Transcriptional regulation by FGF in the switch from pluripotency to skeletal muscle lineage commitment

Laura May-Seen Cowell

PhD

University of York

Biology

September 2023

Abstract

During development, key signalling pathways activate transcription factor regulators that direct pluripotent cells to specific cell fates. The bHLH transcription factor Myod was previously referred to as the "master regulator" of the muscle lineage due to its ability to convert fibroblasts to myoblasts. However, overexpression of Myod in mouse embryonic stem (ES) cells or *Xenopus* pluripotent explants is not sufficient for muscle differentiation. This indicates additional factors are needed for pluripotent cells to become competent to form muscle.

Fibroblast growth factor (FGF) is required for mesodermal gene expression and *Xenopus* skeletal muscle development, and has been implicated in the progression of naïve pluripotency to lineage competence. My hypothesis is that FGF signalling promotes activation of genes required to allow cells to transition from pluripotency to cell lineage commitment.

To investigate the role of FGF, a skeletal muscle inducing protocol was developed in *Xenopus laevis* pluripotent explants expressing Myod and treated with Fgf4. RNA-seq analysis allowed characterisation of the tissues induced by co-expression of Myod + Fgf4. In keeping with my hypothesis, a pattern of gene expression characteristic of fast twitch skeletal muscle was observed. RNA-seq analysis at three developmental time points allowed identification of potential regulators, which were tested for their ability to replace Fgf4 in the skeletal muscle protocol. Using this approach, *tcf12* was identified as an FGF regulated gene that is a promising candidate for a myogenic feedforward transcriptional pathway with Myod, downstream of FGF.

Differentiation of human skeletal muscle progenitors from H9 ES cells revealed that candidate gene expression was conserved in human myogenesis. Furthering our understanding of the regulation of skeletal muscle differentiation may help improve myoblast cell culture methods, or aid identification of potential therapeutic targets for future muscle wasting disease treatments.

Contents

List of Figures

List of Tables

Acknowledgements

Firstly, I would like to thank my supervisors Dr Harv Isaacs and Prof. Betsy Pownall for their invaluable support, guidance and advice throughout my PhD. Whether I was stressed, dwelling on a negative result, or celebrating a successful western, you were always there to encourage me. I could not have asked for better or more supportive supervisors.

Thank you to Prof. David Kent for being a member of my thesis advisory panel – your time, constructive feedback and suggestions have been greatly appreciated.

I would also like to thank my collaborators at the University of Edinburgh, Prof. Jenny Nichols and Dr Lawrence Bates, without whom I would not have had the resources or expertise to test my hypothesis on human ES cells.

In addition, I would like to thank all members of the Frog lab during my PhD, for their help, knowledge, and always making me laugh.

Finally, I would like to thank my partner, Alex, and my family for their support, pep talks, and always believing in me. Your encouragement, hugs and video calls have helped make the completion of this thesis possible.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. With the exception of in situ hybridisation of *tcf12* (carried out by Jennika Bates as part of her MBiol project), work presented in this thesis has not previously been presented for a degree or other qualification at this University, or elsewhere. Animal cap pluripotent cells were dissected by Harv Isaacs and I. Poly(A) library preparation of RNA-seq samples was carried out by the University of York Technology Facility. Initial quality control and mapping of sequences to the *Xenopus laevis* transcriptome was undertaken by Katherine Newling at the University of York Technology Facility. All sources are acknowledged as references.

Chapter 1: Introduction

1.1 Single cell to complex multicellular organism

During embryonic development, a single cell, the fertilised egg, grows into a multicellular organism comprised of many different types of cells. Cells in the early embryo are pluripotent stem cells, with capacity to give rise to all cell types in the developing and adult organism (Young, 2011). Individual cells are progressively restricted in their possible cell fates as development proceeds, and key signalling pathways activate transcription factor regulators to direct pluripotent cells to specific lineages (Gilbert, 2000). Differential gene expression regulates cell interactions, proliferation, specialisation and movement, in order to form an organism with the correct number of each cell type in the appropriate location. The mechanisms required for stem cells to undergo the change from pluripotency to lineage commitment are not fully understood. This thesis focusses on skeletal muscle lineage specification as this is one of the most well studied lineages.

1.2 Discovery of Myod and its role in skeletal muscle lineage commitment

In 1979, it was first shown that treatment of the mouse fibroblast cell line 10T1/2 with the demethylating agent 5-azacytidine can convert fibroblasts to myoblasts (Taylor and Jones, 1979). Genomic DNA isolated from 5-azacytidine-induced myoblasts and transfected into untreated 10T1/2 cells was subsequently shown to also be sufficient for myoblast induction (Lassar et al., 1986). Using this system, Davis, Weintraub and Lassar cloned *myod1* cDNA and demonstrated the ability of this single cDNA to convert a number of cell lines to skeletal muscle, with varying levels of efficiency (Davis et al., 1987). Due to its ability to induce expression of muscle-specific genes in fibroblasts and differentiated melanoma, neuroblastoma, liver, and adipocyte lines, Myod became known as the "master regulator" of the myogenic programme (Davis et al., 1987; Weintraub et al., 1989). It has since been shown that Myod drives chromatin remodelling during transdifferentiation of fibroblasts to myoblasts, and that this precedes transcriptional changes (Dall'agnese et al., 2019). However, while Myod expression was shown to convert 53% of 10T1/2 colonies to muscle, Myod expression or 5-azacytidine treatment in other cell lines showed a reduced frequency of myogenesis (Davis et al., 1987; Taylor and Jones, 1979). 10T1/2 cells provide a permissive environment such that demethylation of the *myod1* gene is sufficient to activate its expression and induce myogenesis, but this is not the case for all cell types.

Overexpressing Myod in embryonic stem (ES) cells leads to transcription of the *myogenin* and *myosin light chain 2* genes but is not sufficient to induce muscle differentiation suggesting that additional factors are needed for effective myogenesis (Dekel et al., 1992). In keeping with this, overexpressing Myod in *Xenopus laevis* (*X. laevis*) embryos leads to wider target gene expression

patterns than normal, however expression is still restricted to the early mesoderm, where FGF is also present (Maguire et al., 2012). Additionally, expression of Myod in *X. laevis* pluripotent animal pole explants (animal caps) leads to transcription of some muscle genes but is not sufficient for muscle differentiation (Hopwood and Gurdon, 1990). It has been proposed that a 'recognition factor' able to interact with both the Myod DNA-binding site and the MyoD basic regions is required for muscle-specific transcriptional activation by Myod (Weintraub et al., 1991). Absence of this recognition factor may be the reason why Myod alone does not activate myogenic genes in particular cell lines or explants.

1.3 Myogenic regulatory factors

Commitment to the skeletal muscle lineage is regulated by the basic helix-loop-helix (bHLH) myogenic regulatory factors (MRFs): Myod, Myf5, Mrf4/Myf6 and Myogenin (Wright et al., 1989; Braun et al., 1989; Rhodes and Konieczny, 1989). MRFs are muscle-specific transcription factors and operate in a highly regulated manner to specify and drive terminal differentiation of muscle in all vertebrates (Pownall et al., 2002). Studies investigating MRF expression patterns in embryos, gene targeting studies in mice, and transcriptional profiling in cell culture, have contributed greatly to our current understanding of their functions and roles (Hernández-Hernández et al., 2017).

In *Xenopus*, *myf5* and *myod* are first expressed during gastrulation in the mesoderm prior to muscle differentiation (Hopwood et al., 1991, 1989) , whereas expression is first detected in the somites of amniote embryos (Pownall et al., 2002). Not only are the MRF genes expressed specifically in the muscle lineage of all vertebrates, these gene are required for the establishment of the myogenic lineage in vivo. Myf5/Myod null mice show a complete loss of skeletal muscle, however expression of either *myf5* or *myod* is sufficient to partially rescue this phenotype (Rudnicki et al., 1993; Haldar et al., 2008). This indicates that Myf5 or Myod activity is required for determination of skeletal muscle and the factors can, at least in part, functionally substitute for each other (Rudnicki et al., 1993). In contrast, Myogenin (Myog) null mice show a severe reduction of all skeletal muscle, despite normal myoblast specification (Hasty et al., 1993; Nabeshima et al., 1993). This suggests that *myog* is responsible for terminal myoblast differentiation rather than early lineage specification (Weintraub, 1993; Rawls et al., 1995).

In *Xenopus*, *myf6* (also known as *mrf4*) is first expressed in differentiated anterior myocytes when muscle-specific *myh4* mRNA is also present (Gaspera et al., 2006). However in mice, *myf6* is expressed in the early ventrolateral dermomyotome, and then again in differentiated muscle fibres (Bober et al., 1991; Summerbell et al., 2002). The role of Myf6 is therefore less well-defined as it

may have roles in both myogenic determination and differentiation in some species (Kassar-Duchossoy et al., 2004; Hinterberger et al., 1991).

1.4 Myogenic regulatory factor binding partners and co-factors

MRFs are transcriptional activators that form heterodimers with members of the E-protein bHLH subfamily (E2A/Tcf3, E2-2/Tcf4 and HEB/Tcf12) through interaction of their HLH domains, and bind consensus E-box sequences (CANNTG) at target promoter regions (Lassar et al., 1991; Murre et al., 1989; Hu et al., 1992). Myod forms relatively stable complexes with paired E-boxes to activate reporters, though co-factors such as Mef2, Pbx and Meis, or Sp1 can functionally substitute for the second E-box (Biesiada et al., 1999; Knoepfler et al., 1999). Interaction with co-factors may contribute to promoter binding specificity and alter the conformation of the Myod complex to expose other regions for interaction (Bengal et al., 1994). For example, Pbx (pre-B-cell leukemia homeobox) proteins form complexes with Meis homeobox proteins and act as pioneer transcription factors to mark genes for activation by Myod (Berkes et al., 2004). Pbx-Meis binding sites are enriched at Myod-specific targets and are required for the myogenic potential of Myod (Fong et al., 2012, 2015).

The Mef2 (myocyte enhancer factor 2) family of MADS-box containing transcription factors contribute to developmental regulation of multiple cell lineages including muscle, neural, blood and immune cells. Mef2c expression is directly activated by MRFs early in development, and subsequently functions synergistically with MRFs to drive myogenesis (Wang et al., 2001; Dodou et al., 2003). For example, Mef2c has been shown to act with Myog to activate the *mrf4* promoter (Naidu et al., 1995). Many Mef2 binding sites are located close to E-boxes at myogenic target genes for cooperative binding with MRFs (Wasserman and Fickett, 1998).

1.5 Regulation of skeletal muscle gene transcription by Myod

Myod directly binds regulatory elements of genes expressed at different times throughout the entire programme of muscle specification and differentiation, functioning via a feed-forward mechanism (Tapscott, 2005). Factors induced by Myod feed-forward to regulate Myod activity at subsequent target genes, and temporally pattern the relative timing of gene expression during myogenesis (Penn et al., 2004; Bergstrom et al., 2002). For example, one model suggests that chromatin arrangement in non-muscle cells makes the *myogenin* E-box unavailable to MRFs. In early differentiation, Myod first activates the p38 MAPK pathway and expression of specific Mef2 isoforms. Myod then interacts with the Pbx complex adjacent to the *myogenin* promoter and recruits a histone deacetyltransferase (HDAC) (Mal and Harter, 2003). The HDAC is replaced with histone acetyltransferase p300 which recruits p300/CBP associated factor (PCAF). PCAF then acetylates Myod at lysine residues near its DNA binding domain which, along with accumulated p38

activity, allows recruitment of the Swi/Snf chromatin remodelling complex (de la Serna et al., 2005; Simone et al., 2004). Remodelling of the locus exposes binding sites for factors such as early Myod target Mef2d and exposes the E-box for stable Myod-DNA binding. Mef2d then recruits RNA polymerase II (Pol II) to transcribe late-stage myogenic genes (Penn et al., 2004). Both Myod and Myf5 have been shown to directly activate muscle lineage gene transcription, in addition to their ability to remodel chromatin to enable Myogenin to drive further transcription (Singh and Dilworth, 2013).

Investigation of the transcriptional targets of MRFs in cultured cells has revealed much about their activity in myogenesis (Blais et al., 2005; Conerly et al., 2016; Fong and Tapscott, 2013). The distinct transcriptional activity of Myf5 compared to Myod indicates an early lineage establishment role for both factors, but a more powerful transcriptional activation activity for Myod because of its ability to strongly recruit Pol II (Conerly et al., 2016). Adding to complexity, there are other bHLH transcription factors which direct the establishment of other cell lineages, for instance Neurod is associated with neurogenesis (Seo et al., 2007a). Myod and Neurod2 bind a 'shared' E-box sequence and a 'private' sequence specific to its associated lineage programme. Swapping the DNA binding domain of Myod to that of Neurod2, allowed the chimeric mutant to gain binding to Neurod2 private sites, but it maintained binding to a subset of Myod-specific genes unless additional point mutations to prevent interaction with Pbx/Meis were introduced (Fong et al., 2015). These data highlight the importance of the interaction of Myod with key partners in order to drive skeletal muscle development.

1.6 FGF signalling and the transition from pluripotency to lineage competence

The ability for a cell to commit to a specific lineage, such as the skeletal muscle cell lineage, requires that it exit the pluripotent state. Embryonic stem cells are pluripotent, self-renewing cells derived from the inner cell mass of the developing blastocyst, with the innate ability to become any cell type within the body (Young, 2011). ES cells can be maintained in culture in a pluripotent self-renewing state or induced to differentiate towards a particular lineage. For example, fibroblast growth factor (FGF) signalling has been shown to be required for mesoderm, endoderm and neural differentiation in mouse ES cells, whereas inhibiting downstream ERK supports self-renewal (Ying et al., 2008; Ma et al., 2016; Villegas et al., 2010).

ES cells must progress from self-renewing, naïve pluripotency to a formative pluripotent state, before transition to primed pluripotency where cells are partially specified (Smith, 2017). The formative pluripotent state involves reconfiguration of the gene regulatory network (GRN) to prepare the cells to acquire competence for multi-lineage induction. The PEA3 (polyoma enhancer

3) subfamily of ETS (E26 transformation-specific) transcription factors (PEA3/Etv4, Etv5 and Etv1), have been implicated in programming the genetic landscape for cell fate determination downstream of FGF-ERK signalling (Garg et al., 2018). For example, there is evidence that progression of naïve pluripotency to lineage competence requires phosphorylation of Etv5 by ERK in order to alter the transcription factor's activity (Kalkan et al., 2019). Phosphorylated Etv5 no longer supports transcription of genes required for self-renewal and instead relocates to activate formative pluripotency enhancers (Kalkan et al., 2019). Therefore, low threshold activation of the FGF-ERK-Etv5 pathway may play a key role in the formative pluripotency transition (Smith, 2017).

Inhibition of FGF signalling in *Xenopus* embryos downregulates expression of at least two genes implicated in regulating pluripotency: *lin28* and *forkhead box d3* (*foxd3*) (Branney et al., 2009). These FGF targets have both been associated with ERK promoting the transition from naïve to primed pluripotency in mouse ES cells (Zhang et al., 2016; Tsanov et al., 2017; Krishnakumar et al., 2016; Respuela et al., 2016).

Additionally, ERK has also been shown to regulate transcription in ES cells by triggering the reversible association and disassociation of Pol II and co-factors from genes and enhancers (Hamilton et al., 2019). Persistent ERK signalling prevents further pluripotency transcription factor expression. As protein turnover continues, pluripotency transcription factor levels decline, allowing irreversible gene silencing and lineage commitment (Hamilton et al., 2019). This project investigates the hypothesis that FGF signalling is required for the competence of embryonic cells to be directed to the skeletal muscle lineage.

1.7 Discovery of the role of FGF signalling in mesoderm induction

One reason for the proposed role of FGF in skeletal muscle lineage specification, is its involvement in mesoderm induction, from which muscle is derived. Early experiments using urodele embryos showed that animal caps taken at blastula stage become ectoderm when cultured on their own, but will give rise to mesoderm when combined with vegetal cells (Nieuwkoop and Ubbels, 1969; Nieuwkoop, 1969). As the mesoderm is derived only from the animal cap cells, this indicates that a signal originating from the vegetal cells changes the fate of responding animal cap cells from ectoderm to mesoderm. This, in combination with other experiments, led to Slack's three-signal model (Figure 1.1). The model involves two signals from the vegetal pole, a ventral (VV) and dorsal (DV) signal, which respectively induce ventral and dorsal mesoderm in the cells above, and a third dorsalising signal from Spemann's organiser (O) which subsequently specifies the different types of mesoderm along the dorsoventral axis (Slack et al., 1987; Boterenbrood and Nieuwkoop, 1973).

FGF and activin were initially proposed as candidates for the endogenous, vegetally localised mesoderm inducers in *X. laevis*, due to their respective abilities to induce ventral and dorsal mesoderm in animal caps (Slack et al., 1987; Isaacs et al., 1992). However it was later shown that FGF acts at the marginal zone (equator) of the embryo as a competence factor, allowing cells to respond to the VV and DV signals (Cornell et al., 1995; Isaacs, 1997). Experiments contributing to this understanding include the fact that expression of a dominant negative truncated FGF receptor (XFD) results in loss of mesoderm and a subset of mesodermal genes (Amaya et al., 1991, 1993; Isaacs et al., 1994; Schulte-Merker and Smith, 1995). While FGF treatment of animal caps or vegetal cells induces expression of mesodermal markers *brachyury* (*tbxt)* and *myod,* activin treatment is only able to induce this expression in animal caps, and not vegetal cells (Cornell et al., 1995). As *tbxt* and *myod* are not normally expressed in vegetal cells, FGF could not be a major component of the endogenous vegetal signal. In addition to this, FGF is required for mesoderm induction by activin, further supporting the model in which FGF allows marginal zone cells to respond to mesoderm induction by vegetal signals (Cornell and Kimelman, 1994; LaBonne and Whitman, 1994). A set of genes expressed in the organiser, including *chordin, noggin* and *goosecoid*, have since been shown to be FGF signalling targets, reinforcing the fact that FGF is not the VV signal (Branney et al., 2009; Delaune et al., 2005; Fletcher and Harland, 2008). The VV and DV signals are now known to be transforming growth factor β (TGFβ) superfamily members: nodal, nodal1, nodal2, VegT, and Vg1 (Agius et al., 2000; Kofron et al., 1999; Birsoy et al., 2006; Weeks and Melton, 1987). Inhibition of FGF signalling has also been shown to prevent activation of specific mesodermal transcripts by VegT (Fletcher and Harland, 2008).

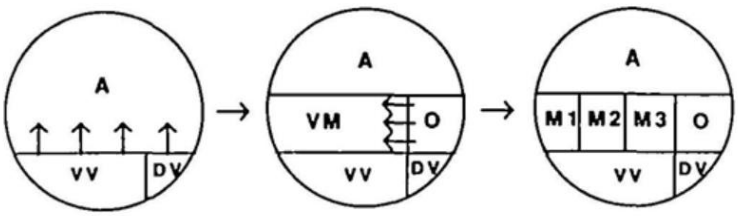


Figure 1.1: The three-signal model. Two mesoderm inducing signals originate from the vegetal hemisphere of the early blastula embryo. The dorsal-vegetal (DV) signal induces Spemann's organiser (O) in the dorsal marginal zone, and the ventral-vegetal (VV) signal induces ventral mesoderm (VM). During gastrula stages, the organiser emits a dorsalising signal to induce different mesodermal fates (M1, M2, M3) along the dorsoventral axis (Smith, 1989; Slack et al., 1987).

FGF is an autocrine regulator of the expression of mesodermal genes such as *tbxt, cdx4,* and the hox genes (Chung et al., 2004; Branney et al., 2009). Several FGF family members and key early mesodermal genes (e.g. *tbxt)* are co-expressed in the presumptive mesoderm in a ring around the blastopore (Lea et al., 2009; Branney et al., 2009; Fletcher and Harland, 2008). Overexpression of

tbxt also activates Fgf4 expression, indicating the existence of an autocatalytic loop involving Fgf4-Tbxt in mesoderm formation in the late blastula embryo (Isaacs et al., 1994; Schulte-Merker and Smith, 1995). Additionally, embryos in which the FGF signalling pathway or *tbxt* expression have been disrupted present similar phenotypes, lacking posterior structures and a differentiated notochord (Herrmann et al., 1990; Amaya et al., 1991; Isaacs et al., 1994). Establishment and maintenance of mesoderm requires a functional FGF signalling pathway as *tbxt* and *myod* expression, and muscle formation, is reduced in XFD embryos (Amaya et al., 1993; Isaacs et al., 1994). Disaggregated blastopore region explants also lose *tbxt* expression in the absence of FGF, indicating that maintenance of *tbxt* expression continues to require FGF during gastrula and neurula stages (Isaacs et al., 1994; Schulte-Merker and Smith, 1995). This reinforces the requirement for FGF cell-cell signalling in the maintenance of mesoderm gene expression and subsequent muscle differentiation (Isaacs et al., 1994; Gurdon et al., 1993; Standley et al., 2001).

1.8 The fibroblast growth factor family

FGF protein was first purified from bovine brain and found to exhibit a mitogenic effect on cultured fibroblasts, before further experiments revealed wider expression and roles during vertebrate development (Gospodarowicz, 1975; Denis Gospodarowicz, 1975; Esch et al., 1985; Isaacs, 1997). The FGF ligand family is comprised of 22 members which can be divided into 7 phylogenetic subfamilies, and further categorised into 3 groups: canonical, intracellular, and endocrine (Itoh and Ornitz, 2004). The canonical FGFs (FGF1-10, 16-18, 20 and 22) signal through the FGF receptor family of tyrosine kinase receptors and share a conserved core of 140 amino acids (Dorey and Amaya, 2010). Canonical FGFs tightly bind heparin sulphate proteoglycans (HSPGs) which limit diffusion through the extracellular matrix, and contribute to FGF receptor (FGFR) specificity and affinity (Ornitz, 2000). Intracellular FGFs (FGF11-14) share a high sequence homology with canonical FGFs, however they are not secreted and do not act via FGFRs (Beenken and Mohammadi, 2009). Endocrine FGFs (FGF19, 21 and 23) have a reduced HSPG binding affinity and instead bind Klotho molecules as receptor binding co-factors, and function in an endocrine manner to regulate adult homeostasis (Beenken and Mohammadi, 2009; Ornitz and Itoh, 2015).

1.9 The fibroblast growth factor receptor family

The four FGFRs (FGFR1-4) consist of a single transmembrane domain, an intracellular domain with tyrosine kinase activity and protein binding sites, and an extracellular region (Böttcher and Niehrs, 2005). The extracellular region is made up of three immunoglobulin-like domains, Ig1, Ig2 and Ig3, which are important for ligand and HSPG binding, and receptor dimerisation (Johnson et al., 1991). Adding to complexity, ligand binding specificity can also be regulated by alternate splicing of Ig3 to generate alternate receptor isoforms of FGFR1-3 (Table 1.1) (Johnson et al., 1991). Binding

specificity is regulated primarily through alternative splicing of Ig3 exons IIIb and IIIc. For example, epithelial cells express FGFR2b and FGFR3b with the IIIb exon, whereas mesenchymal lineages express the IIIc isoforms (Scotet and Houssaint, 1998; Wuechner et al., 1996). FGF ligand-receptor specificity can also be modulated by Ig2, as binding of ligands such as FGF10 elicit a conformational change which increases the number of binding contacts (Yeh et al., 2003).

FGFR isoform	**Ligand specificity**
FGFR1b	FGF1, 2, 3, 10, 22
FGFR1c	FGF1, 2, 4, 5, 6, 19, 20, 21
FGFR2b	FGF1, 3, 4, 6, 7, 10, 22
FGFR2c	FGF1, 2, 4, 5, 6, 8, 9, 17, 18, 19, 21, 23
FGFR3b	FGF1, 9
FGFR3c	FGF1, 2, 4, 8, 9, 17, 18 19, 21, 23
FGFR4	FGF1, 2, 4, 6, 8, 9, 16, 17, 18, 19
FGFRL	FGF2, 3, 4, 8, 10, 22

Table 1.1: Ligand specificity of the fibroblast growth factor receptor family (Tiong et al., 2013; Steinberg et al., 2010)

The final member of the FGFR family is FGFR-like (FGFRL) (Sleeman et al., 2001). The extracellular domain shares up to 50% amino acid identity with the other FGFRs, however FGFRL lacks a tyrosine kinase domain (Steinberg et al., 2010; Ornitz and Itoh, 2015). The receptor has a strong affinity for FGF ligands and was initially thought to function as a decoy receptor to further modulate FGF signalling (Trueb et al., 2003; Steinberg et al., 2010). However, FGFRL has since been shown to increase ERK signalling in pancreatic islet beta cells, via association of SHP-1 phosphatase with the receptors intracellular Src homology 2 (SH2) binding motif (Silva et al., 2013). This indicates that FGFRL is not simply a decoy receptor and instead functions as a non-tyrosine kinase signalling molecule (Silva et al., 2013; Ornitz and Itoh, 2015).

1.10 Canonical fibroblast growth factor signal transduction

Binding of FGF and HSPG co-factors to FGFR1-4 leads to receptor dimerisation and auto- and cross-phosphorylation of tyrosine residues within the intracellular domain (Schlessinger, 2000). Tyrosine phosphorylation creates docking sites for recruitment of signalling complexes in the cytoplasm, and signal transduction via three main pathways: the phospholipase Cγ (PLCγ), phosphoinositide-3 (PI3) kinase or extracellular signal-regulated kinase (ERK) pathway (Figure 1.2) (Yun et al., 2010).

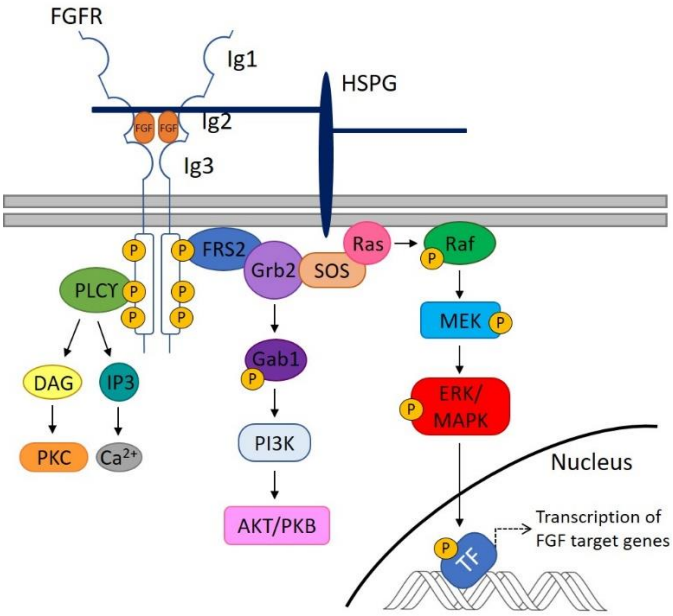


Figure 1.2: Schematic representation of FGF signal transduction via the PLCγ, PI3 kinase and ERK signalling pathways. Binding of extracellular fibroblast growth factors (FGFs) and accessory molecule heparin sulphate proteoglycan (HSPG) causes dimerisation of FGF receptors and autophosphorylation of intracellular tyrosine kinase domains. FGF signal transduction proceeds via 3 main pathways: the phospholipase Cγ (PLCγ), phosphoinositide-3 kinase (PI3K) or extracellular signal-regulated kinase (ERK) pathway (Cowell, 2019).

FGFR substrate 2 (FRS2) is an important target for tyrosine phosphorylation as part of both the PI3 and ERK signal transduction pathways. FRS2 associates with the receptor to allow recruitment of adaptor Grb2 (growth factor receptor bound protein 2) via its SH2 domain (Pawson et al., 1993; Kouhara et al., 1997; Ong et al., 2000; Hadari et al., 2001). Grb2 can recruit Gab1 (growth factor receptor bound protein 2 associated protein 1) for phosphorylation and subsequent recruitment and activation of PI3 kinase (Ong et al., 2001). Downstream of PI3 kinase, proto-oncogene AKT/protein kinase B is an important pathway mediator involved in regulating cell growth and survival (Nicholson and Anderson, 2002).

Alternatively, Grb2 can recruit nucleotide exchange factor son of sevenless (SOS) via its Src homology 3 (SH3) domain (Ong et al., 2000; Pawson et al., 1993). Grb2/SOS activates the small GTP binding protein Ras by promoting dissociation of GDP to allow conversion to the GTP-bound form. Activation of Ras stimulates a cascade of sequential phosphorylation of Raf, Mek and ERK (Schlessinger, 2000). Diphosphorylated ERK (dpERK) induces expression of FGF target genes through direct phosphorylation of transcription factors or indirectly via other kinases (Böttcher and Niehrs, 2005). FGF is the sole activator of the ERK cascade during *Xenopus* early development, a signal transduction pathway critical for normal development (Christen and Slack, 1999).

The PLCγ pathway has a key role in modulation of planar cell polarity for cell movements during gastrulation (Sivak et al., 2005). Binding of PLCγ to a conserved FGFR phosphotyrosine residue

activates the enzyme to hydrolyse phosphotidylinositol-4,5-diphosphate to inositol-1,4,5-trisphosphate (IP_3) and diacylglycerol (DAG) (Mohammadi et al., 1991; Peters et al., 1992). IP_3 stimulates release of calcium ions from the endoplasmic reticulum into the cytoplasm, and DAG activates protein kinase C (PKC).

1.11 Transcription factor effectors

A number of signalling pathways have dedicated downstream transcription factor effectors such as Wnt (TCF), Shh (Gli), and TGFβ/BMP/Activin/Nodal (SMADs) (Basson, 2012). ETS transcription factors have been shown to act as downstream effectors of FGF-ERK signalling (De Launoit et al., 1997; Münchberg and Steinbeisser, 1999). The PEA3 subfamily consists of ETS transcription factors PEA3/Etv4, Etv5 and Etv1, and expression correlates closely with FGF activity in multiple species (Chotteau-Lelièvre et al., 1997; Münchberg and Steinbeisser, 1999; Raible and Brand, 2001). *PEA3* and *etv5* expression is lost in zebrafish embryos treated with FGFR1 inhibitor SU5402, and ectopic expression is seen in the presence of FGF8/3 coated beads (Raible and Brand, 2001). Knockdown of the PEA3 subfamily in zebrafish has also been shown to result in phenotypes resembling FGF deficient embryos with significantly reduced expression of FGF target genes (Znosko et al., 2010). FGF also regulates its own activity through complex negative feedback loops involving the PEA3 subfamily and other ETS factors (Znosko et al., 2010; Garg et al., 2018). For example, Ets2 mediates FGF signalling by directly binding *dual-specificity phosphatase 6* (*dusp6*) promoters to inactivate ERK targets (Ekerot et al., 2008).

However, FGF signal transduction is not exclusively carried out by ETS effectors as dpERK also phosphorylates kinases such as p90RSK (p90 ribosomal S6 kinase) and other transcription factors (Cargnello and Roux, 2011). Both dpERK and p90RSK are able to translocate into the nucleus and alter the activity of transcription factors. For example, dpERK can activate the transcription factor *Elk-1* which is involved in expression of immediate-early genes such as *fos* (Cargnello and Roux, 2011). dpERK and p90RSK can also directly phosphorylate Fos protein, thus stabilising the protein and allowing association with Jun to form transcription factor AP-1 complexes (Murphy et al., 2002). AP-1 induces FGF target *tbxt* and has an important role in the Fgf4-Tbxt autocatalytic loop during mesoderm induction and maintenance (Lee et al., 2011; Isaacs et al., 1994; Schulte-Merker and Smith, 1995). Additionally, a number of transcription factors that are transcribed directly downstream of FGF, such as *tbxt, cdx4* and *myod*, initiate further transcription of target genes in a feedforward mechanism (Smith et al., 1991; Isaacs et al., 1998; Fisher et al., 2002).

Another possible mechanism through which FGF has been shown to regulate target gene transcription is via de-repression. Treatment of pluripotent *X. laevis* animal cap explants with Fgf4

increases expression of FGF targets *myod* and *Cdx1*, *2* and *4*, even in the presence of translational inhibitor cycloheximide (CHX) (Fisher et al., 2002; Keenan et al., 2006). This indicates that activation of a subset of FGF target genes is an immediate early response not reliant on translation of any other proteins. CHX treatment alone activates some transcription of *myod* and *cdx* genes suggesting that target gene expression involves inhibition of a labile transcription repressor, the levels of which rapidly decay in the absence of protein synthesis. One example of this potential mechanism is transcriptional repressor transducin-like enhancer of split 4/groucho-related gene 4 (TLE4/Grg4), which can be partially deactivated by Fgf4 via an ERK consensus site in Grg4 (Burks et al., 2009). Another transcriptional repressor implicated in regulation of a subset of FGF targets is Capicua (CIC). Transcription of PEA3 ETS factors is repressed by CIC in the absence of dpERK, $p90^{RSK}$ and 14-3-3 proteins (Dissanayake et al., 2011; Jiménez et al., 2012).

1.12 FGF4 and skeletal muscle lineage commitment

In addition to its roles promoting transition out of the naïve pluripotent state, mesoderm specification, and maintenance of mesodermal gene expression, FGF is directly implicated in skeletal muscle development in *Xenopus*. For example, inhibition of FGF signalling via XFD leads to a loss of skeletal muscle in *Xenopus* (Amaya et al., 1991, 1993).

Early studies in *Xenopus* animal caps determined that Fgf1, Fgf2, Fgf4 and activin are capable of inducing mesoderm, however these inducing factors vary in potency and expression pattern (Slack et al., 1989; Smith et al., 1990). For instance, activin is at least 40 times more active in inducing muscle than FGF2, and has been shown to directly activate Fgf4 as an immediate early response (Green et al., 1990; Fisher et al., 2002). Fgf4 is also over 100 times more effective at inducing mesoderm than Fgf2 (Isaacs et al., 1994). This may be due to the fact that Fgf4 is secreted more effectively due to possession of a signal sequence (Thompson and Slack, 1992). Unlike Fgf1 and Fgf2, Fgf4 is classically secreted and expression increases in the marginal zone during gastrulation, indicating a role in endogenous mesoderm induction and maintenance (Slack et al., 1989; Isaacs et al., 1992; Lea et al., 2009). Antisense morpholino oligonucleotide inhibition of Fgf4 also showed that Fgf4 is specifically required for the initial activation of *myod* expression in the myogenic cell lineage (Fisher et al., 2002). Cells in the early mesoderm of the *Xenopus* gastrula embryo express high levels of *myod* mRNA and protein before differentiation (Hopwood et al., 1989, 1992). A single cell taken from this zone will differentiate into skeletal muscle when transplanted to a ventral region (Kato and Gurdon, 1993). This indicates that these Fgf4-Myod expressing mesoderm cells represent a population of skeletal muscle precursors already committed to the muscle lineage.

The community effect means that muscle precursor cells will only differentiate and stably express tissue-specific genes if they are in contact with each other (Gurdon et al., 1993). However disaggregated muscle precursor cells are able to maintain high levels of Myod and Myf5, and differentiate when in the presence of myogenic community factor Fgf4 (Standley et al., 2001). However, Fgf2, Wnt8, Bmp4 and Tgfβ2 are not able to substitute for the endogenous community effect in dispersed cells. Additionally, inhibiting FGF signalling in reaggregated muscle precursor cells prevents expression of myogenic genes, despite the fact the cells are in contact with each other (Standley et al., 2001). These data indicate that Fgf4 is a candidate for the competence or recognition factor needed to allow the co-ordinate transcription of the many contractile protein genes required for the differentiation of skeletal muscle.

There is additional evidence for a role for FGF signalling during myogenic lineage specification in other vertebrates including chick (Seed and Hauschka, 1988; Marcelle et al., 1995), zebrafish (Osborn et al., 2020; Groves et al., 2005), and mouse (Han and Martin, 1993). For example, embryonic expression of FGF6 in the mouse is restricted to the skeletal muscle lineage, and FGF8 is required in the zebrafish lateral somite for *myod* expression and terminal differentiation (Han and Martin, 1993; Groves et al., 2005). Zebrafish myogenesis is initiated by FGF-dependent Tbx16 and Tbxta activities, which directly activate *myf5* and *myod* expression (Osborn et al., 2020). For this study, the *Xenopus* animal cap provides a simple and efficient source of pluripotent stem cells to investigate the role of FGF signalling during the acquisition of the skeletal muscle cell fate.

1.13 Project aims

This project uses *X. laevis* pluripotent animal cap explants to investigate the hypothesis that FGF signalling acts as a gatekeeper during cell lineage specification to allow differentiation of pluripotent cells into skeletal muscle cells.

The aims of this project are to:

- Develop a protocol to induce skeletal muscle formation from pluripotent animal cap cells using Fgf4 protein and *myod1.S* mRNA
- Determine the transcriptional output of the myogenic programme initiated by the protocol using RNA-seq analysis and characterise the type of muscle induced
- Investigate the role of Fgf4 in lineage specification during the protocol
- Identify candidate genes potentially involved in skeletal muscle specification and investigate whether they are sufficient to replace Fgf4 in the protocol
- Differentiate skeletal muscle progenitors from human H9 ES cells and test whether the roles of candidate genes are conserved.

Chapter 2: Materials and methods

2.1 Embryological methods

2.1.1 *Xenopus laevis* embryo culture

Embryos were staged according to Nieuwkoop and Faber (1994). *X. laevis* embryos were cultured in 1/10 strength Normal Amphibian Medium (NAM)/10 with microinjections carried out in NAM/3 + 5% ficoll using a Drummond Microinjector.

2.1.2 Animal cap organoids

At late blastula stage (stage 8-9), mounted tungsten needles were used to dissect the animal pole cells of de-membraned *X. laevis* embryos in NAM/2. After approximately 10 minutes in NAM/2, animal caps were transferred to NAM/2 + 5mg/ml BSA, with or without recombinant *Xenopus* Fgf4 protein (Isaacs et al., 1992), until the required stage.

2.1.3 Photography

Images were taken using the SPOT 14.2 Colour Mosaic camera (Diagnostic Instruments Inc.) and SPOT Advanced software. Embryos and organoids were imaged using a Leica MZ FLIII microscope and organoid sections using a LEICA DM2500 microscope. Images were processed using Adobe Photoshop CS3 or Adobe Photoshop Elements 2019.

2.2 Molecular biology methods

2.1.1 Agarose gel electrophoresis

DNA and RNA samples were separated on 0.8-1.5% agarose gels in Tris-Acetate-EDTA buffer (40mM Tris pH7.6, 20mM acetic acid, 1mM EDTA) stained with ethidium bromide. Samples were loaded with a 6x gel loading buffer (NEB) and run alongside 1kb plus DNA ladder (NEB).

2.2.1 Total RNA extraction

Flash frozen *X. laevis* embryos and animal cap organoids were homogenised in 1ml Tri-Reagent (Sigma-Aldrich) and left on ice for 1 minute. Samples were centrifuged at 13,000rpm for 10 minutes at 4°C and the supernatant placed at room temperature for 5 minutes. 200µl chloroform was added to the supernatant and left at room temperature for 5 minutes before being centrifuged at 13,000rpm for 15 minutes at 4°C. The aqueous phase was transferred to a new Eppendorf and 200µl chloroform added before being centrifuged at 13,000rpm for 5 minutes at 4°C. 500µl isopropanol was added to the aqueous phase, vortexed and placed at -20°C for 29 minutes. Samples were centrifuged at 13,000rpm for 15 minutes at 4°C and the supernatant discarded. 200µl ice cold 70% ethanol was added to the RNA pellet, vortexed and centrifuged at 13,000rpm for 10 minutes at 4°C. The supernatant was discarded and the pellet dried by desiccation.

X. laevis RNA for RNA-seq and qPCR was resuspended in 50µl dH_2O and purified using Zymo-spin columns. Samples were DNase I treated for 15 minutes in-column. RNA integrity number (RIN) was measured using the Agilent 2100 Bioanalyzer. Poly(A) mRNA isolation and library preparation was carried out by staff in the University of York Technology Facility using the NEBNext Poly(A) mRNA Magnetic Isolation Module and NEBNext Ultra™ II Directional RNA Library Prep Kit for Illumina. Sequencing was carried out by Illumina NovaSeq with 7.3-20.7 million paired end reads per sample.

2.2.2 cDNA synthesis

cDNA was synthesised from 500ng RNA with 1µl 10µM dNTP (Thermo Scientific), 1µl random hexamers (Promega) made up to 18µl with dH_2O. Reaction mixture was heated to 65°C for 5 minutes before addition of 5µl 5x SSIV buffer (Invitrogen), 1µl 100mM DTT (Invitrogen) and 1µl 200U/µl SuperScript IV Reverse Transcriptase (Invitrogen). The mixture was then incubated at 23°C for 10 minutes, 55°C for 10 minutes and then 80°C for 10 minutes.

2.2.3 Quantitative PCR

For 96-well plate qPCR reactions, 50µl dH_2O was added to 25µl cDNA. *X. laevis* cDNA was amplified using primers listed in Table 2.1. qPCR reactions consisted of 3µl cDNA, 10µl 2X Fast SYBR Green Master Mix (Thermo Fisher Scientific), 0.5µl 10µM forward primer, 0.5µl 10µM reverse primer, and dH_2O up to a total volume of 20µl in a 96-well 0.1ml plate. Reactions were run on QuantStudio™ 3 (Thermo Fisher Scientific) instrument and heated to 95°C for 20 seconds before 40 cycles of 95°C for 1 second and 60°C for 20 seconds, followed by a melt curve analysis.

Gene	**Forward primer sequence 5'-3'**	**Reverse primer sequence 5'-3'**
dicer	GGCTTTTACACATGCCTCTTACC	GTCCAAAATTGCATCTCCAAG
myh4	GTGCGTTGTTTGATTCCCAAT	GCTGGTGGATGAGGAGATGGT
act3	TCACAACAGCTGAAAGGGAGAT	AAGTCCAGAGCCACATAGGC

Table 2.1: Forward and reverse primer sequences for qPCR analysis of *Xenopus* samples

For 384-well plate qPCR reactions, 175µl dH_2O was added to 25µl cDNA. Human cDNA was amplified using primers listed in Table 2.2. qPCR reactions consisted of 3µl cDNA, 5µl 2X Fast SYBR Green Master Mix (Thermo Fisher Scientific), 0.25µl 10µM forward primer, 0.25µl 10µM reverse primer, and dH_2O up to a total volume of 10µl in a 384-well plate. Reactions were run on QuantStudio™7 Pro (Thermo Fisher Scientific) instrument and heated to 95°C for 20 seconds before 40 cycles of 95°C for 1 second and 60°C for 20 seconds, followed by a melt curve analysis.

Gene	Forward primer sequence 5'-3'	Reverse primer sequence 5'-3'
ACTB	GTGGATCAGCAAGCAGGAGT	GCAACTAAGTCATAGTCCGC
MYOD1	AATAAGAGTTGCTTTGCCAG	GTACAAATTCCCTGTAGCAC
MYF5	AATTTGGGGACGAGTTTGTG	CATGGTGGTGGACTTCCTCT
MYH3	ACGACTACCCGTTCATTAGCC	GGATGTCAATGGCGCTGTCT
MYOG	GCTGTATGAGACATCCCCCTA	CGACTTCCTCTTACACACCTTAC
PAX7	CCCCCGCACGGGATT	TATCTTGTGGCGGATGTGGTTA
PAX3	ATCAACTGATGGCTTTCAAC	CAGCTTGTGGAATAGATGTG
TBXT	TGCTTCCCTGAGACCCAGTT	GATCACTTCTTTCCTTTGCATCAAG
NANOG	GACAGTCTCCGTGTGAGGCAT	CCTGTGATTTGTGGGCCTG
TCF12	CCAGCAGTTCACCTTACGTTGC	GCCTTTCCAAGTGCATCACCTG
TCF15	AGGGCCACGGAGATGAGCCT	GGTCCCCCGGTCCCTACACA
MEX3B	GGGCGGCAAGGTTGTAAA	ATGATCTCCCTCCGAGCCAT
SMYD1	GTGAAGAACGCAAGAGGCAGCT	CTCCTTCACCACTTCCTGAGAG
RBFOX2	CCAGCTTTCAAGCAGATGTGTCC	CAAATGGGCTCCTCTGAAAGCG
OCT4	CTTGAATCCCGAATGGAAAGGG	GTGTATATCCCAGGGTGATCCTC

Table 2.2: Forward and reverse primer sequences for qPCR analysis of human samples

2.2.4 Western blot

X. laevis caps flash frozen and stored at -80°C were homogenised in 25µl or 50µl Phosphosafe buffer (Novagen) respectively and centrifuged at 13,000 rpm for 1 minute. Supernatant was added to equal volume of sample buffer (120mM Tris/Cl pH6.8, 20% glycerol, 4% SDS, 0.04% bromophenol blue, 10% β-mercaptoethanol) and heated at 95°C for 5 minutes. Samples were run on 8% or 12% SDS-PAGE gels. Proteins were transferred to Immobilon-P Transfer Membrane (Millipore) by electroblotting wet transfer at 85V for 2 hours at room temperature. Antibodies were used at the dilutions below (Table 2.3). Proteins were detected using BM chemiluminescence blotting substrate (Roche).

Primary antibody	Dilution	Secondary antibody	Dilution
MF20 (21µg/ml DSHB)	1:1,000	Mouse	1:4,000
GAPDH (Sigma)	1:10,000	Mouse	1:4,000
dpERK	1:5,000	Mouse	1:6,000
Total ERK	1:40,000	Rabbit	1:2,000

Table 2.3: Dilutions of antibodies used for western blots.

2.2.5 Subcloning of candidate genes into pCS2+

Clones of candidate genes (Horizon) were grown overnight at 37°C on LB agar plates containing 100µg/ml ampicillin (Table 2.4).

Gene	Image ID	Vector	Species
smyd1	8938913	pCS107	*Xenopus tropicalis*
rbfox2	6986783	pCMV-Sport6ccdB	*Xenopus tropicalis*
mex3b	5156179	pCMV-SPORT6	*Xenopus laevis*
tcf15	5143248	pT7T3D-PacI	*Mus musculus*
pou5f3.2	7557150	pCS108	*Xenopus tropicalis*
ventx2.1-a	8321494	pExpress-1	*Xenopus laevis*
tcf12	5345693	pCMV-SPORT6	*Mus musculus*

Table 2.4: Clones of candidate genes (Horizon)

Colonies were selected from each bacterial plate and subjected to PCR with gene specific primers containing restriction enzyme binding sites for ClaI or XhoI and Kozak sequences (Kozak, 1991) (Table 2.5).

Gene	Primer sequence 5'-3'	
smyd1	Forward	GAGAGAATCGAT**ACCATG**GAGAACGTTGAAATTTTC
	Reverse	GAGAGACTCGAGTGGTTAGGACTTTGCTTCTTGTGG
rbfox2	Forward	GAGAGAATCGAT**ACCATG**GAGAAGAATAAAATGGTTTCG
	Reverse	GAGAGACTCGAGTGGTCAGTACGGAGCAAATCGGCTG
tcf12	Forward	GAGAGAATCGAT**ACCATG**TTCGCTAGCACTTTCTTTATG
	Reverse	GAGAGACTCGAGTGGTTACAGATGACCCATAGGGTTG
mex3b	Forward	GAGAGAATCGAT**ACCATG**CCCAGCTCGCTCTTTGCAGAC
	Reverse	GAGAGACTCGAGTGGTTAGGAGAAGATGCGGATGGC
tcf15	Forward	GAGAGAATCGAT**ACCATG**GCGTTCGCGCTGCTGCGC
	Reverse	GAGAGACTCGAGTGGTCATCGCCGAGGCCCTCGGAG
pou5f3.2	Forward	GAGAGAATCGAT**ACCATG**TACAGCCAACAGCCCTTC
	Reverse	GAGAGACTCGAGTGGTCAACCAATATGGCCGCCCATGGG
ventx2.1	Forward	GAGAGAATCGAT**ACCATG**ACTAAAGCTTTCTCCTCG
	Reverse	GAGAGACTCGAGTGGCTAATAGGCCAGAGGTTGCCC

Table 2.5: Forward and reverse gene specific primer sequences containing restriction enzyme binding sites for ClaI and XhoI respectively. Kozak sequences in bold.

25µl GoTaq Long PCR Master Mix 2x (Promega), 3µl 10µM forward and 3µl 10µM reverse gene specific primers and 18µl dH_2O were used to amplify 1µl colony resuspended in H_2O via PCR. The reaction mixture was heated to 95°C for 2 minutes before 30 cycles of 95°C for 30 seconds, 60°C for 30 seconds and 72°C for 90 seconds before a final elongation at 72°C for 10 minutes.

PCR products were purified using the Monarch PCR and DNA Cleanup Kit (NEB) and eluted in 30µl dH_2O. Clean PCR products and CS2+ plasmid were then cut by ClaI and XhoI restriction enzymes for 3 hours at 37°C. Digest reactions consisted of 15µl clean PCR product or 1µg pCS2+, 1.5µl ClaI restriction enzyme (Bsu15I Thermo Scientific), 1.5µl XhoI restriction enzyme (Promega), 10µl buffer B (Promega), and dH_2O up to a total volume of 100µl. Digests were purified using the Monarch PCR and DNA Cleanup Kit (NEB). Ligations were set up and placed at 18°C overnight with approximately 1:3 ratio of plasmid ends to insert ends, 1µl T4 ligase, 1µl T4 ligase buffer, and dH_2O up to a total volume of 10µl. PCR products ligated into pCS2+ were transformed into dam^-/dcm^- E. coli competent cells (C2925I, NEB). Transformations were grown overnight at 37°C on LB agar plates containing 100µg/ml ampicillin. Colonies were selected from each bacterial plate and subjected to PCR with T7 and SP6 pCS2+ primers to determine which colonies had taken up the insert successfully. The reaction mixture was heated to 95°C for 2 minutes before 30 cycles of 95°C for 30 seconds, 50°C for 30 seconds and 72°C for 90 seconds before a final elongation at 72°C for 10 minutes. Colonies that had successfully taken up the insert were selected and cultured overnight in 3ml LB-broth containing 100µg/ml ampicillin, in a 250rpm shaker at 37°C. Plasmids were isolated from the bacterial cultures using the QIAprep Spin Miniprep Kit and insert sequence confirmed by the Eurofins Genomics postal sequencing service. Plasmids containing correct candidate gene sequences were linearised by NotI restriction enzyme for 3 hours at 37°C. Digest reactions consisted of 2µg plasmid, 2µl NotI (Promega), 10µl buffer D (Promega), and H_2O up to a total volume of 100µl. Digests were purified using the Monarch PCR and DNA Cleanup Kit (NEB). Candidate gene mRNA was then synthesised using the mMESSAGE mMACHINE™ SP6 Transcription Kit using 0.2µg linear plasmid and purified by phenol:chloroform extraction and isopropanol precipitation.

2.2.6 Histology

Organoids and embryos fixed in 10% formaldehyde in PBS + 0.1% Tween were stained overnight in borax carmine. Samples were washed with acid alcohol (70% ethanol, 0.32% HCl) until no pink colour leached into the wash. They were then washed in 70% ethanol before introduction of Histo-Clear II (National Diagnostics) through a graded series of washes. Half of the Histo-Clear II was replaced with pre-melted Paraplast plus (Sigma) and left at 60°C, then replaced with two 1 hour washes of 60°C Paraplast plus. Samples were embedded in moulds with 60°C Paraplast plus and left to cool. Paraplast plus blocks were trimmed, mounted on wooden chucks, cut into 10µm sections and adhered to slides. Two washes in Histo-Clear II, two washes in 100% ethanol, one wash each in 75% ethanol, 50% ethanol, dH_2O, and picro blue black (390ml saturated picric acid + 10ml 1% w/v naphthalene blue black) were carried out and then back again in reverse order dH_2O through to Histo-Clear II. Slides were mounted with DePeX.

2.2.7 Whole-mount immunostaining

MEMFA (0.1M MOPS pH 7.4, 2mM EGTA, 1mM $MgSO_4$, 3.7% formaldehyde) fixed embryos stored in ethanol at -20°C were rehydrated through a graded series of ethanol and washed in PBS. Embryos were then treated with $K_2Cr_2O_7$ in 5% acetic acid at room temperature for 40 minutes. They were then washed in PBS before being bleached for 45 minutes in 5% H_2O_2 in PBS and washed in PBS again. Embryos were blocked in BBT (PBS, 1% BSA, 0.1% Triton X-100) + 5% horse serum for 1 hour before incubation in a 1/10,000 dilution dpERK antibody (Sigma), or 1/100 dilution DSHB antibodies for 12-101 at 4°C overnight. Embryos were washed in BBT then BBT + 5% horse serum for 1 hour before incubation with BBT + 5% horse serum with a 1/1,000 dilution of horse anti-mouse igG-AP conjugated secondary (VectorLab) overnight at 4°C. Immunostaining was visualised with BM purple (Roche).

2.3 Data analysis

2.3.1 Differential expression analysis of RNA-seq data

Initial quality control and mapping of sequences to the *X. laevis* transcriptome was undertaken by Katherine Newling at the University of York Technology Facility. Raw reads for each sample were aligned to the *X. laevis* reference transcriptome using Salmon (http://salmon.readthedocs.io) to produce estimated read counts for each transcript.

Differential gene expression analysis was carried out using the R package Sleuth (http://pachterlab.github.io/sleuth/) by fitting a statistical model to the estimated read counts. Wald tests were performed between each condition within each developmental stage to calculate q values and fold changes/effect sizes. Self-written R scripts used for Sleuth analysis can be found in the appendix.

2.3.2 Gene ontology analysis of RNA-seq data

Genes with mean transcript per million (tpm) values ≥ 1, q value ≤ 0.05, and effect size compared to control ≥ 1.5 from Myod + Fgf4 organoids were submitted to the Protein Analysis Through Evolutionary Relationships (PANTHER) Classification System to identify statistically significant over-represented biological processes at each developmental stage investigated (http://pantherdb.org/) (Mi et al., 2021). The *Mus musculus* gene database was used due to better annotation than the *Xenopus tropicalis* database.

2.3.3 Venn diagram analysis of RNA-seq data

Transcripts with mean tpm values ≥ 1, q value ≤ 0.05, and effect size compared to control ≥ 1.5 for each experimental condition were submitted to the Bioinformatics & Evolutionary Genomics tool (http://bioinformatics.psb.ugent.be/webtools/Venn/) to generate Venn diagrams.

2.3.4 Figures

Heatmaps, skeletal muscle comparison Venn and Euler diagrams, gene ontology bar charts and volcano plots were generated using self-written code in RStudio using R packages including pheatmap, VennDiagram, ggplot2 and EnhancedVolcano. Adobe Photoshop CS3 or Adobe Photoshop Elements 2019 were used to process images of organoids, histology and western blots.

2.4 Tissue culture methods

2.4.1 Culture of H9 cells

H9 embryonic stem cells (Thomson et al., 1998) were maintained on Vitronectin in N2B27 media (24.5ml Neurobasal™ (gibco™ Thermo Fisher), 24.5ml DMEM/F-12 (gibco™ Thermo Fisher), 250µl N2, 500µl B27, 500µl 500mM L-Glutamine, 100µl 50mM β-mercaptoethanol) supplemented with AFX (2ng/ml Activin A2, 12.5ng/ml FGF2 and 2µM tankyrase inhibitor XAV939 (Huang et al., 2009)). Cells were passaged with EDTA.

2.4.2 Differentiation of skeletal muscle progenitors

Protocol as per (Shelton et al., 2016) with the following adaptations: cells cultured on Vitronectin rather than Matrigel, in N2B27 AFX media instead of E8 media, and N2B27 media instead of E6 and StemPro-34 media. As described in Shelton's protocol, cells were pre-treated with 10µM rho kinase inhibitor Y27632 (Watanabe et al., 2007) and dissociated with TrypLE before replating at $1.5x10^5$ cells per well of 12-well plates. After 24 hours, cells were treated with 10µM GSK3 inhibitor CHIR99021 (Kreuser et al., 2020) from day 0-2, and 5ng/ml FGF2 from day 12-20. Due to time restrictions, samples for qPCR and western blot analysis were collected at time points up to day 20 only.

2.4.3 Tissue culture sample collection

Samples were collected for qPCR analysis in 250µl Monarch DNA/RNA Protection Reagent (NEB) and RNA extracted using the Monarch Total RNA Miniprep Kit (NEB). cDNA was synthesised as per 2.2.3 and analysed by 384-well qPCR as per 2.2.4.

Protein samples were collected for western blot in 250µl Phosphosafe buffer (Novagen) and centrifuged at 13,000 rpm for 1 minute. Supernatant was stored at -80°C and analysed by western blot as per 2.2.5.

Chapter 3: Development and optimisation of a skeletal muscle inducing protocol

3.1 Introduction

The ability to investigate muscle differentiation in culture was essential for our current understanding of the molecular basis of the establishment of the myogenic lineage. In the 1960s, myoblasts were isolated from quail embryos and found to clonally expand and retain their ability to differentiate into myotubes (Konigsberg, 1963). From this, important aspects of skeletal muscle differentiation were defined: during differentiation, myoblasts stop proliferating and fuse to form multinucleated myofibres, while at the same time coordinately activating the transcription of contractile protein genes (Devlin and Emerson, 1978; Bucher et al., 1988). It was also shown that myogenesis in culture can be driven by chemically activating the myogenic regulatory genes in mouse fibroblast lines through 5-azacytidine treatment, a protocol that formed the basis of the identification of the MRFs (Taylor and Jones, 1979; Davis et al., 1987). Since then, driving myogenesis in cultured mouse cells has been used to interrogate gene regulatory networks that underpin determination and differentiation of the muscle lineage (Conerly et al., 2016). In this chapter, the development of a protocol to generate skeletal muscle in cultured *Xenopus* explants is presented. In subsequent chapters, this protocol will be used to define the role of FGF in Myod driven myogenesis in *Xenopus*.

3.1.1 Animal cap assays

Xenopus studies have contributed a substantial amount to our knowledge of early vertebrate development since introduction of the model organism in the 1950s (Beck and Slack, 2001). Some of the many benefits of using *Xenopus* include the fact that eggs are laid in large quantities and embryos can withstand extensive manipulation such as dissection of pluripotent animal pole cells, and microinjection of nucleic acids (Slack et al., 1987; Green and Guille, 1999; Vize et al., 1991; Blum and Ott, 2018). Synthetic mRNAs are rapidly translated after injection and DNA constructs are transcribed and translated (Vize et al., 1991). These advantageous qualities made *X. laevis* an ideal choice for a skeletal muscle inducing protocol to further our understanding of how pluripotent cells commit to the skeletal muscle lineage.

During the blastula stage of development, a cavity known as the blastocoel is formed beneath pigmented animal pole cells (Green and Guille, 1999). These animal pole cells (animal caps) can be dissected at developmental stage 8-9 and cultured in simple isotonic buffered salt solution as they contain sufficient yolk for survival (Figure 3.1) (Sudarwati and Nieuwkoop, 1971; Godsave et al., 1988; Asashima et al., 1990; Ariizumi et al., 1991, 2009). Unlike mammalian cell culture media which

can include ill-defined components like serum (which may contain growth factors), the ability to culture frog explants in isotonic saline solution means the presence or absence of experimentally added growth factors is easy to control. Untreated animal caps will differentiate into an ectodermal fate, forming atypical epidermis. However animal cap fate can be diverted to form derivatives of any of the three germ layers via culture in soluble protein, injection of mRNA, or grafting to other tissues (Borchers and Pieler, 2010; Green and Guille, 1999). Culture under various conditions leads to formation of organoids (multicellular structures containing many of the cell types present in an adult organ) with many applications (Ariizumi et al., 2017).

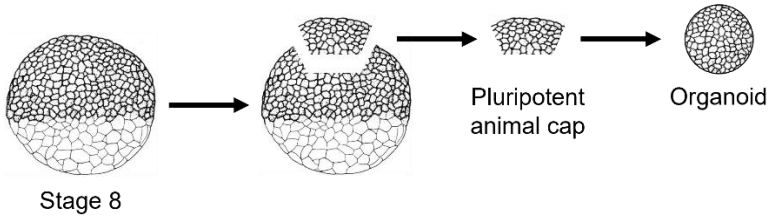


Figure 3.1: Schematic diagram of animal cap assay. Animal pole cells are dissected from blastula stage 8-9 *Xenopus* embryos and cultured in isotonic buffered salt solution until desired stage.

Animal cap assays have been used for many years as a robust and reliable technique for the investigation of cell signalling and lineage specification (Cooke et al., 1987; Lombardo et al., 1998; Angerilli et al., 2018; Brunsdon and Isaacs, 2020). These assays initially rose to prominence when used to demonstrate the mesoderm inducing role of FGF and transforming growth factor β (TGFβ) 'in vivo' (Slack et al., 1987; Kimelman and Kirschner, 1987). In normal development, mesoderm and neural ectoderm cells undergo convergent extension along the dorsal axis leading to elongation of the embryo during gastrulation and neurulation (Keller et al., 2000). If mesoderm is induced in *Xenopus* animal cap organoids, the mesodermal cells undergo the same movements of convergent extension resulting in a clear, elongated phenotype when stage-matched control embryos reach gastrula stage. When stage-matched control embryos reach late tailbud stage, mesoderm-containing organoids subsequently swell to become large vesicles (Slack et al., 1987). This easily observed and well-documented mesodermal phenotype is another benefit of using animal cap assays to study muscle specification.

3.1.2 Aims of this chapter

The aims of this chapter are to:

- Develop a skeletal muscle inducing protocol in organoids derived from *X. laevis* pluripotent animal cap explants using Fgf4 protein and *myod1.S* mRNA
- Optimise protocol conditions to increase the amount of skeletal muscle induced
- Analyse the efficiency of the protocol

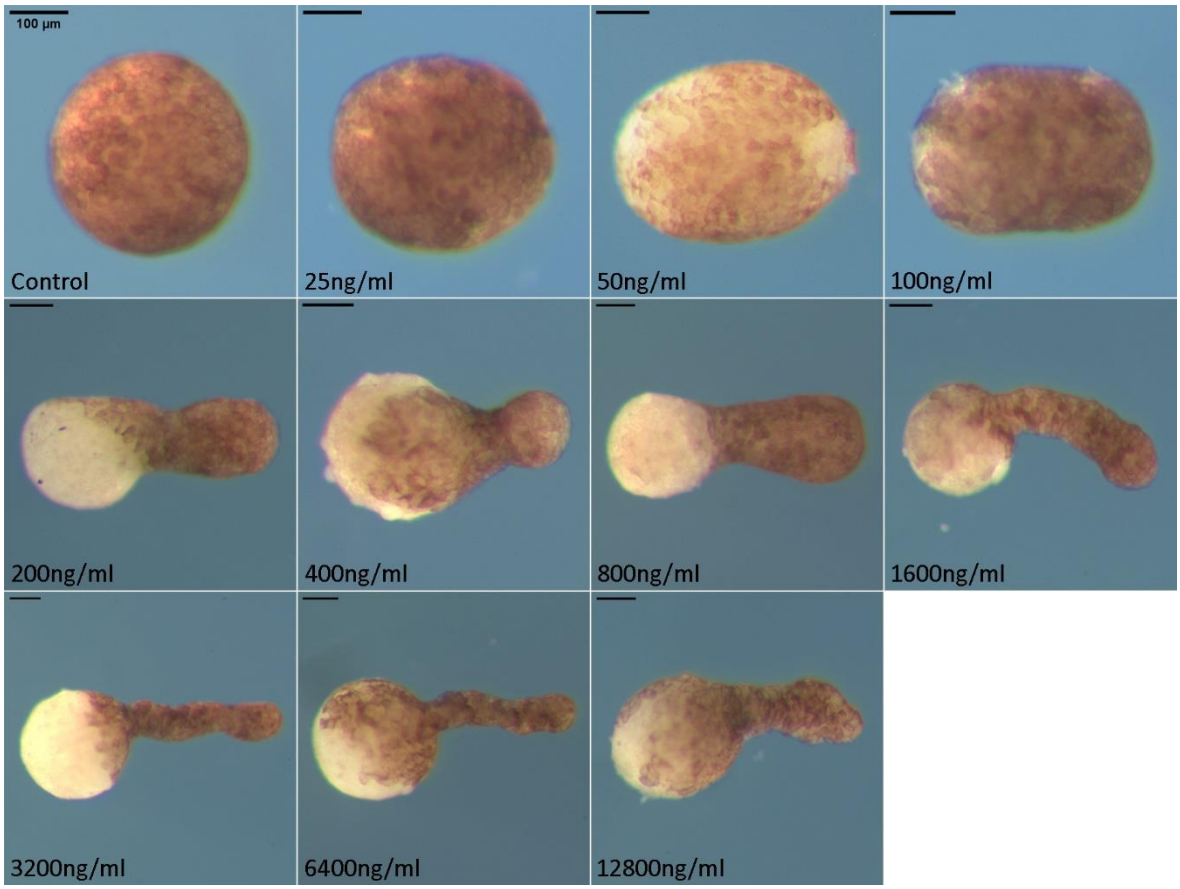
3.2 Results

3.2.1 Fgf4 induces elongation of animal cap organoids

In order to observe the effect of FGF signalling on animal pole cell fate and identify appropriate concentrations for a skeletal muscle inducing protocol, animal cap cells were dissected at blastula stage 8 and cultured in different concentrations of recombinant Fgf4 protein. Culture in solutions with 25ng/ml, 50ng/ml and 100ng/ml *Xenopus* Fgf4 protein (Isaacs et al., 1992) induced slight elongation of animal cap organoids by the time stage-matched control embryos reached gastrula stage 12.5 (Figure 3.2A). Concentrations of Fgf4 between 200ng/ml and 12800ng/ml resulted in a more pronounced phenotype with a single elongated process. Organoid elongation is in keeping with mesoderm induction phenotypes in previous studies (Slack et al., 1987).

When stage-matched control embryos reached late tailbud stage 40, vesicles had formed in all organoids between 50ng/ml and 800ng/ml Fgf4, further indicating that mesoderm was induced (Figure 3.2B). Only one of four organoids at 25ng/ml Fgf4 formed a vesicle suggesting that one unit of activity of the recombinant Fgf4 protein lies between 25ng/ml and 50ng/ml (Cooke et al., 1987; Godsave et al., 1988). Organoids cultured at concentrations above 800ng/ml did not survive till stage 40.

A



B

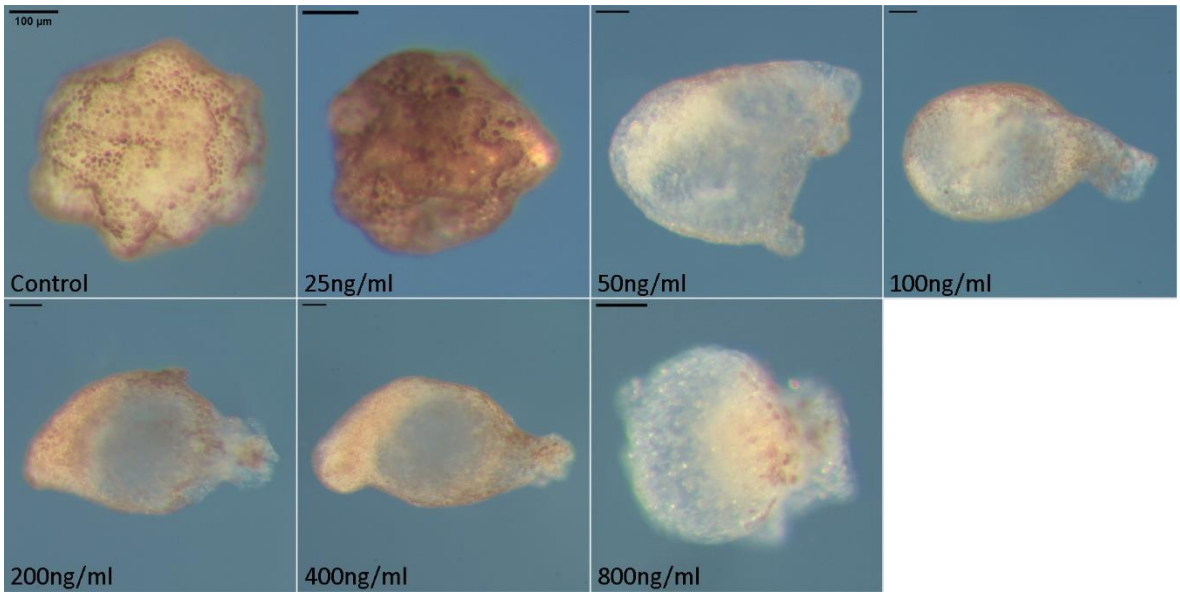


Figure 3.2: Organoids cultured with different concentrations of Fgf4 protein. Animal pole cells dissected from *X. laevis* embryos and cultured with different concentrations of Fgf4 recombinant protein until **A** stage 12.5, **B** stage 40. n=4 for each concentration, n=10 for controls. Scale bars = 100µm.

3.2.2 Overexpression of Myod does not give rise to mesodermal phenotypes in organoids

Previous work has shown that overexpression of Myod in *X. laevis* animal caps leads to transcription of muscle genes but is not sufficient for muscle differentiation (Hopwood and Gurdon, 1990). In order to identify appropriate concentrations for the skeletal muscle protocol, four different concentrations of *myod1.S* mRNA were injected into *X. laevis* embryos bi-laterally at the 2-cell stage. Animal cap cells were dissected at blastula stage 8 and cultured until stage-matched control embryos reached late tailbud stage 40. Injection of 1.6ng, 2.7ng, 4ng or 8ng *myod1.S* mRNA did not induce vesicle formation suggesting that Myod alone is not sufficient for mesoderm induction (Figure 3.3). However, organoids injected with 2.7ng, 4ng and 8ng *myod1.S* mRNA were slightly elongated at stage 40 but not dissimilar in size.

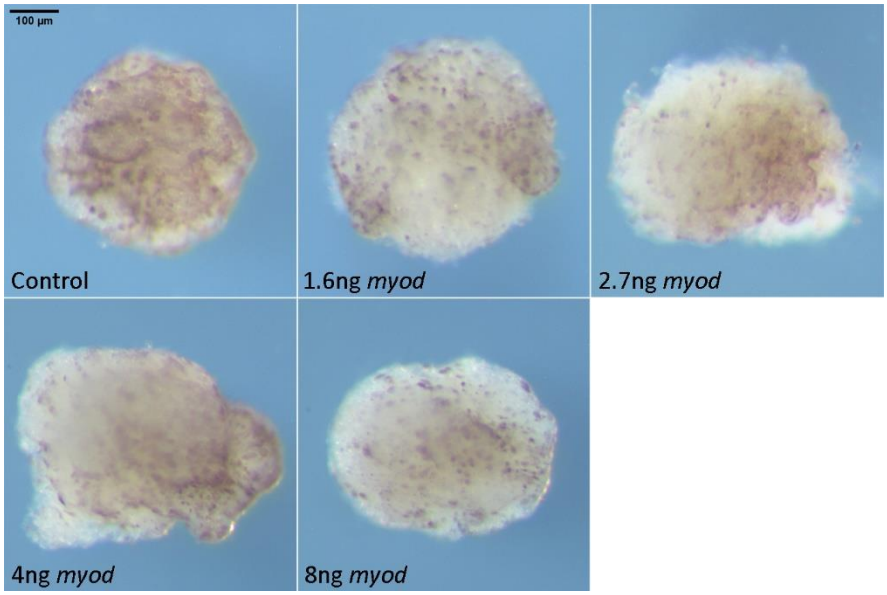


Figure 3.3: Organoids injected with different quantities of *myod1.S* mRNA. Animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1.6ng, 2.7ng, 4ng or 8ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured until stage 40. n=10. Scale bar = 100µm.

3.2.3 Fgf4 and Myod co-expression induces skeletal muscle

Having observed the effects of a range of Fgf4 concentrations and quantities of *myod1.S* mRNA, a preliminary skeletal muscle assay was carried out to test the hypothesis that FGF signalling allows differentiation of pluripotent cells into skeletal muscle. Animal pole cells were dissected from un-injected *X. laevis* embryos and embryos injected with 1.6ng *myod1.S* mRNA. Animal cap explants were cultured with or without 50ng/ml Fgf4 as this was the lowest concentration required to reliably induce a mesodermal phenotype. At gastrula stage 12.5, organoids expressing Myod alone formed round balls of cells similar to controls (Figure 3.4A). Organoids cultured with 50ng/ml Fgf4 were slightly elongated. 1.6ng *myod* + 50ng/ml Fgf4 organoids were substantially more elongated than those expressing Myod or Fgf4 alone.

At stage 40, Myod expressing organoids still resembled untreated controls, and organoids cultured in 50ng/ml Fgf4 had formed large vesicles (Figure 3.4B). Organoids expressing Myod + Fgf4 grew larger than controls but with a denser appearance than the vesicles present in Fgf4 organoids.

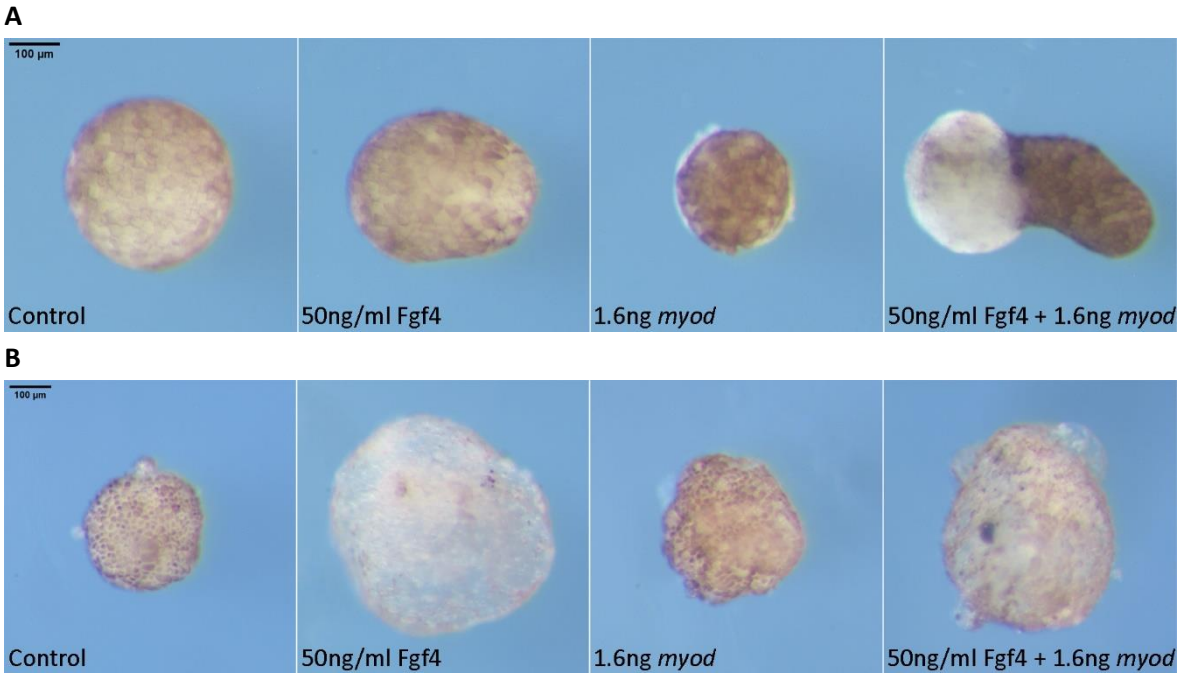


Figure 3.4: Preliminary skeletal muscle assay organoids. Animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1.6ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured without Fgf4, with 25ng/ml Fgf4 or 50ng/ml Fgf4 until **A** stage 12.5, **B** stage 40. n=10. Scale bars = 100µm.

3.2.4 Skeletal muscle protocol optimisation

The quantity of *myod1.S* mRNA injected was then further optimised. 0.4ng, 0.8ng, 1.2ng and 1.6ng *myod1.S* mRNA was injected bilaterally at the 2-cell stage into *X. laevis* embryos. Animal cap cells were dissected from injected and un-injected embryos and cultured with or without 50ng/ml Fgf4. Organoids were sectioned for histological analysis in order to determine which conditions resulted in formation of the most skeletal muscle (Figure 3.5). Control organoids were composed of epidermis with ciliated outer layers and no skeletal muscle cells. Although organoids from embryos injected with *myod1.S* mRNA still appeared ectodermal, they differed from controls in that they had some resemblance to neuroectoderm with neural tube-like structures. Fgf4 treated organoids contained a ring of smooth muscle known as the mesothelium and had large vesicles increasing their size. Skeletal muscle was present in Myod + Fgf4 organoids, with the most skeletal muscle formed with 1.2ng *myod1.S* mRNA. Therefore 1.2ng *myod1.S* + 50ng/ml Fgf4 was selected as the most effective condition for the skeletal muscle protocol. The resulting skeletal muscle inducing protocol methodology is summarised in Figure 3.6.

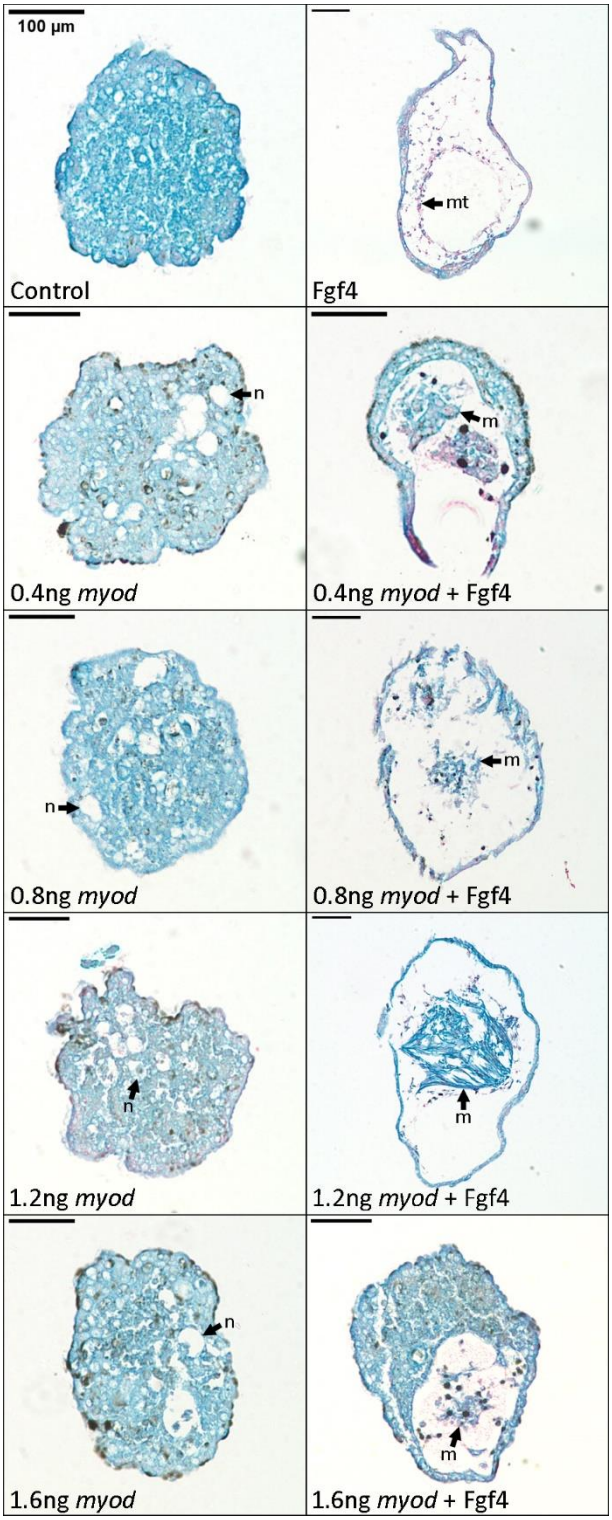


Figure 3.5: Representative skeletal muscle assay organoid sections. 10µm sections of animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 0.4ng, 0.8ng, 1.2ng or 1.6ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured with or without 50ng/ml Fgf4 until stage 41. Sections treated with borax carmine and counterstained with picro blue black. Control organoids are composed of atypical epithelium. Myod expressing organoids have some resemblance to neuroectoderm with neural tube-like structures (n). Fgf4 treated organoids contain smooth muscle mesothelium (mt). *myod* + 50ng/ml Fgf4 organoids contain varying amounts of skeletal muscle (m). Scale bars = 100µm.

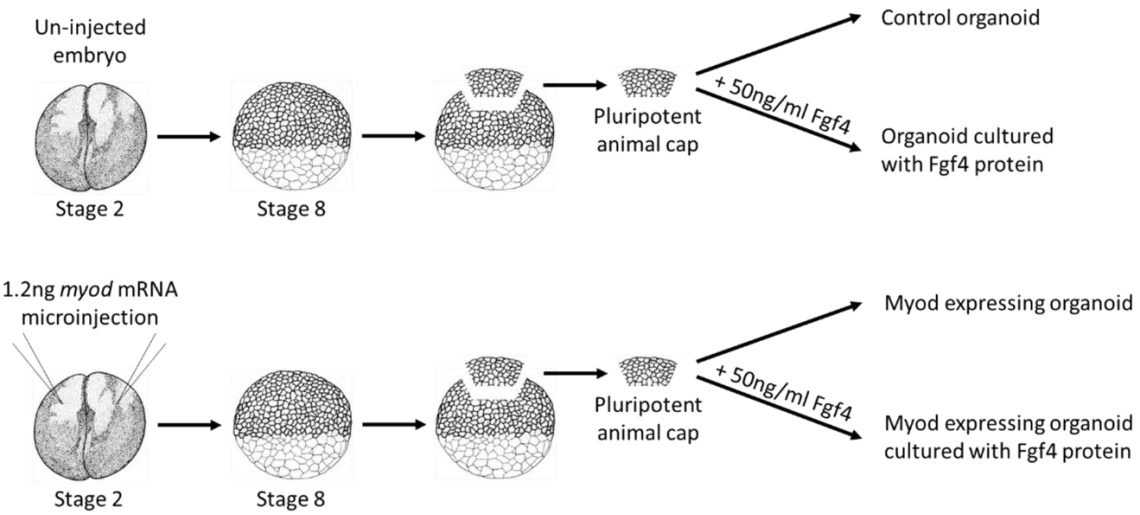


Figure 3.6: Schematic representation of skeletal muscle protocol methodology. *X. laevis* embryos.

3.2.5 Analysis of cell types present in organoid sections

Having identified effective conditions for the skeletal muscle protocol as 1.2ng *myod* with 50ng/ml Fgf4, 10 organoids for each condition were stained with borax carmine, sectioned and counterstained with picro blue black. Each section was then scored according to whether it contained any skeletal muscle, non-muscle mesoderm or ectoderm only. The percentage of sections in each category was recorded for each organoid (Figure 3.7), and the mean average calculated (Figure 3.8).

All control organoids consisted of ectoderm only. A mean average of 83.6% of sections from Fgf4 treated organoids contained large vesicles, a known non-muscle mesoderm phenotype (Figure 3.8A). An average of 6.6% of sections from Fgf4 treated organoids contained a small amount of muscle. Sections towards the start and end of the Fgf4 organoids either side of the vesicles were made up of ectoderm. The majority of Myod expressing organoids were made up of ectoderm (Figure 3.8A) though three myod organoids contained a small amount of muscle (Figure 3.7). An average of 2.9% and 13% of sections from 1.2ng myod organoids contained a small amount of muscle and non-muscle mesoderm respectively (Figure 3.8A). However in Myod + Fgf4 organoids, skeletal muscle was present in a greater number of sections (an average of 47.4%) and typically made up a larger proportion of cells in these sections (Figure 3.8B). Blocks of organised, differentiated skeletal muscle were located within vesicles surrounded by an outer layer of ectoderm (Figure 3.8A).

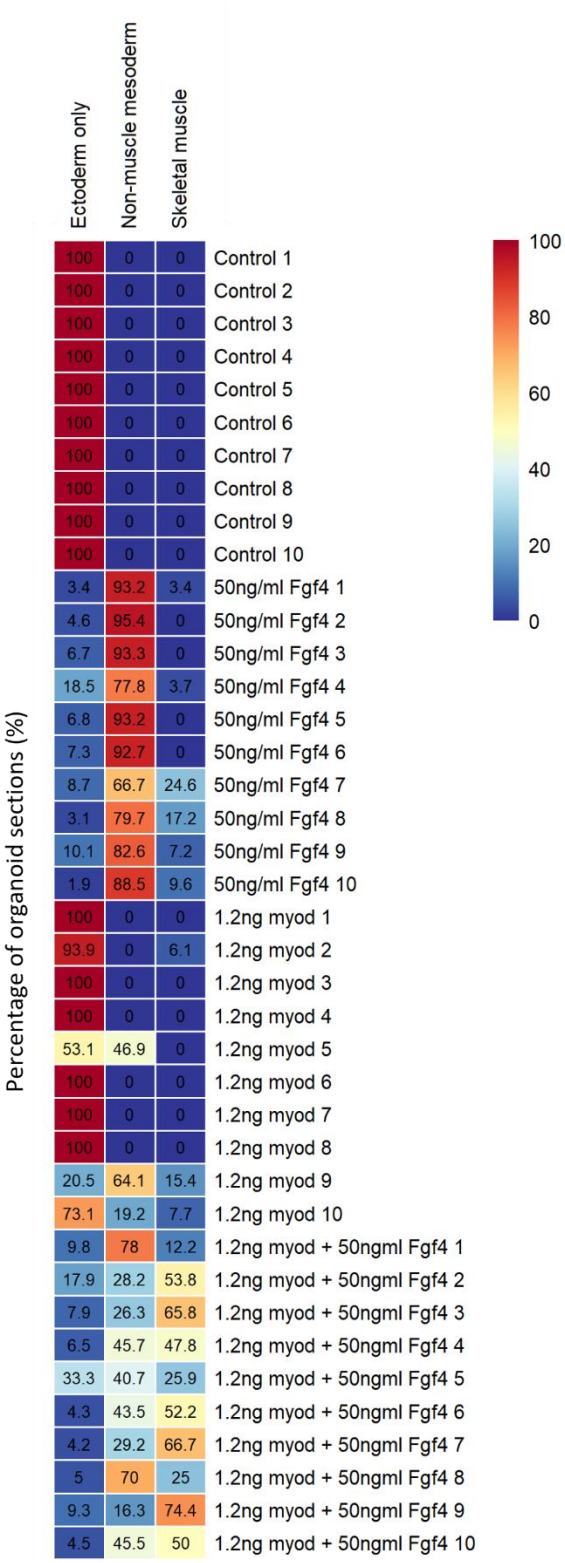
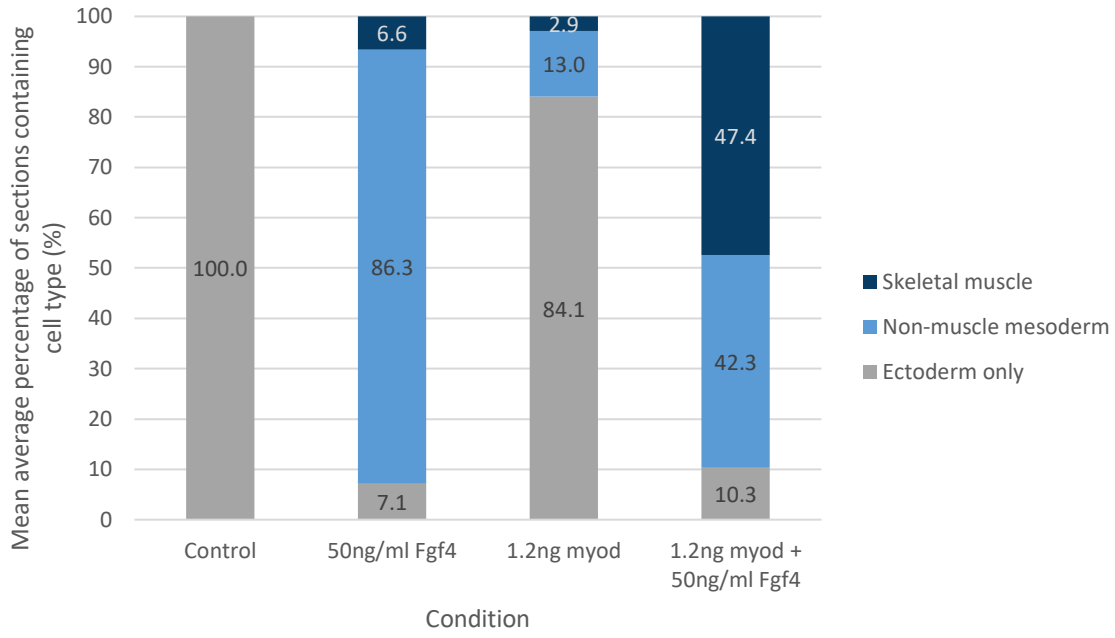


Figure 3.7: Heatmap showing percentage of organoid sections containing skeletal muscle, non-muscle mesoderm or ectoderm. 10µm sections of animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1.2ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured with or without 50ng/ml Fgf4 until stage 41. 10 organoids per condition.

A



B

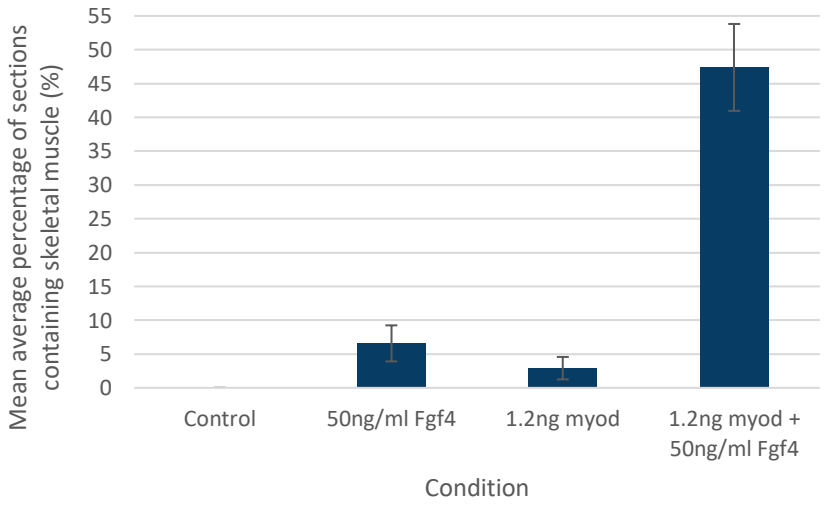


Figure 3.8: Average percentage of organoid sections containing skeletal muscle, non-muscle mesoderm or ectoderm. 10µm sections of animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1.2ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured with or without 50ng/ml Fgf4 until stage 41. Mean average with standard error. n=10

In order to highlight regions of skeletal muscle within an organoid, immunostaining with 12-101 (an antibody against a membrane protein of the sarcoplasmic reticulum) was carried out (Figure 3.9). Slight variations were observed between organoids, though skeletal muscle was typically located in blocks within the organoid surrounded by a vesicle and scattered non-muscle cells.

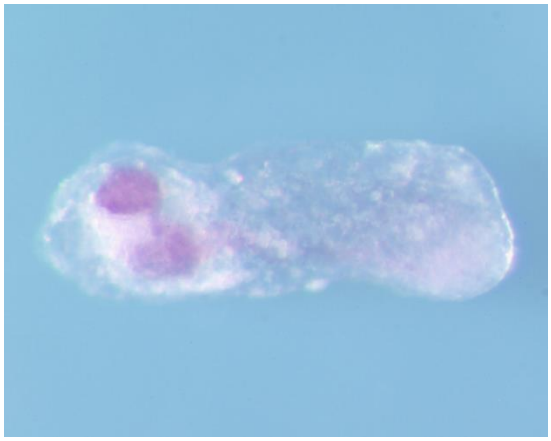


Figure 3.9: Myod + Fgf4 organoid immunostained with 12-101. Animal pole cells dissected from *X. laevis* embryo injected with 1.2ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured with 50ng/ml Fgf4 until stage 41. Immunostained with 12-101 antibody against membrane protein of the sarcoplasmic reticulum. Purple staining indicates skeletal muscle.

3.3 Discussion

3.3.1 FGF signalling induces mesoderm in animal cap organoids

By gastrula stage 12.5, animal cap organoids cultured with Fgf4 protein showed an elongated phenotype, as FGF-induced mesodermal cells replicate the same movements of convergent extension that would occur in the embryo (Slack et al., 1987). Fgf4 treated organoids subsequently developed large vesicles containing a ring of smooth muscle known as the mesothelium. These phenotypes were predicted as Fgf2 overexpression in animal caps has previously been shown to induce primarily ventral mesoderm (Slack et al., 1987; Kimelman and Kirschner, 1987).

Quantification of the amount of skeletal muscle induced showed that sections from a small percentage of Fgf4 treated organoids did contain a small amount of skeletal muscle. This was expected as previous work culturing animal caps in Fgf2 at concentrations between 2ng/ml and 30ng/ml showed similar results with a concentric arrangement of loose mesenchyme, mesothelium and blood cells within an epidermal layer, sometimes containing a few muscle cells (Slack et al., 1987). In keeping with this, animal caps treated with 10ng/ml activin A also differentiate into mesoderm derivatives including blood cells, mesenchyme, coelomic epithelia, and sometimes contain a few muscle cells (Asashima et al., 1990). When animal caps are treated with higher concentrations of Fgf2 or activin A, typically more muscle is formed (Slack et al., 1987; Asashima et al., 1990). As Myod is a direct target of Fgf4, higher concentrations of Fgf4 treatment alone would

also initiate Myod transcription, resulting in more skeletal muscle formation (Fisher et al., 2002). Consequently, a low concentration of 50ng/ml Fgf4 was selected for the protocol in order to preferentially induce mesoderm, and allow Myod to drive myogenesis if the conditions permitted.

Early animal cap studies demonstrated the ability of Fgf1, Fgf2 and Fgf4 to induce mesoderm, though Fgf1 and Fgf2 do not possess a secretory signal sequence, so exhibit a much less potent mesoderm inducing ability than secreted Fgf4 (Slack et al., 1987, 1989; Isaacs et al., 1994). Other molecules such as activin A are capable of inducing mesoderm in animal caps, however this is also FGF dependent as Fgf4 is an immediate early response to activin (Ariizumi et al., 1991; Fisher et al., 2002; Cornell and Kimelman, 1994). Another FGF implicated in mesoderm formation is Fgf8b, though the Fgf8a spliceform plays an important role in neural crest and posterior neural development (Fletcher et al., 2006; Hong et al., 2008). Fgf4 is required for the initial activation of Myod transcription, though use of a translation blocking morpholino against Fgf8 also reduces Myod expression, indicating a role for both in the mesodermal lineage (Fisher et al., 2002; Fletcher et al., 2006). For this study, Fgf4 was selected as it is required for mesoderm maintenance via an autocatalytic regulatory loop with Tbxt, and has been directly implicated in muscle development (Isaacs et al., 1992, 1994; Fisher et al., 2002). For instance, investigation of the community effect revealed that Fgf4 is required for muscle precursor cells to differentiate and stably-express muscle-specific genes, while other factors such as Fgf2 are not sufficient (Standley et al., 2001; Gurdon et al., 1993). However, it would be interesting to replace Fgf4 with Fgf8b in the protocol, to test its ability to induce Myod driven myogenesis in organoids.

3.3.2 Additional regulators are required for Myod to induce skeletal muscle

The phenotypes and histology of organoids cultured as per the skeletal muscle inducing protocol developed in this chapter, support the hypothesis that FGF signalling plays a key role in allowing pluripotent cells to differentiate into muscle. Overexpression of Myod in *Xenopus* animal cap organoids was not sufficient for effective myogenesis. This was expected as Hopwood and Gurdon previously showed that injected *myod* RNA induced expression of actin normally expressed in the myotome, however did not result in muscle differentiation (Hopwood and Gurdon, 1990). Organised skeletal muscle with clear myotubules was formed in 1.2ng *myod* + 50ng/ml Fgf4 organoids. This requirement for more than Myod expression alone aligns with studies in which Myod expression in particular cell lines did not result in skeletal muscle differentiation unless specific environmental changes were made. For example, 10T1/2 cells constitutively expressing Myod will not differentiate in growth medium (40-60% buffalo rat liver-conditioned medium with 15% fetal calf serum and 10^{-4}M β-mercaptoethanol), however they become skeletal muscle cells when cultured in muscle-specific differentiation-promoting medium (10% horse serum and

0.1µg/ml insulin), or if starved (Dekel et al., 1992). Similarly, hepatocytes expressing Myod do not express muscle-specific genes unless fused to fibroblasts, indicating that additional regulators are required (Schäfer et al., 1990). A requirement for FGF in some cultured chick limb muscle was identified by Seed and Haushka (1988), although the interpretation is complicated as FGF also promotes proliferation in cultured cells. The protocol presented in this chapter will be used for a transcriptomic analysis over a developmental time course to identify some of the factors contributing to the ability of Myod to drive myogenesis, and the influence of FGF during this process.

3.3.3 Skeletal muscle protocol limitations

Although Myod + Fgf4 organoids contained a much greater amount of skeletal muscle, a small amount was also present in three Myod organoids. There was also variability within the percentage of skeletal muscle observed between Myod + Fgf4 organoids. The slight variability was not unexpected as previous animal cap studies have also observed some differences in histology between organoids (Asashima et al., 1990). For example, when animal caps were treated with 100ng/ml activin A, various cell types were induced including muscle, notochord and renal tubules in 50%, 33% and 22% of organoids respectively (Asashima et al., 1990).

There are a number of potential contributors to variability within animal cap assays, including the fact that pigmented cells are less responsive to mesoderm inducers than inner blastocoel roof cells (Cooke et al., 1987; Green and Guille, 1999). This may contribute to the appearance of stage 12.5 organoids treated with 200ng/ml or more Fgf4, in which pigmented cells form a layer over the elongating process of inner blastocoel cells. The competence of cells to respond to mesoderm inducers Fgf and activin declines with the onset of gastrulation, thus animal cap cells were dissected before stage 10 (Jones and Woodland, 1987; Green et al., 1990; Slack et al., 1988). Animal pole cells were excised by hand using mounted tungsten needles so some explants may have contained slightly fewer cells than others, or experienced greater cell loss depending on exact dissection time within the stage 8-9 window. A larger explant containing a greater number of cells can sometimes be more susceptible to induction than a smaller explant (Green and Guille, 1999). After dissection, animal cap explants begin to round up and form closed balls of cells. The rate of 'rounding up' varies between egg batches and fully rounded explants are no longer responsive to subsequently applied soluble factors (Green and Guille, 1999). Therefore, larger explants with a greater number of inner blastocoel cells or those which take longer to round up, may be more sensitive to mesoderm induction and skeletal muscle differentiation.

Despite these limitations, animal cap assays remain a well-documented, simple, inexpensive protocol ideal for investigation of cell lineage specification and identification of key genes involved

(Satou-Kobayashi et al., 2021; Borchers and Pieler, 2010). The advantages of these assays continue to be recognised and similar techniques have started to be applied in other model organisms (Simunovic and Brivanlou, 2017). Aggregates of mouse ES cells and blastoderm explants of zebrafish embryos, have been shown to form organoids when cultured appropriately and are now also being used to study embryogenesis and somitogenesis (van den Brink et al., 2020; Schauer et al., 2020). In chapter 4, the skeletal muscle protocol described will be analysed at a transcriptomic level in order to characterise the type of muscle induced.

Chapter 4: Transcriptomic analysis and characterisation of induced skeletal muscle

4.1 Introduction

4.1.1 RNA sequencing

After development and optimisation of the skeletal muscle protocol, the transcriptional output of the myogenic programme induced was investigated. In order to determine gene expression during muscle specification and characterise differentiated muscle, samples were analysed at different time points by Illumina NovaSeq RNA sequencing (RNA-seq). There are several advantages of using RNA-seq over microarrays, including the fact that RNA-seq analysis can detect lower abundance transcripts and has a broader dynamic range (Zhao et al., 2014). Another benefit is that RNA-seq allows analysis of all the transcripts present, whereas microarrays analyse the expression of a biased, pre-determined set of genes only. As this project required identification of both known muscle features and novel transcripts not previously associated with skeletal muscle lineage specification, RNA-seq was the most appropriate method.

4.1.2 Xenopus laevis tetraploidy

X. laevis are tetraploid organisms so this transcriptomic analysis required careful attention when assigning reads to alleles. Ancient polyploidisation events, such as the two whole genome duplication events shared by all vertebrates, were pivotal in the evolution of many eukaryotic genomes (Van De Peer et al., 2009). Polyploidy is still common in today's amphibians, fish and plants (Otto, 2007). The *X. laevis* allotetraploid genome arose via the interspecific hybridisation of diploid progenitors, followed by a genome duplication allowing disomic inheritance and meiotic pairing (Kobel and Du Pasquier, 1986). The two subgenomes evolved asymmetrically and are referred to as long (L) and short (S), with the S subgenome having undergone more deletions and rearrangements over time (Session et al., 2016). The asymmetry between the subgenomes may be due to differences between their diploid progenitors, or the S subgenome remodelling may have been initiated by activation of transposable elements through the L-S merger itself (Session et al., 2016). These pseudoallelles differ by approximately 5-10% at the DNA level, so expression patterns can be similar or differ from each other over time.

4.1.3 Aims of this chapter

The aims of this chapter are to:

- Analyse the transcriptional output of the myogenic programme initiated by the skeletal muscle inducing protocol
- Characterise the type of muscle induced in Myod + Fgf4 organoids

4.2 Results

4.2.1 Sample generation using the skeletal muscle protocol

In order to investigate what is occurring at a transcriptional level to form skeletal muscle during this protocol, samples were collected for RNA-seq. Animal pole cells were dissected from sibling un-injected *X. laevis* embryos and embryos injected with 1.2ng *myod1.S* mRNA bilaterally at the 2-cell stage, and cultured with or without 50ng/ml Fgf4 protein. Animal cap organoids were collected at early neurula stage 14, late neurula stage 20 and late tailbud stage 30 for RNA-seq analysis (Table 4.1). Organoids were also collected at a later stage for validation by western blot.

Condition	**Number of organoids collected per replicate**			
	RNA-seq			**Western blot**
	Stage 14	**Stage 20**	**Stage 30**	**Stage 37-41**
Control organoids	10	10	10	10
50ng/ml Fgf4 organoids	10	10	10	10
1.2ng *myod* organoids	10	10	10	10
1.2ng *myod* + 50ng/ml Fgf4 organoids	10	10	10	10

Table 4.1: Organoids collected for RNA-seq and western blot analysis.

4.2.2 Skeletal muscle myosin is present in Myod + Fgf4 organoids

Western blots were carried out for sarcomere myosin heavy chain (MHC) to confirm that the skeletal muscle protocol was successful. MHC was detected using the MF20 antibody (Figure 4.1). No MHC was present in control or 50ng/ml Fgf4 treated organoids. 1.2ng *myod* + 50ng/ml Fgf4 organoids contained the most MHC in all 4 replicates, and a small amount was also detected in 1.2ng *myod* organoids in 2 replicates.

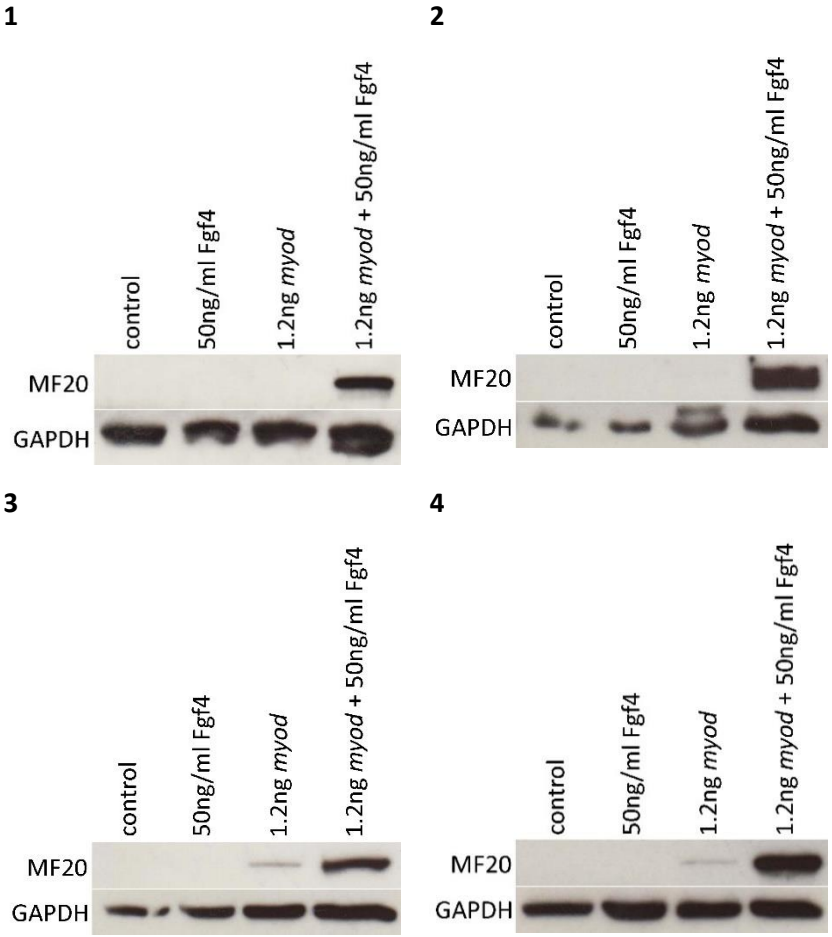


Figure 4.1: Western blots detecting sarcomere myosin heavy chain (MF20) in skeletal muscle protocol organoids. Animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1.2ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured with or without 50ng/ml Fgf4. Animal cap organoids collected between stages 37-41 for assay validation by western blot. 4 replicates (1, 2, 3 and 4).

4.2.3 RNA-seq sample processing and read alignment

Once an MF20 western blot validated the presence of skeletal muscle within a replicate, RNA was extracted from 10 organoids for each condition at stages 14, 20 and 30. Quality control checks (Figure 4.2) and poly(A) library preparation were carried out by the University of York Technology Facility. Samples were then sent for Illumina NovaSeq sequencing, resulting in 7.3-20.7 million paired end reads per sample.

Initial quality control and mapping of sequences to the *X. laevis* transcriptome was undertaken by Katherine Newling at the University of York Technology Facility. Raw reads for each sample were aligned to the *X. laevis* reference transcriptome using Salmon (http://salmon.readthedocs.io) to produce estimated read counts for each transcript.

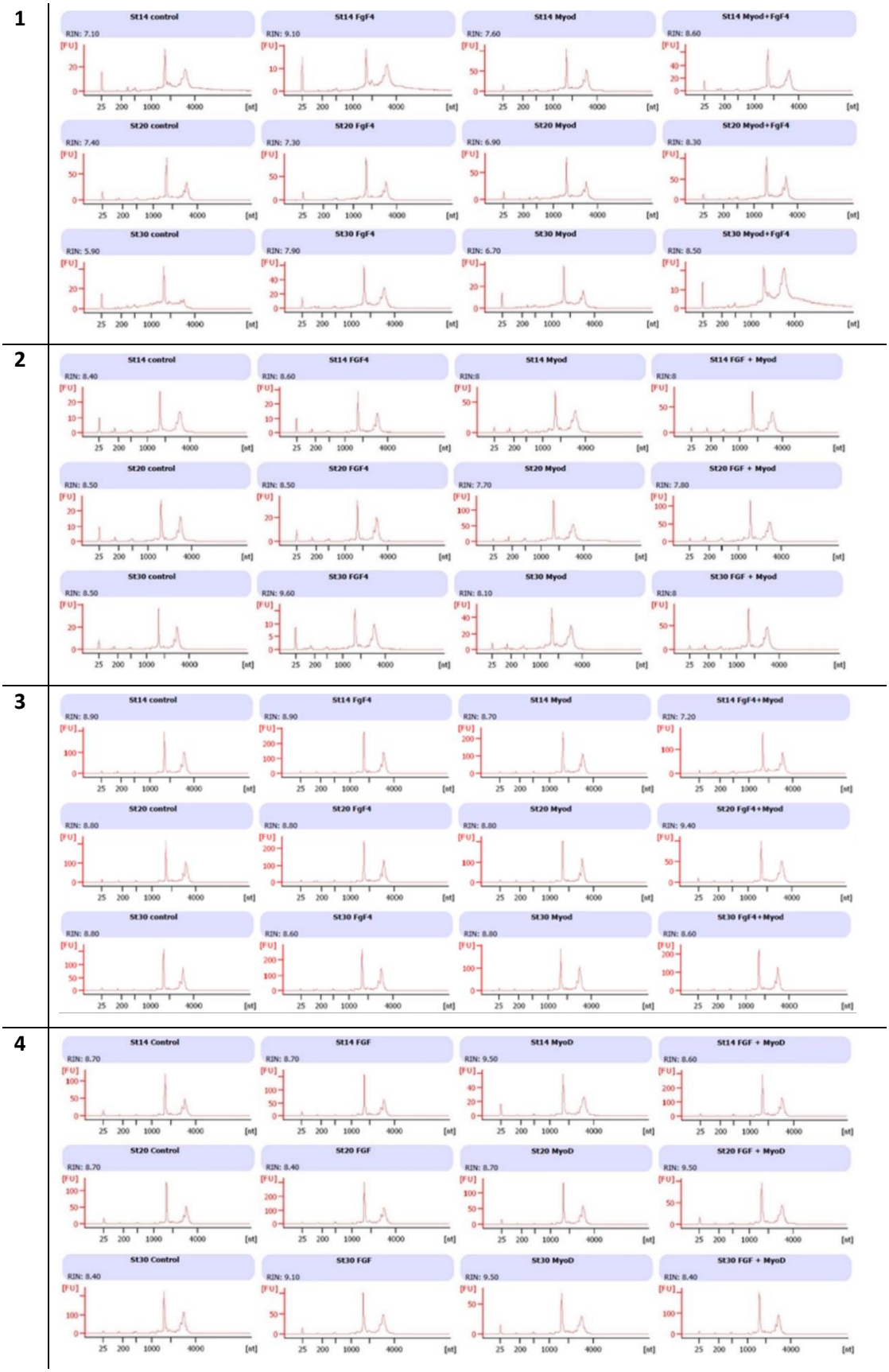


Figure 4.2: Electropherograms and RNA integrity number (RIN) for RNA-seq samples. Animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1.2ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured with or without 50ng/ml Fgf4. RNA extracted from 10 organoids at stages 14, 20 or 30. 4 replicates (1, 2, 3 and 4).

Four replicates were carried out though one set of samples (replicate 1) was later excluded. Replicate 1 varied from the other three replicates as shown in principal component analysis (Figure 4.3). For example, replicate 1 of the Myod expressing condition at stage 30 (S30.M.1) did not cluster with the other stage 30 Myod samples.

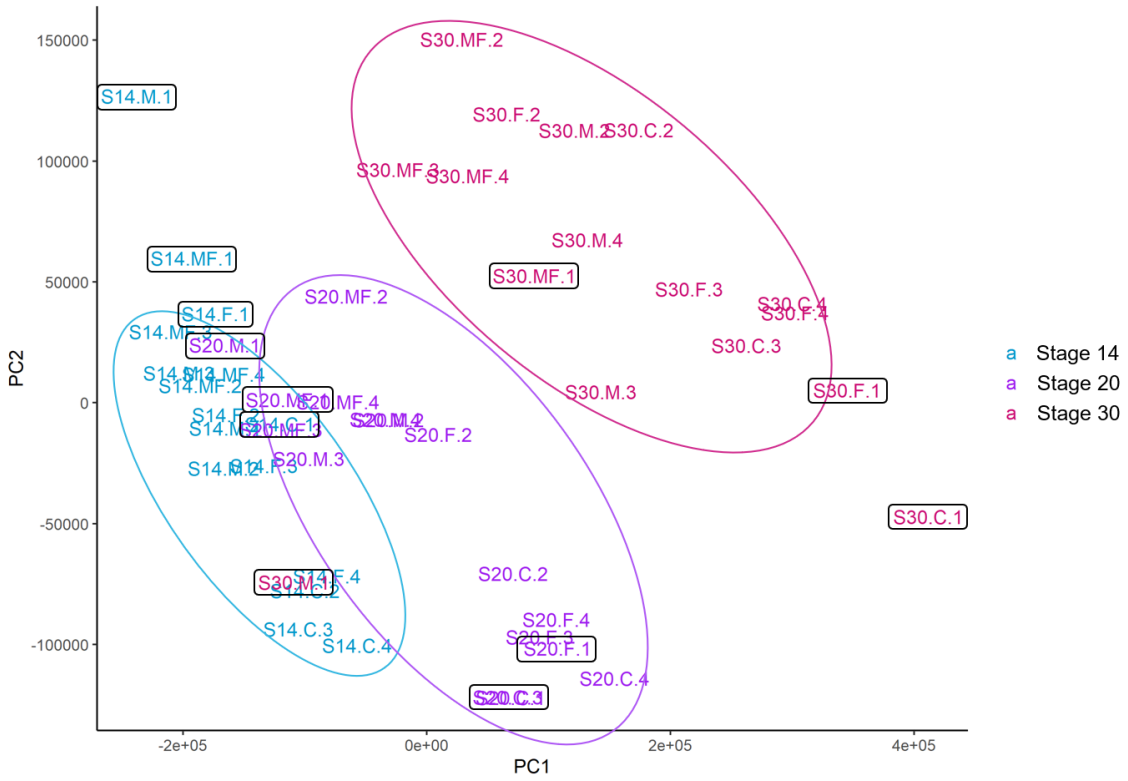
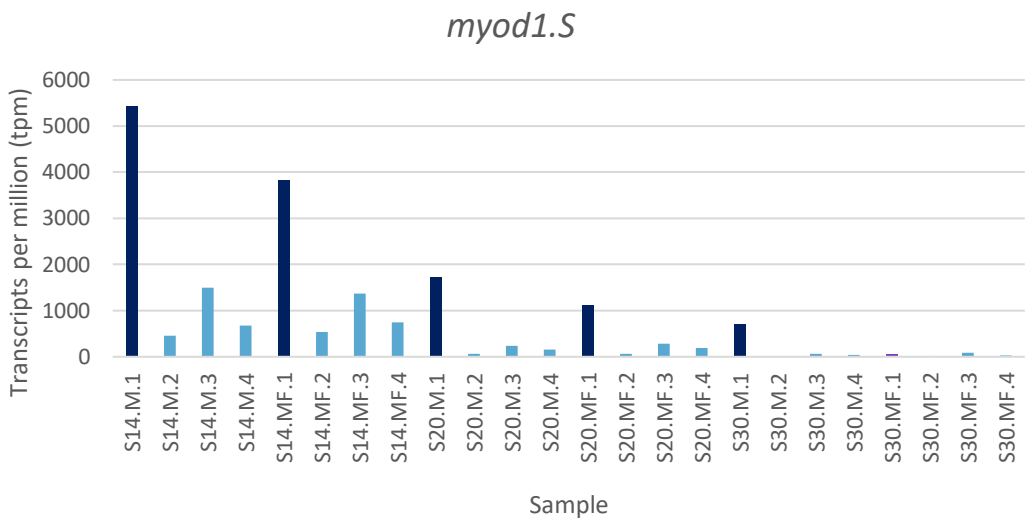


Figure 4.3: Principal component analysis plot. PCA plot of RNA-seq samples at each developmental stage (S14 = stage 14, S20 = stage 20, S30 = stage 30), for each experimental condition (C = control organoids, F = Fgf4 treated organoids, M = Myod expressing organoids, MF = Myod + Fgf4 organoids), in each replicate (1, 2, 3, 4). Replicate 1 (labels highlighted with boxes) excluded from further analysis due to variation from rest of dataset.

Expression of *myod1.S* was higher in replicate 1 Myod and Myod + Fgf4 samples compared to the other replicates, with the greatest differences seen at stage 14 (Figure 4.4A). A one-way ANOVA was performed to determine whether the differences between replicates in *myod1.S* expression for Myod and Myod + Fgf4 samples were significant (Figure 4.4B). ANOVA results revealed a statistically significant difference between at least two replicates ($F(3,20) = 4.07$, $p = 0.02$, $F\ crit = 3.10$). A Tukey-Kramer post-hoc test revealed significant pairwise differences between replicates 1 and 2, and replicates 1 and 4 (Figure 4.4C).

A



B

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Sum of Squares	F value	P value	Critical F value
Between Groups	14835096	3	4945032	4.07	0.02	3.10
Within Groups	24290912	20	1214546			

C

Replicate Comparison	Absolute Mean Difference	Critical Q value	Significant?
1 vs 2	1955.1	1781.7	Yes
1 vs 3	1557.3	1781.7	No
1 vs 4	1841.4	1781.7	Yes
2 vs 3	397.8	1781.7	No
2 vs 4	113.8	1781.7	No
3 vs 4	284.1	1781.7	No

Figure 4.4: *myod1.S* expression. A Transcripts per million (tpm) for *myod1.S* at each developmental stage (S14 = stage 14, S20 = stage 20, S30 = stage 30), for two experimental conditions (M = Myod expressing organoids, MF = Myod + Fgf4 organoids), in each replicate (1 (dark blue), 2, 3, 4). **B** One-way ANOVA. F > Critical F, indicating statistically significant differences between replicates. **C** Tukey-Kramer post-hoc test. Absolute mean difference > Critical Q when comparing replicate 1 with 2 or 4, indicating significant pairwise differences. Replicate 1 excluded from further analysis due to variation from rest of dataset.

4.2.4 Skeletal muscle marker gene expression profiles

In order to characterise skeletal muscle induced by the protocol, gene expression was analysed at developmental stages 14, 20 and 30, and hierarchically clustered by Euclidean distance. Heatmaps were clustered by row in order to identify groups of genes with similar expression profiles within each gene family. Dendrograms represent the Euclidean distance (similarity) between rows and which nodes each gene belongs to as a result of the clustering calculation.

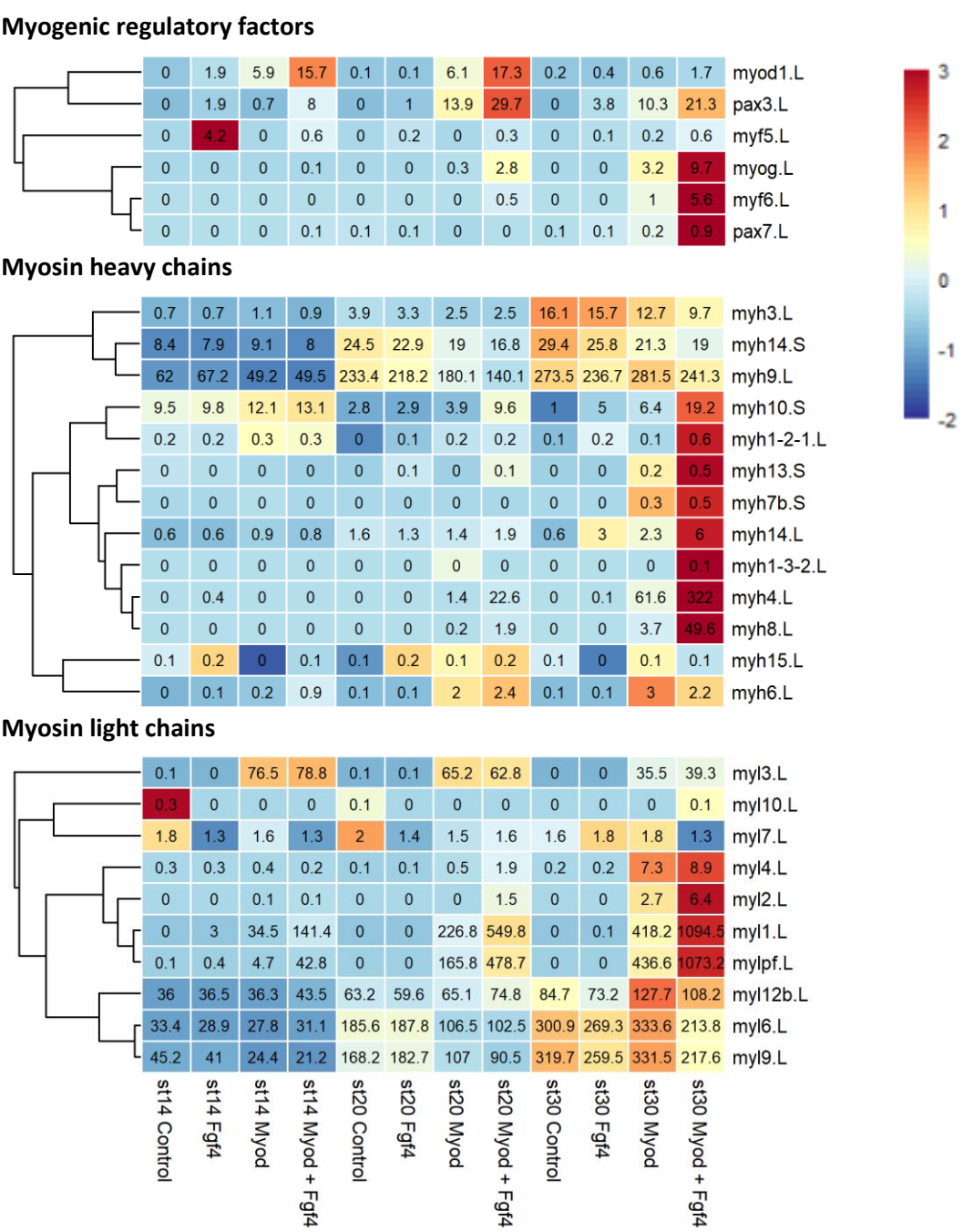
MRFs (Figure 4.5A) and muscle differentiation markers (Figure 4.5H) typically showed a slight increase in expression in Myod organoids, with highest expression in Myod + Fgf4 organoids. Muscle markers were selected with a variety of roles ranging from regulation of early myogenesis to those required for muscle function and structure. For instance, *pax3.L* and *pax7.L* are involved

in the regulation of myogenic progenitor entry into the skeletal muscle differentiation programme during development and regeneration (Buckingham and Relaix, 2007; Ridgeway and Skerjanc, 2001). Other differentiation markers analysed include muscle structural proteins nebulin (*neb.L*) and titin (*ttn*); metabolic enzymes creatine kinase (*ckm.L*) and glycogen phosphorylase (*pygm.L*); and calsequestrin (*casq1/2)* which is the major regulator of calcium ion storage and release in the sarcoplasmic reticulum (Furst et al., 1989; van Deursen et al., 1993; Wang and Michalak, 2020). Myozenin (*myoz1/2*) encodes calsarcin which tethers calcium- and calmodulin-dependent protein phosphatase calcineurin to α-actinin (Frey et al., 2000). Desmin (*des.1.L*) is a muscle-specific intermediate filament required for the formation and maintenance of muscle structure and function (Capetanaki et al., 1997). *myod1.L* and *des.1.L* were present earlier in development before a decrease in expression between stages 20 and 30. Myomaker (*mymk.L*) is a muscle-specific membrane protein which controls myoblast fusion, for which expression was also greater in stage 20 Myod + Fgf4 organoids than at stage 30 (Millay et al., 2013). In contrast, the highest expression for *myf5.L* was observed in stage 14 Fgf4 treated organoids.

Myosins are motor proteins responsible for mechanical force driving processes such as muscle contraction, cell motility or vesicular transport along actin filaments (Weiss et al., 1999). Sarcomeres (the basic contractile units of a myocyte) contain many parallel actin (thin) and myosin (thick) filaments. Globular myosin heads bind actin, contract, release actin, and then reach forward again to repeat the ATP-dependent process (Cooper, 2000). The myosin molecule consists of one or two heavy chains, with one or more light chains associated with each (Korn, 2000). The highest expression for a subset of myosin heavy chains (Figure 4.5B) and myosin light chains (Figure 4.5C) was seen in stage 30 Myod + Fgf4 organoids. Myosin heavy chain 4 (*myh4.L*), myosin light chain 1 (*myl1.L),* and myosin light chain 11 (*mylpf.L)* had significantly higher mean tpm values in these organoids (321.98, 1094.51 and 1073.2 respectively) compared to the expression of other myosins. Actin expression varied with beta actin (*actb.L*), actin gamma 1 (*atcg1.L*) and actin alpha 2 (*acta2.L*) expression increasing over time in all conditions, but particularly in control organoids (Figure 4.5D). The biggest increase in actin expression during the protocol was observed with muscle-specific actin alpha 4 (*acta4.L*) in stage 30 Myod + Fgf4 organoids. Cardiac actin alpha (*actc1.L*) (which is expressed in both the somites and heart during development) was expressed highly in Myod + Fgf4 organoids at each stage.

α-actinins (*actn*) are cytoskeletal actin-binding microfilament proteins. In striated muscle, α-actinins cross-link actin filaments to stabilise the muscle contractile apparatus (Salucci et al., 2015). Tropomyosins (*tpm*) form continuous polymers along most actin filaments and contribute to the regulation of muscle contraction and relaxation (Gunning et al., 2008). Troponins (*tnn*) are found in

striated muscle and control the position of tropomyosin to regulate the interaction of myosin with actin in a calcium-dependent manner (Squire and Morris, 1998). Actinins and tropomyosins followed two types of expression pattern with *actn1.L*, *actn4.L*, *tpm3.L* and *tpm4.L* expressed in all conditions, whereas muscle-specific *actn3.L*, *tpm1.L* and *tpm2.L* were expressed highly in stage 30 Myod + Fgf4 organoids (Figure 4.5E, 4.5F). Troponin expression was typically greatest in stage 30 Myod + Fgf4 organoids with *tnni2.L, tnnc2.L* and *tnnt3.L* presenting the highest mean tpm values (353.9, 439.6 and 305.2 respectively) (Figure 4.5G).



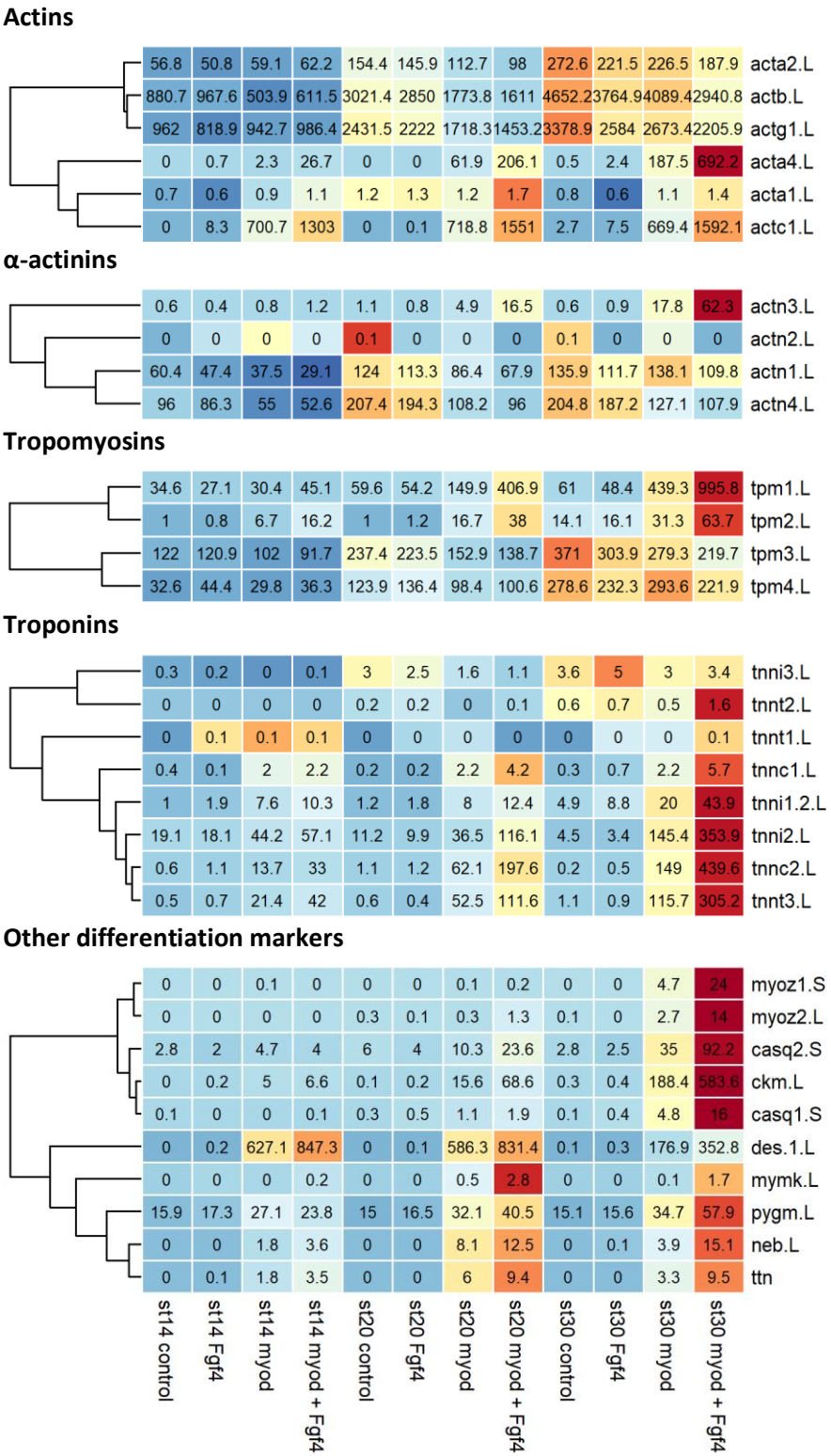


Figure 4.5: Heatmaps showing myogenic gene expression profiles at stages 14, 20 and 30 of the skeletal muscle protocol. Animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1.2ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured with or without 50ng/ml Fgf4. RNA extracted from organoids collected at stages 14, 20 and 30 for RNA-seq. Mean tpm values to 1 decimal place. Heatmaps hierarchically clustered and colours scaled by row. **A** myogenic regulatory factors, **B** myosin heavy chains, **C** myosin light chains, **D** actins, **E** α-actinins, **F** tropomyosins, **G** troponins, **H** other differentiation markers.

4.2.5 Myod co-factor expression

In order to determine whether Myod co-factors were influencing cell fate, the expression patterns of Mef2, Pbx and Meis genes were analysed. Expression of Mef2 and Pbx genes were typically elevated in Myod organoids but highest in Myod + Fgf4 organoids (Figure 4.6A and B). However, *pbx2.L* expression was greatest in stage 14 Myod organoids. The expression pattern of *mef2d.*L was similar to *myod1.L* with the highest tpm in stage 20 Myod + Fgf4 organoids, closely followed by stage 14 Myod + Fgf4 organoids (Figure 4.5A and 4.6A). The highest tpm for Meis expression was observed in stage 14 Fgf4 organoids for *meis3.S*, closely followed by stage 14 controls (Figure 4.6C). The other Meis genes typically had their highest expression in Myod + Fgf4 organoids, though *meis2.L* expression was also high in stage 30 Myod organoids.

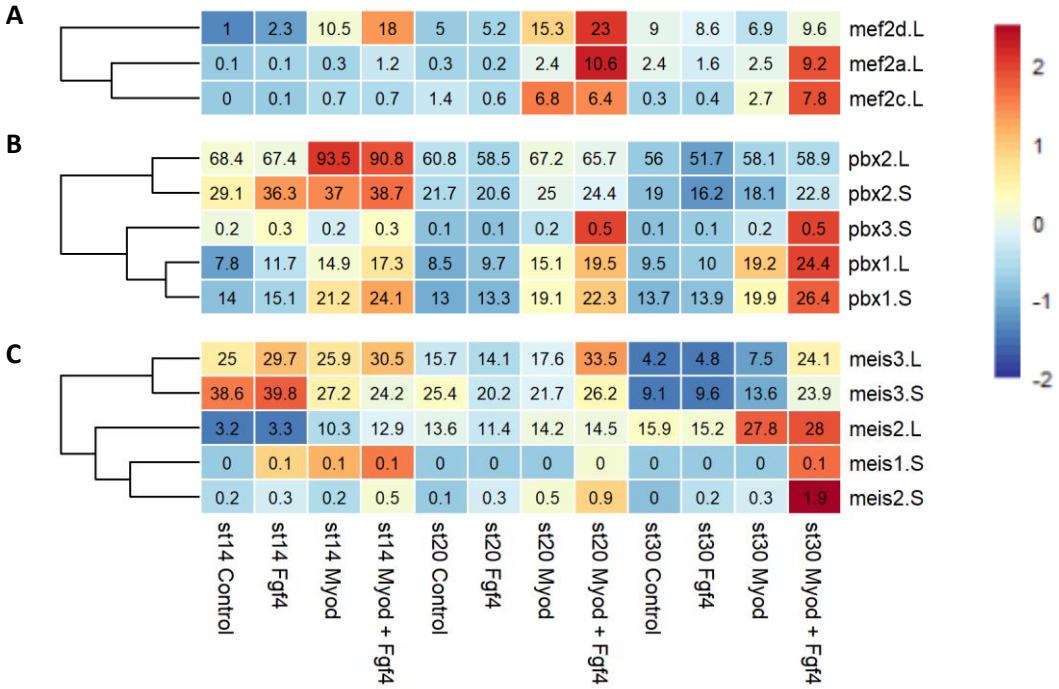


Figure 4.6: Heatmaps showing Myod co-factor expression at stages 14, 20 and 30 of the skeletal muscle protocol. Animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1.2ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured with or without 50ng/ml Fgf4. RNA extracted from organoids collected at stages 14, 20 and 30 for RNA-seq. Mean tpm values to 1 decimal place. Heatmap hierarchically clustered and colours scaled by row. **A** myocyte enhancer factor 2 genes **B** pre-B-cell leukaemia homeobox genes, **C** meis homeobox genes.

4.2.6 Myod + Fgf4 organoids express skeletal muscle specific genes

In order to validate RNA-seq results, quantitative real-time PCR (qPCR) was carried out for two known muscle-specific genes: *myh4.L* and *acta4.S* (Figure 4.7). In both the qPCR and RNA-seq data, a small amount of *myh4.L* was present in the Myod organoids at stage 30, with significantly more in the Myod and Fgf4 organoids. A similar pattern was observed for *acta4.S* expression, though expression began earlier in development.

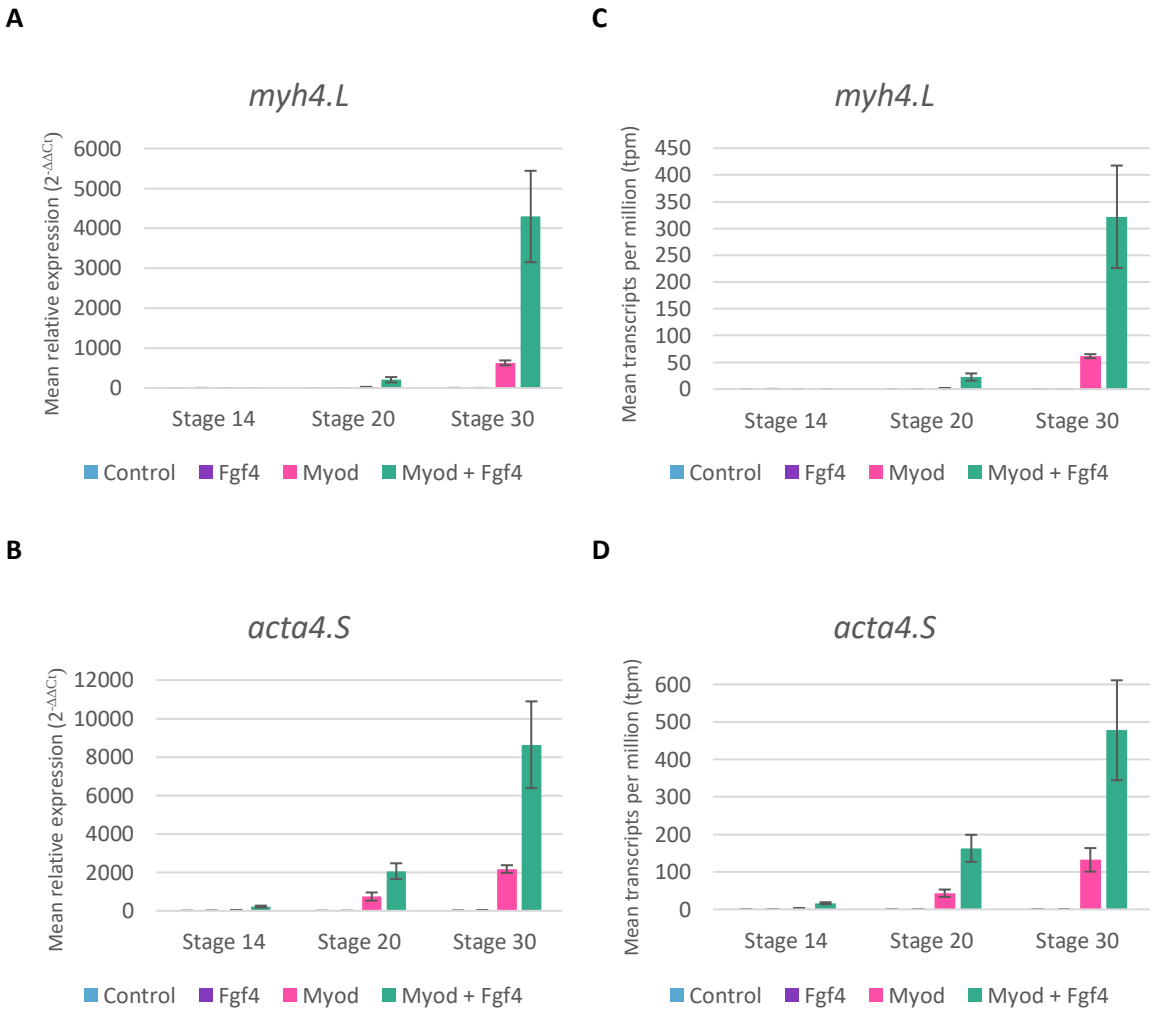


Figure 4.7: qPCR and RNA-seq analysis of *myh4.L* (myosin heavy chain 4) and *acta4.S* (actin alpha 4) expression. A, B mean relative expression by qPCR ($2^{-\Delta\Delta Ct}$). **C, D** mean transcripts per million (tpm) by RNA-seq. Animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1.2ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured with or without 50ng/ml Fgf4. RNA extracted from organoids collected at stages 14, 20 and 30. n=3. Standard error bars.

4.2.7 Hox and Six gene expression profiles

Transcription factors often play an important role in lineage determination due to their ability to regulate expression of lineage-specific genes. To determine whether Homeobox genes played a role in the skeletal muscle protocol, the gene expression profiles of subgroups implicated in muscle lineage commitment were then investigated. Hox genes are involved in mesodermal patterning and somite vertebral fate determination (Carapuço et al., 2005), and sine oculis homeobox (Six) genes are expressed throughout myogenesis with *six1* contributing to regulation of Myod expression (Yajima et al., 2010; Maire et al., 2020; Liu et al., 2013). The majority of Hox genes were most highly expressed in Fgf4 organoids at either stage 20 or 30 (Figure 4.8A). However, a subset of Hox genes had their highest expression in Myod + Fgf4 organoids at stage 14 (*hoxb1.S* and *hoxd1.L*), stage 20 (*hoxb2.L*), and stage 30 (*hoxa2.L*). The expression patterns for Six genes varied with the highest expression of *six1.L* seen in stage 30 Myod + Fgf4 organoids (Figure 4.8B). *six4.L* was expressed in all conditions, with the highest expression in stage 20 control organoids, however *six4.S* expression was highest in stage 14 Fgf4 organoids.

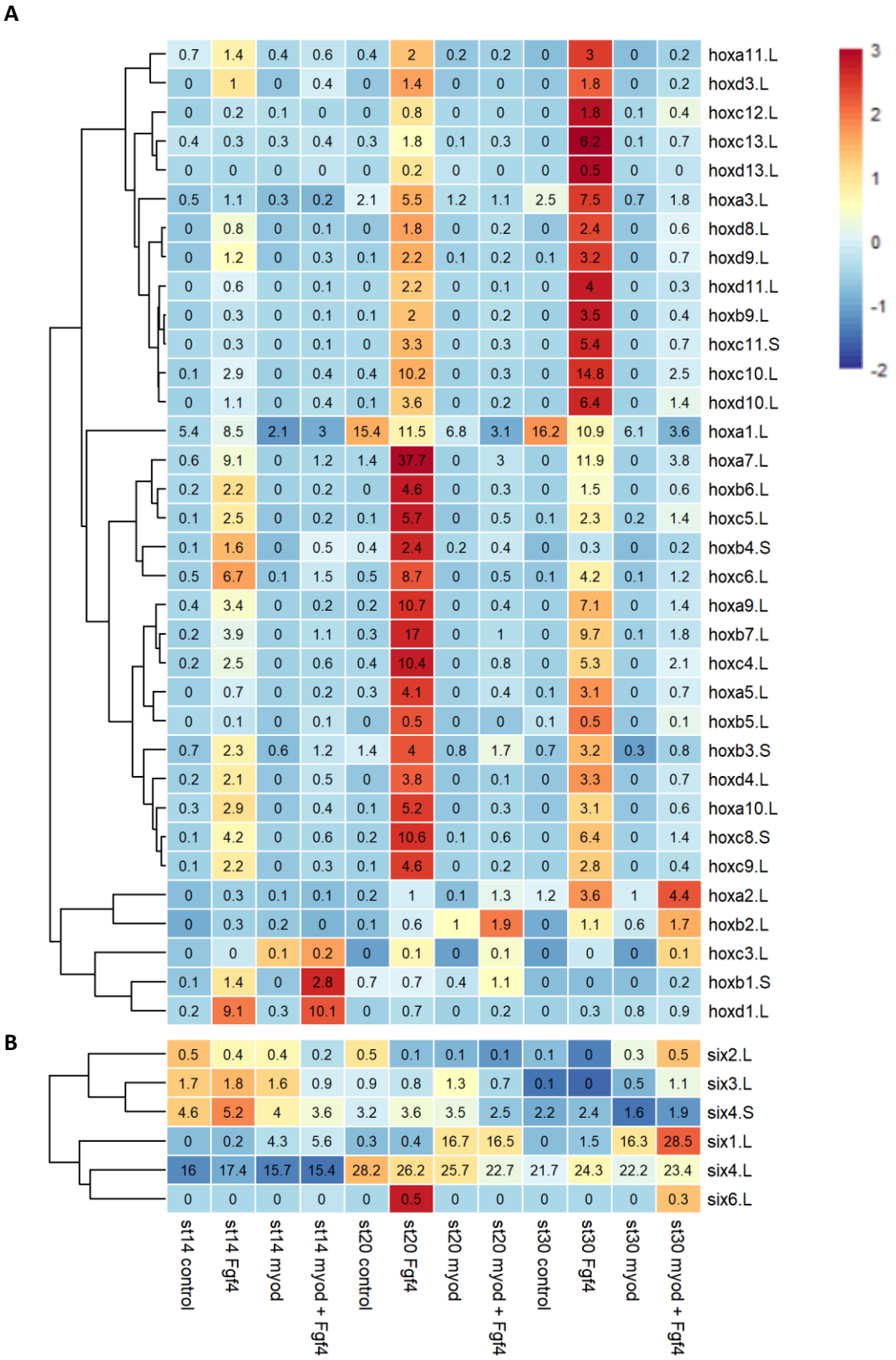


Figure 4.8: Heatmaps of Hox and Six gene expression at stages 14, 20 and 30 of the skeletal muscle protocol. Animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1.2ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured with or without 50ng/ml Fgf4. RNA extracted from organoids collected at stages 14, 20 and 30 for RNA-seq. Mean tpm values to 1 decimal place. Heatmaps hierarchically clustered and colours scaled by row. **A** Hox genes, **B** sine oculis homeobox genes.

4.2.8 Pax and Sox gene expression profiles

As paired box (Pax) and SRY-related HMG-box (Sox) genes also encode transcription factors associated with lineage specification, their expression profiles were then analysed. Sox genes have roles in determination of various cell fates including myogenic progenitors (Meeson et al., 2007), while Pax3 and Pax7 are important for regulation of stem cell entry into the skeletal muscle differentiation programme (Buckingham and Relaix, 2007; Lagha et al., 2008). Pax genes were typically most highly expressed in Myod + Fgf4 organoids, with *pax3.L* exhibiting the highest expression at stage 20 (Figure 4.9A). Sox gene expression patterns varied though many were expressed highly in Myod + Fgf4 organoids (Figure 4.9B). The highest Sox gene tpm value (272.8) was recorded for *sox15.L* in stage 14 Myod + Fgf4 organoids.

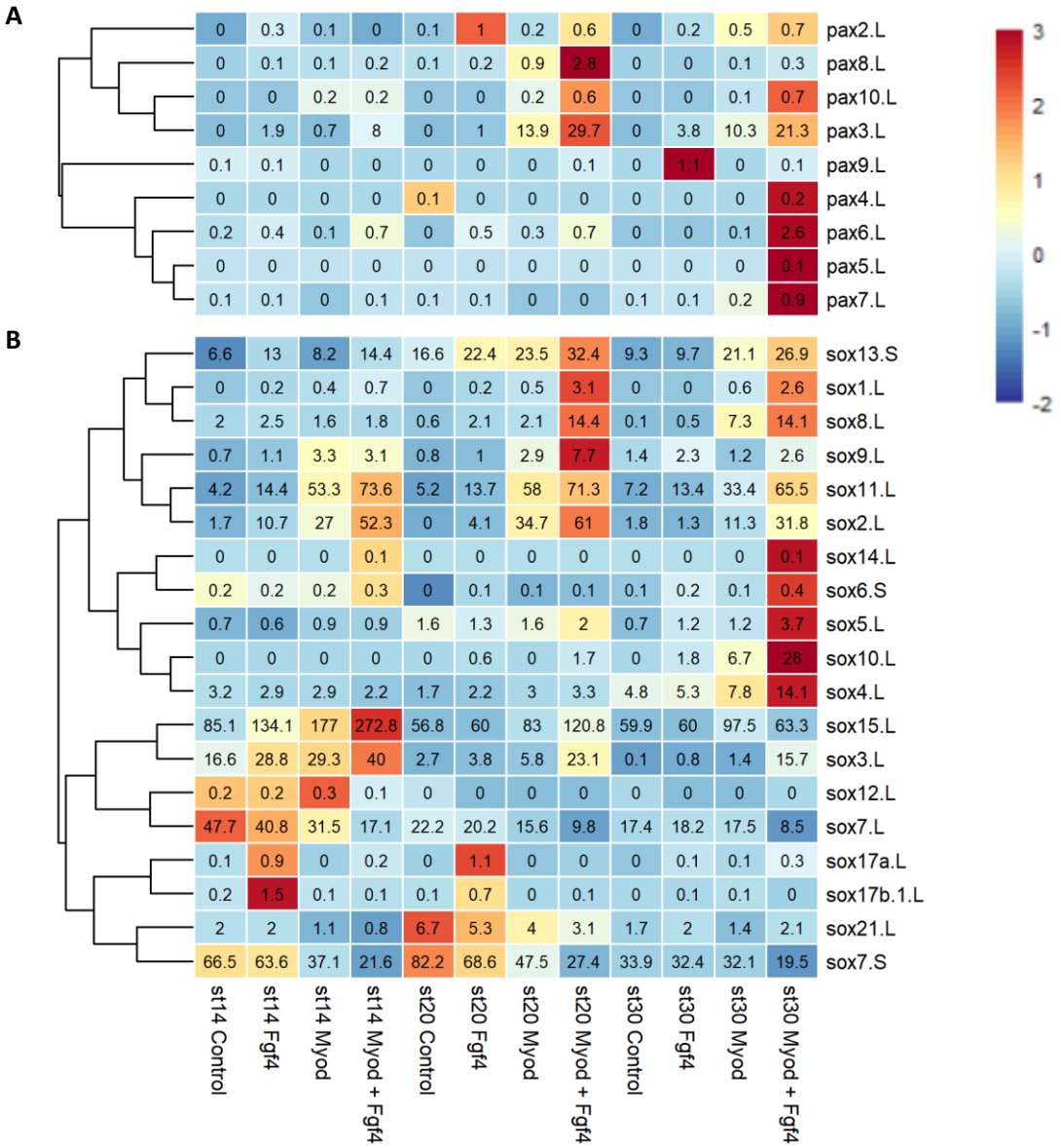


Figure 4.9: Heatmaps of Pax and Sox gene expression at stages 14, 20 and 30 of the skeletal muscle protocol. Animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1.2ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured with or without 50ng/ml Fgf4. RNA extracted from organoids collected at stages 14, 20 and 30 for RNA-seq. Mean tpm values to 1 decimal place. Heatmaps hierarchically clustered and colours scaled by row. **A** paired box genes, **B** SRY-related HMG-box genes.

4.2.9 Celf gene expression profiles

Generation of tissue-specific splice variants is another mechanism which can contribute to lineage commitment and differentiation. For example, the CUGBP Elav-like family member (Celf) RNA-binding proteins are implicated in alternative splicing in striated muscle and brain development, as well as axon regeneration (Ladd et al., 2001; Chen et al., 2016). Celf genes were typically highly expressed in Myod + Fgf4 organoids with expression of *celf1.S, celf2.L* and *celf3.L* highest at stage 20, and *celf1.L* at stage 14 (Figure 4.10).

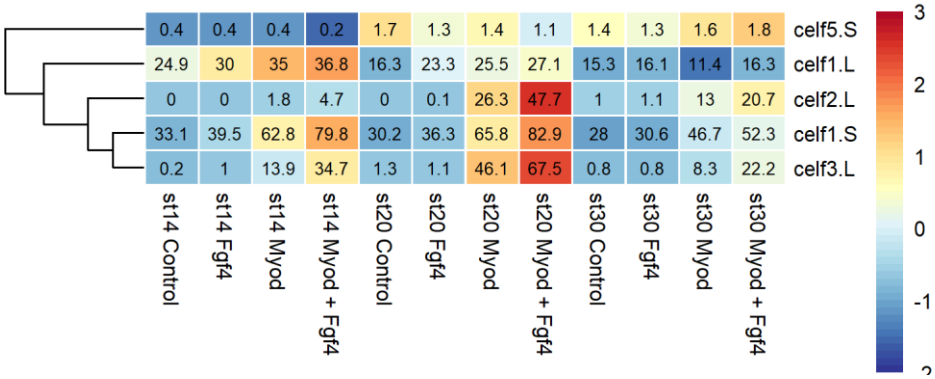


Figure 4.10: Heatmap showing Celf gene expression at stages 14, 20 and 30 of the skeletal muscle protocol. Animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1.2ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured with or without 50ng/ml Fgf4. RNA extracted from organoids collected at stages 14, 20 and 30 for RNA-seq. Mean tpm values to 1 decimal place. Heatmap hierarchically clustered and colours scaled by row.

4.2.10 Pluripotency gene expression profiles

Expression patterns for pluripotency genes *pou5f3* and *ventx1/2* were then analysed to inform when cells exited pluripotency during the protocol. Pou5f3 and Ventx1/2 are *Xenopus* homologs of pluripotency genes Oct and Nanog respectively (Nishitani et al., 2015; Scerbo et al., 2012). The majority of pluripotency factors were downregulated after stage 14, though *ventx* expression remained relatively high in stage 20 Fgf4 organoids (Figure 4.11).

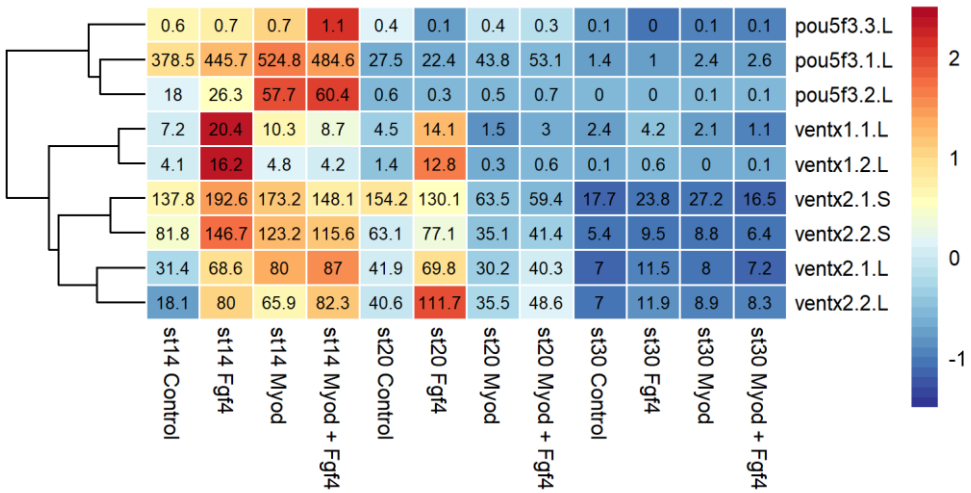


Figure 4.11: Heatmap showing Pou5f3 and Ventx1/2 gene expression at stages 14, 20 and 30 of the skeletal muscle protocol. Animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1.2ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured with or without 50ng/ml Fgf4. RNA extracted from organoids collected at stages 14, 20 and 30 for RNA-seq. Mean tpm values to 1 decimal place. Heatmap hierarchically clustered and colours scaled by row.

4.2.11 Keratin gene expression profiles

As control animal cap organoids differentiate into epidermis, of which epidermal keratins are a key component, the expression of keratin genes was then analysed. Keratin expression varied with *krt18.3.L* and *krt8.2.L* more highly expressed in stage 14 control organoids, while expression of other keratins increased by stage 30 (Figure 4.12). In contrast, *krt18.1.L* was particularly upregulated in stage 30 Myod organoids, as was *krt12.2.L* in stage 30 Fgf4 organoids.

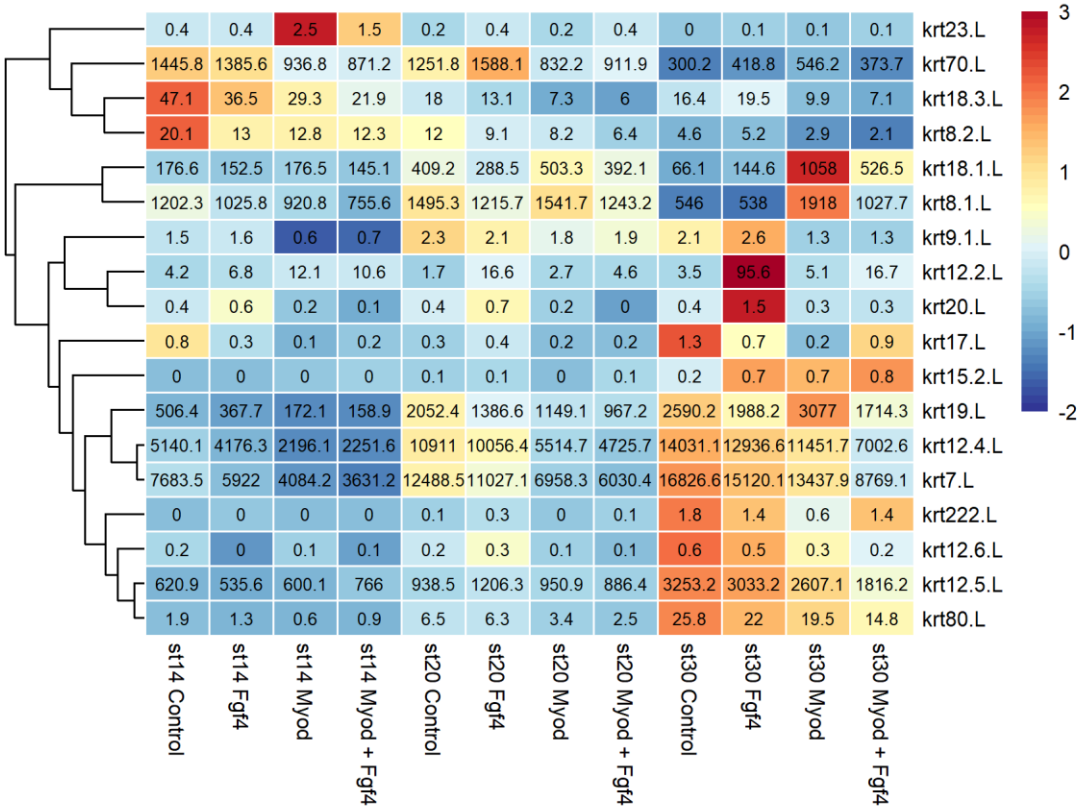


Figure 4.12: Heatmap showing Keratin gene expression at stages 14, 20 and 30 of the skeletal muscle protocol. Animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1.2ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured with or without 50ng/ml Fgf4. RNA extracted from organoids collected at stages 14, 20 and 30 for RNA-seq. Mean tpm values to 1 decimal place. Heatmap hierarchically clustered and colours scaled by row.

4.2.12 Myod + Fgf4 organoids express genes observed in *Xenopus tropicalis* and human skeletal muscle

In order to further characterise the skeletal muscle induced by the protocol, gene expression was compared to existing *Xenopus tropicalis* (*X. tropicalis*) and human skeletal muscle datasets. *X. tropicalis* RNA-seq data was collected by Barbosa-Morais and colleagues (2012) from 20µg RNA samples of skeletal muscle pooled from at least two adult females and two adult males. 93.6% of skeletal muscle genes identified in *X. tropicalis* by RNA-seq (Barbosa-Morais et al., 2012) were also expressed in stage 30 Myod + Fgf4 organoids (Figure 4.13A).

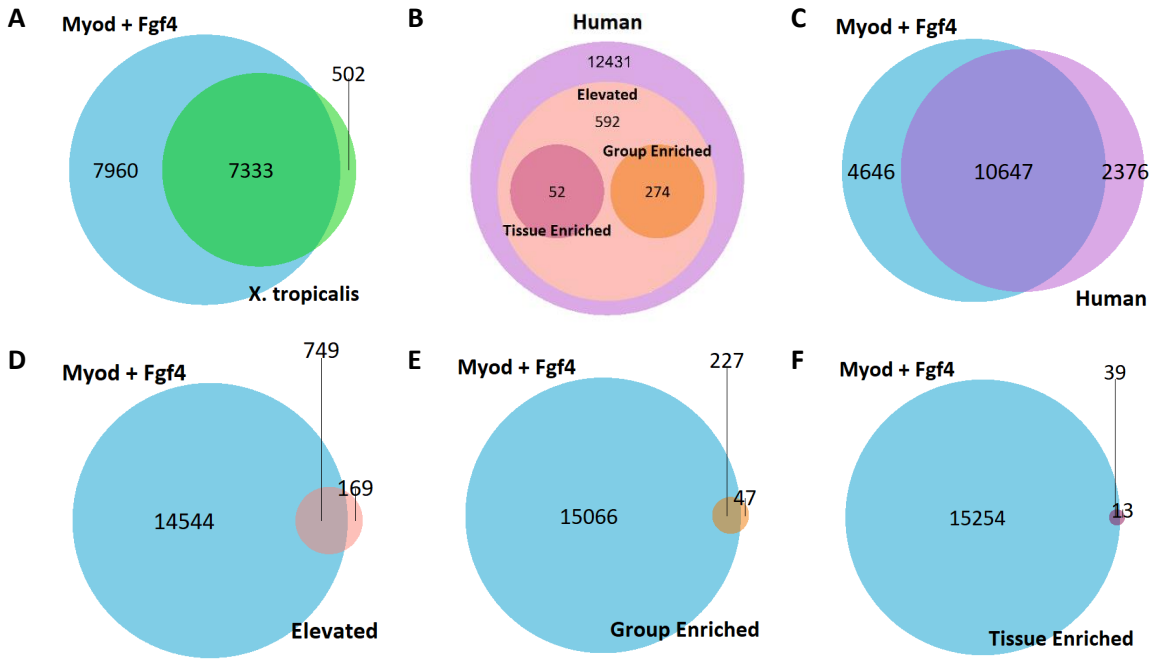


Figure 4.13: Venn and Euler diagrams of genes expressed in stage 30 Myod + Fgf4 organoids, *X. tropicalis* skeletal muscle and human skeletal muscle. A Myod + Fgf4 = genes expressed in animal pole cells dissected from *X. laevis* embryos injected with 1.2ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured with 50ng/ml Fgf4 until stage 30, X. tropicalis = *Xenopus tropicalis* skeletal muscle genes identified in RNA-seq (Barbosa-Morais et al., 2012), **B, C, D, E, F** Human = human skeletal muscle genes, Elevated = human genes elevated in skeletal muscle compared to other tissue types, Group Enriched = human genes with at least four-fold higher average mRNA level in a group of 2-5 tissues compared to any other tissue, Tissue Enriched = human genes with at least four-fold higher mRNA level in skeletal muscle compared to any other tissue. (Human Protein Atlas: Lindskog et al., 2015; Uhlén et al., 2015; Fagerberg et al., 2014).

Human skeletal muscle gene expression data was obtained from the Human Protein Atlas (HPA) in which two transcriptomic datasets, from the HPA and Genotype-Tissue Expression (GTEx) projects, were combined for transcript profiling. This corresponds to a total of 14,590 samples from 50 different normal human tissue types (Lindskog et al., 2015; Uhlén et al., 2015; Fagerberg et al., 2014) (https://www.proteinatlas.org/humanproteome/tissue/skeletal+muscle and https://www.proteinatlas.org/about/assays+annotation#hpa_rna). The HPA dataset analysed 40 human tissues, 1206 cell lines, 18 blood cell types, and total peripheral blood mononuclear cells. The second dataset was part of the GTEx project, in which RNA-seq data was collected from 36

human post-mortem tissue types, including 803 skeletal muscle samples. The combined data was reported in the following categories (Figure 4.13B):

- Expressed in skeletal muscle (13,023 genes)
- Elevated expression compared to other tissue types (918 genes)
- Group enriched: at least four-fold higher average mRNA level in a group of 2-5 tissues compared to any other tissue (274 genes)
- Tissue enriched: at least four-fold higher mRNA level in skeletal muscle compared to any other tissue (52 genes).

81.8% of genes expressed in human skeletal muscle were also expressed in stage 30 Myod + Fgf4 organoids (Figure 4.13C). 81.6%, 82.8% and 75.0% of elevated, group enriched and tissue enriched were expressed in stage 30 organoids respectively (Figures 4.13D-F).

4.2.13 Myod and Myod + Fgf4 organoids express neural genes

While overexpression of *myod* in organoids is not sufficient for effective myogenesis, histological analysis revealed Myod organoids contained structures resembling neuroectoderm (Chapter 3). In order to investigate this further, the expression patterns of neural genes in organoids at stages 14, 20 and 30 were analysed (Figure 4.14).

Neural markers were selected with various roles in neurogenesis including pro-neural bHLH transcription factors *neurod1*, neurogenin (*neurog1/2)* and achaete-scute family bHLH transcription factor 1 (*ascl1.L*); sox genes associated with the neural lineage; and transcription factor forkhead box D3 (*foxd3.L*), transcription factor AP-2 alpha (*tfap2a.L*) and snail family transcriptional repressor 2 (*snai2.S*), which regulate neural crest development (Seo et al., 2007b; Castro et al., 2011; Pevny and Placzek, 2005; Wang et al., 2011; Aybar et al., 2003). Myelin transcription factor 1 (*myt1.L*) encodes a zinc finger protein which promotes neuronal differentiation downstream of Ascl1 (Vasconcelos et al., 2016). Receptors neuropilin 1 and 2 (*nrp1.L* and *nrp2.L*) and their semaphorin ligands (*sema3a.L* and *sema3f.L*) regulate gangliogenesis and axon guidance in the sympathetic nervous system (Schwarz et al., 2009; Maden et al., 2012). Members of the zic family of zinc finger transcription factors are involved in neural tissue and neural crest development, and neural progenitors express nestin (*nes.L*) before a decrease in expression upon differentiation (Merzdorf, 2007; Frederikson and McKay, 1988). Tubulin β 2B class IIb (*tubb2b)* is exclusively expressed in neurons and associated with neuronal migration and axonal guidance, and neural cell adhesion molecule (*ncam1.L*) is a glycoprotein strongly expressed in the nervous system with roles in cell migration, neurite outgrowth and synaptic plasticity (Daume et al., 2022; Romaniello et al., 2012; Rønn et al., 1998).

Neural gene expression varied with several genes showing their highest expression in Myod + Fgf organoids at different stages of development. However, expression of *ascl1.L* and downstream target *myt1.L* was highest in stage 14 Myod organoids, and *snai2.S* and *tfap2a.L* expression was highest in stage 30 Myod organoids.

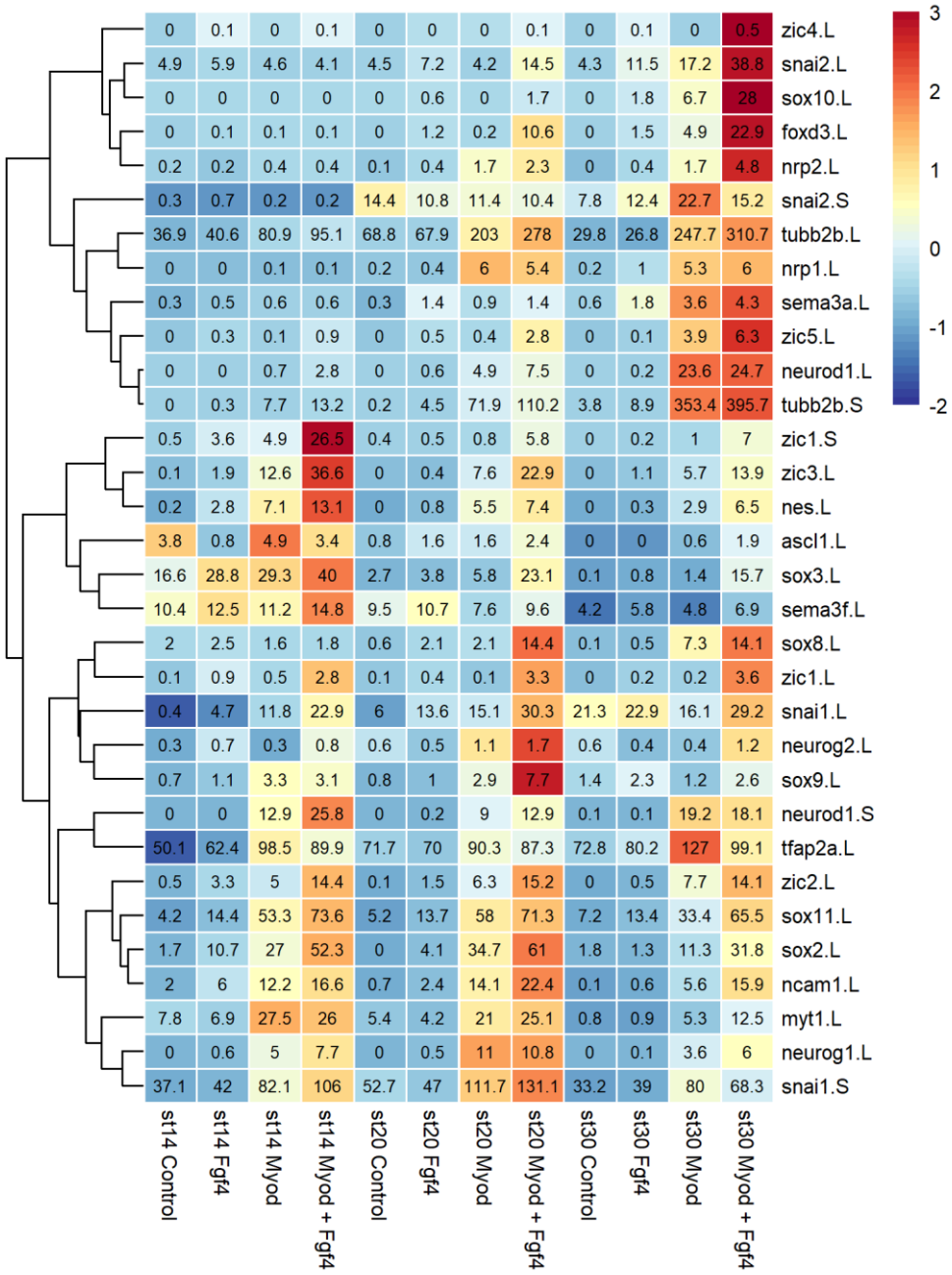


Figure 4.14: Heatmap showing neural gene expression at stages 14, 20 and 30 of the skeletal muscle protocol. Animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1.2ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured with or without 50ng/ml Fgf4. RNA extracted from organoids collected at stages 14, 20 and 30 for RNA-seq. Mean tpm values to 1 decimal place. Heatmap hierarchically clustered and colours scaled by row.

4.3 Discussion

4.3.1 Samples selected for RNA-seq analysis

Skeletal muscle protocol organoids were collected at early neurula stage 14, late neurula stage 20 and late tailbud stage 30 for transcriptomic analysis. Somite formation in the *Xenopus* embryo begins at stage 17 and continues until stage 40, so analysis of organoids at these stages allowed insight into the transcriptional profile during early muscle specification as well as during differentiation (Afonin et al., 2006).

The skeletal muscle protocol was carried out four times in order to generate four replicates of samples at these three stages for RNA-seq. However, replicate 1 was later excluded from the analysis for a number of reasons. Clear differences between replicate 1 and the other samples were observed via PCA plot, with samples such as stage 30 Myod replicate 1 (S30.M.1) not clustering with the other stage 30 samples. In addition to this, the lowest RIN values were observed for replicate 1 samples. It is generally accepted that the ideal RIN for RNA samples for RNA-seq library construction is > 7 (Damaraju et al., 2012). Though studies have been conducted using samples with lower RINs, use of lower quality RNA increases the chance of uneven degradation of different RNA fractions (Puchta et al., 2020). Three replicate 1 samples had RINs lower than 7, with stage 20 Myod, stage 30 control and stage 30 Myod samples recording 6.90, 5.90 and 6.70 respectively. Furthermore, all three of these samples do not cluster well in PCA analysis. Upon inspection of the resulting RNA-seq dataset, *myod1.S* expression was greater in replicate 1 Myod and Myod + Fgf4 samples compared to other replicates. Although all organoids were collected at the same three stages of development, slight differences in collection time may have resulted in differences in *myod.S* expression between replicates. For example, replicate 1 may have been collected earlier, so less injected *myod.S* mRNA had degraded compared to the other replicates. This is supported by the fact that the greatest differences are seen earlier in development. Statistically significant pairwise differences were found between *myod1.S* expression in replicates 1 and 2, and replicates 1 and 4, indicating that replicate 1 was the outlier and other minor differences were not significant. Exclusion of replicate 1 from the analysis resulted in an experiment with three biological replicates and lower q values in subsequent statistical tests, though overall trends remained.

4.3.2 Fgf4 signalling in combination with Myod expression promotes lineage commitment

Transcriptomic data for Myod organoids supports previous work showing that expression of Myod in ES cells and *X. laevis* animal caps leads to some transcription of muscle genes, but is not sufficient for muscle differentiation (Hopwood and Gurdon, 1990; Dekel et al., 1992). Many genes linked to neural lineages were expressed in Myod organoids and at a higher level in Myod + Fgf4 organoids. Regulation of the muscle and neural lineages is complex with different transcriptional profiles

achieved with specific combinations of (sometimes the same) pioneer factors, co-factors, and binding sites.

The highest expression of neural genes *ascl1.L* and *myt1.L* was in stage 14 Myod organoids. This aligns with the fact that Myod organoid histology showed some resemblance to neuroectoderm with the presence of neural tube-like structures (Chapter 3). Zinc finger transcription factor *myt1* is a downstream target of pro-neural bHLH transcription factors Ascl1 (MASH1) and Neurog2 (X-NGNR-1), and has been shown to promote vertebrate neurogenesis (Bellefroid et al., 1996; Vasconcelos et al., 2016). Ascl1 is a pioneer factor required for normal proliferation of neural progenitors and able to convert various cell types to functional neurons (Castro et al., 2011; Vierbuchen et al., 2010; Marro et al., 2011). Ascl1 binds both open and closed chromatin to regulate gene expression and temporal onset of targets required for neurogenesis (Raposo and Vasconcelos, 2015). However, Ascl1 has also been shown to induce some muscle markers (Treutlein et al., 2016; Lee et al., 2020). Many subfamilies of bHLH proteins (such as Myod, Neurod and Ascl) recognise the same canonical sequences, which can result in promiscuous binding of neural or muscle gene promoters (Fong et al., 2012; Lee et al., 2020). Lineage fidelity appears to be achieved through protein-DNA interaction driven chromatin access of these pioneer factors, as well as specific silencing mechanisms of other lineages (Lee et al., 2020; Tapscott, 2005). For example, Myt1l functionally represses the myogenic programme activated by promiscuous binding of Ascl1 to Myod-preferred sites, and the myogenic programme of Myod is dependent on Pbx-Meis interactions (Lee et al., 2020; Mall et al., 2017; Romm et al., 2005; Berkes et al., 2004). Pbx1 is a pioneer transcription factor associated with recruitment of Myod to the *myogenin* promoter, and promoting mesodermal rather than neural fate (Fong et al., 2015). In keeping with this, Pbx genes were typically more highly expressed in Myod + Fgf4 organoids than Myod organoids. However, Pbx1 and Meis1 have also been linked to regulation of the zic3 promoter in the dorsal neural tube, as well as expression of hindbrain and neural crest markers in *Xenopus* (Kelly et al., 2006; Maeda et al., 2002).

Many genes can play a role promoting either myogenic or neuronal lineages depending on specific conditions and co-factor interactions. Another important group of Myod-cooperating transcription factors are the Mef2 genes (Gossett et al., 1989). Myod regulates transcription via a feed-forward circuit and the sequential expression of subsets of genes including the Mef2 genes (Penn et al., 2004; Dodou et al., 2003; Tapscott, 2005). Unlike MRFs, Mef2 family members are expressed widely and are not sufficient to convert fibroblasts to myoblasts on their own (Gossett et al., 1989). However co-operation of Mef2 factors with Myod has been shown to increase efficiency of fibroblast conversion (Molkentin et al., 1995). The highest expression of *mef2a.L, mef2c.L* and

mef2d.L was observed in Myod + Fgf4 organoids. This is in keeping with previous work showing that Myh3 expression requires Mef2d and p38 MAPK activity (Penn et al., 2004). Adding to complexity, Mef2 can also bind Ascl1, as well as myogenic bHLH proteins, in order to direct binding to specific targets (Black et al., 1996).

Upstream of Myod, Pax3 has been shown to induce Six1 and Eya which are regulators of Myod expression (Ridgeway and Skerjanc, 2001; Wardle, 2019). Pax3 and Pax7 are associated with regulation of stem cell differentiation into muscle during development and adult regeneration (Lagha et al., 2008; Buckingham and Relaix, 2007). The highest expression of *six1.L, pax3.L* and *pax7.L* was in Myod + Fgf4 organoids. However Pax3 in combination with Zic1 also drives neural crest specification and determination in *Xenopus* and chicken (Sato et al., 2005; Milet et al., 2013). These data suggest some organoid cells were directed to the neural lineage due to insufficient factors required to promote the muscle lineage or repress the neural lineage.

The highest expression of several neural genes was observed in Myod + Fgf4 organoids, suggesting that Fgf4 may promote lineage commitment even in the context of Myod binding promiscuously to neural gene promoters. It would be interesting to investigate whether FGF signalling also acts to enhance Neurod driven neural differentiation.

4.3.3 Myogenic regulatory factors are expressed in skeletal muscle protocol organoids

The highest expression of *myf5.L* was observed in stage 14 Fgf4 treated organoids. This aligns with previous work in zebrafish in which FGF signalling was shown to act through Tbx16 to directly activate *myf5* and *myod* (Osborn et al., 2020). Although *myf5.L* is not expressed as highly in Myod + Fgf4 organoids, Myf5 is expressed very early and transiently in development and samples were only analysed at 3 developmental stages in this experiment (Ott et al., 1991). While *myf5* and *myod* are both able to activate endogenous *myod*, *myod* is not capable of inducing expression of the endogenous *myf5* gene (Fisher et al., 2003). Therefore it was expected that *myf5.L* would not be expressed in Myod organoids.

myod1.L is expressed at high levels in Myod + Fgf4 organoids at stages 14 and 20, before expression decreases. The highest expression for *myf6.L* (*mrf4*) and *myog.L* is seen in Myod + Fgf4 organoids at stage 30. These expression patterns were as predicted, as Myod (or Myf5) is required for muscle specification, whereas Myogenin is responsible for terminal myoblast differentiation and is expressed later in embryonic development (Rudnicki et al., 1993; Hasty et al., 1993; Weintraub, 1993; Rawls et al., 1995). In mice, Myf6 is expressed early in the ventrolateral dermomyotome, and again in differentiated muscle fibres (Bober et al., 1991; Summerbell et al., 2002). However in

Xenopus, the earliest expression of Myf6 is observed in differentiated anterior myocytes when muscle-specific *myh4* mRNA is also present (Gaspera et al., 2006).

Synthetic *myod1*.S mRNA injected into *X. laevis* embryos is rapidly translated and Myod has been shown to positively autoregulate its own expression (Thayer et al., 1989). However injection of *myod1.S* mRNA and culture in Fgf4 protein resulted in greater expression of myogenic regulatory factors *myod1.L, myf6.L* (*mrf4*), and *myog.L* than organoids injected with *myod1.S* only. This supports the hypothesis that Fgf4 allows differentiation of pluripotent cells into skeletal muscle, as Fgf4 enhances Myod expression in these organoids and thus subsequent feedforward activation of myogenic genes. Fgf4 has been shown to directly induce Myod expression in *X. laevis* (Fisher et al., 2002), however injecting greater quantities of *myod1* mRNA is not sufficient for effective myogenesis, indicating that Fgf4 contributes to lineage specification through additional mechanisms (Hopwood and Gurdon, 1990). For example, another property of Fgf4 is that it is responsible for the community effect in which muscle precursor cells will only differentiate and stably express tissue-specific genes if they are in contact with each other and/or expressing Fgf4 (Standley et al., 2001).

4.3.4 Hox gene expression in protocol organoids

Homeobox genes have been found to be highly overrepresented among hypermethylated genes in the skeletal muscle lineage (Tsumagari et al., 2013). These include Hox genes which play important roles in axial and mesodermal patterning, somite vertebral fate determination and limb bud muscle differentiation (Holland and Garcia-Fernàndez, 1996; Carapuço et al., 2005; Houghton and Rosenthal). Hox genes were previously thought to display temporal collinearity during vertebrate anterior-posterior axis patterning (Monteiro and Ferrier, 2006), however more recent analyses in *Xenopus* indicate that this is not always the case (Kondo et al., 2017, 2019). Though the most anterior *Xenopus* Hox genes are first expressed at early gastrula stage, and the most posterior are expressed last during tailbud stages, the genes located in between (paralogous groups 2-10) show no temporal collinearity (Kondo et al., 2017). This is in keeping with organoid Hox gene expression patterns as the majority do not hierarchically cluster by paralogous group. For example, at the developmental stages analysed, *hoxa2.L* is most highly expressed at stage 30, whereas expression of several more posterior Hox genes peaks at stage 20.

The greatest expression of the majority of Hox genes was in Fgf4 treated organoids, with the highest tpm value attributed to *hoxa7.L* at stage 20. This aligns with previous studies which showed that FGF signalling is required for normal posterior Hox gene expression and regulates Hox gene expression during mesoderm induction, gastrulation and neurulation (Cho and De Robertis, 1990; Pownall et al., 1996). It has also been shown that overexpression of Fgf4 leads to an increase in Hox

gene expression, particularly that of *hoxa7* (Pownall et al., 1996). In *Xenopus, hoxd10, hoxa11, hoxc11* and *hoxa13* have been shown to have sequential anterior expression boundaries in the mesoderm (Lombardo and Slack, 2001). During late tailbud stages 30-32, the mesodermal anterior boundary for *Hoxc11* is around somite level 13, suggesting that 50ng/ml Fgf4 treatment may induce mesoderm resembling that found in the mid-to-posterior trunk (Lombardo and Slack, 2001; Pownall et al., 1996).

The Hox gene with the highest expression at stage 14 was *hoxd1.L* in Myod + Fgf4 organoids. This is in keeping with previous literature as the first Hox gene expressed in *Xenopus* mesoderm is *hoxd1* (Wacker et al., 2004). *hoxa2.L, hoxb2.L* and *hoxb1.S* also showed their highest expression in Myod + Fgf4 organoids. These genes are associated with cardiac muscle in *Drosophila* (Poliacikova et al., 2021). While many aspects of development are conserved between species, *Xenopus* cardiac actin gene *acta4* is initially expressed in skeletal muscle, before being expressed exclusively in cardiac muscle, and this expression pattern may also be observed in other genes (Latinki et al., 2002; Gurdon et al., 1985).

4.3.5 Sox gene expression in protocol organoids

Many Sox genes are highly expressed in Myod + Fgf4 organoids. The first Sox family member (Sry) was discovered as a result of its role in testis determination (Gubbay et al., 1990). Since then, 20 Sox family transcription factors have been identified in human and mouse, each associated with key roles in determination of cell fates including chondrocytes, and neuronal and myogenic progenitor cells (Schepers et al., 2002; Wright et al., 1995; Uwanogho et al., 1995; Meeson et al., 2007). The Sox gene with the highest tpm value was *sox15.L* in stage 14 Myod + Fgf4 organoids, with expression levels then decreasing over time. This is keeping with previous studies as Sox15 is required for myogenic lineage cell determination, with roles both during development and skeletal muscle regeneration from satellite cells (Savage et al., 2009; Meeson et al., 2007; Lee et al., 2004). Sox15 is highly expressed in myoblasts then downregulated to allow myogenic differentiation as overexpression in C2C12 myoblasts has been shown to inhibit Myogenin expression and repress myotube differentiation (Béranger et al., 2000). Several other Sox genes expressed in Myod + Fgf4 organoids have also been associated with the skeletal muscle lineage including *sox11* in mouse satellite cells and *sox5* in trout embryonic development (Oprescu et al., 2023; Rescan and Ralliere, 2010).

4.3.6 Alternative splicing in protocol organoids

In addition to transcription factor gene expression, another developmental regulatory mechanism is alternative splicing to generate required variants (Cooper and Ordahl, 1985). It has been estimated that complex local splicing variants make up over 30% of tissue dependent transcript

variations (Vaquero-Garcia et al., 2016). CELF genes (which encode RNA-binding proteins linked to splicing during striated muscle and brain development) were typically highly expressed in Myod + Fgf4 organoids at stage 14 or 20 (Ladd et al., 2001). Celf1 (also known as CUGBP1) has also been shown to increase translation of Mef2a (Timchenko et al., 2004; Black et al., 1996). The Rbfox (or Fox) genes form another family of RNA-binding proteins associated with alternative splicing in skeletal and cardiac muscle and the brain (Gallagher et al., 2011; Gehman et al., 2012; Das et al., 2007). Celf2 has also been shown to antagonise Rbfox2, yet both were expressed highly in Myod + Fgf4 organoids (Gazzara et al., 2017). However, the RNA-seq analysis was carried out at gene level rather than transcript level and therefore cannot differentiate between spliceforms. In order to detect the presence of novel splicing events, transcript level analysis could be undertaken using the reference transcriptome or by assembling a de novo transcriptome using this dataset.

4.3.7 Fast twitch muscle is induced by the skeletal muscle protocol

Skeletal muscle fibres can be broadly categorised into slow (type I) and fast twitch (type II), with further sub-types within these groups. The concept of fast and slow muscle arose from physiological studies of frog muscles (Peachey and Huxley, 1962). All muscle fibres were initially thought to be able to enter a 'phasic' state with rapid twitches, or a 'tonic' state with prolonged contractions, depending on the experimental conditions (Hess, 1970). In the 1950s, Kuffler and colleagues demonstrated that these two states were exhibited by two distinct fibre populations, and that different compositions of fibre type within an individual muscle determines its characteristics (Kuffler and Vaughan Williams, 1953). Tonic fibres are now often referred to as slow fibres, though they differ from mammalian slow twitch fibres in that they do not fire action potentials (Luna et al., 2015). Slow fibres are typically more resistant to fatigue than fast twitch, partly as a result of a larger mitochondrial volume and greater oxidative enzyme activity, though contractile properties and metabolic capacities vary between species (Jackman and Willis, 1996).

By early tailbud stage 22, *Xenopus* embryos form deep medial fast muscle fibres in anterior somites (Grimaldi et al., 2004). In contrast, the first wave slow fibres arise later in the tail tip of stage 27/28 embryos, initially in the medial somite before becoming more superficial as the bulk of the somite differentiates into fast muscle. By late tailbud stage 35, the posterior half of the embryo contains superficial slow cells around the lateral border of each somite. By tadpole stage 48, the superficial slow fibres form a monolayer at the dorsal and ventral ends of all somites (Grimaldi et al., 2004).

Isoforms of some genes are preferentially expressed in either slow or fast twitch muscle and can be used to aid determination of fibre type. Myod + Fgf4 organoids express high levels of fast twitch skeletal muscle-specific genes including myosins (*myh4.L, myl1.L, mylpf.L*), troponins (*tnnc2.L, tnni2.L, tnnt3.L*), actinin (*actn3.L*), and myozenin (*myoz1.S* which encodes calsarcin-2) (Stuart et al.,

2016; Rasmussen and Jin, 2021; Mills et al., 2001; Frey et al., 2000). Pbx proteins are required for fast twitch muscle differentiation in zebrafish, but not slow (Maves et al., 2007). Enhanced expression of all these genes in Myod + Fgf4 organoids indicates that the skeletal muscle protocol induces fast twitch muscle (Matyushenko et al., 2020). This is in keeping with embryonic myogenesis as fast twitch muscle is formed first in *Xenopus* (Grimaldi et al., 2004).

4.3.8 Myod + Fgf4 organoids express genes observed in *Xenopus tropicalis* and human skeletal muscle

Myod + Fgf4 organoids showed expression of skeletal muscle-specific genes required for lineage regulation, as well as muscle structural proteins and metabolic enzymes. In addition, comparison with existing datasets showed that 93.6% and 81.8% of *X. tropicalis* and human skeletal muscle genes respectively were expressed in stage 30 Myod + Fgf4 organoids. This further confirms the success of the protocol as organoids express the majority of the same genes as in the skeletal muscle of closely related species (Barbosa-Morais et al., 2012; Lindskog et al., 2015; Uhlén et al., 2015; Fagerberg et al., 2014). At least a subset of genes expressed in Myod + Fgf4 organoids but not in the two skeletal muscle datasets, can be attributed to the fact that the organoids are not entirely composed of skeletal muscle cells. Histology presented in Chapter 3 shows that 1.2ng *myod* + 50ng/ml Fgf4 organoids consist of an outer layer of epidermis with an inner block of organised, differentiated skeletal muscle with clear myotubules. Therefore some non-muscle transcripts identified in the RNA-seq, such as keratins, may be exclusively expressed in epidermal cells, and not in other tissue types (Jones et al., 1985; Suzuki et al., 2017). Use of single cell RNA-seq (scRNA-seq) would be beneficial for analysis of specific cell populations within organoids and determining genes expressed in individual cells.

While Myod + Fgf4 organoids contain fast twitch skeletal muscle, and express muscle-specific structural proteins and metabolic enzymes, neural genes are also expressed. This indicates that Fgf4 may have a role promoting lineage specification and differentiation in both the muscle and neural lineages. In chapter 5, the RNA-seq dataset will be used to investigate the role of FGF in skeletal muscle specification, through identification of key signalling pathways and transcriptional targets activated in the protocol.

Chapter 5: Identification and manipulation of potential skeletal muscle lineage regulators

5.1 Introduction

The RNA-seq dataset generated from protocol samples was then analysed to further understand the role of FGF signalling in skeletal muscle lineage commitment, and identify genes with a potential role in skeletal muscle specification. Understanding the precise molecular mechanisms required for the efficient regulation of muscle differentiation may help improve human myoblast cell culture methods, or inform future therapeutics for patients with muscle wasting diseases.

Xenopus models have allowed valuable insight into many developmental processes and diseases, and share a surprising degree of similarity with humans in terms of genome, synteny and anatomy (Tandon et al., 2017; Kostiuk and Khokha, 2021). For instance, it is estimated that the *Xenopus tropicalis* genome contains orthologs of 79% of identified human disease genes (Hellsten et al., 2010).

In recent years, protocols with varying levels of efficiency have been developed to differentiate human skeletal muscle progenitors and satellite cells from pluripotent cells (Shelton et al., 2016; Borchin et al., 2013). These protocols often involve an initial exposure to Chiron, a potent GSK3 inhibitor, to activate Wnt and drive mesoderm specification (Kreuser et al., 2020). This is typically followed by a period of FGF2 treatment, before removal of FGF2, after which cell populations continue to expand for several weeks. In this chapter, an adapted version of Shelton's protocol was used to investigate whether FGF signalling and candidate genes identified in *Xenopus* are conserved in skeletal muscle differentiation of human blastocyst-derived H9 ES cells (Thomson et al., 1998; Shelton et al., 2016).

5.1.1 Aims of this chapter

The aims of this chapter are to:

- Establish the genes and associated biological processes that are significantly upregulated in each condition of the skeletal muscle inducing protocol at stages 14, 20 and 30
- Investigate the role of Fgf4 and other developmental signalling pathways in lineage specification during the skeletal muscle protocol
- Identify candidate genes potentially involved in skeletal muscle lineage specification
- Investigate whether candidate genes are sufficient to replace Fgf4 in the protocol
- Collaborate with Prof. Jenny Nichols (University of Edinburgh) to test whether successful candidate genes have a conserved role in differentiation of human H9 ES cells.

5.2 Results

5.2.1 Differential gene expression analysis of skeletal muscle protocol RNA-seq dataset

Differential gene expression analysis was carried out using the R package Sleuth (http://pachterlab.github.io/sleuth/) by fitting a statistical model to the estimated read counts. Wald tests were performed between each condition within each developmental stage to calculate q values and fold changes/effect sizes (Pimentel et al., 2017). Analysis was carried out at gene level as this is more accurate than transcript level and inferences are thought to be more robust (Soneson et al., 2015). However additional analysis of transcript level abundance estimates would allow differentiation between spliceforms, and could improve differential gene expression results.

Annotated gene transcripts with significant changes in expression were identified by assessing transcripts per million (tpm), q value and effect size for each gene. Tpm is a measure of transcript abundance in each sample adjusted to account for the varying number of reads sequenced for each sample, and the differing expression of transcripts in the whole transcriptome. For example, the tpm value for 'transcript A' would represent the number of transcripts of A that should be observed in one million transcripts sequenced from the whole transcriptome. A q value is a measure of statistical significance of differential expression taking into account false positives. Smaller q values indicate a more significant change in expression and fewer genes expected to be false positives. Effect sizes are calculated by linear models representing the relative change in expression between experimental conditions. For example, an effect size of 1 indicates no change in gene expression between conditions after any batch effects have been taken into account, an effect size of 2 represents a 2-fold upregulation and 0.5 a 2-fold downregulation.

First, mean tpm values for each experimental condition were calculated. A list of genes upregulated by Fgf4 treatment was then compiled for each stage analysed, by identifying genes with mean tpm values for Fgf4 treated samples ≥ 1, q value ≤ 0.05, and effect size compared to control ≥ 1.5. Selection of these parameters was informed by existing knowledge of FGF associated genes in the literature. For example, selection of this effect size cut off meant inclusion of known downstream effector of FGF signalling *etv1.S* (effect size 1.69 in the Fgf4 treated samples at stage 30) in the Fgf4 target gene list (Garg et al., 2018). The same parameters were then applied to the other conditions, using their respective mean tpm values, in order to generate upregulated gene lists for each condition at each stage. Analysis focused on genes upregulated during the protocol as this thesis investigates the role of FGF in Myod driven myogenesis and few genes are negatively regulated by FGF (Branney et al., 2009).

During development, the number of upregulated genes in Fgf4 organoids decreased whereas the number of genes upregulated in Myod and Myod + Fgf4 organoids increased (Table 5.1). Out of 50,487 genes, 321, 282 and 170 fit the criteria for genes upregulated in the Fgf4 treated samples at stages 14, 20 and 30 respectively. 1121, 1795 and 2130 genes were selected for the Myod expressing samples at stages 14, 20 and 30 respectively. The Myod + Fgf4 samples generated the most genes meeting these criteria with 1827, 2867 and 3103 upregulated genes identified at stages 14, 20 and 30 respectively.

Condition	Stage	Number of genes with mean tpm ≥ 1, q value ≤ 0.05, effect size ≥ 1.5
Fgf4	14	321
	20	282
	30	170
Myod	14	1121
	20	1795
	30	2130
Myod + Fgf4	14	1827
	20	2867
	30	3103

Table 5.1: Number of upregulated genes for each condition and stage analysed. Criteria for upregulated gene: mean transcripts per million (tpm) ≥ 1, q value ≤ 0.05 and effect size vs control ≥ 1.5 for each experimental condition at developmental stages 14, 20 and 30.

5.2.2 Gene ontology enrichment analysis of genes upregulated in skeletal muscle organoids

Lists of upregulated genes for the Myod + Fgf4 organoids were submitted to the Protein Analysis Through Evolutionary Relationships (PANTHER) Classification System to identify statistically significant over-represented biological processes at each developmental stage investigated (http://pantherdb.org/) (Mi et al., 2021).

At stage 14, several of the highest enriched gene ontology (GO) terms were associated with various developmental signalling pathways such as the Smoothened, Wnt, MAPK and Notch signalling pathways (Figure 5.1). Histone methylation, protein methylation and protein alkylation were also highly enriched GO terms at this stage. Also within the top 20 enriched GO terms, were muscle associated terms muscle system process, muscle contraction, and actomyosin structure organisation, and neural terms axonogenesis and central nervous system development.

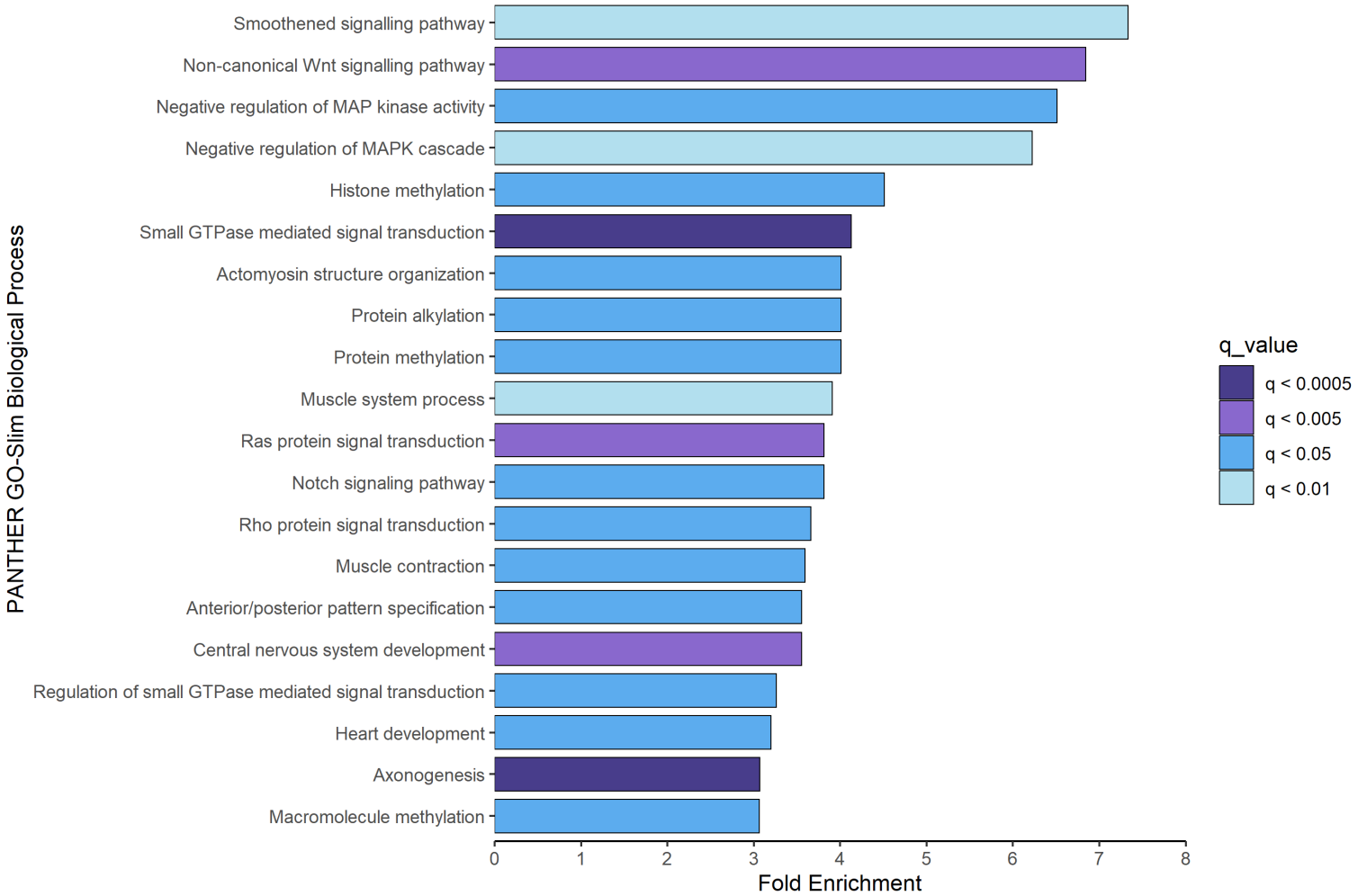


Figure 5.1: Top 20 PANTHER GO-Slim biological processes by fold enrichment over-represented within the upregulated genes in Myod + Fgf4 organoids at stage 14. Genes with mean tpm ≥ 1, q ≤ 0.05, effect size Myod + Fgf4 vs control ≥ 1.5 at stage 14 were uploaded to the PANTHER Classification System. Biological processes identified using PANTHER Fisher's exact statistical overrepresentation test with false discovery rate.

At stage 20, the majority of the highest enriched GO terms were associated with DNA replication and ribosomal RNA maturation and processing (Figure 5.2). Negative regulation of sequestering of calcium ion was another highly enriched GO term at this stage.

By stage 30, highly enriched GO terms in Myod + Fgf4 organoids included striated muscle tissue development and regulation of muscle contraction (Figure 5.3). There was also an increase in highly enriched GO terms related to calcium ions such as: negative regulation of sequestering of calcium ion; release of sequestered calcium ion into cytosol; sequestering of calcium ion; regulation of sequestering of calcium ion; and calcium ion transmembrane import into cytosol. Many of the other top 20 enriched GO terms were related to DNA replication and the cell cycle at this stage.

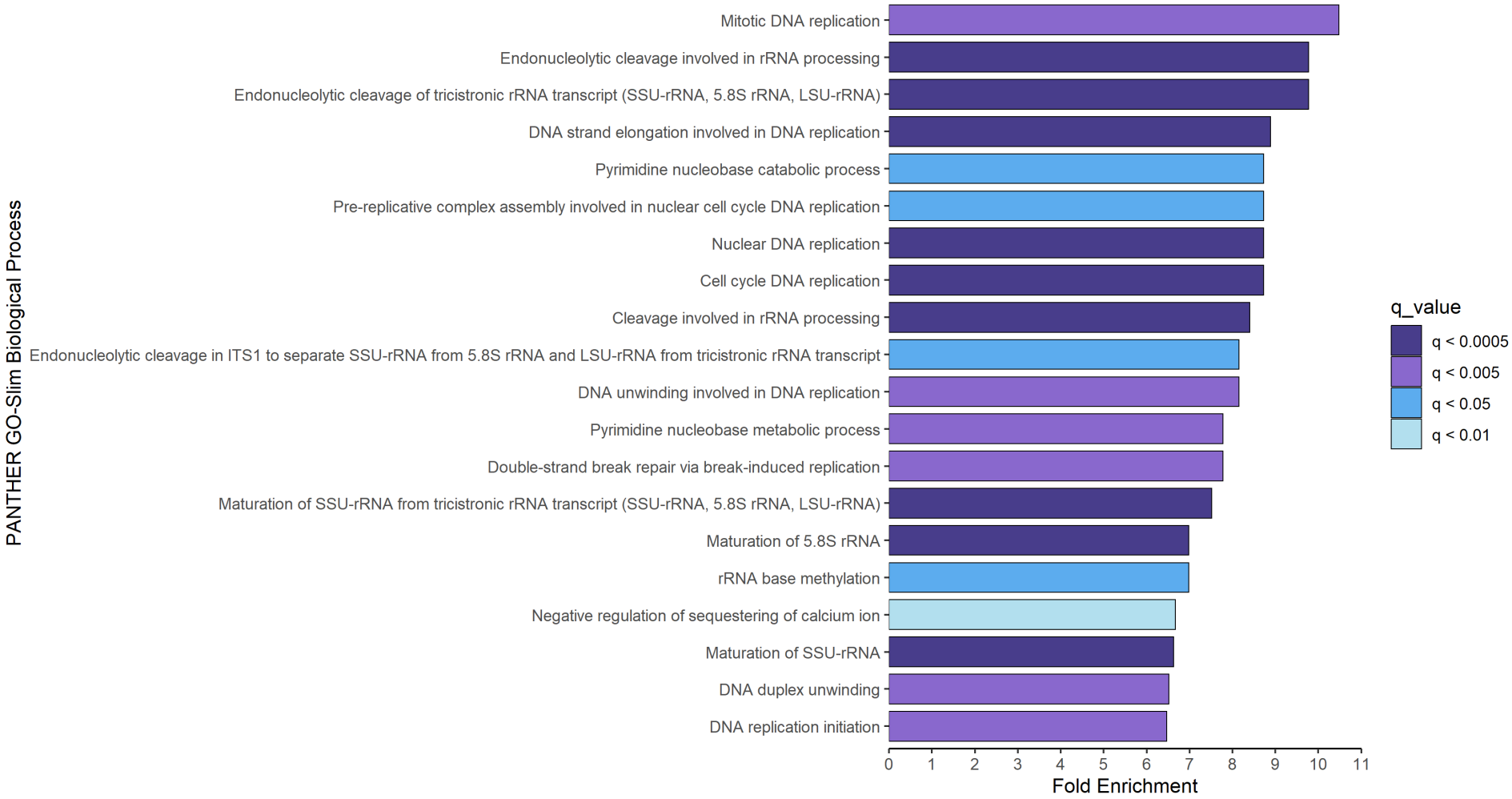


Figure 5.2: Top 20 PANTHER GO-Slim biological processes by fold enrichment over-represented within the upregulated genes in Myod + Fgf4 organoids at stage 20. Genes with mean tpm ≥ 1, q ≤ 0.05, effect size Myod + Fgf4 vs control ≥ 1.5 at stage 20 were uploaded to the PANTHER Classification System. Biological processes identified using PANTHER Fisher's exact statistical overrepresentation test with false discovery rate.

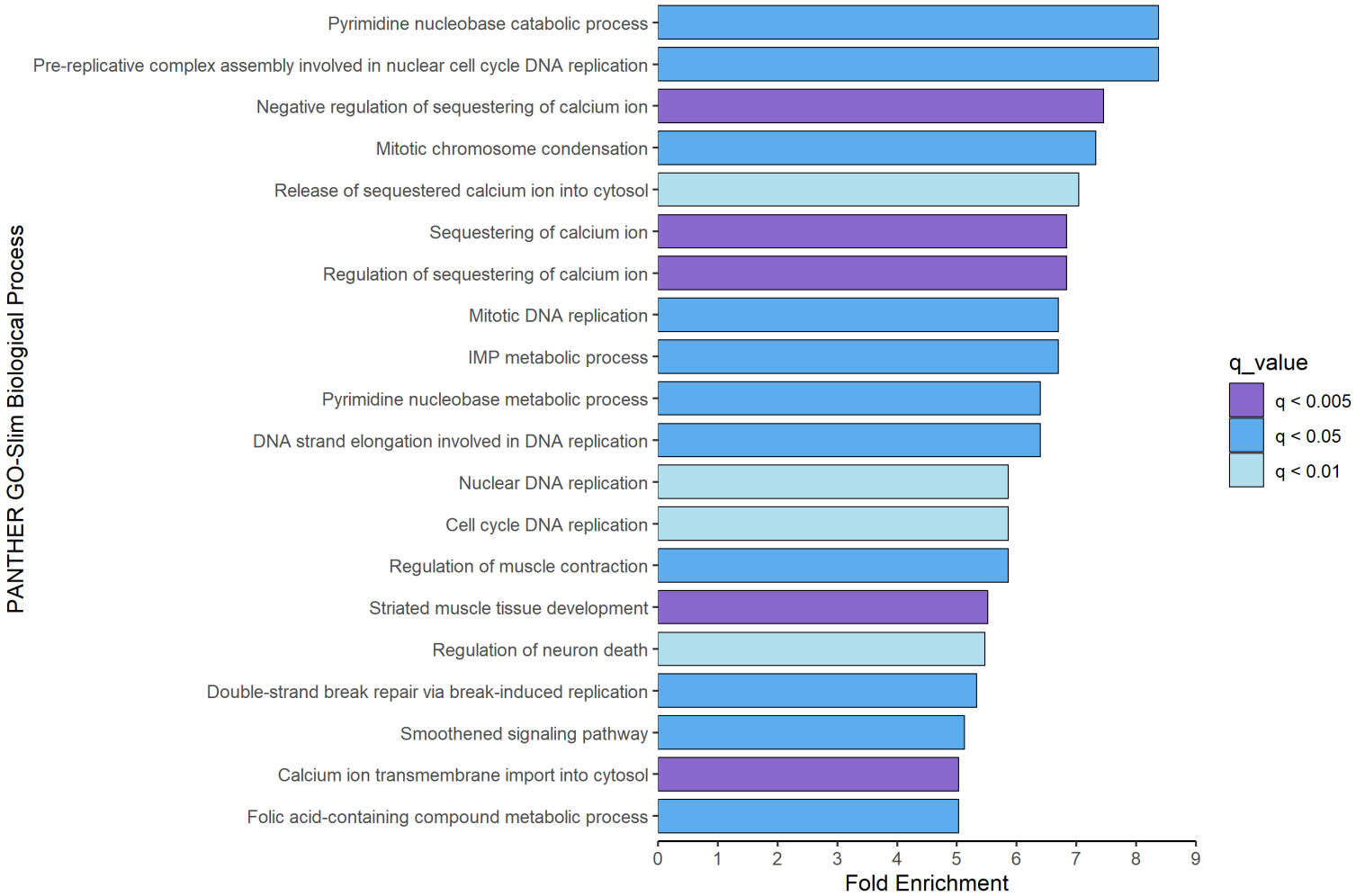


Figure 5.3: Top 20 PANTHER GO-Slim biological processes by fold enrichment over-represented within the upregulated genes in Myod + Fgf4 organoids at stage 30. Genes with mean tpm ≥ 1, q ≤ 0.05, effect size Myod + Fgf4 vs control ≥ 1.5 at stage 30 were uploaded to the PANTHER Classification System. Biological processes identified using PANTHER Fisher's exact statistical overrepresentation test with false discovery rate.

5.2.3 Fgf4 and Wnt signalling are involved in skeletal muscle specification

Having identified over-represented biological processes associated with genes upregulated in Myod + Fgf4 organoids, genes specifically upregulated by Fgf4 were then analysed. To further understand the role of FGF signalling in allowing Myod to promote effective myogenesis, a Venn diagram was created to identify genes upregulated by Fgf4 in Myod + Fgf4 organoids. At stage 14 when the skeletal muscle lineage is specified, 215 genes were upregulated in both Fgf4 and Myod + Fgf4 organoids (Figure 5.4).

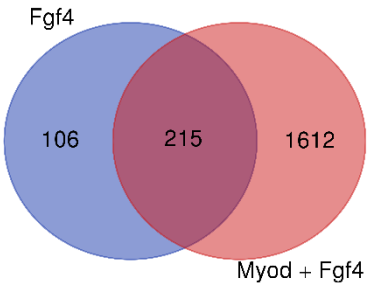


Figure 5.4: Venn diagram of upregulated genes at stage 14 in Fgf4 and Myod + Fgf4 organoids. Genes with mean tpm ≥ 1, q value ≤ 0.05 and effect size vs control ≥ 1.5 for each experimental condition at developmental stage 14.

The 215 genes upregulated in both Fgf4 and Myod + Fgf4 organoids were submitted to the PANTHER Classification System to identify statistically significant over-represented biological processes (Mi et al., 2021). The GO term with the highest fold enrichment was establishment of planar polarity (Figure 5.5). The majority of the most highly enriched GO terms related to non-canonical, canonical or cell-cell Wnt signalling and regulation or negative regulation of the Wnt pathway. The third most enriched GO term was negative regulation of MAPK cascade.

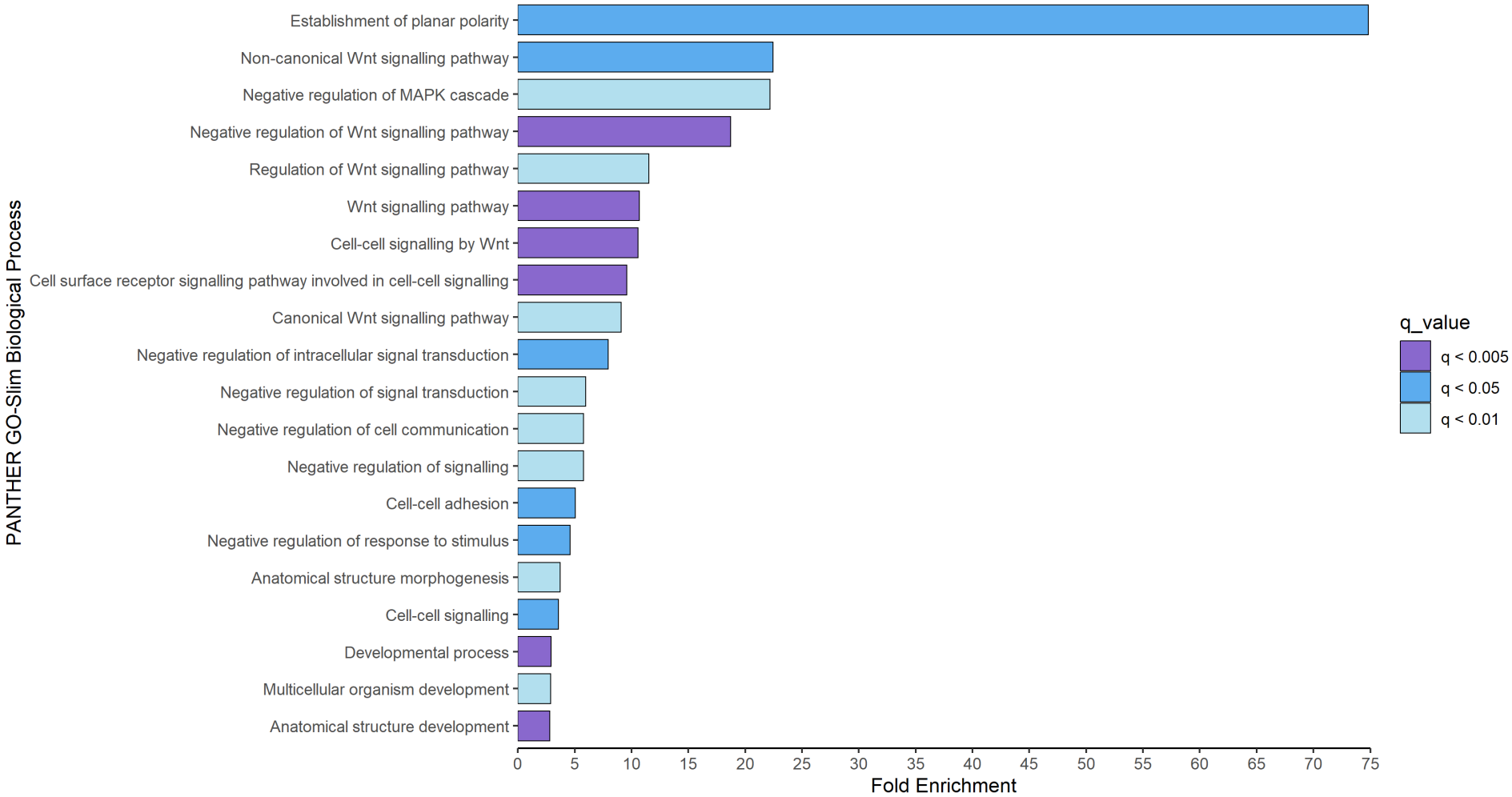


Figure 5.5: Top 20 PANTHER GO-Slim biological processes by fold enrichment over-represented within the genes upregulated in both Fgf4 and Myod + Fgf4 organoids at stage 14. Genes with mean tpm ≥ 1, q ≤ 0.05, effect size Myod + Fgf4 vs control and Fgf4 vs control ≥ 1.5 at stage 14 were uploaded to the PANTHER Classification System. Biological processes identified using PANTHER Fisher's exact statistical overrepresentation test with false discovery rate.

96 genes were upregulated in both Fgf4 and Myod + Fgf4 organoids but not in Myod organoids (Figure 5.6). These genes were submitted to the PANTHER Classification System to identify statistically significant over-represented biological processes and PANTHER pathways. A statistically significant number of genes were associated with the Wnt signalling pathway and the Alzheimer disease-presenilin pathway, such as wnt8a, tcf7, fzd10 and cdh3 (Figure 5.6). Both of these pathways involve regulation of β-catenin, the nuclear effector of canonical Wnt signalling and integral component of cadherin-based adherens junctions (Valenta et al., 2012).

In the absence of canonical Wnt signalling, β-catenin is phosphorylated and tagged for ubiquitination by the "destruction complex" which is made up of glycogen synthase kinase 3 (GSK3), casein kinase 1 (csnk1k1 homolog of CK1), adenomatous polyposis coli (APC), and scaffold protein axin (Amit et al., 2002; Liu et al., 2002; Stamos and Weis, 2013). Binding of Wnt proteins to the transmembrane receptor Frizzled (fzd) and its co-receptors, low density lipoprotein receptor-related protein 5 or 6 (LRP5/6), leads to receptor dimerization, and recruitment of dishevelled (dvl) and the destruction complex (Bhanot et al., 1996; Wehrli et al., 2000; Gao and Chen, 2010). GSK3 activity is then inhibited so that β-catenin is no longer phosphorylated and degraded (Wu et al., 2009). β-catenin therefore accumulates and translocates to the nucleus where it activates Wnt target genes via T-cell factor (TCF) and lymphoid enhancer-binding factor (LEF) transcription factors (Behrens et al., 1996; Molenaar et al., 1996). Secreted Wnt inhibitors such as dickkopf (dkk) interact with and antagonise LRP5/6 activity (Glinka et al., 1998).

Presenilin normally functions as a scaffold mediating β-catenin phosphorylation by GSK3 and CK1 independently of scaffold protein Axin. Therefore mutations in presenilin can lead to increased β-catenin stability and accumulation, as seen in many cases of early-onset familial Alzheimer's disease (Sherrington et al., 1995; Kang et al., 2002). Overexpression of cadherins (cdh) in *Xenopus* and *Drosophila* has been shown to sequester β-catenin at the cell membrane, resulting in insufficient β-catenin availability for nuclear transcriptional activation (Heasman et al., 1994; Sanson et al., 1996).

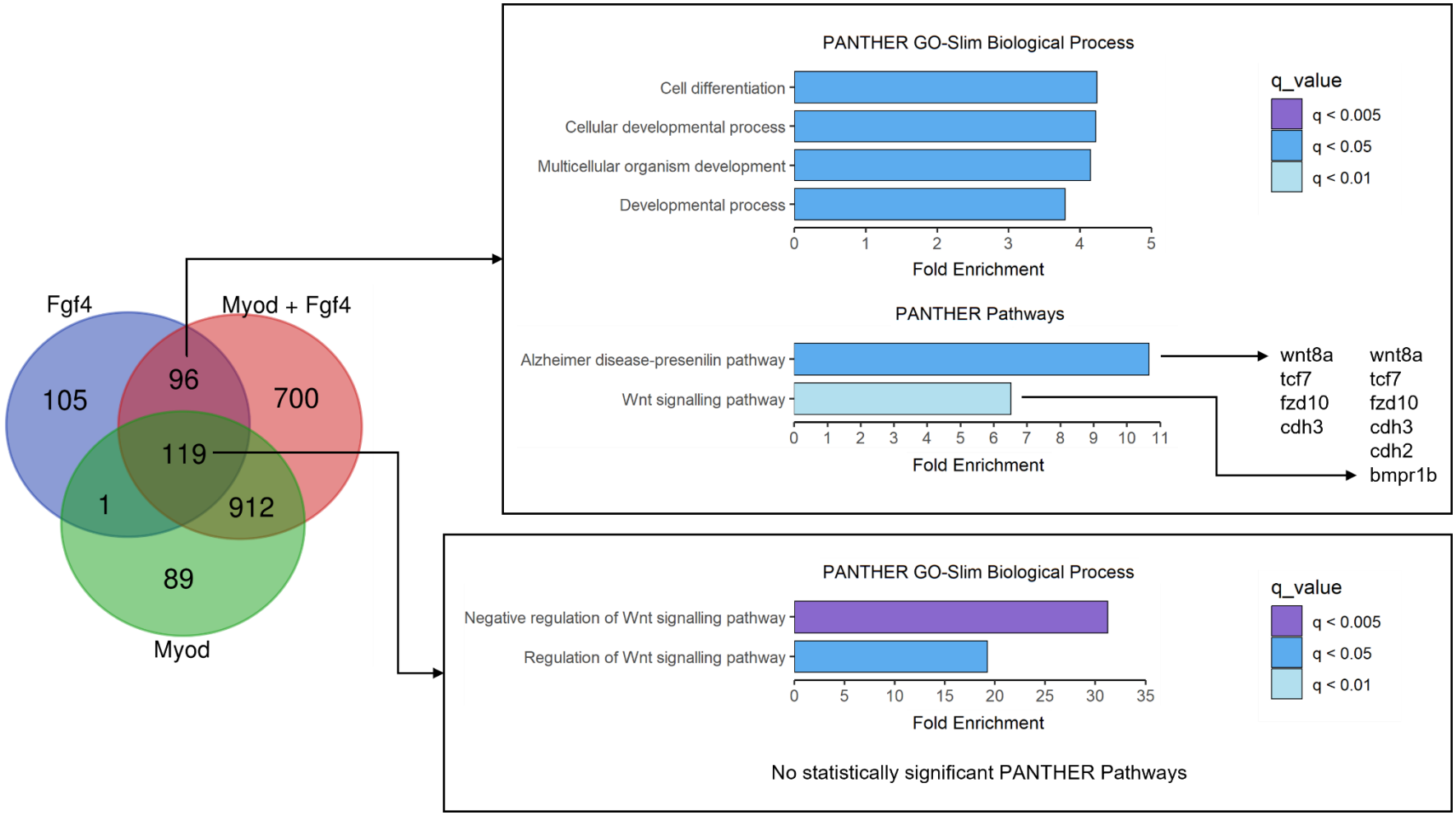
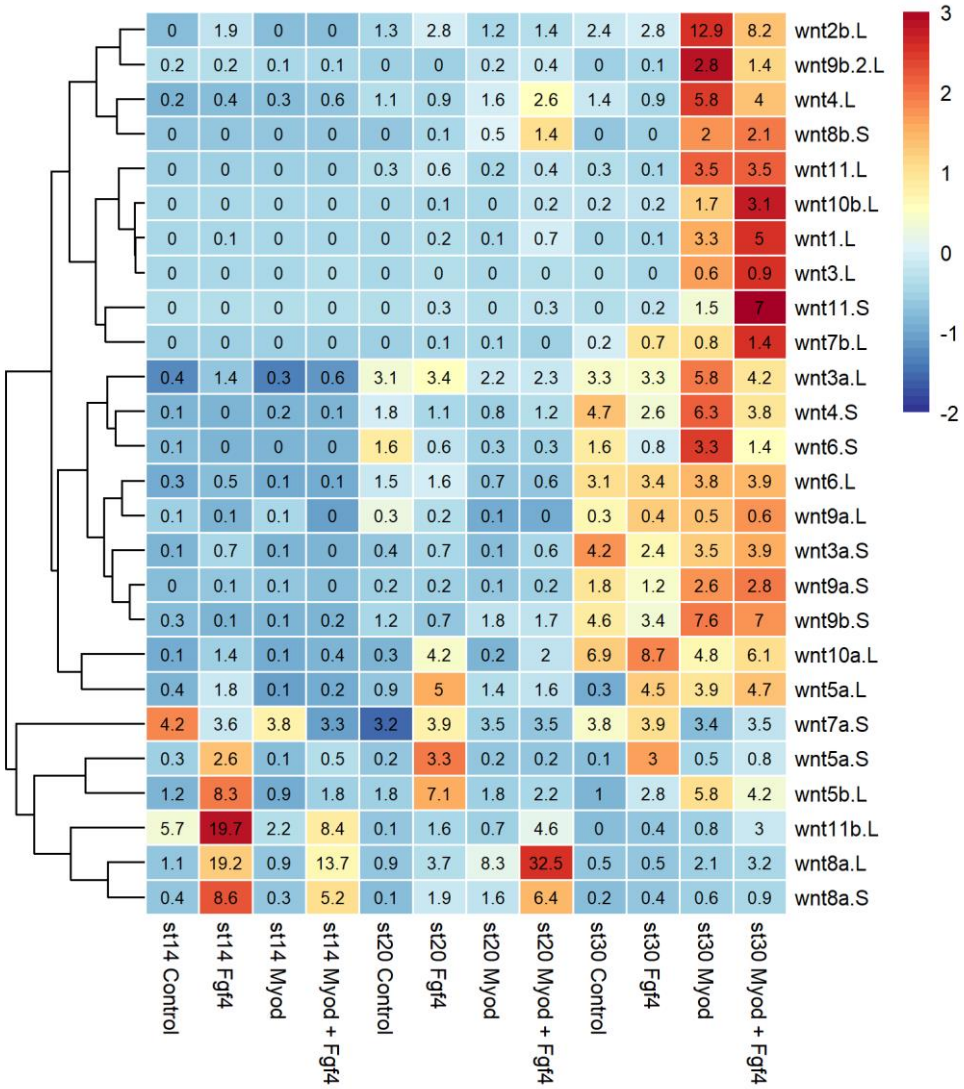


Figure 5.6: Venn diagram of upregulated genes and PANTHER analysis of overlap between Fgf4 and Myod + Fgf4 organoids. Genes with mean tpm ≥ 1, q ≤ 0.05, effect size Fgf4 vs control, Myod + Fgf4 vs control and/or Myod vs control ≥ 1.5 at stage 14 were uploaded to the PANTHER Classification System. Biological processes and pathways identified using PANTHER Fisher's exact statistical overrepresentation test with false discovery rate.

Having identified Wnt signalling as a biological process upregulated by Fgf4 in the skeletal muscle protocol, the expression profiles of wnt genes and downstream signalling components were analysed (Figure 5.7).

The majority of wnts were most highly expressed in stage 30 Myod or Myod + Fgf4 organoids, however the wnt with the highest tpm (32.5) was *wnt8a.L* in stage 20 Myod + Fgf4 organoids (Figure 5.7A). Expression of *wnt11b.L*, *wnt8a.L* and *wnt8a.S* was upregulated in stage 14 Fgf4 and Myod + Fgf4 organoids, and *wnt5a.S* and *wnt5b.L* in Fgf4 organoids. Wnt signalling pathway components were typically most highly expressed in Myod or Myod + Fgf4 organoids at stage 14 or 30 (Figure 5.7B). However, the highest expression of *axin2.S, axin2.L* and *fzd10.L* was in stage 14 Fgf organoids, and the highest expression of *fzd8.L* was in stage 20 Myod organoids.

A



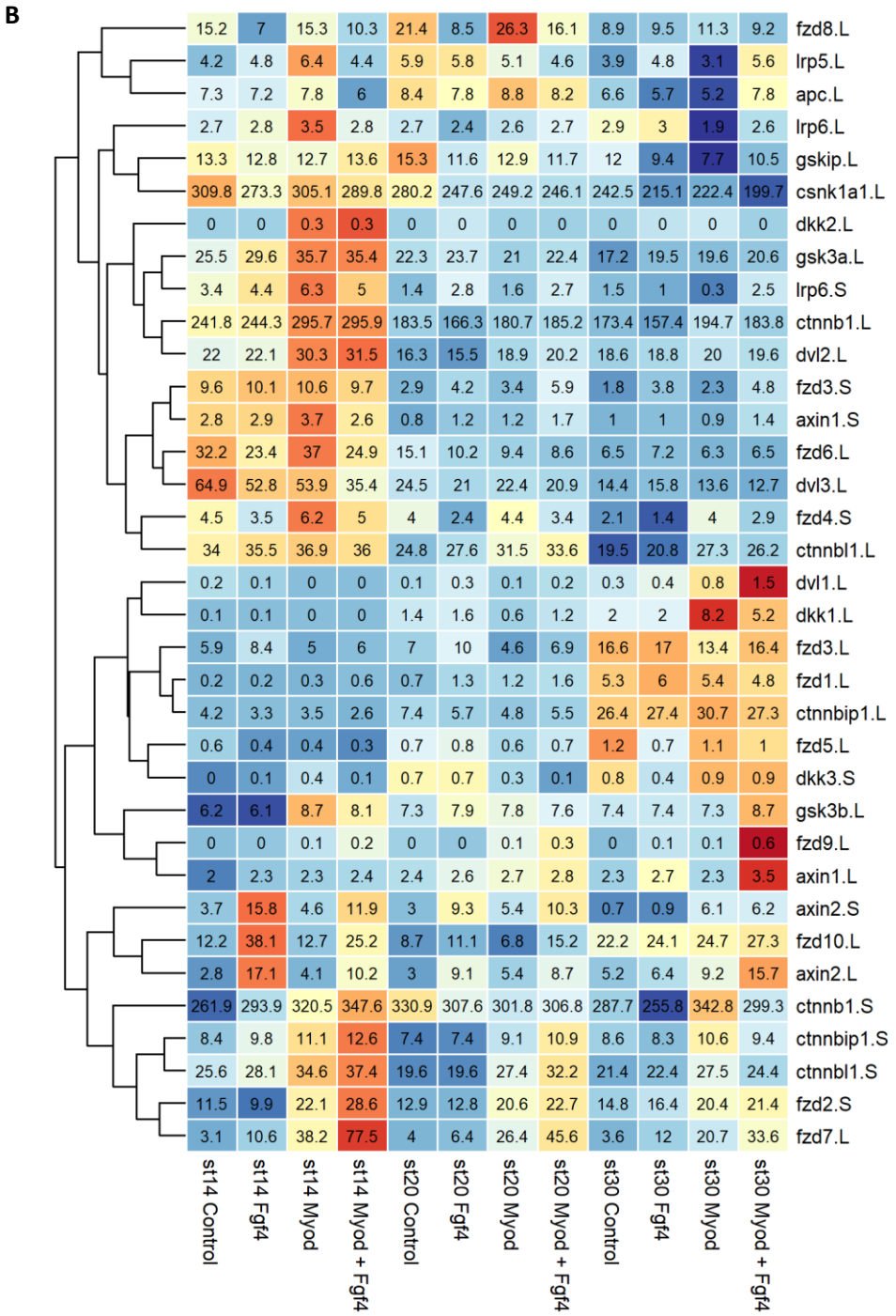


Figure 5.7: Heatmap showing Wnt and canonical Wnt signalling pathway component expression at stages 14, 20 and 30 of the skeletal muscle protocol. Animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1.2ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured with or without 50ng/ml Fgf4. RNA extracted from organoids collected at stages 14, 20 and 30 for RNA-seq. Mean tpm values to 1 decimal place. Heatmap hierarchically clustered and colours scaled by row. **A** Wnt genes, **B** canonical Wnt signalling pathway components.

5.2.4 Co-expression of Myod and Wnt8 is not sufficient to induce skeletal muscle

As Wnt8a expression was upregulated in Fgf4 organoids at stage 14, and in Myod + Fgf4 organoids at stage 14 and 20, the ability of Wnt8 to replace Fgf4 in the skeletal muscle protocol was then tested. 100pg cska Wnt8 was injected into *X. laevis* embryos at the 2-cell stage, with or without 1.2ng *myod1.S* mRNA, and animal cap cells dissected and cultured from stage 9. The presence of skeletal muscle was determined by western blot against MHC using the MF20 antibody. No skeletal muscle was induced by 100pg cska Wnt8 with or without *myod,* as MHC was only present in Myod + Fgf4 organoids (Figure 5.8).

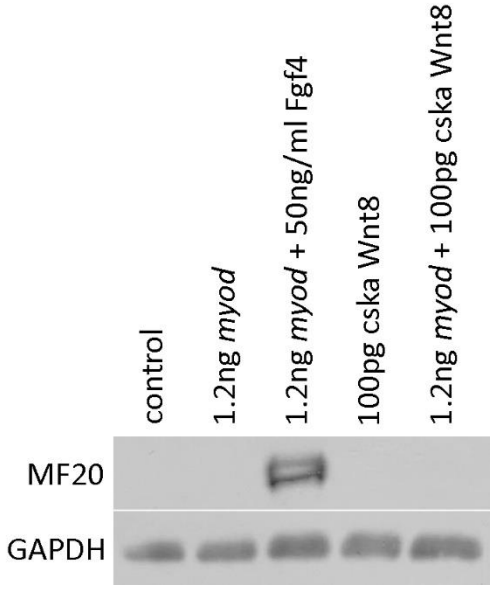


Figure 5.8: Western blot detecting sarcomere myosin heavy chain (MF20) in 1.2ng *myod* + 100pg cska Wnt8 organoids. Animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1.2ng *myod1.S* mRNA, 100pg cska Wnt8 or both, bilaterally at the 2-cell stage and cultured with or without 50ng/ml Fgf4. 10 organoids collected per condition at stage 37. n = 1.

5.2.5 Hedgehog signalling pathway expression profiles

The most enriched GO term in stage 14 Myod + Fgf4 organoids was the smoothened signalling pathway. Smoothened (smo) is a G-protein coupled receptor-like protein and the signal transducer of the hedgehog (Hh) pathway. Therefore, the expression profiles of hedgehog genes and key signalling pathway components were analysed (Figure 5.9).

There are three hedgehog family members in vertebrates: sonic hedgehog (*shh*), Indian hedgehog (*ihh*), and desert hedgehog (*dhh*). Shh is widely expressed and has many important roles in development such as neural cell type specification, limb patterning, and regulation of slow muscle fibre formation (Belgacem et al., 2016; Anderson et al., 2012; Grimaldi et al., 2004). Ihh is associated with skeletal development, proliferation and differentiation of chondrocytes (St-Jacques et al., 1999), and dhh signalling specifies cell fates in the gonads (Yao et al., 2002; Wijgerde et al., 2005). In the canonical hedgehog pathway, Hh binds and inactivates the 12-transmembrane protein

patched (ptch) (Marigo et al., 1996). This abolishes ptch inhibition of smo, as ptch is internalised and degraded (Denef et al., 2000). Smo accumulates in the primary cilium of the cell and activates the Hh downstream signalling cascade, in which Gli proteins translocate to the nucleus and initiate target gene transcription (Carballo et al., 2018). Gli1 is a transcriptional activator, and Gli2 also tends to activate transcription, however Gli2 and Gli3 can act as positive or negative regulators depending on post-transcriptional and post-translational processing (Sasaki et al., 1999). In the absence of Hh signalling, suppressor of fused (sufu) directly binds Gli proteins to inhibit their translocation to the nucleus (Kogerman et al., 1999). The highest expression of *smo.L* and *gli1/2/3* was in Myod + Fgf4 organoids, followed by Myod organoids. However Hh ligands were not expressed highly during the protocol.

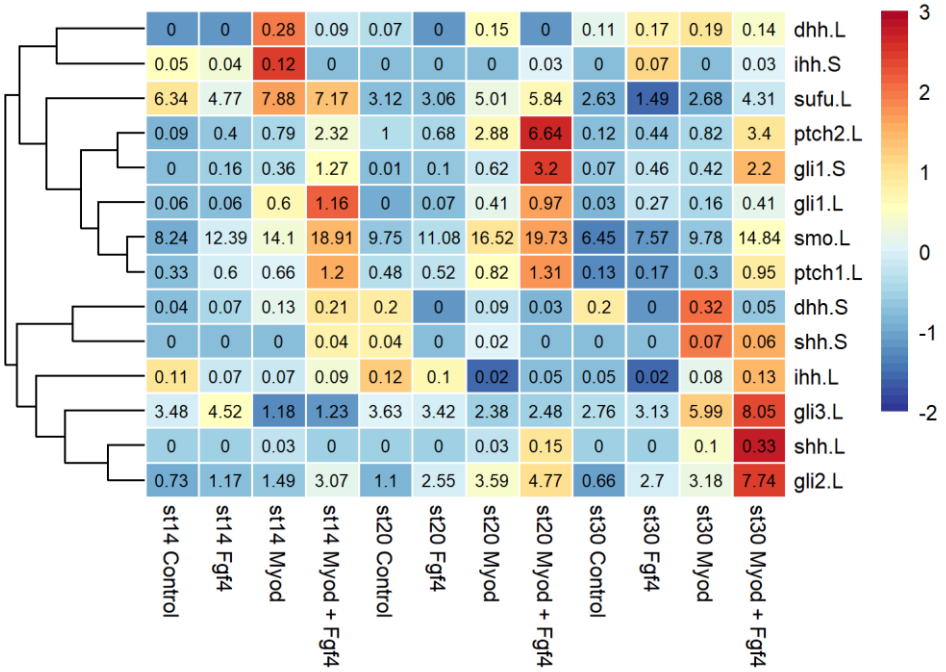


Figure 5.9: Heatmap showing hedgehog and canonical hedgehog signalling pathway component expression at stages 14, 20 and 30 of the skeletal muscle protocol. Animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1.2ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured with or without 50ng/ml Fgf4. RNA extracted from organoids collected at stages 14, 20 and 30 for RNA-seq. Mean tpm values to 2 decimal places. Heatmap hierarchically clustered and colours scaled by row.

5.2.6 Notch – Delta expression profiles

Another highly enriched GO term at stage 14 was the notch signalling pathway. When the cell surface receptor notch interacts with its transmembrane ligand delta (*dll*), its intracellular domain is cleaved and translocated to the nucleus to regulate transcription of target genes (Kopan, 2012). The notch pathway is implicated in regulating cell proliferation, cell fate determination and differentiation in almost every tissue type studied to date (VanDussen et al., 2012; Gioftsidi et al., 2022). For instance, Notch interacts with Wnt and FGF to regulate somitogenesis, and Shh to specify *Xenopus* dorsal midline cell fates (Jen et al., 1999; Wahi et al., 2016; López et al., 2003). Therefore, the expression profiles of notch and delta during the protocol were analysed (Figure 5.10) The highest expression of *notch1.L, notch3.L* and *dll1.L* was in stage 14 Myod + Fgf4 organoids.

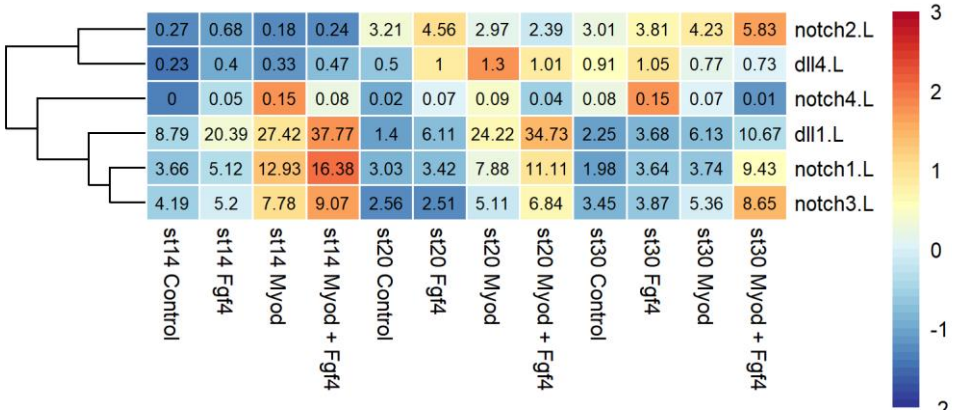


Figure 5.10: Heatmap showing notch and delta expression at stages 14, 20 and 30 of the skeletal muscle protocol. Animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1.2ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured with or without 50ng/ml Fgf4. RNA extracted from organoids collected at stages 14, 20 and 30 for RNA-seq. Mean tpm values to 2 decimal places. Heatmap hierarchically clustered and colours scaled by row.

5.2.7 Fibroblast growth factor expression profiles

In order to determine whether other FGFs were expressed during the protocol, FGF ligands and receptor expression profiles were analysed (Figure 5.11). Expression varied but was typically highest in Fgf4 or Myod + Fgf4 organoids. Expression of *fgfr4.L* was particularly high in stage 14 Myod + Fgf4 organoids, whereas expression of genes such as *fgf13.L, fgf2.L, fgfr3.L* and *fgfr1.S* was highest in stage 30 Myod + Fgf4 organoids.

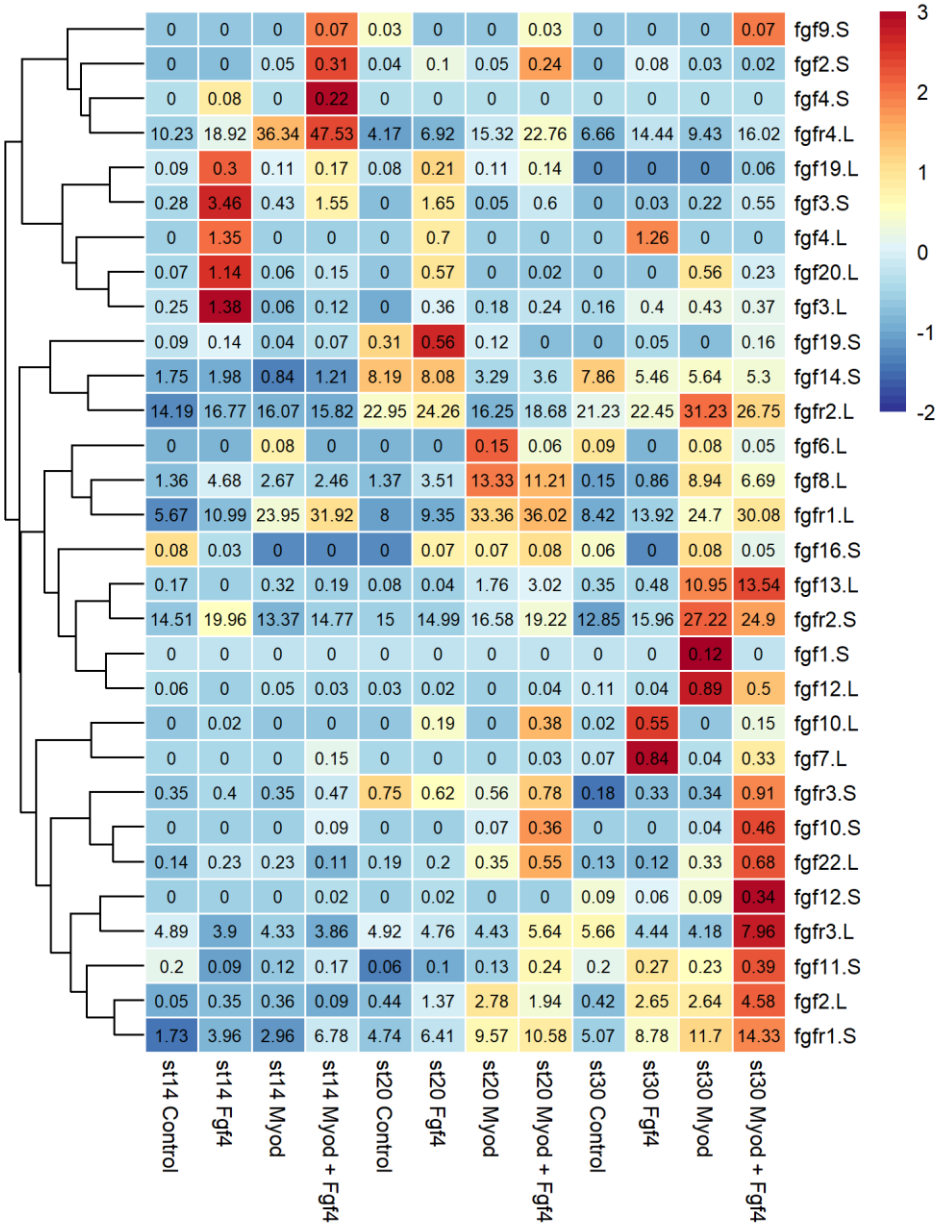


Figure 5.11: Heatmap showing FGF and FGFR expression at stages 14, 20 and 30 of the skeletal muscle protocol. Animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1.2ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured with or without 50ng/ml Fgf4. RNA extracted from organoids collected at stages 14, 20 and 30 for RNA-seq. Mean tpm values to 2 decimal place. Heatmap hierarchically clustered and colours scaled by row.

5.2.8 TGFβ superfamily expression profiles

Activin and TGFβ-like factors are capable of inducing mesoderm and mesodermal derivatives such as muscle in animal caps in a dose-dependent manner (Asashima et al., 1990; Green et al., 1990). Therefore the expression profiles of these genes were analysed to determine whether the signalling pathways were activated during the protocol (Figure 5.12).

The transforming growth factor beta (TGFβ) superfamily is made up of over 30 structurally related growth factors including TGFβ, nodal, bone morphogenetic proteins (BMP), activins (alpha and beta subunits: *inhba* and *inhbb*), and their downstream effectors known as smads (Wrighton et al., 2009; Affolter and Basler, 2007; Smith and Green, 1990; Weiss and Attisano, 2013). Smads belong to one of 3 categories: receptor-regulated smad1/2/3/5/8/9 (R-smads), common partner smad4 (Co-smad), or inhibitory smad6/7 (I-smads) (Masuyama et al., 1999; Kumar et al., 2001; Goto et al., 2007; Hanyu et al., 2001). R-smads can be further divided into those which mediate TGFβ, activin and Nodal signalling (smad2/3), and those which mediate BMP signalling (smad1/2/5/8/9) (Derynck and Zhang, 2003). These signalling pathways have been shown to play a key role in many developmental processes including regulation of cell proliferation and differentiation (Gordon and Blobe, 2008).

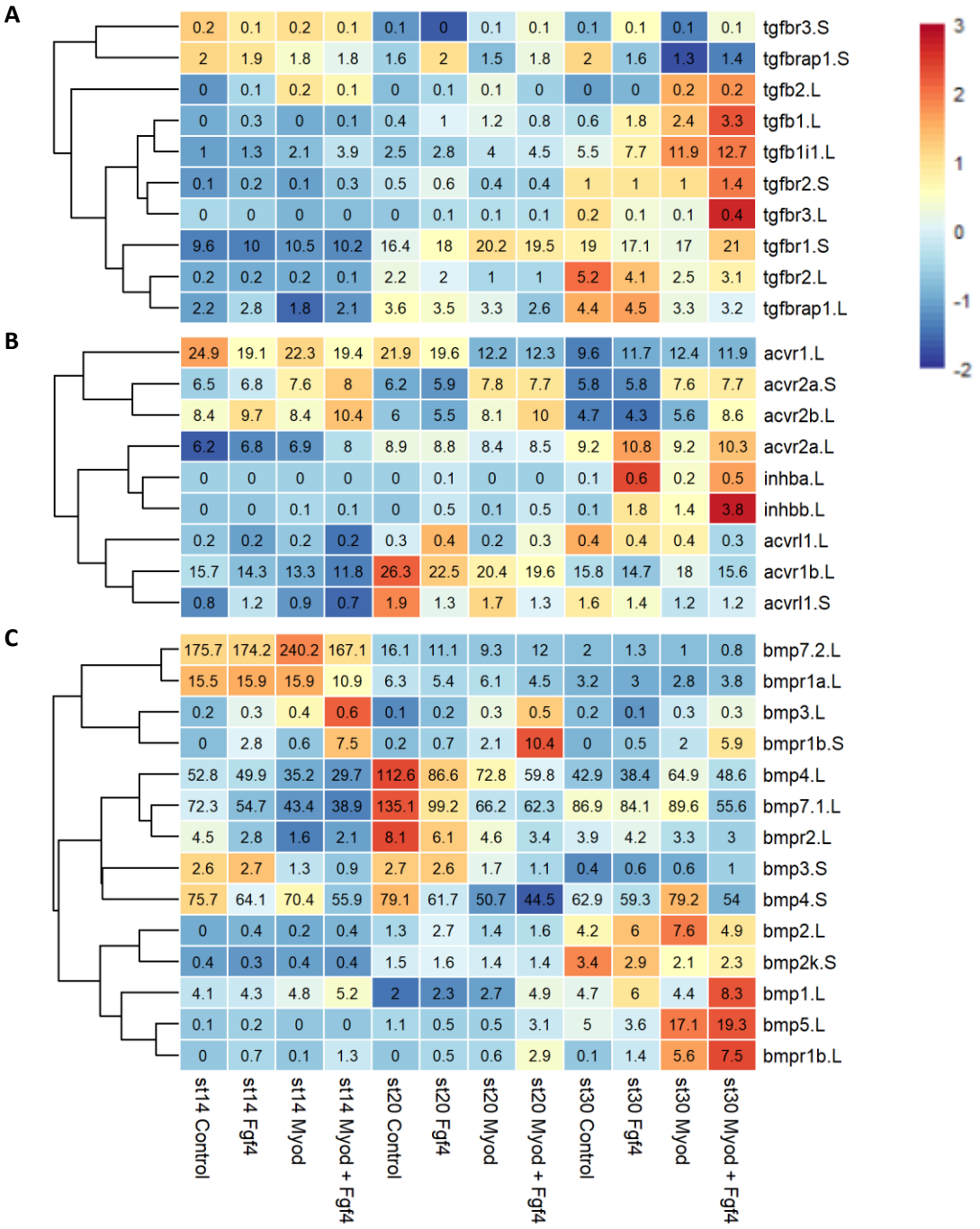
The expression of the majority of TGFβ genes did not vary greatly between conditions during lineage specification at stage 14 (Figure 5.12A). However, *tgfb1.L* and TGFβ1-induced transcript 1 (*tgfb1i1.L*) expression increased during development with the highest expression in stage 30 Myod + Fgf4 organoids. TGFβ receptors *tgfbr1.S, tgfbr2.S* and *tgfb3.L* also showed their highest expression in stage 30 Myod + Fgf4 organoids, whereas *tgfbr2.L* and TGFβ receptor associated protein 1 (*tgfbrap1.L*) were most highly expressed in stage 30 control organoids.

Expression of activin subunit βb *inhbb.L* was highest in stage 30 Myod + Fgf4 organoids, whereas subunit βa inhba.L showed slightly greater expression in stage 30 Fgf4 organoids than Myod + Fgf4 (Figure 5.12B). Activin receptor expression (acvr) varied by type with *acvr1.L, acvr1b.L* and *acvrl1.S* expressed most highly in control organoids, while *acvr2b.L* showed the highest expression in Myod + Fgf4 organoids. acvr2a.L was highly expressed in Myod and Myod + Fgf4 organoids.

BMP expression varied with some family members expressed most highly in stage 20 control organoids (*bmp4.L, bmp7.1.L and bmpr2.L*), whereas others showed their highest expression in stage 30 Myod + Fgf4 organoids (*bmp1.L, bmp5.L* and *bmpr1b.L*) (Figure 5.12C). *bmpr1b.S* was most highly expressed in stage 20 Myod + Fgf4 organoids.

Nodal expression was very low across all conditions and stages analysed, with the highest tpm of 0.29 in stage 20 Myod + Fgf4 organoids for *nodal1.S* (Figure 5.12D).

Smad expression also varied with BMP signalling mediator *smad1.S* and *smad4.1.L* expressed highly in Myod organoids at stage 14, and *smad9.L* and *smad1.L* at stage 30 (Figure 5.12E). *smad4.2.S*. and *smad9.S* showed their highest expression in control organoids at stage 20, and *smad4.2.L* and *smad4.1.*S at stage 30. The greatest expression of inhibitory *smad7.L* was in stage 20 control organoids, while inhibitory *smad6.L* was expressed highly at stage 30 in control, Fgf4, and Myod organoids.



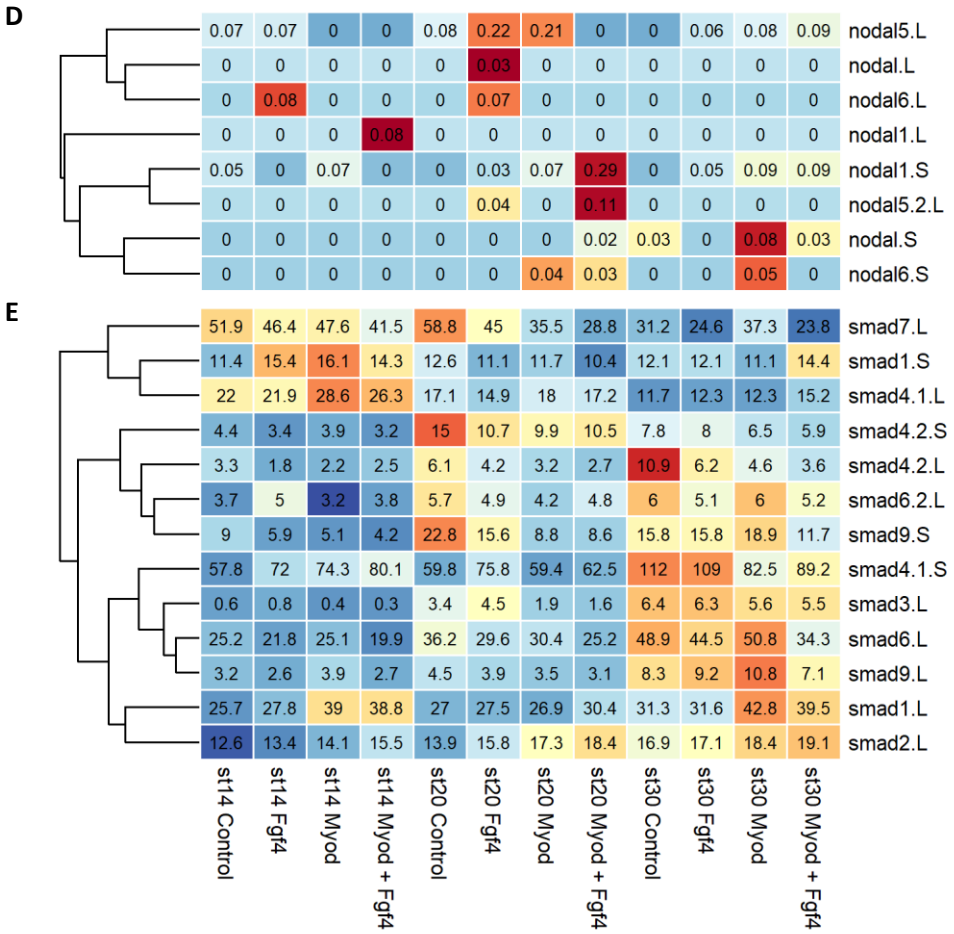


Figure 5.12: Heatmap showing TGFβ superfamily expression at stages 14, 20 and 30 of the skeletal muscle protocol. Animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1.2ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured with or without 50ng/ml Fgf4. RNA extracted from organoids collected at stages 14, 20 and 30 for RNA-seq. Mean tpm values to 1 or 2 decimal places. Heatmap hierarchically clustered and colours scaled by row. **A** transforming growth factor beta genes, **B** activins, **C** bone morphogenetic proteins, **D** nodals, **E** smads.

5.2.9 Candidate genes potentially involved in skeletal muscle lineage specification

In order to identify candidate genes with a potential role in skeletal muscle lineage specification, the dataset was sorted to compile a shortlist of genes fulfilling the following criteria: mean tpm ≥ 1, mean tpm highest in Myod + Fgf4 organoids at stage 14, fold change for Myod + Fgf4 versus control ≥ 1.5, fold change for Myod + Fgf4 versus control highest at stage 14, q values < 0.05. Implementing these criteria reduced the list of genes from 50,487 to 108. The function of each gene from this shortlist was determined from existing literature. Five shortlisted genes were identified as potential candidates: *smyd1.L, tcf15.L*, *tcf12.S, mex3b.S* and *rbfox2.S* (Figure 5.13).

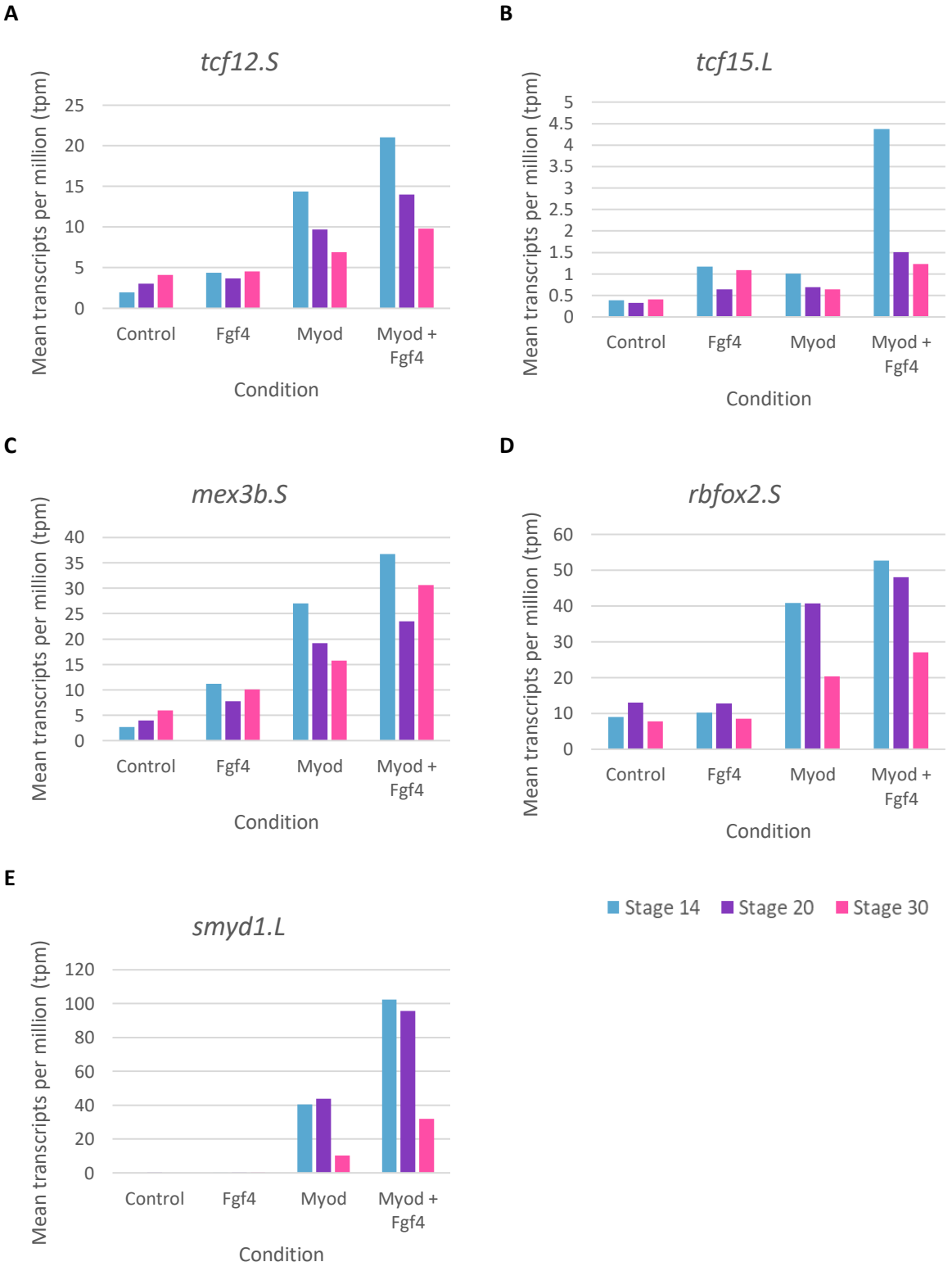


Figure 5.13: Candidate genes and the mean transcripts per million (tpm) for each condition at stages 14, 20 and 30.

SET and MYND domain containing 1 (*smyd1.L)* encodes a histone methyltransferase associated with myogenesis and is specifically expressed in striated muscle (Nagandla et al., 2016). bHLH domain containing transcription factor 15 (*tcf15.L*) is required for paraxial mesoderm development and somitogenesis, and has been shown to prime pluripotent cells for differentiation (Burgess et al., 1996; Davies et al., 2013). Both Tcf15 (paraxis) and Tcf12 (HEB) can act through formation of a heterodimer with another bHLH protein that binds DNA on E-box motifs to activate transcription of target genes. bHLH E-protein Tcf12 has been shown to form such heterodimers with Myod1, Myog, Neurod1 and Twist1 to regulate developmental fates (Parker et al., 2006; Singh et al., 2022; Fan et al., 2023; Hu et al., 1992). Mex-3 RNA binding family member B (*mex3b.S*) encodes an RNA-binding protein involved in post-transcriptional regulation, and has been shown to regulate FGF signalling during patterning of the neural plate (Takada et al., 2009). RNA binding fox-1 homolog 2 (*rbfox2.S*) encodes an RNA-binding protein that regulates alternative splicing events. It is required for myoblast fusion and regulates 30% of splicing transitions associated with muscle differentiation (Singh et al., 2014). Out of the candidate genes selected, *smyd1.L* had the highest fold change versus control and *tcf12.S* the most significant q value (Figure 5.14). As skeletal muscle is induced in Myod + Fgf4 organoids and not in Myod organoids, candidate genes are likely to be downstream targets of Fgf4.

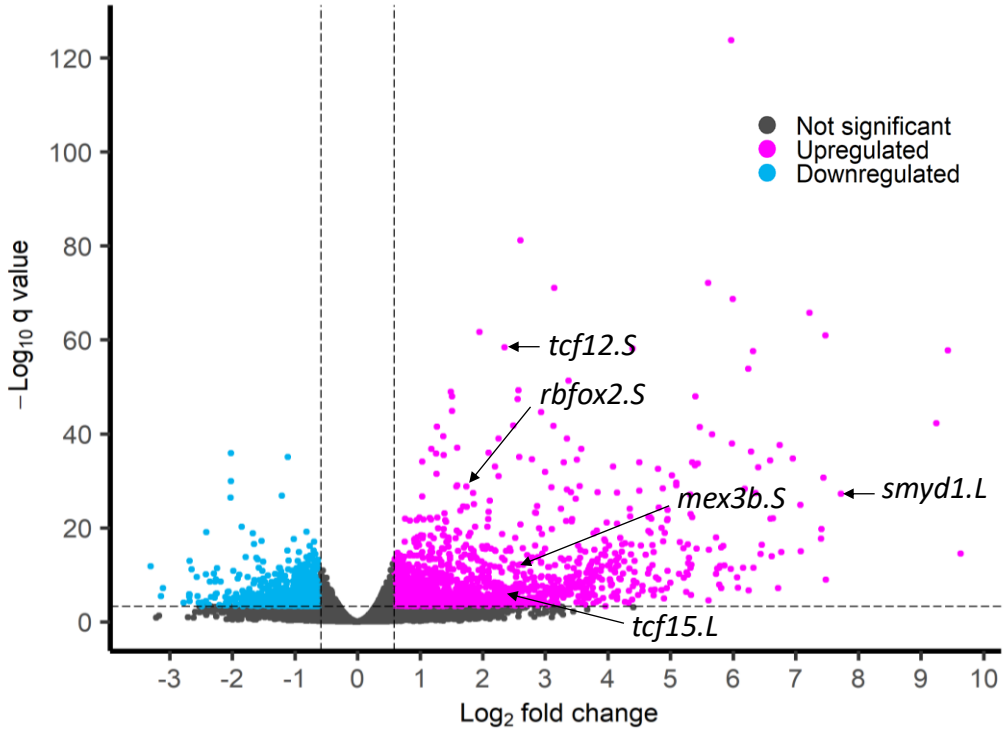


Figure 5.14: Volcano plot showing differential gene expression between myod + Fgf4 organoids and control organoids at stage 14. Log_2 fold change calculated via wald test. Upregulated genes: q value < 0.0005, fold change > 1.5. Downregulated genes: q value < 0.0005, fold change < -1.5.

Known FGF targets *lin28A1, ventx2* and *sp5* were also identified as genes with a potential role in lineage specification (Figure 5.15). Lin28 proteins are required for germ layer specification in *Xenopus*, and have been shown to facilitate the transition of mouse cells from naïve to primed pluripotency (Faas et al., 2013; Zhang et al., 2016; Tsanov et al., 2017; Branney et al., 2009). Sp5 has also been shown to interact with brachyury (Tbxt) downstream of FGF signalling, as well as acting downstream of Wnt signalling during mesoderm and neuroectoderm specification (Harrison et al., 2000; Weidinger et al., 2005; Ossipova et al., 2002; Elsy et al., 2019). Expression of Ventx1 has been shown to be cooperatively activated by FGF/Tbxt and BMP4/Smad1 signalling, and Ventx2 has been shown to regulate dorsoventral patterning of *Xenopus* mesoderm via an autocatalytic loop with BMP4 (Kumar et al., 2018; Onichtchouk et al., 1996; Schuler-Metz et al., 2000).

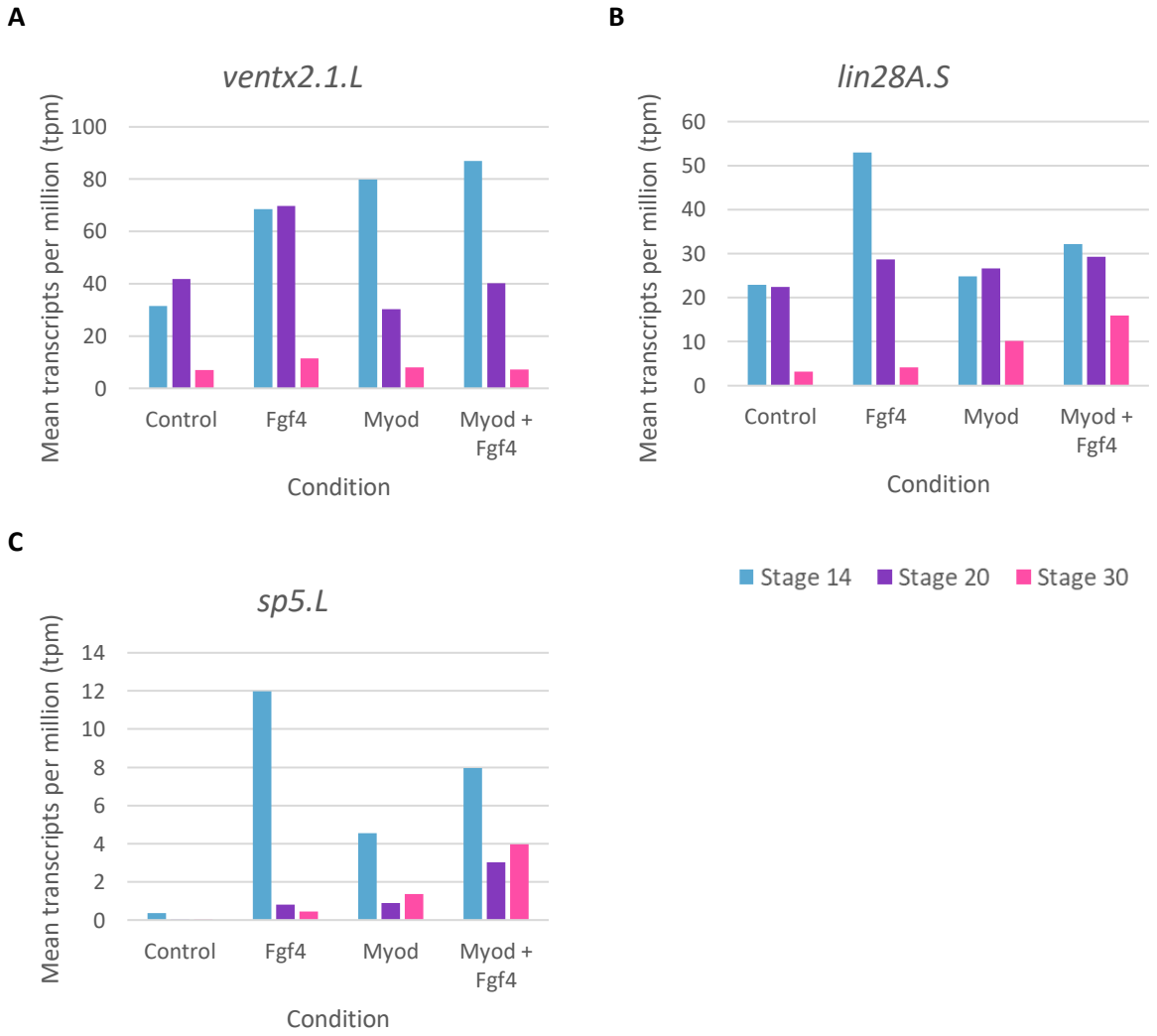


Figure 5.15: Genes with a potential role in lineage specification downstream of FGF and the mean transcripts per million (tpm) for each condition at stages 14, 20 and 30.

5.2.10 Co-expression of Tcf12 and Myod induces skeletal muscle myosin in organoids

Having identified potential downstream FGF target genes able to induce muscle, the ability of these single factors to replace Fgf4 in the protocol was tested. mRNA for each candidate gene was synthesised and 1ng injected into *X. laevis* embryos at the 2-cell stage, with or without 1.2ng *myod1.S* mRNA, and animal cap cells dissected and cultured from stage 9. The presence of skeletal muscle was determined by western blot against MHC using the MF20 antibody (Figure 5.12). MHC was present in 1.2ng *myod* + 50ng/ml Fgf4 organoids and 1ng *tcf12* + 1.2ng *myod* organoids. However MHC was not present in organoids expressing Myod in combination with Tcf15, Rbfox2, Mex3b, Smyd1, Sp5, Lin28A1 or Ventx2 (Figure 5.16).

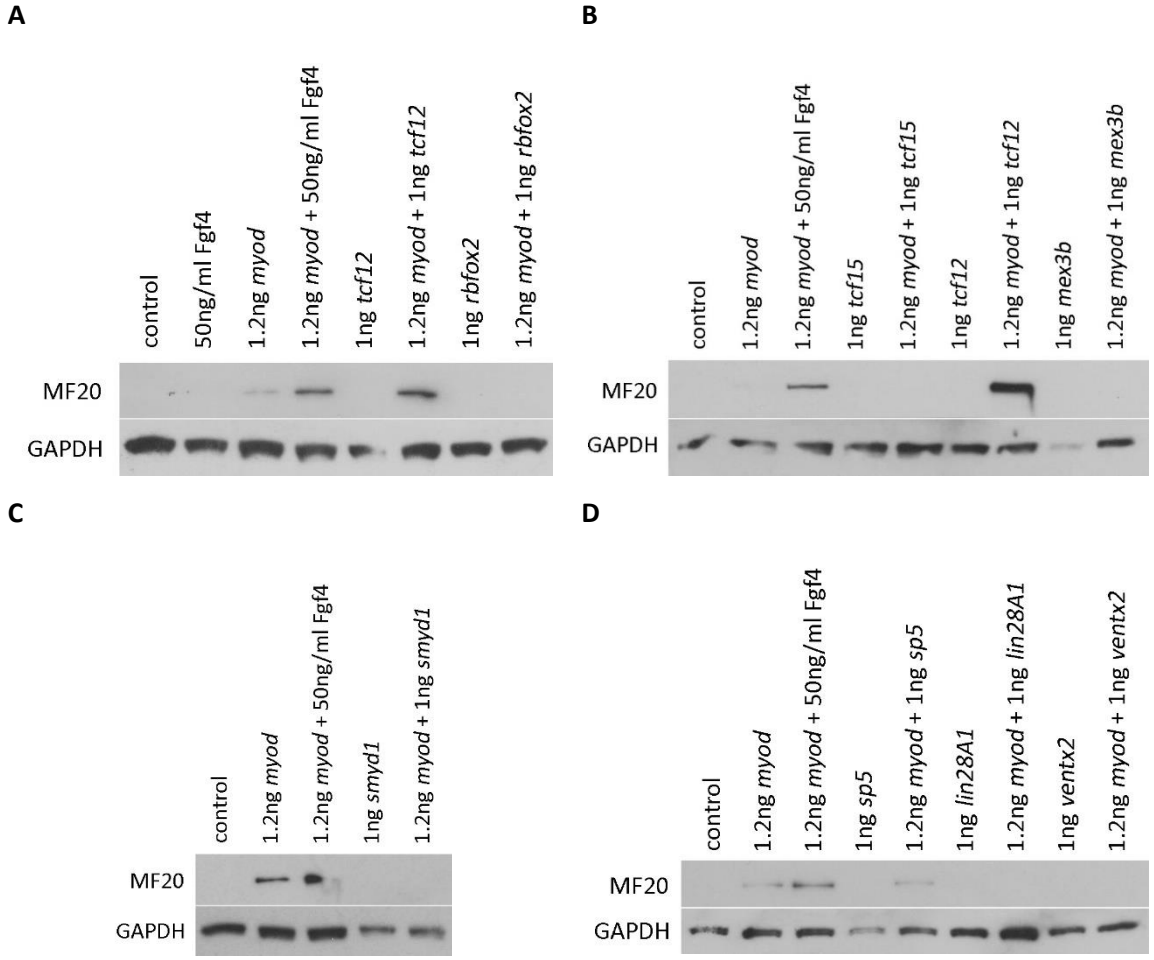


Figure 5.16: Western blots detecting sarcomere myosin heavy chain (MF20) in skeletal muscle assay organoids. Animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1ng of candidate gene mRNA, with or without 1.2ng *myod1.S* mRNA, bilaterally at the 2-cell stage and cultured until stage 37.

5.2.11 Co-expression of Tcf12 and Myod induces skeletal muscle in organoids

In order to confirm the presence of skeletal muscle in Tcf12 + Myod organoids, whole-mount immunostaining was carried out for skeletal muscle marker 12-101. Organoids were then sectioned to determine whether blue staining indicating sarcoplasmic reticulum membrane protein was present. No 12-101 staining was present within control or 1ng *tcf12* organoids, however staining of skeletal muscle cells was observed in 1ng *tcf12* + 1.2ng *myod* organoids (Figure 5.17).

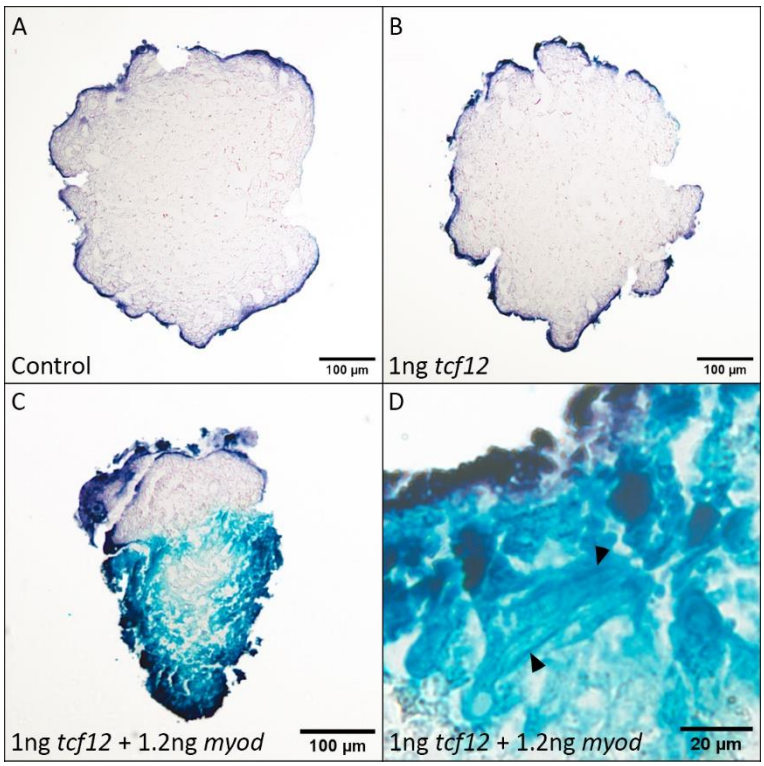


Figure 5.17: Representative sections of 12-101 immunostained organoids. 10µm sections of animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1ng *tcf12* mRNA, with or without 1.2ng *myod1.S* mRNA, bilaterally at the 2-cell stage and cultured until stage 41. Blue staining = 12-101 against sarcoplasmic reticulum membrane protein. 1ng *tcf12* + 1.2ng *myod* organoids contain skeletal muscle. **A**, **B** and **C** scale bar = 100µm, **D** scale bar = 20µm.

5.2.12 Tcf12 is a downstream target of FGF signalling

One of the main downstream effectors of FGF signalling in this system is dpERK. Fgf4 protein in the skeletal muscle protocol phosphorylates ERK, converting it to its active form dpERK, in Fgf4 and Myod + Fgf4 organoids (Figure 5.18).

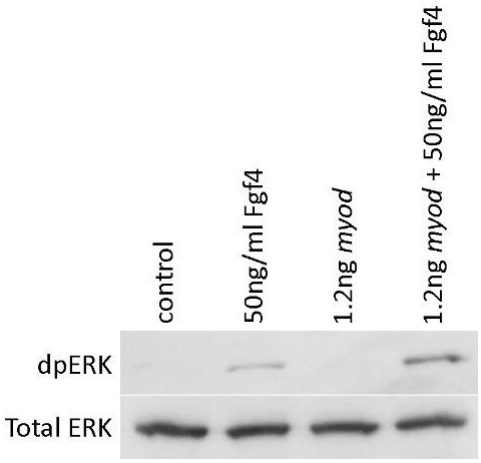


Figure 5.18: Western blot detecting diphosphorylated ERK (dpERK) in skeletal muscle protocol organoids. Animal pole cells dissected from un-injected *X. laevis* embryos and embryos injected with 1.2ng *myod1.S* mRNA bilaterally at the 2-cell stage and cultured with or without 50ng/ml Fgf4. 10 organoids collected per condition at early gastrula stage 10.5. n = 1.

During early gastrulation, FGF-ERK signalling occurs in a ring around the blastopore to induce mesoderm in an embryo (Figure 5.19A). In order to determine whether FGF signalling is required for *tcf12* expression, *X. laevis* embryos were unilaterally injected with a dominant negative truncated FGF receptor (XFD) and in situ hybridisation carried out for *tcf12* by Jennika Bates. In wild-type embryos, *tcf12* is expressed in the same region as endogenous FGF signalling at gastrula stage 10.5 (Figure 5.19B). Inhibition of FGF signalling by XFD results in complete loss of the normal expression of *tcf12,* indicating that *tcf12* expression requires FGF (Figure 5.19C).

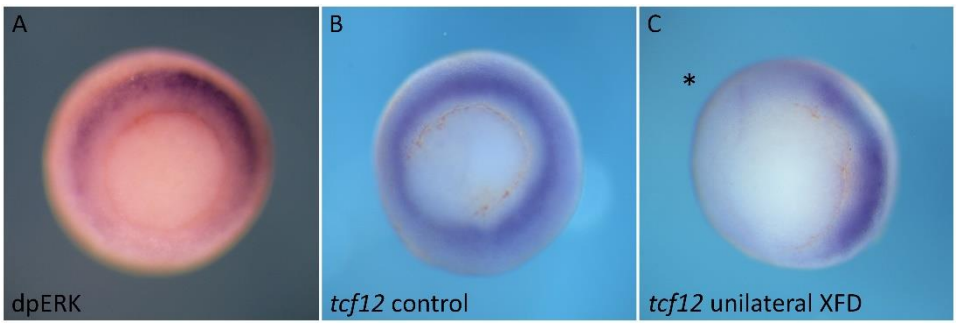


Figure 5.19: dpERK immunostaining and *tcf12* in situ hybridisation. Stage 10.5 *X. laevis* embryos **A** dpERK immunostaining, **B** *tcf12* in situ hybridisation, **C** *tcf12* in situ hybridisation of *X. laevis* embryo unilaterally injected (*) with dominant negative truncated FGF receptor (XFD) at 2-cell stage. In situ hybridisation carried out by Jennika Bates.

5.2.13 FGF promotes differentiation of human skeletal muscle progenitors

Having induced skeletal muscle in *X. laevis* organoids and shown the importance of FGF and *tcf12* in this system, skeletal muscle progenitors were differentiated from human H9 ES cells to determine if a similar pathway is involved. A protocol published by Shelton et al. (2016) was adapted for this purpose, with H9 cells pre-treated with 10µM rho kinase inhibitor Y27632 (Watanabe et al., 2007) and dissociated with TrypLE before replating at $1.5x10^5$ cells per well of 12-well plates. After 24 hours, cells were treated with 10µM CHIR99021 (Chiron) (Kreuser et al., 2020) from day 0-2, and 5ng/ml FGF2 from day 12-20 (Figure 5.20). Chiron is a potent GSK3 inhibitor often used to activate Wnt to drive mesoderm specification in mesoderm-lineage differentiation protocols (Kreuser et al., 2020). At day 0, small clusters and single cells were visible (Figure 5.20A). By day 12, different morphologies had begun to emerge with some cells appearing round and others elongated (Figure 5.20B). A heterogeneous population of cells developed including 3D structures which elongated and formed processes which lengthened over time (Figure 5.20C).

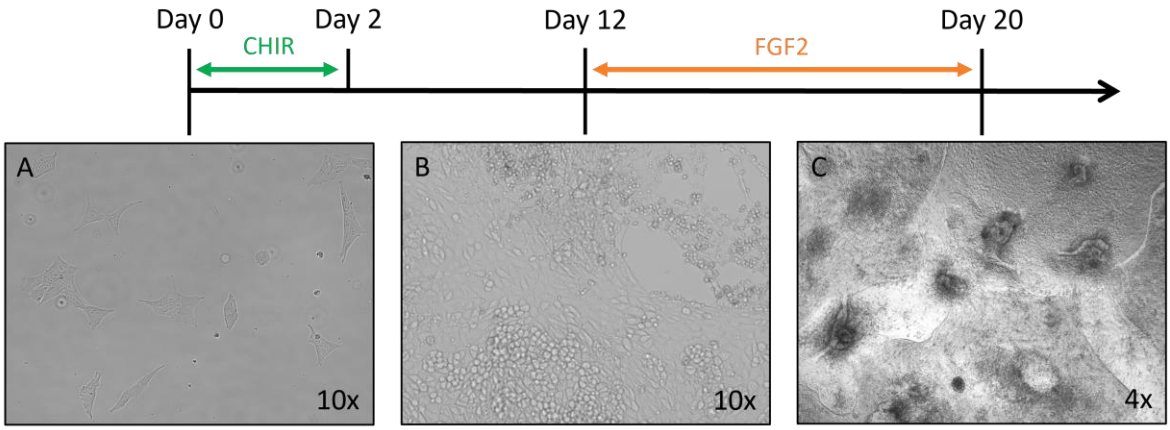


Figure 5.20: Differentiation of skeletal muscle progenitors from H9 ES cells. Human H9 ES cells differentiated following protocol adapted from (Shelton et al., 2016) involving treatment with 10µM CHIR99021 (CHIR) from day 0-2 and 5ng/ml FGF2 from day 12-20.

In order to determine whether this protocol successfully induced skeletal muscle progenitors, RNA samples were collected at the following time points for qPCR analysis: day 0, day 2, day 7, day 12, and day 20 with and without FGF2 treatment. Samples were analysed for expression of pluripotency, mesodermal and myogenic markers. Pluripotency gene *NANOG* was expressed at day 0 but was no longer expressed by day 2 after treatment with Chiron (Figure 5.21A). Within the time points analysed, mesodermal marker *TBXT* (brachyury) was only expressed at day 2 following Chiron treatment (Figure 5.21B). The highest expression of early muscle marker *PAX3* was seen at day 7 (Figure 5.21C), followed by the highest expression of early myogenic regulatory factor *MYF5* at day 12 (Figure 5.21D). Expression of the myogenic progenitor marker *PAX7* was detected at each time point collected with the exception of day 0 (Figure 5.21F). Expression patterns for skeletal

muscle specific genes *MYOD, MYOG* and *MYH3* were similar with the highest expression in cells treated with FGF2 at day 20, and relatively low expression in all other samples (Figure 5.21E, G and H). Day 20 cells not treated with FGF2 showed lower expression of muscle genes suggesting that FGF signalling may have an important role in human skeletal muscle differentiation, in addition to that in *Xenopus*.

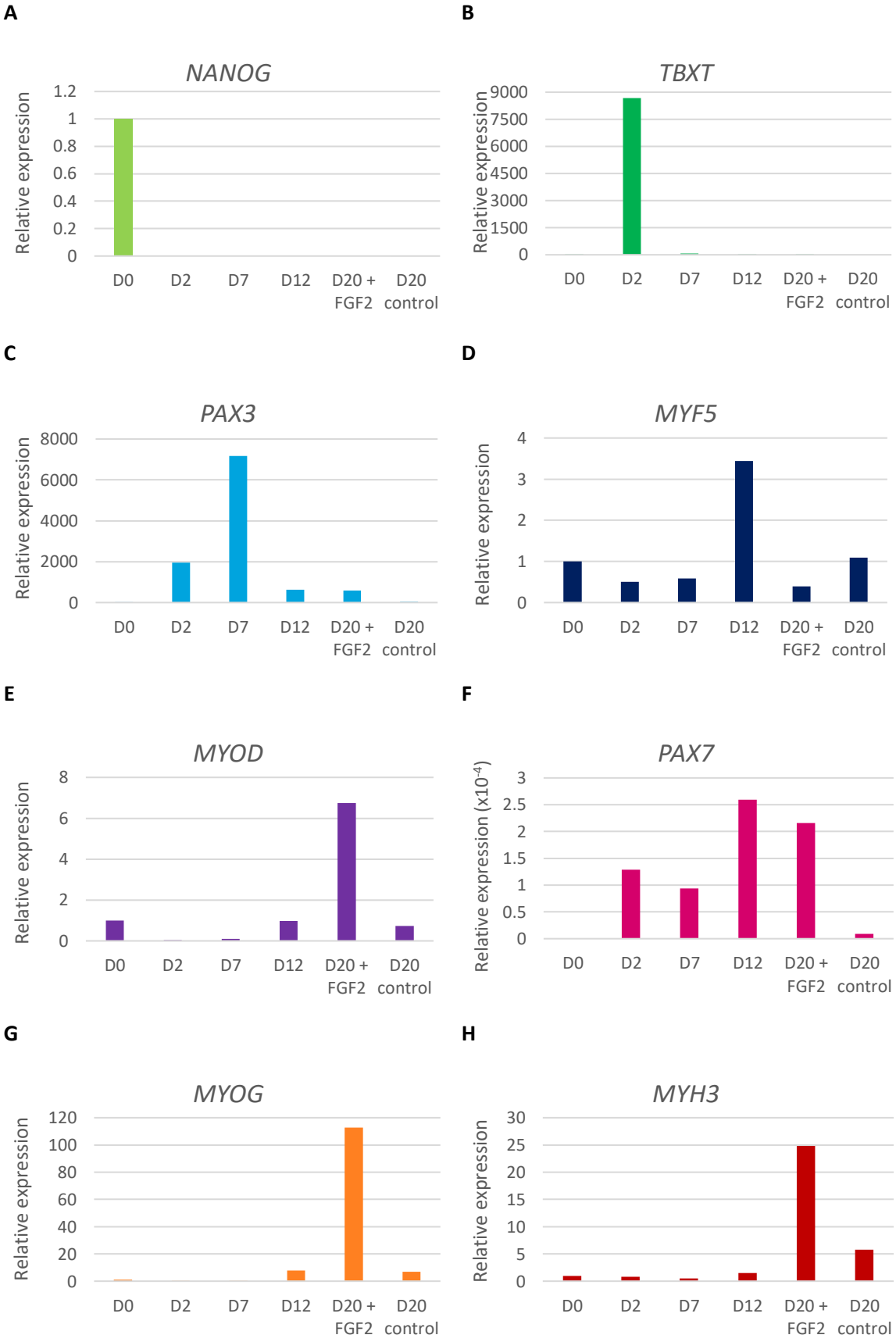


Figure 5.21: qPCR analysis of pluripotency, mesodermal and myogenic markers in differentiating H9 cells. Relative expression normalised to actin beta expression and D0 ($2^{-\Delta\Delta Ct}$). Samples = Day 0, Day 2, Day 7, Day 12, Day 20 treated with 5ng/ml FGF2 for 8 days, Day 20 no FGF2 treatment. n=1.

5.2.14 Candidate gene expression is conserved in human skeletal muscle progenitors

qPCR analysis for candidate genes identified in *X. laevis* was carried out to determine whether gene expression is conserved in human skeletal muscle differentiation. *TCF15* expression was present at a low level in each sample (Figure 5.22E). *TCF12, SMYD1, MEX3B* and *RBFOX2* were expressed most highly in FGF2 treated cells, with reduced expression in day 20 untreated cells (Figure 5.22).

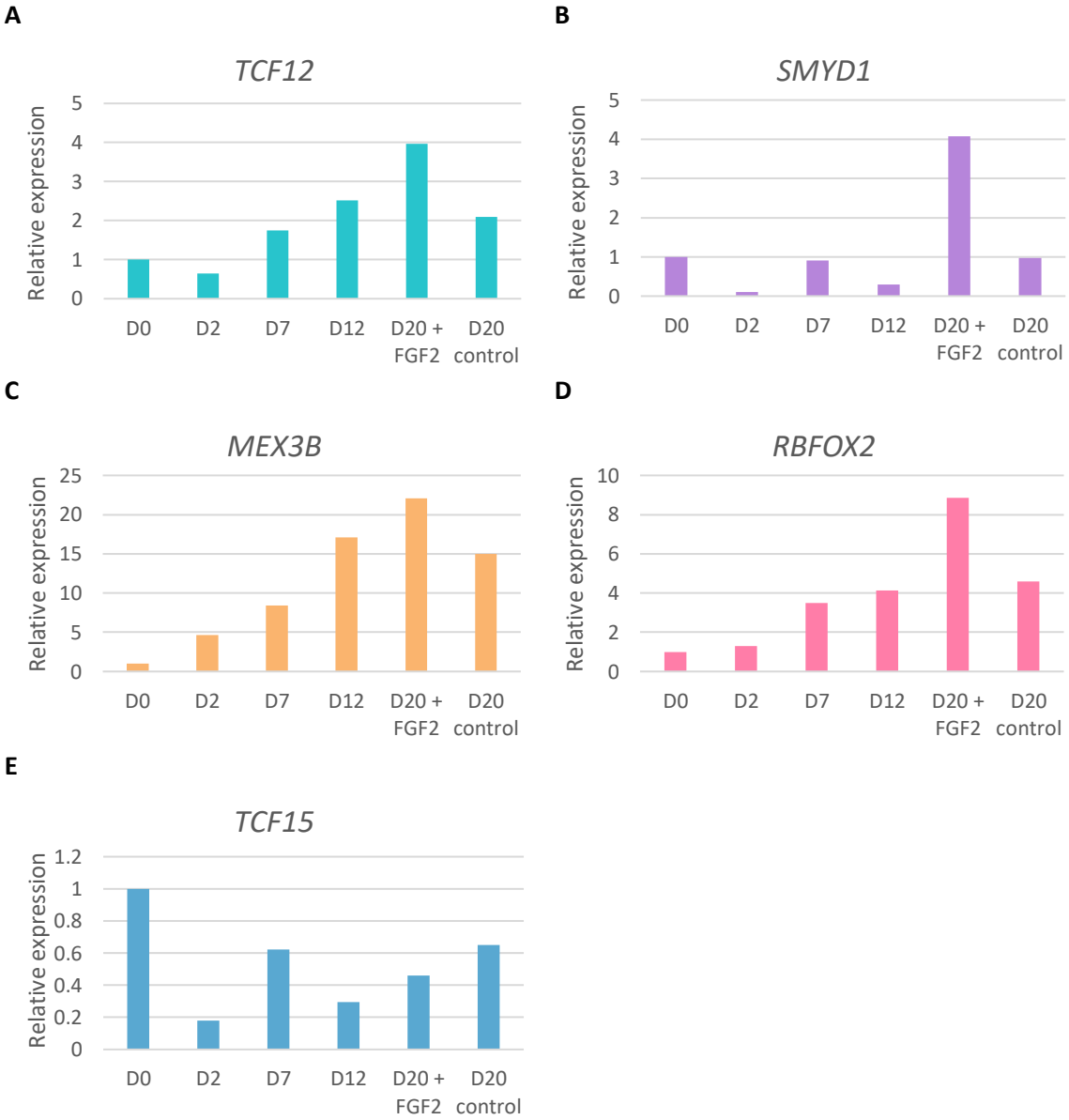


Figure 5.22: qPCR analysis of candidate genes in differentiating H9 cells. Relative expression normalised to actin beta expression and D0 ($2^{-\Delta\Delta Ct}$). Samples = Day 0, Day 2, Day 7, Day 12, Day 20 treated with 5ng/ml FGF2 for 8 days, Day 20 no FGF2 treatment. n=1.

A relatively low level of candidate gene expression, including FGF target *tcf12*, was also observed before addition of FGF2 (Figure 5.22). In order to determine whether additional endogenous or indirect FGF signalling was occurring within the H9 cell culture, dpERK expression was analysed. Protein samples without FGF2 treatment were collected at day 0, day 2, day 4, day 7, day 12 and day 20 for western blot analysis. dpERK expression was highest in day 2 cells following Chiron treatment (Figure 5.23) and coincided with expression of FGF target TBXT (Figure 5.21B). Decreased levels of dpERK were also present in all later time points analysed.

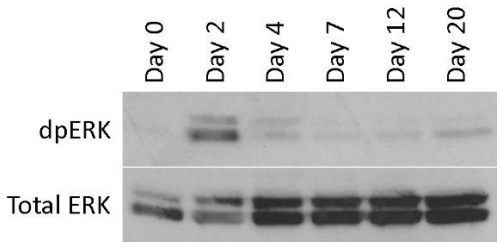


Figure 5.23: Western blot detecting diphosphorylated ERK (dpERK) in differentiating H9 cells.

5.3 Discussion

5.3.1 Developmental signalling pathways and histone methylation are enriched gene ontology terms in stage 14 Myod + Fgf4 organoids

At stage 14, several developmental signalling pathways were identified as highly statistically significant over-represented biological processes in Myod + Fgf4 organoids: Wnt, MAPK/ERK, Smoothened and Notch. At this stage of development, cells have not yet committed to a particular cell fate so it was expected that these biological processes would be enriched.

Histone methylation, protein methylation and protein alkylation were also highly enriched GO terms at this stage. This is in keeping with existing literature as modification of the chromatin landscape to promote enhancer activation and chromatin accessibility often precedes transcription of associated genes during development (Pálfy et al., 2020). It has been shown that methylation and demethylation of both gene-activating and gene-repressive histone marks are important for co-ordinating pluripotency gene regulatory networks, and the establishment and differentiation of cell lineages (Vougiouklakis et al., 2017; Jambhekar et al., 2019). FGF signalling, and downstream effector ERK, have been shown to regulate chromatin accessibility via histone acetylation and dissociation of the polycomb repressive complex (PRC) during specification of different lineages (Semprich et al., 2022; Tambalo et al., 2020). Alkylation of class I histone deacetylases 1, 2 and 3 has also been shown to inhibit their transcriptional repressor function (Doyle and Fitzpatrick, 2010).

As discussed in chapter 4, a subset of neural genes were also expressed in Myod and Myod + Fgf4 organoids. This is reflected in the top 20 GO terms at stage 14 as axonogenesis and central nervous

system development are enriched terms. Three muscle GO terms were already highly enriched, indicating that Myod + Fgf organoids were expressing muscle-related genes by stage 14.

5.3.2 Calcium ion sequestration, striated muscle development and muscle contraction gene ontology terms are enriched during Myod + Fgf4 organoid development

At stage 20, many highly enriched GO terms were associated with ribosomal RNA maturation and processing, and DNA replication. This is in keeping with previous work as modifications that occur during RNA maturation (for example by rRNA m6A methyltransferase METTL5) have been shown to promote cell fate determination and differentiation (Xing et al., 2020; Ignatova et al., 2020). Another highly enriched GO term at this stage was negative regulation of sequestering of calcium ion. When striated muscle is in a resting state, calcium ions are sequestered in the sarcoplasmic reticulum. When stimulated, the calcium ions are released into the sarcoplasm to initiate muscle contraction (Szent-Györgyi, 1975). The fact that genes involved in calcium ion sequestering are enriched suggests that the organoids are continuing towards the muscle fate at stage 20.

By stage 30 in normal embryonic development, tissues have started to differentiate. The GO terms highly enriched at stage 30 in the Myod + Fgf4 organoids reflect this as they include striated muscle tissue development, regulation of muscle contraction and many GO terms related to regulation of calcium ion movement. The combination of enriched muscle contraction genes, and genes associated with calcium ion sequestering and release, further confirms the effectiveness of the skeletal muscle protocol.

5.3.3 Co-expression of Tcf12 and Myod induces skeletal muscle in organoids

Due to elevated expression of candidate genes in stage 14 Myod + Fgf4 organoids (when cell fates are being specified), it was predicted that candidate genes may be able to replace Fgf4 in the skeletal muscle protocol. Out of five candidate genes interrogated, *tcf12.S* proved capable of replacing Fgf4 and still giving rise to muscle.

Tcf12 has previously been implicated in the myogenic pathway (Parker et al., 2006; Wang et al., 2022). bHLH transcription factors can be categorised into two groups: class I proteins such as the widely expressed E-proteins (HEB/Tcf12, E2A/Tcf3, E2-2/Tcf4 and Daughterless), and tissue-specific class II proteins such as the MRFs (Massari and Murre, 2000). Heterodimerisation of MRFs with E-proteins has been shown to modulate MRF DNA-binding specificity (Lassar et al., 1991; Hu et al., 1992). E2A gene products E12 and E47 were initially thought to play a key role in myogenesis due to their ability to heterodimerise with Myod in vitro (Murre et al., 1989, 1991), however it was later demonstrated that E12 and E47 are not required for skeletal muscle differentiation in mouse ES cells (Zhuang et al., 1992). In addition, C2C12 myoblasts, 10T1/2 fibroblasts, and the nuclei of unfused myoblasts and fused myotubes express Tcf12, and not E12 or E47 (Conway et al., 2004;

Perry et al., 2001). It has since been shown that Tcf12 regulates the transcriptional activity of Myod during early skeletal muscle differentiation, and Myogenin activity later in differentiation (Parker et al., 2006; Hu et al., 1992). Tcf12-Myod1 interactions have also been shown to stabilise chromatin and maintain myogenic gene expression in proliferating cells, and inducible deletion of Tcf12 results in defects in mouse muscle development and regeneration (Wang et al., 2022). While this thesis has demonstrated that Tcf12 expression requires FGF, an additional regulatory mechanism has been reported involving non-coding RNAs. MicroRNA miR-7 inhibits Tcf12 expression, but circular RNA circHIPK3 sponges miR-7 to increase Tcf12 expression and promote myoblast proliferation and differentiation (Gao et al., 2021). Tcf12 also directly associates with PRC2 at a subset of developmental promoters in mouse ES cells, including at genes involved in mesoderm specification, and at Hox gene family members (Yoon et al., 2015). Upon differentiation, Tcf12/PRC2-bound promoters switch to become associated with Tcf12/SMAD2/3 indicating that Tcf12 plays a role in both de-repression of lineage specific genes, and Nodal signalling (Yoon et al., 2011, 2015). This suggests that Tcf12 may have been sufficient to replace Fgf4 in the protocol due to its ability to regulate gene expression via formation of Myod heterodimers stabilising myogenic chromatin accessibility, in combination with regulation of repression of other lineage promoters, and interaction with TGFβ superfamily signalling pathways.

In contrast, candidate genes such as RNA-binding proteins *rbfox2.S, mex3b.S* and *lin28A1,* which are involved in post-transcriptional regulation, may require expression of additional FGF targets or signalling pathway activity to help promote effective myogenesis (Takada et al., 2009; Singh et al., 2014). For example, phosphorylation of Lin28A by the ERK signalling pathway in mouse ES cells, has been shown to stabilise Lin28A expression and promote the transition from naïve to primed pluripotency (Tsanov et al., 2017).

Another possibility is that some candidate genes were highly expressed at stage 14 to specify muscle lineage but then downregulated later in development to avoid inhibiting differentiation. An example of this type of expression pattern is sox15 which is expressed highly in myoblasts but antagonises differentiation into myotubes (Béranger et al., 2000). If a candidate gene functioned in this way, and was not involved in (directly or indirectly) initiating the cascade of gene transcription required for differentiation, it would not be sufficient to induce skeletal muscle. Therefore expressing a combination of candidate genes with different mechanisms of action alongside Myod, may have also resulted in skeletal muscle differentiation. Testing the ability of combinations of candidate genes to replace Fgf4 in the protocol, or increasing the amount of mRNA injected, would be a logical next step. For example, replacing Fgf4 with mRNA encoding an RNA-binding protein, histone methyltransferase *smyd1.L,* and transcription factor *tcf15.L,* may be sufficient for

differentiation due to their differing and potentially complementary mechanisms of action (Nagandla et al., 2016; Burgess et al., 1996; Davies et al., 2013). Now that the protocol has been established and a successful candidate identified, a greater number of candidate genes could also be selected and tested.

Interestingly, the histology of Tcf12 + Myod organoids differed from Myod + Fgf4 organoids. While both contain organised skeletal muscle, Myod + Fgf4 organoids typically formed a large vesicle around the block of muscle, which is not present in Tcf12 + Myod organoids. As Tcf12 is a downstream target of FGF signalling, Tcf12 + Myod may result in a more directed differentiation, with less non-muscle mesoderm formed. However, further analysis of the Tcf12 + Myod organoids, such as RNA-seq, would be required to determine any significant differences between the conditions.

5.3.4 Human skeletal muscle progenitors form part of a heterogeneous population

Despite the fact that Tcf12 was the only candidate gene tested that was sufficient to replace Fgf4 in the *Xenopus* protocol, expression of 4 candidate genes increased during early human skeletal muscle progenitor differentiation. In addition, TCF12 has previously been shown to be required for mesoderm development, and NANOG and TGFβ signalling in human ES cells (Li et al., 2017). Disruption of the *TCF12* gene locus resulted in cells retaining pluripotency, and impaired mesodermal development (Li et al., 2017). This suggests that the genes identified may have a conserved role in lineage specification in *Xenopus* and humans.

Expression patterns for lineage markers analysed were in line with previous literature with pluripotency gene NANOG expressed at day 0, before mesodermal marker TBXT, followed by early myogenic genes PAX3 and MYF5, and finally skeletal muscle specific genes MYOD, MYOG and MYH3 (Chambers et al., 2003; Herrmann et al., 1990; Buckingham and Relaix, 2007; Ott et al., 1991; Weintraub, 1993; Hasty et al., 1993; Schiaffino et al., 2015). However, differentiating cell populations showed varying morphologies within each well. This was not unexpected as Shelton and colleagues reported approximately 50% of cells to be myocytes or myotubes after 50 days of their differentiation protocol (Shelton et al., 2016). An adaptation of the Shelton protocol used for this thesis, was to culture H9 cells on vitronectin rather than ill-defined Matrigel. One of the benefits of this is that the heterogeneity of cell culture can be improved by using better defined medium and serum-free conditions without feeders (Marks et al., 2012). The heterogeneity of human and mouse pluripotent cells has been observed in many studies, with expression of genes such as NANOG fluctuating during in vitro culture (Chambers et al., 2007; Singh et al., 2007; Hough et al., 2014). Some human ES cell lines (including H9), have also been shown to acquire epigenetic variants after cell line derivation, leading to mixtures of cells with different epigenetic states (Tanasijevic et

al., 2009). Protocols developed for differentiation of ES and pluripotent cells into specific cell types are typically not 100% efficient due to incomplete understanding of lineage commitment, and the complexities of heterogeneity and signalling pathway interactions.

5.3.5 Multiple signalling pathways interact to promote lineage commitment and differentiation

The role of FGF signalling in cell culture is complex with some human ES cell studies indicating that FGF2 promotes proliferation and represses differentiation, whereas other studies implicate a role in lineage commitment (Xu et al., 2005; Levenstein et al., 2006; Kunath et al., 2007; Ying et al., 2008). For example, in mouse ES cells, high levels of ERK signalling have been linked to differentiation, whereas low levels of ERK promote cell proliferation, cell cycle progression and genomic stability (Ma et al., 2016). Inhibition of ERK signalling in mouse ES cells has also been shown to be essential for maintaining naïve pluripotency (Ying et al., 2008). In contrast, in human ES cells, the Shelton protocol describes the purpose of FGF2 treatment on days 12-20 to promote progenitor proliferation and suppress early MRF expression in order to avoid premature differentiation (Shelton et al., 2016). In spite of this, results presented in this chapter show that FGF2 treated H9 cells have increased expression of MRFs (and some candidate genes) at day 20 when compared with untreated cells.

Adding to complexity, other FGF signalling pathway effectors such as PI3K/Akt have also been associated with regulating differentiation. For instance, PI3K/Akt activity maintains self-renewal of human pluripotent cells by suppressing ERK and canonical Wnt signalling, and promoting activation of self-renewal genes such as NANOG via Activin A/Smad2/3 (Singh et al., 2012; Xu et al., 2008; Vallier et al., 2009). When PI3K/Akt signalling is low, ERK targets GSK3β to activate canonical Wnt signalling, which switches Smad2/3 activity to instead promote differentiation (Singh et al., 2012).

Hh signalling has been shown to promote FGF signalling to pattern anterior mesoderm during gastrulation in mice (Guzzetta et al., 2020), however expression of Hh ligands was low during the protocol. Smoothened and Gli protein expression was highest in Myod + Fgf4 organoids, followed by Myod organoids. This is in keeping with previous studies as Gli2 forms complexes with Myod and Mef2c to enhance Myod activity at myogenic promoters (Voronova et al., 2013). Hh may have been expressed at other stages of development that were not analysed, or Smoothened may be signalling non-canonically. For instance, in drosophila, it has been shown that Smoothened can be intracellularly activated independently of Hh or Patched (Jiang et al., 2018).

TGFβ superfamily members have also been implicated in lineage specification and cell fate commitment. For example, it has been shown that formation of the primitive streak or mesoderm

progenitors from human ES cells requires the cooperative action of Wnt/β-catenin, Activin/Nodal and BMP signalling pathways (Sumi et al., 2008). However TGFβ, Activin and Nodal signalling may not play a key role in skeletal muscle specification during the *Xenopus* differentiation protocol, as expression was not elevated in stage 14 Myod + Fgf4 organoids. *smad2.L* and *smad3.L* function downstream of TGFβ, activin and nodal, and their expression also did not increase until later in development (stage 20 and 30 respectively) (Kumar et al., 2001).

The highest expression of BMP signalling mediator *smad1.S*, co-smad *smad4.1.L* and *bmp7.2.L* was in stage 14 Myod organoids. However *smad1.S, smad1.L, smad4.1.L, smad4.1.S, bmp7.2.L, bmp3.L, bmpr1b.S* and *bmp1.L* also showed elevated expression in stage 14 Myod + Fgf4 organoids, indicating a potential role for BMP signalling in skeletal muscle lineage specification. It would be interesting to inhibit BMP receptors during the protocol to determine whether BMP signalling is required for skeletal muscle differentiation.

The Wnt signalling pathway was a statistically significant enriched GO pathway in genes upregulated in both Fgf4 and Myod + Fgf4 organoids, indicating a more prominent and FGF-regulated role for Wnt in the protocol. *wnt11b.L*, *wnt8a.L* and *wnt8a.S* were highly expressed in stage 14 Fgf4 + Myod and Fgf4 organoids. This aligns with previous studies as these genes are known FGF targets and have been associated with mesoderm induction and myogenesis (Hong et al., 2008). For example, an isoform of Wnt11b has been shown to regulate somite formation, and Wnt8 is a mesoderm patterning factor expressed ventrally in *Xenopus* mesoderm during gastrulation (Dichmann et al., 2015; Hoppler and Moon, 1998; Hong et al., 2008). Introduction of a dominant negative Wnt8 blocks Myod induction in *Xenopus* embryos, and inhibiting Wnt8 signalling reduces skeletal muscle formation (Hoppler et al., 1996; Wu et al., 2000). Wnt8a has also been shown to regulate Tbxt expression and mesoderm induction via canonical signalling in human ES cells (Mazzotta et al., 2016).

In multiple cell culture differentiation protocols, 3μM Chiron has been used to inhibit GSK3, activate Wnt signalling, and induce mesoderm (Borchin et al., 2013; Chal et al., 2015). However, it has since been shown that higher concentrations of 7.5-10μM Chiron improve differentiation efficiency by enhancing paraxial mesoderm gene expression, and reducing expression of ectoderm/neuronal genes (Mendjan et al., 2014; Naujok et al., 2014; Tan et al., 2013). It also significantly increases expression of MSGN1 and TGFβ superfamily genes, particularly NODAL (Shelton et al., 2019). dpERK expression was highest in H9 cells collected on day 2, following a 10μM Chiron pulse, indicating that FGF-ERK activity was also stimulated (Shelton et al., 2016). This further supports the hypothesis that multiple signalling pathways interact in complex ways to determine specific cell fates.

5.3.6 FGF and Wnt signalling interactions promote paraxial mesoderm lineage specification

The Wnt and FGF signalling pathways cooperate to regulate the specification and formation of paraxial mesoderm. In the absence of Wnt or FGF signalling, mouse embryos are truncated and cells normally destined to become posterior paraxial mesoderm form ectopic neural tissue (Boulet and Capecchi, 2012; Ciruna and Rossant, 2001; Takada et al., 1994). Therefore FGF and Wnt signalling are required for differentiation of neuromesodermal progenitors towards a paraxial mesoderm fate rather than neural (Garriock et al., 2015; Jurberg et al., 2014).

During *Xenopus* gastrulation, key FGF8 and Wnt8 targets are expressed in partially overlapping regions around the blastopore (Kjolby et al., 2019). FGF signalling mediator Ets2 has been shown to bind near all Wnt target genes, indicating one mechanism through which these signalling pathways interact (Kjolby et al., 2019). Additionally, FGF has been shown to weaken the ability of co-repressor transducin-like enhancer of split 4/groucho-related gene 4 (TLE4/Grg4) to inhibit canonical Wnt signalling targets (Burks et al., 2009). Therefore FGF can activate expression of a subset of Wnt target genes directly via downstream effectors, or through transcriptional de-repression.

In addition to the protocol presented in this thesis, activation of both Wnt + Fgf signalling is currently the only other way of reliably inducing skeletal muscle in organoids. Stimulation of the Wnt + Fgf pathways together has been shown to induce more skeletal muscle and notochord in organoids than Fgf4 treatment alone (Christian et al., 1992; Slack et al., 1988; Klein and Melton, 1996). Both Fgf4 and Wnt8 signalling are required for *myod* expression as inhibition of either leads to a loss of *myod* transcriptional activation (Hoppler et al., 1996; Fisher et al., 2002). In keeping with this, the combined expression of Wnt8 + Fgf4 in organoids has been shown to enhance Myod and mesodermal Cdx gene expression, compared to Fgf4 alone (Keenan et al., 2006; Burks et al., 2009). However Myod + 100pg cska Wnt8 did not induce skeletal muscle in this study. This may be due to the fact that a greater quantity of Wnt8, or additional FGF signalling targets are required for effective myogenesis. 100pg cska Wnt8 expression alone was not sufficient for mesoderm formation in animal cap organoids, however cska Wnt8 has previously been shown to induce ventral mesoderm (Christian and Moon, 1993). Therefore it is also possible that the cska Wnt8 injected was not transcribed effectively, so this experiment would have been improved by analysis of a direct target of Wnt signalling as a positive control.

Early lineage commitment involves a complicated balance of various signalling pathways interacting such as FGF, Wnt, Activin/Nodal/TGFβ, BMP, Notch and Shh (Sumi et al., 2008; Na et al., 2010; Guzzetta et al., 2020; López et al., 2003). While 50ng/ml Fgf4 alone was not sufficient to induce skeletal muscle, organoids co-expressing Myod + Fgf4 activated the required balance of FGF, Wnt, BMP, Smoothened and Notch signalling targets for effective myogenesis. FGF target Tcf12 was

sufficient to replace Fgf4 in the protocol and still give rise to muscle. RNA-seq analysis of Myod + Tcf12 organoids is in progress and will allow insight into whether the same signalling pathways and transcriptional targets are involved under these conditions. Comparing the Myod + Fgf4 and Myod + Tcf12 datasets, will determine differences in the gene regulatory networks involved in these myogenic programmes, as well as the relative myogenic potency.

Chapter 6: Discussion

6.1 Summary

Much has been discovered about the molecular basis of the establishment of the skeletal muscle lineage over the past 70 years, however the specific mechanisms responsible for regulation of the switch from pluripotency to lineage commitment are not yet fully understood. While Myod was once referred to as the "master regulator" of the muscle lineage due to its ability to convert fibroblasts to myoblasts, it is now clear that additional factors are required for effective myogenesis in other cell lines and pluripotent cells (Davis et al., 1987; Hopwood and Gurdon, 1990; Dekel et al., 1992). FGF signalling is required for skeletal muscle development in *Xenopus*, the expression and maintenance of key mesodermal genes such as *myod* and *tbxt*, and the community effect (Amaya et al., 1991; Isaacs et al., 1994; Fisher et al., 2002; Standley et al., 2001). FGF-ERK-Etv5 signalling has also been implicated in the progression of naïve pluripotency to lineage competence (Kalkan et al., 2019). Therefore, the hypothesis proposed and tested in this thesis was that FGF signalling is the competence factor required to allow the transition from pluripotency to skeletal muscle lineage commitment.

In order to test this, a skeletal muscle inducing protocol was developed in organoids derived from *X. laevis* animal cap explants (chapter 3). In keeping with my hypothesis, overexpression of Myod in organoids was not sufficient to induce muscle, however culture of Myod organoids in 50ng/ml Fgf4 protein induced formation of a block of organised, differentiated skeletal muscle with clear myotubules.

In chapter 4, further analysis of protocol organoids at a transcriptomic level revealed that the skeletal muscle induced in Myod + Fgf4 organoids was fast twitch muscle. Myod + Fgf4 organoids showed expression of MRFs, as well as muscle-specific structural proteins and metabolic enzymes. Additionally, comparison of genes expressed in stage 30 Myod + Fgf4 organoids with existing *X. tropicalis* and human skeletal muscle datasets revealed a high degree of similarity. Expression of the majority of skeletal muscle genes appears to be conserved between protocol organoids, *Xenopus* embryos and humans, suggesting that this protocol may have useful applications in identification of genetic players also involved in human myogenesis.

In chapter 5, analysis of the RNA-seq dataset was used to investigate the role of FGF in Myod driven myogenesis in *Xenopus.* Gene ontology analysis identified the Wnt signalling pathway as statistically significantly enriched as a result of FGF signalling in Myod + Fgf4 organoids. This supports the notion that multiple signalling pathways interact to promote lineage differentiation, with FGF playing a pivotal role. The dataset was also used to identify genes potentially involved in skeletal muscle

specification, such as *tcf12*. Microinjection of 1ng *tcf12* mRNA is sufficient to replace Fgf4 treatment in the protocol, and still give rise to muscle. Inhibition of FGF signalling via XFD expression in *X. laevis* embryos resulted in a complete loss of *tcf12* expression, indicating that endogenous *tcf12* expression requires FGF. FGF transcriptionally activates the expression of *tcf12* (this thesis), as well as *myod* (Fisher et al., 2002), and the Myod itself also activates Tcf12 transcription, as part of a feed-forward mechanism (Figure 6.1). Myod and Myog each form heterodimers with Tcf12 that synergise to activate transcription at E-box elements (Parker et al., 2006). The Myod/Tcf12 heterodimers have also been shown to stabilise chromatin accessibility to maintain myogenic gene expression in mice (Wang et al., 2022). Multiple consensus sites for MAPK phosphorylation have been identified within Tcf12, suggesting that FGF may not only enhance Tcf12 transcription, but also act post-transcriptionally to further regulate Tcf12 protein stability and lineage specification. Differentiation of human skeletal muscle progenitors from H9 ES cells revealed that expression of 4/5 *Xenopus* candidate genes, including *tcf12*, increased during human myogenesis. This suggests that the protocol allowed successful identification of an FGF target with a conserved role in skeletal muscle lineage commitment: *tcf12*.

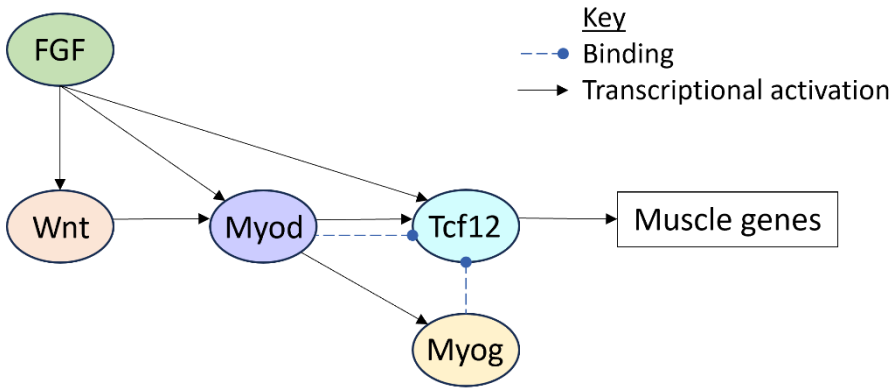


Figure 6.1: Schematic diagram representing FGF regulated feed-forward mechanism involved in skeletal muscle lineage commitment.

6.2 Fibroblast growth factor signalling specifies multiple cell lineages

6.2.1 FGF contributes to specification of neural lineages

While this thesis focused on the skeletal muscle lineage, FGF signalling has also been implicated in specification of other cell fates such as the neural lineage. Early animal cap studies suggested that ectodermal cells adopt a neural fate by default in the absence of cell-cell signalling, as cells disaggregated for 5 hours spontaneously become neural (Godsave and Slack, 1989; Grunz and Tacke, 1989; Weinstein and Hemmati-Brivanlou, 1997). This supported the neural 'default model' which proposed that cells are pre-programmed towards a neural fate, which is inhibited by endogenous BMP signalling. However it was later shown that inhibition of BMP or SMAD1 is not

sufficient for neural induction of ventral ectoderm, unless low levels of Fgf4 or Ras signalling are also present (Delaune et al., 2005). In order to induce the neural fate, FGF has been shown to repress BMP transcription, phosphorylate and inhibit Smad1, as well as act independently of BMP by inducing expression of neural transcription factors such as Zic3 (Pera et al., 2003; Londin et al., 2005; Marchal et al., 2009). FGF signalling plays a key role in neural induction in multiple species including ascidians, amphibians, fish and birds (Marchal et al., 2009; Hongo and Okamoto, 2022; Kudoh et al., 2004; Rentzsch et al., 2004; Alvarez et al., 1998; Storey et al., 1998; Bertrand et al., 2003). FGFs also have major roles in the induction and patterning of the neural plate, neural crest, and nervous system (Geary and Labonne, 2018; Hong et al., 2008; Takemoto et al., 2006). From this, it is clear that FGF is required for specification of other cell fates in addition to the muscle lineage.

6.2.2 FGF contributes to specification of blood lineages

FGF signalling is required for the earliest expression of genes in the nascent mesoderm (Isaacs et al., 1994). It is also required for the cellular response to mesoderm induction, and contributes to specification of dorsal and ventral mesoderm derivatives (Cornell et al., 1995). The blood is a ventral derivative of the mesoderm and FGF is a powerful negative regulator of this lineage. In contrast, SCL/tal1 (stem cell leukaemia) is a bHLH transcriptional regulator of the blood lineage. Depletion of Fgf4 in *Xenopus* embryos results in an expansion of SCL expression; indicating the role of FGF in restricting this ventral cell type (Isaacs et al., 2007). The positive role of FGF in activating and maintaining the expression of Myod to direct the muscle lineage dorsally has been discussed. FGF signalling also delimits the expression domain of BMP to specify cellular identity across the early dorsoventral axis (Fürthauer et al., 2004). In addition, FGF signalling has been shown to mediate a signal transduction pathway between Wnt16 and deltaC (Dlc) to regulate haematopoietic stem cell specification (Lee et al., 2014). FGF signalling has also been implicated in angiogenesis and neovascularisation (Murakami and Simons, 2008). For example, FGF2 beads have been shown to induce endothelial precursors (angioblasts) and pattern vessel formation in quail embryos (Cox and Poole, 2000). Therefore, understanding the role of FGF in muscle lineage specification and its interactions with other cell signalling pathways, could help elucidate the mechanisms involved in commitment to a variety of cell fates.

6.2.3 FGF and skeletal muscle regeneration

As many developmental processes are recapitulated in adult regeneration, furthering our understanding of these mechanisms may help inform approaches in regenerative medicine. Mammals have a relatively good skeletal muscle regenerative capacity due to the existence of muscle stem cells called satellite cells (Bischoff and Heintz, 1994; Morgan and Partridge, 2003). Satellite cells are activated following muscle injury or growth signals, re-enter the cell cycle, and

proliferate via symmetric and asymmetric divisions to either self-renew and maintain the quiescent pool of satellite cells, or expand as myoblasts (McKinnell et al., 2005). Myogenic progenitors differentiate and fuse to each other, or to damaged muscle fibres, to repair muscle integrity and function (Morgan and Partridge, 2003). Loss-of-function dystrophin mutations in Duchenne muscular dystrophy patients affect the regulation of satellite cell polarity and asymmetric division, leading to impaired muscle regeneration (Dumont et al., 2015; Sacco et al., 2010). With age, the efficiency of the regenerative response and self-renewal capacity of satellite cells also decreases, which can lead to sarcopenia (age-related loss of skeletal muscle mass, function and strength) (Bernet et al., 2014; Cosgrove et al., 2014; Fernández-Lázaro et al., 2022; Sousa-Victor and Muñoz-Cánoves, 2016).

As FGF signalling is important in skeletal muscle lineage specification, it is unsurprising that it has also been implicated in muscle regeneration as activated satellite daughter cells commit to the muscle fate. Multiple FGFs have been shown to stimulate rat satellite cell proliferation in culture indicating potential roles in adult myogenesis (Sheehan and Allen, 1999; Kastner et al., 2000). In mice, application of an FGF2-neutralising antibody has been shown to impair muscle regeneration following a crush injury (Lefaucheur and Sébille, 1995b). Wild-type mice also upregulate FGF6 following skeletal muscle injury to promote myoblast proliferation and muscle differentiation (Floss et al., 1997; Armand et al., 2005). Although FGF6 knockout mouse model results are mixed, one model demonstrated a severe regeneration defect with fibrosis and myotube degeneration, which can be rescued by injection of FGF6 (Armand et al., 2003, 2005; Fiore et al., 2000). Loss of FGFR4 has also been shown to result in a severe regeneration defect following toxin-induced muscle injury in mice (Zhao et al., 2006).

Immediate early response gene *fos* is an FGF target rapidly and transiently expressed upon injury. Fos-null satellite cells are slower to activate with a reduced regenerative capacity, indicating a key role for Fos in skeletal muscle repair (Almada et al., 2021). In addition, aged satellite cells show reduced responsiveness to FGF, and ectopic activation of FGFR1 partially rescues compromised satellite cell self-renewal (Bernet et al., 2014). FGF2 and its target miR-29 have also been shown to be required for effective muscle regeneration following injury or exercise (Galimov et al., 2016; Lefaucheur and Sébille, 1995a, 1995b). Muscle-specific deletion of Tcf12 in mice has been shown to reduce muscle weight and myofibre size, and inducible deletion in adult muscle stem cells delayed muscle regeneration (Wang et al., 2022). Therefore FGF signalling and downstream targets are promising key players in muscle regeneration and regulation of satellite cell activation and fate.

6.3 Disease treatment

The *Xenopus* protocol developed in this thesis provides an inexpensive and efficient way to identify transcriptional targets worthy of further investigation, for instance in mammalian models or human cell culture. As the expression of the majority of skeletal muscle genes is conserved between humans and protocol organoids, it is a useful system for preliminary screening and initial candidate gene detection. Identification of key genetic players and interrogation of their mechanisms of action (e.g. downstream target gene expression/de-repression, signalling pathway interactions, chromatin remodelling etc.) regulating myogenesis, may help improve human myoblast cell culture methods or inform development of future therapeutics. This could be applied in terms of improving protocols for more effective PAX7-positive skeletal muscle progenitor culture for patients with age-related or inherited muscle wasting. Cell culture protocols aiming to generate functional PAX7+ skeletal muscle precursors are currently inefficient, though FACS can be used to isolate myogenic cell populations. For example, Borchin and colleagues report >18% of cells cultured as per their protocol as PAX7+ (Borchin et al., 2013), while Shelton et al. show 40% of cells expressing PAX7 following their 50 day differentiation (Shelton et al., 2016).

6.4 Future work

6.4.1 Gene targeting and dominant negative inhibition of Tcf12

This thesis demonstrates the ability of E-protein and bHLH transcription factor Tcf12 to replace Fgf4 in the *Xenopus* skeletal muscle differentiation protocol. The next step is to determine whether Tcf12 is not only sufficient to induce muscle, but also required for myogenesis. One way this will be tested is through microinjection of antisense translation blocking morpholino oligos to disrupt translation of *tcf12.* The morpholino will first be tested on *X. laevis* embryos to determine whether Tcf12 is required for skeletal muscle development in vivo, before being tested in Myod + Fgf4 organoids. An additional, and potentially more effective approach, will be to express a dominant negative Tcf12. This is more likely to result in a phenotype, as E-proteins are capable of compensating for each other (Zhuang et al., 1998). This approach has been successful in previous studies as dominant negative Tcf12 proteins capable of forming non-functional heterodimers have been shown to block T-cell lineage development in mice (Barndt et al., 2000). Due to the fact that muscle specific deletion of Tcf12 in mouse muscle impaired muscle development and regeneration (Wang et al., 2022), inhibition of Tcf12 activity may also have an effect on *Xenopus* myogenesis. It would also be interesting to test if inhibition of TCF12 disrupts human myogenesis in cell culture, as has been shown in mouse models.

6.4.2 Technical improvements for human skeletal muscle differentiation and analysis

Tcf12 was also found to be expressed in human skeletal muscle progenitors, further supporting a conserved role across species. However, the differentiation protocol gave rise to a heterogeneous population of cells within each well. RNA samples have been sent for RNA-seq analysis to gain a better understanding of the genes expressed and the types of cells induced. It would be interesting to run the experiment for 50 days, as described in the original protocol (Shelton et al., 2016), to establish whether the culture conditions preferentially select for the muscle lineage over other cell fates over time. If clearly defined varying cell morphologies were still present, samples could be taken from different cell populations within a well for separate RNA-seq analyses, or ideally scRNA-seq would be carried out. It would also be beneficial to have more replicates of the human samples to increase statistical power and confidence in the reproducibility of the protocol. Immunofluorescent staining was attempted on the human cell populations at different stages for pluripotency, mesoderm and muscle markers, however the results were inconclusive as conditions required further optimisation. Testing conditions such as longer incubation/permeabilisation steps, different antibodies and antibody dilutions, and culturing cells on different coverslip/plate setups could improve antibody binding and image resolution for future samples. Phalloidin and DAPI staining at later stages of the protocol could help determine whether myoblasts have fused to form multinucleated myofibres.

6.4.3 Transcriptomic comparison of Myod + Fgf4 and Tcf12 + Myod organoids

Data presented in this thesis indicates that Tcf12 plays a key role in skeletal muscle lineage commitment, however Tcf12 + Myod organoid histology appeared different to Myod + Fgf4 organoids. In order to investigate this further, 1ng *tcf12* and 1ng *tcf12* + 1.2ng *myod* organoid samples at developmental stages 14, 20 and 30 have been submitted for RNA-seq analysis. This dataset will allow comparison between genes activated by Tcf12 + Myod and those expressed in Myod + Fgf4 organoids, during skeletal muscle specification and differentiation. From this, differences in the gene regulatory networks and myogenic potency could be determined. Tcf12 has also been implicated in other roles as E-proteins can bind various bHLH transcription factors to drive specification of many cell types during myogenesis, neurogenesis and haematopoiesis (Schlaeger et al., 2004; Belle and Zhuang, 2014). Tcf12 in particular has been shown to form heterodimers with Myod1, Myog, Neurod1 and Twist1 to regulate developmental fates (Parker et al., 2006; Singh et al., 2022; Fan et al., 2023; Hu et al., 1992). For example, when Neurod1 expression is induced in mouse ES cells, Tcf12-Neurod1 heterodimers form to drive neuronal migration via a gain of active chromatin and targeted gene expression (Singh et al., 2022). Therefore, it will be interesting to see if neural genes expressed in Myod + Fgf4 organoids are also present in Tcf12 + Myod organoids.

ChIP-seq could also be used to investigate the effect of Tcf12 + Myod expression on chromatin accessibility.

6.4.4 Identification and investigation of additional candidate genes

Development of this protocol has allowed identification of at least one candidate gene involved in skeletal muscle lineage specification. In the future, the protocol could be used to identify further candidates and test their ability to induce muscle. Combinations of mRNA encoding multiple candidates could also be tested to determine whether their combined mechanisms of action are sufficient to promote myogenesis.

6.5 Conclusions and implications

Tight regulation of multiple signalling pathways and target genes is essential for normal development, with FGF, and its targets Tcf12 and Myod, playing key roles in myogenesis. This thesis has demonstrated the role FGF has in activating, and enhancing, transcription of genes that cooperate with Myod to regulate gene transcription, and promote the skeletal muscle lineage.

The protocol developed has the potential to aid identification of key genetic players in muscle lineage specification via a fast and inexpensive assay, before further interrogation in human cell culture. Furthering our understanding of the molecular mechanisms required for the regulation of skeletal muscle differentiation may help improve myoblast cell culture methods, or aid identification of potential therapeutic targets for future muscle wasting disease treatments.

Appendix

R script for differential gene expression analysis using Sleuth

```
# load packages
library(tidyverse)
library(sleuth)

# load files, “samples.txt” file dictates condition to compare to e.g. _C
samples<-read.table("samples.txt", sep="\t", header=TRUE, stringsAsFactors=FALSE)
sample_id <- list.files("salmon")
sal_dirs <- file.path("salmon",sample_id)
transcripts <- read_tsv("transcript_to_gene.tsv")

# differential gene expression analysis for entire dataset - generate sleuth object
so <- sleuth_prep(samples, target_mapping=transcripts, aggregation_column = 'gene_id',
read_bootstrap_tpm = TRUE, gene_mode = FALSE, extra_bootstrap_summary = TRUE)
so$pval_aggregate <- FALSE
so <- sleuth_fit(so, ~sibling_group+stage+condition, 'full')
so <- sleuth_fit(so, ~sibling_group+stage, 'reduced')
so <- sleuth_lrt(so, 'reduced', 'full')
so <- sleuth_wt(so, "conditionF")
so <- sleuth_wt(so, "conditionM")
so <- sleuth_wt(so, "conditionMF")
# view results in shiny app
sleuth_live(so)

# differential gene expression analysis for each developmental stage - filter for required stage and
generate sleuth object
# stage 14 #
samples_S14 <- filter(samples, stage=='_S14')
so_S14 <- sleuth_prep(samples_S14, target_mapping=transcripts, aggregation_column =
'gene_id', read_bootstrap_tpm = TRUE)
so_S14 <- sleuth_fit(so_S14, ~sibling_group+condition, 'full')
so_S14 <- sleuth_fit(so_S14, ~sibling_group, 'reduced')
so_S14 <- sleuth_lrt(so_S14, 'reduced', 'full')
```

```
# save wald test results
so_S14 <- sleuth_wt(so_S14, "conditionF")
results_table_S14_conditionF<- sleuth_results(so_S14, 'conditionF')
write.table(results_table_S14_conditionF,file="3reps_sleuth_wald_test_S14_conditionFvsC.tsv",
sep="\t", row.names=FALSE)
so_S14 <- sleuth_wt(so_S14, "conditionM")
results_table_S14_conditionM<- sleuth_results(so_S14, 'conditionM')
write.table(results_table_S14_conditionM,file="3reps_sleuth_wald_test_S14_conditionM_vsC.tsv
", sep="\t", row.names=FALSE)
so_S14 <- sleuth_wt(so_S14, "conditionMF")
results_table_S14_conditionMF<- sleuth_results(so_S14, 'conditionMF')
write.table(results_table_S14_conditionMF,file="3reps_sleuth_wald_test_S14_conditionMF_vsC.t
sv", sep="\t", row.names=FALSE)

#  stage 20 #
samples_S20 <- filter(samples, stage=='S20')
so_S20 <- sleuth_prep(samples_S20, target_mapping=transcripts, aggregation_column =
'gene_id', read_bootstrap_tpm = TRUE)
so_S20 <- sleuth_fit(so_S20, ~sibling_group+condition, 'full')
so_S20 <- sleuth_fit(so_S20, ~sibling_group, 'reduced')
so_S20 <- sleuth_lrt(so_S20, 'reduced', 'full')

so_S20 <- sleuth_wt(so_S20, "conditionF")
results_table_S20_conditionF<- sleuth_results(so_S20, 'conditionF')
write.table(results_table_S20_conditionF,file="3reps_sleuth_wald_test_S20_conditionFvsC.tsv",
sep="\t", row.names=FALSE)
so_S20 <- sleuth_wt(so_S20, "conditionM")
results_table_S20_conditionM<- sleuth_results(so_S20, 'conditionM')
write.table(results_table_S20_conditionM,file="3reps_sleuth_wald_test_S20_conditionM_vsC.tsv
", sep="\t", row.names=FALSE)
so_S20 <- sleuth_wt(so_S20, "conditionMF")
results_table_S20_conditionMF<- sleuth_results(so_S20, 'conditionMF')
write.table(results_table_S20_conditionMF,file="3reps_sleuth_wald_test_S20_conditionMF_vsC.t
sv", sep="\t", row.names=FALSE)
```

```
# stage 30 #
samples_S30 <- filter(samples, stage=='S30')
so_S30 <- sleuth_prep(samples_S30, target_mapping=transcripts, aggregation_column =
'gene_id', read_bootstrap_tpm = TRUE)
so_S30 <- sleuth_fit(so_S30, ~sibling_group+condition, 'full')
so_S30 <- sleuth_fit(so_S30, ~sibling_group, 'reduced')
so_S30 <- sleuth_lrt(so_S30, 'reduced', 'full')

so_S30 <- sleuth_wt(so_S30, "conditionF")
results_table_S30_conditionF<- sleuth_results(so_S30, 'conditionF')
write.table(results_table_S30_conditionF,file="3reps_sleuth_wald_test_S30_conditionFvsC.tsv",
sep="\t", row.names=FALSE)
so_S30 <- sleuth_wt(so_S30, "conditionM")
results_table_S30_conditionM<- sleuth_results(so_S30, 'conditionM')
write.table(results_table_S30_conditionM,file="3reps_sleuth_wald_test_S30_conditionM_vsC.tsv
", sep="\t", row.names=FALSE)
so_S30 <- sleuth_wt(so_S30, "conditionMF")
results_table_S30_conditionMF<- sleuth_results(so_S30, 'conditionMF')
write.table(results_table_S30_conditionMF,file="3reps_sleuth_wald_test_S30_conditionMF_vsC.t
sv", sep="\t", row.names=FALSE)
```

Abbreviations

Act-	Actin
Actn-	Actinin
Acvr-	Activin receptor
AP-1	Activator protein-1
APC	Adenomatous polyposis coli
Ascl1/MASH1	Achaete-schute family bHLH transcription factor 1
bHLH	Basic Helix-Loop-Helix
BMP	Bone morphogenetic protein
Bmpr-	Bone morphogenetic protein receptor
Casq	Calsequestrin
Cdh-	Cadherin
Cdx-	Caudal type homeobox
Celf-	CUGBP Elav-like family member
CHIR/Chiron	CHIR99021
CHX	Cycloheximide
Ckm	Creatine kinase muscle
csnk1k1/CK1	Casein kinase 1
Ctnnb-	β-catenin
Des	Desmin
Dhh	Desert hedgehog
Dkk-	Dickkopf
Dlc	Delta C
Dll	Delta
dpERK	Diphosphorylated extracellular signal-regulated kinase
Dusp6	Dual-specificity phosphatase 6
Dvl	Dishevelled
Elk-1	E26 transformation-specific like-1
ERK	Extracellular signal-regulated kinase
ES cells	Embryonic stem cells
ETS	E26 transformation-specific transcription factors
Etv-	E26 transformation-specific variant
Eya-	Eya transcriptional coactivator and phosphatase
FGF	Fibroblast growth factor
FGFR	Fibroblast growth factor receptor
FGFRL	Fibroblast growth factor receptor like
Foxd3	Forkhead box d3
FRS2	Fibroblast growth factor receptor substrate 2
Fzd-	Frizzled
Gab1	Growth factor receptor bound protein 2 associated protein 1
GDP	Guanosine diphosphate
GO	Gene ontology
Grb2	Growth factor receptor bound protein 2
Grg4/Tle4	Groucho
GSK3	Glycogen synthase kinase 3

GTP	Guanosine triphosphate
HDAC	Histone deacetyltransferase
Hh	Hedgehog
HMG-box	High Mobility Group-box
HSPG	Heparin sulphate proteoglycan
Ihh	Indian hedgehog
Inhba/b	Inhibin subunit beta a/b
IP_3	Inositol-1,4,5-trisphosphate
Krt-	Keratin
LRP5/6	Low density lipoprotein receptor-related protein 5/6
MAPK	Mitogen-activated protein kinase
Mef2-	Myocyte enhancer factor 2
Meis-	Meis homeobox
Mek	Mitogen-activated protein kinase/ERK kinase
Mex3b	Mex-3 RNA binding family member B
MHC	Sarcomere myosin heavy chain
MRF	Myogenic regulatory factor
Mrf4/Myf6	Myogenic factor 6
Myf5	Myogenic factor 5
Myh-	Myosin heavy chain
Myl-	Myosin light chain
Mymk	Myomaker
Myod1	Myogenic differentiation factor 1
Myog	Myogenin
Myoz	Myozenin/Calsarcin
Myt1	Myelin transcription factor 1
NAM	Normal amphibian medium
Ncam	Neural cell adhesion molecule
Neb	Nebulin
Nes	Nestin
Neurod-	Neuronal differentiation
Neurog-	Neurogenin
Nrp-	Neuropilin
p90RSK	p90 ribosomal S6 kinase
PANTHER	Protein Analysis Through Evoluntionary Relationships
Pax-	Paired box
Pbx-	Pre-B-cell leukemia homeobox
PCA	Principle component analysis
PCAF	p300/CREB-Binding Protein associated factor
PEA3	Polyoma enhancer activator 3
PI3K	Phosphoinositide-3 kinase
PKC	Protein kinase C
PLCγ	Phospholipase Cγ
Pol II	RNA polymerase II
Pygm	Glycogen phosphorylase
qPCR	Quantitative real-time polymerase chain reaction

Rbfox-	RNA-binding fox
RIN	RNA integrity number
RNA-seq	Ribonucleic acid sequencing
RTK	Receptor tyrosine kinase
Sema-	Semaphorin
SH2/3	Src homology domain 2/3
Shh	Sonic hedgehog
Six-	Sine oculis homeobox
Smad-	Mothers against decapentaplegic
Smyd1	SET and MYND domain containing 1
Snai2	Snail family transcriptional repressor 2
SOS	Son of sevenless
Sox-	SRY-related HMG-box
Sufu	Suppressor of fused
SWI/SNF	Switch/Sucrose non-fermentable
Tbxt/Xbra	Brachyury
Tcf-	Transcription factor
TCF/LEF	T-cell factor/LEF
Tfap2a	Transcription factor AP-2 alpha
Tgfbr-	Transforming growth factor β receptor
TGFβ/tgfb-	Transforming growth factor β
Tnn-	Troponin
Tpm	Transcripts per million
Tpm-	Tropomyosin
Ttn	Titin
Tubb2b	Tubulin Beta 2B Class IIb
Wnt-	Wingless
X. laevis	*Xenopus laevis*
X. tropicalis	*Xenopus tropicalis*
XFD/dnFGFR	Dominant negative truncated FGF receptor

References

Affolter, M. and Basler, K. (2007). The Decapentaplegic morphogen gradient: from pattern formation to growth regulation. *Nature Reviews Genetics 2007 8:9*, 8 (9), pp.663–674.

Agius, E. et al. (2000). Endodermal Nodal-related signals and mesoderm induction in Xenopus. *Development (Cambridge, England)*, 127 (6), p.1173.

Almada, A. E. et al. (2021). FOS licenses early events in stem cell activation driving skeletal muscle regeneration. *Cell Reports*, 34 (4), p.108656.

Alvarez, I. S., Araujo, M. and Nieto, M. A. (1998). Neural Induction in Whole Chick Embryo Cultures by FGF. *Developmental Biology*, 199 (1), pp.42–54.

Amaya, E. et al. (1993). FGF signalling in the early specification of mesoderm in Xenopus. *Development*, 118 (2), pp.477–487.

Amaya, E., Musci, T. J. and Kirschner, M. W. (1991). Expression of a dominant negative mutant of the FGF receptor disrupts mesoderm formation in Xenopus embryos. *Cell*, 66 (2), pp.257–270.

Amit, S. et al. (2002). Axin-mediated CKI phosphorylation of β-catenin at Ser 45: a molecular switch for the Wnt pathway. *Genes & Development*, 16 (9), p.1066.

Anderson, C. et al. (2012). Sonic hedgehog acts cell-autonomously on muscle precursor cells to generate limb muscle diversity. *Genes & Development*, 26 (18), pp.2103–2117.

Angerilli, A., Smialowski, P. and Rupp, R. A. W. (2018). The Xenopus animal cap transcriptome: Building a mucociliary epithelium. *Nucleic Acids Research*, 46 (17), pp.8772–8787.

Ariizumi, T. et al. (1991). Concentration-dependent inducing activity of activin A. *Roux's Archives of Developmental Biology*, 200 (4), pp.230–233.

Ariizumi, T. et al. (2009). Isolation and differentiation of Xenopus animal cap cells. *Current Protocols in Stem Cell Biology*, 9 (1).

Ariizumi, T., Michiue, T. and Asashima, M. (2017). In vitro induction of Xenopus embryonic organs using animal cap cells. *Cold Spring Harbor Protocols*, 2017 (12), pp.982–987.

Armand, A. S. et al. (2003). Injection of FGF6 accelerates regeneration of the soleus muscle in adult mice. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1642 (1–2), pp.97–105.

Armand, A. S. et al. (2005). FGF6 regulates muscle differentiation through a calcineurin-dependent pathway in regenerating soleus of adult mice. *Journal of Cellular Physiology*, 204 (1), pp.297–308.

Asashima, M. et al. (1990). Mesodermal induction in early amphibian embryos by activin A (erythroid differentiation factor). *Roux's Archives of Developmental Biology*, 198 (6), pp.330–335.

Aybar, M. J., Nieto, M. A. and Mayor, R. (2003). Snail precedes Slug in the genetic cascade required for the specification and migration of the Xenopus neural crest. *Development*, 130 (3), pp.483–494.

Barbosa-Morais, N. L. et al. (2012). The evolutionary landscape of alternative splicing in vertebrate species. *Science*, 338 (6114), pp.1587–1593.

Barndt, R. J., Dai, M. and Zhuang, Y. (2000). Functions of E2A-HEB Heterodimers in T-Cell Development Revealed by a Dominant Negative Mutation of HEB. *Molecular and Cellular Biology*, 20 (18), pp.6677–6685.

Basson, M. A. (2012). Signaling in cell differentiation and morphogenesis. *Cold Spring Harb. Perspect. Biol.*, 4 (6).

Beck, C. W. and Slack, J. M. (2001). An amphibian with ambition: a new role for Xenopus in the 21st century. *Genome Biology 2001 2:10*, 2 (10), pp.1–5.

Beenken, A. and Mohammadi, M. (2009). The FGF family: biology, pathophysiology and therapy. *Nat. Rev. Drug Discov.*, 8 (3), pp.235–253.

Behrens, J. et al. (1996). Functional interaction of β-catenin with the transcription factor LEF- 1. *Nature*, 382 (6592), pp.638–642.

Belgacem, Y. H. et al. (2016). The Many Hats of Sonic Hedgehog Signaling in Nervous System Development and Disease. *Journal of Developmental Biology*, 4 (4).

Belle, I. and Zhuang, Y. (2014). E Proteins in Lymphocyte Development and Lymphoid Diseases. *Current Topics in Developmental Biology*, 110, pp.153–187.

Bellefroid, E. J. et al. (1996). X-MyT1, a xenopus C2HC-type zinc finger protein with a regulatory function in neuronal differentiation. *Cell*, 87 (7), pp.1191–1202.

Bengal, E. et al. (1994). Positive control mutations in the MyoD basic region fail to show cooperative DNA binding and transcriptional activation in vitro. *Proc. Natl. Acad. Sci. U. S. A.*, 91 (13), pp.6221–6225.

Béranger, F. et al. (2000). Muscle Differentiation Is Antagonized by SOX15, a New Member of the SOX Protein Family. *Journal of Biological Chemistry*, 275 (21), pp.16103–16109.

Bergstrom, D. A. et al. (2002). Promoter-specific regulation of MyoD binding and signal transduction cooperate to pattern gene expression. *Molecular Cell*, 9 (3), pp.587–600.

Berkes, C. A. et al. (2004). Pbx marks genes for activation by MyoD indicating a role for a homeodomain protein in establishing myogenic potential. *Mol. Cell*, 14 (4), pp.465–477.

Bernet, J. D. et al. (2014). P38 MAPK signaling underlies a cell autonomous loss of stem cell self-renewal in aged skeletal muscle. *Nature medicine*, 20 (3), p.265.

Bertrand, V. et al. (2003). Neural Tissue in Ascidian Embryos Is Induced by FGF9/16/20, Acting via a Combination of Maternal GATA and Ets Transcription Factors. *Cell*, 115 (5), pp.615–627.

Bhanot, P. et al. (1996). A new member of the frizzled family from Drosophila functions as a Wingless receptor identified . We show here that cultured Drosophila cells transfected with a novel member of the. *Nature*, 382, pp.225–230.

Biesiada, E. et al. (1999). Myogenic basic helix-loop-helix proteins and Sp1 interact as components of a multiprotein transcriptional complex required for activity of the human cardiac alpha-actin promoter. *Mol. Cell. Biol.*, 19 (4), pp.2577–2584.

Birsoy, B. et al. (2006). Vg1 is an essential signaling molecule in Xenopusdevelopment. *Development*, 133 (1), pp.15–20.

Bischoff, R. and Heintz, C. (1994). Enhancement of skeletal muscle regeneration. *Developmental Dynamics*, 201 (1), pp.41–54.

Black, B. L. et al. (1996). Cooperative transcriptional activation by the neurogenic basic helix- loop-helix protein MASH1 and members of the myocyte enhancer factor-2 (MEF2) family. *Journal of Biological Chemistry*, 271 (43), pp.26659–26663.

Blais, A. et al. (2005). An initial blueprint for myogenic differentiation. *Genes & Development*, 19 (5), p.553.

Blum, M. and Ott, T. (2018). Xenopus: An Undervalued Model Organism to Study and Model Human Genetic Disease. *Cells Tissues Organs*, 205 (5–6), pp.303–313.

Bober, E. et al. (1991). The muscle regulatory gene, Myf-6, has a biphasic pattern of expression during early mouse development. *Journal of Cell Biology*, 113 (6), pp.1255–1265.

Borchers, A. and Pieler, T. (2010). Programming Pluripotent Precursor Cells Derived from Xenopus Embryos to Generate Specific Tissues and Organs. *Genes 2010, Vol. 1, Pages 413-426*, 1 (3), pp.413–426.

Borchin, B., Chen, J. and Barberi, T. (2013). Derivation and FACS-Mediated Purification of PAX3+/PAX7+ Skeletal Muscle Precursors from Human Pluripotent Stem Cells. *Stem Cell Reports*, 1 (6), p.620.

Boterenbrood, E. C. and Nieuwkoop, P. D. (1973). The formation of the mesoderm in urodelean amphibians - V. Its regional induction by the endoderm. *Wilhelm Roux' Archiv für Entwicklungsmechanik der Organismen*, 173 (4), pp.319–332.

Böttcher, R. T. and Niehrs, C. (2005). Fibroblast growth factor signaling during early vertebrate development. *Endocr. Rev.*, 26 (1), pp.63–77.

Boulet, A. M. and Capecchi, M. R. (2012). Signaling by FGF4 and FGF8 is required for axial elongation of the mouse embryo. *Developmental Biology*, 371 (2), pp.235–245.

Branney, P. A. et al. (2009). Characterisation of the fibroblast growth factor dependent transcriptome in early development. *PLoS One*, 4 (3), p.e4951.

Braun, T. et al. (1989). A novel human muscle factor related to but distinct from MyoD1 induces myogenic conversion in 10T1/2 fibroblasts. *The EMBO Journal*, 8 (3), pp.701–709.

van den Brink, S. C. et al. (2020). Single-cell and spatial transcriptomics reveal somitogenesis in gastruloids. *Nature*.

Brunsdon, H. and Isaacs, H. V. (2020). A comparative analysis of fibroblast growth factor receptor signalling during Xenopus development. *Biology of the Cell*, 112 (5), pp.127–139.

Bucher, E. A. et al. (1988). Expression of the troponin complex genes: transcriptional coactivation during myoblast differentiation and independent control in heart and skeletal muscles. *Molecular and Cellular Biology*, 8 (10), p.4134.

Buckingham, M. and Relaix, F. (2007). The Role of Pax Genes in the Development of Tissues and Organs: Pax3 and Pax7 Regulate Muscle Progenitor Cell Functions. *https://doi.org/10.1146/annurev.cellbio.23.090506.123438*, 23, pp.645–673.

Burgess, R. et al. (1996). Requirement of the paraxis gene for somite formation and musculoskeletal patterning. *Nature 1996 384:6609*, 384 (6609), pp.570–573.

Burks, P. J., Isaacs, H. V and Pownall, M. E. (2009). FGF signalling modulates transcriptional repression by Xenopus groucho-related-4. *Biol. Cell*, 101 (5), pp.301–308.

Capetanaki, Y., Milner, D. J. and Weitzer, G. (1997). Desmin in muscle formation and maintenance: Knockouts and consequences. *Cell Structure and Function*, 22 (1), pp.103–116.

Carapuço, M. et al. (2005). Hox genes specify vertebral types in the presomitic mesoderm. *Genes & Development*, 19 (18), p.2116.

Carballo, G. B. et al. (2018). A highlight on Sonic hedgehog pathway. *Cell Communication and Signaling*, 16 (1), pp.1–15.

Cargnello, M. and Roux, P. P. (2011). Activation and Function of the MAPKs and Their Substrates, the MAPK-Activated Protein Kinases. *Microbiology and Molecular Biology Reviews*, 75 (1), pp.50–83.

Castro, D. S. et al. (2011). A novel function of the proneural factor Ascl1 in progenitor proliferation identified by genome-wide characterization of its targets. *Genes & Development*, 25 (9), pp.930–945.

Chal, J. et al. (2015). Differentiation of pluripotent stem cells to muscle fiber to model Duchenne muscular dystrophy. *Nature Biotechnology*, 33 (9), pp.962–969.

Chambers, I. et al. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell*, 113 (5), pp.643–655.

Chambers, I. et al. (2007). Nanog safeguards pluripotency and mediates germline development. *Nature 2007 450:7173*, 450 (7173), pp.1230–1234.

Chen, L. et al. (2016). CELF RNA binding proteins promote axon regeneration in C. elegans and mammals through alternative splicing of syntaxins. *eLife*, 5 (JUN2016).

Cho, K. W. Y. and De Robertis, E. M. (1990). Differential activation of Xenopus homeo box genes by mesoderm-inducing growth factors and retinoic acid. *Genes & development*, 4 (11), pp.1910–1916.

Chotteau-Lelièvre, A. et al. (1997). Differential expression patterns of the PEA3 group transcription factors through murine embryonic development. *Oncogene 1997 15:8*, 15 (8), pp.937–952.

Christen, B. and Slack, J. M. (1999). Spatial response to fibroblast growth factor signalling in Xenopus embryos. *Development*, 126 (1), pp.119–125.

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 and Development*, 7 (1), pp.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. *The EMBO Journal*, 11 (1), p.33.

Chung, H. A. et al. (2004). Screening of FGF target genes in Xenopus by microarray: temporal dissection of the signalling pathway using a chemical inhibitor. *Genes Cells*, 9 (8), pp.749–761.

Ciruna, B. and Rossant, J. (2001). FGF Signaling Regulates Mesoderm Cell Fate Specification and Morphogenetic Movement at the Primitive Streak. *Developmental Cell*, 1 (1), pp.37–49.

Conerly, M. L. et al. (2016). Distinct Activities of Myf5 and MyoD Indicate Separate Roles in Skeletal Muscle Lineage Specification and Differentiation. *Dev. Cell*, 36 (4), pp.375–385.

Conway, K. et al. (2004). The E protein HEB is preferentially expressed in developing muscle. *Differentiation*, 72 (7), pp.327–340.

Cooke, J. et al. (1987). The organization of mesodermal pattern in Xenopus laevis: experiments using a Xenopus mesoderm-inducing factor. *Development*, 101 (4), pp.893–908.

Cooper, G. M. (2000). Actin, Myosin, and Cell Movement. In: *The Cell: A Molecular Approach*. 2nd ed. Sinauer Associates.

Cooper, T. A. and Ordahl, C. P. (1985). A single cardiac troponin T gene generates embryonic and adult isoforms via developmentally regulated alternate splicing. *Journal of Biological Chemistry*, 260 (20), pp.11140–11148.

Cornell, R. A. and Kimelman, D. (1994). Activin-mediated mesoderm induction requires FGF. *Development*, 120 (2), pp.453–462.

Cornell, R. A., Musci, T. J. and Kimelman, D. (1995). FGF is a prospective competence factor for early activin-type signals in Xenopus mesoderm induction. *Development*, 121 (8), pp.2429–2437.

Cosgrove, B. D. et al. (2014). Rejuvenation of the aged muscle stem cell population restores strength to injured aged muscles. *Nature medicine*, 20 (3), p.255.

Cowell, L. M. (2019). *Investigation of the role of Capicua in the FGF signalling pathway and the wound response*. MSc by Research thesis, University of York.

Cox, C. M. and Poole, T. J. (2000). Angioblast Differentiation Is Influenced by the Local Environment: FGF-2 Induces Angioblasts and Patterns Vessel Formation in the Quail Embryo. *Dev Dyn*, 218, pp.371–382.

Dall'agnese, A. et al. (2019). Transcription Factor-Directed Re-wiring of Chromatin Architecture for Somatic Cell Nuclear Reprogramming toward trans-Differentiation. *Molecular Cell*, 76, pp.453–472.

Damaraju, S. et al. (2012). Considerations on Dealing with Tissues and Cell Samples (Include Tissue Banking). *Comprehensive Sampling and Sample Preparation: Analytical Techniques for Scientists*, pp.21–31.

Das, D. et al. (2007). A correlation with exon expression approach to identify cis-regulatory elements for tissue-specific alternative splicing. *Nucleic Acids Research*, 35 (14), p.4845.

Daume, D. et al. (2022). Patterns of tubb2b Promoter-Driven Fluorescence in the Forebrain of Larval Xenopus laevis. *Frontiers in Neuroanatomy*, 16, p.914281.

Davies, O. R. et al. (2013). Tcf15 Primes Pluripotent Cells for Differentiation. *Cell Reports*, 3 (2), pp.472–484.

Davis, R. L., Weintraub, H. and Lassar, A. B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell*, 51 (6), pp.987–1000.

Dekel, I. et al. (1992). Conditional conversion of ES cells to skeletal muscle by an exogenous MyoD1 gene. *The New Biologist*, 4 (3), pp.217–224.

Delaune, E., Lemaire, P. and Kodjabachian, L. (2005). Neural induction in Xenopus requires early FGF signalling in addition to BMP inhibition. *Development*, 132 (2), pp.299–310.

Denef, N. et al. (2000). Hedgehog induces opposite changes in turnover and subcellular localization of patched and smoothened. *Cell*, 102 (4), pp.521–531.

Denis Gospodarowicz, J. S. M. (1975). Initiation of DNA Synthesis in Human Foreskin Fibroblasts Cell Culture Quantitative Determination of [3H] Thymidine Incorporation. *The Journal of Cell Biology*, 66, pp.451–457.

Derynck, R. and Zhang, Y. E. (2003). *Smad-dependent and Smad-independent pathways in TGF-*. 4.

van Deursen, J. et al. (1993). Skeletal muscles of mice deficient in muscle creatine kinase lack burst activity. *Cell*, 74 (4), pp.621–631.

Devlin, R. B. and Emerson, C. P. (1978). Coordinate regulation of contractile protein synthesis during myoblast differentiation. *Cell*, 13 (4), pp.599–611.

Dichmann, D. S., Walentek, P. and Harland, R. M. (2015). The alternative splicing regulator Tra2b is required for somitogenesis and regulates splicing of an inhibitory Wnt11b isoform. *Cell reports*, 10 (4), p.527.

Dissanayake, K. et al. (2011). ERK/p90(RSK)/14-3-3 signalling has an impact on expression of PEA3 Ets transcription factors via the transcriptional repressor capicúa. *Biochem. J*, 433 (3), pp.515–525.

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. *Mechanisms of Development*, 120 (9), pp.1021–1032.

Dorey, K. and Amaya, E. (2010). FGF signalling: diverse roles during early vertebrate embryogenesis. *Development*, 137 (22), pp.3731–3742.

Doyle, K. and Fitzpatrick, F. A. (2010). Redox signaling, alkylation (carbonylation) of conserved cysteines inactivates class I histone deacetylases 1, 2, and 3 and antagonizes their transcriptional repressor function. *Journal of Biological Chemistry*, 285 (23), pp.17417–17424.

Dumont, N. A. et al. (2015). Dystrophin expression in muscle stem cells regulates their polarity and asymmetric division. *Nature Medicine*, 21 (12), pp.1455–1463.

Ekerot, M. et al. (2008). Negative-feedback regulation of FGF signalling by DUSP6/MKP-3 is driven by ERK1/2 and mediated by Ets factor binding to a conserved site within the DUSP6/MKP-3 gene promoter. *Biochem. J*, 412 (2), pp.287–298.

Elsy, M. et al. (2019). Xenopus laevis FGF16 activates the expression of genes coding for the transcription factors Sp5 and Sp5l. *International Journal of Developmental Biology*, 63 (11–12), pp.631–639.

Esch, F. et al. (1985). Primary structure of bovine pituitary basic fibroblast growth factor (FGF) and comparison with the amino-terminal sequence of bovine brain acidic FGF. *Proceedings of the National Academy of Sciences of the United States of America*, 82 (19), pp.6507–6511.

Faas, L. et al. (2013). Lin28 proteins are required for germ layer specification in Xenopus. *Development*, 140 (5), pp.976–986.

Fagerberg, L. et al. (2014). Analysis of the Human Tissue-specific Expression by Genome-wide Integration of Transcriptomics and Antibody-based Proteomics. *Molecular & Cellular Proteomics : MCP*, 13 (2), p.397.

Fan, X. et al. (2023). TWIST1 Homodimers and Heterodimers Orchestrate Lineage-Specific Differentiation. *https://doi.org/10.1128/MCB.00663-19*, 40 (11).

Fernández-Lázaro, D. et al. (2022). Potential Satellite Cell-Linked Biomarkers in Aging Skeletal Muscle Tissue: Proteomics and Proteogenomics to Monitor Sarcopenia. *Proteomes 2022, Vol. 10, Page 29*, 10 (3), p.29.

Fiore, F., Sébille, A. and Birnbaum, D. (2000). Skeletal Muscle Regeneration Is Not Impaired in Fgf6 −/− Mutant Mice. *Biochemical and Biophysical Research Communications*, 272 (1), pp.138–143.

Fisher, M. E. et al. (2003). Cloning and characterisation of Myf5 and MyoD orthologues in Xenopus tropicalis. *Biology of the Cell*, 95 (8), pp.555–561.

Fisher, M. E., Isaacs, H. V and Pownall, M. E. (2002). eFGF is required for activation of XmyoD expression in the myogenic cell lineage of Xenopus laevis. *Development*, 129 (6), pp.1307–1315.

Fletcher, R. B., Baker, J. C. and Harland, R. M. (2006). FGF8 spliceforms mediate early mesoderm and posterior neural tissue formation in Xenopus. *Development*, 133 (9), pp.1703–1714.

Fletcher, R. B. and Harland, R. M. (2008). The role of FGF signaling in the establishment and maintenance of mesodermal gene expression in Xenopus. *Dev. Dyn.*, 237 (5), pp.1243–1254.

Floss, T., Arnold, H. H. and Braun, T. (1997). A role for FGF-6 in skeletal muscle regeneration. *Genes & Development*, 11 (16), p.2040.

Fong, A. P. et al. (2012). Genetic and Epigenetic Determinants of Neurogenesis and Myogenesis. *Developmental Cell*, 22 (4), p.721.

Fong, A. P. et al. (2015). Conversion of MyoD to a neurogenic factor: binding site specificity determines lineage. *Cell reports*, 10 (12), p.1937.

Fong, A. P. and Tapscott, S. J. (2013). Skeletal muscle programming and re-programming. *Current Opinion in Genetics and Development*, 23 (5), Elsevier Current Trends., pp.568–573.

Frederikson, K. and McKay, R. D. G. (1988). Proliferation and differentiation of rat neuroepithelial precursor cells in vivo. *Journal of Neuroscience*, 8 (4), pp.1144–1151.

Frey, N., Richardson, J. A. and Olson, E. N. (2000). Calsarcins, a novel family of sarcomeric calcineurin-binding proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 97 (26), p.14632.

Furst, D. O., Osborn, M. and Weber, K. (1989). Myogenesis in the mouse embryo: differential onset of expression of myogenic proteins and the involvement of titin in myofibril assembly. *Journal of Cell Biology*, 109 (2), pp.517–527.

Fürthauer, M. et al. (2004). Fgf signalling controls the dorsoventral patterning of the zebrafish embryo. *Development*, 131 (12), pp.2853–2864.

Galimov, A. et al. (2016). MicroRNA-29a in Adult Muscle Stem Cells Controls Skeletal Muscle Regeneration During Injury and Exercise Downstream of Fibroblast Growth Factor-2. *Stem Cells*, 34 (3), pp.768–780.

Gallagher, T. L. et al. (2011). Rbfox-regulated alternative splicing is critical for zebrafish cardiac and skeletal muscle functions. *Developmental Biology*, 359 (2), pp.251–261.

Gao, C. and Chen, Y. G. (2010). Dishevelled: The hub of Wnt signaling. *Cellular Signalling*, 22 (5), pp.717–727.

Gao, M. et al. (2021). circHIPK3 regulates proliferation and differentiation of myoblast through the miR-7/TCF12 pathway. *Journal of Cellular Physiology*, 236 (10), pp.6793–6805.

Garg, A. et al. (2018). FGF-induced Pea3 transcription factors program the genetic landscape for cell fate determination. *PLoS Genet.*, 14 (9), p.e1007660.

Garriock, R. J. et al. (2015). Lineage tracing of neuromesodermal progenitors reveals novel wnt-dependent roles in trunk progenitor cell maintenance and differentiation. *Development (Cambridge)*, 142 (9), pp.1628–1638.

Gaspera, B. Della et al. (2006). Spatio-temporal expression of MRF4 transcripts and protein during Xenopus laevis embryogenesis. *Developmental Dynamics*, 235 (2), pp.524–529.

Gazzara, M. R. et al. (2017). Ancient antagonism between CELF and RBFOX families tunes mRNA splicing outcomes. *Genome Research*, 27 (8), pp.1360–1370.

Geary, L. and Labonne, C. (2018). FGF mediated MAPK and PI3K/Akt Signals make distinct contributions to pluripotency and the establishment of Neural Crest. *eLife*, 7.

Gehman, L. T. et al. (2012). The splicing regulator Rbfox2 is required for both cerebellar development and mature motor function. *Genes & Development*, 26 (5), p.445.

Gilbert, S. F. (2000). Developmental Biology. *Development Biology: 6th edition*, (Hall 1988), p.http://www.ncbi.nlm.nih.gov/books/NBK10056/.

Gioftsidi, S., Relaix, F. and Mourikis, P. (2022). The Notch signaling network in muscle stem cells during development, homeostasis, and disease. *Skeletal Muscle 2022 12:1*, 12 (1), pp.1–12.

Glinka, A. et al. (1998). Dickkopf-1 is a member of a new family of secreted proteins and functions in head induction. *Nature 1998 391:6665*, 391 (6665), pp.357–362.

Godsave, S. F., Isaacs, H. V. and Slack, J. M. W. (1988). Mesoderm-inducing factors: a small class of molecules. *Development*, 102 (3), pp.555–566.

Godsave, S. F. and Slack, J. M. W. (1989). Clonal analysis of mesoderm induction in Xenopus laevis. *Developmental Biology*, 134 (2), pp.486–490.

Gordon, K. J. and Blobe, G. C. (2008). Role of transforming growth factor-β superfamily signaling pathways in human disease. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1782 (4), pp.197–228.

Gospodarowicz, D. (1975). Purification of a fibroblast growth factor from bovine pituitary. *Journal of Biological Chemistry*, 250 (7), pp.2515–2520.

Gossett, L. A. et al. (1989). A new myocyte-specific enhancer-binding factor that recognizes a conserved element associated with multiple muscle-specific genes. *Molecular and cellular biology*, 9 (11), pp.5022–5033.

Goto, K. et al. (2007). Selective inhibitory effects of Smad6 on bone morphogenetic protein type I receptors. *Journal of Biological Chemistry*, 282 (28), pp.20603–20611.

Green, J. B. A. et al. (1990). The biological effects of XTC-MIF: quantitative comparison with Xenopus bFGF. *Development*, 108 (1), pp.173–183.

Green, J. and Guille, M. (1999). *The Animal Cap Assay in Molecular Methods in Developmental Biology: Xenopus and Zebrafish*. Springer Science & Business Media.

Grimaldi, A. et al. (2004). Hedgehog regulation of superficial slow muscle fibres in Xenopusand the evolution of tetrapod trunk myogenesis. *Development*, 131 (14), pp.3249–3262.

Groves, J. A., Hammond, C. L. and Hughes, S. M. (2005). Fgf8 drives myogenic progression of a novel lateral fast muscle fibre population in zebrafish. *Development*, 132 (19), pp.4211–4222.

Grunz, H. and Tacke, L. (1989). Neural differentiation of Xenopus laevis ectoderm takes place after disaggregation and delayed reaggregation without inducer. *Cell differentiation and development : the official journal of the International Society of Developmental Biologists*, 28 (3), pp.211–217.

Gubbay, J. et al. (1990). A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. *Nature 1990 346:6281*, 346 (6281), pp.245–250.

Gunning, P., O'Neill, G. and Hardeman, E. (2008). Tropomyosin-based regulation of the actin cytoskeleton in time and space. *Physiological Reviews*, 88 (1), pp.1–35.

Gurdon, J. B. et al. (1985). Actin genes in Xenopus and their developmental control. *Development*, 89 (Supplement).

Gurdon, J. B. et al. (1993). A community effect in muscle development. *Current Biology*, 3 (1), pp.1–11.

Guzzetta, A. et al. (2020). Hedgehog–FGF signaling axis patterns anterior mesoderm during gastrulation. *PNAS*, 117 (27), pp.15712–15723.

Hadari, Y. R. et al. (2001). Critical role for the docking-protein FRS2 alpha in FGF receptor-mediated signal transduction pathways. *Proc. Natl. Acad. Sci. U. S. A.*, 98 (15), pp.8578–8583.

Haldar, M. et al. (2008). Two Cell Lineages, myf5 and myf5-Independent, Participate in Mouse Skeletal Myogenesis. *Developmental cell*, 14 (3), p.437.

Hamilton, W. B. et al. (2019). Dynamic lineage priming is driven via direct enhancer regulation by ERK. *Nature*, 575 (7782), pp.355–360.

Han, J. K. and Martin, G. R. (1993). Embryonic Expression of Fgf-6 Is Restricted to the Skeletal Muscle Lineage. *Developmental Biology*, 158 (2), pp.549–554.

Hanyu, A. et al. (2001). The N domain of Smad7 is essential for specific inhibition of transforming growth factor-β signaling. *The Journal of Cell Biology*, 155 (6), p.1017.

Harrison, S. M. et al. (2000). Sp5, a New Member of the Sp1 Family, Is Dynamically Expressed during Development and Genetically Interacts with Brachyury. *Developmental Biology*, 227 (2), pp.358–372.

Hasty, P. et al. (1993). Muscle deficiency and neonatal death in mice with a targeted mutation in the myogenin gene. *Nature*, 364, pp.501–506.

Heasman, J. et al. (1994). Overexpression of cadherins and underexpression of β-catenin inhibit dorsal mesoderm induction in early Xenopus embryos. *Cell*, 79 (5), pp.791–803.

Hellsten, U. et al. (2010). The Genome of the Western Clawed Frog Xenopus tropicalis. *Science (New York, N.Y.)*, 328 (5978), p.633.

Hernández-Hernández, J. M. et al. (2017). The myogenic regulatory factors, determinants of muscle development, cell identity and regeneration. *Seminars in Cell & Developmental Biology*, 72, pp.10–18.

Herrmann, B. G. et al. (1990). Cloning of the T gene required in mesoderm formation in the mouse. *Nature 1990 343:6259*, 343 (6259), pp.617–622.

Hess, A. (1970). Vertebrate slow muscle fibers. *Physiological Reviews*, 50 (1), pp.40–62.

Hinterberger, T. J. et al. (1991). Expression of the muscle regulatory factor MRF4 during somite and skeletal myofiber development. *Developmental biology*, 147 (1), pp.144–156.

Holland, P. W. H. and Garcia-Fernàndez, J. (1996). HoxGenes and Chordate Evolution. *Developmental Biology*, 173 (2), pp.382–395.

Hong, C. S., Park, B. Y. and Saint-Jeannet, J. P. (2008). Fgf8a induces neural crest indirectly through the activation of Wnt8 in the paraxial mesoderm. *Development (Cambridge, England)*, 135 (23), p.3903.

Hongo, I. and Okamoto, H. (2022). FGF/MAPK/Ets signaling in Xenopus ectoderm contributes to neural induction and patterning in an autonomous and paracrine manner, respectively. *Cells & Development*, 170, p.203769.

Hoppler, S., Brown, J. D. and Moon, R. T. (1996). Expression of a dominant-negative Wnt blocks induction of MyoD in Xenopus embryos. *Genes & development*, 10 (21), pp.2805–2817.

Hoppler, S. and Moon, R. T. (1998). BMP-2/-4 and Wnt-8 cooperatively pattern the Xenopus mesoderm. *Mechanisms of Development*, 71 (1–2), pp.119–129.

Hopwood, N. D. et al. (1992). Expression of XMyoD protein in early Xenopus laevis embryos. *Development*, 114 (1), pp.31–38.

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 (6289), pp.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. *The EMBO Journal*, 8 (1), pp.3409–3417.

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 (2), pp.551–560.

Hough, S. R. et al. (2014). Single-Cell Gene Expression Profiles Define Self-Renewing, Pluripotent, and Lineage Primed States of Human Pluripotent Stem Cells. *Stem Cell Reports*, 2 (6), p.881.

Houghton, L. and Rosenthal, N. *Regulation of a Muscle-Specific Transgene by Persistent Expression of Hox Genes in Postnatal Murine Limb Muscle*.

Hu, J. S., Olson, E. N. and Kingston, R. E. (1992). HEB, a helix-loop-helix protein related to E2A and ITF2 that can modulate the DNA-binding ability of myogenic regulatory factors. *Molecular and cellular biology*, 12 (3), pp.1031–1042.

Huang, S. M. A. et al. (2009). Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. *Nature*, 461 (7264), pp.614–620.

Ignatova, V. V. et al. (2020). The rRNA m6A methyltransferase METTL5 is involved in pluripotency and developmental programs. *Genes & Development*, 34 (9–10), pp.715–729.

Isaacs, H. V. (1997). New perspectives on the role of the fibroblast growth factor family in amphibian development. *Cellular and Molecular Life Sciences*, 53 (4), pp.350–361.

Isaacs, H. V, Deconinck, A. E. and Pownall, M. E. (2007). FGF4 regulates blood and muscle specification in Xenopus laevis. *Biol. Cell*, 99 (3), pp.165–173.

Isaacs, H. V, Pownall, M. E. and Slack, J. M. (1994). eFGF regulates Xbra expression during Xenopus gastrulation. *EMBO J.*, 13 (19), pp.4469–4481.

Isaacs, H. V, Pownall, M. E. and Slack, J. M. (1998). Regulation of Hox gene expression and posterior development by the Xenopus caudal homologue Xcad3. *EMBO J.*, 17 (12), pp.3413–3427.

Isaacs, H. V, Tannahill, D. and Slack, J. M. (1992). Expression of a novel FGF in the Xenopus embryo. A new candidate inducing factor for mesoderm formation and anteroposterior specification. *Development*, 114 (3), pp.711–720.

Itoh, N. and Ornitz, D. M. (2004). Evolution of the Fgf and Fgfr gene families. *Trends in Genetics*, 20 (11), pp.563–569.

Jackman, M. R. and Willis, W. T. (1996). Characteristics of mitochondria isolated from type I and type IIb skeletal muscle. *American Journal of Physiology - Cell Physiology*, 270, pp.673–678.

Jambhekar, A., Dhall, A. and Shi, Y. (2019). Roles and regulation of histone methylation in animal development. *Nature Reviews Molecular Cell Biology 2019 20:10*, 20 (10), pp.625–641.

Jen, W. C. et al. (1999). Periodic repression of Notch pathway genes governs the segmentation of Xenopus embryos. *Genes & Development*, 13 (11), p.1486.

Jiang, K. et al. (2018). An intracellular activation of Smoothened that is independent of Hedgehog stimulation in Drosophila. *Journal of Cell Science*, 131 (1).

Jiménez, G., Shvartsman, S. Y. and Paroush, Z. (2012). The Capicua repressor - a general sensor of RTK signaling in development and disease. *J. Cell Sci.*, 125 (Pt 6), pp.1383–1391.

Johnson, D. E. et al. (1991). The human fibroblast growth factor receptor genes: a common structural arrangement underlies the mechanisms for generating receptor forms that differ in their third immunoglobulin domain. *Mol. Cell. Biol.*, 11 (9), pp.4627–4634.

Jones, E. A. and Woodland, H. R. (1987). The development of animal cap cells in Xenopus: a measure of the start of animal cap competence to form mesoderm. *Development*, 101, pp.557–563.

Jones, E., Sargent, T. D. and Dawid, I. B. (1985). Epidermal keratin gene expressed in embryos of Xenopus laevis. *Proceedings of the National Academy of Sciences of the United States of America*, 82 (16), p.5413.

Jurberg, A. D. et al. (2014). Compartment-dependent activities of Wnt3a/β-catenin signaling during vertebrate axial extension. *Developmental Biology*, 394 (2), pp.253–263.

Kalkan, T. et al. (2019). Complementary Activity of ETV5, RBPJ, and TCF3 Drives Formative Transition from Naive Pluripotency. *Cell Stem Cell*, 24 (5), pp.785-801.e7.

Kang, D. E. et al. (2002). Presenilin couples the paired phosphorylation of β-catenin independent of axin: Implications for β-catenin activation in tumorigenesis. *Cell*, 110 (6), pp.751–762.

Kassar-Duchossoy, L. et al. (2004). Mrf4 determines skeletal muscle identity in Myf5:Myod double-mutant mice. *Nature 2004 431:7007*, 431 (7007), pp.466–471.

Kastner, S. et al. (2000). Gene expression patterns of the fibroblast growth factors and their receptors during myogenesis of rat satellite cells. *Journal of Histochemistry and Cytochemistry*, 48 (8), pp.1079–1096.

Kato, K. and Gurdon, J. B. (1993). Single-cell transplantation determines the time when Xenopus muscle precursor cells acquire a capacity for autonomous differentiation. *Proceedings of the National Academy of Sciences of the United States of America*, 90 (4), pp.1310–1314.

Keenan, I. D., Sharrard, R. M. and Isaacs, H. V. (2006). FGF signal transduction and the regulation of Cdx gene expression. *Dev. Biol.*, 299 (2), pp.478–488.

Keller, R. et al. (2000). Mechanisms of convergence and extension by cell intercalation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 355 (1399), pp.897–922.

Kelly, L. E. et al. (2006). Pbx1 and Meis1 regulate activity of the Xenopus laevis Zic3 promoter through a highly conserved region. *Biochemical and Biophysical Research Communications*, 344 (3), pp.1031–1037.

Kimelman, D. and Kirschner, M. (1987). Synergistic induction of mesoderm by FGF and TGF-beta and the identification of an mRNA coding for FGF in the early Xenopus embryo. *Cell*, 51 (5), pp.869–877.

Kjolby, R. A. S. et al. (2019). Integration of Wnt and FGF signaling in the Xenopus gastrula at TCF and Ets binding sites shows the importance of short-range repression by TCF in patterning the marginal zone. *Development*, 146 (15).

Klein, P. S. and Melton, D. A. (1996). A molecular mechanism for the effect of lithium on development. *Proceedings of the National Academy of Sciences of the United States of America*, 93 (16), p.8455.

Knoepfler, P. S. et al. (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 (18), pp.3752–3761.

Kobel, H. R. and Du Pasquier, L. (1986). Genetics of polyploid Xenopus. *Trends in Genetics*, 2 (C), pp.310–315.

Kofron, M. et al. (1999). Mesoderm induction in Xenopus is a zygotic event regulated by maternal VegT via TGFbeta growth factors. *Development*, 126 (24), pp.5759–5770.

Kogerman, P. et al. (1999). Mammalian Suppressor-of-Fused modulates nuclear – cytoplasmic shuttling of GLI-1. *Nature*, 1 (September), pp.312–319.

Kondo, M. et al. (2017). Comprehensive analyses of hox gene expression in Xenopus laevis embryos and adult tissues. *Development, Growth & Differentiation*, 59 (6), pp.526–539.

Kondo, M. et al. (2019). De novo transcription of multiple Hox cluster genes takes place simultaneously in early Xenopus tropicalis embryos. *Biology Open*, 8 (3).

Konigsberg, I. R. (1963). Clonal analysis of myogenesis. *Science*, 140 (3573), pp.1273–1284.

Kopan, R. (2012). Notch Signaling. *Cold Spring Harbor Perspectives in Biology*, 4 (10).

Korn, E. D. (2000). Coevolution of head, neck, and tail domains of myosin heavy chains. *Proceedings of the National Academy of Sciences of the United States of America*, 97 (23), pp.12559–12564.

Kostiuk, V. and Khokha, M. K. (2021). Xenopus as a platform for discovery of genes relevant to human disease. *Current topics in developmental biology*, 145, p.277.

Kouhara, H. et al. (1997). A lipid-anchored Grb2-binding protein that links FGF-receptor activation to the Ras/MAPK signaling pathway. *Cell*, 89 (5), pp.693–702.

Kozak, M. (1991). Structural features in eukaryotic mRNAs that modulate the initiation of translation. *Journal of Biological Chemistry*, 266 (30), pp.19867–19870.

Kreuser, U. et al. (2020). Initial WNT/β-Catenin Activation Enhanced Mesoderm Commitment, Extracellular Matrix Expression, Cell Aggregation and Cartilage Tissue Yield From Induced Pluripotent Stem Cells. *Frontiers in Cell and Developmental Biology*, 8 (October), pp.1–17.

Krishnakumar, R. et al. (2016). FOXD3 Regulates Pluripotent Stem Cell Potential by Simultaneously Initiating and Repressing Enhancer Activity. *Cell stem cell*, 18 (1), p.104.

Kudoh, T. et al. (2004). Combinatorial Fgf and Bmp signalling patterns the gastrula ectoderm into prospective neural and epidermal domains. *Development*, 131 (15), pp.3581–3592.

Kuffler, S. W. and Vaughan Williams, E. M. (1953). Properties of the 'slow' skeletal muscle fibres of the frog. *J. Physiol.*, 121, pp.318–340.

Kumar, A. et al. (2001). Nodal signaling uses activin and transforming growth factor-β receptor-regulated Smads. *Journal of Biological Chemistry*, 276 (1), pp.656–661.

Kumar, S. et al. (2018). Xbra and Smad-1 cooperate to activate the transcription of neural repressor ventx1.1 in Xenopus embryos. *Scientific Reports*, 8 (1), pp.1–11.

Kunath, T. et al. (2007). FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development*, 134 (16), pp.2895–2902.

de la Serna, I. L. et al. (2005). MyoD targets chromatin remodeling complexes to the myogenin locus prior to forming a stable DNA-bound complex. *Mol. Cell. Biol.*, 25 (10), pp.3997–4009.

LaBonne, C. and Whitman, M. (1994). Mesoderm induction by activin requires FGF-mediated intracellular signals. *Development*, 120 (2), pp.463–472.

Ladd, A. N., Charlet-B., N. and Cooper, T. A. (2001). The CELF Family of RNA Binding Proteins Is Implicated in Cell-Specific and Developmentally Regulated Alternative Splicing. *Molecular and Cellular Biology*, 21 (4), p.1285.

Lagha, M. et al. (2008). Regulation of skeletal muscle stem cell behavior by Pax3 and Pax7. *Cold Spring Harbor Symposia on Quantitative Biology*, 73 (73), pp.307–315.

Lassar, A. B. et al. (1991). Functional activity of myogenic HLH proteins requires hetero-oligomerization with E12/E47-like proteins in vivo. *Cell*, 66 (2), pp.305–315.

Lassar, A. B., Paterson, B. M. and Weintraub, H. (1986). Transfection of a DNA locus that mediates the conversion of 10T1 2 fibroblasts to myoblasts. *Cell*, 47 (5), pp.649–656.

Latinki, B. V. et al. (2002). Distinct Enhancers Regulate Skeletal and Cardiac Muscle-Specific Expression Programs of the Cardiac α-Actin Gene in Xenopus Embryos. *Developmental Biology*, 245 (1), pp.57–70.

De Launoit, Y. et al. (1997). Structure-function relationships of the PEA3 group of Ets-related transcription factors. *Biochemical and Molecular Medicine*, 61 (2), pp.127–135.

Lea, R. et al. (2009). Temporal and spatial expression of FGF ligands and receptors during Xenopus development. *Dev. Dyn.*, 238 (6), pp.1467–1479.

Lee, H.-J. et al. (2004). Sox15 Is Required for Skeletal Muscle Regeneration. *Molecular and Cellular Biology*, 24 (19), p.8428.

Lee, Q. Y. et al. (2020). Pro-neuronal activity of Myod1 due to promiscuous binding to neuronal genes. *Nature Cell Biology 2020 22:4*, 22 (4), pp.401–411.

Lee, S.-Y. et al. (2011). The function of heterodimeric AP-1 comprised of c-Jun and c-Fos in activin mediated Spemann organizer gene expression. *PLoS One*, 6 (7), p.e21796.

Lee, Y. et al. (2014). FGF signalling specifies haematopoietic stem cells through its regulation of somitic Notch signalling. *Nature Communications 2014 5:1*, 5 (1), pp.1-13Lee, Y. et al. (2014). FGF signalling specifie.

Lefaucheur, J. P. and Sébille, A. (1995a). Basic fibroblast growth factor promotes in vivo muscle regeneration in murine muscular dystrophy. *Neuroscience Letters*, 202 (1–2), pp.121–124.

Lefaucheur, J. P. and Sébille, A. (1995b). Muscle regeneration following injury can be modified in vivo by immune neutralization of basic fibroblast growth factor, transforming growth factor β1 or insulin-like growth factor I. *Journal of Neuroimmunology*, 57 (1–2), pp.85–91.

Levenstein, M. E. et al. (2006). Basic Fibroblast Growth Factor Support of Human Embryonic Stem Cell Self-Renewal. *Stem Cells*, 24 (3), pp.568–574.

Li, Y. et al. (2017). Targeted Disruption of TCF12 Reveals HEB as Essential in Human Mesodermal Specification and Hematopoiesis. *Stem Cell Reports*, 9 (3), pp.779–795.

Lindskog, C. et al. (2015). The human cardiac and skeletal muscle proteomes defined by transcriptomics and antibody-based profiling. *BMC Genomics*, 16 (1), p.475.

Liu, C. et al. (2002). Control of β-catenin phosphorylation/degradation by a dual-kinase mechanism. *Cell*, 108 (6), pp.837–847.

Liu, Y. et al. (2013). Six1 Regulates MyoD Expression in Adult Muscle Progenitor Cells. *PLoS ONE*, 8 (6), p.67762.

Lombardo, A., Isaacs, H. V and Slack, J. M. (1998). Expression and functions of FGF-3 in Xenopus development. *Int. J. Dev. Biol.*, 42 (8), pp.1101–1107.

Lombardo, A. and Slack, J. M. W. (2001). Abdominal B-type Hox gene expression in Xenopus laevis. *Mechanisms of Development*, 106 (1–2), pp.191–195.

Londin, E. R., Niemiec, J. and Sirotkin, H. I. (2005). Chordin, FGF signaling, and mesodermal factors cooperate in zebrafish neural induction. *Developmental Biology*, 279 (1), pp.1–19.

López, S. L. et al. (2003). Notch activates sonic hedgehog and both are involved in the specification of dorsal midline cell-fates in Xenopus. *Development*, 130 (10), pp.2225–2238.

Luna, V. M., Daikoku, E. and Ono, F. (2015). 'Slow' skeletal muscles across vertebrate species. *Cell and Bioscience*, 5 (1), pp.1–6.

Ma, X., Chen, H. and Chen, L. (2016). A dual role of Erk signaling in embryonic stem cells. *Experimental Hematology*, 44 (3), pp.151–156.

Maden, C. H. et al. (2012). NRP1 and NRP2 cooperate to regulate gangliogenesis, axon guidance and target innervation in the sympathetic nervous system. *Developmental Biology*, 369 (2), p.277.

Maeda, R. et al. (2002). Xpbx1b and Xmeis1b play a collaborative role in hindbrain and neural crest gene expression in Xenopus embryos. *Proceedings of the National Academy of Sciences of the United States of America*, 99 (8), pp.5448–5453.

Maguire, R. J., Isaacs, H. V and Pownall, M. E. (2012). Early transcriptional targets of MyoD link myogenesis and somitogenesis. *Dev. Biol.*, 371 (2), pp.256–268.

Maire, P. et al. (2020). Myogenesis control by SIX transcriptional complexes. *Seminars in Cell & Developmental Biology*, 104, pp.51–64.

Mal, A. and Harter, M. L. (2003). MyoD is functionally linked to the silencing of a muscle-specific regulatory gene prior to skeletal myogenesis. *Proc. Natl. Acad. Sci. U. S. A.*, 100 (4), pp.1735–1739.

Mall, M. et al. (2017). Myt1l safeguards neuronal identity by actively repressing many non-neuronal fates. *Nature*, 544 (7649), pp.245–249.

Marcelle, C., Wolf, J. and Bronner-Fraser, M. (1995). The in vivo expression of the FGF receptor FREK mRNA in avian myoblasts suggests a role in muscle growth and differentiation. *Developmental Biology*, 172 (1), pp.100–114.

Marchal, L. et al. (2009). BMP inhibition initiates neural induction via FGF signaling and Zic genes. *Proceedings of the National Academy of Sciences of the United States of America*, 106 (41), pp.17437–17442.

Marigo, V. et al. (1996). Biochemical evidence that Patched is the Hedgehog receptor. *Nature 1996 384:6605*, 384 (6605), pp.176–179.

Marks, H. et al. (2012). The Transcriptional and Epigenomic Foundations of Ground State Pluripotency. *Cell*, 149 (3), p.590.

Marro, S. et al. (2011). Direct lineage conversion of terminally differentiated hepatocytes to functional neurons. *Cell stem cell*, 9 (4), p.374.

Massari, M. E. and Murre, C. (2000). Helix-Loop-Helix Proteins: Regulators of Transcription in Eucaryotic Organisms. *Molecular and Cellular Biology*, 20 (2), pp.429–440.

Masuyama, N. et al. (1999). Identification of Two Smad4 Proteins in Xenopus. *Journal of Biological Chemistry*, 274 (17), pp.12163–12170.

Matyushenko, A. M. et al. (2020). Unique functional properties of slow skeletal muscle tropomyosin. *Biochimie*, 174, pp.1–8.

Maves, L. et al. (2007). Pbx homeodomain proteins direct Myod activity to promote fast-muscle differentiation. *Development*, 134 (18), pp.3371–3382.

McKinnell, I. W., Parise, G. and Rudnicki, M. A. (2005). Muscle Stem Cells and Regenerative Myogenesis. *Current Topics in Developmental Biology*, 71, pp.113–130.

Meeson, A. P. et al. (2007). Sox15 and Fhl3 transcriptionally coactivate Foxk1 and regulate myogenic progenitor cells. *The EMBO Journal*, 26 (7), pp.1902–1912.

Mendjan, S. et al. (2014). NANOG and CDX2 pattern distinct subtypes of human mesoderm during exit from pluripotency. *Cell Stem Cell*, 15 (3), pp.310–325.

Merzdorf, C. S. (2007). Emerging roles for zic genes in early development. *Developmental Dynamics*, 236 (4), pp.922–940.

Mi, H. et al. (2021). PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic Acids Research*, 49 (D1), pp.D394–D403.

Milet, C. et al. (2013). Pax3 and Zic1 drive induction and differentiation of multipotent, migratory, and functional neural crest in Xenopus embryos. *Proceedings of the National Academy of Sciences of the United States of America*, 110 (14), pp.5528–5533.

Millay, D. P. et al. (2013). Myomaker is a membrane activator of myoblast fusion and muscle formation. *Nature 2013 499:7458*, 499 (7458), pp.301–305.

Mills, M. A. et al. (2001). Differential expression of the actin-binding proteins, α-actinin-2 and -3, in different species: implications for the evolution of functional redundancy. *Human Molecular Genetics*, 10 (13), pp.1335–1346.

Mohammadi, M. et al. (1991). A tyrosine-phosphorylated carboxy-terminal peptide of the fibroblast growth factor receptor (Flg) is a binding site for the SH2 domain of phospholipase C-gamma 1. *Molecular and cellular biology*, 11 (10), pp.5068–5078.

Molenaar, M. et al. (1996). XTcf-3 transcription factor mediates β-catenin-induced axis formation in xenopus embryos. *Cell*, 86 (3), pp.391–399.

Molkentin, J. D. et al. (1995). Cooperative activation of muscle gene expression by MEF2 and myogenic bHLH proteins. *Cell*, 83 (7), pp.1125–1136.

Monteiro, A. S. and Ferrier, D. E. K. (2006). Hox genes are not always Colinear. *International Journal of Biological Sciences*, 2 (3), p.95.

Morgan, J. E. and Partridge, T. A. (2003). Muscle satellite cells. *The International Journal of Biochemistry & Cell Biology*, 35 (8), pp.1151–1156.

Münchberg, S. R. and Steinbeisser, H. (1999). The Xenopus Ets transcription factor XER81 is a target of the FGF signaling pathway. *Mechanisms of Development*, 80 (1), pp.53–65.

Murakami, M. and Simons, M. (2008). Fibroblast growth factor regulation of neovascularization. *Current opinion in hematology*, 15 (3), p.215.

Murphy, L. O. et al. (2002). Molecular, interpretation of ERK signal duration by immediate early gene products. *Nature Cell Biology*, 4 (8), pp.556–564.

Murre, C. et al. (1989). Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell*, 58 (3), pp.537–544.

Murre, C., Voronova, A. and Baltimore, D. (1991). B-cell- and myocyte-specific E2-box-binding factors contain E12/E47-like subunits. *Molecular and Cellular Biology*, 11 (2), pp.1156–1160.

Na, J., Furue, M. K. and Andrews, P. W. (2010). *Inhibition of ERK1/2 prevents neural and mesendodermal differentiation and promotes human embryonic stem cell self-renewal*.

Nabeshima, Y. et al. (1993). Myogenin gene disruption results in perinatal lethality because of severe muscle defect. *Nature 1993 364:6437*, 364 (6437), pp.532–535.

Nagandla, H. et al. (2016). Defective myogenesis in the absence of the muscle-specific lysine methyltransferase SMYD1. *Developmental Biology*, 410 (1), pp.86–97.

Naidu, P. S. et al. (1995). Myogenin and MEF2 function synergistically to activate the MRF4 promoter during myogenesis. *Molecular and Cellular Biology*, 15 (5), p.2707.

Naujok, O., Diekmann, U. and Lenzen, S. (2014). The Generation of Definitive Endoderm from Human Embryonic Stem Cells is Initially Independent from Activin A but Requires Canonical Wnt-Signaling. *Stem Cell Reviews and Reports*, 10 (4), pp.480–493.

Nicholson, K. M. and Anderson, N. G. (2002). The protein kinase B/Akt signalling pathway in human malignancy. *Cell. Signal.*, 14 (5), pp.381–395.

Nieuwkoop, P. D. (1969). The formation of the mesoderm in urodelean amphibians - II. The origin of the dorso-ventral polarity of the mesoderm. *Wilhelm Roux Archiv für Entwicklungsmechanik der Organismen*, 163 (4), pp.298–315.

Nieuwkoop, P. D. and Faber, J. (1994). *Normal Table of Xenopus Laevis (Daudin) (Daudin : A Systematical and Chronological Survey of the Development from the Fertilized Egg Till the End of Metamorp)*. 1 edition. Routledge.

Nieuwkoop, P. D. and Ubbels, G. A. (1969). The formation of the mesoderm in urodelean amphibians. *Wilhelm Roux' Archiv fur Entwicklungsmechanik der Organismen*, 169 (3), pp.185–199.

Nishitani, E. et al. (2015). Pou5f3.2-induced proliferative state of embryonic cells during gastrulation of Xenopus laevis embryo. *Dev. Growth Differ.*, 57 (9), pp.591–600.

Ong, S. H. et al. (2000). FRS2 proteins recruit intracellular signaling pathways by binding to diverse targets on fibroblast growth factor and nerve growth factor receptors. *Mol. Cell. Biol.*, 20 (3), pp.979–989.

Ong, S. H. et al. (2001). Stimulation of phosphatidylinositol 3-kinase by fibroblast growth factor receptors is mediated by coordinated recruitment of multiple docking proteins. *Proc. Natl. Acad. Sci. U. S. A.*, 98 (11), pp.6074–6079.

Onichtchouk, D. et al. (1996). The Xvent-2 homeobox gene is part of the BMP-4 signalling pathway controling dorsoventral patterning of Xenopus mesoderm. *Development*, 122 (10), pp.3045–3053.

Oprescu, S. N. et al. (2023). Sox11 is enriched in myogenic progenitors but dispensable for development and regeneration of skeletal muscle. *bioRxiv*, p.2023.03.30.534956.

Ornitz, D. M. (2000). FGFs, heparan sulfate and FGFRs: complex interactions essential for development. *Bioessays*, 22 (2), pp.108–112.

Ornitz, D. M. and Itoh, N. (2015). The Fibroblast Growth Factor signaling pathway. *Wiley Interdiscip. Rev. Dev. Biol.*, 4 (3), pp.215–266.

Osborn, D. P. S. et al. (2020). Fgf-driven Tbx protein activities directly induce myf5 and myod to initiate zebrafish myogenesis. *Development (Cambridge, England)*, 147 (8).

Ossipova, O., Stick, R. and Pieler, T. (2002). XSPR-1 and XSPR-2, novel Sp1 related zinc finger containing genes, are dynamically expressed during Xenopus embryogenesis. *Mechanisms of Development*, 115 (1–2), pp.117–122.

Ott, M. O. et al. (1991). Early expression of the myogenic regulatory gene, myf-5, in precursor cells of skeletal muscle in the mouse embryo. *Development*, 111 (4), pp.1097–1107.

Otto, S. P. (2007). The Evolutionary Consequences of Polyploidy. *Cell*, 131 (3), Elsevier B.V., pp.452–462.

Pálfy, M. et al. (2020). Chromatin accessibility established by Pou5f3, Sox19b and Nanog primes genes for activity during zebrafish genome activation. *PLoS Genetics*, 16 (1).

Parker, M. H. et al. (2006). MyoD Synergizes with the E-Protein HEBβ To Induce Myogenic Differentiation. *Molecular and Cellular Biology*, 26 (15), p.5771.

Pawson, T. et al. (1993). Proteins with SH2 and SH3 domains couple receptor tyrosine kinases to intracellular signalling pathways. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 340 (1293), pp.279–285.

Peachey, L. D. and Huxley, A. F. (1962). Structural identification of twitch and slow striated muscle fibers of the frog. *The Journal of cell biology*, 13, pp.177–180.

Van De Peer, Y., Maere, S. and Meyer, A. (2009). The evolutionary significance of ancient genome duplications. *Nature Reviews Genetics*, 10 (10), Nature Publishing Group., pp.725–732.

Penn, B. H. et al. (2004). A MyoD-generated feed-forward circuit temporally patterns gene expression during skeletal muscle differentiation. *Genes Dev.*, 18 (19), pp.2348–2353.

Pera, E. M. et al. (2003). Integration of IGF, FGF, and anti-BMP signals via Smad1 phosphorylation in neural induction. *Genes & Development*, 17 (24), pp.3023–3028.

Perry, R. L. S., Parker, M. H. and Rudnicki, M. A. (2001). Activated MEK1 Binds the Nuclear MyoD Transcriptional Complex to Repress Transactivation. *Molecular Cell*, 8 (2), pp.291–301.

Peters, K. G. et al. (1992). Point mutation of an FGF receptor abolishes phosphatidylinositol turnover and Ca2+ flux but not mitogenesis. *Nature*, 358 (6388), pp.678–681.

Pevny, L. and Placzek, M. (2005). SOX genes and neural progenitor identity. *Current Opinion in Neurobiology*, 15 (1), pp.7–13.

Pimentel, H. et al. (2017). Differential analysis of RNA-seq incorporating quantification uncertainty. *Nature Methods 2017 14:7*, 14 (7), pp.687–690.

Poliacikova, G. et al. (2021). Hox Proteins in the Regulation of Muscle Development. *Frontiers in Cell and Developmental Biology*, 9.

Pownall, M. E. et al. (1996). eFGF, Xcad3 and Hox genes form a molecular pathway that establishes the anteroposterior axis in Xenopus. *Development*, 122 (12), pp.3881–3892.

Pownall, M. E., Gustafsson, M. K. and Emerson Jr, C. P. (2002). Myogenic regulatory factors and the specification of muscle progenitors in vertebrate embryos. *Annu. Rev. Cell Dev. Biol.*, 18, pp.747–783.

Puchta, M., Boczkowska, M. and Groszyk, J. (2020). Low RIN Value for RNA-Seq Library Construction from Long-Term Stored Seeds: A Case Study of Barley Seeds. *Genes*, 11 (10), pp.1–15.

Raible, F. and Brand, M. (2001). Tight transcriptional control of the ETS domain factors Erm and Pea3 by Fgf signaling during early zebrafish development. *Mech. Dev.*, 107 (1–2), pp.105–117.

Raposo, A. A. S. F. and Vasconcelos, F. F. (2015). Ascl1 Coordinately Regulates Gene Expression and the Chromatin Landscape during Neurogenesis. *CellReports*, 10, pp.1544–1556.

Rasmussen, M. and Jin, J. P. (2021). Troponin Variants as Markers of Skeletal Muscle Health and Diseases. *Frontiers in Physiology*, 12, p.747214.

Rawls, A. et al. (1995). Myogenin's Functions Do Not Overlap with Those of MyoD or Myf-5 during Mouse Embryogenesis. *Developmental Biology*, 172 (1), pp.37–50.

Rentzsch, F. et al. (2004). Fgf signaling induces posterior neuroectoderm independently of Bmp signaling inhibition. *Developmental Dynamics*, 231 (4), pp.750–757.

Rescan, P. Y. and Ralliere, C. (2010). A Sox5 gene is expressed in the myogenic lineage during trout embryonic development. *International Journal of Developmental Biology*, 54 (5), pp.913–918.

Respuela, P. et al. (2016). Foxd3 promotes exit from naïve pluripotency through enhancer decommissioning and inhibits germline specification. *Cell stem cell*, 18 (1), p.118.

Rhodes, S. J. and Konieczny, S. F. (1989). Identification of MRF4: A new member of the muscle regulatory factor gene family. *Genes and Development*, 3 (12 B), pp.2050–2061.

Ridgeway, A. G. and Skerjanc, I. S. (2001). Pax3 is Essential for Skeletal Myogenesis and the Expression of Six1 and Eya2. *Journal of Biological Chemistry*, 276 (22), pp.19033–19039.

Romaniello, R. et al. (2012). A novel mutation in the β-tubulin gene TUBB2B associated with complex malformation of cortical development and deficits in axonal guidance. *Developmental Medicine & Child Neurology*, 54 (8), pp.765–769.

Romm, E. et al. (2005). Myt1 family recruits histone deacetylase to regulate neural transcription. *Journal of neurochemistry*, 93 (6), p.1444.

Rønn, L. C. B., Hartz, B. P. and Bock, E. (1998). The neural cell adhesion molecule (NCAM) in development and plasticity of the nervous system. *Experimental Gerontology*, 33 (7–8), pp.853–864.

Rudnicki, M. A. et al. (1993). MyoD or Myf-5 is required for the formation of skeletal muscle. *Cell*, 75 (7), pp.1351–1359.

Sacco, A. et al. (2010). Short Telomeres and Stem Cell Exhaustion Model Duchenne Muscular Dystrophy in mdx/mTR Mice. *Cell*, 143 (7), p.1059.

Salucci, S. et al. (2015). α-Actinin involvement in Z-disk assembly during skeletal muscle C2C12 cells in vitro differentiation. *Micron*, 68, pp.47–53.

Sanson, B., White, P. and Vincent, J.-P. (1996). Uncoupling cadherin-based adhesion from wingless signalling in Drosophila. *Nature*, 383, pp.627–630.

Sasaki, H. et al. (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 (17), pp.3915–3924.

Sato, T., Sasai, N. and Sasai, Y. (2005). Neural crest determination by co-activation of Pax3 and Zic1 genes in Xenopus ectoderm. *Development*, 132 (10), pp.2355–2363.

Satou-Kobayashi, Y. et al. (2021). Temporal transcriptomic profiling reveals dynamic changes in gene expression of Xenopus animal cap upon activin treatment. *Scientific Reports 2021 11:1*, 11 (1), pp.1–12.

Savage, J. et al. (2009). SOX15 and SOX7 differentially regulate the myogenic program in P19 cells. *Stem Cells*, 27 (6), pp.1231–1243.

Scerbo, P. et al. (2012). Ventx factors function as Nanog-like guardians of developmental potential in Xenopus. *PLoS ONE*, 7 (5).

Schäfer, B. W. et al. (1990). Effect of cell history on response to helix–loop–helix family of myogenic regulators. *Nature 1990 344:6265*, 344 (6265), pp.454–458.

Schauer, A. et al. (2020). Zebrafish embryonic explants undergo genetically encoded self-assembly. *Elife*, 9.

Schepers, G. E., Teasdale, R. D. and Koopman, P. (2002). Twenty pairs of Sox: Extent, homology, and nomenclature of the mouse and human Sox transcription factor gene families. *Developmental Cell*, 3 (2), pp.167–170.

Schiaffino, S. et al. (2015). Developmental myosins: Expression patterns and functional significance. *Skeletal Muscle*, 5 (1), pp.1–14.

Schlaeger, T. M. et al. (2004). Decoding Hematopoietic Specificity in the Helix-Loop-Helix Domain of the Transcription Factor SCL/Tal-1. *Molecular and Cellular Biology*, 24 (17), pp.7491–7502.

Schlessinger, J. (2000). Cell signaling by receptor tyrosine kinases. *Cell*, 103 (2), pp.211–225.

Schuler-Metz, A. et al. (2000). The homeodomain transcription factor Xvent-2 mediates autocatalytic regulation of BMP-4 expression in Xenopus embryos. *Journal of Biological Chemistry*, 275 (44), pp.34365–34374.

Schulte-Merker, S. and Smith, J. C. (1995). Mesoderm formation in response to Brachyury requires FGF signalling. *Current Biology*, 5 (1), pp.62–67.

Schwarz, Q. et al. (2009). Neuropilin 1 signaling guides neural crest cells to coordinate pathway choice with cell specification. *Proceedings of the National Academy of Sciences of the United States of America*, 106 (15), pp.6164–6169.

Scotet, E. and Houssaint, E. (1998). Exon III splicing switch of fibroblast growth factor (FGF) receptor-2 and. *Oncogene*, 17 (1), pp.67–76.

Seed, J. and Hauschka, S. D. (1988). Clonal analysis of vertebrate myogenesis: VIII. Fibroblast growth factor (FGF)-dependent and FGF-independent muscle colony types during chick wing development. *Developmental Biology*, 128 (1), pp.40–49.

Semprich, C. I. et al. (2022). ERK1/2 signalling dynamics promote neural differentiation by regulating chromatin accessibility and the polycomb repressive complex. *PLOS Biology*, 20 (12), p.e3000221.

Seo, S. et al. (2007a). Neurogenin and NeuroD direct transcriptional targets and their regulatory enhancers. *EMBO Journal*, 26 (24), pp.5093–5108.

Seo, S. et al. (2007b). Neurogenin and NeuroD direct transcriptional targets and their regulatory enhancers. *The EMBO Journal*, 26 (24), pp.5093–5108.

Session, A. M. et al. (2016). Genome evolution in the allotetraploid frog Xenopus laevis. *Nature*, 538 (7625), pp.336–343.

Sheehan, S. M. and Allen, R. E. (1999). Skeletal muscle satellite cell proliferation in response to members of the fibroblast growth factor family and hepatocyte growth factor. *Journal of Cellular Physiology*, 181 (3), pp.499–506.

Shelton, M. et al. (2016). Robust generation and expansion of skeletal muscle progenitors and myocytes from human pluripotent stem cells. *Methods*, 101, pp.73–84.

Shelton, M. et al. (2019). Gene expression profiling of skeletal myogenesis in human embryonic stem cells reveals a potential cascade of transcription factors regulating stages of myogenesis, including quiescent/activated satellite cell-like gene expression. *PLOS ONE*, 14 (9), p.e0222946.

Sherrington, R. et al. (1995). Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature 1995 375:6534*, 375 (6534), pp.754–760.

Silva, P. N. et al. (2013). Fibroblast Growth Factor Receptor Like-1 (FGFRL1) Interacts with SHP-1 Phosphatase at Insulin Secretory Granules and Induces Beta-cell ERK1/2 Protein Activation. *The Journal of Biological Chemistry*, 288 (24), p.17859.

Simone, C. et al. (2004). p38 pathway targets SWI-SNF chromatin-remodeling complex to muscle-specific loci. *Nat. Genet.*, 36 (7), pp.738–743.

Simunovic, M. and Brivanlou, A. H. (2017). Embryoids, organoids and gastruloids: new approaches to understanding embryogenesis. *Development*, 144 (6), pp.976–985.

Singh, A. et al. (2022). Tcf12 and NeuroD1 cooperatively drive neuronal migration during cortical development. *Development (Cambridge)*, 149 (3).

Singh, A. M. et al. (2007). A Heterogeneous Expression Pattern for Nanog in Embryonic Stem Cells. *Stem Cells*, 25 (10), pp.2534–2542.

Singh, A. M. et al. (2012). Signaling Network Crosstalk in Human Pluripotent Cells: A Smad2/3-Regulated Switch that Controls the Balance between Self-Renewal and Differentiation. *Cell Stem Cell*, 10 (3), pp.312–326.

Singh, K. and Dilworth, F. J. (2013). Differential modulation of cell cycle progression distinguishes members of the myogenic regulatory factor family of transcription factors. *The FEBS Journal*, 280 (17), pp.3991–4003.

Singh, R. K. et al. (2014). Rbfox2-Coordinated Alternative Splicing of Mef2d and Rock2 Controls Myoblast Fusion during Myogenesis. *Molecular Cell*, 55 (4), pp.592–603.

Sivak, J. M., Petersen, L. F. and Amaya, E. (2005). FGF signal interpretation is directed by Sprouty and Spred proteins during mesoderm formation. *Dev. Cell*, 8 (5), pp.689–701.

Slack, J. M. et al. (1987). Mesoderm induction in early Xenopus embryos by heparin-binding growth factors. *Nature*, 326 (6109), pp.197–200.

Slack, J. M. W. et al. (1989). The role of fibroblast growth factor in early Xenopus development. *Development*, 107 (Supplement), pp.141–148.

Slack, J. M. W., Isaacs, H. V. and Darlington, B. G. (1988). Inductive effects of fibroblast growth factor and lithium ion on Xenopus blastula ectoderm. *Development*, 103 (3), pp.581–590.

Sleeman, M. et al. (2001). Identification of a new fibroblast growth factor receptor, FGFR5. *Gene*, 271 (2), pp.171–182.

Smith, A. (2017). Formative pluripotency: the executive phase in a developmental continuum. *Development*, 144 (3), pp.365–373.

Smith, J. C. (1989). Mesoderm induction and mesoderm-inducing factors in early amphibian development. *Development*, 105 (4), pp.665–677.

Smith, J. C. et al. (1990). Identification of a potent Xenopus mesoderm-inducing factor as a homologue of activin A. *Nature*, 345, pp.729–731.

Smith, J. C. et al. (1991). Expression of a Xenopus homolog of Brachyury (T) is an immediate-early response to mesoderm induction. *Cell*, 67 (1), pp.79–87.

Smith, J. C. and Green, J. B. (1990). Graded changes in dose of a Xenopus activin A homologue elicit stepwise transitions in embryonic cell fate. *Nature*, 34 (September), pp.391–394.

Soneson, C., Love, M. I. and Robinson, M. D. (2015). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Research*, 4 (2), p.1521.

Sousa-Victor, P. and Muñoz-Cánoves, P. (2016). Regenerative decline of stem cells in sarcopenia. *Molecular Aspects of Medicine*, 50, pp.109–117.

Squire, J. M. and Morris, E. P. (1998). A new look at thin filament regulation in vertebrate skeletal muscle. *The FASEB Journal*, 12 (10), pp.761–771.

St-Jacques, B., Hammerschmidt, M. and McMahon, A. P. (1999). Indian hedgehog signaling regulates proliferation and differentiation of chondrocytes and is essential for bone formation. *Genes & Development*, 13 (16), p.2072.

Stamos, J. L. and Weis, W. I. (2013). The β-Catenin Destruction Complex. *Cold Spring Harbor Perspectives in Biology*, 5 (1).

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 (8), pp.1347–1357.

Steinberg, F. et al. (2010). The FGFRL1 receptor is shed from cell membranes, binds fibroblast growth factors (FGFs), and antagonizes FGF signaling in Xenopus embryos. *J. Biol. Chem.*, 285 (3), pp.2193–2202.

Storey, K. G. et al. (1998). Early posterior neural tissue is induced by FGF in the chick embryo. *Development*, 125 (3), pp.473–484.

Stuart, C. A. et al. (2016). Myosin content of individual human muscle fibers isolated by laser capture microdissection. *American Journal of Physiology - Cell Physiology*, 310 (5), p.C381.

Sudarwati, S. and Nieuwkoop, P. D. (1971). Mesoderm formation in the anuran Xenopus laevis (Daudin). *Wilhelm Roux' Archiv für Entwicklungsmechanik der Organismen*, 166 (3), pp.189–204.

Sumi, T. et al. (2008). Defining early lineage specification of human embryonic stem cells by the orchestrated balance of canonical Wnt/β-catenin, Activin/Nodal and BMP signaling. *Development*, 135 (17), pp.2969–2979.

Summerbell, D., Halai, C. and Rigby, P. W. J. (2002). Expression of the myogenic regulatory factor Mrf4 precedes or is contemporaneous with that of Myf5 in the somitic bud. *Mechanisms of Development*, 117 (1–2), pp.331–335.

Suzuki, K. ichi T. et al. (2017). Clustered Xenopus keratin genes: A genomic, transcriptomic, and proteomic analysis. *Developmental Biology*, 426 (2), pp.384–392.

Szent-Györgyi, A. G. (1975). Calcium regulation of muscle contraction. *Biophysical Journal*, 15 (7), pp.707–723.

Takada, H. et al. (2009). The RNA-binding protein Mex3b has a fine-tuning system for mRNA regulation in early Xenopus development. *Development*, 136 (14), pp.2413–2422.

Takada, S. et al. (1994). Wnt-3a regulates somite and tailbud formation in the mouse embryo. *Genes and Development*, 8 (2), pp.174–189.

Takemoto, T. et al. (2006). Convergence of Wnt and FGF signals in the genesis of posterior neural plate through activation of the Sox2 enhancer N-1. *Development*, 133 (2), pp.297–306.

Tambalo, M. et al. (2020). Enhancer activation by FGF signalling during otic induction. *Developmental Biology*, 457 (1), pp.69–82.

Tan, J. Y. et al. (2013). Efficient Derivation of Lateral Plate and Paraxial Mesoderm Subtypes from Human Embryonic Stem Cells Through GSKi-Mediated Differentiation. *Stem Cells and Development*, 22 (13), p.1893.

Tanasijevic, B. et al. (2009). Progressive accumulation of epigenetic heterogeneity during human ES cell culture. *Epigenetics*, 4 (5), pp.330–338.

Tandon, P. et al. (2017). Expanding the genetic toolkit in Xenopus: Approaches and opportunities for human disease modeling. *Developmental Biology*, 426 (2), pp.325–335.

Tapscott, S. J. (2005). The circuitry of a master switch: Myod and the regulation of skeletal muscle gene transcription. *Development*, 132 (12), pp.2685–2695.

Taylor, S. M. and Jones, P. A. (1979). Multiple new phenotypes induced in 10T12 and 3T3 cells treated with 5-azacytidine. *Cell*, 17 (4), pp.771–779.

Thayer, M. J. et al. (1989). Positive autoregulation of the myogenic determination gene MyoD1. *Cell*, 58 (2), pp.241–248.

Thompson, J. and Slack, J. M. W. (1992). Over-expression of fibroblast growth factors in Xenopus embryos. *Mechanisms of Development*, 38 (3), pp.175–182.

Thomson, J. A. et al. (1998). Embryonic stem cell lines derived from human blastocysts. *Science (New York, N.Y.)*, 282 (5391), pp.1145–1147.

Timchenko, N. A. et al. (2004). Overexpression of CUG Triplet Repeat-binding Protein, CUGBP1, in Mice Inhibits Myogenesis. *Journal of Biological Chemistry*, 279 (13), pp.13129–13139.

Treutlein, B. et al. (2016). Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq. *Nature 2016 534:7607*, 534 (7607), pp.391–395.

Trueb, B. et al. (2003). Characterization of FGFRL1, a novel fibroblast growth factor (FGF) receptor preferentially expressed in skeletal tissues. *J. Biol. Chem.*, 278 (36), pp.33857–33865.

Tsanov, K. M. et al. (2017). LIN28 phosphorylation by MAPK/ERK couples signalling to the post-transcriptional control of pluripotency. *Nature Cell Biology*, 19 (1), pp.60–67.

Tsumagari, K. et al. (2013). DNA methylation and differentiation: HOX genes in muscle cells. *Epigenetics and Chromatin*, 6 (1), pp.1–17.

Uhlén, M. et al. (2015). Tissue-based map of the human proteome. *Science*, 347 (6220).

Uwanogho, D. et al. (1995). Embryonic expression of the chicken Sox2, Sox3 and Sox11 genes suggests an interactive role in neuronal development. *Mechanisms of Development*, 49 (1–2), pp.23–36.

Valenta, T., Hausmann, G. and Basler, K. (2012). The many faces and functions of β-catenin. *The EMBO Journal*, 31 (12), p.2714.

Vallier, L. et al. (2009). Activin/Nodal signalling maintains pluripotency by controlling Nanog expression. *Development (Cambridge, England)*, 136 (8), p.1339.

VanDussen, K. L. et al. (2012). Notch signaling modulates proliferation and differentiation of intestinal crypt base columnar stem cells. *Development*, 139 (3), pp.488–497.

Vaquero-Garcia, J. et al. (2016). A new view of transcriptome complexity and regulation through the lens of local splicing variations. *eLife*, 5.

Vasconcelos, F. F. et al. (2016). MyT1 Counteracts the Neural Progenitor Program to Promote Vertebrate Neurogenesis. *Cell Reports*, 17 (2), pp.469–483.

Vierbuchen, T. et al. (2010). Direct conversion of fibroblasts to functional neurons by defined factors. *Nature*, 463 (7284), p.1035.

Villegas, S. N., Canham, M. and Brickman, J. M. (2010). FGF signalling as a mediator of lineage transitions - Evidence from embryonic stem cell differentiation. *Journal of Cellular Biochemistry*, 110 (1), pp.10–20.

Vize, P. D. et al. (1991). Assays for Gene Function in Developing Xenopus Embryos. *Methods in Cell Biology*, 36 (C), pp.367–387.

Voronova, A. et al. (2013). Hedgehog signaling regulates MyoD expression and activity. *Journal of Biological Chemistry*, 288 (6), pp.4389–4404.

Vougiouklakis, T., Nakamura, Y. and Saloura, V. (2017). Critical roles of protein methyltransferases and demethylases in the regulation of embryonic stem cell fate. *Epigenetics*, 12 (12), p.1015.

Wacker, S. A. et al. (2004). Timed interactions between the Hox expressing non-organiser mesoderm and the Spemann organiser generate positional information during vertebrate gastrulation. *Developmental Biology*, 268 (1), pp.207–219.

Wahi, K., Bochter, M. S. and Cole, S. E. (2016). The many roles of Notch signaling during vertebrate somitogenesis. *Seminars in Cell & Developmental Biology*, 49, pp.68–75.

Wang, D. Z. et al. (2001). The Mef2c gene is a direct transcriptional target of myogenic bHLH and MEF2 proteins during skeletal muscle development. *Development*, 128 (22), pp.4623–4633.

Wang, Q. and Michalak, M. (2020). Calsequestrin. Structure, function, and evolution. *Cell Calcium*, 90, p.102242.

Wang, S. et al. (2022). Tcf12 is required to sustain myogenic genes synergism with MyoD by remodelling the chromatin landscape. *Communications Biology 2022 5:1*, 5 (1), pp.1–13.

Wang, W. Der et al. (2011). Tfap2a and Foxd3 regulate early steps in the development of the neural crest progenitor population. *Developmental Biology*, 360 (1), pp.173–185.

Wardle, F. C. (2019). Master control: transcriptional regulation of mammalian Myod. *Journal of Muscle Research and Cell Motility 2019 40:2*, 40 (2), pp.211–226.

Wasserman, W. W. and Fickett, J. W. (1998). Identification of regulatory regions which confer muscle-specific gene expression. *Journal of Molecular Biology*, 278 (1), pp.167–181.

Watanabe, K. et al. (2007). A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nature Biotechnology*, 25 (6), pp.681–686.

Weeks, D. L. and Melton, D. A. (1987). A maternal mRNA localized to the vegetal hemisphere in xenopus eggs codes for a growth factor related to TGF-β. *Cell*, 51 (5), pp.861–867.

Wehrli, M. et al. (2000). Arrow encodes an LDL-receptor-related protein essential for Wingless signalling. *Nature*, 407 (6803), pp.527–530.

Weidinger, G. et al. (2005). The Sp1-related transcription factors sp5 and sp5-like act downstream of Wnt/β-catenin signaling in mesoderm and neuroectoderm patterning. *Current Biology*, 15 (6), pp.489–500.

Weinstein, D. C. and Hemmati-Brivanlou, A. (1997). Neural induction in Xenopus laevis: evidence for the default model. *Current Opinion in Neurobiology*, 7 (1), pp.7–12.

Weintraub, H. et al. (1989). Activation of muscle-specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD. *Proceedings of the National Academy of Sciences of the United States of America*, 86 (14), p.5434.

Weintraub, H. et al. (1991). Muscle-specific transcriptional activation by MyoD. *Genes and Development*, 5 (8), pp.1377–1386.

Weintraub, H. (1993). The MyoD family and myogenesis: Redundancy, networks, and thresholds. *Cell*, 75 (7), pp.1241–1244.

Weiss, A. et al. (1999). Organization of human and mouse skeletal myosin heavy chain gene clusters is highly conserved. *Proceedings of the National Academy of Sciences of the United States of America*, 96 (6), pp.2958–2963.

Weiss, A. and Attisano, L. (2013). The TGFbeta Superfamily Signaling Pathway. *Wiley Interdisciplinary Reviews: Developmental Biology*, 2 (1), pp.47–63.

Wijgerde, M. et al. (2005). Hedgehog Signaling in Mouse Ovary: Indian Hedgehog and Desert Hedgehog from Granulosa Cells Induce Target Gene Expression in Developing Theca Cells. *Endocrinology*, 146 (8), pp.3558–3566.

Wright, E. et al. (1995). The Sry-related gene Sox9 is expressed during chondrogenesis in mouse embryos. *Nature Genetics 1995 9:1*, 9 (1), pp.15–20.

Wright, W. E., Sassoon, D. A. and Lin, V. K. (1989). Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD. *Cell*, 56 (4), pp.607–617.

Wrighton, K. H. et al. (2009). Transforming Growth Factor β Can Stimulate Smad1 Phosphorylation Independently of Bone Morphogenic Protein Receptors. *The Journal of Biological Chemistry*, 284 (15), p.9755.

Wu, C. et al. (2000). RGS proteins inhibit Xwnt-8 signaling in Xenopus embryonic development. *Development*, 127 (13), pp.2773–2784.

Wu, G. et al. (2009). Inhibition of GSK3 Phosphorylation of β-Catenin via Phosphorylated PPPSPXS Motifs of Wnt Coreceptor LRP6. *PLOS ONE*, 4 (3), p.e4926.

Wuechner, C. et al. (1996). Developmental expression of splicing variants of fibroblast growth factor receptor 3 (FGFR3) in mouse. *Int. J. Dev. Biol.*, 40 (6), pp.1185–1188.

Xing, M. et al. (2020). The 18S rRNA m6A methyltransferase METTL5 promotes mouse embryonic stem cell differentiation. *EMBO reports*, 21 (10), p.e49863.

Xu, R. H. et al. (2005). Basic FGF and suppression of BMP signaling sustain undifferentiated proliferation of human ES cells. *Nature Methods 2005 2:3*, 2 (3), pp.185–190.

Xu, R. H. et al. (2008). NANOG Is a Direct Target of TGFβ/Activin-Mediated SMAD Signaling in Human ESCs. *Cell Stem Cell*, 3 (2), pp.196–206.

Yajima, H. et al. (2010). Six family genes control the proliferation and differentiation of muscle satellite cells. *Experimental Cell Research*, 316 (17), pp.2932–2944.

Yao, H. H. C., Whoriskey, W. and Capel, B. (2002). Desert Hedgehog/Patched 1 signaling specifies fetal Leydig cell fate in testis organogenesis. *Genes & Development*, 16 (11), pp.1433–1440.

Yeh, B. K. et al. (2003). Structural basis by which alternative splicing confers specificity in fibroblast growth factor receptors. *Proc. Natl. Acad. Sci. U. S. A.*, 100 (5), pp.2266–2271.

Ying, Q. L. et al. (2008). The ground state of embryonic stem cell self-renewal. *Nature 2008 453:7194*, 453 (7194), pp.519–523.

Yoon, S. J. et al. (2011). HEB and E2A function as SMAD/FOXH1 cofactors. *Genes & Development*, 25 (15), pp.1654–1661.

Yoon, S. J., Foley, J. W. and Baker, J. C. (2015). HEB associates with PRC2 and SMAD2/3 to regulate developmental fates. *Nature Communications 2015 6:1*, 6 (1), pp.1–12.

Young, R. A. (2011). Control of Embryonic Stem Cell State. *Cell*, 144 (6), p.940.

Yun, Y. R. et al. (2010). Fibroblast growth factors: Biology, function, and application for tissue regeneration. *Journal of Tissue Engineering*, 1 (1), SAGE Publications., pp.1–18.

Zhang, J. et al. (2016). LIN28 Regulates Stem Cell Metabolism and Conversion to Primed Pluripotency. *Cell Stem Cell*, 19 (1), pp.66–80.

Zhao, P. et al. (2006). Fgfr4 is required for effective muscle regeneration in vivo: Delineation of a MyoD-Tead2-Fgfr4 transcriptional pathway. *The Journal of biological chemistry*, 281 (1), p.429.

Zhao, S. et al. (2014). Comparison of RNA-Seq and Microarray in Transcriptome Profiling of Activated T Cells. *PLOS ONE*, 9 (1), p.e78644.

Zhuang, Y. et al. (1992). Helix-loop-helix transcription factors E12 and E47 are not essential for skeletal or cardiac myogenesis, erythropoiesis, chondrogenesis, or neurogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, 89 (24), pp.12132–12136.

Zhuang, Y. et al. (1998). Functional Replacement of the Mouse E2A Gene with a Human HEB cDNA. *Molecular and Cellular Biology*, 18 (6), pp.3340–3349.

Znosko, W. A. et al. (2010). Overlapping functions of Pea3 ETS transcription factors in FGF signaling during zebrafish development. *Dev. Biol.*, 342 (1), pp.11–25.