# An Analysis of the role of Sox6 in muscle fibre-type specification in the zebrafish (*Danio rerio*)

Thesis presented for the degree of PhD by

Harriet Jackson

MRC Centre for Developmental and Biomedical Genetics, Department of Biomedical Science, The University of Sheffield

Institute of Molecular and Cell Biology, A*STAR, Singapore

September 2013

# Acknowledgements

Firstly, I would like to say a huge thank you to my supervisors, Dr Vincent Cunliffe and Professor Philip Ingham. I thank Vincent for his constant support throughout my PhD and his dedicated help compiling my thesis. I am grateful to Phil for his guidance and encouragement during my time in Singapore.

This thesis would not have been possible without the help of Dr Wang Xingang, Dr Ashish Maurya, Dr Stone Elworthy and Dr Michael Harrison, who have provided me with numerous constructs, fish lines and advice throughout the past four years. Thanks also to the aquaria staff, both in Singapore and Sheffield who have taken great care of my zebrafish lines.

Thanks to all my colleagues and friends, who have guided me through my PhD as well as providing tremendous help in the lab, Wang Xingang, Asha ji Nair, Tom Carney, Benjin, Weixin Niah, Samantha Lee, Ashley Ng, Nicole Kong, Zhou Zhonghua, Yosuke Ono, Kalin Narov, Anjali Verma, Audrey Iyu, Wilson Wang, Raymond Lee, Paola Meza and Stephen Georgiou. A particular thanks to Ashish bhai saab Maurya for often needed distractions throughout the day, and to Claudia for your tremendous friendship in my first few months in Singapore and for making it such a fantastic experience. A huge thanks to Penelope Boyd for being an amazing friend and providing me with the extra hands I needed during my stressful preparations for leaving Singapore, and to Semil Choksi for always dealing patiently with my questions.

I would like to thank the rest of my family and friends, especially my Mum and Dad, for their unwavering love and support throughout the duration of my PhD. Finally, I thank Mark, for keeping me smiling everyday and for joining me on our journey to Singapore.

# Abstract

The transcription factor Sox6 has been implicated in a number of developmental processes in mammals including myogenesis, chondrogenesis, erythropoiesis and gliogenesis. Earlier studies using morpholino-mediated knockdown of zebrafish *sox6*, indicated that *sox6* performed an evolutionarily conserved role in regulating skeletal muscle fibre-type differentiation. Consistent with this role, *sox6* was also identified as a target of the Prdm1a transcription factor, which was previously shown to be a key regulator of slow-twitch fibre type identity. I have utilised the UAS:GAL4 system to misexpress Sox6 in the slow-twitch muscle resulting in the downregulation of slow-twitch muscle genes. To analyse further the role of zebrafish Sox6 in myogenesis, BAC recombineering was used to generate a transgenic line that recapitulates the expression pattern of the endogenous *slow troponin c* (*tnnc1b)* gene. Mutation of potential Sox6 binding sites in the *tnnc1b:GFP* minimal promoter did not result in any ectopic *tnnc1b:GFP* expression in the fast-twitch muscle suggesting that Sox6 represses *tnnc1b* indirectly. Deletion analysis of the *tnnc1b* construct identified a conserved enhancer sequence in the first intron of the *tnnc1b* gene, where the skeletal muscle enhancer in its murine orthologue is also located. I used a zinc-finger nuclease based approach to generate mutant alleles of the *sox6* gene. Consistent with the *sox6* morphant phenotype, *sox6* homozygous mutant embryos show extensive ectopic expression of *slow troponin c* (*tnnc1b*) throughout the fast-twitch muscle. However, other slow-specific muscle genes, such as *slow myosin heavy chain 1 (smyhc1),* are only ectopically expressed in a subset of fast-twitch fibres in *sox6*$^{-/-}$ mutant zebrafish. Interestingly, adult zebrafish *sox6* homozygous mutants are significantly smaller than their sibling counterparts and have an underdeveloped, irregularly shaped myotome. They also exhibit a dramatic scoliosis phenotype providing us with a new experimental model in which to both elucidate the pathogenetic mechanisms underlying this skeletal disorder and further define the biological functions of *sox6.*

Taken together, the results presented in this thesis identify a conserved role for Sox6 in zebrafish myogenesis, but also highlight key differences in the function of this gene between mice and zebrafish during muscle development.

# Contents

**Figures**

**Tables**

# Chapter 1

# Introduction

## 1.1 Vertebrate skeletal muscle

Vertebrate skeletal muscle is composed of two distinct types of fibre that are adapted to perform distinct functions through differences in their physiological and metabolic properties (Schiaffino and Reggiani, 2011). Slow-twitch fibres have a low contraction velocity but are rich in mitochondria and are therefore more efficient at using oxygen to generate ATP, resulting in a high endurance capability. Fast-twitch fibres, by contrast, are more suited to generating short bursts of strength or speed, but they fatigue more rapidly than slow-twitch fibres due to their high contraction velocity. Differing proportions of fibre types underpin the behavioural characteristics of many species. Most fish, for example, have a higher proportion of fast-twitch muscle, which allows them to dart away quickly from predators, whereas terrestrial herbivores, such as cows, have a large proportion of slow-twitch muscle that facilitates the sustained load-bearing necessitated by their grazing life-style. Significant fibre-type variation also occurs within species, reflecting adaptation to different environments and/or activities. In humans, for instance, variation in the proportion of slow-twitch muscle fibres underlies differences in athletic prowess; marathon runners have significantly more slow-twitch fibres than sprinters, allowing them to run for long distances without experiencing overwhelming muscle fatigue (Costill et al., 1976, Gollnick et al., 1972). Indeed, genetic differences in the fast-muscle-specific gene α-actinin-3, found to be mutated in one billion of the world's population of seven billion (North et al., 1999), may be a useful predictor of athletic performance. Although mutations in α-actinin-3 do not cause any disease in humans, homozygosity for mutations in this gene are prevalent in endurance athletes, while at least one wild-type allele is usually present in Olympic sprint athletes (Yang et al., 2003). Moreover, the varying frequency of this mutation in different ethnic groups has led to speculation about whether these variations have evolved to accommodate the differing energy expenditure requirements of particular environments, due to the existence of different selection pressures in these environments (North et al., 1999).

Experiments investigating the effects of electrical stimulation, mechanical loading, exercise and disease states have provided us with good evidence that muscle fibres are plastic and can adapt to changing functional demands. For instance, in several fish species, forced swimming against a water current, results in a series of changes that include increases in the number and size of slow muscle fibres (for review see (Davison, 1997)). These changes are bought about by activation of intracellular signalling pathways which respond to changes in intracellular calcium concentrations, leading to a reprogramming of gene expression and changes in mass, contractile properties, and metabolic states of muscle (for review see Bassel-Duby and Olson, 2006). For example, overexpression of the calcium-binding protein calcineurin, both in vitro and in vivo, induces the de-phosphorylation of the Nuclear Factor of Activated T cells (NFAT) transcription factor, causing its translocation to the nucleus, where it activates a subset of genes associated with slow-twitch myofibres, leading to a fast to slow myofibre transformation (Chin et al., 1998, Naya et al., 2000). Interestingly, recent studies have begun to establish the zebrafish as a model for studying the changes in muscle physiology in response to exercise, creating an emerging field in zebrafish in the study of exercise physiology (Palstra et al., 2010).

The importance of understanding the basis of fibre type specification is underlined by the propensity of certain muscle wasting diseases to affect particular fibre types preferentially. For example, in Duchenne Muscular Dystrophy and Sarcopenia (age-related muscle wasting), fast-twitch fibres are preferentially involved in the dystrophic process, exhibiting more deterioration and atrophy than slow-twitch fibres (Lee et al., 2006, Webster et al., 1988). A fuller understanding of the progression of these diseases and the development of effective therapies will be predicated on developing a comprehensive knowledge of fibre-type ontogeny.

In contrast to the adult fibre-type switching programme, the developmental pathways that allocate muscle progenitors to a particular fibre type have been less well characterised, with some fibre types able to differentiate in the complete absence of innervation in the developing embryo (Butler et al., 1982). The extent to which the developmental and physiological control of fibre type identity share common mechanisms remains elusive, but it appears that many of the signalling pathways activated during adult fibre-type switching are distinct from those deployed during embryonic development. The availability of immunological and molecular genetic probes (Bormioli et al., 1980, Cantini et al., 1980, Lyons et al., 1990) has facilitated the identification of molecular differences between fibres and hence the analysis of changes in fibre type during embryonic development and adult.

### 1.2 Myogenesis in the zebrafish embryo

The zebrafish has emerged as a powerful model system for the genetic analysis of vertebrate myogenesis; not only are their rapidly developing embryos optically transparent, but in contrast to amniotes, the embryonic myotome shows a discrete temporal and spatial separation of fibre type ontogeny that facilitates its analysis (Devoto et al., 1996, van Raamsdonk et al., 1982). Furthermore, unlike amniotes, the fish sclerotome is small; therefore the somites mainly give rise to the muscle, which allows clear visualisation of the fibres.

The process of mesoderm segmentation and differentiation follows a similar path in all vertebrates, although the relative timing and order of specific events may differ between species. During gastrulation, the mesoderm subdivides into the axial mesoderm, derived from dorsally located cells, and the paraxial mesoderm, derived from the more laterally located cells of the mesoderm. The axial mesoderm gives rise to the notochord whereas the paraxial mesoderm will form the somites. The paraxial mesoderm consists of two blocks of tissue, which flank the developing notochord that occupies the midline (Stickney et al., 2000). Mesodermal precursors commit to becoming myoblasts as indicated by the activation of the Myogenic Regulatory Factors (MRFs), *myoD* and *myf5,* members of the basic helix loop helix family of transcription factors (Blagden et al., 1997, Coutelle et al., 2001, Hammond et al., 2007, Weinberg et al., 1996). Inactivation of either *myoD* or *myf5* alone has no major effect on trunk muscle differentiation in zebrafish, although *myoD* mutants lack some head muscles (Hinits et al., 2011). However, simultaneous inactivation of both *myoD* and *myf5* causes a loss of muscle formation, implying significant functional overlap between the two encoded proteins (Hammond et al., 2007). In contrast to higher vertebrates, myogenesis in the zebrafish initiates prior to somite formation in a subpopulation of the presomitic mesoderm located immediately adjacent to the notochord, the so-called adaxial cells. Consistent with these observations, expression of *myoD* in the adaxial cells (Figure 1.1A) begins prior to somite formation, indicating their commitment to the myogenic fate (Coutelle et al., 2001, Weinberg et al., 1996). During adaxial cell differentiation, the majority of these cells elongate and migrate radially outwards to form a subcutaneous layer of mononucleated slow-twitch muscle fibre named the superficial slow-twitch fibres (SSF), (Devoto et al., 1996), (Figure 1.1B and D). A specialized subpopulation of adaxial cells, the muscle pioneers (MPs), retain their medial location and are characterised by their expression of the Engrailed transcription factors (Devoto et al., 1996, Felsenfeld et al., 1991, Hatta et al., 1991). The MPs later form the horizontal myoseptum that subdivides the somites into the dorsal (epaxial) and ventral (hypaxial) compartments. A second wave of *myoD* expression subsequently occurs in the posterior half of each newly formed somite (Figure 1.1A) lateral to the adaxial cells (Weinberg et al., 1996). These cells give rise to the medially located fast-twitch fibres (Stellabotte et al., 2007), which form the bulk of the myotome and begin to differentiate in the wake of the migrating adaxial cells (Henry and Amacher, 2004) (Figure 1.1D). Additional laterally located fast-twitch fibres derive from the Pax3/7 expressing dermomyotome cells (Stellabotte et al., 2007). The fast muscle progenitors mature and fuse with each other to form a multinucleated array of syncytial fibres (Roy et al., 2001) (Figure 1.1C). By 24 hours post fertilization, clear, chevron-shaped blocks of muscle are visible in the embryo.

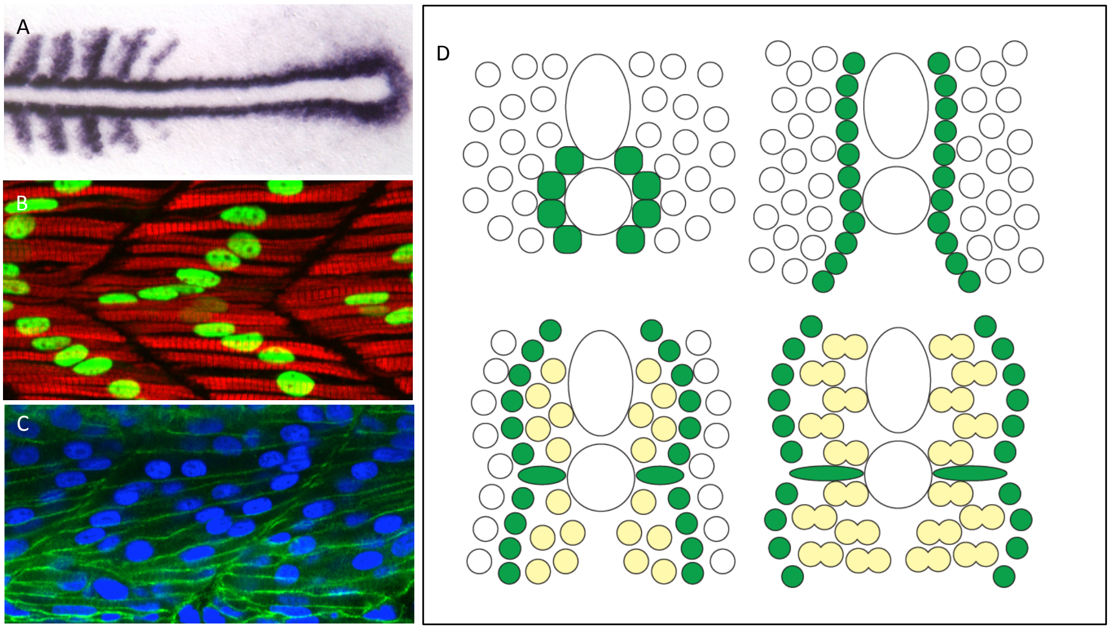


***Figure 1.1: Zebrafish myogenesis.*** *myoD expression reveals that the adaxial cells are the first cells to become committed to the myogenic programme, differentiating into slow-twitch muscle cells. The posterior half of each somite subsequently activates myoD, differentiating into fast-twitch muscle (A). Lateral view of the zebrafish trunk shows the mononucleated (green) slow-twitch fibres (red) running parallel to the notochord (B). In contrast, the fast twitch fibres (outlined in green) lie at an angle to the notochord and are multinucleated (blue) (C). A schematic representing a transverse section through the trunk of the developing myotome at four stages of development (D). Adaxial cells (green) migrate laterally through the myotome, followed by the differentiation of the fast-twitch myoblasts (yellow), which subsequently fuse to form multinucleated fibres. Panel A is from Dr. Wang Xingang and panels B-D are my own images.*

The slow-twitch fibres are the first to undergo the spontaneous contractions seen in the embryo as early as 17 hours post fertilisation (hpf), and that by 21 hpf mediate the stereotypic coiling behaviour that is exhibited by embryos in response to touch (Hirata et al., 2012, Pietri et al., 2009, Saint-Amant and Drapeau, 1998, Saint-Amant and Drapeau, 2000, Saint-Amant and Drapeau, 2001). By 48 hpf the embryo is able to propel itself forward, and will display a darting behaviour in response to touch. This behaviour is primarily reliant on fast muscle contraction, as knockdown of the fast-specific $Ca^{2+}$ release channel, ryanodine receptor 1b, which is essential for fast muscle contraction, results in a significant reduction in burst swimming, whilst coiling remains unperturbed (Naganawa and Hirata, 2011).

### 1.3 Hedgehog signalling and slow muscle development

Secreted signals emanating from the midline are critical regulators of muscle fibre identity in the developing zebrafish embryo. Three Hedgehog genes (Hh) are expressed in axial midline structures of the zebrafish. *Sonic hedgehog-a* (*shh-a*) is expressed both in the notochord and floorplate of the neural tube (Krauss et al., 1993), whereas *indian hedgehog-b (ihh-b;* formerly *echidna hedgehog, ehh)* and *shh-b* (formerly *tiggy-winkle hh, twhh*) are expressed in the notochord and floorplate respectively (Currie and Ingham, 1996, Ekker et al., 1995). In zebrafish, several lines of evidence have indicated that Hh signalling is required for the commitment of muscle precursors to the slow muscle fate (Barresi et al., 2000, Baxendale et al., 2004, Blagden et al., 1997, Du et al., 1997, Hirsinger et al., 2004, Lewis et al., 1999, Roy et al., 2001). A small group of mutants that failed to display a recognisable horizontal myoseptum were originally identified in a large-scale ENU mutagenesis screen (Haffter et al., 1996, van Eeden et al., 1996). The genes affected in this group of mutants were termed the *you-type* genes, due to their U-shaped somites. These genes appeared to play no role in the formation of the notochord, yet interestingly they exhibited defects that were very similar to those observed in mutants lacking the notochord e.g. fewer or no adaxial cells and MPs. It was therefore hypothesised that these genes may be involved in a signalling pathway between the notochord and the somites (van Eeden et al., 1996). Indeed, the mapping of these mutants lead to the discovery that most of these mutants harboured mutations in genes encoding components of the Hh pathway, including *shh-a* itself, as well as essential transducers/effectors of the pathway such as *smoothened* (*smo*), *scube2* (*you*) *dispatched-1* (*con*) and *gli2a* (*yot*). These mutants displayed a dramatic reduction of slow muscle fibres, as assessed by *slow myosin heavy chain 1* (*smyhc1*) and *prox1a* expression (Barresi et al., 2000, Hollway et al., 2006, Karlstrom et al., 1999, Kawakami et al., 2005, Lewis et al., 1999, Nakano et al., 2004, Schauerte et al., 1998, van Eeden et al., 1996, Woods and Talbot, 2005). Conversely, ectopic pathway activation, caused by mutation of the genes encoding the Hh receptors Ptch1 and Ptch2 (Koudijs et al., 2005, Koudijs et al., 2008) or by injection of *shh* mRNA or dominant negative PKA mRNA, a common negative regulator of Hh signalling in vertebrates (Hammerschmidt et al., 1996), was sufficient to induce most muscle precursors to form slow muscle at the expense of fast muscle (Blagden et al., 1997, Du et al., 1997, Kim et al., 2010).

### 1.4 Prdm1

One of the you-type mutants, *ubo,* also exhibited a slow- to fast-twitch fibre transformation phenotype (Baxendale et al., 2004, Liew et al., 2008, Roy et al., 2001, van Eeden et al., 1996, von Hofsten et al., 2008), yet *patched* (*ptc*) (a key regulator and target of the Hh pathway) expression was unaffected in homozygous *ubo* mutant embryos, demonstrating that the capacity of *ubo* mutant embryonic cells to receive a Hh signal was not compromised. The ability of Ubo to rescue mutants with compromised Hh signalling, i.e. *smo* mutants, suggested that Ubo activity is sufficient for slow-muscle development in the zebrafish (Baxendale et al., 2004). Furthermore, the *ubo* mutation acts cell autonomously (Roy et al., 2001), indicating that *ubo* is required for the responding cells to interpret the Hh signal to become slow muscle.

Positional cloning revealed that the protein encoded by *ubo* has substantial similarity to the mammalian SET domain-containing zinc-finger transcription factor B lymphocyte induced maturation protein 1 (Blimp-1) (Baxendale et al., 2004, Liew et al., 2008). Transient expression of Blimp1 (also known as Prdm1) is activated exclusively in the adaxial cells in response to Hh and was identified as the first slow lineage-specific marker to be expressed (Liew et al., 2008). Immunohistochemistry revealed that in the slow muscle precursors of anterior somites, there is reduced Prdm1a expression although expression of Prox1a remains abundant, whereas in the more posterior somites there are high levels of Prdm1a with Prox1a expression barely detectable (Liew et al., 2008). The zebrafish mutant *narrowminded (nrd)* was also found to result from a mutation in the *prdm1a* gene (Hernandez-Lagunas et al., 2005). Both *ubo* and *nrd* homozygous mutant zebrafish lack a horizontal myoseptum and have a reduced number of neural crest cells and their derivatives, Rohon-Beard neurons (Hernandez-Lagunas et al., 2005, Roy and Ng, 2004). The latter phenotype appears to be more severe in *nrd* than that of *ubo*, which can be explained by the differences in the DNA sequence changes caused by these two mutations. While *nrd* is a null allele, resulting from a truncation of the protein at amino acid 154 within the SET domain (Hernandez-Lagunas et al., 2005), *ubo* is predicted to be a hypomorphic allele resulting from a missense mutation in the second zinc-finger domain (Baxendale et al., 2004).

Blimp1 is a member of the PR domain zinc finger protein family, hence its alternative name Prdm1. Prdm family members are characterized by the presences of a PR domain and variable numbers of zinc fingers (Fumasoni et al., 2007). The PR domain is 20-30% identical to the SET domain (Fumasoni et al., 2007), a methyltransferase catalytic domain, however Prdm1 appears to lack this activity and instead recruits chromatin-remodelling enzymes to modulate transcriptional activity (Gyory et al., 2004). Prdm1 contains five DNA binding zinc fingers, two acidic regions, a proline rich region and PEST sequences (a peptide sequence rich in proline, glutamic acid, serine and threonine) (John and Garrett-Sinha, 2009). In mammals, the role of Prdm1 as a key regulator of various non-muscle cell types is well established (John and Garrett-Sinha, 2009). For example in the B-lymphocyte lineage, Prdm1 acts as a transcriptional repressor to drive plasma cell differentiation (Turner et al., 1994), with multiple domains necessary for this repression (N-terminal acidic domain, the proline rich region and the C-terminal acidic domain) (Yu et al., 2000). In most cases, it acts in conjunction with co-repressors such as histone deacetylases, histone methyltransferase enzymes and Groucho family proteins to repress target gene transcription (reviewed by John and Garrett-Sinha, 2009). This mode of activity seems to be conserved in zebrafish muscle, since forced expression of a Prdm1a-EngR fusion is sufficient to rescue slow-twitch fibre differentiation in *ubo* mutants (von Hofsten et al., 2008). Chromatin immunoprecipitation (ChIP) analysis revealed that regulatory sequences from fast-specific muscle genes, *fast myosin light chain 2 (mylz2), fast myosin heavy chain* (f*myhcx), troponin t 3a (tnnt3a)* and *troponin i2 (tnni2)* were enriched in Prdm1-precipitated chromatin verifying that Prdm1a promotes the slow twitch differentiation programme by direct repression of fast-twitch lineage genes (von Hofsten et al., 2008) (Figure 1.2). By contrast, slow-specific genes, *slow myosin heavy chain 1 (smyhc1), slow troponin c 1b (tnnc1b)* and *prox1a* were not enriched, leading to the question of how the expression of slow-twitch genes are activated in slow fibres, in a Prdm1a- dependent manner. One possible explanation was that, if Prdm1a is acting as a repressor, it might be acting indirectly through one or more proteins to activate slow-specific genes. Slow-specific gene expression might be achieved by Prdm1a- mediated repression of a repressor of slow muscle genes. Sox6 was identified as a possible candidate for such a slow-specific transcriptional repressor (von Hofsten et al., 2008) (Figure 1.2).

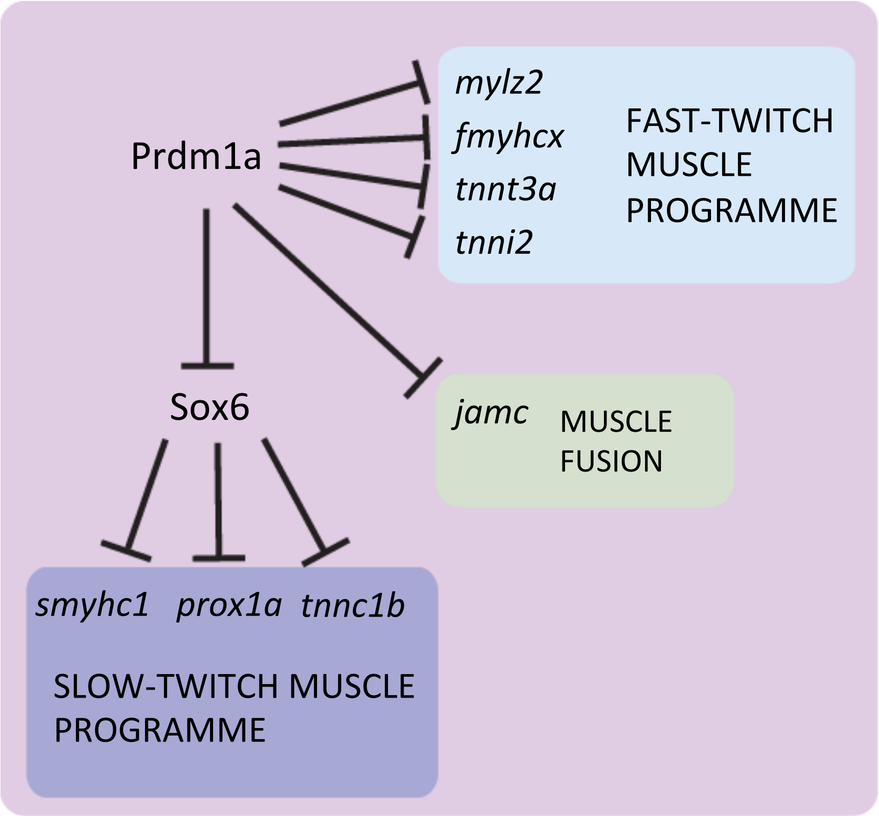


***Figure 1.2: Prdm1a acts as a switch that specifies slow-twitch muscle differentiation in adaxial cells by directly repressing the fast-twitch muscle programme and repressing repressors of the slow-twitch muscle programme.***

### 1.5 Prdm1 in amniotes

The *prdm1* gene is well conserved across the vertebrates and invertebrates. This strong conservation, combined with the fact that mouse Prdm1 protein can rescue the phenotype of a zebrafish *prdm1* mutant suggests an evolutionarily conserved function for this gene (Liew et al., 2008). *Prdm1* is expressed dynamically throughout the mouse embryo during development, including the endoderm, ectoderm and mesoderm (Chang et al., 2002). As in zebrafish, the myotomal expression of mouse *prdm1* is dependent on Hh signalling (Vincent et al., 2012). However, in stark contrast to the zebrafish, tissue-specific knockdown of *prdm1* has no detectable effect on the early formation of the mouse myotome, with normal onset of *smyhc* expression in mutant embryos (Vincent et al., 2012). Loss of *prdm1* in the muscle also had no effect on *sox6* expression, a known repressor of slow muscle in both mouse and zebrafish (Vincent et al., 2012). The exact role of Prdm1 in mouse myogenesis remains obscure. Similarly, Prdm1 expression is not restricted to fast or slow muscle cells in cultured avian somite-derived cells (Beermann et al., 2010). Unlike in zebrafish where there is only transient expression of *prdm1a* in the myotome during development (Baxendale et al., 2004, Liew et al., 2008), chick *prdm1* is expressed both at early stages of myogenesis and in terminally differentiated muscle cells. Moreover, chick *prdm1* it is also required for the expression of both slow and fast MyHC isoforms (Beermann et al., 2010), in contrast to its role in zebrafish, where *prdm1a* is only required for the expression of slow muscle genes. It is notable that even in the zebrafish, Prdm1a activity is not absolutely required for all slow fibre specification. A small population of slow fibres, distinguished by their expression of the *smyhc2* gene, have been shown to differentiate in the absence of Prdm1a function (Elworthy et al., 2008). On the other hand, Prdm1a is required for secondary slow twitch fibres that are specified independently of Hh signalling (Barresi et al., 2000). These results suggest some evolutionary divergence of the roles of Prdm1, in fish, birds and mammals.

### 1.6 Sox6

Members of the Sox family of transcription factors are characterised by the presence of a high mobility group (HMG) box, which is an evolutionarily conserved DNA binding domain consisting of 79 amino acids (Schepers et al., 2002). Unlike the majority of DNA-binding proteins that target the major groove of DNA, Sox proteins interact with the minor groove of the DNA helix and are able to induce a dramatic bend in the DNA molecule (Grosschedl et al., 1994, Wegner, 1999). This characteristic has led to the hypothesis that Sox proteins can function as architectural proteins by assembling other DNA-bound transcription factors and organising local chromatin structure (Werner and Burley, 1997, Wolffe, 1994). Indeed, although the HMG boxes of Sox genes bind to a similar AT-rich sequence motif, they are able to modulate unique sets of target genes in specific cells types through recruitment of different interaction protein partners (Kamachi et al., 2000). Thus, each Sox gene can regulate its own distinct set of target genes during the differentiation of different cell types. Sox proteins are subgrouped based on similarities in their amino acid sequence (Schepers et al., 2002). Sox6 belongs to group D, along with Sox5 and Sox13, which are distinguished by the presence of a leucine zipper and glutamine-rich domain located in the N-terminal half of the protein (Kamachi et al., 2000, Wegner, 1999). The leucine zipper of group D Sox proteins is followed by a glutamine-rich region with which it forms a contiguous coiled-coil domain (Wegner, 1999). This coiled-coil domain is one of the known functional domains of Sox D proteins along with the HMG box (Connor et al., 1995, Lefebvre et al., 1998). The leucine zipper allows homodimerization or heterodimerization between group D Sox proteins, greatly increasing the binding efficiency of the two Sox proteins to adjacent Sox sites, but also serves to mediate interactions with other proteins, examples of which are discussed below.

Sox6 is able to regulate many different processes in a variety of cell types (Hagiwara, 2011). Like other Sox proteins, it achieves specificity as part of multiprotein complexes, additional components of which are unique to different cell types. Importantly the Sox6 protein lacks a transcriptional regulatory domain (i.e. an activator or repressor), which enables it to pair with a number of different partner proteins, resulting in a range of distinct roles in transcriptional regulation, depending on the tissue-type or developmental stage (Kamachi et al., 2000). In both mice and zebrafish the ability of Sox6 to regulate a number of different processes, is reflected in the expression of the gene in a variety of distinct tissues. In zebrafish, expression of *sox6* has been reported in the somites, pectoral fin, otic vesicles, kidney, retina, optic tectum and in other specific regions of the brain (Wang et al., 2011). In some of these tissues, *sox6* plays an essential role, but in others tissues, *sox6* exhibits functional redundancy with other SoxD family members, especially *sox5.* For example during mouse chondrogenesis Sox6, Sox5, and Sox9 are important for the establishment and differentiation of chondrocytes (Hattori et al., 2008, Lefebvre et al., 1998). Sox5 and Sox6 dimerise through their coiled coil domains to greatly increase the binding efficiency of the two Sox proteins to DNA sequences that contain adjacent Sox sites to directly activate chondrocyte-specific gene markers (Lefebvre et al., 1998). In *sox5*$^{-/-}$ and *sox6*$^{-/-}$ single null knockout mice, the mineralization of only a small subset of endochondral elements are affected. However, in double mutant *sox5*$^{-/-}$;*sox6*$^{-/-}$ knockout mice, there is severe underdevelopment of the skeleton and cartilage (Smits et al., 2001). This indicates the existence of functional redundancy between these two transcription factors, although this is not complete, as there is a minor effect on cartilage formation in the *sox5*$^{-/-}$ and *sox6*$^{-/-}$ single mutants (Smits et al., 2001).

During development of the chondrocyte lineage, Sox5 and Sox6 further cooperate with Sox9 to suppress the entry of chondroblasts into the cell cycle (Han and Lefebvre, 2008, Lefebvre et al., 1998). Moreover, chondrogenesis depends on interactions between Sox proteins and histone modifying enzymes. These enzymes are able to change the transcriptional status of a gene reversibly by modifying core histones, so called epigenetic changes, by acetylation or methylation of basic amino acids in their N-terminal domains. Histone Acetyltransferases (HATs) add acetyl groups onto certain histone residues, whereas Histone Deacetylases (HDACs) are able to remove these acetylation marks. Sox9 interacts with HAT Tip60 through its C-terminal transactivation domain in order to regulate the expression of cartilage-specific genes (Hattori et al., 2008). The acetylation of histones by Tip60 may be important for the ability of Sox9 to enhance transcriptional activity of cartilage-specific genes. Similarly, in the pancreas Sox6 regulates glucose-stimulated insulin secretion from β-cells by repressing the Cyclin D1 promoter through recruitment of the histone deacetylase HDAC1 and interacting with promoter-bound β-catenin protein, rather than by binding directly to Cyclin D1 (Iguchi et al., 2007).

In recent years, there has been a vast increase in our understanding of the biological and molecular functions of SoxD proteins in mammals. The role of this group of proteins in zebrafish however, is less well characterised.

## 1.7 Sox6 in myogenesis

Sox6 was first implicated in muscle fibre differentiation through studies in mouse (Hagiwara et al., 2000, Hagiwara et al., 2005, Hagiwara et al., 2007). Microarray analyses performed on mice homozygous for a targeted mutation of the *sox6* gene showed a significant increase in the expression of slow-specific genes as well as a significant decrease in the expression of fast-specific genes compared to wild-type littermates (Hagiwara et al., 2005). These results suggested that the loss of *sox6* expression could relieve the transcriptional repression of slow-specific genes in foetal muscle. In zebrafish, transcription of *sox6* is activated specifically in fast twitch fibre progenitors at the 10-somite stage in response to MyoD and Myf5 activity (Wang et al., 2011). Forced expression of *sox6* in adaxial cells in wild-type embryos inhibited Prox1a expression, whereas morpholino mediated knockdown of Sox6 resulted in a partial rescue of Prox1a*,* complete rescue of *smyhc1* expression in the adaxial cells of *ubo* mutants and caused robust ectopic expression of the slow twitch specific *tnnc1b* gene in fast muscle fibres of genetically wild-type fish (von Hofsten et al. 2008). These findings are consistent with Sox6 acting to repress the slow twitch differentiation programme in fast twitch myoblasts. In *ubo/nrd* mutant embryos, *sox6* is ectopically expressed in adaxial cells (von Hofsten et al. 2008), indicating that its repression by Prdm1a underlies the activation of slow-lineage-specific genes in these cells. However, whether or not *sox6* is a direct target of Prdm1a remains unclear; no enrichment of Sox6 regulatory sequences by Prdm1a ChIP has been reported and mutation of potential Prdm1 binding sites in the *cis*-acting regulatory elements of *sox6* had no effect on reporter gene expression (Wang et al. 2011).

Although these data support a simple model whereby Prdm1a acts to repress fast-specific genes and promote slow-specific gene expression by repressing *sox6*, the inability of the Sox6 morpholino to cause in ectopic expression of all slow muscle genes in the fast muscle implies that either the knockdown of Sox6 was incomplete or other factors are involved in this regulatory network.

## 1.8 Post-transcriptional regulation of fibre-type identity

Another conundrum posed by the transient nature of *prdm1a* expression in adaxial cells, is how the repression of Sox6 activity is maintained in slow-twitch progenitors once *prdm1a* expression has been extinguished. One explanation, given its capacity to direct histone methylation (Ancelin et al., 2006), could be that Prdm1a is able to cause the transcriptional shutdown of *sox6* through chromatin modifications. It is however, unclear whether *sox6* is a direct target of Prdm1a; mutation of potential Prdm1 binding sites in the *cis*-acting regulatory elements of *sox6* had no effect on reporter gene expression (Wang et al., 2011). In any case, the slow-specific repression of *sox6* transcription is lost by early larval stages, arguing against this model (Wang et al., 2011). Notably however, the accumulation of Sox6 protein remains restricted to the fast twitch fibres suggesting that repression is maintained at the post-transcriptional level (Wang et al. 2011). Recent studies have shown that the microRNA miR-499 can regulate *sox6* transcript levels in mammalian myofibres (McCarthy et al. 2009; van Rooij et al 2009). MicroRNAs inhibit translation by binding to the 3' UTR of a gene. Indeed, in zebrafish, the endogenous *sox6* 3'UTR, can restrict reporter gene expression to the fast fibres at 6 dpf (Wang et al., 2011). The zebrafish *sox6* 3'UTR contains consensus recognition sites for the SEED sequence of miR-499, which is encoded by the intron of the slow specific gene *slow myosin heavy chain b (Myh7b)*. When these sites are mutated in the *sox6:eGFP* construct containing the endogenous *sox6* 3'UTR ectopic EGFP expression appears in slow specific fibres. Moreover misexpression of miR-499 in the fast muscle using the UAS;GAL4 system results in a loss of *sox6* in the fast-twitch muscle, revealing miR-499 is sufficient to repress *sox6* ( Wang et al. 2011).

Whether fast-specific gene expression is also regulated by miR-499 remains to be answered; if *sox6* repression needs to be maintained by miR-499 in larvae after the expression of Prdm1a is extinguished at 24 hpf (Liew et al., 2008), this suggests that fast-specific genes may also be repressed by miR-499.

Taken together, these studies suggest a Shh-activated gene regulatory network through which the slow-twitch lineage is established at least in part by Prdm1a-mediated transcriptional repression of Sox6, a repressor of a sub-set of slow-twitch specific genes. This network is then maintained by post-translational repression of *sox6* by miR-499, expression of whose host gene, *myh7b*, is itself negatively regulated by Sox6, establishing a two-way feedback loop (Figure 1.3). The implication of the indirect repression of *sox6* by Prdm1a and the inability of the loss of *sox6* to cause ectopic expression of all slow twitch fibre specific genes indicates that other key components of this network remain to be discovered.

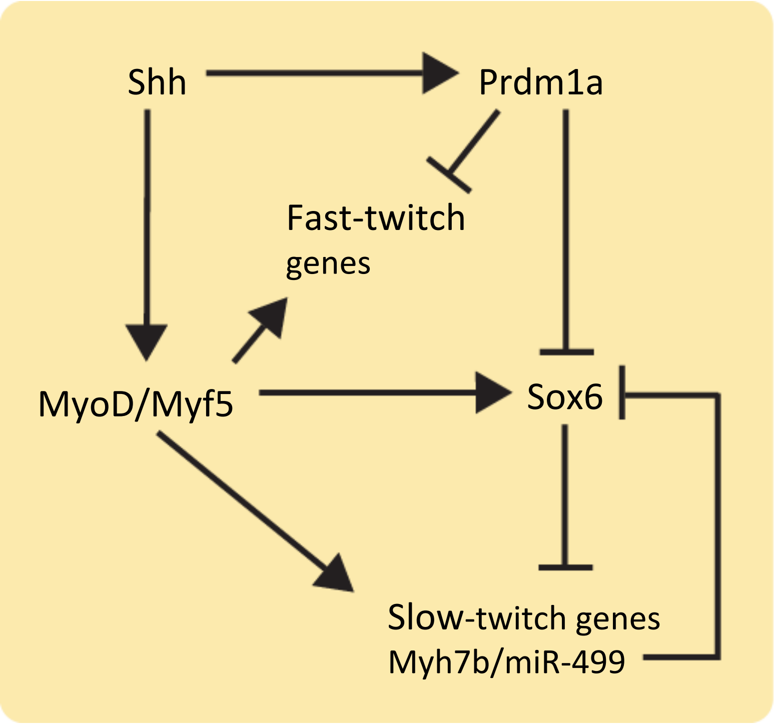


***Figure 1.3: Shh signalling activates a gene regulatory network that specifies and maintains the slow-twitch fibre lineage through transcriptional and post-transcriptional feedback regulation.***

### 1.9 Sox6 in amniotes

The role for Sox6 in myogenesis appears to be at least partially conserved. Sox6 directly represses slow-fibre specific genes in mice, as indicated by ChIP-seq analysis (An et al., 2011), whereas sox6 mutant mice exhibit an increase in slow-twitch fibres and a concomitant decrease in fast-twitch fibres at foetal stages (Hagiwara et al., 2005, Hagiwara et al., 2007). Consistent with this observation, but in contrast to the situation in zebrafish, expression of the slow myosin heavy chain isoform is up-regulated in *sox6* mutant foetal fibres in mice.

A further similarity between mammals and fish is found in the regulation of Sox6 by micro RNAs. As mentioned above, miR-499, encoded by *myh7b*, regulates sox6 transcript levels in both fish and mammalian myofibres. Contrastingly, the microRNA-mediated inhibition of *sox6* in zebrafish leads to the inhibition of translation, whereas in mouse it appears miR-499 promotes the degradation of the transcript (McCarthy et al., 2009, van Rooij et al., 2009).

In addition to its role in regulating the differentiation of fibre types in the mouse embryo, Sox6 also plays an essential role in maintaining fibre -type identity in adult mice (Quiat et al., 2011). The conditional knockout of *sox6* in adult mouse muscle results in a conversion of fast muscle fibres to a slow muscle phenotype. The mechanism for this switch appears to be independent of other known fibre-type switching pathways such as Pgc-1αm, a transcriptional coactivator sufficient to drive slow fibre formation, and phosphorylated AMPK, a regulator of muscle fibre-type transition in response to endurance exercise (Quiat et al., 2011). In muscular atrophy studies in adult mice, miR-499 represses *sox6* such that when miR-499 expression is decreased in response to hindlimb suspension, this results leads to increased expression of *sox6* and a subsequent decrease in *smyhc* isoforms (McCarthy et al., 2009). This example provides rare evidence of a transcription factor acting in both the developmental and physiological programmes of fibre type differentiation. It would be interesting to explore whether this fibre type-switching role for Sox6 is conserved across other species, and whether *sox6* and *miR-499* have roles in atrophy and fibre switching in adult fish.

A perhaps notable difference between fish and mammals is the expression pattern of *sox6*. It is unclear from the literature whether, like zebrafish, sox6 protein is excluded from the slow muscle in the mouse. However, reports suggest that in contrast to zebrafish, *sox6* is expressed in slow fibres of adult mice, albeit at a significantly lower level than its expression level in the fast muscle (Quiat et al., 2011), with its expression in mouse embryos remaining obscure. This raises the question of how repression of slow fibre genes is specific to the fast muscle and whether there is a role for *sox6* in the slow muscle. The importance of Sox6 in both mouse and zebrafish muscle fibre type differentiation and maintenance implies a conserved function, yet key differences are apparent in its regulation.

### 1.10 Adaxial cell migration

Shortly after adaxial cells commit to the slow-twitch lineage, the majority of these cells begin to traverse the myotome before differentiating into the subcutaneous layer of slow-twitch fibres (Devoto et al., 1996). Only the Muscle Pioneers (MPs) remain in their medial location and elongate to span the width of the myotome, so forming the horizontal myoseptum. Relatively little is known about the control of this migration and whether it is directed by a secreted signalling molecule, or if instead it is contact dependent.

Cortes et al (2003) highlighted the dynamic and complementary expression patterns of the calcium-dependent cell adhesion proteins, M-Cadherin and N-Cadherin, implicating them as key mediators of adaxial cell migration. Waves of expression of the genes encoding both proteins pass through the myotome as the adaxial cells migrate, with high-level expression of both genes uniquely defining the migrating cells. According to Cortes et al, it is the homophilic interaction between neighbouring cells promoted by both proteins that drives adaxial cell migration. Consistent with this view, expression of M-cadherin is down-regulated specifically in the non-migratory MPs, while loss of either cadherin or overexpression of N-cadherin disrupts adaxial cell migration (Cortes et al., 2003).

Notably, however, most adaxial cells fail to complete their migration in the absence of Prdm1a function (Baxendale et al., 2004, Roy et al., 2001, von Hofsten et al., 2008). Wild-type slow muscle progenitor transplants are able to differentiate into mononucleated fibres and migrate accordingly in *prdm1a* mutants, indicating that Prdm1a is acting cell-autonomously (Roy et al. 2001). This finding implicates *prdm1a* as a key factor in slow fibre migration. Although Cortes et al noted increased repression of M-Cadherin close to the midline in *prdm1a* mutants, its expression in the majority of the adaxial cell population appears unaffected by this mutation, therefore the spatially regulated expression of cadherins alone is not sufficient to drive migration. The differential expression other genes in SSF and MPs, such as *engrailed* could be accountable for this migratory behaviour, but this possibility remains to be investigated. These observations also raise the issue of whether MPs are specified as such because they don't migrate or rather, they don't migrate because they are specified. Overall, the migration of slow muscle precursor through the zebrafish myotome is a process that remains poorly understood, and more research is needed to identify the protein components of the mechanisms involved in this process.

### 1.11 Fast muscle specification

The majority of the zebrafish myotome is composed of multinucleated fast-twitch muscle fibres. Myoblasts become committed to the fast twitch lineage after their slow-twitch counterparts, and the differentiation of fast-twitch myoblasts begins only after adaxial cells have migrated past them.

The delayed commitment of cells to the fast lineage is reflected in the later onset of MyoD expression in the characteristic "wings" in each somite, from which the fast fibres are derived. MyoD expression is reduced in the *fgf8* mutant *acerebellar* (Reifers et al., 1998) suggesting a possible role for FGFs in activating the Shh-independent expression of *myoD*. In *fgf8* morphants, *myoD* is lost in the lateral fast muscle precursors, concomitant with a loss of fast myosin heavy chain expression (Groves et al., 2005), suggesting a requirement for FGFs in fast muscle differentiation. Evidence suggests that Retinoic Acid (RA) is responsible for activating *fgf8* expression, with inhibition of RA resulting in the downregulation *fgf8* and its target genes (Hamade et al., 2006). Interestingly the more medial fast muscle precursors express *myoD* independently of FGF8 and Hh signalling. The distinction between these FGF8- and Hh- independent fast-twitch fibres and the *eng*-positive Medial Fast Fibres (MFFs) is unclear as *eng* expression was not assessed in these residual fibres. However, these FGF8- and Hh- independent fibres can be ablated by mutations that abolish midline patterning, suggesting dependence on another midline signal (Groves et al. 2005). Although Hh signalling is not required for the early specification of fast muscle fibres, there is a later role for Hh in regulating the differentiation of the dermomyotome into fast muscle fibres (Feng et al., 2006).

The homeodomain transcription factor Six1 has been implicated in the specification of fast muscle fibres in both mice and fish (Bessarab et al., 2008, Grifone et al., 2004, Niro et al., 2010). Expression of the MRF, *myogenin (myog),* follows the expression of *myod* in both fast and slow muscle precursors (Weinberg et al., 1996). In zebrafish *six1a* morphants, early *myog* expression is lost exclusively from fast muscle precursors with a concomitant reduction in expression of fast muscle genes; however the expression of both *myog* and fast muscle genes is restored at later stages (Bessarab et al., 2008). This indicates that although *six1a* is important in controlling the onset of fast muscle differentiation, other factors are also needed for the regulation of *myog* in fast precursors, and also that the expression of *myog* is differentially regulated in fast and slow muscle precursors. A second six gene, *six4a* (*six4.2*) is also expressed in the somites, although unlike *six1a* its expression is not excluded from the adaxial cells (Kobayashi et al., 2000). The role of six4a in zebafish myogenesis remains unknown but the protein could be responsible for two things: (a) it could be responsible for the later activation of *myog* and fast specific genes in six1a morphants, and (b) it could also be necessary for the initial activation of myog in the adaxial cells, a process in which *six1a* does not play a part.

Similarly, the Six proteins appear to play a role in activating the fast-specific network of genes in mice. The knockdown of both *six1* and *six4* results in a downregulation of *myog* and fast-specific genes in the primary myotome, with ChIP analysis revealing a direct interaction between fast genes and Six homeoproteins (Niro et al., 2010). Curiously it appears that the slow-specific gene *tnnc1* is also directly downregulated in six mutants; unlike in adult mice, fast-specific and slow-specific genes are co-expressed in the myocytes present in the primary myotome (Niro et al., 2010).

In adult mice, overexpression of *six1* reveals its ability to transform a slow/oxidative phenotype to a fast/glycolytic phenotype (Grifone et al., 2004). Like Sox6, it appears that Six1 stands out as one of the few transcription factors involved in both embryonic muscle fibre-type specification and fibre-type switching in adults. The conserved roles of Six proteins in activating fast muscle genes during embryogenesis, suggests that as in mice, Six proteins may also have a role in adult fish.

A second set of homeodomain proteins, the Pbx proteins, have also been associated with the activation of fast-specific muscle genes (Maves et al., 2007). Pbx appears to modulate the broad range of MyoD targets, directing cells to a fast twitch fate. Although Pbx proteins are expressed in both the presumptive slow and fast muscle, knockdown of Pbx only results in downregulation of fast specific genes, leaving slow-twitch genes unaffected. More specifically, it is only the lateral fast fibres that are affected, with the *eng* positive fibres appearing normal (Maves et al., 2007). How Pbx proteins facilitate the activation of a specific subset of genes by MyoD remains unknown; one possibility is that they interact with Six proteins to regulate this network of genes.

## 1.12 Myoblast fusion

Fast-twitch muscle precursors fuse with each other to form a multinucleated array of syncytial fibres, in contrast to slow twitch precursors, which are fusion incompetent and mature into mononucleated fibres (Devoto et al., 1996, Roy et al., 2001). In Drosophila, *Dumbfounded*, an Ig-domain transmembrane protein, is essential for fusion of muscle founder cells to other myoblasts and morpholino mediated knockdown of its zebrafish orthologue, *kirrel3l*, indicates this function is conserved across species (Srinivas et al. 2007). Notably, *kirrel3l* is expressed exclusively in fast-twitch progenitors, suggesting that its repression in adaxial cells could underlie the failure of slow-twitch myoblasts to undergo fusion. Surprisingly, however, *kirrel3l* remains repressed in adaxial cells in *ubo* mutants, indicating that its function is not required for their aberrant fusion and implying that that other genes repressed by Prdm1a must be sufficient to promote fusion. A good candidate for such a gene is that encoding the junctional adhesion molecule JamC.

Recently, a novel, vertebrate-specific cell surface receptor pair essential for myocyte fusion has been identified in the zebrafish using an assay developed for identifying brief and weak interactions between extracellular surface proteins (Powell and Wright, 2011). *jamb* and *jamc* are expressed highly in fast muscle precursors and mutations in either gene result in a near complete loss of fast muscle fusion. Although Jamb and Jamc can form homodimers, the heterophilic interaction between them is significantly stronger, and transplant experiments revealed the requirement of this heterophilic interaction for muscle fusion. Notably, in *prdm1a* mutants, *jamc* is ectopically expressed, whereas the expression of *jamb*, similar to *kirrel3l*, remains unchanged. The misexpression of *jamc* in the slow muscle does not however trigger fusion of these fibres, implying that *prdm1a* may repress expression of other factors required for myoblast fusion (Powell and Wright, 2011).

Transducers of the intracellular signalling pathways associated with myoblast fusion also appear to be conserved between flies and fish. The gene *myoblast city* in Drosophila and its homolog in zebrafish, *dedicator of cytokinesis 1 (DOCK1),* are key components of the intracellular fusion pathway (Erickson et al., 1997, Moore et al., 2007). Loss of function of DOCK1 and a closely related protein, DOCK5, results in elongated, mononucleated, fibres that are still capable of expressing fast muscle specific genes. Conversely, in mice, only loss of DOCK1 results in a myoblast fusion defect, with loss of DOCK5 having no effect on fusion (Laurin et al., 2008) . Similarly, knockdown of the adaptor proteins Crk and Crk-like (Crkl), which are known to interact physically with DOCK proteins, also blocks myoblast fusion in the zebrafish (Moore et al. 2007). Interestingly, overexpression of the *crk* or *crkl* mRNA results in an enhancement in fusion, not seen when either DOCK1 or DOCK5 are overexpressed (Moore et al., 2007). A constitutively active form of the GTPase, Rac, the most downstream component of the fusion signal, also results in the formation of giant syncytia (Srinivas et al., 2007). This hyperfusion is dependent on the fusion receptor Kirrel3l, as constitutively active Rac in *kirrel3l* morphants is unable to cause an increase in fusion. The dependence of Crk and Crkl on upstream fusion receptors to cause hyperfusion has not been analysed. Very recently, Casein kinase 2 interacting protein-1 (CKIP-1) has been implicated in zebrafish myoblast fusion (Baas et al., 2012). *In vitro* studies have suggested that the role of CKIP-1 may be conserved in mice, and knockdown of CKIP-1 in both zebrafish and mouse cultured myoblasts blocks myoblast fusion (Baas et al., 2012).

The question still remains as to how all the identified components of zebrafish myoblast fusion integrate to form an organised syncytial array of exclusively fast myoblasts that fuses a finite number of times at a specific developmental timepoint. How is the downregulation of these components regulated in order to terminate the muscle fusion programme? The association and integration of all of these proteins is crucial to understanding the mechanisms of zebrafish myoblast fusion and could provide insights into other cell fusion events, both in teleosts and mammals.

### 1.13 Diversity within lineages

Within both the slow and fast lineage, distinct sub-types of fibre can be distinguished both on the basis of their medial location as well as by their expression of the Engrailed (Eng) family of homeodomain transcription factors (Roy et al., 2001). Muscle Pioneers (MPs) constitute a sub-set of the slow-twitch fibres that derive from the adaxial cells, but they are distinguished from the SSFs by their location as well as by expression of *eng* (Hatta et al., 1991). During somite maturation, the MPs elongate to span the entire width of the somite, separating the dorsal and ventral domains of the myotome by forming the horizontal myoseptum (Figure 1.4). The MPs also provide instructive cues for migrating motor neurons (Melancon et al., 1997) and thus play a critical role in the patterning of the locomotor system. Consistent with their adaxial origin, MPs, like SSFs, are completely absent from embryos that lack Hh pathway activity (Lewis et al., 1999, van Eeden et al., 1996); however, mutant embryos that have reduced Hh pathway activity, lack MPs whilst retaining SSFs (Lewis et al., 1999). This observation implies that the level of Hh activity to which an adaxial cell is exposed determines whether it will follow the SSF or MP differentiation pathway. Consistent with this, high level ectopic activation of the Hh pathway drives differentiation of myoblasts into MPs at the expense of SSFs (Currie and Ingham, 1996, Wolff et al., 2003). This finding poses the question as to how adaxial cells come to be exposed to differing levels of Hh activity. Transient inhibition of the Hh pathway using cyclopamine showed that MP induction requires a longer exposure to Hh activity than is sufficient for SSF induction (Hirsinger et al., 2004), while cell-labelling experiments show that MPs arise from the most medially located adaxial cells (Nguyen-Chi et al., 2012), which might be expected to receive the highest levels of midline-derived Hh activity. The response of myoblasts to Hh signalling can be attenuated by the experimental manipulation of BMP activity (Du et al., 1997, Kawakami et al., 2005, Maurya et al., 2011), suggesting that these two signals may act in concert to specify different cell types. In line with this, high levels of phosphorylated Smad (pSmads) accumulate in the nuclei of most adaxial cells but are excluded from the nuclei of MPs, which is consistent with the possibility that BMP signalling plays an inhibitory role in MP specification (Dolez et al., 2011, Maurya et al., 2011). ChIP analysis suggests that both Gli proteins and pSmads can directly bind to the *eng2a* promoter (Maurya et al. 2012), although no consensus Gli binding site has been found in this genomic region. High BMP signalling from the dorsal and ventral edges of embryos causes accumulation of pSmads in the nucleus (Kawakami et al., 2005, Nguyen-Chi et al., 2012), which might allow the direct inhibition of the *eng2a* promoter. Further to this, laminins appear to also play a key role in establishing the *eng*-positive central domain by sequestering BMPs (Dolez et al. 2011). *eng* expression in the lamininC1 (LamC1) mutant *sly,* is severely reduced or absent in the MPs, with the knockdown of BMP signalling able to rescue it (Dolez et al., 2011). This interaction between BMPs and lamC1 appears to be mediated via heparin sulphate proteoglycans (HSPG) as their inhibition in wild type embryos leads to the ectopic activation of BMP target genes and therefore an inhibition of *eng* expression (Dolez et al., 2011). The crosstalk between these two pathways seems to be regulated at multiple levels: a physical sequestering of BMPs by HSPGs, as well as a regulated balance between the relative levels of the effectors of the two pathways. Remarkably, the nuclear accumulation of pSmad is dependent not only upon BMP activity but also on Hh signalling; in embryos in which the Hh pathway is constitutively activated, the nuclear accumulation of pSmad is significantly attenuated. Various lines of evidence suggest that this regulation occurs downstream of the BMP receptor (Maurya et al., 2011), though exactly how it is meditated remains to be determined.

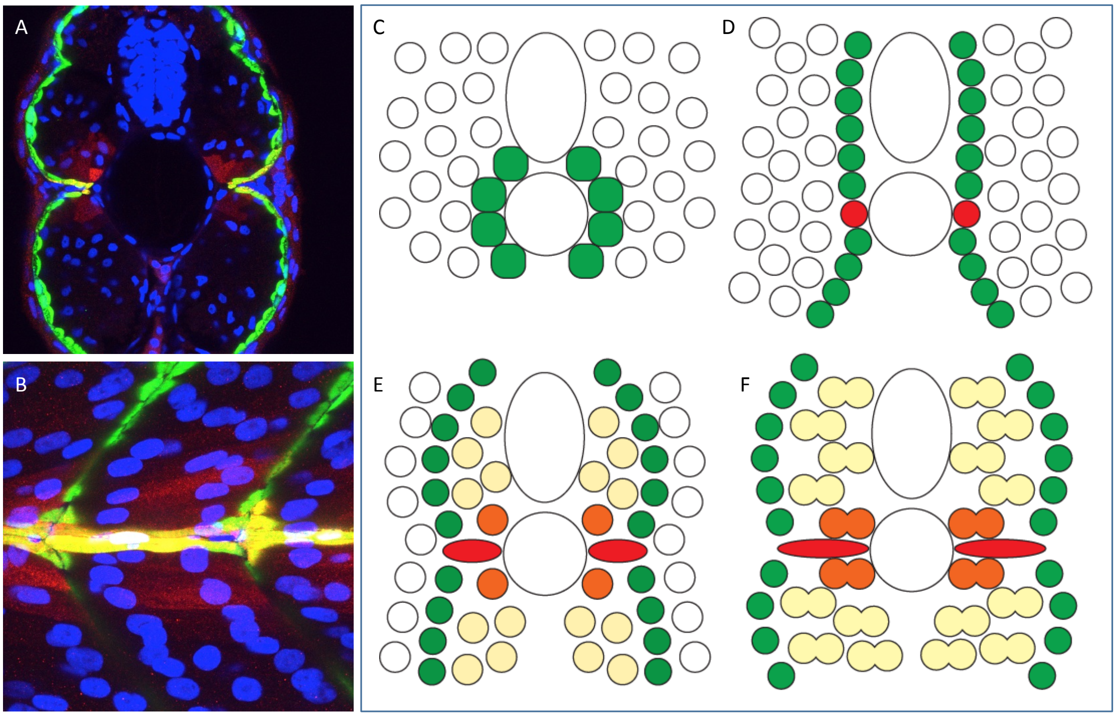


***Figure 1.4: Sub-populations within both the fast and slow-twitch lineages express Engrailed family homeodomain proteins.*** *Eng positive cells (red) can be identified in both the slow muscle, identifiable by the expression of smyhc1:GFP (green); the muscle pioneers (MPs) and in the fast muscle; the medial fast fibres (MFFs) (A and B). The adaxial cells (green) lying closest to the notochord activate expression of Eng proteins (red) (C and D). These eng expressing adaxial cells do not migrate, but instead form muscle pioneers, which elongate to form the horizontal myoseptum. Following the migration of the adaxial cells, the most medial fast-twitch myoblasts are displaced and become juxtaposed to the notochord and also express eng. These eng-expressing cells will differentiate into multinucleated medial fast fibres (E and F). All panels are my own images.*

However, this still does not answer the question as to how the adaxial cells give rise to two different types of slow muscle cells along the anterior posterior boundary, where they are exposed to similar levels of BMP and Hh signalling. Cell tracing experiments have indicated that only the most anterior member of the medial row of adaxial cells is destined to differentiate into an MP (Nguyen-Chi et al., 2012), implying that an additional signal is required to single out this cell. A recent explanation for this implicates the involvement of FGFs in this process. FGF signalling appears to be restricted to the posterior half of the somites by the FGF inhibitor *sprouty4* (Nguyen-Chi et al., 2012). Overexpression of *sprouty4* doubled the number of MPs within the embryo whereas a dominant negative form of *sprouty4* resulted in a loss of MPs (Nguyen-Chi et al., 2012). Contradictory to this, another study has suggested that the numbers of MPs is decreased in the absence of FGF signalling, based on absence of *eng2* expression in the acerebellar mutant (Reifers et al. 1998).

Interestingly, inhibition of FGF signalling by overexpression of *sprouty4*, induced a unique population of *eng* expressing cells that are able to migrate, a behaviour never seen in wild-type embryos (Nguyen-Chi et al., 2012). These observations suggest that FGF signalling may play a role in regulating the migration of myogenic precursors; further experiments are required to establish this definitively.

Another *eng* positive population exist, but in contrast do not express slow specific genes such as *smyhc1*. Instead these cells form fast-twitch fibres, expressing fast myosin isoforms and displaying multinucleated fused fibres. These fibres, named Medial Fast Fibres (MFF), lie lateral to the MPs and are specified at a later stage after slow muscle cell migration (Figure 1.4). Surprisingly, despite their fast-twitch identity, these cells are also Hh responsive, with Hh acting in a concentration dependent manner to induce different cell identities (Wolff et al., 2003). Conditional inhibition of Shh signalling in zebrafish can be achieved by administration of cyclopamine, which revealed that MPs require a higher level of Hh signalling than that of MFFs, which in turn require higher levels than SSFs (Wolff et al., 2003). As described previously, the onset of slow-twitch differentiation precedes that of the fast-twitch lineage programme, therefore the migration of the SSF displaces the fast-twitch precursors leaving them closer to the Shh- secreting notochord (Cortes et al., 2003), which exposes these cells to a higher concentration of Shh, to which they respond by up-regulation of the Hh target gene *ptch2* (Concordet et al., 1996). Like the adaxial cells before them, the majority of these fast–twitch progenitors respond to BMP signalling by accumulating high levels of pSmad in their nuclei. However, the most medially located fast-twitch progenitors lack such pSmad accumulation and it is these cells that activate Eng transcription (Maurya et al., 2011). Thus, as for the MPs, the activation of Eng in MFFs depends upon the antagonistic effect of Hh on BMP activity.

One question is that why MFFs don't differentiate into slow fibres? Although ectopic Hh can induce supernumerary slow fibres, this competence becomes lost at later stages (Wolff et al., 2003), and it is unable to induce *prdm1a* expression. Secondary to that, heatshock induced Prdm1a is unable to increase the slow muscle fibres after 18 hpf (Liew et al. 2008). This indicates that myoblasts lose their competence to respond to the slow twitch lineage programme and express Prdm1a, and instead go on to give rise to the MFFs.

The restricted expression of *eng* and its precise regulation in the myotome raises the question of its function. It could have a role in myogenesis, perhaps by regulating the expression of genes involved in slow muscle migration, although this would be a time-dependent role, as MFF do not migrate. Alternatively, as MPs act as signalling centres for controlling axon guidance (Melancon et al., 1997) and neural crest migration (Honjo and Eisen, 2005), *eng* could have a role in regulating genes involved in these processes. Furthermore, it remains an open question as to whether Eng has the same role in both the MFF and MPs. These two cells types have different gene expression profiles and morphology (mononucleated vs. multinucleated), perhaps indicating distinct roles for this transcription factor in the two populations, or instead it is possible that *eng* fulfils a role in these cells that is unrelated to myogenesis.

### 1.14 Engrailed positive domain in amniote muscle

Hh and BMP also act to regulate the expression of *eng* in amniotes, and this gene expression domain demarcates the epaxial and hypaxial boundary, a structure possibly homologous to the fish myoseptum (Cheng et al., 2004). Shh-soaked beads are able to rescue *eng* expression in chicks with ablated notochords, whereas BMP4-soaked beads are able to extinguish *eng* expression (Cheng et al., 2004). Differences are apparent in these populations of cells between species, namely the timing of signalling events and the exact location of signal sources as well as responsive tissue. The Engrailed-expressing cells are formed at a relatively late timepoint in amniotes, with the expression domain significantly larger, extending to the dermomyotome, myotome and dermatome (Cheng et al., 2004), in contrast to the situation in zebrafish, where this domain is restricted to a thin layer of cells in the midline (Roy et al., 2001). Nevertheless, an *eng*- expressing population exists in both zebrafish and amniotes that originates close to the midline, is dependent on Hh signalling and is antagonised by BMPs, highlighting the striking similarities between the species, and perhaps identifying a homologous population of cells.

## 1.15 Troponins

The basic contractile unit of striated muscle is the sarcomere, which is composed of myosin thick filaments, intercalated with actin thin filament, each of which is associated with a tropomyosin and a troponin regulatory complex. This complex acts in a calcium-responsive manner to regulate contractions in striated muscle and consist of three subunits: troponin c, the calcium-binding subunit, troponin T, the tropomyosin-binding subunit, and troponin I, which inhibits the ATP-ase activity of actomyosin (Greaser and Gergely, 1973). Muscle cell contraction is generated from the cyclic interactions of the myosin heads and the actin subunits of the thin filaments forming a cross-bridge formation. In the absence of calcium, troponin and tropomyosin interact and inhibit interactions between actin and myosin. Upon depolarization of the muscle cell, the calcium released into the cytoplasm binds to Troponin C, causing a conformational change in the tropomyosin-troponin regulatory complex, leading to the derepression of the actin-myosin interactions and subsequent muscle contraction (Lehman et al., 1994).

Muscle fibres can be identified through their distinct gene expression profiles. Each troponin is composed of multiple isoforms that are encoded by distinct genes. These isoforms are expressed in a tissue specific manner, be it slow-twitch, fast-twitch or cardiac muscle, and at a specific developmental stage. In mammals, there is one *fast troponin c* gene that is expressed specifically in the fast muscle, and one *slow troponin c (Tnnc1*) gene that is expressed in both the slow-twitch skeletal muscle and the cardiac muscle (Parmacek and Leiden, 1991). Precise molecular mechanisms regulate the expression of *Tnnc1,* with respect to its tissue specificity. Studies have shown that a skeletal muscle-specific, developmentally regulated transcriptional enhancer is located within the first intron of the mouse *Tnnc1* gene (Parmacek et al., 1994). Conversely, a distinct and independent enhancer located in the immediate 5' flanking region of the gene drives *Tnnc1* expression in cardiac myocytes (Parmacek et al., 1992). In contrast, zebrafish have two paralogous genes, *tnnc1a* and *tnnc1b*, which are expressed specifically in the heart and slow-twitch skeletal muscle, respectively (Sogah et al., 2010). Nothing is known about the regulation of these two different genes in zebrafish.

Tnnc1 is an important gene clinically as mutations in this gene have been linked to certain forms of familial cardiomyopathy (Hoffmann et al., 2001, Mogensen et al., 2004). A mutation in the heart-specific *troponin c* in zebrafish, *tnnc1a*, results in defective ventricular contractility, with the ventricle failing to express ventricular myosin heavy chain (Sogah et al., 2010). The loss of ventricular myosin heavy chain expression suggests that contractile function of the ventricle may be required for the sustained expression of this gene. The high conservation of this gene between zebrafish and humans (91.3%)

means that understanding the regulation and mechanisms of this gene in zebrafish could lead to answers in understanding certain cardiomyopathies.

# Chapter 2

# Materials and Methods

## 2.1 Fish stocks and maintenance

### *2.1.1 Fish maintenance*

Adult fish were maintained on a 14h light/ 10h dark cycle at 28°C in the AVA (Singapore) certificated IMCB zebrafish facility. All experiments were subject to A*STAR IACUC approval. Zebrafish were maintained and staged following standard methods (Kimmel et al., 1995).

### *2.1.2 Mutants*

Mutant strains of zebrafish were maintained as heterozygote and homozygote embryos were obtained by incrossing the heterozygote adults. Zebrafish mutant lines *ubo*$^{tp39}$, *nrd*$^{m805}$ and *smu*$^{b641}$ have been described previously (Barresi et al., 2000, Hernandez-Lagunas et al., 2005, van Eeden et al., 1996, Varga et al., 2001). Homozygous mutants were identified by their U-shaped somites. *sox6* and *sox5* mutant lines were generated for this thesis (described later).

### *2.1.3 Transgenic lines*

Transgenic lines *Tg(smyhc1:eGFP)*$^{i108}$, *Tg(ubo:GFP)*$^{i106}$, *Tg(smyhc1:lyn-tdTomato)*$^{i261}$, *mylz2:GFP, fmyhcx:GFP, sox6:eGFPsox63'UTR* have been previously described (Elworthy et al., 2008, von Hofsten et al., 2008, Wang et al., 2011). *Tg(tnnc1b:eGFP)* and *Tg(tnnc1a:eGFP)* were generated for this thesis (described later).

The *UAS:sox6-GFP;actin:GAL4* line was generated by the crossing the previously-described Tg(actin1β:GAL4)$^{i269}$ line (Scheer and Campos-Ortega, 1999) to the *UAS:sox6-GFP* line generated by Wang Xingang in the Ingham laboratory.

## 2.2 Molecular Biology General procedures

### *2.2.1 PCR (General)*

10ng of plasmid DNA was added to a 20µl reaction containing 10µl 2x GoTaq (Promega) and 10pmol/µl of each primer. The cycling reaction was carried out as follows: 95$^{o}$C 2 minutes 30 seconds, then 30x cycles of (95$^{o}$C 45 seconds, 53-58$^{o}$C for 30 seconds, 72$^{o}$C for 1 minutes per kb of product) and finally 72$^{o}$C for 10 minutes. Annealing temperature was 5$^{o}$C lower than the lowest primer melting temperature.

### *2.2.2 PCR (Cloning)*

100ng of plasmid DNA or1µl of a cDNA synthesis reaction was added to a 50µl reaction volume containing 1µl iproof DNA polymerase (BioRad), high fidelity buffer, 0.2mM dNTP, 10pmol/µl of each primer. The cycling reaction was carried out as follows: 98$^{o}$C 2 minutes

30 seconds then 30x cycles of (98°C for 45 seconds, 58-63°C for 1 minute, 72°C for 30 seconds plus 15 seconds/kb), and finally 72°C for 10 minutes. Annealing temperature was 3°C higher than the lowest primer melting temperature.

### *2.2.3 Restriction endonuclease digestion*

All restriction endonuclease digestions were performed in 20-100µl reaction volumes containing 1 unit of restriction enzyme per µg DNA and reaction buffer diluted in milli-Q water. Reactions were performed at appropriate temperature for the particular enzyme for 2-4 hours.

### *2.2.4 Phenol: chloroform extraction and ethanol precipitation of DNA*

DNA was diluted to a total volume of 400µl using milli-Q water to which 400µl of 1:1 phenol: chloroform pH7 was added. The mixture was emulsified and then centrifuged at 13000rpm for 5 minutes at room temperature. 40µl 3M NaAc (pH 5.2) and 1ml ethanol was then added to the extraction aqueous phase of the solution and then centrifuged at 13000rpm for 30minutes at 4°C. The DNA pellet was washed with 200µl 70%n ethanol, centrifuged for a further 15 minutes at 13000rpm at 4°C, air-dried, and resuspended in the appropriate volume of nuclease-free water.

### *2.2.5 Gel extraction*

DNA fragments were extracted from 0.8%-1% agarose gel and the DNA was extracted using the Axygen "AxyPrep DNA gel extraction kit" according to manufacturer's protocol.

### *2.2.6 Standard cloning*

Primers were designed containing restriction sites at the end. After PCR amplification the product was gel extracted and digested, using the appropriate enzyme, and then purified. 50ng of vector was also digested using the appropriate enzyme, followed by the removal of the phosphate groups from the 5' end of cut plasmid using calf intestinal alkaline phosphatase to prevent recircularization of the linearized vector. The ligation reaction was set up in a total volume of 10µl containing the insert and the vector in an approximately 3:1 molar ratio. The reaction was incubated for 12 hours at 16°C then 2µl was transformed. Colonies where then prepped, digested and sequenced to select for positive inserts.

### *2.2.7 Transformation of chemically competent cells*

50µl of chemically competent Top10 (Invitrogen) cells were used for each transformation. Approximately 10ng of plasmid was incubated with the cells on ice for 5-20 minutes. Cells were then heat-shocked at 42°C for 30 seconds then place immediately on ice for 1 minute. 250µl of LB was added to the cells, which were then placed in a 37°C-shaking incubator for 1 hour. 50µl-200µl of culture was spread onto LB agar plates containing the appropriate antibody. Plates were incubated overnight at 37°C.

### *2.2.8 mRNA synthesis for microinjection*

Capped mRNA was prepared from cDNA clones inserted into plasmids amenable to *in* vitro *transcription*, using the appropriate restriction enzyme and RNA polymerase mRNA Message Machine Kit (Ambion), following the manufacturer's instructions.

### *2.2.9 RNA extraction*

10 embryos were dechorionated. 500µl of Trizol was added and the embryos were homogenized using a homogeniser. A further 500µl of Trizol was added, and after 5 minutes of incubation at room temperature, 200µl of chloroform was added followed by vigorous shaking. After three minutes, the tube was centrifuged at 12,000xg for 15 minutes at 4°C. The aqueous phase was transferred to a new tube and the DNA was precipitated by the addition of isopropanol, and then washed with 75% ethanol. The RNA pellet was resuspended in 50µl of RNase free water.

### *2.2.10 First-strand cDNA synthesis*

1µg of RNA was added to a 13µl reaction containing 0.05ng oligo(dT) and 10ng of each dNTP. The reaction was incubated at 65°C for 5 minutes then transferred to ice for at least one minute. First-strand buffer, 0.1µg DTT, 40units of RNase OUT and 200 units of reverse transcriptase (Superscript III) was added to the reaction. The reaction was then incubated for 1 hour at 50°C and inactivated for 15 minutes at 70°C.

### *2.2.11 Digoxigenin-labelled RNA probe synthesis*

Antisense RNA probes were synthesised by *in vitro* transcription, in a 20µl reaction containing 1µg of linearised DNA, 10x transcription buffer, 10x DIG RNA labelling mix, 2µl RNA polymerase, 1µl ribonuclease inhibitor (RNasin) and nuclease-free water. The reaction was incubated for 2 hours at 37°C followed by the addition of DNase I for a further 15 minutes at 37°C. The RNA products were precipitated by the addition of 130µl nuclease free water, 15µl 3M sodium acetate (pH5.2) and 375µl 100% ethanol and centrifuged at 13000rpm for 30 min at 4°C. The pellet was washed in 70% ethanol, air dried and resuspended in 50µl nuclease free water and 50µl formamide.

## 2.3 Microinjection

Zebrafish embryos were injected with 1nl of mRNA/BAC/DNA/Morpholino (concentrations below) into either the yolk or the cell within the first 20 minutes of them being laid.

### *2.3.1 Capped mRNA*

*In vitro*-synthesised capped mRNA encoding the Tol2 transposase was injected at a concentration of 30ng/µl.

### *2.3.2 BACs*

BACs were injected at a concentration between 50-150ng/µl.

### *2.3.3 Plasmid DNA*

Plasmid DNA was injected at a concentration between 25-50ng/µl

### *2.3.4 Morpholinos*

Morpholinos (MO) were designed and manufactured by Gene Tools, LLC (Philomath, OR, USA)

| **Name** | **Sequence** | **Injection Concentration** |
|---|---|---|
| *prdm1a*-MO | TGGTGTCATACCTCTTTGGAGTCTG | 0.5mM |
| *six1a*-MO | TCTCCTCTGGATGCTACGAAGGAAG | 0.8mg/ml |
| *six4a*-MO | CCGCCTCGTTTCTGCCCAAGCTGAT | 0.8mg/ml |
| *pbx2*-MO2 | CCGTTGCCTGTGATGGGCTGCTGCG | 0.25mg/ml |
| *pbx2*-MO3 | GCTGCAACATCCTGAGCACTACATT | 0.5mg/ml |
| *pbx4*-MO1 | AATACTTTTGAGCCGAATCTCTCCG | 0.5mg/ml |
| *pbx4*-MO2 | CGCCGCAAACCAATGAAAGCGTGTT | 0.5mg/ml |

## 2.4 *In situ* hybridisation

### *2.4.1 Standard in situ hybridisation protocol*

Embryos were fixed with 4% PFA overnight at 4°C and stored in 100% MeOH at -20°C. Fixed embryos were rehydrated in 50% MeOH in PBS, follow by 4 washes in PBST. Embryos were permeabilized in 0.01mg/ml Proteinase K (time dependent on developmental stage) then fixed for 20 minutes in 4% PFA, followed by 5 washes in PBST. Embryos were pre-hybridised for 3 hours in hybridisation buffer at 70°C then hybridised overnight at 70°C in appropriate concentration of DIG-labelled RNA probe. The following day, the embryos were washed at 65°C in 50:50 hybridisation solution:2xSSC for 10 minutes, 2xSSC for 10 minutes, 0.2xSSC 2x 30 minutes and finally 50:50 PBT:0.2xSSC for 10 minutes at room temperature. Embryos were then blocked in PBT for 3 hours at room temperature before being incubated with anti-DIG alkaline phosphatase (Roche, 1/3000) in PBT at 4°C overnight. The embryos were then washed 6x for 20 minutes in PBST and then equilibrated in Staining Buffer (lacking NBT and BCIP) by washing 4x 10 minutes, then stained in staining solution in the dark until a desirable level of staining had been achieved. Embryos were then washed in PBST and fixed in 4% PFA overnight at 4°C followed by storage in 75% glycerol. For embryos that needed to be separated prior to the *in situ* hybridisation procedure and then compared afterward (i.e. Sox6$^{-/-}$ mutants and siblings separated on the basis of their *tnnc1:GFP* expression or *UAS:sox6-GFP;actin:GAL4* embryos based on their GFP expression) the two sets of embryos (i.e. mutant vs. siblings) were treated exactly the same and stained for exactly the same length of time in NBT and BCIP. All images displayed in this thesis are a representative phenotype of all the embryos examined.

### 2.4.2 in situ hybridisation on sections

Sections were defrosted at room temperature for one hour and circled using a PAP pen. The pre-heated probe at the appropriate concentration was added to the sections. Slides were covered with a coverslip and hybridised at 70°C overnight. Slides were washed in pre-warmed wash solution for 4x 30 minutes allowing the coverslip to fall off in the first wash. Slides were transferred to MABT and washed 3x 30 minutes in this solution at room temperature. Sections were blocked in blocking solution for 3 hours at room temperature before being incubated with 1:5000 anti-DIG overnight at room temperature. Slides were washed 5x 20 minutes in MABT at room temperature then 2x 10 minutes in staining buffer lacking NBT/BCIP. Sections were stained in staining buffer overnight at 37°C. Staining was stopped by washing with PBS two times. Glycerol and a cover slip were added to each slide ready for imaging.

### 2.4.3 Fluorescent in situ hybridisation

Fluorescent *in situ* hybridisation was done using the 'TSA™ Plus Cyanine 5 and Fluorescein' and 'TSA™ DNP (HRP) System' from PerkinElmer. Day one is same at normal *in situ* hybridisation. On day two, after final 0.2xSSC wash, embryos were washed in 50:50 0.2xSSC:PBST, followed by a wash in PBST alone for 10 minutes. The endogenous peroxidase was quenched in 2% hydrogen peroxidase ($H_2O_2$) for one hour at room temperature. Embryos were washed 4x 5 minutes in TNT buffer then blocked in TNT block for 3 hours. Anti-DIG-POD (1:500) and anti-GFP (1:500) were added and incubated at 4°C overnight. Embryos were washed 3x 25 minute TNT buffer the next day and DNP-tyramide 1:50 in amplification buffer was added for one hour. A further 3x 25 minute washes in TNT buffer were then performed, followed by blocking in TNT block for 3 hours. The secondary antibody against the DNP substrate, anti-DNP POD (1:500) and secondary antibody against GFP were added and incubated at 4$^{o}$C overnight. The following day embryos were washed 3x 25 minutes in TNT buffer then incubated in cy5-tyramide (1:50) for one hour at room temperature. Embryos were subsequently washed 3x 25 minutes in TNT buffer then put through a glycerol series up to 75% glycerol and imaged.

### *2.4.4 Table of In situ hybridisation probes*

| **Gene name** | **Enzyme and promoter used** | **Reference** |
|---|---|---|
| *sox6* | BamHI/T7 | (von Hofsten et al., 2008) |
| *sox5* | NcoI/SP6 | (Wang et al., 2011) |
| *tnnc1b* | EcoRI/T7 | (von Hofsten et al., 2008) |
| *smyhc1* | Not1/SP6 | (Elworthy et al., 2008) |
| *prox1a* | EcoRI/T7 | (Glasgow and Tomarev, 1998) |
| *tpm2* | SpeI/T7 | This thesis |
| *mylz10* | SpeI/T7 | This thesis |
| *ryr1a* | NcoI/SP6 | (Hirata et al., 2007) |
| *fmyhcx* | HindIII/T7 | (von Hofsten et al., 2008) |
| *mylz2* | BamHI/T7 | (von Hofsten et al., 2008) |
| *myh7b* | NcoI/SP6 | (Wang et al., 2011) |
| *tnni1a* | NcoI/SP6 | This thesis |
| *tnnt1* | NcoI/SP6 | This thesis |
| *tnnc1a* | NcoI/SP6 | This thesis |
| *tnni1b* | SphI/SP6 | This thesis |
| *tnnt2* | NcoI/SP6 | This thesis |

## 2.5 Immunohistochemistry

### *2.5.1 Whole mount immunohistochemistry*

Embryos were fixed overnight in 4% PFA at 4$^{o}$C*, washed twice for 5 min in PBS, the twice for 5 minutes in 50% MeOH and PBS and finally stored in 100% MeOH at -20$^{o}$C for at least 2 hours. For staining, embryos were rehydrated in 50% MeOH/PBTX, then washes 2x

5min in PBTX. For embryos over 2 dpf, embryos were permeabilized in acetone and trypsin (see below), then washed 2x 5min in PBTX. Embryos were then blocked in PBDT 3x 30min. The primary antibody was added to PBDT at the appropriate concentration and incubated overnight at 4°C. The following day embryos were washed 3x 30min in PBDT and the secondary antibody was added at the appropriate concentration for 3 hours at room temperature. Embryos where then washed 3x 30min in PBDT, followed by 2x 5min in PBTX, and put through a glycerol series up until 75% glycerol and imaged.

*When staining using the S58 antibody embryos were fixed in Carnoy's solution (60% ethyl alcohol, 30% chloroform, 10% glacial acetic acid) overnight at room temperature, and then washed as above.

### *2.5.2 Acetone and trypsin permeabilization*

After PBTX washes, embryos were washed for 5 minutes in water. -20°C acetone was then added to the embryos, which were subsequently incubated in at -20°C for 7 minutes. Embryos were then washed 2x 5 minutes then 2x 5 minutes with PBTX. Fish were then either blocked in PBDT as above or a trypsin digest was performed (depended on the antibody that was going to be used and age of embryo).

For the trypsin digest, embryos were incubated in 0.1% trypsin for 30 minutes at room temperature. Embryos were then washed 2x 5 minutes in PBTX then blocked in PBDT as above.

### *2.5.3 Immunohistochemistry on sections*

Slides were thawed for 1 hour on bench then fixed in acetone for 5 minutes at -20°C and allowed to dry. Sections were washed in PBTX 2x 5 minutes then blocked in PBDT for one hour at room temperature. The primary antibody was pipetted on to slide at the appropriate concentration the incubated at 4°C overnight. The following day slides were washed 2x 30 minutes in PBDT then incubate in the appropriate concentration of the secondary antibody for 1 hour at room temperature. Sections were washed in PBDT for 2x 30 minutes, then washed in PBTX for 15 minutes before the addition of Vectorshield and a coverslip.

### *2.5.4 Table of primary antibodies used*

| **Epitope/antibody name** | **Animal developed in** | **Dilution used for staining** | **Provider** |
|---|---|---|---|
| F59 (slow myosin heavy chain) | Mouse | 1/50 | DSHB |

| | | | |
|---|---|---|---|
| F310 (fast myosin light chain) | Mouse | 1/50 | DSHB |
| S58 (slow myosin heavy chain) | Mouse | 1/50 | DSHB |
| GFP | Rabbit | 1/500 | Torrey Pines |
| mCherry | Mouse | 1/500 | Living Colours, Clontech |
| Prox1a | Rabbit | 1/50 (of preabsorbed stock) | Stone Elworthy, Custom made |
| Tnnc1b | Rabbit | 1/250 | Genetex |

### *2.5.5 Table of secondary antibodies used*

| **Epitope/antibody name** | **Dilution used for staining** | **Provider** |
|---|---|---|
| Goat anti-mouse IgG-488 | 1 in 1000 | Molecular Probes, Invitrogen |
| Goat anti-mouse IgG-546 | 1 in 1000 | Molecular Probes, Invitrogen |
| Goat anti-rabbit IgG-488 | 1 in 1000 | Molecular Probes, Invitrogen |
| Goat anti-rabbit IgG-568 | 1 in 1000 | Molecular Probes, Invitrogen |
| Goat anti-mouse IgG-HRP | 1 in 1000 | Molecular Probes, Invitrogen |
| Goat anti-rabbit IgG-HRP | 1 in 1000 | Molecular Probes, Invitrogen |

## 2.6 Cryostat sectioning

Fish were prepared for sectioning by fixing in 4% PFA overnight at 4$^{o}$C. The fish were then mounted in 1.5% low-melt agarose in a 5% sucrose solution. The agarose blocks containing the fish were transferred to 30% sucrose solution overnight at 4$^{o}$C. The agarose blocks were then frozen in (Optimal Cutting Temperature) OTC compound and 20µm thick sections were cut using a cryostat. The slides were baked for at least 2 hours at 37$^{o}$C and stored at -20$^{o}$C.

## 2.7 Cyclopamine treatment

Treatment was done before 50% epiboly using 40µM cyclopamine. Embryos were then fixed as normal in 4% PFA at desired timepoint.

## 2.8 Alcian Blue and Alizarin Red staining

### *2.8.1 Adult zebrafish Alcian Blue and Alizarin Red staining*

Fish were fixed in 4% PFA overnight at room temperature. After the skin and muscle was removed from the fish, a small incision was made in the abdominal cavity and the gut was removed. Fish were then digested in 150mg/ml of trypsin made up in 5% Borax, and incubated overnight at 37°C. Fish were washed twice with water to remove trypsin, then acetone was added for approximately 8 hours until the fish were no longer shiny. If doing a double stain fish were then washed with 50% EtOH then the alcian blue staining solution was added. Fish were stained overnight at room temperature in a glass jar, washed with Borax the following morning, then stained with Alizarin Red for one hour (20µl/ml of 0.5% alizarin red, 1% potassium hydroxide, 50% ethanol). Pigment was then removed through bleaching of the embryo in 1.5% hydrogen peroxide and 1% potassium hydroxide for approximately 20 minutes. Fish were then washed 4x 10 minutes in water and cleared through a glycerol series (25%, 50% and 75%) all with 0.25% KOH.

### *2.8.2 Alizarin Red staining alone on adults*

After washes in acetone, fish were washed in 5% Borax then stained with alizarin red for 1-4 hours at room temperature. Fish were then bleached as above, washed in water and cleared through a glycerol series.

### *2.8.3 Alcian Blue and Alizarin Red staining of embryos and juvenile zebrafish (3 dpf-30 dpf)*

Larvae were fixed in 4% PFA at 4°C overnight. Fish were washed in water and dehydrated in 50% EtOH for 10 minutes at room temperature. 1ml of staining solution was added (either double staining solution or Alizarin Red staining solution) and left over night at room temperature for less than 16 hours. The larvae were washed with water and then bleached in bleaching solution for 20 minutes at room temperature. After three washes with water embryos were incubated in clearing solution 1 for 30 minutes at room temperature, clearing solution 2 for 2 hours at room temperature and finally stored in storage solution at 4°C.

## 2.9 BAC recombineering

### *2.9.1 Identification and preparation of the BAC*

Using the Ensemble Genome Browser (zv9), three BACs were identified which encompassed the *tnnc1b* locus. These BACs were streaked out onto Chloramphenicol plates (12.5µg/ml) and prepared using the Nucleobond PC100 Midi kit. 300ng purified BAC DNA was electroporated at 1.8kV into EL250 *E. coli* cells containing the *red* recombineering system.

### *2.9.2 Ultramer design for eGFP cassette and targeting of cassette*

The heat-shock inducible *red* recombineering system was used to insert an eGFP cassette with an SV40 polyadenylation site into the *tnnc1b* ATG start site (Lee et al., 2001). A DNA cassette containing eGFP-SV40pA-FRT-Kan-FRT (gift from Dr Stone Elworthy) was modified by the addition of 80 nucleotide homology arms for the *tnnc1b* locus, using two long oligonucleotides named ultramers (Tnnc1b ultramer L and Tnnc1b ultramer R in primer list) that had a 20 nucleotide overlap with the reporter cassette (Figure 2.1A and B, and Appendix A1). Following PCR amplification of this cassette, Dpn1 was added to the reaction to digest the plasmid template, and the DNA was then purified by gel electrophoresis and extraction. 300ng of the purified product was then electroporated at 1.8kV into the EL250 cells containing the BAC. Successful integration for the construct was screened for using Kanamycin resistance, the gene for which was then removed by Flp recombinase-mediated removal using the flanking FRT sites (Lee et al., 2001). Primer pairs Tnnc1b left and Tnnc1b right, and Tnnc1b left and GFP right were used to confirm correct insertion into the BAC.

### *2.9.3 iTol2 modification of the BAC*

A further modification to the BAC was made by the replacement of the Chloramphenicol resistance gene with an iTol2-Amp-iTol2 targeting cassette (Figure 2.1C) (Suster et al., 2011). This targeting cassette was then modified by the addition of homology arms (itol2 HA1 and itol2 HA2 in primer list) and electroporated into the cells containing the eGFP modified BAC. Successful integration of the cassette was screened for using Ampicillin resistance. Primer pairs itol2 outside left and inside itol2 were used to confirm correct insertion into the BAC. The modified BAC was then injected into zebrafish embryos along with capped Tol2 transposase mRNA to allow transposition of the BAC into the zebrafish genome. Embryos were grown to sexual maturity then screened for successful integration of the reporter by incrossing.

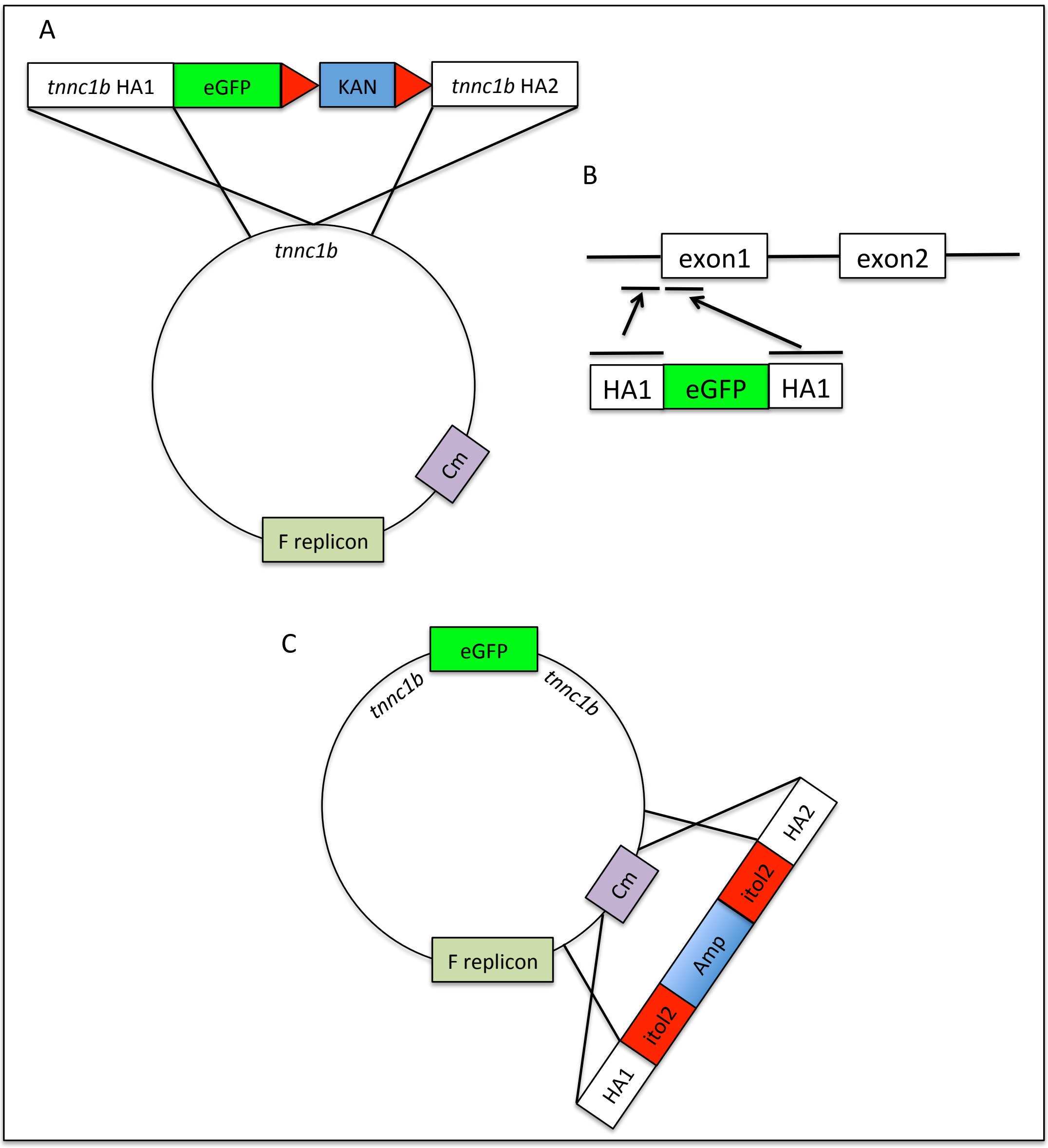


***Figure 2.1: Schematic illustration of the BAC recombineering strategy.*** *An eGFP-FRT-KAN-FRT reporter cassette was inserted into a BAC containing the tnnc1b locus (A). The start site of the GFP cassette replaced the start site of the tnnc1b gene (B). iTol2 sites were added to the eGFP modified BAC using the same strategy, with ampicillin (Amp) replacing the chloramphenicol (Cm) gene (C). Red arrows in A represent FRT sites.*

### *2.9.4 Generating deletion derivatives of the BAC*

DNA sequences lying upstream and downstream from the *tnnc1b* coding region were deleted from the BAC using the *red* recombineering system. Downstream deletions were made by targeting the iTol2-Amp-iTol2 cassette into the eGFP modified BAC but designing the homology arms so that the right arm had homology to the sequence 1kb downstream from the *tnnc1b* stop codon (itol2 ultramer 3 in primer list) and the left arm had homology to the other side of the Chloramphenicol marker (Figure 2.2A, itol2 HA2 in primer list). This resulted in the Chloramphenicol gene being replaced by the Ampicillin gene as before, but also a large amount of downstream sequence was recombined out of the BAC. Primer pairs were then used to confirm the correct insertion of the iTol2 and the deletion of the downstream sequence.

This BAC was further modified by the deletion of upstream sequence. A construct was designed that contained a Kanamycin resistance gene with two flanking homology arms. This construct was targeted to the upstream sequence of *tnnc1b*, leaving intact the F replicon (Figure 2.2B, Reverse ultramer 4 in primer list). The homology arms were designed so they would remove differing amounts of upstream sequence (Forward ultramer 4, +7kb tnnc1b HA forward, +4kb tnnc1b HA forward and +2.5kb tnnc1b HA forward primers were used leaving 10kb, 7kb, 4kb and 2.5kb of upstream sequence respectively). Successful integration was selected for based on kanamycin resistance and PCR was used to confirm the correct insertion of the construct and the deletion of upstream sequence.

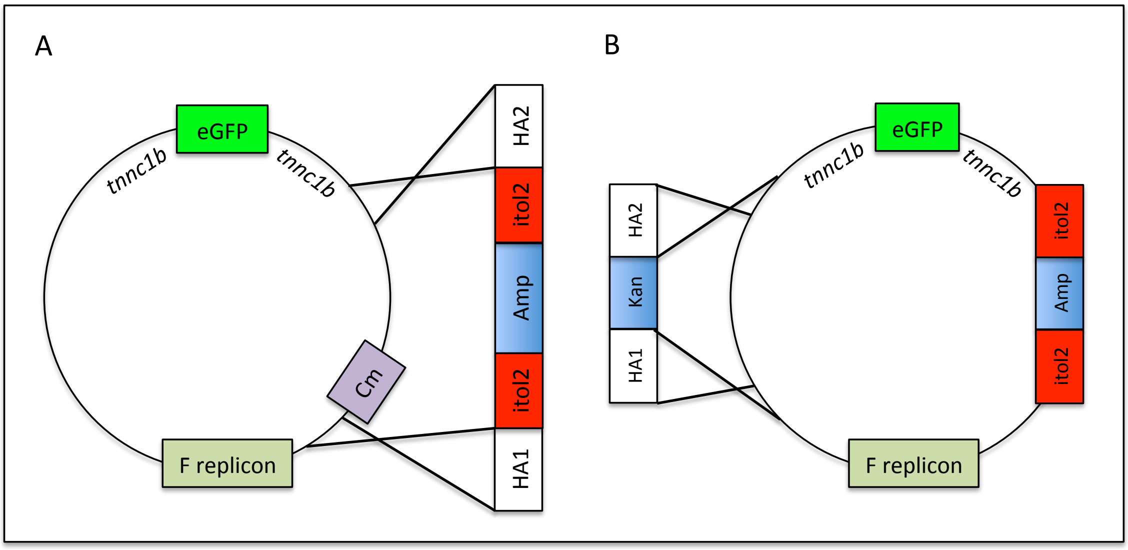


***Figure 2.2: Schematic to illustrate strategy for generating deletion derivative of BAC.*** *The iTol2 cassette was modified so that one of the homology arms would recombine 1kb downstream of the tnnc1b stop site (A). The upstream sequence was deleted by targeting a Kanamycin resistance cassette flanked by homology arms that would target different regions in the upstream sequence (B). HA (homology arm).*

### *2.9.5 Making smaller constructs of tnnc1b:eGFP*

Smaller constructs containing the GFP reporter sequence were PCR amplified out of the BAC and cloned into the pDB739 vector (from Steve Ekker lab; (Balciunas et al., 2006)) that contains iTol2 sites (see 2.2.6. standard cloning).

### *2.9.6 Enhancer analysis using a β-Globin minimal promoter vector*

A β-globin minimal promoter reporter vector was previously generated in the lab by excising β-globin-eGFP-polyA from βg-Egfp-SP72 (MacDonald et al., 2010) using EcoRI and cloned into the same site of pDB739 (Balciunas et a., 2006). Potential enhancers were then cloning into this vector as detailed in "2.2.6. Standard cloning"

### *2.9.7 Mutating potential sox6 binding sites in the tnnc1b:eGFP promoter*

The plasmid containing 2.5kb upstream of the start site of *tnnc1b* as well as the first intron was analysed manually to identify possible Sox6 binding sites. Once these sites were identified they were mutated sequentially and the mutant constructs were injected into zebrafish embryos so establish stable transgenic lines. Sites were mutated by designing forward and reverse primers that had approximately 10bp homology to the sequence 5' of the potential Sox6 binding site and 30bp homology to the sequence 3' of the Sox6 site, with a 4 base pair mismatch in the seed sequence of the sox6 binding site (see Figure 3.7 in Chapter 3). After PCR using these primer pairs, 1µl of Dpn1 restriction enzyme was added to the reaction to digest the methylated plasmid template. After a two-hour incubation at 37°C, the sample was PCR purified (AxyPrep PCR clean-up kit Axygen), transformed and plated onto Ampicillin agar plates. Six colonies were prepped and sequenced. Sequences were analysed using Lasergene SeqMan (DNA STAR) and the mini-preps that were positive for the mutation were then used in the next reaction to mutate the next binding site. Mutated plasmids were injected into zebrafish embryos and stable transgenic lines were created for each construct.

## 2.10 Generation of zinc-finger nuclease Sox6 mutant

A zinc-finger nuclease designed to target the sox6 locus was purchased from Sigma and injected into zebrafish embryos (see Figure 2.3 for design strategy). The *sox6* gene has two transcripts, so the DNA-binding domain of the zinc-finger nuclease was targeted so that a mutation would be generated up stream of the HMG box in both of the *sox6* transcripts (Figure 2.4A and B). The zinc-finger nuclease was injected into AB embryos by Dr Wang Xingang. Once the injected fish had reached maturity they were screened by incrossing. 48 embryos from each pair were placed in separate wells of a 96-well plate and digested in DNA extraction buffer (10mM Tris pH 8, 2mM EDTA, 0.2% Triton X-100, 200µg/ml Proteinase K) overnight at 55°C. Sox6 ZF F and Sox6 ZF R primers were used to amplify the region of interest and Sox6 ZF seq primer was used for sequencing. 2 heterozygous mutant fish were identified in the 40 fish screened. The fish carrying the first allele identified, allele A1, produced 15 out of 48 heterozygous mutant embryos. The fish carrying the second allele identified, allele E2, gave rise to 11 out 48 heterozygous mutant embryos. The remaining embryos from the lay of the positively identified heterozygous mutants were grown up to adulthood. Progeny from allele A1 were fin clipped and genotyped and 8 out of 39 fish were identified which carried the A1 mutant allele. Progeny from allele E2 were genotyped and 6 out of 38 fish were identified which carried the E2 mutant allele. Both mutant alleles introduced a premature stop codon into the sequence (Figure 2.4C). Heterozygous mutants were crossed to generate homozygous sox6 mutants.

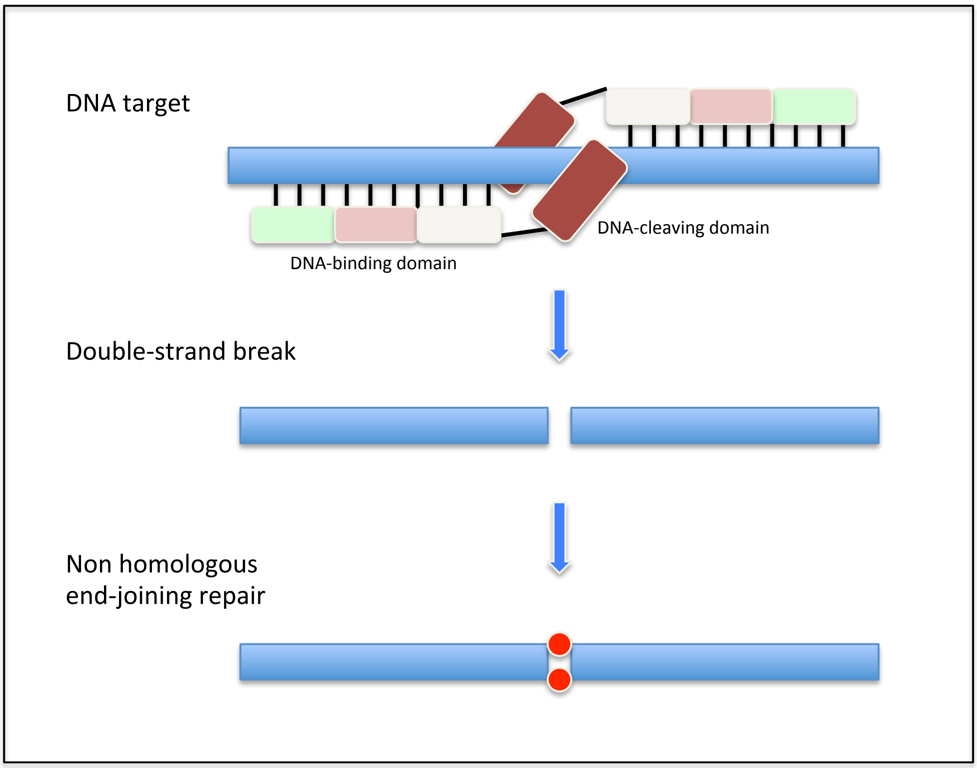


***Figure 2.3: Zinc-finger nucleases.*** *A pair of zinc-finger DNA-binding domains, fused to a DNA-cleavage domain were targeted to a specific region of DNA, creating a double-stranded break in the DNA. This stimulates the cell's endogenous DNA-repair machinery and the DNA is then repaired through non-homologous end joining (NHEJ), which sometimes results in a deletion or insertion of base pairs, producing a fame-shift mutation.*

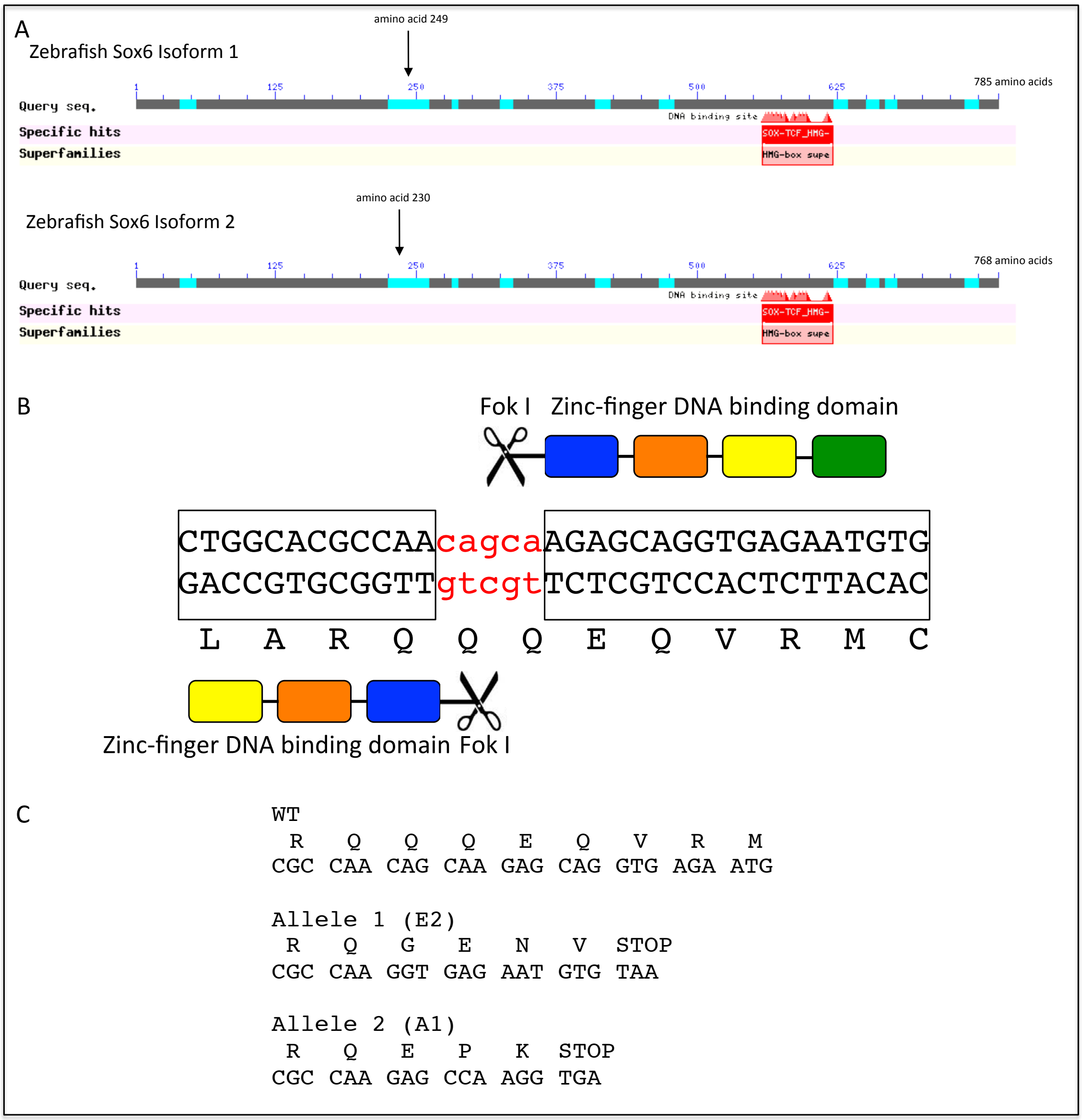


***Figure 2.4: Sox6 zinc-finger nuclease.*** *The zinc-finger nuclease was designed so that it targeted sequences that are present in both isoforms of Sox6, upstream of the HMG box at the position of approximately amino acid 249 and amino acid 230 (A). Schematic representation of the nucleotide sequence in exon 8 of the zebrafish sox6 gene targeted by the zinc-finger nuclease (B). Two mutant alleles were identified, E2 and A1, each of which introduced a premature stop codon in the sequence (C).*

## 2.11 Western Blotting

### *2.11.1 Western blotting*

Embryos were dechorionated and washed in PBS twice. 1ml of Deyolk Buffer was added and embryos were pipetted vigorously until yolk was completely removed from the embryos. Cells were pelleted at 600xg for 3 minutes and washed twice in PBS. Cells were resuspended in Cell Lysis Buffer, pipetted vigorously, and incubated on ice for 15 minutes with regular pipetting every few minutes. The tube was centrifuged at 3500rpm for 5 minutes at 4°C to pellet the nuclei. 1µl of RIPA Buffer was added per 5 embryos and nuclei were pipetted vigorously. After a 10-minute incubation on ice, SDS Loading Buffer was added to the sample and the tube was heated at 95°C for 10 minutes. The sample was run on an SDS-polyacrylamide gel then transferred to a nitrocellulose membrane overnight. The following day the membrane was blocked in Blocking Solution for 1 hour, incubated with the primary antibody for 1 hour and washed 3x 10 minutes in PBST. The membrane was then incubated with a horseradish peroxidase-conjugated secondary antibody for 30 minutes then washed 4x 10 minutes with PBST. The immunoreactive proteins were detected on the nitrocellulose membranes using the Clarity Western ECL substrate detection kit (BioRad) and exposure to X-ray medical film.

### *2.11.2 Table of primary antibodies for western*

| **Name** | **Animal developed in** | **Concentration immunohistochemistry** | **Concentration western** | **Provider** |
|---|---|---|---|---|
| β-actin | Rabbit | N/A | 1/2000 | Cell Signalling |
| GFP | Mouse | N/A | 1/2000 | Living Colours, Clontech |
| Sox6 ab66316 | Rabbit | 1/500<br>1/2000 | 1/5000 | Abcam |
| Sox6 GTX116236 | Rabbit | 1/250<br>1/500 | 1/2500 | Gentex |
| Sox6 A | Rabbit | 1/1000 | 1/2500 | Custom made- Sdix |
| Sox6 2 | Rabbit | 1/1000 | 1/2500 | Custom made- Sdix |
| Sox6 M1 | Rabbit | 1/3000 | 1/10000 | Custom made- Absea |

### *2.11.3 Table of secondary antibodies for western*

| **Name** | **Concentration** | **Provider** |
|---|---|---|
| Anti-Rabbit HRP | 1/5000 | Molecular probes |
| Anti-Mouse HRP | 1/5000 | Molecular probes |

## 2.12 Chromatin Immunoprecipitation (ChIP)

### *2.12.1 Chromatin extraction and sonication*

1000 *UAS:sox6-GFP;actin:GAL4* positive embryos (experiment) or 1000 AB embryos (control) were dechorionated in 1mg/ml Pronase, then washed with cold PBS. Deyolk Buffer was added and embryos were pipetted vigorously to dissolve the yolk. The cells were pelleted by centrifugation at 600g for 3 minutes at 4°C and washed twice in cold PBS. Cells were then fixed in 2mM of the protein crosslinker DSG (Disuccinimidyl glutarte) (as described in An et al., 2011) for 45 minutes, followed by 10 minutes in 1% formaldehyde at room temperature. 1/20th volume of 2.5M glycine was added for 5 minutes to quench the reaction. Cells were then spun down at 1200g for 1 minute at 4°C then washed 3 times in PBS. 1ml of Cell Lysis Buffer was then added to the cells and the mixture was pipetted vigorously to homogenise cells. The tube was incubated on ice for 15 minutes and the mixture was regularly pipetted to ensure cells were completely lysed. The homogenate was then spun down for at 1200g for 5 minutes at 4°C and the pellet was resuspended in 400µl of Nuclear Lysis Buffer. Cells were pipetted up and down to disrupt the nuclei and put on ice for 10 minutes with regular pipetting. The nuclear lysate was then transferred to a 15ml tube and 800µl of cold IP Buffer was added. The lysate was sonicated on ice (10 seconds on, 20 seconds off, 3.00 minutes) and the final volume was made up to 4ml and split into two 2ml tubes. Tubes were either snap frozen or used straight away for ChIP. Sonication efficiency was tested by adding 1µl of 20mg/ml Proteinase K to a 300µl sample of the Chromatin extract, and incubating the sample for 2 hours at 42°C to reverse the protein-DNA crosslinks. The sample was then Phenol-Chloroform extracted and resuspended in 20µl of distilled water and run out onto a 1% agarose gel. If the fragmentation was satisfactory (i.e. most of the DNA fragment are in the size range of 300-700bp), the Chromatin Immunoprecipitation protocol was performed.

### *2.12.2 ChIP using GFP-Trap (ChromoTek)*

Protocol was adapted from Dubin et al., 2010. The chromatin was pre-cleared with 20µl of Sepharose-4B beads (ChromoTek) for 2 hours, on a rotating platform in the cold room. The beads were then spun down at 2500xg for 3 minutes and the supernatant was removed and put in a new tube. 100µl was removed and set aside as the input fraction. 20µl of GFP-Trap beads (ChromoTek) were washed 3 times in wash buffer, then added to the chromatin and incubated on a roller overnight at 4°C. The beads were spun down for 2 minutes at 2500xg at 4°C (if doing IP with western, 100µl was removed as supernatant sample). Beads were then washed twice with Low Salt Buffer, once with High Salt Buffer, twice with LiCl Buffer and finally twice with TE Buffer. 250µl of Elution Buffer was then added to the beads and the samples were incubated at 65°C for 15 minutes with gentle agitation. The supernatant was then transferred to a new tube and 250µl of fresh Elution Buffer was added and this previous stage repeated. To reverse the crosslinking the two elutes were pooled and 15µl of 5MNaCl was added and incubated overnight at 65°C. 400µl of Elution Buffer was also added to the IP with 15µl 5MNaCl and samples were incubated overnight at 65°C. The following day, DNA was precipitated from both the input and IP samples by the addition of EDTA and Tris-HCl pH 6.8 to a final concentration of 40mM and 10mM respectively. Proteinase K was added, and samples were incubated for 2 hours at 55°C. The DNA was then phenol: chloroform extracted and resuspended in 100µl of water.

### *2.12.3 ChIP Western blotting using GFP-Trap*

The chromatin was pre-cleared as above, with 100µl set aside as the input fraction. The GFP-Trap was washed and incubated with the chromatin as above, followed by the removal of a supernatant fraction. The beads were washed as above and then resuspended in 100µl of 2x SDS loading buffer. The sample was then boiled for 10 minutes at 95°C and run on a 10% SDS polyacrylamide gel along with the input and the supernatant fractions. The western blot protocol above was then followed.

## 2.13 Quantitative PCR (qPCR)

Quantitative PCR was performed using an iCycler iQ Real-Time PCR detection system (BioRad Laboratories Ltd) and iQ SYBR Green Supermix (BioRad Laboratories Ltd). All primers were optimized to determine that amplification of the product was linear. Single products were confirmed by dissociation curve analysis. All qPCR reactions were carried out at 95°C for 1 min, followed by 50 cycles (95°C for 10 seconds, 56°C for 20 seconds, 72°C 20 seconds). The quantitative data was analysed with the iCycler iQ software package (BioRad Laboratories Ltd).

The ChIP experiment using the GFP-Trap was repeated three times using both *UAS:Sox6-GFP;actin:GAL4* embryos and wild-type AB embryos. Quantification of PCR products was performed by comparing the threshold cycles (Ct) of all samples run in triplicates. All Cts were normalised against the β-actin housekeeping gene Ct values, using the 2(delta delta Ct) method (Livak and Schmittgen, 2001) to give an expression value relative to the internal control of β-actin expression. The fold changes relative to the β-actin gene were averaged for each ChIP experiment and then the average fold change for the *UAS:sox6-GFP;actin:GAL4* samples were normalised to the average fold change for the wild-type AB control samples to give an expression value relative to the GFP-negative AB control.

## 2.14 Imaging

Light microscopy images were captured using the Carl Zeiss AXIO Zeiss Imager M2 and the AXIO Vision 4.7.2 software. Confocal images were captured using the Olympus FV-1000 confocal microscope and the Olympus FluoView FV-1000 ASW 1.6 confocal software.

## 2.15 Solutions and Buffers

### 2.15.1 General

| Solution or buffer | Composition |
|---|---|
| PBST | PBS with 0.1% Tween 20 |
| DNA Extraction Buffer | 10mM Tris pH 8, 2mM EDTA, 0.2% Triton X-100, 200μg/ml Proteinase K |

### 2.15.2 Solutions and Buffers for in situ hybridisation

| | |
|---|---|
| Hybridisation Solution | 50% Formamide, 5xSSC, 0.1% Tween 20, 50μg/ml Heparin, 500μg/ml tRNA, pH 6.0 with Citric Acid |
| Staining Buffer | 100mM Tris-HCl pH9.5, 50mM $MgCl_2$, 100mM NaCl, 0.1% Tween20, 4.5μg/ml NBT, 3.5μg/ml BCIP |
| PBT | PBS with 0.1% Tween 20, 2% Sheep Serum, 0.2% BSA |

### 2.15.3 Solutions and buffers for in situ hybridisation on sections

| | |
|---|---|
| MABT | 150mM Maleic acid, 100mM NaCl, 0.1% Tween 20, pH7.5 |
| Blocking Solution | 2% Boehringer Blocking Reagent, 20% Sheep Serum in MABT |

### 2.15.4 Solutions and buffers for fluorescent in situ hybridisation

| | |
|---|---|
| TNT Buffer | 0.1M Tris-HCl pH 7.5, 0.15M NaCl, 0.5% Tween20 |
| TNT Block | TNT with 0.5% Perkin-Elmer blocking powder |

### 2.15.5 Solutions and buffers Immunohistochemistry

| | |
|---|---|
| PBTX | PBS with 0.1% TritonX-100 |
| PBDT | PBS with 1% BSA, 1% DMSO and 0.5% |

| | |
|---|---|
| | TritonX-100 |

### 2.15.6 Solutions and buffers for Alizarin red and Alcian blue staining on larvae (3 dpf-30 dpf)

| | |
|---|---|
| Alcian Blue Stock | 0.4% in 70% EtOH (dissolve in 50% then adjusted to 70%) |
| Alizarin Red stock | 0.5% in water |
| Alcian Blue working solution | 50µl Alcian Blue Stock in 70% EtOH with 60mM $MgCl_2$ |
| Double staining solution | 10µl Alizarin Red Stock per 1ml Alcian Blue working solution |
| Alizarin red staining solution | 10µl Alizarin Red Stock in 70% EtOH |
| Bleaching solution | 1.5% $H_2O_2$, 1% KOH |
| Clearing solution 1 | 20% Glycerol, 0.25% KOH |
| Clearing solution 2 | 50% Glycerol, 0.25% KOH |
| Storage solution | 50% Glycerol, 0.1% KOH |

### 2.15.7 Solutions and buffers for Alizarin red and Alcian blue staining on adult fish

| | |
|---|---|
| Alcian Blue staining solution | 40mg Alcian Blue in 30% Acetic Acid, 70% absolute EtOH (dissolve in 50% Ethanol and then adjust to 70%) |
| Alizarin Red stock solution | 0.5% Alizarin Red in water |
| Alizarin Red staining solution | 20µl per 1ml of 1% KOH and 50% EtOH |
| Bleaching solution | 1.5% $H_2O_2$ and 1% KOH |

### 2.15.8 Solutions and buffers for western blotting

Protease inhibitor cocktail tablets was purchased from Roche and dissolved in water to make a 50x stock solution.

| | |
|---|---|
| Deyolk Buffer | PBS + 2x Protease Inhibitor Cocktail |
| Cell Lysis Buffer (western) | 10mM HEPES pH 7.9, 1.5mM MgCl, 10mM KCl, 300mM Sucrose, 0.5% NP-40 + 1x Protease Inhibitor Cocktail |
| RIPA buffer | 50mM Tris HCl pH8, 150mM NaCl, 1% NP-40, 0.5% Sodium Deoxycholate, 0.1% SDS + 1x Protease Inhibitor Cocktail |
| Blocking Solution (western) | 5% Milk Powder, 0.1% Tween-20 in PBS |

### *2.15.9 Solutions and buffers for chromatin extraction*

| | |
|---|---|
| Deyolk Buffer | PBS and 2x Protease Inhibitor Cocktail |
| Cell Lysis Buffer | 10mM Tris-HCl pH 7.5, 10mM NaCl, 0.5% NP-40, 2x Protease Inhibitor Cocktail |
| Nuclear Lysis Buffer | 50mM Tris-HCl pH7.5, 10mM EDTA, 1%SDS, 2x Protease Inhibitor Cocktail |
| IP Buffer | 16.7mM Tris-HCl pH 7.5, 167mM NaCl, 1.2mM EDTA, 0.01% SDS, 2x Protease Inhibitor Cocktail |

(Nuclear Lysis buffer was used instead of RIPA buffer, as previous optimisation for the zebrafish ChIP protocol in the lab had been performed with Nuclear Lysis Buffer.)

### *2.15.10 Solutions and buffers for ChIP with GFP-Trap*

| | |
|---|---|
| Wash Buffer | 10mM Tris-HCl pH 7.5, 150mM NaCl, 0.5mM EDTA |
| Low-Salt Wash Buffer | 150mM NaCl, 0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl pH 8 |
| High-Salt Wash Buffer | 500mM NaCl, 0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl pH 8 |
| LiCl Wash Buffer | 250mM LiCl,  1% NP-40, 1% Na-Deoxycholate, 1mM EDTA, 10mM Tris-HCl pH 8 |
| TE Buffer | 10mM Tris-HCl pH 8, 1mM EDTA pH 8 |
| Elution Buffer | 1% SDS, 100mM $NaHCO_3$ |

## 2.16 Primers

Primers were designed using Primer3 v0.4.0 (Rozen and Skaletsky, 2000) with an optimum melting temperature of 60$^{o}$C.

### *2.16.1 Primers for tnnc1b:eGFP BAC transgenesis*

| | |
|---|---|
| Tnnc1b ultramer L | GTCTGCCGTGAAGAGAGGAGCGCTGGATTACCTGTGAGGAGTTTACTGACCATTTCAAAAAAGGAAAATCTAAACCCGTCGCCACCATGGTGAGCAA |
| Tnnc1b ultramer R | CTGTTAAACATCACATGGGAAATATAAAAATAATAAGAAAAATTTTGTCTTCAACTAACTTTAATTACTGTGCCCATAACTTACCGCTGCTTTATATACATCATCGATATCTGCAGAATTCGCCC |
| Tnnc1b ultramer L end | GTCTGCCGTGAAGAGAGGAG |
| Tnnc1b ultramer R end | CTGTTAAACATCACATGGGA |
| Tnnc1b Left primer 5’ | GTCTGCCGTGAAGAGAGGAG |
| Tnnc1b Right primer 3’ | AGATGTCGAACGCAGCAC |
| eGFP R | GTTTACGTCGCCGTCCAG |

| | |
|---|---|
| Tnnc1b insert 5' | CCCATAACTTACCGCTGCTT |
| Tnnc1b insert 3' | TTTTCTAAGCCATCCATTCCA |

### 2.16.2 Primers for iTol2 BAC transgenesis

| | |
|---|---|
| itol2 HA1 | AAGAGGTTCCAACTTTCACCATAATGAAATAAGATCACTACCGGGCGTATTTTTTGAGTTATCGAGATTCCCTGCTCGAGCCGGGCCC |
| itol2 HA2 | GCCATTCATCCGCTTATTATCACTTATTCAGGCGTAGCAACCAGGCGTTTAAGGGCACCAATAACTATTATGATCCTCTAGATCAGATCT |
| itol2 HA1 end | AAGAGGTTCCAACTTTCACCA |
| itol2 HA2 end | GCCATTCATCCGCTTATTATC |
| itol2 outside left | AAATCCTGGTGTCCCTGTTG |
| itol2 outside right | CATCGAATTTCTGCCATTCA |
| inside itol2R | TTTTGGGGATTTTTACTTTACTTGA |

### 2.16.3 Primers for tnnc1b:eGFP BAC deletions

| | |
|---|---|
| itol2 ultramer 3 (to delete downstream sequence) | CAACAGTGAGGGCATCTCAGTAACGTGACGTTTCATCAAGTCACCAAAACCAAACCATTCAAGTTCAACATTATGATCCTCTAGATCAGATCT |
| itol2 ultramer 3 end | CAACAGTGAGGGCATCTCAG |
| itol2 outside primer 3 | GGCAGATTAGATGGGCAGTT |
| forward test primer itol2 | AGCCCTCCCGTATCGTAGTT |
| Forward ultramer 4 (to delete upstream sequence leaving 10kb) | TCTTCTGTGTCACATGATCCAACAGAAGTCATTTATTATTATGTCAAAAACATTCATGAAGGTTCTCGGAAGGTAGCTTGCAGTGGGCTT |
| Reverse ultramer 4 | GGAGCCACTATCGACTACGCGATCATGGCGACCACACCCGTCCTGTGGATCTACCCACTAGTCAATTCACCTCATACTCTTCCTTTTTCA |
| Forward ultramer 4 end | TCTTCTGTGTCACATGATCCAA |
| Reverse ultramer 4 end | GGAGCCACTATCGACTACGC |
| Outside forward ultramer 4 | TCATGCACTGTATAGGCAAAATG |
| Outside reverse ultramer 4 | TCCGCTCATGAGACAATAACC |
| Test primer R before Kan | TTTGATGGTCCGTTTGTTGA |
| Test primer R in Kan | CTCGTCCTGCAGTTCATTCA |
| +7kb tnnc1b HA Forward | TCCTCCCCCACAGTCAATTATACATCTCAAATCTCTCATTTCACAATAATTAGCCAGGTAGCTTGCAGTGGGCTT |

| | |
|---|---|
| +7kb tnnc1b HA end | TCCTCCCCCACAGTCAATTA |
| F test +7kb tnnc1b | CACAGTTGTTGCCCATTTTG |
| +4kb tnnc1b HA forward | CAGCCGCATTTTGGACTAATTGCAATCTATTCA<br>GGGCATTTTGACTAATACCCACAGGTAGCTTGC<br>AGTGGGCTT |
| +4kb tnnc1b HA end | CAGCCGCATTTTGGACTAAT |
| F test +4kb tnnc1b | AGTCTGATCCCCACTTGCAC |
| +2.5kb tnnc1b HA Forward | TCGCCCAATTACCCTAACCTGCCTAGTTAACCT<br>AATTAACCTAGTTAAGCCTTTAAGGTAGCTTGC<br>AGTGGGCTT |
| +2.5kb tnnc1b HA end | TCGCCCAATTACCCTAACCT |
| F test +2.5kb tnnc1b | CAGAGCAAGGGAATTTTCACA |

### *2.16.4 Primers to clone tnnc1b:eGFP fragments into pDB739*

| | |
|---|---|
| XbaI F tnnc1b -40bp (just after GFP) | CTGCGCTCTAGACCCATAACTTACCGCTGCTT |
| XbaI R +2.5kb | CTGCGCTCTAGAGGTTAATTAGGTTAACTAGGCAGGTT |
| XbaI R +1kb | CTGCGCTCTAGAGACAAGCATGTGACATTTTATGC |
| HindIII F intron 1 | GGGCCCAAGCTTCCAGCGATATGACAGATTTCAA |
| HindIII R +2.5kb | GGGCCCAAGCTTGGTTAATTAGGTTAACTAGGCAGGTT |
| HindIII R +1kb | GGGCCCAAGCTTGACAAGCATGTGACATTTTATGC |
| HindIIItnnc1b+240(2) | GGGCCCAAGCTTCAGGTGACCGGTGGATAAGT |
| HindIII tnnc1b +570 | GGGCCCAAGCTTCACCTATCCAGGCACAGCTT |
| HindIII tnnc1b +760 | GGGCCCAAGCTTAAGCATGCATCCAACGTACA |

### *2.16.5 Primers for cloning zebrafish tnnc1b intron1 into βglobin-GFP minimal promoter vector*

| | |
|---|---|
| F intron1 tnnc1b XbaI 2 | CTGCGCTCTAGAGCGGTAAGTTATGGGCACAG |
| R intron1 tnnc1b XbaI 2 | CTGCGCTCTAGACCATTCTTCTGCTCCTCTGTC |

### *2.16.6 Primers for cloning human tnnc1 intron1 into βglobin-GFP minimal promoter vector*

| | |
|---|---|
| F intron1 tnnc XbaI | GCTCTAGAATCTACAAGGCTGCGGTGAG |
| R intron1 tnnc XbaI | GCTCTAGAGTGACCACTCAATGCCCTTT |

### *2.16.7 Primers for Sox6 site mutations*

| | |
|---|---|
| F Sox6 1 | ATAATGATGTGGGTTTCTATAGACTTTCAGGAAAAAAAAATATAG |

| | |
|---|---|
| R Sox6 1 | GTCTATAGAAACCCACATCATTATACAATAACTCGCCC |
| F Sox6 2 | TCATGTCTGTAAAGGGATAATAGTCAGTAAACTGCAGAAACG |
| R Sox6 2 | TGACTATTATCCCTTTACAGACATGAATTTCATTCTTCTG |
| F Sox6 3 | GGCAATCCAAAAAGGGTTATAAATGGGAGTGAAGCGACG |
| R Sox6 3 | CCATTTATAACCCTTTTTGGATTGCCTTATGGGATAGT |
| F Sox6 4 | TATCCATTCAAAAGGGTGCTGTAACTATAGATTATTAAAAATCACA |
| R Sox6 4 | AGTTACAGCACCCTTTTGAATGGATAAATGCTGTGAAGTG |
| F Sox6 5 | TGAAGCATTTAAAGGGGAGGGTGTCTGGTGTCTGAA |
| R Sox6 5 | AGACACCCTCCCCTTTAAATGCTTCAACCTCAGTGTCT |
| F Sox6 8 | CTGATATGGGGGGTTTGCTTAGGCGGCCAGAAAGAC |
| R Sox6 8 | CCGCCTAAGCAAACCCCCCATATCAGGTTCTAAACTTATCCA |

### *2.16.8 Primers for sequencing Sox6 site mutations*

| | |
|---|---|
| F tnnc1b +1.1kb seq. | GCATGTGACATTTTATGCCATT |
| R tnnc1b +0.3kb seq. | CATTTGCCAATCATTTGCTG |

### *2.16.9 Primers for tnnc1a:eGFP BAC transgenesis*

| | |
|---|---|
| Forward ultramer *tnnc1a:eGFP* | GGCGGTCTCTCTATAAAGCTCATTTGGTTTTGTGTCCTGGTCTCCACTAGTGTCGGCACTTCATCTGATGTCGCCACCATGGTGAGCAA |
| Reverse ultramer *tnnc1a:eGFP* | TGTTTAAAACTCAAAGCACGAAAAAAGACCCAAGACTCACCGCTGCTTTGTAGATGTCGTTGATATCTGCAGAATTCGCCC |
| F ultramer end *tnnc1a:eGFP* | GGCGGTCTCTCTATAAAGCTCA |
| R ultramer end *tnnc1a:eGFP* | TGTTTAAAACTCAAAGCACGAAA |
| F promoter tnnc1a | TGCTCAGTGATCCTGTCTGC |
| R intron 1 tnnc1a | AACATCCTGCCACTTACCAA |
| F tnnc1a +9kb | GCCCACAACCCTAAAGTGAA |
| R tnnc1a +9kb | TGTCATTTGCCTCTCACACC |

### *2.16.10 Primers for tnnc1a:eGFP cloning*

| | |
|---|---|
| Tnnc1a F 2.2kb HindIII | GGGCCCAAGCTTGTGTTCCCCAGCGGTCAG |
| Tnnc1a R 2.2kb HindIII | GGGCCCAAGCTTTCACCGCTGCTTTGTAGATG |
| Tnnc1a F 124bp XbaI | CTGCGCTCTAGATTTTTACAGATTACGTTTTGTTTGA |
| Tnnc1a R 124bp XbaI | CTGCGCTCTAGAGCCGACACTAGTGGAGACCA |

### *2.16.11 Primers for Sox6 zinc-finger mutant genotyping*

| | |
|---|---|
| Sox6 ZF F | GGGTGCAGGGTTGTGAAGTG |
| Sox6 ZF R | ATACATGCACATTACTGCAGGTG |
| Sox6 ZF seq | CTTCCTTCTTCCATTTTGTTC |

After PCR using Sox6 ZF F and Sox6 ZF R, product is sequenced using Sox6 ZF seq

### *2.16.12 Primers for Sox5 zinc-finger mutant genotyping*

| | |
|---|---|
| Lsmallsox5 | GATAGCACATCTCGGGAGGA |
| Rsmallsox5 | ATTTGCTGGCGTTGTTTTTC |

After PCR, product is sequenced using Lsmallsox5

### *2.16.13 Primers for cDNA to make probes*

| | |
|---|---|
| Tnnc1a F | GAACCCTACCCCTGAAGAGC |
| Tnnc1a R | GCTCCTTCACGAGACTCCAC |
| Tnnt1 F | CAGAGTTCGAGGAGGAGGTG |
| Tnnt1 R | CAACCTTGGTTCCCTTTTGA |
| Tnnt2 F | CTGAGCTGGTTTTCCTCTGC |
| Tnnt2 R | TCAGCACAGCATTCACTTCC |
| Tnni1a F | TCACATGAATCATTTAGCAGCA |
| Tnni1a R | TCACTGCCTTGAGGGTCTTT |
| Tnni1b F | GAGGCATTGTTCTAGAGCATCA |
| Tnni1b R | TTTGCAGATTACACAATTTATTGACA |
| Tpm2 F | CAGCCGCCTCTCAGCTACGC |
| Tpm2 R | GCTTCGCCACAGACCTCTCAGC |
| Mylz10 F | GTGTCCCTCAGGCTTTTTGAGACCA |
| Mylz10 R | AGGAACGCAACACACTGGAACACA |

### *2.16.14 Primers for ChIP-qPCR*

| | |
|---|---|
| Tnnc1b qPCR F1 | CAAAGAATGCGTCCTTTTCCT |
| Tnnc1b qPCR R1 | AAAAAGCCCATATTTGAATTTTACA |
| Tnnc1b qPCR F2 | TCCTTTCCTTGCTCCTTTTTAAC |
| Tnnc1b qPCR R2 | GTTCTATCAGCGTTTCTGCAGTT |
| Tnnc1b qPCR F3 | TGCTCAATATGCAAGAAGTACCC |
| Tnnc1b qPCR R3 | CCATGTTCTCATTATCACATGACC |
| Tnnc1b qPCR F4 | AATATAGCACCTATCCAGGCACA |
| Tnnc1b qPCR R4 | AATCAGTGCTTGGTTGAAGGTAA |
| Tnnc1b qPCR F5 | GACCGGTGGATAAGTTTAGAACC |
| Tnnc1b qPCR R5 | GAGGGATGAGTGAAAAAGGAAAG |
| Tnnc1b qPCR F6 | CAGGAAACACTGTCAAAATGGAT |

| | |
|---|---|
| Tnnc1b qPCR R6 | GTAACGGTCTTCAGACACCAGAC |
| Smyhc1 qPCR F1 | CGGTGACTGTAAAGAAGCTGAGA |
| Smyhc1 qPCR R1 | TATAGATGAAGGCGATGGACAAC |
| Smyhc1 qPCR F2 | TAAATTGCATATGCTGGCTCTTT |
| Smyhc1 qPCR R2 | AGCAGTCTATTTTGGATCCCTTC |
| Smyhc1 qPCR F3 | CCATCCTCCTTTAAGCCTATTGT |
| Smyhc1 qPCR R3 | CTCCAGAGCTGGTTCTACCTTTT |
| Smyhc1 qPCR F4 | ACCAGCTCTGCAGTTACAAGGTA |
| Smyhc1 qPCR R4 | CTCTGAGAAAAGGAACAGCACAG |
| Myh7b qPCR F1.2 | CTGGCGTTTACAACAAAGGATT |
| Myh7b qPCR R1.2 | CGGCTCTTTAACATCTCCTCAAT |
| Myh7b qPCR F2 | ATTGCTGCATTCCTGTTATTGTT |
| Myh7b qPCR R2 | TGGCTTTTATTTTCATTCTCATTTT |
| Myh7bF1-XG | ACTCGTTGTGTTTGTTTACTTGTCTG |
| Myh7bR1-XG | CTGCTTTACGCACTGAAGACTGA |
| Myh7bF4-XG | TCAAGAAAACATAAATAGCAAGAGAAGT |
| Myh7bR4-XG | TTTCTGAAACATAATGATAACAGGAGA |
| Prox1a qPCR F1.3 | AACAGATTTTGTTGCCATTTTGT |
| Prox1a qPCR R1.3 | AAATTCTGAGACATGCACCTTGT |
| Prox1a qPCR F2 | GTTTCTTTGGCAAATGGTACATC |
| Prox1a qPCR R2 | CTGCATGCCAAAAGTCTAATTCT |
| Prox1a qPCR F3 | TAATTTCGGAAGACTGGTTCTCA |
| Prox1a qPCR R3 | GCACTTTTACCTAAAGCAAGCTG |
| β-actin1_gDNA_Fwd | AAGATAAGAATTGCAAAGCACCA |
| β-actin1_gDNA_Rev | TTAGCATATTCCGGTTTTTCAGA |

# Chapter 3

# Molecular analysis of the zebrafish *slow troponin c* (*tnnc1b*) promoter

### 3.1 Introduction

It is well established that Sox proteins are able to regulate the transcription of their target genes in various ways, through both direct and indirect mechanisms (Kamachi et al., 2000). Based on previous research, whereby sox6 is able to directly repress MyHCβ expression in the mouse (Hagiwara et al., 2007), it is hypothesised that Sox6 directly represses zebrafish slow muscle genes. In order to characterise the interaction between Sox6 and slow muscle genes, a transgenic line of the slow muscle gene, *tnnc1b,* was made using BAC recombineering to enable us to identify the promoter and *cis*-regulatory regions of the gene that are required to confer slow-twitch specific expression.

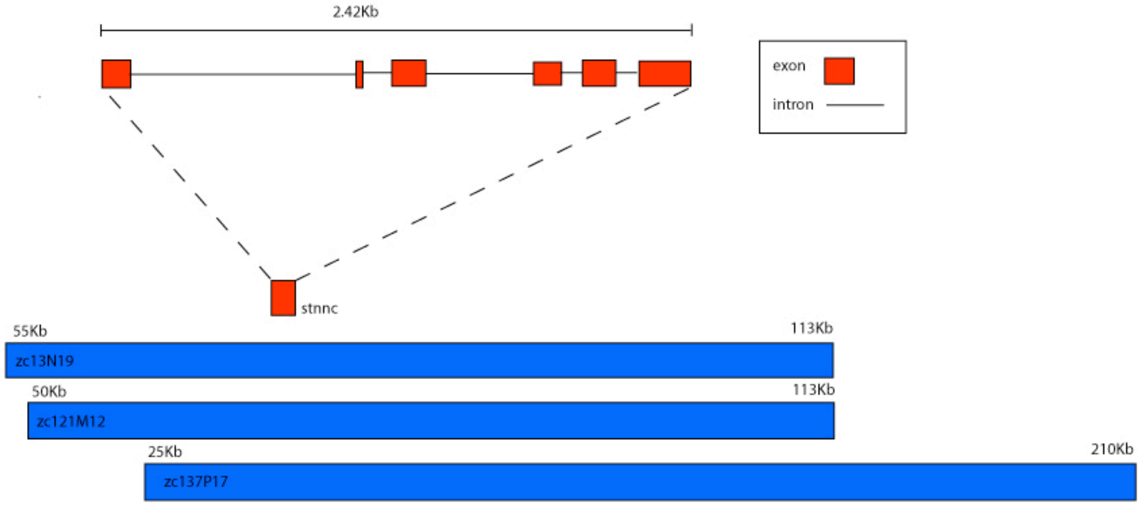


***Figure 3.1: Three BACs encompassing the tnnc1b gene were modified by the addition of a GFP reporter.*** *The three BACs had differing amounts of upstream and downstream sequence from the tnnc1b gene.*

### 3.2 Results

#### 3.2.1 Generation of *tnnc1b:eGFP* reporter line

Three BACs were identified that contained *tnnc1b,* with differing amounts of upstream and downstream sequence flanking the *tnnc1b* exon sequences. The BACs were modified by the insertion of an eGFP reporter gene downstream of the *tnnc1b* promoter (Figure 3.1). The GFP was inserted so that the ATG start codon of the GFP replaced the ATG of the *tnnc1b* gene, enabling the GFP to act as a reporter gene for *tnnc1b* expression. All three modified BACs gave mosaic eGFP expression specifically in the slow muscle when injected into zebrafish embryos, and one of these BACs, zc137P17, was used (after modification with the eGFP cassette) to make a stable transgenic line. The transgenic line *tg*(*tnnc1b:eGFP*) recapitulated expression of the endogenous *tnnc1b* gene as GFP fluorescence was only seen in slow muscle cells (Figure 3.2A-B). This was verified by fluorescent *in situ* hybridisation combined with a GFP antibody stain to show the precise overlap between the expression domains of the endogenous gene and the GFP transgene (Figure 3.2D-F). GFP expression was also observed in the head muscles of transgenic fish from approximately 2dpf onwards (Figure 3.2C). Most head muscles at this embryonic stage of development are a mix of slow and fast-twitch fibres (Hernandez et al., 2005) making it difficult to discern, without the use of another slow-twitch marker such as F59, whether the *tnnc1b:GFP* transgenic line has slow-specific GFP expression in the head. GFP expression in the transgenic line was also visible in neurons of the embryo, most likely the trigeminal ganglion (Figure 3.3A). As another control for the slow-twitch muscle specificity, the *tnnc1b:eGFP* reporter line was crossed to the Tg(*smyhc1:lyntdtom*) reporter line, which recapitulates the expression of *smyhc1.* Both reporter lines overlapped in the slow-twitch muscle of the embryo (Figure 3.3).

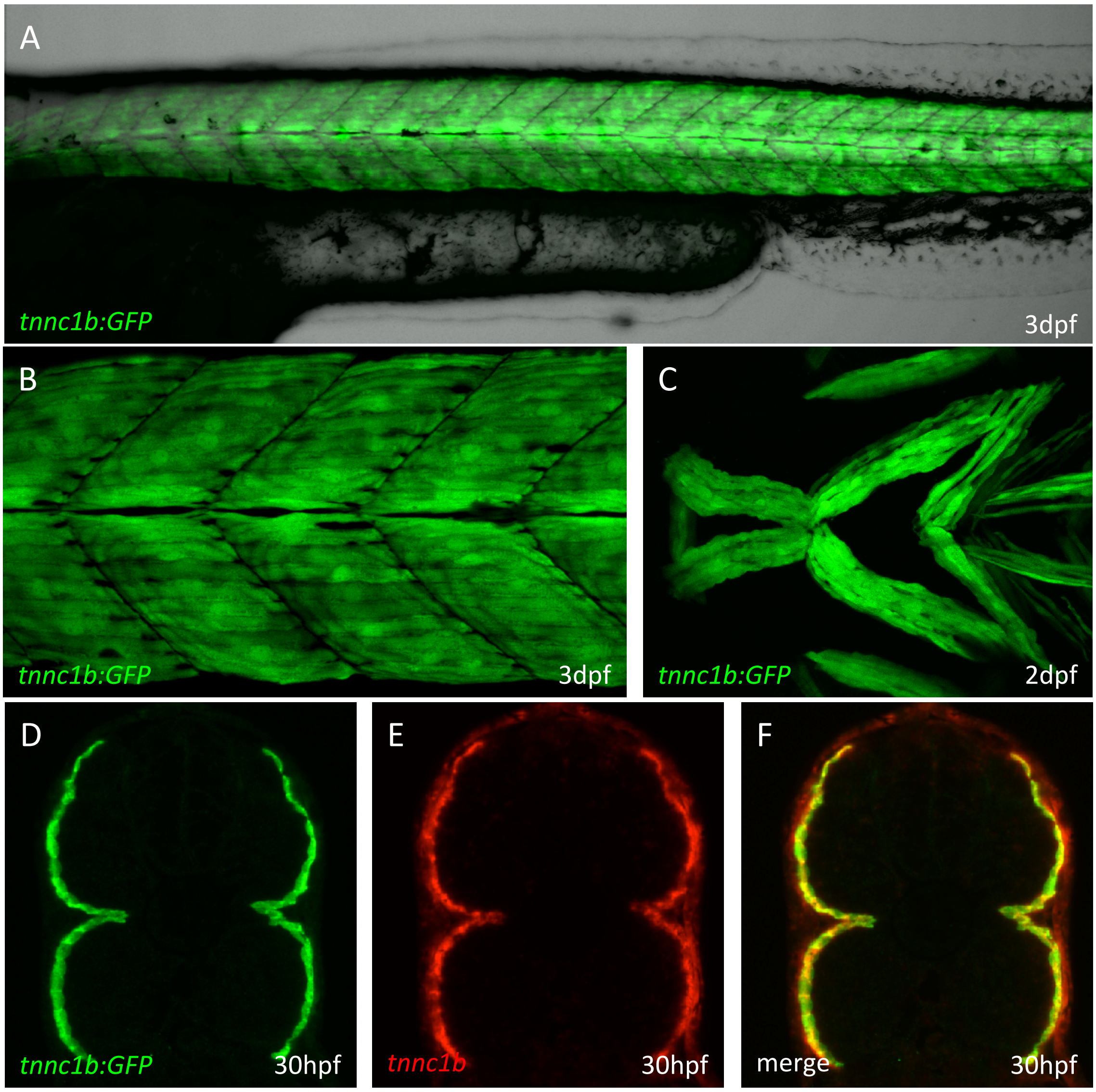


***Figure 3.2: A tnnc1b:eGFP reporter line was generated using BAC recombineering.*** *Expression was seen specifically in the slow-twitch muscle (A-B), and from 2 dpf expression could be seen in jaw muscles of the embryo, specifically the adductor mandibulae intermandibularis anterior, intermandibularis posterior and the sternohyal. (C). The tnnc1b:eGFP transgenic reporter line (Green) replicates the endogenous expression of the gene (Red) as assayed for by fluorescent in situ hybridisation (D-F).*

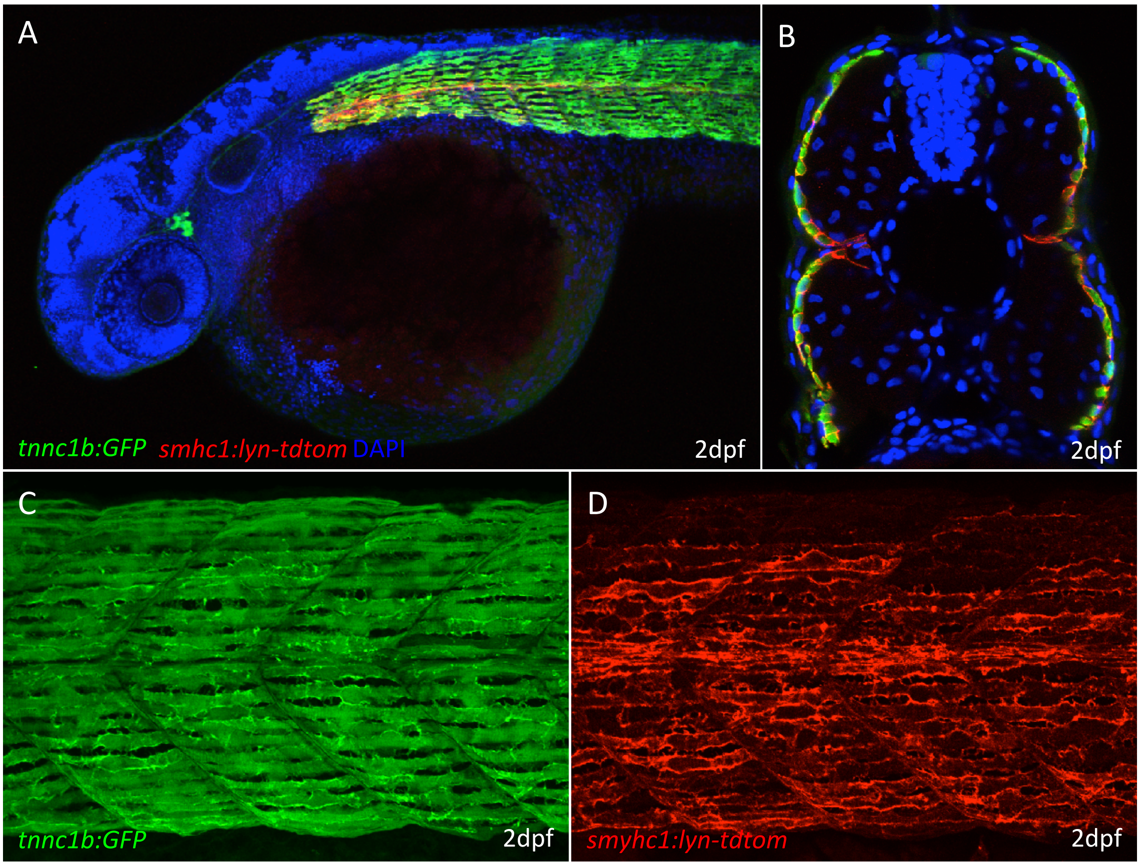


***Figure: 3.3 tnnc1b:eGFP and smyhc1;lyn-tdtom are slow-twitch muscle specific reporter lines*** *(A-D). Expression of tnnc1b:eGFP is also seen in a cluster of neurons in the head of the fish whereas smyhc1;lyn-tdtom is not seen in these cells (A).These neurons most likely represent sensory neurons of the trigeminal ganglion.*

To verify whether *tnnc1b:eGFP* is Hedgehog-responsive, the zc137P17 BAC was injected into *smo* mutants, which have compromised Hedgehog signalling (Barresi et al., 2000, Varga et al., 2001). No GFP expression was seen in the muscle of any of the injected homozygous mutant embryos, indicating that the expression of *tnnc1b* is strictly dependent on Hh signalling (Figure 3.4). Interestingly the GFP expression seen in the neurons of the fish was still present in *smo* mutants (Figure 3.4). This not only acted as a positive control for the injection of the BAC, but also indicated that this neuron-specific expression is not Hedgehog-dependent.

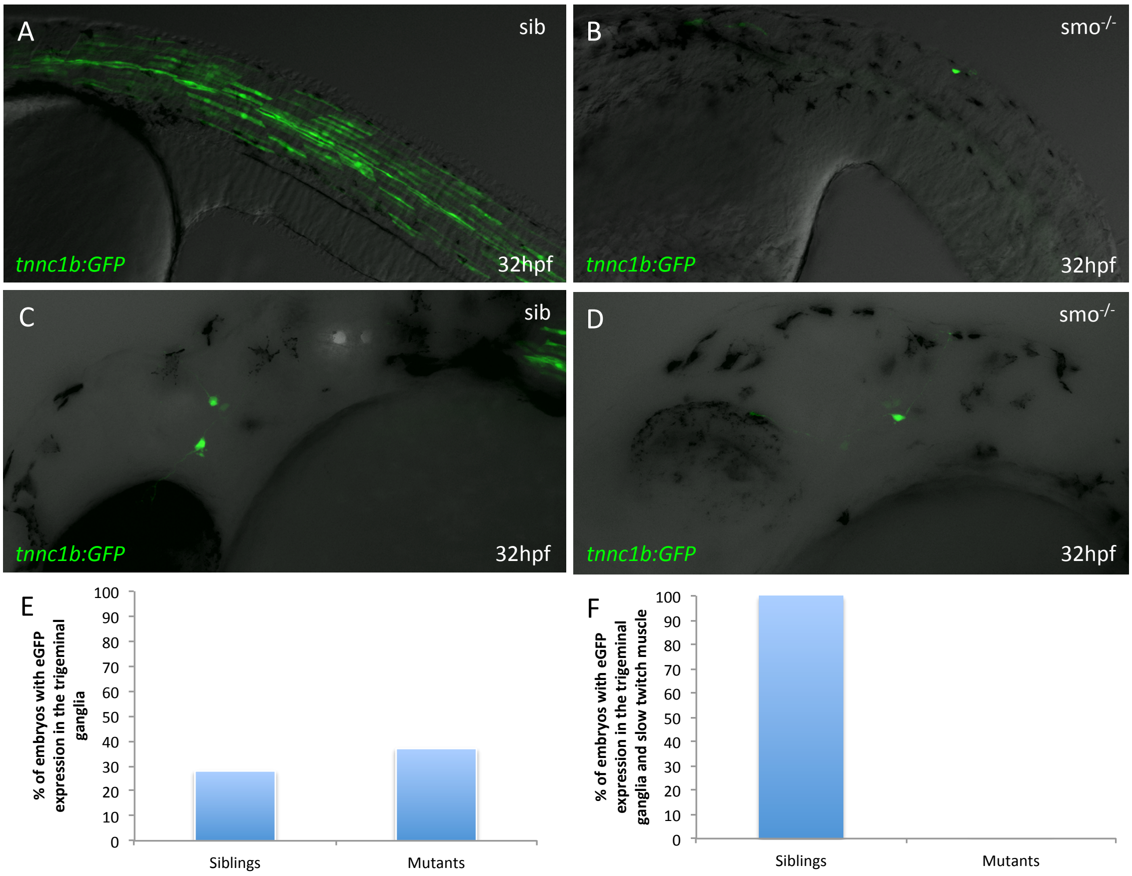


***Figure 3.4: The tnnc1b:eGFP reporter construct is Hedgehog-responsive.*** *The tnnc1b:eGFP BAC construct (zc137P17) with tol2 transposase was injected into embryos from a smo$^{+/-}$ incross. The neuron-specific GFP expression served as a positive control for successful injection of the construct as it was not Hedgehog-responsive (C and D). 28% of siblings and 38% of mutants had GFP expression in one or more of these neurons (siblings n=122, mutants n=43, total n=165) (E). Of the embryos that expressed GFP in these neurons, 100% of the siblings expressed GFP in the muscle (3 or more fibres) (A and F) and no mutants expressed GFP in the muscle (3 or more fibres, with 38/43 fish displaying no muscle specific expression at all) (B and F).*

### 3.2.2 Deletion derivatives of original BAC construct

In order to investigate the *cis*-regulatory elements that mediate slow-specific expression of *tnnc1b,* a series of deletions were made in the original BAC construct by homologous recombination in *E. coli*. Firstly, approximately 209kb of DNA sequence downstream of the last *tnnc1b* exon was removed from BAC zc137P17, leaving 1kb downstream of the stop codon and the 25kb region of upstream sequence intact. The transgenic lines created from the injection of this construct retained their slow-specific expression, therefore a series of deletions were then made in the upstream sequence (Figure 3.5). All embryos that were GFP positive after injection of the constructs were grown to adulthood and screened, whether the GFP expression was observed in the fast-twitch, slow-twitch or both domains.

All transgenic lines generated from the deletion derivatives of the BAC retained slow-specific GFP expression (summarised in Table 3.1).

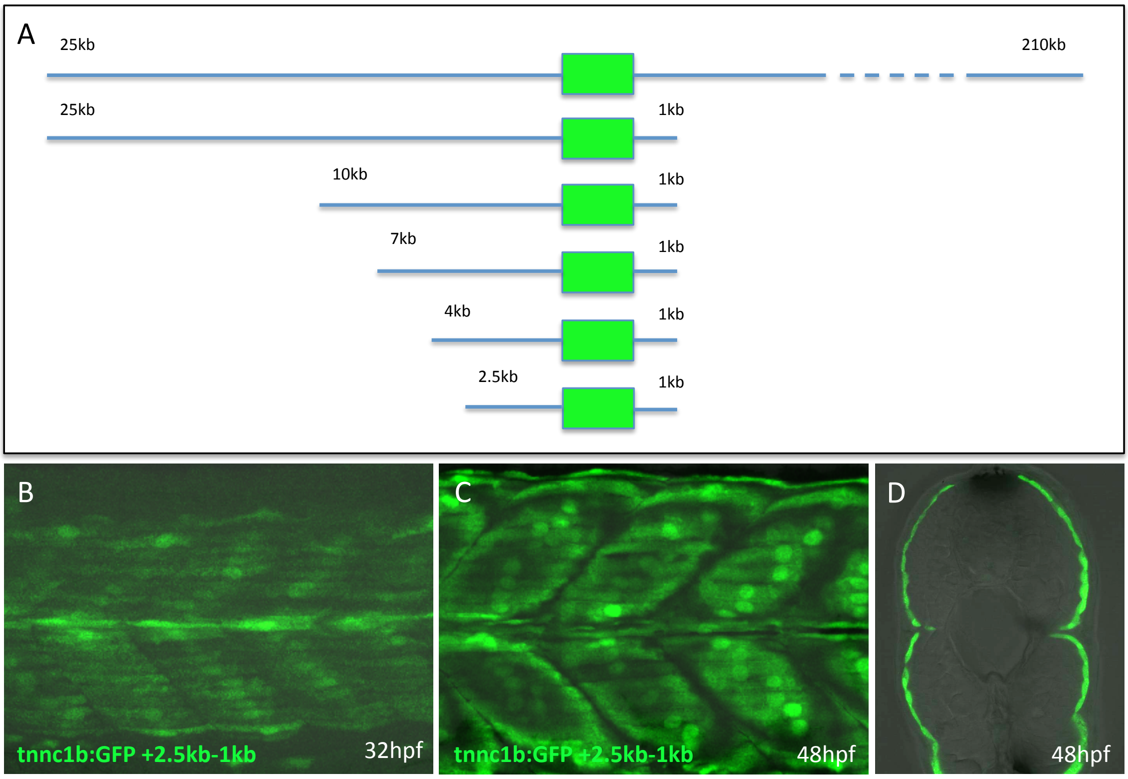


***Figure 3.5: A series of deletions were made to the original BAC construct and all transgenic lines generated from these constructs retained slow specific expression.*** *Summary of deletion derivative made (A). A construct containing 2.5kb of upstream sequence from the translation initiation codon and 1kb of downstream sequence from the translation stop codon, exhibited slow-specific GFP expression (B-D) Green box represents GFP.*

| **Line** | **Founder** | **Expression Pattern** |
|---|---|---|
| *tnnc1b:eGFP* | 1 | Slow-twitch muscle and neurons |
| *tnnc1b:eGFP* +25kb-1kb | 1<br>2 | Slow-twitch muscle and neurons<br>Slow-twitch muscle and neurons |
| *tnnc1b:eGFP* +10kb-1kb | 1<br>2 | Slow-twitch muscle and neurons<br>Slow-twitch muscle and neurons |
| *tnnc1b:eGFP* +7kb-1kb | 1<br>2 | Slow-twitch muscle and heart<br>Slow-twitch muscle and neurons |
| *tnnc1b:eGFP* +4kb-1kb | 1<br>2 | Slow-twitch muscle and neurons<br>Slow-twitch muscle |
| *tnnc1b:eGFP* +2.5kb-1kb | 1<br>2 | Slow-twitch muscle and neurons<br>Slow-twitch muscle and neurons |

***Table 3.1: A summary of the expression characteristics of stable transgenic lines generated from the modified BAC zc137P17 construct.***

Further deletion analysis of this BAC construct was then performed by cloning DNA segments of the BAC into the pDB739 vector that contains iTol2 arms. A series of constructs were made this way and stable transgenic lines were generated with all of them (summarised in Table 3.2). All embryos that were GFP positive after injection with the constructs were grown to adulthood and screened, whether the GFP expression was observed in the fast- or slow-twitch domain or both domains. It appears that that the first intron of *tnnc1b* is vital for GFP expression. When constructs where made which lacked this first intron, no GFP expression was observed, indicating the presence of an enhancer. A construct comprising 560bp upstream of the *tnnc1b* start site plus the first intron, gave rise to GFP expression in the fast muscle in the progeny of one founder. However, in contrast, the progeny of two other founders, generated with the same construct, exhibited slow-specific GFP expression. The offspring of other founders also exhibited ectopic GFP expression in the heart, although no clear trend emerged, making the elucidation of a heart-specific repressor element difficult. A summary of the transgenic lines generated and the expression patterns of their transgenes are shown in Table 3.2.

| Line | Founder | Expression Patter |
|---|---|---|
| *tnnc1b:eGFP* + 2.5kb + intron1 | 1 | Slow-twitch muscle and neurons |
| | 2 | Slow-twitch muscle and neurons |
| *tnnc1b:eGFP* + 2.5kb | N/A | No GFP expression seen in any embryos injected so no fish were grown up |
| *tnnc1b:eGFP* +1kb +inton1 | 1 | Slow-twitch muscle, nose and neurons |
| | 2 | Slow-twitch muscle and neurons |
| | 3 | Slow-twitch muscle and neurons |
| *tnnc1b:eGFP* + 760 + inton1 | 1 | Slow-twitch muscle and heart |
| | 2 | Slow-twitch muscle |
| | 3 | Slow-twitch muscle and neurons |
| *tnnc1b:eGFP* + 560 +intron1 | 1 | Slow-twitch muscle and fast-twitch muscle (weak) |
| | 2 | Slow-twitch muscle, heart and neurons |
| | 3 | Slow-twitch muscle and heart |
| *tnnc1b:eGFP* +270 +intron1 | 1 | Slow-twitch muscle and heart |
| | 2 | Slow-twitch muscle and heart |

***Table 3.2: List of tnnc1b:eGFP pDB739 lines generated including number of founders and expression pattern***

### 3.2.3 Cloning of *tnnc1b* first intron into a *beta-globin* minimal promoter vector

As the loss of the first intron of *tnnc1b* resulted in a complete loss of GFP expression, this region was cloned into a *β-globin* minimal promoter vector to see if it was acting as an enhancer (Figure 3.6D). This vector was made by inserting a *β-globin-eGFP-polyA* cassette into the pDB739 plasmid, which enabled the testing of the regulatory potential of suspected enhancer elements. The GFP expression seen in the progeny of two founder fish for the *β-globin-GFP:intron1-tnnc1b* reporter, was seen not only in the slow-twitch muscle but also throughout the fast muscle (Figure 3.6A and B). Interestingly a skeletal muscle-specific enhancer has been shown to reside in the first intron of the mouse *tnnc1 gene,* indicating the possible presence of a conserved non-coding element (CNE) (Parmacek et al., 1994). Since the mouse and human first intron of *tnnc1* have high sequence conservation (Parmacek et al., 1994), the first intron of human *tnnc1* gene was cloned into the *β-globin* minimal promoter vector and this construct was injected into zebrafish embryos (Figure 3.6D). Two founders were identified which had the same expression pattern as the zebrafish intron1 construct (*β-globin-GFP:intron1-tnnc1b)*, i.e. in which expression was seen in both the slow and fast twitch muscle (Figure 3.6C). This suggests the presence of a conserved skeletal muscle enhancer in the first intron of *tnnc1.*

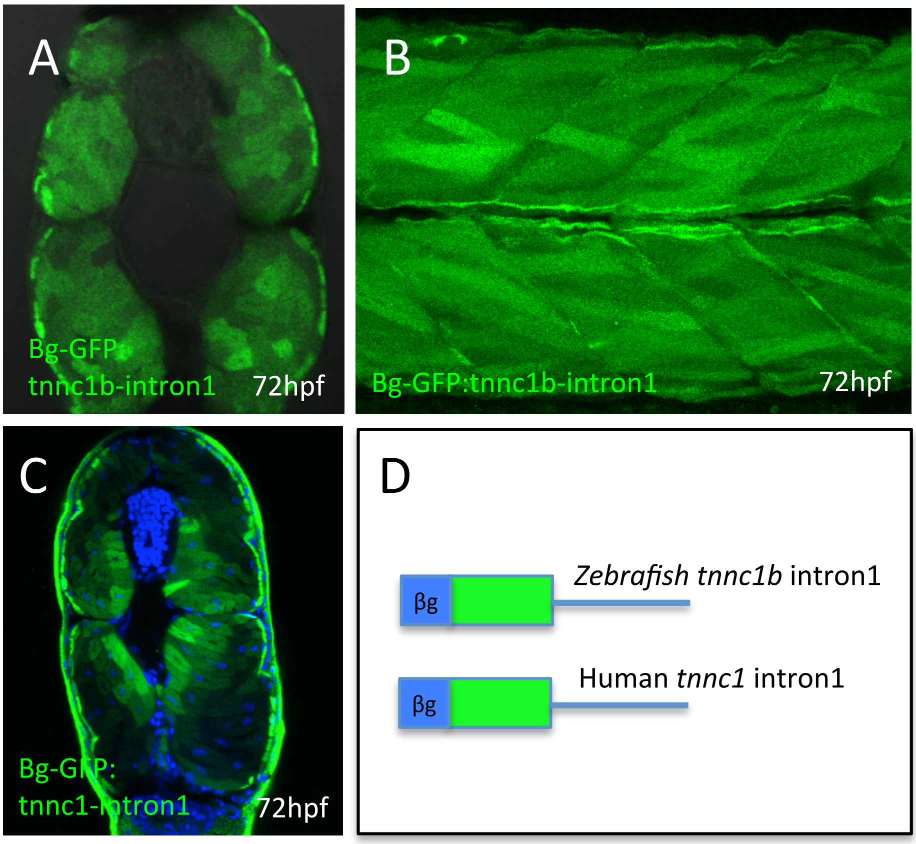


***Figure 3.6: Zebrafish slow troponin c contains a conserved enhancer in intron1.*** *The zebrafish (A and B) and the human (C) first intron of slow troponin c can driver transgene expression in both the fast-twitch and slow-twitch muscle. Schematic of constructs generated (D). (Skin staining in panel C is a staining artefact and was not visible in embryos before immunohistochemistry staining).*

### 3.2.4 Mutating potential Sox6 binding sites

Six potential Sox6 binding sites were identified and mutated in the 2.5kb + intron1 construct (Figure 3.7). These sites were mutated cumulatively, one after another, culminating in the mutation of all six putative Sox6 binding sites. Each of the resulting six constructs was used to make a stable transgenic line (Figure 3.8). No ectopic GFP expression was seen in the fast-twitch muscle of any of the lines (Table 3.3).

```
CAAAGAATGCGTCCTTTTCCTGGTCTTTAAAGTGAATAATCTTAACAAACATAGAAGACTT
......1kb......TTCATGTCTGTAACAATATAATAGTCAGTAAACTGCAGAAACGCTGATAGAACT
TGCTCAATATGCAAGAAGTACCCACTGCTTCCAGCAAGATTCTGAAGCATGCATCCAACGT
ACACATTACTATCCCATAAGGCAATCCAAAACAATTTATAAATGGGAGTGAAGCGACGCAA
CTGACAAGGTCATGTGATAATGAGAACATGGTGGATATAGTACCACTCATAATGCTCATAT
TCACACTACTGTATATACAAGTTCATAAAGTAAATTTAAATTCAATATAGCACCTATCCAG
GCACAGCTTATTTATTTCACTTCACAGCATTTATCCATTCAAACAATTGCTGTAACTATAG
ATTATTAAAAATCACACATTTTCTATATGATTTGGTCAAAGAAACCTACTGGAAGGGCAAT
ATTAAGTTAATTACCTTCAACCAAGCACTGATTGATTGGACCAGTGCTACTAGTCTCTTAG
GTTACAGTCATGTTCATACGTATATGACCAGTGCCTGGCACAGGAATGGAGCTGTGATGCT
GCTCAGCAAATGATTGGCAAATGCAGGTGACCGGTGGATAAGTTTAGAACCTGATATGGGA
TTGTTGCTTAGGCGGCCAGAAAGACCAGTGCCAAGGACTTATTGTCATAACTCGAGTTCTC
CTGCCTGTTGTCTCTGAAATTCCTCCTCTATTTATAGCCCATACCATCCCTAACCCCTTTC
TGCCCTGCTCCCTTTCCTTTTTCACTCATCCCTCGCTCTCTCTGATAGGAAAGGGGAGT
GACAAAAGCCATTGTCTTTTCTAAGCCATCCATTCCAGCCATAAAGGCCAAGGCAGGTGGC
CCTGGGACTCTTTGTCCTGTGTCTGCCGTGAAGAGAGGAGCGCTGGATTACCTGTGAGGAG
TTTACTGACCATTTCAAAAAAGGAAAATCTAAACCCGGGACAGCCATGgtgagcaagggcg
aggagctg......GFP......676bp......tggacgagctgtacaagtaaGATGATGTATATAAAGCAG
CGGTAAGTTATGGGCACAGTAATTA......700bp......ACACTGAGGTTGAAGCATTTAACAATG
AGGGTGTCTGGTGTCTGAAGACCGTTACGCAAGTGACAGCTGTCATTTCTGTAAAGCACAT
ACCTTGGTCGGTCTATGATGAACGACAAGCTATATAGTCTTACAACAAATTGTATTTTTAG
TTCTGAGTCAACTGGAGCCAATTTTGAGCTC


Sox6 site 1  TAACAA
             GGGT

Sox6 site 2  AACAAT
               AGGG

Sox6 site 3  AACAAT
               AGGG

Sox6 site 4  AACAAT
               AGGG

Sox6 site 5  GATTGTT
             GGGT

Sox6 site 6  AACAATG
               AGGG
```

***Figure 3.7: Sox6 sites were mutated in the construct containing 2.5kb upstream of the tnnc1b start site and the first intron.*** *Six sox6 sites were identified and mutated (highlighted in yellow) Red box below represents the nucleotide mutations that were introduced into the construct by PCR. GFP is highlighted in green. Exon1 is highlighted in dark red.*

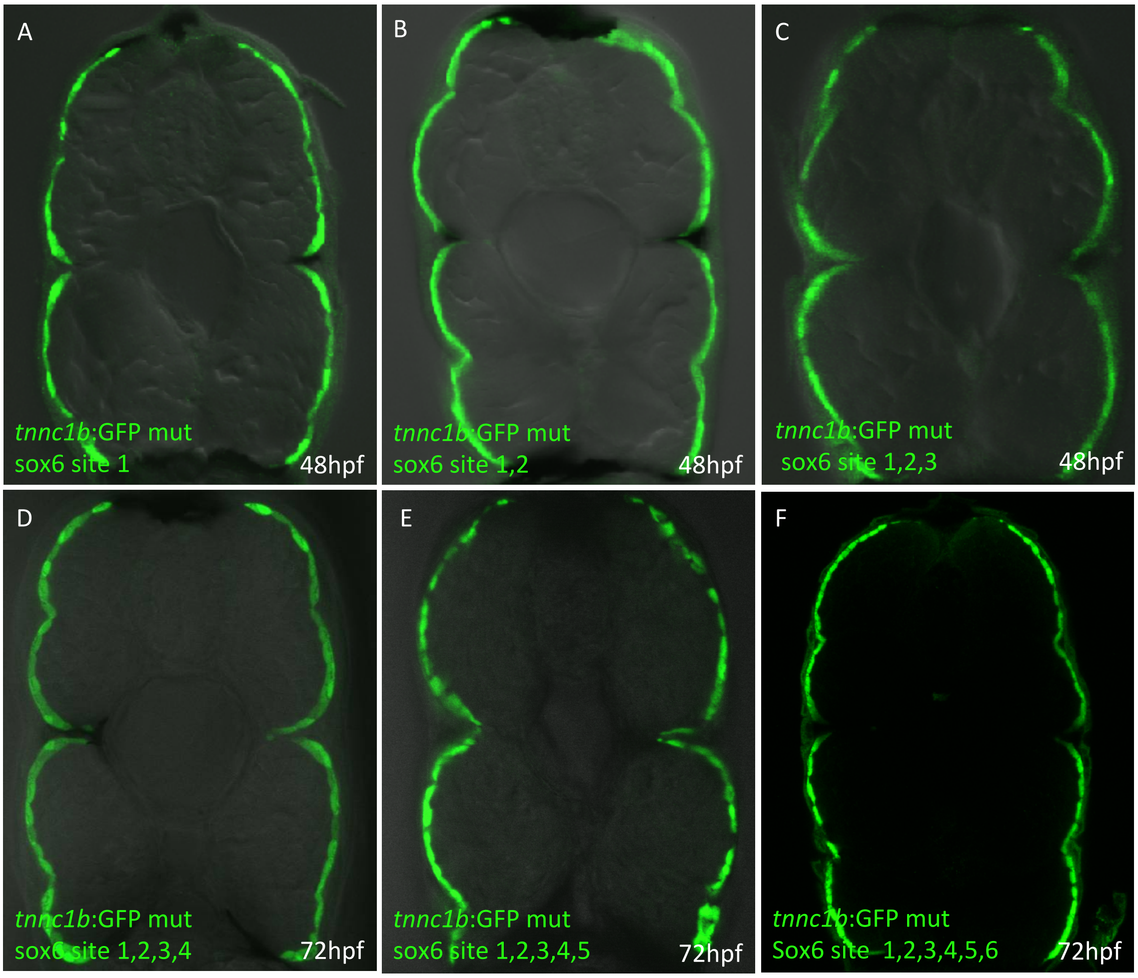


***Figure 3.8: Mutation of potential Sox6 binding sites in the tnnc1:eGFP reporter construct had no effect on GFP expression.*** *GFP expression remained restricted to the slow domain when one or more putative Sox6 binding sites was mutated (A-F).*

| **Line (*tnnc1b:eGFP sox6sitemut*)** | **Founder Number** | **Expression Pattern** |
|---|---|---|
| 1 | 1 | Slow muscle |
| | 2 | Slow muscle (strong) |
| | 3 | Slow muscle, Heart |
| 1,2 | 1 | Slow muscle |
| | 2 | Slow muscle (weaker than 3) |
| | 3 | Slow muscle, heart (very weak) |
| 1,2,3 | 1 | Slow muscle |
| | 2 | Slow muscle, heart, neuron |
| 1,2,3,4 | 1 | Slow muscle (strong), neuron, heart |
| | 2 | Slow muscle weak |
| 1,2,3,4,5 | 1 | Slow muscle (weak), neuron |
| | 2 | Slow muscle (very weak) |
| | 3 | Slow muscle, neuron |
| | 4 | Slow muscle, neuron |
| 1,2,3,4,5,6 | 1 | Slow muscle, neuron, heart |
| | 2 | Slow muscle, neuron |
| | 3 | Slow muscle, neuron |

***Table 3.3: Summary table of the tnnc1b:eGFP Sox6 site mutant lines generated, with the expression patterns for each founder indicated.*** *All lines had slow-specific expression.*

### 3.2.5 Generation of *tnnc1a:eGFP* transgenic line

As was mentioned in Chapter 1, in contrast to mammals, which have a single *slow troponin c (Tnnc1)* gene that is expressed in both slow-skeletal and cardiac muscle, zebrafish have two paralogous genes, *tnnc1a* and *tnnc1b,* which are expressed specifically in the heart and slow-twitch skeletal muscle, respectively (Sogah et al., 2010). Since the results described above suggest that the skeletal muscle enhancer is conserved between zebrafish and humans, it was investigated whether this was also the case for the heart enhancer. Previous studies have localised the mammalian *Tnnc1* heart muscle enhancer to sequences encompassing the transcription start site of the *Tnnc1* gene (bp -123 to +32). The BAC zc117LI9 which encompasses the *tnnc1a* gene along with 125kb of upstream sequence and 35kb of downstream sequence was modified by the addition of a GFP reporter construct so that the start site of the GFP reporter replaced the start site of the *tnnc1a* gene (same strategy that was used for the generation of the *tnnc1b* transgenic reporter line). iTol2 sites were then inserted into the BAC by recombination and a stable transgenic line was then generated which expressed GFP specifically in the heart muscle (Figure 3.9). Expression could be seen clearly throughout the heart from approximately 48 hpf with GFP fluorescence appearing slight stronger in the ventricle.

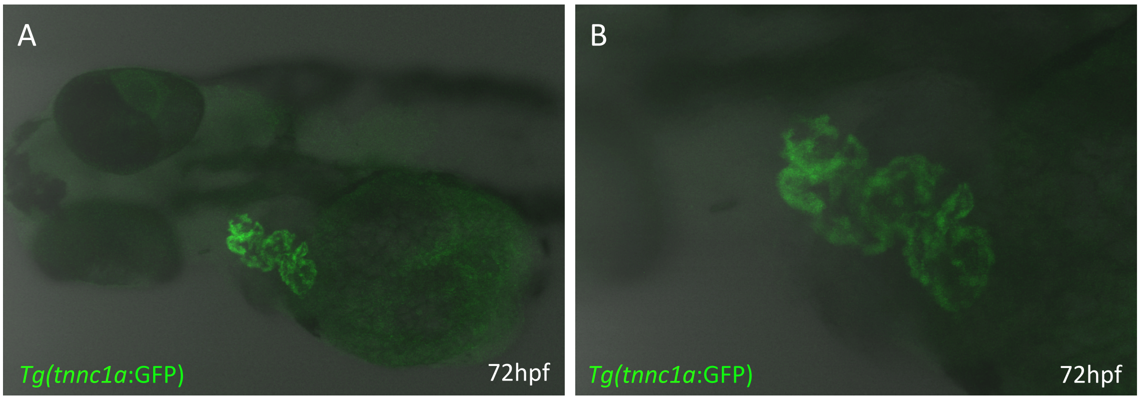


***Figure 3.9 tnnc1a:eGFP is expressed exclusively in the heart muscle and excluded from the skeletal muscle.*** *Two founders were identified which carried the transgene and both had the same expression pattern.*

In order to try and identify the enhancer I cloned approximately 2.2kb upstream of the start site and the GFP into the pDB739 vector and also cloned a 240bp immediate upstream fragment into the βglobin minimal promoter vector. However, when both of these constructs were injected into zebrafish embryos there was no GFP expression. More analysis is needed to determine the cis-regulatory regions of this gene.

## 3.3 Discussion

### 3.3.1 *tnnc1b* promoter analysis

The identification and characterization of the *cis*-regulatory regions of genes is crucial for understanding how the transcriptional programmes and signalling pathways that orchestrate development are functionally integrated (Li and Davidson, 2009). The key components of the transcriptional mechanisms that regulate expression of zebrafish slow-twitch muscle-specific genes are largely unknown. Previous experiments in mice and zebrafish have implicated the transcription factor Sox6 as a negative regulator of these genes (An et al., 2011, Hagiwara et al., 2007, von Hofsten et al., 2008). Although Sox6 is known to be able to directly inhibit slow-specific genes in mice (An et al., 2011), it is unknown whether this is the case in zebrafish. Previous unpublished work in our laboratory has suggested that there is a direct interaction between Sox6 and the *smyhc1* promoter, because mutating the potential Sox6 binding sites in a *smyhc1:GFP* construct, which lie upstream of the *smyhc1* proximal promoter sequences, resulted in the ectopic expression of GFP in several fast-twitch fibres. Moreover, a recent study has indicated that the mutation of Sox6 binding sites in a fugu slow MYH:eGFP reporter cassette results in ectopic GFP expression in the fast-twitch muscle when injected into zebrafish embryos (Asaduzzaman et al., 2013). In order to determine whether another slow-twitch-specific muscle gene exhibited ectopic expression in fast muscle when its regulation by Sox6 was compromised, a transgenic reporter line was generated in which the expression of GFP was placed under the control of the *tnnc1b* promoter. This construct enabled us to identify the minimal promoter and enhancer regions for the *tnnc1b* gene required for slow-twitch specific transcription, and to then identify and mutate any potential Sox6 binding sites in this region in order to establish whether there is a direct functional interaction between the regulatory elements of this gene and the Sox6 transcription factor.

The *tnnc1b:GFP* transgenic reporter line faithfully recapitulated the endogenous *tnnc1b* expression pattern in zebrafish embryos, being restricted to slow-twitch muscle fibres. The transgenic line thus provides us not only with a useful readout of *tnnc1b* expression but also a new slow-twitch muscle-specific fluorescent reporter with which the behaviour of slow-twitch muscle fibres can be studied during embryonic and larval development. The *tnnc1b:GFP* transgenic reporter was hedgehog-responsive, confirming previous reports that describe the dependency of genes involved in slow-twitch muscle development on hedgehog signalling (Barresi et al., 2000, Baxendale et al., 2004, Blagden et al., 1997, Du et al., 1997, Hirsinger et al., 2004, Lewis et al., 1999, Roy et al., 2001).

A common strategy to identify enhancers of genes is to search for conserved non-coding elements (CNEs) bioinformatically, using software tools such as VISTA (Frazer et al., 2004). This type of approach can identify genomic regions within a specified locus that have high nucleotide similarity with putative orthologous regions in other species. An mVISTA analysis (Frazer et al., 2004) was performed that compared the *tnnc1b* gene of zebrafish to its orthologous genes in medaka, stickleback, tetraodon and fugu (all of which have two *tnnc1* genes), as well as to mouse and human *tnnc1* orthologues (Appendix A2). Although the protein coding region of this gene is highly conserved (represented by the three pink peaks) the upstream region of the gene and the introns are not conserved (with a threshold of 50% identity). Since no CNEs were recognised bioinformatically, deletion analysis of the initial reporter construct was then performed. The deletion analysis of the *tnnc1b* promoter led to the discovery that the minimal regulatory elements determining the slow-specific transcription of *tnnc1b* lay within a region encompassing 240bp immediately upstream of the transcription start site, and the first intron of the gene. The loss of the first intron of *tnnc1b* resulted in a complete loss of GFP expression, identifying a possible enhancer in this region. This intron was therefore cloned into a *β-globin* minimal promoter vector containing a GFP reporter and the resulting construct tested for its ability to drive muscle fibre-type-specific gene expression. The results revealed that this intron could drive GFP expression in both the fast- and the slow-twitch muscle, when linked to a *β-globin* minimal promoter. This result suggested that a muscle-specific enhancer may lie in the first intron of *tnnc1b* and further indicated that a repressor element may lie within the 240bp of DNA sequence immediately upstream of the transcription start site. Since previous studies by other researches had shown that a skeletal specific enhancer is located in the first intron of the mouse *Tnnc1* gene, which is highly conserved with the human *Tnnc1* gene (Parmacek et al., 1994), a DNA fragment containing the first intron of the human *Tnnc1* gene was also cloned into the *β-globin* minimal promoter vector, and its ability to drive fibre-type-specific gene expression was tested by microinjection into zebrafish embryos. This construct exhibited GFP expression in both the slow- and fast-twitch muscle, mirroring what is seen with the zebrafish intron1 *tnnc1b* construct. These results suggested the presence of an evolutionarily conserved enhancer in the first intron of the zebrafish and human *tnnc1* genes.

Minimal promoter vectors are a commonly used tool for assessing enhancer activity in zebrafish (Bessa et al., 2009, Fisher et al., 2006b). The minimal promoter vector employed in this study (human *β-globin*), does not drive any pattern of expression on its own when injected into zebrafish embryos ((MacDonald et al., 2010), my data), and only responds specifically to the activity of the enhancer that is cloned into the vector. However, care is needed when interpreting the results of this type of analysis, as problems with this technique can result from the random integration of this vector into the genome, which can often lead to the reporter gene becoming exposed to enhancer activities present in the surrounding genomic regions. Such "Position Effects" (Chung et al., 1993) could result in reporter gene expression patterns that do not result from enhancer activities within the transgene construct itself. This phenomenon was corrected for, as at least two or more stable lines for each construct were generated, which were derived from independent insertions into the genome that had a reproducible GFP expression pattern. To minimise the possibility of an effect of the genomic position at which the transgene was inserted, on the reporter gene expression pattern, one could insert insulator sequences into the vector, which would increase the specificity of the transgene by protecting it from the transcriptional silencing effects of adjacent heterochromation (Bessa et al., 2009, Chung et al., 1993). Another way to optimize the analysis would have been to introduce a positive control for transgenesis into the construct, such as a fluorescent reporter gene under the control of the eye lens-specific crystalline promoter (Kwan et al., 2007). This would mean that constructs which exhibited no GFP expression, e.g. the *tnnc1b* construct that lacked the first intron, could still be injected into zebrafish embryos and a stable line could be created based on expression of the linked reporter in the eye. The lack of GFP expression in the muscle of these fish would then confirm that the first intron is absolutely required for the expression of the gene.

Regardless of the suggested improvements to the transgenic strategy, based on the results of the experiments performed, it is tempting to speculate that the first intron of *tnnc1b* contains a *cis*-regulatory conserved non-coding element (CNE) that is needed to drive the expression of *tnnc1b* in the slow-twitch myotome. The conservation of the enhancer activity in the mammalian *Tnnc1* first intron and the zebrafish *tnnc1b* first intron, despite the absence of any clear-cut DNA sequence conservation, adds to an accumulating body of evidence which indicates that conservation of CNE sequence is not necessary for the conservation of CNE regulatory activity (Fisher et al., 2006a, Friedli et al., 2010, McGaughey et al., 2008, Visel et al., 2008, Weirauch and Hughes, 2010). Although there appears to be little DNA sequence conservation between the first introns of *slow troponin c* in zebrafish and either mouse or human, the ability of the human and zebrafish introns to drive the same GFP expression pattern in zebrafish suggests that there is a functional conservation of regulatory activity. It was previously thought that DNA sequence divergence represents a mechanism for the divergence of function, yet studies such as ours indicate that this relationship does not always hold. For example, when the *cis*-regulatory regions of the human and zebrafish receptor tyrosine kinase gene, *ret,* orthologues were compared, it was found that despite a lack of DNA sequence conservation in the vicinity of the gene in each of these species, the transgene expression pattern conferred by the human *cis*-regulatory sequences closely mimicked that of the zebrafish orthologue (Fisher et al., 2006a).

There are thought to be two types of mechanism that govern how a *cis*-regulatory DNA sequence can be altered without any change in the transcriptional regulation of the gene under its control. Firstly, a transcription factor regulating the expression of a gene may also be changed along with the change in *cis*-regulatory DNA sequence, to compensate for the change in DNA sequence (Weirauch and Hughes, 2010). Secondly, it is possible that the *cis*-regulatory sequences may be rearranged and dispersed into distinct locations and yet retain their collective functionality, due to the ability of enhancers to act independent of position and orientation, relative to the promoter (Weirauch and Hughes, 2010). Further experimental and bioinformatical analysis would be needed to confirm which of these mechanisms is correct.

### 3.3.2 Mutation of sox6 binding sites

Sox6 is able to directly inhibit *Tnnc1* transcription in mice (An et al., 2011). Previous studies in our laboratory have suggested that Sox6 inhibits the expression of *tnnc1b* in zebrafish (von Hofsten et al., 2008), but it remain unclear as to whether this inhibition is a consequences of a direct interaction between Sox6 protein and *tnnc1b cis*-regulatory sequences, or not. To address this question, potential Sox6 binding sites were identified bioinformatically and then mutated in the 2.5kb + intron1 transgene construct, which drives slow-specific expression of GFP. Unexpectedly, mutation of these sites resulted in no ectopic GFP expression in the fast-twitch muscle. This result suggested that Sox6 inhibits the expression of *tnnc1b* through a mechanism that is independent of the conserved putative Sox6 binding sites in intron 1, and potentially indirectly, in direct contradiction to what was previously reported in mice (An et al., 2011).

Another possibility that could help to explain my results is that Sox6 might exert its repressive effects on *tnnc1b* through multiple sites, including one or more cryptic Sox6 binding sites. The binding of Sox proteins to A/T rich sequences is well documented, the preferred consensus sequence being AACAAT (Connor et al., 1995, Harley et al., 1994). However, Sox proteins can bind to variants of this DNA sequence, albeit with a lower affinity (Harley et al., 1994). As described above, the deletion analysis of the *tnnc1b* reporter construct revealed that the first intron of the gene is able to drive GFP expression in both the fast-twitch and slow-twitch muscle fibres. This observation suggests that there is a repressor element present in DNA sequences lying outside of the first intron that is needed for the repression of *tnnc1b* in the fast-twitch muscle. Further deletion of the upstream DNA sequence revealed that 240bp of DNA sequence lying immediately upstream of the transcription start site plus the first intron is sufficient for slow-twitch-specific GFP expression. It may be that cryptic binding sites for Sox6 exist in this region, which allows Sox6 to directly repress gene expression in the fast-twitch muscle. Interestingly, one out of three of the founder fish that I identified carrying the +560bp+intron1 construct exhibited relatively weak GFP expression throughout the slow-twitch and fast-twitch muscle. It may be that this weak expression is caused by a position effect, as only one founder was identified with this weak expression pattern. Another possibility is that there may have been a cryptic Sox6 binding site in the genomic region immediately upstream of the DNA sequences contained within the construct that was lost, leaving fewer, low-affinity, cryptic, Sox6 binding sites in the remaining region, and making it more likely that Sox6 is unable to inhibit the expression of the transgene. However, in order to confirm whether there is a direct or indirect inhibition of the *tnnc1b* promoter by Sox6, extensive chromatin immunoprecipitation analysis of the genomic region encompassing *tnnc1b* would need to be performed, to examine whether there is an enrichment of *tnnc1b*-associated sequences in the Sox6-precipitated chromatin.

### 3.3.3 *tnnc1a* promoter analysis

The zebrafish has a second *slow troponin c* gene, *tnnc1a*, and a preliminary bioinformatics analysis was performed to explore its DNA sequence similarity with zebrafish *tnnc1b* and other vertebrate orthologues. Perhaps unsurprisingly, VISTA analysis indicated that the non-coding regions between the zebrafish *tnnc1a* and *tnnc1b* genes are not conserved (Appendix A3). These differences could be due to the divergence in the expression patterns of these two genes, and in order for each gene to be expressed exclusively in a specific tissue, distinct sets of *cis*-regulatory DNA sequences might be required. Similarly, the heart-specific and the skeletal-specific enhancers in the mammalian *Tnnc1* gene share little sequence similarity (Parmacek et al., 1994), which presumably facilitates the differential tissue-specific expression of this gene in these two types of muscle.

To facilitate the analysis of *cis*-regulatory elements that regulate transcription of *tnnc1a*, a transgenic line was generated using BAC recombineering, in which the *tnnc1a* promoter specifically drove GFP expression in the heart muscle of zebrafish embryos. Further analysis is now needed to determine whether this pattern of GFP expression corresponds precisely to the expression pattern of the endogenous *tnnc1a* gene, but *in situ* hybridization using a *tnnc1a* specific probe revealed that, like the *tnnc1a:eGFP* reporter line, expression of the gene is confined exclusively to the heart. Likewise, previous analysis of the zebrafish *tnnc1a* expression pattern has revealed that by 1 dpf, transcripts are found throughout the heart muscle, and by 2 dpf, expression is mostly restricted to the ventricular chamber (Sogah et al., 2010). The *tnnc1a:eGFP* expression seen in the reporter line appears to be stronger in the ventricle at 3 dpf, yet some expression does also persists in the atrium, although this could be due to the perdurance of residual GFP protein. Taken together, it seems fair to conclude that the *tnnc1a:eGFP* transgenic reporter line faithfully recapitulates endogenous *tnnc1a* expression.

A previous report had localised the *Tnnc1* heart-specific enhancer in the immediate 5' region of the mouse *Tnnc1* promoter (Parmacek et al., 1992). To test whether the equivalent zebrafish genomic region contained a similar heart-specific enhancer, 2.2kb upstream of the *tnnc1a* transcription start site was cloned into the pDB739 vector, which contains itol2 arms for stable integration of the transgene into the zebrafish genome. In addition, a smaller DNA fragment encompassing 240bp immediately upstream of the transcription start site was cloned into the *β-globin* minimal promoter vector. Upon microinjection into embryos, neither construct drove GFP expression in any embryonic tissue, including the heart. There are several possible reasons for this. Firstly, the 2.2kb and 240bp fragments did not contain the 32bp DNA sequence located downstream of the transcription start site that had previously been reported to be part of the heart-specific enhancer in the initial report (Parmacek et al., 1992). Consequently, essential elements needed for the heart-specific transcription of the gene may have been omitted. Secondly, with regards to the 240bp DNA fragment combined with the *β-globin* minimal promoter vector not driving any GFP expression, the previous report that had identified the enhancer located in this region, discovered that it only functioned in concert with the endogenous *Tnnc1* promoter (Parmacek et al., 1992). This observation contrasts with the skeletal muscle enhancer for the mouse *Tnnc1* gene skeletal-specific enhancer located in intron 1, which was able to function with a minimal *SV40* promoter (Parmacek et al., 1994). Although this could also be the case in zebrafish, it is important to note that these experiments were carried out *in vitro* using cultured cells, and therefore care must be taken when analysing these results. *In vitro* studies using cultured cells can sometimes give a misleading impression. For example, enhancers previously identified in the mouse *beta myosin heavy chain* promoter *in vitro* and implicated as being dispensable for the muscle-specific expression of this gene, were found to be essential when analysed *in vivo* (Knotts et al., 1994).

The experiments described in this section could be extended, firstly by taking into account the various issues discussed above, and secondly by analysing further the conservation of *tnnc1a* enhancer function by cloning the human *Tnnc1* heart-specific enhancer into the *β-globin* minimal promoter vector and testing whether it can drive heart-specific expression of the GFP reporter cassette. Increased understanding of the regulation of *tnnc1a* may be clinically relevant as a mutation in the zebrafish *tnnc1a* leads to aberrant blood flow due to a loss in ventricular contraction (Sogah et al., 2010). Further to this mutations in human troponins are linked to certain cardiomyopathies (Gomes and Potter, 2004, Hoffmann et al., 2001, Mogensen et al., 2004) indicating the functional relevance of this gene with regards to heart function.

# Chapter 4

# Misexpression of Sox6 in the slow-twitch muscle using the UAS:GAL4 system

## 4.1 Introduction

*Sox6* expression appears in the somites at approximately the 10-somite stage (Wang et al 2011). The expression of *sox6* is excluded from the adaxial cells and is absent from the slow muscle throughout myogenesis, remaining restricted to the fast muscle domain only. Evidence has suggested that Sox6 protein is able to repress slow-twitch muscle genes, since heatshock-induced expression of sox6 inhibited *prox1a* expression in the adaxial cells (von Hofsten et al., 2008). In order to examine whether Sox6 is capable of inhibiting the expression of slow muscle genes, the UAS:GAL4 system was utilised in order to misexpress Sox6 protein in the slow-twitch muscle.

## 4.2 Results

### 4.2.1 *UAS:sox6-GFP:smyhc1:GAL4*

A *UAS:sox6-GFP* line (made by Wang Xingang) and a *smyhc1:GAL4* line (made by Wang Xingang) were crossed to generate a *UAS:sox6-GFP;smyhc1:GAL4* line, whereby *sox6* would be expressed in the slow-twitch muscle fibres under the control of the *smyhc1* promoter (Figure 4.1A-C). In order to determine whether misexpression of Sox6 in the slow-twitch fibre domain was sufficient to repress slow-specific gene expression, the expression of *smyhc1*, *tnnc1b* and *prox1a* was assessed in *UAS:sox6-GFP;smyhc1:GAL4* fish. For immunohistochemistry analysis embryos were collected from a *UAS:sox6-GFP;smyhc1:GAL4* incross and stained in the same tube with either a Tnnc1b, F59 or Prox1a antibody together with an anti-GFP antibody. Therefore control embryos, as assessed for by their lack of GFP expression consisted of *UAS:sox6-GFP* embryos, *smyhc1:GAL4* embryos, and embryos containing neither transgene. *UAS:sox6-GFP;smyhc1:GAL4* embryos were identified on the basis of their Sox6-GFP expression. For *in situ* hybridisation analysis, embryos were collected from a *UAS:sox6-GFP;smyhc1:GAL4 incross* and separated based on their Sox6-GFP expression; embryos expressing strong Sox6-GFP were collected into one tube and embryos expressing no Sox6-GFP were separated into another tube. Embryos with little or weak Sox6-GFP expression were discarded. Therefore the *in situ* hybridisation reactions were carried out in separate tubes, with the same number of embryos in each and were stained for exactly the same length of time.

*tnnc1b* expression is clearly downregulated in slow-muscle fibres expressing Sox6-GFP, which also exhibit reduced levels of Tnnc1b protein levels (Figure 4.2A-D). Similarly, by comparing slow-twitch fibres that are expressing Sox6-GFP with those that aren't, it is apparent that there is a strong decrease in Prox1a expression in the Sox-GFP positive cells (Figure 4.2I-K). The extent of the downregulation of *smyhc1* in response to the ectopic Sox6 expression seems somewhat variable and unclear. When looking at Smyhc1 protein levels using the F59 antibody, it seems that there may be a slight reduction in its expression with some effect on the fibre morphology, but the transcript levels appear barely affected if at all (Figure 4.2E-H). This data suggests that Sox6 is able to repress the expression of at least some slow muscle genes, and with a particularly strong repressive affect on particular slow-twitch genes, such as *tnnc1b*.

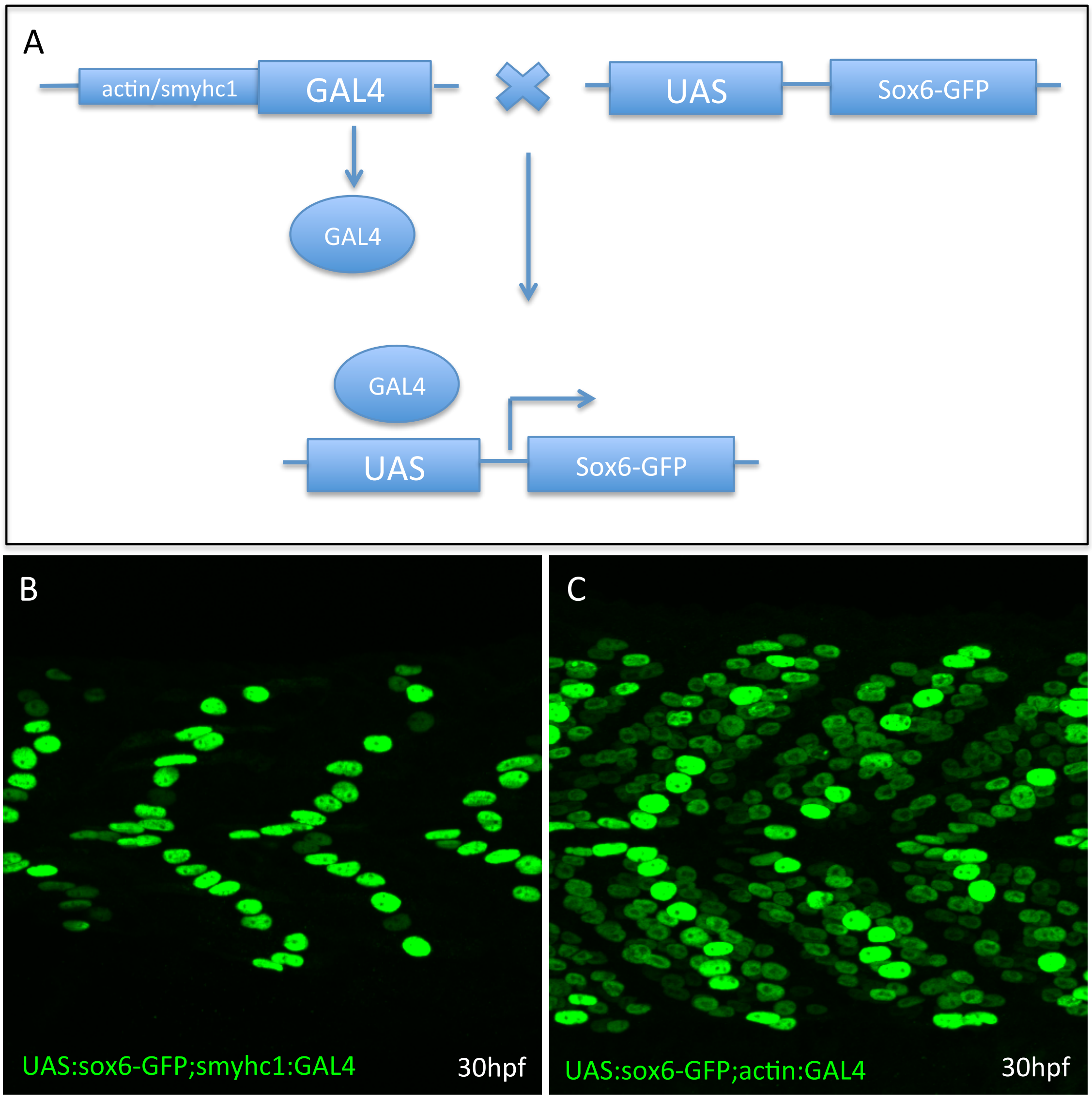


***Figure 4.1: The alpha-actin and smyhc1 promoters drive the expression of GAL4, which then binds to the UAS promoter to drive the expression of sox6-GFP** (A). sox6-GFP was expressed under the control of either the smyhc1 or the actin promoter in order to misexpress sox6 in the slow-twitch muscle (B and C).*

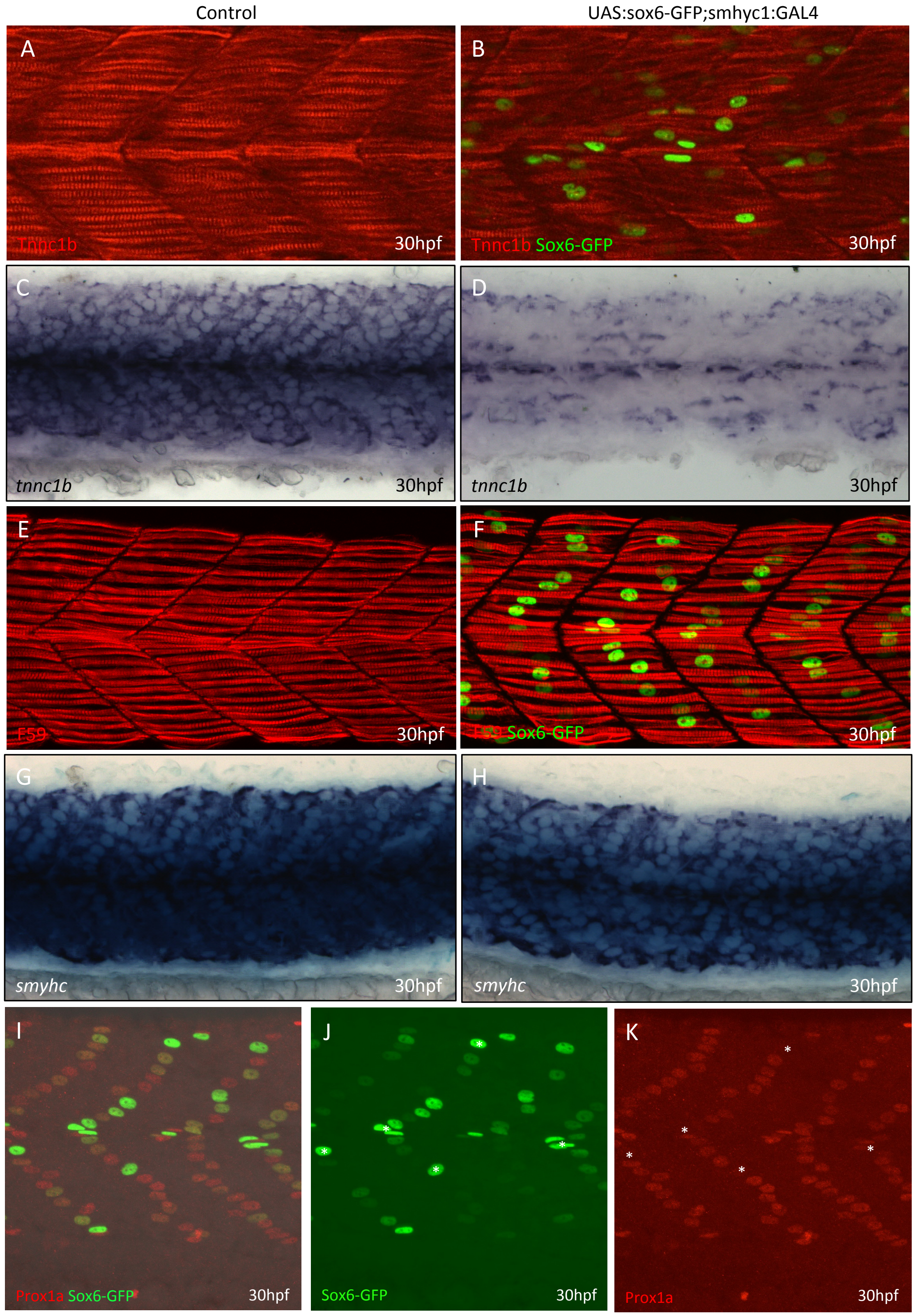


***Figure 4.2: Protein and transcript levels of tnnc1b are strongly repressed in response to ectopic expression of Sox6-GFP under the control of the smyhc1 promoter*** *(A-D). By contrast, ectopic expression of Sox6 in the slow muscle lineage causes only a subtle downregulation in both protein and transcript levels of smyhc1 (E-H). Prox1a protein expression is inhibited by UAS:sox6-GFP;smyhc1:GAL4. In nuclei expressing Sox6-GFP the expression of Prox1a is either completely lost or strongly downregulated (I-K). Each asterisk in J and K indicates a slow muscle cells nucleus that strongly expresses Sox6-GFP (green in panels I and J).*

### 4.2.2 *UAS:sox6-GFP;actin:GAL4*

Since Sox6-GFP expression was placed under the control of the *smyhc1* promoter, Sox6-GFP will only be expressed when *smyhc1* transcription begins, which would obscure the ability to assess whether Sox6 can inhibit *smyhc1* expression. To address this issue, the expression of the Sox6-GFP was placed under the control of regulatory elements from the alpha-actin gene, which is expressed prior to *smyhc1,* by crossing the *UAS:sox6-GFP* line to a line carrying the previously described *actin:GAL4* transgene (Scheer and Campos-Ortega, 1999). Embryos were collected from a *UAS:sox6-GFP;actin:GAL4* incross and sorted for *in situ* hybridisation and immunohistochemistry as described above for the *UAS:sox6-GFP;actin:GAL4* embryos. Fluorescent *in situ* hybridisation confirmed that *tnnc1b* expression is downregulated in response to Sox6, but also indicated that both *smyhc1* and *myh7b* levels are decreased in the slow muscle fibres expressing Sox6-GFP (Figure 4.3). The expression of other slow-specific muscle genes was then examined by *in situ* hybridisation, which revealed that *troponin t1* (*tnnt1), troponin i1a (tnni1a), myosin light chain 10 (mylz10), tropomyosin 2* (*tpm2)* and the *ryanodine receptor 1a* (*ryr1a)* were all downregulated in Sox6-GFP positive fish (Figure 4.4). Whole-mount immunohistochemistry using the F59 and S58 antibodies, which both recognise slow myosin heavy chain protein, further confirmed that the misexpressed Sox6 could down-regulate the protein expression of Smyhc1 (Figure 4.5A-B and D-E).

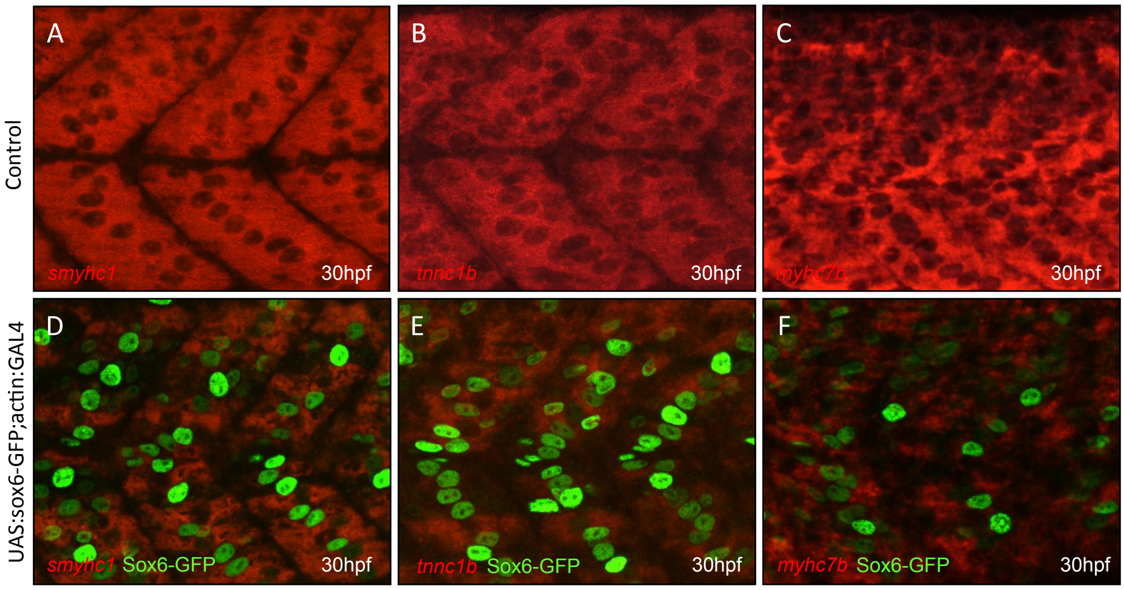


***Figure 4.3: smyhc1, tnnc1b and myh7b expression is downregulated in UAS:sox6-GFP;actin:GAL4 embryos.*** *The expression of three slow-specific genes, smyhc1, tnnc1b and myh7b in control embryos at 30 hpf (A-C). When Sox6 is misexpressed in the slow twitch muscle, the expression of all three genes is downregulated (D-F).*

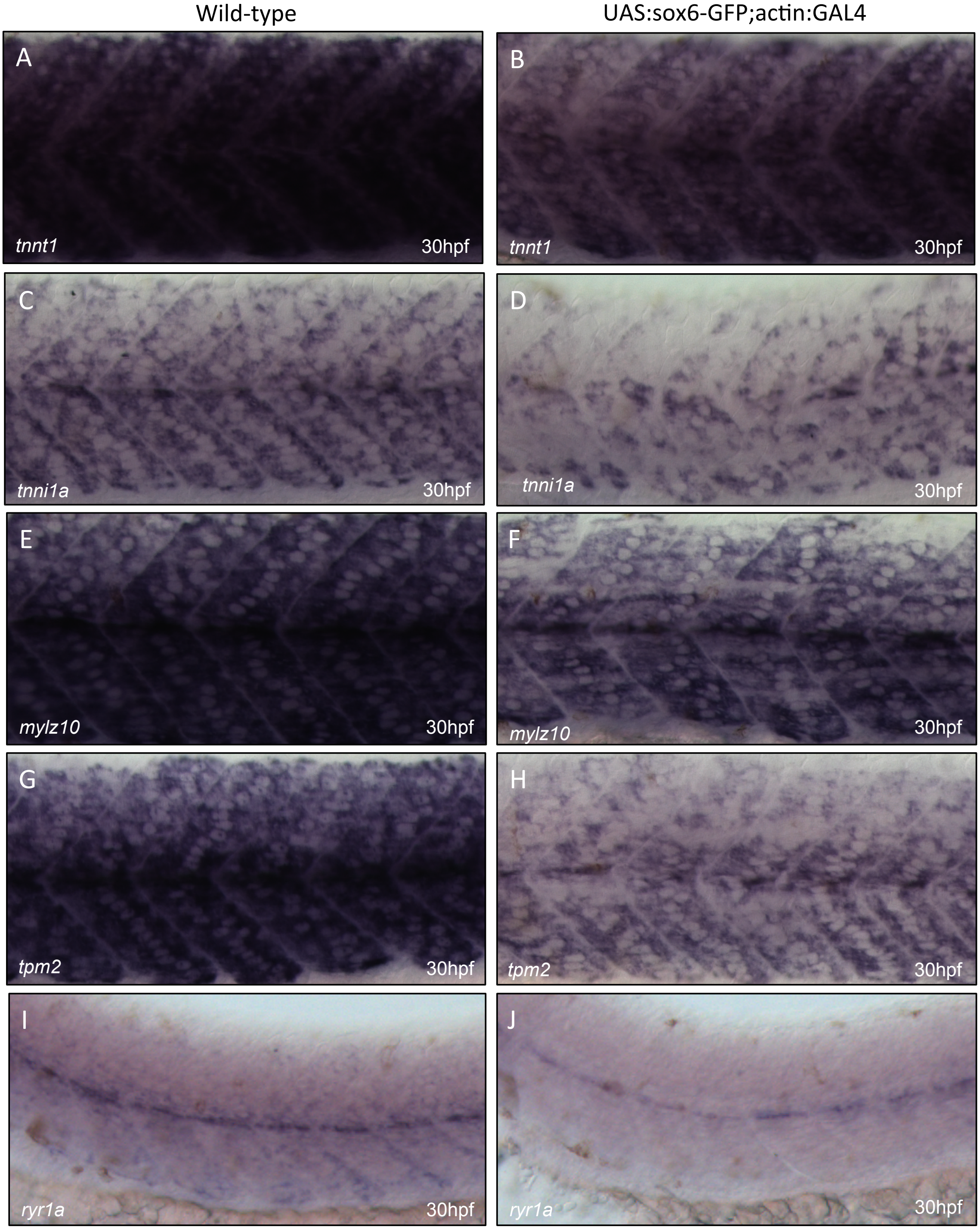


***Figure 4.4: The expression of slow-specific genes is downregulated in UAS:sox6-GFP;actin:GAL4 embryos.*** *The expression tnnt1 (A and B), tnni1a (C and D), mylz10 (E and F), tpm2 (G and H) and ryr1a (I and J) is specifically seen in slow-twitch fibres of control embryos. The expression of these genes is downregulated in slow muscle fibres ectopically expressing Sox6-GFP.*

Additionally to this the *tnnc1b:eGFP* line was crossed to the *UAS:sox6-GFP;actin:GAL4* line. The Sox6-GFP expression repressed the *tnnc1b:eGFP* expression in the slow muscle

(Figure 4.5C and F), confirming the ability of Sox6 to repress this gene and also providing another confirmation of the validity of the *tnnc1b:eGFP* line.

Taken together these results demonstrated that Sox6 is able to repress the expression of slow-twitch muscle genes in zebrafish embryonic trunk muscle.

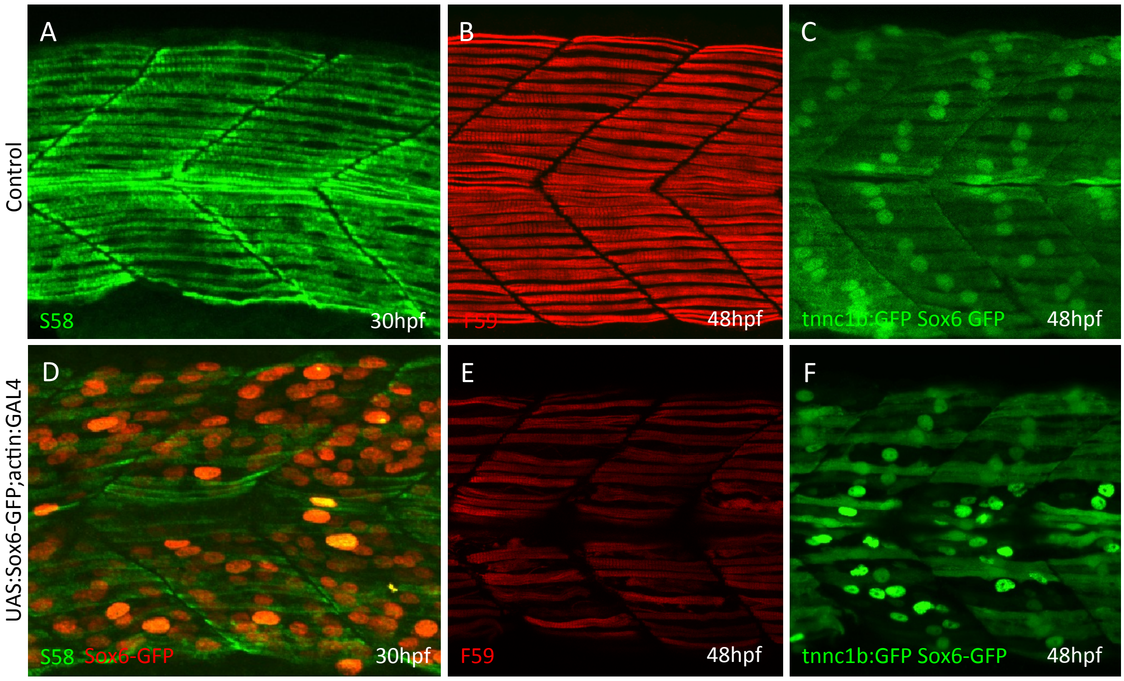


***Figure 4.5: The protein levels of slow-specific genes are decreased in UAS:sox6-GFP;actin:GAL4 embryos.*** *In control embryos S58 and F59 label Smyhc in slow-twitch fibres, nicely identifying the fibre striations (A and B). Ectopic expression of Sox6 under the control of the actin promoter causes a decrease in Smyhc protein expression, which also disrupts the slow-fibre striations (D and E). tnnc1b:eGFP expression is also downregulated when crossed to the UAS:sox6-GFP;actin:GAL4 line (C and F) (Bright green nuclei in panel F indicates localised Sox6-GFP).*

### 4.2.3 Analysis of head muscles

In order to determine whether Sox6 could also repress expression of slow-twitch muscle genes in the head, *UAS:sox6-GFP;actin:GAL4* positive and negative fish were fixed at 4 dpf. At this stage Sox6-GFP was expressed strongly in the head muscle under the control of the actin promoter. Similar to the above experiments, embryos with strong Sox6-GFP expression in the head muscles were sorted from Sox6-GFP negative embryos and in situ hybridisation was carried out in two separate tubes containing the same number of embryos. *In situ* hybridisation to detect *tnnc1b* transcripts indicated that there was no significant difference in the expression of the gene between the Sox6-GFP -positive and -negative embryos (Figure 4.6), suggesting that although Sox6 is able to inhibit the expression of this *tnnc1b* in the trunk muscle it does not have any repressive effect on *tnnc1b,* when ectopically expressed in the head muscle.

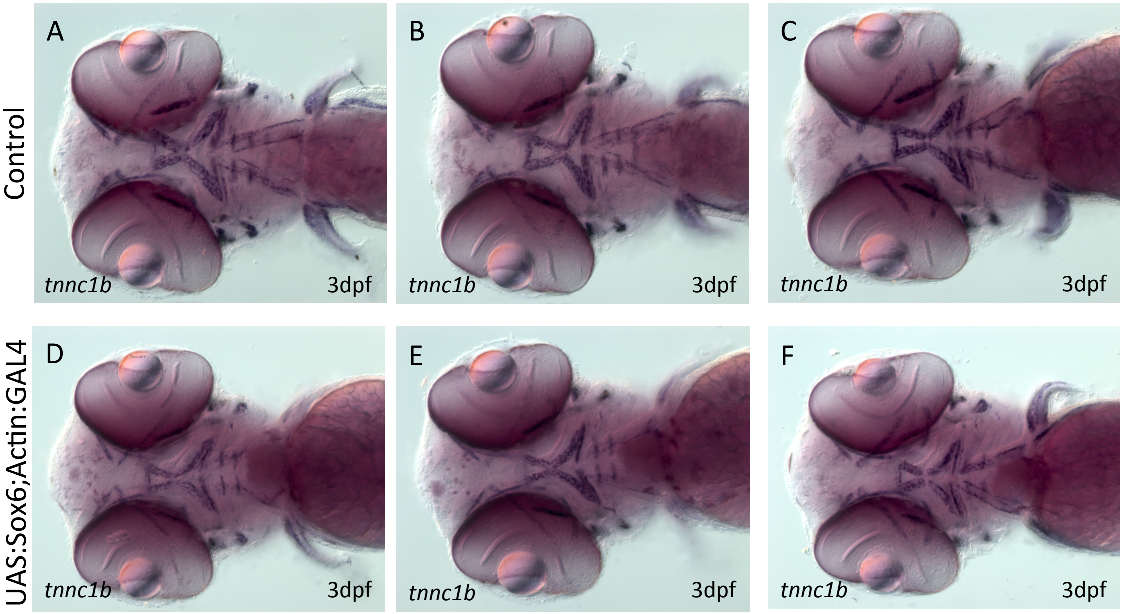


***Figure 4.6: Misexpression of sox6 in head muscles has no effect on the expression of tnnc1b in the head.*** *Panels A-C represent control tnnc1b expression in the head muscles of 3 dpf embryos. Panels D-F represent three examples of tnnc1b expression in embryos that also strongly expressed Sox6-GFP in the head muscles. There appears to be no difference in the expression of tnnc1b between control and UAS:sox6-GFP;actin:GAL4 fish.*

## 4.3 Discussion

These results provide further evidence that Sox6 represses the transcription of slow-twitch muscle genes and thus prevents the expression slow-twitch muscle specific protein isoforms that could be identified with specific antibodies. These findings corroborate previous studies in the zebrafish, including experiments in which forced expression of Sox6 in the adaxial cells of wild-type embryos inhibited Prox1a expression in these cells (von Hofsten et al., 2008). Our results also confirm that Sox6 performs an evolutionarily conserved function as a repressor of slow-twitch fibre differentiation (Hagiwara et al., 2005; Hagiwara et al., 2007), and are consistent with observations in the mouse which demonstrate that Sox6 directly repressed an array of slow-twitch-specific muscle genes in mice (An et al., 2011).

Interestingly, the slow-specific calcium channel *ryanodine receptor 1a* *(ryr1a)* was inhibited by misexpression of sox6, implying that it is not only the structural protein genes that can be inhibited by Sox6. Ryanodine receptors release $Ca^{2+}$ from the sarcoplasmic reticulum into the cytosol, in response to depolarisation of the plasma membrane. The $Ca^{2+}$ can then bind to the troponin complex enabling the sliding of the actin and myosin filaments to produce muscle contraction. This is the first evidence that Sox6 regulates the expression of *ryr1a*.

Recent studies have shown that the microRNA, miR-499, encoded by *Myh7b,* regulates Sox6 expression in both fish and mouse myofibres (McCarthy et al., 2009, van Rooij et al., 2009, Wang et al., 2011). Moreover, *myh7b/miR-499* expression can also be repressed by Sox6 (An et al., 2011, Hagiwara et al., 2007). This reciprocal regulatory relationship had not been confirmed in zebrafish, although the loss of *myh7b* from the adaxial cells in the *prdm1a* mutants was predicted to be due to the inhibitory effect of the ectopic Sox6 in the adaxial cells (Wang et al., 2011). The results presented in this chapter indeed confirm that *myh7b* (and likely *miR-499)* can be repressed by Sox6 *in vivo*, providing evidence for the existence of a regulatory loop of negative feedback between *sox6* and *myh7b/miR-499*.

In *sox6* mutant mice, expression of fast-specific muscle genes is downregulated in response to the increased in expression of slow-twitch muscle genes (Hagiwara et al., 2005, Quiat et al., 2011). Therefore it is possible that an increase in fast-twitch-specific gene expression could compensate for the loss of slow-twitch muscle gene expression in the *UAS:sox6-GFP;actin:GAL4* line. Had time permitted, it would have been interesting to look at the expression of fast-twitch-specific genes in this line to see if indeed this is the case. However, in light of existing knowledge about the transcriptional regulation of fast-twitch muscle genes in development, the regulatory mechanisms are likely to be complicated. For example, in zebrafish, activation of fast-twitch muscle differentiation relies on the expression of genes such as *fgf8* and *six1a,* which are not found in the slow-twitch muscle (Bessarab et al., 2008, Reifers et al., 1998).

Another area of research that would be worthy of further investigation is the topic of whether Sox6 represses expression of slow-specific genes that are expressed in the secondary slow fibres, such as *smyhc2* and *smyhc3* (Elworthy et al., 2008). Sox6 is one of only a few transcription factors that are implicated in both developmental fibre-type specification and adult muscle plasticity (Hagiwara et al., 2005, Hagiwara et al., 2007, McCarthy et al., 2009, Quiat et al., 2011, von Hofsten et al., 2008). For this reason, it seems reasonable to hypothesise that Sox6 is associated with fibre-type switching in juvenile and adult zebrafish.

*Sox6 in situ* hybridisation (Chapter 5, figure 5.1) suggests that *sox6* is not normally expressed in the head muscle of zebrafish embryos, however a more detailed analysis of *sox6* expression in the head is needed. Nevertheless, the forced expression of Sox6 in all head muscles using the *UAS:sox6-GFP;actin:GAL4* line resulted in no change in the expression pattern of *tnnc1b*. This result raises the question as to why Sox6 is able to repress the expression of *tnnc1b* in trunk muscle but has no effect on the same gene in the head muscles. During development, the head mesoderm and the somites express both overlapping and distinct combinations of transcription factor genes, and interestingly, reciprocal ectopic grafting experiments indicate that the head mesoderm and the somites can only respond to their local environmental cues (Mootoosamy and Dietrich, 2002). Sox proteins are known to function in concert with other transcription factors, so it is possible that the other proteins that are needed for Sox6 to efficiently repress *tnnc1b* in the slow-twitch skeletal muscle are not present in the zebrafish embryonic head mesoderm, or it may be that *tnnc1b* is regulated by an entirely different mechanism in the head.

Taken together, the results presented in this chapter demonstrate the ability of Sox6 to repress slow-twitch specific muscle genes in the trunk, however there is not a complete repression of the slow-twitch programme as slow-twitch fibres are still mononucleated, as visualised by the expression of Sox6-GFP in Figure 4.2 panels I-K. Although these experiments reveal that Sox6 protein inhibits the transcription of genes that are associated with slow-twitch fibre differentiation, it still remains unclear as to whether all of the observed effects of Sox6 misexpression are physiologically relevant. Since Sox6 is normally only expressed in the fast-twitch myoblasts, it seems reasonable to infer that the slow-twitch muscle differentiation programme is being repressed by Sox6 in the fast-twitch myoblasts. In order to pursue this question, a zebrafish *sox6*$^{-/-}$ mutant would be invaluable for determining whether loss of *sox6* in the fast muscle results in a derepression of slow-twitch-specific genes in the fast domain.

# Chapter 5

# Functional analysis of *sox6* by zinc-finger nuclease mediated targeted mutagenesis

## 5.1 Introduction

Expression of *sox6* can first be detected in the somites at the 10-somite stage (Wang et al., 2011), where it is restricted to the fast-twitch domain and excluded from adaxial cells. In addition, *sox6* is expressed in other locations during development including the otic vesicles and the brain (Figure 5.1). In order to investigate the function of Sox6 in these tissues more thoroughly, I sought to create mutant alleles of *sox6* using zinc-finger nuclease (ZFN) technology (Meng et al., 2008).

## 5.2 Results

### 5.2.1 Generation of *sox6* zinc-finger nuclease mutant

The ZFN was designed so that both *sox6* transcripts would be mutated upstream of the HMG box at the position of approximately amino acid 249 and amino acid 230 (Figure 2.4A and B). Two alleles, each of which introduces a premature stop codon into the open reading frame that are predicted to encode truncated forms of the protein, were isolated (Figure 2.4C). Homozygous mutants were found to lack full-length Sox6 protein in both fast-twitch fibres and in the forebrain, as assessed by whole-mount immunohistochemistry using a Sox6-specific antibody (Figure 5.2; see Chapter 6 for further details about the antibody). 40 embryos were collected from a *sox6*$^{-/+}$ heterozygous incross and incubated with the Sox6-specific antibody. Of those 40 embryos 32 embryos had specific staining in the nuclei of the fast-twitch fibres and the optic tectum, whereas 8 embryos had no specific staining. These embryos were then genotyped to confirm that those embryos that showed no specific staining were *sox6*$^{-/-}$ mutant embryos. Indeed, all 8 embryos with no staining were *sox6*$^{-/-}$ mutants and no homozygous mutant *sox6* embryos where identified which displayed specific staining.

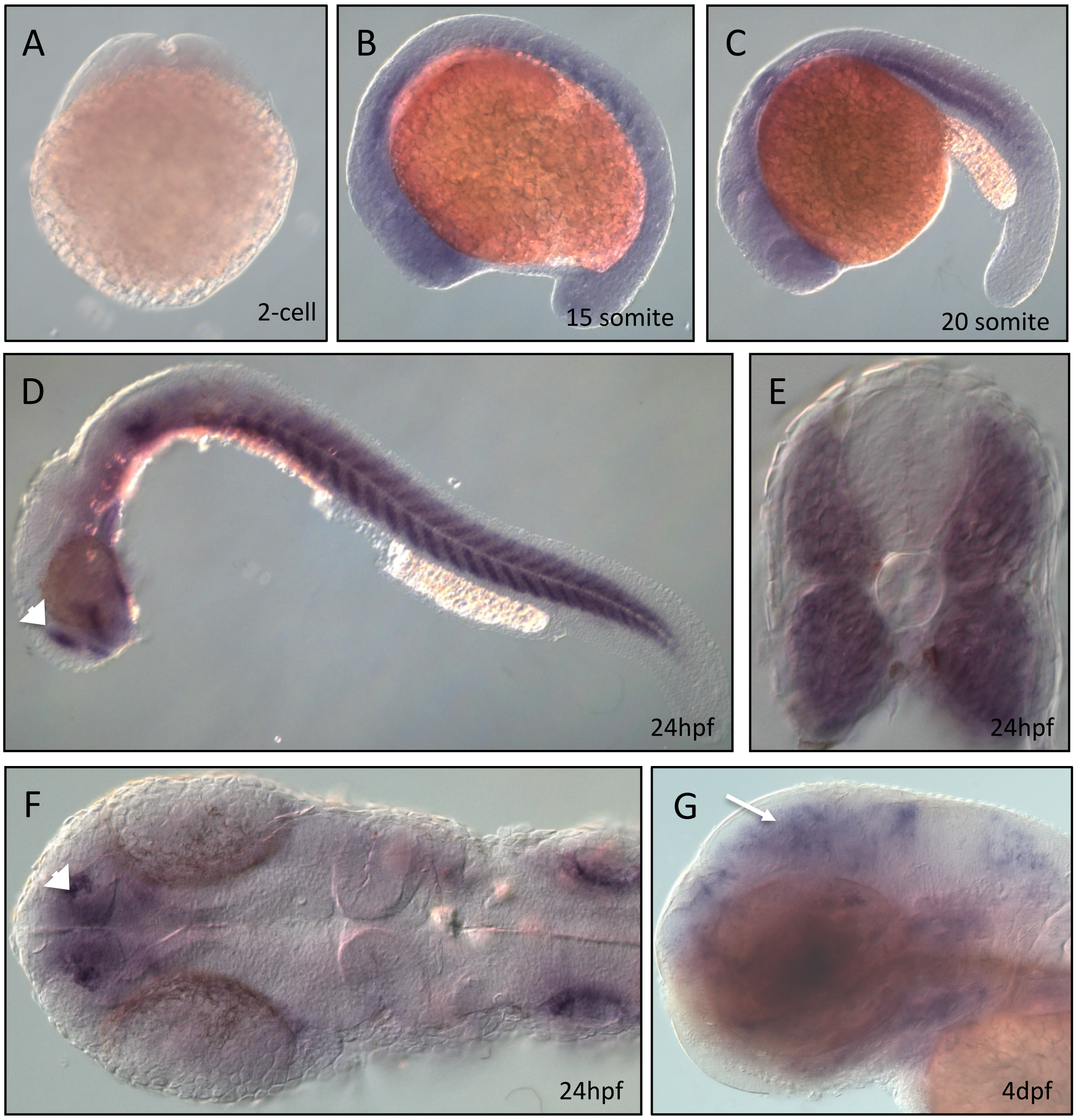


***Figure 5.1: In situ hybridisation to detect sox6 transcripts, reveals expression of this gene throughout zebrafish embryonic development.*** *Sox6 is not maternally expressed (A). The expression of sox6 in the somites is restricted to the fast domain throughout myogenesis (B-E). Sox6 is also expressed in the otic vesicles and brain regions including the telencephalon (arrowhead) and optic tectum (arrow) (D, F, G).*

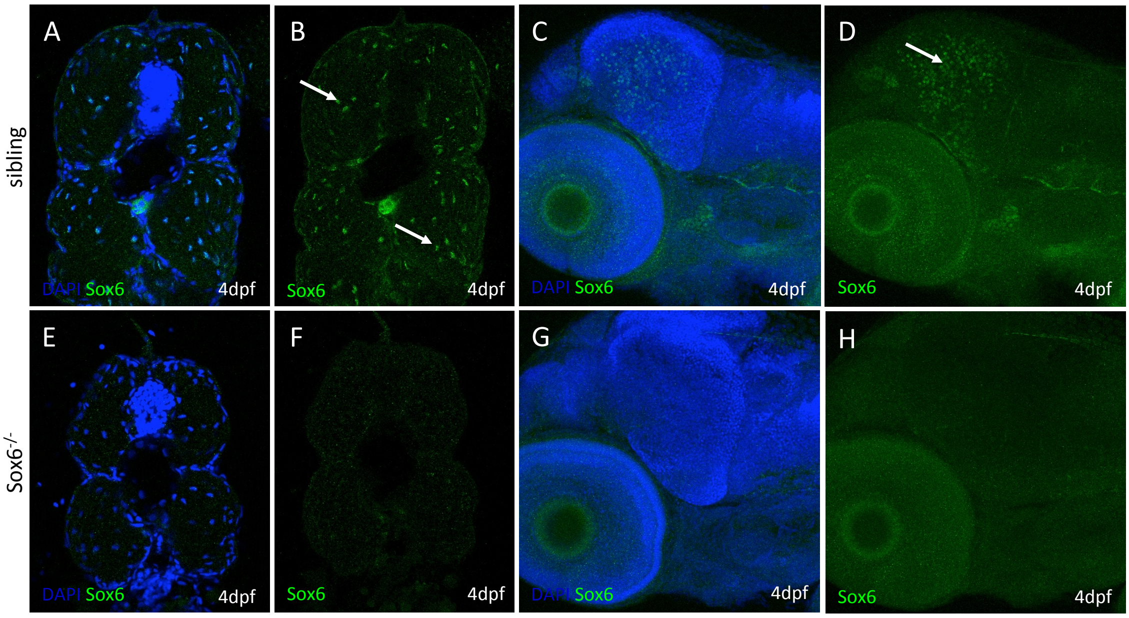


***Figure 5.2: No Sox6 protein is generated in the sox6<sup>-/-</sup> mutant.*** *Immunohistochemistry analysis of Sox6 protein expression using a Sox6-specific antibody reveals that Sox6 protein is detectable in the nuclei of fast muscle fibres (arrows) and the optic tectum (arrows) of wild-type embryos (A-D). In sox6 zinc-finger mutants (E2 allele, sox6<sup>-/-</sup>), the antibody detects no Sox6 protein in either the muscle or the optic tectum indicating the successful generation of a null mutant (E-H). This E2 allele is used in all subsequent experiments.*

### 5.2.2 *tnnc1b* is ectopically expressed in *sox6* homozygous mutants

As discussed in Chapter 4, ectopic expression of Sox6 protein in the slow-twitch muscle of wild-type embryos inhibits the expression of slow-specific genes. It was therefore predicted that loss of *sox6* might cause the derepression of these slow-specific genes in the fast myotome. The *sox6* homozygous mutant embryos show no morphological phenotype; however, *in situ* hybridisation revealed robust ectopic expression of the slow-specific *tnnc1b* gene in the fast-twitch muscle in about a quarter (12 out of 60 embryos, confirmed by genotyping) of embryos at 30hpf generated by a *sox6* heterozygous incross (Figure 5.3B and F). *sox6*$^{-/+}$ heterozygous mutants were crossed to the UAS:sox6-GFP;actin:GAL4 line, enabling the generation of *sox6*$^{-/-}$ homozygous mutants that expressed Sox6-GFP throughout the myotome. The rescue of Sox6-GFP throughout the fast-twitch myotome of *sox6*$^{-/-}$ mutants was able to inhibit the ectopic expression of *tnnc1b* in the fast-muscle (Figure 5.4 A and B). The ectopic *tnnc1b* expression was detectable from approximately 22 hpf onwards and thus provides a means of identifying *sox6* homozygous mutants (Figure 5.3A and E). To take advantage of this in subsequent experiments, I crossed the *tnnc1b:eGFP* transgene described in Chapter 3 into the *sox6* mutant background. Incrossing of the resultant *sox6*$^{-/+}$*;tnnc1b:GFP* fish produced progeny, one quarter (53 out of 230 embryos, confirmed by genotyping) of which exhibited ectopic GFP expression in the fast-twitch muscle (Figure 5.3D, H, I and L). The ectopic GFP expression seen in the *sox6*$^{-/-}$*;tnnc1b:GFP* line allowed sorting of *sox6* homozygous mutants in subsequent experiments. Embryos could be sorted by this method from approximately 24hpf (Figure 5.3 D and H). Whilst still in there chorions the embryos could be manipulated with a needle so that the dorsal side of the embryos was visible under the dissection microscope. Sibling embryos displayed the normal pattern of *tnnc1b:GFP* with a one-fibre-thick layer of slow muscle fibres expressing GFP along the out most edges of the myotome. *Sox6*$^{-/-}$ homozygous mutants however, could be easily identified as the GFP expression had expanded medially and a think line of GFP expression was visible on either side of the notochord. This method proved reliable as embryos that were identified as *sox6*$^{-/-}$ mutants based on their expression of *tnnc1b:GFP* were positively identified as *sox6*$^{-/-}$ mutants by genotyping, with a 100% correlation. Likewise, embryos that displayed no expansion in their *tnnc1b:GFP* expression domain were always identified as either wild-type or *sox6*$^{-/+}$ heterozygous embryos when genotyped.

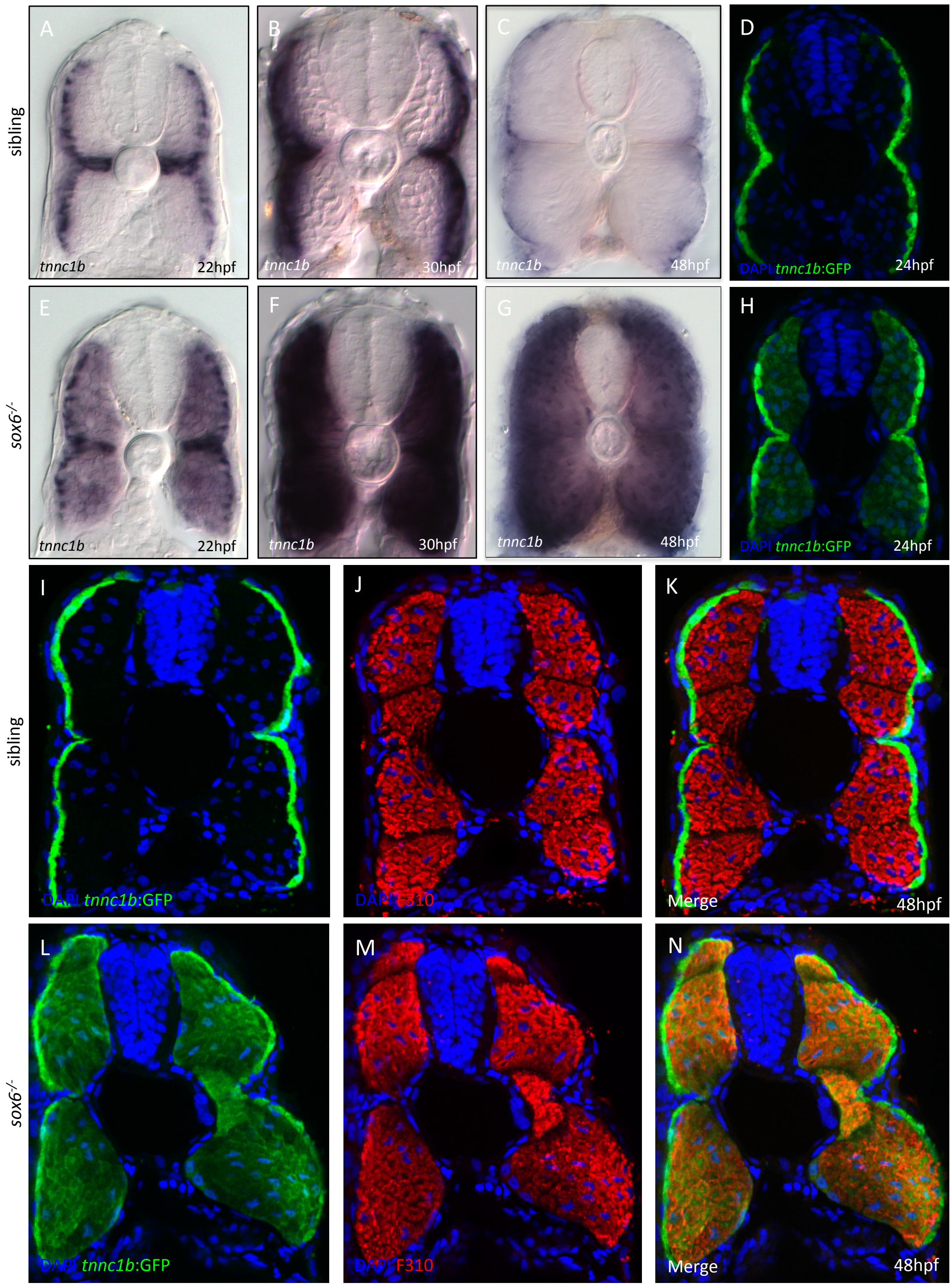


***Figure 5.3: tnnc1b is ectopically expressed in fast-twitch fibres of homozygous sox6 mutants****. Tnnc1b expression is usually restricted to the slow-twitch fibres (A). In situ hybridisation analysis revealed that mutant fast-twitch fibres begin to ectopically express tnnc1b at 22 hpf (E). By 30hpf there is robust expression of tnnc1b throughout the myotome of sox6*$^{-/-}$ *mutants (F) as compared to siblings (B). tnnc1b:eGFP;sox6*$^{-/-}$ *mutant embryos faithfully recapitulate the ectopic expression of tnnc1b observed in sox6*$^{-/-}$ *mutant embryos (C, D, G, H, I and L). tnnc1b:GFP;sox6 embryos can be sorted for genotyping from 24hpf onwards where ectopic GFP expression can be clearly visualised in the fast-twitch domain (D and H). Expression of fast-myosin light chain appears unaffected in tnnc1b:eGFP;sox6*$^{-/-}$ *mutant embryos (J and M) with both fast-myosin light chain and tnnc1b:eGFP co-expressed in the same fibres in sox6*$^{-/-}$ *mutants compared to their segregated expression pattern in wild-type (K and N).*

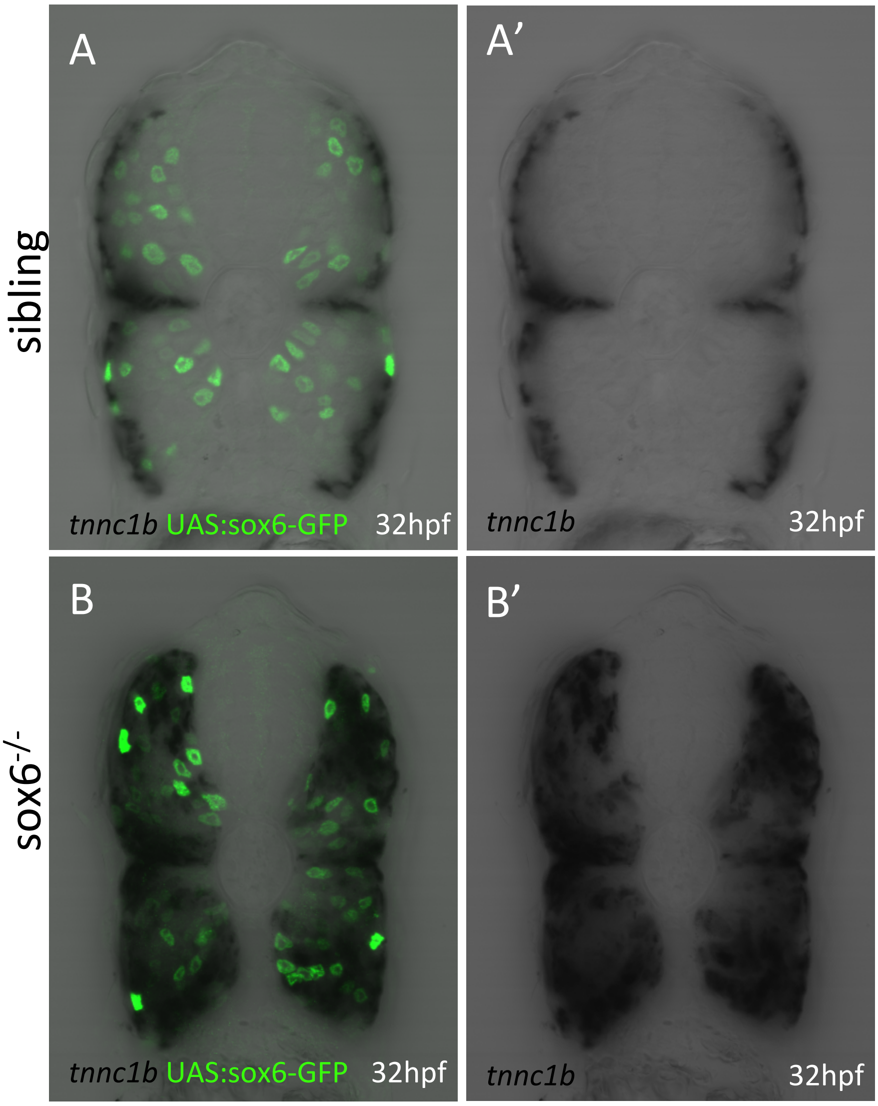


***Figure 5.4 Expression of Sox6-GFP can inhibit the ectopic expression of tnnc1b in sox6$^{-/-}$ mutants.*** *The actin:GAL4 line was crossed to the UAS:sox6-GFP line in order to drive expression of Sox6-GFP throughout the myotome. In sibling embryos the expression of Sox6-GFP could inhibit the expression of tnnc1b in the slow-twitch muscle (A and A’). In sox6$^{-/-}$ mutants, where tnnc1b is expressed throughout the myotome, the expression of Sox6-GFP can inhibit the expression of tnnc1b in both slow- and fast-twitch muscle fibres (B and B’).*

### 5.2.3 Expression of fast-twitch muscle genes are unaffected in sox6$^{-/-}$ mutant fish

Utilising the *tnnc1bGFP;sox6* mutant line to identify *sox6$^{-/-}$* homozygous mutant embryos, the effect of loss of *sox6* function on the expression of fast-specific genes was next examined. The loss of *sox6* function could result in a decrease in fast-specific genes, similar to what is seen in mice (Hagiwara et al., 2005; Hagiwara et al., 2007). Interestingly, it appeared that expression of fast-twitch genes was unaffected in the *sox6$^{-/-}$* mutants. *In situ* hybridizations analysis to detect expression of *fast myosin heavy chain* (*fmyhcx*) and *fast myosin light chain2 (mylz2)* revealed no difference between the *sox6$^{-/-}$* mutants and their siblings (Figure 5.5). These observations were consistent with the results using the F310 antibody to detect fast myosin light chain protein levels, which were also unaffected in the *sox6*$^{-/-}$ mutants (Figure 5.3J and M). These experiments suggest that the expression of fast muscle-specific genes is independent of *sox6* function.

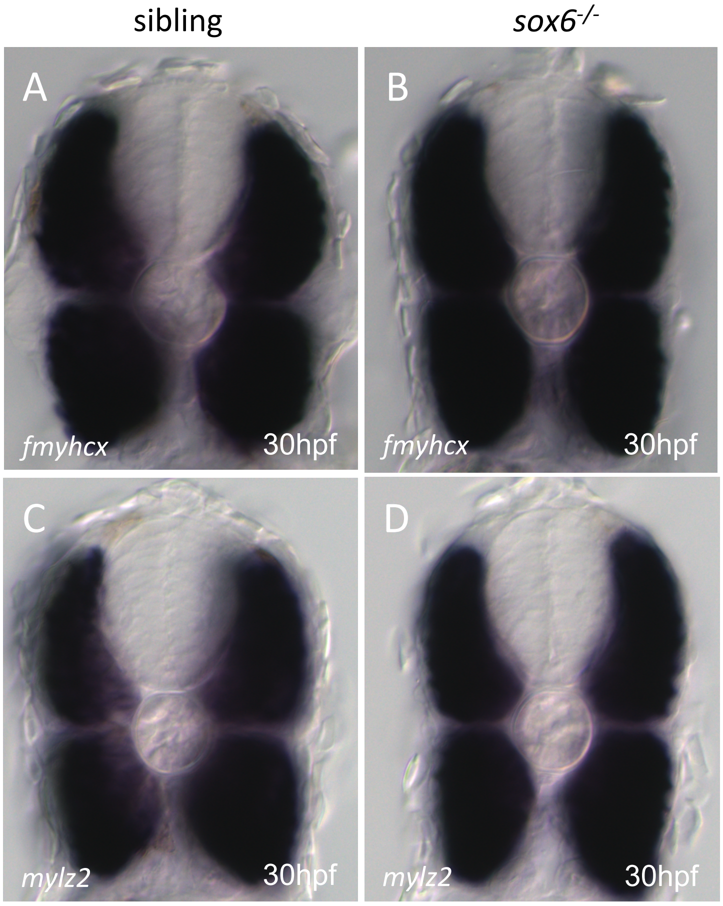


***Figure 5.5: Expression of the fast-specific muscle genes fmyhcx (A and B) and mylz2 (C and D) is unaffected in sox6*$^{-/-}$ *mutant embryos compared to their siblings.***

### 5.2.4 Only a subset of slow-twitch muscle genes are ectopically expressed in *sox6*$^{-/-}$ mutant embryos at 30 hpf

The expression patterns of other slow-twitch muscle genes were next examined using the *tnnc1b:eGFP;sox6* mutant line. Surprisingly, while *troponin t1 (tnnt1*) and the slow-specific *ryanodine receptor 1a* (*ryr1a)* were ectopically expressed throughout the fast-twitch muscle of 30hpf *sox6*$^{-/-}$ mutant embryos, the expression patterns of *troponin i1a* (*tnni1a), myosin light chain 10 (mylz10)* and *tropomyosin 2 (tpm2)* were unaffected at 30 hpf (Figure 5.6).

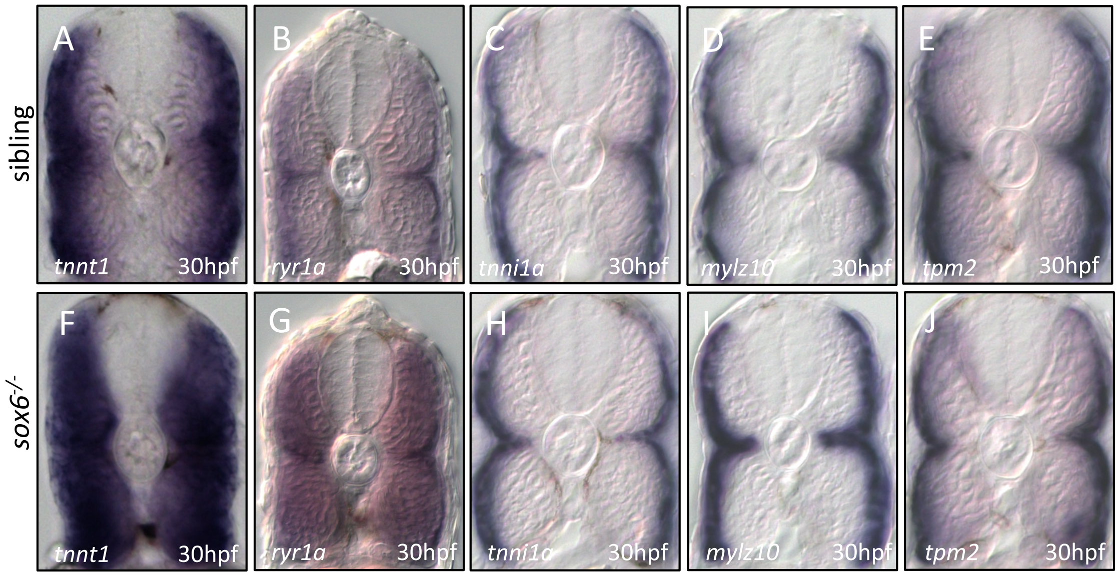


***Figure 5.6: Both tnnt1 (A and F) and ryr1a (B and G) are ectopically expressed in the fast domain of sox6 homozygous mutant embryos at 30 hpf.*** *Conversely expression of tnni1a (C and H), mylz10 (D and I) and tpm2 (E and J) is unaffected at 30 hpf in these mutants.*

The expression of *prox1a* was also analysed by whole-mount *in situ* hybridisation, but these transcripts were difficult to detect at 30 hpf. However, at 22 hpf *prox1a* transcripts were detectable in the slow muscle domain of wild-type and *sox6*$^{-/-}$ mutants, but there was no expression of this gene in the fast domain of *sox6*$^{-/-}$ mutants (Figure 5.7A and E). These observations were further supported by immunostaining embryos using the Prox1a antibody, which verified that no Prox1a protein is expressed in the fast muscle of *sox6* homozygous mutant embryos (Figure 5.7B and F).

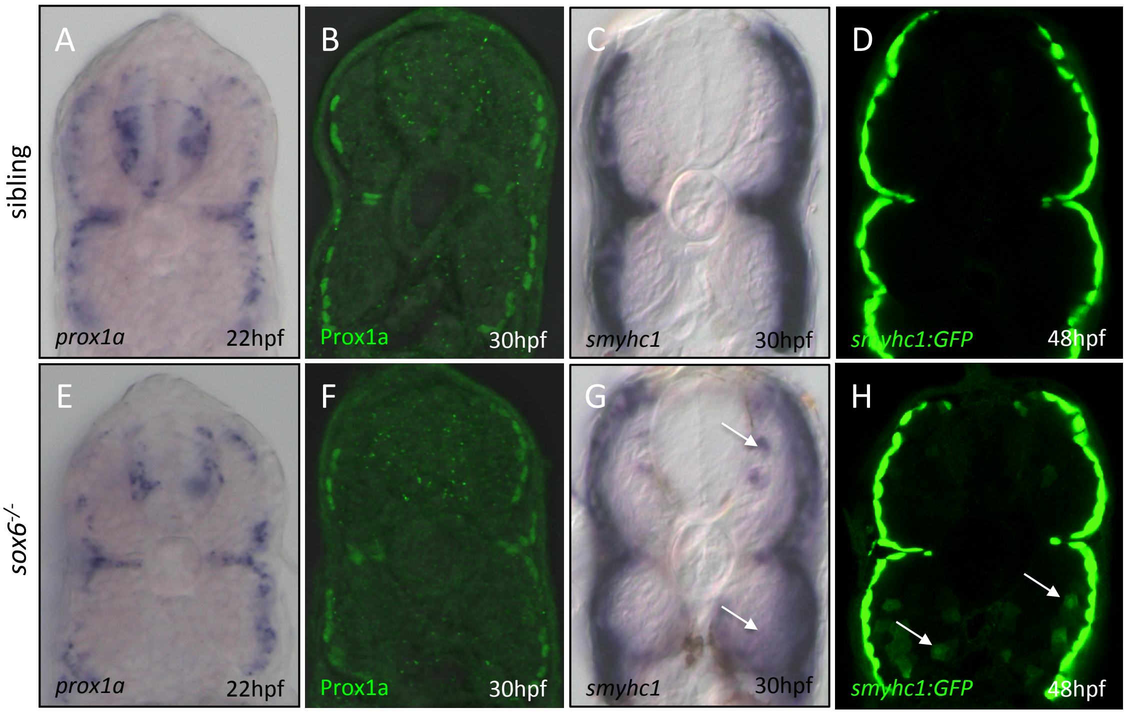


***Figure 5.7: Expression of prox1a and smyhc1 in sox6$^{-/-}$ mutants.*** *prox1a expression is unaffected sox6 mutant embryos. In situ hybridisation and immunohistochemistry revealed no ectopic fibres in the fast domain at 22 hpf (A and E) and 30 hpf (B and F). smyhc1 is expressed in a few fast fibres at 30 hpf as shown by in situ hybridisation (C and G). By 48 hpf the smyhc1:GFP reporter line reveals more fast fibres expressing GFP in the mutant, yet it is still not throughout the whole domain (D and H) Arrows represent fast-twitch fibres ectopically expressing smyhc1.*

Comparative analysis with a probe to detect expression of the slow-specific *smyhc1* gene revealed that *smyhc1* exhibited a small amount of ectopic expression in a few fast fibres per fish at 30 hpf (Figure 5.7C and G). This result was confirmed by crossing the *sox6* mutant to the *smyhc1:GFP* line to generate *sox6* mutants that carried the *smyhc1:GFP* reporter. One quarter of the embryos from a heterozygous incross displayed GFP expression in a small number (2-10) of fast fibres, confirming the results from the *in situ* hybridisation analysis (Figure 5.7D and H). No sibling embryos ever displayed any *smyhc1* positive fast-twitch fibres either by *in situ* hybridisation or reporter gene expression. Previous studies in the lab had identified a series of potential Sox6 binding sites in the *smyhc1* promoter and these were mutated in a *smyhc1:GFP* reporter transgene. Transgenic lines were then generated with this mutated reporter construct. The resulting founder fish displayed a small number of GFP- expressing fast fibres (Stone Elworthy and Jonas von Hofsten, unpublished; Figure 5.8), an effect similar to the ectopic expression of *smyhc1* observed in *sox6*$^{-/-}$ mutant embryos. This result confirms that the ectopic expression of *smyhc1* observed in the fast fibres of *sox6* homozygous mutant embryos is due to the absence of Sox6-mediated direct repression of *smyhc1.*

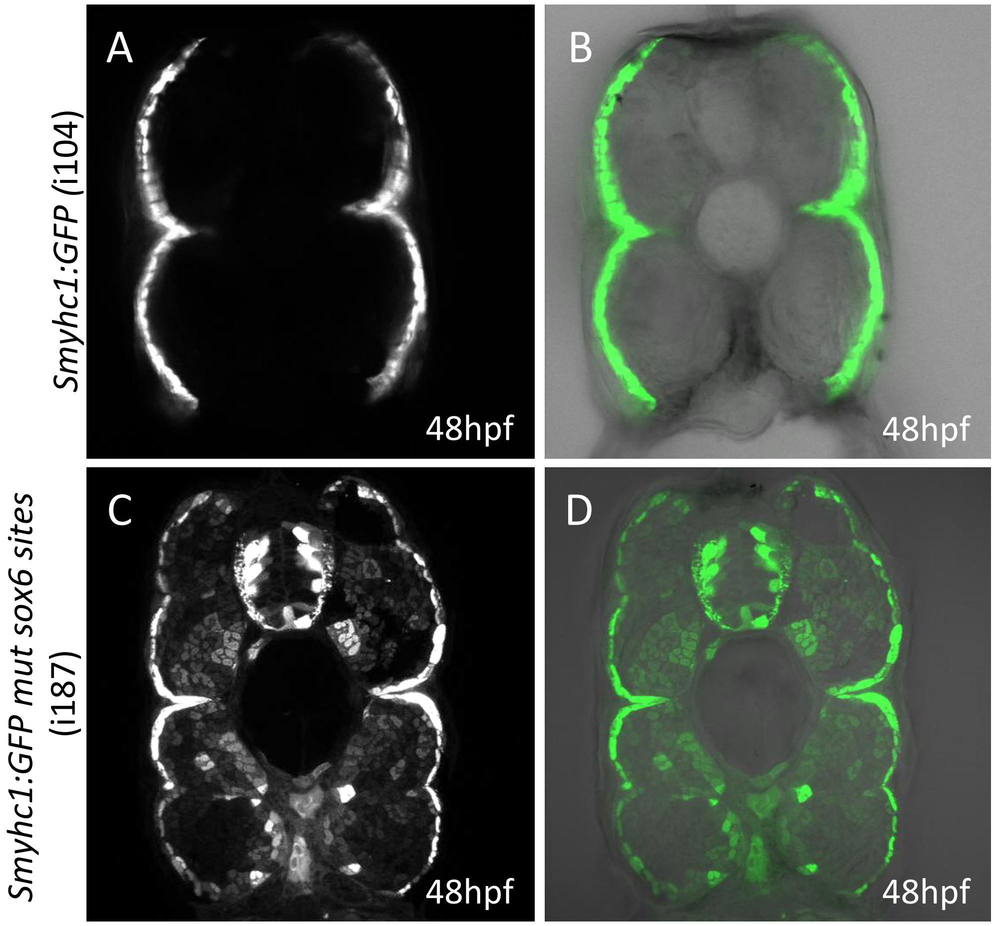


***Figure 5.8: Sox6 restricts expression of a smyhc1:GFP transgene to the slow-muscle fibres of wild-type embryos*** *(A and B). When potential Sox6 binding sites are mutated in the promoter of a smyhc1:GFP transgene, ectopic GFP expression can be readily detected in some fast-twitch muscle fibres as well as in the spinal cord (C and D) (Dr Stone Elworthy; unpublished).*

### 5.2.5 The loss of *sox6* can rescue the expression of slow-specific genes in *prdm1a*$^{-/-}$ mutants

As previously described in Chapter 1 the *prdm1a* mutant alleles *ubo* and *nrd,* show a severe decrease in or complete absence of the expression of slow muscle-specific genes, reflecting the transformation of the slow-twitch lineage to fast-twitch fibres. A previous study from this laboratory had demonstrated that injection of a *sox6*-specific morpholino into *ubo* mutant embryos restored the expression of *smyhc1* and partially rescued the expression of *prox1a* in the slow muscle domain (von Hofsten et al., 2008). In order to confirm and extend these observations, a *prdm1a*-specifc morpholino was injected into *sox6* mutants. The absence of *sox6* function in *prdm1a* morphant embryos was indeed sufficient to rescue expression of *smyhc1* in the slow domain and promote ectopic expression of *smyhc1* in a few fast fibres as well (Figure 5.9A and C). The existence of gaps in the slow domain of *smyhc1* gene expression raised the possibility that the individual cells labelled in the fast domain may actually be slow-twitch myoblasts that have failed to migrate to the surface of the myotome. By contrast to *smyhc1, tnnc1b* was robustly expressed throughout the myotome of *prdm1aMO;sox6*$^{-/-}$ mutant embryos, indicating that expression of this gene was both completely rescued in the slow-twitch domain and ectopically expressed in all of the fast-twitch fibres (Figure 5.9B and D). This is consistent with the notion that Prdm1a activates slow-twitch gene expression indirectly through repression of *sox6* transcription.

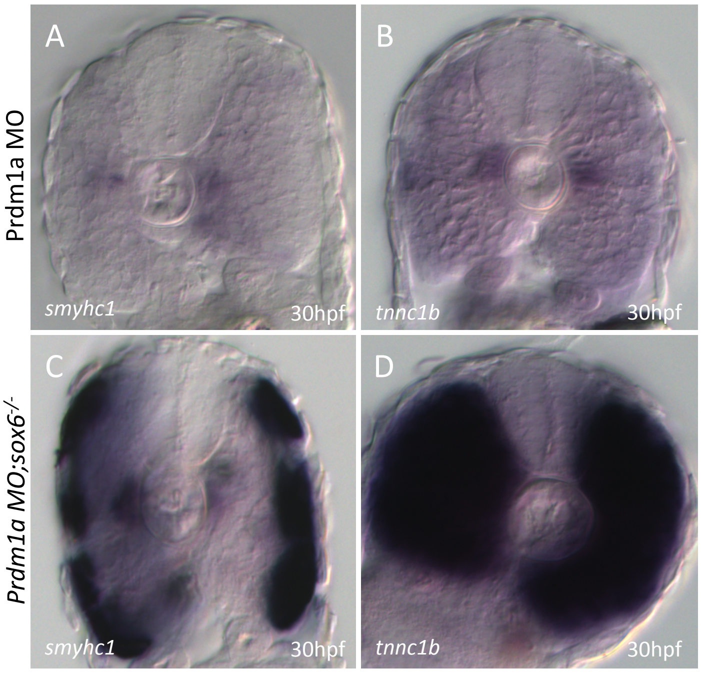


***Figure 5.9: Expression of smyhc1 and tnnc1b is rescued in Prdm1a morphant;sox6*$^{-/-}$ *mutant embryos.*** *Prdm1a morphant embryos exhibit a loss of smyhc1 and tnnc1b (A and B). The loss of sox6 in prdm1a morphants rescues the expression of smyhc1 and tnnc1b in the slow-twitch domain and also causes robust expression of tnncb1 throughout the myotome (C and D).*

As mopholino knockdown of a gene often does not result in the complete loss of expression of the targeted gene, the *nrd* mutant, which harbours a null mutation in *prdm1a,* was crossed to the *sox6* mutant to generate fish doubly mutant for both genes. The *nrd* mutant allele was used instead of the *ubo* mutant allele to ensure complete loss of function of the *prdm1a* gene (Hernandez-Lagunas et al., 2005). In *nrd*$^{-/-}$ homozygous mutant embryos, a complete loss of slow-twitch muscle gene expression was observed. In the *sox6*$^{-/-}$*;nrd*$^{-/-}$ double homozygous mutants, slow-twitch muscle-specific expression was partially rescued for each of the genes examined, and *tnnc1b* was persistently expressed throughout the myotome (Figure 5.10). Notably the expression of *tpm2* in *sox6*$^{-/-}$*;nrd* mutants is observed consistently in the outer fibres of the myotome. This suggests that the rescue of *tpm2* expression occurs in a different population of cells than the rescue of *smyhc1* and *mylz10,* although lineage tracing and double *in situ* hybridisation analysis is required to confirm this. These results indicate that a loss of *sox6* in *prdm1a* mutants can partially rescue the loss of slow-twitch muscle gene expression.

Prdm1a has been identified as a key factor in regulating slow-twitch myoblast migration. Most adaxial cells fail to complete their migration in the absence of Prdm1a (Baxendale et al., 2004; Roy et al., 2001; von Hofsten et al., 2008), whereas transplanted wild-type slow myoblasts are able to differentiate into mononucleated fibres and migrate accordingly in *prdm1a* mutants (Roy et al., 2001).

Notably, the expression pattern of *smyhc1* in *prdm1aMO;sox6*$^{-/-}$ mutant embryos suggests that the loss of sox6 in *prdm1a* morphants rescues the migration of many slow-twitch myoblasts, as *smyhc1* expression is observed mainly in the outer fibres of the myotome. However, in contrast to *prdm1aMO;sox6*$^{-/-}$ mutant embryos, the rescue of slow-specific gene expression in *nrd*$^{-/-}$*;sox6*$^{-/-}$ double mutant embryos, was restricted largely to medial fibres, indicating the failure of these slow-fibres to migrate.

To verify whether the loss of *sox6* in *prdm1a* mutants is able to rescue the migration of slow-twitch fibres, the *sox6*$^{-/-}$*;nrd*$^{-/-}$ double mutant was crossed to the *ubo:GFP* transgenic reporter line. This transgene allowed the migratory behaviour of *prdm1a*-expressing presumptive slow-twitch myoblasts to be compared in double *sox6*$^{-/-}$*;nrd*$^{-/-}$ mutant, single mutant and wild-type embryos. In wild-type embryos, GFP positive fibres were observed only in the fully migrated superficial slow fibres and in the medially located muscle pioneers at 30 hpf (Figure 5.11A and E). By contrast in *nrd*$^{-/-}$ mutants *ubo:GFP* expression was seen throughout the myotome of the embryo, indicating a failure of slow myoblast migration (Figure 5.11C and G). In *sox6*$^{-/-}$*;ubo:GFP* mutant embryos, the pattern of GFP expression is essentially the same as in *sox6*$^{-/+}$*;ubo:GFP* sibling embryos, with only the superficial slow fibres and muscle pioneers expressing GFP (Figure 5.11B and F). In double *sox6*$^{-/-}$*;nrd*$^{-/-}$ mutant embryos, GFP-positive cells were observed throughout the myotome, similar to *nrd*$^{-/-}$*;ubo:GFP* mutant embryos (Figure 5.11D and H). This observation indicates that the loss of *sox6* in *prdm1a* mutants does not rescue slow myoblast migration. Thus, although the loss of *sox6* can rescue the expression of some slow muscle genes in *prdm1a* mutants, rescue of slow muscle fate is incomplete, as adaxial cells fail to complete their radial migration.

In *nrd*$^{-/-}$ embryos, fast muscle genes are expressed ectopically in adaxial cells (Hernandez-Lagunas et al., 2005; Roy et al., 2001). They also lack muscle pioneers, and consequently no horizontal myoseptum is formed, which causes fast-twitch gene expression to be continuous throughout the dorsal-ventral midline. These effects were not rescued by loss of *sox6* activity in *sox6*$^{-/-}$;*nrd*$^{-/-}$ double mutant embryos (Figure 5.12).

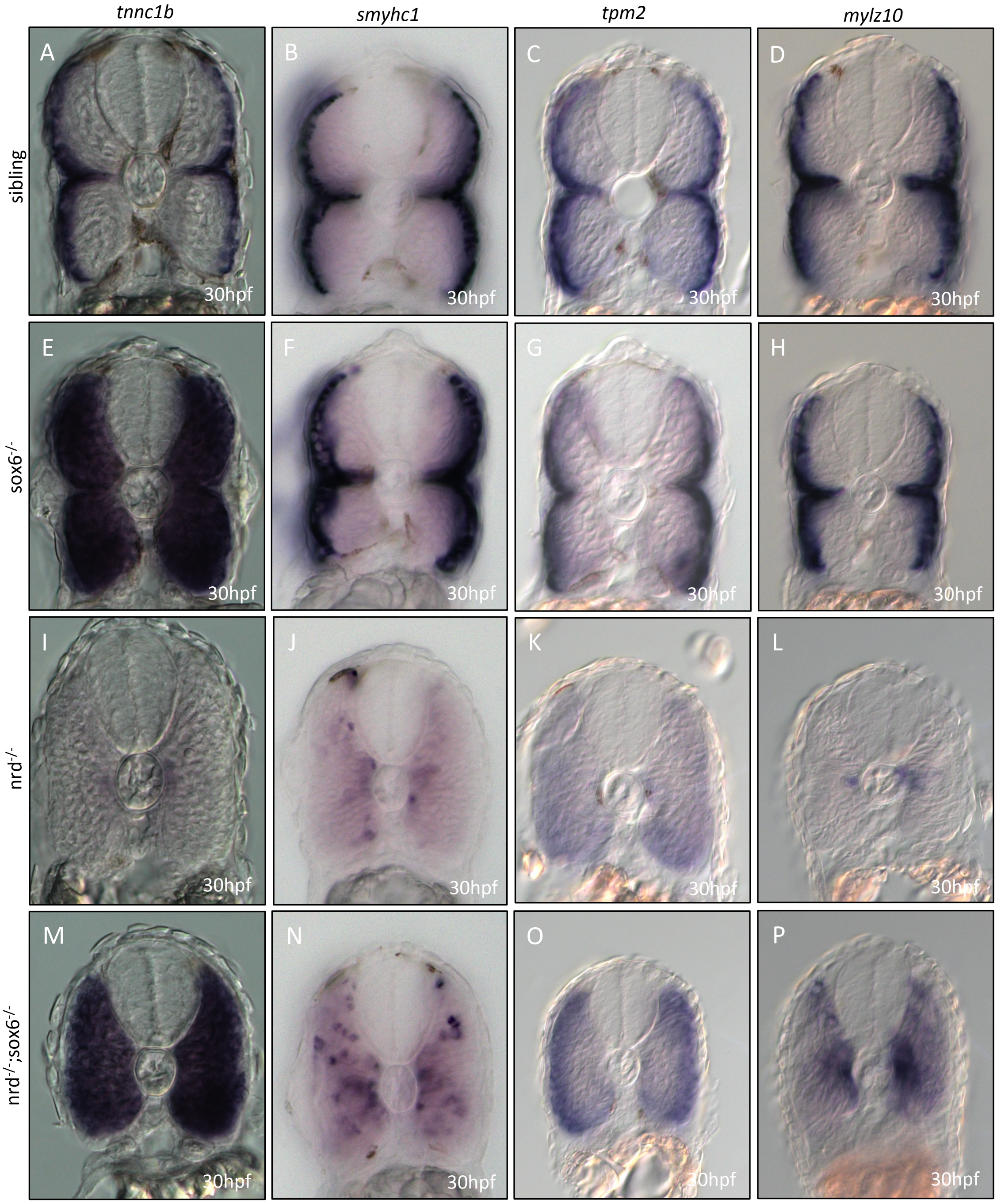


***Figure 5.10: Slow-specific gene expression is rescued in nrd$^{-/-}$;sox6$^{-/-}$ double mutants.*** *tnnc1b, smyhc1, tpm2 and myz10 are slow specific genes which are expressed in the superficial slow-twitch fibres in wildtype embryos at 30hpf (A-D). In sox6$^{-/-}$ mutant embryos the expression of tnnc1b is throughout the myotome, smyhc1 is ectopically expressed in a small subset of fast-twitch fibres and expression of tpm2 and myz10 is unaffected (E-H). In nrd$^{-/-}$ mutants, the expression of slow muscle genes is almost completely lost (I-L). tnnc1b expression is rescued throughout the myotome of double nrd$^{-/-}$;sox6$^{-/-}$ mutants, whereas expression of smyhc1, tpm2 and myz10 is rescued in a small subset of fibres (N-P). Notably the expression of tpm2 appears to be in a different population of cells, when compared to the rescue of smyhc1 and mylz10.*

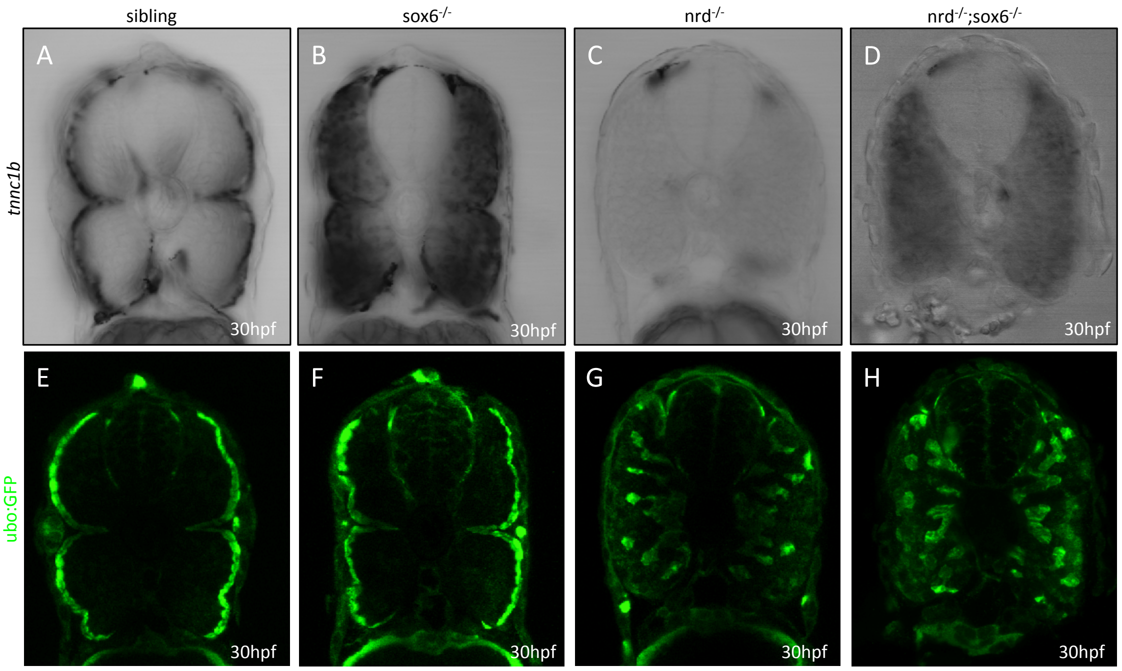


***Figure 5.11: Loss of sox6 in nrd$^{-/-}$ mutants does not rescue slow-twitch muscle fibre migration.*** *ubo:GFP expression is restricted to the slow-twitch fibres by 30 hpf in wild-type embryos and sox6 mutants (E and F). In nrd$^{-/-}$ mutants the expression of ubo:GFP indicates that the cells expressing the reporter gene have not migrated to the outer surface of the myotome (G). The loss of sox6 in nrd mutants does not rescue the migration of these cells to the superficial myotome (H), implying Prdm1a-dependent myoblast migration is regulated independently of sox6.*

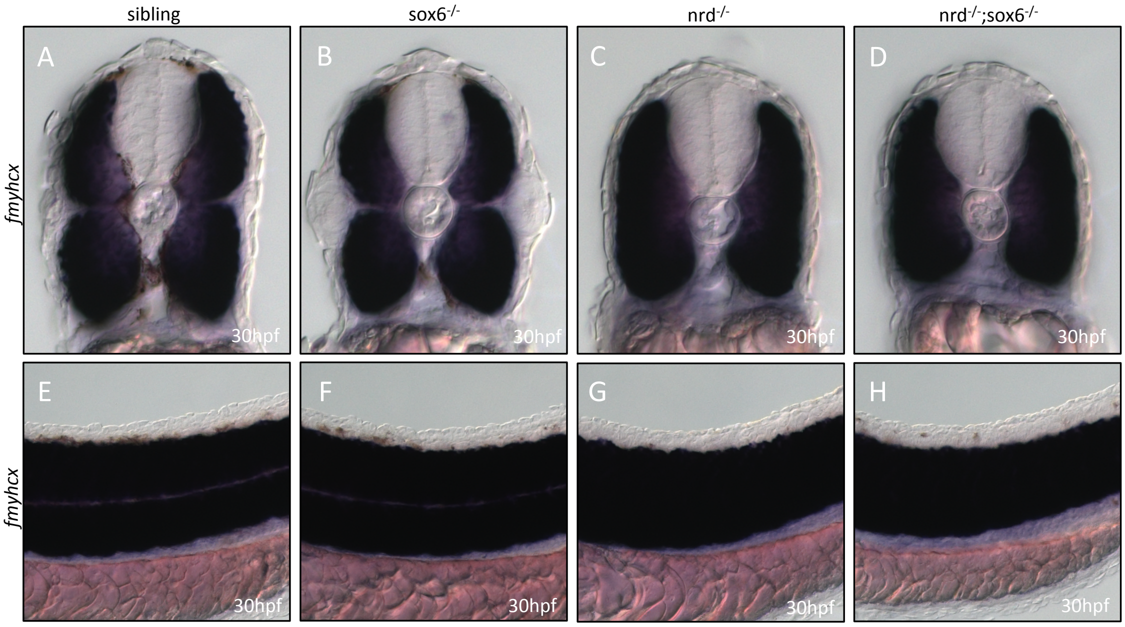


***Figure 5.12: Expression of fast-specific genes is unaffected in sox6$^{-/-}$;nrd$^{-/-}$ double mutants compared to nrd$^{-/-}$ single mutants.*** *In nrd$^{-/-}$ mutant embryos, fmyhcx is expressed throughout the myotome, and neither differentiated slow, nor muscle pioneers, are formed. In double mutant embryos, there is no rescue of the muscle pioneers at the midline, so the expression of fmyhcx still appears to be continuous throughout the myotome.*

### 5.2.6 Fast-twitch gene expression in Cyclopamine treated *sox6$^{-/-}$* mutants

Inhibition of Hedgehog signalling by Cyclopamine treatment causes a failure of slow muscle differentiation in zebrafish embryos and somites appear U-shaped (Wolff et al., 2003). In the proposed model for slow-twitch muscle development, *sox6* transcription is repressed by the Hh-dependent expression of the transcription factor Prdm1a. As *prdm1a* is not expressed in Cyclopamine treated embryos, their muscle phenotype should resemble that of *nrd$^{-/-}$* mutants, with *sox6* being expressed in both the slow and fast- twitch muscle, inhibiting the slow-specific genes throughout the myotome (von Hofsten et al., 2008 and my data). I therefore investigated whether loss of *sox6* could rescue the expression of slow muscle in Cyclopamine-treated embryos. In confirmation of this hypothesis, *sox6$^{-/-}$* mutants treated with Cyclopamine exhibited a rescue in the expression of *tnnc1b* expression throughout the myotome, similar to what is seen in *sox6$^{-/-}$;nrd$^{-/-}$* double mutant embryos (Figure 5.13).

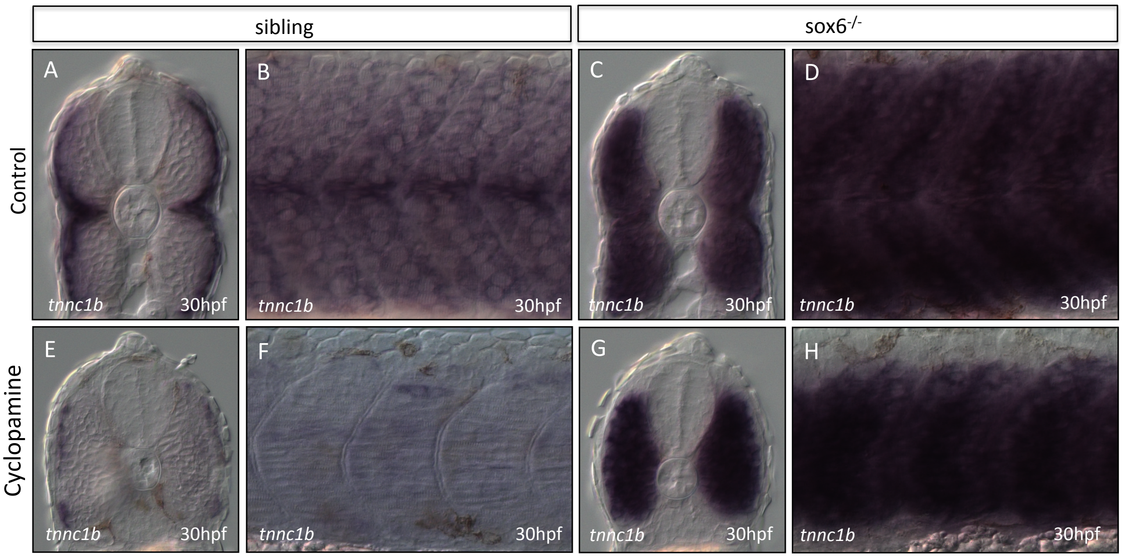


***Figure 5.13: Loss of sox6 in Cyclopamine-treated embryos rescues tnnc1b expression.*** *In Cyclopamine-treated embryos very little tnnc1b is expressed (E and F) when compared to control untreated embryos (A and B). The loss of sox6 in control and Cyclopamine-treated embryos rescues the expression of tnnc1b (C, D, G and H). Rescue of tnnc1b expression was seen in 12 out of 41 embryos treated from a sox6$^{+/-}$ incross.*

### 5.2.7 Knockdown of *six* gene expression in *sox6$^{-/-}$ mutant embryos*

There are many examples of interactions between HMG domain transcription factors and homeodomain transcription factors, which regulate the differentiation of specific cell types during development (for review see Dailey and basilica 2001). One homeodomain-containing transcription factor that has previously been implicated in the transcriptional activation of fast-muscle genes is Six1a (Bessarab et al., 2008; Niro et al., 2010). Since the *six1a* gene is expressed at the same time and in the same population of myoblasts as the HMG domain transcription factor Sox6, I sought to investigate whether these proteins functionally interact during muscle development. Since *sox6$^{-/-}$* mutants exhibit unperturbed, slow-specific expression of *smyhc1* (in contrast to the ectopic expression of the slow-specific gene *tnnc1b*), I hypothesized that *six1a* or another fast-specific transcription factor under the control of *six1a,* might inhibit *smyhc1:GFP* expression in fast muscle. In this view, injecting a *six1a* morpholino (Bessarab et al., 2008) into *sox6$^{-/-}$* mutants would be predicted to cause ectopic *smyhc1:GFP* expression throughout the whole of the fast-twitch domain from 24 hpf onwards. However, injection of the *six1a* morpholino into *sox6* homozygous mutant and sibling embryos had no effect on the expression of *smyhc1:GFP,* when compared to the non-injected *sox6$^{-/-}$* mutants and siblings (Figure 5.14A-B and D-E). This result indicates that the loss of *six1a* is not sufficient to derepress the expression of *smyhc1:GFP* in fast-twitch muscles of wild-type or *sox6$^{-/-}$* mutant embryos. A second *six* gene, *six4a,* is expressed in the adaxial cells as well as in the fast-twitch myoblasts of the zebrafish myotome, suggesting the possibility that the function of this gene might compensate for the loss of *six1a* expression in *six1a* morphants. However, when morpholinos specific for both *six1a* and *six4a* were injected, there was still no increase in ectopic GFP expression in *smyhc1:GFP;sox6$^{-/-}$* embryos (Figure 5.14C and F).

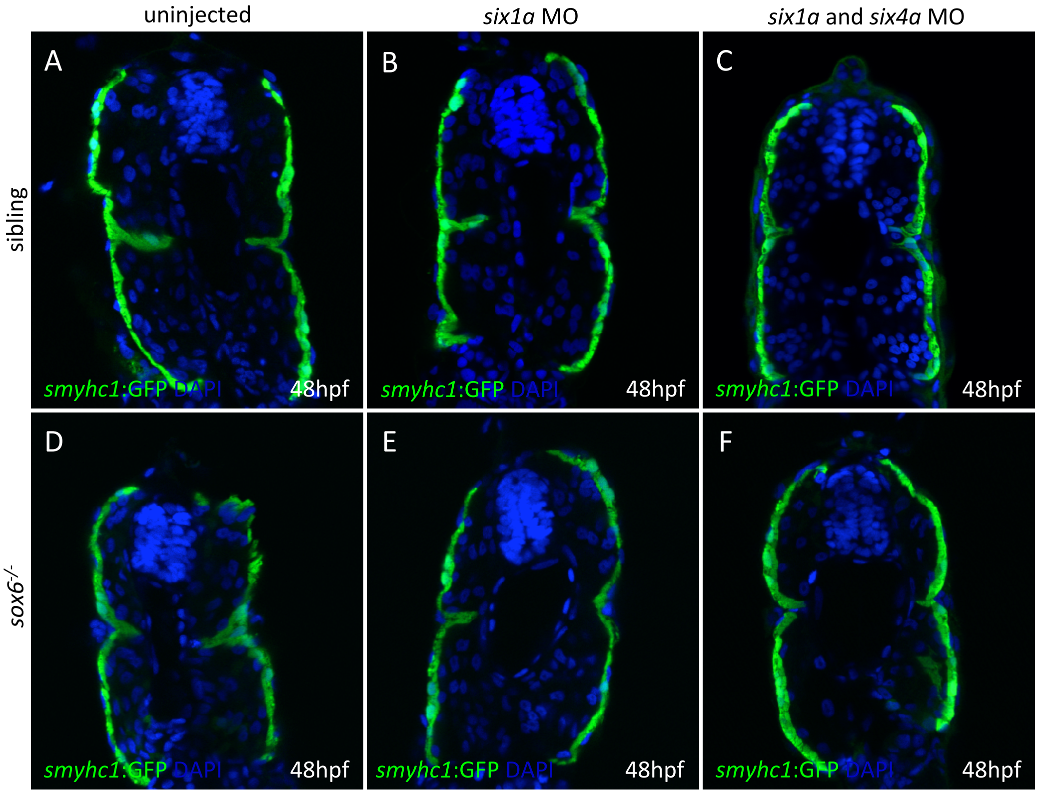


***Figure 5.14: Knockdown of six1a and six4a has no effect on the expression of smyhc1:GFP expression in sox6 sibling or mutant embryos.*** *smyhc1:GFP expression is restricted to the slow fibres in uninjected sox6 siblings (A), six1a morphants (B) and six1a/six4a morphants (C). No increase is observed in the number of ectopic fibres in six1a MO injected and six1a/six4a MO injected sox6$^{-/-}$ mutants (E and F) compared to uninjected sox6$^{-/-}$ mutants (D).*

### 5.2.8 Knockdown of *pbx* gene expression in *sox6*$^{-/-}$ mutant embryos

Pbx proteins are another subtype of homeodomain transcription factor that are involved in transcriptional activation of fast-specific genes (Maves et al., 2007). As with Six1a, I hypothesised that, as this protein is needed for the activation of fast-twitch genes, its expression, or the expression of one of its downstream effectors, could be repressing slow-twitch gene expression in fast-twitch cells. Morpholinos specific for *pbx2* and *pbx4* were injected together into embryos produced by intercrossing *sox6* heterozygous adults. Efficient knockdown of *pbx* was confirmed by the loss of the otic vesicles (Maves et al., 2007). However, there was no effect of either the *pbx2* or the *pbx4* morpholino on *smych1:GFP* expression, in microinjected sox6$^{-/-}$ mutant or sibling embryos, as compared to the uninjected embryos (Figure 5.15). Thus, simultaneous loss of Sox6 and Pbx function is not sufficient to cause derepression of *smyhc1:GFP.*

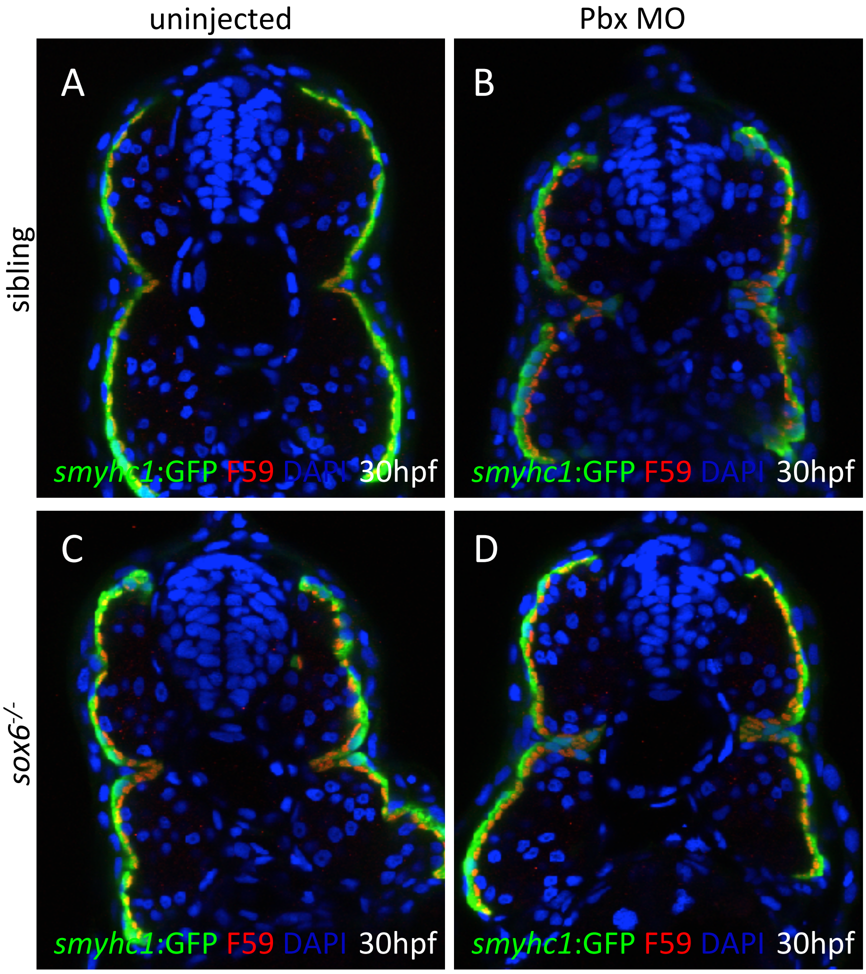


***Figure 5.15: Knockdown of pbx2 and pbx4 has no effect on the expression of smyhc1:GFP expression in sox6 sibling and mutant embryos.*** *Smyhc1:GFP is restricted to the slow twitch fibres in sibling uninjected (A) and Pbx2;Pbx4 morphants (B) In sox6*$^{-/-}$*;Pbx2;Pbx4 morphants (D) the expression of smyhc1:GFP remains restricted to the slow domain with just a few GFP positive fast fibres, similar to what is seen in uninjected sox6*$^{-/-}$ *mutant and sibling embryos (C).*

### 5.2.9 Fast-twitch myoblast fusion in *sox6*$^{-/-}$ mutants

As *sox6* is specifically expressed in the fast-twitch muscle, it may have a role in myoblast fusion, since formation of syncytial, multinucleate fibres is a specific feature of this muscle subtype (Devoto et al., 1996, Roy et al., 2001). Myoblast fusion was assessed by counting the number of nuclei in fast-twitch fibres (demarcated by β-catenin staining) (Figure 5.16A and B). The average number of nuclei in each fast fibre was not significantly different between the *sox6* homozygous mutants and their sibling embryos (Figure 5.16C). This result implies that Sox6 is not required for fast-twitch myoblast fusion.

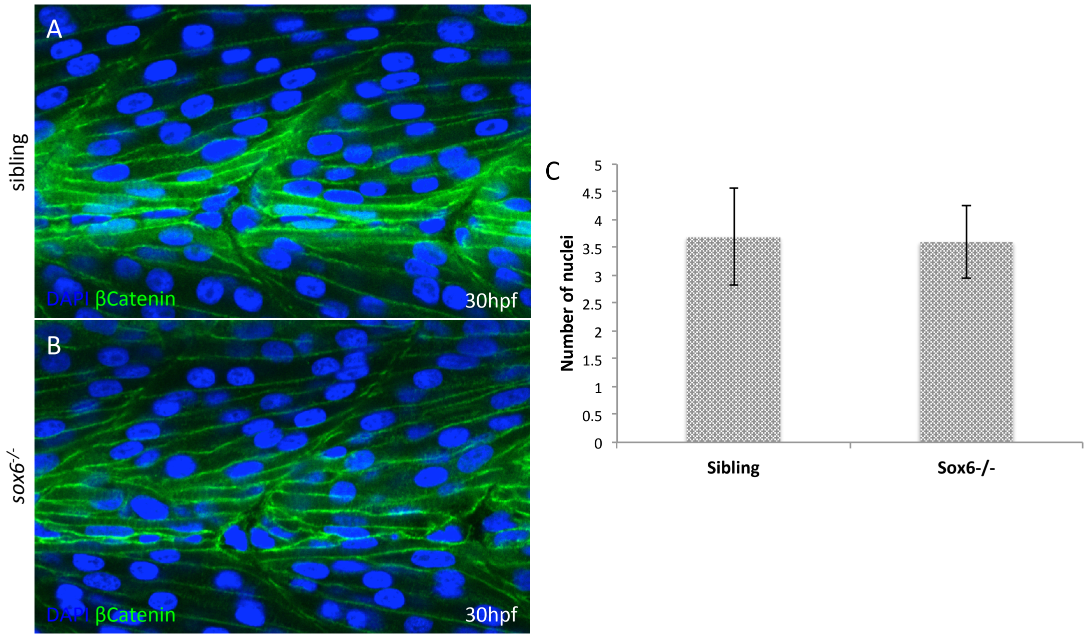


***Figure 5.16: Fast muscles fibres fuse normally in sox6 homozygous mutants*** *(A and B). There is no significant difference between the average number of nuclei per fast muscle fibre in sox6*$^{-/-}$ *mutant and sibling embryos. Student's two-tailed, unpaired t-test P=0.7, error bars represent standard deviation, sibling n=23, mutant n=23 (C).*

### 5.2.10 Analysis of slow-specific muscle gene expression in *sox6*$^{-/-}$ mutants after 3 dpf

The ectopic expression of slow-twitch specific genes seen in *sox6*$^{-/-}$ mutants at 30 hpf varies dramatically between different genes, with the expression pattern of some slow-specific genes not altering at all. In order to see if the expression of some of these slow-specific genes was faithfully restricted to the slow domain of *sox6*$^{-/-}$ mutants at later stages of development, *in situ* hybridisation was performed on *sox6*$^{-/-}$ mutant larvae. Interesting, for each slow-specific gene examined, there were at least some ectopically expressing fibres in the mutants. Thus, by 3 dpf, fast fibres positive for *tpm2* and *mylz10* expression could be readily visualised in *sox6*$^{-/-}$ mutant larvae (Figure 5.17B-C and E-F). Similarly, more fast fibres expressed *smyhc1* in *sox6*$^{-/-}$ mutants by 4 dpf than at 30 hpf (Figure 5.17A and D). Like *smyhc1* expression, it seemed that the expression of *tpm2* and *mylz10* in *sox6*$^{-}$

$^{/-}$ mutant larvae was not detected throughout the whole of the fast-twitch domain, rather this expression was only observed in a few fast fibres. Expression of *smyhc1:GFP* was also compared in *sox6*$^{-/-}$ mutant and sibling larvae at 5 dpf. At this stage, a number of fast-twitch fibres were GFP-positive most of which are located on the ventral side of the embryo (Figure 5.18), comparable to what is seen by *in situ* hybridisation for *smyhc1, tpm2* and *mylz10*. Moreover, it is noticeable that expression of fast myosin light chain, detected with the F310 antibody, is not lost in the fibres expressing *smyhc1:GFP;* instead these two proteins are co-expressed in the same fibres (Figure 5.18B and E). Therefore ectopic expression of slow-twitch muscle genes observed in *sox6*$^{-/-}$ mutant larvae does not occur at the expense of fast-twitch specific gene expression.

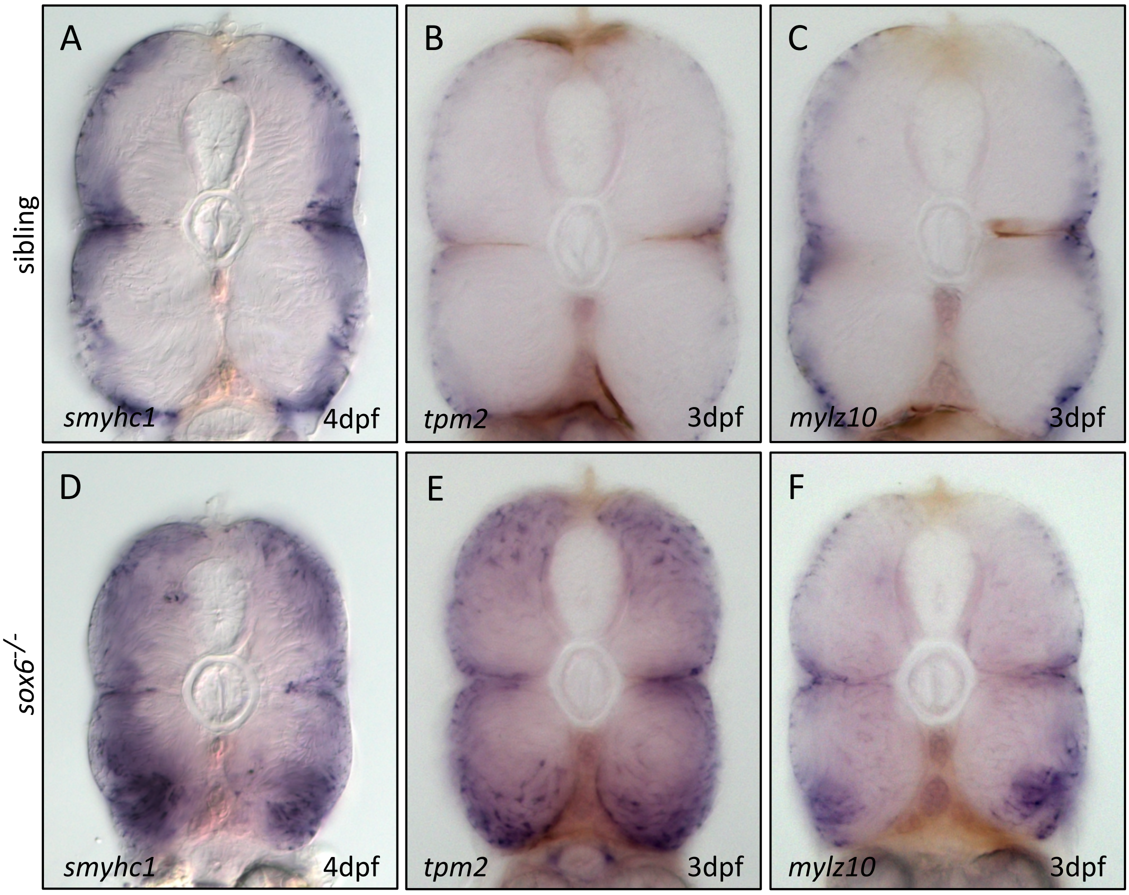


***Figure 5.17: Ectopic expression of slow-twitch genes in the fast-twitch domain of 3 dpf sox6$^{-/-}$ mutant larvae.*** *smych1, tpm2 and mylz10 are restricted to the slow-twitch fibres in wild-type embryos (A-C). In sox6$^{-/-}$ mutants, by 3 dpf some ectopic expression of all three genes is detectable in the fast fibres of the myotome (D-F).*

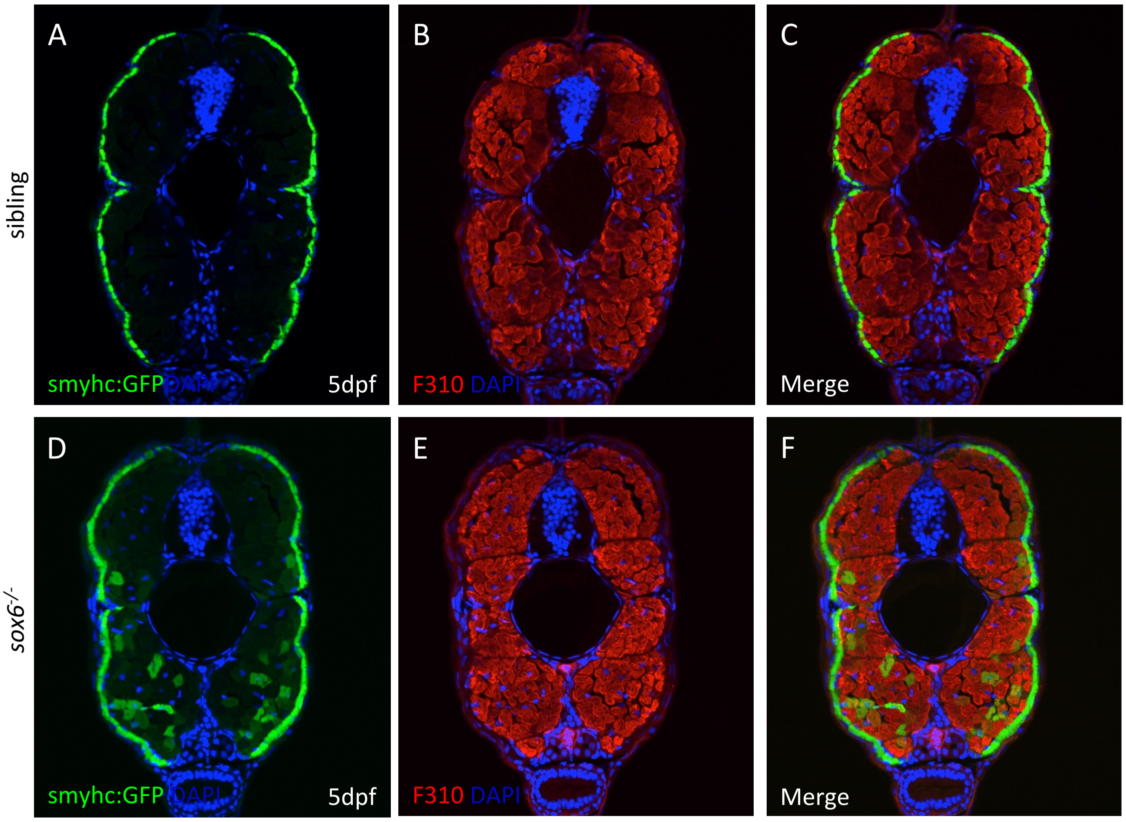


***Figure 5.18: Slow-twitch and fast-twitch specific genes are co-expressed in a subset of muscle fibres located in the fast-twitch domain of sox6$^{-/-}$ mutant larvae at 5 dpf.*** *smyhc1:GFP expression is restricted to the slow fibres at 5 dpf in sibling larvae (A). Analysis of smyhc1:GFP;sox6$^{-/-}$ mutant larvae reveals that at 5 dpf a subset of fast-twitch fibres ectopically express GFP (D). Expression of the fast-twitch marker F310 in the fast domain remains unaffected (B and E), with the same muscle fibres expressing both smyhc1 and fast myosin heavy chain isoforms (C and F).*

### 5.2.11 Bone and cartilage differentiation in *sox6$^{-/-}$* mutants

*Sox6$^{-/-}$* mutants display no observable morphological phenotype during embryogenesis. Moreover, some homozygous sox6$^{-/-}$ mutants are able to survive until adulthood if they are separated from the siblings, but they are not able to reproduce. By approximately 20-30 dpf homozygous mutant fish start to exhibit a slight curvature of the spine, or scoliosis, with the phenotype becoming more severe as they progress into adulthood. The severity of scoliosis varies between each fish, but the majority of fish struggle to swim. This difficulty in swimming is also compounded by a swim bladder defect, as the mutants are unable to regulate their buoyancy and tend to sink to the bottom of the tank. Although some mutants are able to survive until adulthood, the mortality rate is very high, with the vast majority of fish dying before they reach 2 months of age (6 out of 40 homozygotes survived to 2 months). *sox6$^{-/-}$* homozygous mutants are also significantly smaller than their heterozygous or wild-type siblings at one month of age (Figure 5.20F). In order to assess this scoliosis phenotype more closely, the development of cartilage and bone was examined. As described in Chapter 1, Sox6 and Sox5 co-operate with Sox9 to regulate chondrogenesis in mice (Lefebvre et al., 1998). Alcian Blue staining was therefore used to stain the cartilage in *sox6*$^{-/-}$ mutant and sibling fish. There were no consistent morphological differences in cartilage development between the sibling and mutant fish at 6 dpf and 13 dpf (Figure 5.19), although 2 out of the 5 mutants examined at 13 dpf were already noticeably smaller than their siblings and showed a slight delay in cartilage development.

When bone formation was analysed using Alizarin Red staining, scoliosis was readily apparent in all *sox6*$^{-/-}$ mutant fish examined, with a slight curve in the spine noticeable as early as 17 dpf in some fish (Figure 5.20A and C). The scoliosis was a very variable phenotype, with some fish exhibiting only a slight kink in the spine, whereas other fish displayed a massive crumpling of the spine. At 17 dpf the majority of the mutant fish (4/5) showed a delay in the formation of the bone (Figure 5.20A and C). By 27 dpf, *sox6*$^{-/-}$ mutant fish were significantly smaller when compared to siblings, but the structure of the bone appeared mostly normal apart from the scoliosis (Figure 5.20B, D and F). The scoliosis was apparent not only along the antero- posterior axis of the fish, but also along the medio-lateral axis (Figure 5.20 E). By 80 dpf, surviving *sox6*$^{-/-}$ mutant fish all displayed scoliosis to some extent (Figure 5.20 G-J). *sox6*$^{-/-}$ mutants also displayed abnormalities in the formation of the spinal processes, particularly towards their caudal end. Some spinal processes were abnormal in shape, and a subset of these processes displayed branching not seen in sibling fish. As the sample size at this stage was very small (mutants, n=2), it is difficult to infer whether these variations represent natural phenotypic variance within the spinal processes or whether they are caused by loss of *sox6*. Nevertheless, the scoliosis phenotype is clearly linked to the *sox6*$^{-/-}$ mutation, and could be either a primary effect resulting from the loss of *sox6* function in bone, or an indirect secondary consequence of abnormal myogenesis.

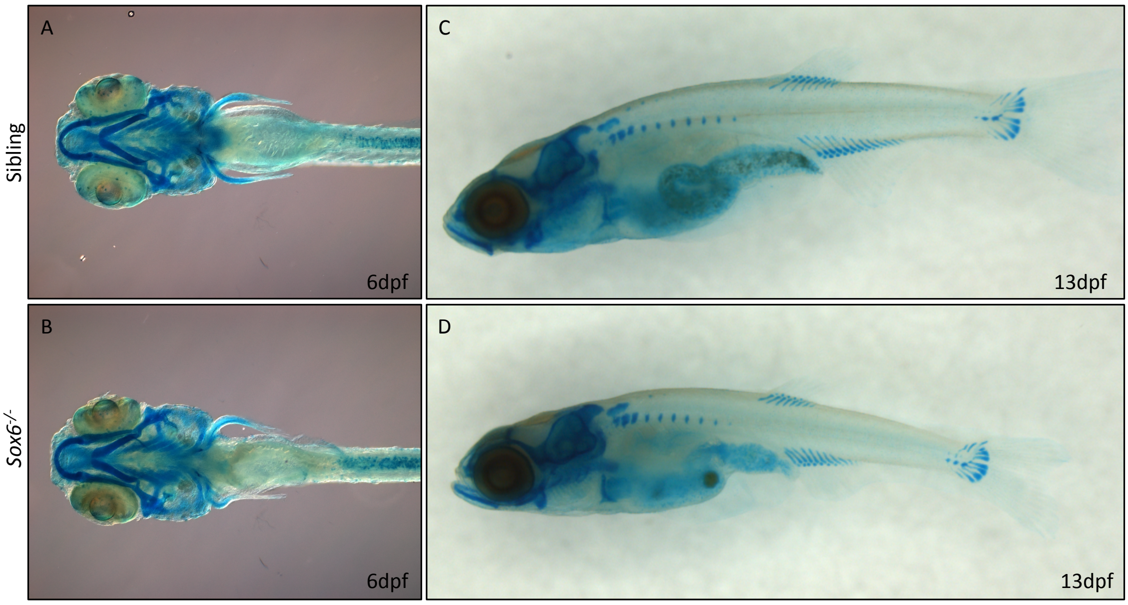


***Figure 5.19: Alcian Blue staining of sox6$^{-/-}$ mutant and sibling fish revealed that there is no apparent defect in cartilage formation at 6 dpf (A and B) and 13 dpf (C and D) in sox6$^{-/-}$ mutants, as compared to sibling fish.***

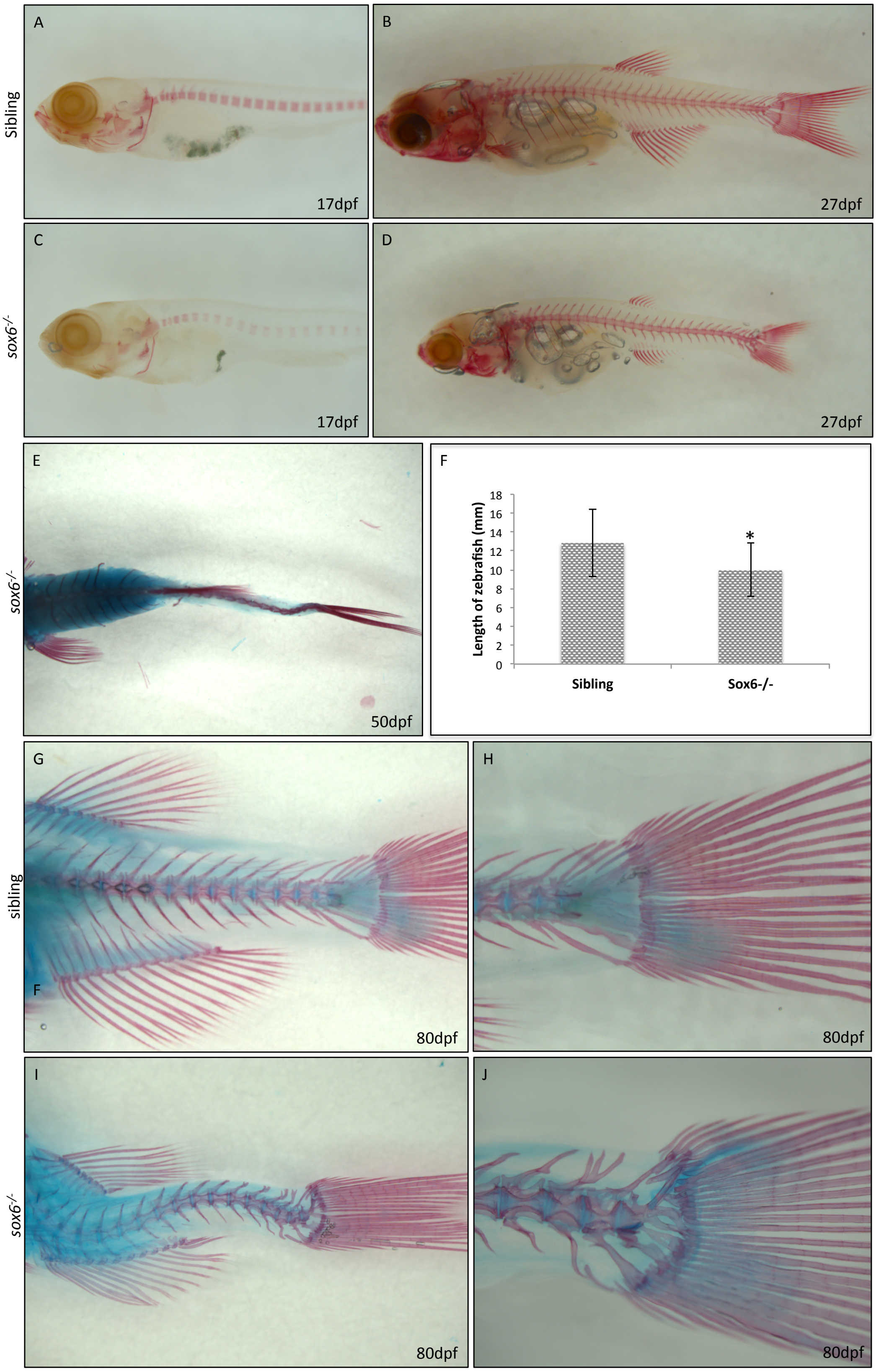


***Figure 5.20: Alizarin Red staining of sox6$^{-/-}$mutant and sibling fish indicated that scoliosis begins in sox6$^{-/-}$ mutants at approximately 17-27 days of age*** *(A-D). At one month of age, sox6$^{-/-}$ mutants are significantly smaller than siblings (indicated by asterisk). Student's two-tailed, unpaired t-test $P<0.05$, error bars represent standard deviation, sibling n=17, mutant n=17 (F). The scoliosis of the spine in sox6$^{-/-}$ mutants is evident along both the anterior-posterior axis as well as the medial-lateral axis (E). By 80 dpf there is clear scoliosis of the spine in all surviving homozygous sox6 mutant fish (G-J).*

### 5.2.12 Analysis of adult *sox6$^{-/-}$* mutant myotome

To address whether the scoliosis phenotype results from the direct loss of *sox6* function or is a secondary consequence of myotomal abnormalities, muscle fibres were examined by visualizing the expression of fast-twitch and slow-twitch markers in sections through the trunk regions of adult *sox6$^{-/-}$* mutants and siblings carrying the *tnnc1b:eGFP* reporter. These revealed that *tnnc1b* expression is robustly overexpressed throughout the myotome of adult *sox6$^{-/-}$* mutants, as well as during development (Figure 5.21). However, the expression of *tnnc1b:eGFP* in the slow domain is clearly stronger than in the fast domain of *sox6$^{-/-}$* mutant fish. The overall morphology of the mutant myotome is very different to that of sibling fish. In cross-section, the *sox6$^{-/-}$* mutant myotome appears much smaller than that of sibling fish and its shape varies between individuals, unlike the highly regular cross-sections of wild-type or heterozyogous siblings (Figure 5.21). The sharp border between slow- and fast-twitch fibres typical of wild-type fish is replaced by an irregular border in *sox6$^{-/-}$* mutants, and the muscle fibres are significantly smaller than those observed in sibling fish, in some cases by as much as 10 times smaller.

The differential expression of *smyhc1:GFP* between siblings and *sox6$^{-/-}$* mutants also appears to be maintained into zebrafish adulthood. As was seen in *sox6$^{-/-}$* mutant embryos, the ectopic expression of *smyhc1* observed in the adult is subtle (Figure 5.22). There is a slight medial expansion of the GFP-positive domain, with some cells in the most medially located muscle fibres faintly expressing GFP, as well as a dorsal and ventral expansion in the expression domain (Figure 5.22B and F). In the fast-twitch domain of *sox6* siblings, no GFP-positive fibres can be identified, whereas in the mutants there are a few fast-twitch fibres faintly expressing GFP (Figure 5.22D and H).

These results show that in sox6$^{-/-}$ mutants the aberrant expression patterns of *tnnc1b:GFP* and *smyhc1:GFP* are maintained into zebrafish adulthood, with *tnnc1b:GFP* being expressed throughout the whole myotome and *smyhc1:GFP* being ectopically expressed in a subset of fast-twitch fibres. *In situ* hybridisation analysis is needed to confirm this result in order to indicate whether the *tnnc1b:GFP* and *smyhc1:GFP* transgenes are recapitulating the endogenous gene expression.

An *in situ* hybridisation using the *smyhc* probe was also carried out on these adult sections. The *smyhc* probe identifies three isoforms of slow myosin heavy chain (Elworthy et al., 2008). Before 48 hpf only *smyhc1* is expressed in the slow myoblasts, therefore the *smyhc* probe will only identify this isoform (Elworthy et al., 2008). *smych2* and *smyhc3* are expressed from 48 hpf and 96 hpf respectively, so potentially all three isoforms could be expressed in adulthood (Elworthy et al., 2008). The *in situ* hybridisation on adult siblings using the *smyhc* probe identified an expression domain larger than the domain identified by the *smyhc1:GFP* transgene, implying that *smyhc2* and/or *smyhc3* are expressed in at a least some distinct muscle cells to *smyhc1* (Figure 5.23). The expression domain identified with the *smyhc* probe overlapped with the *tnnc1b:eGFP* adult expression. In adult *sox6*$^{-/-}$ mutant sections the expression of the *smyhc* genes appears to be extended more medially (Figure 5.23), similar to what is seen in *tnnc1b:eGFP;sox6*$^{-/-}$ mutants (Figure 5.21). This implies that other *smyhc* isoforms, besides *smyhc1*, are derepressed in the absence of Sox6.

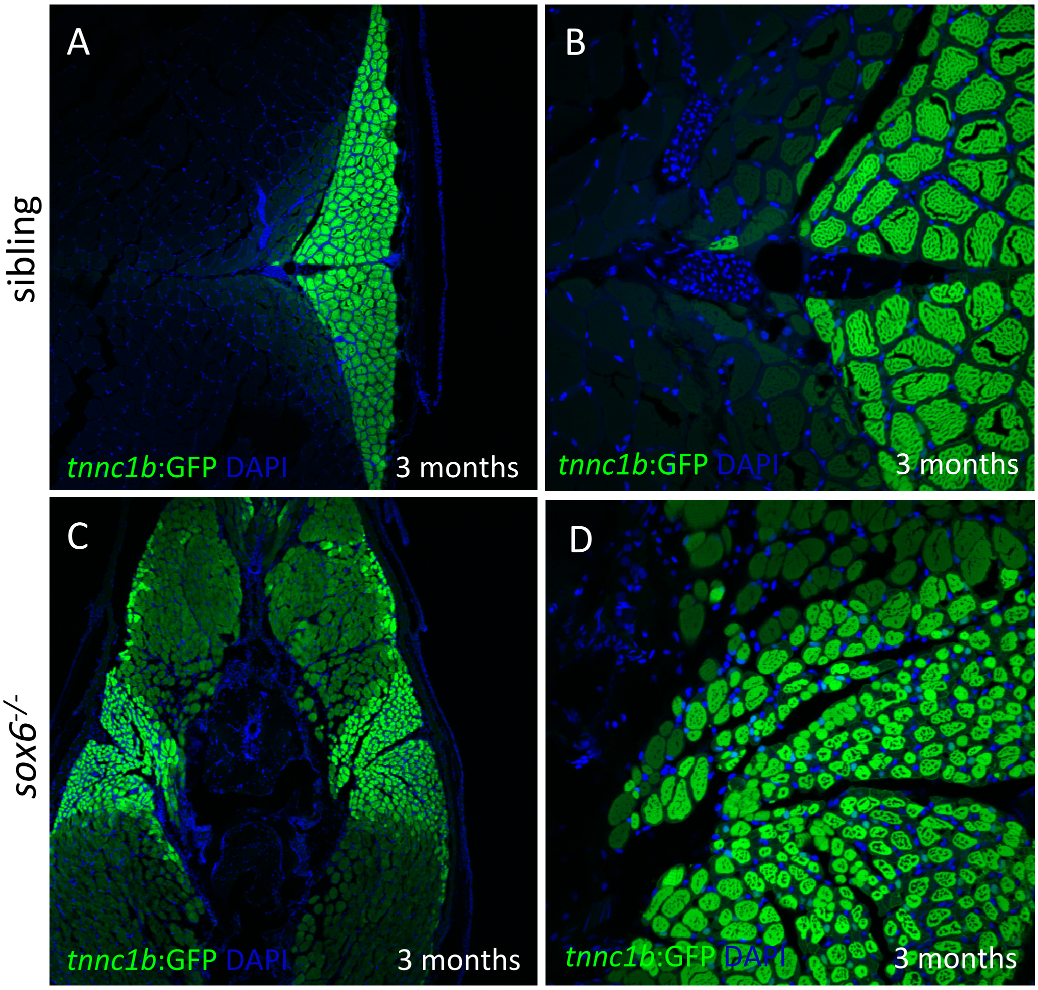


***Figure 5.21: tnnc1b expression persists throughout the myotome of sox6*$^{-/-}$ *mutant adult fish.*** *In wild-type zebrafish adults, tnnc1b:eGFP expression is found in a wedge of myotome, which corresponds to the adult slow-twitch fibre domain (A). In sox6*$^{-/-}$ *mutants, tnnc1b:eGFP expression is seen in both the fast and slow-twitch fibres and the wedge shape appears to encroach further towards the midline of the fish (B). The overall size of the myotome is significantly smaller in sox6*$^{-/-}$ *mutant fish (panels A and C, and B and D are at the same magnification). The size of the individual fibres is also smaller in sox6*$^{-/-}$ *mutant fish (B and D).*

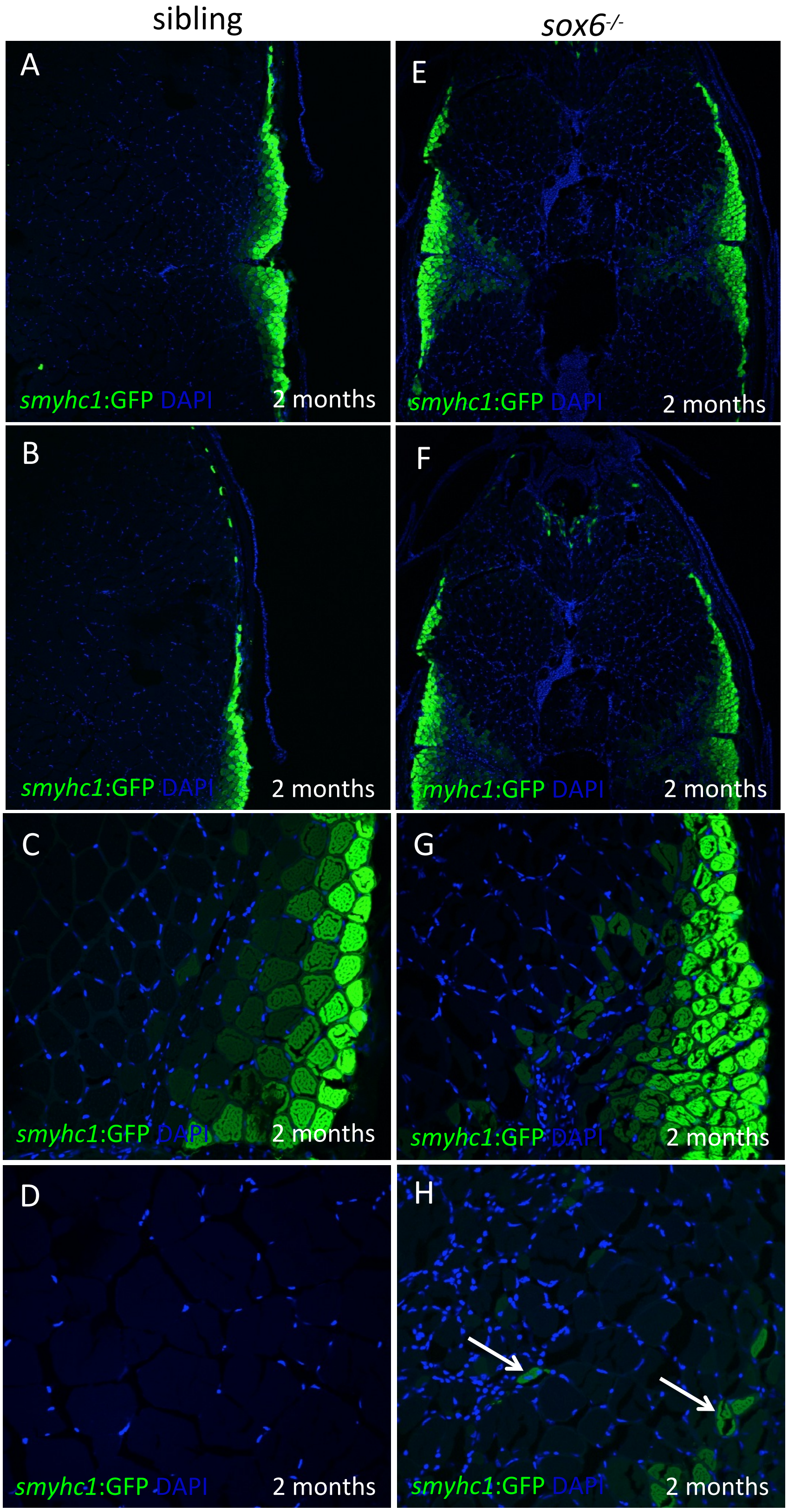


***Figure 5.22: The smyhc1 expression domain is slightly expanded in $sox6^{-/-}$ mutant adults in comparison to the situation in sibling fish****. The expression of smyhc1:GFP is spread further dorsally and ventrally in the $sox6^{-/-}$ mutants (A , B, E and F). There appears to be more fibres expressing GFP in the slow domain of $sox6^{-/-}$ mutants (C and G). In the fast domain of sibling fish there are no GFP- positive fibres (D). Conversely, in $sox6^{-/-}$ mutant fish, a few GFP-positive fibres can be detected (H)(arrows represent ectopic GFP expression in the fast-twitch muscle).*

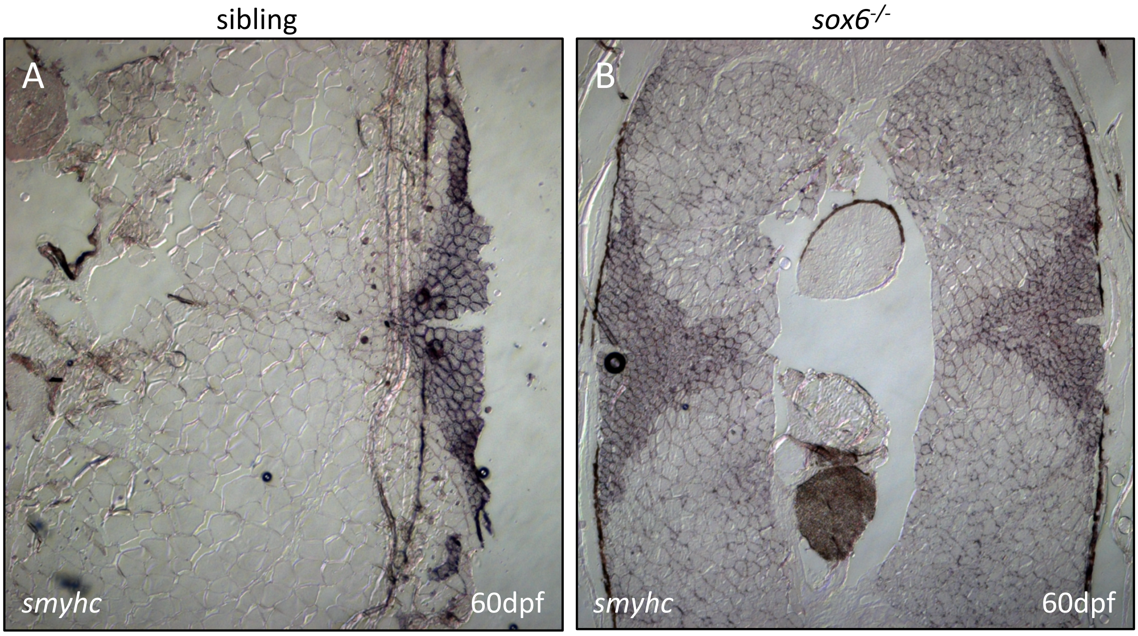


***Figure 5.23: in situ hybridisation to detect smyhc isoforms in sox6$^{-/-}$ mutant and sibling fish.*** *In the sibling fish the expression of smyhc isoforms are restricted to the wedge shaped muscle on either side of the myotome named lateralis superficialis (Devoto et al. 1996; Elworthy et al. 2008) (A)*. *In sox6$^{-/-}$ mutants this expression domain has spread further medially (B), with more smyhc-positive fibres than there are observed in smyhc1:GFP;sox6$^{-/-}$ mutants, indicating that other smyhc isoforms are also derepressed in sox6$^{-/-}$ mutants.*

As described above, the muscle fibres in the sox6$^{-/-}$ mutant are significantly smaller than those seen in the sibling. Intriguingly, despite the fact that Sox6 protein is restricted to the fast-twitch muscle during development, both the slow-twitch and the fast-twitch muscle fibres are reduced in size in the *sox6$^{-/-}$* mutant (Figure 5.24). Moreover, it appears that the slow-twitch fibres are reduced in size to a larger degree than are the fast twitch fibres. One explanation for this could be that in zebrafish adults, *sox6* is expressed throughout the myotome and is no longer restricted to the fast domain, as in earlier stages of development. A perhaps more interesting possibility is that the slow-twitch fibres are undergoing hyperplasia in response to a loss of fast-twitch fibres. Further investigation including counting fibre number is needed to determine if either of these hypotheses is true.

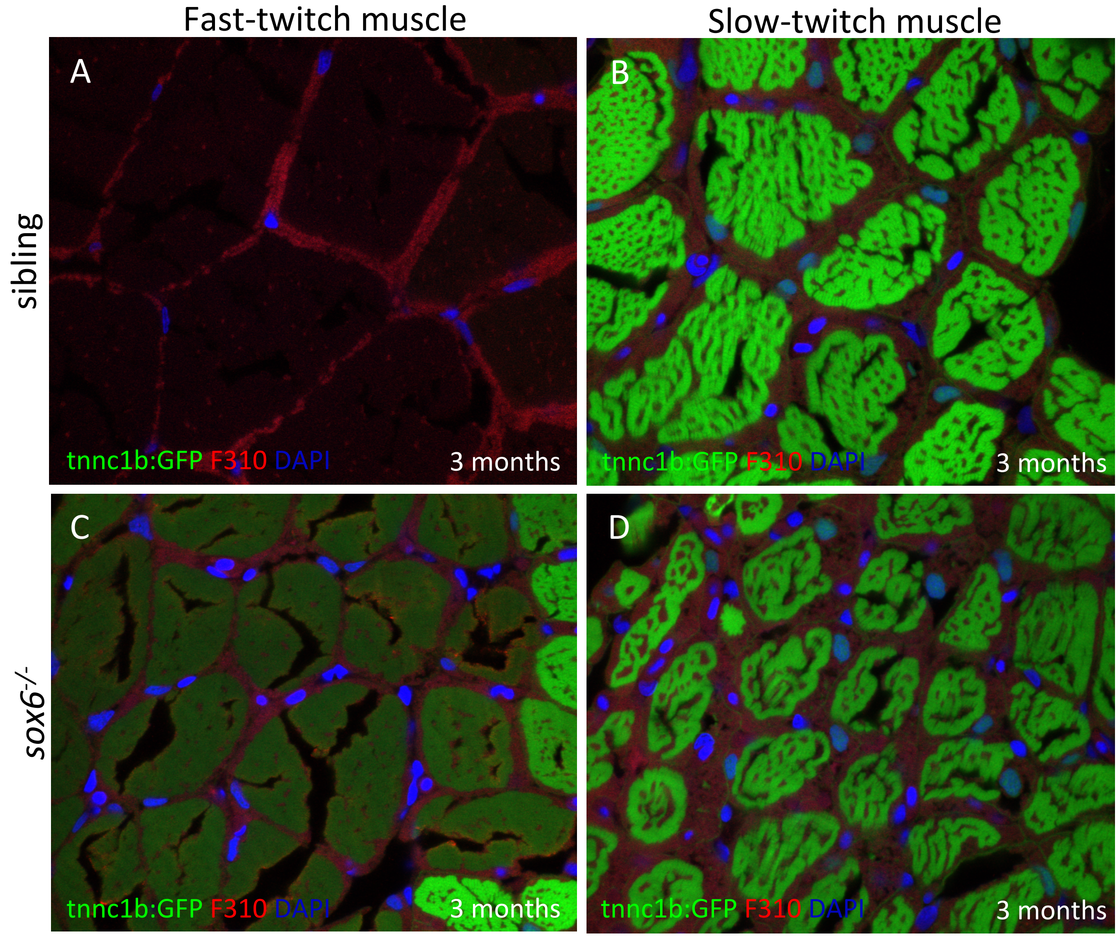


***Figure 5.24: Both slow- (B and D) and fast-twitch (A and C) muscle fibres are smaller in sox6 homozygous mutants in comparison to the muscle fibres of sibling fish.*** F310 staining is non-specific signal, as the staining appears in the sarcolemma instead of the myofibres***.*** However, the staining is useful in this image to outline individual fibres.

### 5.2.13 Sox6 expression in the bone

In order to determine whether the scoliosis phenotype results from a primary effect due to the loss of *sox6* expression in the bone, or a secondary effect due to the loss of *sox6* expression in the muscle, we must first establish whether *sox6* is expressed in zebrafish bone. The opercle, which makes up part of the operculum, is one of the first bones to develop in zebrafish embryos at approximately 4 dpf (Figure 5.25A) (Wada et al., 2010). *In situ* hybridisation performed on wild-type embryos at 4 dpf revealed that *sox6* is not expressed in this bone (Figure 5.25B-D). This gives us an indication that *sox6 is* not expressed in zebrafish bone, although this experiment was done by eye and no bone specific marker was used to label the opercle as a positive control. More work is needed to determine whether it is expressed in other skeletal elements, such as the spinal column, at later stages.

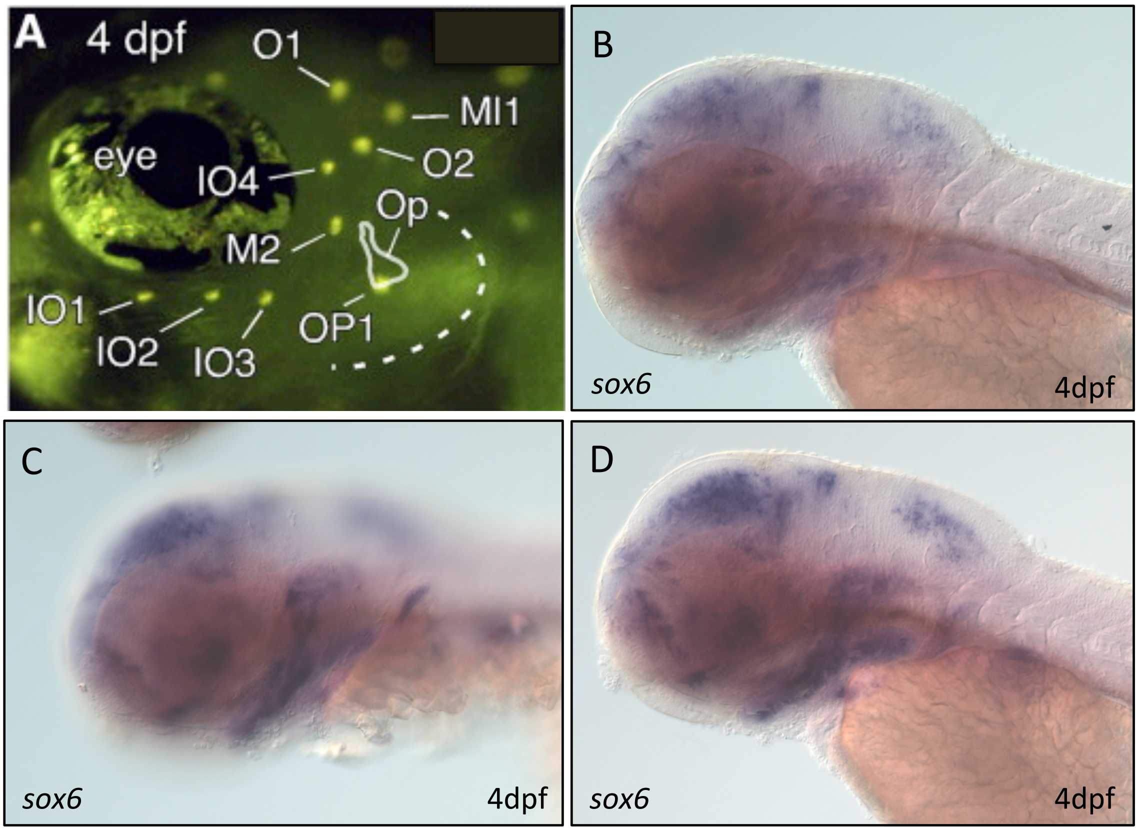


***Figure 5.25: sox6 expression is not expressed in the opercula in wild-type embryos at 4 dpf.*** *The opercle bone is situated posterior to the eye, outlined in panel A. Op- opercle (Panel A taken from Wada et al., 2010). In situ hybridisation revealed no sox6 expression can be detected in the opercle at 4 dpf, implying that sox6 is probably not expressed in the bone (B-D are of the same embryo taken on a different focal plane).*

### 5.2.14 Sox5 zinc-finger nuclease mutant

One possible explanation as to why the loss of *sox6* does not result in ectopic expression of all slow-twitch muscle genes in zebrafish could be due to partial functional redundancy of Sox6 protein with another Sox protein in this organism. One potential source of redundant function could be the closely related *sox5* gene. As described in Chapter 1, *sox5* and *sox6* often function at least partially redundantly with each other during development in mammals, in processes such as chondrogenesis (Lefebvre et al., 1998, Smits et al., 2001). However, although *sox5* is expressed throughout zebrafish embryonic development, it is only weakly expressed in the muscle (Figure 5.26).

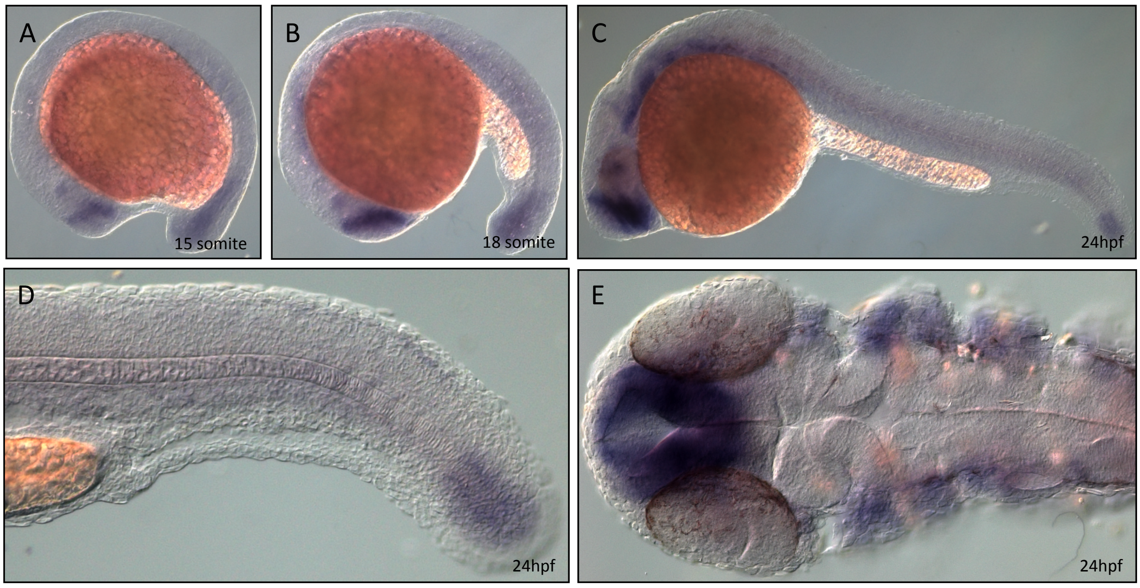


***Figure 5.26: The expression of sox5 during zebrafish development**. sox5 is first specifically expressed in the embryonic tailbud and head regions, most prominently the forebrain, at the 15-somite stage, with expression persisting in these regions until at least 24 hpf. From the approximately 18-somite stage, sox5 is weakly expressed in the somites with this expression persisting until at least 24 hpf.*

### 5.2.15 Elimination of *sox5* does not enhance the *sox6*$^{-/-}$ muscle phenotype

To address the possibility of functional redundancy between the Sox5 and Sox6 proteins in zebrafish myogenesis, a *sox5* mutant was generated using Zinc-Finger Nucleases (Generated by Dr. Stone Elworthy). Although the *sox5*$^{-/-}$ mutants alone had no embryonic phenotype, the *sox5*$^{-/-}$*;sox6*$^{-/-}$ double mutants exhibited a distinct phenotype from 3 dpf (Figure 5.27). The *sox5*$^{-/-}$*;sox6*$^{-/-}$ double homozygous mutants lacked a swim bladder and appeared slightly smaller than the siblings. The pectoral fins projected out at an abnormal angle, and the heads were smaller than sibling. There was some oedema around the head and the lower jaw was absent (Figure 5.27).

The expression of slow-twitch specific genes was compared, first between the siblings and the single *sox5*$^{-/-}$ homozygous fish. In *sox5*$^{-/-}$ embryos, no fast fibres expressing any of the slow-specific genes could be identified (Figure 5.28). Instead, the expression of slow-twitch specific genes in *sox5*$^{-/-}$ mutants remained restricted to the slow domain, as was observed for sibling larvae.

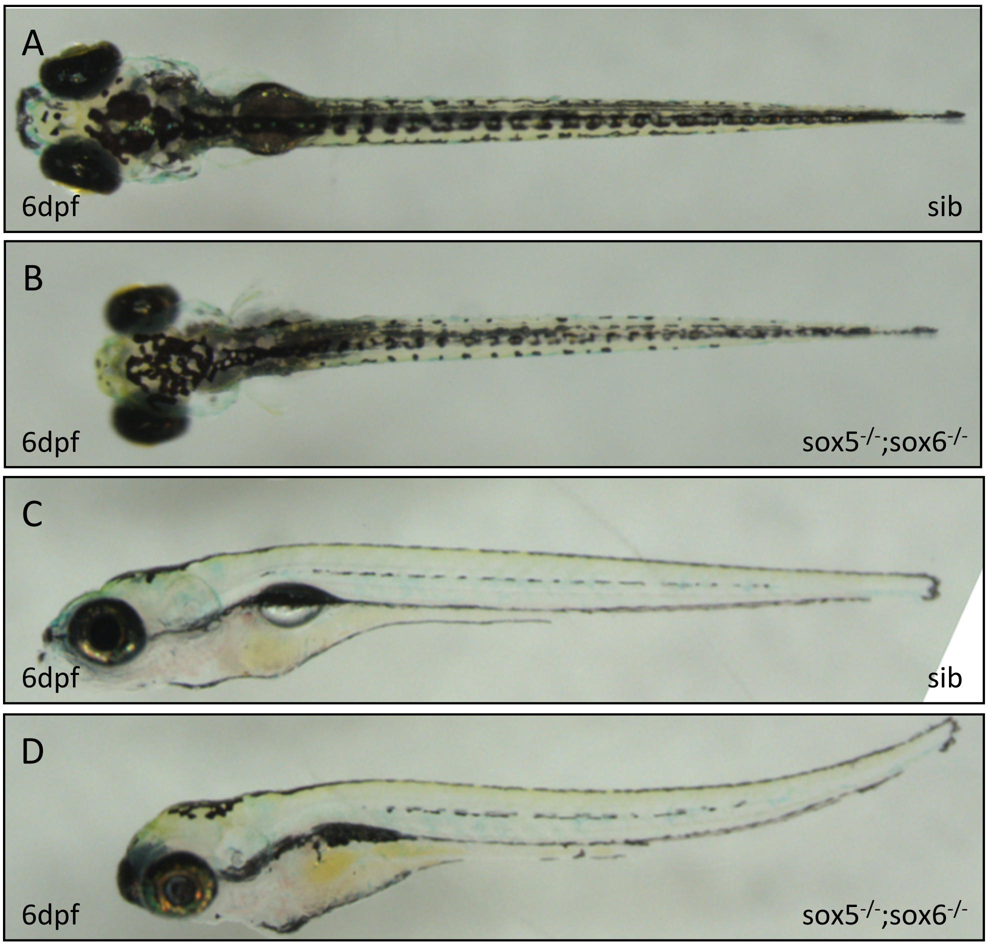


***Figure 5.27: Morphological comparison of sox5$^{-/-}$;sox6$^{-/-}$ mutant and sibling larval phenotypes at 6 dpf.*** *sox5$^{-/-}$;sox6$^{-/-}$ double mutant larvae have a smaller head and oedema around the otic vesicles, and their pectoral fins protrude at an abnormal angle. The sox5$^{-/-}$;sox6$^{-/-}$ double mutant larvae also appear to lack a lower jaw. (The curvature of the body axis seen in panel D is not representative of the double mutants).*

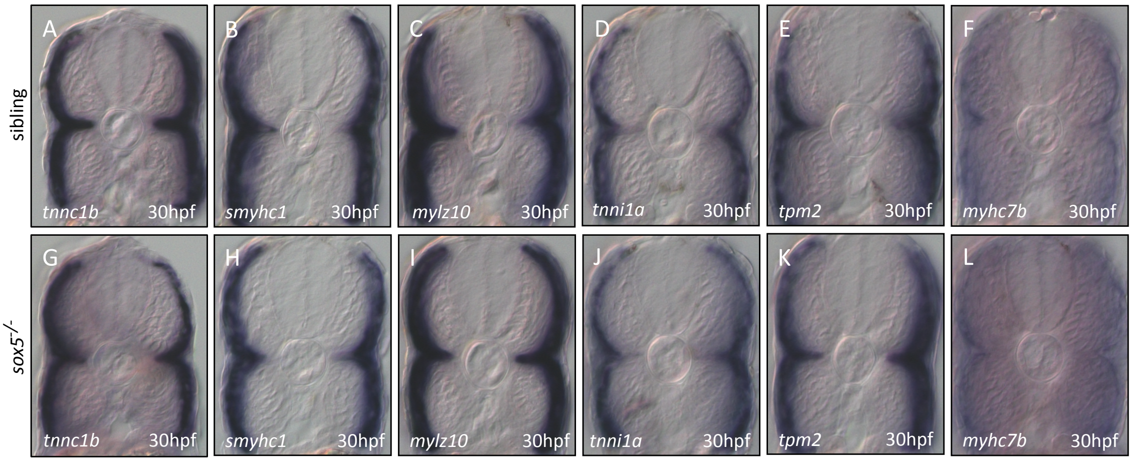


***Figure 5.28: The expression of the slow-twitch specific muscle genes tnnc1b, smyhc1, mylz10, tnni1a, tpm2 and myh7b is unchanged in sox5$^{-/-}$ mutants (G-L) compared to siblings (A-F).***

In order to investigate whether *sox6* and *sox5* might execute partially redundant functions in the myotome, the expression pattern of slow muscle genes was examined in *sox5*$^{-/-}$ single mutant, *sox6*$^{-/-}$ single mutant, *sox5*$^{-/-}$;*sox6*$^{-/-}$ double mutant and sibling embryos at 30 hpf. For the three markers examined, the phenotype of the *sox5*$^{-/-}$;*sox6*$^{-/-}$ double mutant embryos was indistinguishable from that of *sox6*$^{-/-}$ single mutant embryos, and the phenotype of the *sox5*$^{-/-}$ single mutant embryos was indistinguishable from that of sibling embryos (Figure 5.29). Moreover, there was also no significant difference in the number of fast fibres expressing *smyhc1* between the *sox6*$^{-/-}$ embryos and the *sox5*$^{-/-}$;*sox6*$^{-/-}$ double mutants at 3 dpf (double mutants could be separated by phenotype, single mutants were genotyped) (Figure 5.30). Taken together, these results implied that *sox5* does not have a function in muscle fibre-type specification that is redundant with that of *sox6*.

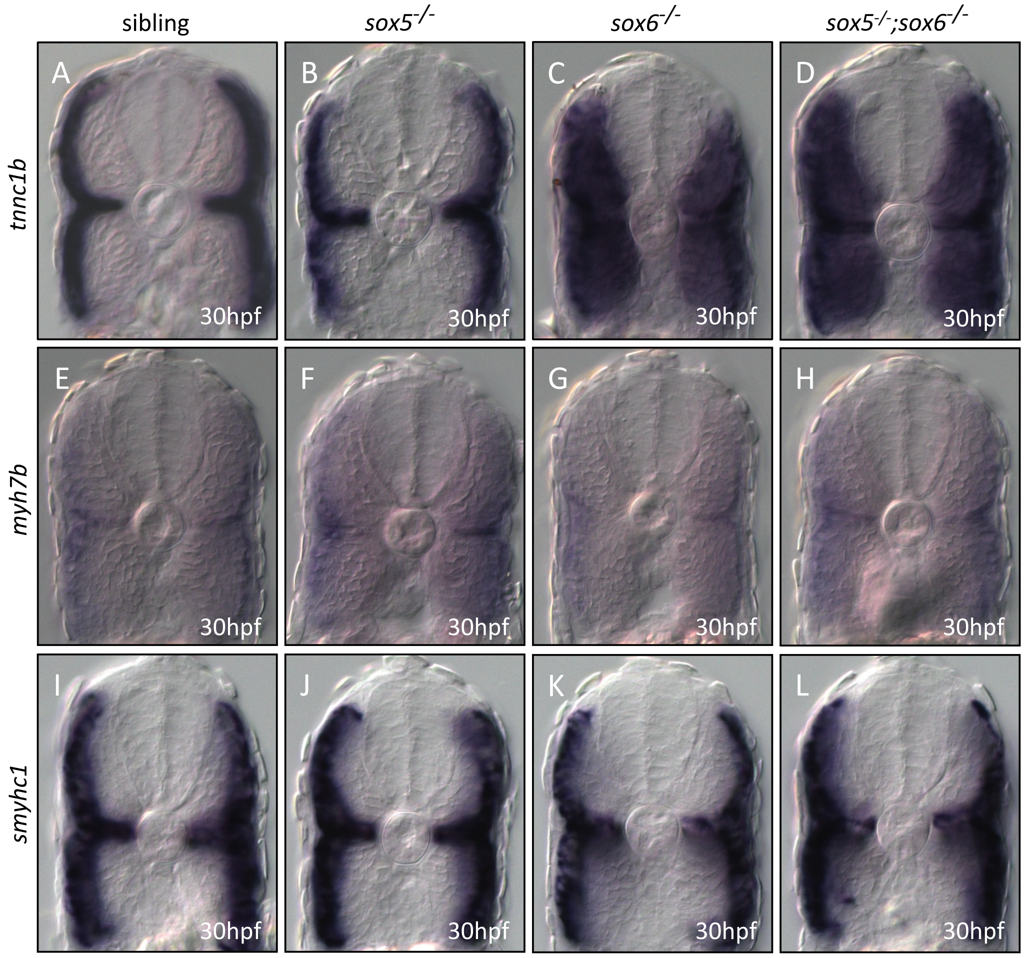


***Figure 5.29: The expression patterns of slow-twitch muscle-specific genes are the same in sox6*$^{-/-}$ *single mutant and sox5*$^{-/-}$*;sox6*$^{-/-}$ *double mutant embryos at 30 hpf.*** *tnnc1b is misexpressed throughout the whole myotome in sox5*$^{-/-}$*;sox6*$^{-/-}$ *double mutant embryos (D), just as is observed in single sox6*$^{-/-}$ *mutants (C). myh7b expression is restricted to the slow fibres in siblings (E), single sox5*$^{-/-}$*mutant (F), sox6*$^{-/-}$ *mutant (G) and the sox5*$^{-/-}$*;sox6*$^{-/-}$ *double mutant (H) embryos. The expression pattern of smyhc1 in sox5*$^{-/-}$*;sox6*$^{-/-}$ *double mutant embryos (L) is the same as that of single sox6*$^{-/-}$ *mutants (K), with just a few fast twitch fibres being smyhc1-positive.*

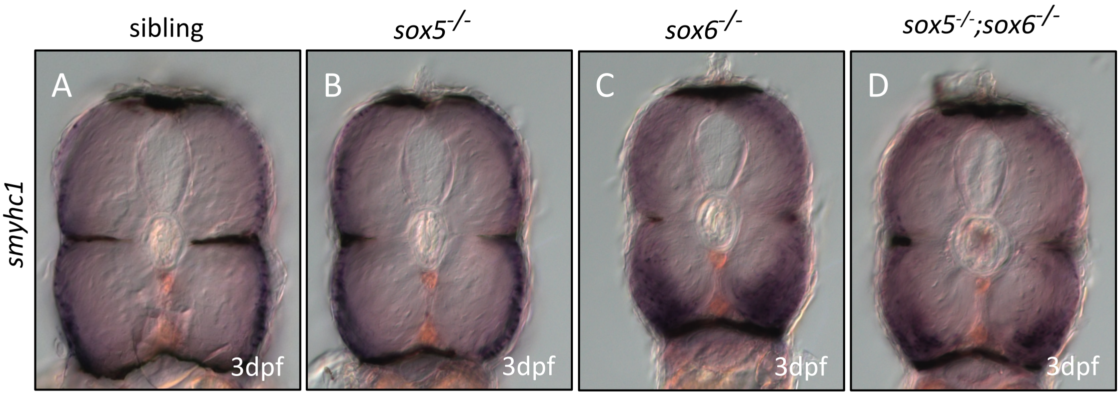


***Figure 5.30: Loss of sox5 function does not change the expression pattern of the slow-twitch muscle-specific marker smyhc1* in *3 dpf larvae.*** *In sox5$^{-/-}$ mutants, smyhc1 expression remains restricted to the slow fibres at 3 dpf (A and B), as observed in wild-type sibling larvae. In sox6$^{-/-}$ mutants the domain of smyhc1 expression is expanded to include the ventral-most fast-twitch fibres at 3 dpf. There is no difference in the expression pattern of this gene in sox6$^{-/-}$ single mutant and sox5$^{-/-}$;sox6$^{-/-}$ double mutant larvae.*

## 5.3 Discussion

The generation of a *sox6* null zebrafish mutant has allowed me to analyse more fully the role of *sox6* in zebrafish myogenesis. Additionally, it has uncovered an interesting adult phenotype associated with the loss of Sox6 function, which needs further elucidation.

The results described in the previous chapter (Chapter 4) have demonstrated that Sox6 is able to repress the expression of slow-twitch muscle genes, a function that is conserved in mice (An et al., 2011; Hagiwara et al., 2005; Hagiwara et al., 2007; Quiat et al., 2011). In *sox6$^{-/-}$* mutant mice, the expression of slow-twitch muscle genes is upregulated at the expense of fast-specific muscle genes, resulting in a slow-to-fast fibre type conversion (Hagiwara et al., 2005; Hagiwara et al., 2007). Although this is at least partly the case in zebrafish, there are some clear and important differences. The *sox6$^{-/-}$* zebrafish mutant has enabled me to elucidate some of these differences, as described in this chapter.

### 5.3.1 Fast-specific gene expression in *sox6$^{-/-}$* mutant zebrafish

One of the starkest differences between the zebrafish and mouse *sox6$^{-/-}$* mutants is the distinct responses of the fast-specific genes to the loss of *sox6* in these two species. In mice, the loss of *sox6* results in a decrease in the expression of several fast-specific genes (Hagiwara et al., 2005). In contrast, the levels of fast-twitch gene expression appear to remain unchanged in zebrafish *sox6$^{-/-}$* mutant embryos, although as in mice, slow-specific genes are upregulated in the zebrafish *sox6$^{-/-}$* mutant. Consequently, the phenotype of zebrafish *sox6$^{-/-}$* mutants is not a fast-twitch to slow-twitch fibre conversion, as seen in mouse. Rather, in zebrafish, both fast-twitch and slow-twitch genes are co-expressed in the same fibres. The decreased expression of fast-specific genes observed in mouse *sox6*$^{-/-}$ mutants could result from two possible mechanisms. Firstly, *sox6* may be necessary for the activation of fast-muscle genes. However, as no peaks of Sox6 binding were found in the promoters of fast-specific genes in mouse ChIP seq analyses (An et al., 2011), it seems likely that this might be achieved indirectly, via indirect activation of the fast-twitch programme. Secondly, the decrease in fast-twitch genes in *sox6*$^{-/-}$ mouse mutants could be a secondary effect due to the increased expression of slow-specific genes, as the activation of these genes may directly or indirectly repress the transcription of fast-twitch genes. In the zebrafish *sox6*$^{-/-}$ mutant embryos, my results indicate that there is no repression fast myosin isoforms, which indicates that *sox6* is not necessary for the activation of these genes. qPCR would be needed to confirm this result and the expression of other fast-specific genes needs to be examined. The reason why fast-myosin isoforms are downregulated in mouse *sox6*$^{-/-}$ mutants, but not in fish *sox6*$^{-/-}$ mutants, may be that a slow-specific gene that is upregulated in mouse, but not in fish, leads to the inhibition of fast-twitch-specific genes transcription. In zebrafish, fast-specific genes are known to be directly inhibited by Prdm1a in slow-twitch muscle (Liew et al., 2008; von Hofsten et al., 2008). Since Prdm1a is an essential gene for slow-twitch muscle specification, it may be that ectopic expression of Prdm1a is needed in the fast-twitch myotome of zebrafish *sox6*$^{-/-}$ mutants to fully convert the fast-twitch domain into a slow-twitch domain. *In situ* hybridization for *prdm1a* in zebrafish *sox6*$^{-/-}$ mutant and sibling embryos would help to determine whether *prdm1a* is ectopically expressed in the fast-twitch muscle of zebrafish *sox6*$^{-/-}$ mutants, although since Prdm1a normally directly represses fast-twitch genes, the expression of fast myosin genes in the fast-domain of *sox6*$^{-/-}$ mutants implies that there is no ectopic expression of *prdm1a.*

### 5.3.2 Slow-twitch muscle differentiation in *nrd*$^{-/-}$;*sox6*$^{-/-}$ double mutant embryos

Prdm1a mutants do not develop slow-twitch muscle, and instead the whole myotome differentiates into multinucleated, fast-twitch fibres (Roy et al., 2001). By contrast, I observed an expansion of the expression domain for a subset of slow-twitch specific genes, in *sox6*$^{-/-}$ mutant embryos. Previous reports have demonstrated that *sox6* morpholino knockdown of *sox6* in *prdm1a*$^{-/-}$ mutants can rescue the expression of slow-twitch muscle genes (von Hofsten et al., 2008). I sought to confirm these results, and to then extend them by determining whether a complete loss of *sox6* gene function in *prdm1a*$^{-/-}$ mutants could completely rescue slow-twitch muscle differentiation. Interestingly, *tnnc1b* expression was both rescued in the slow domain of *sox6*$^{-/-}$;*nrd*$^{-/-}$ double mutant embryos, and ectopically expressed in the fast domain (Figure 5.10). *smyhc1, tpm2* and *mylz10* expression was also rescued in *sox6*$^{-/-}$;*nrd*$^{-/-}$ double mutant embryos, although the ectopic expression of these genes was limited to a few fibres, mainly in the fast domain, and does not extend to all fibres of the myotome (Figure 5.10). However, the slow-twitch differentiation programme is not completely rescued, as slow-twitch muscle fibres fail to migrate in *sox6*$^{-/-}$*;nrd*$^{-/-}$*;ubo:GFP* embryos (Figure 5.11). This does not seem to be the case in *sox6*$^{-/-}$*;prdm1a morphant* embryos, as the rescue in the expression of *smyhc1* was observed in the outer fibres of the myotome (Figure 5.9), most likely due to the incomplete knockdown of *prdm1a* by the morpholino.

### 5.3.3 Expression of slow-twitch-specific genes in the zebrafish *sox6*$^{-/-}$ mutant

Previous studies using a *sox6* morpholino concluded that the knockdown of *sox6* results in no ectopic expression of *smyhc1* in fast-twitch fibres (von Hofsten et al., 2008). However, the complete mutagenic inactivation of *sox6* through the use of zinc-finger nucleases revealed that a small subset of fast fibres do express *smyhc1* in response to the loss of *sox6.* This observation corresponds with what is seen in the mouse, where there is an increased expression of *smyhc* transcripts in *sox6*$^{-/-}$ mutant foetal fibres. It seems to be that by 3 dpf there is a least some ectopic expression of all slow muscle genes examined in the fast-twitch muscle. It is not known why there would be a delay in the ectopic expression of a subset of slow-specific genes in *sox6* mutants. One possibility is that the presence of maternal transcripts of *sox6* present in the embryo at these early stages might repress slow-twitch genes in *sox6* mutants. However, this hypothesis is likely to be incorrect, since *in situ* hybridization analysis at the 2-cell stage revealed that *sox6* is not maternally expressed. Why there is this delay in the expression of a subset of slow-twitch genes in the fast muscle of *sox6*$^{-/-}$ mutants remains an open question.

Unlike *tnnc1b, tnnt1* and *ryr1a*, which are expressed throughout fast domain in *sox6*$^{-/-}$ mutants, other slow-twitch genes such as *smyhc1* and *tpm2* are only ectopically expressed in a subset of fast twitch fibres in homozygous *sox6* mutants, raising the question of why some slow-specific genes appear to be stringently repressed by *sox6,* whereas other do not. Notably the ectopic expression of *smyhc1, tpm2* and *mylz10* at 3dpf and 4dppf appears to be restricted to the dorsal and ventral edges of the myotome. Possible explanations for this could be the presence of an inhibitory signal emanating from the midline, restricting the ectopic expression of these genes to the outer edges. Another possible explanation could be that these cells are newly formed fibres and therefore could express slow-specific genes when they first differentiate but this expression is lost when they mature. This would need to be investigated further, perhaps by observing areas where cell division is occurring, by using Edu staining.

With the recent completion of the zebrafish genome, it has been suggested that one-quarter of zebrafish genes have sister genes with high sequence similarity (Howe et al. 2013), therefore it is possible that a gene that is similar in sequence to *sox6* could compensate for loss of *sox6* in *sox6* null mutant embryos, and maintain repression of slow muscle genes such as *smyhc1* and *tpm2* in the majority of fast-twitch fibres. This possibility is discussed below.

### 5.3.4 Six1a and Pbx transcription factors

The possibility that transcription factors required for the activation of fast-specific genes could also be repressing slow-twitch gene expression in the fast muscle of *sox6*$^{-/-}$mutants, was examined. As discussed in Chapter 1, both Six1a and Pbx homeodomain transcription factors have previously been implicated in the transcriptional activation of fast-specific genes in zebrafish (Maves et al., 2007). The examples of interactions between homeodomain transcription factors and HMG domain transcription factors in regulating transcription (Dailey and Basilico, 2001) led to the hypothesis that Six proteins or Pbx proteins may be cooperating with Sox6 during myogenesis to repress the expression of slow-specific genes in the fast muscle. However, the knockdown of Six proteins or Pbx proteins in the fast-twitch myoblasts of *sox6*$^{-/-}$ mutant embryos, led to no increase in the expression of the *smyhc1:GFP* transgene in the fast-twitch fibres, as compared to the situation for sibling embryos.

### 5.3.5 *sox5*$^{-/-}$ and *sox6*$^{-/-}$ zebrafish mutants

Sox6 is known to have partially redundant functions with Sox5 in processes such as chondrogenesis, due to their high sequence similarity and co-expression in cartilage precursors (Lefebvre et al., 1998, Smits et al., 2001). As *sox5* is also expressed weakly in the somites during development, it is plausible that it may be weakly repressing a subset of slow-twitch muscle genes. The results described in Chapter 3 indicate that *tnnc1b* is not directly repressed by Sox6, therefore it could be that such genes which are indirectly repressed by Sox6, are expressed throughout the whole myotome in *sox6*$^{-/-}$ mutants, whereas genes that are directly repressed such as smyhc1, are expressed only in a subset of fibres as another repressor such as *sox5*, can weakly directly repress the expression of these genes. In order to test this hypothesis, a Zinc-finger nuclease mutant for the *sox5* allele was generated. However, double *sox5*$^{-/-}$;*sox6*$^{-/-}$ zebrafish mutant embryos displayed no increase in the number of *smyhc1*-expressing fast-twitch fibres, than the single *sox6*$^{-/-}$ mutant alone. This result indicates that *sox5* is not compensating for the loss of *sox6* in *sox6*$^{-/-}$ mutants by repressing the expression of slow-twitch genes in the fast-twitch domain.

### 5.3.6 The possibility of a second *sox6* gene

The *sox6* gene is duplicated in most teleost genomes that have been sequenced to date (Alfaqih et al., 2009, Cui et al., 2011, Koopman et al., 2004). The existence of two copies of *sox6* in zebrafish could provide an explanation for the partial derepression of the slow-twitch pathway that was observed in zebrafish *sox6* homozygous mutant embryos. Notably, however, the zebrafish genome seems to be exceptional in lacking a *sox6* duplicate (Zebrafish Zv9 assembly http://www.ensembl.org/index.html). The presence of a second *sox6* gene has been searched for extensively using bioinformatic tools to analyse RNA-seq data and no such duplicate has been identified. Double-conserved synteny (DCS) blocks are runs of genes in the ancestral species lacking a whole genome duplication, that are found on two different chromosomes in the species where whole genome duplication did occur (Kellis et al., 2004). Interestingly, DCS blocks can be observed at the *sox6* locus in teleost species and humans, yet a duplicate *sox6* gene that is observed in other teleost species, appears to be missing at the predicted location in the zebrafish (Appendix A4).

Since the recent publication of the complete zebrafish genome sequence (Howe et al., 2013, Kettleborough et al., 2013), a second *sox6* gene has still not been identified, leaving open the possibility that a distinct repressor is involved in restricting the expression of the *smych1* and other slow-specific genes to the adaxial cells during zebrafish myogenesis.

### 5.3.7 Repression of muscle genes by microRNAs

Myosins have been shown to not only encode the major contractile units of muscle, but they can also influence muscle gene expression by encoding intronic microRNAs, which inhibit the expression of other genes. For example, in zebrafish and mice, *myh7b* encodes miR-499, which inhibits the expression of *sox6,* which in turn inhibits the expression of *myh7b,* forming a negative feedback loop (McCarthy et al., 2009; van Rooij et al., 2009; Wang et al., 2011). Further to this, numerous mouse myosin genes encode related microRNAs within their introns, which regulate myosin expression in both skeletal and heart muscle, and which in turn can lead to changes in fibre-type and therefore muscle performance (Rooij et al., 2009). A similar network of myosin gene-associated miRNAs may also exist in zebrafish. Since the expression of fast-specific myosins are not downregulated in *sox6*$^{-/-}$ mutants it is possible that a microRNA encoded by one of these fast-specific myosins specifically inhibits the expression of *smyhc1* and possibly other slow-specific muscle genes in *sox6*$^{-/-}$ mutants. Conversely, genes such as *tnnc1b* and *ryr1a* could be insensitive to such myosin-encoded miRNAs. To test this hypothesis fast-myosin heavy chain and light chain isoforms would need to be knocked out in the *sox6*$^{-/-}$ mutant background. Any miRNAs encoded in the introns of the fast myosins would therefore not be expressed, and if the hypothesis is correct, there would be no inhibitory effect on slow specific genes by these miRNAs, leading to ectopic expression of all slow-specific genes throughout the myotome.

### 5.3.8 Sox6-independent regulators of slow-twitch muscle fate

The data therefore provides no explanation as to why a subset of slow-muscle genes are ectopically expressed in only a subset of fast-twitch fibres in *sox6*$^{-/-}$ mutants, rather than throughout the whole myotome as was observed for *tnnc1b* expression. This analysis has focused on the assumption that a second repressor of slow-twitch muscle genes prevents the expression a subset of slow-specific genes in the fast-twitch muscle of *sox6*$^{-/-}$ mutants. It may be the case that another gene required for the transcriptional activation of this particular subset of slow-specific genes is not expressed in the fast-twitch domain. However, the rescue of *smyhc1* expression in *nrd*$^{-/-}$;*sox6*$^{-/-}$ mutants suggests that this activator is probably not under the control of Prdm1. An interesting experiment would be to analyse the expression of *smyhc1* in Cyclopamine treated *sox6*$^{-/-}$ mutants. My results have revealed that *tnnc1b* expression can be rescued in Cyclopamine treated *sox6*$^{-/-}$ mutants, indicating that *tnnc1b* can be activated independently of Hh signalling. If *smyhc1* requires an Hh-dependent activator for its expression, then it its expression would not be rescued in Cyclopamine treated *sox6*$^{-/-}$ mutants, indicating two different modes of activation for *tnnc1b* and *smyhc1*. As Prdm1a is thought to act as a switch that activates the slow-twitch differentiation programme in zebrafish, it would be useful to look at differentially expressed genes in wild-type and *prdm1a*$^{-/-}$ mutant embryos, to identify new candidate genes that may be involved in regulating slow-twitch myoblast differentiation.

### 5.3.9 The calcium link

Intriguingly, the three slow-specific genes that are strongly expressed throughout the fast domain in *sox6*$^{-/-}$ mutants at 30 hpf are all genes that regulate calcium signalling in the sarcomere to control muscle contraction. Both Tnnc1b, and Tnnt1 are part of the troponin complex, which regulates actin and myosin interactions in response to changes in intracellular calcium levels (Lehman et al., 1994). Ryr1a*,* is a slow-specific ryanodine receptor that acts as an intracellular calcium channel to mediate the release of calcium ions from the sarcoplasmic reticulum (Meissner, 1994). The alternative combinations of both myosin heavy chain isoforms and troponin T, I and C isoforms regulate the force of contraction and myofibre calcium sensitivity, therefore changes in the isoforms of these genes that a particular myofibre expresses will change the contraction properties (Bottinelli et al., 1991, Geiger et al., 1999). Preliminary experiments in the lab have used a GCaMP calcium reporter gene under the control of the *mylz2* promoter to compare the calcium response in the fast-twitch fibres of sox6$^{-/-}$ mutant fish compared with the fast-twitch fibres of sibling fish (Yosuke Ono, unpublished data). The results of these experiments suggest that *sox6*$^{-/-}$ mutant fast-twitch fibres are able to respond more quickly than sibling fast-twitch fibres. We can speculate that this distinct responsiveness is due to an increase in the number of calcium responding subunits in *sox6*$^{-/-}$ mutant fast-twitch fibres. For example, if a cell expresses both the fast- and slow-specific ryanodine calcium channel, as well as both the fast- and slow- specific troponin c isoforms, then upon depolarization of the muscle cell membrane, both ryanodine receptor isoforms would be able to release calcium from the sarcoplasmic reticulum, and both troponin c isoforms would be able to bind calcium to allow the necessary interactions between actin and myosin, which could result in a faster response time. In order to test this theory, morpholino knockdown of both *ryr1b* and *tnnc1b* in *sox6*$^{-/-}$ mutants could be performed to see if the response time of the fast-twitch muscle returns to a wild-type response time.

Interestingly, Quiat et al. recently reported that *Sox6* mutant mice exhibit increased expression of a gene involved in regulating the calcium response, the calcium uptake channel ATP2a2 (Quiat et al., 2011). This observation indicates that in both zebrafish and mice, Sox6 regulates not only the expression of sarcomeric components of the muscle fibres, but also the calcium-regulatory components. Further to this, mice with conditional loss of *Sox6* in their myofibres have higher endurance level and elevated mitochondrial activity in their muscles, as compared with sibling mice (Quiat et al., 2011). This abnormality is likely due to the increased expression of slow-contractile and calcium-regulatory genes in *Sox6* mutant muscles. It would be interesting to explore whether this is also the case in zebrafish.

### 5.3.10 The phenotype of the *sox6*$^{-/-}$ mutant myotome in adult fish

Sections taken from the trunk of *sox6*$^{-/-}$ mutant and sibling adults revealed that the ectopic expression of both *tnnc1b:eGFP* and *smyhc1:GFP* that was observed in the fast-twitch muscle of mutant embryos also persists in mutant adult fish. Expression of the *tnnc1b:eGFP* transgene is observed throughout the myotome of *sox6*$^{-/-}$ mutant adults, whereas there is only a small expansion of the *smyhc1:GFP* domain in mutant adults, as was also observed in *sox6*$^{-/-}$ mutant embryos, with only a small subset of fast fibres expressing the *smhyc1:GFP* transgene. It also appears that other *smyhc* isoforms are derepressed in a subset of fast-fibres, as the *smyhc in situ* hybridization probe, which recognizes all isoforms of *smyhc,* labels an expression domain larger than that of *smyhc1:GFP* in *sox6*$^{-/-}$ mutants. An intriguing observation common to both the *tnnc1b:eGFP* and the *smyhc* probe expression patterns, is that the wedge-shaped *tnnc1b* and *smyhc* expression domain observed in wild-type adults is still present in *sox6*$^{-/-}$ mutants but it has expanded more medially. Although the expression of *tnnc1b* is seen throughout the muscle of *sox6*$^{-/-}$ mutants, *tnnc1b:eGFP* expression is much stronger in the lateralis superficialis, which appears to have expanded, with a similar change also apparent for the domain of *smyhc* expression. From looking at the differences in the size of the myotome between siblings and *sox6*$^{-/-}$ mutants it may be that this is not an expansion of the lateralis superficialis, but rather a lack of growth in the fast domain, resulting in the medial expression of *smyhc1* and *tnnc1b.* As the expression of *tnnc1b* was only examined using the *tnnc1b* transgenic reporter line, *in situ* hybridisation would need to be performed using the *tnnc1b* probe to determine whether the reporter line is definitely mimicking the endogenous gene expression.

The sustained ectopic expression of slow-twitch muscle genes seen in zebrafish *sox6*$^{-/-}$ mutants is consistent with what is seen in adult mice, since mutant mice that lack *sox6* function in their muscles exhibit an increase in slow-specific genes expression (Quiat et al., 2011).

### 5.3.11 The *sox6*$^{-/-}$ mutant scoliosis adult phenotype

An unexpected phenotype that occurs as a result of the loss of *sox6* function in zebrafish is scoliosis. This phenotype has not been reported in mouse, likely due to the fact that mouse *sox6* mutants die within two weeks of birth (Hagiwara et al., 2001), by which time the mutant mice might still be too young for the scoliosis phenotype to be apparent. Scoliosis is a tree dimensional curvature of the spine and can be classified into three major types: congenital, syndromic, and idiopathic (Altaf et al., 2013). Congenital scoliosis refers to spinal deformities that are present at birth resulting from abnormally-formed vertebrae (Hensinger, 2009). Sox6 and Sox5 function together during chondrogensis, and knockdown of both genes leads to gross loss of cartilage, yet knockdown of *sox6* or *sox5* individually leads to minor cartilage defects (Lefebvre et al., 1998, Smits et al., 2001). The vertebral column is an endochondral bone i.e. bone formed is by the ossification of cartilage (Karsenty and Wagner, 2002), therefore it is possible that a defect in cartilage formation causes abnormal morphogenesis of the vertebral column, leading to scoliosis. However, since the scoliosis observed in *sox6*$^{-/-}$ mutants is not apparent until at least one-month postfertilisation, and the formation of the cartilage appears normal at a younger stages, I would suggest that this is not the form of scoliosis observed in *sox6*$^{-/-}$ mutants, although it cannot be ruled out. The majority of scoliosis observed in humans is idiopathic, meaning that there is no identifiable cause, with the majority of cases not developing until adolescence or adulthood (Lonstein, 1994). The final form of scoliosis, syndromic, arises from a secondary effect on the skeleton that most commonly results from neuromuscular conditions, such as muscular dystrophy (Galasko et al., 1995). The muscle wasting which results from the progression of these diseases often leads to scoliosis because the muscles on either side of the spinal column are weak and therefore unable to maintain the straightness of the vertebral column. Both idiopathic and syndromic scoliosis tend to have a later onset than the congenital form, with curvature of the spine developing during adolescence or adulthood (Galasko et al., 1995; Lonstein et al., 1994). On this basis, I hypothesise that that the scoliosis observed in the *sox6*$^{-/-}$ mutant would most closely resemble these forms, and more phenotypic analysis is now needed to test this possibility.

The scoliosis observed in zebrafish *sox6* mutant adults could arise from either a primary effect due to the loss of *sox6* function in the bone, or as an indirect, secondary consequence of loss of *sox6* function in the muscle. There are several ways in which these two possibilities could be investigated. First, it is important to determine whether *sox6* is expressed in the vertebral column. Preliminary data has suggested that *sox6* is not expressed in the bone, because no *sox6* transcripts could be identified in the zebrafish opercle at 4 dpf. However, *sox6* may be expressed later on during bone development, and this expression may also be specific to the vertebral column. This issue can be addressed by examining the *sox6:eGFP* line, to search for GFP expression in the vertebrae at various stages of development.

If *sox6* is not expressed in the vertebral column, it seems likely that the scoliosis observed in *sox6* mutants could result from the loss of *sox6* in the myotome, and the altered muscle morphogenesis and contractility could be responsible for distorting the vertebral column. In order to test this hypothesis, conditional knockouts of *sox6* could be generated, which specifically knock out the expression of *sox6* in the bone or the muscle. Through the use of transcription activator-like effector nucleases (TALENs) the *sox6* gene can be precisely modified by the addition of *loxP* sites, and a tissue specific promoter can be used to drive the expression of Cre recombinase in either the bone or the muscle generating a conditional knockout of the *sox6* gene (Bedell et al., 2012). If the scoliosis phenotype still developed in fish in which *sox6* expression is knocked out in bone, then the scoliosis could be due to the loss of *sox6* protein in the muscle, which would be confirmed if knockout of *sox6* in the muscle prevented scoliosis. An alternative and perhaps simpler way to resolve this issue would be to rescue the expression of *sox6,* specifically in the muscle of *sox6*$^{-/-}$ mutants. This could be achieved by intercrossing *sox6* heterozygous mutants carrying the *UAS:sox6-GFP;actin:GAL4* transgene. This would drive the expression of *sox6* specifically in the muscle, rescuing its expression in the *sox6* null mutant. If no *sox6*$^{-/-}$;*UAS:sox6-GFP:actin:GAL4* mutant fish developed scoliosis, this would verify that the scoliosis results from a loss of *sox6* specifically in the myotome. As there is no scoliosis phenotype reported in mice with a conditional knockout of *sox6* in the muscle (An et al., 2011; Quiat et al., 2011), this could lead us to believe that the scoliosis phenotype observed in fish is not due to the loss of *sox6* function in the muscle. However this would need to be confirmed by performing the experiments suggested about.

If the scoliosis does result from a loss of *sox6* expression in the myotome, then what is the mechanism behind the scoliosis phenotype? I hypothesize that two distinct mechanisms may be involved, both based on the fact that *sox6*$^{-/-}$ mutant zebrafish are smaller than sibling fish. The cross-sections of trunk tissue taken from adult *sox6*$^{-/-}$ mutant zebrafish, indicate that the *sox6* mutant fibres are significantly smaller than their sibling counterparts, likely resulting in the significant size difference that was observed between *sox6*$^{-/-}$ mutant and sibling fish. Therefore, one hypothesis as to why a loss of *sox6* expression in the muscle results in scoliosis, would be that *sox6* mutant fibres fail to grow at a normal rate, or to a normal size, as compared to sibling fish. However, the vertebral column and other skeletal elements would grow at a normal speed and to a normal size, and with the muscle fibres being unable to keep up with this growth, there would be a crumpling of the spine. If further studies support this idea, then the mechanism of action of *sox6* in growth control would need to be elucidated.

A second hypothesis regarding the mechanism underlying the scoliosis phenotype, which could result from the loss of sox6 in zebrafish muscle, is again based on the smaller size of the *sox6*$^{-/-}$ mutant fibres. The loss *sox6* could, through mechanisms unknown, result in the degeneration of muscle fibres, resulting in the activation of satellite cells and the regeneration of muscle fibres. Satellite cells are normally quiescent in adult muscle, but are able to proliferate in response to degeneration, to give rise to regenerated muscle (Mauro, 1961, Schultz et al., 1978, Seger et al., 2011). The constant degeneration and regeneration of muscle tissue in these mutants would mean that the myotome is not able to grow at the same rate as the spinal column and therefore the spine would crumple. One clue as to why this may be the case is the number of nuclei observed in *sox6*$^{-/-}$ mutant fibres. There appears to be many more nuclei, in both the slow- and fast-twitch fibres of *sox6*$^{-/-}$ mutants, which might reflect increased proliferation driving regeneration. Alternatively, due to the much smaller size of the *sox6*$^{-/-}$ mutant myotome, it may be that there are the same number of nuclei in both the mutant and sibling myotome, but as the mutant myotome is much smaller, the nuclei are closer together and therefore appear more abundant. Nuclei counts would need to be done to confirm which of these hypotheses are correct. To address whether *sox6*$^{-/-}$ mutant fibres are degenerating, leading to an increase in regeneration, a number of experiments could be performed. Firstly, to look for evidence of degeneration, Evans blue dye could be injected into the heart of adult zebrafish. Evans blue dye penetrates degenerating or damaged muscle cells, and can be observed by red auto-fluorescence by fluorescence microscopy (Hamer et al., 2002). In order to examine whether there is an increase in regeneration in the myotome of *sox6*$^{-/-}$ mutants compared to siblings, Edu staining could be utilized to assess the incorporation of these nucleotide analogues into the newly synthesized DNA of replicating cells (Salic and Mitchison, 2008). Additionally, the expression of Pax7, a marker of muscle satellite cells (Seale et al., 2000, Seger et al., 2011), could be examined to determine whether there is an increase in regeneration. An increase in the number of satellite cells, as indicated by an increase in the number of Pax7-expressing cells, would reveal an increase in regeneration.

Another interesting possibility as to why knockdown of *sox6* in zebrafish results in scoliosis is related to the role of *sox5* and *sox6* in notochord formation. Knockout of both *sox5* and *sox6* in mice results in a loss of cells in the notochord (Smits and Lefebvre, 2003), with the consequence of this decline in notochord cell number often leading to scoliosis (Sivakamasundari and Lufkin, 2012). Although there appears to be no effect on notochord formation in single *sox5* and *sox6* mutant mice in zebrafish it may be the case that the loss of *sox6* alone is enough to decrease the number of notochord cells, which could lead to the observed scoliosis phenotype. More analysis is needed to determine if this hypothesis is correct.

Unexpectedly, the slow-twitch muscle fibres as well as the fast-twitch fibres were smaller in *sox6*$^{-/-}$ mutant adults than in siblings. Moreover, the slow-twitch fibres were smaller in size to their sibling counterparts by a larger degree than were the fast-twitch fibres. One explanation that could account for the more pronounced effect on the size of the slow-twitch fibres in *sox6*$^{-/-}$ mutants is that *sox6* expression may no longer be restricted to the fast-twitch domain in juvenile and adult fish. It may be that at some point after development and specification of the myotome, *sox6* has a distinct role in the zebrafish myotome in either growth or maintenance of both slow- and fast-twitch muscle fibres, so its expression would be required continuously throughout the myotome and into adulthood. This would need to be investigated by performing a *sox6 in situ* hybridisation on adult fish. Alternatively, if the degeneration hypothesis is correct, the loss of *sox6* function in the fast-fibres might cause them to degenerate and an excess of slow fibres might then regenerate in order to compensate for the loss of fast-twitch fibres.

The scoliosis and muscle phenotype observed in the *sox6*$^{-/-}$ mutant adults was always observed in the *sox6* homozygous mutants and never observed in the siblings. Nevertheless, it could be the case that the zinc-finger nuclease caused an off-target mutation in a gene other than *sox6* that is linked to the *sox6* locus therefore segregates with it. In order to confirm that the adult scoliosis and muscle phenotype results from the loss of Sox6 protein it would be a good idea to cross the E2 allele to the A1 allele to make a double E2:A1 heterozygous mutant. It would be highly unlikely that the zinc-finger nuclease created the same off-target mutation in both alleles therefore the presence of a scoliosis phenotype in these double heterozygous mutant fish would likely be the result of a loss of Sox6. Secondly, the UAS:Sox6-GFP;actin:GAL4 line can be used to rescue the expression of Sox6 in the muscle and used to identify whether the adult *sox6*$^{-/-}$ muscle phenotype still develops.

### 5.3.12 Future uses of the *sox6*$^{-/-}$*;sox5*$^{-/-}$ zebrafish mutant

The zebrafish *sox6*$^{-/-}$*;sox5*$^{-/-}$ double mutant had an observable, morphological phenotype at 3 dpf. Although no morphological phenotype is observed in single *sox5*$^{-/-}$ or *sox6*$^{-/-}$ mutant embryos at this stage, double mutants lack a lower jaw, have a smaller head with oedema around the otic vesicles and their pectoral fins protrude at an abnormal angle. It is suspected that Sox5 and Sox6 have a number of unknown functions during vertebrate development due to their expression in a wide variety of tissues, at many different developmental time points (Figure 5.1 and Figure 5.26). Further comparative analysis of the *sox5* and *sox6* single and double mutants should help us distinguish between the overlapping and distinct roles these proteins play during embryogenesis, and to identify new functions for these genes that have not been discovered in other vertebrates.

# Chapter 6

# Sox6 Chromatin Immunoprecipitation

### 6.1 Introduction

The results described in Chapter 4 of this thesis demonstrated that Sox6 is able to inhibit the expression of slow-twitch muscle genes when Sox-GFP is misexpressed in slow-twitch fibres. This raised the question of whether this inhibition is direct, via interactions with *cis*-regulatory elements of these genes, or indirect, through interactions with *cis*-regulatory elements of other genes. In mice, ChIP-seq analysis has demonstrated that many slow-specific muscle genes are direct targets of Sox6 (An et al., 2001), but it is still not clear whether this is also the case in zebrafish. Previous studies in the laboratory revealed that mutation of the potential Sox6 binding sites in the promoter of the zebrafish *smyhc1:GFP* reporter gene resulted in ectopic GFP expression in a number of fast-twitch muscle fibres (Chapter 5, Figure 5.8), implying Sox6 normally acts to directly repress *smyhc1* expression in the fast-twitch muscle. In contrast to this, mutation of potential Sox6 binding sites in the *cis*-regulatory regions of the *tnnc1b:eGFP* transgene, had no effect on the expression of this transgene, with eGFP expression remaining restricted to the slow-twitch fibres (Chapter 3, Fig.3.8). These observations suggest that Sox6 represses *tnnc1b* expression either indirectly or alternatively through direct inhibition by binding to an as-yet unidentified, cryptic Sox6 binding site.

Chromatin immunoprecipitation (ChIP) is a technique that can be used to identify the *in vivo* DNA sequence targets of a known transcription factor or other chromatin protein, if an antibody exists that is specific for the protein of interest and which interacts with that protein in its native conformation within chromatin. I therefore sought to use ChIP to determine whether *cis*-regulatory sequences from slow-specific genes are enriched in DNA fragments immunoprecipitated from chromatin using a Sox6-specific antibody.

### 6.2 Results

#### 6.2.1 Identification of a suitable sox6 antibody

In order to perform this experiment, a Sox6-specific antibody that was able to recognise and pull-down Sox6 protein from chromatin extracts, needed to be identified. Two commercially available Sox6 antibodies, ab66316 (Abcam) and GTX116236 (Gentex), were purchased and tested for their specificity by performing Western blotting and whole-mount immunohistochemistry analysis of zebrafish embryos. The molecular weight of Sox6 protein is expected to be approximately 87kDa, but Western blot analysis using these antibodies was unable to identify a specific band of protein of the appropriate molecular weight (Figure 6.1). Furthermore, whole-mount immunohistochemistry using these Sox6 antibodies revealed no specific staining in zebrafish embryos at 30 hpf, confirming that neither of these two commercially supplied antibodies was suitable for use in ChIP experiments.

The ab6316 and the GTX116236 antibodies had been raised with immunogens corresponding to sequences present in the human and mouse Sox6 proteins, therefore it is possible that the antibodies failed in my experiments because their epitopes were not present in the zebrafish Sox6 protein. Therefore, a zebrafish-specific Sox6 protein expression vector was constructed and an antibody was raised by immunising rabbits with recombinant zebrafish Sox6 protein. A fragment of zebrafish *sox6* cDNA was cloned into the His-tag expression vector pET-2lb (by Stone Elworthy), giving rise to the plasmid construct Sox6M. The *sox6* cDNA fragment incorporated into this plasmid coded for 219 amino acids of zebrafish Sox6 protein immediately upstream of the HMG box motif (Appendix A5). Using the His-tagged Sox6 fusion construct, two antibody preparations were obtained from the company Sdix, Sdix-SoxA and Sdix-Sox6B. The two antibodies were generated from the same antigen but in two different rabbits. Neither of these two antibody preparations were able to detect Sox6 protein, because Western blot analysis with each antibody failed to detect a specific protein band (Figure 6.1), and whole-mount immunohistochemistry staining with each antibody gave only diffuse and non-specific staining.

Finally, another Sox6 antibody was generated using the Sox6M expression vector by the company Absea. This antibody, named Sox6M1, specifically labelled the myotome of 30 dpf embryos, as assessed by whole-mount immunohistochemistry, with expression restricted to the fast-muscle nuclei (Figure 6.2A). Furthermore this specific signal was absent in the nuclei of both fast-twitch muscle cells and in the optic tectum of homozygous *sox6* mutants (Chapter 5, Figure 5.2). These results indicated that Sox6M1 antibody might be potentially useful for ChIP experiments. However, a Western blot analysis of protein extracts from *sox6*$^{-/-}$ mutant and *sox6* sibling embryos revealed that the antibody recognises many non-specific epitopes and although there is a protein band of the correct size (~87kD), this band was present in extracts from both *sox6*$^{-/-}$ mutant and sibling embryos (Figure 6.2B). Taken together, these immunohistochemistry and Western blotting results indicate that the Sox6M1 antibody preparation contains a component that recognises Sox6 protein, but there are other components that recognise additional proteins, including some that are of comparable molecular weight to that of Sox6. Notwithstanding these limitations, immunoprecipitation was performed with the Absea Sox6M1 antibody on chromatin samples prepared for ChIP analysis. A band at the correct size for Sox6 protein was observed (Figure 6.2C), but due to the previous Western blot analysis and the high level of non-specific bands, an alternative approach immunoprecipitating Sox6 protein was developed.

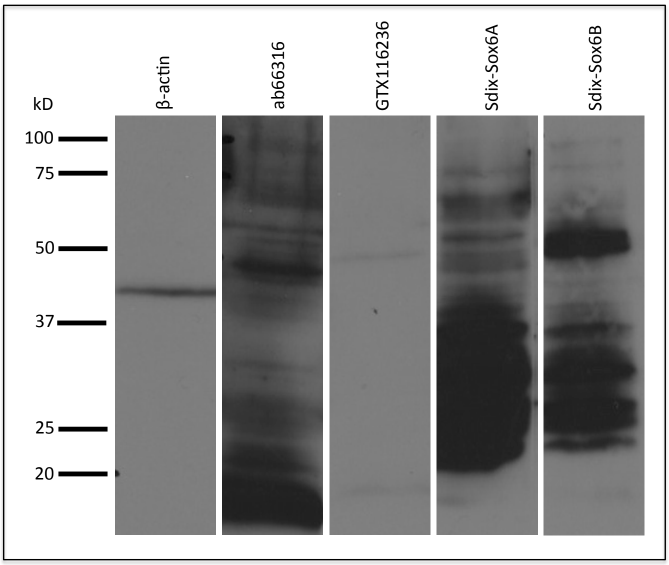


***Figure 6.1: Western blot analysis reveals that neither of the commercially available Sox6 antibodies, ab66316 and GTX116236, specifically recognised Sox6 protein in wild-type embryos at 30 hpf.*** *Moreover, two custom-made antibodies, Sdix-Sox6A and Sdix-SoxB, raised against a zebrafish Sox6 epitope, did not exhibit specificity for Sox6 protein in 30 hpf wild-type embryos.*

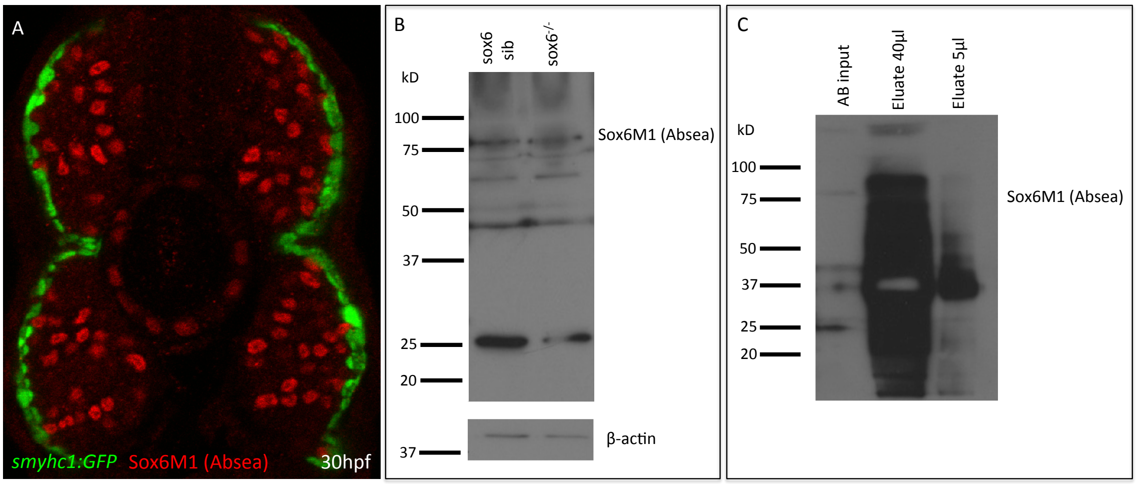


***Figure 6.2: Analysis of the Absea Sox6M1 antibody.*** *The Sox6 antibody, Sox6M1, specifically labels fast-twitch muscle nuclei (red), with no expression seen in the slow-twitch muscle at 30 hpf, identified by expression of smyhc1:GFP (green) (A). Western blot analysis using the Sox6M1 antibody (B) revealed a band at the approximate size of Sox6 (87kDa) in sox6 sibling embryos at 30 hpf. Unfortunately this band was also present in* $sox6^{-/-}$ *mutant embryos, indicating that the observed band is probably not Sox6. Crosslinked, sonicated chromatin was prepared from 30 hpf wild-type zebrafish embryos, and then incubated with the Sox6M1 antibody. Immune complexes were precipitated with Protein G-beads and immunoprecipitated proteins were analysed by SDS-PAGE and Western blotting with the anti-Sox6 antibody, with an equivalent sample of input chromatin run along side. The input fraction revealed a band at the appropriate size for Sox6 proteins, while 40µl of eluate revealed multiple bands with one band at the correct size for Sox6, whereas 5µl of eluate revealed no Sox6 specific band (C).*

### 6.2.2 Pull-down of Sox-GFP using the GFP-Trap

The GFP-Trap (ChromoTek) is a GFP-binding protein coupled to agarose beads that can be used to pull-down GFP fusion proteins (Rothbauer et al., 2008). The *UAS:sox6-GFP;actin:GAL4* line expresses Sox6 fused to GFP, under the control of an *actin* promoter. The results described in Chapter 4 of this thesis demonstrated that this Sox6-GFP protein is functional, as it is able to inhibit the expression of slow-twitch muscle genes. The *actin:GAL4* line was used instead of the *smyhc1:GAL4* line in order to express Sox6-GFP throughout the myotome, including the fast-twitch muscle, where the endogenous *sox6* gene is expressed. Proteins were extracted from *UAS:sox6-GFP;actin:GAL4* embryos at 30 hpf and extracts were then probed with an anti-GFP antibody in a Western blot analysis, which identified several weak bands of protein and one strong protein band (Figure 6.3A). The strong band was thought to likely correspond to the Sox6-GFP fusion protein, although it was slightly larger than the predicted size of 117kDa (87kDa- Sox6 plus 30kDa-eGFP), which could be due to post-translational modifications. The GFP-Trap was used to selectively purify GFP fusion protein-cross-linked chromatin, from *UAS:sox6-GFP;actin:GAL4* embryos at 30 hpf. An extract was also made from wild-type AB embryos and used as a negative control, as these embryos express no GFP protein, and therefore the GFP-Trap was expected to pull down no proteins (or very few) from this sample. The resulting Western blot from this experiment identified two weak bands at 37kDa and 140kDa in the AB input control lane suggesting that the GFP-trap does not specifically recognise GFP proteins under these conditions (Figure 6.3B). The GFP-Trap pull-down on *UAS:sox6-GFP;actin:GAL4* embryos included a protein at the correct size that could be seen in the eluate fraction (indicated by Sox6-GFP label in Figure 6.3B), beneath which was a ladder of smaller protein bands, which could represent degradation products of the full-length Sox6-GFP fusion protein. As might be expected, no full-length fusion protein band was apparent in the supernatant column, although a weaker band of the expected size was present in the input faction, indicating that the GFP-Trap was able to selectively retain the Sox6-GFP protein (Figure 6.3B). With these results confirming that the GFP-Trap could selectively purify the Sox6-GFP fusion protein from chromatin, I proceeded to use the GFP-Trap in ChIP experiments on *UAS:sox6-GFP;actin:GAL4* embryos.

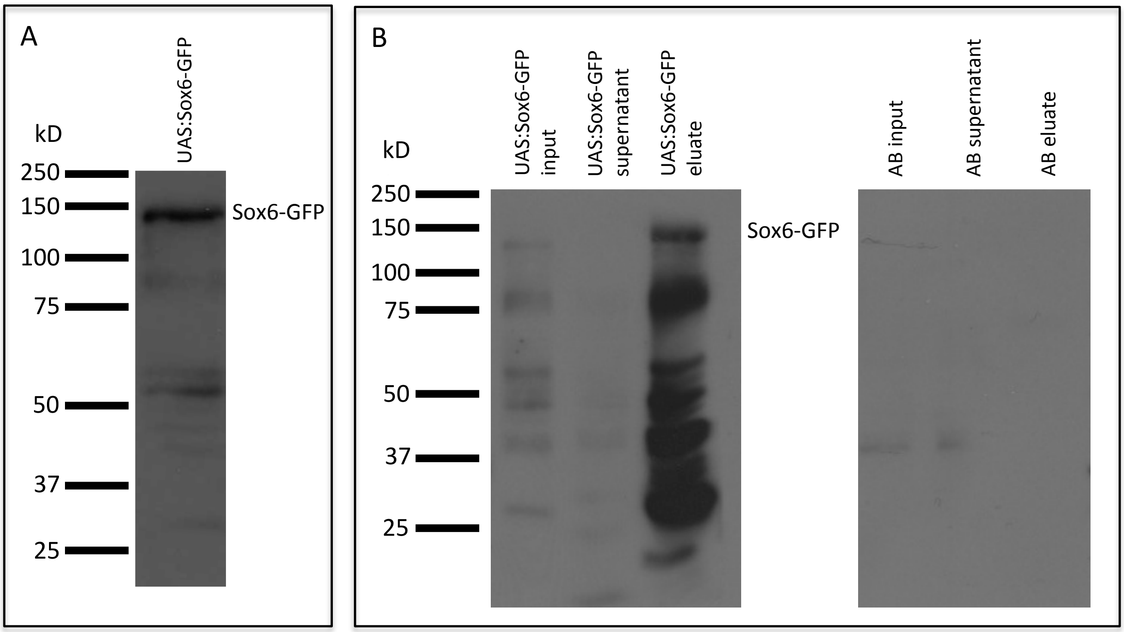


***Figure 6.3: Immunoprecipitation of Sox6-GFP using the GFP-trap.*** *Western blot analysis carried out on UAS:sox6-GFP;actin:GAL4 embryos at 30 hpf using an anti-GFP antibody (A). Cross-linked, sonicated chromatin was prepared from UAS:sox6-GFP;actin:GAL4 embryos and AB embryos at 30 hpf and then incubated with a GFP-Trap (B). Immunoprecipitated proteins and unbound proteins were analysed by SDS-PAGE and Western blotting with an anti-GFP antibody, and an equivalent sample of input chromatin was run alongside. No sox6-GFP protein was immunoprecipitated from the AB control embryos. However, multiple proteins were precipitated from UAS:sox6-GFP;actin:GAL4 embryos, with the largest protein band corresponding to the approximate size of Sox-GFP, which was also present in the input sample.*

### 6.2.3 Chromatin Immunoprecipitation (ChIP) using the GFP-Trap

To investigate whether Sox6 directly binds to the *cis*-regulatory elements of slow-twitch muscle genes, ChIP was performed. Pairs of primers were designed that encompass potential Sox6-binding sites in the proximal regions of *tnnc1b, smyhc1, myh7b* and *prox1a* (see Appendix A6 for location of primer pairs). Given that mutating the Sox6 binding sites in the *smyhc1:GFP* promoter resulted in ectopic GFP expression in the fast-twitch muscle (Chapter 5, Figure 5.8), primer pairs that encompassed these Sox6 binding sites in the *smyhc1* promoter were used as a positive control. Sox6 ChIP-seq analysis performed in mice has previously identified several Sox6 binding sites in the promoters of slow-twitch specific genes (An et al., 2011). Consequently, some primers pairs were designed to encompass the same sites in the zebrafish genes, to test whether Sox6 binding could be detected by ChIP in these regions. Chromatin was extracted from both wild-type AB and *UAS:sox6-GFP;actin:GAL4* embryos at 30 hpf, as significant differential expression of *tnnc1b* was previously observed between Sox6 null mutants and wild-type embryos at this timepoint (Chapter 5, Figure 5.3G), suggesting this may be a suitable timepoint at which to capture Sox6 acting as a transcriptional suppressor of slow-twitch genes. Ectopic expression of *smyhc1* in the fast-muscle of Sox6 mutants had also been observed at this timepoint (Chapter 5, Figure 5.7G), indicating that Sox6 normally acts to repress this gene at this particular stage.

ChIP using the GFP-Trap did not identify any detectable interactions between Sox6-GFP and the *cis*-regulatory regions of the selected slow-twitch genes (Figure 6.4). There was no significant difference between the DNA precipitated from the *UAS:sox6-GFP;actin:GAL4* embryos compared to the DNA precipitated from the AB control embryos, in any of the analysed slow-twitch muscle genes, including the positive control *smyhc1.* As there was also no enrichment in the *smyhc1* promoter region this is thought to be due to either a failure of the GFP-Trap ChIP protocol to selectively retain the Sox6-GFP fusion protein, or dissociation of the Sox6-GFP from its cognate DNA binding sites under the ChIP conditions used. Since mutating Sox6 binding sites in the *smyhc1:GFP* promoter results in a derepression of *smyhc1:GFP* expression in the fast-twitch muscle of 30 hpf embryos, it seems unlikely that Sox6 is not associated with DNA in this region at this stage of development.

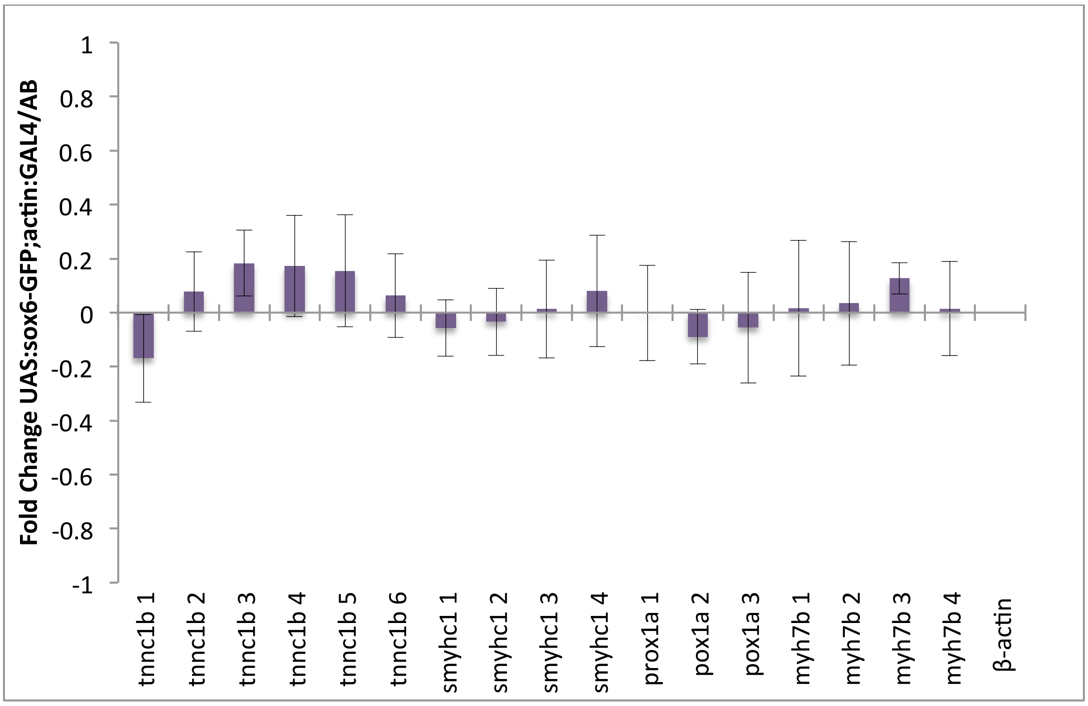


***Figure 6.4: Chromatin immunoprecipitation analysis of Sox6 binding to cis-regulatory regions of slow-twitch muscle specific genes.*** *Chromatin from UAS:sox6-GFP;actin:GAL4 and AB embryos was precipitated with a GFP-Trap at 30 hpf. Sox6-GFP is not detectably associated with the cis-regulatory regions of any of the slow-twitch muscle genes examined. Bar graphs show the mean value of the fold change between UAS:sox6-GFP;actin:GAL4 and AB embryos normalised to the β-actin gene. Results represent experimental and biological triplicates. See Chapter 2, Materials and Methods, section 2.13 Quantitative PCR, for the full method of analysis.*

## Discussion

The results described in Chapter 4 demonstrated the ability of the Sox6-GFP fusion protein to repress expression of slow-twitch specific genes. Chromatin Immunoprecipitation (ChIP) was therefore used to analyse whether Sox6 binds to chromatin encompassing slow-twitch muscle gene promoters. ChIP can be a convenient and effective assay to determine whether a protein directly binds to a particular DNA sequence. In order to successfully carry out this procedure, an antibody is needed that specifically recognises the protein of interest. Two ways to validate that an antibody specifically recognises its epitope, are Western Blot analysis and whole-mount immunohistochemistry. Both of these methods were therefore utilised in order to try to identify a Sox6-specific antibody. Neither of the pre-existing, commercially available antibodies recognised Sox6 protein either in Western blotting or whole-mount immunohistochemistry. This is likely to be due to the fact that these antibodies had been designed to detect mouse and human Sox6 proteins, and therefore there was not enough sequence similarity between the human/mouse epitopes and the corresponding zebrafish sequences within the Sox6 polypeptide, for the antibody to cross-react with zebrafish Sox6. Therefore, it was decided that a Sox6 antibody specific for zebrafish Sox6 protein might be necessary in order to perform ChIP analysis. Polyclonal Sox6 antibodies were generated by two different companies, Sdix and Absea, using the same His-tagged zebrafish Sox6 fusion protein. Whilst the two Sox6 antibodies generated by Sdix did not recognise Sox6 protein in immunohistochemistry and western blot analysis, the Sox6 antibody generated by Absea, did exhibit specificity for fast-twitch muscle nuclei. Polyclonal antibodies are prone to batch-variability, as was observed with the three Sox6 antibodies generated using the same zebrafish Sox-6 fusion protein, making them less than ideal for use in ChIP experiments. Another potential problem when using polyclonal antibodies is the high amount of background cross-reactivity they often exhibit. Although the Sox6M1 (Absea) antibody appeared to be specific in the immunohistochemistry analysis, Western blot analysis revealed that the antibody recognised numerous proteins, with the suspected Sox6 band, also present in *sox6*$^{-/-}$ mutants, indicating that this band did not correspond solely Sox6 protein. The Sox6M1 antibody also failed to specifically pull-down Sox6 protein from cross-linked chromatin, indicating that this antibody was unsuitable for use in ChIP analysis. For this reason an alternative approach was designed, utilising the GFP-Trap, which specifically pulls down GFP from protein extracts (Dubin et al, 2010; Rothbauer et al., 2008). Therefore, this technique was used to selectively precipitate Sox6-GFP from *UAS:sox6-GFP;actin:GAL4* embryos. GFP-Trap-mediated immunoprecipitation performed on cross-linked chromatin extracts obtained from *UAS:sox6-GFP;actin:GAL4* embryos, followed by SDS-PAGE and Western blot analysis, revealed a protein band at the appropriate size for full-length Sox6-GFP, if not a little too high, along with several other bands of lower molecular weight, which in some cases were stronger than the full-length Sox6-GFP band. Reasons for this are unknown, as no bands were observed from the GFP-Trap mediated pull down of cross-linked chromatin extracts from AB embryos, but it is possible that these smaller bands represent proteolytic fragments of the Sox6-GFP fusion protein. The fact that the full-length Sox6-GFP protein band was slightly higher than predicted could be due to the presence of post-translational modifications. Indeed, a recent report in mice has indicated that the E3 ubiquitin ligase, Trip12, recognises and polyubiquitinates Sox6, regulating Sox6 protein levels in skeletal muscle (An et al., 2013). If this is also the case in fish this could explain the higher molecular weight of the expected protein band, but also could explain why there are several smaller molecular weight protein bands, as these may represent the degradation of the ubiquitin labelled protein. Nevertheless, the presence of a band at roughly the correct size in *UAS:sox6-GFP;actin:GAL4* extracts, indicated that ChIP experiments with the GFP-Trap might be successful.

However, *in vivo* ChIP using a GFP-Trap to pull down the Sox6-GFP fusion protein detected no interaction between Sox6 and the selected cis-regulatory regions of slow-specific genes. This result might indicate that Sox6 does not stably and specifically bind to the promoters of slow-twitch muscle genes, or alternatively, it could indicate that there are some unanticipated technical problems with the Sox6-GFP ChIP assay using the GFP-Trap. Previous experiments indicated that Sox6 protein directly represses the *smyhc1* promoter, because mutating Sox6 sites in the *smyhc1:GF*P promoter led to the ectopic expression of GFP in fast-twitch fibres (Chapter 5 Figure 5.8). Primer pairs that yielded a PCR fragment encompassing these mutated Sox6 binding sites were used as a positive control for the binding of Sox6 to the *smyhc1* promoter. However, GFP-Trap-mediated pull down of Sox6-GFP resulted in no enrichment of the *smyhc1* promoter, possibly due to the inability of the GFP-Trap to successfully pull down Sox6-GFP from *UAS:sox6-GFP;actin:GAL4* embryos, or because the Sox6-GFP fusion protein had dissociated from its cognate target sites under the conditions used. The protein band observed in the Western blot analysis from cross-linked chromatin for the GFP-positive embryos is slightly too high to label Sox6-GFP, as the size of the expected protein band is about 117kDA. There were also numerous bands detected of smaller molecular weight than that of the expected Sox6-GFP fusion protein. However, since previous papers have indicated a high specificity of the GFP-Trap to GFP from protein extracts prepared for ChIP analysis (Dubin et al., 2010), this could mean that the GFP-Trap was failing to pull-down Sox6-GFP, or that the Sox6-GFP fusion is dissociating readily from the chromatin under the conditions used. Another possibility for the failure of the ChIP experiment to identify any Sox6 binding regions may be that the level of Sox6 binding to slow-twitch gene promoters *in vivo* is below the threshold for detection in the ChIP assay. However, I feel this latter possibility is unlikely, as the *actin:GAL4* line enables a high level of expression of Sox6-GFP throughout the zebrafish myotome.

The primer pairs that were used in the ChIP analysis only encompassed small regions of the target gene promoters. They were designed based on either the presumption of the conservation of functional Sox6 binding sites between zebrafish and mouse (An et al., 2011), the presence of a Sox6 consensus sequence in the *cis*-regulatory regions for the genes examined, or finally, in the case of *smyhc1*, Sox6 sites that had already been determined as functional sites due to the mutation of these sites in the *smyhc1:GFP* transgene resulting in ectopic GFP expression. It may be the case that there are other Sox6 binding sites in the regulatory region of the slow-twitch genes that have been analysed, which lie outside of the genomic sequences that have been analysed by PCR, therefore ChIP-seq analysis may be a better way to analysis the results from the ChIP experiment as this would allow a genome wide analysis of Sox6 enriched DNA. However the lack of enrichment of DNA in the *smyhc1* promoter pairs suggests that there was an overall failure of the ChIP technique, rather than the failure to identify potential Sox6 binding regions.

Although the above reasons suggest that the ChIP experiment failed, the results may also indicate that Sox6 represses these genes indirectly. It is possible that there is an indirect interaction between Sox6 and *tnnc1b*, since mutation of the canonical Sox6 sites in the *cis*-regulatory regions of the *tnnc1b:eGFP* transgene, some of which are conserved in the mammalian orthologues (An et al., 2011), resulted in no ectopic GFP expression. If this were to be the case, this would indicate an evolutionary divergence in the regulation of this gene between mammals and fish, as previous studies have demonstrated the functional interaction of Sox6 with these regulatory elements in the mouse *tnnc1b* gene (An et al., 2011).

Overall, the results in this chapter remain inconclusive and further analysis is needed to determine whether Sox6 directly inhibits slow-twitch muscle genes in zebrafish. A perhaps different approach to ChIP would be advisable, such as an electrophoretic mobility shift assay (EMSA), given that this does not require the use of a Sox6-specific antibody (Hellman and Fried, 2007, Lane et al., 1992). This would give us an indication of whether Sox6 is able to bind to Sox6 binding sites located within the *tnnc1b* promoter. However the exact location of the binding sites are first needed, as short fragments that are known to interact with Sox6 are needed to perform this technique.

# Chapter 7

# Concluding remarks

Current understanding of the mechanisms underlying the specification of muscle fibre-type during vertebrate embryonic development is relatively limited, in comparison to the extensive body of knowledge about the mechanisms of fibre-type switching in adult vertebrates (Bassel-Duby and Olson, 2006). The zebrafish provides us with a convenient model organism in which to investigate myogenesis, and in particular to elucidate the mechanisms of fibre-type specification, as slow-twitch and fast-twitch fibres develop in a spatio-temporally stereotyped manner. The results described in this thesis have advanced our understanding of the role of Sox6 in myogenesis and identified functions of this gene that are specific to the teleost lineage, as compared to the functions of *sox6* orthologues other vertebrate species.

The results of previous forward genetic screens in zebrafish indicate that mutation of approximately 1 in 10 of all zebrafish genes cause recognisable morphological abnormalities that appear within the first 5 days of development (Haffter et al., 1996, Solnicakrezel et al., 1994). However, this observation does not necessarily imply that the other 90% of zebrafish genes are functionally redundant or irrelevant during these early stages of development, for several reasons. First, this estimate is based on morphological embryonic phenotypes that are readily visualized under the light microscope, and it does not take into account more subtle phenotypes that may nevertheless be embryonic or larval lethal. Second, if a gene is maternally expressed, its maternally deposited mRNA and/or protein may compensate for loss of zygotic gene function caused by homozygosity for a loss-of-function mutation. Third, many genes are not required until the later stages of development, therefore a phenotype may only be apparent after 5dpf. Nevertheless, many genes could have functions that overlap or are redundant with those of closely related duplicates, and which could compensate for the loss of the mutated gene. This situation is relatively common in zebrafish because of the teleost whole genome duplication (Amores et al., 1998). The results described in this thesis indicate that *sox6* is not a functionally redundant gene, as loss of *sox6* function leads to a severe morphological phenotype in adults. As *sox6* mutant fish reach maturity, a severe scoliosis phenotype becomes apparent, indicating that *sox6* is needed for normal development and growth. Moreover, although *sox6* mutant embryos have no morphological phenotype, *in situ* hybridization has revealed aberrant expression of several *slow-twitch* genes in these mutants. My results further indicate that *sox6* is not a maternally expressed gene, and unlike the situation in other teleosts, there is no second copy of *sox6* in the zebrafish genome. I therefore explored the possibility that *sox6* could have an overlapping function with that of the closely related *sox5* gene. My results indicate that *sox5* is not acting redundantly with *sox6* with regards to myogenesis, because double *sox5*$^{-/-}$;*sox6*$^{-/-}$ mutants have no increase in ectopic expression of *smyhc1* as compared to *sox6*$^{-/-}$ mutants alone. However, it seems that *sox5* does function redundantly with *sox6* in other aspects of development, as *sox5*$^{-/-}$;*sox6*$^{-/-}$ double mutant larvae exhibit a severe morphologically abnormal phenotype, whereas *sox5*$^{-/-}$ and *sox6*$^{-/-}$ single mutants do not. This *sox5*$^{-/-}$;*sox6*$^{-/-}$ double mutant line could therefore be utilized to decipher the distinct and overlapping functions of these genes.

In mice, Sox6 appears to be one of the key transcription factors necessary for determining fibre-type fate during embryonic development, with a loss of *sox6*$^{-/-}$ resulting in a switch from fast-twitch fibres to slow-twitch fibres (Hagiwara et al., 2005; Hagiwara et al., 2007; An et al., 2011). However, my analysis revealed that although there is an increase in slow-twitch-specific genes in *sox6*$^{-/-}$ zebrafish mutants, the switch to the slow-twitch fibre programme is incomplete, as ectopic expression of all slow-twitch muscle genes is not observed throughout the whole myotome, fusion of fast-twitch fibres still occurs normally and there appears to be no down-regulation of fast-specific genes, although qPCR is needed to confirm this. This could be because a factor necessary for the activation of the slow-twitch programme is not present in the fast-twitch domain of *sox6*$^{-/-}$ mutants, or alternatively it is possible that another repressor besides *sox6* is able to repress the activation of certain slow-twitch genes.

During muscle atrophy a decrease in miR-499 expression causes a concomitant upregulation of Sox6, implicating the miR-499-associated, myosin-encoded microRNA network in fibre-type plasticity (McCarthy et al., 2009). These results suggest that Sox6 could play a role in both the development and physiology of muscle, and suggest that further studies of zebrafish *sox6* and its regulation, including its potential post-transcriptional control by miRNAs, could provide useful new insights into the mechanisms regulating both the development and the physiological function of muscle. Interestingly, my results show that the adult zebrafish *sox6*$^{-/-}$ mutant myotome is significantly smaller and abnormally shaped compared to that of wild-type siblings, indicating that sox6 may be required for other aspects of muscle development, such as growth and maintenance of the myotome, in addition to its early embryonic role in the specification of fibre-type identity.

Future studies can utilise the tools generated in this thesis to further our understanding of myogenesis as well as other aspects of vertebrate development. The *sox6* mutant in combination with ChIP-seq can be used to identify novel functions for Sox6 in development. The combined use of the *sox6* mutant and calcium reporter proteins such as GCaMP could be used to explore the role of *sox6* in muscle physiology and could provide insights into *sox6* function in adult fibre-type plasticity. The analysis of the *sox6* and *sox5* mutant lines generated in this thesis may reveal novel distinct or overlapping functions of these genes, not yet discovered in vertebrates. The importance of troponins in muscle contraction is highlighted by the fact that mutations in the genes encoding these proteins often result in myopathies (Hoffmann et al., 2001, Mogensen et al., 2004). The generation of the *tnnc1b:GFP* and *tnnc1a:GFP* transgeneic lines provide new troponin-specific reporter lines, which will assist our understanding of the regulation of these genes.

To summarise, I have identified a conserved enhancer element present in the first intron of the *slow troponin c* gene and generated two new transgenic reporter lines, which can be, utilised in future studies. The results in this thesis have furthered our understanding of the conserved role for Sox6 in zebrafish myogenesis, but also identified evolutionary divergence in the function of this gene in vertebrate muscle development. Finally, this thesis has uncovered a novel scoliosis phenotype associated with the loss of *sox6.* Further studies are now needed to investigate the pathogenetic mechanisms involved in the scoliosis phenotype. Rescuing *sox6* expression specifically in the muscle of these mutants by using the UAS:GAL4 system should begin to help us understand whether this defect is a primary consequence resulting from a loss of *sox6* expression in the bone or a secondary consequence due to the loss of *sox6* expression in the muscle.

# Appendix

## A1: Ultramer layer for *tnnc1b* BAC recombineering

```
TAAGTTTAGAACCTGATATGGGATTGTTGCTTAGGCGGCCAGAAAGACCAGTGCCAAGGACTTATTGTCATAACTCGAGTTCTCCTGCCTGTTGTCTCTGAAATTCC

TCCTCTATTTATAGCCCATACCATCCCTAACCCCTTTCTGCCCTGCTCCCTTTCCTTTTTCACTCATCCCTCGCTCTCTCTCTGATAGGAAAGGGGAGTGACAAAAG

CCATTGTCTTTTCTAAGCCATCCATTCCAGCCATAAAGGCCAAGGCAGGTGGCCCTGGGACTCTTTGTCCTGTGTCTGCCGTGAAGAGAGGAGCGCTGGATTACCTG
                                                                         GTCTGCCGTGAAGAGAGGAGCGCTGGATTACCTG

TGAGGAGTTTACTGACCATTTCAAAAAAGGAAAATCTAAACCCGGGACAGCCATGGATGATGTATATAAAGCAGCGGTAAGTTATGGGCACAGTAATTAAAGTTAGT
TGAGGAGTTTACTGACCATTTCAAAAAAGGAAAATCTAAACCCGTCGCCACCATGGTGAGCAA
                                   CCCGCTTAAGACGTCTATAGCTACTACATATATTTCGTCGCCATTCAATACCCGTGTCATTAATTTCAATCA

TGAAGACAAAATTTTTCTTATTATTTTTATATTTCCCATGTGATGTTTAACAGAGGAAGAATTTTCTTCACAGTATTTCCTATTATATCTTATTATCATCTGGAGAA
ACTTCTGTTTTAAAAAGAATAATAAAAATATAAAGGGTACACTACAAATTGTC

AGTCTTTTTCTTCTAGTAAAAAAAAAAAGTTCAACTGCTAAGGTCAATATTATTCGCCCCTTAAGAGATTATTTTATCTTATTCGCTGTTGTCCAGTGACTTCACTA

ATTAACTAAAACATCACTTGGGAAATATTTGAAAACGAATTCAAATTTCACAGTAGAATTACAAATTTTGTCTTCAAA
```

Black – Ultramer sequence
Blue – tnnc1b forward sequence
Red – Kozak sequence
Light green – Starting sequence of eGFP
Dark green – TOPO vector

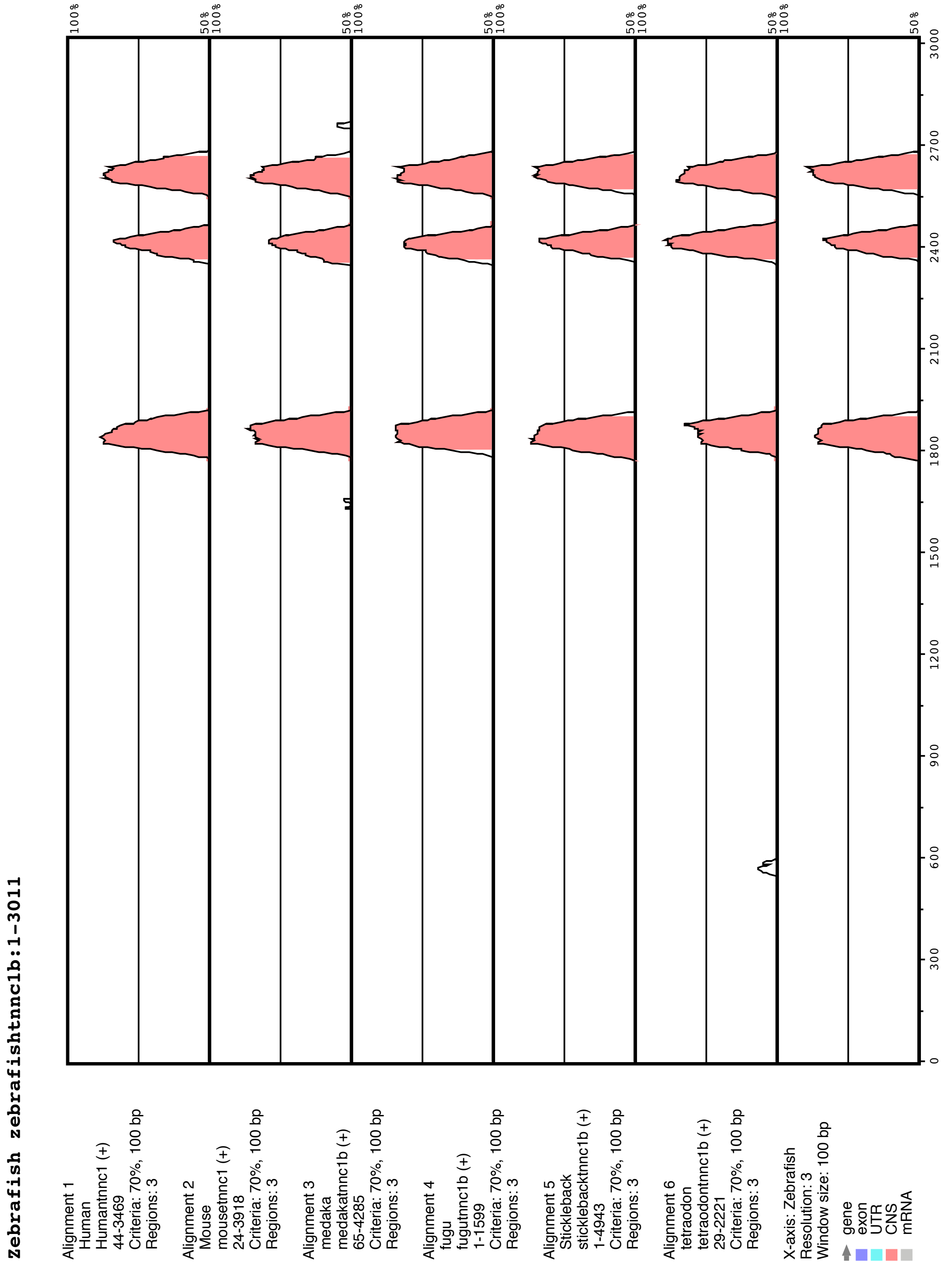
**Forward ultramer**
GTCTGCCGTGAAGAGAGGAGCGCTGGATTACCTGTGAGGAGTTTACTGACCATTTCAAAAAAGGAAAATCTAAACCCGTCGCCACCATGGTGAGCAA

**Reverse ultramer**
CTGTTAAACATCACATGGAAATATAAAAATAATAAGAAAAATTTTGTCTTCAACTAACTTTAATTACTGTGCCCATAACTTACCGCTGCTTTATATACATCATCGA
TATCTGCAGAATTCGCCC

Underlined portions are the end 20mers ordered to amplfy the fragment out of the TOPO vector.

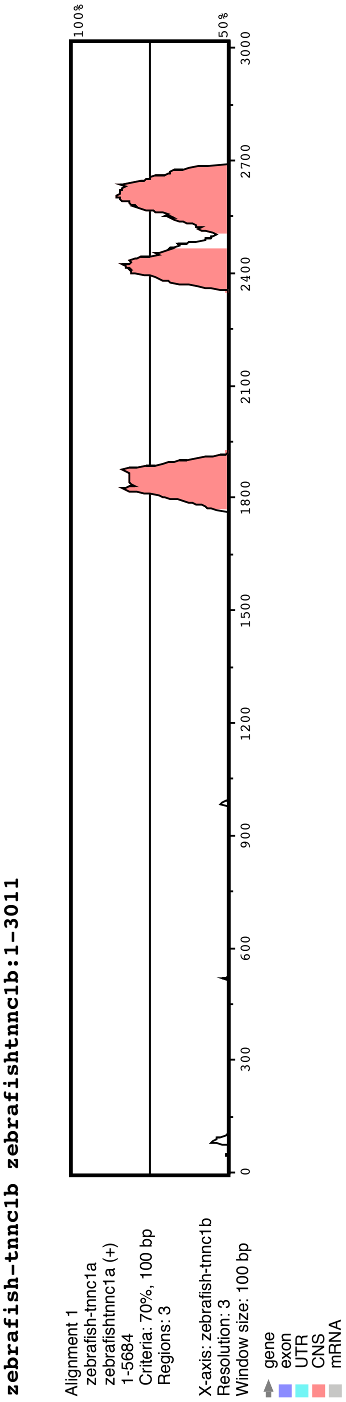
## A2: Zebrafish *tnnc1b* gene sequence comparison with orthologues

VISTA plot alignment and comparison of the zebrafish *tnnc1b* gene with mouse and human *tnnc1,* as well as medaka, fugu, stickleback and tetraodon *tnnc1b. Troponin c* sequences include 500bp of upstream sequence from the first exon. Conserved pink peaks represent protein coding regions. No CNEs can be identified between the zebrafish *tnnc1b* gene and its mammalian and teleost orthologues.



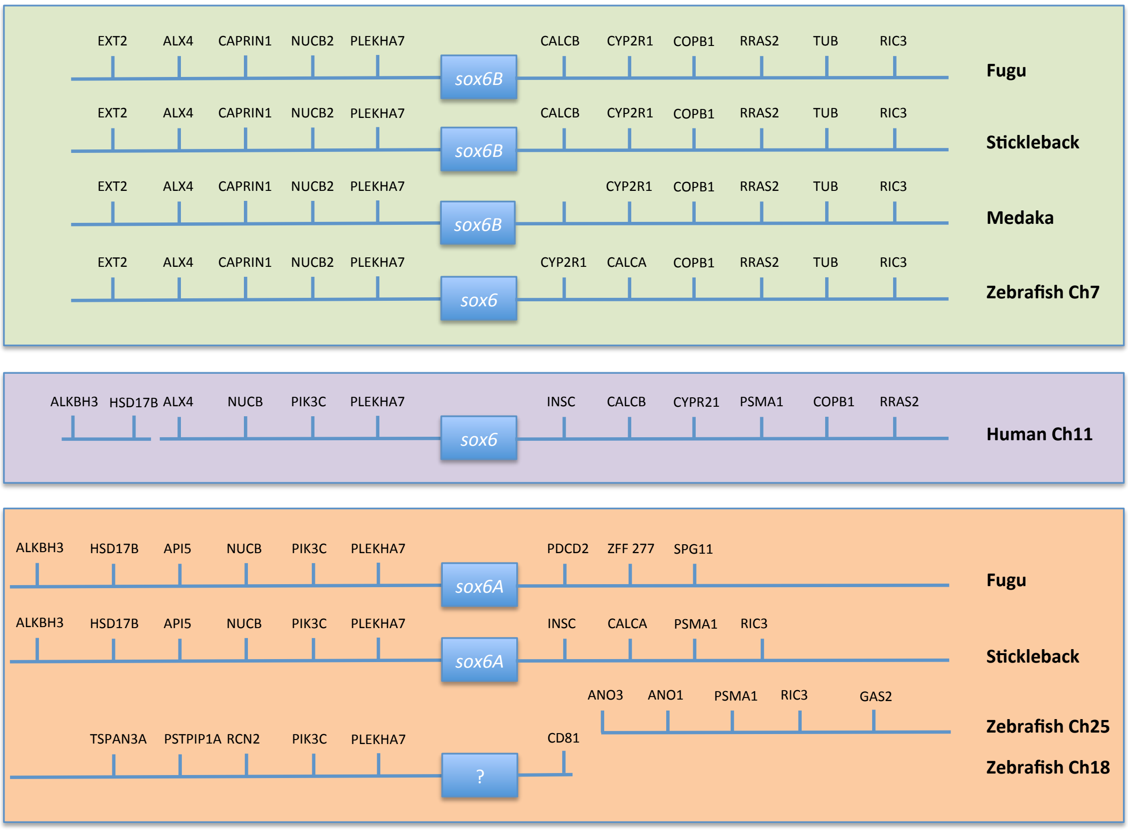
## A3: Zebrafish *tnnc1b* and *tnnc1a* sequence comparison

VISTA plot aligning and comparing the zebrafish *tnnc1b* gene with the zebrafish *tnnc1a* gene*.* The *tnnc1b* and *tnnc1a* sequences include 500bp of upstream sequence from the first exon. Although the protein coding sequences are highly conserved (three pink peaks) there is no conservation between the non-coding regions of these two genes.



## A4: Synteny of the *sox6* gene in vertbrates

The *sox6* gene is duplicated in most teleost genomes including *fugu, medaka* and stickleback. There is conserved synteny of the *sox6A* locus and the *sox6B* locus between teleost species. Although there is conserved synteny between the zebrafish *sox6B* locus it appears that the genes surrounding the *sox6A* locus in *fugu, medaka* and stickleback are split onto two different chromosomes in the zebrafish, with no *sox6A* gene identifiable.



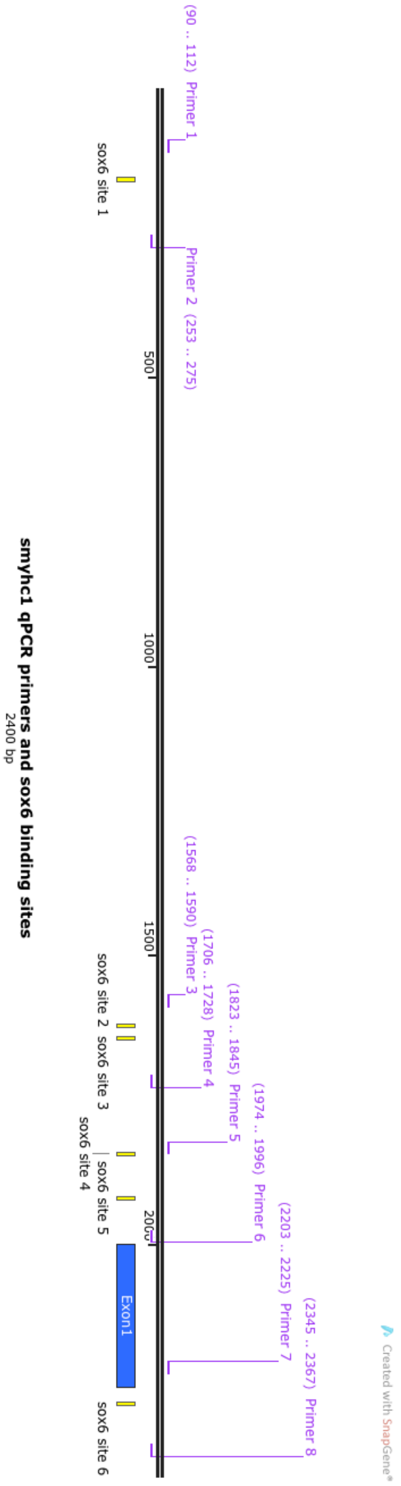
## A5: His-tagged Sox6M fusion protein

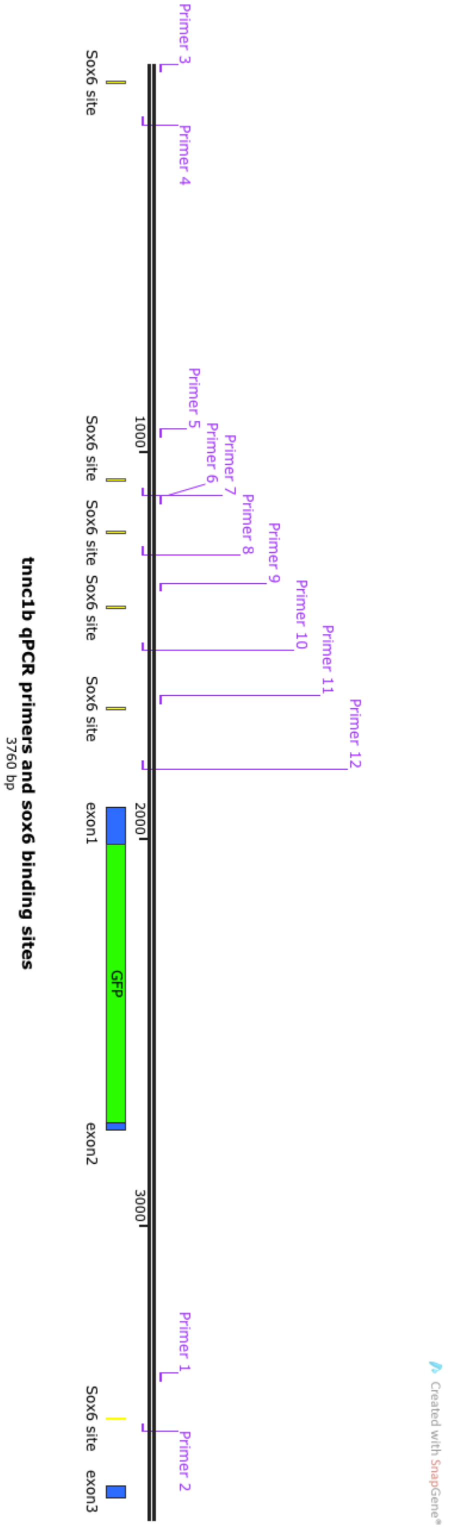
A fragment of *sox6* cDNA, corresponding to the middle of the Sox6 protein, was cloned into the His-tagged expression vector pET-2lb. The zebrafish Sox6 protein sequence is shown below with the Sox6 fragment used in the fusion protein highlighted in red and the HMG-box highlighted in blue.

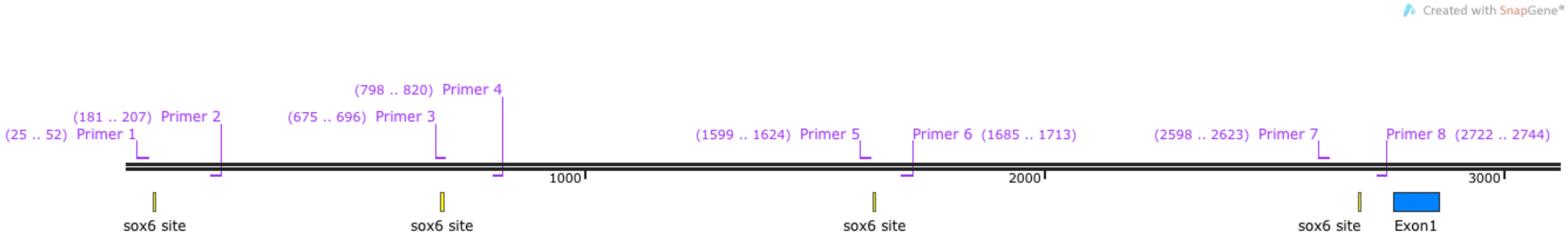
```
MSSKQATSPFTSTPDGGEDGVNQERMPWEKEENSESLVAPQLPLHNLLHNKPPLEELQPISSSV
PPESDWDSLMSAQQRMESDSNKVCSLYSFRNNSTSPHKPEEGARERGDLLSSSAFGTPERRKGS
LADVVDTLKQKKLEEMTKTEQDESSCMEKLLSKDWKEKMERLNTGELLGEIKGTPESLAEKERQ
LSTMITQLISLREQLLAAHDEQKKLAASQMEKQRQQMELARQQQEQIARQQQQLLQQQHKINLL
QQQIQVQGHMPPLMIPIFPHDQRTLAAAAAAQQGFLFPPGMSYKPGDNYPVQFIPSTMAAAAAS
GLNPLQLQQLYAAQLASMQVSPGAKMPPLPQPPNSAGQISPSGLKNEKRSSTPLAQVKEEGTQP
LNLSARPKTAEPIKSPTSPTQSLFPGSKSSPNSLSKSGDIPSPIGGLGRGSSLDILSSLNSTAL
FGDQDAVMKAIQEARKMREQIQREQLQHHQQGMEVKLSALTSMGLNNCRTEKERSHYDNLGHHL
SKLGEDGKIGHRVIDLTRPEDFEGGASTTEARVYREPRGRNSNEPHIKRPMNAFMVWAKDERRK
ILQAFPDMHNSNISKILGSRWKSMTNQEKQPYYEEQARLSKIHLEKYPNYKYKPRPKRTCIIDG
KKLRISEYKQMMRSRRQEMRQFFTVGQQPQTQIPITTSAGVVYPGAITMATTTPSPHMTSDCSS
ASASPEPTIPVIQSTFNMKMEPGTMVPSDAVNGEDEMDMYEDFEDEPKSDYSSENDTHEPVSAN
```

## A6: Primers pairs for ChIP-qPCR

Primer pairs were designed for the *tnnc1b, smyhc1, prox1a* and *myh7b* genes based on the presence of *sox6* sites in the promoter region, or the presence of *sox6* sites previously reported to directly bind to Sox6 in mice (An et al., 2011). Primer pairs for *smyhc1* encompassed the previously mutated sites in the *smyhc1:GFP* transgene.

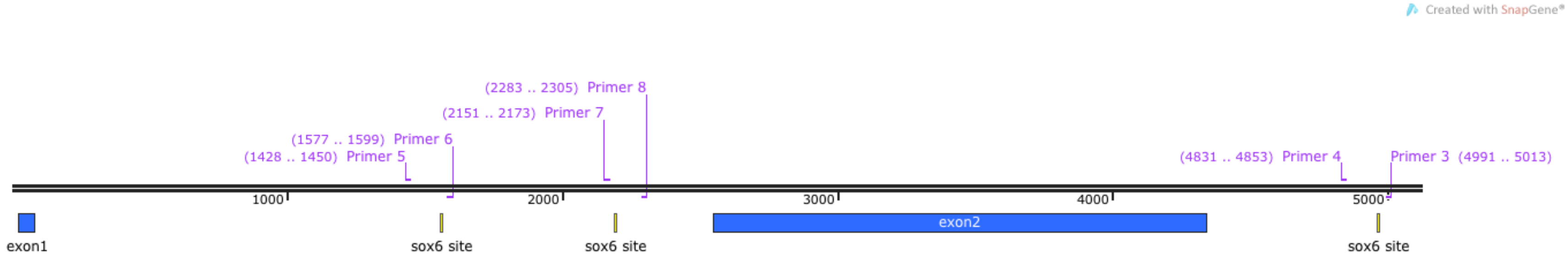






**myh7b qPCR primers and Sox6 binding sites**

3120 bp



**prox1a qPCR primers and sox6 binding sites**

5120 bp

## A7: Expression of heart-specific troponins in sox6$^{-/-}$ mutant embryos

*sox6* transcripts have been detected in the heart tissue of mice and humans (Hagiwara et al., 2011). The loss of *sox6* in mutant mice results in conduction abnormalities in the heart and aberrant expression of heart specific muscle genes, including genes encoding myosin heavy chains and troponins (Hagiwara et al., 2005). As mentioned in Chapter 1, unlike mammals, zebrafish have two slow-specific skeletal muscle troponins, one that is specifically expressed in the heart (*tnnc1a*) and one that is specifically expressed in the skeletal slow-twitch muscle (*tnnc1b*) (Sogah et al., 2010). As loss of *sox6* causes robust ectopic expression of *tnnc1b* in the fast-twitch skeletal muscle, I investigated whether loss of *sox6* had an effect on the expression of *tnnc1a*. In situ hybridisation revealed that there is no change in the expression of *tnnc1a* at 32 hpf in sox6$^{-/-}$ mutants as compared to siblings. Similarly, the expression of two other heart specific troponins, *tnnt2* and *tnni1b,* was unaffected in the mutant compared to the sibling. These results suggest a divergence in the regulation of *tnnc1a* and *tnnc1b*, which complements previous studies identifying two separate enhancers for the mammalian *Tnnc1* gene that are responsible for regulating its tissue specific expression (Parmacek et al., 1994, Parmacek et al., 1992). One question that still needs to be addressed is whether *sox6* is expressed in the zebrafish heart. Expression of *sox6* in the heart has not previously been reported (Wang et al., 2011), and further in situ hybridisation analysis is required, since although its expression appears to be omitted from the heart in Chapter 5 (see Figure 5.1), the images are not clear enough to make a definitive conclusion.

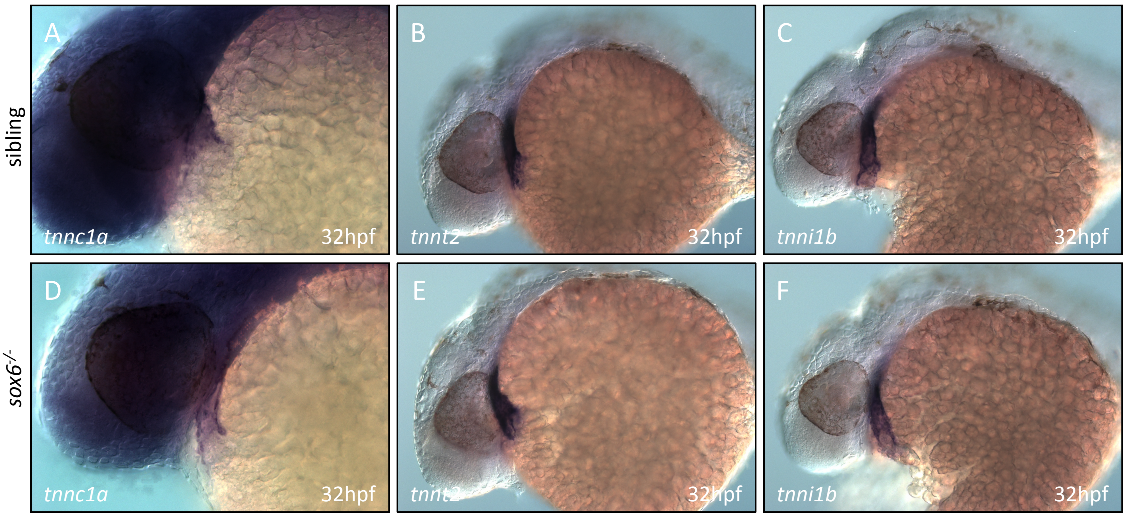


***Figure A5: The expression of the heart specific troponins tnnc1a, tnnt2 and tnni1b is unchanged in the sox6$^{-/-}$ mutant (D-F) when compared to the sibling (A-C) at 32 hpf.***

# Abbreviations

| | |
|---|---|
| Amp | Ampicillin |
| BAC | Bacterial artificial chromosome |
| BCIP | 5-bromo-4-chloro-3-indolyl phosphate |
| βg | Beta-globin |
| bHLH | Basic helix-loop-helix |
| BMP | Bone morphogenetic protein |
| ChIP | Chromatin immunoprecipitiation |
| CKIP-1 | Casein kinase 2 interacting protein-1 |
| CNE | Conserved non-coding element |
| Cm | Chloramphenicol |
| Ct | Threshold cycle |
| DAPI | 4', 6-diamidino-2-phenylindole |
| DCS | Double conserved synteny |
| DIG | Digoxigenin |
| DNase | Deoxyribonuclease |
| Dock1 | Dedicator of cytokinesis |
| Dpf | Days post fertilization |
| DSG | Disuccinimidyl glutarate |
| *E. coli* | *Escherichia coli* |
| Edu | 5-ethynyl-2'-deoxyuridine |
| eGFP | Enhanced green fluorescent protein |
| Eng | Engrailed |
| ENU | N-ethyl-N-nitrosourea |
| EtOH | Ethanol |
| FGF8 | Fibroblast growth factor 8 |
| Flp | Flippase |
| Fmyhcx | Fast myosin heavy chain |
| FRT | Flippase recognition target |
| GFP | Green fluorescent protein |
| HA | Homology arm |
| HATs | Histone acetyltransferases |
| HDACs | Histone deacetylases |
| Hh | Hedgehog |
| HMG | High mobility group |

| | |
|---|---|
| Hpf | Hours post fertilization |
| HRP | Horseradish peroxidase |
| IP | Immunoprecipitation |
| Kan | Kanamycin |
| KOH | Potassium hydroxide |
| MeOH | Methanol |
| MFF | Medial fast fibre |
| MO | Morpholino |
| MPs | Muscle pioneers |
| MRFs | Myogenic regulatory factors |
| Myf5 | Myogenic factor 5 |
| Myh7b | Slow myosin heavy chain 7b |
| Mylz2 | Myosin light chain 2 (Fast-specific) |
| Mylz10 | Myosin light chain 10 (Slow-specific) |
| MyoD | Myoblast determination protein |
| Myog | Myogenin |
| NaAc | Sodium acetate |
| NBT | Nitro blue tetrazolium chloride |
| NFAT | Nuclear factor of activated T cells |
| Nfix | Nuclear factor 1 X-type |
| Nrd | Narrowminded (Prdm1 allele) |
| PBS | Phosphate buffered saline |
| PCR | Polymerase chain reaction |
| PFA | Paraformaldehyde |
| PKA | Protein kinase A |
| POD | Horseradish peroxidase |
| Prdm1 | PR domain zinc-finger protein 1 |
| Prox1 | Prospero homeobox protein 1 |
| Ptc | Patched |
| qPCR | Quantitative polymerase chain reaction |
| RNase | Ribonuclease |
| RNasin | Ribonuclease inhibitor |
| Ryr1a | Ryanodine receptor 1a (Slow-specific) |
| Sib | Sibling |
| SDS | Sodium dodecyl sulfate |
| Shh | Sonic hedgehog |
| Smo | Smoothened |

| | |
|---|---|
| Smyhc | Slow myosin heavy chain |
| SSC | Saline-sodium citrate |
| SSF | Superficial slow-twitch fibres |
| SV40 | Simian vacuolating virus 40 |
| TALENs | Transcription activator-like effector nucleases |
| Tg | Transgene |
| Tpm2 | Tropomyosin 2 (slow-specific) |
| Tnnc1a | Slow troponin c1a (heart-specific) |
| Tnnc1b | Slow troponin c1b |
| Tnni1a | Slow troponin i1a |
| Tnni1b | Troponin i1b (heart-specific) |
| Tnni2 | Fast troponin i2 |
| Tnnt1 | Slow troponin t1 |
| Tnnt2 | Troponin t2 (heart-specific) |
| Tnnt3a | Fast troponin t3a |
| TSA | Tyramide signal amplification |
| Ubo | u-boot (mutant allele of Prdm1) |
| ZFN | Zinc-finger nuclease |

# References

ALFAQIH, M. A., BRUNELLI, J. P., DREW, R. E. & THORGAARD, G. H. 2009. Mapping of five candidate sex-determining loci in rainbow trout (Oncorhynchus mykiss). *BMC Genet,* 10**,** 2.

ALTAF, F., GIBSON, A., DANNAWI, Z. & NOORDEEN, H. 2013. Adolescent idiopathic scoliosis. *British Medical Journal,* 346.

AMORES, A., FORCE, A., YAN, Y. L., JOLY, L., AMEMIYA, C., FRITZ, A., HO, R. K., LANGELAND, J., PRINCE, V., WANG, Y. L., WESTERFIELD, M., EKKER, M. & POSTLETHWAIT, J. H. 1998. Zebrafish hox clusters and vertebrate genome evolution. *Science,* 282**,** 1711-4.

AN, C. I., DONG, Y. & HAGIWARA, N. 2011. Genome-wide mapping of Sox6 binding sites in skeletal muscle reveals both direct and indirect regulation of muscle terminal differentiation by Sox6. *BMC Dev Biol,* 11**,** 59.

AN, C. I., GANIO, E. & HAGIWARA, N. 2013. Trip12, a HECT domain E3 ubiquitin ligase, targets Sox6 for proteasomal degradation and affects fiber type-specific gene expression in muscle cells. *Skelet Muscle,* 3**,** 11.

ANCELIN, K., LANGE, U. C., HAJKOVA, P., SCHNEIDER, R., BANNISTER, A. J., KOUZARIDES, T. & SURANI, M. A. 2006. Blimp1 associates with Prmt5 and directs histone arginine methylation in mouse germ cells. *Nat Cell Biol,* 8**,** 623-30.

ASADUZZAMAN, M., KINOSHITA, S., BHUIYAN, S. S., ASAKAWA, S. & WATABE, S. 2013. Stimulatory and inhibitory mechanisms of slow muscle-specific myosin heavy chain gene expression in fish: transient and transgenic analysis of torafugu MYH(M86-2) promoter in zebrafish embryos. *Exp Cell Res,* 319**,** 820-37.

BAAS, D., CAUSSANEL-BOUDE, S., GUIRAUD, A., CALHABEU, F., DELAUNE, E., PILOT, F., CHOPIN, E., MACHUCA-GAYET, I., VERNAY, A., BERTRAND, S., RUAL, J. F., JURDIC, P., HILL, D. E., VIDAL, M., SCHAEFFER, L. & GOILLOT, E. 2012. CKIP-1 regulates mammalian and zebrafish myoblast fusion. *J Cell Sci,* 125**,** 3790-800.

BALCIUNAS, D., WANGENSTEEN, K. J., WILBER, A., BELL, J., GEURTS, A., SIVASUBBU, S., WANG, X., HACKETT, P. B., LARGAESPADA, D. A., MCIVOR, R. S. & EKKER, S. C. 2006. Harnessing a high cargo-capacity transposon for genetic applications in vertebrates. *PLoS Genet,* 2**,** e169.

BARRESI, M. J., STICKNEY, H. L. & DEVOTO, S. H. 2000. The zebrafish slow-muscle-omitted gene product is required for Hedgehog signal transduction and the development of slow muscle identity. *Development,* 127**,** 2189-99.

BASSEL-DUBY, R. & OLSON, E. N. 2006. Signaling pathways in skeletal muscle remodeling. *Annu Rev Biochem,* 75**,** 19-37.

BAXENDALE, S., DAVISON, C., MUXWORTHY, C., WOLFF, C., INGHAM, P. W. & ROY, S. 2004. The B-cell maturation factor Blimp-1 specifies vertebrate slow-twitch muscle fiber identity in response to Hedgehog signaling. *Nat Genet,* 36**,** 88-93.

BEDELL, V. M., WANG, Y., CAMPBELL, J. M., POSHUSTA, T. L., STARKER, C. G., KRUG, R. G., 2ND, TAN, W., PENHEITER, S. G., MA, A. C., LEUNG, A. Y., FAHRENKRUG, S. C., CARLSON, D. F., VOYTAS, D. F., CLARK, K. J., ESSNER, J. J. & EKKER, S. C. 2012. In vivo genome editing using a high-efficiency TALEN system. *Nature,* 491**,** 114-8.

BEERMANN, M. L., ARDELT, M., GIRGENRATH, M. & MILLER, J. B. 2010. Prdm1 (Blimp-1) and the expression of fast and slow myosin heavy chain isoforms during avian myogenesis in vitro. *PLoS One,* 5**,** e9951.

BESSA, J., TENA, J. J., DE LA CALLE-MUSTIENES, E., FERNANDEZ-MINAN, A., NARANJO, S., FERNANDEZ, A., MONTOLIU, L., AKALIN, A., LENHARD, B., CASARES, F. & GOMEZ-SKARMETA, J. L. 2009. Zebrafish enhancer detection (ZED) vector: a new tool to facilitate transgenesis and the functional analysis of cis-regulatory regions in zebrafish. *Dev Dyn,* 238**,** 2409-17.

BESSARAB, D. A., CHONG, S. W., SRINIVAS, B. P. & KORZH, V. 2008. Six1a is required for the onset of fast muscle differentiation in zebrafish. *Dev Biol,* 323**,** 216-28.

BLAGDEN, C. S., CURRIE, P. D., INGHAM, P. W. & HUGHES, S. M. 1997. Notochord induction of zebrafish slow muscle mediated by Sonic hedgehog. *Genes Dev,* 11**,** 2163-75.

BORMIOLI, S. P., SARTORE, S., VITADELLO, M. & SCHIAFFINO, S. 1980. "Slow" myosins in vertebrate skeletal muscle. An immunofluorescence study. *J Cell Biol,* 85**,** 672-81.

BOTTINELLI, R., SCHIAFFINO, S. & REGGIANI, C. 1991. Force-Velocity Relations and Myosin Heavy-Chain Isoform Compositions of Skinned Fibers from Rat Skeletal-Muscle. *Journal of Physiology-London,* 437**,** 655-672.

BUTLER, J., COSMOS, E. & BRIERLEY, J. 1982. Differentiation of muscle fiber types in aneurogenic brachial muscles of the chick embryo. *J Exp Zool,* 224**,** 65-80.

CANTINI, M., SARTORE, S. & SCHIAFFINO, S. 1980. Myosin types in cultured muscle cells. *J Cell Biol,* 85**,** 903-9.

CHANG, D. H., CATTORETTI, G. & CALAME, K. L. 2002. The dynamic expression pattern of B lymphocyte induced maturation protein-1 (Blimp-1) during mouse embryonic development. *Mech Dev,* 117**,** 305-9.

CHENG, L., ALVARES, L. E., AHMED, M. U., EL-HANFY, A. S. & DIETRICH, S. 2004. The epaxial-hypaxial subdivision of the avian somite. *Dev Biol,* 274**,** 348-69.

CHIN, E. R., OLSON, E. N., RICHARDSON, J. A., YANG, Q., HUMPHRIES, C., SHELTON, J. M., WU, H., ZHU, W., BASSEL-DUBY, R. & WILLIAMS, R. S. 1998. A calcineurin-dependent transcriptional pathway controls skeletal muscle fiber type. *Genes Dev,* 12**,** 2499-509.

CHUNG, J. H., WHITELEY, M. & FELSENFELD, G. 1993. A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila. *Cell,* 74**,** 505-14.

CONCORDET, J. P., LEWIS, K. E., MOORE, J. W., GOODRICH, L. V., JOHNSON, R. L., SCOTT, M. P. & INGHAM, P. W. 1996. Spatial regulation of a zebrafish patched homologue reflects the roles of sonic hedgehog and protein kinase A in neural tube and somite patterning. *Development,* 122**,** 2835-46.

CONNOR, F., WRIGHT, E., DENNY, P., KOOPMAN, P. & ASHWORTH, A. 1995. The Sry-related HMG box-containing gene Sox6 is expressed in the adult testis and developing nervous system of the mouse. *Nucleic Acids Res,* 23**,** 3365-72.

CORTES, F., DAGGETT, D., BRYSON-RICHARDSON, R. J., NEYT, C., MAULE, J., GAUTIER, P., HOLLWAY, G. E., KEENAN, D. & CURRIE, P. D. 2003. Cadherin-mediated differential cell adhesion controls slow muscle cell migration in the developing zebrafish myotome. *Dev Cell,* 5**,** 865-76.

COSTILL, D. L., DANIELS, J., EVANS, W., FINK, W., KRAHENBUHL, G. & SALTIN, B. 1976. Skeletal muscle enzymes and fiber composition in male and female track athletes. *J Appl Physiol,* 40**,** 149-54.

COUTELLE, O., BLAGDEN, C. S., HAMPSON, R., HALAI, C., RIGBY, P. W. & HUGHES, S. M. 2001. Hedgehog signalling is required for maintenance of myf5 and myoD expression and timely terminal differentiation in zebrafish adaxial myogenesis. *Dev Biol,* 236**,** 136-50.

CUI, J., SHEN, X., ZHAO, H. & NAGAHAMA, Y. 2011. Genome-wide analysis of Sox genes in Medaka (Oryzias latipes) and their expression pattern in embryonic development. *Cytogenet Genome Res,* 134**,** 283-94.

CURRIE, P. D. & INGHAM, P. W. 1996. Induction of a specific muscle cell type by a hedgehog-like protein in zebrafish. *Nature,* 382**,** 452-5.

DAILEY, L. & BASILICO, C. 2001. Coevolution of HMG domains and homeodomains and the generation of transcriptional regulation by Sox/POU complexes. *J Cell Physiol,* 186**,** 315-28.

DAVISON, W. 1997. The Effects of Exercise Training on Teleost Fish, a Review of Recent Literature. *Comparative Biochemistry and Physiology Part A: Physiology,* 117**,** 67-75.

DEVOTO, S. H., MELANCON, E., EISEN, J. S. & WESTERFIELD, M. 1996. Identification of separate slow and fast muscle precursor cells in vivo, prior to somite formation. *Development,* 122**,** 3371-80.

DOLEZ, M., NICOLAS, J. F. & HIRSINGER, E. 2011. Laminins, via heparan sulfate proteoglycans, participate in zebrafish myotome morphogenesis by modulating the pattern of Bmp responsiveness. *Development,* 138**,** 97-106.

DU, S. J., DEVOTO, S. H., WESTERFIELD, M. & MOON, R. T. 1997. Positive and negative regulation of muscle cell identity by members of the hedgehog and TGF-beta gene families. *J Cell Biol,* 139**,** 145-56.

DUBIN, M., FUCHS, J., GRAF, R., SCHUBERT, I. & NELLEN, W. 2010. Dynamics of a novel centromeric histone variant CenH3 reveals the evolutionary ancestral timing of centromere biogenesis. *Nucleic Acids Res,* 38**,** 7526-37.

EKKER, S. C., UNGAR, A. R., GREENSTEIN, P., VON KESSLER, D. P., PORTER, J. A., MOON, R. T. & BEACHY, P. A. 1995. Patterning activities of vertebrate hedgehog proteins in the developing eye and brain. *Curr Biol,* 5**,** 944-55.

ELWORTHY, S., HARGRAVE, M., KNIGHT, R., MEBUS, K. & INGHAM, P. W. 2008. Expression of multiple slow myosin heavy chain genes reveals a diversity of zebrafish slow twitch muscle fibres with differing requirements for Hedgehog and Prdm1 activity. *Development,* 135**,** 2115-26.

ERICKSON, M. R., GALLETTA, B. J. & ABMAYR, S. M. 1997. Drosophila myoblast city encodes a conserved protein that is essential for myoblast fusion, dorsal closure, and cytoskeletal organization. *J Cell Biol,* 138**,** 589-603.

FELSENFELD, A. L., CURRY, M. & KIMMEL, C. B. 1991. The fub-1 mutation blocks initial myofibril formation in zebrafish muscle pioneer cells. *Dev Biol,* 148**,** 23-30.

FENG, X., ADIARTE, E. G. & DEVOTO, S. H. 2006. Hedgehog acts directly on the zebrafish dermomyotome to promote myogenic differentiation. *Dev Biol,* 300**,** 736-46.

FISHER, S., GRICE, E. A., VINTON, R. M., BESSLING, S. L. & MCCALLION, A. S. 2006a. Conservation of RET regulatory function from human to zebrafish without sequence similarity. *Science,* 312**,** 276-9.

FISHER, S., GRICE, E. A., VINTON, R. M., BESSLING, S. L., URASAKI, A., KAWAKAMI, K. & MCCALLION, A. S. 2006b. Evaluating the biological relevance of putative enhancers using Tol2 transposon-mediated transgenesis in zebrafish. *Nat Protoc,* 1**,** 1297-305.

FRAZER, K. A., PACHTER, L., POLIAKOV, A., RUBIN, E. M. & DUBCHAK, I. 2004. VISTA: computational tools for comparative genomics. *Nucleic Acids Res,* 32**,** W273-9.

FRIEDLI, M., BARDE, I., ARCANGELI, M., VERP, S., QUAZZOLA, A., ZAKANY, J., LIN-MARQ, N., ROBYR, D., ATTANASIO, C., SPITZ, F., DUBOULE, D., TRONO, D. & ANTONARAKIS, S. E. 2010. A systematic enhancer screen using lentivector transgenesis identifies conserved and non-conserved functional elements at the Olig1 and Olig2 locus. *PLoS One,* 5**,** e15741.

FUMASONI, I., MEANI, N., RAMBALDI, D., SCAFETTA, G., ALCALAY, M. & CICCARELLI, F. D. 2007. Family expansion and gene rearrangements contributed to the functional specialization of PRDM genes in vertebrates. *BMC Evol Biol,* 7**,** 187.

GALASKO, C. S., WILLIAMSON, J. B. & DELANEY, C. M. 1995. Lung function in Duchenne muscular dystrophy. *Eur Spine J,* 4**,** 263-7.

GEIGER, P. C., CODY, M. J. & SIECK, G. C. 1999. Force-calcium relationship depends on myosin heavy chain and troponin isoforms in rat diaphragm muscle fibers. *Journal of Applied Physiology,* 87**,** 1894-1900.

GLASGOW, E. & TOMAREV, S. I. 1998. Restricted expression of the homeobox gene prox 1 in developing zebrafish. *Mech Dev,* 76**,** 175-8.

GOLLNICK, P. D., ARMSTRONG, R. B., SAUBERT, C. W. T., PIEHL, K. & SALTIN, B. 1972. Enzyme activity and fiber composition in skeletal muscle of untrained and trained men. *J Appl Physiol,* 33**,** 312-9.

GOMES, A. V. & POTTER, J. D. 2004. Molecular and cellular aspects of troponin cardiomyopathies. *Ann N Y Acad Sci,* 1015**,** 214-24.

GREASER, M. L. & GERGELY, J. 1973. Purification and properties of the components from tropinin. *J Biol Chem,* 248**,** 2125-33.

GRIFONE, R., LACLEF, C., SPITZ, F., LOPEZ, S., DEMIGNON, J., GUIDOTTI, J. E., KAWAKAMI, K., XU, P. X., KELLY, R., PETROF, B. J., DAEGELEN, D., CONCORDET, J. P. & MAIRE, P. 2004. Six1 and Eya1 expression can reprogram adult muscle from the slow-twitch phenotype into the fast-twitch phenotype. *Mol Cell Biol,* 24**,** 6253-67.

GROSSCHEDL, R., GIESE, K. & PAGEL, J. 1994. HMG domain proteins: architectural elements in the assembly of nucleoprotein structures. *Trends Genet,* 10**,** 94-100.

GROVES, J. A., HAMMOND, C. L. & HUGHES, S. M. 2005. Fgf8 drives myogenic progression of a novel lateral fast muscle fibre population in zebrafish. *Development,* 132**,** 4211-22.

GYORY, I., WU, J., FEJER, G., SETO, E. & WRIGHT, K. L. 2004. PRDI-BF1 recruits the histone H3 methyltransferase G9a in transcriptional silencing. *Nat Immunol,* 5**,** 299-308.

HAFFTER, P., GRANATO, M., BRAND, M., MULLINS, M. C., HAMMERSCHMIDT, M., KANE, D. A., ODENTHAL, J., VAN EEDEN, F. J., JIANG, Y. J., HEISENBERG, C. P., KELSH, R. N., FURUTANI-SEIKI, M., VOGELSANG, E., BEUCHLE, D., SCHACH, U., FABIAN, C. & NUSSLEIN-VOLHARD, C. 1996. The identification of genes with unique and essential functions in the development of the zebrafish, Danio rerio. *Development,* 123**,** 1-36.

HAGIWARA, N. 2011. Sox6, jack of all trades: a versatile regulatory protein in vertebrate development. *Dev Dyn,* 240**,** 1311-21.

HAGIWARA, N., KLEWER, S. E., SAMSON, R. A., ERICKSON, D. T., LYON, M. F. & BRILLIANT, M. H. 2000. Sox6 is a candidate gene for p100H myopathy, heart block, and sudden neonatal death. *Proc Natl Acad Sci U S A,* 97**,** 4180-5.

HAGIWARA, N., MA, B. & LY, A. 2005. Slow and fast fiber isoform gene expression is systematically altered in skeletal muscle of the Sox6 mutant, p100H. *Dev Dyn,* 234**,** 301-11.

HAGIWARA, N., YEH, M. & LIU, A. 2007. Sox6 is required for normal fiber type differentiation of fetal skeletal muscle in mice. *Dev Dyn,* 236**,** 2062-76.

HAMADE, A., DERIES, M., BEGEMANN, G., BALLY-CUIF, L., GENET, C., SABATIER, F., BONNIEU, A. & COUSIN, X. 2006. Retinoic acid activates myogenesis in vivo through Fgf8 signalling. *Dev Biol,* 289**,** 127-40.

HAMER, P. W., MCGEACHIE, J. M., DAVIES, M. J. & GROUNDS, M. D. 2002. Evans Blue Dye as an in vivo marker of myofibre damage: optimising parameters for detecting initial myofibre membrane permeability. *J Anat,* 200**,** 69-79.

HAMMERSCHMIDT, M., BITGOOD, M. J. & MCMAHON, A. P. 1996. Protein kinase A is a common negative regulator of Hedgehog signaling in the vertebrate embryo. *Genes Dev,* 10**,** 647-58.

HAMMOND, C. L., HINITS, Y., OSBORN, D. P., MINCHIN, J. E., TETTAMANTI, G. & HUGHES, S. M. 2007. Signals and myogenic regulatory factors restrict pax3 and pax7 expression to dermomyotome-like tissue in zebrafish. *Dev Biol,* 302**,** 504-21.

HAN, Y. & LEFEBVRE, V. 2008. L-Sox5 and Sox6 drive expression of the aggrecan gene in cartilage by securing binding of Sox9 to a far-upstream enhancer. *Mol Cell Biol,* 28**,** 4999-5013.

HARLEY, V. R., LOVELL-BADGE, R. & GOODFELLOW, P. N. 1994. Definition of a consensus DNA binding site for SRY. *Nucleic Acids Res,* 22**,** 1500-1.

HATTA, K., BREMILLER, R., WESTERFIELD, M. & KIMMEL, C. B. 1991. Diversity of expression of engrailed-like antigens in zebrafish. *Development,* 112**,** 821-32.

HATTORI, T., COUSTRY, F., STEPHENS, S., EBERSPAECHER, H., TAKIGAWA, M., YASUDA, H. & DE CROMBRUGGHE, B. 2008. Transcriptional regulation of chondrogenesis by coactivator Tip60 via chromatin association with Sox9 and Sox5. *Nucleic Acids Res,* 36**,** 3011-24.

HELLMAN, L. M. & FRIED, M. G. 2007. Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nature Protocols,* 2**,** 1849-1861.

HENRY, C. A. & AMACHER, S. L. 2004. Zebrafish slow muscle cell migration induces a wave of fast muscle morphogenesis. *Dev Cell,* 7**,** 917-23.

HENSINGER, R. N. 2009. Congenital Scoliosis Etiology and Associations. *Spine,* 34**,** 1745-1750.

HERNANDEZ, L. P., PATTERSON, S. E. & DEVOTO, S. H. 2005. The development of muscle fiber type identity in zebrafish cranial muscles. *Anat Embryol (Berl),* 209**,** 323-34.

HERNANDEZ-LAGUNAS, L., CHOI, I. F., KAJI, T., SIMPSON, P., HERSHEY, C., ZHOU, Y., ZON, L., MERCOLA, M. & ARTINGER, K. B. 2005. Zebrafish narrowminded disrupts the transcription factor prdm1 and is required for neural crest and sensory neuron specification. *Dev Biol,* 278**,** 347-57.

HINITS, Y., WILLIAMS, V. C., SWEETMAN, D., DONN, T. M., MA, T. P., MOENS, C. B. & HUGHES, S. M. 2011. Defective cranial skeletal development, larval lethality and haploinsufficiency in Myod mutant zebrafish. *Dev Biol,* 358**,** 102-12.

HIRATA, H., WATANABE, T., HATAKEYAMA, J., SPRAGUE, S. M., SAINT-AMANT, L., NAGASHIMA, A., CUI, W. W., ZHOU, W. & KUWADA, J. Y. 2007. Zebrafish relatively relaxed mutants have a ryanodine receptor defect, show slow swimming and provide a model of multi-minicore disease. *Development,* 134**,** 2771-81.

HIRATA, H., WEN, H., KAWAKAMI, Y., NAGANAWA, Y., OGINO, K., YAMADA, K., SAINT-AMANT, L., LOW, S. E., CUI, W. W., ZHOU, W., SPRAGUE, S. M., ASAKAWA, K., MUTO, A., KAWAKAMI, K. & KUWADA, J. Y. 2012. Connexin 39.9 protein is necessary for coordinated activation of slow-twitch muscle and normal behavior in zebrafish. *J Biol Chem,* 287**,** 1080-9.

HIRSINGER, E., STELLABOTTE, F., DEVOTO, S. H. & WESTERFIELD, M. 2004. Hedgehog signaling is required for commitment but not initial induction of slow muscle precursors. *Dev Biol,* 275**,** 143-57.

HOFFMANN, B., SCHMIDT-TRAUB, H., PERROT, A., OSTERZIEL, K. J. & GESSNER, R. 2001. First mutation in cardiac troponin C, L29Q, in a patient with hypertrophic cardiomyopathy. *Hum Mutat,* 17**,** 524.

HOLLWAY, G. E., MAULE, J., GAUTIER, P., EVANS, T. M., KEENAN, D. G., LOHS, C., FISCHER, D., WICKING, C. & CURRIE, P. D. 2006. Scube2 mediates Hedgehog signalling in the zebrafish embryo. *Dev Biol,* 294**,** 104-18.

HONJO, Y. & EISEN, J. S. 2005. Slow muscle regulates the pattern of trunk neural crest migration in zebrafish. *Development,* 132**,** 4461-70.

HOWE, K., CLARK, M. D., TORROJA, C. F., TORRANCE, J., BERTHELOT, C., MUFFATO, M., COLLINS, J. E., HUMPHRAY, S., MCLAREN, K., MATTHEWS, L., MCLAREN, S., SEALY, I., CACCAMO, M., CHURCHER, C., SCOTT, C., BARRETT, J. C., KOCH, R., RAUCH, G. J., WHITE, S., CHOW, W., KILIAN, B., QUINTAIS, L. T., GUERRA-ASSUNCAO, J. A., ZHOU, Y., GU, Y., YEN, J., VOGEL, J. H., EYRE, T., REDMOND, S., BANERJEE, R., CHI, J. X., FU, B. Y., LANGLEY, E., MAGUIRE, S. F., LAIRD, G. K., LLOYD, D., KENYON, E., DONALDSON, S., SEHRA, H., ALMEIDA-KING, J., LOVELAND, J., TREVANION, S., JONES, M., QUAIL, M., WILLEY, D., HUNT, A., BURTON, J., SIMS, S., MCLAY, K., PLUMB, B., DAVIS, J., CLEE, C., OLIVER, K., CLARK, R., RIDDLE, C., ELIOTT, D., THREADGOLD, G., HARDEN, G., WARE, D., MORTIMER, B., KERRY, G., HEATH, P., PHILLIMORE, B., TRACEY, A., CORBY, N., DUNN, M., JOHNSON, C., WOOD, J., CLARK, S., PELAN, S., GRIFFITHS, G., SMITH, M., GLITHERO, R., HOWDEN, P., BARKER, N., STEVENS, C., HARLEY, J., HOLT, K., PANAGIOTIDIS, G., LOVELL, J., BEASLEY, H., HENDERSON, C., GORDON, D., AUGER, K., WRIGHT, D., COLLINS, J., RAISEN, C., DYER, L., LEUNG, K., ROBERTSON, L., AMBRIDGE, K., LEONGAMORNLERT, D., MCGUIRE, S., GILDERTHORP, R., GRIFFITHS, C., MANTHRAVADI, D., NICHOL, S., BARKER, G., WHITEHEAD, S., KAY, M., et al. 2013. The zebrafish reference genome sequence and its relationship to the human genome. *Nature,* 496**,** 498-503.

IGUCHI, H., URASHIMA, Y., INAGAKI, Y., IKEDA, Y., OKAMURA, M., TANAKA, T., UCHIDA, A., YAMAMOTO, T. T., KODAMA, T. & SAKAI, J. 2007. SOX6 suppresses cyclin D1 promoter activity by interacting with beta-catenin and histone deacetylase 1, and its down-regulation induces pancreatic beta-cell proliferation. *J Biol Chem,* 282**,** 19052-61.

JOHN, S. A. & GARRETT-SINHA, L. A. 2009. Blimp1: a conserved transcriptional repressor critical for differentiation of many tissues. *Exp Cell Res,* 315**,** 1077-84.

KAMACHI, Y., UCHIKAWA, M. & KONDOH, H. 2000. Pairing SOX off: with partners in the regulation of embryonic development. *Trends Genet,* 16**,** 182-7.

KARLSTROM, R. O., TALBOT, W. S. & SCHIER, A. F. 1999. Comparative synteny cloning of zebrafish you-too: mutations in the Hedgehog target gli2 affect ventral forebrain patterning. *Genes Dev,* 13**,** 388-93.

KARSENTY, G. & WAGNER, E. F. 2002. Reaching a genetic and molecular understanding of skeletal development. *Developmental Cell,* 2**,** 389-406.

KAWAKAMI, A., NOJIMA, Y., TOYODA, A., TAKAHOKO, M., SATOH, M., TANAKA, H., WADA, H., MASAI, I., TERASAKI, H., SAKAKI, Y., TAKEDA, H. & OKAMOTO, H. 2005. The zebrafish-secreted matrix protein you/scube2 is implicated in long-range regulation of hedgehog signaling. *Curr Biol,* 15**,** 480-8.

KELLIS, M., BIRREN, B. W. & LANDER, E. S. 2004. Proof and evolutionary analysis of ancient genome duplication in the yeast Saccharomyces cerevisiae. *Nature,* 428**,** 617-24.

KETTLEBOROUGH, R. N. W., BUSCH-NENTWICH, E. M., HARVEY, S. A., DOOLEY, C. M., DE BRUIJN, E., VAN EEDEN, F., SEALY, I., WHITE, R. J., HERD, C., NIJMAN, I. J., FENYES, F., MEHROKE, S., SCAHILL, C., GIBBONS, R., WALI, N., CARRUTHERS, S., HALL, A., YEN, J., CUPPEN, E. & STEMPLE, D. L. 2013. A systematic genome-wide analysis of zebrafish protein-coding gene function. *Nature,* 496**,** 494-+.

KIM, H. R., RICHARDSON, J., VAN EEDEN, F. & INGHAM, P. W. 2010. Gli2a protein localization reveals a role for Iguana/DZIP1 in primary ciliogenesis and a dependence of Hedgehog signal transduction on primary cilia in the zebrafish. *BMC Biol,* 8**,** 65.

KIMMEL, C. B., BALLARD, W. W., KIMMEL, S. R., ULLMANN, B. & SCHILLING, T. F. 1995. Stages of embryonic development of the zebrafish. *Dev Dyn,* 203**,** 253-310.

KNOTTS, S., RINDT, H., NEUMANN, J. & ROBBINS, J. 1994. In vivo regulation of the mouse beta myosin heavy chain gene. *J Biol Chem,* 269**,** 31275-82.

KOBAYASHI, M., OSANAI, H., KAWAKAMI, K. & YAMAMOTO, M. 2000. Expression of three zebrafish Six4 genes in the cranial sensory placodes and the developing somites. *Mech Dev,* 98**,** 151-5.

KOOPMAN, P., SCHEPERS, G., BRENNER, S. & VENKATESH, B. 2004. Origin and diversity of the SOX transcription factor gene family: genome-wide analysis in Fugu rubripes. *Gene,* 328**,** 177-86.

KOUDIJS, M. J., DEN BROEDER, M. J., GROOT, E. & VAN EEDEN, F. J. 2008. Genetic analysis of the two zebrafish patched homologues identifies novel roles for the hedgehog signaling pathway. *BMC Dev Biol,* 8**,** 15.

KOUDIJS, M. J., DEN BROEDER, M. J., KEIJSER, A., WIENHOLDS, E., HOUWING, S., VAN ROOIJEN, E. M., GEISLER, R. & VAN EEDEN, F. J. 2005. The zebrafish mutants dre, uki, and lep encode negative regulators of the hedgehog signaling pathway. *PLoS Genet,* 1**,** e19.

KRAUSS, S., CONCORDET, J. P. & INGHAM, P. W. 1993. A functionally conserved homolog of the Drosophila segment polarity gene hh is expressed in tissues with polarizing activity in zebrafish embryos. *Cell,* 75**,** 1431-44.

KWAN, K. M., FUJIMOTO, E., GRABHER, C., MANGUM, B. D., HARDY, M. E., CAMPBELL, D. S., PARANT, J. M., YOST, H. J., KANKI, J. P. & CHIEN, C. B. 2007. The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Dev Dyn,* 236**,** 3088-99.

LANE, D., PRENTKI, P. & CHANDLER, M. 1992. Use of Gel Retardation to Analyze Protein-Nucleic Acid Interactions. *Microbiological Reviews,* 56**,** 509-528.

LAURIN, M., FRADET, N., BLANGY, A., HALL, A., VUORI, K. & COTE, J. F. 2008. The atypical Rac activator Dock180 (Dock1) regulates myoblast fusion in vivo. *Proc Natl Acad Sci U S A,* 105**,** 15446-51.

LEE, E. C., YU, D., MARTINEZ DE VELASCO, J., TESSAROLLO, L., SWING, D. A., COURT, D. L., JENKINS, N. A. & COPELAND, N. G. 2001. A highly efficient Escherichia coli-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. *Genomics,* 73**,** 56-65.

LEE, W. S., CHEUNG, W. H., QIN, L., TANG, N. & LEUNG, K. S. 2006. Age-associated decrease of type IIA/B human skeletal muscle fibers. *Clin Orthop Relat Res,* 450**,** 231-7.

LEFEBVRE, V., LI, P. & DE CROMBRUGGHE, B. 1998. A new long form of Sox5 (L-Sox5), Sox6 and Sox9 are coexpressed in chondrogenesis and cooperatively activate the type II collagen gene. *EMBO J,* 17**,** 5718-33.

LEHMAN, W., CRAIG, R. & VIBERT, P. 1994. Ca(2+)-induced tropomyosin movement in Limulus thin filaments revealed by three-dimensional reconstruction. *Nature,* 368**,** 65-7.

LEWIS, K. E., CURRIE, P. D., ROY, S., SCHAUERTE, H., HAFFTER, P. & INGHAM, P. W. 1999. Control of muscle cell-type specification in the zebrafish embryo by Hedgehog signalling. *Dev Biol,* 216**,** 469-80.

LI, E. & DAVIDSON, E. H. 2009. Building developmental gene regulatory networks. *Birth Defects Res C Embryo Today,* 87**,** 123-30.

LIEW, H. P., CHOKSI, S. P., WONG, K. N. & ROY, S. 2008. Specification of vertebrate slow-twitch muscle fiber fate by the transcriptional regulator Blimp1. *Dev Biol,* 324**,** 226-35.

LIVAK, K. J. & SCHMITTGEN, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods,* 25**,** 402-8.

LONSTEIN, J. E. 1994. Adolescent Idiopathic Scoliosis. *Lancet,* 344**,** 1407-1412.

LYONS, G. E., ONTELL, M., COX, R., SASSOON, D. & BUCKINGHAM, M. 1990. The expression of myosin genes in developing skeletal muscle in the mouse embryo. *J Cell Biol,* 111**,** 1465-76.

MACDONALD, R. B., DEBIAIS-THIBAUD, M., TALBOT, J. C. & EKKER, M. 2010. The relationship between dlx and gad1 expression indicates highly conserved genetic pathways in the zebrafish forebrain. *Dev Dyn,* 239**,** 2298-306.

MAURO, A. 1961. Satellite cell of skeletal muscle fibers. *J Biophys Biochem Cytol,* 9**,** 493-5.

MAURYA, A. K., TAN, H., SOUREN, M., WANG, X., WITTBRODT, J. & INGHAM, P. W. 2011. Integration of Hedgehog and BMP signalling by the engrailed2a gene in the zebrafish myotome. *Development,* 138**,** 755-65.

MAVES, L., WASKIEWICZ, A. J., PAUL, B., CAO, Y., TYLER, A., MOENS, C. B. & TAPSCOTT, S. J. 2007. Pbx homeodomain proteins direct Myod activity to promote fast-muscle differentiation. *Development,* 134**,** 3371-82.

MCCARTHY, J. J., ESSER, K. A., PETERSON, C. A. & DUPONT-VERSTEEGDEN, E. E. 2009. Evidence of MyomiR network regulation of beta-myosin heavy chain gene expression during skeletal muscle atrophy. *Physiol Genomics,* 39**,** 219-26.

MCGAUGHEY, D. M., VINTON, R. M., HUYNH, J., AL-SAIF, A., BEER, M. A. & MCCALLION, A. S. 2008. Metrics of sequence constraint overlook regulatory sequences in an exhaustive analysis at phox2b. *Genome Res,* 18**,** 252-60.

MEISSNER, G. 1994. Ryanodine Receptor Ca2+ Release Channels and Their Regulation by Endogenous Effectors. *Annual Review of Physiology,* 56**,** 485-508.

MELANCON, E., LIU, D. W., WESTERFIELD, M. & EISEN, J. S. 1997. Pathfinding by identified zebrafish motoneurons in the absence of muscle pioneers. *J Neurosci,* 17**,** 7796-804.

MENG, X., NOYES, M. B., ZHU, L. J., LAWSON, N. D. & WOLFE, S. A. 2008. Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. *Nat Biotechnol,* 26**,** 695-701.

MOGENSEN, J., MURPHY, R. T., SHAW, T., BAHL, A., REDWOOD, C., WATKINS, H., BURKE, M., ELLIOTT, P. M. & MCKENNA, W. J. 2004. Severe disease expression of cardiac troponin C and T mutations in patients with idiopathic dilated cardiomyopathy. *J Am Coll Cardiol,* 44**,** 2033-40.

MOORE, C. A., PARKIN, C. A., BIDET, Y. & INGHAM, P. W. 2007. A role for the Myoblast city homologues Dock1 and Dock5 and the adaptor proteins Crk and Crk-like in zebrafish myoblast fusion. *Development,* 134**,** 3145-53.

MOOTOOSAMY, R. C. & DIETRICH, S. 2002. Distinct regulatory cascades for head and trunk myogenesis. *Development,* 129**,** 573-83.

NAGANAWA, Y. & HIRATA, H. 2011. Developmental transition of touch response from slow muscle-mediated coilings to fast muscle-mediated burst swimming in zebrafish. *Dev Biol,* 355**,** 194-204.

NAKANO, Y., KIM, H. R., KAWAKAMI, A., ROY, S., SCHIER, A. F. & INGHAM, P. W. 2004. Inactivation of dispatched 1 by the chameleon mutation disrupts Hedgehog signalling in the zebrafish embryo. *Dev Biol,* 269**,** 381-92.

NAYA, F. J., MERCER, B., SHELTON, J., RICHARDSON, J. A., WILLIAMS, R. S. & OLSON, E. N. 2000. Stimulation of slow skeletal muscle fiber gene expression by calcineurin in vivo. *J Biol Chem,* 275**,** 4545-8.

NGUYEN-CHI, M. E., BRYSON-RICHARDSON, R., SONNTAG, C., HALL, T. E., GIBSON, A., SZTAL, T., CHUA, W., SCHILLING, T. F. & CURRIE, P. D. 2012. Morphogenesis and cell fate determination within the adaxial cell equivalence group of the zebrafish myotome. *PLoS Genet,* 8**,** e1003014.

NIRO, C., DEMIGNON, J., VINCENT, S., LIU, Y., GIORDANI, J., SGARIOTO, N., FAVIER, M., GUILLET-DENIAU, I., BLAIS, A. & MAIRE, P. 2010. Six1 and Six4 gene expression is necessary to activate the fast-type muscle gene program in the mouse primary myotome. *Dev Biol,* 338**,** 168-82.

NORTH, K. N., YANG, N., WATTANASIRICHAIGOON, D., MILLS, M., EASTEAL, S. & BEGGS, A. H. 1999. A common nonsense mutation results in alpha-actinin-3 deficiency in the general population. *Nat Genet,* 21**,** 353-4.

PALSTRA, A. P., TUDORACHE, C., ROVIRA, M., BRITTIJN, S. A., BURGERHOUT, E., VAN DEN THILLART, G. E., SPAINK, H. P. & PLANAS, J. V. 2010. Establishing zebrafish as a novel exercise model: swimming economy, swimming-enhanced growth and muscle growth marker gene expression. *PLoS One,* 5**,** e14483.

PARMACEK, M. S., IP, H. S., JUNG, F., SHEN, T., MARTIN, J. F., VORA, A. J., OLSON, E. N. & LEIDEN, J. M. 1994. A novel myogenic regulatory circuit controls slow/cardiac troponin C gene transcription in skeletal muscle. *Mol Cell Biol,* 14**,** 1870-85.

PARMACEK, M. S. & LEIDEN, J. M. 1991. Structure, function, and regulation of troponin C. *Circulation,* 84**,** 991-1003.

PARMACEK, M. S., VORA, A. J., SHEN, T., BARR, E., JUNG, F. & LEIDEN, J. M. 1992. Identification and characterization of a cardiac-specific transcriptional regulatory element in the slow/cardiac troponin C gene. *Mol Cell Biol,* 12**,** 1967-76.

PIETRI, T., MANALO, E., RYAN, J., SAINT-AMANT, L. & WASHBOURNE, P. 2009. Glutamate drives the touch response through a rostral loop in the spinal cord of zebrafish embryos. *Dev Neurobiol,* 69**,** 780-95.

POWELL, G. T. & WRIGHT, G. J. 2011. Jamb and jamc are essential for vertebrate myocyte fusion. *PLoS Biol,* 9**,** e1001216.

QUIAT, D., VOELKER, K. A., PEI, J., GRISHIN, N. V., GRANGE, R. W., BASSEL-DUBY, R. & OLSON, E. N. 2011. Concerted regulation of myofiber-specific gene expression and muscle performance by the transcriptional repressor Sox6. *Proc Natl Acad Sci U S A,* 108**,** 10196-201.

REIFERS, F., BOHLI, H., WALSH, E. C., CROSSLEY, P. H., STAINIER, D. Y. & BRAND, M. 1998. Fgf8 is mutated in zebrafish acerebellar (ace) mutants and is required for maintenance of midbrain-hindbrain boundary development and somitogenesis. *Development,* 125**,** 2381-95.

ROTHBAUER, U., ZOLGHADR, K., MUYLDERMANS, S., SCHEPERS, A., CARDOSO, M. C. & LEONHARDT, H. 2008. A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. *Mol Cell Proteomics,* 7**,** 282-9.

ROY, S. & NG, T. 2004. Blimp-1 specifies neural crest and sensory neuron progenitors in the zebrafish embryo. *Curr Biol,* 14**,** 1772-7.

ROY, S., WOLFF, C. & INGHAM, P. W. 2001. The u-boot mutation identifies a Hedgehog-regulated myogenic switch for fiber-type diversification in the zebrafish embryo. *Genes Dev,* 15**,** 1563-76.

SAINT-AMANT, L. & DRAPEAU, P. 1998. Time course of the development of motor behaviors in the zebrafish embryo. *J Neurobiol,* 37**,** 622-32.

SAINT-AMANT, L. & DRAPEAU, P. 2000. Motoneuron activity patterns related to the earliest behavior of the zebrafish embryo. *J Neurosci,* 20**,** 3964-72.

SAINT-AMANT, L. & DRAPEAU, P. 2001. Synchronization of an embryonic network of identified spinal interneurons solely by electrical coupling. *Neuron,* 31**,** 1035-46.

SALIC, A. & MITCHISON, T. J. 2008. A chemical method for fast and sensitive detection of DNA synthesis in vivo. *Proc Natl Acad Sci U S A,* 105**,** 2415-20.

SCHAUERTE, H. E., VAN EEDEN, F. J., FRICKE, C., ODENTHAL, J., STRAHLE, U. & HAFFTER, P. 1998. Sonic hedgehog is not required for the induction of medial floor plate cells in the zebrafish. *Development,* 125**,** 2983-93.

SCHEER, N. & CAMPOS-ORTEGA, J. A. 1999. Use of the Gal4-UAS technique for targeted gene expression in the zebrafish. *Mech Dev,* 80**,** 153-8.

SCHEPERS, G. E., TEASDALE, R. D. & KOOPMAN, P. 2002. Twenty pairs of sox: extent, homology, and nomenclature of the mouse and human sox transcription factor gene families. *Dev Cell,* 3**,** 167-70.

SCHIAFFINO, S. & REGGIANI, C. 2011. Fiber types in mammalian skeletal muscles. *Physiol Rev,* 91**,** 1447-531.

SCHULTZ, E., GIBSON, M. C. & CHAMPION, T. 1978. Satellite cells are mitotically quiescent in mature mouse muscle: an EM and radioautographic study. *J Exp Zool,* 206**,** 451-6.

SEALE, P., SABOURIN, L. A., GIRGIS-GABARDO, A., MANSOURI, A., GRUSS, P. & RUDNICKI, M. A. 2000. Pax7 is required for the specification of myogenic satellite cells. *Cell,* 102**,** 777-86.

SEGER, C., HARGRAVE, M., WANG, X., CHAI, R. J., ELWORTHY, S. & INGHAM, P. W. 2011. Analysis of Pax7 expressing myogenic cells in zebrafish muscle development, injury, and models of disease. *Dev Dyn,* 240**,** 2440-51.

SIVAKAMASUNDARI, V. & LUFKIN, T. 2012. Bridging the Gap: Understanding Embryonic Intervertebral Disc Development. *Cell Dev Biol,* 1.

SMITS, P. & LEFEBVRE, V. 2003. Sox5 and Sox6 are required for notochord extracellular matrix sheath formation, notochord cell survival and development of the nucleus pulposus of intervertebral discs. *Development,* 130**,** 1135-48.

SMITS, P., LI, P., MANDEL, J., ZHANG, Z., DENG, J. M., BEHRINGER, R. R., DE CROMBRUGGHE, B. & LEFEBVRE, V. 2001. The transcription factors L-Sox5 and Sox6 are essential for cartilage formation. *Dev Cell,* 1**,** 277-90.

SOGAH, V. M., SERLUCA, F. C., FISHMAN, M. C., YELON, D. L., MACRAE, C. A. & MABLY, J. D. 2010. Distinct troponin C isoform requirements in cardiac and skeletal muscle. *Dev Dyn,* 239**,** 3115-23.

SOLNICAKREZEL, L., SCHIER, A. F. & DRIEVER, W. 1994. Efficient Recovery of Enu-Induced Mutations from the Zebrafish Germline. *Genetics,* 136**,** 1401-1420.

SRINIVAS, B. P., WOO, J., LEONG, W. Y. & ROY, S. 2007. A conserved molecular pathway mediates myoblast fusion in insects and vertebrates. *Nat Genet,* 39**,** 781-6.

STELLABOTTE, F., DOBBS-MCAULIFFE, B., FERNANDEZ, D. A., FENG, X. & DEVOTO, S. H. 2007. Dynamic somite cell rearrangements lead to distinct waves of myotome growth. *Development,* 134**,** 1253-7.

STICKNEY, H. L., BARRESI, M. J. & DEVOTO, S. H. 2000. Somite development in zebrafish. *Dev Dyn,* 219**,** 287-303.

SUSTER, M. L., ABE, G., SCHOUW, A. & KAWAKAMI, K. 2011. Transposon-mediated BAC transgenesis in zebrafish. *Nat Protoc,* 6**,** 1998-2021.

TURNER, C. A., JR., MACK, D. H. & DAVIS, M. M. 1994. Blimp-1, a novel zinc finger-containing protein that can drive the maturation of B lymphocytes into immunoglobulin-secreting cells. *Cell,* 77**,** 297-306.

VAN EEDEN, F. J., GRANATO, M., SCHACH, U., BRAND, M., FURUTANI-SEIKI, M., HAFFTER, P., HAMMERSCHMIDT, M., HEISENBERG, C. P., JIANG, Y. J., KANE, D. A., KELSH, R. N., MULLINS, M. C., ODENTHAL, J., WARGA, R. M., ALLENDE, M. L., WEINBERG, E. S. & NUSSLEIN-VOLHARD, C. 1996. Mutations affecting somite formation and patterning in the zebrafish, Danio rerio. *Development,* 123**,** 153-64.

VAN RAAMSDONK, W., VAN'T VEER, L., VEEKEN, K., HEYTING, C. & POOL, C. W. 1982. Differentiation of muscle fiber types in the teleost Brachydanio rerio, the zebrafish. Posthatching development. *Anat Embryol (Berl),* 164**,** 51-62.

VAN ROOIJ, E., QUIAT, D., JOHNSON, B. A., SUTHERLAND, L. B., QI, X., RICHARDSON, J. A., KELM, R. J., JR. & OLSON, E. N. 2009. A family of microRNAs encoded by myosin genes governs myosin expression and muscle performance. *Dev Cell,* 17**,** 662-73.

VARGA, Z. M., AMORES, A., LEWIS, K. E., YAN, Y. L., POSTLETHWAIT, J. H., EISEN, J. S. & WESTERFIELD, M. 2001. Zebrafish smoothened functions in ventral neural tube specification and axon tract formation. *Development,* 128**,** 3497-509.

VINCENT, S. D., MAYEUF, A., NIRO, C., SAITOU, M. & BUCKINGHAM, M. 2012. Non conservation of function for the evolutionarily conserved prdm1 protein in the control of the slow twitch myogenic program in the mouse embryo. *Mol Biol Evol,* 29**,** 3181-91.

VISEL, A., PRABHAKAR, S., AKIYAMA, J. A., SHOUKRY, M., LEWIS, K. D., HOLT, A., PLAJZER-FRICK, I., AFZAL, V., RUBIN, E. M. & PENNACCHIO, L. A. 2008. Ultraconservation identifies a small subset of extremely constrained developmental enhancers. *Nat Genet,* 40**,** 158-60.

VON HOFSTEN, J., ELWORTHY, S., GILCHRIST, M. J., SMITH, J. C., WARDLE, F. C. & INGHAM, P. W. 2008. Prdm1- and Sox6-mediated transcriptional repression specifies muscle fibre type in the zebrafish embryo. *EMBO Rep,* 9**,** 683-9.

WADA, H., GHYSEN, A., SATOU, C., HIGASHIJIMA, S., KAWAKAMI, K., HAMAGUCHI, S. & SAKAIZUMI, M. 2010. Dermal morphogenesis controls lateral line patterning during postembryonic development of teleost fish. *Dev Biol,* 340**,** 583-94.

WANG, X., ONO, Y., TAN, S. C., CHAI, R. J., PARKIN, C. & INGHAM, P. W. 2011. Prdm1a and miR-499 act sequentially to restrict Sox6 activity to the fast-twitch muscle lineage in the zebrafish embryo. *Development,* 138**,** 4399-404.

WEBSTER, C., SILBERSTEIN, L., HAYS, A. P. & BLAU, H. M. 1988. Fast muscle fibers are preferentially affected in Duchenne muscular dystrophy. *Cell,* 52**,** 503-13.

WEGNER, M. 1999. From head to toes: the multiple facets of Sox proteins. *Nucleic Acids Res,* 27**,** 1409-20.

WEINBERG, E. S., ALLENDE, M. L., KELLY, C. S., ABDELHAMID, A., MURAKAMI, T., ANDERMANN, P., DOERRE, O. G., GRUNWALD, D. J. & RIGGLEMAN, B. 1996. Developmental regulation of zebrafish MyoD in wild-type, no tail and spadetail embryos. *Development,* 122**,** 271-80.

WEIRAUCH, M. T. & HUGHES, T. R. 2010. Conserved expression without conserved regulatory sequence: the more things change, the more they stay the same. *Trends Genet,* 26**,** 66-74.

WERNER, M. H. & BURLEY, S. K. 1997. Architectural transcription factors: proteins that remodel DNA. *Cell,* 88**,** 733-6.

WOLFF, C., ROY, S. & INGHAM, P. W. 2003. Multiple muscle cell identities induced by distinct levels and timing of hedgehog activity in the zebrafish embryo. *Curr Biol,* 13**,** 1169-81.

WOLFFE, A. P. 1994. Architectural transcription factors. *Science,* 264**,** 1100-1.

WOODS, I. G. & TALBOT, W. S. 2005. The you gene encodes an EGF-CUB protein essential for Hedgehog signaling in zebrafish. *PLoS Biol,* 3**,** e66.

YANG, N., MACARTHUR, D. G., GULBIN, J. P., HAHN, A. G., BEGGS, A. H., EASTEAL, S. & NORTH, K. 2003. ACTN3 genotype is associated with human elite athletic performance. *Am J Hum Genet,* 73**,** 627-31.

YU, J., ANGELIN-DUCLOS, C., GREENWOOD, J., LIAO, J. & CALAME, K. 2000. Transcriptional repression by blimp-1 (PRDI-BF1) involves recruitment of histone deacetylase. *Mol Cell Biol,* 20**,** 2592-603.