal a Shh-activated gene regulatory network that acts through both transcriptional and post-transcriptional mechanisms to establish and maintain lineage-specific gene expression patterns in the skeletal muscle of the zebrafish (Fig. 4K, L). Notably, the involvement of miR-499 and Sox6 in this network has been conserved from fish to mammals, although a role for miR-499 in inhibiting sox6 translation had not previously been demonstrated. The role of Prdm1a in this network, by contrast, appears to be telost specific, as there is no evidence implicating it in mammalian fibre type specification. Although our data confirm the role of Prdm1a in repressing sox6 transcription, they imply that this function is likely to be indirect, indicating that other key components of the network remain to be discovered.

Acknowledgements
We thank Stone Elworthy for the His-tagged Sox6 fusion construct, Augustine Cheong Weng Onn for affinity purification of the anti-Sox6 antibody, Byrappa Venkatesh and Ravi Vydianathan for advice on multi-species sequence alignments and members of the P.W.I. laboratory for helpful comments and advice.

Funding
This work was supported by the Agency for Science Technology and Research (A*STAR).

Competing interests statement
The authors declare no competing financial interests.

Supplementary material
Supplementary material for this article is available at http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.070516/-/DC1

References

Genes Dev. 11, 2163-2175.


