Investigation into the FGF-mediated regulation and function of *rasl11b* during early *Xenopus* development

Matthew Edward Barrett Broadsmith

Master of Science (by research)

University of York

Biology

August 2022

Abstract

Fibroblast growth factor (FGF) signalling is essential for the initiation and regulation of multiple developmental processes including gastrulation, mesoderm induction and anteroposterior patterning. Despite an extensive understanding of FGF signal transduction via receptor tyrosine kinases (RTKs), the exact mechanism of FGF target gene expression has yet to be fully elucidated.

Capicua (CIC) is a labile transcriptional repressor of RTK target genes, with evidence suggesting that CIC may act in the FGF signalling pathway. Therefore, we hypothesise that the expression of a subset of FGF target genes is reliant upon the ERK-mediated relief of CIC transcriptional repression.

In this study, gene-level RNA-Seq analysis of CIC knockdown and FGF overexpressing *Xenopus tropicalis* embryos was undertaken, which identified a statistically significant overlap of genes upregulated in both sets of embryos.

One of the genes significantly upregulated following CIC knockdown and FGF overexpression was *rasl11b* which encodes a Ras-like small GTPase with a currently uncharacterised role in development. This study has defined the domains of *rasl11b* expression during *X. tropicalis* development, particularly within the pre-somitic mesoderm (PSM). Manipulation of FGF signalling shows that during gastrulation *rasl11b* expression in the early mesoderm is FGF-dependent, but later in development, *rasl11b* expression in the PSM is regulated independently of FGF signalling. Furthermore, this study suggests that Rasl11b may upregulate FGF activity in the PSM during somitogenesis.

Dysregulation of FGF signalling and/or CIC transcriptional repression is associated with a range of disorders and cancers. Consequently, understanding the molecular mechanisms of FGF target gene expression and function will contribute to the development of more effective therapeutics for developmental disorders and cancer.

Contents

List of Figures

List of Tables

Acknowledgments

I would like to thank my supervisors Dr Harv Isaacs and Prof Betsy Pownall for their invaluable support, guidance, and feedback throughout the project. I would also like to thank Alex, Laura and Helena for their continued help, advice, and humour during my time in the lab. Additional thanks to Prof James Chong for his time and constructive feedback as a member of my thesis advisory panel.

Declaration

I declare that this thesis is a presentation of original work conducted under the supervision of Dr Harv Isaacs and Prof Betsy Pownall at the University of York, and that I am the sole author. RNA-Seq data was collected by Michael King, a previous PhD student in the Isaacs' lab and processed by members of staff at the University of York Technology Facility. Initial quality control of RNA-Seq data and mapping of sequences to the *Xenopus tropicalis* transcriptome was undertaken by Dr John Davey at the University of York Technology Facility. Except for initial RNA-Seq gene lists, work in this thesis has not previously been presented for an award at this, or any other, University. All sources are acknowledged as references.

Chapter 1: Introduction

1.1 The fibroblast growth factor family

The fibroblast growth factor (FGF) family is comprised of two spatially and functionally distinct groups of polypeptides: intracellular and secreted FGFs (Itoh and Ornitz, 2004, Itoh and Ornitz, 2008). The intracellular FGFs act as essential regulators of neuronal and myocardial excitability and have been shown to interact directly with various proteins including members of the voltage-gated sodium channel family (Goldfarb, 2005). However, these co-factors are not secreted and their role during normal embryonic development has yet to be determined (Ornitz and Itoh, 2015). The secreted FGFs are subdivided into two groups according to their mode of action: endocrine and canonical FGFs (Itoh, 2010). The endocrine FGFs act as endocrine factors with essential roles in the adult where they regulate phosphate, bile acid, carbohydrate, and lipid metabolism (Potthoff et al., 2012, Hu et al., 2013). In contrast, the canonical FGFs act as autocrine or paracrine factors and function to control cell proliferation, differentiation, and survival (Powers et al., 2000, Borello et al., 2008).

Irrespective of their mode of action, all secreted FGFs function as small polypeptide signalling molecules which signal through cell-surface receptor tyrosine kinases (RTKs) to regulate critical development processes (Schlessinger, 2000). Gastrulation, mesoderm induction and anteroposterior patterning are all processes dependent upon the action of downstream effectors of the FGF signalling pathway for their normal developmental progression (Böttcher and Niehrs, 2005, Itoh, 2007, McIntosh et al., 2000). The importance of FGF signalling in these processes has been well documented in *Xenopus*, a developmental model popular for its large, externally developing embryos which are resilient to extensive surgical manipulation (Grainger, 2012, Harland and Grainger, 2011). Using this model, the importance of FGF signalling has been highlighted by the aberrant phenotypes exhibited upon FGF knockdown, including failure of cells to migrate during gastrulation, loss of mesodermal markers, and posterior truncation of the anteroposterior (AP) axis (Fletcher and Harland, 2008, Amaya et al., 1991, Amaya et al., 1993, Isaacs et al., 1994, Pownall et al., 1996).

1.2 Fibroblast growth factor signalling

Activation of the FGF signalling pathway by canonical FGFs requires synergistic interaction between the FGF ligand, heparan sulphate proteoglycan (HSPG) and the inactive FGF receptor (Rapraeger et al., 1991, Zhang et al., 2006). HSPGs are a class of cell-surface transmembrane proteins, glycerophosphatidylinositide-anchored proteins, and diffusible extracellular matrix (ECM) proteins which all present heparan sulphate glycosaminoglycan (GAG) at their surface (Aviezer et al., 1994).

HSPGs are essential for successful FGF receptor activation as they mediate simultaneous yet independent interactions with both the receptor and the FGF ligand (Matsuo and Kimura-Yoshida, 2013, Lin, 2004). Such interactions promote the increases in binding affinities required for a 1:1:1 FGF-HSPG-FGF receptor complex to form (Belov and Mohammadi, 2013). Formation of this complex induces a conformational change resulting in receptor dimerization and stabilisation of a 2:2:2 FGF-HSPG-FGF receptor complex (Ornitz and Itoh, 2015). Dimerization activates the receptor leading to subsequent auto- and trans-phosphorylation of the intracellular tyrosine-kinase domains (Schlessinger, 2000). Once phosphorylated, these domains act as docking sites for the binding of signalling molecules via their Src homology 2 (SH2) domains, allowing for the recruitment and activation of different signalling complexes (Pawson et al., 1993, Liu et al., 2012, Ong et al., 2000). The three main signalling pathways proceeding the activation of the FGF receptor are: the phospholipase Cγ (PLCγ), phosphoinositide-3 kinase/AKT (PI3K/AKT), or MAP kinase/ERK (MAPK/ERK) pathway (Figure 1) (Eswarakumar et al., 2005, Gotoh, 2008).

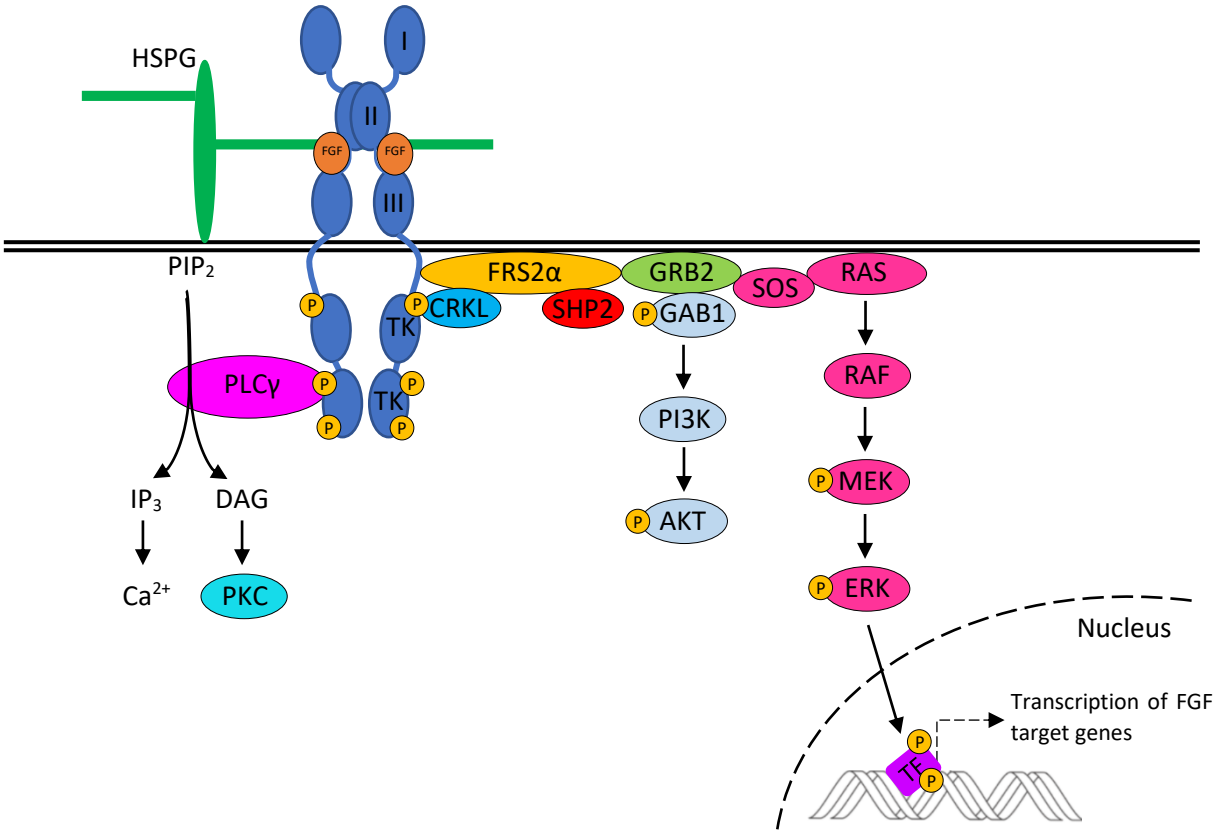


Figure 1: Schematic representation of FGF signalling via the phospholipase Cγ, phosphoinositide-3 kinase/AKT, and MAP kinase/ERK pathway.
Synergistic interaction between fibroblast growth factor (FGF) ligands, heparan sulphate proteoglycans (HSPGs) and the FGF receptor, stimulates the activation and dimerization of the receptor. Receptor activation leads to the auto- and trans-phosphorylation of intracellular tyrosine kinase domains that bind signalling molecules capable of transducing the extracellular signal via three pathways: the PLCγ, PI3K/AKT, or MAPK/ERK pathway.

1.3 The MAP Kinase/ERK signalling pathway

One common signalling pathway utilised in development is the MAPK/ERK pathway, which is initiated upon phosphorylation and activation of the adaptor protein, FGF receptor substrate 2α (FRS2α) (Ong et al., 2000). FRS2α is constitutively associated with the juxtamembrane region of the FGF receptor and interacts with CRKL, an adaptor protein which binds a phospho-tyrosine docking site formed upon receptor activation (Kouhara et al., 1997). This interaction is proposed to be necessary but not sufficient for the activation of FRS2α and later ERK phosphorylation (Seo et al., 2009). Activated FRS2α binds the membrane anchored adaptor protein, growth factor receptor bound 2 (GRB2) and the tyrosine phosphatase SHP2, before recruiting the guanine nucleotide exchange factor, son of sevenless (SOS) (Hadari et al., 1998). The recruitment of SOS to the plasma membrane activates Ras, a membrane-bound GTPase, by promoting the exchange of GDP for GTP (Böttcher and Niehrs, 2005). Interactions of activated Ras with Raf stimulates the serine/threonine kinase to phosphorylate and activate MEK, which in turn, phosphorylates ERK at key threonine and tyrosine residues of its activation loop (Figure 1) (Schlessinger, 2000, Roberts and Der, 2007).

1.4 Transcriptional regulation in the MAP kinase/ERK pathway

Phosphorylated and activated ERK (diphosphorylated ERK or dpERK) is the downstream effector of the MAPK/ERK pathway and acts to regulate the expression of FGF target genes. ERK can induce the expression of target genes through either the direct phosphorylation of transcription factors, or by phosphorylating and activating kinases to do so (Zhao et al., 2003, Cargnello and Roux, 2011). ERK has been shown to mediate the expression of a diverse range of FGF target genes through the activation of the E26 transformation specific (ETS) and activator protein 1 (AP-1) transcription factors (Charlot et al., 2010, Foulds et al., 2004). ETS transcription factors are monomers, whereas AP-1 transcription factors are heterodimers which comprise members of the Fos and Jun families (Lee et al., 2011, Okazaki and Sagata, 1995). ERK activates the transcription of genes encoding the ETS and AP-1 transcription factors and directly phosphorylates the resulting proteins (Brent and Tabin, 2004, Raible and Brand, 2001). This phosphorylation activates the transcription factors to interact with DNA and regulate the expression of FGF target genes (Tsang and Dawid, 2004). However, in the absence of ERK signalling, the transcription of the PEA3 subfamily of ETS transcription factors (Pea3, Etv5 and Etv1) is found to be repressed by the high mobility group box repressor Capicua (CIC) (Dissanayake et al., 2011).

One important transcriptional target of FGF signalling is *egr1* (*early growth response 1*) which encodes a zinc finger transcription factor expressed in the dorsal marginal zone – a key embryonic tissue of the early gastrula stage embryo involved in body axis organisation and neural induction

(Branney et al., 2009, Panitz et al., 1998). Expression of this gene requires ERK phosphorylation of the ETS transcription factor ELK-1 (Nentwich et al., 2009). Although the expression of many FGF target genes, like *egr1*, are regulated by the actions of the ETS and AP-1 transcription factors, the exact mechanism of FGF target gene expression has yet to be elucidated. This is exemplified by the one of the most well documented targets of FGF signalling in Xenopus, the T-box transcription factor Brachyury (*tbxt)*. Brachyury is essential for mesoderm formation in vertebrates and since its discovery as an FGF target gene over two decades ago, it is still unclear how ERK mediates its expression (Showell et al., 2004, Isaacs et al., 1994).

The well characterised FGF target gene *myod1*, encodes a basic helix-loop-helix transcription factor capable of inducing skeletal muscle cell differentiation by the direct regulation of muscle-specific genes (Tapscott, 2005). When *Xenopus laevis* animal caps were treated with FGF4, *myod1* was found to be activated, even in the presence of the translational inhibitor cycloheximide (CHX) (Fisher et al., 2002). This indicated that the activation of *myod1* is an immediate early response, independent of transcription and translation. Given this, and the finding that CHX treatment alone activated some transcription of *myod1*, it was proposed that the regulation of FGF target gene expression likely involved inhibition by a labile transcriptional repressor, the levels of which decrease in the absence of protein synthesis (Fisher et al., 2002).

1.5 The transcriptional repressor Capicua and ERK signalling

The high mobility group box repressor CIC has been proposed as a candidate for the role as a labile transcriptional repressor and has been linked to the transcriptional regulation of RTK target genes in multiple organisms (Jiménez et al., 2000, Jin et al., 2015). In the absence of RTK signalling, CIC binds the promoter and enhancer sequences of RTK target genes which results in their transcriptional repression (Jiménez et al., 2012). In *Drosophila* development, RTK signalling activates ERK which is been found to directly phosphorylate CIC, causing CIC degradation and the transcriptional activation of RTK target genes (Keenan et al., 2020, Astigarraga et al., 2007). Given this, the clear suitability of CIC as a labile transcriptional repressor, and that CIC is already found to repress some FGF target genes in the absence of ERK signalling , our lab hypothesised that a subset of FGF target genes is regulated by CIC transcriptional repression. We proposed that FGF signalling activates ERK to directly phosphorylate CIC and alleviate its transcriptional repression of this subset of FGF target genes (Figure 2).

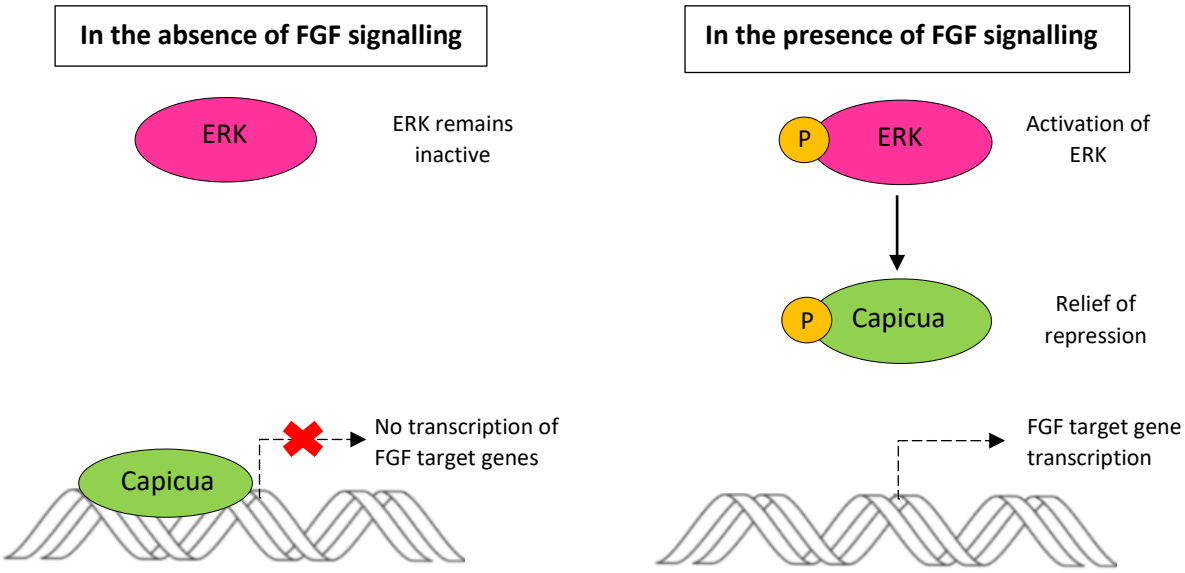


Figure 2: Hypothesised model for the transcriptional regulation of FGF target genes. Under normal conditions, Capicua is thought to bind FGF target genes and repress their transcription. However, upon FGF signalling and subsequent ERK activation, Capicua transcriptional repression of FGF target genes is relieved by its direct phosphorylation by ERK.

Our lab performed a RNA-Seq in triplicate on *Xenopus tropicalis* embryos in which CIC was knocked-down or FGF4 was overexpressed (King, 2019). CIC knockdown was achieved using transcription activator-like effector nucleases (TALENs) and FGF was overexpressed by microinjection of FGF4. CIC knockdown or FGF overexpressing embryos exhibited similar phenotypes: CIC knockdown embryos exhibited a loss of head structures and FGF overexpressing embryos resembled a similar, previously described posteriorized phenotype (Isaacs et al., 1994). In RNA-Seq transcript-level analysis, gene transcripts with a q-value <0.05 and an effect size >1.5 (for upregulated genes) or <0.75 (for downregulated genes) were identified (Cowell, 2019). Statistically significant overlaps between upregulated and downregulated gene transcripts were observed in CIC knockdown and FGF overexpressing embryos, with 44 transcripts upregulated and 21 transcripts downregulated in both sets of embryos (Cowell, 2019). Taken together, these data support our hypothesis that CIC operates in the same pathway as FGF.

1.6 The small GTPase rasl11b

Rasl11b is a highly conserved protein among vertebrates yet remains a poorly documented member of the Ras small GTPase family (Colicelli, 2004, Pézeron et al., 2008). However, according to transcript-level RNA-Seq analysis, this relatively unstudied GTPase was largely upregulated upon CIC knockdown or FGF overexpression, exhibiting some of the highest effect sizes (2.60 for CIC knockdown and 1.99 for FGF overexpression) and lowest q-values (3.45×10^{-7} for CIC knockdown and 0.00145 for FGF overexpression) of all genes analysed (Cowell, 2019). This robust level of

upregulation provides good evidence that Rasl11b is a FGF target gene that is likely transcriptionally repressed by CIC in the absence of ERK signalling. Our lab has also demonstrated preliminary evidence for FGF regulation of rasl11b expression, through in-situ hybridisations illustrating rasl11b expression in the posterior presomitic mesoderm (Cowell, 2019). This region is a known domain of FGF expression with the colocalization of these expression patterns further evidencing a possible link between FGF and rasl11b (Delfini et al., 2005, Lea et al., 2009).

1.7 Project aims

This study presents three main project aims:

1) To validate transcript-level RNA-Seq analysis by identifying a subset of FGF target genes transcriptionally regulated by CIC
2) To define the domains of *rasl11b* expression and investigate whether the gene is an FGF target subject to CIC regulation
3) To investigate the function of *rasl11b* during somitogenesis

Chapter 2: Materials and methods

2.1 Embryological methods

All experimental procedures carried out in this thesis are in accordance with the UK Home Office Animals (Scientific Procedures) Act 1986 under the project license PPL POF245295 held by Prof Betsy Pownall and were approved by the University of York AWERB.

2.1.1 *Xenopus tropicalis* in vitro fertilization and embryo culture

Ovulation was induced in female *Xenopus tropicalis* by subcutaneous injection of 10 units of human chronic gonadotropin (hCG: Chorulon) 20 hours prior to booster injection with 100 units of hCG. *X.* tropicalis injections were undertaken by another lab member with a personal license issued by the UK Home Office. Eggs were collected 4 hours following booster injection, on 60mm dishes coated with L-15 medium (Sigma-Aldrich) + 10% foetal calf serum. Fresh suspensions of macerated testes, obtained from culled males, in L-15 medium + 10% foetal calf serum were washed over eggs, which were then incubated at room temperature for 10 minutes to allow fertilisation. Dishes of fertilised eggs were flooded with MRS/9 (1/9th Modified Ringer's Solution: 11.11mM NaCl, 0.2mM KCl, 0.22mM $CaCl_2$, 0.11mM $MgCl_2$, 5mM HEPES/NaOH pH 7.6 (Tindall et al., 2007)) and de-jellied at 30 minutes post fertilisation by swirling in a solution of 3% L-cysteine in distilled water (pH 7.8) for 5-10 minutes as required. De-jellied embryos were washed thoroughly in distilled water and cultured before the onset of gastrulation in MRS/20 (1/20th MRS: 5mM NaCl, 0.09mM KCl, 0.1mM $CaCl_2$, 0.05mM $MgCl_2$, 5mM HEPES/NaOH pH 7.6) at 21.5-27°C on 1.5% agarose-coated 60mm dishes. Embryonic stage was determined according to the criteria for *X. laevis* (Nieuwkoop and Faber, 1994).

2.1.2 *Xenopus laevis* in vitro fertilization and embryo culture

Ovulation was induced in female *Xenopus laevis* by subcutaneous injection of 250-300 units of hCG 16 hours prior to egg collection on 60mm dishes. *X. laevis* injections were carried out by another lab member with a personal license issued by the UK Home Office. Fresh suspensions of macerated testes in distilled water, obtained from culled males, were washed over eggs, which were incubated at room temperature for 10 minutes to allow fertilisation. Dishes of fertilised eggs were flooded with NAM/3 (1/3rd Normal Amphibian Medium: 36.7mM NaCl, 0.67mM KCl, 0.33mM $Ca(NO_3)_2$, 0.33mM $MgSO_4$, 33.3µM EDTA, 5mM HEPES pH7.4, 1mM $NaHCO_3$ (Slack and Forman, 1980)) and de-jellied at 30 minutes post fertilisation prior to cleavage in a solution of 3% L-cysteine hydrochloride monohydrate in distilled water (pH 7.8) for 5-10 minutes as required. De-jellied embryos were washed thoroughly in distilled water and cultured before the onset of gastrulation in NAM/10 (1/10th NAM: 11mM NaCl, 0.2mM KCl, 0.1mM $Ca(NO_3)_2$, 0.1mM $MgSO_4$, 1µM EDTA,

0.1mM HEPES pH 7.4 (Slack and Forman, 1980) on 1.5% agarose-coated 60mm dishes. *X. laevis* embryos were staged according to Nieuwkoop and Faber (1967).

2.1.3 Drug treatments

Cleavage stage 8 and neurula stage 15 *X. tropicalis* embryos were treated in 1ml of MRS/20, 0.2% Dimethyl sulfoxide (DMSO) in MRS/20, and 200µM SU5402 (Mohammadi et al., 1997) (Sigma-Aldrich) or 25µM PD0325901 (Anastasaki et al., 2012, Sebolt-Leopold and Herrera, 2004) (Cell Guidance Systems) in MRS/20 with 0.2% DMSO. Embryos were treated in 6-well plates coated thinly with 1.5% agarose to minimise drug absorption and incubated at 28°C for approximately 2 hours, until stages 10 and 19 had been reached.

2.1.4 Microinjection

X. tropicalis embryos were unilaterally injected at the 2-cell stage in a solution of 3% Ficoll in MRS/9 using a gas PM 1000 Cell Microinjector with pulled needles (Narishige). *X. laevis* embryos were unilaterally injected at the 2-cell stage in a solution of 5% Ficoll in NAM/3 using the Drummond Microinjector with pulled needles (Drummond). Injected embryos were cultured in these Ficoll solutions to allow healing before transfer into MRS/20 and NAM/10 respectively. *X. tropicalis* and *X. laevis* injections were undertaken by Dr Harv Isaacs.

2.1.5 Embryo fixation

X. tropicalis or *X. laevis* embryos required for RNA extraction and reverse transcription (quantitative) polymerase chain reaction (RT-(q)PCR) were snap frozen on dry ice at the desired stages of development then stored at -70°C. Embryos required for in-situ hybridisation or immunostaining were fixed by rolling in 10ml of MEMFA (0.1M MOPS pH 7.4, 2mM EGTA, 1mM $MgSO_4$, 3.7% formaldehyde (Guille, 1999)) in 20ml glass scintillation vials (Fisher Scientific) for 1 hour. Fixed embryos were rolled in 20ml of 100% methanol for 5 minutes to remove excess MEMFA then stored in 10ml of 100% methanol at -20°C.

2.1.6 Photography and image manipulation

Embryos were imaged using a Leica MZ FLIII microscope with a SPOT 14.2 Colour Mosaic camera (Diagnostic Instruments Inc.) and SPOT Advanced software. Adobe Photoshop CS3 and Adobe Photoshop Elements 2022 were used to optimise digital images.

2.2 Molecular biology methods

2.2.1 Extraction of total RNA

For each RNA extraction, 5 frozen *X. tropicalis* or *X. laevis* embryos were thawed on ice before being homogenised by pipetting in 1ml of Tri-Reagent (Sigma-Aldrich). Samples were left on ice for 1

minute then centrifuged at 13,000rpm for 10 minutes at 4°C. Supernatants were transferred into new Eppendorf tubes and incubated at room temperature for a maximum of 5 minutes, before 200µl of chloroform was added. Tubes were vigorously shaken by hand for 15 seconds, left at room temperature for 5 minutes then centrifuged at 13,000rpm for 15 minutes at 4°C. The aqueous phase was transferred into a new Eppendorf tube and 200µl of chloroform was added, before being centrifuged at 13,000rpm for 5 minutes at 4°C. The fresh aqueous phase was transferred into a new Eppendorf tube, 500µl of isopropanol was added, then tubes were vortexed briefly and placed at -20°C for 29 minutes. Samples were centrifuged at 13,000rpm for 15 minutes at 4°C, the supernatant was removed and 200µl of ice cold 70% ethanol was added to the RNA pellet. Tubes were vortexed, centrifuged at 13,000rpm for 10 minutes at 4°C and the supernatant discarded. RNA pellets were dried by desiccation then resuspended in 100µl of nuclease-free water. 100µl of LiCl solution (4M LiCl, 20mM Tris pH 7.4, 10mM EDTA) was added and samples were placed at -70°C overnight to precipitate RNA. Samples were centrifuged at 13,000rpm for 20 minutes at 4°C, the supernatant discarded and 200µl of ice cold 70% ethanol was added to the RNA pellet. Tubes were vortexed then centrifuged at 13,000rpm for 5 minutes at 4°C and the supernatant discarded again. The RNA pellet was dried by desiccation and resuspended in 20µl of nuclease-free water.

RNA samples were cleaned up using the Zymo RNA Clean & Concentrator kit as per the manufacturer's protocol. This was inclusive of the optional DNase I treatment before RNA purification and samples were eluted in 15µl of nuclease-free water. RNA samples were used immediately for cDNA synthesis or stored at -70°C.

2.2.2 Measurement of DNA and RNA concentrations

Concentrations of DNA and RNA samples were quantified by measuring absorbance at 260nm using a Nanodrop-8000 Spectrophotometer (Thermo Fisher Scientific).

2.2.3 Agarose gel electrophoresis

DNA and RNA samples were run on 1-2.5% gel in Tris-Acetate-EDTA buffer (TAE: 40mM Tris, 20mM acetic acid, 1mM EDTA pH 8.0) and stained with ethidium bromide. DNA samples were run at 150-160V for 20 minutes and RNA samples were run at 190V for 8-20 minutes. The 1kb plus DNA ladder (New England Biolabs) was run alongside samples to predict fragment sizes.

2.2.4 First-strand cDNA synthesis

cDNA was synthesised using 1µg of total RNA with 1µl of 50µM random hexamers (Invitrogen) and 1µl of 10mM dNTP mix (Roche) made up to 13µl with nuclease-free water. Reaction mixtures were vortexed briefly, incubated at 65°C for 5 minutes then placed on ice for 1 minute before the addition of 4µl of 5x SSIV buffer (Invitrogen), 1µl of 100mM Dithiothreitol (DTT) (Invitrogen), 1µl of 200U/µl

SuperScript IV Reverse Transcriptase and 1µl of nuclease-free water. For -RT controls, reverse transcriptase was substituted with an equal volume of nuclear-free water. Mixtures were then incubated at 23°C for 10 minutes, 55°C for 10 minutes and 80°C for 10 minutes. All cDNA was used immediately for PCR amplification or stored at -70°C.

2.2.5 Confirmation of Rasl11b knockdown

To confirm that the *rasl11b* antisense morpholino (AMO) efficiently blocks splicing in *X. tropicalis*, cDNA from *rasl11b* AMO injected embryos was analysed using RT-PCR. NCBI Primer Blast (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) was used to design gene-specific RT-PCR primers for *rasl11b* in *X. tropicalis* (Table 1).

Table 1: Forward and reverse gene specific primer sequences for PCR amplification of different regions of *rasl11b* in *X. tropicalis*.

Amplicon position	Forward primer sequence 5'-3'	Reverse primer sequence 5'-3'	Amplicon size (bp)
Spanning exon 1/2 boundary	ATGCGGCTGATCCAGAACAT	GCGCTTGGTAAGGAATCTCAC	171
Spanning region of exon 4	CCGATGCTGTTGTGATCGTG	GCGAGTTTTTCCGCCTTTCA	318

PCR reactions contained 10µl of 2X PCR Master Mix (Promega), 1µl of 10µM forward and 1µl of 10µl reverse gene-specific primers, 6µl of nuclease-free water and 2µl of cDNA. For H_2O controls, cDNA was substituted with an equal volume of nuclease-free water. Reaction mixtures were heated to 95°C for 2 minutes before 30 cycles of 95°C for 30 seconds, 52°C for 30 seconds and 72°C for 43 seconds (for primers spanning exon 1/2 boundary) or 22 seconds (for primers spanning a region of exon 4), before a final elongation at 72°C for 10 minutes.

PCR products were checked via agarose gel electrophoresis and purified using the Monarch DNA Gel Extraction Kit (New England Biolabs) as per the manufacturer's protocol. Products were eluted in 10µl of DNA Elution Buffer before sequencing.

2.2.6 RT-qPCR

NCBI Primer Blast was also used to design *X. laevis* gene-specific RT-qPCR primers with amplicons sized between 50-150bp. At least one primer in each pair was designed to span an exon-exon boundary to limit amplification from possible genomic DNA contaminant (Table 2).

Table 2: Forward and reverse gene specific primer sequences for qPCR amplification of *rasl11b*, *tbxt* and *dicer1* in *X. laevis*.

Gene	Forward primer sequence 5'-3'	Reverse primer sequence 5'-3'
rasl11b	CGACTATGAAAGGAATGCAGGAAA	CTGAACTTGTATGGCAAGATTCGT
tbxt	CACCACCTACCTCAAGTCAGT	GAGATGAGTAATGAGGTGTAGAGC
dicer1	GGCTTTTACACATGCCTCTTACC	GTCCAAAATTGCATCTCCAAG

For each pair of RT-qPCR primers, serial dilutions of cDNA were used to perform a qPCR standard curve assay to ensure that efficiencies fell between 90-110%.

RT-qPCR reactions were set up in 96-well 0.1ml plates and contained 0.5µl of 10µM forward and 0.5µl of 10µM reverse gene-specific primers, 10µl of 2X SYBR Green PCR Master Mix (Applied Biosystems), 6µl of nuclease-free water and 3µl of cDNA. For non-template controls (NTC), cDNA was substituted with an equal volume of nuclease-free water. RT-qPCR reactions were ran in a QuantStudio 3 Real-Time PCR System (Thermo Fisher Scientific) using the Fast Run mode and Comparative CT (ΔΔCT) settings.

Each qPCR analysis was carried out using three technical replicates with the mean cycle threshold (Ct) value normalised to the Ct of the reference gene *dicer1* to generate a delta-Ct (ΔCt) value for genes of interest. The ΔCt value of cDNA from injected embryos was subtracted from the ΔCt value of cDNA from control embryos to give a delta-delta-Ct (ΔΔCt) value. From this, the relative expression of each gene was calculated using 2^(-ΔΔCt). RStudio was then used to generate grouped bar graphs to show the changes in relative expression between genes in control and injected embryos (Team, 2021).

2.2.7 PCR-based cloning of full-length *rasl11b*

Gene-specific RT-PCR primers for amplification of the full-length of *X. tropicalis rasl11b* were designed with restriction sites, Xho1 and Xba1, to allow in-frame insertion into the pCS2+ plasmid (Table 3). The forward primer was also designed to include a partial Kozak sequence to improve the efficiency of mRNA translation.

Table 3: Forward and reverse gene specific primer sequences for amplification of full-length *rasl11b* in *X. tropicalis*.
Xho1 and Xba1 restriction sites are shown in red and blue respectively. Bold indicates a partial Kozak sequence.

Primer	Sequence 5'-3'
rasl11b-Xho1 forward	AGAGAGCTCGAG**ACCATG**CGGCTGATCCAGAACATG
rasl11b-Xba1 reverse	AGAGAGTCTAGAGACGGAAGTGGCAGTCCTGAC

To increase the probability that *rasl11b* would be successfully amplified, two PCR reactions were set up using different DNA polymerases, Taq and Pfu. The first reaction was a conventional PCR, comprised of 10µl of 2X GoTaq Long PCR Master Mix (Promega), 1µl of 10µM forward and 1µl of 10µM reverse gene specific primers , 2µl of cDNA and 6µl of nuclease-free water. The second reaction was a hot start PCR which initially contained 2.5µl of 10X Pfu DNA Polymerase Buffer (Promega), 0.5µl of 10µM forward and 0.5µl of 10µM reverse gene specific primers, 0.5µl of 10mM dNTP mix (Promega), 1µl of cDNA and 19.25µl of nuclease-free water. Both reaction mixtures were heated to 95°C for 2 minutes before the immediate addition of 0.75µl of Pfu DNA Polymerase (Promega) to the hot start PCR reaction. Both PCR reactions were then subject to 30 cycles of 95°C for 30 seconds, 61-65°C for 30 seconds and 72°C for 1 minute 35 seconds , before a final elongation at 72°C for 10 minutes.

PCR products were checked via agarose gel electrophoresis and those demonstrating the highest level of amplification at a size consistent with the coding sequence of *rasl11b* were products amplified using an annealing temperature of 60-61°C and the Taq polymerase. These PCR products were cleaned up using the Monarch DNA Gel Extraction Kit (New England Biolabs). Each PCR product was made up to a total volume of 100µl with nuclease-free water before the addition of 400µl of Gel Dissolving Buffer. Samples were loaded into the provided spin columns inside collection tubes. Tubes were centrifuged at 13,000rpm for 1 minute before discarding the flow-through. 200µl of DNA Wash Buffer was added and tubes were centrifuged again at 13,00rpm for 1 minute and the flow-through discarded. The DNA Wash Buffer step was repeated before the spin column was transferred to new Eppendorf tube and 25µl of DNA Elution Buffer was added. Tubes were left for 1 minutes at room temperature before centrifuging at 13,000rpm for 1 minute to elute the clean PCR product.

To generate complementary sticky ends between the full-length *rasl11b* insert and the pCS2+ vector, both were digested respectively using Xho1 and Xba1 restriction enzymes. Restriction digests of full-length *rasl11b* contained 15µl of the insert, 10µl of 10X buffer H (Promega), 1.5µl of Xho1 (Promega), 1.5µl of Xba1 (Promega) and 72µl of nuclease-free water. Restriction digests of pCS2+ comprised 5µl of the vector, 10µl of 10X buffer H, 1.5µl of Xho1, 1.5µl of Xba1 and 82µl of nuclease-free water. Digests were incubated at 37°C for 2 hours and complete digestion was confirmed by agarose gel electrophoresis. Following digestion, both the insert and vector were purified using the Monarch DNA Gel Extraction Kit (New England Biolabs) as previously, with agarose gel electrophoresis used to confirm that neither had not been lost during clean up.

Ligation reactions contained 7µl of full-length *rasl11b* insert, 1µl of vector, 1µl of 10X ligase buffer (Promega) and 1µl of T4 ligase (Promega). For control ligations, the full-length *rasl11b* insert was substituted with an equal volume of nuclease-free water. Reaction mixtures were incubated at 18°C for 20 hours before being used for bacterial transformation. DNA was extracted from resulting colonies using Miniprep, checked by colony PCR and sequenced.

2.2.8 Bacterial transformation

dam^{-}*/dcm*$^{-}$ competent *E. coli* cells (New England Biolabs) were thawed from -70°C on ice before 25µl of cells were added to 5µl of 1ng/µl plasmid DNA. Bacteria were chilled on ice for a further 30 minutes prior to heat shock at 42°C for 30 seconds, then cells were immediately replaced on ice for 2 minutes. 1ml of SOC Outgrowth Media (New England Biolabs) was added to each transformation reaction and incubated at 37°C for 1 hour, with shaking (250rpm). Lysogeny broth (LB) agar plates containing 100µg/ml ampicillin (amp) were used to plate out 100µl of transformed culture. The remaining 900µl of culture was centrifuged at 6500rpm for 1 minute and 800µl of supernatant was discarded. The bacterial pellet was gently resuspended with the final 100µl of transformed culture also plated onto LB-amp agar plates. All plates were incubated at 37°C overnight with resulting colonies used for colony PCR (for full-length *rasl11b* transformations) or to directly inoculate 3ml of liquid LB-amp and incubated at 37°C overnight (for re-transformation of in-situ hybridisation probe templates).

2.2.9 Colony PCR of full-length *rasl11b*

Following transformation, colonies were screened for the presence of the full-length *rasl11b* insert using colony PCR. 2µl of nuclease-free water was added to each colony on the plate, mixed, and the resulting suspension was added to the PCR mix. The pipette tip was subsequently used to inoculate 3ml of liquid LB-amp which was incubated at 37°C overnight. Colony PCR reactions contained 2µl of colony/water suspension, 10µl of 2X PCR Master Mix (Promega), 1µl of 10µM SP6 primer, 1µl of 10µM reverse gene specific primer (Table 3) and 6µl of nuclease-free water. Reaction mixtures were heated to 90°C for 1 minute 30 seconds before 30 cycles of 98°C for 30 seconds, 50°C for 30 seconds and 72°C for 1 minute 35 seconds, before a final elongation at 72°C for 10 minutes. PCR products were checked by agarose gel electrophoresis and colonies yielding products of a size consistent with the coding sequence of *X. tropicalis rasl11b* were selected for DNA extraction by Miniprep.

2.2.10 DNA minipreps

Plasmid DNA was isolated from 3ml bacterial cultures using the QIAprep Spin Miniprep Kit (Qiagen) as per the manufacturer's protocol.

2.2.11 Sequencing

Purified plasmid DNA and PCR products were sequenced in both directions using the Eurofins Genomics LightRun Tube custom DNA sequencing service. PCR products were sequenced using the respective forward and reverse gene-specific primers (Table 1) and plasmid DNA was sequenced using SP6 and T7 primers.

2.2.12 Plasmid insert digestion

pGEM-T Easy and pCS107 plasmids ligated with cDNA fragments of *X. tropicalis rasl11b* and *ripply2.2* respectively, were provided by Prof Betsy Pownall and would act as templates for the synthesis of antisense in-situ hybridisation probes. To confirm the presence of these inserts in purified plasmids, an EcoR1 digest was used to excise *rasl11b* from pGEM-T Easy and an EcoR1 and Not1 digest was used to excise *ripply2.2* from pCS107. Restriction digests contained 1-3µg of plasmid DNA, 2µl of 10X buffer H (Promega) and 1µl of each of the appropriate restriction enzymes (Promega) made up to a total volume of 20µl with nuclease-free water. Digests were incubated at 37°C for 2 hours and digestion product sizes were checked via agarose gel electrophoresis.

2.2.13 Plasmid Linearisation and purification

To generate the templates from which full-length *rasl11b* mRNA and different gene-specific in-situ hybridisation probes could be synthesised, plasmids were linearised using restriction enzymes. Linearisation reactions contained 1-3µg of plasmid DNA, 10µl of 10X buffer (Promega) and 2µl of the appropriate enzyme (Table 4) made up to a total volume of 100µl with nuclease-free water. Reactions were incubated at 37°C for 2 hours and agarose gel electrophoresis was used to check that complete digestion had occurred.

Table 4: Plasmids, restriction enzymes, buffers and polymerases used for synthesis of in-situ hybridisation probes and full-length *rasl11b* mRNA.
Enzymes and buffers used to linearise plasmids ligated with *X. tropicalis* full-length *rasl11b* cDNA and cDNA fragments of *rasl11b* and *ripply2.2.* Polymerases used to generate corresponding full-length *rasl11b* mRNA and antisense RNA probes following linear plasmid purification.

Plasmid	**Linearisation enzyme**	**Buffer**	**Polymerase**
pCS2+ *rasl11b* (full-length)	Not1	D	SP6
pGEM-T easy *rasl11b*	Nco1	H	SP6
pCS107 *ripply2.2*	EcoR1	H	T7

Linearised plasmids were made up to a total volume of 400µl with nuclease-free water before 40µl of 3M sodium acetate and 400µl of phenol-chloroform were added. Samples were vortexed for 1 minute then centrifuged at 13,000 rpm for 5 minutes at 4°C. The resulting aqueous layer was

transferred into a new 1.5ml Eppendorf tube and precipitated in 1ml of ice cold 100% ethanol for 1-2 hours at -20°C. DNA precipitates were centrifuged at 13,000rpm for 15 minutes at 4°C with the resulting supernatant discarded. 100µl of ice cold 70% ethanol was added to pellets, vortexed for 1 minute then centrifuged again at 13,000 rpm for 5 minutes at 4°C. Supernatants were discarded and pellets were dried by vacuum desiccation before resuspension in 20µl of nuclease-free water. Agarose gel electrophoresis was used to confirm that linearised plasmids had not been lost during clean up.

2.2.14 In-vitro transcription of synthetic *rasl11b* mRNA

Synthetic *rasl11b* mRNA was transcribed using the mMESSAGE mMACHINE SP6 Transcription Kit (Thermo Fisher Scientific) as per the manufacturer's protocol. Capped transcription reactions contained 0.36µg of linearised plasmid, 10µl of 2X NTP/CAP, 2µl of 10X Reaction Buffer and 2µl of Enzyme Mix made up to a total volume of 20µl with nuclease-free water. Reactions were incubated at 37°C for 2 hours before the optional TURBO DNase step was followed. To recover capped synthetic full-length *rasl11b* mRNA, the phenol chloroform extraction and isopropanol precipitation step was followed and resulting mRNA was resuspended in 20µl of nuclease-free water. Agarose gel electrophoresis was used to check mRNA before it was stored at -70°C

2.2.15 In-situ hybridisation probe synthesis and purification

Digoxigenin (DIG) and Fluorescein (FLU) labelled antisense RNA probes used for in-situ hybridisation were synthesised using the appropriate RNA polymerases (Table 4). Reaction mixtures contained 1µg of plasmid DNA, 4µl of 5X transcription buffer (Thermo Fisher Scientific), 2µl of 100mM DTT (Invitrogen), 2µl of 10X DIG or FLU NTP mix (Roche) and 1µl of polymerase (Thermo Fisher Scientific) made up to a total volume of 20µl with nuclease-free water. Reactions were incubated at 37°C overnight and RNA probe synthesis was confirmed by agarose gel electrophoresis.

Probes were purified by precipitation at -20°C overnight in 50µl of 5M ammonium acetate, 300µl of 100% ethanol and 50µl of nuclease-free water. Precipitates were centrifuged at 13,000rpm for 15 minutes at 4°C, with the resulting supernatant discarded and replaced with 100µl of 70% ethanol. Samples were vortexed for 1 minute and centrifuged at 13,000 rpm for 5 minutes at 4°C. Supernatants were discarded and pellets were dried by vacuum desiccation before resuspension in 50µl of nuclease-free water. The concentration of the RNA probe in samples following purification was estimated by agarose gel electrophoresis.

2.2.16 Modified wholemount in-situ hybridisation

Whole-mount in-situ hybridisation was modified from as previously described (Harland, 1991). MEMFA fixed embryos with vitelline membranes were rehydrated by washing in a series of

gradually decreased concentrations of ethanol in PBS with 0.1% Tween (PBSAT: 137mM NaCl, 2.7mM KCl, 10mM NA_2HPO_4, 1.8mM KH_2PO_4, 0.1% Tween). Embryos were washed at room temperature for 10 minutes in 20ml of 100%, 75% and 50% ethanol in PBSAT respectively, then washed for a further 5 minutes in 20ml PBSAT. They were then rolled in 10ml of $K_2Cr_2O_7$ in 5% acetic acid for 40 minutes at room temperature and washed in 20ml of PBS 3 times for 5 minutes then 3 times for 20 minutes. Embryos were bleached under bright light in 10ml of 5% H_2O_2 in PBS for 45 minutes and washed 3 times in 20ml of PBS for 10 minutes. They were then washed twice in 5ml of 0.1M Triethanolamine (pH 7.8) for 5 minutes with 12.5µl of acetic anhydride added to the second wash. This was followed by a further addition of 12.5µl of acetic anhydride before the embryos were swirled continuously for 5 minutes. Embryos were washed 3 times in 20ml of PBSAT for 5 minutes then transferred into 1.5ml screw-top Eppendorf tubes. 1ml of PBSAT with 250µl of prewarmed pre-hybridisation buffer (50% formamide, 5x SSC pH 7.0, 100µg/ml heparin, 1x Denhart's, 0.1% Tween, 0.1% CHAPS, 10mM EDTA) at 60°c was used to equilibrate embryos for 1 minute at room temperature. They were then incubated at 60°C, first in 1ml of pre-hybridisation buffer for 10 minutes, then in 1ml of hybridisation buffer (pre-hybridisation buffer with 1mg/ml total yeast RNA) on a horizontal tube rocker for 2 hours. Depending on the concentration of DIG-labelled antisense RNA probe, 1-6µl of probe was added to embryos in 1ml of fresh hybridisation buffer and incubated overnight at 60°C with rocking.

Embryos were heated at 60°C for: two 10-minute washes in 1ml of hybridisation buffer, three 20-minute washes in 1ml of 2X SSC with 0.1% Tween, and three 30-minute washes in 1ml of 0.2X SSC with 0.1% Tween. They were then washed twice in 1ml of MAB with 0.1% Tween (MABT: 100mM maleic acid, 150mM NaCl, 0.1% Tween, pH 7.8) at room temperature for 15 minutes. Embryos were equilibrated in 1ml of MABT with 2% BMB for 30 minutes and then blocked in 1 ml of MABT with 2% BMB and 20% heat treated lamb serum at room temperature for 2 hours with rocking. The solution was replaced with 1ml of fresh MABT with 2% BMB and 20% heat treated lamb serum, before a 1 in 2,000 dilution of sheep anti-DIG antibody coupled to alkaline phosphatase (Roche) was added and rolled at 4°C overnight.

Embryos were washed three times in 1ml of MABT for 5 minutes at room temperature then transferred back into 20ml glass scintillation vials and washed three times in 20ml of MABT for 1 hour. They were then washed twice in 5ml of alkaline phosphatase buffer (AP buffer: 100mM Tris pH 9.5, 50mM $MgCl_2$, 100mM NaCl, 0.1% Tween), first for 3 minutes then for 10 minutes at room temperature. Embryos were stained in 1ml of 1 in 3 diluted BM purple (Roche) in AP buffer for 12-72 hours then washed twice in 20ml of PBSAT for 15 minutes and stored in MEMFA.

2.2.17 Modified wholemount double in-situ hybridisation for *ripply2.2* and *rasl11b*

Wholemount double in-situ hybridisation was carried out following all steps listed in the modified single in-situ hybridisation method above, up to the addition of the DIG-labelled antisense RNA probe. For this step, depending on the concentration of each probe, 1-6µl of DIG-labelled *rasl11b* and FLU-labelled *ripply2.2* antisense RNA probes were added to embryos in 1ml of fresh hybridisation buffer and incubated overnight at 60°C with rocking.

Embryos were heated at 60°C for: two 10-minute washes in 1ml of hybridisation buffer, three 20-minute washes in 1ml of 2X SSC with 0.1% Tween, and three 30-minute washes in 1ml of 0.2X SSC with 0.1% Tween. They were then washed twice in 1ml of MABT at room temperature for 15 minutes. Embryos were equilibrated in 1ml of MABT with 2% BMB for 30 minutes and then pre-incubated in 1 ml of MABT with 2% BMB and 20% heat treated lamb serum at room temperature for 2 hours. The solution was replaced with 1ml of fresh MABT with 2% BMB and 20% heat treated lamb serum, before a 1 in 2,000 dilution of sheep anti-FLU antibody coupled to alkaline phosphatase (Roche) was added and rolled at 4°C overnight.

Embryos were washed three times in 1ml of MABT for 5 minutes at room temperature then transferred back into 20ml glass scintillation vials and washed three times in 20ml of MABT for 1 hour. They were then washed twice in 5ml of AP buffer, first for 3 minutes then for 10 minutes at room temperature. Embryos were stained in 175µg of BCIP (5-Bromo-4-chloro-3-indolyl phosphate) in 1ml of AP buffer for 2-5 hours depending on the desired intensity of staining. Embryos were then transferred back into 1.5ml screw-top Eppendorf tubes and AP was inactivated by incubating embryos in 1ml of MABT with 10mM EDTA for 10 minutes at 65°C with rocking. They were washed twice in 1ml of 100% methanol for 5 minutes then 3 times in 1ml of MABT for 5 minutes, before blocking in 1ml of MABT with 2% BMB and 20% heat treated lamb serum at room temperature with rocking. The solution was replaced with 1ml of fresh MABT with 2% BMB and 20% heat treated lamb serum, before a 1 in 2,000 dilution of sheep anti-DIG antibody coupled to alkaline phosphatase (Roche) was added and rolled at 4°C overnight.

Embryos were washed three times in 1ml of MABT for 5 minutes at room temperature then transferred back into 20ml glass scintillation vials and washed three times in 20ml of MABT for 1 hour. They were then washed twice in 5ml of AP buffer, first for 3 minutes then for 10 minutes at room temperature. Embryos were stained in 175µg of Magenta Phosphate (5-bromo-6-chloro-3-indolyl phosphate) in 1ml of AP buffer for 20 hours so the underlying BCIP stain could still be visualised. Stained embryos were washed twice in PBSAT for 15 minutes and stored in MEMFA.

2.2.18 Wholemount immunostaining

Wholemount immunostaining was carried as previously described (Christen and Slack, 1999). MEMFA fixed embryos were rehydrated by washing in a series of gradually decreased concentrations of ethanol in PBS. Embryos were washed at room temperature for 10 minutes in 20ml of 100%, 75% and 50% ethanol in PBS respectively, then washed once more in 20ml of PBS for 10 minutes. They were then treated with 10ml of $K_2Cr_2O_7$ in 5% acetic acid for 40 minutes at room temperature and washed in 20ml of PBS 3 times for 5 minutes then 3 times for 30 minutes. Embryos were bleached in 10ml of 5% H_2O_2 for 45 minutes and washed 3 times in 20ml of PBS for 15 minutes, then transferred into 1.5ml screw-top Eppendorf tubes. They were then equilibrated by washing twice in 1ml of BBT (PBS, 1% BSA, 0.1% Triton X-100) for 1 hour at room temperature before blocking in 1ml of BBT with 5% horse serum for 1 hour. Embryos were rolled in 1ml of 1 in 10,000 diluted mouse anti-dpERK1+2 primary antibody (Sigma) in BBT with 5% horse serum overnight at 4°C.

Embryos were washed 5 times for 1 hour at room temperature, 4 times in 1ml of BBT then once in 1ml of BBT with 5% horse serum. The solution was replaced with 1ml of 1 in 1,000 diluted horse anti-mouse igG-AP conjugated secondary antibody in BBT with 5% horse serum and rolled overnight at 4°C. Embryos were re-transferred into 20ml glass scintillation vials then washed 5 times for 1 hour at room temperature, once in 20ml of BBT and 4 times in 20ml of PBSAT. They were then twice washed in 5ml of AP buffer, first for 3 minutes then for 10 minutes at room temperature. Embryos were stained in 1ml of 1 in 3 diluted BM purple (Roche) in AP buffer for 12-72 hour and washed twice in 20ml of PBS for 15 minutes. They were then bleached in 10ml of 5% H_2O_2 in PBS for approximately 4 hours to help visualise staining.

2.3 Bioinformatic methods

2.3.1 Differential expression analysis of RNA-Seq data

X. tropicalis mRNA samples of CIC knockdown by TALENs, water injected and FGF4 overexpressing embryos were collected in triplicate and total RNA was analysed using the bioanalyzer to confirm RNA quality. Library preparation and Illumina sequencing was performed by staff at the Bioscience Technology Facility at the University of York. Samples were sequenced using a single lane of the Illumina HiSeq 2000 platform and first strand cDNA synthesis used random hexamers and reverse transcriptase to construct the cDNA library. Each cDNA was then sequenced in a high-throughput manner to obtain a read count. The number of reads for each transcript was used to calculate transcripts per million (TPM). TPM is a measure of the abundance of each transcript in each sample adjusted for the varying number of reads for each sample, and the varying expression of transcripts

across the whole transcriptome. For example, the TPM value for a given transcript should represent the number of times that transcript would be observed if one million transcripts were sequenced at random from the entire transcriptome.

Initial RNA-Seq bioinformatics analysis was performed at the Bioscience Technology Facility at the University of York. Raw reads for each sample were aligned to the *X. tropicalis* reference transcriptome (https://www.xenbase.org/entry/displayJBrowse.do?data=data/xt9_1) (genome v9.1) using Salmon (https://salmon.readthedocs.io/en/latest/) to produce estimated read counts for each transcript in each sample. All of the above sample preparation, RNA-Sequencing and preliminary analysis contributed to the thesis of Michael King, a previous PhD student in the Isaacs' lab (King, 2019). Subsequent analysis was undertaken for the purposes of this project.

Sleuth (http://pachterlab.github.io/sleuth/) was used in 'gene mode' to aggregate estimated read counts for transcripts of each gene before they were fit to a statistical model to calculate differential gene expression (q-values and effect sizes). The expression of transcripts from 16,549 genes was detected before genes with a TPM < 1.5 were removed to filter out those with very low levels of expression. The remaining genes were presented in volcano plots which were generated using Rstudio (Team, 2021).

2.3.2 Gene ontology analysis of RNA-Seq data

Genes with a q-value < 0.1 and an effect size > 1.75 in CIC knockdown and/or FGF4 overexpressing embryos were selected for gene ontology (GO) analysis using the Protein Analysis Evolutionary Relationships (PANTHER) classification tool (Mi et al., 2019) (http://www.pantherdb.org/). The *Xenopus tropicalis* reference list was used to match genes to their respective PANTHER GO-Slim Biological Processes. Genes were classified according to the PANTHER GO-Slim subset of GO terms which were identified during GO Phylogenetic Annotation (PAINT) and were judged in expert review to be informative of function and evolutionarily conserved (Mi et al., 2021).

2.3.3 Gene overlap analysis

GeneVenn (http://genevenn.sourceforge.net/) was used to identify genes that were either upregulated (q-value < 0.1 and effect size > 1.75) or downregulated (q-value < 0.1 and effect size < -1.75) in both CIC knockdown and FGF overexpressing gene lists (Pirooznia et al., 2007). The GeneOverlap package in RStudio was used to calculate the statistical significance of observed gene overlaps using Fisher's Exact test (Shen L, 2022, Team, 2021).

2.3.4 Morpheus temporal expression analysis

Genes upregulated in both CIC knockdown and FGF overexpressing embryos were mapped to their respective TPM values in a previous RNA-Seq temporal expression study (Owens et al., 2016). This study provided TPM values for each gene in WT *X. tropicalis* embryos at 30-minute intervals up to 24 hours post fertilisation (Owens et al., 2016). This allowed a dataset to be created containing temporal expression profiles for each gene upregulated in both CIC knockdown and FGF overexpressing embryos, omitting those absent in data from Owens et al., 2016. The dataset was uploaded to Morpheus (https://software.broadinstitute.org/morpheus/) to generate a heatmap illustrating these data.

2.3.5 Capicua binding site analysis

The Find Individual Motif Occurrences (FIMO) (https://meme-suite.org/meme/tools/fimo) tool in MEME Suite 5.4.1 used UCSC reference genomes (https://genome.ucsc.edu/index.html) to scan for CIC consensus binding sites in the genomic loci of *Xenopus tropicalis, Mus musculus* and *Homo Sapiens rasl11b* genes (Kent et al., 2002, Grant et al., 2011, Bailey et al., 2015). FIMO uses a dynamic programming algorithm to convert log-odds scores into p-values, assuming a zero-order background model. A threshold p-value of $p < 0.001$ was used to report results.

Chapter 3: Gene-level transcriptomic analysis of Capicua knockdown and FGF overexpression

3.1 Introduction

As discussed in chapter 1, the literature presents an expanding body of evidence which demonstrates that CIC acts a transcriptional repressor of RTK target genes. The high mobility group box repressor has been shown to exhibit the same spatial and temporal expression pattern as some FGF ligands (King, 2019, Lea et al., 2009). Treatment of *X. tropicalis* embryos with FGF4 or FGF8 leads to the post-transcriptional modification and degradation of the CIC protein (King, 2019). In addition to this, the phenotype produced by CIC knockdown by TALENs in *X .tropicalis* embryos strongly resembles that of the posteriorized phenotype produced when FGF is overexpressed (Isaacs et al., 1994, King, 2019). Collectively, this evidence is supportive of the hypothesis that CIC acts downstream of the FGF signalling pathway but still does not demonstrate a direct link between CIC and FGF signalling. One method to establish if such a link exists is to identify genes that are commonly regulated by both CIC and FGF signalling.

To do this, an RNA-Seq experiment was carried out in triplicate by a PhD student in the Isaacs' lab (King, 2019) on control, FGF overexpressing and CIC knockdown *X. tropicalis* embryos. CIC knockdown was facilitated using transcription activator-like effector nucleases (TALENs) which comprise a non-specific DNA-cleaving nuclease (Fok1) fused to a DNA-binding domain (Lei et al., 2012). The DNA binding domain can be easily engineered to target any sequence, at which the dimerization of forward and reverse TALENs induces DNA double strand breaks (Lei et al., 2012, Kim et al., 1996, Joung and Sander, 2013). This can be utilised for genome-editing since DNA repair mechanism such as non-homologous end joining (NHEJ) introduce variable length insertion/deletion (indel) mutations at the site of the break (Joung and Sander, 2013, Mani et al., 2005, Urnov et al., 2010). Indel mutations typically cause frameshifts resulting in nonsense mediated decay of mRNA or non-functional gene products (Chang et al., 2007, Lin et al., 2017). In this experiment, TALENs were designed to target the HMG-box of the *CIC* gene, since this is the domain which facilitates binding of the protein to octameric T(G/C)AATG(A/G)A sequences in the promoters and enhancers of target genes (Jiménez et al., 2012).

91.8% (74/81) of *X. tropicalis* embryos injected with a total of 1ng of forward and reverse TALENs mRNA produced a mutant phenotype ranging in severity from reduced eye size and pigmentation or cyclopia (50.8%), to a complete loss of head structures (16.1%) (King, 2019). Sequencing of DNA extracted from 8 TALENs injected embryos at the late tailbud stage 40-41, demonstrated the

introduction of indel mutations in sequences adjacent to the target site, indicating a TALENs targeting efficiency of 100%. (King, 2019). The range of phenotypes observed here is likely attributed to the mosaic nature of TALENs targeting. Once injected, TALEN mRNAs are translated rapidly using the cell's own translational machinery, but it takes time for protein levels to accumulate to the threshold necessary for targeting to occur. During this time, cell division can take place resulting in some cells that escape targeting, whilst those that are targeted experience variability in indel sizes. Additionally, TALENs can a disrupt both alleles of a gene, leading to the production of two different mutant alleles each with a different sized indel following NHEJ in somatic cells. Since indels are usually located in spacer sequences between sites bound by a TALEN pair, TALENs can also rebind, thus introducing additional mutations during development (Lei et al., 2012).

Until the midblastula transition (MBT), the zygotic genome remains quiescent, and development is controlled by maternal factors deposited in the cytoplasm (Lee et al., 2014). Consequently, TALEN disrupted *CIC* loci were not transcribed and CIC knockdown was not initiated until after the MBT, with any CIC expression prior to this derived from maternally deposited mRNA. To ensure that FGF overexpression and CIC knockdown occurred at the same stage of development, FGF4 was injected into embryos in a CSKA plasmid, this drives the expression of FGF4 off a β-actin promoter which only becomes active following the MBT (Isaacs et al., 1994). Embryos were injected either CSKA-FGF4 plasmid or CIC TALEN mRNA at the 1-2 cell stage and were collected at neurula stage 14 for RNA-Seq analysis using Illumina HiSeq.

Since FGF4 overexpression results in the upregulation of FGF target genes (Branney et al., 2009) and evidence suggests that CIC acts a transcriptional repressor downstream of the FGF signalling pathway (Isaacs et al., 1994, King, 2019), we hypothesise that CIC knockdown should lead to the upregulation of a subset of FGF target genes. In RNA-Seq transcript level analysis, the abundance of 43,558 transcripts for 23,635 genes was analysed for 3 batches of FGF4 injected, CIC knockdown and water injected embryos (Cowell, 2019). Gene transcripts with a q-value < 0.05 and an effect size > 1.5 were identified as upregulated, with 331 and 81 transcripts upregulated in CIC knockdown and FGF overexpressing embryos respectively (Cowell, 2019). 44 of the 87 transcripts upregulated in FGF overexpression were also upregulated in CIC knockdown, with this overlap identified as highly statistically significant (Cowell, 2019). The large proportion of transcripts upregulated in both FGF overexpressing and CIC knockdown embryos supports the notion that CIC acts in the same pathway as FGF.

Despite the identification of transcripts upregulated in both CIC knockdown and FGF overexpressing embryos, a subset of genes commonly regulated by CIC and FGF signalling has yet to be described. To address this, the same RNA-Seq dataset (King, 2019) was used to analyse the differential expression of transcripts from CIC knockdown and FGF overexpressing embryos at the level of the gene.

The aims of this chapter are:

- Utilise gene-level RNA-Seq data to analyse the differential expression of genes in CIC knockdown and FGF4 overexpressing embryos
- Undertake gene ontology enrichment analysis to identify biological processes associated with upregulated genes
- Identify the genes upregulated and downregulated by both CIC knockdown and FGF4 overexpression
- Establish the statistical significance of the overlap in genes upregulated and downregulated by both CIC knockdown and FGF4 overexpression
- Analyse the temporal expression pattern of genes upregulated upon CIC knockdown and FGF overexpression during early *X. tropicalis* development

3.2 Results

3.2.1 Gene-level transcriptomics reveal gene expression changes following CIC knockdown and FGF overexpression

Initial quality control of RNA-Seq data and mapping of sequences to the *X. tropicalis* transcriptome was undertaken by Dr John Davey at the University of York Technology Facility. Aggregated read counts for transcripts of each gene were fitted to linear models to calculate the differential expression of genes between CIC knockdown or FGF4 overexpressing embryos and water injected controls. The biological significance of changes in gene expression were assessed according to the q-value and effect size calculated for each gene. A q-value is an adjusted p-value which accounts for the false discovery rate and is an established measure of the statistical significance of differential gene expression (Storey and Tibshirani, 2003). Effect size is a measure of the relative change in gene expression between treated and control samples. The smaller the q-value the more significant the change in gene expression and the lower the chance this change is due to false positives. The more positive or more negative the effect size, the greater the increase or decrease in gene expression respectively.

Initial gene-level RNA-Seq analysis detected the expression of transcripts from a total of 16,549 genes in both CIC knockdown and FGF4 overexpressing embryos, before genes with a TPM < 1.5 were removed to filter out those with very low levels of expression. The remaining 13,054 and 12,923 genes expressed in CIC knockdown and FGF4 overexpressing embryos respectively, exhibit large variation in q-values and effect sizes (Figure 3A-B). To distinguish the biologically significant changes in gene expression, genes with a q-value < 0.1 and an effect size > 1.75 (for upregulated genes) or < -1.75 (for downregulated genes) were identified (green plots, Figure 3A-B). The selection of these significance thresholds was biologically validated given that well characterised FGF target genes including *egr1* and *dusp6* were identified as significantly upregulated in FGF4 overexpressing embryos (Figure 3A) (Gómez et al., 2005, Panitz et al., 1998, Branney et al., 2009).

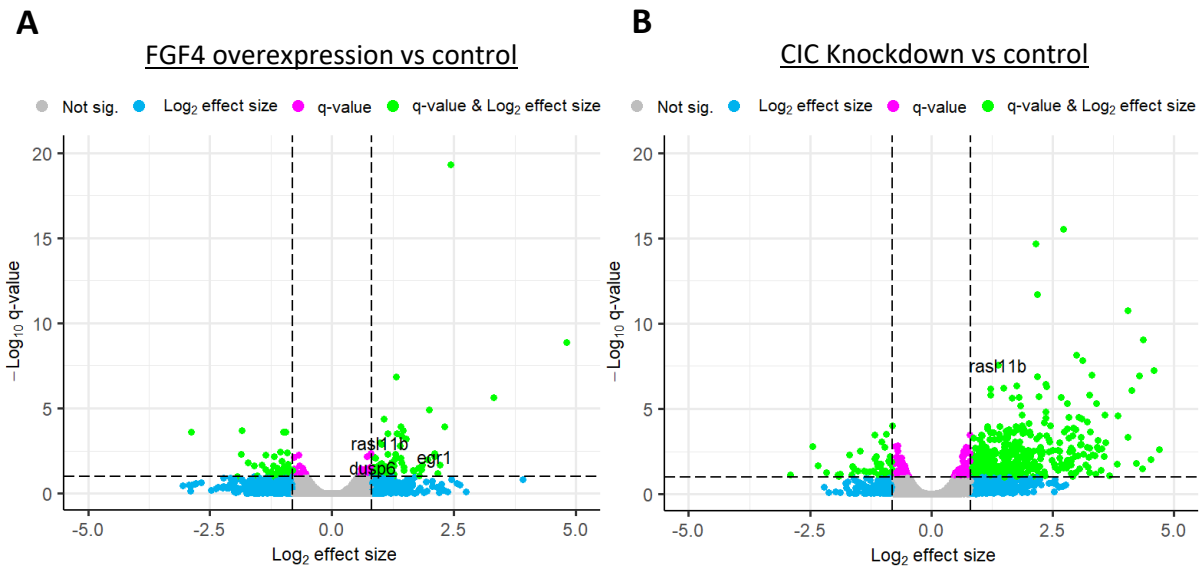


Figure 3: Gene-level transcriptomic analysis demonstrates changes in gene expression upon FGF4 overexpression or CIC knockdown in *X. tropicalis* embryos.
Volcano plots showing the differential expression of genes relative to water injected controls in embryos (A) overexpressing FGF4 from the CSKA-FGF4 plasmid or (B) knocked down for CIC using TALENs. Effect sizes (log_2) were plotted against q-values ($-log_{10}$) for each gene. Dashed lines represent significance thresholds. Genes in grey did not meet either the q-value or effect size thresholds for significance. Genes in pink only meet the q-value threshold of < 0.1. Genes in blue only meet the effect size thresholds of > 1.75 for upregulation or < -1.75 for downregulation. Genes in green meet both the q-value and an effect size threshold and were deemed to show significant changes in gene expression.

A similarly low number of genes exhibit a significant reduction in gene expression between the two treatments, with 41 genes downregulated in CIC knockdown embryos compared to 34 genes in FGF4 overexpressing embryos (Figure 3A-B). However, 367 genes were identified as significantly upregulated upon CIC knockdown compared to 62 genes upon FGF4 overexpression (Figure 3A-B). The greater number of genes upregulated by CIC knockdown than FGF4 overexpression supports the notion that CIC acts as a transcriptional repressor. The complete gene lists produced in RNA-

Seq gene-level analysis are available on Google Drive (https://docs.google.com/spreadsheets/d/1Bjraw6AQ-DgC6TQvkPEPe8XUoDNXXTDxkeReQ5jmPS8/edit?usp=sharing).

3.2.2 Biological processes associated with genes upregulated following CIC knockdown and FGF overexpression

To identify the biological processes associated with genes upregulated upon FGF4 overexpression and CIC knockdown, gene ontology (GO) enrichment analysis was undertaken to classify genes according to their functional characteristics. The PANTHER classification tool (Mi et al., 2019) was initially used to identify biological processes associated with each gene upregulated following CIC knockdown and FGF4 overexpression. The tool was then used to perform a statistical overrepresentation test, which utilised Fisher's exact test and calculated false discovery rates (FDRs) to identify which biological processes are overrepresented by the genes upregulated upon FGF4 overexpression (Table 5) and CIC knockdown (Table 6) respectively. Genes were classified according to the PANTHER GO-Slim subset of GO terms which were judged in expert review to be informative of function and evolutionarily conserved (Mi et al., 2021).

Table 5: Biological processes associated with genes upregulated in FGF4 overexpressing *X. tropicalis* embryos.
PANTHER gene ontology processes identified as statistically overrepresented by genes upregulated in FGF4 overexpressing embryos. Fisher's exact test with false discovery rate (FDR) was used to determine significance using an FDR threshold of < 0.05. +/– indicates whether a process is overrepresented or underrepresented respectively. GO-Slim biological processes are a subset of gene ontology terms identified by the PANTHER classification system which indicate the biological processes to which a protein contributes.

Panther Go-Slim Biological Process	Number of genes	Expected	Fold enrichment	+/-	Raw P-value	FDR
Negative regulation of intracellular signal transduction	3	0.14	21.64	+	4.01×10^{-4}	3.86×10^{-2}
Regulation of cell population proliferation	4	0.22	18.00	+	7.97×10^{-5}	2.21×10^{-2}
Cell population proliferation	4	0.23	17.05	+	9.78×10^{-5}	1.97×10^{-2}
Cellular response to growth factor stimulus	4	0.24	16.55	+	1.09×10^{-4}	1.62×10^{-2}
Intracellular signal transduction	7	1.40	5.01	+	4.39×10^{-4}	4.05×10^{-2}
System development	7	1.45	4.82	+	5.52×10^{-4}	4.89×10^{-2}
Unclassified	10	22.55	0.44	-	8.04×10^{-5}	1.78×10^{-2}

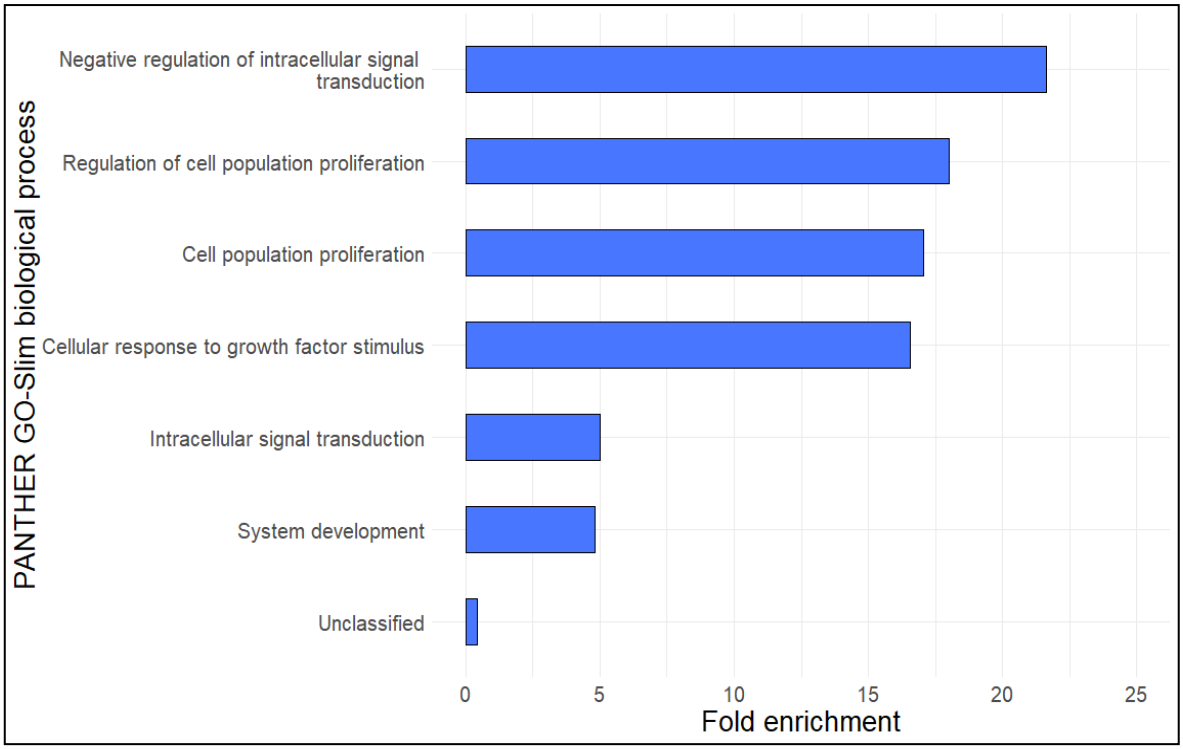


Figure 4: Fold enrichment of biological processes associated with genes upregulated in FGF4 overexpressing *X. tropicalis* embryos.
Bar chart showing the fold enrichment of biological processes identified as statistically overrepresented by genes upregulated in FGF4 overexpressing embryos. Fisher's exact test with false discovery rate (FDR) was used to determine significance using an FDR threshold of < 0.05. GO-Slim biological processes are a subset of gene ontology terms identified by the PANTHER classification system which indicate the biological processes to which a protein contributes.

PANTHER GO analysis identified 6 biological processes associated with genes significantly upregulated in FGF4 overexpressing embryos (Figure 4). The GO terms with the highest fold enrichment were negative regulation of intracellular signal transduction, regulation of cell population proliferation, cell population proliferation and cellular response to growth factor stimulus. These biological processes are consistent with the action of FGF4, since inherently as a growth factor, FGF can induce cell population proliferation, but it is also known that FGF can transcriptionally induce negative regulators of the FGF signalling pathway (Laestander and Engström, 2014, Ekerot et al., 2008). Given that these GO terms are known FGF-mediated biological processes, PANTHER enrichment analysis supports the notion that genes identified to be upregulated in FGF4 overexpressing embryos are FGF target genes.

Table 6: Biological processes associated with genes upregulated in CIC knockdown *X. tropicalis* embryos.

PANTHER gene ontology processes identified as statistically overrepresented by genes upregulated in CIC knockdown embryos. Fisher's exact test with false discovery rate (FDR) was used to determine significance using an FDR threshold of < 0.05. +/– indicates whether a process is overrepresented or underrepresented respectively. GO-Slim biological processes are a subset of gene ontology terms identified by the PANTHER classification system which indicate the biological processes to which a protein contributes.

Panther Go-Slim Biological Process	Number of genes	Expected	Fold enrichment	+/-	Raw P-value	FDR
Regulation of cell cycle	10	2.16	4.62	+	9.26×10^{-5}	5.13×10^{-2}
Cell cycle	15	4.69	3.20	+	1.05×10^{-4}	4.67×10^{-2}
Cellular catabolic process	22	7.63	2.88	+	1.25×10^{-5}	2.76×10^{-2}
Regulation of metabolic process	45	24.57	1.83	+	6.66×10^{-5}	4.92×10^{-2}

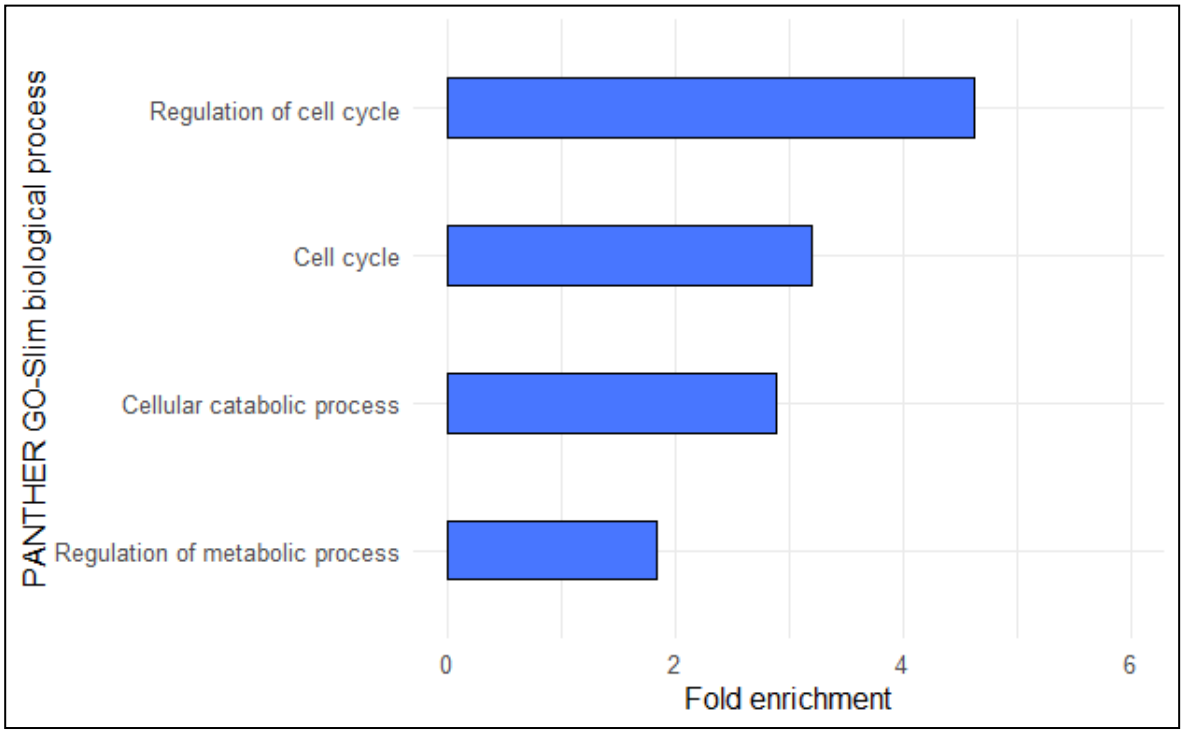


Figure 5: Fold enrichment of biological processes associated with genes upregulated in CIC knockdown *X. tropicalis* embryos.

Bar chart showing the fold enrichment of biological processes identified as statistically overrepresented by genes upregulated in CIC knockdown embryos. Fisher's exact test with false discovery rate (FDR) was used to determine significance using an FDR threshold of < 0.05. GO-Slim biological processes are a subset of gene ontology terms identified by the PANTHER classification system which indicate the biological processes to which a protein contributes.

PANTHER GO analysis identified 4 biological processes associated with genes significantly upregulated in CIC knockdown embryos (Figure 5). The two GO terms with the highest fold

enrichment were regulation of cell cycle and cell cycle. This is encouraging since FGF has been shown to directly regulate cell cycle progression (Lobjois et al., 2004, Zhang et al., 2004). The two GO terms with the lowest fold enrichment were cellular catabolic process and regulation of metabolic process. Some secreted FGFs are important in the regulation of a variety of metabolic and catabolic processes in the adult (Yan et al., 2011, Nies et al., 2015, Tezze et al., 2019). Given that these GO terms are biological processes that can be mediated by FGF signalling, PANTHER enrichment analysis supports the hypothesis that some genes upregulated upon CIC knockdown are likely FGF targets.

3.2.3 Analysing overlap between upregulated and downregulated genes in FGF overexpressing and CIC knockdown embryos

A strong overlap between significantly upregulated and downregulated genes was identified in CIC knockdown and FGF4 overexpressing embryos with 27 genes upregulated and 9 downregulated in both sets of embryos (Figure 6A-B). This demonstrates that 44% of the genes upregulated and 26% of the genes downregulated in FGF4 overexpression were also upregulated and downregulated in CIC knockdown respectively (Figure 6A-B).

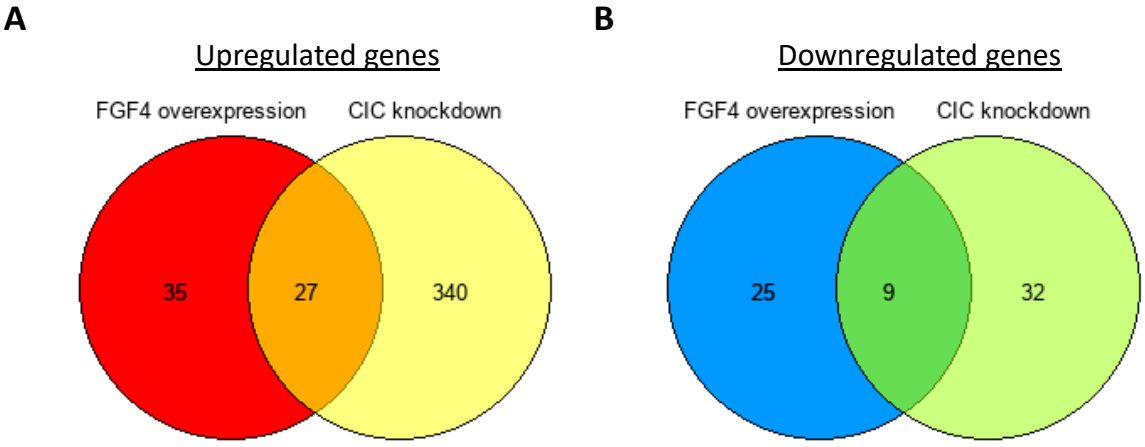


Figure 6: Venn diagrams showing the overlap between significantly upregulated and downregulated genes in CIC knockdown and FGF4 overexpressing *X. tropicalis* embryos.
Genes were identified in RNA-Seq analysis of embryos overexpressing FGF4 from the CSKA-FGF4 plasmid or knocked down for CIC using TALENs. (A) overlap of upregulated genes: RNA-Seq q-value < 0.1 and effect size > 1.75 (B) overlap of downregulated genes: RNA-Seq q-value < 0.1 and effect size < -1.75.

To investigate the statistical significance of the observed gene overlaps, Fisher's exact test was used to calculate the probability of these overlaps occurring by chance. The overlaps observed between upregulated and downregulated genes exhibit p-values of $< 9.4 \times 10^{-33}$ and $< 2.8 \times 10^{-18}$ respectively and are thus deemed highly significant. The statistical significance of the high proportion of genes

upregulated in FGF4 overexpression that are also upregulated in CIC knockdown supports the hypothesis that a subset of FGF target genes is transcriptionally regulated by CIC.

Table 7: Genes upregulated in both CIC knockdown and FGF4 overexpressing *X. tropicalis* embryos.

Genes identified as upregulated in RNA-Seq analysis of embryos overexpressing FGF4 from the CSKA-FGF4 plasmid or knocked down for CIC using TALENS. Genes with an RNA-Seq q-value < 0.1 and an effect size > 1.75 were classed as upregulated.

Gene	Capicua knockdown		FGF4 overexpression	
	q-value	Effect size	q-value	Effect size
adcy4	0.024	2.10	0.010	2.44
arrdc2	6.61×10^{-7}	2.34	1.51×10^{-7}	2.50
azin2	0.007	3.58	0.034	3.43
cbx4	0.012	2.22	6.07×10^{-4}	2.87
fam83c	5.34×10^{-4}	1.84	4.24×10^{-5}	2.08
fgd3	1.36×10^{-7}	4.54	1.25×10^{-5}	3.97
fos	2.20×10^{-15}	4.43	5.21×10^{-20}	5.38
frzb	5.67×10^{-4}	4.49	0.049	3.21
ier3	6.41×10^{-4}	2.63	0.015	2.33
lgals9c	4.54×10^{-7}	3.38	0.053	2.01
mmp1	3.12×10^{-5}	8.19	0.022	4.65
mmrn2	0.034	2.99	0.022	3.62
nfkbiz	0.002	2.12	0.017	2.02
rasl11b	2.79×10^{-8}	2.60	0.001	1.99
rgl2	2.94×10^{-16}	6.59	0.002	2.62
sgk1	2.46×10^{-4}	2.42	1.17×10^{-4}	2.66
smpdl3a	0.032	2.91	0.046	3.16
tnfrsf10b	2.09×10^{-12}	4.52	0.063	1.92
trim2	0.032	2.09	0.060	2.15
wnt8a	0.030	1.83	0.005	2.20
LOC100486038	0.025	2.24	0.009	2.66
LOC101730746	0.029	1.80	0.018	1.99
LOC101731765	0.054	1.76	3.02×10^{-4}	2.51
LOC101733948	0.039	2.22	0.015	2.68
LOC105945272	0.053	3.24	0.095	3.44
LOC105945708	0.022	1.96	1.97×10^{-4}	2.75
LOC105947813	1.48×10^{-5}	5.11	1.15×10^{-4}	4.95

Of the 27 genes upregulated in both FGF4 overexpression and CIC knockdown (Table 7), the most significant increases in expression upon CIC knockdown were demonstrated by *rgl2*, *fos*, *tnfrsf10b*, and *rasl11b*, which exhibit q-values between $2.94x10^{-16}$ and $2.79x10^{-8}$. Of the 9 genes downregulated in both FGF4 overexpression and CIC knockdown (Table 8), the most significant decreases in expression upon CIC knockdown were demonstrated by *spib*, *cebpa*, *pax6* and *cygb*. These genes exhibit a much higher range of q-values between 0.002 and 0.011. CIC knockdown mediating such highly significant upregulation of some genes also upregulated in FGF4 overexpression also supports the hypothesis that a subset of FGF targets is transcriptionally regulated by CIC.

Table 8: Genes downregulated in both CIC knockdown and FGF4 overexpressing *X. tropicalis* embryos.
Genes identified as downregulated in RNA-Seq analysis of embryos overexpressing FGF4 from the CSKA-FGF4 plasmid or knocked down for CIC using TALENS. Genes with an RNA-Seq q-value < 0.1 and an effect size < -1.75 were classed as downregulated.

Gene	Capicua knockdown		FGF4 overexpression	
	q-value	Effect size	q-value	Effect size
axl	0.013	-1.76	0.044	-1.75
cebpa	0.003	-2.76	$1.97x10^{-4}$	-3.61
cygb	0.011	-1.83	0.004	-2.08
nkain1	0.016	-1.81	0.036	-1.83
pax6	0.005	-3.22	0.005	-3.62
pdp2	0.044	-2.30	0.065	-2.45
rasgef1a	0.025	-2.15	0.064	-2.16
slc7a2.1	0.016	-2.17	0.051	-2.14
spib	0.002	-5.43	$2.62x10^{-4}$	-7.37

3.2.4 The temporal expression patterns of genes upregulated upon CIC knockdown and FGF overexpression

To analyse the temporal expression profiles of genes upregulated by both FGF4 overexpression and CIC knockdown, upregulated genes were mapped to their temporal expression data from a previous RNA-Seq study (Owens et al., 2016) and their relative expression from 4-24 hours post fertilisation (hpf) was presented by heat map (Figure 7).

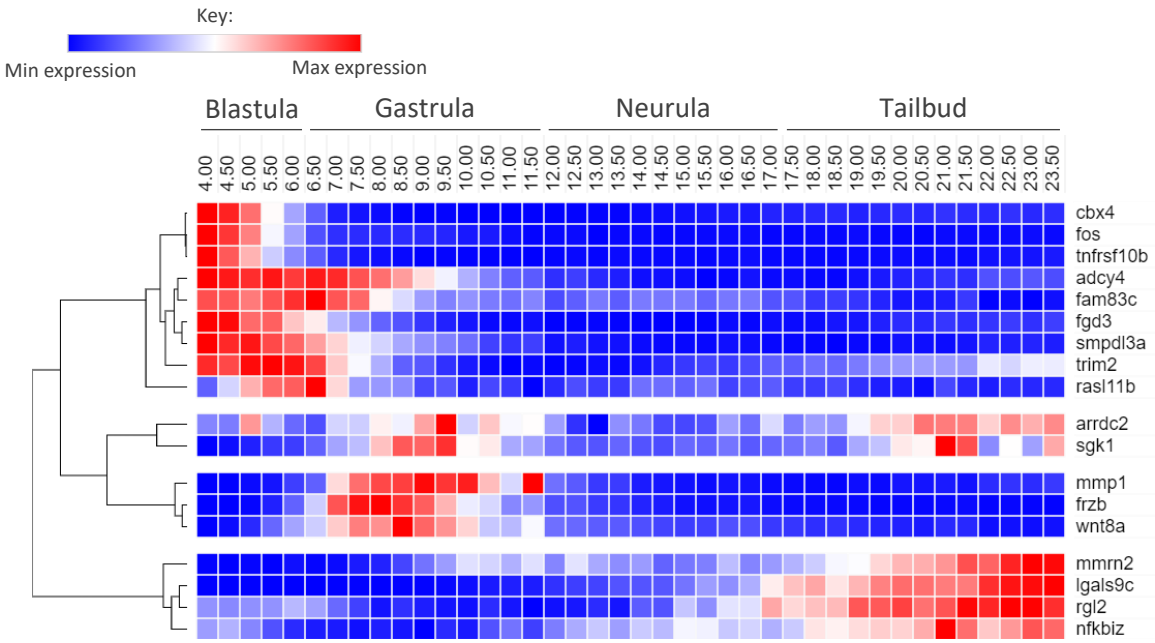


Figure 7: Temporal expression patterns of genes upregulated by both CIC knockdown and FGF4 overexpression throughout early X. tropicalis development.
Heat map showing the temporal expression patterns of genes identified as upregulated by FGF4 overexpression and CIC knockdown from 4-24 hours post fertilisation. Genes were identified in RNA-Seq analysis of embryos overexpressing FGF4 from the CSKA-FGF4 plasmid or knocked down for CIC using TALENs. Genes with an RNA-Seq q-value < 0.1 and an effect size > 1.75 were classed as upregulated. Temporal expression profiles for each gene were generated using data from a previous RNA-Seq experiment by Owens et al. (2016).

Genes upregulated upon both CIC knockdown and FGF4 overexpression group into 4 hierarchical clusters based on the similarity of their temporal expression patterns (Figure 7). Each gene cluster exhibits peak expression levels at different stages of development demonstrating that the expression of these genes is not confined to a common developmental period (Figure 7). For example, the largest gene cluster exhibits peak expression at blastula stages, the second largest at tailbud stages, and the third largest at gastrula stages.

Genes upregulated by both CIC knockdown and FGF4 overexpression also demonstrate dynamic expression patterns throughout early *X. tropicalis* development (Figure 7). Genes within the largest cluster exhibit rapid and marked changes in expression within as few 30 minutes, for example *cbx4* and *fos*. The smallest cluster of genes, *arrdc2* and *sgk1*, show highly variable temporal expression patterns, these genes alternative between high and low-level expression rapidly at multiple stages of development (Figure 7). The dynamic and rapid changes in gene expression demonstrated by genes upregulated by both CIC knockdown and FGF4 overexpression supports the notion that these genes are regulated by a labile transcriptional repressor.

3.3 Discussion

3.3.1 CIC knockdown upregulates notably more genes than FGF4 overexpression

Gene-level transcriptomic analysis showed that CIC knockdown mediated the upregulation of almost sixfold more genes than FGF4 overexpression. The upregulation of genes following CIC knockdown is supportive of the notion that CIC acts a transcriptional repressor. However, given that vastly more genes were upregulated by CIC knockdown than FGF4 overexpression, it is likely that in addition to FGF signalling, CIC transcriptional repression may be alleviated by FGF independent signalling pathways. For example, ERK is activated in the wound response following the rapid influx of calcium ions (Dieckgraefe et al., 1997, Matsubayashi et al., 2004, Yang et al., 2004) and may relieve CIC transcriptional repression of genes required for this FGF independent process following injury.

3.3.2 CIC knockdown and FGF4 overexpression upregulate genes involved in FGF-mediated biological processes

PANTHER gene ontology enrichment analysis revealed that genes significantly upregulated in FGF4 overexpressing embryos were associated with biological processes that are consistent with the downstream effects of FGF signalling. The processes most enriched for these genes include the negative regulation of intracellular signal transduction, regulation of cell population proliferation and cellular response to growth factor stimulus. FGF ligands are well established growth factors known to mediate cell proliferation through the activation of the FGF signalling pathway (Laestander and Engström, 2014, Yun et al., 2010). FGF signalling is also known to transcriptionally induce negative regulators of the pathway including *dusp6* and *MKP1* to provide negative feedback (Lewis et al., 1995, Ekerot et al., 2008). PANTHER enrichment analysis therefore demonstrates that genes upregulated by FGF4 overexpression are involved in FGF-mediated biological processes which supports the hypothesis that these genes are FGF targets.

PANTHER GO enrichment analysis demonstrated that genes significantly upregulated in CIC knockdown embryos were associated with biological processes including the regulation of cell cycle, cell cycle, cellular catabolic process, and regulation of metabolic process. FGF has been shown to directly regulate cell cycle progression (Lobjois et al., 2004, Zhang et al., 2004) and a variety of metabolic and catabolic processes in the adult (Yan et al., 2011, Nies et al., 2015, Tezze et al., 2019). This demonstrates that genes upregulated by CIC knockdown are involved in processes which can be mediated by FGF, thus supporting the notion that some of these genes are FGF targets. The notably higher number of genes upregulated by CIC knockdown than FGF4 overexpression likely

explains the difference in biological processes associated with upregulated genes from each set of embryos, since it is apparent that CIC transcriptional repression is not just confined to FGF targets.

3.3.3 Overlap between upregulated and downregulated genes in CIC knockdown and FGF overexpressing embryos

Gene-level RNA-Seq analysis showed highly significant overlap between changes in gene expression in CIC knockdown and FGF4 overexpressing embryos, but not all genes were upregulated or downregulated in both sets of embryos. Of the 367 genes upregulated in CIC knockdown embryos, 340 were not upregulated in FGF4 overexpression. As previously discussed, these 340 genes likely represent genes whose transcriptional repression by CIC is alleviated by FGF independent signalling pathways. Of the 62 genes upregulated by FGF4 overexpression, 35 were not upregulated by CIC knockdown. These genes are likely downstream targets of the PLCγ and PI3K/AKT signalling pathways, which are other two main pathways, excluding the MAPK/ERK cascade, through which FGF signal transduction occurs (Schlessinger, 2000). However, the highly significant overlap of genes upregulated in both FGF4 overexpression and CIC knockdown supports the hypothesis that a subset of FGF target genes is transcriptionally regulated by CIC.

3.3.4 Temporal expression patterns of genes upregulated upon CIC knockdown and FGF overexpression

Genes upregulated upon both CIC knockdown and FGF4 overexpression exhibit peak expression levels at different stages of development, which demonstrates that the expression of these genes is not confined to a common developmental period. This suggests that CIC transcriptional repression is not restricted to genes involved in a specific developmental process, for example, gastrulation. Additionally, the rapid and dynamic changes in expression reflected by some genes is consistent with rapid gene activation following the alleviation of a labile transcriptional repressor. This is supported by the observation that the immediate early gene *fos was* upregulated following CIC knockdown and FGF4 overexpression, which can be transcribed within minutes of stimulation (Hoffman et al., 1993).

Chapter 4: Investigating the expression pattern of *rasl11b* and its regulation by the FGF signalling pathway

4.1 Introduction

As demonstrated in chapter 3, gene-level transcriptomic analysis revealed a statistically significant overlap between genes upregulated in CIC knockdown and FGF4 overexpressing *X. tropicalis* embryos. Of these genes, *rasl11b* is one of the most significantly upregulated in CIC knockdown, presenting a q-value of $< 2.79 \times 10^{-8}$. In FGF4 overexpression, *rasl11b* also demonstrated highly significant upregulation with a q-value of < 0.001. Collectively, this evidence supports the hypothesis that *rasl11b* is an FGF target gene that is transcriptionally regulated by CIC.

Rasl11b is a member of the Ras family of small GTPases with a currently uncharacterised role in development (Colicelli, 2004). The small GTPase is largely understudied, with present research reflecting a limited understanding of the regulation and expression of the protein. However, one study has demonstrated that *rasl11b* expression has both a maternal and zygotic component, and that expression of the GTPase is dynamic throughout zebrafish development (Pézeron et al., 2008). At late blastula (dome) stage, *rasl11b* is expressed ubiquitously within the blastoderm (Pézeron et al., 2008). Then, at mid-gastrula (shield) stage, *rasl11b* expression is enriched in the embryonic shield (Pézeron et al., 2008), a structure considered orthologous to Spemann's organiser (Oppenheimer, 1936, Spemann and Mangold, 1924). This is in contrast to *Xenopus*, where *rasl11b* is expressed in the marginal zone (Hufton et al., 2006), with some reports indicating enrichment in the ventral region (Popov et al., 2017). Later in zebrafish development, during somitogenesis and organogenesis, *rasl11b* is expressed in the tailbud, tailtip, posterior spinal cord and several head structures including the forebrain, hindbrain, and isthmic organiser (Pézeron et al., 2008). Whereas, in *Xenopus* at mid-tailbud stage, *rasl11b* is expressed in the posterior presomitic mesoderm (Dickinson et al., 2006). By zebrafish larval stage, *rasl11b* is expressed exclusively in the otic vesicles (Pézeron et al., 2008).

Members of the TGF-β (transforming growth factor-beta) family of signalling molecules have been shown to regulate *rasl11b* expression in multiple types of mesodermally derived adult cells (Stolle et al., 2007, Luo et al., 2020). During zebrafish development, *rasl11b* expression in the embryonic shield is modulated by the TGF-β/Nodal signalling pathway. This was exemplified in shield stage embryos, where overexpression of the activated TGF-β type 1 receptor Taram-A was able to induce ectopic expression of *rasl11b* (Pézeron et al., 2008). In contrast, nodal signalling deficient mutants which lack both maternal and zygotic components of *oep* (*one-eyed pinhead*), show a complete loss

of *rasl11b* expression in the embryonic shield (Pézeron et al., 2008). However, mutants for other components of the TGF-β/Nodal signalling pathway demonstrate minimal change in *rasl11b* expression (Pézeron et al., 2008). Therefore, regulation of *rasl11b* in zebrafish development is specifically dependent upon signalling via the *oep* coreceptor, the orthologue of *tdgf1.3* in *Xenopus,* through a likely uncharacterised branch of the TGF-β/nodal signalling pathway.

A preliminary investigation into *rasl11b* expression during *Xenopus* development was undertaken by Cowell (2019), although this data has not been published. In addition to this, gene-level transcriptomic analysis suggests that *rasl11b* is regulated by FGF signalling, and that some domains of *rasl11b* expression appear to correspond to regions of FGF activity. For example, *rasl11b* is expressed in the *Xenopus* marginal zone and presomitic mesoderm, both of which are known regions of FGF4 activity (Isaacs et al., 1995, Lea et al., 2009). Therefore, using the *Xenopus* developmental model, this chapter presents a stage series of *rasl11b* expression analysis from early neurula to late tailbud stages of development. A sequence of experiments is also presented in which FGF and MAPK/ERK signalling is manipulated to investigate whether *rasl11b* is regulated by FGF.

The aims of this chapter are:

- Define the domains of *rasl11b* expression during neurula and tailbud stages of *X. tropicalis* development
- Determine whether domains of *rasl11b* expression correspond to regions of FGF activity in neurula and tailbud stage embryos
- Explore the effects of chemical inhibitors of FGF receptor and MEK signalling on *rasl11b* expression in gastrula and neurula stage embryos
- Analyse the differential expression of *rasl11b* in gastrula stage *X. laevis* embryos overexpressing the dominant negative FGF receptor and untreated controls
- Investigate the effect of FGF4 injection on the spatial expression pattern of *rasl11b* during neurula stages
- Identify putative CIC binding sites within *X. tropicalis*, mouse and human *rasl11b* genes

4.2 Results

4.2.1 Analysing *rasl11b* expression during neurula and tailbud stages of *X. tropicalis* development

To define the domains of *rasl11b* expression, in-situ hybridisation was performed to analyse the expression of *rasl11b* during neurula and tailbud stages of *X. tropicalis* development (Figure 8-9). The spatiotemporal expression pattern of *rasl11b* largely resembles that of the segmentally expressed gene *ripply2.2* (*ripply2 homolog, gene 2*), specifically within the presomitic mesoderm (PSM) (Figure 8-9). *ripply2.2* is required during somitogenesis for the formation of somite boundaries and is expressed in PSM at the anterior halves of the S-I to S-III somitomeres (Kondow et al., 2007, Hitachi et al., 2009). The expression pattern of *ripply2.2* has been well characterised in *Xenopus* development, hence its visualisation here by in-situ hybridisation (Figure 8-9), serves to validate our assignment of *rasl11b* expression to the correct embryonic domains. Given that FGF signals via the MAPK/ERK pathway, regions of active FGF signalling are indicated by the activated form of ERK (dpERK). Therefore, to determine whether domains of *rasl11b* expression correspond to regions of FGF activity, in-situ hybridisations for *rasl11b* were compared to immunostaining for dpERK at each stage of development (Figure 8-9).

At neurula stage 13, *rasl11b* is expressed in a ring around the closing blastopore, which corresponds to a domain of dpERK activity (Figure 8F and 8K). During neurula stages 13-19, *rasl11b* expression is considerably less widespread in the PSM than *ripply2.2*, which is expressed in three clear domains either side of the neural tube corresponding to somitomeres S-I to S-III (Figure 8A-C, 8F-H) (Hitachi et al., 2009). At these stages, *rasl11b* mRNA appears to overlap with two of these domains but extends much less laterally into the PSM than *ripply2.2* (Figure 8A-C, 8F-H). This overlapping expression suggests the *rasl11b* is expressed in two of the S-I to S-III somitomeres, though it is not clear which. Meanwhile, throughout neurula stages, a large region of active FGF signalling indicated by dpERK activity is detected in the posterior (Figure 8K-M). Whereas the two anterior regions of dpERK activity adjacent to the neural tube at stages 13-14, are later found in the dorsal midline at stage 19 (Figure 8K-M).

At early tailbud stage 22, *rasl11b* in the PSM appears to continue to overlap with two of the three domains of *ripply2.2* expression (Figure 8D and 8I). At this stage, rasl11b mRNA continues to be expressed around the closed blastopore and is also detected in the otic vesicles (Figure 8I). Stage 26 embryos show that *ripply2.2* and *rasl11b* continue to be expressed in the PSM, with some minor expression of *ripply2.2* detected in the endoderm (Figure 8E and 8J). *rasl11b* is also expressed in the otic vesicles with low-level expression present in the developing heart (Figure 8J). Throughout

early tailbud stages, dpERK is active in the tailbud, cement gland, heart anlagen and branchial arch region (Figure 8O).

During mid-tailbud stages 28-29, *rasl11b* continues to be expressed in similar regions of the PSM as *ripply2.2* (Figure 9A and 9D). At these stages, levels of *rasl11b* expression in the heart and otic vesicles are notably higher than at stage 26, with high levels of *rasl11b* mRNA now expressed in the branchial arches (Figure 9D). In addition, there is also widespread expression of *ripply2.2* in the endoderm at these stages (Figure 9A).

Later in development, from stages 31-36, *rasl11b* and *ripply2.2* continue to exhibit high levels of expression in the PSM, with *rasl11b* mRNA still detected in the branchial arches and otic vesicles at stage 36 (Figure 9B-C, 9E-F). During these stages, *ripply2.2* expression in the endoderm gradually decreases until at stage 36, where expression is confined to ventral surface of the embryo (Figure 9C). Throughout mid to late tailbud stages, dpERK activity increases in the embryo and is found in a variety of structures including the tailbud, cement gland, otic vesicles, branchial arches, notochord, and somites (Figure 9G-I). The overlap of some regions of dpERK activity, in areas of active FGF signalling, with domains of *rasl11b* expression supports the hypothesis that *rasl11b* is regulated by the FGF signalling pathway.

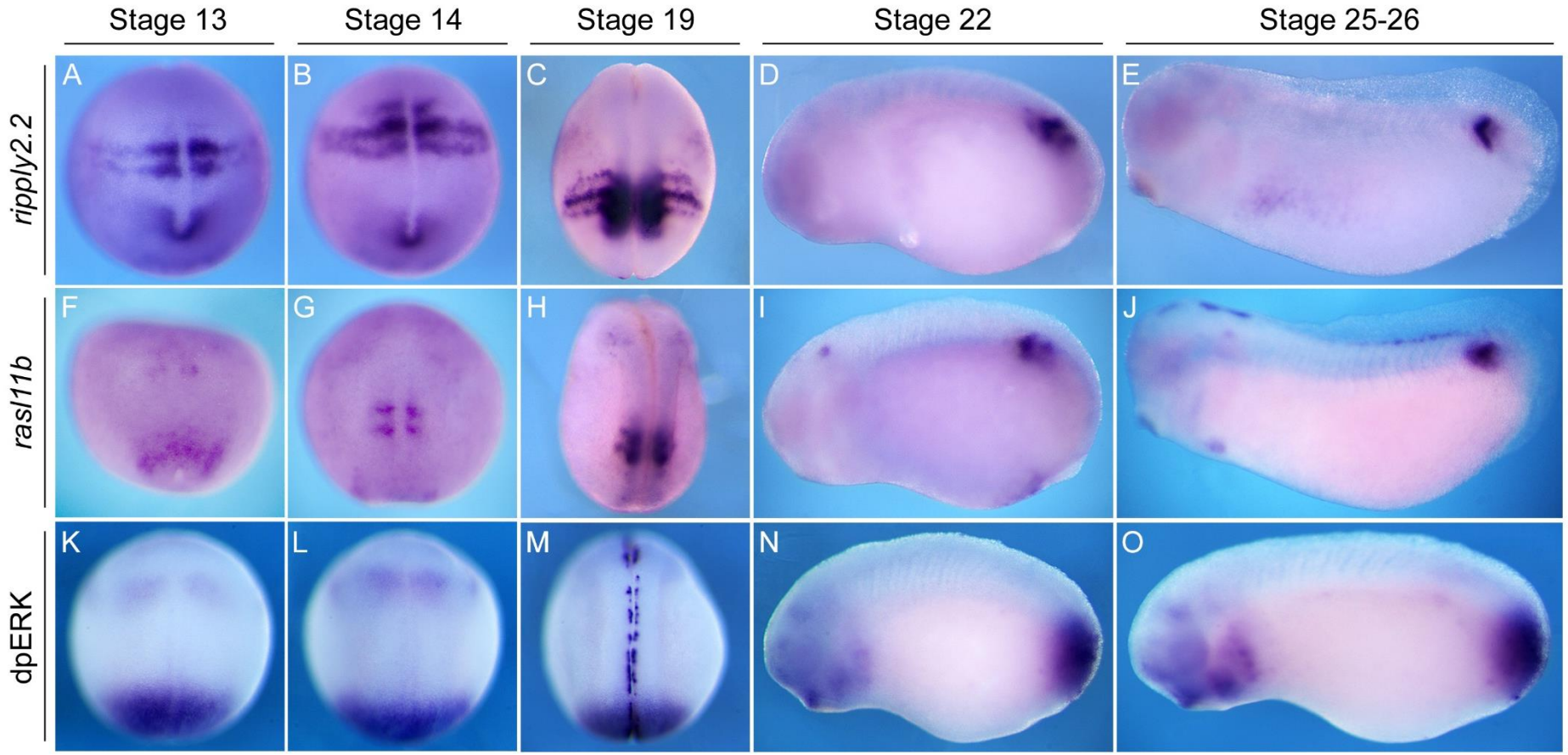


Figure 8: Spatiotemporal analysis of *ripply2.2*, *rasl11b* and dpERK during neurula and early tailbud stages of *X. tropicalis* development.
Expression patterns of (A-E) *ripply2.2* and (F-J) *rasl11b* was analysed by in-situ hybridisation and compared to regions of active FGF signalling indicated by immunostainings for dpERK (K-O). Embryos were analysed at stage 13-26. Stage 13-14 embryos show posterior-dorsal views, stage 19 embryos show dorsal views with anterior upwards, and stage 22-26 embryos show lateral views with anterior to left. Embryos were staged was according to Nieuwkoop and Faber (1994).

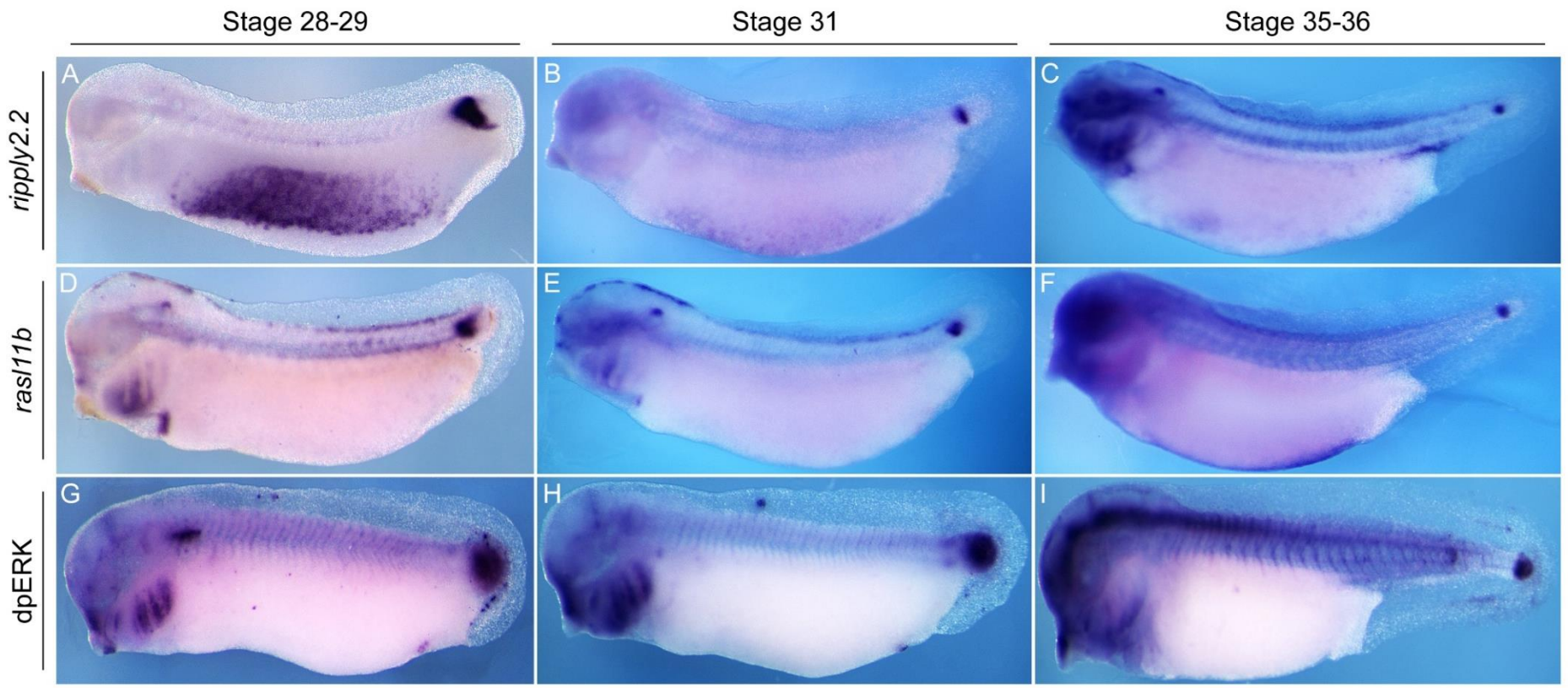


Figure 9: Satiotemporal analysis of *ripply2.2*, *rasl11b* and dpERK during mid to late tailbud stages of *X. tropicalis* development.
Expression of (A-C) *ripply2.2* and (D-F) *rasl11b* was analysed by in-situ hybridisation and compared to regions of active FGF signalling indicated by immunostainings for dpERK (G-I). Embryos were analysed at stages 28-36 and show lateral views with anterior to left. Embryos were staged according to Nieuwkoop and Faber (1994).

During *X. tropicalis* development, *rasl11b* and *ripply2.2* were shown to exhibit highly similar and seemingly overlapping expression patterns within the PSM. Therefore, to precisely define the domains of *rasl11b* expression, double in-situ hybridisation was performed to analyse the expression of both *rasl11b* and *ripply2.2* in neurula and tailbud stage *X. tropicalis* embryos (Figure 10). At neurula stage 19, *ripply2.2* is expressed in three domains either side of the neural tube known to correspond to the S-I to S-III somitomeres (Figure 10A) (Hitachi et al., 2009). In comparison at stage 19, *rasl11b* is expressed around the closed blastopore and in two domains on each side of the neural tube (Figure 10C). This contrasts with tailbud stage 26, when *ripply2.2* and *rasl11b* are both expressed in two domains adjacent to the neural tube (Figure 10B and 10D), with *rasl11b* appearing to be expressed strongly in a posterior domain and weakly in a more anterior domain of the PSM (Figure 10D).

Double in-situ hybridisation for both *rasl11b* and *ripply2.2* in stage 19 embryos demonstrates three domains of expression either side of the neurula tube (Figure 10E). The two more posterior domains of expression appear bolder and darker than those seen in analysis of *ripply2.2* expression alone (Figure 10A and 10E), indicating that *rasl11b* expression overlaps with *ripply2.2* at somitomeres S-II and S-III. Analysis of *ras11b* and *ripply2.2* in stage 26 embryos demonstrates two domains of expression adjacent to the neurula tube (Figure 10F), both of which appear bolder and darker than those seen in analysis of *rasl11b* and *ripply2.2* respectively (Figure 10B, 10D, 10F). This indicates that expression of *rasl11b* and *ripply2.2* in the PSM overlap entirely at stage 26.

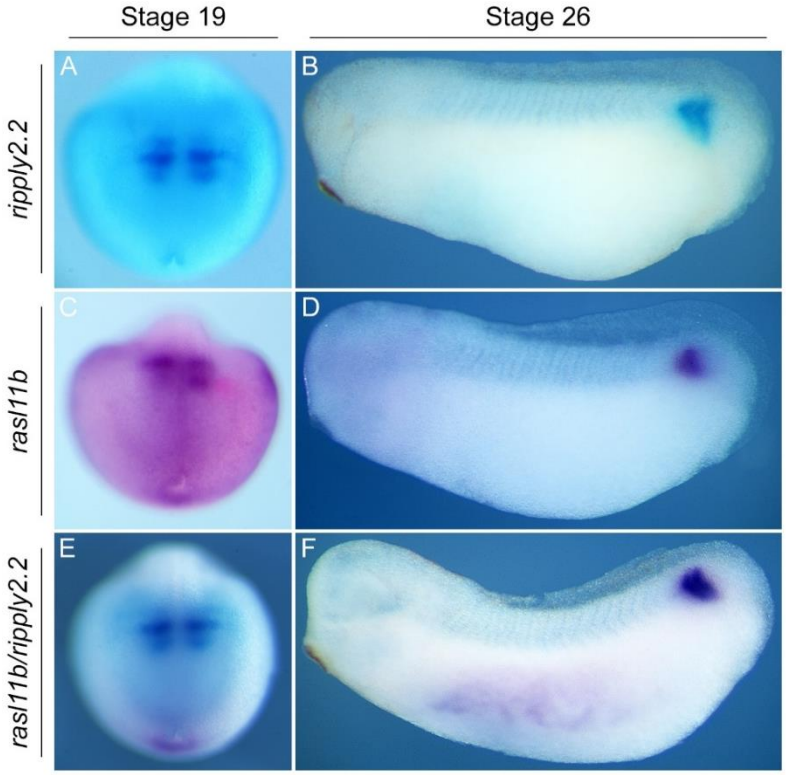


Figure 10: Overlap of *ripply2.2* and *rasl11b* expression in the presomitic mesoderm of neurula and tailbud stage *X. tropicalis* embryos.
Expression patterns of (A-B) *ripply2.2*, (C-D) *rasl11b* and (E-F) both *rasl11b* and *ripply2.2* were analysed by in-situ hybridisation in stage 19 and stage 26 embryos. Domains *ripply2.2* and *rasl11b* were marked by BCIP (blue) and Magenta Phosphate (magenta) respectively. Stage 19 embryos show posterior-dorsal views and stage 26 embryos show lateral views with anterior to left. Embryos were staged according to Nieuwkoop and Faber (1994).

4.2.2 Investigating *rasl11b* regulation via manipulation of FGF and MAPK/ERK signalling

To attempt to elucidate whether *rasl11b* expression is regulated by FGF signalling via the MAPK/ERK pathway during gastrulation, gastrula stage *X. tropicalis* embryos were treated with chemical inhibitors of FGFR and MEK signalling respectively. SU5402 is an inhibitor of FGFR1 signalling and inhibits signal transduction by most FGF ligands expressed during gastrulation (Mohammadi et al., 1997, Lea et al., 2009). PD0325901 is an inhibitor of MEK signalling and inhibits signal transduction via the MAPK/ERK pathway (Anastasaki et al., 2012, Sebolt-Leopold and Herrera, 2004). The expression of *rasl11b* in embryos treated with each inhibitor was analysed using in-situ hybridisation. This was compared to dpERK immunostaining which demonstrate the effect of each inhibitor on the downstream effector of FGF signalling.

Gastrula stage 10.5 embryos treated with normal growth media (MRS/20) or vehicle (DMSO) demonstrate a ring of dpERK activity around the blastopore (Figure 11A-B, 11E-F), corresponding to a region of active FGF signalling in the early mesoderm (Isaacs et al., 1995, Isaacs et al., 1994). As expected, treatment with MRS/20 or DMSO has no effect on the circumblastoporal expression of *rasl11b* in the early mesoderm (Figure 11I-J, 11M-N). However, treatment with either SU5402 or PD0325901 leads to a complete loss of dpERK activity around the blastopore (Figure 11C-D, 11G-H) indicating complete inhibition of FGF signalling via the MAPK/ERK pathway. Promisingly, stage 10.5 embryos treated with SU5402 show a total loss of *rasl11b* expression around the blastopore (Figure 11K and 11O). Furthermore, two thirds of embryos treated with PD0325901 demonstrate complete inhibition of *rasl11b* expression in the early mesoderm (Figure 11L and 11P) with the remaining embryos showing a large reduction in *rasl11b* expression around the blastopore (Figure 11P). Taken together, these results support the hypothesis that during gastrulation *rasl11b* expression is regulated by FGF signalling via the MAPK/ERK pathway.

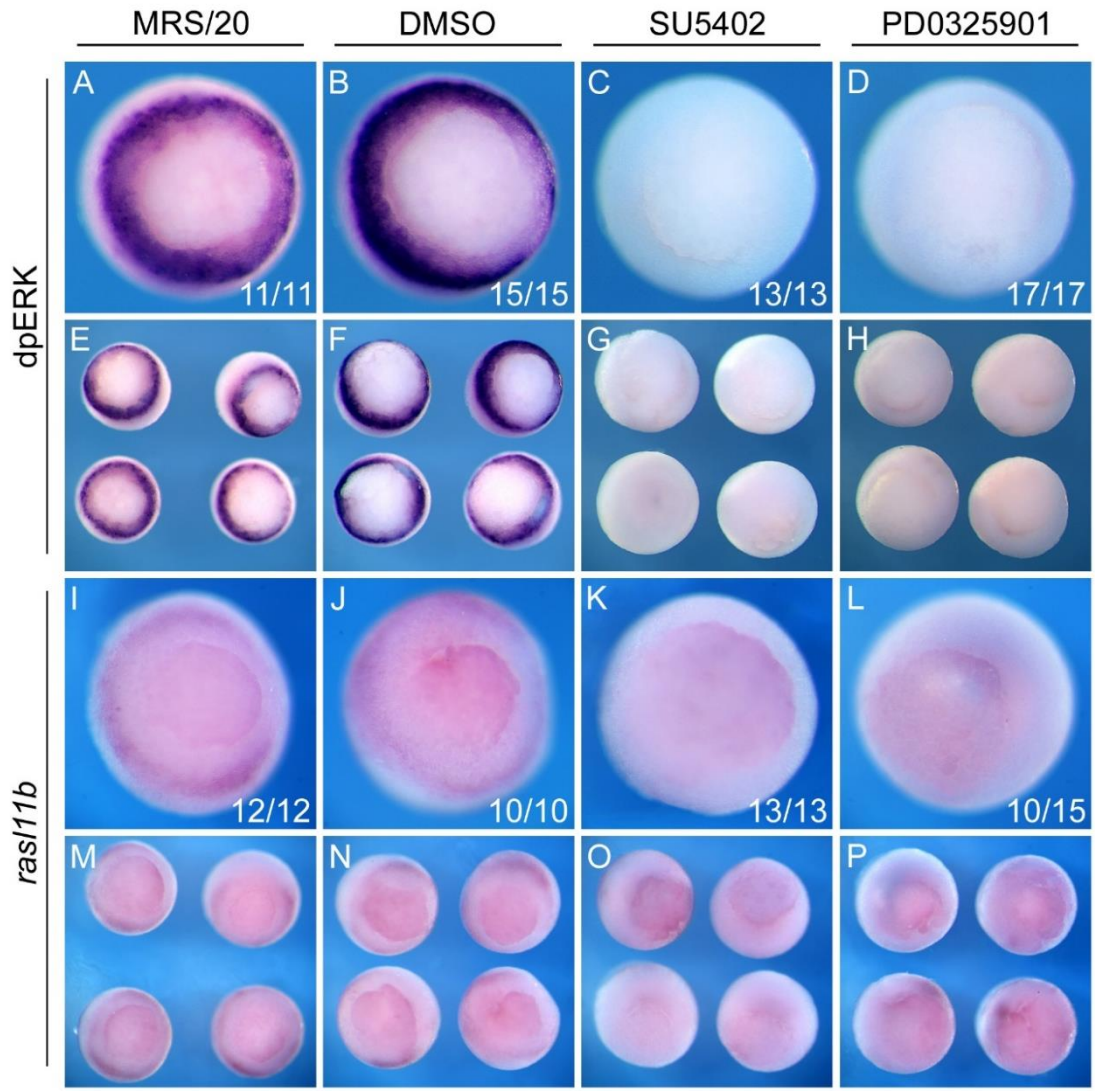


Figure 11: Gastrula stage *X. tropicalis* embryos treated with FGFR and MEK inhibitors analysed by immunostaining for dpERK and in-situ hybridisation for *rasl11b*.
Embryos were treated with MRS/20, 0.2% DMSO, 200µM FGFR inhibitor (SU5402) or 25µM MEK inhibitor (PD0325901) for 2 hours. At stage 10.5, embryos were analysed for *rasl11b* expression by in-situ hybridisation and for dpERK activity by immunostaining. Expression of *rasl11b* was compared to dpERK immunostainings which demonstrate the effect of each inhibitor on the downstream effector of the FGF signalling pathway. Embryos show vegetal views and were staged according to Nieuwkoop and Faber (1994).

To attempt to further support the notion that *rasl11b* is an FGF target gene, the expression of *rasl11b* in *X. laevis* embryos overexpressing a dominant negative FGFR (XFD) was compared to untreated controls. XFD is a truncated form of the FGFR which lacks catalytic activity and forms non-functional heterodimers with endogenous FGFRs to effectively inhibit FGF signalling (Amaya et al., 1993, Amaya et al., 1991). *X. laevis* embryos injected with XFD mRNA at the two-cell stage were analysed by RT-qPCR for the expression of *rasl11b* and *tbxt* at gastrula stages 10.5 and 12, with expression normalised to the housekeeping gene *dicer1* (Figure 12A and 12B). Expression was analysed at late gastrula stage 12 to mitigate for the possible confounding effect of maternal

rasl11b contributing to zygotic expression of the gene at stage 10.5. Given that FGF signalling is required for the maintenance of *tbxt* expression during gastrulation (Isaacs et al., 1994), *tbxt* expression was used to demonstrate the efficacy of XFD at inhibiting FGF signalling.

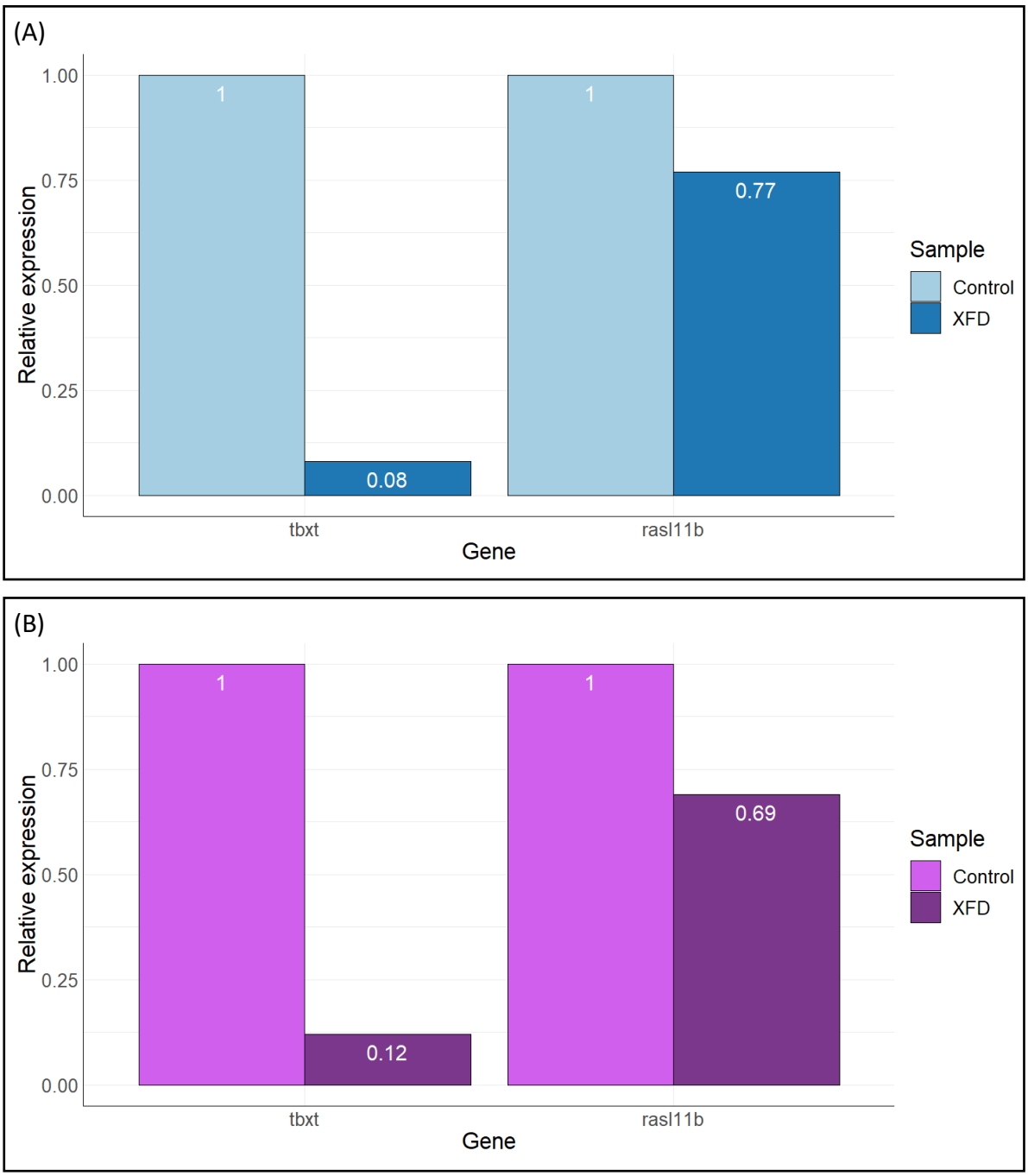


Figure 12: Differential expression of *tbxt* and *rasl11b* in gastrula stage *X. laevis* embryos overexpressing a dominant negative FGF receptor.
RT-qPCR was used to analyse the differential expression of *tbxt* and *rasl11b* in (A) stage 10.5 and (B) stage 12 embryos injected with dominant negative FGF receptor (XFD) relative to uninjected controls. Expression was normalised to the housekeeping gene *dicer1*. Embryos were injected with XFD mRNA at the two-cell stage.

At gastrula stages, XFD overexpression dramatically downregulates *tbxt*, with XFD injected *X. laevis* embryos exhibiting 92 and 88% reductions in *tbxt* expression relative to uninjected controls at stages 10.5 and 12 respectively (Figure 12A and 12B). This indicates that XFD overexpression was highly effective at inhibiting FGF signalling during gastrulation. Promisingly, *rasl11b* expression was also downregulated in XFD overexpressing embryos which show 23 and 31% reductions in *rasl11b* expression relative to uninjected controls at stages 10.5 and 12 respectively. The substantial downregulation of *rasl11b* seen by inhibiting FGF signalling during gastrulation supports the notion that *rasl11b* expression is at least, in part, regulated by FGF signalling during gastrula stages.

To attempt to elucidate whether *rasl11b* expression is regulated by FGF signalling via the MAPK/ERK pathway later in development, neurula stage *X. tropicalis* embryos were treated with the same chemical inhibitors of FGFR and MEK signalling as previously described. The expression of *rasl11b* in embryos treated with each inhibitor was analysed using in-situ hybridisation. This was compared to dpERK immunostainings which demonstrate the effect of each inhibitor on the downstream effector of FGF signalling.

Neurula stage 19 embryos treated with MRS/20 or DMSO demonstrate regions of dpERK activity along the dorsal midline (Figure 13A and 13B) and in the posterior surrounding the closed blastopore (Figure 13E and 13F). These regions of dpERK activity are known to correspond to domains of active FGF signalling (Isaacs et al., 1995, Isaacs et al., 1994). As expected, treatment with MRS/20 or DMSO has no effect on the expression of *rasl11b* in the PSM (Figure 13I and 13J) or around the closed blastopore (Figure 13M and 13N). However, treatment with SU5402 has no effect on ERK activation along the dorsal midline (Figure 13C) but leads to a complete loss of dpERK activity in the posterior domain (Figure 13G). In contrast, embryos treated with PD0325901 show no activation of ERK in either the dorsal midline or the posterior (Figure 13D and 13H). This indicates that dpERK activity in the posterior is regulated by FGF signalling, whereas activation of ERK along the dorsal midline is via an FGF independent signalling pathway.

Embryos treated with either SU5402 or PD0325901 demonstrate no change in *rasl11b* expression in the PSM (Figure 13K and 13L) or around the closed blastopore (Figure 13O and 13P) relative to MRS/20 controls (Figure 13I and 13M). This demonstrates that during neurula stages *rasl11b* expression can be regulated independently of both FGF signalling and the MAPK/ERK signalling pathway.

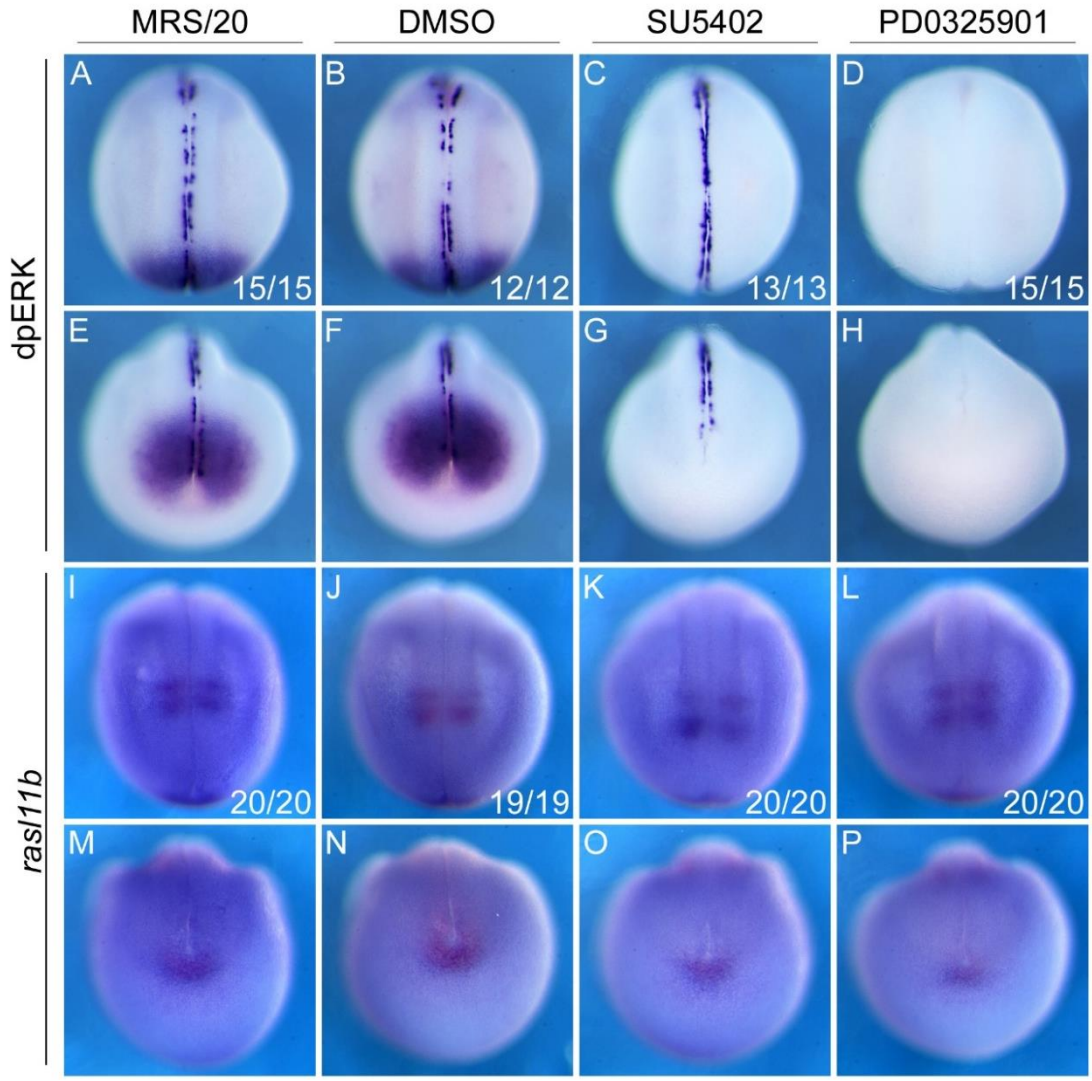


Figure 13: Neurula stage 19 *X. tropicalis* embryos treated with FGFR and MEK inhibitors analysed by immunostaining for dpERK and in-situ hybridisation for *rasl11b*.
Embryos were treated with MRS/20, 0.2% DMSO, 200µM FGFR inhibitor (SU5402) or 25µM MEK inhibitor (PD0325901) for 2 hours. At stage 19, embryos were analysed for *rasl11b* expression by in-situ hybridisation and for dpERK activity by immunostaining. Expression of *rasl11b* was compared to dpERK immunostainings which demonstrate the effect of each inhibitor on the downstream effector of the FGF signalling pathway. Embryos show dorsal (A-D, I-L) and posterior (E-H, M-P) views and were staged according to Nieuwkoop and Faber (1994).

As shown previously, gene-level transcriptomic analysis revealed that *rasl11b* was significantly upregulated in FGF4 overexpressing and CIC knockdown *X. tropicalis* embryos at neurula stage 14. To attempt to validate these results and investigate the effect of FGF4 overexpression on the spatial expression pattern of *rasl11b* at neurula stages, *X. tropicalis* embryos at the two-cell stage were unilaterally injected with CSKA-FGF4 plasmid and analysed for the expression of *rasl11b* at stage 14. In-situ hybridisation was used to analyse *rasl11b* expression whereas immunostaining for dpERK was used to demonstrate the effect of FGF4 overexpression on the downstream effector of FGF signalling.

A pronounced upregulation of dpERK activity is seen upon FGF4 overexpression (Figure 14A and 14B). Embryos overexpressing FGF4 demonstrate an expansion of the posterior domain of dpERK activity, particularly on the injected side which shows widespread but mosaic activation of ERK along the AP axis (Figure 14A and 14B). This observation is consistent with the mosaicism of FGF4 expression from the CSKA plasmid (Isaacs et al., 1994), though given injection of the plasmid induced widespread ERK activation, FGF4 overexpression was clearly facilitated.

As previously shown at stage 14, *rasl11b* is expressed in the PSM in two somitomeres on each side of the neural stube, and in the posterior around the closed blastopore (Figure 8G and 14C). However, upon FGF4 injection, *rasl11b* expression appears stronger and more widespread in the posterior (Figure 14D, white arrows), with perhaps some minor upregulation in the PSM (Figure 14D). The apparent upregulation of *rasl11b* expression in stage 14 embryos supports the notion that *rasl11b* is a target of FGF signalling at neurula stages.

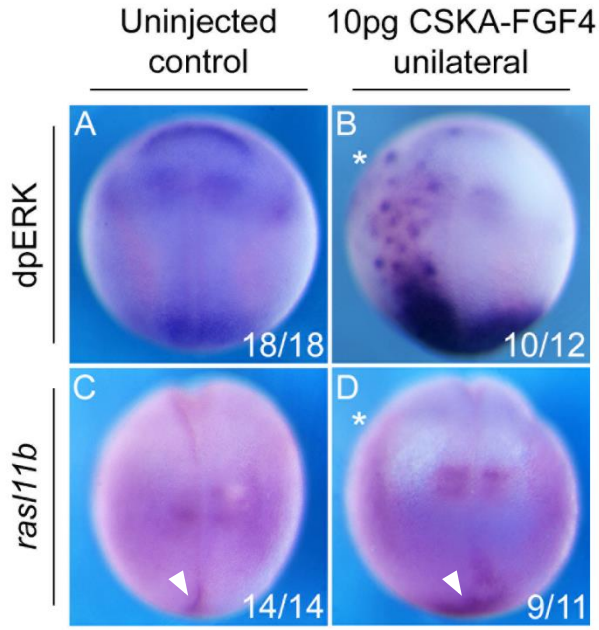


Figure 14: CSKA-FGF4 injected neurula stage 14 *X. tropicalis* embryos analysed for dpERK by immunostaining and in-situ hybridisation for *rasl11b*.
Two-cell stage *X. tropicalis* embryos were unilaterally injected with 10pg CSKA-FGF4. At stage 14, embryos were analysed for (A-B) dpERK activity by immunostaining and (C-D) *rasl11b* by in-situ hybridisation. Expression of *rasl11b* was compared to dpERK immunostainings which demonstrate the effect of FGF4 overexpression on the downstream effector of the FGF signalling pathway. Embryos show a dorsal view with anterior upwards and were staged according to Nieuwkoop and Faber (1994). Asterisks indicate the injected side and white arrowheads indicate posterior *rasl11b* expression.

To further investigate the effect of FGF4 overexpression on the spatial expression pattern of *rasl11b* during neurula stages, *X. tropicalis* embryos were injected with CSKA-FGF4 at the two-cell stage then analysed for the expression of *ripply2.2* and *rasl11b* by in-situ hybridisation at stage 19. The

expression of *rasl11b* was compared to *ripply2.2* to see the effect of FGF4 overexpression on *rasl11b* relative to a marker of somitomeres S-I to S-III (Hitachi et al., 2009), domains shown to partially overlap with wild-type expression of *rasl11b* (Figure 10E and 10F).

As previously shown, stage 19 embryos demonstrate clearly segmented domains of *rasl11b* and *ripply2.2* expression in the PSM (Figure 15A and 15B). However, upon FGF4 injection, both *rasl11b* and *ripply2.2* present the same two classes of expression pattern on the injected side of *X. tropicalis* embryos (Figure 15B-C, 15E-F). The first class of expression pattern shows no clear difference to uninjected controls (Figure 15C and 15F). However, the second class of expression pattern demonstrates disruption to individual domains of *rasl11b* and *ripply2.2* expression (Figure 15B and 15E). Given that *ripply2.2* marks individual somitomeres, disruption to *ripply2.2* expression likely corresponds to a disruption to somite segmentation. Since *rasl11b* expression appears to partially overlap with *ripply2.2* and was disrupted upon FGF4 overexpression, this suggests that FGF signalling either directly or indirectly regulates the spatial expression pattern of *rasl11b* in the PSM.

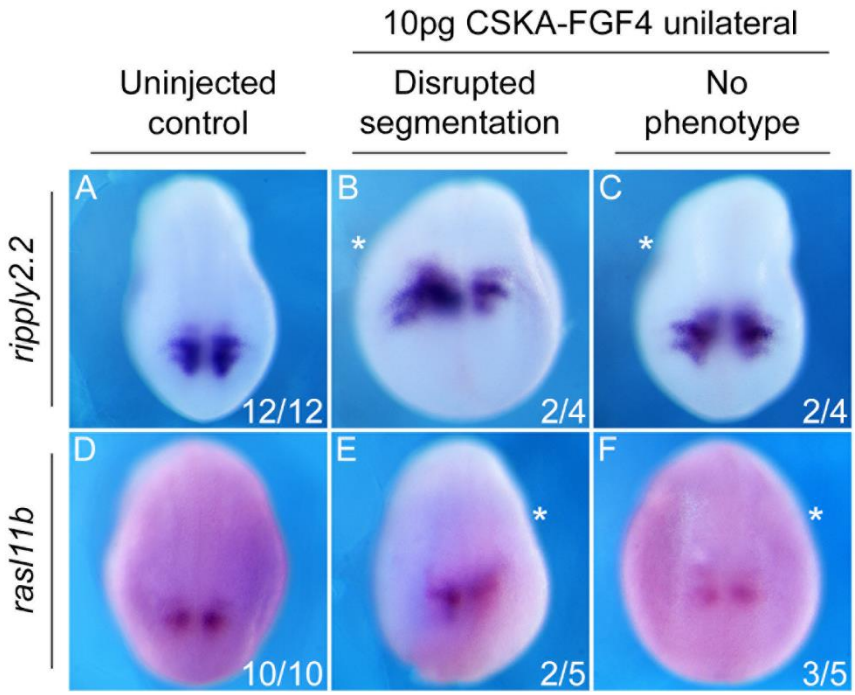


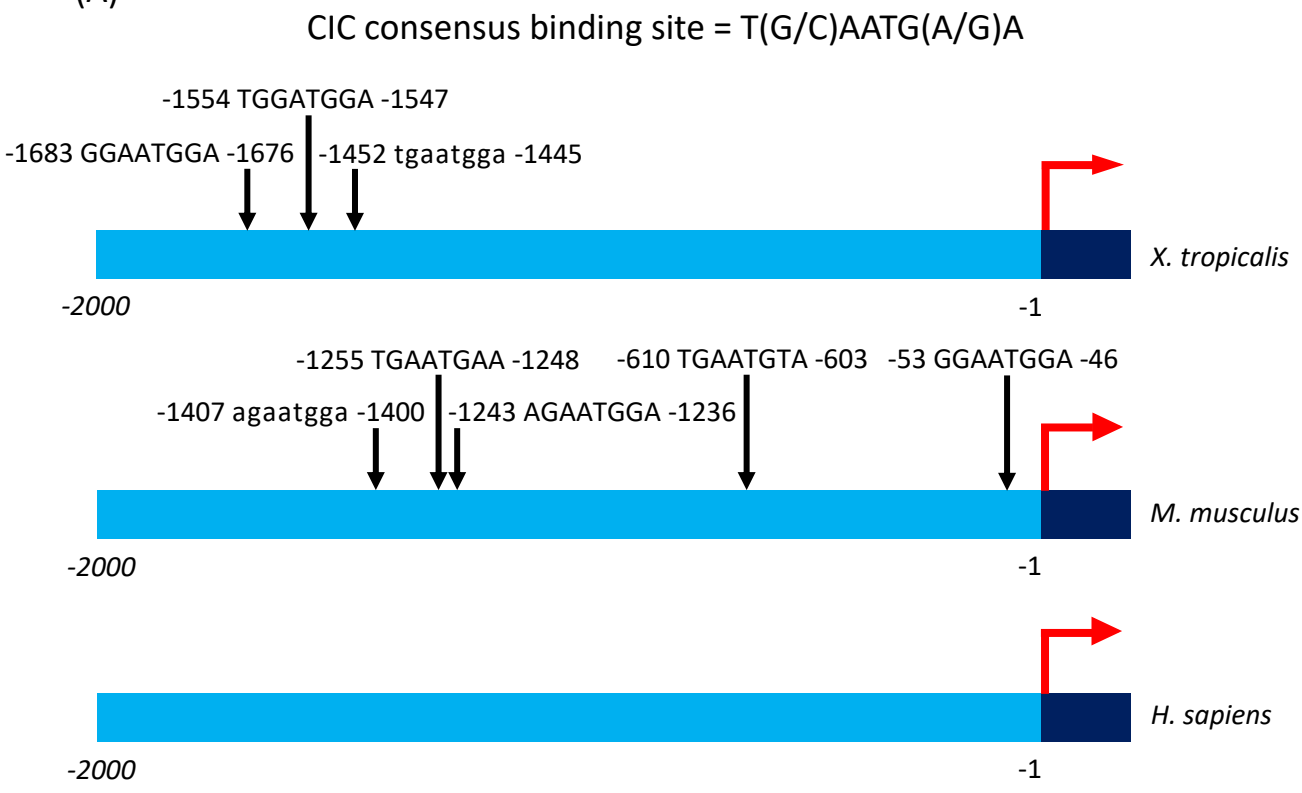
Figure 15: CSKA-FGF4 injected neurula stage 20 *X. tropicalis* embryos analysed by in-situ hybridisation for *ripply2.2* and *rasl11b*.
Two-cell stage *X. tropicalis* embryos were unilaterally injected with 10pg CSKA-FGF4. At stage 20, embryos were analysed by in-situ hybridisation for (A-C) *ripply2.2* and (D-F) *rasl11b*. Expression of *rasl11b* was compared to *ripply2.2* to see the effect of FGF4 overexpression on *rasl11b* relative to a marker of somitomeres S-I to S-III. Embryos show a dorsal view with anterior upwards and were staged according to Nieuwkoop and Faber (1994). Asterisks indicate the injected side.

4.2.3 Identification of CIC binding sites within *rasl11b* gene sequences

This chapter so far has presented evidence in attempt to substantiate the notion that *rasl11b* is an FGF target gene. However, this forms only part of our hypothesis, which states that *rasl11b* is also transcriptionally repressed by CIC in the absence of ERK activation. Therefore, given that CIC has been shown to bind octameric T(G/C)AATG(A/G)A sites in the promoters and enhancers of target genes (Jiménez et al., 2012), the FIMO tool in MEME Suite 5.4.1 (Bailey et al., 2015, Grant et al., 2011) was used to identify these sequences within *X. tropicalis*, mouse (*M. musculus*) and human (*H. sapiens*) *rasl11b* genes (Figure 16). FIMO uses a dynamic programming algorithm to convert log-odds scores into p-values, assuming a zero-order background model. Only sequences with a p-value < 0.001 were identified.

Analysis of the 2kb upstream region of *X. tropicalis* and mouse *rasl11b* genes identified 3 and 5 putative CIC binding sites respectively, with no such sites found in this region in humans (Figure 16A). However, in analysis of the transcribed region of *rasl11b* at least one putative CIC binding site was identified in each of these vertebrates (Figure 16B). In the 2kb downstream region of *rasl11b*, 4 CIC binding sites were identified in *X. tropicalis,* whereas 3 sites were found in mouse and humans (Figure 16C). The identification of putative CIC binding sites in *rasl11b* genes of multiple vertebrates supports the notion that CIC binds *rasl11b* to transcriptionally regulate its expression.

(A)



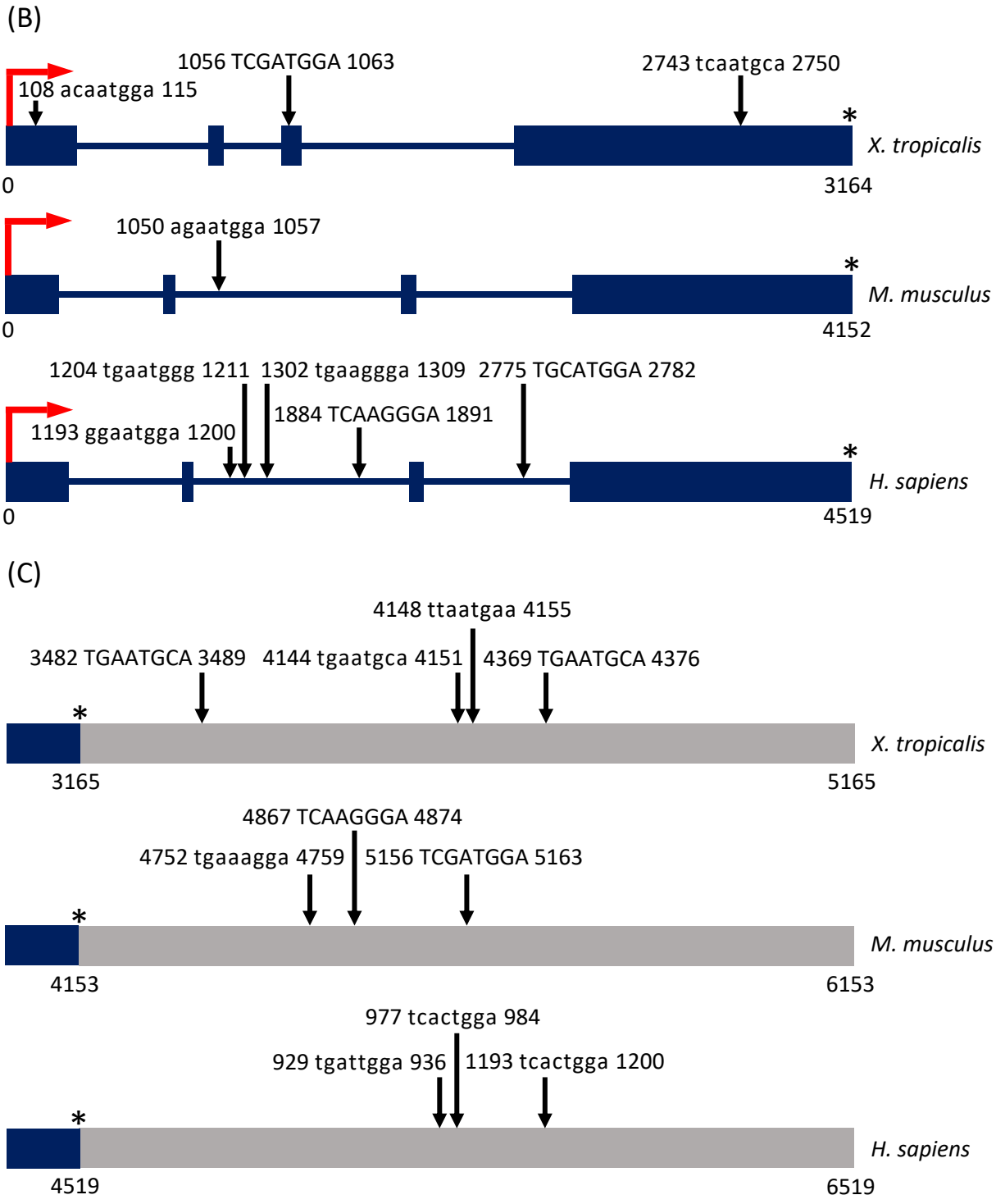


Figure 16: Schematic representation of putative CIC binding sites within different regions of *X. tropicalis*, *M. musculus* and *H. sapiens rasl11b* genes.

The FIMO (Find Individual Motif Occurrences) tool in MEME Suite 5.4.1 was used to identify putative CIC binding sites within the: (A) 2kb upstream region, (B) transcribed region and (C) 2kb downstream region of the *rasl11b* gene in *X. tropicalis*, mouse and human. Only sequences with a p-value < 0.001 were identified and are indicated by black arrows. Upper and lower-case sequences indicate detection on the forward and reverse strands respectively. Red arrows indicate transcriptional start sites and asterisks indicate stop codon positions.

4.3 Discussion

4.3.1 Domains of *rasl11b* expression correspond to regions of active FGF signalling throughout early *X. tropicalis* development

In-situ hybridisation has demonstrated that during gastrulation *rasl11b* is expressed in a ring around the blastopore. This was shown to correspond to a region of dpERK activity known to be stimulated by active FGF signalling in the early mesoderm (Isaacs et al., 1994, Isaacs et al., 1995, Christen and Slack, 1999).

During neurula stages, *rasl11b* is additionally expressed in the PSM, in two domains either side of the neural tube. The overlapping expression of *rasl11b* and *ripply2.2* seen by double in-situ hybridisation suggests that these domains correspond to the S-II and S-III somitomeres. Immunostainings demonstrate high levels of dpERK activity in the posterior, which appear to decrease anteriorly, with the boundary of visible dpERK activity encroaching on domains of *rasl11b* expression in the PSM. This is consistent with the gradient of FGF activity known to be established along the AP axis, with high levels of FGF in the posterior decreasing anteriorly (Dubrulle et al., 2001, Dubrulle and Pourquié, 2004, Sawada et al., 2001). At a critical threshold of FGF signalling, somitomeres are formed, and the spatial position along the AP axis where this is occurs is named the determination front, which is marked by the expression of genes required for somite formation, including *ripply2.2* (Pownall and Isaacs, 2010, Hitachi et al., 2009, Dubrulle et al., 2001). The expression of *rasl11b* in somitomeres at the determination front corresponds to a known region of FGF activity and suggests a role for *rasl11b* in somite segmentation.

Throughout tailbud stages, *rasl11b* continues to be expressed in two somitomeres in the PSM, but shows additional expression in the otic vesicles, heart, and branchial arches. Previous reports have demonstrated that various FGF ligands are expressed in each of these structures and corroborate the immunostainings shown here which indicate that these domains correspond to regions of dpERK activity (Christen and Slack, 1999, Lea et al., 2009, Deimling and Drysdale, 2011). Taken together, these results demonstrate that throughout *X. tropicalis* development *rasl11b* is consistently expressed in regions of active FGF signalling, thus supporting the hypothesis that *rasl11b* is regulated by the FGF signalling pathway.

4.3.2 Inhibition of FGF and MAPK/ERK signalling affects *rasl11b* expression at gastrula stages

Gastrula stage 10.5 embryos treated with chemical inhibitors of FGFR and MEK signalling demonstrate almost complete inhibition of *rasl11b* expression around the blastopore when analysed by in-situ hybridisation. Treatment with either inhibitor leads to a total loss of dpERK activity in embryos analysed by immunostaining, indicating that treatment was effective in

inhibiting activation of the downstream effector of the MAPK/ERK pathway (Eswarakumar et al., 2005). This suggests that *rasl11b* expression in the early mesoderm is regulated by FGF signalling via the MAPK/ERK pathway during gastrulation.

However, when FGF signalling is inhibited in gastrula stage embryos using the dominant negative FGF receptor, qPCR analysis shows only a minor reduction in total *rasl11b* expression compared to *tbxt*, which is almost completely lost. During gastrula stages, *tbxt* is solely regulated by FGF signalling in the early mesoderm (Isaacs et al., 1994), hence the drastic reduction in *tbxt* expression following XFD injection demonstrates that XFD is highly effective at inhibiting FGF signalling. This suggests that when FGF signalling is inhibited in gastrula stage embryos, the visual loss of mesodermal *rasl11b* expression shown by in-situ hybridisation corresponds to only a minor reduction in total *rasl11b* expression, when quantified by RT-qPCR. Regardless, the minor reduction in total *rasl11b* expression seen upon XFD injection suggests that during gastrulation *rasl11b* is at least, in part, regulated by the FGF signalling pathway.

The high levels of *rasl11b* expression remaining following XFD injection suggest that *rasl11b* during gastrula stages may also be regulated by an FGF independent signalling pathway. Given that the *oep* nodal coreceptor modulates *rasl11b* expression in shield stage zebrafish (Pézeron et al., 2008), and that some nodal ligands are expressed during gastrula stages in *Xenopus* (Luxardi et al., 2010), nodal signalling may also regulate *rasl11b* expression during gastrulation in *Xenopus*.

4.3.3 *rasl11b* expression is unaffected by inhibiting FGF and MAPK/ERK signalling during neurula stages

Neurula stage 19 embryos treated with MEK inhibitor exhibit a complete loss of dpERK activity in embryos analysed by immunostaining, indicating that treatment was effective in inhibiting the MAPK/ERK signalling pathway. Embryos treated with the FGFR inhibitor also exhibit a total loss of dpERK activity, except for along the dorsal midline. Given that activation of ERK was inhibited in known regions of active FGF signalling including the posterior (Isaacs et al., 1995, Christen and Slack, 1999), then the FGFR inhibitor was clearly effective in inhibiting FGF signalling, indicating that dpERK activity along the dorsal midline is likely FGF independent.

Neurula stage embryos treated with both inhibitors demonstrate no clear change in *rasl11b* expression in the PSM or around the closed blastopore when analysed by in-situ hybridisation. This demonstrates that *rasl11b* expression is regulated independently of FGF signalling and the MAPK/ERK pathway at neurula stages. Taken together, these results suggest that during gastrulation *rasl11b* expression in the early mesoderm is FGF-dependent, but later in development *rasl11b* expression is regulated by an alternative signalling pathway.

4.3.4 FGF overexpression affects the spatial expression pattern of *rasl11b* during neurula stages

Neurula stage 14 embryos unilaterally injected with CSKA-FGF4 demonstrate a pronounced upregulation of dpERK activity when analysed by immunostaining. Injected embryos exhibit an expansion of the posterior domain of dpERK activity and widespread but mosaic activation of ERK across the AP axis on the injected side. This observation is consistent with the mosaicism of FGF4 expression from the CSKA plasmid, a phenomenon commonly observed with expression from injected DNA constructs in amphibian embryos (Isaacs et al., 1994). Overexpression of FGF4 lead to seemingly stronger and more widespread expression of *rasl11b* expression in the posterior, with potentially some minor upregulation in the PSM. The upregulation of *rasl11b* in FGF4 overexpressing stage 14 embryos analysed by in-situ hybridisation, validates gene-level RNA-Seq analysis and supports the notion that *rasl11b* is a target of FGF signalling during early neurula stages.

Later in development, at stage 19, unilateral injection of CSKA-FGF4 visibly alters the spatial expression pattern of *rasl11b* in the PSM for half of the embryos injected. These embryos exhibit an expression pattern which displays clear disruption to individual domains of *rasl11b* expression. The same proportion of embryos unilaterally injected with CSKA-FGF4 demonstrate the same class of expression pattern when analysed for expression of *ripply2.2* by in-situ hybridisation. Given that *ripply2.2* normally marks the positions of the S-I to S-III somitomeres (Hitachi et al., 2009), this expression pattern likely reflects disrupted segmentation between them. Since *rasl11b* was earlier shown to exhibit overlapping expression with *ripply2.2,* it therefore appears that CSKA-FGF4 injection has affected the spatial expression of *rasl11b* by disrupting somite boundary positioning. One possible reason for this is that the mosaic expression of FGF4 from the CSKA-FGF4 plasmid disrupts the gradient of FGF activity along the AP axis (Dubrulle et al., 2001, Isaacs et al., 1994). The uneven distribution of FGF activity would present the critical threshold of FGF signalling needed for specifying somite positions unevenly along the AP axis.

4.3.5 Putative CIC binding sites are found in *rasl11b* genes of multiple vertebrates

Putative CIC binding sites were identified in the 2kb upstream region of the *X. tropicalis* and mouse *rasl11b* genes. These regions of DNA contain the promoters of the *rasl11b* gene in each of these vertebrates and are the loci at which most transcriptional repressors would bind to modulate *rasl11b* expression. However, CIC may bind regions other than the promoter to repress the transcription of *rasl11b*, and multiple putative CIC binding sites were identified within the transcribed region and 2kb downstream region of the *rasl11b* genes in *X. tropicalis*, mouse and humans. The high number of possible CIC binding sites in each of these regions in multiple

vertebrates supports the hypothesis that *rasl11b* is transcriptionally regulated by CIC. Although binding of CIC to these sites could be confirmed using chromatin immunoprecipitation sequencing (ChIP-Seq).

Chapter 5: Investigating the function of *rasl11b* during somitogenesis in *X. tropicalis*

5.1 Introduction

In chapter 4, it was revealed that *rasl11b* demonstrates a similar and partly overlapping expression pattern to the segmentally expressed gene *ripply2.2* within the PSM. Specifically, *rasl11b* was shown to overlap with *ripply2.2* expression within somitomeres S-II and S-III at the determination front. Genes are expressed at the determination front in response to a critical threshold of FGF signalling, thought to be established by the actions of multiple opposing signalling gradients along the AP axis of the PSM (Pownall and Isaacs, 2010, Aulehla and Pourquié, 2010, Diez del Corral et al., 2003). High levels of Wnt and FGF signalling in the posterior decrease anteriorly as they are met by an opposing gradient of retinoic acid (RA) signalling which is highest in the anterior (Dubrulle and Pourquié, 2004, Delfini et al., 2005, Aulehla et al., 2003, Moreno and Kintner, 2004). Genes expressed at the determination front are typically required for somite development (somitogenesis), including *ripply2.2* which is essential for the formation of somite boundaries (Kondow et al., 2007). Ripply2.2 interacts with the transcriptional corepressor Tle4 to inhibit the expression of genes required for somite formation, this terminates the somite-forming program, thus establishing a somite boundary (Hitachi et al., 2009, Kondow et al., 2007).

Given that *rasl11b* appears to be expressed at the determination front, this suggests that the GTPase has a role in somitogenesis. However, current research presents a limited understanding of *rasl11b* function during development, with only one study reflecting an inhibitory role of the GTPase during endoderm and prechordal plate development in zebrafish (Pézeron et al., 2008). Therefore, to attempt to elucidate the function of *rasl11b* during somitogenesis, *rasl11b* was overexpressed and knocked down in neurula stage *X. tropicalis* embryos, with the resulting effect on developing somites visualised by in-situ hybridisation for *ripply2.2*. Rasl11b overexpression was facilitated by injection of full-length *rasl11b* mRNA which contained a partial Kozak sequence to improve the efficiency of mRNA translation in vivo.

Knock down of *rasl11b* was achieved using antisense morpholino oligos (AMOs), which are short single-stranded DNA analogues with a modified phosphorodiamidate backbone (Summerton, 1999). AMOs sterically block the binding of molecules to their mRNA target and can be used to facilitate gene knockdown via two mechanisms. Translation-blocking AMOs block binding of the translation initiation complex to target mRNA until it has been degraded, thus preventing protein synthesis, leading to knockdown of the target gene (Nasevicius and Ekker, 2000). Splice-blocking

AMOs block the binding of splicing machinery to a specified splice site within target pre-mRNA (Moulton, 2007). This prevents splicing at this site, which can lead to the activation of cryptic splice sites, exon skipping, or intron inclusion in mature target mRNA (Giles et al., 1999, Draper et al., 2001). These modifications to the splicing process can result in effects such as the nonsense mediated decay of target mRNA and the synthesis of a truncated protein, both of which lead to a knockdown of gene function (Moulton, 2007). Splicing modifications can easily be detected by RT-PCR hence Rasl11b knockdown was facilitated using splice-blocking AMOs.

The aims of this chapter are:

- Analyse the effect of Rasl11b overexpression on somite development in neurula stage *X. tropicalis* embryos
- Explore the effect of Rasl11b knockdown on somitogenesis in neurula to early tailbud stage *X. tropicalis* embryos

5.2 Results

5.2.1 Analysing the effect of Rasl11b overexpression on somite development

As previously shown, during neurula stages *ripply2.2* is expressed in the PSM in three domains either side of the neural tube (Figure 17A). Given that these domains are known correspond to the S-I to S-III somitomeres (Hitachi et al., 2009), *ripply2.2* expression was analysed by in-situ hybridisation to visualise the effect of Rasl11b overexpression on somite development. Neurula stage 20-21 *X. tropicalis* embryos unilaterally injected with full-length *rasl11b* mRNA at the two-cell stage exhibit 3 classes of expression pattern on the injected side (Figure 17B-C). The first class of expression pattern demonstrates an anterior shift of segmented *ripply2.2* expression (Figure 17B). However, this segmentation is disrupted in the second class of expression pattern (Figure 17C), which reflects notably similar *ripply2.2* expression to that seen in embryos injected with CSKA-FGF4 (Figure 15B). The final class of expression demonstrates an extension of *ripply2.2* more laterally than seen in uninjected controls (Figure 17D). The similarity of *ripply2.2* expression observed between embryos injected with CSKA-FGF4 and *rasl11b* mRNA suggests that *rasl11b* may function in the FGF signalling pathway during somitogenesis.

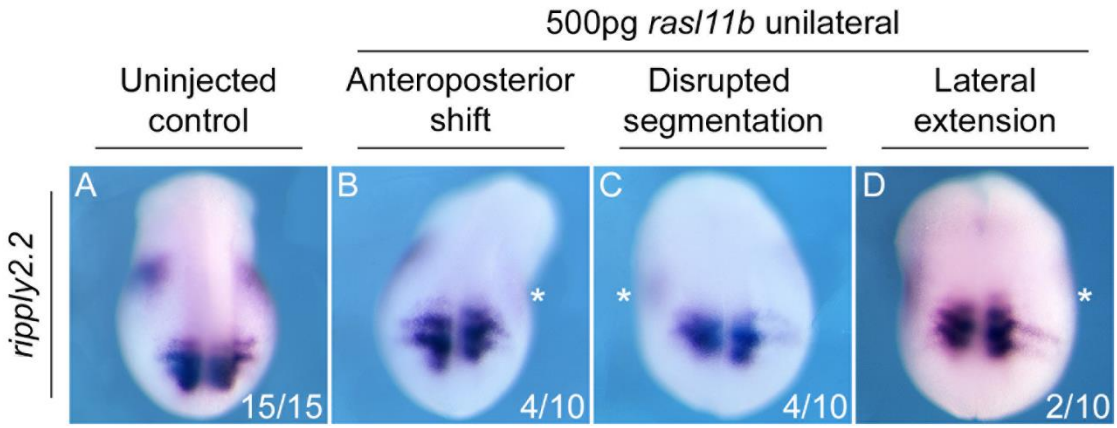


Figure 17: Neurula stage X. tropicalis embryos injected with *rasl11b* mRNA analysed by in-situ hybridisation for *ripply2.2*.
Two-cell stage embryos were unilaterally injected with 500pg full-length *rasl11b* mRNA and at stage 20-21 were analysed by in-situ hybridisation for *ripply2.2*. Expression of *ripply2.2* was used to demonstrate the effect of *rasl11b* overexpression on somite development. Embryos show a dorsal view with anterior upwards and were staged according to Nieuwkoop and Faber (1994). Asterisks indicate the injected side.

5.2.2 Exploring the effect of Rasl11b knockdown on somitogenesis

To achieve Rasl11b knockdown, neurula stage *X. tropicalis* embryos were unilaterally injected with a splice-blocking AMO specific for *rasl11b* mRNA (*rasl11b* AMO). The *rasl11b* AMO targets the splice site between exon 1 and intron 1 of *rasl11b* pre-mRNA, which should block the splicing of intron 1, causing it to be retained in mature mRNA (Figure 18A). The retention of intron 1 leads to the introduction of a premature stop codon in *rasl11b* mRNA (Figure 18A), which when translated should give rise to a truncated non-functional protein (Moulton, 2007).

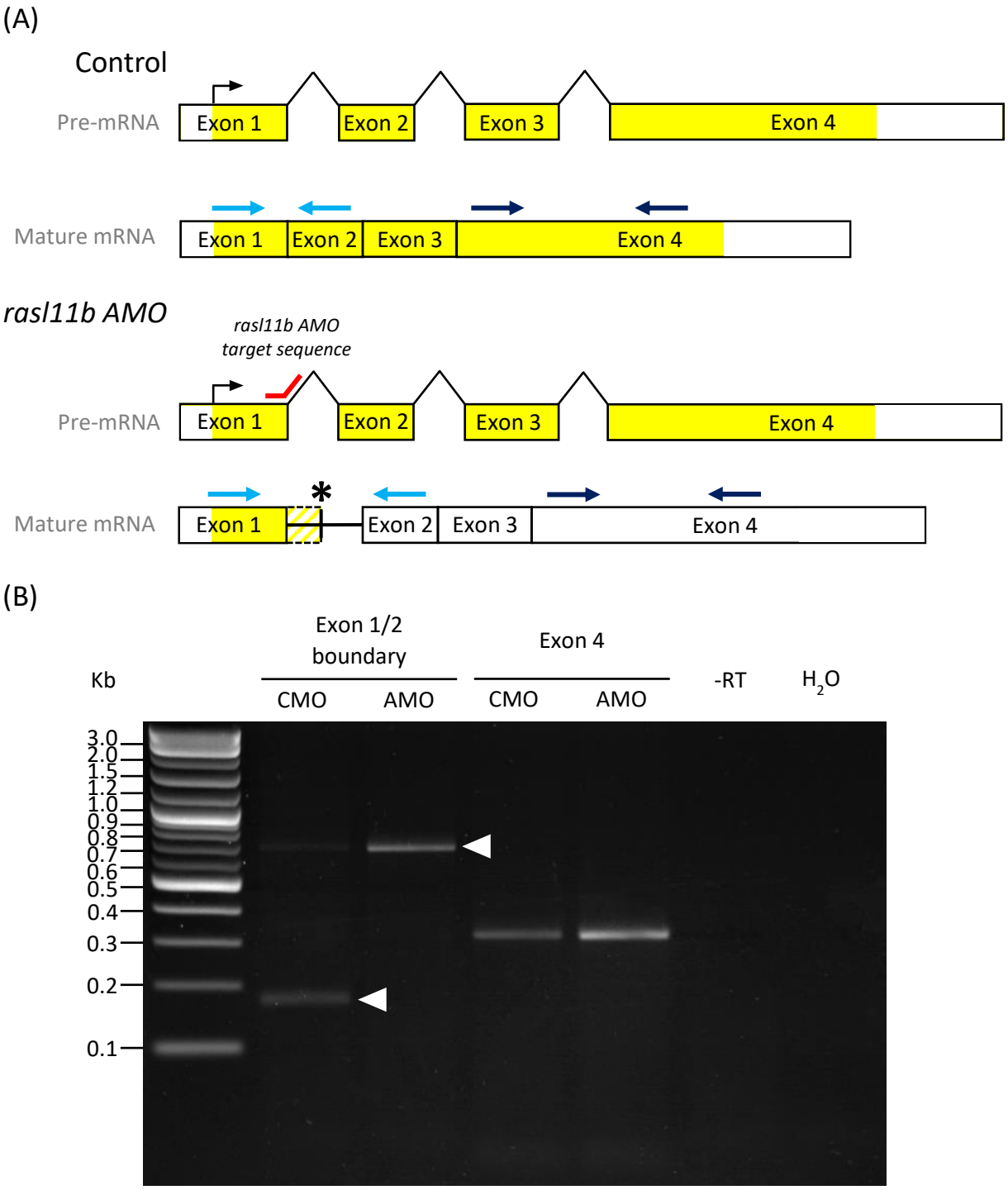


Figure 18: rasl11b AMO blocks splicing of intron 1 from rasl11b mRNA

(A) Schematic representation showing the mode of action of the rasl11b splice-blocking antisense morpholino oligo (*rasl11b* AMO). *rasl11b* AMO targets the splice site between exon 1 and intron 1 of *rasl11b* pre-mRNA, which blocks the splicing of intron 1 causing it to be retained in mature *rasl11b* mRNA. Retention of intron 1 introduces a premature stop codon (*) into the mature mRNA and results in the formation of a truncated protein. Inclusion or exclusion of intron 1 can be detected using RT-PCR with primers which flank the boundary between exon 1 and 2 (light blue arrows), primers which flank a region of exon 4 are used as controls (dark blue arrows). (B) RT-PCR analysis of neurula stage *X. tropicalis* embryos unilaterally injected with *rasl11b* AMO and control morpholino oligo (CMO). Embryos were injected with 10ng *rasl11b* AMO or CMO at the two-cell stage and analysed by RT-PCR at stage 15. *rasl11b* cDNA was amplified using primers which flank

the exon1/2 boundary or a region of exon 4. White arrowheads indicate the change in PCR product size following retention of intron 1. Embryos were staged according to Nieuwkoop and Faber (1994).

To confirm that *rasl11b* AMO effectively blocks splicing, neurula stage embryos unilaterally injected with *rasl11b* AMO or non-specific control morpholino oligo (CMO) were analysed using RT-PCR. *rasl11b* cDNA was amplified using primers which either flank the boundary between exons 1 and 2 (Figure 18A, light blue arrows), or a region of exon 4 (Figure 18A, dark blue arrow). As expected, the size of the product amplified from exon 4 is the same between embryos injected with CMO and *rasl11b* AMO (Figure 18B). However, amplification of the region between exons 1 and 2 in CMO injected embryos generates two products, a more abundant smaller product amplified from spliced mature *rasl11b* mRNA, and a less abundant larger product amplified from unspliced *rasl11b* pre-mRNA (Figure 18B and Table 9). Only the larger PCR product was amplified from embryos injected with *rasl11b* AMO, with sequencing confirming that this product contained intron 1 (Table 9). This indicates that intron 1 was retained in all *rasl11b* mRNA detected in embryos injected with the *rasl11b* AMO. Furthermore, sequencing of PCR products identified multiple premature stop codons in each reading frame throughout intron 1 of *rasl11b* (Table 9), demonstrating that even splicing at a cryptic site should not give rise to a functional protein.

Table 9: Sequences of RT-PCR products showing retention of intron 1 in *rasl11b* cDNA extracted from *rasl11b* AMO injected embryos.

rasl11b cDNA was amplified using primers which flank the boundary between exons 1 and 2. Red and blue indicate 15bp intronic and exonic sequences flanking the intron 1 splice site respectively. Yellow highlight indicates the *rasl11b* antisense morpholino oligo (*rasl11b* AMO) target sequence.

RT-PCR product	Sequence
Rasl11b intron 1 (reference genome sequence)	GTGAATGTCCCTGGGGCACAGGCAGGGAGGGGGCAGAGGGAATGCTTCCTCA AGGCTGTAATTAAATGGCACAATATAATGAGTGATCGCTGCCTGCACAATGTAT ATGGTTCTGCTTTATTAAAAGGAGTTTAAAATGGGGGGGCATTATACTAGTGCT GATAGATTCTGAAAGGCTTCATGCCAGTGCAGGGGGCAGTGCCAGCTCTGACT CTATATGTTTGCCCTGGTACTGGGACTTGAATTAGATTGTACCCATACAGGACC CTGCCAATGCAGATCATTTGATTCTGTCAAGTGAGTGAGATGCACAGTGCCCAT AGTGTTATACAGGGTTTGGCAGAACTACAGCTCTCAGAACCCCCCTGACAGTCA CTGGCTTCTAGGAGCTGCTGAGAACTTTGCTTTTAGCTTAAAGGAGATATAGTA TGTGGGTGTTAAACATTATTTACTTTAAGCAGTCAACTGACCTTTATTTTGTTTTT CCTTGCAG
Intron-inclusive product amplified from embryos injected with *rasl11b* AMO	CACCACACGGAATGTGCCAGCAGCAGTGCGGGCACCGCCTCTAGCCGGGTCAT CAAGATCGCTGTGGTCGGGGGCAGTGGCGTGGGCAAGACAGGTGAATGTCCC TGGGGCACAGGCAGGGAGGGGGCAGAGGGAATGCTTCCTCAAGGCTGTAATT AAATGGCACAATATAATGAGTGATCGCTGCCTGCACAATGTATATGGTTCTGCT TTATTAAAAGGAGTTTAAAATGGGGGGGCATTATACTAGTGCTGATAGATTCTG AAAGGCTTCATGCCAGTGCAGGGGGCAGTGCCAGCTCTGACTCTATATGTTTGC CCTGGTACTGGGACTTGAATTAGATTGTACCCATACAGGACCCTGCCAATGCAG ATCATTTGATTCTGTCAAGTGAGTGAGATGCACAGTGCCCATAGTGTTATACAG GGTTTGGCAGAACTACAGCTCTCAGAACCCCCCTGACAGTCACTGGCTTCTAGG AGCTGCTGAGAACTTTGCTTTTAGCTTAAAGGAGATATAGTATGTGGGTGTTAA ACATTATTTACTTTAAGCAGTCAACTGACCTTTATTTTGTTTTTCCTTGCAGCCCT GGTGGTGAGATTCCT
Intron-spliced product amplified from embryos injected with CMO	TGCGGGCACCGCCTCTAGCCGGGTCATCAAGATCGCTGTGGTCGGGGGCAGTG GCGTGGGCAAGACAGCCCTGGTGGTGAGATTCCTTACCAAGCGCAAA
Unspliced pre-mRNA product amplified from embryos injected with CMO	AATGTGCCAGCAGCAGTGCGGGCACCGCCTCTAGCCGGGTCATCAAGATCGCT GTGGTCGGGGGCAGTGGCGTGGGCAAGACAGGTGAATGTCCCTGGGGCACAG GCAGGGAGGGGGCAGAGGGAATGCTTCCTCAAGGCTGTAATTAAATGGCACA ATATAATGAGTGATCGCTGCCTGCACAATGTATATGGTTCTGCTTTATTAAAAGG AGTTTAAAATGGGGGGGCATTATACTAGTGCTGATAGATTCTGAAAGGCTTCAT GCCAGTGCAGGGGGCAGTGCCAGCTCTGACTCTATATGTTTGCCCTGGTACTGG GACTTGAATTAGATTGTACCCATACAGGACCCTGCCAATGCAGATCATTTGATTC TGTCAAGTGAGTGAGATGCACAGTGCCCATAGTGTTATACAGGGTTTGGCAGA ACTACAGCTCTCACAACCCCCCTGACAGTCACTGGCTTCTAGGAGCTGCTGAGA ACTTTGCTTTTAGCTTAAAGGAGATATAGTATGTGGGTGTTAAACATTATTTACT TTAAGCAGTCAACTGACCTTTATTTTGTTTTTCCTTGCAGCCCTGGTGGTGAGATT CCTTACCAAGCGC

Given that the *rasl11b* AMO is seemingly effective in facilitating Rasl11b knockdown, the AMO was used to explore the effect of Rasl11b knockdown on somitogenesis. To do this, mid-neurula stage and late neurula to early tailbud stage embryos unilaterally injected with Rasl11b were analysed by in-situ hybridisation for *ripply2.2* (Figure 19). As before, *ripply2.2* expression in the PSM was used to visualise any effects on somite development. However, following Rasl11b knockdown, almost all embryos at both stages demonstrate no visible change in *ripply2.2* expression relative to uninjected controls (Figure 19A, C,D,F). Whereas a shift of *ripply2.2* expression anteriorly along the AP axis was only observed in one embryo at each stage (Figure 19B and 19E). This suggest that *rasl11b* has no significant role during somitogenesis in *X. tropicalis*.

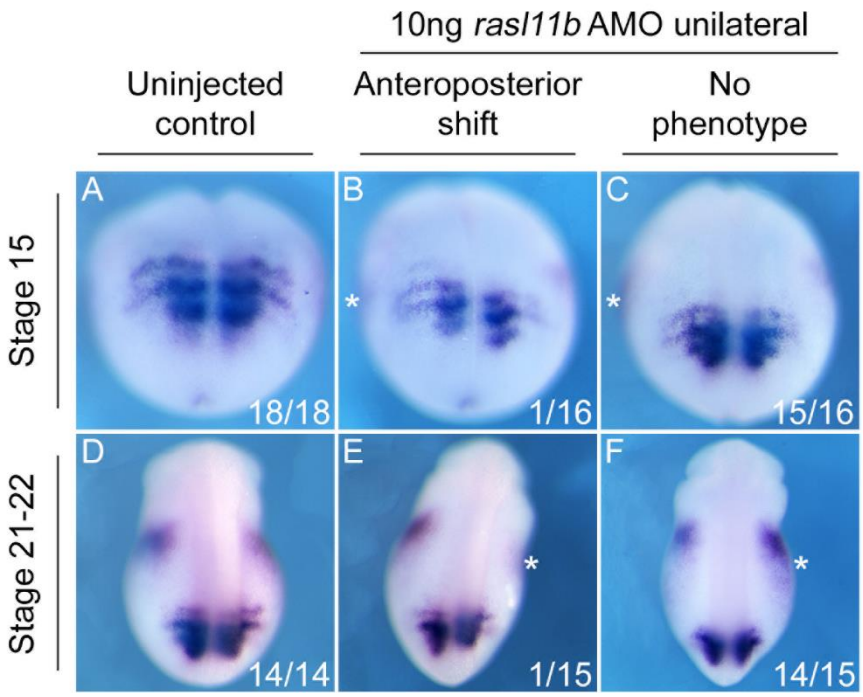


Figure 19: Neurula to early tailbud stage *X. tropicalis* embryos injected with *rasl11b* AMO analysed by in-situ hybridisation for *ripply2.2*.
Two-cell stage embryos were unilaterally injected with 10ng *rasl11b* antisense morpholino oligo (*rasl11b* AMO). At stages 15 and 21-22 embryos were analysed by in-situ hybridisation for *ripply2.2*. Expression of *ripply2.2* was used to demonstrate the effect of *rasl11b* knockdown on somite development. Embryos show a dorsal view with anterior upwards and were staged according to Nieuwkoop and Faber (1994). Asterisks indicate the injected side.

5.3 Discussion

5.3.1 FGF and Rasl11b overexpression mediate similar changes to *ripply2.2* expression during somitogenesis

As revealed in the previous chapter, some neurula stage *X. tropicalis* embryos unilaterally injected with CSKA-FGF4 show disruption to individual domains of *ripply2.2* expression relative to uninjected controls. Notably similar disruption to *ripply2.2* expression is observed in some neurula stage

embryos unilaterally injected with full-length *rasl11b* mRNA. Given that *ripply2.2* expression is a marker of developing somites (Hitachi et al., 2009), then it appears that FGF and Rasl11b overexpression have a similar effect on somite development. This suggests that Rasl11b may function in the FGF signalling pathway during somitogenesis.

Furthermore, in other embryos unilaterally injected with *rasl11b* mRNA, segmented *ripply2.2* expression is shifted anteriorly along the AP axis on the injected side. A similar shift in *ripply2.2* expression was shown to correspond to an anterior expansion of dpERK activity in embryos knocked down for Sulf1, an inhibitor of FGF signalling during somitogenesis (Freeman et al., 2008). Given that an anterior shift of *ripply2.2* expression is known to correspond to an anterior expansion of FGF signalling, then rasl11b overexpression appears to upregulate FGF activity along the AP axis during somitogenesis. This further supports the notion that Rasl11b functions in the FGF signalling pathway during somite development.

5.3.2 Rasl11b knockdown has a minor effect on *ripply2.2* expression during somitogenesis

Almost all neurula to early tailbud stage embryos unilaterally injected with *rasl11b* AMO demonstrate no change in r*ipply2.2* expression relative to uninjected controls, with only one embryo at each stage showing an anterior shift of *ripply2.2* on the injected side. This indicates that Rasl11b knockdown has only a minor effect on *ripply2.2* expression, and since *ripply2.2* marks developing somites (Hitachi et al., 2009), this suggests that either Rasl11b has no significant role in somitogenesis or there is redundancy in Rasl11b function. However, Rasl11b overexpression induced visible disruption to somite development in a manner consistent with the upregulation of FGF signalling (Freeman et al., 2008). Taken together, these results suggest that Rasl11b has a redundant function in the FGF signalling pathway during somitogenesis.

Chapter 6: General Discussion

6.1 Summary

Despite an extensive understanding of FGF signal transduction, the exact mechanism of FGF target gene expression has yet to be fully elucidated. An expanding body of evidence presents the role of CIC as a labile transcriptional repressor of RTK target genes, with previous studies by our lab providing strong evidence that CIC acts downstream of the FGF signalling pathway (King, 2019, Cowell, 2019). Therefore, we hypothesise that the expression of a subset of FGF target genes is reliant upon the ERK-mediated relief of CIC transcriptional repression.

In this study, gene-level transcriptomic analysis of CIC knockdown and FGF overexpressing *X. tropicalis* embryos was undertaken, with findings supportive of the notion that a subset of FGF target genes is transcriptionally regulated by CIC. These findings include the identification of a statistically significant overlap between genes upregulated in both CIC knockdown and FGF overexpressing embryos. In temporal expression analysis of these overlapping genes, rapid and dynamic changes in gene expression were reflected, which is consistent with regulation by a labile transcriptional repressor (Keenan et al., 2020).

One of the genes identified as significantly upregulated in both CIC knockdown and FGF overexpressing embryos was *rasl11b*. A gene which encodes a small Ras-like GTPase, though current research reflects a limiting understanding of the regulation and expression of this gene during development. Therefore, this study presents a stage series of *rasl11b* expression throughout *X. tropicalis* development and demonstrates that *rasl11b* is expressed in two somitomeres in the PSM. Domains of *rasl11b* expression were shown to correspond to regions of active FGF signalling so FGF signalling via the MAPK/ERK pathway was manipulated to further investigate *rasl11b* regulation by FGF.

Inhibition of FGF signalling revealed that *rasl11b* expression in the early mesoderm is FGF-dependent at gastrula stages. However, during neurula stages, FGF signalling is not required for *rasl11b* expression in the PSM. Although, the spatial expression pattern of *rasl11b* was affected by FGF overexpression during neurula stages. Additionally, CIC binding site analysis identified putative CIC binding sites throughout the *rasl11b* locus in multiple vertebrates. Taken together, these results support the notion that *rasl11b* is a target of FGF signalling that may be transcriptionally regulated by CIC.

Given that *rasl11b* was shown to be expressed at the determination front and genes expressed in this region typically present roles in somite development, this study also investigated the function of *rasl11b* during somitogenesis. Knockdown of *rasl11b* reveals no significant effect on somite development. However, overexpression of *rasl11b* demonstrates a disruption to somite development consistent with the upregulation of FGF signalling. These results suggest that Rasl11b has a redundant function in the FGF signalling pathway during somitogenesis.

6.2 FGF signalling, Capicua and Rasl11b in development and disease

6.2.1 Fibroblast growth factor signalling

The FGF signalling pathway is essential for the initiation and regulation of a variety of developmental processes including gastrulation, mesoderm induction, limb development and anteroposterior patterning (Böttcher and Niehrs, 2005, Itoh, 2007, McIntosh et al., 2000, ten Berge et al., 2008). Normal development necessitates the correct regulation of this pathway, which is which is exemplified by the range of abnormalities that arise when FGF signalling is dysregulated. Different mutations in FGFRs can result in multiple types of cancer and skeletal disorders including thanatophoric dysplasia, achondroplasia, and Crouzon syndrome (Passos-Bueno et al., 1999, Aviezer et al., 2003, Yeh et al., 2013, Wesche et al., 2011). Whereas mutations in FGF ligands can lead to lacrimo-auriculo-dento-digital syndrome, bilateral renal aplasia and hypophosphatemic rickets (Milunsky et al., 2006, Barak et al., 2012, White et al., 2001). This study has presented evidence in attempt to further our understanding of the molecular mechanism of FGF target gene expression. An enhanced understanding of this mechanism would be invaluable in the development of treatments for disorders associated with the dysregulation of FGF signalling.

This study identified many genes upregulated in both CIC knockdown and FGF overexpressing embryos that demonstrate rapid changes in gene expression, which is consistent with the action of the labile transcriptional repressor CIC. Following ERK activation, CIC rapidly dissociates from DNA and is subsequently degraded, hence CIC mediates rapid changes in gene expression following ERK activation (Keenan et al., 2020, King, 2019). The immediate early gene *fos* was upregulated in both CIC knockdown and FGF overexpression and is known to be expressed following ERK activation in the wound response (Cowell, 2019, Dieckgraefe and Weems, 1999). This suggests that some of the subset of FGF target genes transcriptionally regulated by CIC may be involved in immediate early response, such as the wound response.

6.2.2 Capicua transcriptional repression

During *Drosophila* development, CIC is known to transcriptionally repress the expression of RTK target genes in the absence of ERK activation (Jiménez et al., 2012, Keenan et al., 2020). In this

investigation, CIC knockdown lead to the upregulation of markedly more genes than FGF overexpression in *X. tropicalis* embryos. The upregulation of genes following CIC knockdown is consistent with the alleviation of transcriptional repression, which supports the notion that CIC acts as a transcriptional repressor during *X. tropicalis* development. However, since vastly more genes were regulated by CIC knockdown than FGF overexpression, this demonstrates that CIC regulation is not just limited to FGF target genes, which is consistent with reports that CIC acts as downstream of multiple RTK signalling pathways (Lee, 2020)

Given that CIC is known to regulate the expression of RTK target genes, CIC dysregulation is implicated in diseases typically characterised by uncontrolled RTK signalling, particularly cancer and neurodegenerative disorders (Jiménez et al., 2012). Loss of function mutations in CIC are prevalent in lung, stomach, and prostate cancers (Tanaka et al., 2017, Okimoto et al., 2017), while mutations in the HMG-box of CIC have been shown to prevent DNA binding in oligodendroglioma (Forés et al., 2017). In mammals, CIC has been reported to form a nuclear complex with the polyglutamate repeat protein ATXN1, dysregulation of this interaction induces the upregulation of RTK target genes, which is thought to contribute to spinocerebellar ataxia type 1 (SCA1) neurodegeneration (Jiménez et al., 2012). This study has sought to improve our understanding of the developmental signalling pathways modulated by CIC transcription repression, which may aid in the development of more targeted therapeutics for diseases associated with CIC dysregulation.

6.2.3 Rasl11b function

Current research presents a limited understanding of the function of Rasl11b during development, with one study suggesting the GTPase inhibits *oep* signalling during zebrafish endoderm and prechordal plate development (Pézeron et al., 2008). However, this study has illustrated that *rasl11b* is expressed in developing somites and that the GTPase appears to upregulate FGF activity along the AP axis of the PSM. Rasl11b possesses similar inherent GTPase activity as membrane-bound Ras (Colicelli, 2004), which catalyses hydrolysis of GTP to GDP to active Raf, the upstream kinase of the MAPK/ERK cascade (Schlessinger, 2000, Roberts and Der, 2007). This suggests that Rasl11b may activate the MAPK/ERK pathway via interactions with Raf, which would be consistent with Rasl11b appearing to upregulate FGF activity, since Raf is activated downstream of FGF signalling.

Rasl11b is a member of the Ras family of small GTPases and exhibits similar structural and catalytic properties to membrane-bound Ras, a well characterised oncogene (Colicelli, 2004, Pylayeva-Gupta et al., 2011). Activating mutations can render Ras proteins constitutively active resulting in dysregulated MAPK/ERK signalling and uncontrolled cell proliferation (Gillies et al., 2020). Rasl11b

is known to be upregulated in neuroblastoma, a disease typically characterised by dysregulated MAPK/ERK signalling (Liu and Li, 2019, Eleveld et al., 2015). Hence, if Rasl11b functions in the MAPK/ERK pathway, aberrant upregulation of the protein may lead to oncogenesis. Consequently, it is important that we understand how Rasl11b functions during normal and disease conditions to be able to develop effective therapeutics.

6.3 Future work

This investigation has presented evidence in support of the notion that a subset of FGF target genes is transcriptionally regulated by CIC. However, gene-level transcriptomic analysis could be validated using qPCR to quantify the expression of more genes found to be upregulated following CIC knockdown and FGF overexpression. Alternatively, in-situ hybridisation could be used to compare the spatial expression of these genes between wildtype, CIC knockdown and FGF overexpressing embryos at multiple developmental stages.

This study has also attempted to define the domains of *rasl11b* expression during *X. tropicalis* development, particularly within the PSM. The assignment of *rasl11b* expression to the correct somitomeres could be validated by immunofluorescence microscopy on mid neurula stage sections using different fluorescent antibodies specific for Rasl11b and Ripply2.2. Additionally, evidence for the regulation of *rasl11b* by FGF signalling could also be reinforced by repeating qPCR to quantify *rasl11b* expression in multiple uninjected and XFD injected biological replicates. This would allow the statistical significance to be calculated which would validate true biological effects.

FIMO binding site analysis identified multiple putative CIC binding sites within the *rasl11b* gene of multiple vertebrates. Binding of CIC to these genomic loci could be confirmed by chromatin immunoprecipitation sequencing (ChIP-Seq) using tagged constructs. The output from this experiment could be compared to gene-level transcriptomic analysis to determine whether CIC binds the genomic loci of other genes upregulated in both CIC knockdown and FGF overexpression.

Finally, this study also suggested that Rasl11b may upregulate MAPK/ERK signalling, this could be confirmed by analysing embryos injected with full-length *rasl11b* mRNA by western blot using antibodies specific for dpERK.

Abbreviations

Akt/PKB	Protein kinase B
AMO	Antisense morpholino oligo
AP-1	Activator protein 1
BCIP	5-Bromo-4-chloro-3-indolyl phosphate
ChIP-Seq	Chromatin immunoprecipitation sequencing
CHX	Cycloheximide
CIC	Capicua
CMO	Control morpholino oligo
CRKL	CRK-like protein
DAG	Diacylglycerol
DIG	Digoxigenin
DMSO	Dimethyl sulfoxide
dpERK	Diphosphorylated ERK
ECM	Extracellular matrix
Egr1	Early growth response 1
ERK	Extracellular signal-related kinase
ETS	E26 transformation-specific transcription factors
FDR	False discovery rate
FGF	Fibroblast growth factor
FGFR	Fibroblast growth factor receptor
FIMO	Find individual motif occurrences
FLU	Fluorescein
FRS2α	Fibroblast growth factor receptor substrate 2 alpha
GDP	Guanosine diphosphate
GO	Gene ontology
GRB2	Growth factor bound 2
GTP	Guanosine triphosphate
HCG	Human chronic gonadotropin
HMG-box	High mobility group box
HSPG	Heparan sulphate proteoglycan
IP_3	Inositol-1,4,5-trisphosphate

MBT	Mid-blastula transition
MEK/MAPKK	Mitogen-activated protein kinase kinase
MRS	Modified ringer's solution
Myod1	Myogenic differentiation 1
NAM	Normal amphibian medium
NHEJ	Non-homologous end joining
Oep	One-eyed pinhead
PANTHER	Protein Analysis Evolutionary Relationships
PI3K	Phosphoinositide 3-kinase
PLCγ	Phospholipase C gamma
PSM	Presomitic mesoderm
RA	Retinoic acid
Raf/MAPK	Mitogen-activated protein kinase
Rasl11b	Ras like family 11 member B
Ripply2.2	Ripply2 homolog, gene 2
RTK	Receptor tyrosine kinase
SCA1	Spinocerebellar ataxia type 1
SH2	Src homology 2
SHP2	Src homology 2 containing protein tyrosine phosphatase 2
SOS	Son of Sevenless
TALENs	Transcription activator-like effector nucleases
Tbxt	T-box transcription factor T / Brachyury
TGF-β	Transforming growth factor beta
TPM	Transcripts per million
XFD	Dominant negative fibroblast growth factor receptor

Appendix

Appendix 1: Rstudio script

```
########## Transcript level analysis ##########
#Setting working directory
setwd("~/RNA-seq analysis")
#Clearing R memory
rm(list=ls())
#Installing the 'rhdf5' package since this is required for the
#installation of 'sleuth'
if (!requireNamespace("BiocManager", quietly = TRUE))
  install.packages("BiocManager")
BiocManager::install()
BiocManager::install(c("rhdf5"))
biocLite("rhdf5")
#Installing the 'devtools' package since this is also required for the
#installation of 'sleuth'
install.packages("devtools")
#Installing the 'sleuth' package
devtools::install_github("pachterlab/sleuth")
#Loading the 'devtools' package into the local environment
library(devtools)
#Intsalling tidyverse package
install.packages("tidyverse")
#Loading tidyverse
library(tidyverse)
#Loading the 'sleuth' package into the local environment
library(sleuth)
suppressMessages(library("sleuth"))
#The first step in a sleuth analysis is to specify where the kallisto
#results are stored. A variable is created for this purpose with:
```

```
sample_id <- dir(file.path("data/salmon"))
sample_id
#A list of paths to the kallisto results indexed by the sample IDs is
#collated with
kal_dirs <- file.path("data/salmon", sample_id)
kal_dirs
#The next step is to load an auxillary table that describes the
#experimental design and the relationship between the kallisto
#directories and the samples:
s2c <- read.table(file.path("data", "metadata", "hiseq_info.txt"),
          header = TRUE, stringsAsFactors=FALSE)
s2c <- dplyr::select(s2c, sample, batch, condition)
s2c
#Now the directories must be appended in a new column to the table
#describing the experiment. This column must be labeled path,
#otherwise sleuth will report an error. This is to ensure that
#samples can be associated with kallisto quantifications.
s2c <- dplyr::mutate(s2c, path = kal_dirs)
#It is important to check that the pairings are correct:
print(s2c)
#Constructing the sleuth object#
#The sleuth object must first be initialized with:
so <- sleuth_prep(s2c, read_bootstrap_tpm=TRUE, extra_bootstrap_summary = TRUE)
so
#Then the full model is fit with
so <- sleuth_fit(so, ~condition+batch, 'full')
#What this has accomplished is to “smooth” the raw kallisto abundance estimates
#for each sample using a linear model with a parameter that represents the
#experimental condition. To test for transcripts that are differential expressed
#between the conditions, sleuth performs a second fit to a “reduced” model that
#presumes abundances are equal in the two conditions. To identify differential
```

```
#expressed transcripts sleuth will then identify transcripts with a significantly
#better fit with the “full” model.
#The 'reduced' model is fit with
so <- sleuth_fit(so, ~batch, 'reduced')
so
#and the test is performed with
so <- sleuth_lrt(so, 'reduced', 'full')
so
#In general, sleuth can utilize the likelihood ratio test with any pair of models
#that are nested, and other walkthroughs illustrate the power of such a framework
#for accounting for batch effects and more complex experimental designs.
#The models that have been fit can always be examined with the models() function.
models(so)
#The results of the test can be examined with
sleuth_table <- sleuth_results(so, 'reduced:full', 'lrt', show_all = FALSE)
sleuth_significant <- dplyr::filter(sleuth_table, qval <= 0.05)
head(sleuth_significant, 20)
#Opening sleuth web interface to see results of the test
sleuth_live(so)
#Using the wald test to measure the effect of FGF treatment relative to the
#control treatment (water)
so <- sleuth_wt(so, "conditioneFGF")
models(so)
#using the wald test to measure the effect of CIC knockout relative to the
#control treatment (water)
so <- sleuth_wt(so, "conditionCIC_KO")
models(so)
#Opening sleuth web interface to see the effect of the treatments on each
#transcript relative to the control - select 'settings' - 'wald test' -
#test table. This will show p-values, q-values and effect sizes for
#transcripts for each treatment
```

```
sleuth_live(so)
################ gene level analysis ################
#reading in text file assigning transcripts to their genes
transcripts <- read.table("data/metadata/transcript_to_gene.txt", header=T)
#creating a new sleuth object for gene level analysis
so <- sleuth_prep(s2c, target_mapping=transcripts,
                  aggregation_column = 'gene_id',
                  read_bootstrap_tpm = TRUE, gene_mode = TRUE,
                  extra_bootstrap_summary = TRUE)
#fitting full model to gene level counts
so <- sleuth_fit(so, ~condition+batch,"full")
#fitting reduced model to gene level counts
so <- sleuth_fit(so, ~batch, "reduced")
#running the likelihood test
so <- sleuth_lrt(so, "reduced", "full")
so
#The models fit are examined with:
models(so)
#The results of the test can be examined with
sleuth_table <- sleuth_results(so, 'reduced:full', 'lrt', show_all = FALSE)
sleuth_significant <- dplyr::filter(sleuth_table, qval <= 0.05)
head(sleuth_significant, 20)
#Opening sleuth web interface to see results of the test
sleuth_live(so)
#Using the wald test to measure the effect of FGF treatment relative to the
#control treatment (water)
so <- sleuth_wt(so, "conditioneFGF")
models(so)
#using the wald test to measure the effect of CIC knockout relative to the
#control treatment (water)
so <- sleuth_wt(so, "conditionCIC_KO")
```

```
models(so)
sleuth_live(so)
#saving current gene level sleuth object
sleuth_save(so, "rnaseq_sleuth_object.rds")
#loading the sleuth object for next use
so <- readRDS("rnaseq_sleuth_object.rds")
sleuth_live(so)
#show test table
sleuth_results(so, "conditioneFGF")
#retrieving FGF gene level table from shiny app
fgf_gene_level_table <- sleuth_results(so, "conditioneFGF")
write.table(fgf_gene_level_table, file="fgf_gene_level_table.txt", sep="\t")
#retrieving CIC knock down gene level table from shiny app
CIC_gene_level_table <- sleuth_results(so, "conditionCIC_KO")
write.table(CIC_gene_level_table, file="CIC_gene_level_table.txt", sep="\t")
#retrieving kallisto TPM gene level counts from shiny app
kallisto_counts <- kallisto_table(so, use_filtered = TRUE, normalized = TRUE,
                          include_covariates = FALSE)
kallisto_counts <- as.data.frame(kallisto_counts)
kallisto_counts
#subsetting MK1 MK4 and MK10 samples and merging TPMs into CIC knockdown
#gene level analysis
MK1 <- subset(kallisto_counts, kallisto_counts$sample =="MK1")
CIC_gene_level_table <- merge(CIC_gene_level_table, MK1, by='target_id')
MK4 <- subset(kallisto_counts, kallisto_counts$sample =="MK4")
CIC_gene_level_table <- merge(CIC_gene_level_table, MK4, by='target_id')
MK10 <- subset(kallisto_counts, kallisto_counts$sample =="MK10")
CIC_gene_level_table <- merge(CIC_gene_level_table, MK10, by='target_id')
#creating text file containing CIC knockdown gene level analysis with
#TPM for CIC knockdown samples (MK1,MK4 and MK10)
write.table(CIC_gene_level_table, file="CIC_gene_level_table.txt", sep="\t")
```

```
#subsetting MK3 MK6 and MK12 samples and merging TPMs into FGF overexpression
#gene level analysis
MK3 <- subset(kallisto_counts, kallisto_counts$sample =="MK3")
fgf_gene_level_table <- merge(fgf_gene_level_table, MK3, by='target_id')
MK6 <- subset(kallisto_counts, kallisto_counts$sample =="MK6")
fgf_gene_level_table <- merge(fgf_gene_level_table, MK6, by='target_id')
MK12 <- subset(kallisto_counts, kallisto_counts$sample =="MK12")
fgf_gene_level_table <- merge(fgf_gene_level_table, MK12, by='target_id')
#creating text file containing FGF overexpression gene level analysis with
#TPM for FGF overexpressing samples (MK3 MK6 and MK12)
write.table(fgf_gene_level_table, file="fgf_gene_level_table.txt", sep="\t")
#In excel opened txt files and added column 'mean_TPM' for both CIC knockdown
#and FGF ovexpressing analyses. Then filtered out gene with <1.5 TPMs.
#This is to remove the genes with very low levels of expression.
fgf_data <- read.table("fgf_gene_level_table_post_filter.txt", header=T)
fgf_data <- fgf_data[complete.cases(fgf_data), ]
CIC_data <- read.table("CIC_gene_level_table_post_filter.txt", header=T)
CIC_data <- CIC_data[complete.cases(CIC_data), ]
### making basic volcano plots ###
library(ggplot2)
fgf_data <- read.table("fgf_gene_level_table_post_filter.txt", header=T)
fgf_data <- fgf_data[complete.cases(fgf_data), ]
p <- ggplot(data=fgf_data, aes(x=b, y=-log10(qval))) + geom_point() +
  theme_minimal()
p2 <- p + geom_vline(xintercept=c(-1.75, 1.75), col="red") +
  geom_hline(yintercept=log2(1.75), col="red")
print(p2)
##### enhanced volcano plot - figure #####
#installing enhanced volcano
if (!requireNamespace('BiocManager', quietly = TRUE))
  install.packages('BiocManager')
```

```
BiocManager::install('EnhancedVolcano')
#installing development version
devtools::install_github('kevinblighe/EnhancedVolcano')
#loading enhanced volcano package into r session
library(EnhancedVolcano)
#plotting most basic volcano plot
EnhancedVolcano(fgf_data,
        lab = rownames(fgf_data),
        x = 'b',
        y = 'qval')
#adjusting p-value threshold (q-value) and fold change threshold (effect size)
#point sizes and label sizes
EnhancedVolcano(fgf_data,
        lab = fgf_data$target_id,
        x = 'b',
        y = 'qval',
        title = 'FGF overexpressing',
        pCutoff = 0.05,
        FCcutoff = 1.5,
        pointSize = 3.0,
        labSize = 6.0)
#changing colours of points
EnhancedVolcano(fgf_data,
        lab = fgf_data$target_id,
        x = 'b',
        y = 'qval',
        title = 'FGF overexpression',
        pCutoff = 0.05,
        FCcutoff = 1.5,
        pointSize = 3.0,
        labSize = 6.0,
```

```
        col=c('grey', 'deepskyblue2', 'magenta', 'green'),
        colAlpha = 1)
#Here I have set the q-value threshold to 0.1 and a effect size threshold of 0.8
#this is for the FGF overexpressing data
EnhancedVolcano(fgf_data,
        lab = fgf_data$target_id,
        x = 'b',
        y = 'qval',
        xlab = bquote(~Log[2]~ 'effect size'),
        ylab = bquote(~-Log[10]~ 'q-value'),
        xlim = c(-5,5),
        ylim = c(0, 20),
        pCutoff = 0.1,
        FCcutoff = log2(1.75),
        cutoffLineWidth = 0.8,
        pointSize = 3.0,
        labSize = 6.0,
        col=c('grey', 'deepskyblue2', 'magenta', 'green'),
        colAlpha = 1,
        legendLabels=c(bquote(' Not sig.'), bquote(~Log[2]~'effect size'),
                bquote(' q-value'),
                bquote(' q-value &' ~Log[2]~'effect size')),
        legendPosition = 'right',
        legendLabSize = 16,
        legendIconSize = 5.0,
        selectLab = c('rasl11b', 'dusp6', 'egr1'))
#For the CIC knockdown data
EnhancedVolcano(CIC_data,
        lab = CIC_data$target_id,
        x = 'b',
        y = 'qval',
```

```
        xlab = bquote(~Log[2]~ 'effect size'),
        ylab = bquote(~-Log[10]~ 'q-value'),
        xlim = c(-5,5),
        ylim = c(0, 20),
        pCutoff = 0.1,
        FCcutoff = log2(1.75),
        cutoffLineWidth = 0.8,
        pointSize = 3.0,
        labSize = 6.0,
        col=c('grey', 'deepskyblue2', 'magenta', 'green'),
        colAlpha = 1,
        legendLabels=c(bquote(' Not sig.'), bquote(~Log[2]~'effect size'),
                bquote(' q-value'),
                bquote(' q-value &' ~Log[2]~'effect size')),
        legendPosition = 'right',
        legendLabSize = 16,
        legendIconSize = 5.0,
        selectLab = c('rasl11b'))
#Remember that in th excel files from the sleuth output the q-values are
#not transformed and volcano enhanced transforms them here. However, the
#effect sizes (b) are log2 transformed hence you need to log2 your effect
#size threshold here since enhanced volcano does not transform them here.
#There the threshold I have chosen are (TPM>1.5), q-value < 0.1 (-log10 of
#0.1 is 1 hence the threshold of the -log10 transformed q values on the
#plot is 1), effect size > log2(1.75) or effect < log2(-1.75) (log2 of
#1.75 is ~0.8 hence the threshold is set at 0.8 and -0.8 on the plot).
#subsetting FGF overexpression genes with a q-value less than 0.1
fgf_subset <- subset(fgf_data, fgf_data$qval < 0.1)
#Keeping only genes with an effect size > log2(1.75) in the subset
fgf_upregulated <- subset(fgf_subset, fgf_subset$b > log2(1.75))
#This should show all the upregulated genes with an effect size >0.8 and
```

```
#q-value < 0.1
fgf_upregulated
#creating table with all of the upregulated genes with these thresholds
write.table(fgf_upregulated, file="fgf_upregulated_genes.txt", sep="\t")
#keeping only genes with an effect size < -0.8 in the subset
fgf_downregulated <- subset(fgf_subset, fgf_subset$b < -log2(1.75))
fgf_downregulated
write.table(fgf_downregulated, file="fgf_downregulated_genes.txt", sep="\t")
#subsetting CIC knockdown genes with a q-value less than 0.1
CIC_subset <- subset(CIC_data, CIC_data$qval < 0.1)
#Keeping only genes with an effect size > 0.8 in the subset
CIC_upregulated <- subset(CIC_subset, CIC_subset$b > log2(1.75))
##This should show all the upregulated genes with an effect size >0.8 and
#q-value < 0.1
CIC_upregulated
#creating table with all of the upregulated genes with these thresholds
write.table(CIC_upregulated, file="CIC_upregulated_genes.txt", sep="\t")
#keeping only genes with an effect size < -0.8 in the subset
CIC_downregulated <- subset(CIC_subset, CIC_subset$b < -log2(1.75))
CIC_downregulated
write.table(CIC_downregulated, file="CIC_downregulated_genes.txt", sep="\t")
##############################################################################
#Making bar chart figure for panther gene ontology analysis of the genes
#upregulated in embryos following FGF overexpression
#reading in table containing panther analysis
panther_FGF <- read.table("FGF_upregulated_panther_list.txt", header=T)
# Basic barplot
p<-ggplot(data=panther_FGF, aes(x=Fold_enrichment,
        y=Panther_Go.Slim_Biological_Process)) +
 geom_bar(stat="identity")
#viewing basic barplot
```

```
p
#Ordering the Go-Slim Biological process by fold enrichment from
#smallest to largest
panther_FGF$Panther_Go.Slim_Biological_Process <-
 factor(panther_FGF$Panther_Go.Slim_Biological_Process,
       levels = c("Unclassified", "System_development",
                  "Intracellular_signal_transduction",
                  "Cellular_response_to_growth_factor_stimulus",
                  "Cell_population_proliferation",
                  "Regulation_of_cell_population_proliferation",
                  "Negative_regulation_of_intracellular_signal_transduction"))
#Creating bar graph
p<-ggplot(data=panther_FGF,
        aes(x=Fold_enrichment,
            y=Panther_Go.Slim_Biological_Process)) +
 scale_y_discrete(labels=c("Unclassified", "System development",
        "Intracellular signal transduction",
        "Cellular response to growth factor stimulus",
        "Cell population proliferation",
        "Regulation of cell population proliferation",
        "Negative regulation of intracellular signal
        transduction"))+
 geom_bar(stat="identity", fill="royalblue1", color="black", width=0.5)+
 ylab("PANTHER GO-Slim biological process")+
 xlab("Fold enrichment")+
 theme_minimal() +
 theme(axis.text = element_text(size = 10))+
 theme(axis.title = element_text(size = 15))+
 xlim(0,25)
#Viewing bar graph
p
```

```
##Making bar chart figure for panther gene ontology analysis of the genes
#upregulated in embryos following CIC knockdown
#Reading in table containing panther analysis
panther_CIC <- read.table("CIC_upregulated_panther_list.txt", header=T)
#Basic barplot
q<-ggplot(data=panther_CIC, aes(x=Fold_enrichment,
                        y=Panther_Go.Slim_Biological_Process)) +
 geom_bar(stat="identity")
#View basic barplot
q
#Ordering the Go-Slim Biological process by fold enrichment from
#smallest to largest
panther_CIC$Panther_Go.Slim_Biological_Process <-
 factor(panther_CIC$Panther_Go.Slim_Biological_Process,
       levels = c("Regulation_of_metabolic_process",
                  "Cellular_catabolic_process",
                  "Cell_Cycle",
                  "Regulation_of_cell_cycle"))
#Creating bar graph
q<-ggplot(data=panther_CIC,
        aes(x=Fold_enrichment,
            y=Panther_Go.Slim_Biological_Process)) +
 scale_y_discrete(labels=c("Regulation of metabolic process",
                           "Cellular catabolic process",
                           "Cell cycle",
                           "Regulation of cell cycle"))+
 geom_bar(stat="identity", fill="royalblue1", color="black", width=0.4)+
 ylab("PANTHER GO-Slim biological process")+
 xlab("Fold enrichment")+
 theme_minimal() +
 theme(axis.text = element_text(size = 10))+
```

```
 theme(axis.title = element_text(size = 12))+
 xlim(0,6)
#Viewing bar graph
q
################################################################################
#Calculating the significance of gene overlaps between FGF overexpressing and
#CIC knockdown embryos
#Recall list of genes upregulated in embryos following FGF overexpression
fgf_upregulated
#Recall list of genes upregulated in embryos following CIC knockdown
CIC_upregulated
#Installing BiocManager packaged required for to load GeneOverlap package
if (!require("BiocManager", quietly = TRUE))
 install.packages("BiocManager")
#Installing the GeneOverlap package
BiocManager::install("GeneOverlap")
#Loading the GeneOverlap package
library(GeneOverlap)
#defining the number of genes in Xenopus tropicalis genome
genome <- 23635
#constructing the GeneOverlap object for upregulated genes
go.obj <- newGeneOverlap(fgf_upregulated$target_id,
                         CIC_upregulated$target_id,
                         genome.size=genome)
#viewing the GeneOverlap object
go.obj
#Testing the significance of the association
go.obj <- testGeneOverlap(go.obj)
go.obj
#GeneOverlap identified an overlap of 27 genes that were both upregulated in CIC
#knockdown and FGF overexpressing embryos. Fisher's exact test was then used to
```

```
#identify the statistical significance of the observed overlap. The test identified
#that this overlap is very highly significant (P=9.4x10^-33) so it is unlikely
#that this number of genes was likely to have been upregulated in both FGF
#overexpressing and CIC knockdown embryos by chance.
#Getting additional information from Fisher's exact test
print(go.obj)
#The Fisher' exact test also calculates an Odds ratio which is indicative of the
#strength of association between two variables (the more greater than 1 the odds
#ratio is, the strongr the association). The odds ratio for the lists of genes
#upregulated in FGF overexpressing and CIC knockdown embryos is 52.6 which
#demonstrates a very strong association between the lists and is consistent with
#the large number of genes overlapping between them.
#Recall list of genes downregulated in embryos following FGF overexpression
fgf_downregulated
#Recall list of genes downregulated in embryos following CIC knockdown
CIC_downregulated
#constructing the GeneOverlap object for downregulated genes
go.obj1 <- newGeneOverlap(fgf_downregulated$target_id,
                          CIC_downregulated$target_id,
                          genome.size=genome)
#viewing the GeneOverlap object
go.obj1
#Testing the significance of the association
go.obj1 <- testGeneOverlap(go.obj1)
go.obj1
#GeneOverlap identified an overlap of 9 genes that were both downregulated in CIC
#knockdown and FGF overexpressing embryos. Fisher's exact test was then used to
#identify the statistical significance of the observed overlap. The test identified
#that this overlap is very highly significant (P=2.8x10^-18) so it is unlikely
#that this number of genes was likely to have been downregulated in both FGF
#overexpressing and CIC knockdown embryos by chance.
```

```
#Getting additional information from Fisher's exact test
print(go.obj1)
#The odds ratio for the lists of genes downregulated in FGF overexpressing and CIC
#knockdown embryos is 264.2 which demonstrates a very strong association between
#the lists and is consistent with the large number of genes overlapping between
#them.
#Generating lists containing the genes upregulated and downregulated
#in CIC knockdown and FGF overexpressing embryos respectively
FGF_genes_upregulated <- fgf_upregulated$target_id
FGF_genes_downregulated <- fgf_downregulated$target_id
CIC_genes_upregulated <- CIC_upregulated$target_id
CIC_genes_downregulated <- CIC_downregulated$target_id
FGF_genes <- list("FGF upregulated"=FGF_genes_upregulated,
            "FGF downregulated"=FGF_genes_downregulated)
CIC_genes <- list("CIC upregulated"=CIC_genes_upregulated,
            "CIC downregulated"=CIC_genes_downregulated)
#Generating matrix plot visualising the Odds Ratio and P-value of
#Fisher's exact tests between the different gene lists
#Constructing basic matrix plot
gom.obj <- newGOM(FGF_genes, CIC_genes,
            + genome)
dev.off()
drawHeatmap(gom.obj)
drawHeatmap(gom.obj, log.scale = T, adj.p=F, ncolused=6,
        grid.col = "Blues", note.col="red")
#Constructing complete matrix plot
gom.obj3 <- newGOM(FGF_genes, CIC_genes + genome)
dev.off()
drawHeatmap(gom.obj3, adj.p=TRUE, cutoff=1,
        ncolused=5, grid.col="Blues", note.col="black")
drawHeatmap(gom.obj3)
```

```
################################################################################
#Creating bar graph of RT-qPCR data of Stage 10.5 X. tropicalis embyros
#injected with XFD
#Reading in text file containing RT-qPCR data
XFD_10.5 <- read.table("XFD_stage_10.5.txt", header=T)
#loading ggplot2
library(ggplot2)
#Specifying the order that genes will be displayed on the bar plot
XFD_10.5$Gene <- factor(XFD_10.5$Gene, levels = c("tbxt", "rasl11b"))
#constructing basic bar plot
XFD <- ggplot(data=XFD_10.5, aes(x=Gene, y=Fold_change,
           fill=Sample)) +
 geom_bar(stat="identity", color="black",
        position = position_dodge())+
 geom_text(aes(label=Fold_change),
        position = position_dodge(0.9), color="white", size=8,
        vjust = 1.6)+
 theme_minimal() +
 ylab("Relative expression")+
 xlab("Gene")+
 theme(axis.title = element_text(size = 25))+
 theme(axis.text = element_text(size = 20))+
 theme(legend.title = element_text(size=25))+
 theme(legend.text = element_text(size=20))+
 theme(legend.key.size = unit(1.2, 'cm'))+
 scale_fill_brewer(palette="Paired")+
theme(axis.line = element_line(size = 0.5, colour = "black",
                       linetype=1))
#Viewing bar plot
XFD
#Creating bar graph of RT-qPCR data of Stage 12 X. tropicalis embyros
```

```
#injected with XFD
#Reading in text file containing RT-qPCR data
XFD_12 <- read.table("XFD_stage_12.txt", header=T)
#Specifying the order that genes will be displayed on the bar plot
XFD_12$Gene <- factor(XFD_12$Gene, levels = c("tbxt", "rasl11b"))
#Constructing bar plot
XFD1 <- ggplot(data=XFD_12, aes(x=Gene, y=Fold_change,
                          fill=Sample)) +
 geom_bar(stat="identity", color="black",
         position = position_dodge())+
 geom_text(aes(label=Fold_change),
          position = position_dodge(0.9), color="white", size=8,
          vjust = 1.6)+
 theme_minimal() +
 ylab("Relative expression")+
 xlab("Gene")+
 theme(axis.title = element_text(size = 25))+
 theme(axis.text = element_text(size = 20))+
 theme(legend.title = element_text(size=25))+
 theme(legend.text = element_text(size=20))+
 theme(legend.key.size = unit(1.2, 'cm'))+
 scale_fill_manual(values=c("mediumorchid2", "mediumorchid4"))+
 theme(axis.line = element_line(size = 0.5, colour = "black",
                          linetype=1))
#Viewing bar plot
XFD1
```

References

AMAYA, E., MUSCI, T. J. & KIRSCHNER, M. W. 1991. Expression of a dominant negative mutant of the FGF receptor disrupts mesoderm formation in Xenopus embryos. *Cell,* 66**,** 257-70.

AMAYA, E., STEIN, P. A., MUSCI, T. J. & KIRSCHNER, M. W. 1993. FGF signalling in the early specification of mesoderm in Xenopus. *Development,* 118**,** 477-87.

ANASTASAKI, C., RAUEN, K. A. & PATTON, E. E. 2012. Continual low-level MEK inhibition ameliorates cardio-facio-cutaneous phenotypes in zebrafish. *Dis Model Mech,* 5**,** 546-52.

ASTIGARRAGA, S., GROSSMAN, R., DÍAZ-DELFÍN, J., CAELLES, C., PAROUSH, Z. & JIMÉNEZ, G. 2007. A MAPK docking site is critical for downregulation of Capicua by Torso and EGFR RTK signaling. *EMBO J,* 26**,** 668-77.

AULEHLA, A. & POURQUIÉ, O. 2010. Signaling gradients during paraxial mesoderm development. *Cold Spring Harb Perspect Biol,* 2**,** a000869.

AULEHLA, A., WEHRLE, C., BRAND-SABERI, B., KEMLER, R., GOSSLER, A., KANZLER, B. & HERRMANN, B. G. 2003. Wnt3a plays a major role in the segmentation clock controlling somitogenesis. *Dev Cell,* 4**,** 395-406.

AVIEZER, D., GOLEMBO, M. & YAYON, A. 2003. Fibroblast growth factor receptor-3 as a therapeutic target for Achondroplasia--genetic short limbed dwarfism. *Curr Drug Targets,* 4**,** 353-65.

AVIEZER, D., HECHT, D., SAFRAN, M., EISINGER, M., DAVID, G. & YAYON, A. 1994. Perlecan, basal lamina proteoglycan, promotes basic fibroblast growth factor-receptor binding, mitogenesis, and angiogenesis. *Cell,* 79**,** 1005-13.

BAILEY, T. L., JOHNSON, J., GRANT, C. E. & NOBLE, W. S. 2015. The MEME Suite. *Nucleic Acids Res,* 43**,** W39-49.

BARAK, H., HUH, S. H., CHEN, S., JEANPIERRE, C., MARTINOVIC, J., PARISOT, M., BOLE-FEYSOT, C., NITSCHKÉ, P., SALOMON, R., ANTIGNAC, C., ORNITZ, D. M. & KOPAN, R. 2012. FGF9 and FGF20 maintain the stemness of nephron progenitors in mice and man. *Dev Cell,* 22**,** 1191-207.

BELOV, A. A. & MOHAMMADI, M. 2013. Molecular mechanisms of fibroblast growth factor signaling in physiology and pathology. *Cold Spring Harb Perspect Biol,* 5.

BORELLO, U., COBOS, I., LONG, J. E., MCWHIRTER, J. R., MURRE, C. & RUBENSTEIN, J. L. 2008. FGF15 promotes neurogenesis and opposes FGF8 function during neocortical development. *Neural Dev,* 3**,** 17.

BRANNEY, P. A., FAAS, L., STEANE, S. E., POWNALL, M. E. & ISAACS, H. V. 2009. Characterisation of the fibroblast growth factor dependent transcriptome in early development. *PLoS One,* 4**,** e4951.

BRENT, A. E. & TABIN, C. J. 2004. FGF acts directly on the somitic tendon progenitors through the Ets transcription factors Pea3 and Erm to regulate scleraxis expression. *Development,* 131**,** 3885-96.

BÖTTCHER, R. T. & NIEHRS, C. 2005. Fibroblast growth factor signaling during early vertebrate development. *Endocr Rev,* 26**,** 63-77.

CARGNELLO, M. & ROUX, P. P. 2011. Activation and function of the MAPKs and their substrates, the MAPK-activated protein kinases. *Microbiol Mol Biol Rev,* 75**,** 50-83.

CHANG, Y. F., IMAM, J. S. & WILKINSON, M. F. 2007. The nonsense-mediated decay RNA surveillance pathway. *Annu Rev Biochem,* 76**,** 51-74.

CHARLOT, C., DUBOIS-POT, H., SERCHOV, T., TOURRETTE, Y. & WASYLYK, B. 2010. A review of post-translational modifications and subcellular localization of Ets transcription factors: possible connection with cancer and involvement in the hypoxic response. *Methods Mol Biol,* 647**,** 3-30.

CHRISTEN, B. & SLACK, J. M. 1999. Spatial response to fibroblast growth factor signalling in Xenopus embryos. *Development,* 126**,** 119-25.
COLICELLI, J. 2004. Human RAS superfamily proteins and related GTPases. *Sci STKE,* 2004**,** RE13.
COWELL, L. 2019. *Investigation of the role of Capicua in the FGF signalling pathway and wound response.* PhD, University of York.
DEIMLING, S. J. & DRYSDALE, T. A. 2011. Fgf is required to regulate anterior-posterior patterning in the Xenopus lateral plate mesoderm. *Mech Dev,* 128**,** 327-41.
DELFINI, M. C., DUBRULLE, J., MALAPERT, P., CHAL, J. & POURQUIÉ, O. 2005. Control of the segmentation process by graded MAPK/ERK activation in the chick embryo. *Proc Natl Acad Sci U S A,* 102**,** 11343-8.
DICKINSON, K., LEONARD, J. & BAKER, J. C. 2006. Genomic profiling of mixer and Sox17beta targets during Xenopus endoderm development. *Dev Dyn,* 235**,** 368-81.
DIECKGRAEFE, B. K. & WEEMS, D. M. 1999. Epithelial injury induces egr-1 and fos expression by a pathway involving protein kinase C and ERK. *Am J Physiol,* 276**,** G322-30.
DIECKGRAEFE, B. K., WEEMS, D. M., SANTORO, S. A. & ALPERS, D. H. 1997. ERK and p38 MAP kinase pathways are mediators of intestinal epithelial wound-induced signal transduction. *Biochem Biophys Res Commun,* 233**,** 389-94.
DIEZ DEL CORRAL, R., OLIVERA-MARTINEZ, I., GORIELY, A., GALE, E., MADEN, M. & STOREY, K. 2003. Opposing FGF and retinoid pathways control ventral neural pattern, neuronal differentiation, and segmentation during body axis extension. *Neuron,* 40**,** 65-79.
DISSANAYAKE, K., TOTH, R., BLAKEY, J., OLSSON, O., CAMPBELL, D. G., PRESCOTT, A. R. & MACKINTOSH, C. 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**,** 515-25.
DRAPER, B. W., MORCOS, P. A. & KIMMEL, C. B. 2001. Inhibition of zebrafish fgf8 pre-mRNA splicing with morpholino oligos: a quantifiable method for gene knockdown. *Genesis,* 30**,** 154-6.
DUBRULLE, J., MCGREW, M. J. & POURQUIÉ, O. 2001. FGF signaling controls somite boundary position and regulates segmentation clock control of spatiotemporal Hox gene activation. *Cell,* 106**,** 219-32.
DUBRULLE, J. & POURQUIÉ, O. 2004. fgf8 mRNA decay establishes a gradient that couples axial elongation to patterning in the vertebrate embryo. *Nature,* 427**,** 419-22.
EKEROT, M., STAVRIDIS, M. P., DELAVAINE, L., MITCHELL, M. P., STAPLES, C., OWENS, D. M., KEENAN, I. D., DICKINSON, R. J., STOREY, K. G. & KEYSE, S. M. 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**,** 287-98.
ELEVELD, T. F., OLDRIDGE, D. A., BERNARD, V., KOSTER, J., COLMET DAAGE, L., DISKIN, S. J., SCHILD, L., BENTAHAR, N. B., BELLINI, A., CHICARD, M., LAPOUBLE, E., COMBARET, V., LEGOIX-NÉ, P., MICHON, J., PUGH, T. J., HART, L. S., RADER, J., ATTIYEH, E. F., WEI, J. S., ZHANG, S., NARANJO, A., GASTIER-FOSTER, J. M., HOGARTY, M. D., ASGHARZADEH, S., SMITH, M. A., GUIDRY AUVIL, J. M., WATKINS, T. B., ZWIJNENBURG, D. A., EBUS, M. E., VAN SLUIS, P., HAKKERT, A., VAN WEZEL, E., VAN DER SCHOOT, C. E., WESTERHOUT, E. M., SCHULTE, J. H., TYTGAT, G. A., DOLMAN, M. E., JANOUEIX-LEROSEY, I., GERHARD, D. S., CARON, H. N., DELATTRE, O., KHAN, J., VERSTEEG, R., SCHLEIERMACHER, G., MOLENAAR, J. J. & MARIS, J. M. 2015. Relapsed neuroblastomas show frequent RAS-MAPK pathway mutations. *Nat Genet,* 47**,** 864-71.
ESWARAKUMAR, V. P., LAX, I. & SCHLESSINGER, J. 2005. Cellular signaling by fibroblast growth factor receptors. *Cytokine Growth Factor Rev,* 16**,** 139-49.
FISHER, M. E., ISAACS, H. V. & POWNALL, M. E. 2002. eFGF is required for activation of XmyoD expression in the myogenic cell lineage of Xenopus laevis. *Development,* 129**,** 1307-15.

FLETCHER, R. B. & HARLAND, R. M. 2008. The role of FGF signaling in the establishment and maintenance of mesodermal gene expression in Xenopus. *Dev Dyn,* 237**,** 1243-54.
FORÉS, M., SIMÓN-CARRASCO, L., AJURIA, L., SAMPER, N., GONZÁLEZ-CRESPO, S., DROSTEN, M., BARBACID, M. & JIMÉNEZ, G. 2017. A new mode of DNA binding distinguishes Capicua from other HMG-box factors and explains its mutation patterns in cancer. *PLoS Genet,* 13**,** e1006622.
FOULDS, C. E., NELSON, M. L., BLASZCZAK, A. G. & GRAVES, B. J. 2004. Ras/mitogen-activated protein kinase signaling activates Ets-1 and Ets-2 by CBP/p300 recruitment. *Mol Cell Biol,* 24**,** 10954-64.
FREEMAN, S. D., MOORE, W. M., GUIRAL, E. C., HOLME, A. D., TURNBULL, J. E. & POWNALL, M. E. 2008. Extracellular regulation of developmental cell signaling by XtSulf1. *Dev Biol,* 320**,** 436-45.
GILES, R. V., SPILLER, D. G., CLARK, R. E. & TIDD, D. M. 1999. Antisense morpholino oligonucleotide analog induces missplicing of C-myc mRNA. *Antisense Nucleic Acid Drug Dev,* 9**,** 213-20.
GILLIES, T. E., PARGETT, M., SILVA, J. M., TERAGAWA, C. K., MCCORMICK, F. & ALBECK, J. G. 2020. Oncogenic mutant RAS signaling activity is rescaled by the ERK/MAPK pathway. *Mol Syst Biol,* 16**,** e9518.
GOLDFARB, M. 2005. Fibroblast growth factor homologous factors: evolution, structure, and function. *Cytokine Growth Factor Rev,* 16**,** 215-20.
GOTOH, N. 2008. Regulation of growth factor signaling by FRS2 family docking/scaffold adaptor proteins. *Cancer Sci,* 99**,** 1319-25.
GRAINGER, R. M. 2012. Xenopus tropicalis as a model organism for genetics and genomics: past, present, and future. *Methods Mol Biol,* 917**,** 3-15.
GRANT, C. E., BAILEY, T. L. & NOBLE, W. S. 2011. FIMO: scanning for occurrences of a given motif. *Bioinformatics,* 27**,** 1017-8.
GUILLE, M. 1999. *Molecular Methods in Developmental Biology Xenopus and Zebrafish,* Totowa, NJ, Humana Press.
GÓMEZ, A. R., LÓPEZ-VAREA, A., MOLNAR, C., DE LA CALLE-MUSTIENES, E., RUIZ-GÓMEZ, M., GÓMEZ-SKARMETA, J. L. & DE CELIS, J. F. 2005. Conserved cross-interactions in Drosophila and Xenopus between Ras/MAPK signaling and the dual-specificity phosphatase MKP3. *Dev Dyn,* 232**,** 695-708.
HADARI, Y. R., KOUHARA, H., LAX, I. & SCHLESSINGER, J. 1998. Binding of Shp2 tyrosine phosphatase to FRS2 is essential for fibroblast growth factor-induced PC12 cell differentiation. *Mol Cell Biol,* 18**,** 3966-73.
HARLAND, R. M. 1991. In situ hybridization: an improved whole-mount method for Xenopus embryos. *Methods Cell Biol,* 36**,** 685-95.
HARLAND, R. M. & GRAINGER, R. M. 2011. Xenopus research: metamorphosed by genetics and genomics. *Trends Genet,* 27**,** 507-15.
HITACHI, K., DANNO, H., TAZUMI, S., AIHARA, Y., UCHIYAMA, H., OKABAYASHI, K., KONDOW, A. & ASASHIMA, M. 2009. The Xenopus Bowline/Ripply family proteins negatively regulate the transcriptional activity of T-box transcription factors. *Int J Dev Biol,* 53**,** 631-9.
HOFFMAN, G. E., SMITH, M. S. & VERBALIS, J. G. 1993. c-Fos and related immediate early gene products as markers of activity in neuroendocrine systems. *Front Neuroendocrinol,* 14**,** 173-213.
HU, M. C., SHIIZAKI, K., KURO-O, M. & MOE, O. W. 2013. Fibroblast growth factor 23 and Klotho: physiology and pathophysiology of an endocrine network of mineral metabolism. *Annu Rev Physiol,* 75**,** 503-33.
HUFTON, A. L., VINAYAGAM, A., SUHAI, S. & BAKER, J. C. 2006. Genomic analysis of Xenopus organizer function. *BMC Dev Biol,* 6**,** 27.
ISAACS, H. V., POWNALL, M. E. & SLACK, J. M. 1994. eFGF regulates Xbra expression during Xenopus gastrulation. *EMBO J,* 13**,** 4469-81.

ISAACS, H. V., POWNALL, M. E. & SLACK, J. M. 1995. eFGF is expressed in the dorsal midline of Xenopus laevis. *Int J Dev Biol,* 39**,** 575-9.
ITOH, N. 2007. The Fgf families in humans, mice, and zebrafish: their evolutional processes and roles in development, metabolism, and disease. *Biol Pharm Bull,* 30**,** 1819-25.
ITOH, N. 2010. Hormone-like (endocrine) Fgfs: their evolutionary history and roles in development, metabolism, and disease. *Cell Tissue Res,* 342**,** 1-11.
ITOH, N. & ORNITZ, D. M. 2004. Evolution of the Fgf and Fgfr gene families. *Trends Genet,* 20**,** 563-9.
ITOH, N. & ORNITZ, D. M. 2008. Functional evolutionary history of the mouse Fgf gene family. *Dev Dyn,* 237**,** 18-27.
JIMÉNEZ, G., GUICHET, A., EPHRUSSI, A. & CASANOVA, J. 2000. Relief of gene repression by torso RTK signaling: role of capicua in Drosophila terminal and dorsoventral patterning. *Genes Dev,* 14**,** 224-31.
JIMÉNEZ, G., SHVARTSMAN, S. Y. & PAROUSH, Z. 2012. The Capicua repressor--a general sensor of RTK signaling in development and disease. *J Cell Sci,* 125**,** 1383-91.
JIN, Y., HA, N., FORÉS, M., XIANG, J., GLÄßER, C., MALDERA, J., JIMÉNEZ, G. & EDGAR, B. A. 2015. EGFR/Ras Signaling Controls Drosophila Intestinal Stem Cell Proliferation via Capicua-Regulated Genes. *PLoS Genet,* 11**,** e1005634.
JOUNG, J. K. & SANDER, J. D. 2013. TALENs: a widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol,* 14**,** 49-55.
KEENAN, S. E., BLYTHE, S. A., MARMION, R. A., DJABRAYAN, N. J., WIESCHAUS, E. F. & SHVARTSMAN, S. Y. 2020. Rapid Dynamics of Signal-Dependent Transcriptional Repression by Capicua. *Dev Cell,* 52**,** 794-801.e4.
KENT, W. J., SUGNET, C. W., FUREY, T. S., ROSKIN, K. M., PRINGLE, T. H., ZAHLER, A. M. & HAUSSLER, D. 2002. The human genome browser at UCSC. *Genome Res,* 12**,** 996-1006.
KIM, Y. G., CHA, J. & CHANDRASEGARAN, S. 1996. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc Natl Acad Sci U S A,* 93**,** 1156-60.
KING, M. 2019. *Investigating the role of Capicua in mediating FGF transcriptional regulation X. tropicalis.* PhD, University of York.
KONDOW, A., HITACHI, K., OKABAYASHI, K., HAYASHI, N. & ASASHIMA, M. 2007. Bowline mediates association of the transcriptional corepressor XGrg-4 with Tbx6 during somitogenesis in Xenopus. *Biochem Biophys Res Commun,* 359**,** 959-64.
KOUHARA, H., HADARI, Y. R., SPIVAK-KROIZMAN, T., SCHILLING, J., BAR-SAGI, D., LAX, I. & SCHLESSINGER, J. 1997. A lipid-anchored Grb2-binding protein that links FGF-receptor activation to the Ras/MAPK signaling pathway. *Cell,* 89**,** 693-702.
LAESTANDER, C. & ENGSTRÖM, W. 2014. Role of fibroblast growth factors in elicitation of cell responses. *Cell Prolif,* 47**,** 3-11.
LEA, R., PAPALOPULU, N., AMAYA, E. & DOREY, K. 2009. Temporal and spatial expression of FGF ligands and receptors during Xenopus development. *Dev Dyn,* 238**,** 1467-79.
LEE, M. T., BONNEAU, A. R. & GIRALDEZ, A. J. 2014. Zygotic genome activation during the maternal-to-zygotic transition. *Annu Rev Cell Dev Biol,* 30**,** 581-613.
LEE, S. Y., YOON, J., LEE, H. S., HWANG, Y. S., CHA, S. W., JEONG, C. H., KIM, J. I., PARK, J. B., LEE, J. Y., KIM, S., PARK, M. J., DONG, Z. & KIM, J. 2011. The function of heterodimeric AP-1 comprised of c-Jun and c-Fos in activin mediated Spemann organizer gene expression. *PLoS One,* 6**,** e21796.
LEE, Y. 2020. Regulation and function of capicua in mammals. *Exp Mol Med,* 52**,** 531-537.
LEI, Y., GUO, X., LIU, Y., CAO, Y., DENG, Y., CHEN, X., CHENG, C. H., DAWID, I. B., CHEN, Y. & ZHAO, H. 2012. Efficient targeted gene disruption in Xenopus embryos using engineered transcription activator-like effector nucleases (TALENs). *Proc Natl Acad Sci U S A,* 109**,** 17484-9.

LEWIS, T., GROOM, L. A., SNEDDON, A. A., SMYTHE, C. & KEYSE, S. M. 1995. XCL100, an inducible nuclear MAP kinase phosphatase from Xenopus laevis: its role in MAP kinase inactivation in differentiated cells and its expression during early development. *J Cell Sci,* 108 (Pt 8)**,** 2885-96.

LIN, M., WHITMIRE, S., CHEN, J., FARREL, A., SHI, X. & GUO, J. T. 2017. Effects of short indels on protein structure and function in human genomes. *Sci Rep,* 7**,** 9313.

LIN, X. 2004. Functions of heparan sulfate proteoglycans in cell signaling during development. *Development,* 131**,** 6009-21.

LIU, B. A., ENGELMANN, B. W., JABLONOWSKI, K., HIGGINBOTHAM, K., STERGACHIS, A. B. & NASH, P. D. 2012. SRC Homology 2 Domain Binding Sites in Insulin, IGF-1 and FGF receptor mediated signaling networks reveal an extensive potential interactome. *Cell Commun Signal,* 10**,** 27.

LIU, J. & LI, Y. 2019. Upregulation of MAPK10, TUBB2B and RASL11B may contribute to the development of neuroblastoma. *Mol Med Rep,* 20**,** 3475-3486.

LOBJOIS, V., BENAZERAF, B., BERTRAND, N., MEDEVIELLE, F. & PITUELLO, F. 2004. Specific regulation of cyclins D1 and D2 by FGF and Shh signaling coordinates cell cycle progression, patterning, and differentiation during early steps of spinal cord development. *Dev Biol,* 273**,** 195-209.

LUO, Y., WANG, A. T., ZHANG, Q. F., LIU, R. M. & XIAO, J. H. 2020. gene enhances hyaluronic acid-mediated chondrogenic differentiation in human amniotic mesenchymal stem cells via the activation of Sox9/ERK/smad signals. *Exp Biol Med (Maywood),* 245**,** 1708-1721.

LUXARDI, G., MARCHAL, L., THOMÉ, V. & KODJABACHIAN, L. 2010. Distinct Xenopus Nodal ligands sequentially induce mesendoderm and control gastrulation movements in parallel to the Wnt/PCP pathway. *Development,* 137**,** 417-26.

MANI, M., SMITH, J., KANDAVELOU, K., BERG, J. M. & CHANDRASEGARAN, S. 2005. Binding of two zinc finger nuclease monomers to two specific sites is required for effective double-strand DNA cleavage. *Biochem Biophys Res Commun,* 334**,** 1191-1197.

MATSUBAYASHI, Y., EBISUYA, M., HONJOH, S. & NISHIDA, E. 2004. ERK activation propagates in epithelial cell sheets and regulates their migration during wound healing. *Curr Biol,* 14**,** 731-5.

MATSUO, I. & KIMURA-YOSHIDA, C. 2013. Extracellular modulation of Fibroblast Growth Factor signaling through heparan sulfate proteoglycans in mammalian development. *Curr Opin Genet Dev,* 23**,** 399-407.

MCINTOSH, I., BELLUS, G. A. & JAB, E. W. 2000. The pleiotropic effects of fibroblast growth factor receptors in mammalian development. *Cell Struct Funct,* 25**,** 85-96.

MI, H., EBERT, D., MURUGANUJAN, A., MILLS, C., ALBOU, L. P., MUSHAYAMAHA, T. & THOMAS, P. D. 2021. PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic Acids Res,* 49**,** D394-D403.

MI, H., MURUGANUJAN, A., HUANG, X., EBERT, D., MILLS, C., GUO, X. & THOMAS, P. D. 2019. Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). *Nat Protoc,* 14**,** 703-721.

MILUNSKY, J. M., ZHAO, G., MAHER, T. A., COLBY, R. & EVERMAN, D. B. 2006. LADD syndrome is caused by FGF10 mutations. *Clin Genet,* 69**,** 349-54.

MOHAMMADI, M., MCMAHON, G., SUN, L., TANG, C., HIRTH, P., YEH, B. K., HUBBARD, S. R. & SCHLESSINGER, J. 1997. Structures of the tyrosine kinase domain of fibroblast growth factor receptor in complex with inhibitors. *Science,* 276**,** 955-60.

MORENO, T. A. & KINTNER, C. 2004. Regulation of segmental patterning by retinoic acid signaling during Xenopus somitogenesis. *Dev Cell,* 6**,** 205-18.

MOULTON, J. D. 2007. Using morpholinos to control gene expression. *Curr Protoc Nucleic Acid Chem,* Chapter 4**,** Unit 4.30.

NASEVICIUS, A. & EKKER, S. C. 2000. Effective targeted gene 'knockdown' in zebrafish. *Nat Genet,* 26**,** 216-20.
NENTWICH, O., DINGWELL, K. S., NORDHEIM, A. & SMITH, J. C. 2009. Downstream of FGF during mesoderm formation in Xenopus: the roles of Elk-1 and Egr-1. *Dev Biol,* 336**,** 313-26.
NIES, V. J., SANCAR, G., LIU, W., VAN ZUTPHEN, T., STRUIK, D., YU, R. T., ATKINS, A. R., EVANS, R. M., JONKER, J. W. & DOWNES, M. R. 2015. Fibroblast Growth Factor Signaling in Metabolic Regulation. *Front Endocrinol (Lausanne),* 6**,** 193.
NIEUWKOOP, P. D. & FABER, J. 1994. *Normal table of Xenopus laevis (Daudin) : a systematical and chronological survey of the development from the fertilized egg till the end of metamorphosis,* New York, Garland Pub.
OKAZAKI, K. & SAGATA, N. 1995. The Mos/MAP kinase pathway stabilizes c-Fos by phosphorylation and augments its transforming activity in NIH 3T3 cells. *EMBO J,* 14**,** 5048-59.
OKIMOTO, R. A., BREITENBUECHER, F., OLIVAS, V. R., WU, W., GINI, B., HOFREE, M., ASTHANA, S., HRUSTANOVIC, G., FLANAGAN, J., TULPULE, A., BLAKELY, C. M., HARINGSMA, H. J., SIMMONS, A. D., GOWEN, K., SUH, J., MILLER, V. A., ALI, S., SCHULER, M. & BIVONA, T. G. 2017. Inactivation of Capicua drives cancer metastasis. *Nat Genet,* 49**,** 87-96.
ONG, S. H., GUY, G. R., HADARI, Y. R., LAKS, S., GOTOH, N., SCHLESSINGER, J. & LAX, I. 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**,** 979-89.
OPPENHEIMER, J. M. 1936. Transplantation experiments on developing teleosts (Fundulus and Perca). *Journal of Experimental Zoology,* 72**,** 409-437.
ORNITZ, D. M. & ITOH, N. 2015. The Fibroblast Growth Factor signaling pathway. *Wiley Interdiscip Rev Dev Biol,* 4**,** 215-66.
OWENS, N. D. L., BLITZ, I. L., LANE, M. A., PATRUSHEV, I., OVERTON, J. D., GILCHRIST, M. J., CHO, K. W. Y. & KHOKHA, M. K. 2016. Measuring Absolute RNA Copy Numbers at High Temporal Resolution Reveals Transcriptome Kinetics in Development. *Cell Rep,* 14**,** 632-647.
PANITZ, F., KRAIN, B., HOLLEMANN, T., NORDHEIM, A. & PIELER, T. 1998. The Spemann organizer-expressed zinc finger gene Xegr-1 responds to the MAP kinase/Ets-SRF signal transduction pathway. *EMBO J,* 17**,** 4414-25.
PASSOS-BUENO, M. R., WILCOX, W. R., JABS, E. W., SERTIÉ, A. L., ALONSO, L. G. & KITOH, H. 1999. Clinical spectrum of fibroblast growth factor receptor mutations. *Hum Mutat,* 14**,** 115-25.
PAWSON, T., OLIVIER, P., ROZAKIS-ADCOCK, M., MCGLADE, J. & HENKEMEYER, M. 1993. Proteins with SH2 and SH3 domains couple receptor tyrosine kinases to intracellular signalling pathways. *Philos Trans R Soc Lond B Biol Sci,* 340**,** 279-85.
PIROOZNIA, M., NAGARAJAN, V. & DENG, Y. 2007. GeneVenn - A web application for comparing gene lists using Venn diagrams. *Bioinformation,* 1**,** 420-2.
POPOV, I. K., KWON, T., CROSSMAN, D. K., CROWLEY, M. R., WALLINGFORD, J. B. & CHANG, C. 2017. Identification of new regulators of embryonic patterning and morphogenesis in Xenopus gastrulae by RNA sequencing. *Dev Biol,* 426**,** 429-441.
POTTHOFF, M. J., KLIEWER, S. A. & MANGELSDORF, D. J. 2012. Endocrine fibroblast growth factors 15/19 and 21: from feast to famine. *Genes Dev,* 26**,** 312-24.
POWERS, C. J., MCLESKEY, S. W. & WELLSTEIN, A. 2000. Fibroblast growth factors, their receptors and signaling. *Endocr Relat Cancer,* 7**,** 165-97.
POWNALL, M. E. & ISAACS, H. V. 2010. FGF Signalling in Vertebrate Development.
POWNALL, M. E., TUCKER, A. S., SLACK, J. M. & ISAACS, H. V. 1996. eFGF, Xcad3 and Hox genes form a molecular pathway that establishes the anteroposterior axis in Xenopus. *Development,* 122**,** 3881-92.
PYLAYEVA-GUPTA, Y., GRABOCKA, E. & BAR-SAGI, D. 2011. RAS oncogenes: weaving a tumorigenic web. *Nat Rev Cancer,* 11**,** 761-74.

PÉZERON, G., LAMBERT, G., DICKMEIS, T., STRÄHLE, U., ROSA, F. M. & MOURRAIN, P. 2008. Rasl11b knock down in zebrafish suppresses one-eyed-pinhead mutant phenotype. *PLoS One,* 3**,** e1434.

RAIBLE, F. & BRAND, M. 2001. Tight transcriptional control of the ETS domain factors Erm and Pea3 by Fgf signaling during early zebrafish development. *Mech Dev,* 107**,** 105-17.

RAPRAEGER, A. C., KRUFKA, A. & OLWIN, B. B. 1991. Requirement of heparan sulfate for bFGF-mediated fibroblast growth and myoblast differentiation. *Science,* 252**,** 1705-8.

ROBERTS, P. J. & DER, C. J. 2007. Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer. *Oncogene,* 26**,** 3291-310.

SAWADA, A., SHINYA, M., JIANG, Y. J., KAWAKAMI, A., KUROIWA, A. & TAKEDA, H. 2001. Fgf/MAPK signalling is a crucial positional cue in somite boundary formation. *Development,* 128**,** 4873-80.

SCHLESSINGER, J. 2000. Cell signaling by receptor tyrosine kinases. *Cell,* 103**,** 211-25.

SEBOLT-LEOPOLD, J. S. & HERRERA, R. 2004. Targeting the mitogen-activated protein kinase cascade to treat cancer. *Nat Rev Cancer,* 4**,** 937-47.

SEO, J. H., SUENAGA, A., HATAKEYAMA, M., TAIJI, M. & IMAMOTO, A. 2009. Structural and functional basis of a role for CRKL in a fibroblast growth factor 8-induced feed-forward loop. *Mol Cell Biol,* 29**,** 3076-87.

SHEN L, S. I. 2022. *GeneOverlap: Test and visualize gene overlaps* . R package version 1.32.0.

SHOWELL, C., BINDER, O. & CONLON, F. L. 2004. T-box genes in early embryogenesis. *Dev Dyn,* 229**,** 201-18.

SLACK, J. M. & FORMAN, D. 1980. An interaction between dorsal and ventral regions of the marginal zone in early amphibian embryos. *J Embryol Exp Morphol,* 56**,** 283-99.

SPEMANN, H. & MANGOLD, H. 1924. über Induktion von Embryonalanlagen durch Implantation artfremder Organisatoren. *Archiv für mikroskopische Anatomie und Entwicklungsmechanik,* 100**,** 599-638.

STOLLE, K., SCHNOOR, M., FUELLEN, G., SPITZER, M., CULLEN, P. & LORKOWSKI, S. 2007. Cloning, genomic organization, and tissue-specific expression of the RASL11B gene. *Biochim Biophys Acta,* 1769**,** 514-24.

STOREY, J. D. & TIBSHIRANI, R. 2003. Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A,* 100**,** 9440-5.

SUMMERTON, J. 1999. Morpholino antisense oligomers: the case for an RNase H-independent structural type. *Biochim Biophys Acta,* 1489**,** 141-58.

TANAKA, M., YOSHIMOTO, T. & NAKAMURA, T. 2017. A double-edged sword: The world according to Capicua in cancer. *Cancer Sci,* 108**,** 2319-2325.

TAPSCOTT, S. J. 2005. The circuitry of a master switch: Myod and the regulation of skeletal muscle gene transcription. *Development,* 132**,** 2685-95.

TEAM, R. C. 2021. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

TEN BERGE, D., BRUGMANN, S. A., HELMS, J. A. & NUSSE, R. 2008. Wnt and FGF signals interact to coordinate growth with cell fate specification during limb development. *Development,* 135**,** 3247-57.

TEZZE, C., ROMANELLO, V. & SANDRI, M. 2019. FGF21 as Modulator of Metabolism in Health and Disease. *Front Physiol,* 10**,** 419.

TINDALL, A. J., MORRIS, I. D., POWNALL, M. E. & ISAACS, H. V. 2007. Expression of enzymes involved in thyroid hormone metabolism during the early development of Xenopus tropicalis. *Biol Cell,* 99**,** 151-63.

TSANG, M. & DAWID, I. B. 2004. Promotion and attenuation of FGF signaling through the Ras-MAPK pathway. *Sci STKE,* 2004**,** pe17.

URNOV, F. D., REBAR, E. J., HOLMES, M. C., ZHANG, H. S. & GREGORY, P. D. 2010. Genome editing with engineered zinc finger nucleases. *Nat Rev Genet,* 11**,** 636-46.

WESCHE, J., HAGLUND, K. & HAUGSTEN, E. M. 2011. Fibroblast growth factors and their receptors in cancer. *Biochem J,* 437**,** 199-213.

WHITE, K. E., CARN, G., LORENZ-DEPIEREUX, B., BENET-PAGES, A., STROM, T. M. & ECONS, M. J. 2001. Autosomal-dominant hypophosphatemic rickets (ADHR) mutations stabilize FGF-23. *Kidney Int,* 60**,** 2079-86.

YAN, D., CHEN, D., COOL, S. M., VAN WIJNEN, A. J., MIKECZ, K., MURPHY, G. & IM, H. J. 2011. Fibroblast growth factor receptor 1 is principally responsible for fibroblast growth factor 2-induced catabolic activities in human articular chondrocytes. *Arthritis Res Ther,* 13**,** R130.

YANG, L., CRANSON, D. & TRINKAUS-RANDALL, V. 2004. Cellular injury induces activation of MAPK via P2Y receptors. *J Cell Biochem,* 91**,** 938-50.

YEH, E., FANGANIELLO, R. D., SUNAGA, D. Y., ZHOU, X., HOLMES, G., ROCHA, K. M., ALONSO, N., MATUSHITA, H., WANG, Y., JABS, E. W. & PASSOS-BUENO, M. R. 2013. Novel molecular pathways elicited by mutant FGFR2 may account for brain abnormalities in Apert syndrome. *PLoS One,* 8**,** e60439.

YUN, Y. R., WON, J. E., JEON, E., LEE, S., KANG, W., JO, H., JANG, J. H., SHIN, U. S. & KIM, H. W. 2010. Fibroblast growth factors: biology, function, and application for tissue regeneration. *J Tissue Eng,* 2010**,** 218142.

ZHANG, X., IBRAHIMI, O. A., OLSEN, S. K., UMEMORI, H., MOHAMMADI, M. & ORNITZ, D. M. 2006. Receptor specificity of the fibroblast growth factor family. The complete mammalian FGF family. *J Biol Chem,* 281**,** 15694-700.

ZHANG, Y., LIN, Y., BOWLES, C. & WANG, F. 2004. Direct cell cycle regulation by the fibroblast growth factor receptor (FGFR) kinase through phosphorylation-dependent release of Cks1 from FGFR substrate 2. *J Biol Chem,* 279**,** 55348-54.

ZHAO, J., YUAN, X., FRÖDIN, M. & GRUMMT, I. 2003. ERK-dependent phosphorylation of the transcription initiation factor TIF-IA is required for RNA polymerase I transcription and cell growth. *Mol Cell,* 11**,** 405-13.