Investigating the coordinated transcriptional networks regulating *Xenopus* myogenesis

Caitlin McQueen

PhD

University of York

Biology

September 2018

Abstract

Myogenesis is the process by which mesodermal progenitor cells are committed to a myogenic fate and differentiate to form myofibers via a complex transcriptional network largely controlled by Myogenic Regulatory Factors (MRFs). One such MRF, MyoD, has been extensively investigated and its regulatory network shown to include not only protein-coding mRNAs, but non-coding RNAs and epigenetic modifications too, leading to highly coordinated transcriptional activation of target genes. The advances in high-throughput and small-RNA sequencing have revealed that transcriptional output by RNA Polymerase III, like Polymerase II, is differentially regulated between tissues and during development. This indicates that developmental transcriptional networks are potentially more complex than previously thought, involving cross-talk between multiple polymerases.

Polr3G is a subunit unique to RNA Polymerase III, is associated with maintaining pluripotency and proliferation in hESCs and exists in 2 distinct forms in mammals and in *Xenopus*. This thesis presents the distinct expression profiles of Polr3G and its paralog Polr3gL during *Xenopus tropicalis* development, the specific expression of Polr3G in the skeletal muscle lineage and investigation into its regulation by myogenic factors. Custom tRNA *Xenopus* microarrays show the regulation of tRNAs by Polr3G at the earliest point of embryonic transcriptional activation, and the differential regulation of tRNAs in the muscle lineage in response to changes in expression of Polr3G and Polr3gL and expression changes to RNA Polymerase III targets and transcriptional machinery in animal caps undergoing myogenic differentiation. This thesis also presents finding from an RNA-seq study analysing founder (F0) embryos where MyoD was disrupted using CRISPR/Cas9 methods to determine early gene targets of MyoD. Together, this thesis proposes that a conserved downregulation of RNA Polymerase III activity during differentiation is shared between this model and previous cell culture studies but that the novel factors, Polr3G and Polr3gL, may be implicated in its mediation.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I'd like to thank my supervisor Betsy Pownall for all of the guidance, support during late-night lab stints and new protocols, beers and laughs over the past five years. Her enthusiasm and positive attitude has helped keep the faith even during the hardest points. I'd also like to thank my co-supervisor Bob for his thoughtful suggestions, tea and biscuits during my PhD. Thanks also to Harv Isaacs for all of his help with large datasets and helpful tips and ideas during lab meetings.

I'd like to also thank the rest of the Frog (and Fish) lab for their friendship and teamwork throughout my PhD. To Gideon and Lewis for their cakes, TV suggestions and general chatter. To Ali for his quiet company during tough months of injections and for his awesome teamwork with western coordination. And to Mike, for his running advice, encouragement and positive attitude. My friends Rosie, for being an excellent running partner; Jaydene, for sending positive vibes from Suffolk; and Jordan, for his patience with my occasional uselessness over the last few years.

To my parents and sister who have given their constant support, and encouragement in the background throughout my time at York, have listened to my lab-related rambles, and chipped in even if they didn't follow everything and always made sure I have rested when needed. Thanks for making me feel smart no matter what. Finally, this thesis is dedicated to Cameron Brown. Without your love and companionship, this PhD would not have been possible. You are my world and I thank you a million times over for everything that you do to make our home a safe and comforting place to return to, and for keeping me positive even when I didn't want to be. #TeamCamlin.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author.

This work has not previously been presented for an award at this, or any other,

University. All sources are acknowledged as References.

Signed…….

1. Introduction

The central dogma of molecular biology is that the flow of genetic information from DNA to Protein involves a two-step process of transcription and translation. Transcriptional and translational outputs define a cell's function, and there are multiple levels of regulation controlling these processes, that regulate and define overall gene expression levels within a cell.

During embryonic development, cells become restricted to particular lineages by activating the expression of specific sets of genes. In animals, 3 RNA Polymerases transcribe DNA; RNA Polymerase I transcribes Ribosomal RNAs; RNA Polymerase II transcribes protein coding mRNAs, microRNAs and other long non-coding RNAs and RNA Polymerase III synthesises small structural RNAs including 5S ribosomal RNA, U6 spliceosomal RNA and tRNAs. Much is known about the regulation of RNA Polymerase II activity during development, and recent advances in high-throughput sequencing has meant that transcriptional networks can be studied even at the single cell level (Briggs et al., 2018; Ramsköld et al., 2012). The regulation of the activity of RNA Polymerases I and III and their roles in development and differentiation however, have received less attention- particularly with regards to their role in development.

1.1. Regulation of Myogenesis by RNA Polymerase II products

Myogenesis is the process of converting multipotent mesodermal cells into terminally differentiated myofibers through a complex network of transcriptional regulation. This program of differentiation is driven largely by a family of closely related transcription factors known as Myogenic Regulatory Factors; MyoD, Myf5, MRF4 and Myogenin (Pownall et al., 2002). Extensive study of the regulation of myogenesis by these four factors has revealed the level of complexity through which their activity is activated, enhanced and also modulated in early development.

Identification of MRFs

Transfection experiments on 10T1/2 fibroblasts were used to determine the factors capable of converting them to myoblasts and revealed that transfection of a single locus could convert cells to a myogenic fate (Pinney et al., 1988) and other studies identified a single cDNA, MyoD, capable of converting a number of different cell types to myoblasts, categorizing it as the first candidate master regulator of myogenesis (Davis et al., 1987).

MyoD is a member of a family of transcription factors including Myf5, MRF4 and Myogenin that all containing a basic helix-loop-helix (bHLH) domain (Pownall et al., 2002). These factors show highly regulated temporal and spatial expression patterns, specific to the muscle lineage, and function in the specification of myogenic lineage and driving terminal differentiation of skeletal muscle.

Knock-out studies in mice have been particularly important in characterising the roles of these factors and during development as well as revealing their partial redundancy. Myf5 and MyoD both act as specification factors of multipotent myogenic progenitors located in the somites. Typically, Myf5 is the earliest MRF to be expressed in the muscle lineages (Hopwood et al., 1991; Ott et al., 1991) and is expressed early in the dorsal-medial somites whilst MyoD expression is induced soon after in the dorsal-lateral somites, giving rise to the trunk/intercostal muscles and body wall/limb muscles respectively. Myf5/MyoD null mice fail to generate skeletal muscle, however, expression of either Myf5 or MyoD alone is sufficient to partially restore determination of skeletal muscle myoblasts, although both individual null mutations result in defects within different muscle lineages (Haldar et al., 2008; Rudnicki et al., 1993).

More recent cell culture studies have also shown that Myf5 and MyoD genome-wide binding patterns are largely similar, and that binding of both factors is associated with histone acetylation. But because Myf5 is expressed first, their model predicts that binding of Myf5 is more associated with histone acetylation and chromatin reorganisation rather than transcriptional activation of targets, whilst ChIP revealed that MyoD recruits RNA Polymerase II to target genes and activates gene expression. This suggests that Myf5 may act prior to MyoD in order to specify the myogenic programme before myogenic differentiation is activated by MyoD once it is expressed (Conerly et al., 2016).

In contrast, Myogenin null mice show defects in development and differentiation of skeletal muscle tissues, despite normal specification of myoblasts (Hasty et al., 1993; Nabeshima et al., 1993). This evidence suggests that Myogenin is crucial in driving terminal differentiation of myoblasts, but is not required for initial

specification of myogenic progenitors. Myogenin is activated later in development than the other factors, is directly activated by MyoD and acts to drive terminal differentiation.

The role of Mrf4 is less well defined- initially being expressed in mice soon after induction of Myf5, its expression decreases before second induction later during myogenesis (Jennings, 1992). Therefore permitting it to potentially have roles in both myogenic determination and differentiation. The expression of an Mrf4 transgene controlled by a Myogenin promoter partially restores full myogenesis supporting its' role in terminal differentiation (Zhu and Miller, 1997). Moreover, the close localisation of Mrf4 to the Myf5 locus meant that in initial null experiments mutation of Myf5 also resulted in disruption of the *mrf4* gene and in Myf5/MyoD null mice with maintained expression of Mrf4, skeletal muscle differentiation is preserved (Kassar-Duchossoy et al., 2004). This indicates that, whilst the MRFs have distinct roles during myogenesis, there is also a level of redundancy among factors (Figure 1.1).

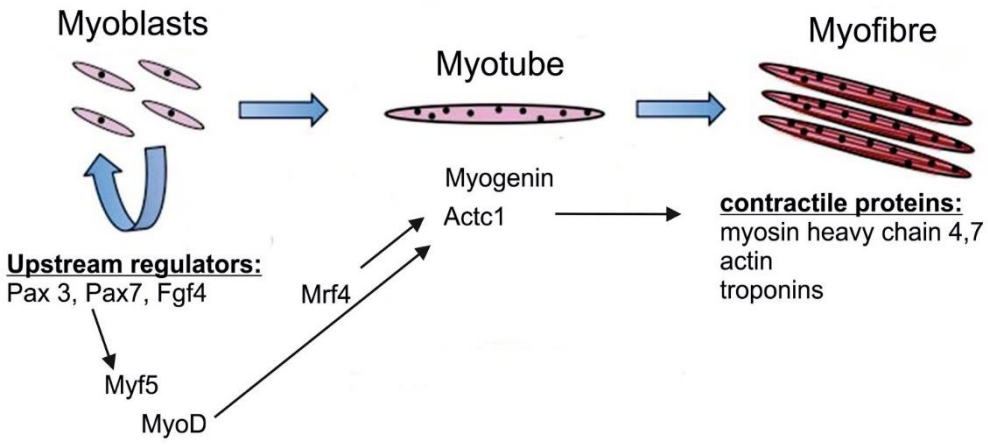


Figure 1.1. Schematic overview of myogenesis and the key bHLH factors that drive specification and differentiation as "master regulators".

As well as their distinct temporal expression patterns during development; MRFs have different molecular features within their functional domains, defining them as either specification factors or differentiation factors. Substitution of the Myogenin coding sequence into the Myf5 locus alone is not sufficient to restore specification of myoblasts and normal development of skeletal muscle indicating that timing alone is not a sole determinant of the role of an MRF (Wang and Jaenisch, 1997). A later study revealed that structural differences between factors enable MRFs to act as specification factors or differentiation factors. A specific motif within the MyoD protein with high sequence conservation throughout vertebrate species was revealed to be key to its function as a myogenic specification factor. The motif is located within the carboxy-terminal domain of alpha Helix III of MyoD, and is also conserved to Myf5. Whilst Helix III of Myogenin has some sequence similarity to this motif, substitution of the Myogenin sequence into a chimeric MyoD protein is not able to recapitulate MyoD activity. Moreover, substitution of the MyoD motif is sufficient to convert Myogenin into a specification factor (Bergstrom and Tapscott, 2001). Therefore, whilst some redundancy is present between these factors, clear structural differences enable them to function distinctly and work collaboratively to induce a wide myogenic transcriptional programme.

1.1.1. The MyoD transcriptional network.

Interaction between MRFs and E-proteins

As part of the bHLH family of transcription factors, the MRFs require dimerization in order to bind to DNA at target promoter regions. They can form both homo-dimers with themselves, or hetero-dimers with other factors such as E-proteins,

which are expressed throughout development in different tissues. The basic domain of bHLH factors is required for DNA binding at target sites whilst the HLH domain facilitates dimerization with partnering E proteins. Together, these dimers bind the specific consensus sequence CANNTG in the DNA, known as E boxes. During Myogenesis, MyoD associates with E12/E47 resulting in activation of transcription of myogenic genes (Lluís et al., 2005). The crucial role of dimerization between MyoD and E47 in the activation of MyoD target genes was shown by tethering of these proteins to force dimerization. Transfection of the MyoD-E47 fusion into $10T_{1/2}$ cells is sufficient to both overcome negative regulation and greatly enhances expression of myogenic genes (Neuhold and Wold, 1993). Importantly, binding at two E-box sites located in close proximity to one another is necessary for activation of transcription by MyoD (Weintraub et al., 1990).

Surprisingly however, it has been reported in C2C12 cells that MyoD binding sites are detected at E-boxes throughout the genome, and not just specifically at myogenic target gene promoters (Cao et al., 2010). Moreover, another study revealed that the overall genome binding patterns of MyoD identified through ChIP-Seq are conserved between endogenous MyoD in myotubes and overexpressed lenti-MyoD expressed in $MyoD^{-/-}$ $Myf5^{-/-}$ MEFs (Yao et al., 2013b). Additionally, overexpression of MyoD in MEFs does not result in promiscuous binding of additional sites in the genome from the reported binding in myotubes, and the introduced MyoD protein still requires the formation of heterodimers with E-proteins in order to bind DNA. This indicates that there is a biological measure to ensure that MyoD binding at the genome-wide sites observed in C2C12 cells is regulated (Yao et al., 2013b).

Interestingly, MyoD binding is not always associated with activation of gene transcription (Cao et al., 2010). This therefore indicates that MyoD occupancy alone is not sufficient to drive gene expression, and that therefore additional regulation exists to activate expression exclusively at target genes. Comparisons of the binding patterns of MyoD and the bHLH neurogenic factor NeuroD2 identified that, whilst both factors bind to shared E-box sites in the genome with a GC core of the consensus, only factor specific binding at private E-box sites is accompanied by activation of lineage specific gene expression and driving of differentiation programs. For MyoD the sequences are CAGGTG and CAGCTG whereas for NeuroD2 this sequence is CAGATG. Private binding at these specific sites is driven both by chromatin accessibility and the presence of co-factor binding motifs i.e. Meis sites are located at MyoD preferred sites (Fong et al., 2012). In addition, by replacing the bHLH domain of MyoD with that of NeuroD2, its role becomes that of a Neurogenic factor, activating NeuroD2 target genes, indicating that MyoD's function as a master regulator of myogenic differentiation depends largely on its protein sequence defining binding site specificity (Fong et al., 2015).

Id proteins (Id1, Id2, and Id3) are HLH factors capable of binding both bHLH proteins and their co-factors and preventing the formation of dimer complexes. Id proteins therefore act as a mechanism of repression of bHLH transcriptional activation (Benezra et al., 1990; Wang and Baker, 2015). In myogenesis, Id proteins interact with both MRFs and E proteins to inhibit differentiation. Evidence exists to suggest that all three Id proteins bind to E-proteins but only two proteins interact with MRFs, specifically Myf5 and MyoD. Therefore suggesting that the

regulation by these proteins is important for early myogenic specification rather than terminal differentiation (Langlands et al., 1997). Expression of Id1 is depleted in terminally differentiated muscle cells, and its overexpression in C2C12 cells causes delays in myogenic differentiation (Jen et al., 1992). Therefore, direct inhibition of E protein-MRF dimerisation is a possible negative regulator of MRF activity during development. However, this requires further investigation**.**

Interaction between MyoD and co-factors

Alone, MyoD is sufficient to drive myogenic specification in many cell types and in subpopulations of myoblasts *in vivo* (Haldar et al., 2008), but full differentiation and expression of contractile proteins requires its interaction with additional factors (Hopwood and Gurdon, 1990; Tapscott, 2005). Therefore, in order to direct full myogenic differentiation, MyoD must activate a wide network of target genes through interaction with assistance through a number of co-factors.

Mef2

The Mef2 family of MADS-box containing transcription factors assists the MRFs in activating transcription of myogenic genes during muscle differentiation. Some factors belonging to this family, *mef2c* for example, are activated directly by MRFs early in myogenic differentiation and subsequently act in coordination with MRFs to drive full myogenic differentiation (Dodou et al., 2003; Naidu et al., 1995; Wang et al., 2001). Unlike MRFs, MEF expression is not restricted to skeletal muscle and Mef overexpression alone cannot convert fibroblasts (Gossett et al., 1989). Despite this, in co-operation with Mef factors, MyoD increases its efficiency of conversion to myoblasts (Molkentin et al., 1995) and at myogenic target genes,

Mef2 binding sites are located close to E-boxes (Wasserman and Fickett, 1998). Activation of target genes by MyoD binding shows distinct temporal regulation, which is in part, controlled by interaction of MyoD with primary target genes, such as Mef2 (Penn et al., 2004).

Pbx

Another family shown to interact with MyoD family members; the Pbx TALE-class homeodomain proteins, form complexes with an additional homeodomain protein, Meis, and act as ‘pioneer” factors which are able to bind to DNA in repressive chromatin in order to mark MyoD target genes such as Myogenin and enable activation transcription by MyoD (Berkes et al., 2004). This is achieved through binding of the protein complex to DNA at specific motifs of target gene promoters (of note the Myogenin promoter), prior to the activation of transcription, and the recruitment of MyoD through cooperative binding via its conserved bHLH H/C and Helix III domains (Knoepfler et al., 1999; Sartorelli, V; Caretti, 2005). In studies comparing the binding activities of MyoD and NeuroD2, MyoD specific binding sites are shown to be enriched for Pbx/Meis binding sites, whilst sites specific for NeuroD2, a neurogenic bHLH factor, are enriched for a different Pbx complex (Fong et al., 2012). Pbx/Meis sites are crucial to the maintenance of MyoD as a myogenic factor, and only in deleting these sequences is MyoD converted to a fully neurogenic factor (Fong et al., 2015). It has been shown that Pbx proteins have a particular role in the differentiation of fast-muscle by MyoD (Maves et al., 2007; Yao et al., 2013a). The Pbx-Meis complex also has roles in other differentiation pathways and interacts with other bHLH factors all through their conserved domains. Thus indicating a conserved mechanism of regulating bHLH factor activity in development.

p38

Factors such as p38 Mitogen Activated Protein Kinase (MAPK) can act as limitation factors to MyoD-dependent transcription, resulting in waves of target gene activation. This is achieved through p38-dependent phosphorylation of E47 which promotes heterodimerisation with MyoD and occupancy at target gene promoters, like that of muscle creatine kinase (MCK) (Lluís et al., 2005). p38 also facilitates MyoD/Mef2 complex assembly and binding to late target gene promoters, which can be activated early through overexpression of p38. This is a regulatory mechanism known as a 'feed-forward' circuit. The MyoD-induced Feed-Forward circuit means some direct targets are activated within 6 hours of MyoD expression, while later targets don't show activation of expression for days after, despite still being directly bound by MyoD (Bergstrom et al., 2002; Tapscott, 2005). The existence of multiple waves of transcriptional activation by MyoD was also shown during myogenic differentiation of P19 cells. Early targets activated by MyoD were identified as premyogenic mesoderm factors Meox1, Six1, and Pax7 and their expression preceded activation of both Mef2c and Myf5. MyoD was shown to bind directly to both early and late targets such as Myogenin, indicating its ability to temporally regulate gene expression through induction of other transcriptional regulators (Gianakopoulos et al., 2011). Together, this data shows that MyoD is capable of activating a complex transcriptional network by interacting with a multitude of cofactors to direct myogenesis in a tightly regulated temporal pattern.

1.2. MicroRNAs

In addition to protein coding genes, RNA Polymerase II activity in myogenesis also results in specific expression of regulatory non-coding RNAs, microRNAs (miRs). These include miR-1, miR-206 and miR-133. MiRs are typically around 22bp long and act to negatively regulate target genes through binding of miRs to sites in UTRs of target genes resulting in either mRNA degradation or translational disruption. miR-1, miR-206 and miR-133 are expressed in the somites of Chick, *Xenopus* and Zebrafish embryos during development (Chen et al., 2006; Goljanek-Whysall et al., 2012; Mishima et al., 2009; Sweetman et al., 2006; Sweetman et al., 2008). Studies in Chick revealed that miR-1 and miR-206 are positively regulated by both Myf5 and Myogenin, whilst MyoD and Mrf4 upregulate miR-206 (Sweetman et al., 2008). The disruption of expression of any one of these miRs leads to disorganisation of muscle structures and downregulated expression of contractile protein genes. Importantly, miR-133 has recently been shown to have a role in regulation of the Shh related effector, Gli3 during chick myogenesis (Mok et al., 2018). In amniotes, Shh signals from axial structures induce somitic expression of myogenic factors. In miR-133 knockdown embryos, Myogenin expression in posterior somites is reduced, and many components of the Shh pathway are significantly downregulated. This effect can be rescued by activation of Shh signalling, therefore indicating that miR-133 is an important mediator of myogenic signals. Together, these data identify that non-coding RNAs act cooperatively with protein-coding RNAs as important regulators of muscle development and further investigation into the roles of additional families of non-coding RNAs might reveal novel mechanisms regulating myogenesis.

1.3. Transcription by RNA Polymerase I

RNA Polymerase I is a 14-subunit complex that transcribes 18S, 28S and 5.8S ribosomal RNAs. These ribosomal RNAs play a catalytic roles within the core of the ribosome and interact with transcripts of RNA Polymerase III, 5S rRNAs (Laferté et al., 2006). As ribosomal RNAs make up the majority of all RNAs within a cell (60%), the rate of RNA Polymerase I activity has a great deal of control over cellular growth and proliferation, and dysregulation of activity is linked with oncogenesis.

Enhanced ribosomal biogenesis is a common feature of transformed cells (White, 2004) and RNA polymerase I activity is activated by oncogenes such as c-Myc, and mitogenic signals such as Erk (White, 2008). It has been shown in paediatric patients with Rhabdomyosarcoma, a cancer affecting the connective tissue, that levels of pre-rRNA transcripts increase in cancer progression giving an indication for tumour prognosis (Williamson et al., 2006). Moreover, whilst RNA Polymerase I activity is required for the maintenance of cell growth and survival, it is regulated by multiple tumour suppressors in normal development. This indicates the importance of tight regulation on RNA Polymerase I activity and ribosome biogenesis for normal cellular function. However, regulation of RNA Polymerase I activity in myogenesis is outside of the scope of this thesis.

1.4 Transcription by RNA Polymerase III

The third of the three nuclear RNA Polymerases in Eukarya, RNA Polymerase III, consists of 17 subunits and is responsible for the transcription of tRNAs, ribosomal 5S RNAs, U6 snRNA, 7SK RNA as well as many other small non-coding RNAs (White, 2011).

The core promoter machinery of Polymerase III consists of a transcription factor TFIIIB. TFIIIB consist of 3 subunits; TATA binding protein (TBP) also utilised by RNA Polymerase I and II, BDP1 and Brf1 or Brf2. Brf1 is required for transcription of tRNAs, which have promoter elements located within the transcribed region, whilst Brf2 is required for transcription of targets such as U6 RNA with promoter elements upstream of the transcription initiation site. ChIP-Seq studies support this, as they show no overlap in the peaks of Brf1 and Brf2 binding (Moqtaderi et al., 2010). It is also interesting to note that Brf1 targets, tRNAs, are much more abundant than Brf2.

In addition to TFIIIB, the transcription factor complex TFIIIC is also required for transcription activation. TFIIIC consists of six subunits, and transcription is dependent on molecular interactions between TFIIIB subunits Brf1/Brf2 with TFIIIC and with RNA Polymerase III itself. The TFIIIC complex recognizes A and B block sequences at target gene promoters, binds and then recruits TFIIIB via protein-protein interactions (Paule and White, 2000).

Protein-protein interactions of these subunits and the recruitment of transcriptional machinery are regulated by a number of embryonic signalling pathways and proteins, and modulates the transcription of RNA Polymerase III target genes during development and differentiation.

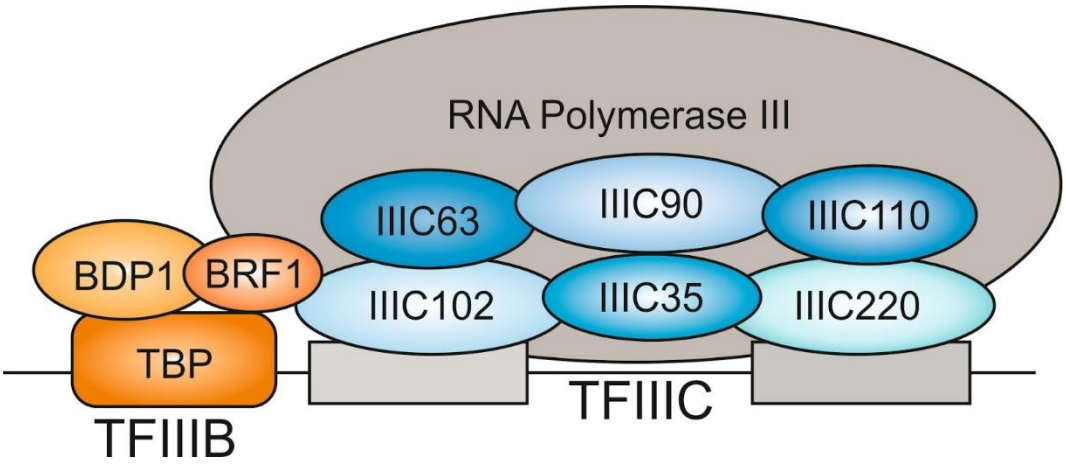


Figure 1.2. Basal transcriptional machinery of RNA Polymerase III at tRNA gene promoters. The RNA Polymerase III enzyme interacts with 2 transcription factor complexes TFIIIB (orange) and TFIIIC (blue). TFIIIC recognizes target gene promoters, binds the DNA, and recruits TFIIIB. TFIIIB consists of the TATA Binding Protein (TBP) which binds TATA boxes recognized at transcriptional target promoters, BDP1 and either Brf1 or Brf2. In activation of transcription at tRNA genes, TFIIIB recruits Brf1 to its complex

1.4.1. Regulation of RNA Polymerase III transcription

mTOR

The protein kinase Target of Rapamycin (TOR) is a regulator of a number of cellular processes and responds to extracellular signals such as stress induced signals, growth factors (IGFs) and nutrient levels. It is conserved throughout eukaryotes, in yeast as TOR1 and TOR2 proteins, through to mammals (mTOR). The activation of mTOR through a number of signalling pathways and external signals initiates a complex downstream signal cascade, resulting in wide cellular responses including altered transcription, enhanced ribosome biogenesis, changes in the microtubule organisation and the actin cytoskeleton, and mitochondrial metabolism (Laplante and Sabatini, 2009; Laplante and Sabatini, 2012).

TOR is a conserved positive regulator of RNA Polymerase III activity throughout eukaryotes. In *Drosophila*, TOR signalling is inhibited through starvation of larvae and results in decreased expression of RNA Polymerase III targets including tRNAs, 5S rRNA and 7SL RNA. This was matched in TOR-null drosophila larvae. In *Drosophila* larvae overexpressing TOR, increases in expression of targets was observed along with cell growth. TOR directly interacts with the Brf subunit of TFIIIB in *Drosophila* to mediate these changes in response to nutrient levels (Marshall et al., 2012). TOR regulation of RNA Polymerase III outputs in response to nutrient availability is a conserved mechanism also found in yeast (Therapy et al., 1999).

In mammals, the homologue mTOR is found bound to tRNA and 5S rRNA genes therefore indicating that conserved TOR mediated regulation of RNA Polymerase III is also important for cell growth control in mammalian development (Kantidakis et al., 2010; Tsang et al., 2010). In mammals however, mTOR associates with TFIIIC in the nucleus, and it is indicated that TFIIIC recruits mTOR to promoters of target genes where it then activates transcription indirectly via phosphorylation-induced inhibition of RNA Polymerase III repressors (Kantidakis et al., 2010; Shor et al., 2010). This mechanism of regulation of RNA Polymerase III activty is crucial for the modulation of cell growth in response to extracellular signals such as stress.

Maf1

RNA Polymerase III transcription is negatively regulated by the conserved protein Maf1 (Kantidakis et al., 2010; Shor et al., 2010). In its active form, Maf1 binds directly to RNA Polymerase III, altering the conformation of its clamp domain sub-

complex, and also binds Brf1 preventing assembly of the TFIIIB complex and recruitment of Polymerase III (Reina et al., 2006). Disruption of basal transcriptional machinery interactions leads to reduced association of RNA Polymerase III with, and activation of its target genes (Desai et al., 2005; Vannini et al., 2010). Repression by Maf1 occurs as a response to multiple cellular signalling pathways including DNA damage and nutrient depletion and in Maf1 mutant cells this repression of RNA Polymerase III targets is abolished (Upadhya et al., 2002).

The ability of Maf1 to repress Polymerase III transcription is regulated by the phosphorylation state of Maf1. The phosphorylated form of Maf1 is unable to relocate to the nucleus and interact with Polymerase III machinery. PKA and TOR signalling both have roles in regulation of Maf1 repression by facilitating the phosphorylation of Maf1 at its site used for RNA Polymerase III interaction (Kantidakis et al., 2010; Moir et al., 2006; Oficjalska-Pham et al., 2006; Shor et al., 2010). Therefore, Polymerase III transcription is closely regulated under normal cellular conditions and is altered accordingly to a number of environmental triggers such as nutrient availability and this becomes dysregulated in cancers (Laplante and Sabatini, 2012; Shor et al., 2010). This indicates that transcription by Polymerase III is under tight regulation during the cell cycle and development, and that this regulation is well coordinated with cellular cues from RNA Polymerase II transcriptional products.

MAPK-ERK signalling

The MAPK-ERK pathway also activates RNA Polymerase III transcription. MAPK is activated by a wide number of external signals and promotes cellular growth, and embryo patterning much like TOR signalling. Inhibition of MAPK-ERK signalling reduced expression of RNA Polymerase III targets whilst inducible overexpression of Raf alone (a component of the MAPK signalling pathway) is sufficient to stimulate enhanced expression of RNA polymerase III transcripts in 3T3 cells (Felton-Edkins et al., 2003a). This is achieved through promotion of TFIIIB interaction with both TFIIIC and the RNA Polymerase III complex, specifically through phosphorylation of the Brf1 subunit, enhancing its activity (Felton-Edkins et al., 2003a). A recent study identified that MAPK-ERK signalling also promoted Brf1-dependent tRNA synthesis in *Drosophila* cultured S2 cells, and this upregulation enhanced rates of protein synthesis and proliferation in *Drosophila* epithelial and stem cells. The study revealed that MAPK-ERK signalling also upregulates RNA Polymerase III activity through inhibition of nuclear localisation of the RNA Polymerase III inhibitor Maf1 and its repressor activity (Sriskanthadevan-Pirahas et al., 2018). This indicates that Erk signalling may have multiple roles promoting the regulation of RNA Polymerase III activity during development, however, the context of these roles during embryogenesis and differentiation has still not been investigated. This is an interesting route of investigation for this thesis as the earliest MAPK-ERK activity present during *Xenopus* development is the result of FGF signalling, and results in mesoderm induction and, amongst other differentiation programmes, the activation of myogenic differentiation. It could therefore be that MAPK-ERK mediated activity of RNA Polymerase III has a role indirecting myogenesis.

1.5. Regulation of RNA Polymerase III activity is important for normal cell growth and division

1.5.1. Regulation of TFIIIB- interaction with Rb and p53

Rb

Transcription by RNA Polymerase III is regulated by a number of signalling pathways and directly by transcription factors. Importantly, the tumour suppressors Rb and p53, regulated by mitogenic signalling pathways, have been shown to directly regulate RNA Polymerase III activity by disrupting protein-protein interactions between components of the basal transcription machinery.

Rb is a cell cycle regulator with a role in regulating cell division through G1 phase progression, and also has a role in differentiation pathways (Giacinti and Giordano, 2006). Rb is part of a tumour-suppressor pathway which is dysregulated so frequently in cancers that it is recognized as a “hallmark” of oncogenesis (Hanahan and Weinberg, 2000; Sherr and McCormick, 2002). Rb has been shown to bind directly to the BRF1 subunit of TFIIIB (Felton-Edkins et al., 2003b; Sutcliffe et al., 2000), resulting in disruption of protein-protein interactions between the RNA Polymerase III enzyme and its transcriptional machinery, causing reduced transcription of target genes (Figure 1.3.).

p53

The transcription factor p53 is induced in response to DNA damage and stress and acts as a driver of both cell cycle arrest and/or apoptosis. p53 regulates the cell cycle through inducing expression of the CDK inhibitor $p21^{Cip1}$, and activates expression of pro-apoptotic proteins (Shaw, 1996; Sherr and McCormick, 2002).

p53 interacts with the TBP subunit of TFIIIB and disrupts its' interaction with TFIIIC (Figure 1.3). TFIIIB recruitment to target genes is therefore diminished, leading to reduced TFIIIB occupancy at tRNA genes and reduced target gene expression- an effect that can be reversed by overexpression of TBP (Crighton et al., 2003; Felton-Edkins et al., 2003b; Sutcliffe et al., 2000).

The activity of both of these tumour suppressors results in decreased transcription by RNA Polymerase III in normal cells. In most cancers, one or both of these proteins are mutated, which may underlie the increased activity of Polymerase III observed in cancer cells and could explain why activity of RNA Polymerase III is deregulated and more active in most, if not all cancers.

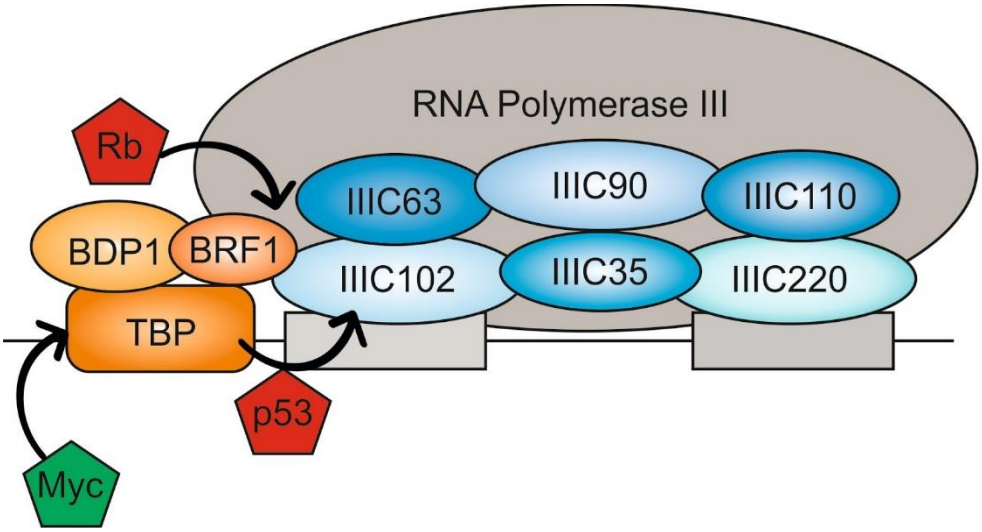


Figure 1.3. Regulation of RNA Polymerase III transcription by tumour suppressors Rb and p53 and the proto-oncogene Myc. Induced by mitogenic signals (anti-growth), transcription by RNA Polymerase III is modulated by Rb and p53. Rb directly associates with Brf subunits of TFIIIB and p53 interacts with the TBP subunit to disrupt protein-protein interactions with the Polymerase III complex and TFIIIC respectively. Myc has a role in the promotion of transcription by RNA Polymerase III during cancer transformation and recruits RNA Polymerase III to target genes through interaction with TFIIIB. Red pentagons represent negative regulation and green indicate positive regulation.

1.5.2. Dysregulation of the RNA Polymerase III core promoter machinery in cancer

In transformed cells and cancers, dysregulation of RNA Polymerase III activity and the hyper-activation of RNA Polymerase III target genes is common. This can be achieved through overexpression and enhanced activity of core promoter elements. The proto-oncogene c-Myc is a transcription factor with roles in cell growth and proliferation. Under normal conditions, developmental signalling pathways tightly regulate activity of c-Myc, and the unstable nature of c-Myc, means that sustained positive regulation of expression is required (Miller et al., 2012). However, dysregulation of c-Myc through over-activation, mutations of the *c-myc* gene itself, and chromosomal translocations, is found in a number of different cancers (Dang, 1999; Schmidt, 1999).

c-Myc directly interacts with TFIIIB *in vivo* (Figure 1.3), resulting in enhanced recruitment of RNA Polymerase III to target promoters and activation of RNA Polymerase III transcription- overexpression of tRNA and 5S rRNA targets. This is supported by the observation using ChIP, that endogenous c-Myc is located at RNA Polymerase III target gene promoters such as tRNAs and 5S rRNA (Felton-Edkins et al., 2003b; Gomez-Roman et al., 2003). However, not all genes occupied by RNA Polymerase III were also occupied by c-Myc. Therefore indicating that, whilst these factors interact directly, occupancy by RNA Polymerase III does not completely dictate c-Myc occupancy.

Other studies have also shown that components of the core RNA Polymerase III transcription machinery have elevated activity in transformed cells and in primary cancer cells (Marshall and White, 2008; White, 2004). The Brf1 and Brf2 subunits of TFIIIB are differentially regulated in some cancer cell lines, and whilst Brf2

activity is enhanced specifically much more-so than Brf1, Brf1 dependent expression of VA1 RNA is elevated higher than Brf2-dependent transcription of U6 RNA (Cabarcas et al., 2008). U6 RNA is however, upregulated in some cancers such as breast cancer (Appaiah et al., 2011), and as a Brf2 target of RNA Polymerase III, this could therefore be a direct effect of enhanced Brf2 activity.

It has been found in some human ovarian cancers that the activity of subunit complex TFIIIC is also upregulated (Winter et al., 2000). Although the study was limited to only nine samples, TFIIC was upregulated in all nine, and all 5 tested subunits of TFIIIC were overexpressed. TFIIIC binds to DNA at target gene sites and recruits TFIIIB to activate transcription, therefore suggesting that elevated activity of this factor alone also results in increased and dysregulated transcriptional output. In breast cancer for example, analysis of tRNA expression levels shows significantly elevated overall levels when compared with non-cancer samples (Pavon-Eternod et al., 2009).

1.5.3. Elevated RNA Polymerase III activity drives cancer progression

Moreover, it has been indicated that overexpression of RNA Polymerase III activity and expression of particular target genes drives cancer progression, with wider transcriptional effect. Overexpression of the tRNA initiator Methionine ($tRNA^{iMet}$) has been shown to alter expression of other tRNAs when overexpressed in cancer cell lines, increasing expression of tRNAs for charged amino acids, and also increases cellular proliferation relative to non-tumour derived cells (Pavon-Eternod et al., 2013). Although this study was based on a single overexpression clone, these data suggest that RNA Polymerase III transcription is necessary for correct

cellular function, however must be modulated in order to maintain normal proliferation levels.

A more recent study identified particular tRNA isoacceptors as promoters of cancer metastasis when overexpressed. $tRNA^{Glu}_{UUC}$ and $tRNA^{Arg}_{CCG}$ in particular were elevated in metastatic cells compared with parental cells. When these tRNAs were overexpressed in non-metastatic cells, increased cell invasiveness and promotion of metastasis was observed (Goodarzi et al., 2016). These results indicate that dysregulation of tRNAs is important for multiple stages of cancer progression, therefore making RNA Polymerase III activity a good target for novel therapeutics in treating cancer (Grewal, 2015).

1.6. tRNAs and the genetic code

Quantitatively, the most significant role of RNA Polymerase III activity is the transcription of tRNAs. tRNAs are small non-coding RNAs generally consisting of between 70-100 base pairs in length. tRNAs contain an amino acceptor stem, which functions to recruit amino acids to translation machinery and importantly, an anticodon loop, which pairs with mRNA using Watson-Crick base pairing enabling tRNAs to convert transcriptional messages into translational output.

There are 64 codons in the genetic code of which 61 encode amino acids and three encode stop signals. Due to wobble pairing, whereby the 3rd base in an mRNA codon can undergo non-Watson-Crick base pairing with the 1st base of a tRNA anticodon, the number of codons outweighs the number of specific tRNAs encoded by a genome. Multiple tRNA anticodon families ***(isoacceptors)*** encode each amino acid isotype- in *Xenopus* there are 54 isoacceptor families encoding

the 20 standard amino acid isotypes. Each anticodon isoacceptor in turn is encoded for by multiple tRNA genes. In *Xenopus tropicalis*, it is estimated that there are 2638 tRNA genes (Chan and Lowe, 2016; Chan, P.P. & Lowe, 2009). As a result of gene duplication, up to hundreds of tRNA genes with identical sequences can encode the same triplet anticodon.

Isotype	tRNA Gene Count By Isoacceptor						Total genes
Ala	AGC 67	GGC	CGC 36	TGC 47			150
Arg	ACG 33	GCG 3	CCG 8	TCG 33	CCT 27	TCT 105	209
Asn	ATT 1	GTT 47					48
Asp	ATC	GTC 59					59
Cys	ACA 2	GCA 92					94
Gln	CTG 40	TTG 28					68
Glu	CTC 51	TTC 84					135
Gly	ACC	GCC 85	CCC 105	TCC 80			270
His	ATG 1	GTG 42					43
Ile	AAT 63	GAT 5	TAT 51				119
Leu	AAG 38	GAG 2	CAG 32	TAG 25	CAA 22	TAA 46	165
Lys	CTT 97	TTT 72					169
Met	CAT 138						138
Phe	AAA 2	GAA 58					60
Pro	AGG 60	GGG 1	CGG 36	TGG 53			150
Ser	AGA 40	GGA 3	CGA 23	TGA 41	ACT	GCT 99	206
Thr	AGT 151	GGT	CGT 29	TGT 91			271
Trp	CCA 33						33
Tyr	ATA	GTA 107					107
Val	AAC 55	GAC	CAC 53	TAC 36			144

Table 1.1. Summary of predicted tRNA isoacceptors encoding the 20 amino acid isotypes predicted by the *Xenopus tropicalis* genome (adapted from Lowe et al. gtRNA db)

1.7. Regulation of tRNA transcription by RNA Polymerase III during development

1.7.1. mRNA Codon Usage Bias

Usage bias of mRNA codons refers to the selective usage of specific codons more than others in translation. This phenomenon is evolutionarily conserved, however particular codon usage biases vary between species (Behura and Severson, 2013). mRNA codon usage bias can arise through genetic mutation and selection in particular species whereby weak selective pressures drive usage of translationally favoured codons.

One of the most striking examples of codon-usage bias in invertebrates is that of the silkworm *bombyx mori*. The silk glands of *B.mori* develop rapidly and produces almost exclusively the silk proteins fibroin and sericin. There are two parts to the silk gland, the middle and the posterior. The middle produces the protein sericin which is made up of 31% Serine whereas the posterior produces fibroin which is made up of 46% Glycine, 29% Alanine and 12% Serine. Fibroin sequence analysis revealed the amino acids sequence is significantly enriched for Serine, Glycine and Alanine codons. The mRNAs expressed in the posterior part of the silk gland were shown to be enriched in particular codons for each amino acid: GCU for Alanine; CGA and GGU for Glycine and UCA for Serine (Hentzen et al., 1981; Suzuki and Brown, 1972).

Codon usage bias is also suggested in *Caenorhabditis elegans* whereby the amount of bias in codon usage within genes is skewed by the level at which a gene is expressed. Low levels of usage bias occur in genes with low expression levels whilst high levels of bias is present in genes that are most highly expressed.

Moreover, this bias appears to select only a few translationally favourable mRNA codons (Stenico et al., 1994).

1.7.2. Evidence of tRNA usage bias - lessons from prokaryotes

A common theme throughout both eukaryotes and prokaryotes is that codon usage is driven by selective pressure from translation- the notion that translation efficiency and tRNA levels would drive selection of a few codons that are most favourable for translation through wobble-pairing. The coordination of mRNA codon usage with abundance of specific tRNA pools has been observed in a number of models including both prokaryotes and eukaryotes, however, understanding of the drivers and responders within this relationship is still unclear. In *E.coli*, biased abundance of particular tRNA isoacceptors was observed throughout different growth rates and a correlation with mRNA codon usage frequency was identified (Dong et al., 1996). tRNAs most abundant at slowest growth rates appear to show increased expression moreso than tRNAs expressed at low levels, though this appears to be modest in size of change. The study also revealed that tRNA genes, even when existing in a single operon, showed different transcriptional control in response to changes in codon usage with growth rate increase.

A relationship between codon usage and tRNA levels was also observed in another bacterial species *Bacillus subtilis*, but it was observed that tRNA levels correlated with their gene copy number rather than mRNA codon usage changes. This study argued that mRNA codon usage was in fact determined by a drive for translational efficiency, determined by tRNA abundance (Kanaya et al., 1999).

1.7.3. Evidence of tRNA usage bias - Lessons from invertebrates

The silk gland of the silkworm *Bombyx mori,* due to its striking mRNA codon usage patterns, has been studied in order to elucidate possible correlation between the codon usage observed from mRNAs, and the expression of particular tRNA isoacceptor families. The tRNA isoacceptor families enriched in *B.mori* during silk gland development correlate with the amino acids most abundant in the silk proteins expressed. It was also found that specific Serine tRNA isoacceptors with more favourable translational capacity due to wobble pairing were more highly abundant (Garel and Hentzen, 1974; Hentzen et al., 1981). These data identify that tRNA transcription is modulated by a driving force of amino acid usage and that particular isoacceptors capable of encoding multiple anticodons by wobble pairing rules is favoured to ensure optimal translational efficiency.

A study in *C.elegans* revealed that co-ordination of tRNA abundance with amino acid codon usage bias within mRNA transcripts is present in other organisms too, but that gene copy number of tRNAs also has influence over tRNA abundance (Duret, 2000; Stenico et al., 1994).

In *Drosophila*, all amino acids contribute to codon usage bias and show preference to one of their synonymous codons. This usage bias correlated with predicted highest translational efficiency based on availability of tRNA anticodons. Moreover, when tRNA isoacceptors are more significantly changed in abundance across *Drosophila* development, their corresponding mRNA codons show less biased usage (Moriyama and Powell, 1997). This data suggests that, in lower eukaryotic systems, a number of factors determine tRNA regulation and mRNA codon usage bias but that an overriding driver of coordination is selective pressure from the

need for efficient translation, which is altered under different conditions or at different developmental stages.

1.7.4. Evidence of tRNA usage bias - Lessons from vertebrates

In humans, evidence exists to support the notion that tRNA expression is altered during development to best match the need for efficient translation, and that this demand is set by mRNA codon usage. An initial study into codon usage bias in human genes revealed that tissue specific mRNAs can be distinguished from each other based on their synonymous codon usage (Plotkin et al., 2004). Studies utilising customized tRNA microarrays have shown that tRNA expression is also differentially regulated between tissue types. Comparing reproductive tissues (testes, ovary) with immune tissue (thymus and lymph node) with expression levels in the brain as a control revealed that expression of tRNAs encoding isoacceptors for hydrophobic amino acids are greatly increased and charged amino acid isoacceptor expression is decreased. Reproductive tissues however show notably decreased expression of hydrophobic amino acid isoacceptors and increased expression of those encoding small amino acids (Dittmar et al., 2006). When a comparison to the predicted codon-usage of highly translated tissue specific mRNAs was carried out, a significant correlation was also found in a few anticodon families. This study therefore indicates that RNA Polymerase III activity is differentially regulated between tissues, but that this is independent of mRNA codon bias.

A more recent study used tRNA microarrays coupled with RNA-Seq analysis of mRNAs to determine the differences in regulation between proliferative cells from

primary tumours, cancer cell lines and immortalized cell lines and differentiated normal tissue and embryonic stem cells post-induction of differentiation (Gingold et al., 2014). The results showed that expression of specific tRNA genes is significantly altered in proliferative cells over differentiated, and that some tRNA genes of the same isoacceptor family show opposing trends. This change happens gradually as shown by transitioning cells from proliferative to senescent indicating a steady alteration of RNA Polymerase III activity in distinct cellular programs. Interestingly, mRNAs highly expressed in either proliferating or differentiated cells are enriched in different codons respectively and these codons match with the anticodons of tRNA genes upregulated in each cell type. Together, these results suggest that in human tissues, there exists a coordinated regulation of mRNAs and tRNAs such that changes in tRNA levels match the demand of mRNA codon bias to ensure the high expression of required proteins in each cell type (Topisirovic and Sonenberg, 2014).

1.7.5 Evidence of a stable tRNA-mRNA interface

A recent study using ChIP-Seq to analyse RNA Polymerase III binding in multiple mammalian species determined the level of conservation in binding patterns throughout millions of years of evolution. In the species studied (from mouse to marsupial), RNA Polymerase III binding at individual tRNA genes varied by several orders of magnitude, but this variation was less significant at the isoacceptor level where overall binding appears directly correlated with the number of other isoacceptor families encoding the same amino acid isotype. Patterns of overall occupancy at the isotype level were conserved across the six species, and even across different tissues within a single species. Accompanying RNA-Seq data

revealed that mRNA codon usage across the mammalian species is also well conserved and is unvarying for overall abundance of individual codons (Kutter et al., 2011). This evidence indicates an evolutionarily conserved, coordinated regulation of the mRNA-tRNA interface, in which relative tRNA abundance at individual gene level is significantly varied between diverse tissues throughout mammalian species. Yet, as a result of redundancy of the genetic code and regulatory cross-talk between isoacceptor families, this converts to a stable pool of tRNAs coding for each amino acid isotype between diverse tissues and species throughout millions of years of evolution. This conservation is closely mirrored in mRNA codon usage, therefore suggesting that an underlying mechanism of coordinated regulation is crucial for development and correct cellular function.

During development, regulation of tRNA gene expression shows dynamic regulation. An analysis of tRNA expression in mouse brain and liver at different developmental stages showed that individual genes undergo significant changes in expression throughout their development. However, on observations of overall binding at the isoacceptor level, this is equilibrated as increased expression of a tRNA gene matches a simultaneous decreased expression of another gene in the same isoacceptor family, resulting in a stable pool of tRNAs for each amino acid throughout development (Schmitt et al., 2014). This therefore suggests that developmental cues act to regulate specific tRNA genes to maintain an equilibrated pool of tRNA isotypes rather than to change the expression of tRNA isoacceptor families. This work is supported by a more recent study, which showed that whilst mRNA codon usage differs between proliferative and quiescent cells, tRNA anticodon abundance did not significantly correlate with or differ

between the two cell types (Rudolph et al., 2016). Together, this data shows that as yet, no clear relationship between RNA Polymerase II activity and RNA Polymerase III transcription has been determined across vertebrates. Yet, as many signals, which result in altered transcription of RNA Polymerase II, also lead to changes in the activity of RNA polymerase III, exploring this relationship *in vivo* is of interest.

During myogenesis, a transcriptional network is activated by the MRFs, resulting in the rapid upregulation of multiple large proteins; most significantly Myosins, Actins and Troponins. The transcription of these factors occurs almost simultaneously during tailbud stages in *Xenopus* development and it is therefore likely that myogenic cells undergo a change in demand from translation as a result of the expression of these genes. Therefore, *Xenopus* myogenesis provides a useful model in which to investigate the role of novel factors in regulating both mRNA and tRNA expression to drive differentiation of muscle cells *in vivo*.

1.8. Polr3G. A Polymerase III subunit with distinct developmental regulation

RNA Polymerase III consists of 17 subunits, which are highly evolutionarily conserved from yeast to humans. Most subunits show structural similarity to subunits of RNA Polymerases I and II, with 5 subunits showing no homology to any subunits of RNA Polymerase I or II. Three of these subunits in humans (RPC32, RPC39 and RPC62) form a dissociable sub-complex which directs gene specific transcription initiation through interaction between RPC39 and TFIIIB (Wang and Roeder, 1997).

The subunit RPC32 (Polr3G) is unique to Polymerase III, showing no homology to any Polymerase I or II subunits. It interacts with RPC62 via two alpha-Winged-Helices, and binding causes conformational change to expose Polr3G to the remaining Polymerase III subunits, forming a stable transcriptional complex (Boissier et al., 2015).

As a result of gene duplication, RPC32 exists as two isoforms; RPC32α (Polr3G) and RPC32β (Polr3gL). Proteins with significant homology to Polr3G were identified throughout Eukaryotic species, however only one gene exists in species such as *Saccharomyces cerevisiae* and *C.elegans,* which more closely resembles Polr3gL than Polr3G. It was concluded therefore that Polr3G evolved as a result of a gene duplication event. However, surprisingly it appears that in Chicken (*Gallus gallus*), only one Polr3G gene exists with homology closer to Polr3G than Polr3gL (Renaud et al., 2014).

1.8.1. Distinct functions of Polr3G and Polr3gL

The existence of two distinct isoforms of Polymerase III; α or β, was determined by the inclusion of either Polr3G or Polr3gL into its dissociable sub-complex. Polr3G and Polr3gL show different expression profiles, as Polr3G is expressed at higher levels in embryonic stem cells, and expression then decreases during differentiation. Conversely, Polr3G expression also increases during cell transformation. Polr3gL levels, on the other hand, remain constant throughout differentiation and transformation. siRNA knockdown of each subunit and cell growth assays in hESCs revealed that Polr3gL expression is essential for cell survival, whilst Polr3G is non-essential and that the two factors were non-

redundant (Haurie et al., 2010). These distinct Polymerase III isoforms may have different roles in development, with distinct repertoires of target genes.

Each isoform appears to have a distinct function in embryonic stem cells; ectopic expression of Polr3G results in increased resistance to retinoic acid-induced differentiation (Wong et al., 2011), and its targeting by siRNAs results in decreased anchorage-independent cell growth. Conversely, Polr3gL targeting results in inhibited general cell growth, and ectopic expression results in decreased expression of the cell cycle genes Aurora A kinase, Cyclin E and p27 (Haurie et al., 2010). This suggests that Polr3G is a proliferative, pluripotent isoform while Polr3gL is a differentiation related isoform. However, global ChIP-Seq revealed that, these subunits appear to show few unique targets, and their binding patterns were largely overlapping leaving the defined role of each isoform unclear (Renaud et al., 2014).

More recently, a study in cancer and hESC lines showed that telomerase reverse transcriptase (TERT), a key promoter of oncogenesis and proliferation, directly interacts with the Polr3G subunit of RNA Polymerase III through both its N and C-terminal domains. Interaction with TERT in transformed cells enhanced RNA Polymerase III recruitment to tRNA genes, resulting in increased expression of tRNAs. Moreover, expression levels of TERT positively correlate with increased expression of tRNA genes in cancer cells (Khattar et al., 2016) and transformation of IMR90 cells results in the upregulation of Polr3G along with few other subunits of RNA Polymerase III (Durrieu-Gaillard et al., 2017). These results suggest that Polr3G has a direct role in regulating tRNA gene expression during stem cell proliferation and differentiation, and is dysregulated in cancer through interaction with key oncogenes.

1.8.2. Regulation of Polr3G in cell lines

There has been little study into the regulators of Polr3G expression during vertebrate development and differentiation, with studies limited to hESC cell culture only. Wong et al. conducted promoter analysis using binding site prediction matrices and located two Nanog target sites upstream of the Polr3G transcriptional start site. Subsequent siRNA targeting of Oct4 and Nanog in hESCs resulted in downregulation of Polr3G therefore suggesting that Polr3G is a downstream target of the pluripotency factors it is coexpressed alongside (Wong et al., 2011). Using the ERK signalling inhibitor U0126, evidence also indicated that Polr3G is positively regulated by the MAPK signalling pathway. However, as this pathway is used by multiple signalling factors, identification of specific regulators to place this regulation into developmental context requires more in-depth study.

Polr3G is also negatively regulated by the micro-RNA miR-1305 in stem cells. Overexpression of miR-1305 in hESCs results in upregulated expression of differentiation markers from the primary germ layers (notably, endoderm markers appear downregulated), which was coordinated with downregulation of pluripotency factors. Conversely, inhibition of miR-1305 results in larger colony formation consistent with maintained high proliferation rates. This indicates that expression of miR-1305 is a positive driver of differentiation and reduced cell division in stem cell populations (Jin et al., 2016). Polr3G was identified through target scanning software as a predicted target of miR-1305, which was confirmed through RT-PCR and Western blotting in miR-1305 overexpressing and siRNA targeted cells. In cells overexpressing miR-1305, Polr3G expression is downregulated and Polr3G is upregulated when miR-1305 is inhibited. Further experiments identified that the role for this regulation is in promoting hESC

differentiation as knocking-down of miR-1305 and Polr3G together resulted in increased expression of differentiation markers, whereas miR-1305 knock-down alone did not (Jin et al., 2016).

Together, these findings indicate that Polr3G expression is regulated by different factors during pluripotency and differentiation programs, however it is unclear how cell culture findings relate to *in vivo* differentiation, and if so, whether regulation of Polr3G differs between cell lineages.

1.9. This thesis

The central hypothesis under investigation in this thesis is whether transcription by RNA Polymerase III is regulated during myogenic differentiation, and if the well-known RNA Polymerase II regulator, MyoD, contributes to coordinated regulation of RNA Polymerase II and III activity during skeletal muscle differentiation.

Polr3G is a highly regulated RNA Polymerase III subunit and is the focus of three of the five results chapters presents (Chapters 4, 5 and 6). Chapter 3 investigates transcriptional targets of MyoD using CRISPR/Cas9 targeting and acts as a standalone chapter. Chapter 7 investigates the regulation of RNA polymerase III transcription during myogenic differentiation in animal caps.

The core aims of this thesis are:

1. To determine whether the use of CRISPR/Cas9 as a means of gene targeting in founder populations was sufficiently effective to use as an alternative to, and to enhance results of Antisense Morpholino Oligos in *Xenopus* studies using MyoD as an example.
2. To characterise the temporal and spatial expression profiles of Polr3G and Polr3gL during *Xenopus* development.
3. To characterise the regulation of Polr3G and Polr3gL by mesodermal signalling pathways (FGF, Wnt) and myogenic factors (MyoD) *in vivo*.
4. To determine the regulation of tRNAs by Polr3G and Polr3gL during *Xenopus* development and myogenesis.
5. To determine the regulation of RNA polymerase III targets in response to growth factor induced myogenic differentiation in animal caps.

2. Materials and Methods

2.1. Molecular Biology Methods

2.1.1. Gel electrophoresis

DNA and RNA samples were run on 1-2% Agarose gels stained with Ethidium Bromide in either TAE (50X: 242g/L Tris, 57.1ml/L Glacial Acetic acid, 100ml/L 0.5M EDTA pH8.0) or TBE (10X: 1X buffer at 150-180V. Samples were loaded with 6X gel loading buffer (NEB) alongside 4µl of 2-log DNA ladder (NEB).

2.1.2. Standard RT-PCR setup

Total RNA was extracted and purified as outlined in 2.1.8. cDNA was synthesised using 1µg of RNA and Superscript IV (Thermo-Fisher) as per manufacturer's instructions with random hexamers.

20µl reactions are assembled using 2X PCR Mastermix (Promega):

10µl 2X PCR Mastermix

1µl 10µM Forward Primer

1µl 10µM Reverse Primer

1µl cDNA

7µl H_2O

A 30 cycle PCR is carried out after initial denaturation of 2 minutes at 95˚C

Cycling conditions: 95˚C – 30 seconds

__˚C – 30 seconds (5˚c below Tm˚)

72˚C – 60 seconds/kb product

A final extension of 72˚c for 10 minutes is included.

Primers are listed:

Name	Forward	Reverse	Use
MyoD genomic	TTACTTTGCGCCGTTGCTAT	GTTGCGCAAAATCTCCACTT	CRISPR validation
Polr3G	GGGCTGCCTTTACATTTGACAT	CTCGCCCTCCTCCTCTTCC	RT-PCR and ISH Probe
Polr3gL	AGCGGTCGGGGCCAACTGACCTT	TTTGGGGAGGACAACGGCGG	RT-PCR and ISH Probe
Polr3G-HA	GAGAGACTCGAGACCATGGACTACCCATACGATGTTCCAGATTACGCTGCCAAAGGGAGGGGAAGG	AGAGAGTCTAGATTAGTAGGTGCTTCATCCAT	Full length cloning CS2+
Polr3gL-HA	GAGAGAGAATTCACCATGGACTACCCATACGATGTTCCAGATTACGCTGCTGGAAAAGCAGAGGTAGC	AGAGAGTCTAGATCAGTAGGTGCTTCATCCAT	Full length cloning CS2+
tRNA^{iMet}	AGCAGAGTGGCGCAGCGGAAG	TAGCAGAGGATGGTTTCGATCCATCGA	RT-PCR
tRNA^{eMet}	GCCTCGTTAGCGCAGTCGGTA	TGCCCCGTGTGAGGATCGAAC	RT-PCR
tRNA^{Leu}	GTCAGGATGGCCGAGTGGTCT	TGTCAGAAGTGGGATTTGAACCCA	RT-PCR
tRNA^{Tyr}	CCTTCGATAGCTCAGCTGGTA	TCCTTCGAGCCGGAATTGAAC	RT-PCR
tRNA$^{Ala(AGC)}$	GGGGGATTAGCTCAAATGG	TGGAGGATGCAGGCATCG	RT-PCR(q)
tRNA$^{Ala(CGC)}$	GGGGATGTAGCTCAGTGGT	TGGAGATGCCGGGGATTG	RT-PCR(q)
tRNA$^{Ala(TGC)}$	GGGGATGTAGCTCAGTGGT	TGGAGATGCTGGGGATTGA	RT-PCR(q)
tRNA$^{Ser(GCT)}$	GACGAGGTGGCCGAGTGG	GACGAGGATGGGATTCGAAC	RT-PCR(q)
5S rRNA	GCCATACCACCCTGAAAG	AGGTATTCCCAGGCGGTCT	RT-PCR(q)
U6 RNA	GTGCTTGCTTCGGCAGCACAT	AAAAATATGGAACGCTTCACGAAT	RT-PCR
28S rRNA	GGCGCCCCGGCTGAGGTG	TGAGATCGTTTCGGCCCCAAGAC	RT-PCR
Wt1	AGGCACACAGGCATTAAACC	TGTTGTGGTGACGAACCAAT	RT-PCR

Table 2.1. Primer sequences for RT-PCR analysis of *Xenopus tropicalis* embryos.

2.1.3. Plasmid Transformation

Dcm$^-$/Dam$^-$ competent cells were used for transformation as per the manufacturer's instructions (heat shock at exactly 42˚C for 30 seconds, incubate on ice for 2 minutes, incubate at 37˚C for 1 hour). The media used was LB broth (1L: 5g yeast extract, 10g tryptone, 10g NaCl).

2.1.4. DNA minipreps

Plasmid DNA was prepared by overnight growth of single transformed colonies in 3ml LB with 1µg/ml ampicillin. Colonies were incubated at 37˚C with agitation. Preparations were incubated on ice for a minimum of 20 minutes prior to centrifugation at 10,000rpm for 10 minutes to pellet the bacteria and plasmid extraction was carried out using the QiaPrep® Spin Miniprep kit (Qiagen). DNA was eluted in 50µl volume and concentration was determined by Nanodrop® ND1000.

2.1.5. Purification of DNA

The volume of sample was made up to 200ul using molecular grade H_2O. An equal volume of water-saturated phenol-chloroform was added and the samples were vortexed briefly. Samples were centrifuged for 5 minutes at 13,000rpm and the top aqueous phase was collected. An equal volume of chloroform-isoamyl was added, vortexed, and a second 5 minute spin was carried out. The top phase was collected in a new tube, 0.1 volumes of 3M sodium acetate and 2.5 volumes of 100% ethanol were added and the sample vortexed to mix. The samples were left at -20˚C° for 1 hour-overnight to precipitate. Precipitated DNA was centrifuged for

20 minutes at 4˚C and supernatant was removed. Resulting pellet was washed in 70% ethanol and vacuum dried for 5-10 minutes and resuspended in the required volume of H_2O (30µl-50µl) and stored at -20˚C.

2.1.6. Quantification of DNA, RNA and gRNA by spectrophotometry

The concentration was determined using the Nanodrop® ND1000 spectrophotometer. After initialisation with water, the equipment was blanked on the DNA, RNA or ssDNA setting for DNA, RNA and gRNA samples respectively. 1µl of product was measured to determine concentration in ng/µl.

2.1.7. Western Blot analysis

For analysis of protein expression both endogenous and overexpressed, 20-25 *X.tropicalis* embryos were collected and homogenised in 30µl Phosphosafe™ Extraction Reagent by pipetting. 5µl of 7X cOmplete™, Mini EDTA-free Protease Inhibitor cocktail (Roche) was added to the mix and embryos were immediately incubated at -80˚C for 10 minutes. Samples were centrifuged at 13,000rpm for 20 minutes and resulting supernatants (~40µl) were collected into new tubes and boiled at 95˚C for 5 minutes in 1X volume 2X sample buffer to denature the protein. After boiling, samples were vortexed and particular embryo values were loaded onto 15% Acrylamide/SDS PAGE gels (For 20ml 15% resolving gel: 4.6ml H_2O, 5ml 1.5M Tris-HCl pH8.8, 10.05ml 30% Acrylamide mix, 200µl 10% SDS, 100µl 10% APS, 50µl TEMED) (For 10ml 5% stacking gel: 6.1ml H_2O, 2.5ml 0.5M Tris-HCl pH6.8, 1.3ml 30% Acrylamide mix, 100µl 10% SDS, 50µl 10% APS, 25µl TEMED).

Embryo values:

6. Overexpression/HA tag/GAPDH: 5 embryos
7. Endogenous/Morpholino/GAPDH: 10-12.5 embryos

Samples were run out at 180V for 2 hours alongside 10µl PageRuler™ Prestained Protein Ladder 26616 (ThermoFisher) in Tris-Glycine gel running buffer (3g/l Tris, 14.4g/l Glycine, 10ml/l 10% SDS). After completion, gels were soaked for 10-20 minutes in gel Tris-Glycine transfer buffer (3g/l Tris, 14.4g/l Glycine, 10ml/l 10% SDS, 100ml/l methanol). Gels were then transferred at 90V for 2.5hours at room temperature, or 30V overnight at 4˚C onto a PVDF membrane pre-soaked in methanol for 15 seconds and equilibrated in transfer buffer. Membranes were blocked for 1 hour in 5% Marvel milk powder/PBST (1L: 100ml PBS, 5ml 20% Tween). Primary antibody dilutions were made up in blocking solution and incubation was carried out by rolling at 4˚C overnight. Membranes were washed four times in PBST for 15 minutes and a second block for 30 minutes preceded 1 hour incubation with secondary antibodies at room temperature. Membranes were washed 4 times for 15 minutes and developed using the BM® Chemiluminescence (Roche) substrate kit. The membranes were exposed to Hyperfilm™ECL® (Amersham) until expression was visualised.

Once developed, membranes were stripped by rocking for 30-45 minutes at 55˚C in stripping buffer (For 100ml buffer: 20ml 10% SDS, 12.5ml 0.5M Tris-HCl pH6.8, 800µl β-Mercaptoethanol, 67.5ml H_2O). Membranes were washed extensively in H_2O, followed by multiple (5 minimum) 5 minute PBST washes. Membranes were then placed back into 5% Marvel/PBST blocking solution for 1 hour.

Target	Primary antibody dilution	Secondary antibody dilution
Polr3G (Eurogentec) (EIEAERKLQREWT)	1:10,000 affinity purified	1:2000 anti-rabbit HRP
GAPDH	1:10,000 (Sigma)	1:4000 anti-mouse HRP
Anti-HA	1:1000 (Sigma)	1:4000 anti-mouse HRP

Table 2.2. Primary and secondary antibody dilutions for protein detection by western blot of endogenous and overexpressed samples.

2.1.8. Total RNA extraction

For extraction of total RNA from tissue samples, 10-15 embryos/20 explants were snap frozen using dry ice. Frozen samples were homogenised in a sterile RNase free tube using a tissue homogeniser treated with RNaseZAP™ (Sigma) on ice. Samples were centrifuged for 10 minutes at 13,000rpm to remove excess fatty tissue and yolk. Approximately 1ml supernatant was collected in a new tube. A 5 minute room temperature incubation preceded addition of 200µl chloroform (Sigma). The samples were shaken for 15 seconds and left to stand for 10 minutes. Samples were centrifuged at 13,000rpm for 15 minutes to separated RNA, DNA and protein phases and the RNA phase (top aqueous) was collected. An equal volume of isopropanol was added to the samples, followed by vortexing and 5-10 minute incubation at room temperature to precipitate the RNA. Samples were centrifuged at 13,000rpm for 10 minutes to pellet the RNA. Pellets were washed with 75% ethanol and centrifuged for 10 minutes at 13,000rpm. RNA was dried for 5-10 minutes under vacuum and resuspended in 40µl of nuclease free water. RNA was then purified using the Zymo-Spin™ Clean and Concentrator™ Kit as per the manufacturer's instructions.

2.1.8.2. DNase I treatment of RNA

To remove DNA, RNA samples were incubated at room temperature for 20 minutes with 5ul DNA digestion buffer and 5ul DNase I (Zymo). 100ul RNA binding buffer and 150µl 100% ethanol were mixed with samples and samples were run through Zymo-Spin™ columns. Columns were washed with 400µl RNA Wash Buffer followed by a 30 second centrifuge at 13,000rpm before an additional in-column DNase treatment was carried out by adding 5µl DNase I and 75µlDNA digestion buffer to the column for more rigorous removal of DNA. RNA purification was then continued as per the manufacturer's instructions. Final elution was in 20-50µl H_2O as needed and RNA was stored at -80˚C until required.

2.1.8.3. Zymo-Spin™ Clean and Concentrator™ purification of total RNA

DNase I in-tube treated samples (50µl) were mixed with 100µl (2X volume) of RNA binding buffer and 150µl (1X total volume) 100% ethanol. Samples were loaded into collection columns and centrifuged for 30 seconds at 13,000rpm. After in column DNase I treatment (as above), 400µl RNA Prep buffer was added to the samples and tubes were centrifuged at 13,000rpm for 30 seconds. Columns were washed twice with RNA Wash Buffer, 700µl and 400µl consecutively and the second wash step was followed by a 2 minute spin to ensure the columns were completely dry. Samples were eluted in 20µl of nuclease free H_2O. Sample concentration was measured using the ND-1000 spectrophotometer. For RNA-Seq and microarray samples, RNA Integrity Numbers were measured by running samples on an Agilent® 2100 Bioanalyzer™. RIN scores of 6.5 (Microarray) or 7.0 (RNA-Seq) were determined suitable quality for high-throughput methods.

2.1.9. Design and *in vitro* synthesis of a MyoD gRNA.

The design tool ChopChop (https://chopchop.rc.fas.harvard.edu/) was used to scan the input sequence for suitable Cas9 target sequences including a PAM site. Any off-targets with up to 2 mismatches in the first 20bp of sequence were searched for using the *Xenopus tropicalis* genome version (xenTro3/GCA_000004195.1).

2.1.9.1. gRNA template synthesis

A sequence spanning 71-89bp in the first exon of the XtMyoD coding sequence (5'-TCGTCGTAGAAGTCATCGG-3') on the reverse strand was selected as no off-targets were predicted for this sequence. The 5' primer was designed to include an increased efficiency promoter for transcription by T7 RNA Polymerase as described by (Nakayama et al. 2014), and an added 5' G nucleotide to fit requirements of the T7 polymerase. The resulting forward primer consisted of the sequence:

5′-GCAGCTAATACGACTCACTATAGG TCGTCGTAGAAGTCATCGG GTTTTAGAGCTAGAAATA-3′

An additional gRNA for the Tyrosinase was also designed and synthesised using the forward primer sequence:

5'-GCAGCTAATACGACTCACTATAGGGAAAGGAACATGGTCCCTC GTTTTAGAGCTAGAAATA-3'

The reverse primer used is common to all gRNAs:

5'-AAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCC TTATTT TAACTTGCTATTTCTAGCTCTAAAAC-3' (Nakayama et al., 2014)**.**

2.1.9.2. gRNA transcription

Phusion polymerase was used to amplify the template using 5µM of each primer, annealing at 60°C and extending for 15 seconds for 35 cycles. The resulting templates were checked on a 2% Agarose/TBE gel and 2µl were taken directly from the PCR reaction for in vitro transcription. Transcription reactions were carried out using the Megashortscript® T7 Transcription Kit (Life Technologies) following the manufacturers guidelines:

2µl-8µl template DNA

2µl ATP

2µl CTP

2µl UTP

2µl GTP

2µl T7 enzyme mix

2µl 10X transcription buffer

_µl H_2O (to final volume of 20µl)

Transcription reactions were incubated at 37˚C overnight and were followed by a 15 minute TURBO DNase treatment at 37˚C. gRNA transcripts were purified by phenol-chloroform extraction (add 180µl of H_2O and 200µl phenol-chloroform, centrifuge at 13,000rpm for 5 minutes.) and chloroform extraction (add equal volume of chloroform-isoamyl to samples and centrifuge at 13,000rpm for 5 minutes). Samples were precipitated using NH_4OAc (0.1 volume)/ethanol (2

volumes). gRNA was resuspended in a final volume of 20µl. RNA quality was measured by gel electrophoresis and spectrophotometry reading using the ND1000 analyser. (NB- An optimal concentration of 1.8µg/µl is desired for Cas9 co-injection mixtures).

2.1.9.3 Cas9 protein production (carried out by Olga Moroz, Wilson Lab, Department of Chemistry, York)

Cas9 protein was made using a plasmid (Addgene) by expression in Rosetta-2 cells (Novagen) at 30˚C in kanamycin/chloramphenicol substituted autoinduction media (Studier F.W., Protein Expression and Purification, 2005). Pelleted cells were resuspended in lysis buffer (buffer A (50mM Tris-HCl pH 8.0, 500mM NaCl, 10mM imidazole) with cOmplete™,Mini EDTA-free Protease Inhibitor cocktail (Roche) and lysed by sonication. GE Healthcare HiTrap Nickel NTA column was equilibrated in 50mM Tris-HCl pH 8.0, 500mM NaCl, 10mM imidazole) and washed before a gradient of Buffer B (Buffer A + 500mM Imidazole) was applied to elute the protein. The peak eluted between 7- 30% buffer B. The fractions containing Cas9 were pooled and concentrated before size-exclusion chromatography in 20mM Tris, 200mM KCl, 10mM $MgCl_2$ (Superdex 200 16/60; Amersham Pharmacia Biotech). The purified Cas9 was concentrated to 50 mg/ml using ultrafiltration in Amicon centrifugation filter units (Millipore). Aliquots were flash-frozen and stored at −80°C.

2.1.10. Amplification and cloning of genomic DNA for CRISPR sequencing

To assess the efficiency of Cas9 targeting, single embryos were collected at NF Stage 25 and transferred to 0.5ml PCR tubes containing 200ul of TAD lysis buffer (50mM Tris pH7.0, 50mM NaCl, 5mM EDTA, 0.5% SDS, 10% Chelex, fresh 250ug/ml Proteinase K) and incubated at 55℃ for 1 hour followed by 95℃ for 15 minutes to deactivate the Proteinase K. Samples were vortexed to disrupt the cell membranes. Samples were centrifuged for 10 minutes at 13,000rpm, and the top 100µl of clear supernatant (genomic DNA) was collected. The DNA was diluted 1:10 and amplified by PCR reaction using primers flanking the Cas9 target site (Primers included in table in section 2.1).

PCR products were cloned into the pGEM®-T Easy vector system as per manufacturer's guidelines and transformed into E.coli. Minipreps of individual clones were sequenced by GATC Biotech using an SP6 primer. Sequences from 3-10 clones from each embryo were aligned alongside the wild type amplicon sequence using DNAStar SeqMan to identify INDELs.

2.1.11. Linearisation of plasmid DNA

Plasmid linearization was carried out by restriction digest to produce templates for sense transcription of mRNA by SP6 and antisense labelled RNA for in situ hybridisation. For both cases, a 100µl reaction was set up:

1-5µg plasmid DNA

10µl necessary restriction enzyme buffer

2µl restriction enzyme (see table)

_µl (to 100 total) H_2O

Restriction reactions were incubated at 37˚C for a minimum of 90 minutes and linearization was checked by gel electrophoresis alongside 0.5µl of uncut product. Resulting products were then purified as per 2.1.5.

Clone	Enzyme	Polymerase	Source	Use
pSP64T- MyoD	XbaI	SP6	Cloned in lab	mRNA
eFGF	BamHI	SP6	Cloned in lab	mRNA
CS2+-Polr3G	NotI	SP6	Cloned in lab	mRNA
CS2+-Polr3gL	NotI	SP6	Cloned in lab	mRNA
pSP73-XMyoD	BamHI	SP6	Harvey 1991	In situ probe
α-cardiac actin	EcoRI	SP6	Gurdon 1985.	In situ probe
Polr3G	NcoI	SP6	Cloned in lab	In situ probe
Polr3gL	SalI	T7	Cloned in lab	In situ probe

Table 2.3. Clone information, restriction enzyme and polymerases used for in situ probes and synthetic mRNA transcripts.

2.1.12. In vitro transcription of synthetic mRNA

Full length clones were linearised as per 2.1.11. Both MEGAscript® SP6 and mMESSAGEmMACHINE® SP6 kits were used as per manufacturer's instructions:

20ul MEGAscript® SP6 reaction:

2µl 10X MEGAscript transcription buffer

2µl 40mM ATP

2µl 40mM CTP

2µl 40mM UTP

2µl 5mM GTP

2.5µl 40mM Methyl-GTP (Ambion)

2µl SP6 enzyme mix

1-2µl template DNA

_µl H_2O (to 20µl total)

20µl mMESSAGEmMACHINE® SP6 reaction:

3µl template DNA

10µl 2X NTP/Cap

2µl 10X Reaction buffer

2µl SP6 enzyme mix

_µl H_2O (to 20µl total)

Transcription reactions took place at 37˚C for the specified time as per manufacturer's instructions. Transcriptions were checked by gel electrophoresis of

a 1-2µl sample and successful reactions were treated for 15 minutes with TURBO™ DNase at 37˚C followed by purification and isopropanol precipitation. RNA was resuspended in 20µl and stored at -80˚C in 2µl aliquots until use.

2.1.13. Digoxigenin labelled *in vitro* transcription of antisense probes

Labelled probes were synthesised for use in In situ hybridisation. The 20µl reaction was assembled as follows:

4µl 4X Transcription buffer (Promega)

2µl 10X DIG NTP mix (Roche)

2µl 100mM Dithiothreitol (DTT)

1µl Polymerase (SP6, T7, T3) (Thermo-Scientific)

3µl DNA template

7µl H_2O

Transcriptions were incubated overnight at 37˚C.

After checking for product, samples were purified by ethanol/ammonium acetate precipitation at -20˚C for 1 hour minimum (add 50µl H_2O, 50µl 5M Ammonium Acetate and 300µl 100% ethanol). Samples were centrifuged at 13,000rpm for 15 minutes and pellets were washed in 70% ethanol. Samples were dried under vacuum and resuspended in 50ul H_2O. Probes were checked on a 1.5% Agarose/TAE gel run at 180V. Probes were stored at -80˚C until use.

2.1.14. Polr3G antisense morpholino oligonucleotide (AMO) design

Translation blocking Morpholinos 1 and 2 were designed by Gene Tools® LLC. The sequences used for design were

Polr3G MO1:

CTAGTTTTTCATTTTGTTTCTCTTCCCATGCATTAG[TA(ATG)GCTGCCAAAGG GAGGGGAAG]GG

Polr3G MO2:

CCACGCGTCCGGCAGATTTGGTGAATCAGCGAGTTCTATGTTTGATTCATCA GCGAGAAATTATATCAGA[CTGAAGTA(ATG)GCTGCCAAAGGGAG]GGGAAGG

Sequences are of the sense strand. (ATG) indicates the transcriptional start site of the mRNA coding sequence and the sequences targeted by the MOs are in square brackets. Lower case indicates the coding sequence.

Morpholinos were ordered with the sequence reading 5' to 3':

AMO 1: CTTCCCCTCCCTTTGGCAGCCATTA

AMO 2: CTCCCTTTGGCAGCCATTACTTCAG

A splice-blocking Morpholino was also designed to disrupt splicing of the intron between coding exons 2 and 3:

AMO 3: AATGCGTACACAGTACCTACTTTGT

In addition, Gene Tools gives a 5 base-pair mismatch control morpholino (CMO) sequence for each design to use as a specificity control. The CMO for MO2 was used with the sequence:

CMO: GACAACAAAGTTCATCTTCGCTTGA

Morpholinos were diluted to stock concentrations of 3mM in nuclease free H_2O and stored at 4˚c. They were diluted to working concentrations immediately before injection.

2.1.15. qRT-PCR

Total RNA was extracted and purified as outlined in 2.1.8. cDNA was synthesised using 1µg of RNA and Superscript IV (Thermo-Fisher) as per manufacturer's instructions with random hexamers.

Primers were designed using Integrated DNA Technologies PrimerQuest tool. Products were designed to be between 75 and 110 base pairs in length and primers were designed across exon junctions to prevent amplification of any residual genomic DNA.

A list of primers is given:

Name	Forward	Reverse
Xt MyoD	CCGATGGCATGACAGACTAT	ATTTGGGCTGTCGCTGTA
Xt Rbm24	GGCTATGGCTTTGTCACGAT	ACATTTGCTTTCCTGCCATCTA
Xt FoxC1	CAGGGCTTCAGTGTGGAT	TGTCCTGGAAGAGGAGATGA
Xt Rbm20	GAACTCAATGACTTTCATGGC	AACTCCCAGTCCTTCAGATTG
Xt Gli2	GCAGAAGTGGCCCATGA	GCCATTTGGTGGCAGTATTC
Xt Zeb2	GAGGAAGATGAACTGAGGGAAAG	CCTCCTTCATGTCATCAGAACC
Xt Polr3G	AGCTGACAAACAAAGCATTGAA	CAGCTTTCATCTCCCTTGGAA
Xt Polr3gL	TGTTCCCGAACCTGGAATAC	CTTGATGAAGTACGGGAGAGTC
Xt Dicer	GGCTTTTACACATGCCTCTTACC	GTCCAAAATTGCATCTCCAAG
Xt α-cardiac	GTCACCAGGGTGTCATGG	TATCCTGACCCTGAAGTACCC

actin		
Xl Polr3G	GCTGCCTTCACATTTGACATC	TCTATAGAAGGAAAGATTGGCAAGG
Xl Polr3gL	GTTGTTCCCGAGCCTAGAATAC	GCTCCTCGAAGTTCCTGTTT
Xl α-cardiac actin	CGTACCACAGGTATCGTTCTTG	GGGCAGAGCATAACCTTCATAG
Xl ODC	AAAGCTTGTTCTACGCATAGCAACT	AGGGTGGCACCAAATTTCAC
Xl MyoD	AAGGCCGCCACTATGAGGGAGAG	GCTGGGGCTTTGGGTGGAG

Table 2.4. Primer sequences for qPCR analysis of *X.tropicalis* and *X.laevis* samples

Analysis was carried out on a minimum of three biological replicates. Average Ct normalisation to the housekeeping gene dicer, pairwise t-tests were carried out comparing the mean relative expression for control and experimental sets for each gene. When more than two conditions were compared, ANOVA analysis was carried out. Graphs were constructed in GraphPad Prism5. Error bars represent SEM, * = $p<0.05$, ** = $p<0.01$.

2.1.16. ChIP-PCR *(Adapted from Maguire et al., 2012)*

50 tailbud stage embryos were demembraned and fixed for 1 hour in 5ml of 1% formaldehyde/NAM media (NAM/10 premixed with 135 µL 37% formaldehyde) with gentle agitation. Embryos were quenched in 125mM glycine/NAM solution for 10 minutes at room temperature, snap frozen on dry ice and stored at -80°C until use.

Embryos were homogenised in 600µl sonication buffer (20 mM Tris–HCl pH 8.0, 70 mM KCl, 1 mM EDTA, 10% glycerol, 5 mM DTT, 0.125% Nonidet P40 and PIC) and were centrifuged for 10 minutes at 13,000rpm to pellet DNA. Pellets were washed gently in sonication buffer before resuspension in 2ml of sonication buffer.

Chromatin was sonicated in 30 bursts of 20 seconds (50% on/off, 35% max power) with incubation on ice for 30 seconds in between rounds. After sonication, samples were split into two 1.5ml tubes and centrifuged at 13,000rpm for 5 minutes. Supernatant was collected in 100µl aliquots, snap frozen on dry ice and stored at -80°C until use.

80µl of chromatin was incubated in 100µl ChIP incubation buffer (50 mM Tris–HCl pH 8.0, 100 mM NaCl, 2 mM EDTA, 1 mM DTT, 1% Nonidet P40, and PIC) and either 200µl D7F2 MyoD or 10µl goat anti-mouse IgG for 2 hours at 4°C. 20µl of chromatin was kept as input sample and stored at -20°C. Protein G magnetic beads were prepared during this time. 50µl bead suspension was added to 50µl of ChIP incubation buffer + 0.1% BSA and incubated at 4°C with rotation for 1 hour. Beads were collected and washed twice in 1ml ChIP incubation buffer for 5 minutes at 4°C. Supernatant was removed from beads and beads were added to ChIP samples. Samples were incubated at 4°C overnight with rotation.

Beads were collected and washed in 1ml wash buffer 1 (50 mM Tris–HCl (pH 8.0), 100 mM NaCl, 2 mM EDTA, 1 mM DTT, 1% Nonidet P40, 0.1% deoxycholate, and PIC) for 5 minutes at 4°C with rotation. Beads were collected and wash step was repeated with 1ml wash buffer 2 (50 mM Tris–HCl pH 8.0, 500 mM NaCl, 2 mM EDTA, 1 mM DTT, 1% Nonidet P40, 0.1% deoxycholate, and PIC), 1ml wash buffer 3 (50 mM Tris–HCl pH 8.0, 100 mM NaCl, 250 mM LiCl, 2 mM EDTA, 1 mM DTT, 1% Nonidet P40, 0.1% deoxycholate, and PIC), 1ml wash buffer 1, and 1ml TE buffer (10 mM Tris–HCl pH 8.0, 1 mM EDTA). Beads were collected, resuspended in 200µl Elution buffer (0.1 M NaHCO3 pH 8.8 and 1% SDS) and incubated on rotation at room temperature for 15 minutes. Supernatant was collected and Elution buffer step was repeated. 200µl supernatant was added to

first collected supernatant. Input samples were thawed on ice and 380µl elution buffer was added. 400µl of elution buffer was also added to collected beads. 16µl of 5M NaCl was added to all samples and reverse crosslinking was carried out at 65°C for 5 hours with agitation.

Samples with beads were collected and supernatant moved to new 1.5ml tube. 400µl phenol-chloroform was added to samples and samples were centrifuged at 13,000rpm for 5 minutes. 400µl Chloroform-Isoamyl alcohol (24:1) was added to aqueous phase and samples were centrifuged at 13,000rpm for 5 minutes. Aqueous phase was collected to new 1.5ml tubes and precipitated at -20°C overnight in 40µl 3M NaOAc, and 1200µl 100% ethanol with 2µl glycoblue added for pellet visualisation.

Samples were centrifuged at 13,000rpm for 25 minutes. Supernatant was removed and pellets washed gently in 500µl 70% ethanol for 10 minutes. Pellets were vacuum dried for 5 minutes and pellets were resuspended in 30µl nuclease-free H_2O.

ChIP-PCR analysis was carried out as below. For visualisation of ChIP samples, nested PCR was required. All primers are included in table 2.5.

1µl final sample was added to 20µl PCR reaction mix:

10µl Promega PCR mastermix (2X)

1µl 10µM forward primer (1)

1µl 10µM reverse primer (2)

1µl ChIP sample

7µl H_2O

A 30 cycle PCR was carried out with the conditions:

95°C 30 seconds

40-50°C 30 seconds

72°C 20 seconds

5µl of PCR product was added to the following PCR mix:

10µl Promega PCR mastermix (2X)

1µl 10µM forward primer (3)

1µl 10µM reverse primer (4)

3µl H_2O

An additional 30 cycle PCR reaction was carried out with the conditions:

95°C 30 seconds

40-50°C 30 seconds

72°C 15 seconds

6µl product was run out on a 2% Agarose/TBE gel.

Polr3G_ Region 1	(1)CCCATTGTGGCTTTAAAAGGA (3)GGAAAAAGTTAACATTCAGATG	(2)CAACATTAAGCAGATCCCAGTG (4)ACTATCAGCAGTTAGCTC
Polr3G_ Region 2	(1)TGGCAGACATCCTTTTGAAT (3)ATGTTGCCTTTGCCCCTTC	(2)CCCTTTAAGTGGCCACAATG (4)CCTGCTGGGGTCTGCACCT
Polr3G_ Region 3	(1)GAGCGACCGCTAAGGTTTAAT(3)CAGTATGGAACACCAATATGCA C	(2)GTGCTGCCCTGTCCTGTTA (4)CTACCCAGTTCCCTGAACCA
Polr3G_ Region 4	(1)TGGTGTGCTTCTGGATTTTG (3)AGTGCAGATGTGCAAACAGG	(2)TGTGTTATAAGGTTTGTGCAGGA (4)TTGAGCAATCAAGTGGCTTC
Polr3G_ Region 5	(1)CTCCAAAGAGTTGGGTTTGA (3)TATTCCTGGCCCCTAACTGA	(2)GCCATGGGAATTATCGTTCA (4)CCCCTTTGAAGCAACATT
Polr3G_ Region 6	(1)TTCCTTGGAACAGCTTTGCT (3)TCTAGTACAACTAGGGTCCT	(2)GATAACAGGCCCCATACCTG (4)TTCCAAGGATATGTTAAGAC
Polr3G_ Region 7	(1)CGACTTCCAGGAGATTCTAG (3)TGGTGTGCTTAAATAGGTG	(2)GGTGTGCCCTAAATGCAGC (4)GCTCTTCAATTACATAGCTT
Polr3G_ Region 8	(1)GAACTGCACAGGAGACGTTG (3)GTAACAGCACTCAACATCAC	(2)AGGTCCTGAGCATTCTGGAT (4)GGATAACAGATCCTATACCTG
Rbm24 _Regio n 1	(1) GGAGCGCCACGACTTGT AT (3)CAGCCTCATGTTTTTATATG	(2)GTGTCCTCGAAGTTATTGAGGA AT (4)CTGTGACAATATCTGGACTG
Rbm24 _Regio n 2	(1)CCGTCGCTCAACACTCACT (3)GAGACAAAGGGGTGGGGAA	(2)GGGTAACCAGAGAGCCACAC (4)GCTCCTCCCACCAAAGGGG

Rbm24	(1)AACAAGACCAGTTCCCCAAG	(2)GGGATTAAGTGGGGTGGTTC
_Regio	(3)TAGAAGGCAGGCTTTTGAGC	(4)AGCTCTGGGGGTTAGAGAGC
n 3		

Table 2.5. Primer sequences for ChIP-PCR analysis of *X.laevis* embryos.

2.1.17. Northern Blot Analysis

Total RNA was extracted from a *Xenopus tropicalis* embryo stage series using Sigma Tri Reagent. 5µg of RNA was added to a mix of H_2O to a final volume of 10µl and equal volume of NEB 2X loading buffer was added. Samples were incubated at 90˚C for 5minutes before loading onto 10% Acrylamide/7.5M Urea/TBE gels.

10% Acrylamide/ 7.5M Urea/ TBE gel mix:

to 500ml H_2O (filter)

225g Urea

125ml 40% Acrylamide (19:1)

50ml 10XTBE

(Make 20ml with 120µl 10%APS (fresh), 40µl TEMED) Gels were pre-run at 170V for 15minutes before sample loading and running at 180V for 65-70minutes until Xylene blue is approximately 1cm from the bottom of the gel. Gels were stained in Ethidium Bromide solution (10ul 10mg/ml EtBr in 200ml TBE) for 15minutes. Gels were photographed with a ruler zeroed at the top to measure the distance the bands of the molecular weight ladder migrated (in *Xenopus* rRNA bands run at: 28S/4kb, 18S/1.8kb, 5S/121bases).

RNA was transferred by electroblotting onto a HydrobondN membrane in 1X TBE for 75minutes at 100V and crosslinked using Stratalinker UV crosslinker for 2 bursts of autocrosslinking, followed by baking at 80˚C.

Church and Gilbert phosphate buffer for hybridisations (2X PB):

to 500ml H_2O

67g Na_2HPO_4

2ml H_3PO_4

Membranes were pre-hybridised at 40˚C for 30minutes in a 50ml Falcon tube. (30ml Hybridisation buffer: 15ml 2X PB (see above), 2.1g SDS, 0.3g BSA, to 30ml with H_2O, 300µl 10mg/ml ssDNA (boil+add to warm buffer)).

Radioactive probes were synthesised by reaction at 37˚C for 45min, the reaction was stopped by adding 1µl 0.5M EDTA and heating to 60˚C.

50µl probe reaction:

1µl 24-mer oligo 100ng/µl (complimentary to 3'end or intron seq)

5µl kinase buffer

5µl P^{32} gamma ATP (3000mCi/ml)

2µl T4 polynucleotide kinase (Promega)

37µl H_2O

Name	Sequence 5'-3'
$tRNA^{Tyr}$ intron	ACCTAAGGATTGCTGTATCACACC
$tRNA^{Tyr}$ 3'	TCCTTCGAGCCGGAATTGAACCAG
$tRNA^{Leu}$ 3'	TGTCAGAAGTGGGATTCGAACCCA

Table 2.6. 24-mer probe sequences for Northern Blot analysis of *X.tropicalis*

Excess P^{32} was removed from probes using a G25 spin column. 1µl of probe was blotted onto paper to determine cpm (should be at least 1 million/ml). Probes were heated to 80-90˚C and added to hybridisation buffer. Blots were hybridised overnight at 40˚C. Blots were washed twice in 2X SSC/0.05% SDS for 30minutes at room temperature and exposed to film overnight and at different lengths of time as necessary.

2.2. Embryological methods

2.2.1. Fertilisation and culture of *Xenopus tropicalis* embryos.

Females were primed overnight with a subcutaneous injection of 10 units of Human Chorionic Gonadotropin (HCG; Chorulon). Females were induced using 100units HCG 3-4 hours prior to laying. The eggs were fertilised on L15+10% foetal calf serum layered plates with sperm suspension obtained from male. The testes were dissected and homogenised in 1ml L15+10% foetal calf serum. Embryos were flooded with MRS/9 and cultured. Media is replaced 3-4 hours post-fertilisation with MRS/20. Embryos were kept in petri dishes lined with 3ml 1.5% agarose. Prior to first cleavage (25-30 minutes post-fertilisation), embryos were

de-jellied in 3% L-Cysteine (Sigma) solution in MRS/9, pH7.8-7.9. Embryos were kept between 22˚C and 27˚C and staged according to Nieuwkoop and Faber (NF stage).

2.2.2. Fertilisation and culture of *Xenopus laevis* embryos.

Females were induced overnight by subcutaneous injection of 250 units of HCG and kept in the dark at 19˚c. Eggs were fertilised with a suspension of testes crushed in water. Embryos were cultured in NAM/10 (Normal Amphibian Medium) at 14-15˚C in Petri dishes lined with 1.5% agarose. Prior to first cleavage, embryos were de-jellied in 2.5% L-cysteine hydrochloride monohydrate (Sigma) pH7.8. Embryos were staged according to NF.

2.2.3. Microinjections

X.tropicalis embryos were injected at the one to two cell stage (in both blastomeres at two cell) into either the marginal zone or animal poles with maximum of 2nl per cell of mRNA at the appropriate concentration. Embryos were injected in MRS/9+3% ficoll and transferred into MRS/20 2-3 hours after injection.

For gRNA/Cas9 injections, *X.tropicalis* embryos were injected as soon as possible at the one cell stage with 1nl of 300pg gRNA/ 1ng Cas9/300mM KCl mixture.

2.2.4. Animal caps

Animal caps were dissected using mounted tungsten needles to remove the animal pole cells at late blastula stage (NF Stage 8-9) of *X.laevis* or *X.tropicalis* embryos. The dissections were carried out in NAM/2 and animal caps were then cultured in NAM/2 until the required stage.

2.2.5. NF Stage 25 somite explants

X.tropicalis Embryos were left to develop in MRS/20 until NF Stage 25. At stage 25, embryos were transferred into NAM/2 for dissections. Embryos were removed from their vitelline membranes and allowed to rest briefly. Somite explants were dissected using mounted tungsten needles by removing the head and ventral sections. Sections were snap-frozen using dry ice and stored at -80˚C until use. Confirmation of somite presence was performed by carrying out RT-PCR for MyoD expression in all three sections.

2.2.6. Whole-mount *in situ* hybridisation

Antisense Digoxigenin (DIG) labelled probes were synthesised as described in 2.1.14. Vitelline membranes of *X.tropicalis* embryos were removed manually using forceps and embryos were pierced in the ventral cavity or animal hemisphere to prevent collection of probe. Embryos were fixed in MEMFA (0.1M MOPS pH7.4, 2mM EGTA, 1mM $MgSO_4$, 3.7% formaldehyde) for one hour and stored in 100% methanol at -20˚C until use.

Embryos were rehydrated through a series of washes of a gradient of methanol/PBST (75% methanol/ 25% PBST, 50% methanol/ 50% PBST, 100%

PBST). Embryos were then permeabilised with PBST + 10mg/ml Proteinase K (Roche). Embryos were washed twice in 5ml Triethanolamine pH7.8 for 5 minutes per wash, 12.5µl acetic anhydride (Sigma) was added to the second wash for 5 minutes and then a second 12.5µl was added for another 5 minutes. Embryos were then washed in PBST for 5 minutes and refixed in 10% formaldehyde for 20 minutes before a series of 5 minute PBST washes. Embryos were then transferred into 50% PBSAT/ 50% hybridisation buffer (50% formamide, 5X SSC, 1mg/ml total yeast RNA, 100µg/ml heparin, 1X Denhart's, 0.1% Tween-20, 0.1% CHAPS, 10mM EDTA) at 60˚C for 10 minutes. The solution was replaced with 100% hybridisation buffer and rocked in the hybridisation over for two hours minimum. Embryos were then transferred into 1ml Hybridisation buffer + 2-3µl DIG probe and rocked overnight at 60˚C.

Embryos were washed twice in hybridisation buffer for 10 minutes at 60˚C. They were then washed three times for 20 minutes in 2X SSC + 0.1% Tween-20, and three times for 30 minutes in 0.2X SSC + 0.1% Tween-20 at 60˚c. The embryos were then washed twice for 15 minutes at room temperature in Maleic Acid Buffer (MAB: 100mM maleic acid, 150mM NaCl, 0.1% Tween-20 ph7.8). They were preincubated in 1ml of MAB + 20% Heat treated lamb serum (60˚C for 30 minutes) + 2% Boehringer Mannheim Blocking Reagent (BMB) for two hours. This was replaced with fresh blocking solution containing a 1:2000 dilution of affinity purified anti-DIG fragments coupled to alkaline phosphatase (Roche) and rolled overnight at 4˚C.

Embryos were washed briefly three times in MAB. Embryos were subsequently washed three times for one hour in 20ml scintillation vials in MAB. MAB solution

was replaced with 1ml of BM Purple™ staining solution (Roche) and left until colour developed. Embryos were fixed in 10% formalin and photographed.

2.2.7. Photography

Embryos were photographed using a SPOT SP401-230 camera (Diagnostic Instruments Inc.) attached to a Leica MZFLIII microscope and a PC running SPOT advanced software. Image files were then formatted using Adobe Photoshop™ CS3® (Adobe Systems Incorporated).

2.3. RNA-seq analysis

RNA integrity was measured by the Bioanalyzer-2000 in the Technology Facility. mRNA libraries were prepared for sequencing using the NEBNext® UltraTM RNA Library Prep Kit for Illumina, using the NEBNext Poly(A) mRNA Magnetic Isolation Module to isolate Poly(A) mRNA from total RNA. HiSeq3000 2 x 150 bp paired end sequencing was performed by the University of Leeds Next Generation Sequencing Facility.

2.3.1. Computational analysis of RNA-seq data

Illumina deep sequencing resulted in ~440 million reads across the 6 samples. Raw RNA-Seq reads were mapped using the *Xenopus tropicalis* genome version 9.0 (Xenbase). FPKM (fragments per kilobase of transcript per million mapped reads) values were calculated to normalise the number of reads per fragment to the length of the fragment in order to avoid bias towards longer fragments. FPKM

values for three biological replicates were analysed by pairwise t-tests comparing expression in control and MyoD CRISPR-targeted samples. 655 genes showed differential expression based on paired t-tests ($p<0.05$). Expression data for target genes was extracted from RNA-seq data available (Tan et al 2013) and uploaded to https://software.broadinstitute.org/morpheus/ creating a heatmap of expression and hierarchical clustering. Euclidean distance was used as the metric of linkage for complete samples.

2.4. Microarray methods

2.4.1. Microarray Design

A custom microarray for tRNA isoacceptor families and also some mRNAs of interest was designed using the Agilent eArray software to create 8X15K arrays that measure tRNAs and mRNAs simultaneously. Human and *Cricetulus griseus* cell line (CHO-K1) tRNAs were included as controls for species specificity and for use in a collaborative study.

2.4.2. *Xenopus* tRNAs probe sequences.

2638 predicted tRNA genes within the *Xenopus tropicalis* genome version 4.0 were identified from the Lowe Lab GtRNAdb (Chan, P.P. & Lowe, 2009) http://gtrnadb.ucsc.edu/. These were downloaded as a single FASTA file.

Using Agilent eArray design tools, 60mer sense probes with eArray's Base Composition and Best Probe Methodology were designed. 5 probes were designed per sequence for 48 of the 54 tRNA anticodon isoacceptor families as

well as probes for suppressor tRNA and Selenocysteine, resulting in a total of 285 probes. The probe sequences were filtered against Agilent's *Xenopus tropicalis* transcriptome which acts as a similarity index to remove probes with significant cross-hybridisation potential. 6 isoacceptor families did not pass filtering by Agilent, however; in order to include them in the design, 30 additional probe sequences were designed manually and uploaded to eArray.

2.4.3. CHO-K1 tRNA probe sequences.

As eArray did not contain transcriptome data for the CHO-K1 cell line, 3 60mer probes per tRNA anticodon isoacceptor family were designed manually for the 5' end, 3' end and intermediate sequences. For some isoacceptor families with greater variation in sequences, or with introns, 3 additional probes were designed.

The CHO transcriptome was uploaded manually as a similarity index and the probe quality control check identified a high number (254) of probes with cross-hybridisation potential; however, no probes were indicated as poor quality. 3 probes for 13 human tRNA isoacceptor families were also designed as controls.

2.4.4. *Xenopus* mRNA and other Polymerase III target probe sequences.

In addition to tRNA sequences, 5 probes each for 244 selected mRNAs and small ncRNAs were designed using the same criteria as *Xenopus* tRNAs. Coding sequences were downloaded to form a single FASTA file for each probeset for upload to eArray. mRNAs included myogenic genes, bHLH transcription factors, known FGF targets, and mesodermal genes.

A total of 2058 probes were selected for Agilent's standard 8X15K array type design. Agilent's eArray probe check highlighted 102 *Xenopus* probes with cross-hybridisation potential, but no probes were classified as poor in the quality control check. To fill the array, 7 replicates for each probe group were selected.

2.4.5. Microarray sample preparation, RNA labelling and array scanning.

Total RNA was extracted from whole embryos or explants at the desired NF stage using Tri Reagent followed by DNase I in tube and on column treatment using ZymoSpin RNA Clean and Concentrator columns. RNA was labelled using the Agilent Low Input Quick Amp WT labelling kit as per the manufacturer's instructions. 100ng of DNase treated RNA was combined with spike-in mix, WT primer mix and First strand synthesis was carried out to make cDNA by combing the mix with 5x First Strand Buffer, 0.1M DTT, dNTP mix and was incubated for 2 hours at 40°C before 70°C for 15 minutes. Cyanine 3-CTP Labelled cRNA was synthesised from the cDNA and was subsequently purified prior to hybridisation using ZymoSpin columns. Array hybridisation was carried out for 17 hours at 65°C. The microarray was processed using a high resolution laser scanner resulting in a TIFF image. The TIFF file was uploaded to Agilent's Feature Extraction software for quality control assessment and quantification of signal intensity for each individual spot compared to background signals (Processed Signal Values).

The replicate spot intensity for each individual probe across the array were averaged using Agilent's GeneSpring software to give a Raw Processed Signal for each Probe ID (gProcessedSignal). Raw values and values normalised to both 75th centile shift and between individual samples for each probe across the

arrays. (gProcessedSignal (normalized)) were exported into an Excel spreadsheet. Mean values for control and experimental samples, and fold changes were calculated from log_2 normalised values between control and experimental samples from sibling matched biological replicates.

2.4.6. Microarray result processing and statistical analyses

GraphPad Prism5 software was used to carry out statistical tests (t-tests) on the mean fold change for each isoacceptor family and each amino acid isotype for paired samples, and to construct graphs of mean fold changes. It was decided that a threshold of 1.2 and 0.8 would be used as a guide for up/downregulation of transcription of tRNA families. Each column includes error bars indicating SEM and * indicates the level of significance from the statistical analyses after Bonferroni Correction.

2.5. Correlating tRNA expression with mRNA codon usage (Katherine Newling, Technology Facility, University of York)

'codon usage scores' were determined by calculating the number of times a particular codon appears in the coding sequence of all mRNAs included on the microarray, and multiplying it by the intensity value for that mRNA. A sum of these values was then calculated for each mRNA *(number of codons in particular mRNA X intensity value of mRNA)* for each mRNA that has a microarray intensity value of above 10. tRNA expression values were taken as the intensity values on the array. Correlation was calculated using Spearman's Rho assessment for monotonic

relationship between variables. Correlation coefficients between 0 and 1 were calculated and p-values presented.

3. Analysis of MyoD transcriptional targets using CRISPR/Cas9 targeting

3.1. Introduction

3.1.1. Methods of targeting gene expression in *Xenopus*

Post-transcriptional methods of inhibiting gene activity

Genes regulated by MyoD have been previously investigated using myoblasts derived from mouse knockouts and targets in both myoblasts and myocytes have been identified (Cao et al., 2010; Yao et al., 2013b). These studies were carried out in cell culture and may not identify all genes regulated by MyoD i*n vivo*. This chapter describes an in vivo transcriptional analysis of genes that require MyoD in the *Xenopus* gastrula.

SiRNAs and Morpholino Antisense Oligos are widely used methods to post-transcriptionally inhibit gene function and morpholinos are mostly used in frog and fish studies. They are nucleotide based and pair to target mRNA through complimentary base-pairing, but possess a morpholine ring structure which is incorporated rather than the sugar-phosphate backbone of nucleotides. The morpholine structure is resistant to nucleases and therefore persists in a cell for relatively long time periods after introduction.

MOs are designed to block either translation or splicing from occurring (Wagner et al., 2004).Translation-blocking Morpholino Oligos are usually designed as 25mers and are complementary to a sequence of the mRNA including the translational start site (ATG). MOs bind to the target sequence with high affinity and interferes with translational machinery (Summerton and Weller, 1997).

Splice blocking morpholinos in contrast, are designed to target splice donor/splice acceptor sites of pre-mRNA, resulting in a mis-spliced mRNA and a truncated protein of interest (Draper et al., 2001).

Morpholino Oligos have been particularly useful in the study of *Xenopus* gene functions due to the method of microinjections as a means of introduction. MOs offered a specific knock-down protocol for *Xenopus* which was previously unavailable to use, limiting studies to overexpression methods only (Heasman et al., 2000; Rana et al., 2006). Morpholinos are non-toxic to embryos and long lasting during development meaning they can be applicable to both maternal and zygotic mRNAs. One of the many advantages of *Xenopus* as a model organism is the ease, and the speed at which experiments can be carried out. Direct micro-injection of morpholinos perfectly compliment overexpression methods and can be carried out in the same time windows. For almost two decades therefore, morpholinos have remained the most commonly used targeting method in *Xenopus* and are considered reliable, providing the correct experimental controls are included (Eisen and Smith, 2008; Heasman, 2002).

3.1.2. Identifying novel targets of MyoD *in vivo* using Morpholino Oligos.

A previous study carried out in the Pownall lab (Maguire et al., 2012) utilised microarray technology alongside morpholino targeting of MyoD at NF Stage 11.5 (mid-gastrula) in order to identify genes regulated by MyoD at this timepoint. This early timepoint was selected as MyoD is present in the embryo, but myogenic differentiation has not started and analysing this time window would identify early genes involved in myogenic determination. Morpholino targeting in *Xenopus* is particularly advantageous as similar experiments can also be carried out using mRNA rescue experiments in order to validate identified target genes from screens. Genes found to be significantly downregulated in targeted embryos were subsequently also shown to have developmentally relevant spatial expression

patterns and many were co-expressed with MyoD. This resulted in the identification of novel target genes regulated by MyoD early during myogenesis. Identified targets included genes involved in promotion of muscle differentiation (Rbm24) and somitogenesis (FoxC1, Esr1 and Esr2). These targets have roles in somitogenesis which was supported by observed disruption of somite formation in MyoD morpholino targeted embryos. However, the requirement for MyoD for the expression of the identified targets was argued to be transient as their expression at neurula stages was largely recovered.

This study showed that during gastrulation, MyoD directly regulated genes required for somitogenesis and novel genes with roles in myogenesis and highlights in importance of *in vivo* study of complex transcriptional networks. Despite extensive cell culture study into the regulation of myogenesis by MyoD, novel targets that are only identified *in vivo* may therefore still be unknown. Recent advances in targeting technologies now also calls for the defence of morpholino studies due to the availability of genome targeting agents which are both cheap and simple to use in externally developing embryos like *Xenopus* and Zebrafish.

This chapter presents the results from a PhD enhancement award that was granted for me to reassess the targets of MyoD at this same early stage using CRISPR/Cas9 for gene targeting and an RNA-Seq analysis. Therefore, it sits as a stand-alone chapter and it was published in Mechanisms of Development in 2017 (McQueen and Pownall, 2017).

3.1.3. CRISPR/Cas9 as a gene-targeting tool in *Xenopus*

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) are used by bacteria in archaea as an adaptive immune method against invading viruses. In particular, the bacteria *Streptococcus pyogenes* protect themselves from foreign nucleic acids (i.e. from viruses) through the induction of CRISPR RNAs (crRNAs). These are produced from incorporation of foreign DNA into the genome through exposure to an invading virus and they are located in proximity to CRISPR. This invading “protospacer DNA” is transcribed producing crRNAs specific to the invading viral DNA sequence. crRNAs form a complex with another RNA known as the trans-activating RNA (tracrRNA) and Cas9 endonuclease and through base pairing with invading DNA, effectively recruit Cas9 to targets DNA. When target sequences are adjacent to Proto- spacer Adjacent Motifs (PAMs), Cas9 induces a double-strand break in the invading DNA by cleavage, thus inactivating the viral DNA (Sander and Joung, 2014). Since the understanding that Cas9 gains specificity for recruitment to a target sequence through the interaction with two dual RNAs, and that this in fact can be replaced by a single programmed RNA designed to incorporate elements of both crRNA:tracrRNA (Jinek et al., 2012) known now as the guideRNA, there has been a great potential for CRISPR based techniques in gene targeting and genome-therapies.

Recent developments in exploiting this gene targeting technology have now made directed targeting at the DNA level an easy and efficient option in a variety of model systems. Including the potential to create stable founder population mutants in *Xenopus* (Guo et al., 2014; Nakayama et al., 2013).

Genome editing by CRISPR/Cas9 requires minimally the guide RNA (gRNA) sequence and Cas9 mRNA/Protein. In *Xenopus*, Cas9 protein may be co-injected

with gRNA to induce more rapid cleavage of the target site and higher efficiency targeting rates (Guo et al., 2014; Hwang et al., 2013; Nakayama et al., 2013). The availability of CRISPR/Cas9 as a targeting method offers many advantages over use of Morpholinos, especially the limited side-effects and higher specificity of gene targeting strategies compared with the effects of Morpholinos (Gentsch et al., 2018). Moreover, as differences in phenotype are now being observed when repeats of Morpholino studies are carried out with CRISPR/Cas9, the validity of results from some Morpholino based studies is being questioned (Kok et al., 2014; Rossi et al., 2015; Schulte-Merker and Stainier, 2014). The injection of gRNA and Cas9 results in a population of mosaic founder individuals (F0s) and it is yet to be assessed whether the level of targeting and resulting mutations are sufficient to cause effects both at the level of the phenotype and at the molecular level.

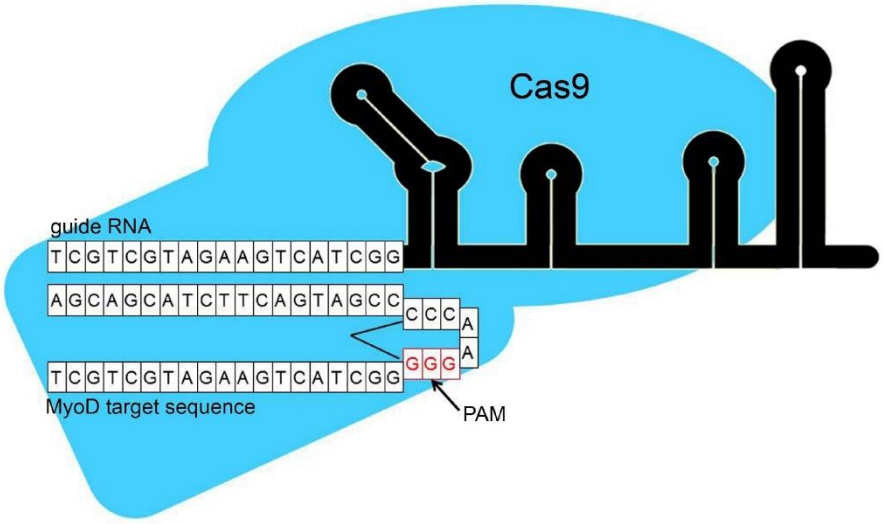


Figure 3.1: Schematic of Cas9 guide RNA base pairing with endogenous MyoD target sequence and recruitment of Cas9. The PAM NGG site is indicated in red.

3.1.4. Aims of this chapter

Previous *in vivo* analysis of MyoD transcriptional activity in *Xenopus* used MOs to knock-down MyoD and Affymetrix gene-chip to identify targets (Maguire et al., 2012). However, recent studies using other gene-targeting methods attempting to replicate results of antisense morpholinos in genetic mutants have shown discrepancies, presumed to be due to off-target effects means that targets identified using morpholino targeting require validation (El-Brolosy et al., 2018; Gentsch et al., 2018; Rossi et al., 2015).

The aim of this chapter is to determine novel targets of MyoD during early myogenesis, which will be achieved through the targeting of MyoD by CRISPR/Cas9- now readily available and economical to use in *Xenopus*. This chapter will also assess whether the use of CRISPR/Cas9 in F0 embryos is effective method of gene targeting and advantageous over Morpholino studies when paired with global transcriptomic analyses (i.e. can new targets be indentified from mutation methods that are not found in MO studies). The analysis is carried out at the earliest NF stage that MyoD protein is found localised to the nucleus and therefore active (Hopwood et al., 1989), as was also used in Maguire et al., 2012. Comparison of previously identified target genes will be used to validate the previous Morpholino study and additional target genes will be identified by RNA-Seq.

- To assess the effectiveness of CRISPR/Cas9 activity in *Xenopus* founder populations using *tyrosinase* as a target.
- To disrupt expression MyoD in *Xenopus* embryos using CRISPR/Cas9.

- To evaluate at the individual sequence level, the extent and types of mutations in F0 embryos
- To use CRISPR/Cas9 targeting with RNA-Seq to identify novel targets of MyoD during the earliest stages of myogenic differentiation.

3.2. Results

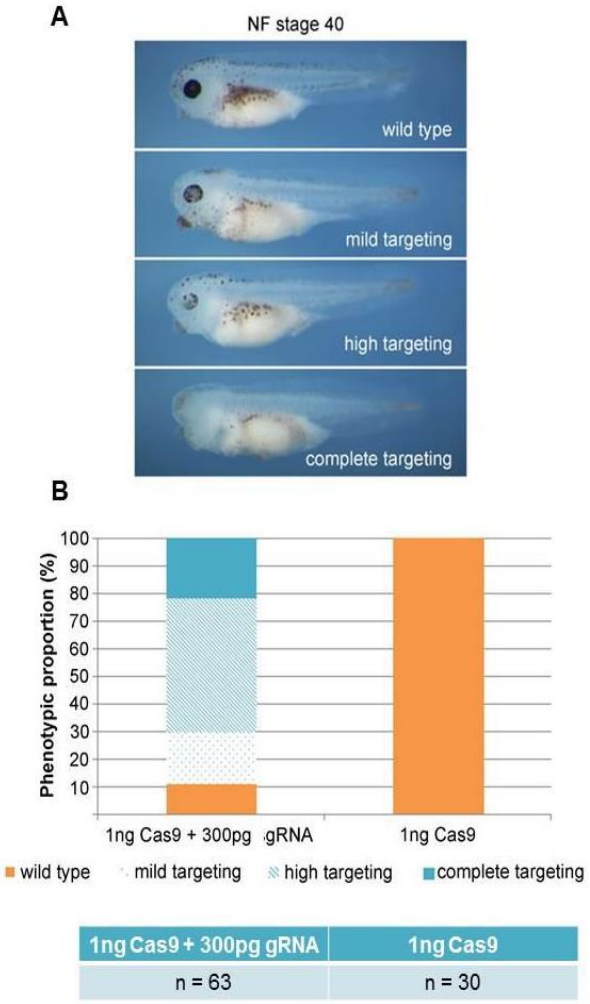
3.2.1. Validation of Cas9 effectiveness through F0 targeting of tyrosinase control.

Previous studies investigating the effectiveness of Cas9 targeting in mosaic F0 embryos, both in *Xenopus* and Zebrafish, have used genes with observable mutant phenotypes (Blitz et al., 2013; Hwang et al., 2013; Irion et al., 2014; Nakayama et al., 2014). The pigmentation gene *tyrosinase* has been commonly utilised in *Xenopus* CRISPR studies due to its distinctive albino phenotype (Guo et al., 2014; Nakayama et al., 2013).The tyrosinase gRNA sequence as designed in (Nakayama et al., 2014) with high efficiency 5' primer was synthesised *in vitro* by PCR reaction followed by a transcription reaction with T7.

To determine the effectiveness of Cas9 protein produced in house, 300pg of tyrosinase gRNA was injected alongside 1ng of Cas9 protein in solution. gRNA injected embryos cultured to NF Stage 40 (n=67) revealed effective targeting, 90% mutant phenotype observed. However, tyrosinase targeting also highlighted the mosaicism present between targeted individuals of the F0 population as a series of phenotypes were observed (Figure 3.2. A) from mild targeting [some pigmentation loss observed in the eye], to complete targeting [full loss of

pigmentation in both eye and skin resulting in an albino individual]. The largest percentage of individuals however showed high targeting by which most pigmentation in the eye is lost (Figure 3.2. B).

As targeting of tyrosinase resulted in clear phenotypic changes, this was then injected to a small number of individuals additionally to each round of Cas9/MyoD gRNA injections to ensure successful targeting.



1ng Cas9 + 300pg gRNA	1ng Cas9
n = 63	n = 30

Figure 3.2. Assessment of Cas9 targeting efficiency using tyrosinase as a visual marker. (A) NF stage 40 embryos were assessed for loss of pigment and graded as per the severity scale. (B) Phenotypic proportions of individuals (n=67) categorised within the four classes were calculated and compared with individuals (n=30) injected with Cas9 only.

3.2.2. Gene targeting *X. tropicalis* MyoD using CRISPR/Cas9

3.2.2.1. Detecting genetic disruption of MyoD in embryos

Xenopus tropicalis is a diploid frog and, as such, genetic methods are simplified using this model. *X. tropicalis* MyoD was targeted using CRISPR/Cas9 in order to identify genes that require MyoD for their expression in the early mesoderm, prior to myogenic differentiation. A synthetic guide RNA (gRNA) was designed against a sequence in exon 1 coding for the amino terminal part of the bHLH domain such that any disruptive mutation would result in truncation before the DNA binding domain. The gRNA was co-injected with Cas9 into 1- to 4-cell embryos.

In order to measure mutagenesis, single embryos were collected for sequencing analysis. Non-homologous end-joining (NHEJ) that repairs DNA after cleavage by Cas9 results in random insertions or deletions (INDELs), therefore genomic DNA was extracted from individual embryos and the targeted region of the *MyoD* gene was amplified by PCR and cloned such that different mutations in a single embryo could be identified. A total of 35 embryos were collected and 3-15 clones were sequenced from each individual. Of these 35 embryos, 31 contained at least one mutated sequence (targeting of 88.6%; Figure 3.3. A), indicating a high efficiency of gene targeting. As expected, each embryo differed in the proportion of mutant sequences, some returning all mutated sequences, and others showing only 50% mutant sequences.

3.2.2.2. Characterising alleles

Each sequence was further characterised to determine whether an insertion, a deletion or a point mutation had occurred. The average targeting efficiency throughout sequenced embryos shows that 78% of returned sequences were mutated (Figure 3.3. B), of which the majority were deletions (77%) rather than insertions or point mutations (Figure 3.3. C). The level of mosaicism within a single embryo was high with several different mutations identified in a single F0 embryo. 10 sequences from a single embryo were aligned to the predicted wild type *MyoD* sequence (Figure 3.3. D). As predicted, all mutation events occur and the near the protospacer-adjacent motif (PAM) where NHEJ results in many different alleles. Figure 3.3.E shows the proportion of the alleles that code for frameshift mutations as a result of either a deletion or an insertion. Only a small proportion of sequences (2.3%) represent in-frame insertions. Frameshift mutations represent the majority of mutated sequences identified, however, this equates to less than half of all sequences returned (43.7%). This highlights a caveat when using F0 embryos for genetic analyses; although CRISPR/Cas9 targeting results in a very high proportion of mutated alleles in an individual embryo, in this case, less than half of these mutations will result in a truncated protein or a genetic null. Mutations causing indels in multiples of 3 were categorised as in-frame deletions/insertions. 25.9% sequences returned were confirmed wild type, 35.1% showed frame shift deletions, 8.6% showed frame shift insertions, 24.1% showed in frame deletions, 2.3% showed in frame insertions and 4% showed missense mutations.

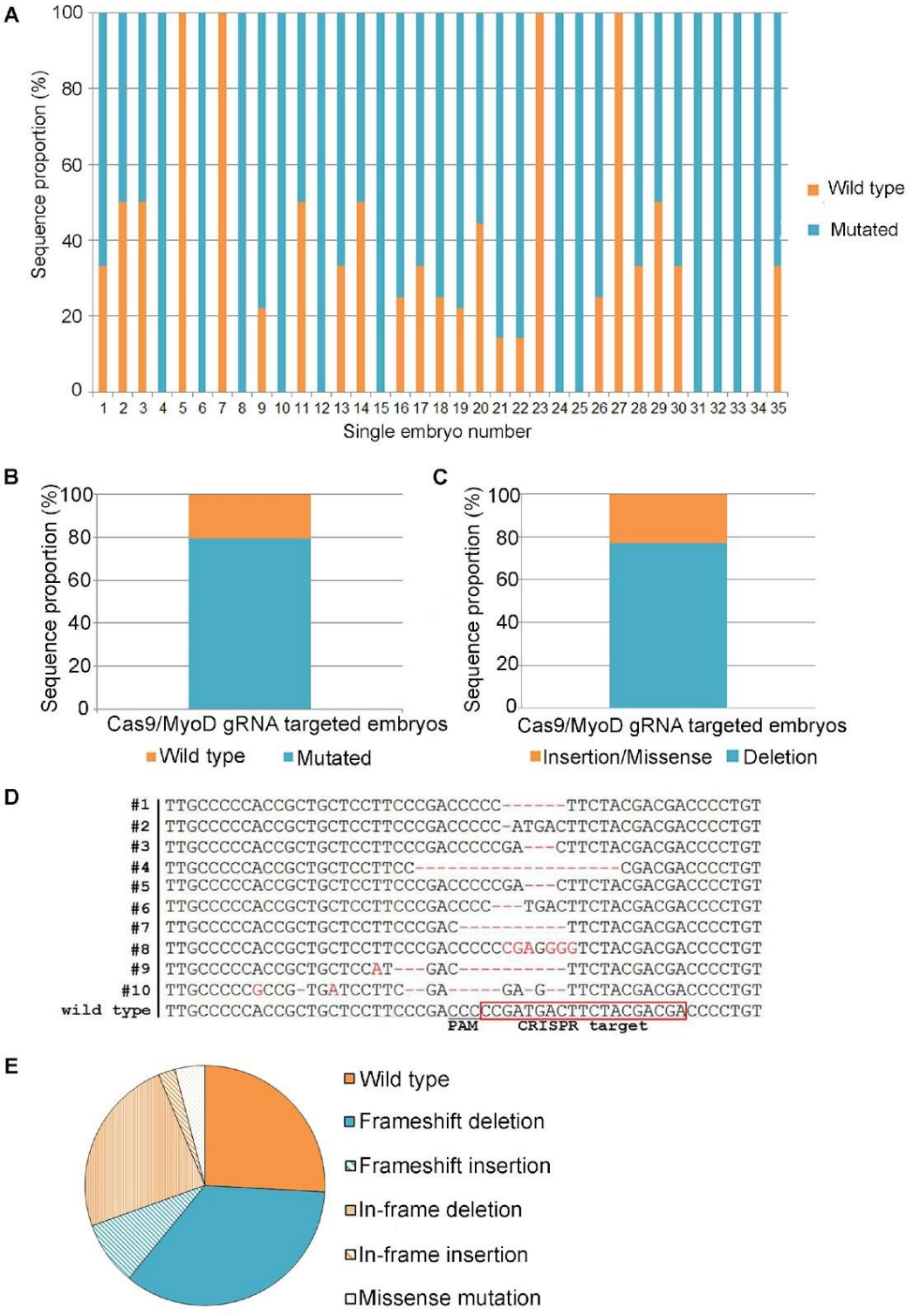


Figure 3.3. Assessment of CRISPR/Cas9 targeting efficiency of MyoD1 through genotyping. At NF stage 25, genomic DNA was extracted and a 432bp region including the predicted CRISPR target site was amplified by PCR and cloned into pGEM T-Easy. 3-15 clones per embryo were sequenced (a total of 153 sequences were analysed). (A) Proportions of mutated vs wild type sequences within 35 individual embryos.

(B) Overall proportion of wild type and mutated sequences in confirmed Cas9 mutated embryos. (C) Characterisation of mutation types in confirmed Cas9 mutated embryos. (D) 10 sequences from a single Cas9 targeted embryo indicates the level of mosaicism in F0 individuals. Cas9 PAM sequence is indicated in red underline. (E) Characterising the sequence categories present in sequenced F0 embryos. (Blue indicates mutation or disruption; orange indicates wild type sequence or silent mutation).

3.2.2.3. Determining the presence of off-target effects.

In other systems using CRISPR/Cas9 to generate mutations, the embryos are raised to maturity and out-crossed at least twice. This significantly reduces the likelihood of off-target mutations being carried forward. This study uses F0, so any potential off-target effects need to be considered. The MyoD target sequence identified in the application ChopChop (http://chopchop.cbu.uib.no) was selected due to its location within the functional domain of MyoD and also the prediction of zero mismatch off-targets elsewhere in the *X.tropicalis* genome.

However, MRFs have a highly conserved functional domain, defining them as regulators of myogenesis, and Myf5 and MyoD have partially redundant roles within very similar time-points of development. At the NF stage analysed, the only other MRF active is Myf5. Therefore, to ensure that any effects of targeting are due to MyoD mutation and not due to off-target effects, the amino terminal of the Myf5 bHLH domain was also sequenced to ensure no mutations occur as a result of Cas9 targeting (Figure 3.4).

Of the 14 sequences returned, no targeting was observed. Cas9 targeting of MyoD has been highly efficient and therefore it was concluded that off-target mutation of other MRFs would not be the cause of any changes in gene expression.

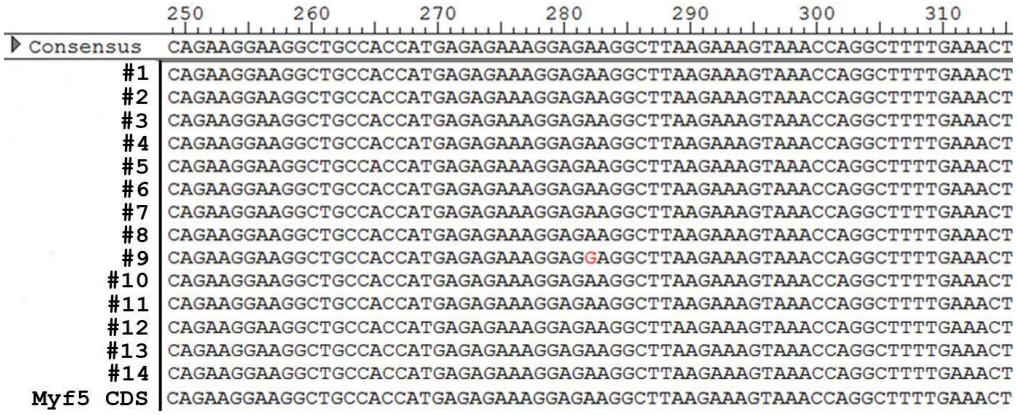


Figure 3.4. Analysis of potential off-target mutagenesis in the bHLH domain of Myf5. Genomic DNA from 5 embryos targeted for MyoD was also amplified using primers designed against the amino terminal of the bHLH domain of Myf5 and sequenced. Sequence reads were aligned to the wild type Myf5 sequence and analysed for mutations.

3.2.3. Analysis of transcripts in MyoD-targeted embryos

3.2.3.1. Disruption of MyoD gene transcription and MyoD activity

To further characterise embryos targeted by CRISPR/Cas9 and to allow comparison of targets with those identified from previous studies (Maguire et al., 2012), mRNA was extracted from groups of ten embryos at the same stage as before (NF Stage 11.5) and qRT-PCR was used to analyse the expression of *MyoD* and its known target gene identified in the previous screen *Rbm24* (*Seb4*) (Li et al., 2010, Maguire et al., 2012). There was a significant decrease in *MyoD* expression ($P < 0.01$), (presumably from non-sense mediated decay) and *Rbm24* ($P<0.05$) when embryos injected with MyoD gRNA + Cas9 protein are compared to those injected with embryos injected with the same amount of Cas9 protein alone. The results were calculated as relative proportions of expression and repeated for three biological replicates and relative expression of *MyoD* in targeted embryos compared with controls was reduced to 0.57 and *Rbm24* is reduced to 0.77 (Figure 3.5. A). In addition, RNA sequences were sequenced to determine

the mutation rate and 6 distinct mutations were identified from one sample set (Figure 3.5. B). The total proportions for each of the mutations shown are as follows: -1= 175 (15%), -2= 14 (1%) -3= 219 (18%), -4= 70 (6%), -6= 70 (6%), -8= 40 (3%). 49% total RNA-Seq reads for MyoD show mutation, however, almost half of the mutated sequences result in no frameshift of coding sequence and are therefore unlikely to result in non-functional protein.

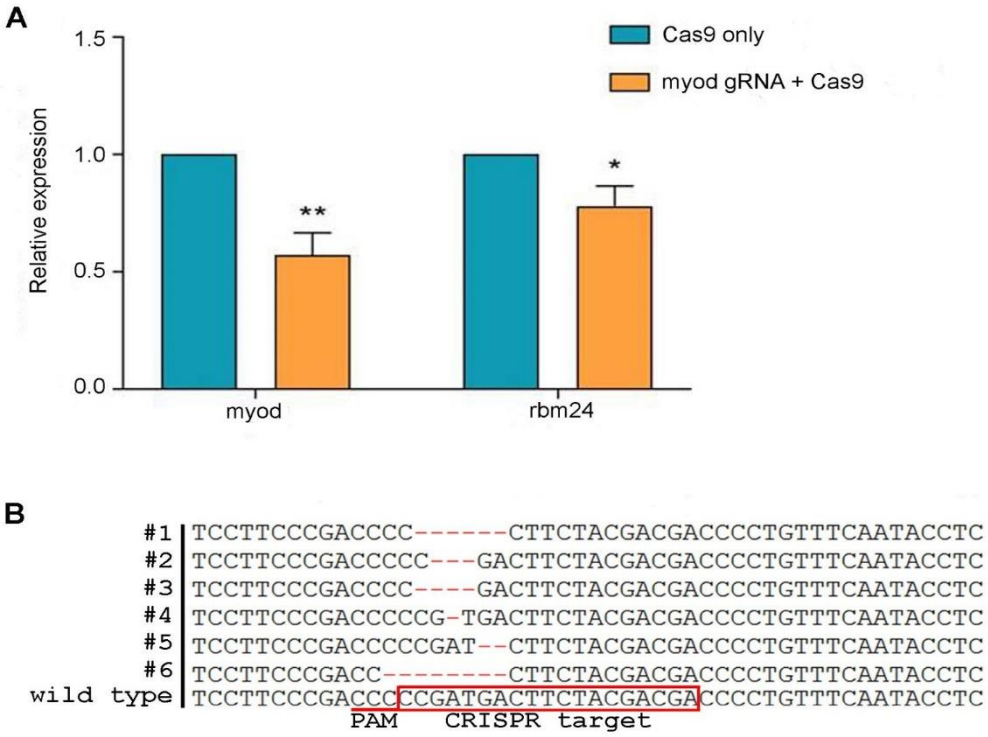


Figure 3.5. Validating samples sent for RNA-Seq. To determine the mRNA levels of MyoD in gRNA injected embryos, qRT-PCR analysis was carried out for MyoD and the known MyoD target Rbm24. Pair-wise t-tests were carried out comparing relative expression for Cas9 only and gRNA injected sets. Error bars represent SEM, * = $p<0.05$, ** = $p<0.01$. (B) MyoD RNA sequences returned from mapping raw RNA-Seq reads to the *Xenopus tropicalis* MyoD1 gene. A total of 1200 sequences were extrapolated across the three biological replicates sent for RNA-Seq and proportions of reads showing each mutation type were calculated from. Sequences were aligned against the sequence for wild type MyoD.

3.2.3.2. Identifying genes that require MyoD using RNA-Seq analysis

Three biological repeats for *X.tropicalis* experimental embryos (targeted for *MyoD*) and sibling controls (Cas9-only injected) were collected at NF Stage 11.5 and mRNA was extracted for RNA-seq. cDNA libraries were prepared and Illumina deep sequencing resulted in 440 million reads across the 6 samples. RNA-Seq reads were mapped using the *Xenopus tropicalis* genome version 9.0 (Xenbase.org) and FPKM values were established for all genes. Transcripts that align to *MyoD* were analysed for INDELs, and Figure 3.6.B shows that a significant proportion of the reads have deletions in the expected target site adjacent to the PAM. Insertions are less likely to be detected, as they would fail to align with reference genome.

To produce an overview of the significance of fold changes observed in CRISPR targeted samples, a volcano plot for (log_2) fold change vs ($–log_{10}$) paired t-test P-value was constructed using Python script (Figure 3.6). Each individual point represents a gene and the dotted line represents a p-value of <0.05. Points in red indicate genes with a fold change of less than 1; that is, where the average FPKM value of experimental samples have not doubled or halved compared to that of the control. Blue points represent genes with a fold change greater than 1 but a P-value of >0.05, so not statistically significant. Yellow points represent genes with a fold change greater than 1 and a P-value of <0.05. The majority of points show a fold change of less than 1 and P-values of >0.05, indicating no significant change in gene expression at NF Stage 11.5 in response to CRISPR/Cas9 targeting of MyoD. However, 1165 genes mapped to the *X.tropicalis* genome display significant change and are further analysed in sections 3.2.4 and 3.2.5.

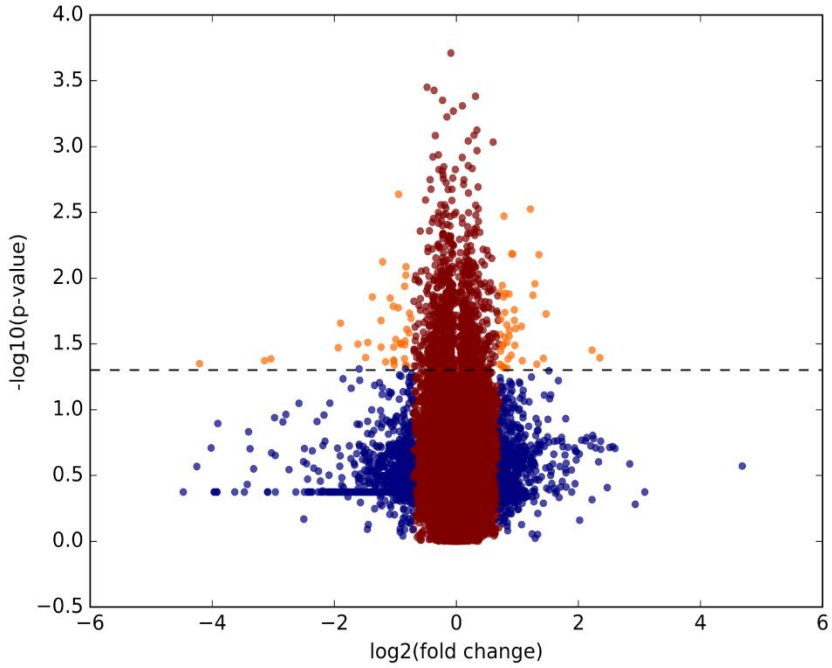


Figure 3.6. Initial analysis of RNA-Seq data fold changes and t-test significance. RNA-Seq reads were mapped using the *Xenopus tropicalis* genome version 9.0 (Xenbase). FPKM (fragments per kilobase of transcript per million mapped reads) values were calculated in order to avoid bias towards longer fragments by normalising the number of reads per fragment to the length of the fragment. FPKM values for three biological replicates were analysed by pairwise t-tests comparing expression in control and MyoD CRISPR-targeted samples. A volcano plot showing t test significance value (-log10 p-value) vs fold change (log2) was constructed in Python. Genes in blue indicate a fold change of greater than 1, genes in yellow indicate a fold change greater than 1 and a p value of <0.05.

3.2.4. Computational analysis of early genetic targets of MyoD

Of the 1165 genes found to be significantly altered in the absence of MyoD, some showed very low expression levels. Therefore, a minimum expression threshold of 5.0 FPKM average for the control samples was applied. In addition, as MyoD expression in targeted samples showed a fold change of 0.71, therefore genes with fold changes in this same range (between 0.68 and 0.91) were selected for

further analysis. During manual curation, two genes that fell just outside the criteria cut-off: *FoxC1* (0.84; P<.058) and *Pbx2* (0.93; P<.02) were also included. This resulted in a short list of 100 potential target genes (See Appendix, Table 1). Notably, previously identified target genes of MyoD, ESR1 and Esr2, Delta and Tbx6 were not identified in this screen, however, targets FoxC1, FoxC2 and Rbm24 were identified as in Maguire et al., 2012.

3.2.4.1. Temporal expression analysis of potential target genes

To further investigate whether the identified genes are expressed at a time consistent with activation by MyoD, temporal expression profiles were analysed. As MyoD protein is first detected in the mesoderm at NF Stage 11, candidate genes with expression prior to these stages are less likely to be bonafide target genes of MyoD. Furthermore, as the analysis was carried out at Stage 11.5, genes coding for contractile proteins or other differentiation specific genes are not expected to be identified by this study, however, the expression profile of Actc1 was included as a reference for this class of genes. RNA-Seq data from a development time course of *Xenopus tropicalis* is available (Tan et al., 2013) and these expression profiles were used for hierarchical cluster analysis of target genes. To do this, expression data was extracted for the 100 short-listed genes (Appendix Table1) and used to create a heat map of expression levels over a developmental time course (Appendix Figure 1).

33 genes aligning with reference profiles for determination (MyoD1) and differentiation (Actc1) were selected for further heat mapping and cluster analyses (shown in Table 3.1 and Figures 3.7. and 3.8.). Euclidean distance was used as the metric of linkage for complete samples and a selected cutting point for the

clustered dendrogram resulted in the formation of 5 clusters of distinct expression profiles; this is shown in the heat map where orange boxes represent highest expression levels (Figure 3.8.). To determine the expression patterns observed within the 5 clusters, profiles of relative expression for each stage were constructed from the expression data used in the heat map and clustering.

Gene symbol	ENSEMBL ID	Average FPKM Control	Average FPKM Experimental	Experimental Relative expression	p-value
bmpr1b	ENSXETG00000019220	8.01	5.45	0.68	0.02
pgp	ENSXETG00000016097	6.12	4.52	0.74	0.03
gbx2.2	ENSXETG00000003293	42.10	31.57	0.75	0.01
sp8	ENSXETG00000030115	12.27	9.28	0.76	0.05
nkx6-2	ENSXETG00000023614	20.42	15.51	0.76	0.02
tsfm	ENSXETG00000009653	5.08	3.96	0.78	0.05
decr2-like	ENSXETG00000010329	12.93	10.08	0.78	0.04
rbm20	ENSXETG00000025245	7.04	5.51	0.78	0.04
zeb2	ENSXETG00000000237	18.67	14.88	0.80	0.01
foxc2	ENSXETG00000016387	80.68	65.42	0.81	0.03
babam1	ENSXETG00000025571	9.14	7.47	0.82	0.02
gli2	ENSXETG00000011189	12.23	10.04	0.82	0.01
sp5	ENSXETG00000025407	64.04	53.11	0.83	0.03
pmm2	ENSXETG00000004549	45.85	38.21	0.83	0.04
pygm	ENSXETG00000034136	123.57	103.71	0.84	0.04
foxc1	ENSXETG00000000594	73.17	61.58	0.84	0.06
fstl1	ENSXETG00000018009	25.29	21.31	0.84	0.02
pex16	ENSXETG00000001027	8.44	7.15	0.85	0.05
slc13a4	ENSXETG00000008163	18.80	15.99	0.85	0.03
ak6	ENSXETG00000018174	14.39	12.24	0.85	0.03
pgk1	ENSXETG00000007447	19.96	17.15	0.86	0.02

mrps30	ENSXETG00000017716	11.89	10.23	0.86	0.04
flvcr2	ENSXETG00000027282	8.63	7.47	0.87	0.04
cdx1	ENSXETG00000010282	77.18	66.87	0.87	0.03
pdlim7	ENSXETG00000007240	13.77	11.97	0.87	0.01
msi1	ENSXETG00000012216	55.14	47.91	0.87	0.00
rnf7	ENSXETG00000014753	99.16	86.80	0.88	0.04
pcdh8.2	ENSXETG00000008792	73.31	64.29	0.88	0.01
herpud2	ENSXETG00000013111	5.58	4.91	0.88	0.04
rnf157	ENSXETG00000019548	5.75	5.16	0.90	0.05
epn1	ENSXETG00000022662	42.43	38.19	0.90	0.04
dnajc24	ENSXETG00000008179	14.14	12.76	0.90	0.03
pbx2	ENSXETG00000005223	169.56	158.17	0.93	0.02

Table 3.1. Shortlisted target genes identified from RNA-Seq analysis. After initial heatmapping and cluster analysis, only genes located within clusters showing developmentally relevant expression profiles were shortlisted as early MyoD targets. The list was manually curated using existing spatial expression profiles and literature to result in 33 shortlisted early target genes of MyoD1.

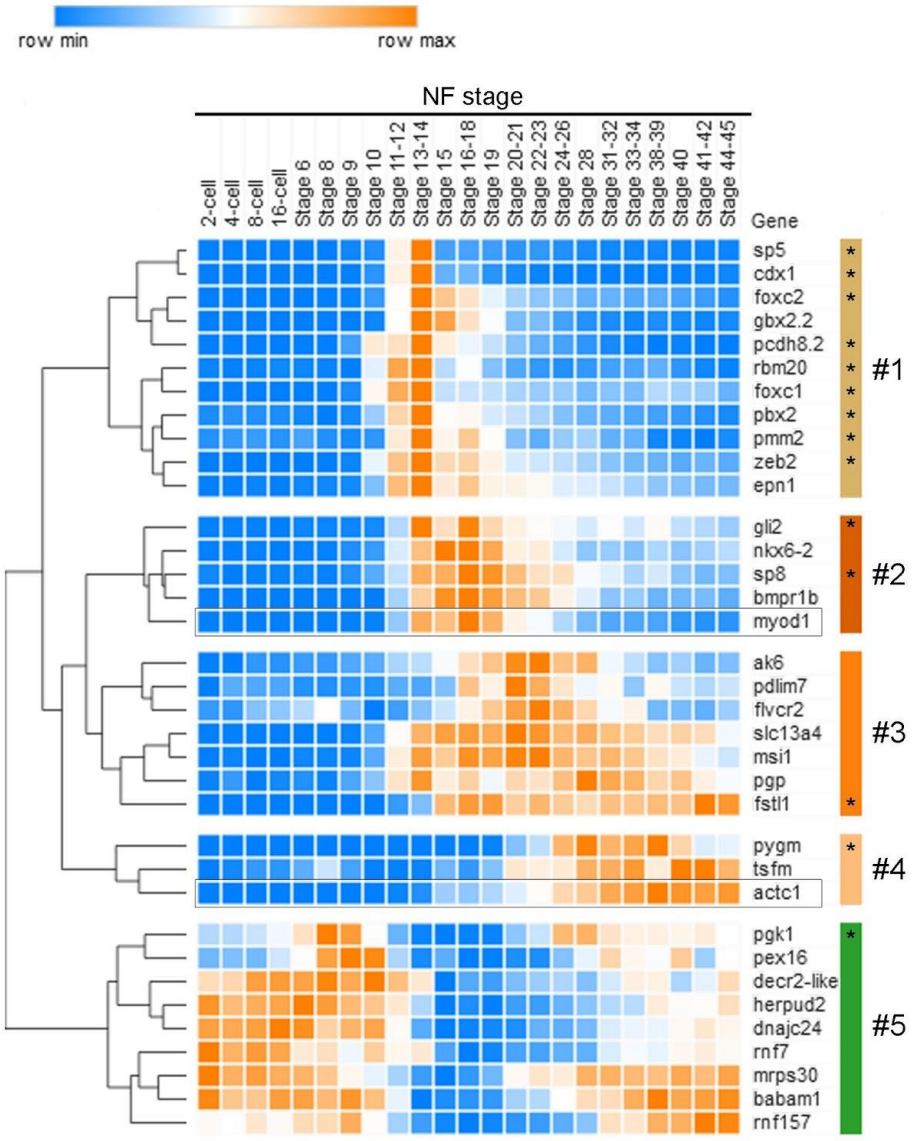


Figure 3.7. Hierarchical clustering of shortlisted early MyoD target genes. Expression data from (Tan et al., 2013) was transformed to relative expression data and uploaded to https://software.broadinstitute.org/morpheus/ for heat map conversion. Relative expression is shown as a scale of low (blue) to high (orange). Euclidean distance was used as the metric for hierarchical clustering of complete samples, which resulted in 5 clusters showing distinct expression profiles. (*) indicates genes which have known or predicted roles in muscle development, or interaction with MyoD. MyoD and Actc1 were included in the analysis in order to highlight relevant clusters of interest.

Each individual cluster was further analysed by time-course profiling (Figure 3.8.). Clusters 1 and 2 show similar overall expression profiles: genes within these clusters have very low or no maternal expression with earliest notable expression at stage 10 (Figure 3.8. A and Figure 3.8. B). Expression in both clusters increases during gastrula and neurula stages, however, in Cluster 1 expression increase is more rapid, as highest expression is observed at Stages 13-14, whilst genes in Cluster 2 show peak expression at Stages 16-18. Both clusters then show decreases in gene expression in later stages. Notably, *MyoD* itself is allocated to Cluster 2. Genes located in Cluster 3 also show low or no expression prior to mid-blastula transition (MBT) and the overall expression trend shows increasing expression until early tailbud Stages 20-22 (Figure 3.8. C.). Expression then decreases in later stages. Individual gene expression within this cluster however, is more varied than in other clusters. Cluster 4 is a much smaller cluster containing the known MyoD target *alpha-cardiac actin* (Actc1). Gene expression for this cluster shows delayed gene activation with increases occurring from Stage 14 onwards, this increasing expression is maintained through tailbud stages and only decreases slightly in the later tadpole stages (Figure 3.8.D). Cluster 4 contains the muscle glycogen phosphorylase pygm. Cluster 5 is distinct from all other clusters in that the genes located within this cluster show maternal expression (Figure 3.8.E.). Expression decreases rapidly after MBT and is at lowest levels during neurula stages, then increases again during tailbud Stages (20-28) through to later tadpole Stages (31-45).

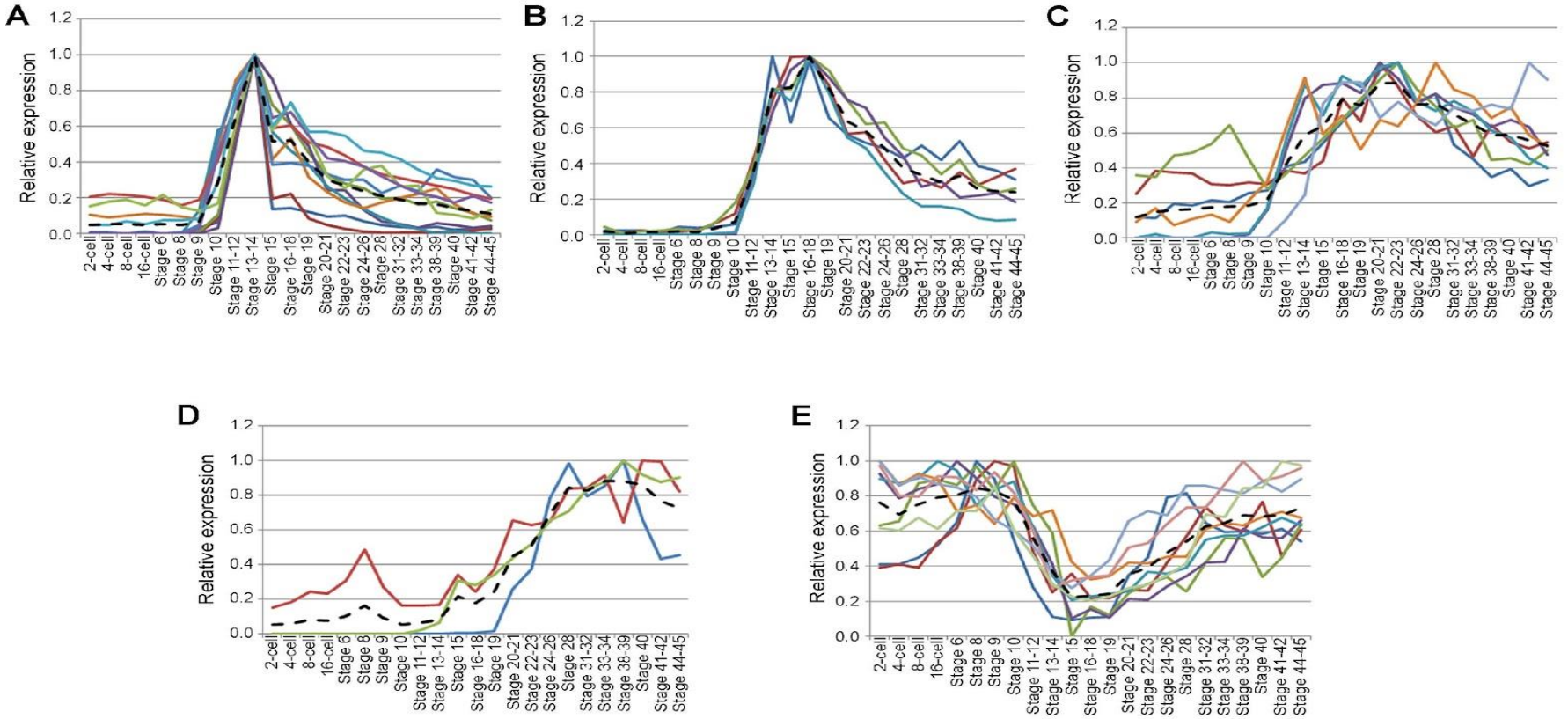


Figure 3.8. MyoD target gene cluster analysis. Expression data from (Tan et al., 2013) was used in a time course analysis of whole embryonic development. Target genes with developmentally relevant expression profiles were identified from the initial heatmap. After hierarchical clustering, profiles of gene relative expression over time for each cluster reveals distinct expression profiles between clusters. Mean relative expression is shown for each cluster along with expression profiles for each gene. (A-E) represent clusters 1-5 respectively.

3.2.5. Validation of identified target genes

Hierarchical cluster analysis was used to highlight genes that show developmentally relevant expression profiles, and strengthen their status as candidates for early genetic targets of MyoD; clusters 1 and 2 include genes identified in other studies as myogenic or pre-myogenic genes. In order to validate whether any of these genes require MyoD for their expression during gastrula stages, gRNA targeting MyoD together with Cas9 protein was injected and the expression of several candidate genes in these embryos at Stage 11.5 was assayed as compared to Cas9 only injected embryos using qRT-PCR (Figure 3.9). qRT-PCR analysis confirms *Rbm20*, *Rbm24*, G*li2*, *FoxC1*, and *Zeb2* (aka *XSip1*), as well as *MyoD* itself, are all significantly down-regulated in targeted embryos. This validation supports the notion that gene targeting and transcriptomic analysis of founder embryos has provided a robust list of candidates genes regulated by MyoD prior to the onset of skeletal muscle differentiation.

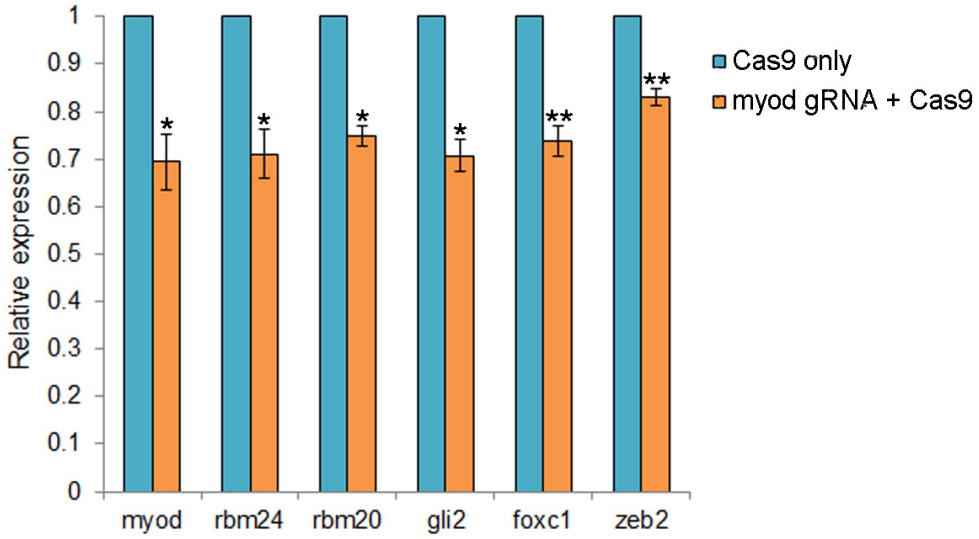


Figure 3.9. qRT-PCR analysis of identified early targets of MyoD at NF Stage 11.5. Analysis shows the expression of MyoD and the known MyoD target gene Rbm24 alongside predicted target genes Rbm20, Gli2, Foxc1 and Zeb2. Pair-wise t-tests were carried out for the mean relative expression of three biological replicates for Cas9 only and Cas9 plus gRNA injected sets for each gene. Error bars represent SEM, * = $p<0.05$, ** = $p<0.01$.

3.3. Discussion

MyoD is known to direct several different sub-programmes of gene expression during myogenesis (Bergstrom et al., 2002; Blais et al., 2005; Soleimani et al., 2012) consistent with its role as an essential determination gene for the proliferative myoblast (Rudnicki et al., 1993). However, MyoD is also a robust initiator of transcriptional targets during myogenic differentiation, distinguishing itself in this way from Myf5 (Conerly et al., 2016). It is an interesting proposition that one transcription factor can activate distinct panels of genes at two different stages of cell lineage specification. Indeed, this notion of promoter swapping is supported by MyoD binding analysis using chromatin immunoprecipitation (ChIP-seq) protocols that have shown that MyoD functions as a transcriptional regulator during both myogenic determination and differentiation by binding and activating distinct sets of genes (Soleimani et al., 2012). This chapter successfully utilised CRISPR/Cas9 techniques to identify both known and novel MyoD target genes at the earliest stages of myogenesis- currently, no known RNA-Seq analysis has been carried out using F0 populations of embryos and therefore this chapter is of importance when deciding on future techniques in both this thesis and in the wider *Xenopus* community.

3.3.1. Using F0 CRISPR/Cas9 targeted embryos

CRISPR/Cas9 gene editing very effectively targeted the MyoD gene in embryos, however not all INDELs result in alleles that would generate a disrupted protein. Approximately 80% of injected embryos are successfully targeted and the penetrance of mutation in each individual is also very high. However, because

Cas9 can act on one or both (or neither) alleles in cells as the early embryo divides, and the nature of NHEJ is that it leads to random INDELs, the resulting F0 embryos are inherently genetically mosaic. This leads to a population of F0s with ill-defined genotypes, with less than half of alleles analysed carrying a disruptive mutation. In zebrafish, it is standard practice to outcross founder fish and breed to a known mutant genotype (Li et al., 2016); however this technique is not feasible for *Xenopus* as outcrossing frogs requires more space and time.

Nevertheless, it has been established that using founder embryos from gene editing protocols in transcriptional analyses is both feasible and valuable. CRISPR/Cas9 targeting of MyoD results in a significant reduction of MyoD transcripts overall and a high percentage of these with INDELs (Figure 3.3.); moreover, the known target gene Rmb24 (Seb4) is significantly down regulated in these samples. RNA-Seq analysis has provided a shortlist of genes that require MyoD, *in vivo*, prior to myogenic differentiation (Table 3.1.).

Analysis of putative MyoD targets in the context of a published time course of gene expression during *Xenopus tropicalis* development (Tan et al., 2013) provided a way of curating genes on the basis of temporal expression, however spatial restriction of expression is also an important factor to consider.

3.3.2. Rbm24

Rbm24 has a single N-terminal RNA Recognition Motif (RRM) which is conserved throughout both invertebrates (*C.elegans*) and vertebrates. MyoD and Rbm24 (aka Seb4) share a very close expression pattern, both temporally and spatially, with the notable exception that Rbm24 is expressed in the cardiac as well as the

skeletal muscle cell lineage (Fetka et al., 2000; Li et al., 2010a; Maguire et al., 2012). It was found that Rbm24 expression is at low levels in myoblasts expressing MyoD, then accumulates over time and is expressed in all skeletal muscle lineages (Grifone et al., 2014). Rbm24 (Seb4) is an RNA binding protein essential for cardiac and skeletal muscle specific alternative splicing (Cardinali et al., 2016; Yang et al., 2014). In mice, inducible Rbm24 mutants show defects in skeletal muscle M-band formation (Yang et al., 2014) and in C2C12 cells, Rbm24 promotes myogenic differentiation through upregulation of Myogenin. This is achieved through interaction of Rbm24 with the 3' UTR of Myogenin, promoting its stability within the cells. Knock-down of Rbm24 in C2C12 cells results in reduced levels of Myogenin and inhibited differentiation (Jin et al., 2010). Supporting this, downregulation of Rbm24 by the micro-RNA miR-222 causes defects in muscle differentiation which can be rescued through Rbm24 overexpression (Cardinali et al., 2016), whilst it has also been shown previously to be a direct target of MyoD (Li et al., 2010a). Recently, it has been found that Rbm24 is also essential for normal somitogenesis in fish (Maragh et al., 2014) consistent with the findings in *Xenopus*.

3.3.3. Rbm20

Rbm20, a related gene with similar functions specifically in driving cardiac development by directing cell specific alternative splicing (Li et al., 2013; Paquette et al., 2014), was also identified as a target in the analyses. Mutated rbm20 has been previously identified as a causative factor of Cardiomyopathy, and regulates genes related to cardiac development through repression of alternative splicing (Li et al., 2010b). A notable gene regulated by Rbm20 is the sarcomeric protein Titin

(Guo et al., 2012; Li et al., 2013). The spatial expression pattern of Rbm20 has not been examined in *Xenopus*, however, in chick embryos it shows very early (yet transient) expression in somites with persistent expression the heart (Geisha.arizona.edu).

3.3.4. Zeb2

Zeb2 codes for an E-box binding repressor, which could act like Snail repressors in modulating 'enhancer swapping', where MyoD binds to regulatory sequences in different genes in myoblasts as compared with myotubes (Soleimani et al., 2012). Moreover, during gastrulation Zeb2 is expressed in the dorsal marginal zone with some more lateral mesodermal expression overlapping with MyoD.

3.3.5. Gli2

Gli2 is a downstream effector of the Shh signalling pathway which has roles in many developmental programs, including myogenesis in amniotes (Pan et al., 2006; Sasaki et al., 1999). In amniotes it has been shown previously that Gli2 and Gli3 effectors of Shh signalling promote skeletal muscle determination through regulation of MRFs. Whilst both Gli2 and Gli3 are sufficient to induce expression of Myf5 in somites (Borycki et al., 1999; McDermott et al., 2005), Gli2 has also been indicated as an essential factor for MyoD expression and promotes MyoD activity through direct interaction with MyoD and Mef2C at target gene promoters (Voronova et al., 2013).

In *Xenopus*, Gli2, along with Zeb2 (aka XSip1) is expressed in the neuroectoderm just after gastrulation (Aguero et al., 2012; Papin et al., 2002). The mesoderm, where MyoD is active as a transcriptional regulator, provides signals that instruct

the overlying ectoderm to become neural tissue, so it is possible that there is an indirect regulation of these genes by MyoD. There is however, no in situ hybridization data for gastrula specific expression of Gli2.

This reflects one limitation of this type of study where the genes identified are not necessarily direct targets, particularly when a transcriptional network such as that downstream of MyoD, is so wide. Further investigation into the expression pattern of Gli2 at this stage might reveal in more detail, the potential that MyoD regulated effectors of Shh signalling. This is of particular interest as, in amniotes, Shh signalling from the floorplate activates expression of MyoD. In *Xenopus*, as MyoD is expressed much earlier than the formation of the floorplate, this regulatory pathway may be more complex, and result in downstream effects on MyoD expression at later stages.

3.3.6. FoxC1 and FoxC2

FoxC1 shows both early mesodermal and later somitic expression in *Xenopus*, and like Rbm24, FoxC1 was also identified as a direct target of MyoD in a previous analysis using morpholino oligos (Maguire et al., 2012). And chromatin immunoprecipitation (ChIP) of upstream promoter elements showed direct binding of MyoD at identified E-boxes.

FoxC1 and FoxC2 are significant targets in the analyses and are known to be expressed in the paraxial mesoderm amniotes (Kume et al., 1998), as well as fish and frogs (Köster et al., 1998; Maguire et al., 2012; Topczewska et al., 2001). FoxC1/C2 are essential for somitogenesis (Kume et al., 2001; Topczewska et al., 2001), and identified as transcriptional targets of MyoD in previous studies (Gianakopoulos et al., 2011; Maguire et al., 2012). In the early paraxial mesoderm,

FoxC1/C2 are co-expressed with the early muscle regulator Pax3, however later in somitogenesis FoxC1/2 regulate the endothelial lineage (Lagha et al., 2009; Mayeuf-Louchart et al., 2014) but their expression is nonetheless essential for the normal migration of muscle precursor cells to the limb (Mayeuf-Louchart et al., 2016). Interestingly, Gli2 has been found to act upstream of FoxC1/2 in the induction of myogenesis in P19 cells (Savage et al., 2010).

3.3.7. Pbx2

Pbx2 is also an interesting target as this family of TALE-class homeodomain proteins are associated with myogenesis (Berkes et al., 2004; Maves et al., 2007) and binding sites for Pbx transcription factors are found in regions of the genome associated with MyoD binding (Fong et al., 2015); it is thought that Pbx proteins help 'pioneer' or establish the myogenic programme (Yao et al., 2013a). This model fits well with *Pbx* genes being early targets of MyoD *in vivo*.

3.3.8. Sp8 and Sp5

Genes coding for the zinc finger transcription factors Sp8 and Sp5 were also identified as targets in a screen for early targets of Wnt signalling (Nakamura et al., 2016) and as downstream regulators promoting FGF signalling (Branney et al., 2009; Kasberg et al., 2013). Identifying the genes coding for Sp5/8 as early MyoD targets is consistent with the important role for FGF (Fisher et al., 2002) and Wnt (Hoppler et al., 1996) signalling in activating MyoD in *Xenopus*. The fact that some regulators are identified that are known to act with MyoD (such as Pbx and Sp5/8) as downstream targets of MyoD is not surprising as the analyses focus on a

window of time very early during the specification of the myogenic lineage when transcriptional feed-forward pathways are being established.

More surprising is that later targets of MyoD are detected at such an early developmental stage: the skeletal muscle specific protocadherins (pcdh8) and kinases (pgk1 and pygm), and follistatin (fstl1) are expressed in somites (Berti et al., 2015) and are notable as they are identified as MyoD targets at such an early time-point. Of particular interest is that previous study had identified fstl1 as being negatively regulated by MyoD through the induction of miR-206 expression (Rosenberg et al., 2006), yet in CRISPR/Cas9 targeted embryos it appears downregulated. To determine whether this is due to differential regulation by MyoD in different model systems or at different stages of myogenic differentiation, further investigation is needed.

3.4. Conclusions

Overall, this chapter illustrates the ease of genetically modifying *Xenopus* embryos using CRISPR/Cas9 techniques. Cas9 specifically disrupts at the target sequence with no off-targeting of other related MRFs present at the time of the analysis. F0 embryos show high targeting rates, as determined by T-cloning and sequencing, and there is significantly downregulated expression is detected of MyoD and its target genes. However, despite high targeting rates, less than half of the mutated sequences produced resulted in a frame-shift in the coding sequence of MyoD. Therefore while there is a significant loss of MyoD expression, it is not likely to be as complete as seen by targeting using MOs (Maguire Figure 1).

Due to the mosaicism of F0 embryos, achieving a high level of protein reduction would require out-crossing of mutants selected for specific mutations to produce $MyoD^{+/-}$ heterozygous and $MyoD^{-/-}$ homozygous populations of frogs. Even so, this chapter has identified novel transcriptional targets of MyoD in the frog gastrula and future adaptation of the CRISPR/Cas9 protocol such as injection of 2 gRNAs simultaneously to induce a larger deletion, or by targeting promoter elements to prevent transcriptional activation would make CRISPR a powerful tool in F0 embryos. Nevertheless, this study provided an insight into the early transcriptional activity of MyoD, *in vivo*, prior to myogenic differentiation and has supported some of the findings of previous Morpholino experiments.

4. Characterising the expression of Polr3G and Polr3gL in vivo

4.1. Introduction

Unlike RNA Polymerase I and RNA Polymerase II, RNA Polymerase III exists as two distinct forms depending on the alternate presence of the small subunit Polr3G or Polr3gL (RPC32α or RPC32β) (Haurie et al., 2010).

Polr3G has been associated with the proliferative state; transformed cells and pluripotent human stem cells have been shown to express higher levels of Polr3G than Polr3gL with expression of Polr3G, with its expression being downregulated during induced differentiation programmes such as exposure to Retinoic Acid (Wong et al., 2011). Polr3G overexpression has been shown to render ES cells more resistant to differentiation programmes of all three germ layers making it a novel factor of pluripotency during development. Polr3gL has been less studied in the literature, but is knockdown in stem cells leads to loss of cell survival, and its expression is maintained during differentiation (Haurie et al., 2010). These data suggest very different potential roles for these factors *in vivo*.

This chapter presents the cloning and expression analysis of these two isoforms in *Xenopus tropicalis*.

4.1.1. Aims of this chapter:

1. To identify Polr3G and Polr3gL in *Xenopus.*
2. To characterise the temporal and spatial expression profiles of Polr3G and Polr3gL *in vivo*.
3. To determine whether the expression of Polr3G and/or Polr3gL is regulated by the myogenic factor MyoD.

4.2. Results

4.2.1. Identification of Polr3G and Polr3gL genes in *Xenopus tropicalis.*

In order to identify predicted Polr3G and Polr3gL genes in *X.tropicalis,* protein and mRNA sequences were identified from NCBI databases. Nucleotide and Protein BLAST analysis was carried out for the identified sequences in order to identify similar sequences in other organisms. Nucleotide BLAST analysis was also carried out for the published *X.tropicalis* version 8.0 genome to ensure the sequences were present. To determine that the annotated sequences were correctly identified and either Polr3G or Polr3gL in *Xenopus*, Polr3G and Polr3gL protein sequences for multiple vertebrate model organisms were identified in NCBI and a phylogenetic tree was constructed in MegAlign (Figure 4.1). The protein sequence for *X.tropicalis* Polr3G clustered closely with *X.laevis* Polr3G and was also more closely related to mammalian Polr3G than amphibian Polr3gL. All predicted Polr3gL sequences were also more closely related to each other than Polr3G sequences. Therefore, the annotated predicted genes were confirmed as Polr3G and Polr3gL.

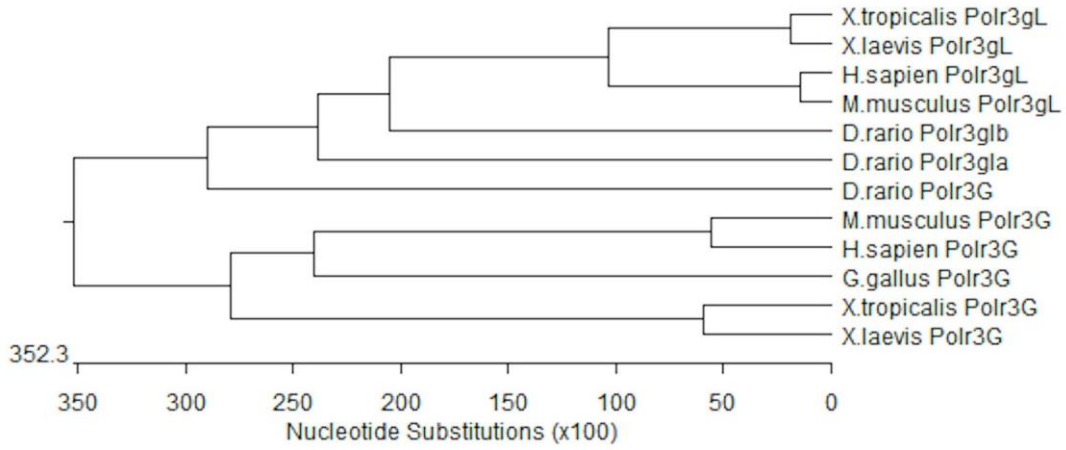


Figure 4.1. The molecular phylogenetic analysis of Polr3G and Polr3gL proteins across vertebrates. The phylogenetic reconstruction was produced using the Jotun-Hein Method using DNAStar MegAlign software. The scale bar represents a genetic distance of 100 nucleotide substitutions.

Polr3G (and possibly Polr3gL) interacts with another RNA Polymerase III subunit, RPC62, via a core domain predicted to include amino acid residues 50 to 100 (Boissier et al., 2015). Transcription initiation by RNA Polymerase III is dependent on the interaction of these subunits, along with one other to form a tertiary complex which is then incorporated into the RNA Polymerase III complex, suggesting that this domain may be important for protein function. Alignment of *X.tropicalis* Polr3G and Polr3gL protein sequences revealed that the paralogues shown only 48.4% identity (Figure 4.2. Unshaded indicates sequence difference). Low conservation was observed for much of the protein structure including both the core domain implicated in RPC62 interaction (Figure 4.2. Blue underline) and the peptide recognition sequence for a *Xenopus tropicalis* specific Polr3G antibody located at sites 91-103 (Figure 4.2. orange box).

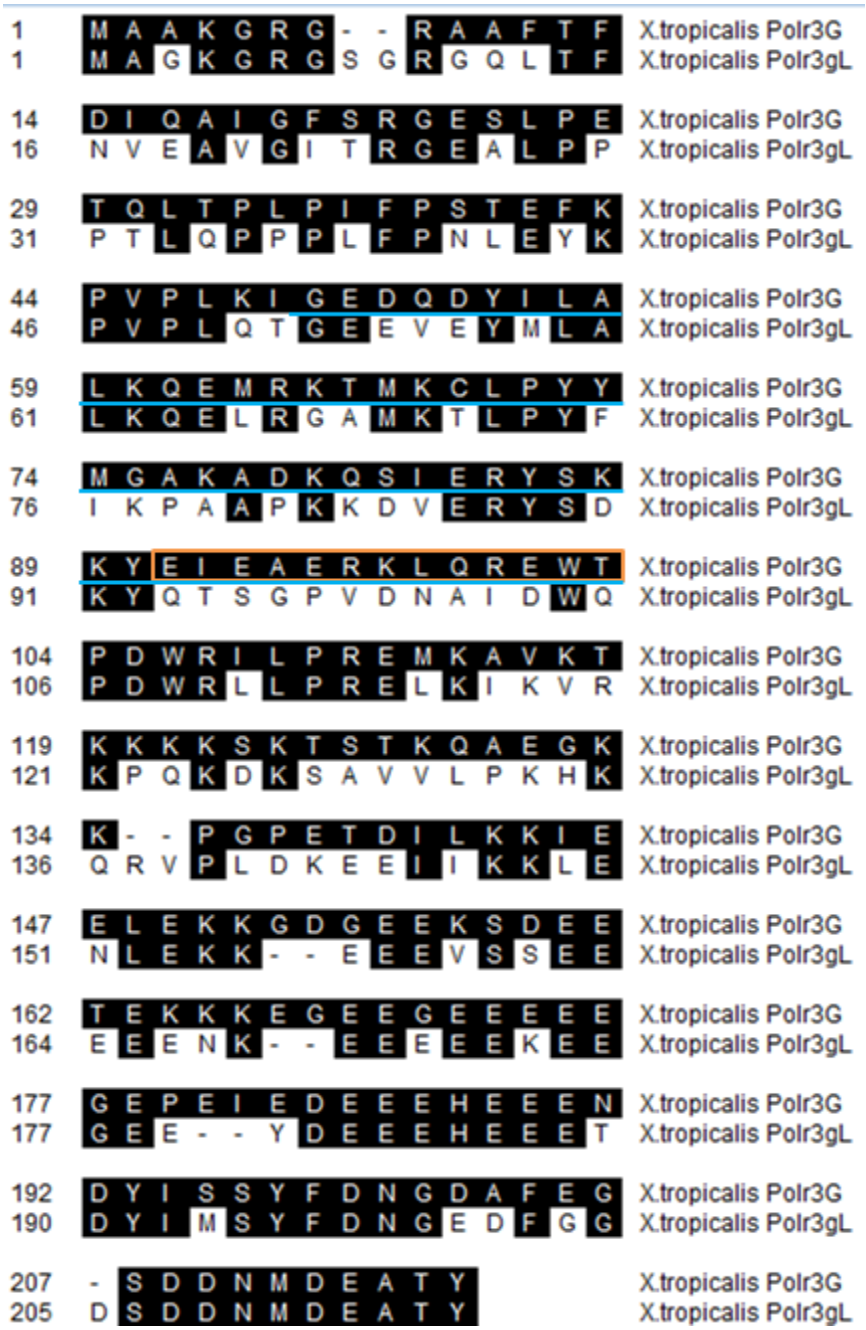


Figure 4.2. Alignment of *X.tropicalis* Polr3G and Polr3gL protein sequences. Protein sequences were aligned in MegAlign using the Clustal W method. Divergence from Polr3G sequence is indicated in white. Peptide recognition sequence for custom-produced antibody against Polr3G is indicated by the orange box. The core domain implicated in interaction with RPC62 is indicated by blue underline (amino acids 50-100).

4.2.2. Cloning of regions of *Xenopus tropicalis* Polr3G and Polr3gL coding sequence for in situ probe synthesis

Sequence differences at the nucleotide level allow the design of specific RT-PCR primers and allow distinct probes for in situ hybridisation analysis of both subunits. Figure 4.3 shows the nucleotide sequence alignment of *X.tropicalis* Polr3G and Polr3gL coding sequences. Primers are indicated in red for Polr3G sequence and green for Polr3gL. Sequence divergence in the regions amplified by these primers is sufficient to distinguish between the two paralogues in *X.tropicalis*, enabling RT-PCR and in situ hybridisation to be carried out to compare expression profiles of the subunits.

Figure 4.3. Alignment of *X.tropicalis* Polr3G and Polr3gL coding sequences. Sequences were aligned in DNAStar MegAlign using the Clustal W method. Sequence differences from Polr3G sequence are indicated in white. Primers designed for use RT-PCR and subsequent amplification of coding sequence regions for in situ hybridisation probe synthesis are indicated by red (Polr3G) and green (Polr3gL) arrows.

In order to characterise expression profiles of Polr3G and Polr3gL in *X.tropicalis* during development, RT-PCR and in situ hybridisation (ISH) assays were used. To generate probes, 400bp-600bp sections of the Polr3G and Polr3gL coding sequences were amplified from unique regions of each cDNA to ensure specific recognition of the different mRNAs. These sections were also used for RT-PCR analysis (Figure 4.5.A). The amplified products (Figure 4.4. A) were subsequently T-cloned into the pGEM T-Easy vector system for in situ hybridisation probe synthesis (Figure 4.4. B). The plasmids were sequenced to determine the orientation of insertion and antisense cRNAs were transcribed incorporating DIG-UTP.

4.2.3. Cloning full length Polr3G and Polr3gL mRNA for in vivo overexpression.

In order to investigate gene function in *Xenopus*, synthetic mRNA injection to overexpress a protein of interest is often used as an initial functional assay. The full length sequence of Polr3G and Polr3gL mRNAs, including 3' UTR, were collected from the NCBI mRNA database. Full length transcripts were amplified from cDNA at NF Stage 4 (Polr3G) and NF Stage 25 (Polr3gL) and were engineered to have restriction sites at 5' and 3' ends to allow for directional cloning into $CS2^{+}$ (Figure 4.4. C) in addition to a KOZAK translation initiation sequence.

PCR products were T-cloned and subsequently excised from pGEM vectors using the designed restriction sites. Excised fragments were ligated into $CS2^{+}$ vectors (Figure 4.5. D) digested with BamHI and XbaI and EcoRI and XbaI for Polr3G and

Polr3gL respectively. Plasmids were transformed into competent *E.coli*. Purified plasmid preps were then digested using the same restriction enzymes to confirm successful ligation before sequencing to ensure the predicted full length transcripts were successfully inserted (Figure 4.4. C).

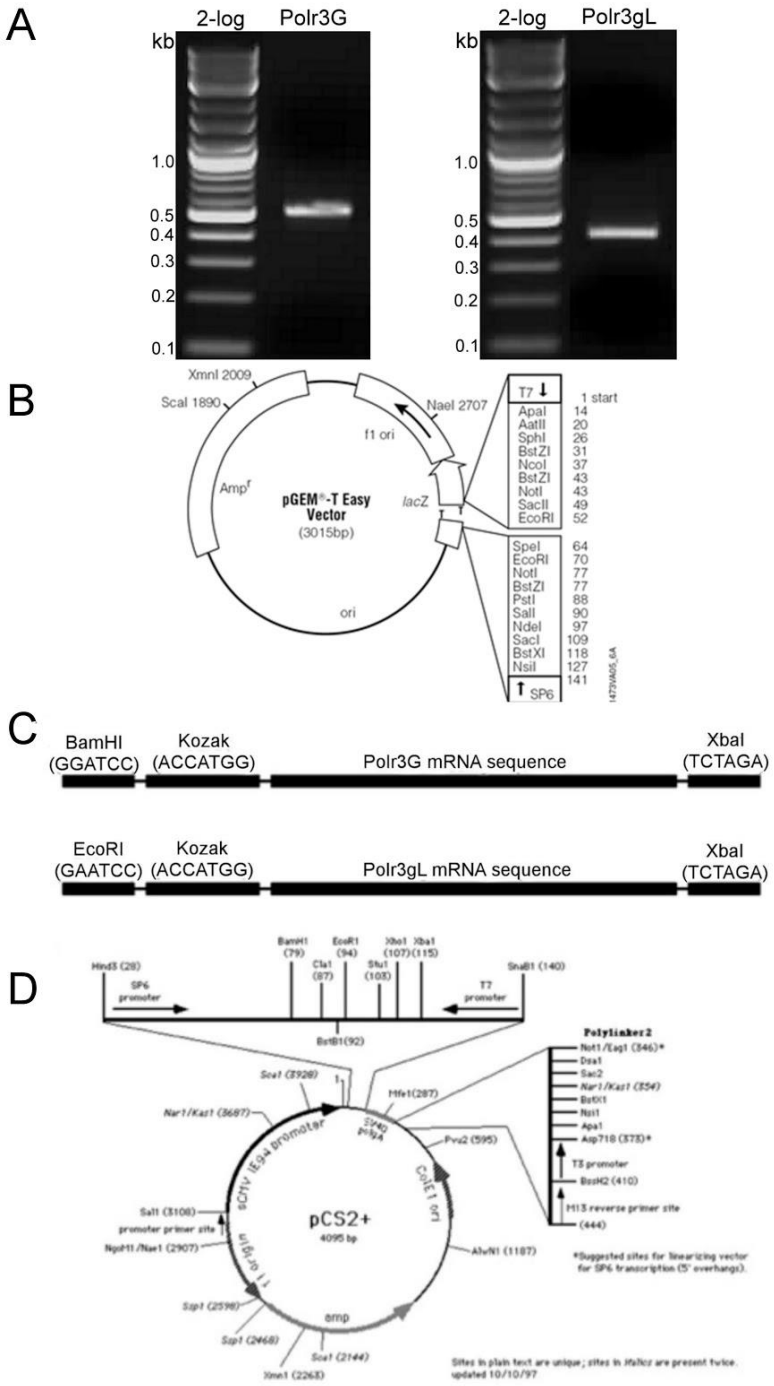


Figure 4.4. PCR amplification of Polr3G and Polr3gL coding sequence regions for in situ probe and SP6-based mRNA synthesis. (A) Regions of Polr3G and Polr3gL CDS amplified by PCR using cDNA from embryos at NF stage 4 and NF stage 25 respectively. (B) Plasmid map of pGEM-TEasy vector. PCR product were cloned into pGEM and sequenced to determine the necessary linearisation and transcription enzymes to synthesise antisense probes. (C) Schematic showing restriction sites used for cloning the full-length Polr3G and Polr3gL CDS into the expression vector $CS2^{+}$. (D) Plasmid map of the $CS2^{+}$ vector.

4.2.4. Expression patterns of Polr3G and Polr3gL during *Xenopus* development.

4.2.4.1. Temporal expression patterns of Polr3G and Polr3gL.

To determine the RNA expression patterns of Polr3G and Polr3gL during *Xenopus tropicalis* development, a stage series RT-PCR was carried out on the paralogues across all major stages of early embryogenesis (Figure 4.5. A). Polr3G expression is highest maternally, early in development, prior to Midblastula Transition (MBT) at NF stage 8.5. After the activation of embryonic transcription, Polr3G expression decreases rapidly to undetectable levels by NF stage 11. Polr3gL is not expressed maternally and is activated after MBT. It becomes detectable by NF stage 10 and is at highest expression levels at tailbud stages. This indicates that the two paralogues have different roles during development and that Polr3G is less active after MBT. RNA-Seq data supporting this was also published online (Figure 4.5. C). Polr3G protein expression as detected by western blot analysis is highest in embryos prior to MBT (Figure 4.5. B), however, protein is still detectable in later developmental stages indicating that Polr3G protein is relatively stable *in vivo*.

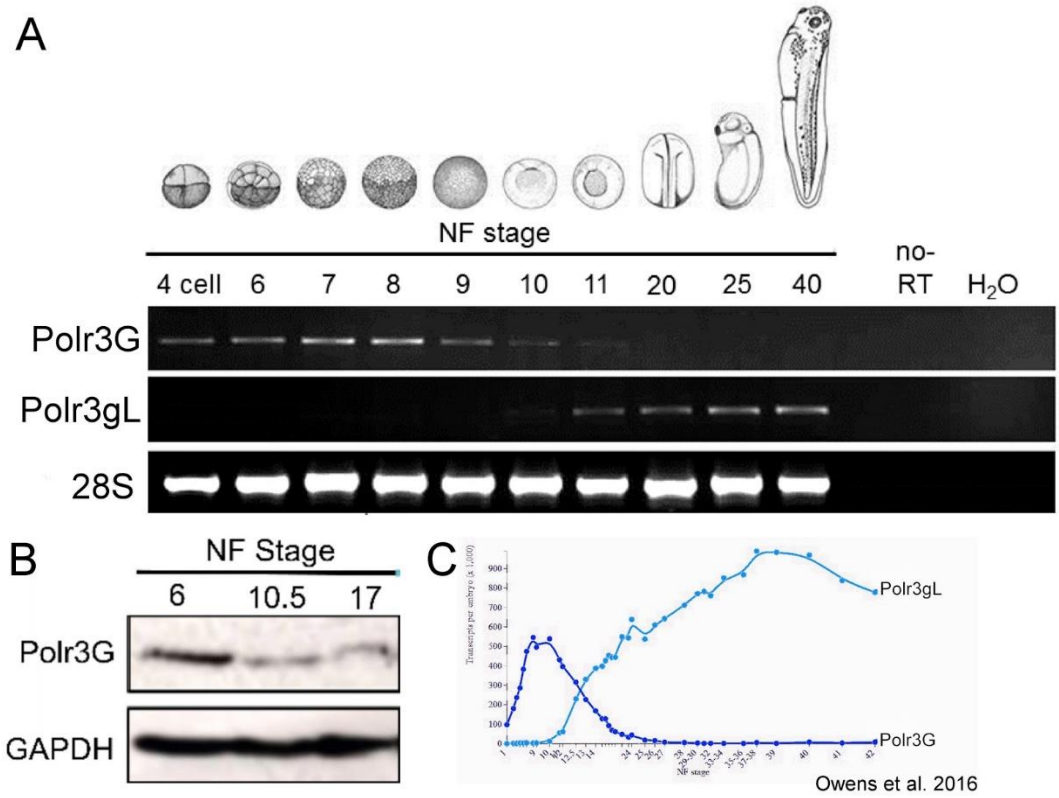


Figure 4.5. Temporal expression analysis of Polr3G and Polr3gL expression during *Xenopus tropicalis* development. (A) RT-PCR analysis of Polr3G and Polr3gL mRNA expression across a stage series of *Xenopus tropicalis* development. Total RNA was extracted from embryos at the desired NF stage using Tri-Reagent. 28S rRNA was used as an endogenous control. (B) Western blotting of temporal protein expression profile for Polr3G during development. GAPDH was used as a loading control. (C) RNA-Seq data for Polr3G and Polr3gL across *Xenopus* development carried out by Owens et al. 2016 and deposited online via Xenbase.org.

4.2.4.2. Spatial expression profiles of Polr3G and Polr3gL.

Spatial localisation of mRNAs can sometimes indicate the tissues or developmental pathways a protein is functional in. In situ hybridisation analysis was therefore carried out to determine whether Polr3G and Polr3gL are localised to specific regions within *Xenopus tropicalis* embryos during development (Figure 4.6.A). In situ specimens were left to develop for a longer period of time (24hours)

at 25°C to allow staining of Polr3G, expressed at low levels, to be detected. Polr3G is expressed maternally and is localised to animal hemispheres of cells during cleavage stages. At NF Stage 20, Polr3G expression is more localised and is detected within paraxial somitic mesoderm (Figure 4.6. B). By tailbud NF Stage 23, Polr3G is detectable only at very low levels and is localised just to somites (Figure 4.6. C-D). This is supported by the RT-PCR analysis indicating high levels of Polr3G expression in early developmental stages which then decreases in later stages. At NF Stage 34, somitic expression of Polr3G is no longer detected and instead, expression is localised to the pronephros and structures of the eye (Figure 4.6. E). This indicates that somitic expression is transient and limited to early tailbud stages of development.

In contrast, Polr3gL is not detectable by in situ hybridisation analysis until later stages. At NF Stage 20, Polr3gL expression is detected in the lateral plate mesoderm and is enriched in both anterior and posterior structures at NF Stage 20 (Figure 4.7. B,), but overlaps with Polr3G as shown in dorsal views (Figure 4.7. B*). By NF Stage 23, Polr3gL is largely detectable only in anterior regions (Figure 4.7. C), branchial arches and hindbrain although some expression is still detectable in the lateral plate mesoderm (Figure 4.7. C-D). Interestingly, whilst expression is detected in the mesoderm, no somitic expression of Polr3gL is detected (Figure 4.7. D). By NF Stage 30, anterior expression has resolved to expression in the branchial arches and hindbrain (Figure 4.7. E, arrows, ba). Expression is also detected in the blood and posterior mesoderm of the tail (Figure 4.7. E, b and pm).

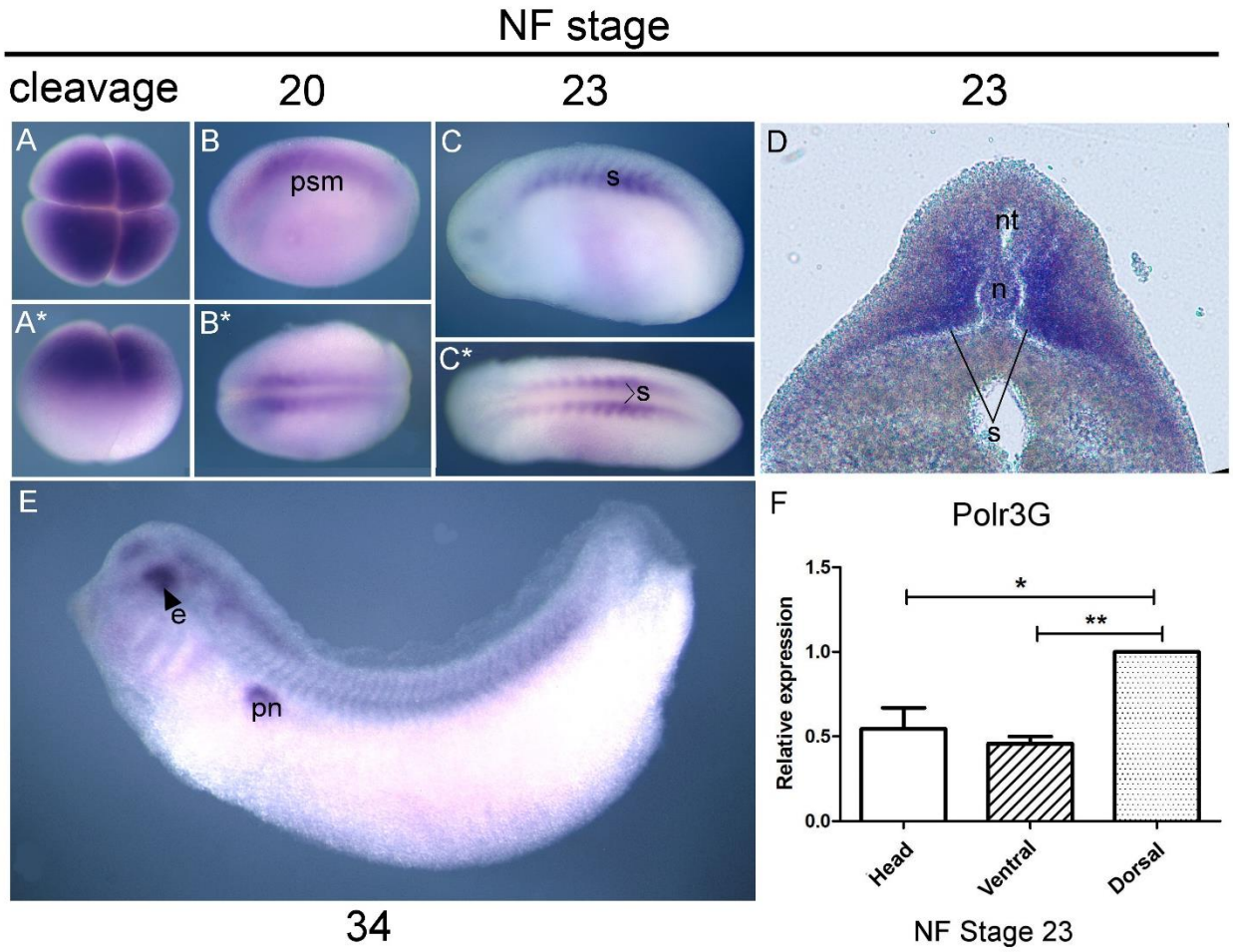


Figure 4.6. Spatial localisation of Polr3G mRNA expression during *Xenopus* development. (A-E) In situ hybridisation of *Xenopus tropicalis* stage series for Polr3G. Embryos were collected at cleavage (4-cell) (A-A*), late neural (St20) (B-B*), tailbud (St23) (C-D) and tadpole (E) stages. In situ specimens were developed in BM purple substrate alongside a MyoD positive control. (D) Vibratome cross-section through St23 embryos probed for Polr3G expression. (s) somites, (n) notochord, (nt) neural tube, (psm) presomitic mesoderm, (pn) pronephros, (e) eye. (F) qRT-PCR for Polr3G (C) expression in tailbud (St23) sections. Head, Ventral and Dorsal sections were collected. Ct values were normalised to dicer. All expression values were calculated relative to somite expression. One-way ANOVA statistical analysis was carried out on relative expression of each section for each gene (*) represents the level of significance.

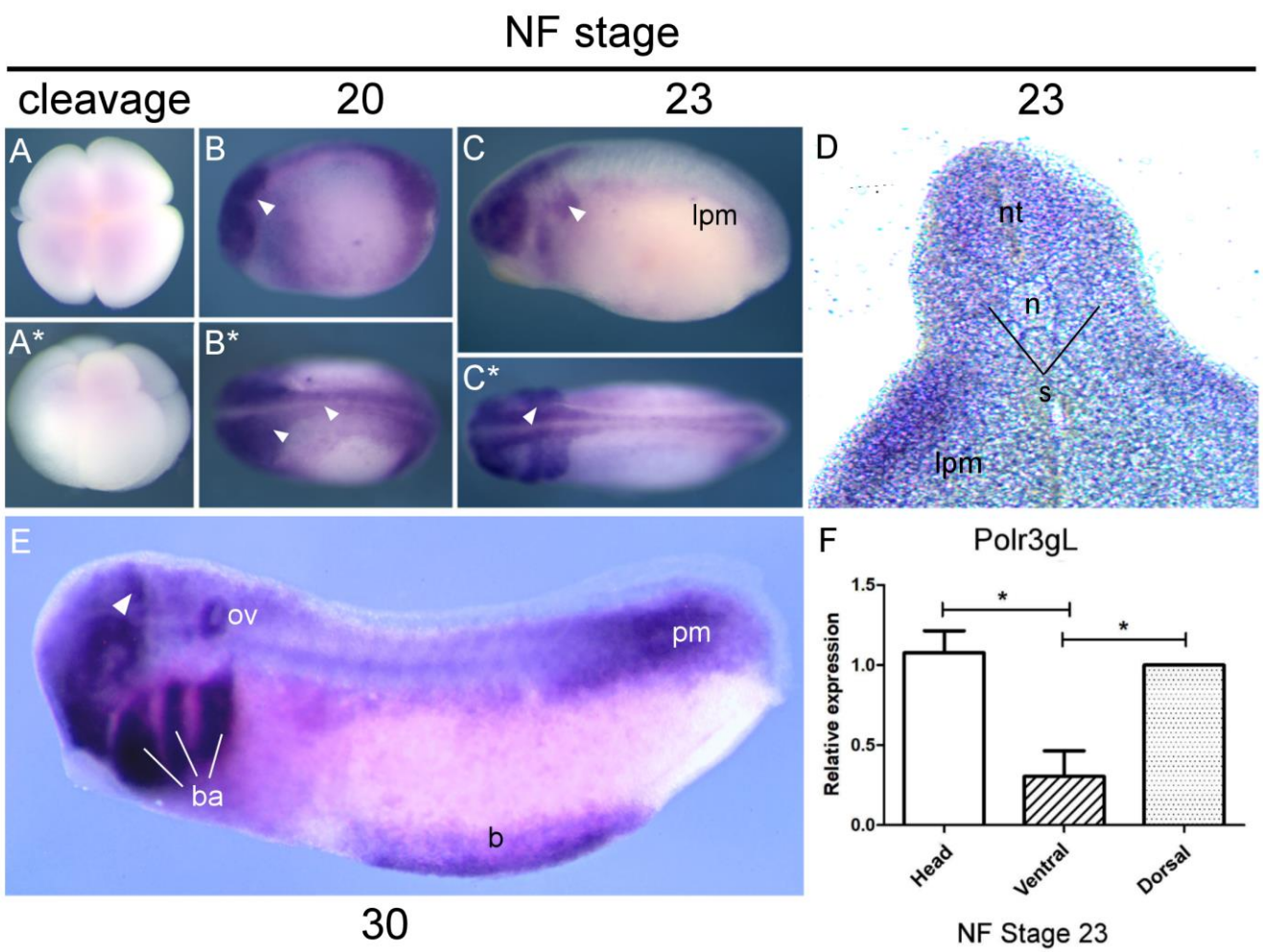


Figure 4.7. Spatial localisation of Polr3gL mRNA expression during *Xenopus* development. (A-E) In situ hybridisation of *Xenopus tropicalis* stage series for Polr3gL. (D) Vibratome cross-section through St23 embryos probed for Polr3gL expression. (s) somites, (lpm) lateral plate mesoderm, (n) notochord, (nt) neural tube, (ba) branchial arches, (pm) posterior mesoderm, (ov) otic vesicle, (b) blood. Arrows in B and B* indicate anterior structures and lateral plate mesoderm. Arrows in C, C* and E indicate hindbrain and branchial arches (C). (F) qRT-PCR for Polr3gL expression in tailbud sections. Head, Ventral and Dorsal sections were collected. All normalised expression values were calculated relative to somite expression. One-way ANOVA statistical analysis was carried out on relative expression of each section for each gene (*) represents the level of significance.

To confirm the overall spatial localisation patterns of Polr3G and Polr3gL in tailbud stages, embryo dissections were carried out at NF Stage 23. Anterior (head), dorsal (somite) and ventral sections were collected and quantitative RT-PCR analysis was carried out on sections to measure mRNA expression levels of Polr3G and Polr3gL relative to levels in the dorsal sections (Figure 4.6. And 4.7. F). Polr3G expression is highest in the somite sections as indicated by in situ hybridisation analysis with expression being significantly higher than both head and ventral sections (Figure 4.6.F). There is no significant difference in Polr3G expression between head and ventral sections. Polr3gL expression is significantly higher in somite sections than ventral sections, supporting in situ hybridisation data, however, there is no significant difference in expression between head and somite sections (Figure 4.7.F). Polr3gL expression is significantly higher in head sections compared with ventral sections. These results confirm expression localisation indicated from the in situ hybridisation analysis and suggest that Polr3G and Polr3gL have distinct temporal and spatial expression patterns during *Xenopus* development.

4.2.5. Regulation of somitic expression of Polr3G by myogenic factors.

4.2.5.1. Locating predicted MyoD binding sites within the polr3g promoter region

Polr3G expression at NF Stage 25 overlaps in the somites with expression of the myogenic regulatory factors MyoD and Mrf4. MyoD is known as the “Master Regulator” of myogenesis in *Xenopus* as it is the first myogenic gene to be induced in the mesoderm by Fgf signalling and MyoD expression alone is sufficient to induce expression of α-actin (skeletal muscle) in animal caps. bHLH factors such as MyoD and Mrf4 bind to DNA sequences with the consensus CANNTG known as E-boxes. Therefore, to determine the possibility of MyoD regulating Polr3G directly, the genomic region surrounding the *polr3g* gene was scanned for sites matching this consensus. As the genome sequence surrounding *Xenopus tropicalis* was incomplete, *Xenopus laevis polr3g* was analysed. *Xenopus laevis polr3g* is similar in genomic structure, with 7 short coding exons and long introns. In the genomic region spanning 2Kb upstream of the 5’ UTR through to coding exon 2, a high number of E box sequences were located (Figure 4.8). This indicates that MyoD may be able to regulate Polr3G directly.

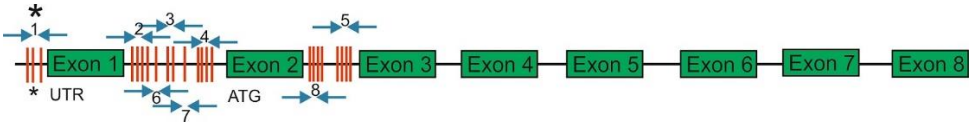


Figure 4.8. Overview of E-box locations within the *Xenopus laevis polr3g* genomic sequence. Green boxes represent exons 1-8. Red boxes indicate the presence of E box consensus sequence CANNTG. Arrows indicate regions of analysis in figure 4.9 and 4.10. (*) represents positive for MyoD binding.

MyoD requires occupancy at two sites within close proximity in order to activate gene expression (Weintraub et al., 1990). Therefore, E-box sites located in the *Xenopus laevis polr3g* promoter region likely to be used for transcriptional regulation would be more likely to be situated in close proximity to another E-box site. The promoter region was therefore more closely analysed in order to identify any closely located E-box sites.

Multiple E-box locations were identified with at least 2 sites in close proximity (Figure 4.9. A-E). Paired E-boxes are first located upstream of the 5' UTR (Figure 4.9.A) whereby a pair of E boxes with the sequences CAGATG and CATTTG are identified. An additional E-Box sequence is located less than 100bp downstream, also with the sequence CATTTG. Four E-boxes are located 855-975bp downstream of non-coding exon 1 with sequences CAGCTG, CAGCTG, CAGGTG and CACCTG respectively (Figure 4.9.B). Any of these E-boxes might be paired with another within the region. Two potentially paired E-box

sequences were located 2156bp and 2176bp downstream of the exon 1 and 704bp and 684bp upstream of the first coding exon and translational start site, both of the sequence CATTTG (Figure 4.9.C). Multiple E-boxes are located around and within the first coding exon with sequences CAGGTG, CACTTG, CATTTG, CAATTG (Figure 4.9.D) and the final region observed was located in the intron between exons 1 and 2 (Figure 4.9.E). This region contained four E-box sequences CATTTG, CAGCTG, CAGGTG and CACATG. This evidence suggests that Polr3G regulatory regions contain DNA elements required for regulation by MyoD and therefore Polr3G could be a target gene activated by MyoD.

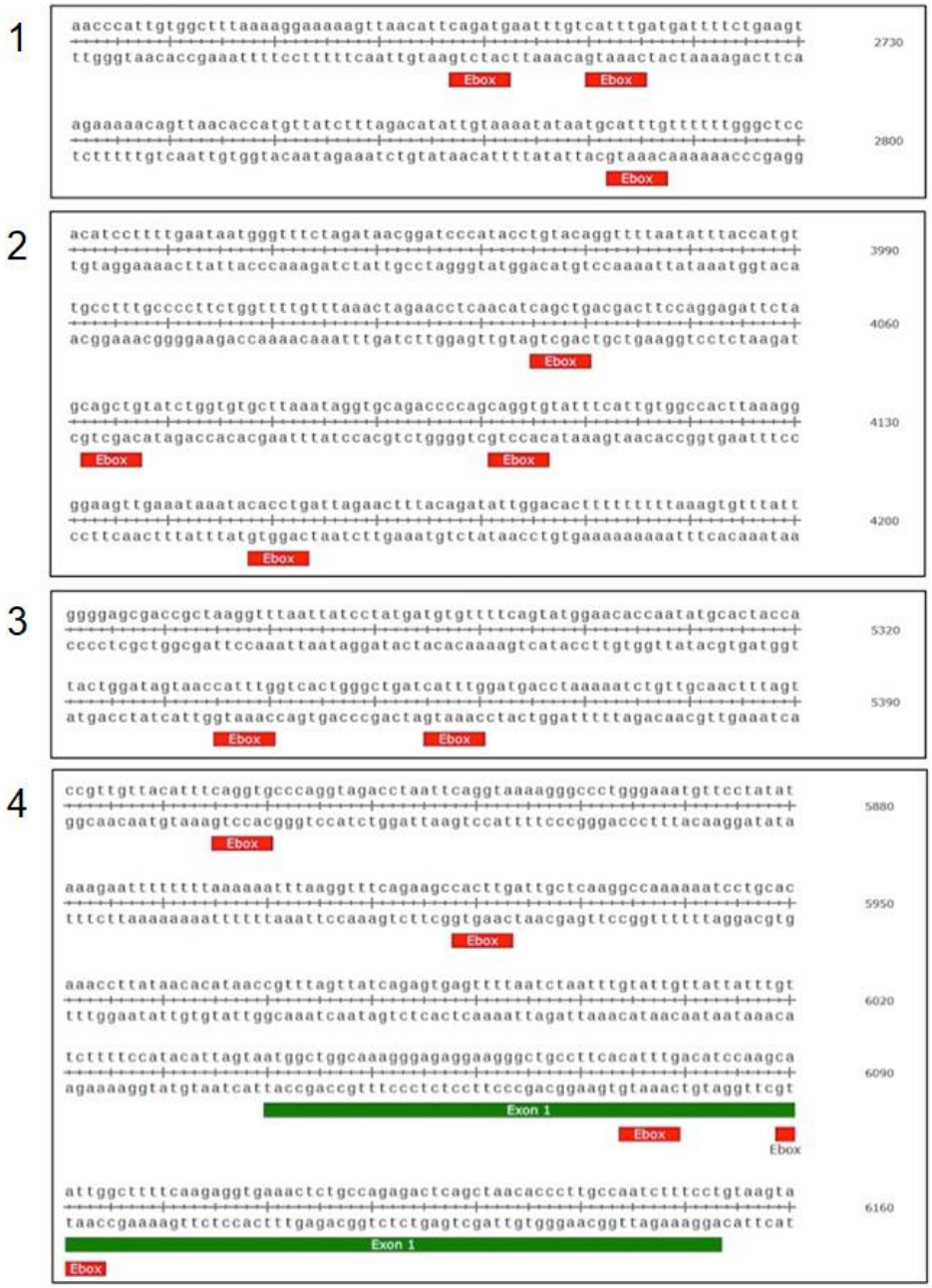

1
aacccattgtggctttaaaaggaaaaagttaacattcagatgaatttgtcatttgatgattttctgaagt
ttgggtaacaccgaaatttttccttttttcaattgtaagtctacttaaacagtaaactactaaaagacttca
Ebox
Ebox
2730
agaaaaacagttaacaccatgttatctttagacatattgtaaaatataatgcatttgtttttttgggctcc
tctttttgtcaattgtggtacaatagaaatctgtataacattttatattacgtaaacaaaaaacccgagg
Ebox
2800
2
acatccttttgaataatgggtttctagataacggatcccatacctgtacaggttttaatatttaccatgt
tgtaggaaaacttattacccaaagatctattgcctagggtatggacatgtccaaaattataaatggtaca
3990
tgcctttgccccttctggttttgtttaaactagaacctcaacatcagctgacgacttccaggagattcta
acggaaacggggaagaccaaaacaaatttgatcttggagttgtagtcgactgctgaaggtcctctaagat
Ebox
4060
gcagctgtatctggtgtgcttaaataggtgcagaccccagcaggtgtatttcattgtggccacttaaagg
cgtcgacatagaccacacgaatttatccacgtctggggtcgtccacataaagtaacaccggtgaatttcc
Ebox
Ebox
4130
ggaagttgaaataaatacacctgattagaactttacagatattggacacttttttttaaagtgtttatt
ccttcaactttatttatgtggactaatcttgaaatgtctataacctgtgaaaaaaaaatttcacaaataa
Ebox
4200
3
ggggagcgaccgctaaggtttaattatcctatgatgtgttttcagtatggaacaccaatatgcactacca
cccctcgctggcgattccaaattaataggatactacacaaaagtcataccttgtggttatacgtgatggt
5320
tactggatagtaaccatttggtcactgggctgatcatttggatgacctaaaaatctgttgcaactttagt
atgacctatcattggtaaaccagtgacccgactagtaaacctactggatttttagacaacgttgaaatca
Ebox
Ebox
5390
4
ccgttgttacatttcaggtgcccaggtagacctaattcaggtaaaagggccctgggaaatgttcctatat
ggcaacaatgtaaagtccacgggtccatctggattaagtccatttcccgggacccctttacaaggatata
Ebox
5880
aaagaatttttttttaaaaaatttaaggtttcagaagccacttgattgctcaaggccaaaaaatcctgcac
tttcttaaaaaaaatttttttaaattccaaagtcttcggtgaactaacgagttccggtttttaggacgtg
Ebox
5950
aaaccttataacacataaccgtttagttatcagagtgagttttaatctaatttgtattgttattatttgt
tttggaatattgtgtattggcaaatcaatagtctcactcaaaattagattaaacataacaataataaaca
6020
tcttttccatacattagtaatggctggcaaagggagaggaagggctgccttcacatttgacatccaagca
agaaaaggtatgtaatcattaccgaccgtttccctctccttcccgacggaagtgtaaactgtaggttcgt
Exon 1
Ebox
Ebox
6090
attggcttttcaagaggtgaaactctgccagagactcagctaacacccttgccaatctttcctgtaagta
taaccgaaaagttctccactttgagacggtctctgagtcgattgtgggaacggttagaaaggacattcat
Exon 1
Ebox
6160

Figure 4.9. Genomic paired E-box locations within the *polr3g* genomic region.

4.2.5.2. Analysis of MyoD binding at Polr3G regulatory sites.

To determine whether MyoD might regulate Polr3G expression in the somites through binding at any of the identified consensus sequence regions, ChIP-PCR analysis was carried out on *X.laevis* NF Stage 25 embryos overexpressing MyoD. In addition to the regions containing multiple E-box sites, with possible paired sites, three regions with single identified E-boxes were also included (Figure 4.8. Regions 6-8). These regions contained E-boxes with the either the sequence CAGCTG or CAGGTG, previously identified as the sequences more closely associated with activated transcription by MyoD in cell lines (Fong et al., 2012).

Three regions of the *rbm24* promoter were also included as a positive control for the ChIP assay. They were identified by a previous study in the Pownall lab in *X.tropicalis* and equating regions within *X.laevis* were located.

MyoD binding was identified in one region of the *polr3g* promoter containing multiple E-box sequences (Figure 4.10. A). Region 1 is located upstream of the first non-coding exon and contains three E-boxes (Figure 4.9. A). No other region with multiple E-boxes showed the presence of MyoD above background levels.

In addition, no region of the *polr3g* promoter with a single E-box sequence was positive for MyoD binding (Figure 4.10. B). MyoD binding was located at region 7, however this was not above background levels shown in the IgG control.

Region 2 of the *rbm24* promoter was shown to be positive for MyoD binding (Figure 4.10. C). Located upstream of the first coding exon, the region 2 paired E-box sequences were identified as CAGCTG and CAGCTG.

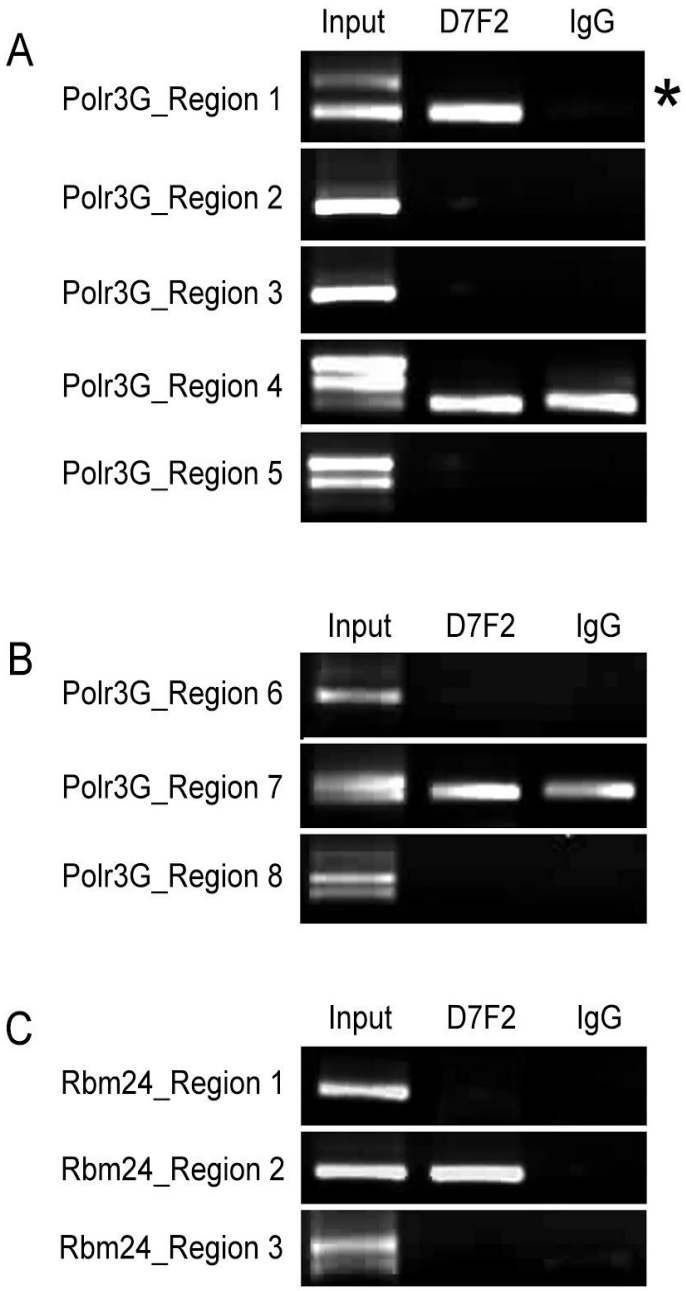


Figure 4.10. ChIP-PCR analysis of MyoD binding at Polr3G promoter sites. (A) Regions in which multiple E boxes were located in close proximity and potentially paired. Regions can be found in Figure 4.8. Above. (B) Regions where a single E-box sequence was located but with the consensus CAGCTG identified as associated with MyoD binding and transcriptional activation. (C) Regions identified within the known MyoD target gene Rbm24 promoter as having multiple E-box sequences in close proximity. Note that nested PCR was used to visualise MyoD binding.

4.2.6. Transcriptional regulation of Polr3G by MyoD

4.2.6.1. Expression of Polr3G/Polr3gL in animal caps overexpressing MyoD protein

The MRFs are responsible for the activation of many myogenic genes. To determine whether expression of Polr3G in the somites is regulated by the MRFs, animal caps from embryos overexpressing MyoD were collected at NF Stage 25 and qRT-PCR analysis was carried out for Polr3G and Polr3gL expression. The known MyoD target gene Actc1 was also included. In caps overexpressing MyoD, Polr3G was significantly upregulated in comparison with control caps (Figure 4.11.). Polr3gL showed no change in expression. The known MyoD target gene Actc1 was also included.

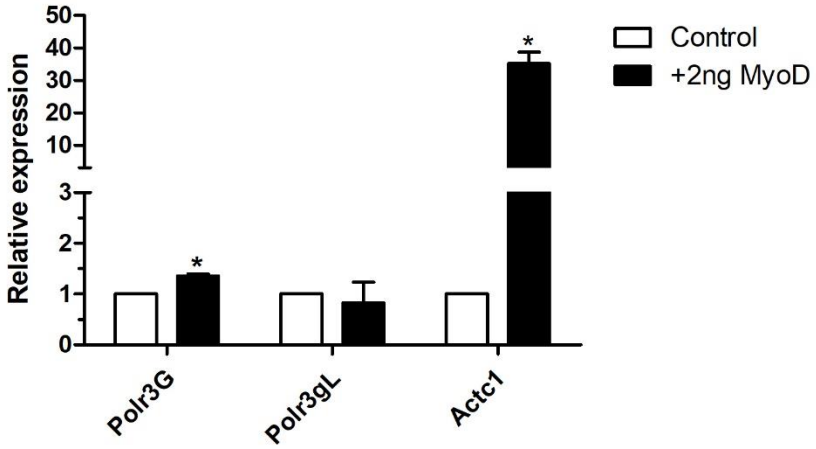


Figure 4.11. qRT-PCR analysis of Polr3G activation by MyoD in *Xenopus tropicalis* animal caps. Analysis was carried out on total RNA for control animal cap and animal caps overexpressing 2ng MyoD for three biological replicates. Ct values were normalised to Dicer and expression values were calculated relative to control using the ΔΔCt method. t-tests were carried out to determine the statistical significance between the mean relative expression of control and MyoD overexpression (*) indicate level of significance.

4.2.6.2. Expression of Polr3G/Polr3gL in MyoD gRNA injected embryos

To confirm regulation of Polr3G by MyoD at NF Stage 25, embryos targeted for MyoD expression by CRISPR/Cas9 were analysed for expression of Polr3G and Polr3gL by qRT-PCR (Figure 4.12.). Whilst MyoD and Rbm24 showed significant downregulation in targeted embryos, Polr3G and Polr3gL both showed modest changes in expression. Polr3G expression was downregulated to 0.91 control levels and Polr3gL was upregulated to 1.05 control expression levels. However, due to the mosaicism of F0 embryos, these smaller expression changes are more likely to be in line with the expected values from overexpression studies.

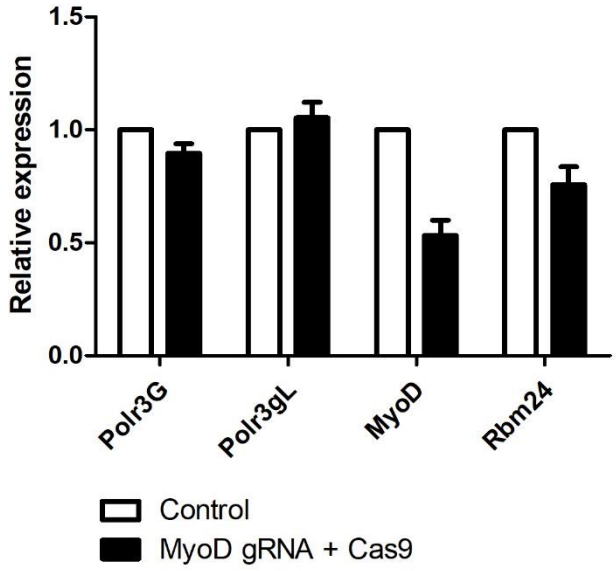


Figure 4.12. qRT-PCR analysis of Polr3G and Polr3gL expression in F0 CRISPR/Cas9 embryos targeted for MyoD. NF Stage 25 embryos targeted by CRISPR/Cas9 were collected and RNA extracted was analysed for expression of Porl3G and Polr3gL alongside MyoD and Rbm24 as control genes. Data was normalised to Dicer and relative expression values were calculated using the ΔΔCt method. Error bars represent SEM (n=2).

4.3. Discussion

4.3.1. Polr3G and Polr3gL have distinct expression profiles during *Xenopus* development

Peak expression of Polr3G in the earliest stages of *Xenopus* development (Figure 4.5) indicate a role during the fast, synchronous cell divisions of cleavage stages. This supports previous data from cell line studies which identify Polr3G as a stem cell regulator. Polr3G is co-expressed with pluripotency factors such as Oct4 in ESCs and is downregulated during induced differentiation. Polr3G is also increasingly upregulated during staged oncogenic transformation of fibroblasts.

In contrast with cell culture studies, Polr3gL is not expressed during early *Xenopus* development and is only activated after the onset of embryonic transcription at MBT (Figure 4.5). This is notable due to the fact that cell culture studies have shown that Polr3gL is essential for cell viability and survival. In human hepatocarcinoma cells, Polr3gL expression levels are maintained during differentiation and Polr3G is unable to compensate for loss of Polr3gL through siRNA targeting (Haurie et al. 2010) despite the maintenance of RNA Polymerase III activity by the Polr3G isoform. This suggests that, unlike cells, Polr3gL is not essential for cell growth and survival during early embryogenesis. It is possible that Polr3G is an essential factor for early developmental viability, however, as Polr3G mRNA and protein is expressed prior to MBT and likely to be maternally deposited into oocytes, depletion of Polr3G expression at this stage is not possible.

After MBT, the expression profiles of Polr3G and Polr3gL remain strikingly different in *Xenopus* embryos. Expression of Polr3G, surprisingly, is isolated to the somites during late neural and early tailbud stages (Figure 4.6. C-D). In *Danio*

rerio, expression of Polr3G genes have also been characterised. The RNA Polymerase III gene *polr3gla* in *D.rerio* is also expressed in somites at the 20-25 somite stage (Thisse et al., 2001) suggesting an evolutionarily conserved role in the muscle lineage for a subunit of the RNA Polymerase III complex during early myogenesis. However, the protein sequence of Polr3gla is distinct from Polr3G sequences of other species and through phylogenetic analysis is identified as Polr3gL (Figure 4.1). Therefore, muscle specific expression of Polr3G in *X.tropicalis* could be due to an acquired developmental role. As previous analysis of Polr3G and Polr3gL activity revealed that both subunits occupy the same target genes, but that the *polr3g* promoter but not the *polr3gl* promoter is occupied by Myc in P493-6 cells (Renaud et al., 2014), it is likely that these two subunits are regulated by different factors during development.

4.3.2. Polr3G expression is regulated by myogenic factors

Polr3G expression at tailbud stages overlaps with expression of key myogenic transcription factors MyoD and Mrf4. Previous analysis of early targets of MyoD revealed that MyoD binding was present at E box sequences in the promoter regions (Maguire et al., 2012). Multiple potential binding sites were located upstream of the *polr3g* coding sequence, and, as needed for MyoD activation, many sites were located in pairs. Unlike previous studies in cell lines, the site CAGGTG most associated with activation of transcription by MyoD (Fong et al., 2012; Fong et al., 2015), were largely not located in the *polr3g* genetic region. This may mean that the predictions made in *in vitro* studies are not replicated in *in vivo* studies. ChIP analysis located a site in the Polr3G promoter region at which MyoD binding was positive. However, unlike other target genes, binding at only one

region was detected whereas MyoD binding at multiple sites was identified in previous studies. Notably, when the promoter region of Polr3gL was also examined, a number of E boxes were also identified (data not shown), yet MyoD overexpression did not lead to upregulated expression of Polr3gL in animal caps (Figure 4.11.). It is therefore possible that other bHLH factors including other MRFs, the neurogenic factor NeuroD2 and possible negative regulators may instead regulate Polr3gL expression via these E box sites. Additionally, as the MRF antagonist family of Id proteins are also bHLH factors, they may bind E box sites in the Polr3gL promoter to prevent MyoD binding (Benezra et al., 1990; Wang and Baker, 2015).

MyoD overexpression in animal caps induces a conversion to a more myogenic state as indicated by the expression of skeletal muscle actin, Actc1. However, alone MyoD is not sufficient to drive full myogenic differentiation and expression of Myogenin and contractile protein genes (Hopwood and Gurdon, 1990). This indicates that MyoD animal cap cells expressing Actc1 are still in a progenitor-like state and, characteristic of early MyoD/Myf5 positive myoblasts, are still proliferative (Cossu and Butler-Browne, 1999). Polr3G expression was significantly upregulated in MyoD induced animal caps, and CRISPR/Cas9 targeting of MyoD resulted in downregulated expression of Polr3G indicating that the expression of Polr3G observed in the somites at tailbud regions is likely to be regulated, at least in part, by MyoD. As Polr3G is expressed at very low levels in post-MBT stages of development, and and RNA Polymerase III complex containing either Polr3G or Polr3gL is predicted to be required for the transcriptional initiation of target genes (Haurie et al., 2010), Polr3G may be an important limiting factor for the rate of transcription by RNA Polymerase III. As such, even modest increases in its

expression are likely to have a significant effect on the expression of its target genes. Interestingly, Polr3gL shows little change in expression in response to MyoD, suggesting that, these factors have different roles in both development and in myogenesis. It is a possibility that Polr3gL is regulated by later myogenic terminal differentiation factors such as Myogenin. However, this was not investigated in this thesis

4.4. Conclusions

This chapter has identified that Polr3G and Polr3gL have distinct temporal and spatial expression profiles during *Xenopus tropicalis* development. Polr3G is expressed at high levels pre-MBT and decreases during gastrulation as Polr3gL is activated. Polr3G expression during tailbud stages is enriched in early skeletal muscle structures and is co-localised with the myogenic factors MyoD and Mrf4 at this timepoint. This expression is transient, and later Polr3G is expressed in the pronephros indicating perhaps that Polr3G has different mechanism of regulation later in development. ChIP-PCR revealed that MyoD is able to bind regulatory regions of the *polr3g* gene similar to the observed binding in previous studies (Fong et al., 2012; Maguire et al., 2012) and increase its expression in animal caps alongside the myogenic marker gene Actc1. As MyoD acts as a myogenic determination factor, it is likely that the observed expression of Polr3G during myogenic differentiation is timed with commitment of cells to become proliferative myoblasts rather than terminal differentiation of skeletal muscle. Nevertheless, this chapter indicates that novel factors of RNA Polymerase III transcriptional regulation may have important developmental roles in *Xenopus*.

5. Regulation of tRNA transcription through the RNA Polymerase III subunit Polr3G

5.1. Introduction

5.1.1. RNA Polymerase III studies in *Xenopus*

Xenopus laevis were utilised as an important model for studying transcriptional regulation of all three RNA Polymerases and took prevalence in the 1970s. The vast numbers of synchronously developing *Xenopus* embryos accessible for biochemical and molecular studies enabled the study of transcriptional control of both RNA Polymerase II and RNA Polymerase III transcripts during oogenesis, the subsequent silencing of transcription during early development, and the reactivation of transcription at MBT (Newport and Kirschner, 1982a; Newport and Kirschner, 1982b; Wormington and Brown, 1983).

It was found that during oogenesis, single oocytes accumulate around 90ng of tRNA over a number of months after which RNA Polymerase III activity, like RNA Polymerase II, is rapidly silenced (Gilbert, 2000). Studies of this kind led to identification of developmentally regulated tRNA and 5S rRNA genes. $tRNA^{Tyr}$ and 5S rRNA have specific genes that are expressed in the oocytes prior to fertilisation and up to MBT, when transcription of zygotically expressed populations is activated (Stutz et al., 1989; Wolffe and Brown, 1988). At this point, the dissociation of transcription machinery from the oocyte-type genes causes the repression of these genes whilst stable interaction of transcription machinery and zygotic genes maintains their expression throughout development. These findings give early indications of the usefulness of *Xenopus* as a model for elucidating the complex mechanisms regulating initiation of transcription by Polymerase III.

Xenopus laevis have also been used to analyse the genomic organisation of RNA Polymerase III targets identifying large clusters in which many tRNA genes are located, and revealing that these loci of >3kb are repeated hundreds of times in

the genome. Moreover, these clusters do not contain continuous tRNA sequences, but these sequences are separated by spacers of substantial size often much larger than the tRNA sequences themselves. This information better characterised possible regulatory regions of tRNAs, and suggested very early, the vast copy numbers of tRNA genes encoded by the *Xenopus* genome (Clarkson et al., 1973; Narayanswami et al., 1995; Rosenthal and Doering, 1983).

The use of northern blot analysis has been common practice in molecular biology for measuring the expression of tRNAs in *Xenopus* and other model organisms and utilises the hybridisation of P^{32} labelled probes complementary to part of the tRNA sequence as a means of detecting expression both prior to and beyond onset of embryonic transcription (Stutz et al., 1989). However, the advent of reverse transcriptase-based methods such as quantitative PCR (qRT-PCR) provided easier and more quantitative methods for analysing transcription.

An online genomic tRNA database predicts the number and genomic location of tRNA genes encoded within the genomes of many model organisms by predicting genome sequences that form cloverleaf structure through an adapted algorithm. This analysis has been carried out to include the *Xenopus tropicalis* genome (Chan and Lowe, 2016; Chan, P.P. & Lowe, 2009). Eukaryotic genomes contain multiple genes encoding the same anticodon and their high sequence homology within families, and synteny between species suggests that this is due to duplication events (Lloyd et al., 2012). It is predicted that in *Xenopus tropicalis* a total of 2638 genes encode tRNAs, this adds complication to looking at tRNA expression by methods such as RNA-Seq and ChIP-Seq whereby genomic mapping is key to the analysis. Multiple genes encode each anticodon and share very high sequence identity. For example, the AlaAGC anticodon is encoded for by

64 genes (Figure 5.1.) and much of the 73bp sequence is identical between all of the genes, making individual genes indistinguishable by transcriptomic methods.

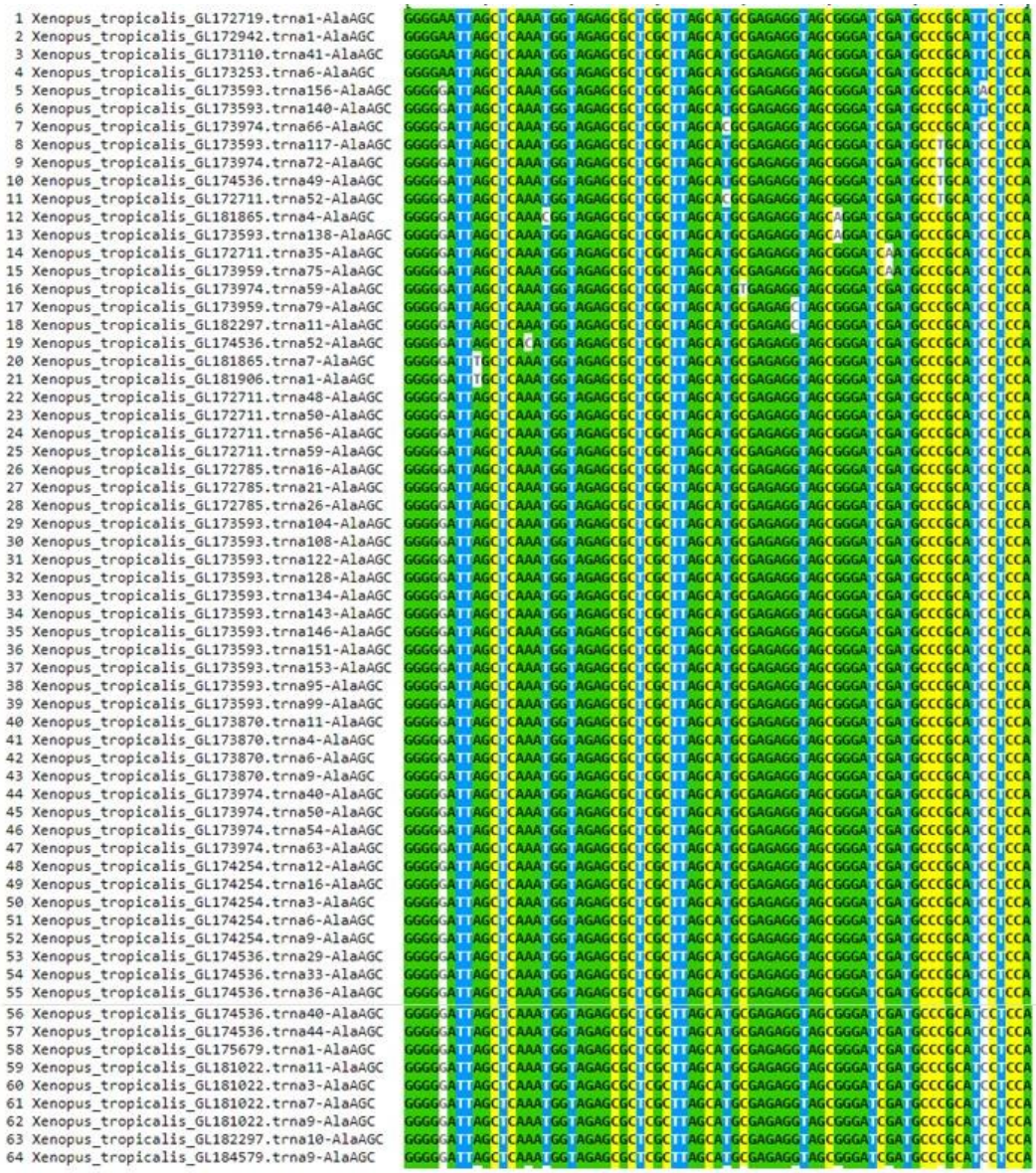


Figure 5.1 Alignment of tRNA Ala^{AGC} genes predicted to be encoded by the *Xenopus tropicalis* genome by the GtRNAdb algorithm. FASTA sequences were aligned using Clustal Omega software. Colours indicate nucleotide identity.

5.1.2. tRNA assays

tRNAs have often been considered to be ‘housekeeping’ RNAs with many publications using their expression as endogenous loading controls (Geslain and Pan, 2011). However, both ChIP-Seq and Microarray studies have indicated that tRNAs are tightly regulated with specific expression profiles amongst different tissue types and cell characteristics (Dittmar et al., 2006; Schmitt et al., 2014; Topisirovic and Sonenberg, 2014). This regulation has been shown to be important in the case of cancers with even small increases in $tRNA^{iMet}$ expression resulting in dysregulated cell division and changes in global tRNA expression patterns (Pavon-Eternod et al., 2013). Therefore, the developmental regulation of tRNA expression could be crucial to defining their potential role in differentiation and disease.

Even now, techniques to effectively measure tRNA transcription at both the nascent and mature levels are not yet developed and tRNAs pose many problems for traditional high-throughput and quantitative methods. Problems with traditional sequencing methods come from tRNA molecules tight tertiary structure, making their sequence largely inaccessible for cDNA synthesis and adapter ligation for RNA-Seq methods. Extensive post-transcriptional modifications, particularly N1-methyladenosine (m1A) and N1-methylguanosine (m1G) (Figure 5.2.), prevent primer extension by traditional Reverse Transcriptase enzymes. Recently, attempts to remove these modifications have been made to improve sequencing of tRNAs (Wilusz, 2015; Zheng et al., 2015), however this technique is not yet widely used with no standard protocols.

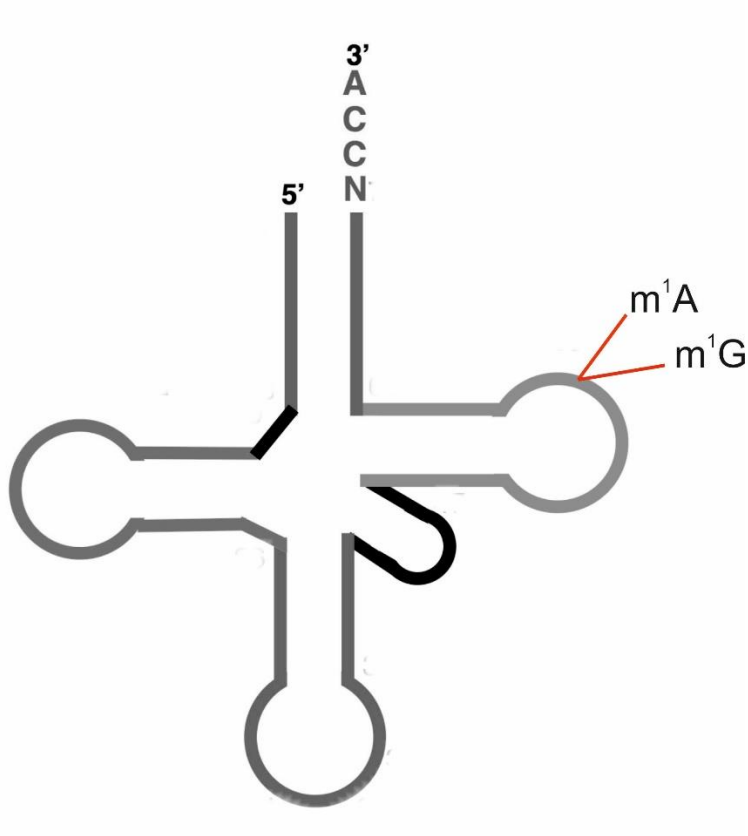


Figure 5.2. tRNA modifications known to be inhibitory for Reverse transcriptase first strand synthesis. m^1A is a methyladenosine modification. m^1G is a methylguanosine modification.

Other high-throughput methods have been used to investigate tRNA expression. ChIP-Seq using an antibody for the catalytic subunit of RNA Polymerase III, Polr3A has been carried out to determine global localisation of Polymerase III at target sites. These studies argue that the presence of RNA Polymerase III binding represents transcriptional activation of target genes and that binding dynamics can show changes in expression. This technique has led to multiple publications investigating global tRNA transcriptional regulation (Kutter et al., 2011; Rudolph et al., 2016; Schmitt et al., 2014). However, ChIP-Seq requires extensive genome annotation as tRNAs are very repetitive in sequence. There are predicted to be over 2000 tRNA sequences in the *Xenopus tropicalis* genome, compared with

approximately 250 in the human genome (Chan & Lowe 2016). This adds complications to analysis as multiple copies of almost identical sequence are located in different genomic locations. In addition, published versions of the *Xenopus* genome, repetitive elements are masked and therefore it is likely that tRNA clusters are removed from genomic data. Moreover, it has been shown in cases of mRNA regulation that the presence of RNA Polymerase II and transcription factors at a site in the genome, is not always associated with activated transcription (Cao et al., 2010; Nakamura et al., 2016). Therefore, ChIP-Seq may not accurately reflect transcription and expression levels of tRNAs.

5.1.3. tRNA microarrays

Another technique for tRNA expression analysis, the tRNA microarray, is based on hybridisation of labelled cRNA to 40-70mer probes which are immobilised onto a glass slide. RNA samples are labelled with the fluorescent dye Cy3 or, to compare two samples simultaneously, Cy3 and Cy5 and hybridised to the slide overnight at a temperature ranging between 42°c and 70°c. Slides are then scanned to produce a high resolution TIF image for which fluorescent intensity can be used as a measure of gene expression. Previous studies have used microarrays designed for human tRNAs in order to look at expression differences between different tissue and cell types (Dittmar et al., 2006; Gingold et al., 2014; Topisirovic and Sonenberg, 2014). Due to the high conservation of tRNA sequences, individual genes are unable to be measured, however isoacceptor families can be distinguished from one-another.

Dittmar et al. used a microarray designed for 49 anticodon families encoding all amino acids in humans plus an initiator methionine sequence (Dittmar et al.,

2006). This study then compared expression of tRNA families between human liver, vulva, testes, ovary, thymus, lymph node and spleen versus levels in the brain. This study showed that different tissues are enriched for distinct tRNA families carrying amino acids of a particular biochemical property. Moreover, by comparing this with mRNA codon usage, expression of tRNA isoacceptors was shown to correlate with abundance of corresponding codons. tRNA microarrays were also used more recently, coupled with RNA-Seq, in order to determine whether proliferative cancer cells contained distinct tRNA profiles from senescent cells (Gingold et al., 2014; Topisirovic and Sonenberg, 2014). This study found that proliferative cells show upregulation of distinct families of tRNAs and that biased mRNA codon usage in proliferation genes match to the specific upregulated tRNAs.

5.1.4. Aims of this chapter

1. To measure tRNA expression *in vivo* during *Xenopus* development.
2. To identify tRNA families and mRNAs activated or repressed by Polr3G activity *in vivo*.
3. To determine if Polr3G overexpression has wider transcriptional effects by analysing mRNA expression in embryos overexpressing Polr3G.

5.2. Results

5.2.1. Overexpression of Polr3G *in vivo.*

In order to investigate gene function in *Xenopus*, synthetic mRNA injection to overexpress a protein of interest is often used as an initial functional assay. A cDNA coding for full length Polr3G mRNA was cloned into the CS2+ expression vector and mRNA was synthesised *in vitro*. Embryos were injected at the 1 or 2-cell stage and allowed to develop to NF Stage 25 when endogenous Polr3G protein levels are downregulated. To show that injection of synthetic mRNA from constructs resulted in increased protein expression, a western blot was carried out for endogenous Polr3G protein and confirmed that protein expression in embryos injected with Polr3G mRNA was significantly upregulated in comparison to uninjected control embryos at NF Stage 25 (Figure 5.3).

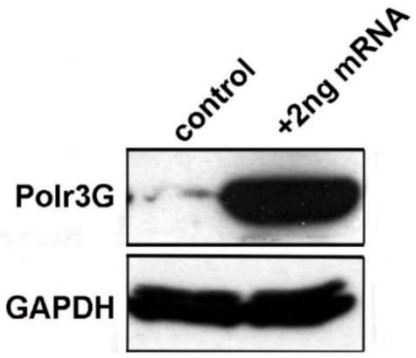


Figure 5.3. Overexpression of Polr3G in *Xenopus tropicalis*. Western blot analysis of endogenous Polr3G protein at NF Stage 25 in control uninjected samples versus embryos injected with 2ng synthetic Polr3G mRNA. A GAPDH loading control was also included.

5.2.2. Regulation of transcription by Polr3G at the onset of transcription

To determine whether Polr3G overexpression has an effect on tRNA expression during development, 2ng of synthetic mRNA was injected into *X.tropicalis* embryos at the 1 or 2-cell stage and these were left to develop until Midblastula Transition (MBT) and the activation of transcription. MBT occurs at NF Stage 8.5 in *Xenopus*, so NF Stages 7, 8 and 9 were collected for analysis.

RT-PCR analysis was carried out for selected tRNA families to determine their expression levels across a timecourse of stages. At the earliest stages of development in control embryos, expression of tRNAs was not detected. In comparison with control uninjected embryos, embryos injected with Polr3G showed increased transcription of four tRNA families with $tRNA^{Leu}$ transcripts detectable at NF Stage 7, and both $tRNA^{Tyr}$ and $tRNA^{iMet}$ detected at NF Stage 8 (Figure 5.4. B). tRNA transcripts are undetectable at NF Stage 7 in control embryos (Figure 5.4. A) and both $tRNA^{Tyr}$ and $tRNA^{iMet}$ are undetectable until NF Stage 9. However, U6 RNA expression levels appears unchanged in embryos overexpressing Polr3G suggesting that Polr3G selectively regulates a subset of RNA Polymerase III target genes. As mature processed tRNAs are not detectable by cDNA/PCR methods, only nascent transcripts from active transcription prior to modification are detected in these analyses. This enables the identification of early changes to tRNA transcription dynamics.

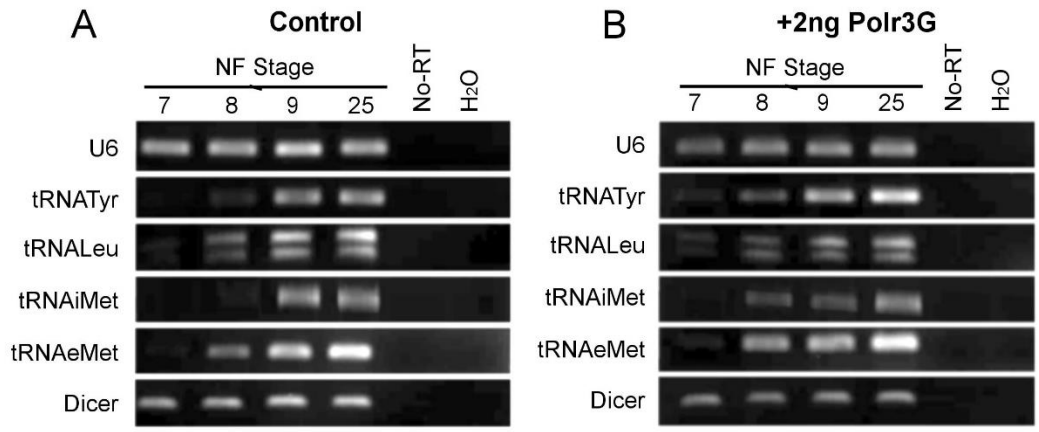


Figure 5.4. RT-PCR analysis of tRNA transcription dynamics in response to overexpression of Polr3G at MBT. *X.tropicalis* embryos were injected at the 1-2 cell stage with 2ng Polr3G mRNA and left to develop until desired stages. RT-PCR for selected families of tRNAs was carried out against control uninjected embryos at the same NF Stages. tRNA families Tyr (GTA), Leu (CAA), iMet and eMet (CAT) were included in the analyses. The RNA polymerase III target U6 RNA was also included as a measure of altered wider Polymerase III transcription and the control gene Dicer was also included. (A) RT-PCR analysis of uninjected control embryos. (B) RT-PCR analysis of embryos overexpressing Polr3G.

5.2.3. Detection of tRNA transcripts across development by Northern Blot

The failure to detect tRNAs before the onset of transcription at MBT by RT-PCR was surprising and as early embryos are utilised to make protein in overexpression studies, it was predicted that this observation was the result of a PCR artefact. In order to test this, northern blot analysis was to measure expression of both $tRNA^{Leu}$ and $tRNA^{Tyr}$ across development. Urea Acrylamide gels denature the tertiary structure of tRNAs and allow hybridisation of P^{32} labelled probes. Mature, processed $tRNA^{Leu}$ and $tRNA^{Tyr}$ transcripts can be detected by 3' probes by bands at 84nt and 73nt respectively, at all stages of development (Figure 5.5. A bottom and middle panel). An additional probe was designed to include the intron sequence of $tRNA^{Tyr}$ genes to allow measurement of nascent

tRNA transcripts and pre-tRNA. Intron probe hybridisation was only detectable in stages post-MBT when tRNA transcription by RNA Polymerase III is activated (Figure 5.5. A top panel). Two bands are detected, one at 103nt representing nascent tRNA transcripts including 5' and 3' sequences and an additional band at 86nt representing pre-tRNA including intron only (Figure 5.5. B). The intron-containing intermediate tRNA transcript is also detected using 3' probes (Figure 5.5. A middle panel- note band of higher molecular weight). Taking into account the correlation of RT-PCR products and the detection of nascent tRNA transcripts at post-MBT stages using Northern Blotting, these results indicate that while RT-based methods such as PCR, RNA-Seq and microarray analyses do not detect mature tRNAs, they can be used to quantify nascent tRNA transcripts in *Xenopus*.

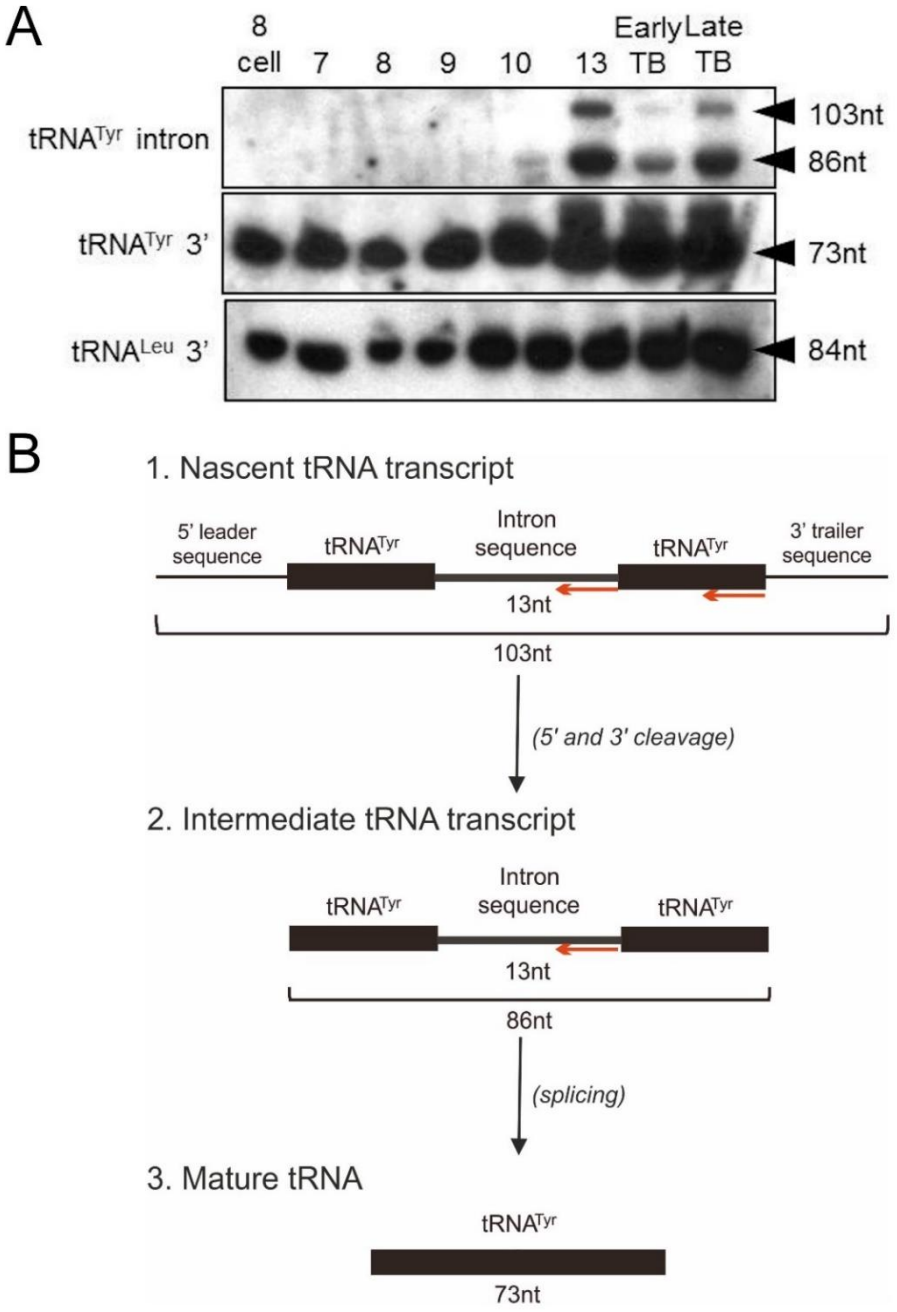


Figure 5.5. Northern blot analysis of both nascent and mature tRNA transcripts and processing of nascent tRNA Tyrosine transcripts. (A) Total RNA was extracted from 10 *Xenopus tropicalis* embryos at the desired developmental stage. RNA was run on an Acrylamide/Urea gel. Probes were labelled with P^{32}. (B) Processing of $tRNA^{Tyr}$ transcripts indicating the bands identified from Northern Blot analysis. Note: This experiment was carried out with my supervisor Dr ME Pownall who has a license to use P^{32}.

5.2.4. Microarray analysis of Polr3G regulation of tRNA expression dynamics at NF Stage 9

5.2.4.1. Microarray design

As active transcription of tRNAs can be detected using cDNA-based methods, in order to look more globally at tRNA expression dynamics in *Xenopus*, I designed a custom microarray by manually compiling tRNA sequences for *Xenopus,* Human and CHO genomes, to include probes for all tRNA isoacceptor families. Selected mRNAs for myogenic and contractile protein genes, as well as other bHLH factors, mesodermal genes and known targets of FGF signalling were also included. Briefly- 5 60mer probes were designed for each tRNA isoacceptor family passing quality checks using Agilent eArray software. Additional manual probes were designed for 5', 3' and middle sections of isoacceptor family sequences not included in the Agilent design. To determine probe specificity, sequences for CHO and Human tRNA isoacceptor families were also manually designed (see methods in Chapter 2 for full design details). Embryos were collected at NF stage 9, timed just after activation of embryonic tRNA transcription as detected by RT-PCR, for control and Polr3G overexpression samples. After quality control was carried out using the Agilent Bioanalyzer 2000 to determine RNA Integrity Numbers (RINs) (Figure 5.6), RNA samples were used to synthesise cDNA and subsequently labelled with Cy3. cRNA samples were hybridised for 17 hours at 65°c and resulting slides were scanned to create a high resolution TIFF file for processing of expression data (see methods 2.4.4).

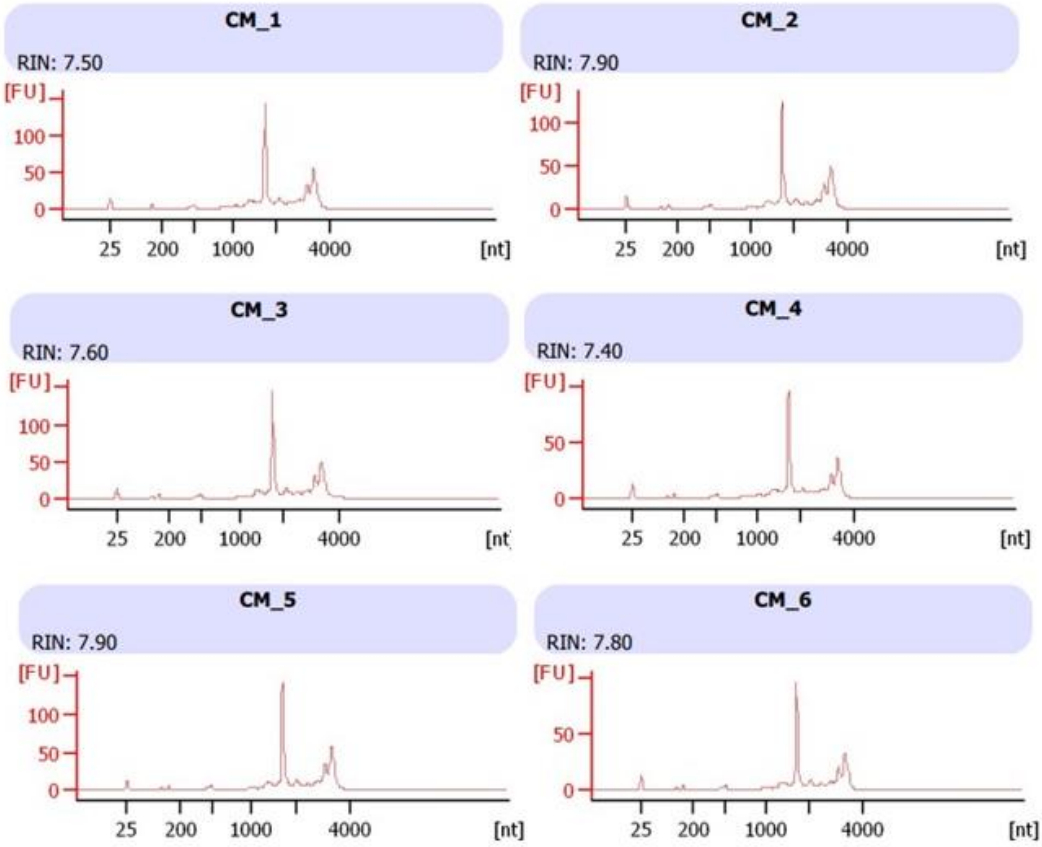


Figure. 5.6. Bioanalyzer results for microarray RNA sample quality control. Samples 1, 3 and 5 refer to control embryo samples at NF Stage 9. Samples 2, 4 and 6 are stage-matched embryos overexpressing Polr3G.

5.2.4.2. Analysing Fold Changes

Analysis of both mean fold change and differences between the means of control and experimental samples were carried out for tRNAs by isoacceptor and by amino acid and summarised in Figure 5.7. Replicate probes gave highly coordinated fold changes and mean normalised expression for each anticodon family and additionally, t-testing of mean expression for Polr3G injected samples versus controls showed that tRNA expression levels were statistically different for many isoacceptor families.

To ensure that changes in expression were sizably different whilst accounting for strong predicted regulation of tRNA genes, fold change thresholds of 0.8 and 1.2 were used to define families as downregulated or upregulated respectively. Interestingly, more isoacceptor families appear upregulated than downregulated (Figure 5.7.A). When isoacceptor families are combined by amino acid isotype, fewer families show changes in levels exceeding the expression change thresholds (Figure 5.7.B).

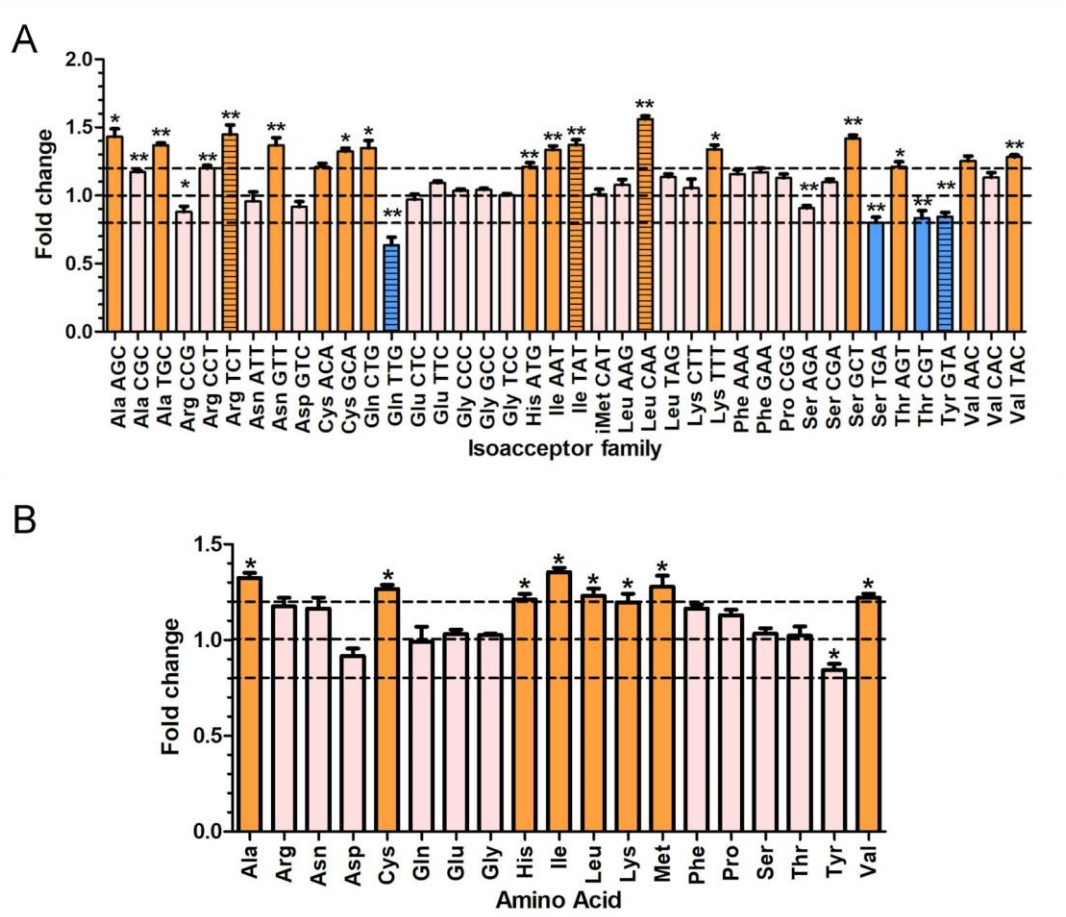


Figure 5.7. Overall tRNA expression changes in response to Polr3G overexpression. Expression fold change summary of tRNA isoacceptor families in *X.tropicalis* NF Stage 9 whole embryos injected with 2ng Polr3G mRNA versus control uninjected across 3 biological replicates. Signal data was processed using GeneSpring to determine normalised signal values for each probe within each array, and a normalised intensity reading by comparing probes across the arrays. Normalised values were used to determine expression fold changes for each biological replicate to give an average fold change for each probe across the three replicates. Control and Polr3G injected mean expressions for each of the 5 probes designed per family were then analysed using t-tests to determine statistical significance of the change in expression for each anticodon isoacceptor family and amino acid isotype. To account for multiple comparisons, Bonferroni Correction was applied. (A) Summary chart of fold changes for isoacceptor families with raw signals of >10. Orange indicates up-regulated transcription, blue indicates down-regulated transcription, and pink indicates no change in expression. SEM is presented as error bars and * indicates the level of statistical significance. (B) Summary chart of fold changes for all isoacceptor family probes per amino acid family. Orange indicates up-regulated transcription, blue indicates down-regulated transcription.

Particular subsets of amino acid isotype families show distinct regulation patterns between the isoacceptor families as a result of Polr3G overexpression. $tRNA^{Gln}$ has two isoacceptor families detected at Stage 9 and whilst Gln^{CTG} is significantly upregulated by overexpression of Polr3G, expression of Gln^{TTG} is decreased to 0.64 relative to control expression levels (Figure 5.8. A). A similar pattern of expression changes occurs in $tRNA^{Thr}$ families (Figure 5.8. C). $tRNA^{Ile}$ shows upregulated expression of their isoacceptors which is conserved through to the amino acid level (Figure 5.7. B, Figure 5.8. B). $tRNA^{Ser}$ isoacceptor families show variable regulation by Polr3G; Ser^{AGA} and Ser^{GCA} families show no change in expression, Ser^{GCT} is significantly upregulated and Ser^{TGA} is significantly down-regulated in embryos overexpressing Polr3G (Figure 5.8. D). Polr3G therefore specifically regulates a subset of tRNA isoacceptor families during early development and selectively activates or represses transcription. Despite significantly altered expression of $tRNA^{Gln}$, $tRNA^{Ser}$, and $tRNA^{Thr}$ isoacceptor families, at the amino acid family level, these changes equilibrate so that Glutamine, Serine and Threonine show no significant changes in level. This suggests that tRNA isoacceptor families equilibrate overall expression changes so that the relative levels of their amino acid remain constant, as suggested in a previous study (Schmitt et al., 2014).

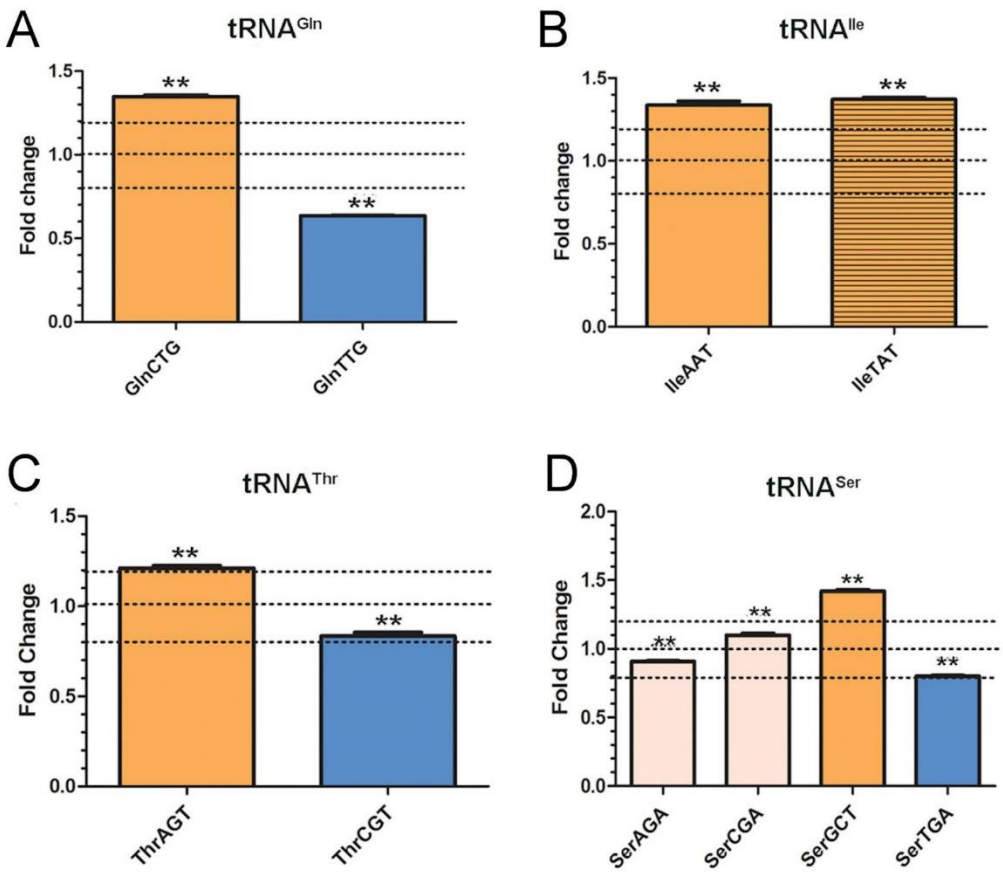


Figure 5.8. Analysis of tRNA expression changes in response to Polr3G overexpression by amino acid. (A) Isoacceptors encoding families belonging to $tRNA^{Gln}$ (B) $tRNA^{Ile}$ (C) $tRNA^{Met}$ (D) $tRNA^{Ser}$. Orange indicates expression is significantly up-regulated. Blue indicates expression is significantly down-regulated.

5.2.5. Validation of Microarray analysis

The results from microarray analysis show some discrepancies from the initial analysis carried out by RT-PCR in Figure 5.4. Therefore, to determine possible reasons for these differences, and to validate the use of microarrays as a tool for measuring tRNA expression, variation analysis and an additional qRT-PCR analysis for $tRNA^{iMet}$, $tRNA^{eMet}$, $tRNA^{Tyr}$, and $tRNA^{Leu}$ were carried out on NF Stage 9 control and Polr3G samples. Figure 5.9 shows that even within experimental sample groups, gene expression is largely variable. This may

account for the difference in expression change in the initial RT-PCR experiment versus later microarray experiments.

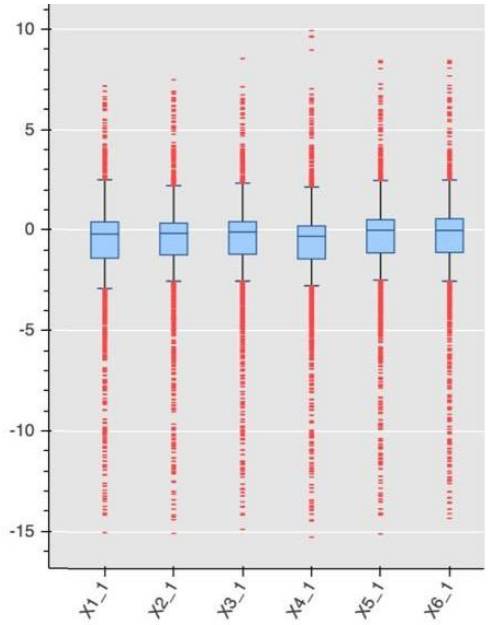


Figure 5.9. Variation of gene expression within NF Stage 9 microarray samples. X1-3 represent control samples. X4-6 represent samples overexpressing Polr3G. Samples are paired such that sibling embryos are present in samples X1 and X4, X2 and X5 and X3 and X6 respectively.

qRT-PCR analysis was carried out to quantify gene expression changes observed between additional control and Polr3G overexpression samples at NF Stage 9 in order to validate microarray results for the initially analysed tRNA families (Figure 5.10.). Some variation can be observed in $tRNA^{Leu}$ expression changes as in the microarray this was significantly upregulated whilst qRT-PCR analysis indicates in embryos overexpressing Polr3G, $tRNA^{Leu}$ expression is moderately (but not statistically significantly) downregulated. $tRNA^{Tyr}$ is downregulated in both microarray and qRT-PCR samples, but is not statistically different from controls in qRT-PCR results. $tRNA^{iMet}$ again shows no change in expression compared with control embryos in response to Polr3G overexpression and $tRNA^{emet}$ also shows a

similar response. tRNA^{eMet} was only detected at very low levels in the microarray samples and was included only as a comparison with tRNA^{iMet}. This data suggests that the initial RT-PCR results were possibly outliers in the analysis as qRT-PCR and microarray data show similar results.

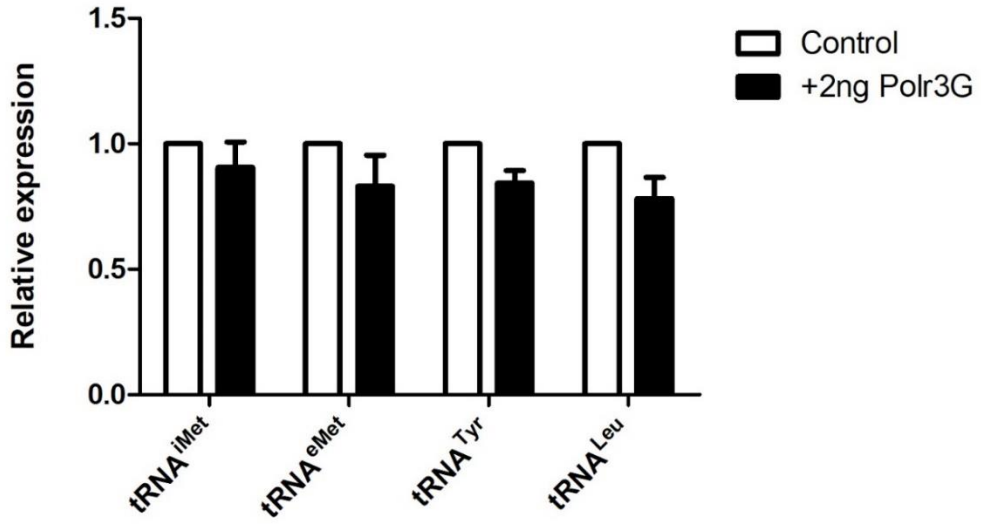


Figure 5.10. Validation of microarray analysis by qRT-PCR of selected tRNA families. Total RNA was extracted from control embryos and embryos overexpressing Polr3G protein at NF Stage 9 and rigorously DNase I treated using Zymo Clean and Concentrator-5 columns. qRT-PCR was carried out for tRNA^{iMet}, tRNA^{eMet}, tRNA^{Tyr} and tRNA^{Leu} using primers previously used in RT-PCR in Figure 5.2. Relative expression values were calculated using the ΔΔCt method with Dicer used as an endogenous control.

5.2.6. Analysis of RNA Polymerase III targets and mRNAs

To detect wider transcriptional changes as a response to Polr3G overexpression, probes for additional RNA polymerase III targets and selected mRNAs of interest were also analysed for changes in expression. Interestingly, RNA polymerase III targets somatic and oocyte 5S rRNAs were significantly downregulated (fold changes of 0.64 and 0.63 respectively) in samples overexpressing Polr3G compared with control embryos. 7SK RNA showed a fold change down to 0.81 expression levels and 7SL RNA also was downregulated to 0.8 expression levels. Whilst U6 RNA appeared to show some level of upregulation (fold change 1.15), no RNA Polymerase III target was shown to be upregulated above the threshold cut-off (Figure 5.11.A).

Several mRNAs selected for analyses were also shown to have significant changes in expression levels compared with controls. Many mRNAs showing significant changes beyond thresholds however (See Appendix. Table 2), were excluded due to low expression levels including all mRNAs down-regulated below threshold. mRNAs shown to be significantly upregulated include a number of key regulators of early embryonic patterning (Figure 5.11.B); the BMP antagonist and dorsalising factor chordin (chrd), the forkhead box transcriptional repressor *foxd5*; the bHLH transcriptional repressor id3; wnt8a, a ligand involved in ventral mesoderm patterning; the Spemann Organiser patterning gene (gsc); a regulator of mesoderm differentiation during gastrulation, eomesodermin (eomes); the mesendodermal patterning gene nodal (xnr4); and the left-right determination factor lefty.

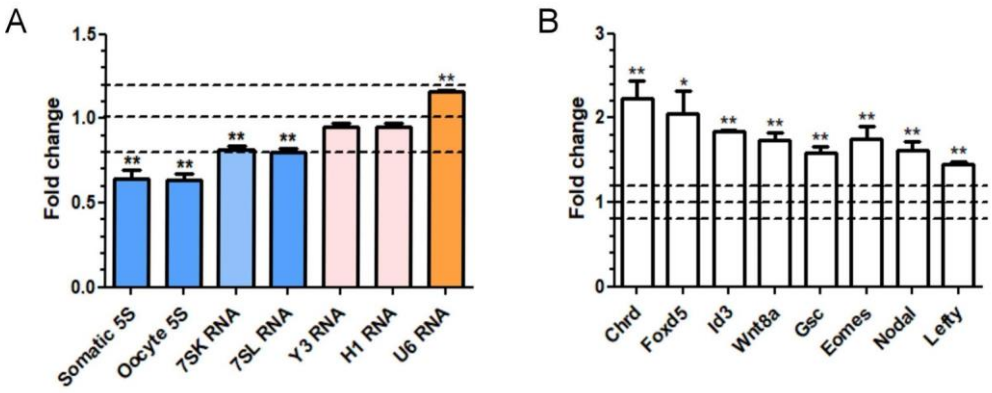


Figure 5.11. Wider transcriptional effects of Polr3G overexpression. Expression fold change summary of RNA Polymerase III target genes and upregulated mRNAs in *X.tropicalis* NF Stage 9 whole embryos injected with 2ng Polr3G mRNA versus control uninjected across 3 biological replicates. Normalised values were used to determine expression fold changes for each biological replicate and combined to give an average fold change for each individual probe across the three replicates. Control and Polr3G injected mean normalised expression for each of the 5 probes designed per gene were then analysed using t-tests to determine statistical significance after Bonferroni Correction of the change in expression. (A) Summary chart of fold changes for RNA Polymerase III target genes with raw signals of >10. Orange indicates upregulated transcription, blue indicates down-regulated transcription, pink indicates no change in expression. SEM is presented as error bars and * indicates the level of statistical significance. (B) Summary chart of fold changes for genes upregulated in Polr3G overexpression samples.

5.2.7. Correlating tRNA anticodon family expression levels to tRNA anticodon family gene copy number.

Previous studies in *C.elegans* have indicated that tRNA gene copy number (the number of genes encoded for within a genome) has a direct effect on the expression levels of particular tRNAs. The study showed that tRNA isoacceptor families with higher numbers of encoding genes were expressed at higher levels than tRNA isoacceptor families with lower gene copy numbers (Duret, 2000). With the microarray and GtRNAdb data, it is possible to test this hypothesis in Xenopus. Therefore, in order to determine whether gene copy number also has an effect in *Xenopus*, the correlation tRNA expression values from control NF Stage 9 samples and gene copy numbers predicted from the GtRNAdb database (http://gtrnadb.ucsc.edu/) was analysed using Pearson's correlation coefficient for linear relationships (Figure 5.12). The plot shows average tRNA isoacceptor expression values from the microarray against the corresponding gene copy number predicted for *Xenopus tropicalis*. A positive correlation is indicated by R^2 values close to the value of 1. No correlation between the factors was identified from the analysis (R^2= 0.0081, p = 0.52). This indicates that in *Xenopus*, gene copy number does not affect expression levels of tRNA isoacceptor families.

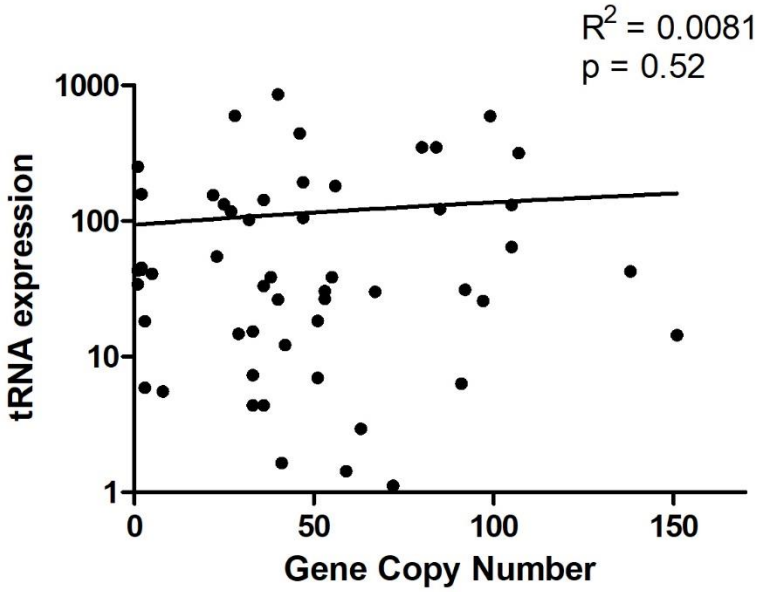


Figure 5.12. Linear correlation of tRNA expression levels to tRNA isoacceptor gene copy number. tRNA expression was calculated using average intensity values from NF Stage 9 control embryos. Linear correlation was determined using Pearson's correlation coefficient R^2= 0.0081. The x-axis values indicate tRNA anticodon gene copy number as predicted from http://gtrnadb.ucsc.edu, the y-axis values indicate log tRNA expression values.

5.3. Discussion

5.3.1. Upregulation of Polr3G results in global changes in tRNA expression

Polr3G overexpression at blastula stages, just after the onset of zygotic transcription, results in dynamic alterations in the expression of tRNA anticodon families, with many families showing significantly altered expression, both upregulated and downregulated. Interestingly, no previously identified tRNA families possibly regulated by Polr3G were shown to be significantly upregulated at NF Stage 9. Most notably, $tRNA^{iMet}$ showed no significant change in expression in *Xenopus* embryos in response to overexpression of Polr3G (Figure 5.7.), whilst in cell culture studies is significantly upregulated (Haurie et al., 2010). In cell studies, positive regulation of $tRNA^{iMet}$ by Polr3G also fits with the observations that both enhanced levels Polr3G expression, and upregulation of $tRNA^{iMet}$ expression are associated with increased proliferation and oncogenic transformation (Haurie et al., 2010; Pavon-Eternod et al., 2013). Therefore, in *Xenopus* early development, it is possible that Polr3G acts on different tRNA target genes to those in ESCs.

One problem of looking at only nascent transcripts is that the total tRNA pools can not be measured and so overall tRNA expression may be quite different to the changes measured by microarray and differences in tRNA expression may be more significant than detected. One way to measure mature tRNA transcripts is by northern blot. The denaturing gel opens the cloverleaf structure of mature tRNAs allowing hybridisation. However, this method is not as quantitative as microarray analysis. Recently, modifying enzymes have been used to remove methylation marks on tRNAs which prevent elongation by reverse transcriptase, allowing total tRNA expression to be measured by reverse-transcriptase techniques such as

qPCR and high-throughput sequencing (Wilusz, 2015; Zheng et al., 2015). Whilst this method proves promising for the study of tRNAs in future, as yet no standard protocol has been developed for this treatment and therefore the use of modifying enzymes would not have been feasible in the scope of this thesis.

5.3.2. A possible mechanism for the upregulation of RNA Polymerase II targets in embryos overexpressing Polr3G.

It is surprising that overexpression of Polr3G results in the selected upregulation of a few RNA Polymerase II regulated genes with roles in early development (Figure 5.11. B), as Polr3G is involved in the regulation of transcription by RNA Polymerase III. Moreover, these identified genes have not been previously identified in cell culture studies. Polr3G overexpression has been shown to upregulate RNA Polymerase II regulated factors involved in pluripotency, however this has only been shown to be indirect effect of Polr3G expression enhancing resistance to differentiation programmes in ESCs rather than direct regulation (Wong et al., 2011). Therefore, it is possible that Polr3G overexpression results in changes to RNA Polymerase II target genes through an indirect mechanism in *Xenopus*.

One potential mechanism for the upregulation of these early patterning genes is the timing of MBT. All of the genes upregulated in embryos overexpressing Polr3G are activated rapidly at Stage 9 just after MBT and accumulate to maximum expression levels by gastrula stages (Figure 5. 13). In *Xenopus,* MBT is preceded by rapid synchronous cell divisions known as cleavage. During these stages, embryos do not increase in size, meaning that cells become smaller throughout consecutive divisions (Newport and Kirschner, 1982a). Previous studies in

Xenopus have shown that MBT is timed by the titration of four limiting replication factors (RecQ4, Treslin, Cut5 and Drf1) through the rapid cell divisions of cleavage stages, and the increase in Nuclear:Cytoplasmic (N:C) ratio (Collart et al., 2013; Newport and Kirschner, 1982b). Therefore, it may be possible that MBT and activation of transcription could be brought forward by more rapid cell division, increasing the N:C ratio more rapidly.

Cell culture studies have indicated that Polr3G increases cell proliferation rates in cancer cell lines and increased tumour growth (Khattar et al., 2016). Furthermore, overexpression of Polr3G in ESCs results in the upregulation of positive cell cycle regulators such as CyclinE and Aurora Kinase A (Haurie et al., 2010). Therefore, it is possible that overexpression of Polr3G in *Xenopus* results in increased rates of cell division- this might result in premature activation of transcription of zygotic transcripts such as the patterning genes identified by the microarray as timing of MBT is brought forward.

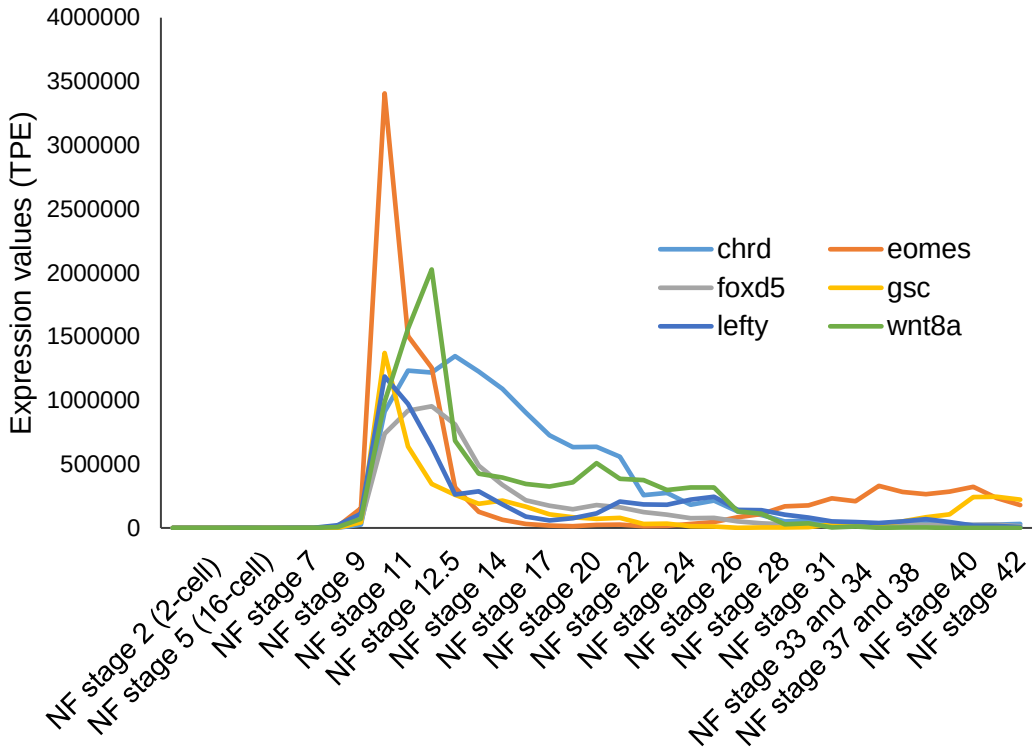


Figure 5.13. RNA-Seq expression values across *Xenopus tropicalis* development of early patterning genes identified as upregulated in embryos overexpressing Polr3G. Genes identified by microarray analysis are included with the exclusion of Id3 due to scaling differences (Owens et al., 2016). Transcripts per embryo are displayed as unit of measure.

5.3.3. Upregulation of Polr3G results in upregulation of key early developmental regulators

Overexpression of Polr3G at NF Stage 9 results in the upregulation of a number of key early developmental and patterning genes. In cell studies, upregulation of Polr3G results in ectopic expression of cell cycle regulators Aurora Kinase A and Cyclin E and maintained expression of pluripotency factors in response to triggered differentiation (Haurie et al., 2010; Wong et al., 2011). Aurora Kinase A is a cell cycle serine threonine kinase with implications in oncogenesis (Fu et al., 2007; Goldenson and Crispino, 2015), it is maternally expressed in *Xenopus* and is rapidly degraded post-MBT (Bischoff et al., 1998). Interestingly, in NF Stage 9

embryos overexpressing Polr3G, no upregulation of either Aurora Kinase A or Cyclin E was detected suggesting that Polr3G may have a different role during early development to stem cells.

5.3.3.1. Chordin

Chordin (Chrd) is a known BMP antagonist with roles in dorso-ventral patterning and neural induction in the early embryo (François and Bier, 1995). Chrd, in collaboration with noggin, act as potent dorsalising factors during early embryogenesis (Smith, 1995). Chrd is activated by organiser homeobox genes such as goosecoid (Gsc) (Sasai et al., 1994) and its overexpression in ventral cells results in induction of a second dorsal axis (Oelgeschläger et al., 2003). During neural induction, chordin inhibits the action of a member of the Bone Morphogenetic Protein family, BMP4 through direct interaction resulting in inhibition of ventral signals and dorsalised mesodermal fate (Piccolo et al., 1996; Sasai et al., 1995). Its upregulation in embryos overexpressing Polr3G may indicate that Polr3G has an early role in mesoderm induction in the early embryo. However, because Polr3G regulates tRNAs rather than mRNAs, this effect is likely to be indirect, possibly due to the mechanism described above.

5.3.3.2. Wnt8a

Wnt8 (wnt8a) is a secreted ligand that acts as a mesoderm patterning factor in *Xenopus*. Wnt8a is expressed ventrally in *Xenopus* gastrula mesoderm, excluded from the organiser region. Wnt8a is able to cooperate with the mesoderm inducing factor activin, and with mesodermal genes such as Xbra to promote a dorsal mesoderm (notochord and neural) fate (Cunliffe and Smith, 1994; Otte and Moon,

1992; Smith and Harland, 1991; Sokol and Melton, 1992). Co-expression of Wnt8a with both Xbra and Fgf4 enhances formation of muscle in ectodermal explant assays but is not sufficient alone to cause dorsal mesoderm differentiation (Christian et al., 1992; Sokol, 1993). There is also evidence that Wnt8a works with dorsalising signals from the Spemann Organiser acting on lateral mesoderm, resulting in a gradient response in dorso-ventral axis formation (Christian and Moon, 1993). Polr3G based upregulation supports the potential that RNA Polymerase III based transcription is able to indirectly regulate mesoderm induction during early development. Later expression of Polr3G in the somites may therefore be a downstream effect of early activity acting as a positive feedback loop.

5.3.3.3. Id3

Id3 acts as a transcriptional repressor of bHLH factor activity. Id proteins regulate the activation of networks by both the MRFs and NeuroD2 by prevention of dimerization of bHLH factors with their co-factors such as E proteins. Id3 acts downstream of Myc to promote cellular proliferation and survival in neural precursors (Kee and Bronner-Fraser, 2005; Light et al., 2005). Id proteins are also inhibitors of muscle differentiation and have been implicated in cancer (Jen et al., 1992; Lasorella et al., 2014). This may indicate a conserved role for Polr3G in maintaining a proliferative and undifferentiated cell state in *Xenopus* development and in cells.

5.3.3.4. Foxd5

Foxd5a (FoxD5) is a member of the forkhead box family of transcription factors and acts as a transcriptional repressor in neural fate specification (Fetka et al., 2000; Sölter et al., 1999). Foxd5 is expressed maternally and zygotically in the neural ectoderm and paraxial mesoderm at gastrula stages and regulates a transcriptional network to promote expression of neural Sox and Zic genes (Yan et al., 2009a). Foxd5 expression in animal caps results in promotion of both mesodermal and neural fate. In whole embryos, dorsal injection of Foxd5 results in an expanded neural plate whilst ventral expression of Foxd5 results in a partial duplicated axis without secondary neural tissue. Foxd5, via notch signalling, maintains an undifferentiated cell fate within the neural plate regions by repression of differentiation genes and upregulation of early proneural genes NeuroD2 and Xnrgn1 (Sullivan et al., 2001; Yan et al., 2009b). Foxd5 is also thought to regulate convergent extension during gastrula stages (Fetka et al., 2000). Upregulation of Foxd5 by Polr3G may infer a role for Polr3G in maintaining an undifferentiated cell state during early *Xenopus* development supporting the evidence from cell studies (Wong et al. 2011).

5.3.3.6. Eomes

Eomesodermin (Eomes) is another transcription factor with roles in early determination of mesoderm (Russ et al., 2000; Ryan et al., 1996). Part of the T-box family of transcription factors, Eomes expressed in response to Activin signalling (Ryan et al., 1996; Ryan et al., 2000) from MBT and is highest at gastrula stages before decreasing during neural stages of development. Expression of Eomes is localised to mesoderm at gastrula stages, is present in a

concentration gradient from dorsal-ventral regions and overlaps with factors such as Gsc but prior to Chrd and Wnt8. Moreover, expression of Eomes in in animal caps is sufficient to induce the expression of mesodermal genes Chrd, Xbra, Wnt8 and Mix1 and in D.rerio has a role in regulating formation of the Spemann Organiser (Bruce et al., 2003; Ryan et al., 1996). This supports the evidence from Chrd and Wnt8 that Polr3G may have an indirect role in early mesoderm determination through upregulation of key mesoderm inducing factors. However, interestingly in ESC culture studies Eomes expression was downregulated when Polr3G was overexpressed (Wong et al., 2011). This may therefore infer that Polr3G has different roles *in vivo* and *in vitro*.

5.3.4. Discrepancies between initial tRNA RT-PCR and Microarray analysis.

One potential problem with the data analysis is that the microarray samples appear to show different results to Polr3G overexpression at early stages of development compared with initial RT-PCR data carried out for tRNAs iMethionine (CAT), eMethionine (CAT), Leucine (CAA) and Tyrosine (GTA) (Figure 5.4.). In particular, $tRNA^{Tyr}$ was downregulated in response to Polr3G overexpression in microarray samples whereas in the PCR data it appears to be upregulated and promiscuously activated at NF Stage 8.

One possibility for this difference is that samples collected at NF Stage 9 have just undergone MBT and the activation of transcription of many mRNAs (Newport and Kirschner, 1982a; Newport and Kirschner, 1982b). Variation analysis revealed that samples collected for microarray show high levels of variability in expression such that expression values in one sample is less likely to representative of another.

Another possible reason for the differences seen is that embryos remain at NF Stage 9 for a long period of time. Whilst embryos are collected at roughly the same time post fertilisation each day of sample collections, embryos collected that look the same morphologically could be much further along in NF Stage 9 development. Therefore, samples collected for PCR analysis may be collected earlier in NF Stage 9 where transcriptional regulation is altered to downregulate an initial burst of tRNA transcription than microarray samples, which were all collected at the same time post-fertilisation when incubated at the exact same temperature.

A third possible reason for discrepancies is that there may be DNA contamination artefacts in PCR data. PCR methods involve the exponential amplification of DNA fragments, usually cDNA collected from RNA samples. mRNA PCR can be adapted to amplify RNA only by designing primers that either span exons with large introns between and using an extension temperature only suitable for the desired small product, or by designing primers across exon junctions again ensuring no intron sequence from DNA molecules is amplified. tRNAs are small in size ranging from just 70-100 base pairs, and their introns, included in just a small number of tRNA families, are very short usually consisting of around 10-15bp. This therefore means that PCR primers cannot be designed to exclude DNA. Moreover, in *Xenopus tropicalis* there are predicted to be over 2600 genes encoding tRNAs, therefore any DNA contamination is much more of a problem for tRNAs compared with mRNAs. Microarray analysis, whilst based on cDNA synthesis too, uses hybridisation as opposed to amplification to measure results. Therefore DNA contamination is much more likely to alter results of tRNA PCR analysis than microarray. Moreover, whilst initial PCR analyses were crried out after a single DNase treatment, microarray samples were treated twice with DNase, meaning

these samples are much less likely to contain significant amounts of DNA. Whilst no band is detected in the No-RT sample, this represents only NF Stage 7 RNA and does not exclude the possibility of contamination in later stage samples.

Finally, microarray analyses were carried out in triplicate with quantitative readouts as a measure. Thhis therefore makes the results from the microarray analyses more robust than the gel-based RT-PCR analysis as statistical analysis could be carried out to ensure significant differences between samples. Therefore, the results from the microarray analysis can be considered the more accurate readout of Polr3G regulatory activity, though RT-PCR analysis can be improved to measure tRNAs more accurately by carrying out qPCR.

5.4. Conclusions

Overall, this chapter has shown that, whilst mature tRNAs can only be detected through Northern Blot analysis, it is possible to detect tRNA expression dynamics across development and between control and experimental samples by cDNA based RT-PCR and microarray based techniques. This chapter has shown that during early stages of development, Polr3G overexpression results in diverse changes to tRNA isoacceptor families, and possible activation of tRNA transcription prior to MBT. Polr3G overexpression may also resulted in the upregulation of crucial early transcriptional pathways involved in embryonic patterning. These results therefore suggest that Polr3G may act as an early developmental regulator in *Xenopus* and that dysregulated RNA Polymerase III activity has wider transcriptional effect *in vivo*.

6. Characterising a myogenic tRNA profile in *Xenopus*

6.1. Introduction

6.1.1. Putative roles of Polr3G and Polr3gL *in vivo.*

The incorporation of Polr3G or Polr3gL into the RNA Polymerase III complex forms two distinct isoforms of the enzyme. Polr3G and Polr3gL are thought to have diverged as the result of a gene duplication event with Polr3gL appearing to be the ancestral gene copy (Renaud et al., 2014). As these subunits were revealed by ChIP-Seq to have largely overlapping binding patterns *in vitro* but showed different regulation by the transcription factor c-Myc, it was determined that the isoforms adopted distinct functions through their interaction with regulators, rather than their capacity to activate specific subsets of target genes (Renaud et al., 2014). In cell lines, as discussed previously, Polr3G appears to associate more with proliferative capacity and stemness, whilst Polr3gL expression is maintained throughout induced differentiation leading to the hypothesis that this subunit has roles in later developmental commitment (Haurie et al., 2010; Wong et al., 2011). Chapter 4 indicates that, as in cell studies Polr3G is downregulated during development and essentially replaced by Polr3gL expression (Figure 4.5). However transient expression of Polr3G is localised to the somites during early myogenesis and Polr3G is upregulated by the early MRF, MyoD (Figure 4.11). Therefore, it is possible that Polr3G and Polr3gL have different roles during early embryonic development through modified interaction with key transcriptional regulators and that Polr3G has a cooperative role with MRFs such as MyoD in the specification of early muscle lineages.

6.1.2. Early muscle lineages in *Xenopus*.

In *Xenopus,* the early myogenic lineages are specified from mesoderm inducing signals in the dorsal marginal zone of blastula and gastrula stage embryos and the first expression of an MRF is detectable at gastrula stages. MyoD is expressed around the blastopore (excluding dorsal-most regions which will form the notochord) at gastrula stages and is localised to presomitic mesoderm during neurula stage development. Myf5, also induced by FGF signalling, shows more limited expression even within gastrula stages where it is expressed in the dorso-lateral marginal zone of the blastopore (Hopwood et al., 1991; Maguire et al., 2012). By late neurula and tailbud stages, expression of MyoD and Mrf4 are localised to segmenting somites along the dorsal axis of the embryo, (Della Gaspera et al., 2012; Hopwood et al., 1989). Somites contain precursor cells for all muscle lineages in the trunk and arise as paired epithelial condensations of paraxial mesoderm located either side of the neural tube, and arise from anterior to posterior positions (Emerson, 1993; Pownall et al., 2002). Proliferative myoblasts are localised to the somites, their fate maintained by axial signals and the eFGF-mediated community effect, and migrate out to the trunk of tadpoles during differentiation (Emerson, 1993; Standley et al., 2001). Therefore, during *Xenopus* development, muscle lineage specification and differentiation is localised to the dorsal axis of the embryo, marked by the expression of MRFs.

6.1.3. Aims of this chapter:

1. To determine the effects of Polr3G or Polr3gL overexpression on the transcription of tRNAs in dorsal regions of *Xenopus* embryos.
2. To identify any unique transcriptional targets of Polr3G and Polr3gL in the context of muscle differentiation.

6.2 Results

6.2.1. Regulation of expression of tRNAs in the dorsal region by Polr3G.

To determine the effect of overexpression of Polr3G on the regulation of tRNAs within the muscle lineage, embryos were injected at the 1 and 2-cell stage with equal amounts of synthetic mRNA coding for either HA-tagged Polr3G or Polr3gL. Western blot analysis on Stage 25 embryos was used to measure protein expression as a result of injections by HA detection (Figure 6.1. A) and indicated that Polr3G and Polr3gL were expressed at equal amounts in all three biological replicates for the microarray. To ensure that the dorsal regions collected (Figure 6.1. B) were correctly enriched for both Polr3G and MyoD as predicted by in situ hybridisation analysis, qRT-PCR analysis was carried out for expression of Polr3G, Polr3gL and MyoD in collected head, ventral and dorsal sections from microarray sibling embryo samples (Figure 6.1. C). Both Polr3G and MyoD show significantly upregulated expression in dorsal sections compared with head and ventral sections. Polr3gL expression in dorsal regions is significantly higher than ventral sections but no significant difference from expression in head sections. This high expression in dorsal regions reflects the expression of Polr3gL in the

lateral plate mesoderm detected in the in situ hybridisation analysis (Figure 4.7. C-D). Polr3G expression is low post-MBT but is transiently enriched in the somites at tailbud stages, so overexpression analysis is used here to reveal transcript specific effects.

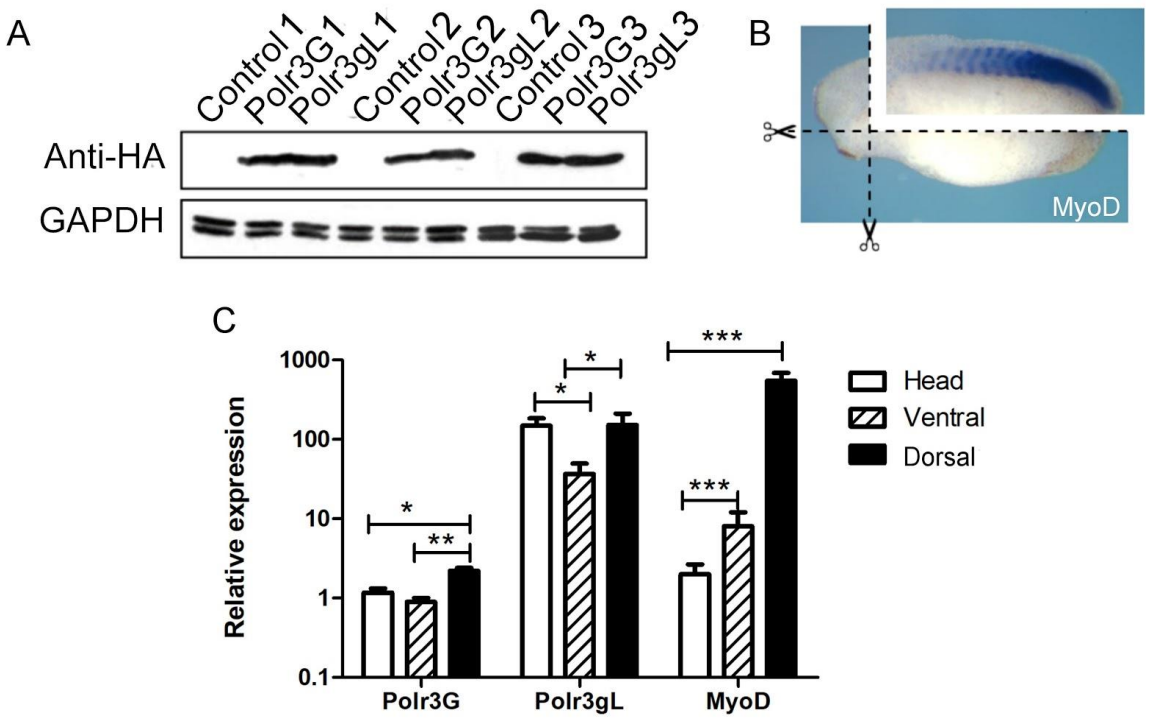


Figure 6.1. NF Stage 25 overexpression of Polr3G and Polr3gL in *Xenopus tropicalis* embryos. (A) 2ng of HA-tagged Polr3G and Polr3gL mRNA was injected into *Xenopus tropicalis* embryos at the 1 to 2-cell stage of development and embryos were allowed to develop until NF Stage 25. Western blot analysis of 10 embryos alongside 10 control embryos was carried out for detection of HA-tag expression. A GAPDH loading control was also used. (B) Collection of dorsal regions was carried out to enrich for myogenic tissue at this stage of development. (C) qRT-PCR analysis was carried out for control sections for three biological to determine enrichment of MyoD and Polr3G as predicted by in situ hybridisation. Ct values were normalised to Dicer and expression values were calculated relative to Polr3G expression levels in ventral sections. One-way ANOVA statistical analysis of expression values was carried out for each gene between sections. (*) indicates level of significance * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Note that scale on Y axis is logarithmic.

The same microarray approach that was used in Chapter 5 was used in this chapter to investigate tRNA expression changes in response to overexpression of either subunit. Total DNase treated mRNA was extracted from dorsal sections of control and experimental sibling embryos at the desired NF stage and labelled cRNA was synthesised for microarray hybridisation. Samples were collected such that control, Polr3G overexpression and Polr3gL overexpression samples were all collected from sibling embryos for 3 biological replicates. Computational analysis of the data was carried out as previously, with mean fold change versus control samples summarised in Figure 6.2. In comparison with NF Stage 9 data, relatively fewer tRNA isoacceptor families show significantly altered expression in response to overexpression of Polr3G. Only two isoacceptor families are significantly upregulated beyond the threshold of a 1.2 fold change. 6 families show significantly decreased expression and there appears to be fewer amino acid families showing different isoacceptor expression patterns. This may indicate a specific subset of tRNA families regulated within the muscle lineage by Polr3G.

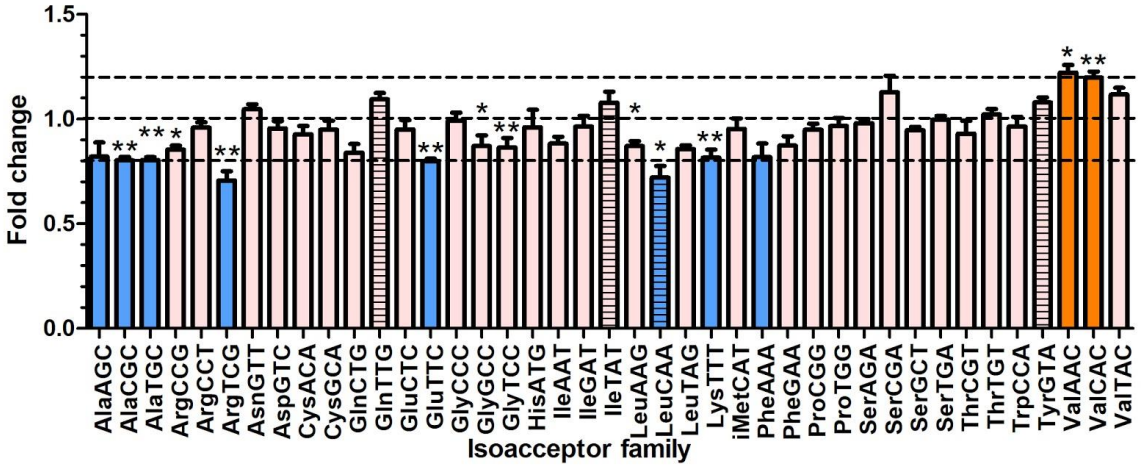


Figure 6.2. Overexpression of Polr3G in NF Stage 25 dorsal regions results in limited changes to tRNA gene expression. Expression fold change summary of tRNA isoacceptor families in *X.tropicalis* NF Stage 25 dorsal sections injected with 2ng Polr3G mRNA versus control uninjected across 3 biological replicates. Isoacceptor families with raw signals of >10 were included in the analysis. Orange indicates up-regulated transcription, blue indicates down-regulated transcription, and pink indicates no change in expression. SEM is presented as error bars and * indicates the level of statistical significance between control and injected samples after Bonferroni Correction.

Some amino acids do show differential regulation of their isoacceptor families in dorsal regions as summarised in Figure 6.2. $tRNA^{Ala}$ isoacceptor families are all downregulated in response to overexpression of Polr3G in dorsal regions (Figure 6.3. A, right-hand panel), contrasting with upregulation observed in NF Stage 9 embryos overexpressing the subunit (Figure 6.3. A, left-hand panel). Similarly, $tRNA^{Arg}$ and $tRNA^{Leu}$ isoacceptors show altered expression changes in response to Polr3G overexpression at the two different stages. Arg^{CCG} and Arg^{CCT} expression does not change in response to overexpression of Polr3G at tailbud stages (Figure 6.3. C, right-hand panel), but Arg^{CCT} is significantly upregulated in Stage 9 samples (Figure 6.3. B, left-hand panel). Interestingly, Arg^{TCG} detected and upregulated at Stage 9 in Polr3G overexpression samples, is not detected at

stage 25. Instead Arg^{TCG} is detectable and shows downregulated expression in Polr3G samples. Leu^{CAA} is also shown to be upregulated in Polr3G overexpression samples at NF Stage 9 but downregulated in NF Stage 25 samples (Figure 6.3. C). These results indicate that Polr3G may have different roles in regulating tRNA expression throughout development.

Contrastingly, $tRNA^{Val}$ isoacceptors appear to show very similar response to Polr3G overexpression at both stages analysed (Figure 6.3. D). Therefore, Polr3G may have a maintained role in the modulation of some tRNA isoacceptor families throughout embryonic development and tissue differentiation.

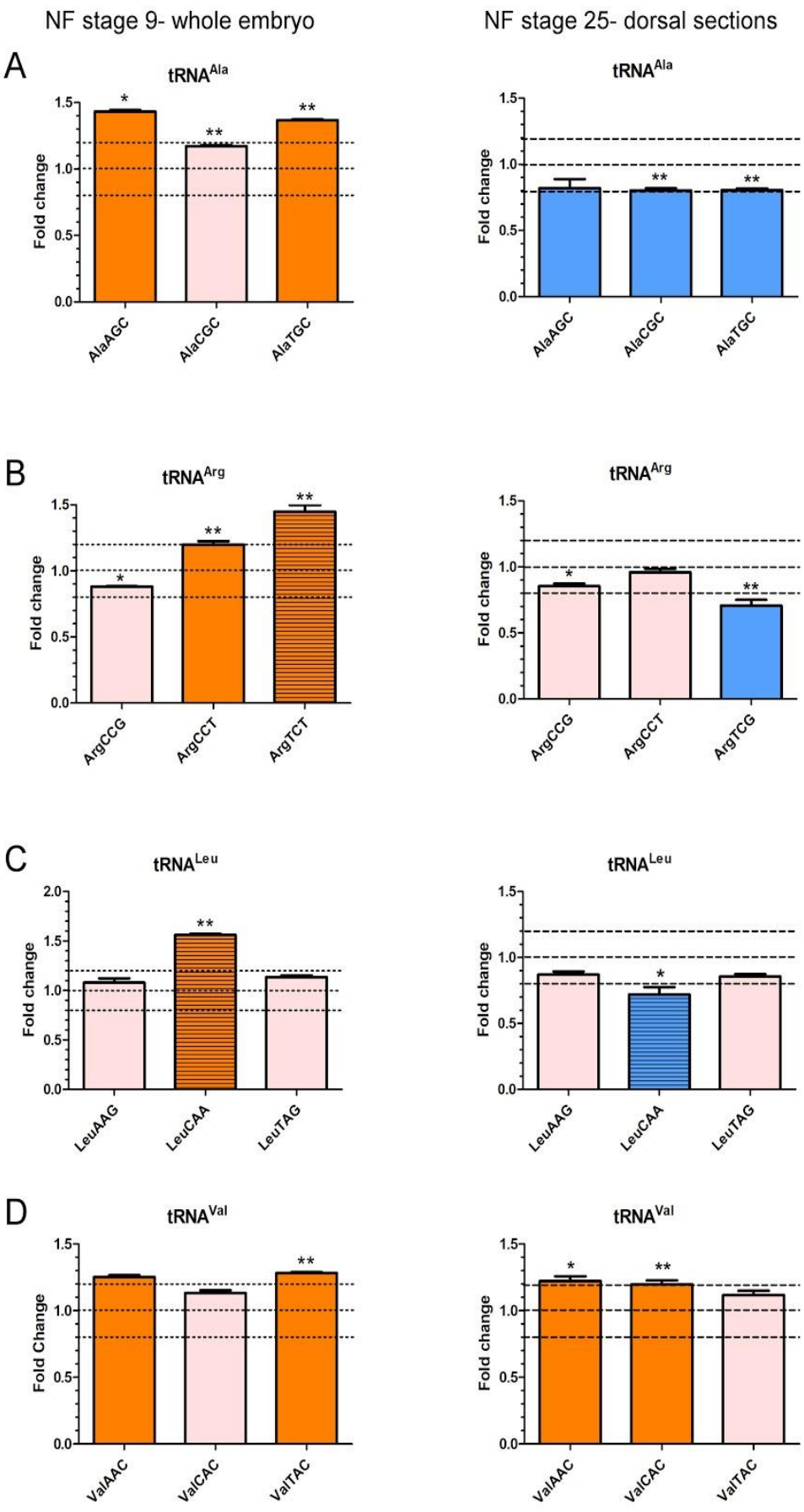

NF stage 9- whole embryo
NF stage 25- dorsal sections
A
tRNA^Ala
Fold change
AlaAGC
AlaCGC
AlaTGC
B
tRNA^Arg
ArgCCG
ArgCCT
ArgTCT
ArgTCG
C
tRNA^Leu
LeuAAG
LeuCAA
LeuTAG
D
tRNA^Val
Fold Change
ValAAC
ValCAC
ValTAC

Figure 6.3. Highlighting tRNA isoacceptor family expression changes between NF Stage 9 embryos overexpressing Polr3G and NF Stage 25 dorsal sections overexpressing Polr3G. Relative fold change of amino acids by isoacceptor family in Polr3G injected samples versus controls. Left-hand panels show fold changes in NF Stage 9 samples, right-hand panels show fold changes in NF Stage 25 samples. (A-D) Expressed isoacceptors encoding (A) $tRNA^{Ala}$ (B) $tRNA^{Arg}$ (C) $tRNA^{Leu}$ and (D) $tRNA^{Val}$. Orange indicates expression is significantly up-regulated. Blue indicates expression is significantly down-regulated. SEM is represented as error bars on the graphs and (*) indicates the level of statistical significance after Bonferroni Correction.

6.2.2. Regulation of expression of tRNAs in the dorsal region by Polr3gL.

To determine whether Polr3G and Polr3gL show different transcriptional activities within the early myogenic lineages, sibling samples from embryos overexpressing Polr3gL were also collected and analysed by microarray. Figure 6.4.A shows a summary for all isoacceptor families with expression values of above 10. Polr3gL overexpression, like Polr3G also results transcriptional changes to limited isoacceptor families in NF Stage 25 dorsal regions. However, Polr3gL appears to regulate similar families to Polr3G. All $tRNA^{Ala}$ families, Arg^{TCG} and Leu^{CAA} all show significant downregulation in both microarrays (Figure 6.3, Figure 6.4. A). However, Pol3gL appears to downregulate a number of other families not significantly decreased in Polr3G samples for example Ile^{AAT}, Leu^{AAG} and Ser^{GCT}. Moreover, whilst overexpression of Polr3gL results in significant upregulation of Tyr^{GTA}, and no significant changes to $tRNA^{Val}$ families, Polr3G overexpression results in no change in Tyr^{GTA} expression but significant upregulation of Val^{AAC} and Val^{CAC} at this stage. This indicates that Polr3G and Polr3gL may have distinct transcriptional activities within dorsal lineages.

To determine if Polr3G and Polr3gL differentially regulate tRNAs in dorsal sections, a comparison of normalised expression of tRNA isoacceptors in samples expressing each factor was carried out. Figure 6.4.B shows a scatterplot of expression data for each tRNA isoacceptor with detectable expression in both samples. The X-axis shows normalised expression in Polr3G samples and the Y-axis shows normalised expression in Polr3gL samples. Points that fall outside of the dashed lines indicate that a tRNA is expressed by less than half in Polr3G or Polr3gL samples. In comparison to the expression levels in sibling embryos overexpressing Polr3G, tRNA gene expression is largely reduced in Polr3gL samples (Figure 6.4. B). The isoacceptors showing higher expression in embryos overexpressing Polr3gL (Figure 6.4. B, red circle) are Asp^{GTC}, Tyr^{GTA} and Ser^{CGA}, however, most data points show greater expression in Polr3G expressing samples. This indicates that at NF Stage 25, in dorsal regions, Polr3G expression activates or maintains higher levels of tRNA expression moreso than Polr3gL, but the two subunits may have different target genes.

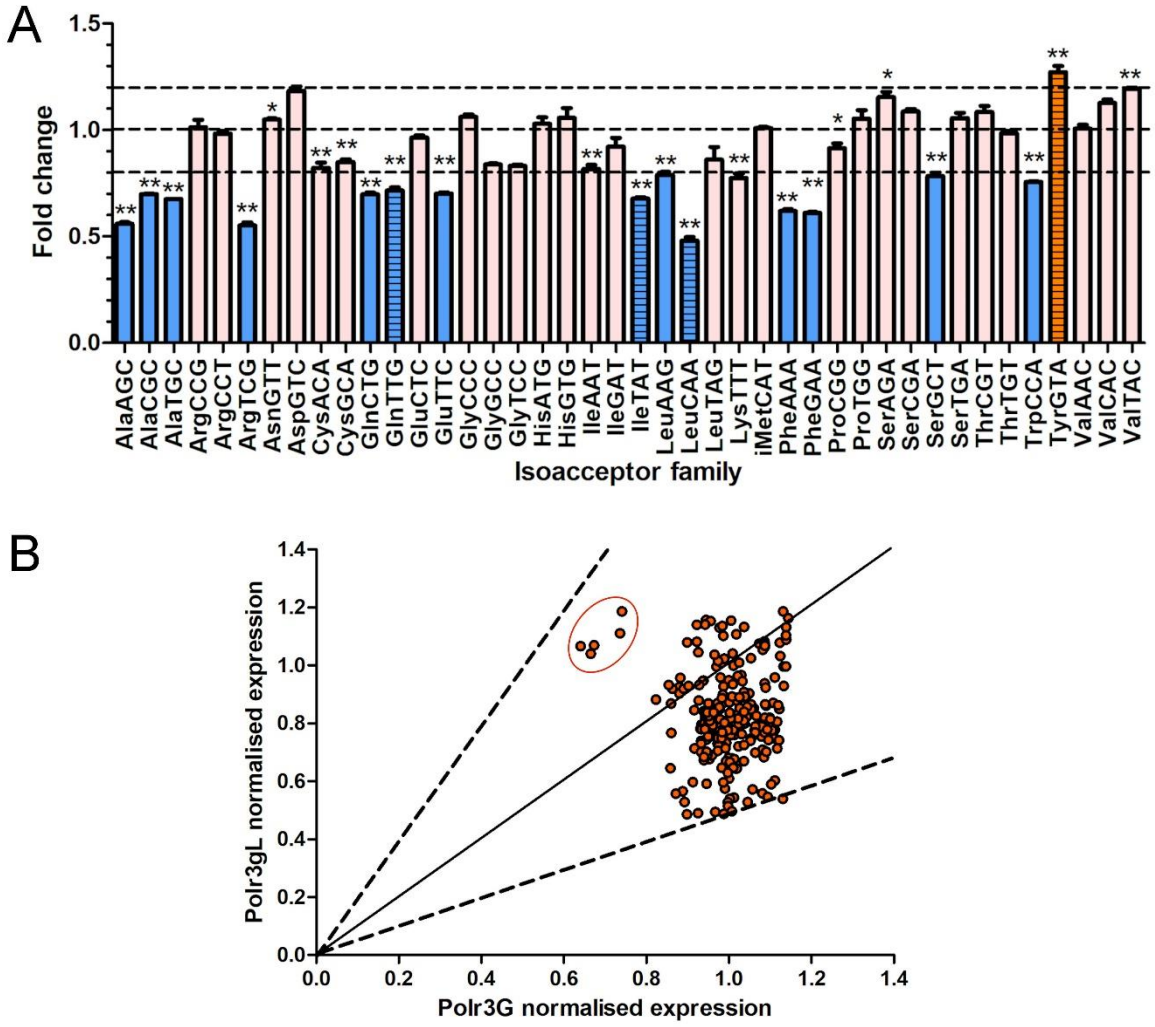


Figure 6.4. Overexpression of Polr3gL in NF Stage 25 dorsal regions results in greater changes to tRNA gene expression than Polr3G. Expression fold change summary of tRNA isoacceptor families in *X.tropicalis* NF Stage 25 dorsal sections injected with 2ng Polr3gL mRNA versus control uninjected across 3 biological replicates. Isoacceptor families with raw signals of >10 were included in the analysis. Orange indicates up-regulated transcription, blue indicates down-regulated transcription, and pink indicates no change in expression. SEM is presented as error bars and * indicates the level of statistical significance between control and injected samples after Bonferroni Correction. (B) Scatterplot comparing normalised expression of tRNA isoacceptor families in embryos overexpressing Polr3G with Polr3gL. Solid line indicates no expression difference. Dashed lines indicate expression fold change of 2. Polr3G values are on the X-axis and Polr3gL values are on the Y-axis.

6.2.3. Validation of microarray results by qRT-PCR analysis.

To confirm the results from the microarray, selected tRNA isoacceptor families were analysed by qRT-PCR in control and sibling matched embryos overexpressing either Polr3G or Polr3gL. All $tRNA^{Ala}$ families were selected due to their conserved downregulation in both Polr3G and Polr3gL overexpression samples. In contrast, Ser^{GCT} was selected as it showed greater downregulation in samples overexpressing Polr3gL than Polr3G in the microarray analysis. All $tRNA^{Ala}$ families show downregulated expression in Polr3gL samples (Figure 6.5.) In contrast to the microarray data, Ala^{AGC} shows no downregulation in response to Polr3G overexpression but is significantly downregulated in Polr3gL. Ser^{GCT} is moderately downregulated in Polr3G samples, but is downregulated to a higher extent in response to Polr3gL overexpression. This data supports the findings from the microarray that overexpression of Polr3gL has a stronger downregulating effect on similar transcriptional targets as Polr3G.

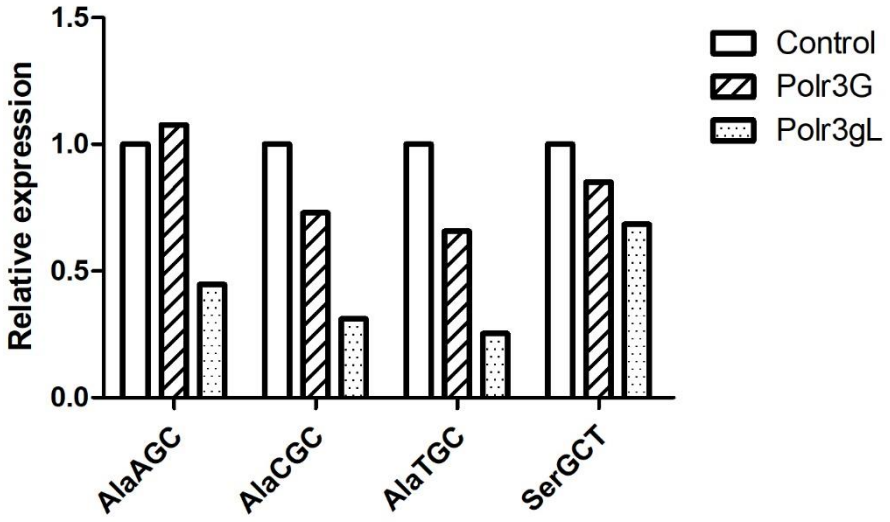


Figure 6.5. Validation of microarray data by qRT-PCR analysis of selected tRNA isoacceptor families.

Total RNA was extracted from control dorsal sections and sections overexpressing Polr3G or Polr3gL protein at NF Stage 25 and was rigorously DNase I treated using Zymo Clean and Concentrator-5 columns. qRT-PCR was carried out for tRNAs Ala and Ser, Ala^{AGC}, Ala^{CGC} and Ala^{TGC} and Ser^{GCT} using primers designed to amplify the full tRNA sequence (between 70 and 100bp). Relative expression values were calculated using the ΔΔCt method with Dicer used as an endogenous control.

6.2.4. Correlation of tRNA abundance with mRNA codon usage during muscle differentiation *(in collaboration with the Technology Facility at the University of York)*.

Previous studies in humans have suggested that tRNA expression may be correlated with mRNA codon usage for specific tissues (Dittmar et al., 2006; Gingold et al., 2014). In order to assess this relationship in *Xenopus*, Codon Usage Estimate values were calculated by multiplying usage numbers for particular codons included in coding sequences of all mRNAs included on the arrays by the gene expression value of each mRNA (all three biological replicates included so replicate data points were expected to be located in close proximity to one-another). tRNA anticodon isoacceptor intensity values from the array were used as a measure of expression. Spearman's correlation analysis was carried out on control samples used in the Stage 9 and Stage 25 microarrays for all mRNAs included on the array and then for a subset of mRNAs involved in muscle differentiation, not assuming a linear relationship.

The scatterplots produced are shown in Figure 6.6 and indicate a weak correlation between tRNA expression and predicted codon usage at both stages , the correlation in NF Stage 25 dorsal samples is increased to 0.349 (Figure 6.6. B) from 0.175 at NF Stage 9 (Figure 6.6. A). This may indicate that throughout development, tRNA transcription is regulated increasingly to match mRNA codon usage. Very similar correlations are observed even if mRNA codon usage is calculated only for mRNAs encoding muscle specific (contractile protein) genes (Figure 6.6. C and D). Green data points to the far right of plots B and D, representing high tRNA expression values and high mRNA codon usage. This points were identified as tRNAs Glu^{CTC} and Glu^{TTC} (Note that these points also fall

to the right hand side of plots A and C alongside other isoacceptors) and red data points represent Leu^{CAG}. A single data point is also located at the top left of plots B and D, representing high tRNA expression for low mRNA codon usage. This plot was identified as Thr^{CGT}, however, only one biological replicate of the three included in the analysis was represented, therefore this is likely to be an anomalous result.

In NF Stage 9 plots (Figure 6.6. A and C), grey data points representing low tRNA expression values but high mRNA codon counts were identified as Lys^{CTT}. Lys^{TTT} data points are located at similar mRNA codon counts, but higher tRNA expression (Figure 6.6. C, red circle). This result may indicate that Lys^{TTT} expression is favoured over Lys^{CTT} at these stages as it enables more efficient translation.

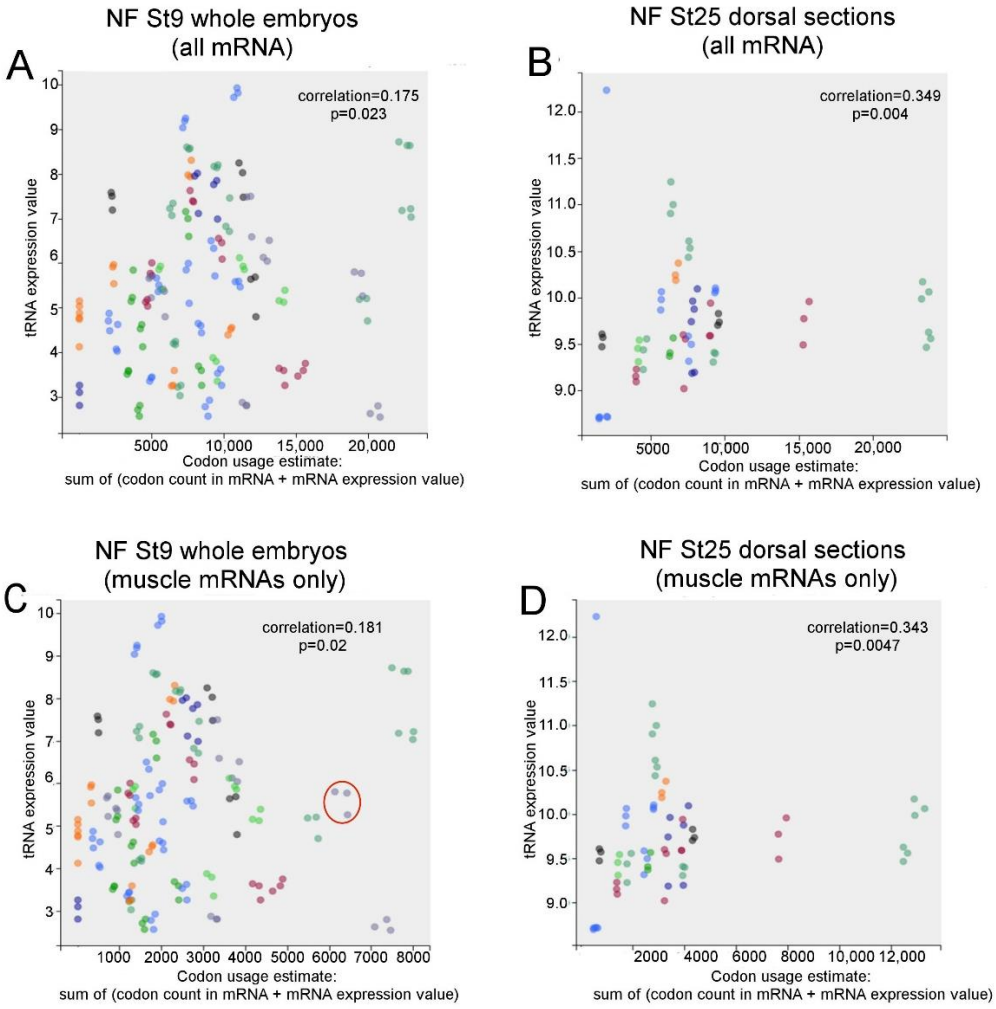


Figure 6.6. Correlation between mRNA codon usage and tRNA expression. (A-B) correlation plots for tRNA expression versus codon usage using Spearman correlation coefficient for control embryos at NF Stage 9 (A) and dorsal sections at Stage 25 (B) including all mRNAs on the microarray design. Each data point indicates a tRNA isoacceptor family expressed above threshold levels in the microarray. Isoacceptor families encoding the same amino acid are coloured the same. Data points were included for all three biological replicates rather than an average of the three in order to identify outliers in one of the sets. (C-D) correlation plots for samples included in charts A and B but including only mRNAs for contractile protein genes.

6.2.5. Targeting Polr3G expression in *Xenopus tropicalis* by Antisense Morpholino Oligo (AMO) knock-down.

Gene targeting or knock-downs are a crucial method to determine a gene's in vivo function. There is a high level of maternally stored Polr3G transcript, therefore in order to determine the function of Polr3G during development and myogenesis in more detail, post-transcriptional targeting by antisense morpholino knockdown appeared the most direct way-forward. CRISPR/Cas9 methods would not deplete the maternal store of Polr3G mRNAs. Antisense morpholinos can be effective to inhibit translation of mRNAs or to block correct splicing of a newly transcribed mRNA. Antisense Morpholinos (AMOs) targeting translation initiation through binding and blocking of the ATG start site were designed using the coding sequence and upstream sequences collected from EST data (Figure 6.7 A). 10ng of each AMO were injected alongside 10ng of a 5 base-pair mismatch control morpholino to ensure that any effects observed were not artefacts of AMO injection alone. Embryos were left to develop until NF Stage 25, the stage at which Polr3G mRNA expression is localised to the somites, and the efficiency of the translation blocking AMOs was assessed by western blot (Figure 6.7. B). Western blot analysis showed no decrease in protein levels of Polr3G at Stage 25 compared with control embryos.

An alternative AMO disrupting correct mRNA splicing of Polr3G was also designed to affect splicing of exons 2 and 3 within the coding sequence (Figure 6.7. C). The morpholino target site included the 3' end of exon 2 and the intron 3' to it. The

alternative splicing product would include an additional 748bp intron sequence between exons 2 and 3 and a premature stop codon 293bp downstream of exon 2 (*) resulting in a truncated protein missing exons 3-7 of coding sequence. To determine whether mRNA splicing can be targeted by this morpholino, primers for RT-PCR were designed to amplify a control region within exon 2 and the alternatively spliced region using primers spanning exon 2 and exon 3. RT-PCR analysis was carried out on embryos injected with either control morpholino, or with increasing concentrations of splice-blocking morpholino (Figure 6.7. D). If splice-blocking had occurred, an additional PCR product of 847bp would be present in reactions using exon 2:3 spanning primers (Figure 6.7. D bottom panel, top arrow). No additional product was observed in any of the morpholino targeted embryo samples, therefore indicating that the splice-blocking morpholino is unable to target Polr3G mRNA in *Xenopus tropicalis* embryos.

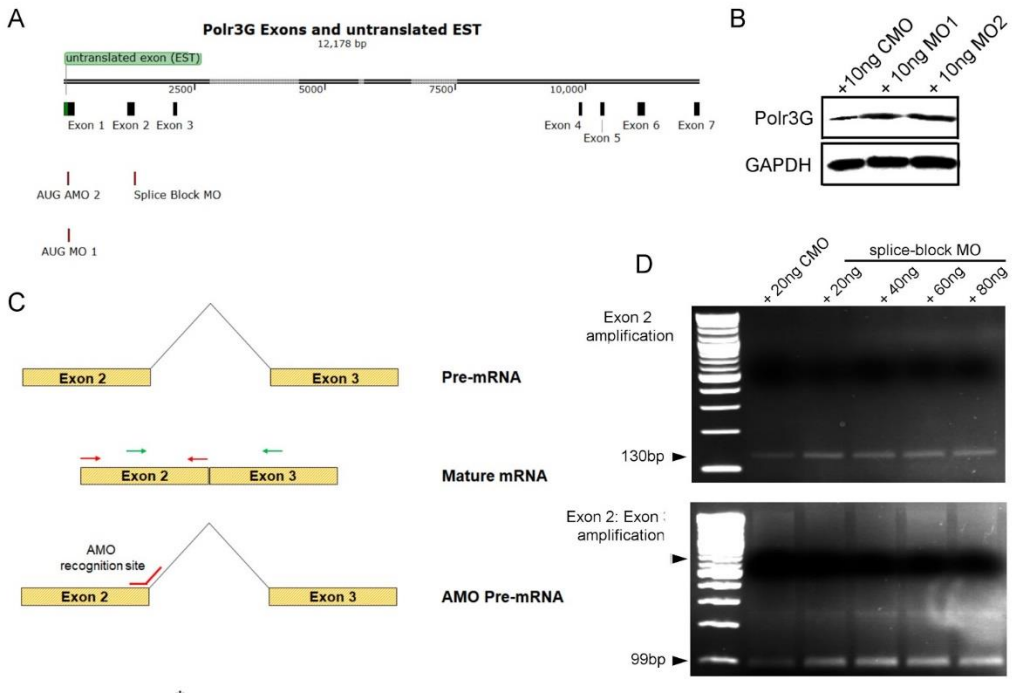


Figure 6.7. Antisense Morpholino (AMO) targeting of Polr3G in X*enopus tropicalis*. (A) Schematic diagram of Polr3G genomic sequence and coding exons. Green marks 5' untranslated exon identified through EST analysis, Antisense Morpholino Oligo (AMO) target sequences are also identified in red. (B) Western blot for NF St25 embryos injected with 10ng Control Morpholino (CMO) or 10ng of translation blocking morpholinos. (C) Schematic diagram showing splice blocking morpholino targeting and primers designed for validation PCR. (D) RT-PCR analysis of Stage 25 embryos injected with CMO of a series of concentrations (20ng, 40ng, 60ng, 80ng) of splice-blocking MO. arrows mark the predicted control products. Lower panel upper arrow indicated predicted alternative splice product arrow.

6.3. Discussion

6.3.1. Known roles of Polr3G and Polr3gL

Polr3G overexpression has different results on tRNA expression at NF Stage 9 and NF Stage 25. Targets such as $tRNA^{Ala}$ and $tRNA^{Arg}$ families upregulated at NF Stage 9 were shown to have reduced expression in response to Polr3G overexpression at NF Stage 25 (Figure 6.3. A and B). Moreover, targets with no change in expression at NF Stage 9 also had downregulated expression in NF Stage 25 samples. These results indicate the potential that Polr3G has different transcriptional activity throughout development and regulates different target genes at different stages and in different embryonic regions. tRNA expression in embryos overexpressing Polr3G are higher than embryos overexpressing Polr3gL. This may be due to the association of Polr3G expression with pluripotency and cell proliferation (Haurie et al., 2010). RNA Polymerase III targets are downregulated during differentiation programmes (White et al., 1989), and the downregulated expression of tRNAs in embryos overexpressing Polr3gL may be a result of increased RNA Polymerase III complexes incorporating Polr3gL to regulate transcription in a differentiated manner, whilst Polr3G maintains a more proliferative and pluripotent cell state with higher levels of RNA Polymerase III transcription.

Changes in transcriptional activity can be a result of altered interactions with binding partners or co-factors or by different signalling environments. Polr3G is known to interact with another subunit of RNA Polymerase III, RPC62, which is involved in activation of transcription of target genes (Boissier et al., 2015; Wang and Roeder, 1997). It has not been established whether Polr3gL also interacts

with this subunit and therefore it is possible, due to low sequence identity with Polr3G, that Polr3gL interacts with other subunits of RNA Polymerase III to change overall binding activity. Moreover, as the negative regulator Maf1 binds to this sub-complex in order to repress transcription by RNA Polymerase III (Vannini et al., 2010), the differences in sequence may be sufficient to alter its regulation by Maf1 *in vivo*.

RNA Polymerase III is also regulated by a number of developmental signalling pathways (Erk, TOR) and cellular regulators (p53, Rb, Myc, Maf1) which alters its activity in vitro (Felton-Edkins et al., 2003a; Felton-Edkins et al., 2003b; Marshall et al., 2012). As signalling environments change during development with the activation of Fgf, Wnt, Bmp pathways as specification of germ layers and tissue differentiation occurs, it is likely that RNA Polymerase III activity will also change. This may therefore explain the change in transcriptional regulation observed between NF Stage 9 and NF Stage 25 samples across the microarrays. It does however, appear that changes in expression of particular tRNA families results in a weak correlative relationship with mRNA codon usage predicted at NF Stage 25 (Figure 6.6. produced by Katherine Newling). The correlation between mRNA codon usage and tRNA expression has been previously described in human tissue and cell culture studies and is suggested to be due to altered translational pressure set between different cell types. Therefore RNA Polymerase III activity may be correlated with a drive fromtranslational pressure set by RNA Polymerase II as shown in previous studies (Dittmar et al., 2006; Gingold et al., 2014; Hentzen et al., 1981).

6.3.2. Gene-targeting of Polr3G

Strategies to knockdown Polr3G using morpholinos were unsuccessful. A large maternal store of both mRNA and protein has prevented this important functional analysis of Polr3G *in vivo*. The only way to deplete maternal stores would be by CRISPR/Cas9 targeting of Polr3G and outcrossing of mosaic F0 individuals to produce populations of mutant individuals. However, this also has limitations which made the method unavailable within the scope of this thesis- the first being that *Xenopus tropicalis* require 6 months to reach sexual maturity and would require multiple rounds of offspring generation in order to produce heterozygous and homozygous mutant populations. The second limitation to the genetic targeting of Polr3G is that previous study has shown that depletion of Polr3gL in cell lines by siRNA targeting results in loss of cell survival and viability (Haurie et al., 2010). As, unlike in cell studies, Polr3G is the only paralogue expressed during early stages of *Xenopus* development until the activation of embryonic transcription at NF Stage 8, it is therefore possible that Polr3G would be essential for embryo survival before the expression of Polr3gL.

As Polr3gL is only expressed after the activation of embryonic transcription, it would be possible to target its expression in *Xenopus* using morpholinos. This could also indicate whether Polr3G would be able to act to replace the loss of Polr3gL due to a level of redundancy between the two subunits which was not observed in cell studies (Haurie et al., 2010). However, as the focus of this thesis is to characterise the regulation of RNA Polymerase III transcription during myogenesis, and Polr3gL is not expressed in the somites at early stages of this differentiation programme, this analysis was beyond the scope of this thesis.

6.4. Conclusions

Overall, the data in this chapter has revealed that overexpression of either Polr3G or Polr3gL in dorsal regions of tailbud *Xenopus* embryos results in similar changes to tRNA expression levels, but overexpression of Polr3gL appears to have a greater effect in downregulating expression. Data comparison for control embryos/sections of the two stages analysed by microarray revealed that a stronger correlation between tRNA expression and predicted mRNA codon usage is present in NF Stage 25 dorsal sections than NF Stage 9 embryos, possibly due to coordination of mRNA ‘demand’ and tRNA ‘supply’. However, as knockdown by morpholinos was unsuccessful, in order to fully characterise the role of Polr3G during myogenesis, disruption of protein function by gene targeting (CRISPR/Cas9) is required in future study.

7. Determining the regulation of RNA Polymerase III activity during myogenic differentiation

7.1. Introduction

7.1.1. MyoD is regulated by Wnt and FGF signalling pathways

7.1.1.1. Wnt signalling

Wnt ligands are a family of secreted glycoproteins that bind to Frizzled receptors on the surfaces of cell membranes to induce a downstream signal cascade. Wnt ligands have roles in regulating a multitude of cellular processes including cell polarity, cell fate specification, proliferation and the maintenance of stem cell populations. Wnts signal by distinct types of signalling, the canonical Wnt/β-catenin pathway (Figure 7.1. A) and by non-canonical Planar Cell polarity (PCP) and Calcium dependent (PKC) pathways (Figure 7.1 B).

In amniotes, Wnts are known as lateral signals that activate the myogenic genes MyoD and Myf5. Early experiments in chick identified that signals from the floorplate and neural tube were essential in establishing early myogenic lineages in unspecified somites (Münsterberg and Lassar, 1995). These signals were identified as a combination of Sonic Hedgehog (Shh) and Wnt ligands 1 and 3. The combination of Shh with one of these ligands was sufficient to induce expression of MyoD in dissected paraxial mesoderm in vitro (Münsterberg et al., 1995). In mouse, Wnt signals from the neural tube and dorsal ectoderm activate both MyoD and Myf5, however with differential preference dependent on particular Wnt ligands. It was found that cultures of medial paraxial mesoderm, which *in vivo* gives rise to the epaxial somite, expressed Myf5 initially in response to Shh and Wnt signals from the neural tube acting via Wnt1 (Borycki et al., 1999; Cossu et al., 1996; Tajbakhsh et al., 1998). Cells cultured from lateral paraxial mesoderm however were shown to express MyoD prior to Myf5 via Wnt7 signalling

(Tajbakhsh et al., 1998). This indicates that multiple Wnt ligands have a role in early specification of myogenic lineages in amniote development.

In *Xenopus*, Wnt signalling also regulates MyoD expression. The inhibition of Wnt8 through introduction of a dominant negative Wnt8 ligand inhibits expression of MyoD, whilst overexpression of Wnt8 results in ectopic MyoD expression (Hoppler et al., 1996). This regulation is via the β-catenin pathway as inhibition of GSK3β by introduction of LiCl resulted in expanded expression of MyoD around the blastopore at gastrula stages, whilst inhibition through introduction of a truncated TCF3 protein (N-XTcf-3) results in reduction of MyoD expression at gastrula stages (Hamilton et al., 2001). Moreover, evidence exists to suggest that Wnt8 enhances induction of dorsal mesoderm by FGF signalling resulting in increased MyoD expression in animal caps and greater muscle formation compared with similar overexpression levels of FGF4 alone (Burks et al., 2009; Slack et al., 1988).

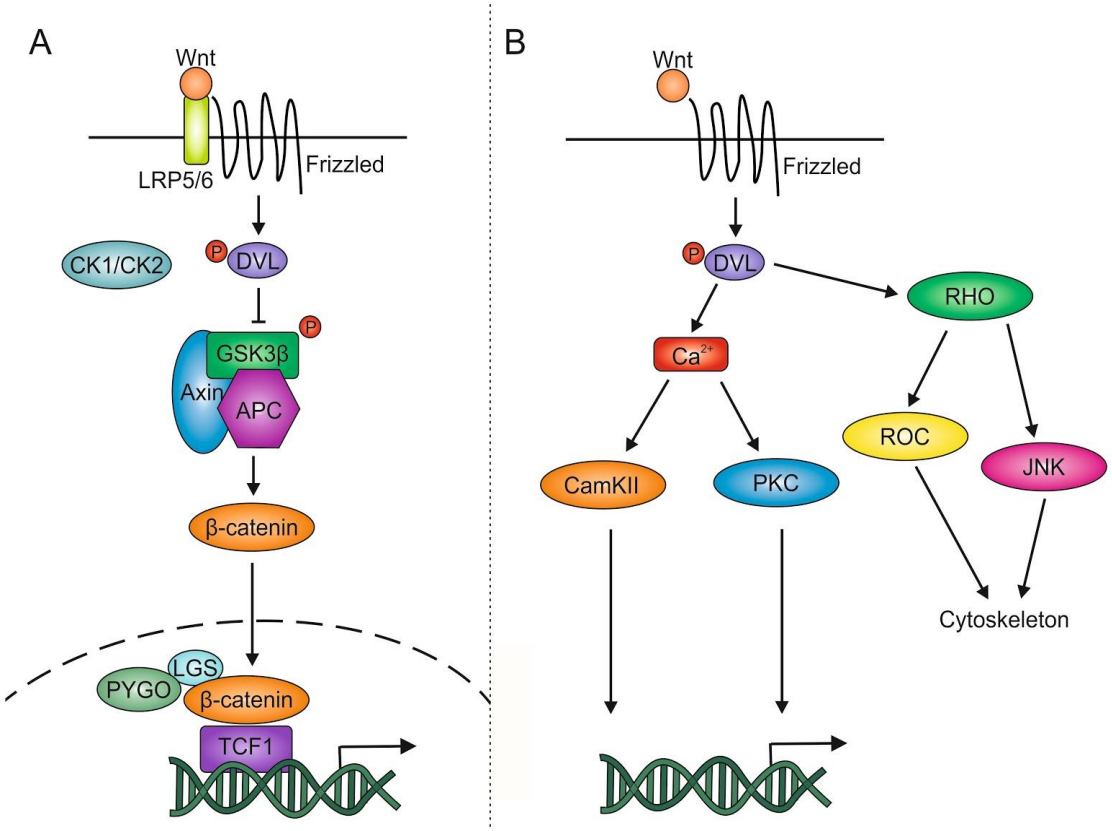


Figure 7.1. The Canonical and Non-Canonical Wnt signalling pathways. (A) The canonical Wnt/β-catenin pathway. Wnt ligands bind to either LRP5/6 and signal via the 7 transmembrane domain Frizzled which phosphorylates dishevelled (DVL). DVL disrupts the function of the GSK3β/PC/Axin complex which targets β-catenin in the absence of Wnt signalling to the proteasome for degradation. β-catenin translocates to the nucleus and activates gene expression. (B) The non-canonical PCP and Ca^{2+} pathways. Wnt ligands again bind to Frz receptors on the cell surface to activate dishevelled, however in the PCP pathway, DVL activates a Rho GTPase based signalling cascade and in Ca^{2+} dependent signalling, DVL signals via CamKII and PKC to activate gene expression.

7.1.1.2. FGF signalling

FGF is another known regulator of MyoD expression. There are 22 members of the Fgf family of secreted ligands, which all contain a 140 amino acid conserved "core" (Dorey et al., 2010; Pownall and Isaacs, 2010). Most members of the Fgf family, with the exception of intracellular ligands, signal through surface bound tyrosine kinase receptors (FGFRs 1-4) (Figure 7.2). Fgf ligands bind to FGFRs together with heparan sulfate (HS) in a 2:2:2 ratio. The dimerisation of FGFRs results in cross-phosphorylation of intracellular tyrosine residues and activates multiple signal transduction pathways including the PLCγ pathway which regulates cellular morphology and migration, the PI3K/PKB pathway regulating cell survival and the MAPK/Erk signalling pathway which has roles in proliferation and cell fate determination, including induction of mesoderm in the early embryo.

Unlike amniotes, in *Xenopus,* both MyoD and Myf5, are activated prior to somite formation (Hopwood et al., 1989). MyoD protein is detectable by immunohistochemistry at Stage 11 (Hopwood et al., 1992), and lineage tracing of gastrula stage *Xenopus* embryos revealed that expression of MyoD was exclusively localised to mesodermal progenitor cells fated to the myotomal lineage and would give rise to skeletal muscle. Therefore, MyoD has been identified as a marker of muscle lineages in *Xenopus* development (Pownall et al., 2002). FGF4 is expressed in the marginal zone as a response to mesoderm inducing signals, including Xnr1 and Xnr2 (Jones et al., 1995). Once expressed, FGF4 activates a wide variety of mesodermal genes including MyoD (Branney et al., 2009; Fisher et al., 2002; Isaacs et al., 1994). The ability of FGF4 to activate MyoD in the presence of the translation inhibitor cycloheximide shows that MyoD is in fact a

direct target of FGF4, rather than an indirect downstream target (Fisher et al., 2002).

During gastrula stages of *Xenopus* development, FGF4 mediates a phenomenon known as ‘the community effect’, whereby groups of muscle precursor cells in close proximity to one another are able to maintain myogenic specific gene expression (MyoD and Myf5), and are able to coordinate their differentiation (Standley et al., 2001). Therefore FGF4 is an important early mediator of myogenic specification, regulating the activity of MyoD and Myf5.

As in tissue culture models, MyoD coordinates the activation of a wide network of target genes to direct full myogenic differentiation. *In* vivo studies have revealed a number of novel early target genes with roles in both myogenesis and somitogenesis that were significantly downregulated in *Xenopus laevis* gastrulae after Morpholino knockdown of MyoD. These targets had not been previously identified through studies in cell cultures and therefore, the network of direct MyoD target genes required for full myogenic differentiation may be far wider than previously thought (Maguire et al., 2012). However, MyoD alone is not always sufficient to drive full myogenic differentiation. Expression of MyoD mRNA in ectodermal explants (animal caps) results in expression of a-actin representing a change in fate to myogenic lineages, but no expression of Myogenin or contractile protein genes indicating full myogenic differentiation (Hopwood and Gurdon, 1990). This was also shown in cell culture whereby ectopic expression of MyoD in hESCs alone was not sufficient to activate skeletal muscle differentiation and additionally required the chromatin remodelling factor BAF60C (Albini et al., 2013).

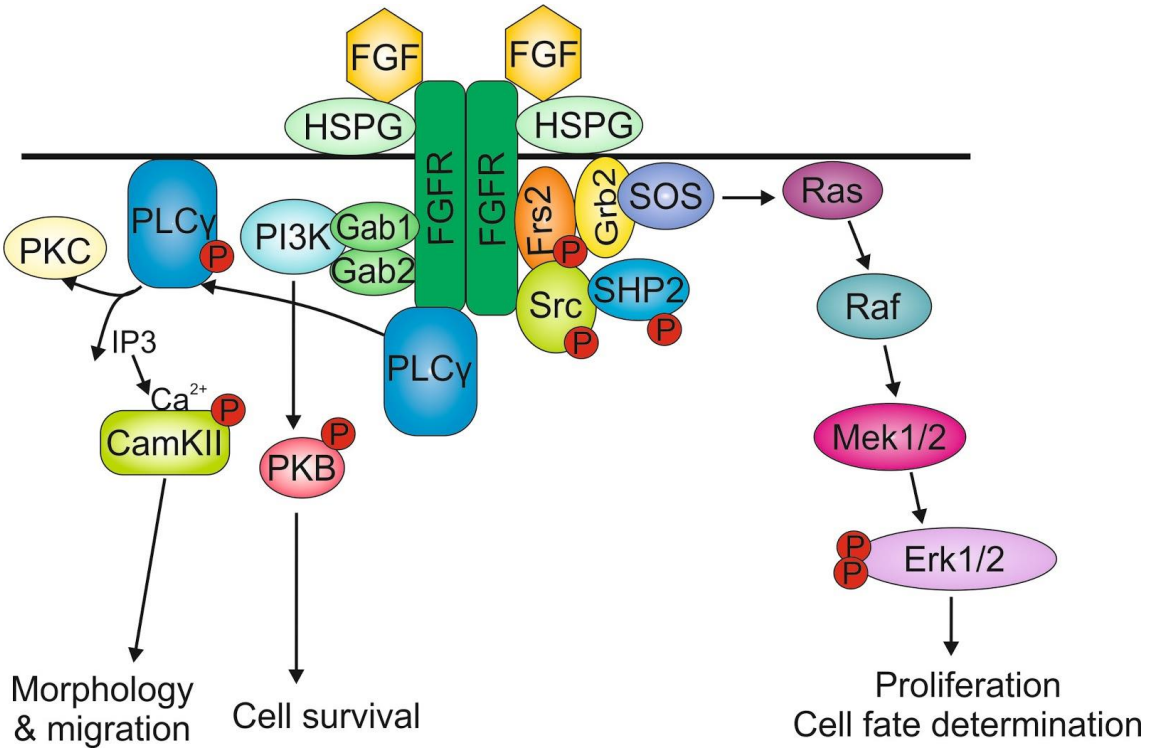


Figure 7.2. FGF signalling has diverse roles in regulating development. Binding of FGF ligands to surface FGFRs leads to dimerisation and autophosphorylation of the intracellular receptor domains. The phosphorylated domains are bound by signal transduction proteins and the activation of intracellular PLCγ, PI3K/PKB and Erk signalling cascades.

7.1.2. The animal cap assay- a way to make muscle

The *Xenopus* animal cap is a powerful tool for generating mesoderm from a naïve ectodermal tissue. As described above, Fgf and Wnt treated animal caps, make muscle. The technique makes use of pluripotent cells of the blastula stage embryo (NF Stage 8-9) and their competence to respond to factors both injected into the embryo, or diffusible factors in external culture media in order to change cell fate. Caps are grown to a desired stage and can then be used in morphological, RNA expression and protein expression analyses (Figure 7.3).

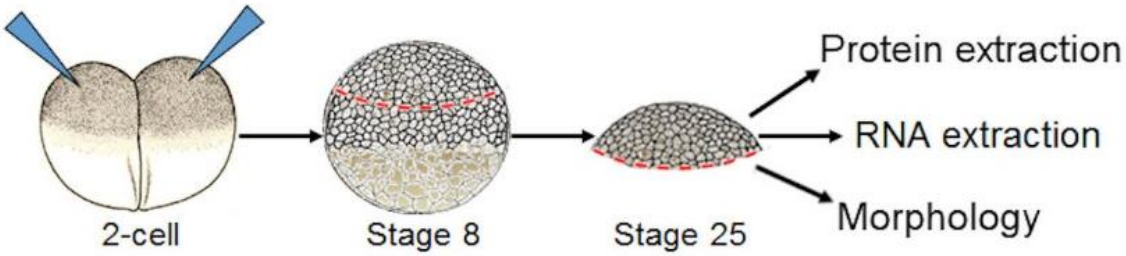


Figure 7.3. Schematic diagram of the animal cap explant assay. Control or injected embryos are left to develop until blastula Stages 8-9. They are removed of their membranes and the animal hemisphere cells dissected. Animal caps are grown in isolation in high salt media until the desired stage for collection. After which, caps are collected for RNA extraction or western blot analysis.

The animal cap is so named due to the region of the embryo from which the explant is collected, the pigmented animal hemisphere. The term cap is used because these animal cells cover the blastocoel, a cavity formed during blastula stages of cell division. The cells of the *Xenopus* embryo contain sufficient yolk for survival so the caps are cultured in simple isotonic saline solution and no tissue culture type media supplement or growth factor serum is required as is the case in tissue culture studies. Isolated caps will differentiate into ectodermal tissue (skin and nervous system), forming atypical epidermis. However, animal caps are competent to form derivatives from all three germ layers in response to injected mRNAs, culture in soluble protein or grafting to other tissues, can be diverted to a new fate.

Many cell fate changes also result in morphological changes such as cap elongation and many characterised growth factors in *Xenopus* development give recognisable phenotypic changes in the assay. It was found that overexpression of Fgf4 resulted in the formation of vesicles with differentiated mesodermal tissues muscle, blood and mesothelium (Slack et al., 1987). Fgf4 and activated Wnt

signalling (culture in LiCl or injection of Wnt8) in combination drives a more dorsal fate, muscle induction is increased compared with caps overexpressing equivalent levels of Fgf4 only and caps also show some induction of neural tissue, consistent with the role of the dorsal mesoderm as a neural inducer (Slack et al., 1988).

To further investigate the expression of Polr3G in the myogenic lineage, an animal cap based protocol was used to generate skeletal muscle using Wnt8a (Wnt8) and Fgf4. This myogenic tissue could be directly compared to the untreated ectoderm from which it was derived.

7.1.3. Aims of this chapter

1. To determine whether combined FGF + Wnt signalling regulates Polr3G and Polr3gL expression.
2. To determine whether induction of muscle differentiation by Fgf4 + Wnt8 in animal cap explants results in transcriptional changes to RNA Polymerase III targets.
3. To determine whether Fgf4/Wnt8 are regulators of RNA Polymerase III transcription.

7.2. Results

7.2.1. Characterising an induced myogenic tRNA profile

7.2.1.1. Modelling myogenesis in vivo through the animal cap assay.

To test the possibility that there is a muscle specific tRNA profile in *Xenopus*, the animal cap assay was used as a model of myogenic differentiation. Overexpression of Fgf4 alone in animal caps is sufficient to induce a change in fate to form primarily ventral mesoderm (Slack et al., 1987), and the induction of dorsal mesoderm (and muscle in particular) can be enhanced through the co-injection of Fgf4 with Wnt8 as previously shown by enhanced expression of MyoD (Burks et al., 2009). In order to determine whether a myogenic-specific tRNA profile exists, and implicate any role for somitic Polr3G in this process, a model of myogenic differentiation in animal caps was analysed by microarray.

Xenopus tropicalis animal cap explants were dissected at blastula stages (NF Stage 8/9) and cultured until tailbud stages (NF Stage 25) for analysis by custom microarray. To confirm the induction of muscle by the injection of Fgf4 and Wnt8, a set of animal caps were allowed to develop until NF Stage 40 and were processed for histological sectioning and staining with Borax-Carmine and Picro-Blue/Naphthalene-Black. Figure 7.4. shows a fully differentiated control animal cap (Figure 7.4. A), while Figure 7.4. B and C are two examples of Fgf4/Wnt8 caps at Stage 40. A comparison of control and Fgf4/Wnt8 induced caps at NF Stage 40 indicated that whilst control caps formed ciliated (Figure 7.4. A, arrow) atypical epidermis as predicted, caps overexpressing Fgf4 and Wnt8 were induced to form differentiated muscle (Figure 7.4. B) surrounded by a layer of mesothelium (mt). In

addition to muscle, some Fgf4/Wnt8 induced caps showed evidence of the differentiation of neural tissue (n) (Figure 7.4. C) and the organisation of tissue into neural tube-like structures (Figure 7.4. C, arrow). This is a result of the induction of dorsal mesoderm by Fgf4 and Wnt8 and its subsequent interaction with ectoderm also present in the caps mimicking the neural inductive signalling events that occur in whole embryos. Therefore, caps are sufficiently induced by Fgf/Wnt signalling to form differentiated muscle and can be used as a tool in microarray experiments to model myogenesis.

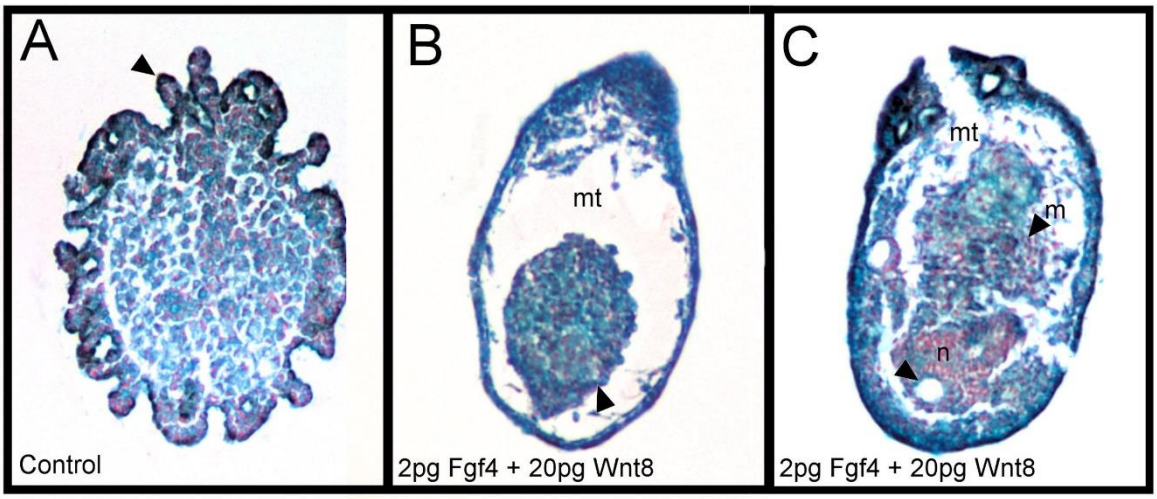


Figure 7.4. Induction of myogenic differentiation in *Xenopus tropicalis* animal caps by co-injection of Fgf4 and Wnt8. Animal caps of control uninjected embryos and embryos overexpressing both fgf4 wnt8 were collected at NF Stage 8-9 and left to develop until the desired stage (NF Stage 40) after which they were fixed overnight in paraformaldehyde and processed for histological sectioning with borax carmine and picro-blue. (A) Control animal caps NF Stage 40 form ciliated outer layers with inner layers of atypical epithelium. (B) Animal caps overexpressing both Fgf4 and Wnt8 at NF Stage 40. Arrow marks formation of differentiated muscle, surrounded by (mt) mesothelium. C) Animal cap overexpressing both Fgf4 and Wnt8 at NF Stage 40. (mt) mesothelium, (m) differentiated muscle, (n) neural tissue, arrow marks formation of neural tube-like structures.

Whilst observation at NF Stage 40 indicates that differentiated muscle is formed as a result of co-injection of Fgf4 and Wnt8, the transcriptional changes driving this change in fate and morphology are initiated earlier in development. Therefore, animal caps collected at NF Stage 25 were also collected for imaging and processed for histology (Figure 7.5). Whilst differentiation of muscle is not yet obvious at this stage, movements of the dorsal mesodermal cells induced through injection of these factors is clear from the elongated morphology of caps in comparison with control caps (Figure 7.5. B). Sections also highlight this change in morphology with control cap sections being more uniformed and round (Figure 7.5. C) whilst caps induced by Fgf4/Wnt8 show elongation and empty spaces within the sections as a result of cell movements (Figure 7.5.D, arrow). Moreover, expression of MyoD and Actc1 as measure of myogenic induction at this stage were shown significantly increased in caps induced with both Fgf4 only and with both Fgf4 and Wnt8 (Figure 7.5. E) Therefore, the downstream transcriptional effects of mesoderm induction are likely to already be significant at NF Stage 25 and therefore observing changes using microarray analysis should identify tRNAs and mRNAs transcribed during myogenesis.

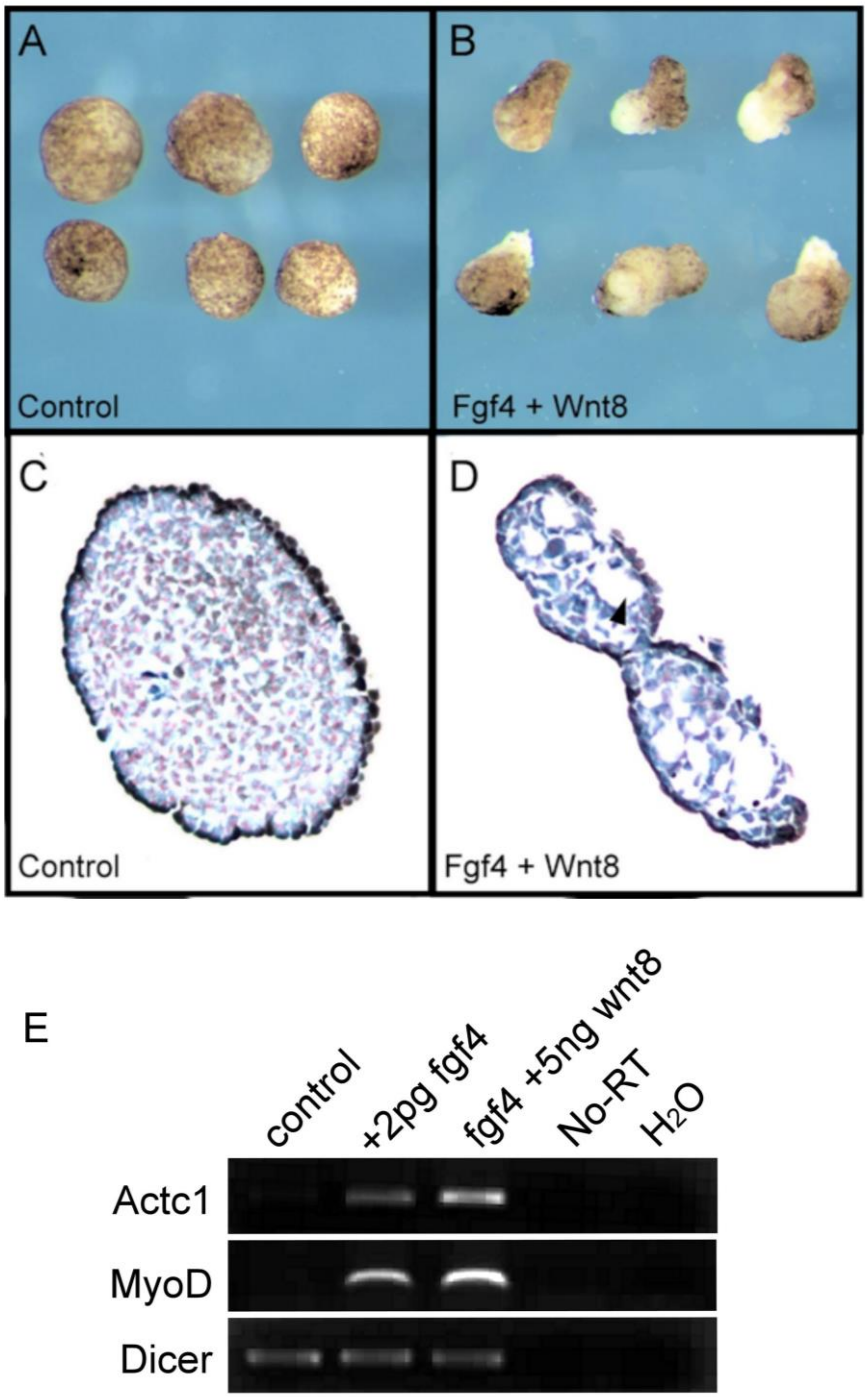


Figure 7.5. Fgf4 + Wnt8 induces mesoderm formation in *Xenopus tropicalis* animal caps with morphological effects observable by NF Stage 25. (A) *Xenopus tropicalis* control animal caps collected at NF Stage 8-9 and allowed to develop until NF Stage 25. (B) Caps injected with 2pg Fgf4 and 20pg Wnt8 at the 1- 2 cell stage. (C-D) Histological sections of control and Fgf4+Wnt8 injected animal caps at NF Stage 25 stained with Borax Carmine and Picro-Blue. (C) Representative control cap (D) Fgf4 + Wnt8 injected cap. (E) RT-PCR of Actc1 and MyoD1 in control animal caps and caps induced with Fgf4 only or both Fgf4 and Wnt8.

7.2.2. Microarray analysis of NF Stage 25 Fgf4 + Wnt8 induced animal caps

7.2.2.1. tRNA expression in control animal caps and animal caps induced by Fgf4 + Wnt8.

To determine the regulation of tRNAs by FGF/Wnt signals, a custom Agilent microarray described in previous chapters (p80, p160) was used to analyse transcription in NF Stage 25 animal caps. Strikingly, when Fgf4/Wnt8 caps were analysed for expression in comparison to control caps, all tRNA isoacceptor families detected on the microarray showed significantly reduced expression in response to injection of Fgf4 + Wnt8 (Appendix. Table 3, Figure. 7.6. A). When other RNA Polymerase III targets were analysed, it was clear that they too were significantly downregulated in Fgf4 + Wnt8 samples, suggesting an overall repression of RNA Polymerase III transcription in response to these inductive signals.

In contrast, and as expected, a number of contractile protein genes and the myogenic regulatory factor (MRF) genes were significantly upregulated in Fgf4/Wnt8 caps (Figure 7.6. B and C). Notably, the early myogenic marker gene Actc1 was highly upregulated, indicating that the muscle lineages induced were in the early stages of differentiation rather than terminally differentiated as expected at this stage. Of the myogenic factors, Myf5 was the most highly upregulated in response to Fgf4 + Wnt8 with a mean fold change of 69.6 (Figure 7.6. C), followed by MyoD (fold change- 23.9). Mrf4 and Myogenin, factors more associated with terminal differentiation, were less highly upregulated (fold change 12.4 and 6.4 respectively).

Cell cycle regulators known for their role in proliferation (Ccne1, Ccne2, Mycl and Mycn) were upregulated in response to Fgf4 + Wnt8. Moreover, the cell cycle inhibitor p27kip1 was significantly downregulated (Figure 7.6. D). This indicates that induced animal caps at NF Stage 25 still proliferative. However, other markers of pluripotency such as Klf4, Klf6, Myc and Aurka show either unchanged or downregulated expression in Fgf4/Wnt8 caps, suggesting a loss of potency. Determined myoblasts expressing Myf5 and MyoD are still proliferative despite their commitment to a myogenic fate (Emerson, 1990; Pownall et al., 2002), therefore indicating that the cells in the Fgf4/Wnt8 induced animal caps are likely to be myoblasts, as shown in previous studies.

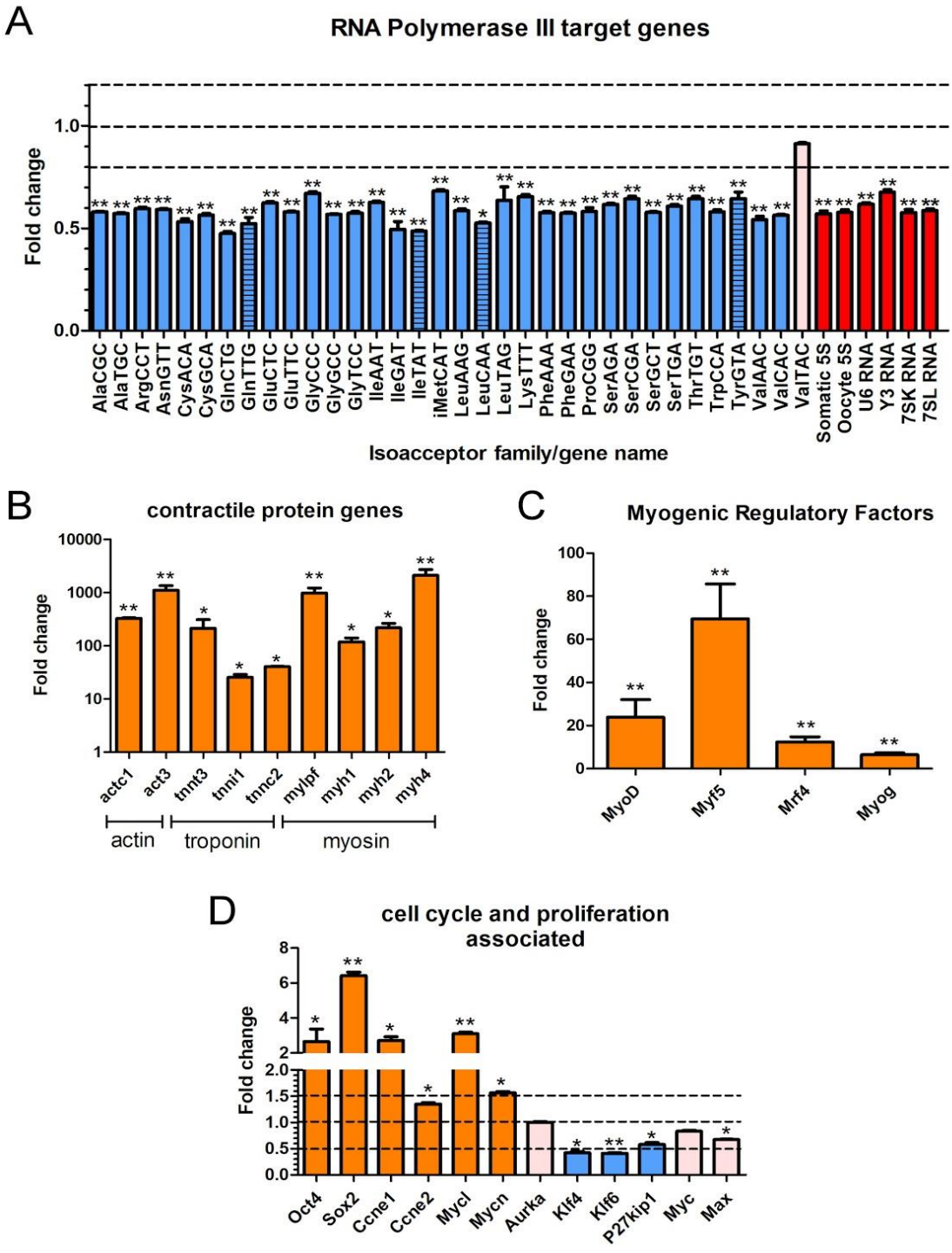


Figure 7.6. Overexpression of Fgf4 + Wnt8 in NF Stage 25 *Xenopus tropicalis* animal cap explants results in changes to altered RNA Polymerase III transcription and differentiation. Expression fold change summary of *X.tropicalis* NF Stage 25 animal cap explants injected with 2pg Fgf4 and 20pg Wnt8 mRNA versus control uninjected caps across 3 biological replicates. RNAs with raw signal intensities of >10 were included in the analyses. (A) RNA Polymerase III target genes (significantly altered tRNA isoacceptor families are indicated in blue, other RNA Polymerase III targets are included in red). (B) Contractile protein

genes showing significant upregulation in Fgf4 + Wnt8 caps. Note that Y-axis scale is logarithmic (C) Myogenic Regulatory Factors. (D) Selected genes associated with pluripotency and cell cycle regulators. Pink in all panels indicates no significant change in expression. SEM is presented as error bars and * indicates the level of statistical significance between control and injected samples after Bonferroni Correction.

7.2.2.2 Regulation of the Polymerase III transcription machinery by Fgf4 + Wnt8 signalling.

To determine possible reasons for the downregulation of RNA Polymerase III transcription, expression levels of known subunits of Polymerase III and its transcriptional machinery were also analysed. Significantly, many components of RNA Polymerase III machinery were significantly downregulated in response to Fgf4 + Wnt8 (Figure 7.7.)- notably, the components of the TFIIIB complex, Brf1 and Bdp1 which are important for transcription initiation and have been implicated in F9 cell studies to be important for regulation of RNA Polymerase III transcription during differentiation (Athineos et al., 2010). The majority of RNA polymerase III subunits were also downregulated with the exception of Polr3A, Polr3C, Polr3K and Polr3gL (Figure 7.7.). These data indicate a downregulation of RNA Polymerase III activity at the transcriptional machinery level during myogenic differentiation.

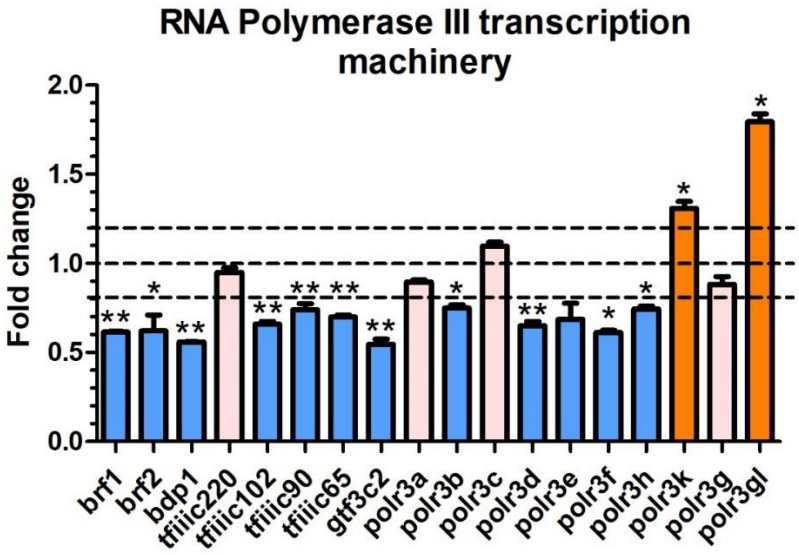


Figure 7.7. Overexpression of Fgf4 + Wnt8 in NF Stage 25 animal cap explants results in downregulation of RNA Polymerase III core transcription machinery. Expression fold change summary of *X.tropicalis* NF Stage 25 animal cap explants injected with 2pg Fgf4 and 20pg Wnt8 mRNA versus control uninjected caps across 3 biological replicates. RNAs with raw signal intensities of >10 were included in the analyses.

7.2.3. Validation of RNA Polymerase III downregulation by qPCR analysis

To validate the global downregulation of RNA Polymerase III transcription in Fgf4/Wnt8 induced animal caps, selected tRNAs and the RNA Polymerase III target 5S rRNA were analysed for expression by qRT-PCR. In support of the microarray analyses, qPCR for selected RNA Polymerase III target genes showed decreased expression levels in induced caps compared with control levels (Figure 7.8.). The confirmation that not only tRNA transcription is affected by the induction of a myogenic differentiation programme, but that wider RNA Polymerase III transcription is also downregulated. This is striking in comparison to the dramatic upregulation of a set of RNA Polymerase II genes.

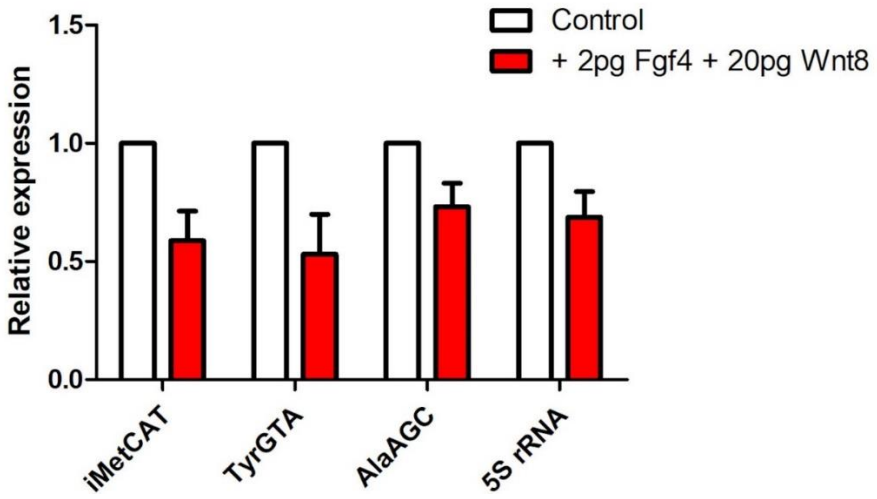


Figure 7.8. Validation of RNA Polymerase III transcriptional changes in Fgf4 + Wnt8 induced animal caps by qRT-PCR. qRT-PCR analysis of expression of selected tRNAs and 5S rRNA in control and Fgf4 + Wnt8 induced animal caps was carried out at NF Stage 25. Relative expression levels compared with control cap expression was calculated using the ΔΔCt method using Dicer as an endogenous control for 3 biological replicates. Mean relative expression is presented on the chart with error bars representing SEM.

7.2.4. Determining the regulation of Polr3G and Polr3gL by myogenic inducing growth factors

When analysed separately, the three biological replicates included in the microarray showed different extents of induction of MRFs and contractile protein gene expression. Despite all being statistically significant, the extent of expression changes in experimental set 1 were much lower than in the other 2 sets. (Figure 7.9. A- note logarithmic scale). This could be due to sample 1 being slightly younger than the other two, this is consistent with higher levels of Myf5 and lower levels of other contractile protein genes. Therefore, when analysing Polr3G and Polr3gL expression in Fgf4/Wnt8 induced caps by qRT-PCR, additional samples were collected to ensure staging of samples was representative. In caps induced

by Fgf4 + Wnt8, no significant change in Polr3G expression was identified, however Polr3gL was found to be significantly upregulated (Figure 7.9. B). Actc1 was included as a control gene to ensure variation across the three samples was limited and did not affect statistical analyses. This data suggests that, unlike MyoD alone, Fgf4/Wnt8 induction of animal caps does not lead to upregulation of Polr3G but does result in increased expression of Polr3gL, the subunit included in the isoform of RNA Polymerase III more associated with differentiation. Despite sample 1 showing a less differentiated transcriptional profile, the qRT-PCR analysis validates the findings of the microarray where Polr3gL is upregulated along with the myogenic differentiation genes.

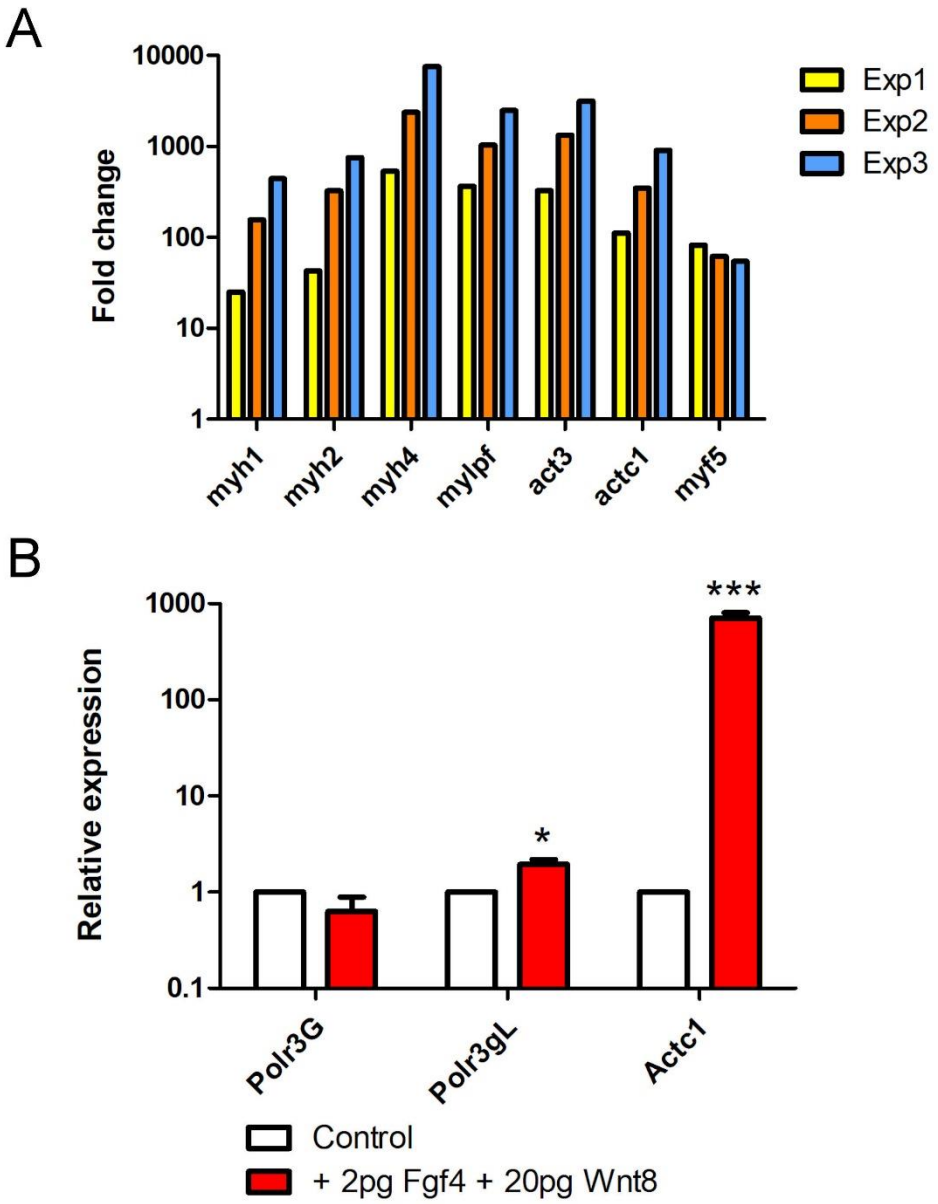


Figure 7.9. qRT-PCR analysis of Polr3G and Polr3gL expression levels in Fgf4 + Wnt8 induced animal caps at NF Stage 25. (A) Expression of contractile protein genes in the three biological replicates used for microarray analysis. Note the log scale for fold change. (B) qRT-PCR analysis was carried out for control animal caps and animal caps induced by Fgf4 and Wnt8 at NF Stage 25 for two biological replicates (2 and 3) included on the microarray analysis and an additional pair collected (set 4), but not analysed on the microarray, to account for lower values in experimental set 1. Ct values were normalised to the control gene Dicer and expression values were calculated relative to Polr3G expression levels in control caps. Paired t-tests were carried out for the data (*) indicates level of significance * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Note that scale on Y axis is logarithmic.

7.2.5. Investigating Maf1 and Polr3G as regulators of RNA Polymerase III downregulation during myogenic differentiation

Maf1 is a known negative regulator of RNA Polymerase III activity. Nuclear Maf1 inhibits interaction of the transcription initiation complex of RNA Polymerase III to inhibit transcription (Vannini et al., 2010). As Maf1 was not included on the microarray, an additional animal cap experiment was designed and carried out on a single biological replicate to investigate (1) whether Maf1 is regulated during myogenic differentiation induced by Fgf4/Wnt8 and therefore if Maf1-based negative regulation could be a possible mechanism of the downregulation of RNA polymerase III activity observed in animal caps and (2) whether expression of Polr3G is sufficient to rescue the downregulation of RNA Polymerase III target genes in Fgf4/Wnt8 animal caps.

Animal caps were injected either with combined 2pg Fgf4 + 20pg Wnt8 as per the microarray analysis or were injected with a combination 2pg Fgf4 + 20pg Wnt8 + 1ng Polr3G mRNA and allowed to develop until NF Stage 25. qRT-PCR analysis was carried out on selected RNA Polymerase III target genes as well as Maf1 and Actc1 (to determine induction of myogenic differentiation). Table 7.1 shows the relative fold changes of expression of each gene analysed in both experimental sets of caps versus expression in uninjected control caps. Polr3G overexpression in animal caps induced with Fgf4/Wnt8 was not sufficient to rescue the downregulation of RNA Polymerase III target genes observed in the microarray and Fgf4/Wnt8 only caps. However, overexpression of Polr3G reduced the expression of Actc1 in Fgf4/Wnt8 caps. Previous study in ESCs showed that Polr3G overexpression results in a resistance of ESCs to induced differentiation by Retinoic Acid (Wong et al., 2011). This result therefore suggests that Polr3G may

be able to inhibit induction of myogenic differentiation in animal caps by Fgf4/Wnt8 but is not sufficient to rescue the downregulation of Polymerase III activity.

Maf1 expression is not significantly upregulated in animal caps induced by Fgf4/Wnt8 (Table 7.1, middle column). Moreover, Erk-based phosphorylation of Maf1 in Drosophila results in the inhibition of nuclear localisation of Maf1, and the upregulation of RNA Polymerase III activity (Sriskanthadevan-Pirahas et al., 2018). Therefore, Maf1-based inhibition is unlikely to be the mechanism of downregulation observed during myogenic differentiation of animal caps.

	Control	Fgf4 + Wnt8	Fgf4 + Wnt8 + Polr3G
iMet	1.000	0.153	0.184
Tyr	1.000	0.224	0.266
Ala	1.000	0.697	0.621
5S	1.000	0.325	0.386
Actc1	1.000	77.849	35.300
Maf1	1.000	1.134	0.691

Table 7.1. Investigation of Maf1 as a potential regulator of RNA Polymerase III activity in differentiating animal caps. Relative expression values of selected RNA Polymerase III target genes, the RNA Polymerase III regulator Maf1 and Actc1 as a marker of myogenic differentiation in control animal caps, animal caps induced with Fgf4 and Wnt8 and animal caps expressing Fgf4, Wnt8 and Polr3G.

7.3. Discussion

7.3.1. Downregulation of Polymerase III activity

The results from the microarray analysed in this chapter indicate that, whilst myogenic and contractile protein genes are significantly upregulated in response to induction of myogenesis in animal caps by Fgf4 and Wnt8 (Figure 7.6. B and C), a global downregulation of RNA Polymerase III transcription is evident (Figure 7.6. A and D). tRNAs and other RNA Polymerase III target genes show greater downregulation as a response to Fgf4 and Wnt8 than any other overexpression analysis carried out in this thesis with most tRNA expression levels downregulated to 0.6 or under compared with control cap levels. This was surprising due to the expectation that RNA Polymerase III activity may correlate to mRNA codon usage demand set by differentiated cells as observed in other studies in human tissue and the silkworm silk gland (Dittmar et al., 2006; Garel and Hentzen, 1974; Hentzen et al., 1981).

However, downregulation of RNA Polymerase III activity in response to differentiation has also been observed in F9 Embryonal Carcinoma (EC) cell culture studies. An initial study using the cell line described revealed that when F9 cells are induced to differentiate into endoderm by exposure to Retinoic Acid, transcription of tRNAs, alongside other RNA Polymerase III targets were dramatically downregulated. This study also showed that the downregulation effect was due to reduced activity of the TFIIIB complex (White et al., 1989). Subsequent studies confirmed this finding and determined that the loss of TFIIIB activity was due specifically to decreased levels of the Brf1 and Bdp1 subunits of TFIIIB accompanied with downregulated c-Myc expression and downregulation of Erk phosphorylation (Athineos et al., 2010; Dumay-Odelot et al., 2010). Surprisingly, in

this analysis, Fgf4/Wnt8 induced caps show no change in Myc (c-Myc) expression and indeed, upregulated Mycn and Mycl expression was observed. However, consistent with the previous studies in F9 cells, Brf1 and Bdp1 (along with most other subunits of the RNA Polymerase III transcriptional machinery) were downregulated in response to Fgf4/Wnt8 induced myogenic differentiation (Figure 7.7). It is possible that *in vivo* differentiation programmes modulate RNA Polymerase III activity using these same mechanisms here and that induction of cell differentiation or loss of pluripotency drives downregulation of RNA Polymerase III activity.

7.3.2. Regulation of Pluripotency and Proliferation during Fgf4/Wnt8 induced myogenesis

7.3.2.1. Proliferation Markers

The downregulation of RNA Polymerase III during F9 cell differentiation was found to be a response of reduced cellular proliferation identified by to the reduced levels of positive regulators of TFIIIB activity c-Myc and phosphorylated Erk (Athineos et al., 2010). These mechanisms are unlikely to be working in Fgf4/Wnt8 animal caps. In control animal caps, no activation of Erk signalling is induced. Fgf4 signals through the MAPK/Erk signalling pathway during mesoderm induction, thus increasing levels of dpErk in animal caps (Branney et al., 2009). Therefore, a reduction in activated Erk levels compared with control animal caps would not be likely in this context. Moreover, while the regulatory phosphorylation state of Myc is unknown in this context (Macek et al., 2018; Welcker et al., 2004), transcriptional levels of Myc remain unchanged (Figure 7.6. D). Overall the conclusion is that the regulation of these factors is unlikely to be a conserved

mechanism for the downregulation of RNA Polymerase III targets observed in this analysis. As some other cell cycle genes were included on the microarray, these were analysed in order to try to determine proliferation levels. Surprisingly, despite activation of myogenesis, the cell cycle progression markers Ccne1 and Ccne2 (Caldon and Musgrove, 2010) were upregulated in the growth factor induced caps (Figure 7.6. D). Moreover, the cell cycle inhibitor and known promoter of neuronal differentiation p27kip was significantly downregulated (Carruthers et al., 2003). These data suggest that cells in Fgf4/Wnt8 induced caps have increased proliferation rates despite activating both myogenic and neurogenic differentiation programmes. MyoD and Myf5 are expressed in determined proliferative myoblasts prior to full myogenic differentiation (Pownall and Emerson, 1992). Therefore, in myogenic lineages, differentiation programmes are activated and proceed in the presence of cell proliferation, before the myoblasts exit the cell cycle during terminal differentiation and formation of mature muscle. Therefore, it may be that *in vivo* regulation of RNA Polymerase III activity during myogenesis is independent of proliferation levels and that a more decisive signal dominates expression of Polymerase III products. The data in this chapter is suggestive of a mixed population of cells, some of which are still proliferative and expressing MyoD and Myf5, whilst others have exited the cell cycle and express differentiation genes such as Myogenin and Myhs.

Much regulation of RNA Polymerase III activity is directed by protein-protein interactions and modifications (Felton-Edkins et al., 2003b; Sutcliffe et al., 2000; Vannini et al., 2010; Welcker et al., 2004). Therefore, investigation of the post-transcriptional regulation of RNA Polymerase III basal transcriptional machinery and its regulators of activity such as Maf1 and c-Myc may elucidate in more detail

the mechanism of the downregulation observed in animal caps. This would give the model of downregulation constructed from cell culture studies an *in vivo* developmental context. Moreover, as cell growth rates appear to determine transcriptional activity of RNA Polymerase III in other model organisms (Dong et al., 1996; Dumay-Odelot et al., 2010; Sriskanthadevan-Pirahas et al., 2018; White, 2005), a closer look at proliferation rates using PCNA as a marker of proliferative cells may help to explain the slightly contradictory cell cycle results from the microarray and whether all cells are dividing or whether only a subset is responsible for this result.

7.3.2.2. Pluripotency markers

Some markers of pluripotency (Klf4 and Klf6) were significantly downregulated in Fgf4/Wnt8 induced animal caps (Figure 7.6. D) (Nandan and Yang, 2009). However, other factors indicative of a pluripotent state were upregulated. Oct4 and Sox2 are typically markers of pluripotent stem cells and are coexpressed with Polr3G in ESCs (Wong et al., 2011). Therefore it is surprising for them to be upregulated in Fgf4/Wnt8 caps. However, animal caps induced by these growth factors also differentiate into neural tissues in addition to myogenic lineages (Figure 7.4.). Sox2 is a marker of neuromesodermal progenitors, the neural crest and is expressed in the neural tube, a structure observed in differentiated Fgf4/Wnt8 induced animal caps (Chal and Pourquié1, 2017; Graham et al., 2003; Monsoro-Burq et al., 2005; Takemoto et al., 2011). Moreover, the homeobox gene Oct4 is known to interact with Sox family proteins during neurogenesis and its inhibition results in repression of neural induction (Archer et al., 2011). This

therefore may explain why an increase in expression of these factors is observed in the microarray, despite induction of differentiation.

7.3.3. Regulation of Polr3G by Fgf4 and Wnt8

Unlike MyoD expression alone, Fgf4 and Wnt8 combined overexpression in animal caps leads to no change in Polr3G expression but increased expression of Polr3gL (Figure 7.7. and Figure 7.9. B). The aim of this chapter was to capture early stages of the myogenic differentiation programme, in which proliferative mesodermal cells become specified to the skeletal muscle lineage and express the early myogenic regulatory factors MyoD and Myf5. The data collected indicates that MyoD and Myf5 were significantly upregulated in Fgf4/Wnt8 caps compared with control levels and that Myf5 was the most significantly upregulated. It is widely accepted that Myf5 is the earliest expressed MRF and a marker of undifferentiated myoblasts, while MyoD is activated just after Myf5 playing a role in terminal differentiation as well as specification of progenitor cells (Conerly et al., 2016; Hopwood et al., 1991; Tajbakhsh et al., 1998). At first glance, this indicates that the Fgf4/Wnt8 caps collected were in the earliest stages of myogenesis. However, there was also upregulated expression of Mrf4 and Myogenin too (12.4 and 6.4 fold change respectively) (Figure 7.6. C), but to a lesser extent. This observation, along with the downregulation of some genes associated with pluripotency (Klf4 and Klf6) and upregulation of some contractile protein genes, indicates that the Fgf4/Wnt8 caps were committed to terminal myogenic differentiation. This programme of full myogenic differentiation cannot be induced by MyoD alone in animal caps, or indeed not in ES cells or P19 cells (Albini et al., 2013; Fong et al., 2012; Hopwood and Gurdon, 1990). The transient somitic

expression of Polr3G observed in tailbud stages of *Xenopus* development, and the ability of MyoD to activate Polr3G expression in animal caps (Figure 4.6, Figure 4.11), suggests that there is a specific early timepoint at which myoblasts express Polr3G, possibly whilst still proliferative, before further differentiation. The experiment described in this chapter did not capture that timepoint.

During differentiation of ES cells, Polr3G expression is downregulated and the expression of Polr3gL is maintained, indicating that the Polr3gL isoform of RNA Polymerase III is more associated with differentiated, mature cell types (Haurie et al., 2010). This is consistent with the upregulation of Polr3gL observed in the microarray study. In this study the early specification of cells to the myogenic lineage has already occurred in the NF Stage 25 Fgf4/Wnt8 caps, and the now committed myoblasts are differentiating and activating the expression of differentiation factors such as Myogenin and contractile protein genes. This data may also explain the observation that overexpression of Polr3gL in embryos results in a greater downregulation of tRNAs compared with Polr3G (Figure 6.4). That Polr3gL, in response to differentiation factors, acts to promote myogenic differentiation in embryos, and the corresponding downregulation of RNA Polymerase III activity.

The overexpression of Polr3G in ESCs resulted in reduced expression of many differentiation markers after induced differentiation by Retinoic Acid (Wong et al., 2011), suggesting that Polr3G actively promotes pluripotency in the presence of driving signals of differentiation. Therefore, it was possible that relieving limitations to expression levels of Polr3G might be able to rescue the downregulation of tRNAs and other target genes during myogenic differentiation of animal caps. In this study, Polr3G alone was not sufficient to upregulate expression of tRNAs and

5S rRNA in animal caps compared with caps induced by Fgf4/Wnt8 only. It therefore may be that, at this stage during myogenic differentiation, Polr3gL, as its expression exceeds that of Polr3G and the two factors appear to regulate the same target genes, competes with and is able to overcome the activity of Polr3G in the regulation of RNA Polymerase III target genes. However, the reduced expression of Actc1 in the presence of Polr3G suggests that, as in ESCs, Polr3G can increase resistance of cells to fully differentiate and maintain a progenitor-like state even in the presence of inducing signals.

7.4. Conclusions

This chapter has shown that Fgf4 and Wnt8 induces a myogenic transcriptional profile in animal caps, characterised by upregulation in MRFs (primarily Myf5) and contractile protein genes (Myh). The overexpression of Fgf4 and Wnt8 in animal caps also results in a downregulation of many subunits of RNA Polymerase III transcriptional machinery, of particular interest Brf1 and Bdp1, and this downregulation is coupled with global downregulation of RNA Polymerase III transcriptional targets, consistent with early studies in F9 cells (Athineos et al., 2010; White et al., 1989). Fgf4/Wnt8 induction of animal caps resulted in no change in expression of Polr3G at the stage analysed but showed upregulated expression of Polr3gL, the paralog expressed later in development and maintained during differentiation. This is possibly due to the activation of a terminal myogenic differentiation programme by Mrf4 and Myogenin that are also found to be upregulated in the Fgf4/Wnt8 animal caps. Together these results suggest that

Polr3G expression in the somites, regulated at least in part by MyoD (Figures 4.10- 4.12) has transient activity and does not have a role in full myogenic differentiation. Polr3gL is the relevant isoform at the stage analysed here at the onset of full myogenic differentiation.

8. Discussion

8.1. Summary

The aim of this PhD was to determine whether transcription by RNA Polymerase III is regulated during myogenic differentiation in *Xenopus tropicalis*, and whether the alternate subunit Polr3G or Polr3gL has a role in this regulation. The data presented in chapters 4-7 characterise the distinct expression profiles of these subunits during development and show that Polr3G is transiently expressed in the somites during early myogenic differentiation, identifying the subunit as a potential novel player in the regulation of myogenesis. I have shown that Polr3G is at least partly regulated by MyoD *in vivo,* but that an induced model of full myogenic differentiation in animal caps by expression of Fgf4 and Wnt8 results in upregulated levels of Polr3gL instead. I have also shown through custom-designed tRNA microarrays, that modulation of Polr3G expression leads to dynamic changes in expression of tRNAs during early development and in myogenic lineages, and that Polr3gL overexpression results in a greater downregulation of tRNAs compared with overexpression of Polr3G. This suggests that these RNA Polymerase III subunits have overlapping target genes, but different levels of activity when expressed in *Xenopus*. This thesis is the first study to characterise these subunits using a model organism.

I have also shown in chapter 7 that wider effects on RNA Polymerase III activity are imposed during the induction of skeletal muscle differentiation in animal caps as a result of growth factor signalling. This chapter shows that Fgf4/Wnt8 induced differentiation results in dramatic downregulation of tRNAs and other RNA Polymerase III target genes, as well as subunits of RNA Polymerase III transcriptional machinery. This thesis is the first study to identify a conserved

effect of differentiation programmes on the activity of RNA Polymerase III in cell culture studies and in animal caps.

8.2. The ever expanding network of transcriptional regulation by MyoD

When MyoD was first discovered, its importance in regulating myogenesis was determined by a simple experiment whereby driving its expression was sufficient to convert a number of cell types to a myogenic fate (Choi et al., 1990; Davis et al., 1987; Pinney et al., 1988). By the 1990s, MyoD had been coined a master regulator of myogenesis (Weintraub et al., 1991); its expression in fibroblasts sufficient to drive expression of myosin heavy chain and its expression was found localised to the somites of many model organisms (Davis et al., 1987; Pownall et al., 2002). In *Xenopus,* MyoD expression was shown localised early to myogenic lineages in both the early mesoderm and the somites (Hopwood et al., 1989; Pownall et al., 2002). However, shortly after it was determined that MyoD alone could not convert *Xenopus* ectodermal explants to a myogenic fate (Hopwood and Gurdon, 1990), suggesting and so a wider transcriptional network was required. The advances in transcriptomic analyses have enabled, in great detail, the characterisation of the MyoD transcriptional network. MyoD has been shown to interact with a number of cofactors such as E proteins, Mefs and Pbx2 (Berkes and Tapscott, 2005; Tapscott, 2005) that generate a feed-forward circuit for MyoD transcriptional activity. MyoD has also been shown to interact with different genomic regulatory elements in myoblasts compared with differentiated myofibers. This process, known as enhancer switching, is regulated by the presence or absence of the inhibitory complex Snai1-HDAC1/2 which, when bound to

enhancer sites, prevents MyoD binding (Soleimani et al., 2012). This prevents activation of late differentiation genes. These experiments showed that transcriptional regulation by MyoD was regulated temporally (Penn et al., 2004). More recently, ChIP studies in P19 cells have shown MyoD is required not only for activation of myoblast genes during differentiation of stem cells to muscle, but initially regulates pre-myogenic mesoderm genes such as Meox1, Pax7 and Six1 (Gianakopoulos et al., 2011).

The enhancement of genome-wide analyses such as RNA-Seq and ChIP-Seq in the 2000s, resulted in the genome-wide mapping of specific DNA sequence elements which are more likely to be used by MyoD in order to activate gene expression, though MyoD was found to bind many intergenic regions associated with no gene activation (Cao et al., 2010). Studies have also shown that MyoD binding is associated with chromatin remodelling and epigenetic modifications in the absence of gene expression (Conerly et al., 2016; De La Serna et al., 2005; Oler et al., 2010). Together, these data indicate that MyoD not only directly activates gene expression of target genes through binding of enhancers in their regulatory regions, but may also have a facilitative role in enabling other factors to activate target genes, thus regulating in an independent manner too. High-throughput, RNA-sequencing has also been useful in the identification of muscle-specific non-coding RNAs in chick (Rathjen et al., 2009). This thesis describes a genome-wide analysis in *Xenopus* revealing novel targets of MyoD within the early myogenic lineage.

Maguire et al. 2012, highlighted the importance of including *in vivo* studies when characterising gene networks through identification of many novel MyoD target

genes, with roles in somitogenesis in addition to genes involved in myogenic differentiation such as Myf5. This study on an Affymetrix microarray containing 14,000 features identified over 300 genes differentially regulated in embryos targeted with MyoD antisense morpholinos. This study revealed a novel function of MyoD to drive two processes important for the formation of skeletal muscle simultaneously (Maguire et al., 2012). Another important finding from this study was the evidence that Myf5 expression during gastrula stages of *Xenopus* development is dependent on MyoD. Previous studies in other vertebrates and in *Xenopus* identify Myf5 as the earliest MRF expressed (Hopwood et al., 1991; Pownall et al., 2002; Rudnicki et al., 1993), yet perhaps MyoD expression is required for establishment of Myf5 expression in mesoderm to establish a myogenic lineage much earlier than the formation of somites.

A role for MyoD in the regulation of RNA Polymerase III activity during myogenesis in Xenopus.

This thesis has revealed that transcription by RNA Polymerase III may also be an important node in the MyoD transcriptional network. Like microRNAs, tRNAs are small, non-coding RNAs important for post-transcriptional regulation of gene expression. As mediators of translation, their expression levels are highly coordinated with cell growth and proliferation rates (Dumay-Odelot et al., 2010; White, 2005) and evidence exists to suggest that their expression levels also correlate with demand from mRNA codon usage (Dittmar et al., 2006; Gingold and Pilpel, 2011). However, *in vivo* development is highly complex and it is likely that many factors influence activity of RNA Polymerase III. This thesis has shown that MyoD can regulate Polr3G, a developmentally regulated subunit in *Xenopus,* and that through induced myogenic differentiation, Fgf4/Wnt8 induction of animal caps

activates the related gene Polr3gL. This results in a strong downregulation of other RNA Polymerase III transcription factors and targets of RNA Polymerase III. In contrast to findings in cell culture studies, evidence in Chapter 7 however, suggests that, as indicated by the upregulation of key cell cycle genes, that RNA Polymerase III activity is downregulated independently of cell proliferation.

It is not yet clear whether a conserved determinant of tRNA expression levels is present throughout eukaryotes. Studies in invertebrates suggest that tRNA expression is dependent on other controlling factors such as gene copy number (Duret, 2000). In this study, tRNA isoacceptors with greater number of encoding genes are more highly expressed than those with lower gene numbers. However, data from Chapter 5 suggests that this is not the case in *Xenopus* (Figure 5.12.). Data in this thesis instead suggests that the conversion of lineage and transcriptional changes in protein coding genes during myogenic differentiation may have a dominant regulatory effect on RNA Polymerase III activity *in vivo*. Chapter 6 indicates an increasing correlation between mRNA codon usage and tRNA isoacceptor expression from the microarray results in NF Stage 25 embryos compared with NF Stage 9 embryos but that during early stages of myogenic differentiation, a weak correlation still exists (Figure 6.6). It is not yet clear if the resulting levels of mature tRNA pools correlates at all to mRNA codon usage in muscle cells. Previous study in human tissues has shown correlation between codon usage bias and tRNA expression in human tissues, however only a small number of isoacceptor families showed a significant and strong correlation (Dittmar et al., 2006). It is possible therefore, that by inclusion of all isoacceptors expressed above threshold level may mask a subset of families with a stronger correlation to mRNA codon usage. It would therefore be useful to extend this

analysis to look at a more differentiated stage of myogenesis (i.e. NF Stage 40) and to analyse individual amino acids for correlation to determine if a subset show enhanced coordination to RNA Polymerase II activity as in human studies. Studies in bacteria support this finding by suggesting that the correlation of tRNA expression to corresponding mRNA codons enhances translational efficiency of a cell, this would be favourable in cells undergoing differentiation as enhanced translation would enable the cell to produce new proteins faster (Dittmar et al., 2006; Gingold et al., 2014; Saikia et al., 2016). However, contrasting with this evidence, other mammalian studies suggest that during differentiation of muscle and liver, a dominant regulation mechanism exists to maintain stable levels of tRNAs throughout development, as upregulation of one gene of an isoacceptor family results in corresponding downregulation of a different gene within the same family (Rudolph et al., 2016; Schmitt et al., 2014). This evidence therefore suggests that tRNA levels are stable, despite changing translational needs. This thesis, whilst contrasting with invertebrate studies, does not yet clearly define a dominant determinant of tRNA expression *in vivo*.

8.3. Polr3G regulation by other myogenic factors

This thesis investigates the role of MyoD in regulation the RNA Polymerase III subunits Polr3G and Polr3gL during myogenesis. However, in *Xenopus* Myf5 and Mrf4 are also important regulators of myogenesis with overlapping expression profiles. In amniotes, MyoD independent lineages within somites express Myf5, and the two appear to have partially redundant roles during myogenic determination (Haldar et al., 2008; Rudnicki et al., 1993). Myf5 and MyoD are also known markers of proliferative myoblasts (Emerson, 1990; Pownall and Emerson, 1992), and Polr3G is known to be associated with proliferation in cell culture studies and cancer(Durrieu-Gaillard et al., 2017; Haurie et al., 2010). It is therefore possible that during early embryo development and myogenic specification, Polr3G is regulated by Myf5 in addition to MyoD. Additionally, in *Xenopus*, Mrf4 has a very similar expression pattern to MyoD during tailbud stages and has been shown to have roles in both determination of myogenic lineage and in myogenic differentiation (Della Gaspera et al., 2012; Jennings, 1992; Kassar-Duchossoy et al., 2004). Therefore, transient expression of Polr3G in the somites at tailbud stages might be a result of combined regulation by MyoD and Mrf4.

Muscle-specific miRs have been identified in vertebrates and shown to have important roles in the regulation of myogenesis. miR-1, miR-133 and miR-206 are all induced by MRFs in chick and expressed in the somites during myogenesis (Sweetman et al., 2006; Sweetman et al., 2008). miR-206 has been shown to promote differentiation of muscle cells by downregulation of inhibitory factors such as MyoR, an antagonist of MyoD activity; Id3, an HLH factor predicted to antagonise bHLH protein function through interaction with E proteins and other factors associated with precursor cells such as Utrn (Kim et al., 2006; Rosenberg et al., 2006). In cell culture studies, Polr3G downregulation during differentiation of

hESCs has been shown to be under the control of microRNAs. miR-1305 has been shown to promote differentiation of all three germ layers in ESCs. Overexpression of miR-1305 resulted in upregulation of markers of all germ layers including the mesoderm markers Gata4 and Foxa1, whilst downregulation resulted in upregulation of Oct4 and Nanog (Jin et al., 2016). miR-1305 promotes differentiation through its downregulation of Polr3G and restoration of pluripotency can be achieved through overexpression of Polr3G or knockdown of miR-1305 indicating that interactions between small non-coding RNAs are important determinants of cell differentiation. As Polr3G downregulation by miR-1305 did not discriminately upregulate markers of any one particular germ layer, and miR-1305 is not known to be expressed in the muscle lineage, it is possible that regulation of Polr3G by cell type-specific miRs may be a determinant in many differentiation programmes. As muscle miRs are known, their regulation of Polr3G may be an important missing link between MRF activity and transcriptional regulation RNA Polymerase III. This is unlikely to be the mechanism that regulates the transient expression of Polr3G in the somites, as MiR binding sites are very specific (only miR-203 binding sites identified by TargetScan)(Agarwal et al., 2015) . However any, role for somite specific MiRs in regulating Polr3G may be an interesting avenue of investigation.

8.4. Post-transcriptional regulation of RNA Polymerase III activity

The central dogma of molecular biology is that the conversion of genetic information to protein output involves a two-step process: transcription and translation. Regulation of gene expression can occur at many stages during these processes and therefore, transcriptional output is by no means the full picture in studying gene expression. It is well established that RNA Polymerase III activity is regulated post-transcriptionally through disruption or promotion of protein-protein interactions between subunits and components of the transcriptional machinery, and through inactivation of factors through post-translational modifications such as phosphorylation.

This thesis studies the transcriptional regulation of RNA Polymerase III during muscle differentiation, and the possibility that myogenic factors play a role in mediating this activity. It has been shown that the MRF MyoD at least in part regulates expression of Polr3G and that myogenic signalling pathways regulate expression of both targets and subunits of RNA Polymerase III, however further study is required to determine the full mechanism of this regulation.

The proto-oncogene c-Myc is a bHLH transcription factor and a key regulator of RNA Polymerase III activity. c-Myc is an activator of RNA polymerase III activity and promotes transcription of tRNAs through recruitment of RNA Polymerase III, via interaction with TFIIIB, to target genes (Gomez-Roman et al., 2003). Previous studies have shown that downregulation of Polymerase III subunits Brf1 and Bdp1, and associated target genes during F9 cell differentiation are accompanied by downregulation of c-Myc protein expression (Athineos et al., 2010). Depletion of c-Myc is sufficient to inhibit RNA Polymerase III activity thus providing an explanation for this result (Felton-Edkins et al., 2003b). Phosphorylation of c-Myc is known to regulate its activity through protein destabilisation (Welcker et al.,

2004). More recently, P21 (RAC1)–activated kinase 2 (Pak2) mediated phosphorylation has also been shown to modulate c-Myc activity through its bHLH domain by phosphorylation of residue Ser-373 in the bHLH domain of c-Myc, inhibiting its interaction with its binding partner Max; this results in inhibition of Myc DNA binding at target genes (Macek et al., 2018). It may be that Fgf4 activity results in the phosphorylation of c-Myc in a residue crucial for its interaction with the TFIIIB complex which, whilst does not result in reduced expression of c-Myc, does result in reduced c-Myc activity and the downregulation of RNA Polymerase III target genes.

Binding of Fgf ligands to target receptors results in an intracellular signalling cascade marked by phosphorylation of Erk. Erk signalling has been shown to phosphorylate and negatively regulate the activity of the RNA Polymerase III repressor Maf1. Maf1 inhibits RNA Polymerase III transcription through the disruption of the interaction between the Polr3G subcomplex and other initiation factors of the RNA Polymerase III transcriptional machinery Brf1 and TBP (Vannini et al., 2010). Maf1 has recently been shown to promote differentiation of mouse embryonic stem cells to adipocytes (Chen et al., 2018). Inhibition of Maf1 in mESCs results in reduced differentiation of mesoderm and upregulated expression RNA Polymerase III target genes. In addition, targeting of Brf1 in mESCs resulted in enhanced adipogenesis, suggesting that downregulation of RNA Polymerase III mediated by Maf1 is important for differentiation. This downregulation of RNA Polymerase III target genes is consistent with the results in Chapter 7. In *Drosophila*, Ras/Erk signalling promotes cell proliferation rates by inhibiting nuclear localisation of Maf1 to enhance tRNA synthesis in both stem cells and epithelial lineages (Sriskanthadevan-Pirahas et al., 2018). However, as Erk signalling when activated through Fgf4, in the context of mesoderm induction,

promotes cells to differentiate into muscle, and this results in global downregulation of RNA Polymerase III transcriptional machinery and target genes. It is therefore more likely that induction of muscle differentiation in animal caps by Fgf4/Wnt8 resulted in the upregulation of Maf1 activity through an indirect mediator, or possibly through phosphorylation of Maf1 at a different position to enhance its repressor activity. Maf1 expression was shown to not be increased in animal caps induced by Fgf4/Wnt8, so transcriptional regulation of this factor in animal caps is unlikely to be the mechanism of repression of RNA Polymerase III activity in contrast with previous studies. However, post-transcriptional activity, particularly the nuclear localisation of Maf1, may still be an interesting avenue of investigation for future study.

8.5. Can a potential role for Polr3G in muscle cell progenitors be translated to satellite cells?

Polr3G is expressed in undifferentiated stem cells and, during *Xenopus* development, is present at highest levels in the earliest few stages prior to activation of transcription and induction of tissue differentiation programmes. However, a role for Polr3G and its expression in populations of adult stem cells has not yet been characterised. During development, mesodermal cells become determined to the myogenic lineage through embryonic signals such as Fgf and Wnt and differentiate into proliferative myoblasts expressing MyoD and Myf5 (Emerson, 1990; Pownall and Emerson, 1992). Subsequently, through activation of a myogenic transcriptional programme directed by MRFs, myoblasts exit the cell cycle and differentiate to form myotubes and myofibers.

Muscle is stable throughout embryogenesis and into adulthood. It is maintained in response to natural cell turnover resulting from daily use but also has the capacity to regenerate in response to injury (i.e. from mechanical strain) due to the presence of satellite cells within the organisation of myofibers (Schmalbruch and Lewis, 2000; Yin et al., 2013). Satellite cells are so named because they are located at the periphery of myofibers. They are mononucleated and under normal conditions, non-dividing cells. However, they retain the capacity to rapidly re-enter the cell cycle after injury, forming proliferative myoblasts (Yin et al., 2013). Satellite cells can both self-renew through asymmetric cell division (and can divide symmetrically too) (Kuang et al., 2007), and can differentiate into functional myoblasts and myofibers. These characteristics define them as a population of muscle-specific adult stem cells. Pax7 is the conserved marker of satellite cell populations and is required for their specification (Seale et al., 2000) and most quiescent satellite cells also express Myf5 (Beauchamp et al., 2000). However, subsets of satellite cells do not express Myf5, and 25% of satellite cell populations are also positive for expression of MyoD. In addition, $MyoD^{-/-}$ cells show reduced differentiation capacity (Cornelison and Wold, 1997; Megeney et al., 1996; Sabourin et al., 1999; Wood et al., 2013) indicating that, as with embryonic myogenesis, MRFs have essential roles in regulation of adult muscle maintenance and induced myogenesis of muscle specific-stem cells.

Day-to-day turnover of muscle cells can be repaired without cell death or inflammatory signalling responses. However, large injuries inflicted by trauma or as a result of genetic mutation result in necrosis of myofibers and inflammatory signalling, leading to activation, proliferation and subsequent differentiation of satellite cells. Satellite cells are essential for myogenic regeneration, as ablation of all Pax7+ cells results in total loss of regeneration in response to injury (Lepper et

al., 2011). Satellite cells have also been shown to have a role in regeneration of *Xenopus* tadpole tail muscle (Chen et al. 2006) and, as a side note, satellite cells were originally identified in frogs (Mauro, 1961).

Importantly, adult tissues lose their regenerative capacity and ageing muscle is less able to repair after injury, either day-to-day or more traumatic. The ageing population is at increased risk of muscle atrophy and the degeneration of skeletal muscle (Sarcopenia) (Walston, 2012), which is correlated to a progressive loss in satellite cell function and reduced numbers of $Pax7^{+}$ cells (Alway et al., 2014). Moreover, satellite cells are defective in some patients with muscular dystrophy, where these cells lose their proliferative capacity, and due to a lack of defined cell polarity, these cells lose their capacity to self-renew, thus depleting satellite cell populations (Blau et al., 1983; Chang et al., 2016) . Additionally, stem cell "exhaustion" of satellite cells in mice, caused by defective telomerase activity and resulting telomere shortening through successive rounds of degeneration and regeneration, resulted in Muscular Dystrophy phenotypes similar to that shown in human patients (Sacco et al., 2010). Thus far, effective treatment of these conditions has been evasive, but enhancement of satellite cell populations could be a promising therapeutic option.

Knockdown of Polr3G in ESCs results in upregulation of many differentiation markers (Wong et al., 2011), including mesoderm marks Brachyury and Eomes. As this thesis has revealed that Polr3G is transiently expressed in myoblasts during early stages of myogenic differentiation, it is also possible that Polr3G is expressed in satellite cell populations along with Pax7 and either MyoD or Myf5. In addition, Polr3G interacts with telomerase in transformed cells and enhances its recruitment to and activation of target genes (Khattar et al., 2016). Dysfunction of telomerase and shortened telomeres have been identified in

muscular dystrophy patients, resulting in non-functional satellite cell populations (Sacco et al., 2010; Tichy et al., 2017). The shortening of telomeres results in the inability of satellite cells to replenish damaged muscle fibres as they become less able to self-renew and replenish their population in ageing muscle. Another common problem with ageing satellite cells is the disruption of asymmetric cell division meaning that the daughter cells produced are then both committed, and can no longer act to replenish the reservoir of muscle stem cells (Blau et al., 2015). The role of RNA Polymerase III in satellite cells has not been investigated, however, Polr3G overexpression in ESCs results in increased expression of pluripotency markers and decreased expression of lineage commitment marker genes in response to differentiation by Retinoic Acid (Wong et al., 2011). Polr3G, if overexpressed in ageing satellite cell populations, may result in resistance to lineage commitment of stem cell populations, meaning that reservoirs can replenish muscle fibres for longer than normal. Moreover, Polr3G and its interaction with TERT is implicated in the increased cell proliferation observed in cancers, and is predicted to be due to the enhanced activation of tRNAs and other RNA Polymerase III targets observed in resulting cell populations (Khattar et al., 2016). As telomerase dysfunction is implicated in satellite cell “exhaustion”, and the inability to replenish populations of stem cells in muscle, it is possible therefore that expansion of satellite cell populations in patients with muscular dystrophy, or age-related muscle degeneration, may be achieved through overexpression of Polr3G and TERT in combination. These two factors would tackle both reduced population replenishment and reduced ‘stemness’ in satellite cells in combination.

8.6. Future Work

8.6.1. Gene targeting of Polr3G/Polr3gL

The targeting of Polr3G by antisense morpholinos was not possible in this thesis due to the high level of maternal RNA and protein stored in the embryo. Nevertheless, knockdown study is a crucial component of gene functional characterisation and future study would benefit from targeting using CRISPR/Cas9 and out-crossing offspring to form heterozygous and homozygous populations. In ESCs, Polr3G was found to be non-essential for cell growth and viability, and was unable to compensate for the loss of Polr3gL in siRNA knockdown lines suggesting that Polr3gL is essential in ESCs for survival (Haurie et al., 2010). However, as Polr3G is expressed in oogenesis and cleavage stages, this may not be possible as this subunit may be essential for viability in the early embryo. Genetic compensation has been shown to be induced in CRISPR/Cas9 mutagenesis studies (Rossi et al., 2015), whereby genes of similar function to the target gene are upregulated as a result of mutagenesis, and reduce the effects of the knockdown. It is not known exactly how this mechanism of compensation works, but it is thought to involve the detection of products formed from nonsense-mediated decay of mutated mRNA transcripts. Nonsense-mediated decay is facilitated by mRNA surveillance genes which detect premature stop codons upstream of exon-exon junctions. A recent study has shown that mutating the decay factor Upf1 results in decreased genetic compensation in CRISPR/Cas9 mutants (El-Brolosy et al., 2018). It is therefore possible that CRISPR/Cas9 based targeting of Polr3G would upregulate expression of Polr3gL in order to compensate; this kind of redundancy between related factors is seen in MyoD$^{-/-}$

mutant mice where the increase in Myf5 expression allows skeletal muscle to develop despite the absence of MyoD (Rudnicki et al., 1992; Rudnicki et al., 1993).

El-Brolosy also showed that mutations resulting in the complete inhibition of transcription of the target gene results in reduced genetic compensation and enhancement of mutant phenotypes (El-Brolosy et al., 2018). Therefore, targeting the promoter regions of target genes could overcome this artefact. As analysis of tRNA modulation in response to Polr3gL overexpression at MBT stages was not carried out in this thesis, it is yet to be determined whether both subunits regulate the same target genes as seen in the NF Stage 25 analysis carried out in this thesis.

In addition to full characterisation of Polr3G, which might not be possible, Polr3gL could be targeting using morpholinos due to its expression post-MBT. This would be a more immediate way to determine whether Polr3G and Polr3gL can compensate for each other *in vivo* during development as mRNA rescue experiments can be easily coupled with morpholinos targeting in *Xenopus*. As Polr3gL is induced by Fgf4 and Wnt8 in animal caps, it is also possible that Polr3gL has a role in differentiation of myoblasts. Therefore, if morpholino targeting resulted in loss of myogenic differentiation, it would reveal a subunit of RNA Polymerase III with a role in myogenesis with a unique repertoire of RNA polymerase III (and possibly RNA Polymerase II) target genes.

8.6.2. Analysis of Regulation of Polr3G/Polr3gL by other myogenic factors

Extended analysis of regulation of Polr3G by other MRFs during the early stages of muscle differentiation may determine whether RNA Polymerase III is also regulated independently of MyoD (Myf5, Mrf4). Myf5, whilst the earliest expressed myogenic factor in *Xenopus* is expressed in a more limited number of cells even from gastrula stages. The localised expression to the dorsal-lateral marginal zone at gastrula stages is followed by expression in only the outer regions of somites at tailbud stages (Hopwood et al., 1991; Maguire et al., 2012). Whilst Polr3G is not expressed in a similar pattern to Myf5 at these stages, it is possible that MyoD and other MRFs co-regulate both Polr3G and Polr3gL and that this regulation occurs at earlier developmental stages. An existing model in cell culture studies predicts that Myf5 acts earlier in myogenesis to remodel chromatin through histone modifications, and that MyoD recruits RNA Polymerase II to target genes and activates transcription (Conerly et al., 2016). Therefore, co-expression of MyoD with Myf5 may result in greater activation of Polr3G than MyoD achieves alone. The role of Mrf4 is not as clearly defined as Myf5 and MyoD as it can act as both a specification factor, compensating for the loss of Myf5 (Kassar-Duchossoy et al., 2004) and as a differentiation factor alongside Myogenin (Chanoine et al., 2004; Pownall et al., 2002). Its expression in the somites at tailbud stages of *Xenopus* development overlaps with MyoD (Della Gaspera et al., 2012) which could mean they act together in regulating the same target genes. Polr3G expression is located in the same regions as both MyoD and Mrf4. If the expression of Mrf4 acts as a molecular switch to convert determined myoblasts to differentiation, it may be that factors such as Mrf4 and Myogenin act as negative regulators of Polr3G, although MRFs are widely accepted to be transcriptional activators, so this regulation would most likely be indirect.

8.6.3. Post-transcriptional regulation of RNA Polymerase III subunits, cofactors, regulators

Much regulation of RNA Polymerase III activity occurs post-transcriptionally. To fully characterise the regulation of RNA Polymerase III during Fgf4 and Wnt8 induced myogenesis, protein based analyses should be carried out to determine the post-transcriptional effects in addition to the results observed in the microarray in Chapter 7. Protein levels of Brf1 and Bdp1 are reduced during F9 cell differentiation and transcriptional downregulation of these factors is also observed in Fgf and Wnt induced animal caps. Determining the result on protein expression would fully characterise this observation in caps to determine a conserved regulation of RNA Polymerase III during different differentiation programmes. Moreover, analysis of post-translational modifications may also determine the mechanism by which RNA Polymerase III activity is modulated during myogenic differentiation. As no transcriptional downregulation of the positive Polymerase III regulator c-Myc is observed, it may be that regulation of its expression or activity is due to protein targeting or through disruption of function (Macek et al., 2018; Salghetti et al., 1999). Phosphorylation of c-Myc as a result of Fgf4 and Wnt8 signalling may determine the downregulation of RNA Polymerase III activity observed.

In addition, as no clear negative regulator of RNA Polymerase III was identified by transcriptomic analyses to explain the downregulation of both RNA Polymerase III targets and transcription factors, it may be that expression of regulators such as Rb are higher in Fgf4/Wnt8 induced caps. This would be detectable by western blot analysis.

8.7. Conclusions

This discussion presents evidence that the known MyoD transcriptional network is becoming more complex with the improvements to whole-genome sequencing and transcriptomic analyses, and that this network is not limited to protein coding mRNAs, but also includes epigenetic modifications and regulation of non-coding RNAs (MicroRNAs). Evidence of the regulation of Polr3G by MyoD in *Xenopus* also indicates that RNA Polymerase III transcriptional activity may also be a node in this network and the role of Polr3G in promotion of pluripotency may also have a role in regulation of adult stem cells. With a potential role in myogenic progenitors, overexpression of Polr3G and Telomerase in satellite cell populations may act as a novel therapeutic strategy for treatment of age-related muscle degeneration and muscular dystrophy.

Furthermore, this thesis has suggested that the downregulation of RNA Polymerase III activity appears to be conserved between cell lines and in animal cap explants, with the induction of myogenic differentiation resulting in global downregulation of RNA Polymerase III target genes. This thesis also suggests that, in addition to the subunits Brf1 and Bdp1, most other transcription factors of RNA Polymerase III are also downregulated during myogenic differentiation of animal caps, with the exception of Polr3gL. In contrast to findings in cell culture, where this effect is attributed to reduced proliferation, the animal cap forms 3D tissues with multiple cell types, where induction of myogenic differentiation may first lead to the determination of proliferative myoblasts. Therefore the downregulation of RNA Polymerase III transcription may be an effect of post-translational modifications to key RNA Polymerase III regulators, requiring further investigation.

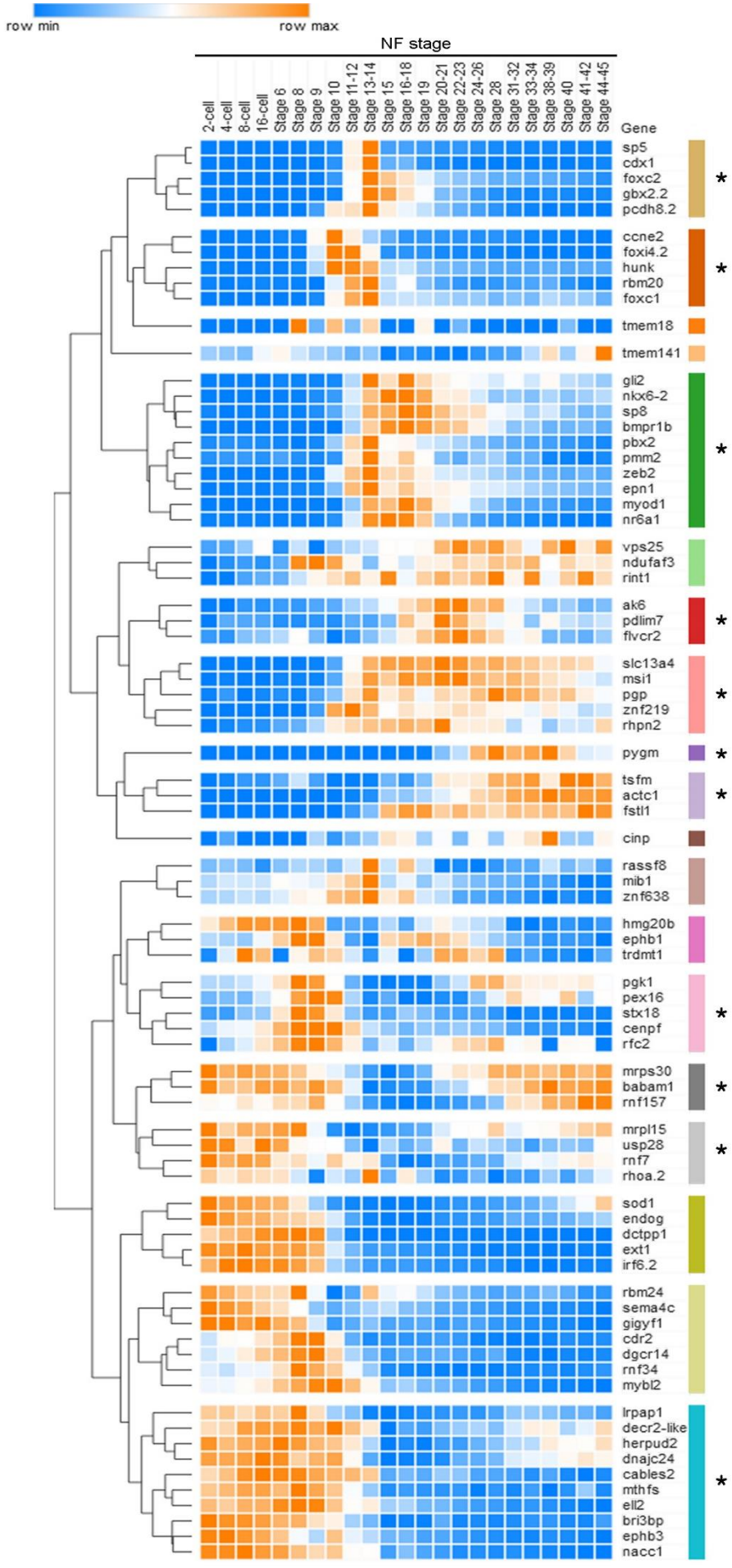
Appendix

Gene symbol	ENSEMBL ID	mean control	mean experimental	fold change	p-value
bmpr1b	ENSXETG00000019220	8.01	5.45	0.68	0.02
myod1	ENSXETG00000001320	90.10	64.29	0.71	0.02
ndufaf3	ENSXETG00000024755	15.31	11.02	0.72	0.03
tmem141	ENSXETG00000003949	8.53	6.25	0.73	0.04
pgp	ENSXETG00000016097	6.12	4.52	0.74	0.03
trdmt1	ENSXETG00000027671	5.01	3.72	0.74	0.04
rassf8	ENSXETG00000023490	11.13	8.34	0.75	0.04
gbx2.2	ENSXETG00000003293	42.10	31.57	0.75	0.01
nodal		6.12	4.62	0.76	0.03
sp8	ENSXETG00000030115	12.27	9.28	0.76	0.05
ndufb4		16.41	12.42	0.76	0.01
nkx6-2	ENSXETG00000023614	20.42	15.51	0.76	0.02
tsfm	ENSXETG00000009653	5.08	3.96	0.78	0.05
decr2-like	ENSXETG00000010329	12.93	10.08	0.78	0.04
rbm20	ENSXETG00000025245	7.04	5.51	0.78	0.04
zeb2	ENSXETG00000000237	18.67	14.88	0.80	0.01
mespa		9.87	7.87	0.80	0.05
znf638	ENSXETG00000015780	6.74	5.38	0.80	0.02
sod1	ENSXETG00000007350	28.27	22.64	0.80	0.05
foxc2	ENSXETG00000016387	80.68	65.42	0.81	0.03
babam1	ENSXETG00000025571	9.14	7.47	0.82	0.02
gli2	ENSXETG00000011189	12.23	10.04	0.82	0.01
arl10		9.70	7.98	0.82	0.01
ccne2	ENSXETG00000006660	72.36	59.78	0.83	0.02
mrpl15	ENSXETG00000008103	12.68	10.49	0.83	0.05
sp5	ENSXETG00000025407	64.04	53.11	0.83	0.03
lrrn1-like.1		6.57	5.46	0.83	0.05
mthfs	ENSXETG00000024293	17.77	14.80	0.83	0.01
pmm2	ENSXETG00000004549	45.85	38.21	0.83	0.04
c4orf32		26.81	22.39	0.83	0.04
endog	ENSXETG00000025614	9.73	8.14	0.84	0.04
pygm	ENSXETG00000034136	123.57	103.71	0.84	0.04
foxc1	ENSXETG00000000594	73.17	61.58	0.84	0.06
sowahc		5.89	4.96	0.84	0.01
fstl1	ENSXETG00000018009	25.29	21.31	0.84	0.02
sema4c	ENSXETG00000001251	22.06	18.66	0.85	0.04
pex16	ENSXETG00000001027	8.44	7.15	0.85	0.05
fahd1-like		8.00	6.79	0.85	0.05
ephb1	ENSXETG00000013722	7.89	6.71	0.85	0.03
slc13a4	ENSXETG00000008163	18.80	15.99	0.85	0.03
AK6	ENSXETG00000018174	14.39	12.24	0.85	0.03

rbm24	ENSXETG00000024618	30.00	25.56	0.85	0.03
tmem18	ENSXETG00000027985	42.69	36.39	0.85	0.01
hunk	ENSXETG00000007352	41.66	35.63	0.86	0.00
cdr2	ENSXETG00000016662	6.64	5.69	0.86	0.01
dctpp1	ENSXETG00000018900	12.09	10.36	0.86	0.03
pgk1	ENSXETG00000007447	19.96	17.15	0.86	0.02
mrps30	ENSXETG00000017716	11.89	10.23	0.86	0.04
dgcr14	ENSXETG00000022387	6.43	5.54	0.86	0.02
jam3		5.05	4.36	0.86	0.05
stx18	ENSXETG00000016051	14.61	12.60	0.86	0.02
cables2	ENSXETG00000002013	24.42	21.07	0.86	0.00
zbtb8a.1		5.58	4.82	0.86	0.04
cinp	ENSXETG00000010255	7.06	6.10	0.86	0.03
bicd2-like.1		38.26	33.10	0.87	0.04
flvcr2	ENSXETG00000027282	8.63	7.47	0.87	0.04
cdx1	ENSXETG00000010282	77.18	66.87	0.87	0.03
vps25	ENSXETG00000024599	15.21	13.19	0.87	0.05
mib1	ENSXETG00000003146	14.11	12.25	0.87	0.03
sept6-like		12.70	11.03	0.87	0.02
pdlim7	ENSXETG00000007240	13.77	11.97	0.87	0.01
msi1	ENSXETG00000012216	55.14	47.91	0.87	0.00
znf219	ENSXETG00000016157	19.87	17.28	0.87	0.05
rnf7	ENSXETG00000014753	99.16	86.80	0.88	0.04
rint1	ENSXETG00000023226	9.88	8.66	0.88	0.02
hmg20b	ENSXETG00000022092	12.48	10.94	0.88	0.04
pcdh8.2	ENSXETG00000008792	73.31	64.29	0.88	0.01
lrpap1	ENSXETG00000005500	11.86	10.41	0.88	0.02
ext1	ENSXETG00000019136	47.42	41.62	0.88	0.03
wipf2-like		24.34	21.36	0.88	0.02
nr6a1	ENSXETG00000008578	71.16	62.60	0.88	0.05
herpud2	ENSXETG00000013111	5.58	4.91	0.88	0.04
plekhg4		10.65	9.42	0.88	0.04
cenpf	ENSXETG00000023124	34.80	30.79	0.89	0.00
rnf34	ENSXETG00000020965	7.57	6.71	0.89	0.04
dhx32-like		17.74	15.73	0.89	0.05
bri3bp	ENSXETG00000001207	16.21	14.38	0.89	0.01
ggnbp2		24.95	22.16	0.89	0.05
ell2	ENSXETG00000013296	18.61	16.58	0.89	0.05
vdac3		17.56	15.66	0.89	0.05
ephb3	ENSXETG00000017293	42.52	37.95	0.89	0.02
rhog-like.1		14.00	12.52	0.89	0.05
pttg1ip.2		5.42	4.86	0.90	0.05
rfc2	ENSXETG00000018234	11.70	10.49	0.90	0.01
rhoa.2	ENSXETG00000009241	445.54	399.57	0.90	0.02

rnf157	ENSXETG00000019548	5.75	5.16	0.90	0.05
prr13-like		65.55	58.85	0.90	0.04
mybl2	ENSXETG00000012125	51.14	45.92	0.90	0.01
gigyf1	ENSXETG00000018415	23.73	21.32	0.90	0.03
ilvbl		6.74	6.07	0.90	0.05
epn1	ENSXETG00000022662	42.43	38.19	0.90	0.04
nacc1	ENSXETG00000005594	44.47	40.05	0.90	0.03
irf6.2	ENSXETG00000018661	13.55	12.22	0.90	0.02
rhpn2	ENSXETG00000018271	9.65	8.71	0.90	0.02
fmnl3-like		15.30	13.80	0.90	0.02
slc30a1		26.69	24.09	0.90	0.04
dnajc24	ENSXETG00000008179	14.14	12.76	0.90	0.03
usp28	ENSXETG00000022958	10.14	9.16	0.90	0.03
nsd1-like		7.39	6.68	0.90	0.02
cdc42se2-like.1		6.09	5.51	0.90	0.01
pbx2	ENSXETG00000005223	169.56	158.17	0.93	0.02
actc1	ENSXETG00000012911	-	-	-	-

Appendix Table 1. **100 shortlisted genes significantly downregulated in embryos targeted with MyoD gRNA.** After paired t-test analysis, 1165 genes were identified as showing significant differences in expression in Cas9 only controls and MyoD gRNA targeted samples. Of these genes, 100 were shortlisted as potential target genes of MyoD showing a fold change of 0.68-0.91 and an average control FPKM of >5. In addition, foxc1 and pbx2 were manually added to the list despite not meeting criteria cutoffs due to their implicated roles as MyoD interactants/targets. As another reference gene for expression profiles, the MyoD target actc1 was also included in further analysis.


row min
row max
NF stage
2-cell
4-cell
8-cell
16-cell
Stage 6
Stage 8
Stage 9
Stage 10
Stage 11-12
Stage 13-14
Stage 15
Stage 16-18
Stage 19
Stage 20-21
Stage 22-23
Stage 24-26
Stage 28
Stage 31-32
Stage 33-34
Stage 38-39
Stage 40
Stage 41-42
Stage 44-45
Gene
sp5
cdx1
foxc2
gbx2.2
pcdh8.2
ccne2
foxi4.2
hunk
rbm20
foxc1
tmem18
tmem141
gli2
nkx6-2
sp8
bmpr1b
pbx2
pmm2
zeb2
epn1
myod1
nr6a1
vps25
ndufaf3
rint1
ak6
pdlim7
flvcr2
slc13a4
msi1
pgp
znf219
rhpn2
pygm
tsfm
actc1
fstl1
cinp
rassf8
mib1
znf638
hmg20b
ephb1
trdmt1
pgk1
pex16
stx18
cenpf
rfc2
mrps30
babam1
rnf157
mrpl15
usp28
rnf7
rhoa.2
sod1
endog
dctpp1
ext1
irf6.2
rbm24
sema4c
gigyf1
cdr2
dgcr14
rnf34
mybl2
lrpap1
decr2-like
herpud2
dnajc24
cables2
mthfs
ell2
bri3bp
ephb3
nacc1

Appendix Figure 1. Hierarchical clustering using Euclidean distance of developmental expression profiles of 73 target genes shortlisted from the initial RNA-Seq results filtering. Genes included are those which had ENSEMBL IDs and were included in the dataset from (Tan et al., 2013). Myod1 and actc1 were also included as reference genes. Clusters were manually assessed to shortlist developmentally relevant clusters for further analysis (*).

Gene name	Control_1/ Polr3G_1	Control_2 /Polr3G_2	Control_3/ Polr3G_3	Average FC	p-value
bmp4	1.543689	1.416169	1.456241	1.472033	0.0026
bmp7.2	1.165607	1.081415	1.140376	1.129133	0.0121
chrd	1.771032	2.130559	2.786807	2.229466	0.0360
eomes	1.481903	1.919614	1.780796	1.727438	0.0369
foxd5	1.420048	2.049841	2.462223	1.977371	0.0058
fzd1	1.127759	0.950777	1.234178	1.104238	0.1810
gsc	1.484293	1.613645	1.644941	1.58096	0.0212
id3	1.945837	1.761897	1.827762	1.845165	0.0098
lefty	1.594944	1.344384	1.47802	1.472449	0.0090
nodal	1.630057	1.813672	1.491113	1.644947	0.0063
nodal 2a like	1.316108	1.101156	1.122252	1.179839	0.0283
nodal3a	1.293065	1.14928	1.359913	1.267419	0.0231
nodal3c	1.322936	1.17073	1.333059	1.275575	0.0141
nog	1.445232	1.286689	1.667153	1.466358	0.0334
oct4	1.032665	1.211562	1.241664	1.235147	0.0898
polr3G	777.7382	324.8658	373.2499	491.9513	0.0055
sox17	1.271671	1.047167	1.498247	1.272362	0.1138
sox17b.1	1.356134	1.126055	1.178551	1.220247	0.0044
sox17b.2	1.376717	1.095506	1.164169	1.21213	0.0143
wnt8a	1.628576	1.768179	1.823384	1.740046	0.0152

Appendix Table 2. Curated genelist of mRNAs with significantly upregulated expression in embryos overexpressing Polr3G at NF stage 9.

GeneName	Average Control	Average Fgf4/Wnt8	Fold change
AlaCGC	1.118668	0.649607	0.580501
AlaTGC	1.030837	0.593304	0.571457
ArgCCT	1.108241	0.659931	0.595372
AsnGTT	1.135748	0.67515	0.591122
CysACA	1.351646	0.715613	0.532681
CysGCA	1.361717	0.768962	0.564647
GlnCTG	1.309711	0.618362	0.474755
GlnTTG	1.110973	0.681004	0.612963
GlnTTG	1.30932	0.564803	0.431448
GluCTC	1.110934	0.691623	0.623652
GluTTC	1.309936	0.76054	0.580613
GlyCCC	1.055806	0.717104	0.670402
GlyGCC	1.109038	0.629752	0.567774
GlyTCC	1.098989	0.631297	0.573973
IleAAT	1.105668	0.692492	0.626242
IleTAT	1.323536	0.643281	0.486073
LeuAAG	1.253989	0.674589	0.537955
LeuCAA	1.796093	0.950372	0.525164
LeuTAA	1.437447	0.786479	0.552508
LysTTT	1.067345	0.701763	0.654048
PheAAA	1.309667	0.753875	0.576305
PheGAA	1.322102	0.758373	0.57363
SerAGA	1.054232	0.651607	0.616673
SerCGA	1.095853	0.705067	0.643278
SerGCT	1.080569	0.617578	0.577136
SerTGA	1.076661	0.634404	0.589072
SerTGA	1.036689	0.648076	0.62629
ThrTGT	1.093908	0.704634	0.644398
TrpCCA	1.251981	0.730306	0.579566
TyrGTA	1.046715	0.764457	0.729131
TyrGTA	1.175729	0.65707	0.559161
ValAAC	1.229314	0.665965	0.541949
ValCAC	1.271793	0.715918	0.563035
ValTAC	0.91455	0.834654	0.912627

Appendix Table 3. Fold changes induced by Fgf4/Wnt8 in animal caps for all tRNA isoacceptors with expression values of >10 included on the custom Agilent microarray.

References

Agarwal, V., Bell, G. W., Nam, J. and Bartel, D.P. (2015). Predicting effective microRNA target sites in mammalian mRNAs. *eLife.* **4**:e05005.

Aguero, T. H., Fernandez, J. P., Vega Lopez, G. A., Tribulo, C. and Aybar, M. J. (2012). Indian hedgehog signaling is required for proper formation, maintenance and migration of Xenopus neural crest. *Dev. Biol.* **364**, 99–113.

Albini, S., Coutinho, P., Malecova, B., Giordani, L., Savchenko, A., Forcales, S. V. and Puri, P. L. (2013). Epigenetic Reprogramming of Human Embryonic Stem Cells into Skeletal Muscle Cells and Generation of Contractile Myospheres. *Cell Rep.* **3**, 661–670.

Alway, S. E., Myers, M. J. and Mohamed, J. S. (2014). Regulation of Satellite Cell Function in Sarcopenia. *Front. Aging Neurosci.* **6**,.

Appaiah, H. N., Goswami, C. P., Mina, L. A., Badve, S., Sledge, G. W., Liu, Y. and Nakshatri, H. (2011). Persistent upregulation of U6:SNORD44 small RNA ratio in the serum of breast cancer patients. *Breast Cancer Res.* **13**, R86.

Archer, T. C., Jin, J. and Casey, E. S. (2011). Interaction of Sox1, Sox2, Sox3 and Oct4 during primary neurogenesis. *Dev. Biol.* **350**, 429–440.

Athineos, D., Marshall, L. and White, R. J. (2010). Regulation of TFIIIB during F9 cell differentiation. *BMC Mol. Biol.* **11**, 2–9.

Beauchamp, J. R., Heslop, L., Yu, D. S., Tajbakhsh, S., Kelly, R. G., Wernig, A., Buckingham, M. E., Partridge, T. A. and Zammit, P. S. (2000). Expression of CD34 and Myf5 defines the majority of quiescent adult skeletal muscle satellite cells. *J. Cell Biol.* **151**, 1221–34.

Behura, S. K. and Severson, D. W. (2013). Codon usage bias: causative factors, quantification methods and genome-wide patterns: with emphasis on insect genomes. *Biol. Rev.* **88**, 49–61.

Benezra, R., Davis, R. L., Lockshon, D., Turner, D. L. and Weintraub, H. (1990). The protein Id: A negative regulator of helix-loop-helix DNA binding

proteins. *Cell* **61**, 49–59.

Bergstrom, D. A. and Tapscott, S. J. (2001). Molecular distinction between specification and differentiation in the myogenic basic helix-loop-helix transcription factor family. *Mol. Cell. Biol.* **21**, 2404–12.

Bergstrom, D. A., Penn, B. H., Strand, A., Perry, R. L. S., Rudnicki, M. A. and Tapscott, S. J. (2002). Promoter-specific regulation of MyoD binding and signal transduction cooperate to pattern gene expression. *Mol. Cell* **9**, 587–600.

Berkes, C. A. and Tapscott, S. J. (2005). MyoD and the transcriptional control of myogenesis. *Semin. Cell Dev. Biol.* **16**, 585–95.

Berkes, C. A., Bergstrom, D. A., Penn, B. H., Seaver, K. J., Knoepfler, P. S. and Tapscott, S. J. (2004). Pbx Marks Genes for Activation by MyoD Indicating a Role for a Homeodomain Protein in Establishing Myogenic Potential. *Mol. Cell* **14**, 465–477.

Berti, F., Nogueira, J. M., Wöhrle, S., Sobreira, D. R., Hawrot, K. and Dietrich, S. (2015). Time course and side-by-side analysis of mesodermal, pre-myogenic, myogenic and differentiated cell markers in the chicken model for skeletal muscle formation. *J. Anat.* **227**, 361–382.

Bischoff, J. R., Anderson, L., Zhu, Y., Mossie, K., Ng, L., Souza, B., Schryver, B., Flanagan, P., Clairvoyant, F., Ginther, C., et al. (1998). A homologue of Drosophila aurora kinase is oncogenic and amplified in human colorectal cancers. *EMBO J.* **17**, 3052–3065.

Blais, A., Tsikitis, M., Acosta-Alvear, D., Sharan, R., Kluger, Y. and Dynlacht, B. D. (2005). An initial blueprint for myogenic differentiation. *Genes Dev.* **19**, 553–69.

Blau, H. M., Webster, C. and Pavlath, G. K. (1983). Defective myoblasts identified in Duchenne muscular dystrophy. *Proc. Natl. Acad. Sci. U. S. A.* **80**, 4856–60.

Blau, H. M., Cosgrove, B. D. and Ho, A. T. V (2015). The central role of muscle stem cells in regenerative failure with aging. *Nat. Med.* **21**, 854–862.

Blitz, I. L., Biesinger, J., Xie, X. and Cho, K. W. Y. (2013). Biallelic genome modification in F(0) Xenopus tropicalis embryos using the CRISPR/Cas system. *Genesis* **51**, 827–34.

Boissier, F., Dumay-Odelot, H., Teichmann, M. and Fribourg, S. (2015). Structural analysis of human RPC32β - RPC62 complex. *J. Struct. Biol.* **192**, 313–319.

Borycki, A. G., Brunk, B., Tajbakhsh, S., Buckingham, M., Chiang, C. and Emerson, C. P. (1999). Sonic hedgehog controls epaxial muscle determination through Myf5 activation. *Development* **126**, 4053–63.

Branney, P. A., Faas, L., Steane, S. E., Pownall, M. E. and Isaacs, H. V (2009). Characterisation of the Fibroblast Growth Factor Dependent Transcriptome in Early Development. *PLoS One* **4**, 1–17.

Briggs, J. A., Weinreb, C., Wagner, D. E., Megason, S., Peshkin, L., Kirschner, M. W. and Klein, A. M. (2018). The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution. *Science* **360**, eaar5780.

Bruce, A. E. E., Howley, C., Zhou, Y., Vickers, S. L., Silver, L. M., King, M. Lou and Ho, R. K. (2003). The maternally expressed zebrafish T-box gene eomesodermin regulates organizer formation. *Development* **130**, 5503–17.

Burks, P. J., Isaacs, H. V. and Pownall, M. E. (2009). FGF signalling modulates transcriptional repression by Xenopus groucho-related-4. *Biol. Cell* **101**, 301–308.

Cabarcas, S., Jacob, J., Veras, I. and Schramm, L. (2008). Differential expression of the TFIIIB subunits Brf1 and Brf2 in cancer cells. *BMC Mol. Biol.* **9**, 74.

Caldon, C. and Musgrove, E. A. (2010). Distinct and redundant functions of cyclin E1 and cyclin E2 in development and cancer. *Cell Div.* **5**,.

Cao, Y., Yao, Z., Sarkar, D., Lawrence, M., Sanchez, G. J., Parker, M. H., MacQuarrie, K. L., Davison, J., Morgan, M. T., Ruzzo, W. L., et al. (2010). Genome-wide MyoD binding in skeletal muscle cells: a potential for broad cellular reprogramming. *Dev. Cell* **18**, 662–74.

Cardinali, B., Cappella, M., Provenzano, C., Garcia-Manteiga, J. M., Lazarevic, D., Cittaro, D., Martelli, F. and Falcone, G. (2016). MicroRNA-222 regulates muscle alternative splicing through Rbm24 during differentiation of skeletal muscle cells. *Cell Death Dis.* **7**, e2086.

Carruthers, S., Mason, J. and Papalopulu, N. (2003). Depletion of the cell-cycle inhibitor p27Xic1 impairs neuronal differentiation and increases the number of ElrC+ progenitor cells in Xenopus tropicalis. *Mech. Dev.* **120**, 607–616.

Chal, J. J. and Pourquié1, O. P. (2017). Making muscle: skeletal myogenesis in vivo and in vitro. *Development* **15**, 2104–2122.

Chan, P. P. and Lowe, T. M. (2016). GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Res.* **44**, D184–D189.

Chan, P.P. & Lowe, T. M. (2009). GtRNAdb: A database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Res.* **37**, 93–97.

Chang, N. C., Chevalier, F. P. and Rudnicki, M. A. (2016). Satellite Cells in Muscular Dystrophy - Lost in Polarity. *Trends Mol. Med.* **22**, 479–496.

Chanoine, C., Della Gaspera, B. and Charbonnier, F. (2004). Myogenic regulatory factors: redundant or specific functions? Lessons from Xenopus. *Dev. Dyn.* **231**, 662–70.

Chen, J.-F., Mandel, E. M., Thomson, J. M., Wu, Q., Callis, T. E., Hammond, S. M., Conlon, F. L. and Wang, D.-Z. (2006). The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nat. Genet.* **38**, 228–233.

Chen, C. Y., Lanz, R. B., Walkey, C. J., Chang, W. H., Lu, W. and Johnson, D. L. (2018). Maf1 and Repression of RNA Polymerase III-Mediated Transcription Drive Adipocyte Differentiation. *Cell Rep.* **24**, 1852–1864.

Choi, J., Costa, M. L., Mermelstein, C. S., Chagas, C., Holtzer, S. and Holtzer, H. (1990). MyoD converts primary dermal fibroblasts, chondroblasts, smooth muscle, and retinal pigmented epithelial cells into striated mononucleated myoblasts and multinucleated myotubes. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 7988–7992.

Christian, J. L. and Moon, R. T. (1993). Interactions between Xwnt-8 and Spemann organizer signaling pathways generate dorsoventral pattern in the embryonic mesoderm of Xenopus. *Genes Dev.* **7**, 13–28.

Christian, J. L., Olson, D. J. and Moon, R. T. (1992). Xwnt-8 modifies the character of mesoderm induced by bFGF in isolated Xenopus ectoderm. *EMBO J.* **1**, 33–41.

Clarkson, S. G., Birnstiel, M. L. and Purdom, I. F. (1973). Clustering of transfer RNA genes of Xenopus laevis. *J. Mol. Biol.* **79**, 411–29.

Collart, C., Allen, G. E., Bradshaw, C. R., Smith, J. C. and Zegerman, P. (2013). Titration of Four Replication Factors is Essential for the Xenopus laevis Midblastula Transition. *Science (80-.).* **341**, 893–896.

Conerly, M. L., Yao, Z., Zhong, J. W., Groudine, M. and Tapscott, S. J. (2016). Distinct Activities of Myf5 and MyoD Indicate Separate Roles in Skeletal Muscle Lineage Specification and Differentiation. *Dev. Cell* **36**, 375–385.

Cornelison, D. D. W. and Wold, B. J. (1997). Single-Cell Analysis of Regulatory Gene Expression in Quiescent and Activated Mouse Skeletal Muscle Satellite Cells. *Dev. Biol.* **191**, 270–283.

Cossu, G. and Butler-Browne, G. (1999). In vivo satellite cell activation via Myf5 and MyoD in regenerating mouse skeletal muscle. *J. Cell Sci.* **112**, 2895–2901.

Cossu, G., Kelly, R., Tajbakhsh, S., Di Donna, S., Vivarelli, E. and Buckingham, M. (1996). Activation of different myogenic pathways: myf-5 is induced by the neural tube and MyoD by the dorsal ectoderm in mouse paraxial mesoderm. *Development* **122**, 429–37.

Crighton, D., Woiwode, A., Zhang, C., Mandavia, N., Morton, J. P., Warnock, L. J., Milner, J., White, R. J. and Johnson, D. L. (2003). p53 represses RNA polymerase III transcription by targeting TBP and inhibiting promoter occupancy by TFIIIB. *EMBO J.* **22**, 2810–20.

Cunliffe, V. and Smith, J. C. (1994). Specification of mesodermal pattern in Xenopus laevis by interactions between Brachyury, noggin and Xwnt-8. *EMBO J.* **13**, 349–59.

Dang, C. H. I. V (1999). c-Myc Target Genes Involved in Cell Growth , Apoptosis , and Metabolism. *Mol. Cell. Biol.* **19**, 1–11.

Davis, R. L., Weintraub, H. and Lassar, A. B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51**, 987–1000.

De La Serna, I. L., Ohkawa, Y., Berkes, C. A., Bergstrom, D. A., Dacwag, C. S., Tapscott, S. J. and Imbalzano, A. N. (2005). MyoD targets chromatin remodeling complexes to the myogenin locus prior to forming a stable DNA-bound complex. *Mol. Cell. Biol.* **25**, 3997–4009.

Della Gaspera, B., Armand, A.-S., Sequeira, I., Chesneau, A., Mazabraud, A., Lécolle, S., Charbonnier, F. and Chanoine, C. (2012). Myogenic waves and myogenic programs during Xenopus embryonic myogenesis. *Dev. Dyn.* **241**, 995–1007.

Desai, N., Lee, J., Upadhya, R., Chu, Y., Moir, R. D. and Willis, I. M. (2005). Two steps in Maf1-dependent repression of transcription by RNA polymerase III. *J. Biol. Chem.* **280**, 6455–62.

Dittmar, K. A., Goodenbour, J. M. and Pan, T. (2006). Tissue-specific differences in human transfer RNA expression. *PLoS Genet.* **2**, e221.

Dodou, E., Xu, S.-M. and Black, B. L. (2003). mef2c is activated directly by myogenic basic helix-loop-helix proteins during skeletal muscle development in vivo. *Mech. Dev.* **120**, 1021–1032.

Dong, H., Nilsson, L. and Kurland, C. G. (1996). Co-variation of tRNA Abundance and Codon Usage in Escherichia coli at Different Growth Rates. *J. Mol. Biol.* **260**, 649–663.

Dorey, K., Amaya, E., Harpal, K., Yamaguchi, T. P., Rossant, J. and Papalopulu, N. (2010). FGF signalling: diverse roles during early vertebrate embryogenesis. *Development* **137**, 3731–42.

Draper, B. W., Morcos, P. A. and Kimmel, C. B. (2001). Inhibition of zebrafish fgf8 pre-mRNA splicing with morpholino oligos: A quantifiable method for gene knockdown. *Genesis* **30**, 154–156.

Dumay-Odelot, H., Durrieu-Gaillard, S., Da Silva, D., Roeder, R. G. and

Teichmann, M. (2010). Cell growth- and differentiation-dependent regulation of RNA polymerase III transcription. *Cell Cycle* **9**, 3687–3699.

Duret, L. (2000). tRNA gene number and codon usage in the C. elegans genome are co-adapted for optimal translation of highly expressed genes. *Trends Genet.* **16**, 287–289.

Durrieu-Gaillard, S., Dumay-Odelot, H., Boldina, G., Tourasse, N. J., Allard, D., André, F., Macari, F., Choquet, A., Lagarde, P., Drutel, G., et al. (2017). Regulation of RNA polymerase III transcription during transformation of human IMR90 fibroblasts with defined genetic elements. *Cell Cycle* **4101**, 1–34.

Eisen, J. S. and Smith, J. C. (2008). Controlling morpholino experiments: don't stop making antisense. *Development* **135**, 1735–43.

El-Brolosy, M., Rossi, A., Kontarakis, Z., Kuenne, C., Guenther, S., Fukuda, N., Takacs, C., Lai, S.-L., Fukuda, R., Gerri, C., et al. (2018). Genetic compensation is triggered by mutant mRNA degradation. *bioRxiv* 328153.

Emerson, C. P. (1990). Myogenesis and developmental control genes. *Curr. Opin. Cell Biol.* **2**, 1065–1075.

Emerson, C. P. (1993). Embryonic signals for skeletal myogenesis: arriving at the beginning. *Curr. Opin. Cell Biol.* **5**, 1057–1064.

Felton-Edkins, Z. A., Fairley, J. A., Graham, E. L., Johnston, I. M., White, R. J. and Scott, P. H. (2003a). The mitogen-activated protein (MAP) kinase ERK induces tRNA synthesis by phosphorylating TFIIIB. *EMBO J.* **22**, 2422–2432.

Felton-Edkins, Z. A., Kenneth, N. S., Brown, T. R. P., Daly, N. L., Gomez-Roman, N., Grandori, C., Eisenman, R. N. and White, R. J. (2003b). Direct regulation of RNA polymerase III transcription by RB, p53 and c-Myc. *Cell Cycle* **2**, 181–4.

Fetka, I., Doederlein, G. and Bouwmeester, T. (2000). Neuroectodermal specification and regionalization of the Spemann organizer in Xenopus. *Mech. Dev.* **93**, 49–58.

Fisher, M., Isaacs, H. and Pownall, M. (2002). eFGF is required for activation of

XmyoD expression in the myogenic cell lineage of Xenopus laevis. *Development* **129**, 1307–1315.

Fong, A. P., Yao, Z., Zhong, J. W., Cao, Y., Ruzzo, W. L., Gentleman, R. C. and Tapscott, S. J. (2012). Genetic and epigenetic determinants of neurogenesis and myogenesis. *Dev. Cell* **22**, 721–35.

Fong, A. P., Yao, Z., Zhong, J. W., Johnson, N. M., Farr, G. H., Maves, L. and Tapscott, S. J. (2015). Conversion of MyoD to a neurogenic factor: binding site specificity determines lineage. *Cell Rep.* **10**, 1937–46.

François, V. and Bier, E. (1995). Xenopus chordin and Drosophila short gastrulation genes encode homologous proteins functioning in dorsal-ventral axis formation. *Cell* **80**, 19–20.

Fu, J., Bian, M., Jiang, Q. and Zhang, C. (2007). Roles of Aurora kinases in mitosis and tumorigenesis. *Mol. Cancer Res.* **5**, 1–10.

Garel, J. P. and Hentzen, D. (1974). Codon responses of tRNAAla , tRNAGly and tRNASer from the posterior part of the silkgland of *Bombyx mori* L. *FEBS Lett.* **39**, 359–363.

Gentsch, G. E., Spruce, T., Monteiro, R. S., Owens, N. D. L., Martin, S. R. and Smith, J. C. (2018). Innate Immune Response and Off-Target Mis-splicing Are Common Morpholino-Induced Side Effects in Xenopus. *Dev. Cell* **44**, 597–610.

Geslain, R. and Pan, T. (2011). tRNA: Vast reservoir of RNA molecules with unexpected regulatory function. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 16489–90.

Giacinti, C. and Giordano, A. (2006). RB and cell cycle progression. *Oncogene* **25**, 5220–7.

Gianakopoulos, P. J., Mehta, V., Voronova, A., Cao, Y., Yao, Z., Coutu, J., Wang, X., Waddington, M. S., Tapscott, S. J. and Skerjanc, I. S. (2011). MyoD directly up-regulates premyogenic mesoderm factors during induction of skeletal myogenesis in stem cells. *J. Biol. Chem.* **286**, 2517–2525.

Gilbert, S. (2000). *Developmental Biology. 6th Edition*.

Gingold, H. and Pilpel, Y. (2011). Determinants of translation efficiency and accuracy. *Mol. Syst. Biol.* **7**, 481.

Gingold, H., Tehler, D., Christoffersen, N. R., Nielsen, M. M., Asmar, F., Kooistra, S. M., Christophersen, N. S., Christensen, L. L., Borre, M., Sørensen, K. D., et al. (2014). A Dual Program for Translation Regulation in Cellular Proliferation and Differentiation. *Cell* **158**, 1281–1292.

Goldenson, B. and Crispino, J. D. (2015). The aurora kinases in cell cycle and leukemia. *Oncogene* **34**, 537–545.

Goljanek-Whysall, K., Sweetman, D. and Münsterberg, A. E. (2012). microRNAs in skeletal muscle differentiation and disease. *Clin. Sci.* **123**, 611–25.

Gomez-Roman, N., Grandori, C., Eisenman, R. N. and White, R. J. (2003). Direct activation of RNA polymerase III transcription by c-Myc. *Nature* **421**, 290–4.

Goodarzi, H., Nguyen, H. C. B., Zhang, S., Dill, B. D., Molina, H. and Tavazoie, S. F. (2016). Modulated Expression of Specific tRNAs Drives Gene Expression and Cancer Progression. *Cell* **165**, 1416–1427.

Gossett, L. A., Kelvin, D. J., Sternberg, E. A. and Olson, E. N. (1989). A New Myocyte-Specific Enhancer-Binding Factor That Recognizes a Conserved Element Associated with Multiple Muscle-Specific Genes. *Mol. Cell. Biol.* **9**, 5022–5033.

Graham, V., Khudyakov, J., Ellis, P. and Pevny, L. (2003). SOX2 Functions to Maintain Neural Progenitor Identity. *Neuron* **39**, 749–765.

Grewal, S. S. (2015). Why should cancer biologists care about tRNAs? tRNA synthesis, mRNA translation and the control of growth. *Biochim. Biophys. Acta* **1849**, 898–907.

Grifone, R., Xie, X., Bourgeois, A., Saquet, A., Duprez, D. and Shi, D. L. (2014). The RNA-binding protein Rbm24 is transiently expressed in myoblasts and is required for myogenic differentiation during vertebrate development. *Mech. Dev.* **134**, 1–15.

Guo, W., Schafer, S., Greaser, M. L., Radke, M. H., Liss, M., Govindarajan, T., Maatz, H., Schulz, H., Li, S., Parrish, A. M., et al. (2012). RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing. *Nat. Med.* **18**, 766–773.

Guo, X., Zhang, T., Hu, Z., Zhang, Y., Shi, Z., Wang, Q., Cui, Y., Wang, F., Zhao, H. and Chen, Y. (2014). Efficient RNA/Cas9-mediated genome editing in Xenopus tropicalis. *Development* **141**, 707–14.

Haldar, M., Karan, G., Tvrdik, P. and Capecchi, M. R. (2008). Two cell lineages, myf5 and myf5-independent, participate in mouse skeletal myogenesis. *Dev. Cell* **14**, 437–45.

Hamilton, F. S., Wheeler, G. N. and Hoppler, S. (2001). Difference in XTcf-3 dependency accounts for change in response to beta-catenin-mediated Wnt signalling in Xenopus blastula. *Development* **128**, 2063–2073.

Hanahan, D. and Weinberg, R. A. (2000). The Hallmarks of Cancer. *Cell* **100**, 57–70.

Hasty, P., Bradley, a, Morris, J. H., Edmondson, D. G., Venuti, J. M., Olson, E. N. and Klein, W. H. (1993). Muscle deficiency and neonatal death in mice with a targeted mutation in the myogenin gene. *Nature* **364**, 501–506.

Haurie, V., Durrieu-Gaillard, S., Dumay-Odelot, H., Da Silva, D., Rey, C., Prochazkova, M., Roeder, R. G., Besser, D. and Teichmann, M. (2010). Two isoforms of human RNA polymerase III with specific functions in cell growth and transformation. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 4176–81.

Heasman, J. (2002). Morpholino Oligos: Making Sense of Antisense? *Dev. Biol.* **243**, 209–214.

Heasman, J., Kofron, M. and Wylie, C. (2000). βCatenin Signaling Activity Dissected in the Early Xenopus Embryo: A Novel Antisense Approach. *Dev. Biol.* **222**, 124–134.

Hentzen, D., Chevallier, A. and Garel, J.-P. (1981). Differential usage of iso-accepting tRNASer species in silk glands of Bombyx mori. *Nature* **290**, 267–269.

Hoppler, S., Brown, J. D. and Randall, T. M. (1996). Expression of a dominant-

negative Wnt blocks induction of MyoD in Xenopus embryos. *Genes Dev.* **10**, 2805–2817.

Hopwood, N. D. and Gurdon, J. B. (1990). Activation of muscle genes without myogenesis by ectopic expression of MyoD in frog embryo cells. *Nature* **347**, 197–200.

Hopwood, N. D., Pluck, A. and Gurdon, J. B. (1989). MyoD expression in the forming somites is an early response to mesoderm induction in Xenopus embryos. *EMBO J.* **8**, 3409–17.

Hopwood, N. D., Pluck, A. and Gurdon, J. B. (1991). Xenopus Myf-5 Marks Early Muscle-Cells and Can Activate Muscle Genes Ectopically in Early Embryos. *Development* **111**, 551–560.

Hopwood, N. D., Pluck, A., Gurdon, J. D. and Dilworth, S. M. (1992). Expression of XMyoD protein in early Xenopus laevis embryos. *Development* **114**, 31–38.

Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., Peterson, R. T., Yeh, J.-R. J. and Joung, J. K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 227–9.

Irion, U., Krauss, J., Nüsslein-Volhard, C., Auer, T. O., Bene, F. Del, Auer, T. O., Duroure, K., Cian, A. De, Concordet, J.-P., Bene, F. Del, et al. (2014). Precise and efficient genome editing in zebrafish using the CRISPR/Cas9 system. *Development* **141**, 4827–30.

Isaacs, H. V, Pownall, M. E. and Slack, J. M. (1994). eFGF regulates Xbra expression during Xenopus gastrulation. *EMBO J.* **13**, 4469–4481.

Jen, Y., Weintraub, H. and Benezra, R. (1992). Overexpression of Id protein inhibits the muscle differentiation program: In vivo association of Id with E2A proteins. *Genes Dev.* **6**, 1466–1479.

Jennings, C. G. (1992). Expression of the myogenic gene MRF4 during Xenopus development. *Dev. Biol.* **151**, 319–32.

Jin, D., Hidaka, K., Shirai, M. and Morisaki, T. (2010). RNA-binding motif protein 24 regulates myogenin expression and promotes myogenic differentiation.

Genes to Cells **15**, 1158–1167.

Jin, S., Collin, J., Zhu, L., Montaner, D., Armstrong, L., Neganova, I. and Lako, M. (2016). A Novel Role for miR-1305 in Regulation of Pluripotency-Differentiation Balance, Cell Cycle, and Apoptosis in Human Pluripotent Stem Cells. *Stem Cells* **34**, 2306–2317.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. and Charpentier, E. (2012). A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science (80-.).* **337**, 816–822.

Jones, C. M., Kuehn, M. R., Hogan, B. L., Smith, J. C. and Wright, C. V (1995). Nodal-related signals induce axial mesoderm and dorsalize mesoderm during gastrulation. *Development* **121**, 3651–3662.

Kanaya, S., Yamada, Y., Kudo, Y. and Ikemura, T. (1999). Studies of codon usage and tRNA genes of 18 unicellular organisms and quantification of Bacillus subtilis tRNAs: gene expression level and species-specific diversity of codon usage based on multivariate analysis. *Gene* **238**, 143–155.

Kantidakis, T., Ramsbottom, B. A., Birch, J. L., Dowding, S. N. and White, R. J. (2010). mTOR associates with TFIIIC, is found at tRNA and 5S rRNA genes, and targets their repressor Maf1. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 11823–8.

Kasberg, A. D., Brunskill, E. W. and Potter, S. S. (2013). SP8 regulates signaling centers during craniofacial development Abigail. *Dev. Biol.* **381**, 312–323.

Kassar-Duchossoy, L., Gayraud-Morel, B., Gomès, D., Rocancourt, D., Buckingham, M., Shinin, V. and Tajbakhsh, S. (2004). Mrf4 determines skeletal muscle identity in Myf5:Myod double-mutant mice. *Nature* **431**, 466–71.

Kee, Y. and Bronner-Fraser, M. (2005). To proliferate or to die: Role of Id3 in cell cycle progression and survival of neural crest progenitors. *Genes Dev.* **19**, 744–755.

Khattar, E., Kumar, P., Liu, C. Y., Akıncılar, S. C., Raju, A., Lakshmanan, M., Jean, J., Maury, P., Qiang, Y., Li, S., et al. (2016). Telomerase reverse

transcriptase promotes cancer cell proliferation by augmenting tRNA expression. *J. Clin. Invest.* **126**, 4045–4060.

Kim, H. K., Lee, Y. S., Sivaprasad, U., Malhotra, A. and Dutta, A. (2006). Muscle-specific microRNA miR-206 promotes muscle differentiation. *J. Cell Biol.* **174**, 677–687.

Knoepfler, P. S., Bergstrom, D. A., Uetsuki, T., Dac-korytko, I., Sun, Y. H., Wright, W. E., Tapscott, S. J. and Kamps, M. P. (1999). A conserved motif N-terminal to the DNA-binding domains of myogenic bHLH transcription factors mediates cooperative DNA binding with Pbx – Meis1 / Prep1. *Nucleic Acids Res.* **27**, 3752–3761.

Kok, F. O., Shin, M., Ni, C.-W., Gupta, A., Grosse, A. S., van Impel, A., Kirchmaier, B. C., Peterson-Maduro, J., Kourkoulis, G., Male, I., et al. (2014). Reverse Genetic Screening Reveals Poor Correlation between Morpholino-Induced and Mutant Phenotypes in Zebrafish. *Dev. Cell* **32**, 97–108.

Köster, M., Dillinger, K. and Knöchel, W. (1998). Expression pattern of the winged helix factor XFD-11 during Xenopus embryogenesis. *Mech. Dev.* **76**, 169–173.

Kuang, S., Kuroda, K., Le Grand, F. and Rudnicki, M. A. (2007). Asymmetric Self-Renewal and Commitment of Satellite Stem Cells in Muscle. *Cell* **129**, 999–1010.

Kume, T., Deng, K.-Y., Winfrey, V., Gould, D. B., Walter, M. A. and Hogan, B. L. . (1998). The Forkhead/Winged Helix Gene Mf1 Is Disrupted in the Pleiotropic Mouse Mutation congenital hydrocephalus. *Cell* **93**, 985–996.

Kume, T., Jiang, H., Topczewska, J. M. and Hogan, B. L. (2001). The murine winged helix transcription factors, Foxc1 and Foxc2, are both required for cardiovascular development and somitogenesis. *Genes Dev.* **15**, 2470–2482.

Kutter, C., Brown, G. D., Gonçalves, A., Wilson, M. D., Watt, S., Brazma, A., White, R. J. and Odom, D. T. (2011). Pol III binding in six mammals shows conservation among amino acid isotypes despite divergence among tRNA genes. *Nat. Genet.* **43**, 948–55.

Laferté, A., Favry, E., Sentenac, A., Riva, M., Carles, C. and Chédin, S. (2006). The transcriptional activity of RNA polymerase I is a key determinant for the level of all ribosome components. *Genes Dev.* **20**, 2030–2040.

Lagha, M., Brunelli, S., Messina, G., Cumano, A., Kume, T., Relaix, F. and Buckingham, M. E. (2009). Pax3:Foxc2 Reciprocal Repression in the Somite Modulates Muscular versus Vascular Cell Fate Choice in Multipotent Progenitors. *Dev. Cell* **17**, 892–899.

Langlands, K., Yin, X., Anand, G. and Prochownik, E. V (1997). Differential interactions of Id proteins with basic-helix-loop-helix transcription factors. *J. Biol. Chem.* **272**, 19785–93.

Laplante, M. and Sabatini, D. M. (2009). mTOR signaling at a glance. *J. Cell Sci.* **122**, 3589–94.

Laplante, M. and Sabatini, D. M. (2012). mTOR signaling in growth control and disease. *Cell* **149**, 274–293.

Lasorella, A., Benezra, R. and Iavarone, A. (2014). The ID proteins: master regulators of cancer stem cells and tumour aggressiveness. *Nat. Rev. Cancer* **14**, 77–91.

Lepper, C., Partridge, T. A., Fan, C.-M. and Gruss, P. (2011). An absolute requirement for Pax7-positive satellite cells in acute injury-induced skeletal muscle regeneration. *Development* **138**, 3639–46.

Li, H.-Y., Bourdelas, A., Carron, C. and Shi, D.-L. (2010a). The RNA-binding protein Seb4/RBM24 is a direct target of MyoD and is required for myogenesis during Xenopus early development. *Mech. Dev.* **127**, 281–291.

Li, D., Morales, A., Gonzalez-Quintana, J., Norton, N., Siegfried, J. D., Hofmeyer, M. and Hershberger, R. E. (2010b). Identification of novel mutations in RBM20 in patients with dilated cardiomyopathy. *Clin. Transl. Sci.* **3**, 90–97.

Li, S., Guo, W., Dewey, C. N. and Greaser, M. L. (2013). Rbm20 regulates titin alternative splicing as a splicing repressor. *Nucleic Acids Res.* **41**, 2659–2672.

Li, M., Zhao, L., Page-McCaw, P. S. and Chen, W. (2016). Zebrafish Genome Engineering Using the CRISPR–Cas9 System. *Trends Genet.* **32**, 815–827.

Light, W., Vernon, A., Lasorella, A., Lavarone, A. and LaBonne, C. (2005). Xenopus Id3 is required downstream of Myc for the formation of multipotent neural crest progenitor cells. *Development* **132**, 1831–1841.

Lluís, F., Ballestar, E., Suelves, M., Esteller, M. and Muñoz-Cánoves, P. (2005). E47 phosphorylation by p38 MAPK promotes MyoD/E47 association and muscle-specific gene transcription. *EMBO J.* **24**, 974–84.

Lloyd, R. E., Foster, P.G., Guille, M. and Littlewood, D.T.J. (2012). Next generation sequencing and comparitive analyses of *Xenopus* mitogenomes. BMC Genomics. **13**, 496

Macek, P., Cliff, M. J., Embrey, K. J., Holdgate, G. A., Nissink, J. W. M., Panova, S., Waltho, J. P. and Davies, R. A. (2018). Myc phosphorylation in its basic helix-loop-helix region destabilizes transient α-helical structures, disrupting Max and DNA binding. *J. Biol. Chem.* **293**, 9301–9310.

Maguire, R. J., Isaacs, H. V and Pownall, M. E. (2012). Early transcriptional targets of MyoD link myogenesis and somitogenesis. *Dev. Biol.* **371**, 256–68.

Maragh, S., Miller, R. A., Bessling, S. L., Wang, G., Hook, P. W. and McCallion, A. S. (2014). Rbm24a and Rbm24b are required for normal somitogenesis. *PLoS One* **9**,.

Marshall, L. and White, R. J. (2008). Non-coding RNA production by RNA polymerase III is implicated in cancer. *Nat. Rev. Cancer* **8**, 911–4.

Marshall, L., Rideout, E. J. and Grewal, S. S. (2012). Nutrient / TOR-dependent regulation of RNA polymerase III controls tissue and organismal growth in Drosophila. *EMBO J.* **31**, 1916–1930.

Mauro, A. (1961). Satellite cell of skeletal muscle fibers. *J. Cell Biol.*

Maves, L., Waskiewicz, A. J., Paul, B., Cao, Y., Tyler, A., Moens, C. B. and Tapscott, S. J. (2007). Pbx homeodomain proteins direct Myod activity to promote fast-muscle differentiation. *Development* **3382**, 3371–3382.

Mayeuf-Louchart, A., Lagha, M., Danckaert, A., Rocancourt, D., Relaix, F.,

Vincent, S. and Buckingham, M. (2014). Notch regulation of myogenic versus endothelial fates of cells that migrate from the somite to the limb. *Proc. Natl. Acad. Sci.* **111**, 1–6.

Mayeuf-Louchart, A., Montarras, D., Bodin, C., Kume, T., Vincent, S. D. and Buckingham, M. (2016). Endothelial cell specification in the somite is compromised in Pax3-positive progenitors of Foxc1/2 conditional mutants, with loss of forelimb myogenesis. *Development* **143**, dev.128017-.

McDermott, A., Gustafsson, M., Elsam, T., Hui, C.-C., Emerson, C. P. and Borycki, A.-G. (2005). Gli2 and Gli3 have redundant and context-dependent function in skeletal muscle formation. *Development* **132**, 345–57.

McQueen, C. and Pownall, M. E. (2017). An analysis of MyoD-dependent transcription using CRISPR/Cas9 gene targeting in Xenopus tropicalis embryos. *Mech. Dev.* **146**, 1–9.

Megeney, L. A., Kablar, B., Garrett, K., Anderson, J. E. and Rudnicki, M. A. (1996). MyoD is required for myogenic stem cell function in adult skeletal muscle. *Genes Dev.* **10**, 1173–83.

Miller, D. M., Thomas, S. D., Islam, A., Muench, D. and Sedoris, K. (2012). c-Myc and Cancer Metabolism. *Clin. Cancer Res.* **18**, 5546–5554.

Mishima, Y., Abreu-Goodger, C., Staton, A. A., Stahlhut, C., Shou, C., Cheng, C., Gerstein, M., Enright, A. J. and Giraldez, A. J. (2009). Zebrafish miR-1 and miR-133 shape muscle gene expression and regulate sarcomeric actin organization. *Genes Dev.* **23**, 619–632.

Moir, R. D., Lee, J., Haeusler, R. A., Desai, N., Engelke, D. R. and Willis, I. M. (2006). Protein kinase A regulates RNA polymerase III transcription through the nuclear localization of Maf1. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 15044–9.

Mok, G. F., Lozano-Velasco, E., Maniou, E., Viaut, C., Moxon, S., Wheeler, G. and Münsterberg, A. (2018). miR-133-mediated regulation of the Hedgehog pathway orchestrates embryo myogenesis. *Development* **145**, dev159657.

Molkentin, J. D., Black, B. L., Martin, J. F. and Olson, E. N. (1995). Cooperative Activation of Muscle Gene Expression by MEF2 and Myogenic bHLH Proteins. *Cell* **83**, 1125–1136.

Monsoro-Burq, A.-H., Wang, E. and Harland, R. (2005). Msx1 and Pax3 Cooperate to Mediate FGF8 and WNT Signals during Xenopus Neural Crest Induction. *Dev. Cell* **8**, 167–178.

Moqtaderi, Z., Wang, J., Raha, D., White, R. J., Snyder, M., Weng, Z. and Struhl, K. (2010). Genomic binding profiles of functionally distinct RNA polymerase III transcription complexes in human cells. *Nat. Struct. Mol. Biol.* **17**, 635–40.

Moriyama, E. N. and Powell, J. R. (1997). Codon Usage Bias and tRNA Abundance in Drosophila. *J. Mol. Evol.* **45**, 514–523.

Münsterberg, A. and Lassar, A. (1995). Combinatorial signals from the neural tube, floor plate and notochord induce myogenic bHLH gene expression in the somite. *Development* **121**, 651–660.

Münsterberg, A. E., Kitajewski, J., Bumcrot, D. A., McMahon, A. P. and Lassar, A. B. (1995). Combinatorial signaling by Sonic hedgehog and Wnt family members induces myogenic bHLH gene expression in the somite. *Genes Dev.* **9**, 2911–2922.

Nabeshima, Y., Hanaoka, K., Hayasaka, M., Esumi, E., Li, S., Nonaka, I. and Nabeshima, Y. (1993). Myogenin gene disruption results in perinatal lethality because of severe muscle defect. *Nature* **364**, 532–5.

Naidu, P. S., Ludolph, D. C., To, R. Q., Hinterberger, T. J. and Konieczny, S. F. (1995). Myogenin and MEF2 Function Synergistically To Activate the MRF4 Promoter during Myogenesis. *Mol. Cell. Biol.* **15**, 2707–2718.

Nakamura, Y., de Paiva Alves, E., Veenstra, G. J. C. and Hoppler, S. (2016). Tissue- and stage-specific Wnt target gene expression is controlled subsequent to β-catenin recruitment to cis-regulatory modules. *Development* **143**, 1914–1925.

Nakayama, T., Fish, M. B., Fisher, M., Oomen-Hajagos, J., Thomsen, G. H. and Grainger, R. M. (2013). Simple and efficient CRISPR/Cas9-mediated targeted mutagenesis in Xenopus tropicalis. *Genesis* **51**, 835–43.

Nakayama, T., Blitz, I. L., Fish, M. B., Odeleye, A. O., Manohar, S., Cho, K. W. Y. and Grainger, R. M. (2014). Cas9-based genome editing in Xenopus

tropicalis. *Methods Enzymol.* **546**, 355–75.

Nandan, M. O. and Yang, V. W. (2009). The role of Krüppel-like factors in the reprogramming of somatic cells to induced pluripotent stem cells. *Histol. Histopathol.* **24**, 1343–1355.

Narayanswami, S., Doering, J. L., Fokta, F. J., Rosenthal, D. S., Nguyen, T.-N. and Hamkalo, B. A. (1995). Chromosomal locations of major tRNA gene clusters of Xenopus laevis. *Chromosoma* **104**, 68–74.

Neuhold, L. A. and Wold, B. (1993). HLH forced dimers: Tethering MyoD to E47 generates a dominant positive myogenic factor insulated from negative regulation by Id. *Cell* **74**, 1033–1042.

Newport, J. and Kirschner, M. (1982a). A Major Developmental Transition in Early Xenopus Embryos: I. characterization and timing of cellular changes at the Midblastula Stage. *Cell* **30**, 675–686.

Newport, J. and Kirschner, M. (1982b). A Major Developmental Transition in Early Xenopus Embryos: II. Control of the Onset of Transcription. *Cell* **30**, 687–696.

Oelgeschläger, M., Kuroda, H., Reversade, B. and De Robertis, E. M. (2003). Chordin Is Required for the Spemann Organizer Transplantation Phenomenon in Xenopus Embryos. *Dev. Cell* **4**, 219–230.

Oficjalska-Pham, D., Harismendy, O., Smagowicz, W. J., Gonzalez de Peredo, A., Boguta, M., Sentenac, A. and Lefebvre, O. (2006). General repression of RNA polymerase III transcription is triggered by protein phosphatase type 2A-mediated dephosphorylation of Maf1. *Mol. Cell* **22**, 623–32.

Oler, A. J., Alla, R. K., Roberts, D. N., Wong, A., Hollenhorst, P. C., Chandler, K. J., Cassiday, P. A., Nelson, C. A., Hagedorn, C. H., Graves, B. J., et al. (2010). Human RNA polymerase III transcriptomes and relationships to Pol II promoter chromatin and enhancer-binding factors. *Nat. Struct. Mol. Biol.* **17**, 620–8.

Ott, M. O., Bober, E., Lyons, G., Arnold, H. and Buckingham, M. (1991). Early expression of the myogenic regulatory gene, myf-5, in precursor cells of skeletal muscle in the mouse embryo. *Development* **111**, 1097–1107.

Otte, A. P. and Moon, R. T. (1992). Ectopic induction of dorsal mesoderm by overexpression of Xwnt-8 elevates the neural competence of Xenopus ectoderm. *Dev. Biol.* **152**, 184–187.

Owens, N. D. L., Blitz, I. L., Lane, M. A., Patrushev, I., Overton, J. D., Gilchrist, M. J., Cho, K. W. Y. and Khokha, M. K. (2016). Measuring Absolute RNA Copy Numbers at High Temporal Resolution Reveals Transcriptome Kinetics in Development. *Cell Rep.* **14**, 632–647.

Pan, Y., Bai, C. B., Joyner, A. L. and Wang, B. (2006). Sonic hedgehog Signaling Regulates Gli2 Transcriptional Activity by Suppressing Its Processing and Degradation Sonic hedgehog Signaling Regulates Gli2 Transcriptional Activity by Suppressing Its Processing and Degradation †. *Mol. Cell. Biol.* **26**, 3365–3377.

Papin, C., van Grunsven, L. A., Verschueren, K., Huylebroeck, D. and Smith, J. C. (2002). Dynamic regulation of Brachyury expression in the amphibian embryo by XSIP1. *Mech. Dev.* **111**, 37–46.

Paquette, J., Tokuyasu, T., Salzberg, S., Altherr, M., Bredt, D., Blaxall, B., Adami, E., Rintisch, C., Dauksaite, V., Radke, M. H., et al. (2014). RNA-binding protein RBM20 represses splicing to orchestrate cardiac pre-mRNA processing. *Bioinformatics.* **26**, 285–286.

Paule, M. R. and White, R. J. (2000). Transcription by RNA polymerases I and III. *Nucleic Acids Res.* **28**, 1283–1298.

Pavon-Eternod, M., Gomes, S., Geslain, R., Dai, Q., Rosner, M. R. and Pan, T. (2009). tRNA over-expression in breast cancer and functional consequences. *Nucleic Acids Res.* **37**, 7268–80.

Pavon-Eternod, M., Gomes, S., Rosner, M. R. and Pan, T. (2013). Overexpression of initiator methionine tRNA leads to global reprogramming of tRNA expression and increased proliferation in human epithelial cells. *RNA* **19**, 461–6.

Penn, B. H., Bergstrom, D. A., Dilworth, F. J., Bengal, E. and Tapscott, S. J. (2004). A MyoD-generated feed-forward circuit temporally patterns gene expression during skeletal muscle differentiation. *Genes Dev.* **18**, 2348–53.

Piccolo, S., Sasai, Y., Lu, B. and De Robertis, E. M. (1996). Dorsoventral Patterning in Xenopus: Inhibition of Ventral Signals by Direct Binding of Chordin to BMP-4. *Cell* **86**, 589–598.

Pinney, D., Pearson-White, S., Konieczny, S., Latham, K. and Emerson, C. P. (1988). Myogenic lineage determination and differentiation: Evidence for a regulatory gene pathway. *Cell* **53**, 781–793.

Plotkin, J. B., Robins, H. and Levine, A. J. (2004). Tissue-specific codon usage and the expression of human genes. *Proc. Natl. Acad. Sci.* **101**, 12588–12591.

Pownall, M. E. and Emerson, C. P. (1992). Sequential activation of three myogenic regulatory genes during somite morphogenesis in quail embryos. *Dev. Biol.* **151**, 67–79.

Pownall, M. E. and Isaacs, H. V. (2010). *FGF Signalling in Vertebrate Development*. Morgan & Claypool Life Sciences.

Pownall, M. E., Gustafsson, M. and Emerson, C. P. (2002). Myogenic Regulatory Factors and the Specification of Muscle Progenitors in Vertebrate Embryos. *Annu. Rev. Cell Dev. Biol.* **18**, 747–783.

Ramsköld, D., Luo, S., Wang, Y.-C., Li, R., Deng, Q., Faridani, O. R., Daniels, G. A., Khrebtukova, I., Loring, J. F., Laurent, L. C., et al. (2012). Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat. Biotechnol.* **30**, 777–782.

Rana, A. A., Collart, C., Gilchrist, M. J. and Smith, J. C. (2006). Defining synphenotype groups in Xenopus tropicalis by use of antisense morpholino oligonucleotides. *PLoS Genet.* **2**, 1751–1772.

Rathjen, T., Pais, H., Sweetman, D., Moulton, V., Munsterberg, A. and Dalmay, T. (2009). High throughput sequencing of microRNAs in chicken somites. *FEBS Lett.* **583**, 1422–1426.

Reina, J. H., Azzouz, T. N. and Hernandez, N. (2006). Maf1, a new player in the regulation of human RNA polymerase III transcription. *PLoS One* **1**, e134.

Renaud, M., Praz, V., Vieu, E., Florens, L., Washburn, M. P., L'Hôte, P. and

Hernandez, N. (2014). Gene duplication and neofunctionalization: POLR3G and POLR3GL. *Genome Res.* **24**, 37–51.

Rosenberg, M. I., Georges, S. A., Asawachaicharn, A., Analau, E. and Tapscott, S. J. (2006). MyoD inhibits Fstl1 and Utrn expression by inducing transcription of miR-206. *J. Cell Biol.* **175**, 77–85.

Rosenthal, D. S. and Doering, J. L. (1983). The genomic organization of dispersed tRNA and 5 S RNA genes in Xenopus laevis. *J. Biol. Chem.* **258**, 7402–10.

Rossi, A., Kontarakis, Z., Gerri, C., Nolte, H., Hölper, S., Krüger, M. and Stainier, D. Y. R. (2015). Genetic compensation induced by deleterious mutations but not gene knockdowns. *Nature* **524**, 230–233.

Rudnicki, M. A., Braun, T., Hinuma, S. and Jaenisch, R. (1992). Inactivation of MyoD in mice leads to up-regulation of the myogenic HLH gene Myf-5 and results in apparently normal muscle development. *Cell* **71**, 383–390.

Rudnicki, M. A., Schnegelsberg, P. N. J., Stead, R. H., Braun, T., Arnold, H.-H. and Jaenisch, R. (1993). MyoD or Myf-5 is required for the formation of skeletal muscle. *Cell* **75**, 1351–1359.

Rudolph, K. L. M., Schmitt, B. M., Villar, D., White, R. J., Marioni, J. C., Kutter, C. and Odom, D. T. (2016). Codon-Driven Translational Efficiency Is Stable across Diverse Mammalian Cell States. *PLoS Genet.* **12**, e1006024.

Russ, A. P., Wattler, S., Colledge, W. H., Aparicio, S. A. J. R., Carlton, M. B. L., Pearce, J. J., Barton, S. C., Surani, M. A., Ryan, K., Nehls, M. C., et al. (2000). Eomesodermin is required for mouse trophoblast development and mesoderm formation. *Nature* **404**, 95–99.

Ryan, K., Garrett, N., Mitchell, A. and Gurdon, J. . (1996). Eomesodermin, a Key Early Gene in Xenopus Mesoderm Differentiation. *Cell* **87**, 989–1000.

Ryan, K., Garrett, N., Bourillot, P.-Y., Stennard, F. and Gurdon, J. . (2000). The Xenopus Eomesodermin promoter and its concentration-dependent response to activin. *Mech. Dev.* **94**, 133–146.

Sabourin, L. A., Girgis-Gabardo, A., Seale, P., Asakura, A. and Rudnicki, M.

A. (1999). Reduced differentiation potential of primary MyoD-/- myogenic cells derived from adult skeletal muscle. *J. Cell Biol.* **144**, 631–43.

Sacco, A., Mourkioti, F., Tran, R., Choi, J., Llewellyn, M., Kraft, P., Shkreli, M., Delp, S., Pomerantz, J. H., Artandi, S. E., et al. (2010). Short Telomeres and Stem Cell Exhaustion Model Duchenne Muscular Dystrophy in mdx/mTR Mice. *Cell* **143**, 1059–1071.

Saikia, M., Wang, X., Mao, Y., Wan, J. I., Pan, T. A. O. and Qian, S. (2016). Codon optimality controls differential mRNA translation during amino acid starvation. *RNA* **22**, 1–9.

Salghetti, S. E., Kim, S. Y., Tansey, W. P., Matsumoto, M., Nakamichi, I., Kitagawa, K., Shirane, M., Tsunematsu, R., Tsukiyama, T., Ishida, N., et al. (1999). Destruction of Myc by ubiquitin-mediated proteolysis: cancer-associated and transforming mutations stabilize Myc. *EMBO J.* **18**, 717–726.

Sander, J. D. and Joung, J. K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* **32**, 347–55.

Sartorelli, V; Caretti, G. (2005). Mechanisms underlying the transcriptional regulation of skeletal myogenesis. *Curr. Opin. Genet. Dev.* **15**, 528–535.

Sasai, Y., Lu, B., Steinbeisser, H., Geissert, D., Gont, L. K. and De Robertis, E. M. (1994). Xenopus chordin: A novel dorsalizing factor activated by organizer-specific homeobox genes. *Cell* **79**, 779–790.

Sasai, Y., Lu, B., Steinbeisser, H. and De Robertis, E. M. (1995). Regulation of neural induction by the Chd and Bmp-4 antagonistic patterning signals in Xenopus. *Nature* **376**, 333–336.

Sasaki, H., Nishizaki, Y., Hui, C., Nakafuku, M. and Kondoh, H. (1999). Regulation of Gli2 and Gli3 activities by an amino-terminal repression domain: implication of Gli2 and Gli3 as primary mediators of Shh signaling. *Development* **126**, 3915–3924.

Savage, J., Voronova, A., Mehta, V., Sendi-Mukasa, F. and Skerjanc, I. S. (2010). Canonical Wnt signaling regulates Foxc1/2 expression in P19 cells☆. *Differentiation* **79**, 31–40.

Schmalbruch, H. and Lewis, D. M. (2000). Dynamics of nuclei of muscle fibers and connective tissue cells in normal and denervated rat muscles. *Muscle Nerve* **23**, 617–26.

Schmidt, E. V (1999). The role of c- myc in cellular growth control. *Oncogene* **18**, 2988–2996.

Schmitt, B. M., Rudolph, K. L. M., Karagianni, P., Fonseca, N. A., White, R. J., Talianidis, I., Odom, D. T., Marioni, J. C. and Kutter, C. (2014). High-resolution mapping of transcriptional dynamics across tissue development reveals a stable mRNA-tRNA interface. *Genome Res.* **24**, 1797–1807.

Schulte-Merker, S. and Stainier, D. Y. R. (2014). Out with the old, in with the new: reassessing morpholino knockdowns in light of genome editing technology. *Development* **141**, 3103–4.

Seale, P., Sabourin, L. A., Girgis-Gabardo, A., Mansouri, A., Gruss, P. and Rudnicki, M. A. (2000). Pax7 Is Required for the Specification of Myogenic Satellite Cells. *Cell* **102**, 777–786.

Shaw, P. H. (1996). The role of p53 in cell cycle regulation. *Pathol. Res. Pract.* **192**, 669–75.

Sherr, C. J. and McCormick, F. (2002). The RB and p53 pathways in cancer. *Cancer Cell* **2**, 103–112.

Shor, B., Wu, J., Shakey, Q., Toral-Barza, L., Shi, C., Follettie, M. and Yu, K. (2010). Requirement of the mTOR Kinase for the Regulation of Maf1 Phosphorylation and Control of RNA Polymerase III-dependent Transcription in Cancer Cells. *J. Biol. Chem.* **285**, 15380–15392.

Slack, J. M. W., Darlington, B. G., Heath, J. K. and Godsave, S. F. (1987). Mesoderm induction in early Xenopus embryos by heparin-binding growth factors. *Nature* **326**, 197–200.

Slack, J. M., Isaacs, H. V. and Darlington, B. G. (1988). Inductive effects of fibroblast growth factor and lithium ion on Xenopus blastula ectoderm. *Development* **103**, 581–590.

Smith, J. (1995). Mesodern-inducing factors and mesodermal patterning. *Curr.*

Opin. Cell Biol. **7**, 856–861.

Smith, W. C. and Harland, R. M. (1991). Injected Xwnt-8 RNA acts early in Xenopus embryos to promote formation of a vegetal dorsalizing center. *Cell* **67**, 753–765.

Sokol, S. Y. (1993). Mesoderm formation in Xenopus ectodermal explants overexpressing Xwnt8: evidence for a cooperating signal reaching the animal pole by gastrulation. *Development* **118**,.

Sokol, S. Y. and Melton, D. A. (1992). Interaction of Wnt and activin in dorsal mesoderm induction in Xenopus. *Dev. Biol.* **154**, 348–355.

Soleimani, V. D., Yin, H., Jahani-Asl, A., Ming, H., Kockx, C. E. M., van Ijcken, W. F. J., Grosveld, F. and Rudnicki, M. A. (2012). Snail Regulates MyoD Binding-Site Occupancy to Direct Enhancer Switching and Differentiation-Specific Transcription in Myogenesis. *Mol. Cell* **47**, 457–468.

Sölter, M., Köster, M., Hollemann, T., Brey, A., Pieler, T. and Knöchel, W. (1999). Characterization of a subfamily of related winged helix genes, XFD-12/12′/12″ (XFLIP), during Xenopus embryogenesis. *Mech. Dev.* **89**, 161–165.

Sriskanthadevan-Pirahas, S., Deshpande, R., Lee, B. and Grewal, S. S. (2018). Ras/ERK-signalling promotes tRNA synthesis and growth via the RNA polymerase III repressor Maf1 in Drosophila. *PLOS Genet.* **14**, e1007202.

Standley, H. J., Zorn, A. M. and Gurdon, J. B. (2001). eFGF and its mode of action in the community effect during Xenopus myogenesis. *Development* **128**, 1347–57.

Stenico, M., Lloyd, A. T. and Sharp, P. M. (1994). Codon usage in Caenorhabditis elegans: delineation of translational selection and mutational biases. *Nucleic Acids Res.* **22**, 2437–46.

Stutz, F., Gouilloud, E. and Clarkson, S. G. (1989). Oocyte and somatic tyrosine tRNA genes in Xenopus laevis. *Genes Dev.* **3**, 1190–1198.

Sullivan, S. A., Akers, L. and Moody, S. A. (2001). foxD5a, a Xenopus Winged Helix Gene, Maintains an Immature Neural Ectoderm via Transcriptional Repression That Is Dependent on the C-Terminal Domain. *Dev. Biol.* **232**,

439–457.

Summerton, J. and Weller, D. (1997). Morpholino Antisense Oligomers: Design, Preparation, and Properties. *Antisense Nucleic Acid Drug Dev.* **7**, 187–195.

Sutcliffe, J. E., Brown, T. R. P., Allison, S. J., Scott, P. H. and White, R. J. (2000). Retinoblastoma Protein Disrupts Interactions Required for RNA Polymerase III Transcription. *Mol. Cell. Biol.* **20**, 9192–9202.

Suzuki, Y. and Brown, D. D. (1972). Isolation and Identification of the messenger RNA for silk fibroin from Bombyx mori. *J. Mol. Biol.* **63**, 409–429.

Sweetman, D., Rathjen, T., Jefferson, M., Wheeler, G., Smith, T. G., Wheeler, G. N., Münsterberg, A. and Dalmay, T. (2006). FGF-4 signaling is involved in mir-206 expression in developing somites of chicken embryos. *Dev. Dyn.* **235**, 2185–2191.

Sweetman, D., Goljanek, K., Rathjen, T., Oustanina, S., Braun, T., Dalmay, T. and Münsterberg, A. (2008). Specific requirements of MRFs for the expression of muscle specific microRNAs, miR-1, miR-206 and miR-133. *Dev. Biol.* **321**, 491–499.

Tajbakhsh, S., Borello, U., Vivarelli, E., Kelly, R., Papkoff, J., Duprez, D., Buckingham, M. E. and Cossu, G. (1998). Differential activation of Myf5 and MyoD by different Wnts in explants of mouse paraxial mesoderm and the later activation of myogenesis in the absence of Myf5. *Development* **125**, 4155–4162.

Takemoto, T., Uchikawa, M., Yoshida, M., Bell, D. M., Lovell-Badge, R., Papaioannou, V. E. and Kondoh, H. (2011). Tbx6-dependent Sox2 regulation determines neural or mesodermal fate in axial stem cells. *Nature* **470**, 394–398.

Tan, M. H., Au, K. F., Yablonovitch, A. L., Wills, A. E., Chuang, J., Baker, J. C., Wong, W. H. and Li, J. B. (2013). RNA Sequencing Reveals a Diverse and Dynamic Repertoire of the Xenopus tropicalis Transcriptome Over Development. *Genome Res.* **23**, 201–216.

Tapscott, S. J. (2005). The circuitry of a master switch: Myod and the regulation of skeletal muscle gene transcription. *Development* **132**, 2685–95.

Therapy, G., Vol, M. B. and Schultz, M. C. (1999). Target of rapamycin (TOR) signaling coordinates tRNA and 5S rRNA gene transcription with growth rate in yeast Review Article. *Gene Ther. Mol. Biol.* **4**, 339–348.

Thisse, B., Pflumio, S., Fürthauer, M., Loppin, B., Heyer, V., Degrave, A., Woehl, R., Lux, A., Steffan, T. and Charbonnier, X. Q. (2001). Expression of the zebrafish genome during embryogenesis. *ZFIN Direct Data Submiss.* **http://zfin.org**,.

Tichy, E. D., Sidibe, D. K., Tierney, M. T., Stec, M. J., Sharifi-Sanjani, M., Hosalkar, H., Mubarak, S., Johnson, F. B., Sacco, A. and Mourkioti, F. (2017). Single Stem Cell Imaging and Analysis Reveals Telomere Length Differences in Diseased Human and Mouse Skeletal Muscles. *Stem Cell Reports* **9**, 1328–1341.

Topczewska, J. M., Topczewski, J., Shostak, A., Kume, T., Solnica-Krezel, L. and Hogan, B. L. M. (2001). The winged helix transcription factor Foxc1a is essential for somitogenesis in zebrafish. *Genes Dev.* **15**, 2483–2493.

Topisirovic, I. and Sonenberg, N. (2014). Distinctive tRNA Repertoires in Proliferating versus Differentiating Cells. *Cell* **158**, 1238–9.

Tsang, C. K., Liu, H. and Zheng, X. F. S. (2010). mTOR binds to the promoters of RNA polymerase I- and III-transcribed genes. *Cell Cycle* **9**, 953–7.

Upadhya, R., Lee, J. and Willis, I. M. (2002). Maf1 Is an Essential Mediator of Diverse Signals that Repress RNA Polymerase III Transcription. *Mol. Cell* **10**, 1489–1494.

Vannini, A., Ringel, R., Kusser, A. G., Berninghausen, O., Kassavetis, G. A. and Cramer, P. (2010). Molecular basis of RNA polymerase III transcription repression by Maf1. *Cell* **143**, 59–70.

Voronova, A., Coyne, E., Al Madhoun, A., Fair, J. V, Bosiljcic, N., St-Louis, C., Li, G., Thurig, S., Wallace, V. A., Wiper-Bergeron, N., et al. (2013). Hedgehog signaling regulates MyoD expression and activity. *J. Biol. Chem.* **288**, 4389–4404.

Wagner, D. S., Dosch, R., Mintzer, K. A., Wiemelt, A. P. and Mullins, M. C. (2004). Maternal control of development at the midblastula transition and

beyond: mutants from the zebrafish II. *Dev. Cell* **6**, 781–90.

Walston, J. D. (2012). Sarcopenia in older adults. *Curr. Opin. Rheumatol.* **24**, 623–627.

Wang, L.-H. and Baker, N. E. (2015). E Proteins and ID Proteins: Helix-Loop-Helix Partners in Development and Disease. *Dev. Cell* **35**, 269–280.

Wang, Y. and Jaenisch, R. (1997). Myogenin can substitute for Myf5 in promoting myogenesis but less efficiently. *Development* **124**, 2507–2513.

Wang, Z. and Roeder, R. G. (1997). Three human RNA polymerase III-specific subunits form a subcomplex with a selective function in specific transcription initiation. *Genes Dev.* **11**, 1315–26.

Wang, D., Valdez, M. R., Mcanally, J., Richardson, J. and Olson, E. N. (2001). The Mef2c gene is a direct transcriptional target of myogenic bHLH and MEF2 proteins during skeletal muscle development. *Development* **128**, 4623–4633.

Wasserman, W. W. and Fickett, J. W. (1998). Identification of regulatory regions which confer muscle-specific gene expression. *J. Mol. Biol.* **278**, 167–181.

Weintraub, H., Davis, R., Lockshon, D. and Lassar, A. (1990). MyoD binds cooperatively to two sites in a target enhancer sequence: occupancy of two sites is required for activation. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 5623–7.

Weintraub, H., Davis, R., Tapscott, S., Thayer, M., Krause, M., Benezra, R., Blackwell, T. K., Turner, D., Rupp, R., Hollenberg, S., et al. (1991). The myoD gene family: nodal point during specification of the muscle cell lineage. *Science* **251**, 761–6.

Welcker, M., Orian, A., Jin, J., Grim, J. E., Grim, J. A., Harper, J. W., Eisenman, R. N. and Clurman, B. E. (2004). The Fbw7 tumor suppressor regulates glycogen synthase kinase 3 phosphorylation-dependent c-Myc protein degradation. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 9085–9090.

White, R. J. (2004). RNA polymerase III transcription and cancer. *Oncogene* **23**, 3208–16.

White, R. J. (2005). RNA polymerases I and III, growth control and cancer. *Nat. Rev. Mol. Cell Biol.* **6**, 69–78.

White, R. J. (2008). RNA polymerases I and III, non-coding RNAs and cancer. *Trends Genet.* **24**, 622–9.

White, R. J. (2011). Transcription by RNA polymerase III: more complex than we thought. *Nat. Rev. Genet.* **12**, 459–63.

White, R. J., Stott, D. and Rigby, P. W. (1989). Regulation of RNA polymerase III transcription in response to F9 embryonal carcinoma stem cell differentiation. *Cell* **59**, 1081–1092.

Williamson, D., Lu, Y.-J., Fang, C., Pritchard-Jones, K. and Shipley, J. (2006). Nascent pre-rRNA overexpression correlates with an adverse prognosis in alveolar rhabdomyosarcoma. *Genes, Chromosom. Cancer* **45**, 839–845.

Wilusz, J. E. (2015). Removing roadblocks to deep sequencing of modified RNAs. *Nat. Methods* **12**, 821–2.

Winter, A. G., Sourvinos, G., Allison, S. J., Tosh, K., Scott, P. H., Spandidos, D. A. and White, R. J. (2000). RNA polymerase III transcription factor TFIIIC2 is overexpressed in ovarian tumors. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 12619–24.

Wolffe, A. and Brown, D. (1988). Developmental regulation of two 5S ribosomal RNA genes. *Science (80-.).* **241**, 1626–1632.

Wong, E. H. M., Smith, D. K., Rabadan, R., Peiris, M. and Poon, L. L. M. (2010). Codon usage bias and the evolution of influenza A viruses. Codon Usage Biases of Influenza Virus. *BMC Evol. Biol.* **10**, 253.

Wong, R. C.-B., Pollan, S., Fong, H., Ibrahim, A., Smith, E. L., Ho, M., Laslett, A. L. and Donovan, P. J. (2011). A novel role for an RNA polymerase III subunit POLR3G in regulating pluripotency in human embryonic stem cells. *Stem Cells* **29**, 1517–27.

Wood, W. M., Etemad, S., Yamamoto, M. and Goldhamer, D. J. (2013). MyoD-expressing progenitors are essential for skeletal myogenesis and satellite cell development. *Dev. Biol.* **384**, 114–127.

Wormington, W. M. and Brown, D. D. (1983). Onset of 5 S RNA gene regulation during Xenopus embryogenesis. *Dev. Biol.* **99**, 248–57.

Yan, B., Neilson, K. M. and Moody, S. A. (2009a). Notch signaling downstream of foxD5 promotes neural ectodermal transcription factors that inhibit neural differentiation. *Dev. Dyn.* **238**, 1358–1365.

Yan, B., Neilson, K. M. and Moody, S. A. (2009b). foxD5 plays a critical upstream role in regulating neural ectodermal fate and the onset of neural differentiation. *Dev. Biol.* **329**, 80–95.

Yang, J., Hung, L.-H., Licht, T., Kostin, S., Looso, M., Khrameeva, E., Bindereif, A., Schneider, A. and Braun, T. (2014). RBM24 Is a Major Regulator of Muscle-Specific Alternative Splicing. *Dev. Cell* **31**, 87–99.

Yao, Z., Farr, G. H., Tapscott, S. J. and Maves, L. (2013a). Pbx and Prdm1a transcription factors differentially regulate subsets of the fast skeletal muscle program in zebrafish. *Biol. Open* **2**, 546–55.

Yao, Z., Fong, A. P., Cao, Y., Ruzzo, W. L., Gentleman, R. C. and Tapscott, S. J. (2013b). Comparison of endogenous and overexpressed MyoD shows enhanced binding of physiologically bound sites. *Skelet. Muscle* **3**,.

Yin, H., Price, F. and Rudnicki, M. A. (2013). Satellite cells and the muscle stem cell niche. *Physiol. Rev.* **93**, 23–67.

Zetser, A., Gredinger, E. and Bengal, E. (1999). p38 Mitogen-activated Protein Kinase Pathway Promotes Skeletal Muscle Differentiation. *J. Biol. Chem.* **274**, 5193–5200.

Zheng, G., Qin, Y., Clark, W. C., Dai, Q., Yi, C., He, C., Lambowitz, A. M. and Pan, T. (2015). Efficient and quantitative high-throughput tRNA sequencing. *Nat. Methods* **12**, 835–7.

Zhu, Z. and Miller, J. B. (1997). MRF4 can substitute for myogenin during early stages of myogenesis. *Dev. Dyn.* **209**, 233–241.